Journal of Difference Equations and Applications
Vol. 00, No. 00, January 2008, 1–23

Some dynamical properties of Benford sequences

Arno Berger*

Mathematical and Statistical Sciences
University of Alberta, Edmonton, Canada

(*Received 25 August 2009; in revised form 30 October 2009*)

Dedicated to Peter E. Kloeden on the occasion of his sixtieth birthday

Numerical data generated by dynamical processes often obey Benford's Law of logarithmic mantissa distributions. For nonautonomous difference equations $x_n = T_n(x_{n-1})$, $n = 1, 2, \ldots$, this article presents necessary as well as sufficient conditions for (x_n) to conform with Benford's Law in its strongest form: The proportion of values in $\{x_0, x_1, \ldots, x_n\}$ with base b mantissa less than t tends to $\log_b t$ as $n \to \infty$, for all integer bases $b > 1$. The assumptions on (T_n), viz. asymptotic convexity and eventual expansivity on average, are very mild and met e.g. by practically all polynomial, rational, and exponential maps and any combinations thereof. The results complement, extend and unify previous work.

Keywords: Benford's Law, Benford sequence, uniform distribution mod 1, nonautonomous dynamical system, shadowing.

AMS Subject Classification: 11K06, 37A45, 37B55.

1. Introduction

Benford's Law (*BL*) is the probability distribution for the mantissa with respect to base $b \in \mathbb{N} \setminus \{1\}$ given by

$$\mathbb{P}(\text{mantissa}_b \le t) = \log_b t\,, \quad \forall t \in [1, b[\,; \tag{1}$$

the most well-known special case is that

$$\mathbb{P}(\text{first significant digit}_b = d) = \log_b\left(1 + d^{-1}\right), \quad \forall d = 1, \ldots, b-1\,.$$

First recorded by Newcomb [20] and popularised by Benford [1], BL is increasingly attracting the interest of scientists across a wide range of disciplines [7]. Empirical data following (1) have been discussed extensively, for instance in real-life data (e.g. physical constants, stock market indices, tax returns [14, 18, 23, 25]), in stochastic processes (e.g. sums and products of random variables [12, 23]), and in deterministic sequences (e.g. $(n!)$ and Fibonacci numbers [1, 10, 11, 16]). For dynamical systems, be they autonomous or nonautonomous, deterministic or stochastic, discrete or continuous, a thorough mathematical analysis of BL has recently been initiated, and it has been demonstrated that, in one way or the other, the asymptotic distribution of orbits and trajectories obeys (1) surprisingly often [3, 5, 6, 8, 12, 26, 27].

*Email: aberger@math.ualberta.ca

ISSN: 1023-6198 print/ISSN 1563-5120 online

DOI: 10.1080/1023619YYxxxxxxxx
http://www.informaworld.com

Most of these results are rather limited in their scope. It is the purpose of the present article to more comprehensively elucidate when, and when not, dynamical systems should be expected to follow BL.

Concretely, given x_0 let the sequence (x_n) of real numbers be generated iteratively by

$$x_n = T_n(x_{n-1}), \quad n = 1, 2, \ldots, \tag{2}$$

with measurable maps $T_n : \mathbb{R} \to \mathbb{R}$. Assuming that (x_n) conforms with a strong form of BL, can anything interesting at all be said about (T_n)? For the autonomous case, that is, for T_n not depending on n, Section 2 provides a complete answer to this question, by way of Theorems 2.5 and 2.10. Also, it is demonstrated through examples that the conclusions of both theorems are best possible in general. Section 3 is devoted to the strengthening of known BL results for (2) by embedding them into a more comprehensive and general setting. For a concrete example, consider (T_n) according to

$$T_n(x) = \begin{cases} e^x + 1 & \text{if } n \text{ is an odd prime number,} \\ x^2 - 2 & \text{if } n \text{ is even,} \\ x + 3 & \text{otherwise.} \end{cases}$$

For this system, the results in [4, 5] do not apply. However, it follows from the main result of Section 3, Theorem 3.7, that for Lebesgue almost every $x_0 \in \mathbb{R}$ the sequence (x_n) generated by (2) follows BL for every base b. The key ingredients in the proof of the very general Theorem 3.7 are versatile notions of *asymptotic convexity* and *eventual expansivity on average* for (T_n), together with a powerful nonautonomous shadowing technique that enables the application of standard uniform distribution tools. Simple examples show how the results presented here extend and unify previous work.

Throughout the article, the following, mostly standard terminology and notation will be used: The sets of natural, non-negative integer, integer, rational, positive real, and real numbers are symbolised by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}^+$, $\mathbb{R}$, respectively. For every $a, x \in \mathbb{R}^+$ with $a \neq 1$, $\log_a x$ denotes the base a logarithm of x; if used without a subscript, log is understood as the natural logarithm. For notational convenience, $\log_a 0 := 0$ for all $a \in \mathbb{R}^+ \backslash \{1\}$. The integers $\lfloor x \rfloor$ and $\lceil x \rceil$ denote, respectively, the largest integer not larger, and the smallest integer not smaller than $x \in \mathbb{R}$; the fractional (or non-integer) part of x is $[\![x]\!] := x - \lfloor x \rfloor$. The cardinality of the finite set A is $\#A$, and λ symbolises Lebesgue measure on $\mathbb{R}$ (or parts thereof). The indicator function of any set B is denoted by $\mathbb{1}_B$, that is, $\mathbb{1}_B(x)$ equals 1 or 0 depending on whether $x \in B$ or $x \notin B$. If $a \in \mathbb{R}$ and $B \subset \mathbb{R}$ then $a + B := \{a + b : b \in B\}$. For every $a \in \mathbb{R}$, the Dirac measure at a is δ_a, hence $\delta_a(B) = \mathbb{1}_B(a)$ for every $B \subset \mathbb{R}$. The support of any Borel (probability) measure μ on $\mathbb{R}$, symbolised as $\operatorname{supp} \mu$, by definition is the smallest closed set with full measure. Thus for example $\operatorname{supp} \delta_a = \{a\}$. As usual, the terms *absolutely continuous* and *almost every(where)* are abbreviated as *a.c.* and *a.e.*, respectively, and are understood relative to some reference measure that usually is clear from the context and not specified explicitly. If $f : X \to \mathbb{R}$ is measurable and μ is any (probability) measure on X then $f\mu$, defined via $f\mu(A) := \mu\big(f^{-1}(A)\big)$, is a (probability) measure on $\mathbb{R}$. For example, given any (probability) measure μ on $\mathbb{R}$, 2^μ and $\lfloor \mu \rfloor$ are (probability) measures concentrated on, respectively, $\mathbb{R}^+$ and $\mathbb{Z}$. The map $T : X \to X$ *preserves* μ, or μ is T-*invariant*, if $T\mu = \mu$.

2. Benford sequences: basic properties and examples

Given any natural number b larger than one, every $x \in \mathbb{R}^+$ can be written uniquely as $x = \langle x \rangle_b b^k$ with $1 \le \langle x \rangle_b < b$ and the appropriate $k \in \mathbb{Z}$. The number b is usually referred to as a *base*, and $\langle \cdot \rangle_b : \mathbb{R}^+ \to [1, b[$ is the (*base b*) *mantissa function.* Note that $\langle x \rangle_b = b^{[\![\log_b x]\!]}$ for all $x \in \mathbb{R}^+$. Also, $\lfloor \langle x \rangle_b \rfloor \in \{1, \ldots, b-1\}$ is called the *first significant digit* (*base b*) of x, henceforth symbolised as $\mathrm{fsd}_b x$. For notational convenience, define $\langle 0 \rangle_b := 0$, and hence $\mathrm{fsd}_b 0 = 0$ for every base b.

A most basic way of generating numerical data is through explicitly or recursively defined *sequences.* A sequence may conform with BL in a rather strong sense.

Definition 2.1 A sequence $(x_n)_{n \in \mathbb{N}_0}$ of real numbers is b-*Benford* if

$$\lim_{n\to\infty} \frac{\#\{j < n : \langle |x_j| \rangle_b \le t\}}{n} = \log_b t\,, \quad \forall t \in [1, b[\,;$$

it is (*strictly*) *Benford* if it is b-Benford for every $b \in \mathbb{N}\backslash\{1\}$.

The following correspondence between the b-Benford property and uniform distribution modulo one is well known. The term *uniformly distributed modulo one* is henceforth abbreviated as *u.d.* mod 1.

Proposition 2.2 ([11].) *A sequence* $(x_n)_{n \in \mathbb{N}_0}$ *of real numbers is b-Benford if and only if the sequence* $(\log_b |x_n|)_{n \in \mathbb{N}_0}$ *is u.d.* mod 1.

Example 2.3 Using Proposition 2.2 and standard results from uniform distribution theory (as can be found e.g. in [17]), examples of Benford sequences are easily produced.

(i) The sequence $(n!)$ is Benford, see [5, Exp.5.4(ii)] or [11, Thm.3].

(ii) Let (x_n) be generated by $x_n = x_{n-1} + x_{n-2}$, $n \ge 2$. The sequence thus defined is Benford except for the trivial case $(x_0, x_1) = (0, 0)$. In particular, the sequence of Fibonacci numbers, corresponding to $(x_0, x_1) = (1, 1)$, is Benford [3, Exp.3.5(i)].

(iii) Define (x_n) recursively as $x_n = x_{n-1}^2 + 1$, $n \ge 1$. For (Lebesgue) almost every $x_0 \in \mathbb{R}$, the sequence (x_n) is Benford. There are, however, also exceptional points: [5, Exp.4.3] explicitly gives x_0 such that $\mathrm{fsd}_{10} x_n = 9$ for all n. Another example is

$$x_0 = \lim_{n\to\infty} \sqrt{\ldots \sqrt{\sqrt{10^{\frac{2}{3}\cdot 2^n} - 1} - 1} \ldots - 1} = 4.53002223124696101566\ldots\,,$$

for which $(\mathrm{fsd}_{10}\, x_n) = (4, 2, 4, 2, \ldots)$ is 2-periodic. It is an open problem whether (x_n) is b-Benford for any b if $x_0 = 0$.

(iv) The sequence (2^n) is b-Benford precisely if $\log_b 2$ is irrational, i.e., if and only if b is not of the form 2^j for some $j \in \mathbb{N}$.

(v) For every $a \in \mathbb{R}^)$, the sequence (n^a), and similarly the sequence of prime numbers (p_n), is not b-Benford for any b, see [5]. Note, however, that these sequences conform with BL in a weaker sense [24].

Some elementary properties of Benford sequences follow directly from Proposition 2.2.

Proposition 2.4 (cf. [13].) *Let* (x_n) *be a sequence of real numbers, and* $b \in \mathbb{N}\backslash\{1\}$. *Then:*

(i) *If* (x_n) *is b-Benford then so is* (αx_n^j) *for every non-zero* $\alpha \in \mathbb{R}$ *and* $j \in \mathbb{N}$.
(ii) *If* $x_n \ne 0$ *for all n then* (x_n) *is b-Benford if and only if* (x_n^{-1}) *is b-Benford.*
(iii) *If* (x_n) *is* b^j*-Benford for some* $j \in \mathbb{N}$ *then* (x_n) *is b-Benford.*

Intuitively, it is plausible that if a sequence is to be Benford then this puts severe constraints on its distributional behaviour. To make this more formal, denote by $\mathcal{P}(\mathbb{R})$ and $\mathcal{P}(\overline{\mathbb{R}})$ the set of all Borel probability measures on the real line $\mathbb{R}$ and the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, respectively. Recall that $\mathcal{P}(\mathbb{R})$ and $\mathcal{P}(\overline{\mathbb{R}})$ are complete metrizable spaces when endowed with the topology of weak convergence, and $\mathcal{P}(\overline{\mathbb{R}})$ is compact. Also recall that if (μ_n) is a sequence in $\mathcal{P}(\overline{\mathbb{R}})$ with $\mu_n(\mathbb{R}) \equiv 1$ then $\mu_n \xrightarrow{w} \mu$ in $\mathcal{P}(\overline{\mathbb{R}})$ means that

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu_n \;\to\; \int_{\mathbb{R}} f \, \mathrm{d}\mu + \mu(\{-\infty\}) f_- + \mu(\{+\infty\}) f_+$$

holds for every continuous function $f : \mathbb{R} \to \mathbb{R}$ for which the limits $f_- := \lim_{x \to -\infty} f(x)$ and $f_+ := \lim_{x \to +\infty} f(x)$ exist. Given any real sequence (x_n), consider the empirical averages

$$\eta_n := \frac{1}{n+1} \sum\nolimits_{j=0}^{n} \delta_{x_j} \in \mathcal{P}(\mathbb{R}) \, .$$

By compactness, (η_n) has a (weak) accumulation point in $\mathcal{P}(\overline{\mathbb{R}})$. Denote by $\mathcal{A}[(x_n)] \subset \mathcal{P}(\overline{\mathbb{R}})$ the (non-empty) set of all accumulation points of (η_n). It is readily confirmed that $\mathcal{A}[(x_n)]$ is compact and connected (though not necessarily path-connected), with $\mu\big(\overline{\{x_j : j \in \mathbb{N}_0\}}\big) = 1$ for every $\mu \in \mathcal{A}[(x_n)]$. If (x_n) is Benford then each element of $\mathcal{A}[(x_n)]$ has a very simple structure.

Theorem 2.5. *Assume that (x_n) is b-Benford for infinitely many $b \in \mathbb{N}\backslash\{1\}$. Then for every $\mu \in \mathcal{A}[(x_n)]$, there exist numbers $p_-, p_0, p_+ \geq 0$ with $p_- + p_0 + p_+ = 1$ such that $\mu = p_- \delta_{-\infty} + p_0 \delta_0 + p_+ \delta_{+\infty}$.*

Proof. Assume $\mu \in \mathcal{A}[(x_n)]$ and let $\eta_{n_k} \xrightarrow{w} \mu$. Given $0 < s < t$ and $0 < \varepsilon < 1$, choose b so large that (x_n) is b-Benford and also

$$\frac{|\log s| + |\log t|}{\log b} < \varepsilon \, .$$

Note that this choice implies $s > b^{-1}$ and $t < b$. Assume first that $s \geq 1$, hence $\langle x \rangle_b = x$ for every $s < x < t$, and

$$\begin{aligned} \mu(]s,t[) &\leq \underline{\lim}_{k\to\infty} \eta_{n_k}(]s,t[) = \underline{\lim}_{k\to\infty} \frac{1}{n_k+1} \sum\nolimits_{j=0}^{n_k} \delta_{x_j}(]s,t[) \\ &= \underline{\lim}_{k\to\infty} \frac{\#\{j \leq n_k : s < x_j < t\}}{n_k+1} = \underline{\lim}_{k\to\infty} \frac{\#\{j \leq n_k : s < \langle x_j \rangle_b < t\}}{n_k+1} \\ &= \log_b t - \log_b s \leq \frac{|\log s| + |\log t|}{\log b} < \varepsilon \, . \end{aligned}$$

Similarly, if $t \leq 1$ then $\langle x \rangle_b = xb$ for every $s < x < t$, and

$$\mu(]s,t[) \leq \underline{\lim}_{k\to\infty} \frac{\#\{j \leq n_k : sb < \langle x_j \rangle_b < tb\}}{n_k+1} = |\log_b s| - |\log_b t| < \varepsilon \, .$$

Finally, if $s < 1 < t$ then by the above

$$\mu(]s,t[) = \mu(]s,1[) + \mu(\{1\}) + \mu(]1,t[) \le |\log_b s| + \lim_{k\to\infty} \eta_{n_k}(\{1\}) + \log_b t$$
$$= \frac{|\log s| + \log t}{\log b} < \varepsilon\,.$$

Since $\varepsilon > 0$ was arbitrary, $\mu(]s,t[) = 0$. Analogously, $\mu(]s,t[) = 0$ whenever $s < t < 0$. Overall therefore $\mu(\mathbb{R}\backslash\{0\}) = 0$, hence $\mu(\{-\infty, 0, +\infty\}) = 1$, and the claim follows. □

REMARK 2.6 In a wider context unrelated to BL, Theorem 2.5 is a consequence of the following simple observation: Let ν be a probability measure on $[0,1]$ with $\nu(\{0,1\}) = 0$, and let $\ell : \mathbb{N}\backslash\{1\} \times \mathbb{R}^+ \to \mathbb{R}$ have the property that

$$\lim_{b\to\infty} \big|\ell(b,x)\big| = 0\,, \quad \text{locally uniformly on } \mathbb{R}^+\,.$$

Given a real sequence (x_n), assume that $\big([\![\ell(b,|x_n|)]\!]\big)$ is distributed according to ν for infinitely many b. Then, by an argument very similar to the one above, every $\mu \in \mathcal{A}[(x_n)]$ is a convex combination of $\delta_{-\infty}$, δ_0, and $\delta_{+\infty}$. Evidently, Theorem 2.5 corresponds to the special case $\ell(b,x) = \log_b x$ and $\nu = \lambda|_{[0,1]}$

Theorem 2.5 provides a simple necessary condition for (x_n) to be Benford. This condition is clearly not sufficient in general, as can be seen for instance from $(x_n) = (n)$ for which $\mathcal{A}[(x_n)] = \{\delta_{+\infty}\}$ yet (x_n) is not b-Benford for any b. On the other hand, every compact and connected subset of $\mathcal{P}(\overline{\mathbb{R}})$ allowed by Theorem 2.5 equals $\mathcal{A}[(x_n)]$ for some Benford sequence (x_n), as demonstrated by Example 2.9 below. Also, contrary to $\mathcal{A}[(x_n)] \subset \mathcal{P}(\overline{\mathbb{R}})$, the set of accumulation points of (x_n) or $(\frac{1}{n+1}\sum_{j=0}^{n} x_j)$ in $\overline{\mathbb{R}}$ may have a less trivial structure, even if (x_n) is Benford. Finally, Example 2.11 shows that Theorem 2.5 may fail if (x_n) is b-Benford only for finitely many bases b. As a statement about Benford sequences, therefore, Theorem 2.5 is best possible in that in general neither its assumptions can be weakened nor its conclusions strengthened. In preparation for these examples, a simple fact about uniform distribution will be formulated. To this end let (N_j), (L_j), and (ξ_j) be sequences of natural, non-negative integer, and real numbers, respectively, and consider any sequence (x_n) with

$$x_n = \xi_j + n\varepsilon_j\vartheta\,, \qquad \forall n : 1 \le n - (N_1 + L_1 + \ldots + N_{j-1} + L_{j-1}) \le N_j\,, \tag{3}$$

where $\vartheta \in \mathbb{R}$, $\varepsilon_j \in \{-1, 1\}$, and $N_0 := L_0 := 0$. Thus (x_n) consists of arithmetic progressions of lengths $N_1, N_2, \ldots$ with the same increment ϑ (in absolute value) but possibly different off-sets ξ_j, interspersed with arbitrary segments of lengths $L_1, L_2, \ldots$, that is, (x_n) has the form

$$(x_n) = \Big(\underbrace{\xi_1 + \varepsilon_1\vartheta, \xi_1 + 2\varepsilon_1\vartheta, \ldots, \xi_1 + N_1\varepsilon_1\vartheta}_{N_1\text{-term arithmetic progression}}, \underbrace{x_{N_1+1}, \ldots, x_{N_1+L_1}}_{L_1 \text{ arbitrary terms}},$$
$$\underbrace{\xi_2 + (N_1+L_1+1)\varepsilon_2\vartheta, \ldots, \xi_2 + (N_1+L_1+N_2)\varepsilon_2\vartheta}_{N_2\text{-term arithmetic progression}},$$
$$\underbrace{x_{N_1+L_1+N_2+1}, \ldots, x_{N_1+L_1+N_2+L_2}}_{L_2 \text{ arbitrary terms}}, \ldots\Big)\,.$$

In the special case $L_j \equiv 0$, $\xi_j \equiv 0$, and $\varepsilon_j \equiv 1$ the sequence (x_n) simply equals

$(n\vartheta)$ and hence is u.d. mod 1 if and only if $\vartheta \in \mathbb{R}\backslash\mathbb{Q}$. The following is a mild generalisation of this familiar fact.

Lemma 2.7. *Let (N_j), (L_j), (ε_j), and (ξ_j) be sequences in, respectively, $\mathbb{N}$, $\mathbb{N}_0$, $\{-1,1\}$, and $\mathbb{R}$, with*

$$\lim_{j\to\infty} \frac{L_1+\ldots+L_j}{N_1+\ldots+N_j} = 0\,.$$

Assume the sequence (x_n) obeys (3) with some $\vartheta \in \mathbb{R}$. Then:

(i) *If* $\lim_{j\to\infty} \dfrac{N_1+\ldots+N_j}{j} = +\infty$ *and $\vartheta \in \mathbb{R}\backslash\mathbb{Q}$ then (x_n) is u.d.* mod 1.

(ii) *If* $\overline{\lim}_{j\to\infty} \dfrac{N_j}{N_1+\ldots+N_j} > \frac{1}{2}$ *and $\vartheta\in\mathbb{Q}$ then (x_n) is not u.d.* mod 1.

Proof. To prove (i), let ϑ be irrational and assume $1 \le n-(N_1+L_1+\ldots+N_{m-1}+L_{m-1}) \le N_m$ for some $m\in\mathbb{N}$. Then, for every $h\in\mathbb{Z}\backslash\{0\}$,

$$\begin{aligned}\left|\sum\nolimits_{j=1}^{n} e^{2\pi i h x_j}\right| = \Big|\sum\nolimits_{j=1}^{m-1} e^{2\pi i h \xi_j} \sum\nolimits_{k=1}^{N_j} e^{2\pi i h(k+N_1+L_1+\ldots+N_{j-1}+L_{j-1})\varepsilon_j\vartheta} + \\ + \sum\nolimits_{j=1}^{m-1}\sum\nolimits_{k=1}^{L_j} e^{2\pi i h x_{k+N_1+L_1+\ldots+N_{j-1}+L_{j-1}+N_j}} + \\ + e^{2\pi i h\xi_m} \sum\nolimits_{j=1+N_1+L_1+\ldots+N_{m-1}+L_{m-1}}^{n} e^{2\pi i h j \varepsilon_m \vartheta}\Big| \\ \le \frac{2m}{|e^{2\pi i h\vartheta}-1|} + L_1+\ldots+L_{m-1}\,.\end{aligned}$$

For $1 \le n-(N_1+L_1+\ldots+N_{m-1}+L_{m-1}+N_m) \le L_m$ the same estimate holds with m replaced by $m+1$. With $\lim_{j\to\infty}\dfrac{j}{N_1+\ldots+N_j} = 0$, it follows that $\lim_{n\to\infty} n^{-1}\sum_{j=1}^n e^{2\pi i h x_j} = 0$, and since h was arbitrary, (x_n) is u.d. mod 1.

To verify (ii), assume that $\vartheta = p/q$ with $(p,q)\in\mathbb{Z}\times\mathbb{N}$ and p,q relatively prime. For $h=q$, the same explicit computation as before shows that

$$\begin{aligned}\sum\nolimits_{j=1}^{n} e^{2\pi i q x_j} = \sum\nolimits_{j=1}^{m-1} e^{2\pi i q\xi_j} N_j + \sum\nolimits_{j=1}^{m-1}\sum\nolimits_{k=1}^{L_j} e^{2\pi i q x_{k+N_1+L_1+\ldots+N_{j-1}+L_{j-1}+N_j}} + \\ + e^{2\pi i q \xi_m}\big(n-(N_1+L_1+\ldots+N_{m-1}+L_{m-1})\big)\,,\end{aligned}$$

and consequently, for $n = N_1+L_1+\ldots+N_m+L_m$,

$$\left|\frac{1}{n}\sum\nolimits_{j=1}^{n} e^{2\pi i q x_j} - \frac{\sum_{j=1}^m e^{2\pi i q\xi_j}N_j}{N_1+\ldots+N_m}\right| \le 2\frac{L_1+\ldots+L_m}{N_1+\ldots+N_m}\,.$$

The assumption on (N_j) implies

$$\begin{aligned}\overline{\lim}_{n\to\infty}\left|\frac{1}{n}\sum\nolimits_{j=1}^{n} e^{2\pi i q x_j}\right| \ge \overline{\lim}_{m\to\infty}\left|\frac{\sum_{j=1}^m e^{2\pi i q\xi_j}N_j}{N_1+\ldots+N_m}\right| \\ \ge 2\,\overline{\lim}_{j\to\infty}\frac{N_j}{N_1+\ldots+N_j} - 1 > 0\,,\end{aligned}$$

showing that $(n^{-1}\sum_{j=1}^{n} e^{2\pi i q x_j})$ does not converge to 0, and hence (x_n) is not u.d. mod 1. □

Remark 2.8 The assumptions in Lemma 2.7(i) are best possible in that the assertion may fail if either $\underline{\lim}_{j\to\infty}\dfrac{N_1+\ldots+N_j}{j} < +\infty$ or $\overline{\lim}_{j\to\infty}\dfrac{L_1+\ldots+L_j}{N_1+\ldots+N_j} > 0$. On the other hand, the assumption on (N_j) in (ii) clearly is restrictive and can be relaxed considerably in special situations, that is, with further conditions imposed on (ε_j) and (ξ_j). Note, however, that if merely $\dfrac{N_1+\ldots+N_j}{j} \to +\infty$, i.e. in the setting of (i), the rationality of ϑ does not necessarily rule out uniform distribution of (x_n). A simple example is $(N_j) = (j)$, $L_j \equiv 0$, $\varepsilon_j \equiv 1$, $\xi_j = j\xi$ with some irrational ξ, and $\vartheta = 1$. In this case

$$\left|\sum_{j=1}^{n} e^{2\pi i h x_j}\right| \le (1+\sqrt{2n})\left(1+\frac{2}{|e^{2\pi i h\xi}-1|^2}\right)$$

holds for all $n \in \mathbb{N}$ and $h \in \mathbb{Z}\backslash\{0\}$, and hence (x_n) is u.d. mod 1.

Example 2.9 Given any probability vector $p = (p_-, p_0, p_+) \in \mathbb{R}^3$, that is, $p_-, p_0, p_+ \ge 0$ and $p_- + p_0 + p_+ = 1$, there exists a sequence (x_n) that is b-Benford for infinitely many but not all bases b, and $\mathcal{A}[(x_n)] = \{\nu_p\}$ where $\nu_p = p_-\delta_{-\infty} + p_0\delta_0 + p_+\delta_{+\infty}$. To explicitly define such a sequence, assume w.l.o.g. $p_-, p_0, p_+ > 0$ and let $\mathbb{N} = J_- \cup J_0 \cup J_+$ with

$$\begin{aligned}
J_- &= \bigcup_{j=1}^{\infty} \big(M_1+\ldots+M_{j-1} + [1, \lfloor p_- M_j\rfloor]\big) \cap \mathbb{N}\,,\\
J_0 &= \bigcup_{j=1}^{\infty} \big(M_1+\ldots+M_{j-1} + \lfloor p_- M_j\rfloor + [1, \lfloor p_0 M_j\rfloor]\big) \cap \mathbb{N}\,,\\
J_+ &= \mathbb{N}\backslash(J_- \cup J_0)\,,
\end{aligned}$$

where (M_j) is a strictly increasing sequence of natural numbers yet to be specified further, and $M_0 := 0$. Thus $\mathbb{N}$ is partitioned into $\lfloor p_- M_1\rfloor$ elements of J_-, followed by $\lfloor p_0 M_1\rfloor$ elements of J_0, then $M_1 - \lfloor p_- M_1\rfloor - \lfloor p_0 M_1\rfloor$ elements of J_+, followed in turn by $\lfloor p_- M_2\rfloor$ elements of J_- etc. It is easy to check that each J_i indeed has density p_i, provided that $\lim_{j\to\infty}\dfrac{M_j}{M_1+\ldots+M_j} = 0$. The latter will be assumed from now on. For the sequence (x_n) defined according to

$$x_n = \begin{cases} -2^n & \text{if } n \in J_-\,,\\ 2^{-n} & \text{if } n \in J_0\,,\\ 2^n & \text{if } n \in J_+\,,\end{cases}$$

it is clear that $\eta_n \xrightarrow{w} \nu_p$ and therefore $\mathcal{A}[(x_n)] = \{\nu_p\}$. For every b not a power of 2, that is, for $b \notin \{2^j : j \in \mathbb{N}\}$, Lemma 2.7(i) with

$$N_j = \begin{cases} \lfloor p_- M_{\frac{j+2}{3}}\rfloor & \text{if } j \in 3\mathbb{N}_0+1\,,\\ \lfloor p_0 M_{\frac{j+1}{3}}\rfloor & \text{if } j \in 3\mathbb{N}_0+2\,,\\ M_{\frac{j}{3}} - \lfloor p_- M_{\frac{j}{3}}\rfloor - \lfloor p_0 M_{\frac{j}{3}}\rfloor & \text{if } j \in 3\mathbb{N}\,,\end{cases} \qquad \varepsilon_j = \begin{cases} -1 & \text{if } j \in 3\mathbb{N}_0+2\,,\\ 1 & \text{otherwise}\,,\end{cases}$$

as well as $L_j \equiv 0$ and $\vartheta = \log_b 2$, shows that (x_n) is b-Benford. While Lemma 2.7(ii) does not apply if $b = 2^j$ for some j, it is obvious from the definition of (x_n)

that $(j \log_b |x_n|)$ is a sequence of *integers* in this case and hence $(\log_b x_n)$ cannot be u.d. mod 1. Thus (x_n) is b-Benford exactly if $b \notin \{2^j : j \in \mathbb{N}\}$.

By applying the above idea repeatedly, it is not hard to show that, given any compact connected set $D \subset \mathcal{P}(\overline{\mathbb{R}})$ contained in the convex hull of $\delta_{-\infty}, \delta_0, \delta_{+\infty}$, a sequence (x_n) can be constructed with $\mathcal{A}[(x_n)] = D$ that is b-Benford unless $b \in \{2^j : j \in \mathbb{N}\}$.

From Example 2.9 it may also be deduced that (x_n) being b-Benford for infinitely many bases b does generally not allow any structural conclusions to be drawn about the set of accumulation points of (x_n) in $\overline{\mathbb{R}}$ or of (δ_{x_n}) in $\mathcal{P}(\overline{\mathbb{R}})$. Neither, for that matter, of $\left(\dfrac{x_1 + \ldots + x_n}{n}\right)$ in $\overline{\mathbb{R}}$, as the example

$$(x_n) = (e^1, e^2, \ldots, e^{N_1}, e^{-1}, e^{-2}, \ldots, e^{-N_2}, e^1, e^2, \ldots, e^{N_3}, e^{-1}, \ldots)$$

shows: If $N_j \to +\infty$ then (x_n) is Benford by Lemma 2.7(i). However, given any (possibly one-point) interval $[s,t] \subset [0, +\infty]$, the set of accumulation points of $\left(\dfrac{x_1 + \ldots + x_n}{n}\right)$ can be made to equal $[s,t]$ simply by choosing (N_j) appropriately. On the other hand and in accordance with Theorem 2.5, every $\mu \in \mathcal{A}[(x_n)]$ is a convex combination of δ_0 and $\delta_{+\infty}$.

As indicated in the introduction, this article focuses on the Benford property of sequences defined recursively by

$$x_j = T_j(x_{j-1}), \quad j \in \mathbb{N}, \tag{4}$$

where (T_j) is a sequence of measurable maps of the real or extended real line (or parts thereof) into itself. Relation (4) may be interpreted as a *nonautonomous dynamical system* (in discrete time). For every $n \in \mathbb{N}$, denote by T^n the composition $T^n := T_n \circ \ldots \circ T_1$, and $T^0 := \mathrm{id}$. The symbol $O_T(x)$, referred to as the *orbit* of x under (T_j), denominates the sequence generated by (4) subject to the initial condition $x_0 = x$; thus $O_T(x) = \big(T^n(x)\big)_{n \in \mathbb{N}_0}$. For any function $f : \mathbb{R} \to \mathbb{R}$ the symbol $f\big(O_T(x)\big)$ is understood to mean $\big(f(T^n(x))\big)$; thus for example $[\![O_T(x)]\!] = ([\![T^n(x)]\!])$. Note that the interpretation of $O_T(x_0)$ as a *sequence* differs from terminology sometimes used in dynamical systems theory (e.g. [15]) according to which the orbit of x_0 is the mere *set* $\{x_n : n \in \mathbb{N}_0\}$.

The remainder of this section is devoted to the *autonomous* version of (4), i.e. to T_j independent of j. In this case, Theorem 2.5 has a corollary worth noting.

Theorem 2.10. *Let X be a Borel subset of $\overline{\mathbb{R}}$ and assume the map $T : X \to X$ preserves a (Borel) probability measure μ, that is $T\mu = \mu$. Then*

$$\mu\big(\{x \in X : O_T(x) \text{ is Benford}\}\big) = 0. \tag{5}$$

Proof. W.l.o.g. assume that $X = \overline{\mathbb{R}}$. Let $B \subset \mathbb{N}\backslash\{1\}$ be an infinite set of bases and

$$X_B := \big\{x \in X : O_T(x) \text{ is } b\text{-Benford for all } b \in B\big\}.$$

It will now be shown that $\mu(X_B) = 0$. Clearly this implies (5).

Note first that if μ is ergodic then $\mathcal{A}\big[O_T(x)\big] = \{\mu\}$ for μ-almost every x. In this case, if $\mu = p_-\delta_{-\infty} + p_0\delta_0 + p_+\delta_{+\infty}$ then every $y \in \{-\infty, 0, +\infty\}$ is periodic. Clearly $\mu(X_B) = 0$ in this case. If, on the other hand, μ is not a convex combination of $\delta_{-\infty}$, δ_0, $\delta_{+\infty}$ then, by Theorem 2.5, $\mu(X_B) = 0$ as well. Thus $\mu(X_B) = 0$ holds whenever

μ is ergodic. In general, let $(\Omega, (R_\omega, \nu_\omega)_{\omega\in\Omega}, \mathbb{P})$ be an ergodic decomposition of μ, see e.g. [19, Sec.II.6]. Then $\nu_\omega(X_B) = 0$ for every $\omega \in \Omega$, hence

$$\begin{aligned}\mu(X_B) &= \int_{\mathbb{R}} \mathbb{1}_{X_B}(r)\,\mathrm{d}\mu(r) = \int_\Omega \int_{\mathbb{R}} \mathbb{1}_{X_B\cap R_\omega}(r)\,\mathrm{d}\nu_\omega(r)\,\mathrm{d}\mathbb{P}(\omega)\\ &= \int_\Omega \nu_\omega(X_B \cap R_\omega)\,\mathrm{d}\mathbb{P}(\omega) = 0\,,\end{aligned}$$

and the proof is complete. □

Example 2.11 Let $T : [0, 1] \to [0, 1]$ be the symmetric tent map $T(x) = 1-|2x-1|$. It is well known (and easy to check) that T is ergodic w.r.t. $\lambda|_{[0,1]}$. Define two maps $\tau_1, \tau_2 : [0, 1] \to [0, 1]$ as

$$\tau_1(x) = \tfrac{1}{2}x\,, \quad \tau_2(x) = 1 - \tfrac{1}{2}x\,.$$

Then $T \circ \tau_1(x) = T \circ \tau_2(x) = x$ for all $x \in [0, 1]$, and τ_1, τ_2 can be used for a symbolic description of the dynamics of T in a standard way. To this end, denote by Σ_2 the space of all sequences in $\{1, 2\}$, that is $\Sigma_2 = \{1, 2\}^{\mathbb{N}_0}$, which is a compact metrizable space when endowed with the product topology. The map

$$h : \begin{cases} \Sigma_2 & \to [0, 1] \\ (s_n) & \mapsto \lim_{j\to\infty} \tau_{s_0} \circ \tau_{s_1} \circ \ldots \circ \tau_{s_j}\left(\tfrac{1}{2}\right)\end{cases}$$

is well defined, continuous and onto. Every $x \in [0, 1]$ has at most two pre-images under h, and in fact $\#h^{-1}(\{x\}) = 1$ unless x is a dyadic rational, i.e., unless $2^j x$ is an integer for some $j \in \mathbb{N}$. Moreover, $T \circ h = h \circ \sigma$ holds with σ denoting the standard (left) shift on Σ_2 given by $\sigma\big((s_n)\big) = (s_{n+1})$ for every $(s_n) \in \Sigma_2$.

(i) Since T is ergodic w.r.t. Lebesgue measure, the Birkhoff ergodic theorem implies that for every base b and almost every $x \in [0, 1]$,

$$\begin{aligned}\lim_{n\to\infty} \frac{\#\{j < n : \langle T^j(x)\rangle_b \le t\}}{n} &= \lim_{n\to\infty} \frac{1}{n} \sum\nolimits_{j=0}^{n-1} \sum\nolimits_{k=1}^{\infty} \mathbb{1}_{[b^{-k}, tb^{-k}]}\big(T^j(x)\big)\\ &= \int_0^1 \sum\nolimits_{k=1}^{\infty} \mathbb{1}_{[b^{-k}, tb^{-k}]}(x)\,\mathrm{d}x = \sum\nolimits_{k=1}^{\infty} (t-1)b^{-k}\\ &= \frac{t-1}{b-1}\,.\end{aligned}$$

For almost every x, therefore, $\big(\langle T^n(x)\rangle_b\big)$ is *uniformly* rather than logarithmically distributed, and hence a typical orbit (in the sense of Lebesgue measure) is not b-Benford for any b. It will now be shown that given any *finite* set B of bases, the map T is conjugate and isomorphic (i.e., dynamically equivalent both in a topological and measure-theoretic sense) to a continuous map $S : [0, 1] \to [0, 1]$ preserving a probability measure equivalent to $\lambda|_{[0,1]}$ such that $O_S(x)$ is b-Benford for all $b \in B$ and almost all $x \in [0, 1]$.

To construct S, first choose an a.c. probability measure μ with $\operatorname{supp}\mu = [0, 1]$ such that $[\![\log_b \mu]\!]$ is uniform for all $b \in B$. Such a measure is easily found as follows. For every $a < 0$, denote by $U_{(a,0)}$ the uniform distribution on $[a, 0]$. Recall that the

Fourier transform of $U_{(a,0)}$ is

$$\widehat{U_{(a,0)}}(x) = \int_{\mathbb{R}} e^{itx}\, \mathrm{d}U_{(a,0)}(t) = -\frac{1}{a}\int_a^0 e^{itx}\mathrm{d}t = \frac{\sin(\frac{1}{2}ax)}{\frac{1}{2}ax}e^{\frac{1}{2}iax}\,.$$

Also, let ν_∞ be the negative exponential distribution, i.e. $\nu_\infty(]-\infty, x]) = \min(e^x, 1)$ for all $x \in \mathbb{R}$, the Fourier transform of which is

$$\widehat{\nu_\infty}(x) = \int_{-\infty}^0 e^{itx}e^t\,\mathrm{d}t = \frac{1}{1+ix}\,.$$

Let $\beta := \min B \geq 2$ and consider the convolution of the $\#B$ uniform distributions $U_{(-\log_\beta b,0)}$, $b \in B$ and ν_∞, i.e.

$$\nu := \ast_{b\in B} U_{(-\log_\beta b,0)} \ast \nu_\infty\,;$$

here $\ast$ denotes the usual convolution of probabilities on the real line. Clearly, ν is a.c. w.r.t. λ, and $\operatorname{supp}\nu =]-\infty, 0]$. The Fourier transform of ν is

$$\widehat{\nu}(x) = e^{-\frac{1}{2}ix\sum_{b\in B}\log_\beta b}\,\frac{1}{1+ix}\prod\nolimits_{b\in B}\frac{\sin(\frac{1}{2}x\log_\beta b)}{\frac{1}{2}x\log_\beta b}\,.$$

Since $\widehat{\nu}(2k\pi\log_b\beta) = 0$ for all $k \in \mathbb{Z}\backslash\{0\}$ and $b \in B$, the probability measure $[\![(\log_b\beta)\nu]\!]$ equals $\lambda|_{[0,1]}$ for every $b \in B$. (This uses the readily confirmed fact that $\mu \in \mathcal{P}(\mathbb{R})$ satisfies $[\![\mu]\!] = \lambda|_{[0,1]}$ if and only if $\widehat{\mu}(2\pi k) = 0$ for all $k \in \mathbb{Z}\backslash\{0\}$.) Define now $\mu := \beta^\nu$, that is

$$F_\mu(x) := \mu([0,x]) = \nu(]-\infty, \log_\beta x])\,, \quad \forall x > 0\,.$$

From the above it is clear that the measure μ thus constructed has all the desired properties: It is equivalent to $\lambda|_{[0,1]}$, and $[\![\log_b\mu]\!] = [\![(\log_b\beta)\nu]\!] = \lambda|_{[0,1]}$ for all $b \in B$. The distribution function F_μ of μ maps $[0,1]$ homeomorphically onto itself, with $F_\mu(0) = 0$ and $F_\mu(1) = 1$. Since $0 < F'_\mu(x) < +\infty$ for all $0 < x < 1$, $F_\mu|_{]0,1[}$ is a diffeomorphism. With these preparations, define $S : [0,1] \to [0,1]$ as

$$S := F_\mu^{-1}\circ T\circ F_\mu\,.$$

It is easy to see that S is a continuous, tent-like map with $S(0) = S(1) = 0$ and $S\big(F_\mu^{-1}(\frac{1}{2})\big) = 1$. By construction, (S,μ) is topologically conjugate and isomorphic to (T,λ). Hence dynamically T and S are identical, up to the change of variables brought about by the homeomorphism F_μ. Moreover, [2, Lem.4.10] shows that S is ergodic w.r.t. μ. For Lebesgue almost every x, therefore, $O_S(x)$ is b-Benford for all $b \in B$. Thus Theorem 2.5 and 2.10 may fail if the b-Benford property is stipulated only for *finitely many* bases b.

A similar approach can be carried out whenever T is ergodic w.r.t. a probability measure equivalent to λ. A popular example in this regard is the logistic map $Q_4 : x \mapsto 4x(1-x)$. Almost no orbit of such a map typically is b-Benford for any b, but given any finite set B of bases, there exists a dynamically equivalent map S such that $O_S(x)$ is b-Benford for every $b \in B$ and Lebesgue almost all x. One may conclude that for an autonomous system the b-Benford property of most of its orbits and finitely many b does not have much dynamical significance. As evidenced through Theorem 2.5 and 2.10, the situation is completely different if infinitely many bases are considered.

(ii) It was explained above why in the case of the symmetric tent map the orbit $O_T(x)$ is, for almost every $x \in [0,1]$, not b-Benford for any b. Thus it is natural to ask whether $O_T(x)$ is b-Benford at all for *some* x and b. This question will now be answered in the affirmative. For this purpose, pick a sequence (N_j) in $\mathbb{N}$ with $N_j \to +\infty$ and consider the symbolic sequence

$$s^* = (s_n^*) = \big(\underbrace{1,\ldots,1}_{N_1 \text{ times}}, 2, \underbrace{1,\ldots,1}_{N_2 \text{ times}}, 2, \underbrace{1,\ldots,1}_{N_3 \text{ times}}, 2, 1\ldots\big) \in \Sigma_2\,. \tag{6}$$

Define $x^* := h(s^*) \in [0,1]$. With $I_1 = [0, \frac{1}{2}]$, $I_2 = [\frac{1}{2}, 1]$, note that $T^n(x^*) \in I_{s_n^*}$ for all $n \in \mathbb{N}_0$. Consequently, $O_T(x^*)$ stays in I_1 for the first N_1 steps, then makes a one-step excursion to I_2, then remains in I_1 for N_2 steps etc. Since for all $x \in I_1$,

$$\log_b T(x) = \log_b(2x) = \log_b 2 + \log_b x\,,$$

Lemma 2.7(i) applies whenever $\log_b 2$ is irrational, and consequently $O_T(x^*)$ is b-Benford for every $b \notin \{2^j : j \in \mathbb{N}\}$. As (6) provides uncountably many *different* points x^*, the set

$$\big\{x \in [0,1] : O_T(x) \,\text{is}\, b\text{-Benford for all } b \in \mathbb{N}\backslash\{2^j : j \in \mathbb{N}\}\big\}$$

is uncountable; clearly it is also dense in $[0,1]$. Under the appropriate additional assumption on (N_j), however, $O_T(x^*)$ is not 2^j-Benford for any $j \in \mathbb{N}$. If for instance

$$\overline{\lim}_{j\to\infty} \frac{N_j}{N_1 + \ldots + N_j} > \frac{1}{2}\,, \tag{7}$$

then this follows directly from Lemma 2.7(ii). On the other hand, if

$$\lim_{j\to\infty} \frac{|N_2 - N_1| + \ldots + |N_j - N_{j-1}|}{N_1 + \ldots + N_j} = 0\,, \tag{8}$$

then $O_T(x^*)$ cannot be 2^j-Benford either. This will be shown for $j = 1$ here; the case $j > 1$ can be dealt with similarly. (Note that the conditions (7) and (8) are mutually exclusive, postulating a fairly rapid and a rather slow, uniform growth of (N_j), respectively. Either condition allows for uncountably many different choices in (6).)

To show that $O_T(x^*)$ cannot be 2-Benford whenever (8) holds, first define $S : \mathbb{R}^+ \to \mathbb{R}$ as

$$S(y) := -\log_2 T(2^{-y}) = \begin{cases} y - 1 - \log_2(2^y - 1) & \text{if } 0 < y < 1\,, \\ y - 1 & \text{if } y \geq 1\,, \end{cases}$$

so that $S^n(-\log_2 x_0) = -\log_2 T^n(x_0)$ for every $x_0 \in]0,1[\backslash \bigcup_{j\in\mathbb{N}} T^{-j}(\{0\})$ and $n \in \mathbb{N}$, and hence $O_T(x_0)$ is 2-Benford if and only if $O_S(-\log_2 x_0)$ is u.d. mod 1. The map S is ergodic w.r.t. the probability measure with density $2^{-y}(\log 2)\mathbb{1}_{[0,+\infty[}(y)$. Denote by $\widetilde{S}$ the map induced by S on $]0,1[$, that is $\widetilde{S}(y) = [\![S(y)]\!]$ for all $0 < y < 1$. The map $\widetilde{S}$ is ergodic w.r.t. the induced measure on]0,1[, which has $2^{1-y}(\log 2)\mathbb{1}_{[0,1[}(y)$ as its density. Also, let $y^* := -\log_2 x^*$ as well as $\widetilde{y}^* := [\![y^*]\!]$. The orbit modulo one of y^* under S and the orbit of $\widetilde{y}^*$ under $\widetilde{S}$ are related in that

$$[\![O_S(y^*)]\!] = \big(\underbrace{\widetilde{y}^*, \ldots, \widetilde{y}^*}_{N_1 \text{ times}}, \underbrace{\widetilde{S}(\widetilde{y}^*), \ldots, \widetilde{S}(\widetilde{y}^*)}_{N_2 \text{ times}}, \underbrace{\widetilde{S}^2(\widetilde{y}^*), \ldots, \widetilde{S}^2(\widetilde{y}^*)}_{N_3 \text{ times}}, \ldots\big) .$$

Suppose now that $O_T(x^*)$ was 2-Benford. Then $O_S(y^*)$ would be u.d. mod 1, that is $\eta_n \stackrel{w}{\to} \lambda|_{[0,1]}$, where $\eta_n := \frac{1}{n+1}\sum_{j=0}^n \delta_{[\![S^j(y^*)]\!]}$. In particular, therefore,

$$\eta_{N_1+\ldots+N_n-1} = \frac{1}{N_1+\ldots+N_n}\sum\nolimits_{j=1}^n N_j \delta_{\widetilde{S}^{j-1}(\widetilde{y}^*)} \stackrel{w}{\to} \lambda|_{[0,1]} \quad \text{as } n \to \infty .$$

For every continuous function $f : [0,1] \to \mathbb{R}$,

$$\begin{aligned} &\Big|(N_1+\ldots+N_n)\int f \,\mathrm{d}(\widetilde{S}\eta_{N_1+\ldots+N_n-1}) - \\ &\qquad\qquad - (N_1+\ldots+N_{n-1}+2N_n)\int f\,\mathrm{d}\eta_{N_1+\ldots+N_{n-1}+2N_n-1}\Big| \\ &\quad \le \big|N_n f\big(\widetilde{S}^n(\widetilde{y}^*)\big) - N_1 f(\widetilde{y}^*)\big| + \sum\nolimits_{j=1}^{n-1} |N_{j+1}-N_j|\big|f\big(\widetilde{S}^j(\widetilde{y}^*)\big)\big| + \\ &\qquad\qquad + \sum\nolimits_{j=1}^{N_n} |f([\![S^{N_1+\ldots+N_n+j}(y^*)]\!])| \\ &\quad \le \big(N_1 + 2N_n + \sum\nolimits_{j=1}^{n-1}|N_{j+1}-N_j|\big)\|f\|_\infty , \end{aligned}$$

which, together with (8), shows that

$$\lim\nolimits_{n\to\infty}\Big|\int f\,\mathrm{d}(\widetilde{S}\eta_{N_1+\ldots+N_n-1}) - \int f\,\mathrm{d}\eta_{N_1+\ldots+N_{n-1}+2N_n-1}\Big| = 0 ,$$

and hence $\widetilde{S}\eta_{N_1+\ldots+N_n-1} \stackrel{w}{\to} \lambda|_{[0,1]}$ as well. As a consequence, $\lambda|_{[0,1]}$ would be $\widetilde{S}$-invariant. (Note that $\widetilde{S}$ is not continuous but its points of discontinuity form a sequence converging to 0.) This however is impossible because $\widetilde{S}$ is ergodic w.r.t. a measure equivalent to, yet different from $\lambda|_{[0,1]}$. Thus $O_T(x^*)$ cannot possibly be 2-Benford. At the time of writing, the author does not know whether there exists *any* $x \in [0,1]$ at all for which $O_T(x)$ is 2-Benford.

As was the case with (i), the argument above carries over to other maps. The same strategy can for instance be used to demonstrate the existence of uncountably many points $x \in [0,1]$ for which $O_{Q_4}(x)$ is b-Benford for every b not a power of 2. Note, however, that unlike the tent map considered above, Q_4 is not linear near 0, which makes the analysis slightly more involved as a shadowing argument has to be employed before Lemma 2.7 can be applied, see Example 2.12 below for some details in an essentially equivalent context.

Example 2.12 A tacit but important assumption in Theorem 2.10 is that μ be *finite*. To see how that theorem may fail if T preserves an *infinite* measure, consider the C^∞ unimodal map $T_\alpha : [0,1] \to [0,1]$ defined according to

$$T_\alpha(x) := \begin{cases} 1 - e^{\frac{1}{4}\alpha(1-(2x-1)^{-2})} & \text{if } x \neq \frac{1}{2} , \\ 1 & \text{if } x = \frac{1}{2} , \end{cases}$$

where $\alpha > 6$ is a parameter. A characteristic feature of T_α is its flat critical point, i.e. $T_\alpha^{(n)}(\frac{1}{2}) = 0$ for all $n \in \mathbb{N}$. In [28] the ergodic theory of such maps is developed in detail. In particular, it is shown that T_α is conservative and ergodic w.r.t. an infinite but σ-finite measure μ equivalent to $\lambda|_{[0,1]}$. Moreover,

$$\frac{1}{n+1}\sum\nolimits_{j=0}^{n} \delta_{T_\alpha^j(x)} \xrightarrow{w} \delta_0 \tag{9}$$

holds for almost every $x \in [0,1]$. Note that $T'_\alpha(0) = \alpha$. It will now be shown that $O_{T_\alpha}(x)$ is Benford for almost every x, provided that $\log_b \alpha$ is irrational for all bases b. (The latter condition is satisfied for all but countably many α; one may e.g. choose $\alpha = e^2$.) To prepare for the argument, fix an α with the latter property and let $\tau_\alpha : [0,1[\to [0,\frac{1}{2}[$ be the smooth map with $\tau_\alpha(0) = 0$ and $T_\alpha \circ \tau_\alpha(x) \equiv x$. Clearly, $\tau_\alpha(x) = x/\alpha\big(1 + f(x)\big)$ for some smooth f with $f(0) = 0$. It is readily confirmed that

$$h(x) := \lim\nolimits_{n\to\infty} \alpha^n \tau_\alpha^n(x) = x \prod\nolimits_{j=0}^{\infty}\big(1 + f\circ\tau_\alpha^j(x)\big)$$

defines a smooth function $h : [0,1[\to \mathbb{R}$ with $h(0) = 0$, $h'(0) = 1$, and $h(x) > x$ for all $0 < x < 1$. Also, $h \circ \tau_\alpha = \alpha^{-1}h$ and hence

$$h \circ T_\alpha^n(x) = \alpha^n h(x)$$

holds for all x and n, provided that $\{x, T_\alpha(x), \ldots, T_\alpha^n(x)\} \subset [0,\frac{1}{2}[$. Given any base b and $\varepsilon > 0$, pick $N \in \mathbb{N}$ so large that, with $\delta = \tau_\alpha^N(\frac{1}{2})$,

$$\left|\log_b \frac{h(x)}{x}\right| < \varepsilon\,, \quad \forall x \in]0,\delta]\,,$$

and consequently

$$|n\log_b \alpha + \log_b h(x) - \log_b T_\alpha^n(x)| = \left|\log_b \frac{h\circ T_\alpha^n(x)}{T_\alpha^n(x)}\right| < \varepsilon\,, \tag{10}$$

whenever $\{x, \ldots, T_\alpha^n(x)\} \subset]0,\delta[$. For every $x \in [0,1]\setminus\bigcup_{j=0}^{\infty} T_\alpha^{-j}(\{\frac{1}{2}\})$ the point $T_\alpha^n(x)$ lies, for each $n \in \mathbb{N}$, in exactly one of the two intervals $I_1 = [0,\delta]$ and $I_2 = [\delta,1]$. Associate with each such x two integer sequences $\big(N_j(x)\big)$ and $\big(L_j(x)\big)$ such that

$$\begin{aligned} &x, T_\alpha(x), \ldots, T_\alpha^{N_1-1}(x) \in I_1\,,\\ &T_\alpha^{N_1}(x), \ldots, T_\alpha^{N_1+L_1-1}(x) \in I_2\,,\\ &T_\alpha^{N_1+L_1}(x), \ldots, T_\alpha^{N_1+L_1+N_2-1}(x) \in I_1\,, \quad \text{etc.,} \end{aligned} \tag{11}$$

where all numbers N_j, L_j are positive, except perhaps $N_1 = 0$. The sequences $\big(N_j(x)\big)$, $\big(L_j(x)\big)$ are uniquely determined by (11). It follows from the Hopf ratio ergodic theorem that

$$\lim\nolimits_{j\to\infty} \frac{N_1(x) + \ldots + N_j(x)}{j} = +\infty\,, \quad \lim\nolimits_{j\to\infty} \frac{L_1(x) + \ldots + L_j(x)}{N_1(x) + \ldots + N_j(x)} = 0\,, \tag{12}$$

holds for almost every x. Let (y_n) be any sequence with

$$y_n = \big(n - (N_1 + L_1 + \ldots + N_{j-1} + L_{j-1})\big) \log_b \alpha + \log_b h\big(T_\alpha^{N_1+L_1+\ldots+N_{j-1}+L_{j-1}}(x)\big)$$
$$\forall n : 0 \le n - (N_1 + L_1 + \ldots + N_{j-1} + L_{j-1}) < N_j \,.$$

By (12) and Lemma 2.7(i), (y_n) is u.d. mod 1 for almost every x. Moreover, it follows from (10) that $|y_n - \log_b T_\alpha^n(x)| < \varepsilon$ whenever $n \in J_\delta^x := \{j \in \mathbb{N} : T_\alpha^j(x) \in [0, \delta]\}$. According to (9), for almost every x, J_δ^x has density one, that is $\lim_{n\to\infty} n^{-1}\#(J_\delta^x \cap \{1, 2, \ldots, n\}) = 1$. Overall, for some set $A_\varepsilon \subset [0, 1]$ with $\lambda(A_\varepsilon) = 1$, and every $x \in A_\varepsilon$, the sequence $\big(\log_b T_\alpha^n(x)\big)$ differs, on a set of density one, by less than ε from the uniformly distributed sequence (y_n). It follows from [3, Lem.2.3] that for every $x \in A_0 := \bigcap_{j\in\mathbb{N}} A_{1/j}$ the sequence $\big(\log_b T_\alpha^n(x)\big)$ is u.d. mod 1. Clearly, $\lambda(A_0) = 1$. Finally, taking the (countable) intersection over all A_0 corresponding to all $b \in \mathbb{N}\backslash\{1\}$ shows that $O_{T_\alpha}(x)$ is Benford for μ-almost every $x \in [0, 1]$. Thus Theorem 2.10 may fail drastically if μ is not finite.

3. An application of nonautonomous shadowing

The goal of this section is to provide mild conditions on the sequence of maps (T_j) such that (4) generates many Benford sequences. For this, eventual expansivity of (T_j) is a crucial property and will be formalised using the following tailor-made terminology.

Definition 3.1 A sequence (a_n) of real numbers is *eventually positive on average*, abbreviated henceforth as *eposa*, if

$$\underline{\lim}_{m,n\to\infty} \frac{1}{n} \sum\nolimits_{j=1}^{n} a_{m+j} > 0\,.$$

Thus (a_n) is eposa if and only if there exists $\alpha > 0$ and $N_0 \in \mathbb{N}$ such that $n^{-1}\sum_{j=1}^n a_{m+j} > \alpha$ for all $m, n \ge N_0$. Clearly, (a_n) is eposa whenever $\underline{\lim}_{n\to\infty} a_n > 0$, but neither does $a_n > 0$ for all n imply that (a_n) is eposa, nor does $\underline{\lim}_{n\to\infty} a_n = -\infty$ rule this out. For example, (n^{-1}) is not eposa while $\big((-1)^n n + 1\big)$ is. If (a_n) is eposa then $\underline{\lim}_{n\to\infty} n^{-1}\sum_{j=1}^n a_j > 0$. On the other hand, even if $(n^{-1}\sum_{j=1}^n a_j)$ converges to a positive limit, (a_n) may not be eposa.

To identify b-Benford sequences generated by (4), define the maps $S_j : y \mapsto \log_b T_j(b^y)$ so that $\log_b T^n(x) = S^n(\log_b x)$ holds whenever both sides are defined. It is natural, therefore, to first seek conditions ensuring that $O_S(y)$ is u.d. mod 1. Recall that if a map $S : \mathbb{R}^+ \to \mathbb{R}$ is convex then it is a.c., and its right-hand derivative S' exists everywhere and is non-decreasing.

Lemma 3.2. *Assume the maps $S_j : \mathbb{R}^+ \to \mathbb{R}^+$ satisfy, for some $y_0 > 0$ and all $j \in \mathbb{N}$, the following conditions:*

(i) *S_j is convex on $[y_0, +\infty[$;*
(ii) *$S_j'(y_0) > 0$.*

If $\big(\log S_j'(y_0)\big)$ is eposa then there exists $y_1 \ge y_0$ such that $O_S(y)$ is u.d. mod 1 *for (Lebesgue) almost every $y \ge y_1$. However, the set $\big\{y \ge y_1 : O_S(y) \text{is u.d. mod } 1\big\}$ is meagre (i.e., a countable union of nowhere dense sets) and hence its complement is uncountable and dense in $[y_1, +\infty[$.)*

Proof. Let $\beta_j := S_j'(y_0)$ and note first that, as $(\log \beta_j)$ is eposa, there exist $\gamma > 0$

and $N_0 \in \mathbb{N}$ such that $\beta_{m+1} \cdot \ldots \cdot \beta_{m+n} \geq e^{\gamma n}$ holds whenever $m, n \geq N_0$. Since S_j is a.c.,

$$S_j(y) - S_j(y_0) = \int_{y_0}^{y} S_j'(u)\,\mathrm{d}u \geq \beta_j(y - y_0)\,, \quad \forall y \geq y_0, j \in \mathbb{N}\,.$$

In particular, therefore, if $y \geq y_0(1 + \beta_1^{-1})$ then $S_1(y) \geq y_0$, but also

$$S^2(y) = S_2 \circ S_1(y) \geq S_2\big(S_1(y_0) + \beta_1(y - y_0)\big) \geq S_2 \circ S_1(y_0) + \beta_2\beta_1(y - y_0)\,.$$

Let $y_1 := y_0(1 + \beta_1^{-1} + \beta_1^{-1}\beta_2^{-1} + \ldots)$. By induction, if $y \geq y_1$ then $S^n(y) \geq \beta_1 \cdot \ldots \cdot \beta_n(y - y_0)$ and $S^n(y) \geq y_0$ for all $n \in \mathbb{N}$. Define the measurable functions $f_n : [y_1, +\infty[\to \mathbb{R}$ as $f_n(y) = S^n(y)$ and note that $f_n'(y) \geq \beta_1 \cdot \ldots \cdot \beta_n$ for all $y \geq y_1$. Moreover, for $n > m > N_0$ and $n - m \geq N_0$,

$$\begin{aligned} f_n'(y) - f_m'(y) = & \\ &= (S_n' \circ S^{n-1}(y) \cdot \ldots \cdot S_{m+1}' \circ S^m(y) - 1)S_m' \circ S^{m-1}(y) \cdot \ldots \cdot S_2' \circ S_1(y)S_1'(y) \\ &\geq (\beta_n \cdot \ldots \cdot \beta_{m+1} - 1)\beta_m \cdot \ldots \cdot \beta_1 \\ &\geq (e^{\gamma(n-m)} - 1)e^{\gamma m}e^{-\gamma N_0}\beta_{N_0} \cdot \ldots \cdot \beta_1 \\ &= (e^{\gamma(n-m)} - 1)e^{\gamma m}C > 0\,, \end{aligned}$$

with some positive constant C. Since $f_n - f_m$ has an increasing derivative, given any $h \in \mathbb{Z}\backslash\{0\}$ and $s, t \in \mathbb{R}$ with $y_1 \leq s < t$,

$$\left| \int_s^t e^{2\pi i h(f_n(y) - f_m(y))}\mathrm{d}y \right| \leq \frac{e^{-\gamma \min(m,n)}}{|h|(e^{\gamma|n-m|} - 1)C}$$

holds, see [17, Lem.2.1]. Consequently, the exponential sum

$$I_h(N) := \frac{1}{N^2}\sum\nolimits_{m,n=1}^{N} \int_s^t e^{2\pi i h(f_n(y) - f_m(y))}\mathrm{d}y$$

can, for all $N > N_0$, be bounded below and above as

$$\begin{aligned} 0 \leq N^2 I_h(N) &\leq 2\sum\nolimits_{\substack{m,n=1 \\ n \geq m}}^{N} \left| \int_s^t e^{2\pi i h(f_n(y) - f_m(y))}\mathrm{d}y \right| \\ &\leq 4(t - s)NN_0 + \frac{2}{|h|C}\sum\nolimits_{\substack{m,n=N_0+1 \\ n-m \geq N_0}}^{N} \frac{e^{-\gamma m}}{e^{\gamma(n-m)} - 1} \\ &\leq N\left(4(t - s)N_0 + \frac{2}{|h|C(e^{\gamma N_0} - 1)(e^{\gamma} - 1)}\right)\,. \end{aligned}$$

Hence $\sum_{N=1}^{\infty} I_h(N)/N$ converges, and by [17, Thm.4.2] the sequence $\big(f_n(y)\big)$ is u.d. mod 1 for almost every $y \in [s, t]$, that is, $\lambda([s, t]\backslash U) = 0$ where

$$U = \big\{ y \geq y_1 : O_S(y) \text{ is u.d. mod } 1 \big\}\,.$$

Thus $[y_1, +\infty[\backslash U = \bigcup_{j \in \mathbb{N}}\big([y_1 + j - 1, y_1 + j]\backslash U\big)$ has measure zero as well.

It remains to show that U, though of full measure in $[y_1, +\infty[$, is nevertheless meagre. To this end, choose $0 < \delta < \frac{1}{4}$ so small that

$$6\delta < \frac{\gamma}{2\gamma - \log \delta} < \frac{1}{N_0} ,$$

and let $M_\delta := \lceil -\gamma^{-1} \log \delta \rceil + 1$. Given any $y > y_1$ and $0 < \varepsilon < \min(1, y - y_1)$,

$$e^{N_\varepsilon \gamma} \varepsilon > 1$$

holds for $N_\varepsilon := \lceil -\gamma^{-1} \log \varepsilon \rceil + 1$, showing that there exists an interval $I_0 \subset [y - \varepsilon, y + \varepsilon]$ such that $S^{N_\varepsilon}(I_0) = [m_1, m_1 + 1]$ for some $m_1 \in \mathbb{N}$. Within I_0, an interval I_1 can be found with

$$S^{N_\varepsilon}(I_1) = [m_1 + \tfrac{1}{2} - \delta, m_1 + \tfrac{1}{2} + \delta] ,$$

but also

$$S^{N_\varepsilon + M_\delta}(I_1) \supset [m_2, m_2 + 1]$$

for some $m_2 \in \mathbb{N}$. Hence there exists $I_2 \subset I_1$ with

$$S^{N_\varepsilon}(I_2) \subset [m_1 + \tfrac{1}{2} - \delta, m_1 + \tfrac{1}{2} + \delta] \quad \text{and} \quad S^{N_\varepsilon + M_\delta}(I_2) = [m_2 + \tfrac{1}{2} - \delta, m_2 + \tfrac{1}{2} + \delta] .$$

Continuing this process yields a sequence of nested non-empty compact intervals $I_0 \supset I_1 \supset I_2 \supset \ldots$ and therefore the existence of a point $y^* \in [y - \varepsilon, y + \varepsilon]$ with the property that

$$S^{N_\varepsilon + (j-1)M_\delta}(y^*) \in [m_j + \tfrac{1}{2} - \delta, m_j + \tfrac{1}{2} + \delta] , \quad \forall j \in \mathbb{N} ,$$

holds with the appropriate sequence (m_j) of natural numbers. It follows that

$$\underline{\lim}_{n\to\infty} \frac{\#\{j < n : [\![S^j(y^*)]\!] \in [\frac{1}{2} - \delta, \frac{1}{2} + \delta]\}}{n} \geq \frac{1}{M_\delta} > \frac{\gamma}{2\gamma - \log \delta} > 6\delta ,$$

which in turn shows that $O_S(y^*)$ is *not* u.d. mod 1. Let $\varphi : [0,1] \to [0,1]$ be a continuous, piecewise linear function with $\varphi(y) = 0$ whenever $|y - \frac{1}{2}| \geq 2\delta$, and $\varphi(y) \equiv 1$ on $[\frac{1}{2} - \delta, \frac{1}{2} + \delta]$. Since

$$\frac{1}{n} \sum\nolimits_{j=0}^{n-1} \varphi\big([\![S^j(y^*)]\!]\big) - \int_0^1 \varphi(y)\,\mathrm{d}y \geq \frac{\#\{j < n : [\![S^j(y^*)]\!] \in [\frac{1}{2} - \delta, \frac{1}{2} + \delta]\}}{n} - 4\delta > 2\delta$$

for all sufficiently large n, it follows that

$$\underline{\lim}_{n\to\infty} \left| \frac{1}{n} \sum\nolimits_{j=0}^{n-1} \varphi\big([\![S^j(y^*)]\!]\big) - \int_0^1 \varphi(y)\,\mathrm{d}y \right| \geq 2\delta .$$

For every $m \in \mathbb{N}$ define

$$U_m := \left\{ y \geq y_1 : \left| \frac{1}{n} \sum\nolimits_{j=0}^{n-1} \varphi\big([\![S^j(y^*)]\!]\big) - \int_0^1 \varphi(u)\,\mathrm{d}u \right| \leq \delta \ \forall n \geq m \right\} .$$

Note that $U_1 \subset U_2 \subset \ldots$. If $O_S(y)$ is u.d. mod 1 then $y \in U_m$ for all sufficiently large m, hence $U \subset \bigcup_{m\in\mathbb{N}} U_m$. By the continuity of φ and S^j, every set U_m is closed and, by the argument above, has empty interior. Overall, therefore $U \subset \bigcup_{m\in\mathbb{N}} U_m$ where $U_m = \overline{U_m}$ and $\operatorname{int} U_m = \emptyset$. □

Lemma 3.2 directly implies a result on Benford sequences generated by (4) which significantly extends [5, Thm.5.5].

THEOREM 3.3. *Assume the maps* $T_j : \mathbb{R}^+ \to \mathbb{R}^+$ *satisfy, for some* $x_0 > 0$ *and all* $j \in \mathbb{N}$, *the following conditions:*

(i) $x \mapsto \log T_j(e^x)$ *is convex on* $[x_0, +\infty[$;
(ii) $xT_j'(x)/T_j(x) \geq \beta_j > 0$ *for all* $x \geq x_0$.

If $(\log \beta_j)$ *is eposa then there exists* $x_1 \geq x_0$ *such that* $O_T(x)$ *is Benford for almost every* $x \geq x_1$. *However, the set* $\{x \geq x_1 : O_T(x) \text{is } b\text{-Benford for some } b\}$ *is meagre.*

Proof. Fix any base b and let $S_j(y) := \log_b T_j(b^y)$ for every $j \in \mathbb{N}$, so that $S^n(y) = \log_b T^n(b^y)$ for all n. The maps (S_j) satisfy all the assumptions of Lemma 3.2 with $y_0 = x_0/\log b$, and hence $O_S(y)$ is u.d. mod 1 for almost every $y \geq y_1$ for some $y_1 \geq y_0$. Note that y_1 may depend on b. However, an inspection of the proof of Lemma 3.2 shows that one can simply take $y_1 = x_0(1 + \beta_1^{-1} + \beta_1^{-1}\beta_2^{-1} + \ldots)/\log b$. With this, $O_T(x)$ is b-Benford for almost every $x \geq e^{x_0(1+\beta_1^{-1}+\beta_1^{-1}\beta_2^{-1}+\ldots)} =: x_1$. Note that x_1 is independent of b. Since the countable union of null-sets is a null-set, $O_T(x)$ is Benford for almost all $x \geq x_1$. As $y \mapsto b^y$ maps $[y_1, +\infty[$ homeomorphically onto $[x_1, +\infty[$, the set $X_b := \{x \geq x_1 : O_T(x) \text{ is } b\text{-Benford}\}$ is meagre, and so is $\bigcup_b X_b$. □

REMARK 3.4 (i) Unlike [5, Thm.5.5] neither does Theorem 3.3 require $x \mapsto x^{-1}\log T_j(e^x)$ to be non-decreasing, nor does it stipulate that $\inf_j \beta_j > 1$. Note that $(\beta_j - 1)$ is eposa whenever $(\log \beta_j)$ is. In view of [5, Thm.5.5] one might suspect that the assumption that $(\log \beta_j)$ be eposa could be weakened to $(\beta_j - 1)$ being eposa. This, however, is not the case: Theorem 3.3 may fail with this latter assumption, as can be seen from the simple example $T_j(x) = x^{\beta_j}$ with $\beta_j = e^{2(-1)^{j+1}}$ for which $(\beta_j - 1)$ is eposa yet $O_T(x)$ is 2-periodic for every $x > 0$.

(ii) It is readily confirmed that Theorem 3.3 also implies (the reciprocal version of) [5, Thm.4.4].

EXAMPLE 3.5 (i) Let $T_j(x) = x^{p_j}$ where $p_j > 0$ for all $j \in \mathbb{N}$. If $(\log p_j)$ is eposa then $O_T(x)$ is Benford for almost all $x > 0$. In [5, Exp.5.9] the latter was proved only under the stronger assumption $\inf_j p_j > 1$. Thus for example, if p_j is chosen according to

$$p_j = \begin{cases} 1 & \text{if } j \text{ is a prime number,} \\ 2 & \text{otherwise,} \end{cases}$$

then Theorem 3.3 applies whereas [5, Thm.5.5] does not. On the other hand, for

$$p_j = \begin{cases} 2 & \text{if } j \text{ is a prime number,} \\ 1 & \text{otherwise,} \end{cases}$$

$(\log p_j)$ is not eposa, and hence Theorem 3.3 does not apply. That $O_T(x)$ is nevertheless Benford for a.e. x for this system as well follows from [4, Thm.3.1]. The latter result is tailor-made for the power-like setting and contains an assumption on (β_j) that is less restrictive than the requirement that $(\log \beta_j)$ be eposa. What makes Theorem 3.3 interesting in comparison is that unlike [4, Thm.3.1] it does not contain any structural assumptions on (T_j) beyond convexity and expansivity on average.

(ii) For $T_j(x) = x^{2j}+1$, Theorem 3.3 implies that $O_T(x)$ is Benford for almost every $x \in \mathbb{R}$. Note that [5, Thm.5.5] does not apply in this case as $x \mapsto x^{-1}\log T_j(e^x)$ is decreasing.

(iii) If (T_j) is chosen as $T_j(x) = x^2 + (-1)^j$ then Theorem 3.3 does not apply because $x \mapsto \log T_j(e^x)$ is not convex for odd j. In this example one is tempted to expect that $O_T(x)$ is nevertheless Benford for a.e. large x because T_j does not differ too much from $\widetilde{T}_j(x) \equiv x^2$ to which Theorem 3.3 applies. This line of argument will be made rigorous by Theorem 3.7 below.

As Example 3.5(iii) shows, it is mainly the convexity assumption (i) that may restrict the practical applicability of Theorem 3.3. Fortunately, this assumption can be weakened considerably by what is essentially a nonautonomous shadowing argument. Again, the result will first be stated in uniform distribution terms. To this end, assume that (S_j) satisfies assumptions (i) and (ii) of Lemma 3.2, and that in addition

$$\frac{S_j'(y+\rho)}{S_j'(y)} - 1 \le C\rho \tag{13}$$

holds for all $j \in \mathbb{N}$, $y \ge y_0$ and $0 < \rho < \rho_0$, where C, ρ_0 are positive constants. Furthermore, let $\widetilde{S}_j : \mathbb{R}^+ \to \mathbb{R}^+$ be a.c. maps that differ from S_j by little in the sense that, for a.e. $y \ge y_0$

$$\big|\widetilde{S}_j(y) - S_j(y)\big| + \big|\widetilde{S}_j'(y) - S_j'(y)\big| \le \beta_j \Gamma(y)\,, \quad \forall j \in \mathbb{N}\,, \tag{14}$$

with some non-increasing function $\Gamma : \mathbb{R}^+ \to \mathbb{R}^+$. Under mild conditions the uniform distribution of $O_S(y)$ for a.e. sufficiently large y carries over to $O_{\widetilde{S}}(y)$.

Lemma 3.6. *Let (S_j) satisfy assumptions (i) and (ii) of Lemma 3.2, as well as (13). Also, assume that $(\widetilde{S}_j)$ satisfies (14) with $\beta_j := S_j'(y_0)$. If $(\log \beta_j)$ is eposa,*

$$\inf\nolimits_j \beta_j > 0 \quad \textit{and} \quad \sum\nolimits_{j=1}^{\infty} \Gamma(\beta_j \cdot \ldots \cdot \beta_1) < +\infty\,, \tag{15}$$

then there exists $y_1 \ge y_0$ such that $\{y \ge y_1 : O_{\widetilde{S}}(y)$ is u.d. mod $1\}$ has full measure in $[y_1, +\infty[$ yet is meagre.

Proof. Note first that, as $(\log \beta_j)$ is eposa, $\inf_j \beta_j > 0$ implies that

$$1 < C_1 := \sup\nolimits_j \left(1 + \beta_j^{-1} + \beta_{j+1}^{-1}\beta_j^{-1} + \beta_{j+2}^{-1}\beta_{j+1}^{-1}\beta_j^{-1} + \ldots\right) < +\infty\,.$$

Using this and (14) together with the convexity of S_j, it is readily confirmed that

$$\widetilde{S}_{j+m-1} \circ \ldots \circ \widetilde{S}_j(y) \ge \max\big(S_{j+m-1} \circ \ldots \circ S_j(y_0), \beta_{j+m-1} \cdot \ldots \cdot \beta_j y_0, y_0\big)$$

holds for all $j, m \in \mathbb{N}$, provided that $y \ge y_1 := C_1\big(2y_0 + \Gamma(y_0)\big)$. Moreover, the function $g_{j,m}$ with

$$g_{j,m}(y) = S_j^{-1} \circ \ldots \circ S_{j+m-1}^{-1} \circ \widetilde{S}_{j+m-1} \circ \ldots \circ \widetilde{S}_j(y)$$

is well-defined for $y \geq y_1$, and

$$|g_{j,m}(y) - y| \leq \Gamma(y) + \sum\nolimits_{k=0}^{m} \frac{\Gamma \circ \widetilde{S}_{j+k} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+k} \cdot \ldots \cdot \beta_j} < C_1 \Gamma(y_0) \, .$$

For $M > m$, it follows from

$$\begin{aligned} |g_{j,M}(y) - g_{j,m}(y)| &\leq \\ &\leq \frac{\left| S_{j+m}^{-1} \circ \ldots \circ S_{j+M-1}^{-1} \circ \widetilde{S}_{j+M-1} \circ \ldots \circ \widetilde{S}_{j+m} - \mathrm{id} \right| \circ \widetilde{S}_{j+m-1} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+m-1} \cdot \ldots \cdot \beta_j} \\ &\leq \frac{\Gamma \circ \widetilde{S}_{j+m-1} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+m-1} \cdot \ldots \cdot \beta_j} + \sum\nolimits_{k=0}^{M-m} \frac{\Gamma \circ \widetilde{S}_{j+m+k} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+m+k} \cdot \ldots \cdot \beta_j} \\ &\leq \frac{\Gamma(y_0)}{\beta_{j+m-1} \cdot \ldots \cdot \beta_j} \left(1 + \sum\nolimits_{k=0}^{\infty} \beta_{j+m+k}^{-1} \cdot \ldots \cdot \beta_j^{-1} \right) \\ &\leq \frac{C_2}{\beta_{j+m-1} \cdot \ldots \cdot \beta_j} \, , \end{aligned}$$

where C_2 is a positive constant not depending on j, m, M or y, that $\big(g_{j,m}(y)\big)_{m \in \mathbb{N}}$ is, for every j and y, a Cauchy sequence. Hence for each j, the function

$$g_{j,\infty}(y) := \lim\nolimits_{m \to \infty} g_{j,m}(y)$$

is well-defined and continuous on $[y_1, +\infty[$, and

$$|g_{j,\infty}(y) - y| \leq \Gamma(y) + \sum\nolimits_{k=0}^{\infty} \frac{\Gamma \circ \widetilde{S}_{j+k} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+k} \cdot \ldots \cdot \beta_j} \, , \quad \forall j \in \mathbb{N}, y \geq y_1 \, .$$

Note that $S_j \circ g_{j,m} = g_{j+1,m-1} \circ \widetilde{S}_j$ and thus also $S_j \circ g_{j,\infty} = g_{j+1,\infty} \circ \widetilde{S}_j$ for all j. In particular, therefore, $S^n \circ g_{1,\infty} = g_{n+1,\infty} \widetilde{S}^n$ and

$$\left| \widetilde{S}^n(y) - S^n \circ g_{1,\infty}(y) \right| = | \, \mathrm{id} - g_{n+1,\infty} | \circ \widetilde{S}^n(y) \leq \Gamma \circ \widetilde{S}^n(y) + \sum\nolimits_{k=0}^{\infty} \frac{\Gamma \circ \widetilde{S}^{n+1+k}(y)}{\beta_{n+1+k} \cdot \ldots \cdot \beta_{n+1}} \, ,$$

which in turn implies that $\lim_{n \to \infty} \left| \widetilde{S}^n(y) - S^n \circ g_{1,\infty}(y) \right| = 0$ holds whenever $y \geq y_1$. Thus $O_{\widetilde{S}}(y)$ is u.d. mod 1 if and only if $O_S\big(g_{1,\infty}(y)\big)$ is, and the proof will be complete once it has been verified that $g_{1,\infty}$ is a homeomorphism on $[y_1, +\infty[$, and $g_{1,\infty}^{-1}$ is a.c. To show this, explicitly compute $g_{1,m}'$ as

$$g_{1,m}'(y) = \prod\nolimits_{j=1}^{m} \frac{\widetilde{S}_j'}{S_j' \circ g_{j,m+1-j}} \circ \widetilde{S}^{j-1}(y) = \prod\nolimits_{j=1}^{m} \big(1 + a_{j,m}(y) + b_{j,m}(y) \big) \circ \widetilde{S}^{j-1}(y) \, ,$$

where

$$a_{j,m}(y) = \frac{\widetilde{S}_j'(y) - S_j'(y)}{S_j' \circ g_{j,m+1-j}(y)} \quad \text{and} \quad b_{j,m}(y) = \frac{S_j'(y)}{S_j' \circ g_{j,m+1-j}(y)} - 1 \, .$$

It follows from

$$|a_{j,m}(y)| \le \beta_j^{-1} \big|\widetilde{S}'_j(y) - S'_j(y)\big| \le \Gamma(y)$$

and $\widetilde{S}^j(y) \ge \beta_j \cdot \ldots \cdot \beta_1 y_0$ as well as (15) that

$$\sum\nolimits_{j=1}^m \big|a_{j,m}\circ\widetilde{S}^{j-1}(y)\big| \le \Gamma(y) + \sum\nolimits_{j=1}^{m-1}\Gamma(\beta_j\cdot\ldots\cdot\beta_1 y_0) \le \sum\nolimits_{j=0}^{\infty}\Gamma(\beta_j\cdot\ldots\cdot\beta_1 y_0) < +\infty$$

holds for all $m \in \mathbb{N}$ and $y \ge y_1$. Similarly, (13) implies that

$$\begin{aligned} |b_{j,m}(y)| &\le \max\left(\frac{S'_j(y)}{S'_j(y - |g_{j,m+1-j}(y) - y|)}, \frac{S'_j(y + |g_{j,m+1-j}(y) - y|)}{S'_j(y)}\right) - 1 \\ &\le C_2 |g_{j,m+1-j}(y) - y| , \end{aligned}$$

showing that the estimate

$$|b_{j,m}(y)| \le C_2 \left(\Gamma(y) + \sum\nolimits_{k=0}^{\infty} \frac{\Gamma \circ \widetilde{S}_{j+k} \circ \ldots \circ \widetilde{S}_j(y)}{\beta_{j+k}\cdot\ldots\cdot\beta_j}\right)$$

is valid for all $j, m \in \mathbb{N}$, $y \ge y_1$. Consequently,

$$\begin{aligned} \sum\nolimits_{j=1}^m \big|b_{j,m}\circ\widetilde{S}^{j-1}(y)\big| &\le C_2 \sum\nolimits_{j=1}^m \left(\Gamma\circ\widetilde{S}^{j-1}(y) + \sum\nolimits_{k=0}^{\infty}\frac{\Gamma\circ\widetilde{S}^{j+k}(y)}{\beta_{j+k}\cdot\ldots\cdot\beta_j}\right) \\ &\le C_2 \sum\nolimits_{k=0}^{\infty} \Gamma\circ\widetilde{S}^k(y)\left(1 + \sum\nolimits_{j=1}^k \beta_k^{-1}\cdot\ldots\cdot\beta_j^{-1}\right) \\ &\le C_3 \sum\nolimits_{k=0}^{\infty}\Gamma(\beta_k\cdot\ldots\cdot\beta_1 y_0) < +\infty , \end{aligned}$$

with some $C_3 > 0$ independent of m and y. Overall, therefore, there exists a positive constant C_4 such that

$$C_4^{-1} \le g'_{1,m}(y) \le C_4 , \quad \forall m \in \mathbb{N}, y \ge y_1 .$$

Thus $g_{1,\infty}$ is a homeomorphism on $[y_1, +\infty[$, and both $g_{1,\infty}$ and $g_{1,\infty}^{-1}$ are a.c. □

As was the case with Lemma 3.2, it is straightforward to translate Lemma 3.6 into a statement about the generation of nonautonomous Benford orbits.

THEOREM 3.7. *Let the maps $T_j : \mathbb{R}^+ \to \mathbb{R}^+$ satisfy the assumptions* (i) *and* (ii) *of Theorem* 3.3 *and, in addition, for all* $0 < \rho < \rho_0$,

$$\left|\frac{T'_j\big(x(1+\rho)\big)/T_j\big(x(1+\rho)\big)}{T'_j(x)/T_j(x)} - 1\right| \le C\rho , \quad \forall j \in \mathbb{N}, x \ge x_0 , \tag{16}$$

with positive constants C, ρ_0. Assume that the maps $\widetilde{T}_j : \mathbb{R}^+ \to \mathbb{R}^+$ are a.c. on $[x_0, +\infty[$ and that

$$\left|\log\frac{\widetilde{T}_j(x)}{T_j(x)}\right| + x\left|\frac{\widetilde{T}'_j(x)}{\widetilde{T}_j(x)} - \frac{T'_j(x)}{T_j(x)}\right| \le \beta_j\Delta(x) , \quad \forall j \in \mathbb{N}, x \ge x_0 , \tag{17}$$

holds with some non-increasing function $\Delta : \mathbb{R}^+ \to \mathbb{R}^+$. *If* $(\log \beta_j)$ *is eposa,* $\inf_j \beta_j > 0$ *and* $\sum_{j=1}^{\infty} \Delta(e^{\beta_j \cdot \ldots \cdot \beta_1}) < +\infty$ *then there exists* $x_1 \ge x_0$ *such that* $O_{\widetilde{T}}(x)$ *is Benford for a.e.* $x \ge x_1$*, yet the set* $\{x \ge x_1 : O_{\widetilde{T}}(x) \text{ is } b\text{-Benford for some } b\}$ *is meagre.*

Proof. It is immediate to check that, given any base b, the maps $S_j(y) = \log_b T_j(b^y)$ and $\widetilde{S}_j(y) = \log_b \widetilde{T}(b^y)$ satisfy all the assumptions of Lemma 3.6 with $\Gamma(y) = \Delta(b^y)/\log 2$ and $y_0 = x_0/\log b$. Thus there exists $x_{1,b} \ge x_0$ and $N_b \subset [x_{1,b}, +\infty[$ with $\lambda(N_b) = 0$ such that $O_{\widetilde{T}}(x)$ is b-Benford whenever $x \in [x_{1,b}, +\infty[\setminus N_b$. Since $\Delta(x) \to 0$ as $x \to +\infty$, it is possible to choose $x_1 \ge \max(x_0, x_{1,2})$ such that $\widetilde{T}_j'(x) > 0$ for all $j \in \mathbb{N}$ and $x \ge x_1$. Note that x_1 is independent of b. Thus for every j the map $\widetilde{T}_j$ is non-singular on $[x_1, +\infty[$, see e.g. [9, Prop.2.3.2]. From $\lim_{n\to\infty} \widetilde{T}^n(x) = +\infty$ for all $x \ge x_1$, it follows that $[x_1, +\infty[\subset \bigcup_{n=0}^{\infty} \widetilde{T}^{-n}(N_b) \cap [x_1, +\infty[$, and $O_{\widetilde{T}}(x)$ is b-Benford unless $x \in \bigcup_{n=0}^{\infty} \widetilde{T}^{-n}(N_b) \cap [x_1, +\infty[$. By non-singularity, the latter set has measure zero. Consequently, $X_b := \{x \ge x_1 : O_T(x) \text{ is } b\text{-Benford}\}$ has full measure in $[x_1, +\infty[$, and the remaining argument is identical with the one proving Theorem 3.3. □

Example 3.8 The assumptions of Theorem 3.7 may appear somewhat technical. They are, however, naturally satisfied by a wide variety of examples for which Benford behaviour has hitherto been established only in special cases.

(i) As observed in Example 3.5(iii), Theorem 3.3 does not apply to $T_j(x) = x^2 + (-1)^j$. However, letting $T_j(x) \equiv x^2$ and $\widetilde{T}_j(x) = x^2 + (-1)^j$, Theorem 3.7 applies with $x_0 = 2$, $C = 1$, $\rho_0 = 1$, and $\Delta(x) = 2/x$, showing that $O_{\widetilde{T}}(x)$ is Benford for a.e. sufficiently large x.

To see that this approach works in much greater generality, let each map $\widetilde{T}_j$ be a polynomial of degree $m_j \in \mathbb{N}$, i.e.

$$\widetilde{T}_j(x) = a_{j,m_j} x^{m_j} + a_{j,m_j-1} x^{m_j-1} + \ldots + a_{j,1} x + a_{j,0}\,, \quad j \in \mathbb{N}\,, \tag{18}$$

with $a_{j,k} \in \mathbb{R}$ for all j and $0 \le k \le m_j$, and $a_{j,m_j} \ne 0$. To avoid obvious pathologies, it is natural to assume that

$$\sup_j \left(|a_{j,m_j}|^{-1} + \max_{k=0}^{m_j} |a_{j,k}| \right) < +\infty\,. \tag{19}$$

Also, let $T_j(x) = a_{j,m_j} x^{m_j}$ and assume w.l.o.g. that $a_{j,m_j} > 0$. Clearly, (T_j) satisfies all assumptions of Theorem 3.3 with $\beta_j = m_j$ as well as (16) with $C = 1$ and $\rho_0 = 1$. Moreover, (17) holds with $\Delta(x) = D/x$ for any sufficiently large constant D. Note that $\inf_j \beta_j \ge 1 > 0$, and the condition $\sum_j \Delta(e^{\beta_j \cdot \ldots \cdot \beta_1}) < +\infty$ is trivially met, provided that $(\log m_j)$ is eposa. Whenever the latter is the case, therefore, for any sequence of polynomials (18) satisfying (19) the nonautonomous orbit $O_{\widetilde{T}}(x)$ is Benford for a.e. sufficiently large x. This fact has been proved in [4] merely under the additional assumption that $m_j \ge 2$ for all j, and with a weaker conclusion concerning the exceptional points.

(ii) Assume that each $\widetilde{T}_j$ is of the form $\widetilde{T}_j(x) = e^{x^2} P_j(x)$ where P_j is a polynomial of degree $m_j \in \mathbb{N}_0$ satisfying (19). Under the additional assumption that $\sup_j m_j < +\infty$ this system has been analysed in [4] by means of an ad-hoc shadowing argument. The results of the present article enable a more systematic treatment without this additional assumption. On the one hand, it is easily confirmed that with $T_j(x) = e^{x^2} a_{j,m_j} x^{m_j}$ all the conditions of Theorem 3.7 are met. Still easier, however, is noting that Theorem 3.3 applies directly. Indeed, with $T_j(x) = e^{x^2} P_j(x)$ condition (i) of Theorem 3.3 is satisfied for every sufficiently large x_0, independent of j, and so is (ii) with $\beta_j = 2 + m_j$. Hence $(\log \beta_j)$ is eposa, and $O_T(x)$ is Benford for a.e. sufficiently large x. Note that this argument does require neither $m_j \geq 1$ nor $\sup_j m_j < +\infty$.

Remark 3.9 (i) If (S_j) and (T_j) do not satisfy (13) and (16), respectively, then Lemma 3.6 and Theorem 3.7 do not apply. (For example, (13) does not hold for $S_j(y) = e^{e^{jy}}$.) In concrete cases, however, the underlying shadowing idea may well be salvaged as it merely requires that $\sum_{j=1}^m \left| b_{j,m} \circ \widetilde{S}^{j-1}(y) \right|$ remain bounded uniformly in m and y, which can be guaranteed by imposing more restrictive conditions on the decay of Γ.

(ii) Theorem 3.7 specifies conditions under which the nonautonomous orbit $O_{\widetilde{T}}(x)$ is Benford for a.e. sufficiently large x, that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{[1,t]}\left(\langle \widetilde{T}^j(x)\rangle_b\right) = \log_b t\,, \quad \forall b \in \mathbb{N}\backslash\{1\}, 1 < t < b\,. \tag{20}$$

It is natural ask for a finer analysis of the convergence in (20). Analogously to well-known results on almost sure behaviour of ergodic sums [21, Thm.3.2.3], this convergence can be arbitrarily slow. It seems plausible, however, that in the setting of Theorem 3.7 a better quantitative description of (20) can be achieved by means of a (nonautonomous) *almost sure invariance principle* for $\left(\langle \widetilde{T}^j \rangle_b\right)$, cf. [22]. No pertinent results in this regard are known to the author.

References

[1] F. Benford, *The law of anomalous numbers*, Proceedings of the American Philosophical Society **78** (1938), 551–572.
[2] A. Berger, *Chaos and Chance*, de Gruyter, Berlin New York, 2001.
[3] A. Berger, *Multi-dimensional Dynamical Systems and Benford's Law*, Discrete Contin. Dyn. Syst. **13** (2005), 219–237.
[4] A. Berger, *Benford's Law in power-like dynamical systems*, Stoch. Dyn. **5** (2006), 587–607.
[5] A. Berger, L.A. Bunimovich, and T.P. Hill, *One-dimensional dynamical systems and Benford's law*, Trans. Amer. Math. Soc. **357** (2005), 197–219.
[6] A. Berger and T.P. Hill, *Newton's Method obeys Benford's Law*, Amer. Math. Monthly **114** (2007), 588–601.
[7] A. Berger and T.P. Hill, *Benford Online Bibliography* at `www.benfordonline.net`.
[8] A. Berger and S. Siegmund, *On the distribution of mantissae in nonautonomous difference equations*, J. Difference Equ. Appl. **13** (2007), 829–845.
[9] A. Boyarsky and P. Góra, *Laws of Chaos. Invariant Measures and Dynamical Systems in One Dimension*, Birkhäuser, 1997.
[10] J. Brown and R. Duncan, *Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences*, Fibonacci Quarterly **8** (1970), 482–486.
[11] P. Diaconis, *The distribution of leading digits and uniform distribution mod 1*, Ann. Probab. **5** (1979), 72–81.
[12] G. Gottwald and M. Nicol, *On the nature of Benford's Law*, Phys. A **303** (2001), 387–396.
[13] T.P. Hill, *Base-Invariance Implies Benford's Law*, Proc. Amer. Math. Soc. **123** (1995), 887-895.
[14] T.P. Hill, *A statistical derivation of the significant-digit law*, Statistical Science **10** (1996), 354–363.
[15] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.
[16] A. Kontorovich and S. Miller, *Benford's Law, values of L-functions and the 3x+1 problem*, Acta Arithmetica **120** (2005), 269-297.
[17] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.
[18] E. Ley, *On the Peculiar Distribution of the U.S. Stock Indices Digits*, Amer. Stat. **50** (1996), 311–313.
[19] R. Mané, *Ergodic Theory and Differentiable Dynamics*, Springer, Berlin Heidelberg New York, 1987.
[20] S. Newcomb, *Note on the frequency of use of the different digits in natural numbers*, Amer. J. Math. **4** (1881), 39-40.
[21] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1983.
[22] W. Philipp and W.F. Stout, *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables*, Memoirs of the Amer. Math. Soc. 161, 1975.
[23] R. Raimi, *The first digit problem*, Amer. Math. Monthly **102** (1976), 322–327.
[24] P. Schatte, *On H_∞-summability and the uniform distribution of sequences*, Math. Nachr. **113** (1983), 237-243.
[25] P. Schatte, *On mantissa distributions in computing and Benford's law*, J. Information Processing and Cybernetics **24** (1998), 443–455.

[26] M. Snyder, J. Curry, and A. Dougherty, *Stochastic aspects of one-dimensional discrete dynamical systems: Benford's law*, Physical Review E **64** (2001), 1–5.

[27] C. Tolle, J. Budzien, and R. LaViolette, *Do dynamical systems follow Benford's law?*, Chaos **10** (2000), 331–337.

[28] R. Zweimüller, *S-unimodal Misiurewicz maps with flat critical points*, Fund. Math. **181** (2004), 1–25.