Benford's Law in power-like dynamical systems

ARNO BERGER*

Fachbereich Mathematik
Johann Wolfgang Goethe-Universität
D-60325 Frankfurt a. Main, GERMANY
aberger@math.uni-frankfurt.de

Received 11 February 2005

A generalized shadowing lemma is used to study the generation of Benford sequences under non-autonomous iteration of power-like maps $T_j : x \mapsto \alpha_j x^{\beta_j}(1 - f_j(x))$, with $\alpha_j, \beta_j > 0$ and $f_j \in C^1$, $f_j(0) = 0$, near the fixed point at $x = 0$. Under mild regularity conditions almost all orbits close to the fixed point asymptotically exhibit Benford's logarithmic mantissa distribution with respect to all bases, provided that the family (T_j) is contracting *on average*, i.e. $\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}\log\beta_j > 0$. The technique presented here also applies if the maps are chosen at random, in which case the contraction condition reads $\mathbb{E}\log\beta > 0$. These results complement, unify and widely extend previous work. Also, they supplement recent empirical observations in experiments with and simulations of deterministic as well as stochastic dynamical systems.

Keywords: Dynamical system, shadowing, Benford's law, uniform distribution mod 1.

AMS 2000 subject classification: 37A50, 60A10, 37B55

1. Introduction

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

$$\mathbb{P}(\text{mantissa}_b \le t) = \log_b t\,, \quad \forall t \in [1, b[\,; \tag{1.1}$$

the most well-known special case is that with respect to base $b = 10$

$$\mathbb{P}(\text{first significant digit} = d) = \log_{10}\left(1 + d^{-1}\right), \quad \forall d = 1, \ldots, 9\,.$$

Examples of empirical data sets following (1.1) have been discussed extensively, for instance in real-life data (e.g., physical constants, stock market indices, tax returns [11,15,17,19,21]), in stochastic processes (e.g., sums and products of random variables [10,19]), and in deterministic sequences (e.g., $(n!)$ and Fibonacci numbers [2,6,9]). It was only recently that a thorough mathematical analysis of BL for dynamical systems has been initiated [4,5,10,20,22]. Following physical experiments

*Permanent address: Department of Mathematics and Statistics, University of Canterbury, Christchurch, NEW ZEALAND (`arno.berger@canterbury.ac.nz`)

and numerical simulations, it has been shown that the asymptotic mantissa distribution of orbits of (autonomous and non-autonomous, one- and higher-dimensional) dynamical systems equals Benford's distribution (1.1) surprisingly often.

Dynamical systems, both deterministic and stochastic, are widely used as models for real-world phenomena. If the latter exhibit, on an empirical level, a striking statistical property like (1.1) – as often they do – then it is natural to ask for a rigorous manifestation of this property in the underlying mathematical model. For example, affine processes

$$x_{j+1} = a_{j+1}x_j + b_{j+1}\,, \tag{1.2}$$

are a standard tool in econometrics and financial mathematics (and in many other disciplines), and there is abundant evidence of BL in econometric and financial data [12,15,17]. A rigorous mathematical analysis of (1.2) under the perspective of BL was carried out in [4,5,10]. These results show that under mild, reasonable assumptions the logarithmic mantissa distribution is in fact inherent to the mathematical model. Moreover, as emphasized in [10], the coefficients in the recursion (1.2) do not have to be stochastic at all for (1.1) to emerge. Rather, BL turns out to be a phenomenon ubiquitous both in deterministic and stochastic systems.

The astonishing ubiquity of BL will become apparent also through the present article. As mentioned earlier, a fairly comprehensive analysis with respect to BL of *linearly dominated* dynamics has been given recently, showing that the emergence of (1.1) is indeed typical in this case regardless of the initial data. For dynamics brought about by truly *non-linear* maps, however, the situation may be more complicated, as detailed in [4,5] for rather special classes of systems. It is natural to ask for a similar analysis for more general, non-uniform systems. For example, maps of the type

$$T_j(x) = \alpha_j x^{\beta_j}(1-x)^{\beta_j}\,, \tag{1.3}$$

have been studied extensively as models describing aspects of economic growth and socio-spatial dynamics [8,16]. The main purpose of the present work is to provide a rigorous analysis of systems like (1.3) with respect to BL. More generally, Benford properties of orbits under iteration of maps $T_j : \mathbb{R}^+ \to \mathbb{R}^+$ will be studied which are *power-like* near 0, i.e., T_j can be written locally near 0 as

$$T_j(x) = \alpha_j x^{\beta_j}\big(1 - f_j(x)\big)\,, \quad x > 0\,,$$

with $\alpha_j, \beta_j > 0$ and $f_j \in C^1$ with $f_j(0) = 0$. As for (1.3), the results in [5] do for instance not allow to conclude that for

$$T_j(x) = x^2 + (-1)^j x^4 + x^6\,, \quad j \in \mathbb{N}\,, \tag{1.4}$$

the non-autonomous orbit $(x_n)_{n\in\mathbb{N}_0}$ generated by

$$x_{j+1} = T_{j+1}(x_j)\,, \quad j \in \mathbb{N}_0\,,$$

is a Benford sequence (i.e., the asymptotic mantissa distribution of this sequence is given by (1.1) for *every* base b) for Lebesgue almost all x_0 sufficiently close to either the origin or infinity. (A set of numerical data obeys (1.1) if and only if the set of reciprocals follows (1.1) as well; thus, with respect to BL, 0 and infinity are equivalent and will be considered interchangeably.) Notice that both T_{2j} and T_{2j+1} are in a sense close to $Q : x \mapsto x^2 + x^6$ near 0 and infinity. For the map Q the result is known [5], and one may expect that an improved shadowing technique will work also for (1.4); see Theorem 3.2 below.

By its very nature, [5, Thm.5.5] can only deal with maps which are sufficiently contracting near 0 (or expanding near infinity). If for instance

$$T_j(x) = \begin{cases} \sqrt{x}\,(1+x^2) & \text{if } j \in J\,, \\ x^3(1+x^3) & \text{if } j \notin J\,, \end{cases} \tag{1.5}$$

with some set $J \subset \mathbb{N}$, then clearly T_j is expanding rather than contracting near the fixed point at 0 whenever $j \in J$, and there is no way of applying the aforementioned result, even though T_j has good Benford properties for $j \notin J$. However, assuming that J has a density, that is,

$$\rho(J) = \lim\nolimits_{N\to\infty} \frac{\#(J \cap \{1, \ldots, N\})}{N} \tag{1.6}$$

exists, it is plausible that contraction near 0 (and thus also the generation of Benford sequences) will prevail on average, provided that $(1/2)^{\rho(J)} 3^{1-\rho(J)} > 1$ or, equivalently,

$$\rho(J) \log \frac{1}{2} + \big(1 - \rho(J)\big) \log 3 > 0\,, \tag{1.7}$$

which clearly imposes an *upper* bound on $\rho(J)$. It turns out that condition (1.7), together with a mild assumption on the rate of convergence in (1.6), does indeed guarantee the generation of Benford sequences through (1.5) from almost all initial points near 0; see Corollary 3.2 below.

The key tool developed in the next section is a shadowing lemma which significantly improves the corresponding result in [5]. Combining this tool and standard techniques from the theory of uniform distribution modulo 1 enables a straightforward analysis not only of systems like (1.4) and (1.5) but also of randomized versions thereof. For example, if

$$T_{\xi_j}(x) = x^{\xi_j} e^{-x^2}\,, \quad x > 0\,; j \in \mathbb{N}\,,$$

where (ξ_j) is an i.i.d. sequence of positive random variables such that $\log \xi_j$ has finite variance, then for almost all x close to 0, the stochastic orbit $\big(T_{\xi_n} \circ \ldots \circ T_{\xi_1}(x)\big)_{n\in\mathbb{N}}$ is a Benford sequence with probability one, provided that

$$\mathbb{E} \log \xi_j > 0\,,$$

which is the natural analogue of (1.7) in the stochastic setting (see Theorem 3.3). As will become clear shortly, it is the versatility of the shadowing technique which allows to treat all these apparently different problems in a unifying manner.

2. Preliminaries and basic shadowing result

The sets of natural, integer and real numbers are denoted by $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$, respectively; $\mathbb{N}_0$ symbolizes the non-negative integers, and $\mathbb{R}^+$ stands for the positive reals. For every real x, the number $\lfloor x \rfloor$ denotes the largest integer not larger than x. The symbol λ is used for Lebesgue measure on $\mathbb{R}$ (or subsets thereof).

Only the most basic definitions concerning BL in dynamical systems are restated here for the reader's convenience; for details on notation and terminology see [5]. Throughout, b denotes a natural number larger than one (referred to as a *base*). Every positive real x can be written uniquely as $x = M_b(x)b^k$ with $M_b(x) \in [1, b[$ and the appropriate integer k. The function $M_b : \mathbb{R}^+ \to [1, b[$ is called the (*base* b) *mantissa function*; for convenience set $M_b(0) = 0$ for all b. The integer $\lfloor M_b(x) \rfloor \in \{1, \ldots, b-1\}$ is referred to as the *first significant* b*-digit* of x. For a given base b, $\log_b$ stands for the logarithm with respect to b; if used without a subscript, the log symbol denotes the natural logarithm. Furthermore, for $n \in \mathbb{N}$, $\log_{(n)} x$ is meant to signify the n-fold iterated (natural) logarithm of $x > 0$, i.e.

$$\log_{(1)} x = \log x \,, \quad \log_{(n+1)} x = \log\bigl(\log_{(n)} x\bigr) \,,$$

whenever this quantity is defined.

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

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

it is a *(strict) Benford sequence* if it is a b-Benford sequence for every $b \in \mathbb{N} \setminus \{1\}$.

It is well known [5,9] that the sequence (x_n) is b-Benford if and only if $(\log_b |x_n|)$ is uniformly distributed mod 1 (henceforth abbreviated as *u.d.* mod 1). A few tools from uniform distribution theory will be used in the sequel; for an authoritative reference to this field the reader is referred to [14].

The present article studies the Benford property of sequences generated recursively by

$$x_{j+1} = T_{j+1}(x_j) \,, \quad j \in \mathbb{N}_0 \,, \tag{2.1}$$

where the maps (T_j) are assumed to be *power-like* near 0, i.e., for some $\delta > 0$,

$$T_j(x) = \alpha_j x^{\beta_j}\bigl(1 - f_j(x)\bigr) \,, \quad \forall x : 0 \le x < \delta \,, \tag{2.2}$$

with $\alpha_j, \beta_j > 0$, and $f_j \in C^1[0, \delta]$ with $f_j(0) = 0$ for all $j \in \mathbb{N}$. Relation (2.1) may be interpreted as a *non-autonomous dynamical system*. For $n \in \mathbb{N}$, T^n denotes the composition $T^n = T_n \circ T_{n-1} \circ \ldots \circ T_2 \circ T_1$, and $T^0 = id$. The symbol $O_T(x)$, called the *orbit* of x under (T_j), will denote the sequence generated by (2.1) subject to the initial condition $x_0 = x$; with the above notation, $O_T(x) = \bigl(T^n(x)\bigr)_{n \in \mathbb{N}_0}$. Note that this interpretation of the orbit as a *sequence* differs from the standard terminology in dynamical systems theory (e.g., [13]) where the orbit of x is the mere *set* $\{x_n : n \in \mathbb{N}_0\}$.

On a logarithmic scale, i.e. by setting $x_j = b^{-y_j}$, recursion (2.1) for the power-like maps (2.2) becomes

$$y_{j+1} = -\log_b T_{j+1}(b^{-y_j}) = -\log_b \alpha_{j+1} + \beta_{j+1} y_j - \log_b\bigl(1 - f_{j+1}(b^{-y_j})\bigr) .$$

Thus $y_{j+1} = S_{j+1}(y_j)$ with (S_j) denoting a family of perturbed affine maps of the form

$$S_j(y) = \gamma_j + \beta_j y + g_j(y) , \quad j \in \mathbb{N} , \tag{2.3}$$

where γ_j and $\beta_j \neq 0$ are real coefficients, and g_j is a continuous function which satisfies

$$|g_j(y)| \leq \Gamma(|y|) , \quad \forall |y| \geq \xi \, ; \, j \in \mathbb{N} ,$$

with a non-increasing function Γ. To study shadowing properties of the maps (2.3) let $B_n = \prod_{j=1}^n \beta_j$ and $B_0 = 1$. Also, for each $j \in \mathbb{N}$ define the affine map $\widetilde{S}_j : y \mapsto \gamma_j + \beta_j y$ which may be thought of as a simplified version of S_j. The inverse of $\widetilde{S}_j$ is $\widetilde{S}_j^{-1}(y) = (y - \gamma_j)/\beta_j$; consequently, $\widetilde{S}^{-n} = (\widetilde{S}^n)^{-1} = \widetilde{S}_1^{-1} \circ \ldots \circ \widetilde{S}_n^{-1}$. Up to reindexing, the key shadowing result in [5] is a special case of

Theorem 2.1. *Suppose the maps $S_j : \mathbb{R} \to \mathbb{R}$ are continuous, and*

$$|S_j(y) - (\gamma_j + \beta_j y)| \leq \Gamma(|y|) , \quad \forall |y| \geq \xi \, ; \, j \in \mathbb{N} ,$$

where (γ_j), (β_j), ξ are real numbers with $\beta_j \neq 0$, $\xi > 0$, and $\Gamma : [0, +\infty[\to \mathbb{R}^+$ is a non-increasing function. Let $\widetilde{S}_j(y) = \gamma_j + \beta_j y$. Then, with an appropriate $\eta \geq \xi$:

(i) If the series $\sum_{j=0}^\infty |B_{j+1}|^{-1} \Gamma(t|B_j|)$ and $\sum_{j=1}^\infty |B_j|^{-1} |\gamma_j|$ both converge for some $t > 0$, then $h(y) = \lim_{n\to\infty} \widetilde{S}^{-n} \circ S^n(y)$ exists for all $|y| \geq \eta$. Moreover, h is a continuous function, and $\sup_{|y|\geq\eta} |h(y) - y| < \infty$.

(ii) If, in addition,

$$\limsup_{n\to\infty} |B_n| \sum\nolimits_{j=n}^\infty |B_{j+1}|^{-1} \Gamma(t|B_j|) < \infty , \tag{2.4}$$

then for each y with $|y| \geq \eta$ there exists precisely one point $\overline{y}$ such that the sequence $\bigl(|\widetilde{S}^n(\overline{y}) - S^n(y)|\bigr)_{n\in\mathbb{N}_0}$ is bounded. In fact, $\overline{y} = h(y)$ with the function h from (i).

Proof. From (2.3) it follows that for all $n \in \mathbb{N}_0$,

$$\widetilde{S}^{-n} \circ S^n(y) = y + \sum\nolimits_{j=0}^{n-1} B_{j+1}^{-1} \bigl(S^{j+1}(y) - \gamma_{j+1} - \beta_{j+1} S^j(y)\bigr) . \tag{2.5}$$

The assumption $\sum_{j=0}^\infty |B_{j+1}|^{-1}\Gamma(t|B_j|) < \infty$ implies that $|B_n| \to \infty$ as $n \to \infty$. Choose $\eta_1 \geq t$ sufficiently large to ensure that

$$\sum\nolimits_{j=0}^\infty |B_{j+1}|^{-1} \Gamma(\eta_1 |B_j|) < \eta_1 ,$$

and also $|B_n|\eta_1 \geq \xi$ for all $n \in \mathbb{N}_0$. Setting $\eta = 2\sum_{j=1}^{\infty}|B_j|^{-1}|\gamma_j| + 4\eta_1$, it is easily checked that for all $n \in \mathbb{N}_0$,

$$|S^n(y)| \geq \tfrac{1}{2}|B_n||y| + \eta_1|B_n| \qquad \text{and} \qquad \left|\frac{S^n(y)}{B_n} - y\right| < \tfrac{1}{2}\eta\,,$$

provided that $|y| \geq \eta$. Thus

$$g(y) := y + \sum_{j=0}^{\infty} B_{j+1}^{-1}\big(S^{j+1}(y) - \gamma_{j+1} - \beta_{j+1}S^j(y)\big) \tag{2.6}$$

defines a continuous function for $|y| \geq \eta$, for which (2.5) and (2.6) yield

$$\big|\widetilde{S}^{-n} \circ S^n(y) - g(y)\big| \leq \sum_{j=n}^{\infty}|B_{j+1}|^{-1}\Gamma\big(|S^j(y)|\big) \leq \sum_{j=n}^{\infty}|B_{j+1}|^{-1}\Gamma(\eta_1|B_j|) \tag{2.7}$$

for all $n \in \mathbb{N}_0$. The right-hand side of (2.7) tends to 0 as $n \to \infty$. Therefore

$$g(y) = \lim_{n\to\infty}\widetilde{S}^{-n} \circ S^n(y) = h(y)\,, \tag{2.8}$$

which concludes the proof of (i), because

$$|g(y) - y| \leq \sum_{j=0}^{\infty}|B_{j+1}|^{-1}\Gamma(\eta_1|B_j|) \;<\; \eta_1\,, \quad \forall |y| \geq \eta\,.$$

Assertion (ii) follows immediately from (2.4) and the identity

$$\widetilde{S}^n \circ g(y) - S^n(y) = B_n\sum_{j=n}^{\infty} B_{j+1}^{-1}\big(S^{j+1}(y) - \gamma_{j+1} - \beta_{j+1}S^j(y)\big)\,.$$

Since $|B_n| \to \infty$ as $n \to \infty$, the point $\overline{y} = g(y) = h(y)$ is the only point with the desired property. □

In the setting of Theorem 2.1 (ii), for each $|y| \geq \eta$ there exists precisely one point, namely $h(y)$, whose non-autonomous orbit $O_{\widetilde{S}}\big(h(y)\big)$ resembles ("shadows") $O_S(y)$. This explains why this result is referred to as a (non-autonomous) *shadowing lemma*, a terminology particularly appropriate if the lim sup in (2.4) actually vanishes. (See for instance [13,18] for a detailed account on shadowing in dynamical systems.) Clearly, one expects $O_{\widetilde{S}}(y)$ to be more easily analyzed than $O_S(y)$.

3. Benford's law: Results and examples

As indicated earlier, Theorem 2.1 is a key tool for revealing the Benford properties of orbits under the non-autonomous power-like maps

$$T_j : x \mapsto \alpha_j x^{\beta_j}\big(1 - f_j(x)\big)\,, \quad j \in \mathbb{N}\,, \tag{3.1}$$

near the fixed point at the origin. Without substantial loss of generality assume $\alpha_j, \beta_j > 0$. Also, recall the abbreviation $B_n = \prod_{j=1}^n \beta_j$.

Theorem 3.1. *Let (T_j) be given by (3.1) with C^1 functions (f_j) which satisfy $f_j(0) = 0$ and $|f_j'(x)| \leq C$ for all $0 < x \leq \delta$ and $j \in \mathbb{N}$ with constants $C > 0$, $\delta > 0$. Assume that*

$$\sum_{j=1}^{\infty} B_j^{-1}\,|\log\alpha_j| < \infty\,, \tag{3.2}$$

and also that for some $\kappa > 0$,

$$\lim_{n\to\infty} B_n \sum_{j=n}^{\infty} B_{j+1}^{-1}\kappa^{B_j} = 0\,. \tag{3.3}$$

Then $O_T(x)$ *is a* b*-Benford sequence for almost all* $x > 0$ *sufficiently close to* 0 *if and only if the same is true for* $O_{\widetilde{T}}(x)$ *where* $\widetilde{T}_j : x \mapsto \alpha_j x^{\beta_j}$.

Proof. Fix a base $b \in \mathbb{N}\backslash\{1\}$ and consider the maps

$$S_{j,b}(y) = -\log_b T_j(b^{-y}) = -\log_b \alpha_j + \beta_j y + g_{j,b}(y)\,,$$

with the functions $g_{j,b}$ given by

$$g_{j,b}(y) = -\log_b\bigl(1 - f_j(b^{-y})\bigr)\,.$$

By the assumptions on (f_j), there exist constants $\eta_1 > 0$, $D > 0$ independent of b and j such that the function $g_{j,b}$ is C^1 on $[\eta_1/\log b, +\infty[$, and

$$|g_{j,b}(y)| \le D b^{-y}/\log b =: \Gamma_b(y)\,.$$

From (3.3) it follows that

$$\lim_{n\to\infty} B_n \sum_{j=n}^{\infty} B_{j+1}^{-1}\Gamma_b(tB_j) = 0\,,$$

whenever $t > \max(-\log_b \kappa, 0)$. By Theorem 2.1 there exists a continuous map h_b with $\sup_{y\ge\eta_2/\log b}|h_b(y) - y| < \infty$ for some $\eta_2 \ge \eta_1$ not depending on b, and also

$$\lim_{n\to\infty}\bigl|\widetilde{S}_b^n\bigl(h_b(y)\bigr) - S_b^n(y)\bigr| = 0\,, \quad \forall y \ge \eta_2/\log b\,; \tag{3.4}$$

here $\widetilde{S}_b^n = \widetilde{S}_{n,b} \circ \ldots \circ \widetilde{S}_{1,b}$ with $\widetilde{S}_{j,b}(y) = -\log_b \alpha_j + \beta_j y$ depends on b, as is indicated by a subscript. Thus, for $y \ge \eta_2/\log b$ the sequence $O_{S_b}(y)$ is u.d. mod 1 if and only if $O_{\widetilde{S}_b}\bigl(h_b(y)\bigr)$ is u.d. mod 1. Furthermore, using the chain rule and termwise differentiation of the formula

$$h_b(y) = y + \sum_{j=0}^{\infty} B_{j+1}^{-1} g_{j+1,b} \circ S_b^j(y)\,,$$

it can be checked that for $y \ge \eta_3/\log b$, with some $\eta_3 \ge \eta_2$ which does not depend on b, the function h_b is C^1 with positive derivative and thus maps sets of measure zero onto sets of measure zero; also, η_3 can be chosen such that $|h_b(y)-y|+|h_b^{-1}(y)-y| \le \eta_3/\log b$ for all $y \ge \eta_3/\log b$. Since $\widetilde{S}_b^n(y) = -\log_b \widetilde{T}^n(b^{-y})$, the orbit $O_T(x)$ is b-Benford for almost all $x < e^{-2\eta_3}$ if and only if the same is true for $O_{\widetilde{T}}(x)$. $\square$

Corollary 3.1. *Under the hypotheses of Theorem 3.1 the non-autonomous orbit* $O_T(x)$ *is a (strict) Benford sequence for almost all* x *close to* 0 *if and only if the same is true for* $O_{\widetilde{T}}(x)$.

Proof. For each b there exist sets N_b, $\widetilde{N}_b$ of measure zero such that $O_T(x)$ and $O_{\widetilde{T}}(\widetilde{x})$ are b-Benford whenever $x \in [0,\delta]\backslash N_b$ and $\widetilde{x} \in [0,\delta]\backslash\widetilde{N}_b$, respectively; according to the proof of Theorem 3.1, $\delta = e^{-2\eta_3}$ does not depend on b. Since $\bigcup_b N_b$ and $\bigcup_b \widetilde{N}_b$ both have measure zero, the proof is complete. $\square$

Remark 3.1. (i) The results in [5], when applied to power-like maps, are much weaker than Theorem 3.1. This can be seen for example from

$$\alpha_j \equiv 1\,, \quad \beta_j = \begin{cases} 1 & \text{if } j = 1\,, \\ 1 + j^{-1}(\log j)^{-t} & \text{if } j \geq 2\,, \end{cases}$$

with $t \in \mathbb{R}$, for which the hypotheses of [5,Thm.5.5] are not satisfied at all since $\lim_{j\to\infty} \beta_j = 1$. However, Theorem 3.1 applies for $t \leq 1$. If $t < 1$ then $O_T(x)$ is a Benford sequence for almost all x (or even *all* x if $0 < t < 1$) close to 0. For $t = 1$, no orbit is b-Benford for any b. Finally, (3.3) fails if $t > 1$. This corresponds nicely to the fact that, at least in the case $f_j \equiv 0$, there are no b-Benford sequences whatsoever, because for each $x > 0$ the iterates $T^n(x)$ tend to a non-zero limit. Note also that, even if it were applicable, [5,Thm.5.5] would impose additional, restrictive conditions on the functions (f_j).

(ii) Theorem 3.1 and Corollary 3.1 assert that under certain conditions, the Benford properties of orbits under (3.1) are determined by the respective dominant terms $\widetilde{T}_j : x \mapsto \alpha_j x^{\beta_j}$. As the proof shows, deciding whether $O_T(x)$ is typically b-Benford reduces to verifying that $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all y. The example in (i) suggests that Theorem 3.1 is close to optimal: if $(B_n y)$ is increasing and u.d. mod 1 for some $y > 0$ then $(B_n/\log n)$ is unbounded [5]; on the other hand, (3.3) is satisfied whenever $(B_n/\log n)$ is eventually bounded below by a positive number. From a practical point of view, Theorem 3.1 will thus usually detect the generation of Benford sequences under (T_j) whenever typical orbits under $(\widetilde{T}_j)$ are Benford.

(iii) It is hardly surprising that *some* condition in the spirit of (3.2) has to be imposed on the sequence (α_j) lest it dominates the non-autonomous dynamics. For a very simple example of the latter situation, consider the family of polynomials

$$T_j(x) = 2^{-2^j} x^2 + 1\,, \quad j \in \mathbb{N}\,,$$

which are power-like near infinity. Each map T_j has infinity as an attracting fixed point, yet it is easily checked that $\lim_{n\to\infty} T^n(x) = 1$ for every $x \in \mathbb{R}$, so that $O_T(x)$ never is a Benford sequence; this is in accordance with (3.2) as the series $\sum_{j=1}^{\infty} B_j^{-1}\, |\log \alpha_j| = \sum_{j=1}^{\infty} \log 2$ diverges. (Notice, however, that $\big(T^n(x) - 1\big)$ may be b-Benford for all x; this can be seen to be the case if and only if $\log_b 2$ is 2-normal [14].)

(iv) In the autonomous case, i.e. for T_j not depending on j, condition (3.3) requires $\beta_j \equiv \beta > 1$, and rescaling x by a factor $\alpha^{(\beta-1)^{-1}}$ allows to make $\alpha_j \equiv \alpha > 0$ disappear from (3.1) without affecting the Benford properties of $O_T(x)$. There is thus no restriction at all on α, which is clearly reflected by the fact that condition (3.2) is void in the autonomous case.

Despite its technical appearance, Theorem 3.1 has a number of interesting applications when combined with standard results from uniform distribution theory. A few of these applications will be discussed below by means of examples. It should be emphasized, however, that notwithstanding its versatility Theorem 3.1 does rely heavily on the particular structure (3.1) of the maps under consideration. An even more comprehensive statement on BL in dynamical systems, including for instance random fluctuations or more general maps, may require refined or even completely different methods (see, however, Examples 4 and 5 below).

Example 1: Iterating polynomials

Consider a family (P_j) of real polynomials of degree at least two,

$$P_j(x) = a_{j,m_j} x^{m_j} + a_{j,m_j-1} x^{m_j-1} + \ldots + a_{j,1} x + a_{j,0}\,, \quad j \in \mathbb{N}\,, \tag{3.5}$$

with $m_j \in \mathbb{N}\backslash\{1\}$ and $a_{j,k} \in \mathbb{R}$, $a_{j,m_j} \neq 0$. Since (x_n) is a Benford sequence if and only if (x_n^{-1}) is a Benford sequence, one can study the non-autonomous dynamics under (3.5) near infinity by instead analyzing

$$T_j(x) = \big(P_j(x^{-1})\big)^{-1} = a_{j,m_j}^{-1} x^{m_j}\big(1 - f_j(x)\big) \tag{3.6}$$

near the fixed point at 0; here

$$f_j(x) = \frac{a_{j,m_j-1} x + \ldots + a_{j,0} x^{m_j}}{a_{j,m_j} + a_{j,m_j-1} x + \ldots + a_{j,0} x^{m_j}}\,.$$

Obviously, $f_j(0) = 0$ for all j; if $|a_{j,m_j}|^{-1} + \max_{k=0}^{m_j} |a_{j,k}| \leq C$ holds uniformly in j then also $|f_j'(x)| \leq 4(C+1)^4$ for all $j \in \mathbb{N}$ whenever $|x| \leq (C+2)^{-4}$. With $B_n = m_1 \cdot \ldots \cdot m_n$, the remaining conditions in Theorem 3.1 are also met. The natural numbers B_n are distinct, and therefore

$$\big|(\widetilde{S}_b^{n_1})'(y) - (\widetilde{S}_b^{n_2})'(y)\big| = |B_{n_1} - B_{n_2}| \geq 1\,, \quad \forall y \in \mathbb{R}\,,$$

provided that $n_1 \neq n_2$. By Koksma's Theorem [14,Thm.4.3] the sequence $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all y. Theorem 3.1 thus leads to the following

Theorem 3.2. *Let the maps P_j be real polynomials,*

$$P_j(x) = a_{j,m_j} x^{m_j} + a_{j,m_j-1} x^{m_j-1} + \ldots + a_{j,1} x + a_{j,0}\,,$$

with $m_j \in \mathbb{N}\backslash\{1\}$ and $a_{j,k} \in \mathbb{R}$, $a_{j,m_j} \neq 0$ for all j; $0 \leq k \leq m_j$. If

$$\sup\nolimits_{j\in\mathbb{N}}\big(|a_{j,m_j}|^{-1} + \max\nolimits_{k=0}^{m_j} |a_{j,k}|\big) < \infty\,, \tag{3.7}$$

then $O_P(x)$ is a Benford sequence for almost all $x \in \mathbb{R}\backslash[-K,K]$ with K sufficiently large. However, $\mathbb{R}\backslash[-K,K]$ also contains an uncountable dense set of exceptional points, i.e. points whose orbits are not Benford sequences.

Proof. Only the assertion concerning the exceptional points remains to be proved. Without loss of generality assume $a_{j,m_j} > 0$ for all j. The family $(\widetilde{T}_j)$ associated with (3.6) is $\widetilde{T}_j : x \mapsto a_{j,m_j}^{-1} x^{m_j}$, and therefore $\widetilde{S}_{j,b}(y) = \log_b a_{j,m_j} + m_j y$ as in the proof of Theorem 3.1. Since $B_j \geq 2^j$ for all j, condition (3.7) implies that $\sum_{j=1}^{\infty} B_j^{-1} \log_b a_{j,m_j}$ converges. If $m_j = 2$ for all but finitely many j, let (k_j) be any sequence in $\{0,1\}$ which does not contain two consecutive digits 1, i.e. $k_j k_{j+1} = 0$ for all $j \in \mathbb{N}$. Otherwise let (k_j) be any sequence of integers such that $0 \leq k_j \leq \lfloor m_j/3 \rfloor$

for all $j \in \mathbb{N}$. In both cases, there are uncountably many different points $\overline{y}$ defined through

$$\overline{y} := \sum\nolimits_{j=1}^{\infty} B_j^{-1}(k_j - \log_b a_{j,m_j}) .$$

By construction, $\left(B_n \sum_{j=1}^{\infty} B_j^{-1} k_j\right)$ is not u.d. mod 1 because $B_n \sum_{j=1}^{\infty} B_j^{-1} k_j \in [0, \frac{3}{4}[+\mathbb{N}_0$ for all n. Next, choose b sufficiently large to ensure that

$$\left|\widetilde{S}_b^n(\overline{y}) - B_n \sum\nolimits_{j=1}^{\infty} B_j^{-1} k_j\right| \leq B_n \sum\nolimits_{j=n+1}^{\infty} B_j^{-1} |\log_b a_{j,m_j}|$$

is less than $\frac{1}{9}$. Then $\left(\widetilde{S}_b^n(\overline{y})\right)$ is not u.d. mod 1 either, and $O_{\widetilde{T}}(b^{-\overline{y}})$ is not b-Benford.

The foregoing argument remains unchanged if finitely many k_j are replaced by arbitrary integers. This allows making $\overline{y}$ sufficiently large for the shadowing argument in Theorem 3.1 to apply; it also shows that exceptional points are actually dense near 0. In terms of the original family (P_j), this means that exceptional points form an uncountable dense set in $\mathbb{R}\backslash[-K, K]$ for some K sufficiently large. $\square$

Note that, in particular, Theorem 3.2 applies if (P_j) is a family of *finitely* many different polynomials, which is only rudimentally covered by the results in [5]. It can be formulated analogously for rational maps with degree (i.e. difference of the degrees of the numerator and denominator polynomial, respectively) at least two. Also, (P_j) may contain some affine maps, i.e. polynomials with degree $m_j = 1$, as long as this does not happen too often; details are left to the interested reader.

Example 2: The robustness of BL

The assertion of Theorem 3.2 is intuitively plausible since each individual map P_j is expanding near the attracting fixed point at infinity, and $O_P(x)$ is a Benford sequence for almost all large x. But even if T_j is predominantly chosen to be a map not generating any Benford sequence, $O_T(x)$ may still be strict Benford for almost all x, which can be seen as follows.

Fix the parameter $\lambda > 1$, and set $j_n = \lfloor 2n^{\lambda} \rfloor$; the sequence (j_n) is strictly increasing. Let

$$J_\lambda := \{j_n : n \in \mathbb{N}\} \subset \mathbb{N} .$$

For each N there exists a unique index n_N such that $j_{n_N} \leq N$ and $j_{n_N+1} > N$; the estimate $n_N = \mathcal{O}\left(N^{\lambda^{-1}}\right)$ as $N \to \infty$ is immediate. Clearly, J_λ has density zero, that is

$$\rho(J_\lambda) = \lim\nolimits_{N\to\infty} \frac{\#(J_\lambda \cap \{1, \ldots, N\})}{N} = \lim\nolimits_{N\to\infty} \frac{n_N}{N} = 0 .$$

Let the maps T_j be chosen according to

$$T_j(x) = \begin{cases} x^2 & \text{if } j \in J_\lambda , \\ x\left(1 + e^{-x^2}\right) & \text{if } j \notin J_\lambda . \end{cases}$$

Although not important in the sequel, it is readily confirmed that for some x_λ with $0 < x_\lambda < 1$, the iterates $T^n(x)$ converge to 0 or infinity, depending on whether $|x| \le x_\lambda$ or $|x| > x_\lambda$, respectively. It was proved in [5] that for $x \mapsto x\left(1 + e^{-x^2}\right)$, no orbit is a b-Benford sequence for any b. With $\alpha_j \equiv 1$, and β_j equal to 2 or 1 for $j \in J_\lambda$ or $j \notin J_\lambda$, respectively, the assumptions of Theorem 3.1 are satisfied, since for all sufficiently large n

$$c_1 n^{\lambda^{-1}} \le \log_2 B_n = \#(J_\lambda \cap \{1, \ldots, n\}) \le c_2 n^{\lambda^{-1}} ,$$

with constants $c_2 \ge c_1 > 0$. To see that $(B_n y)$ is u.d. mod 1 for almost all y, for every $N \in \mathbb{N}$ let

$$e_N = \#\left\{(m,n) \in \mathbb{N}^2 : 1 \le m, n \le N \, , \, B_m = B_n\right\} ,$$

i.e., e_N is the number of ordered pairs (m, n) in $\{1, \ldots, N\}^2$ with $B_m = B_n$. Consequently,

$$e_N = (N - j_{n_N} + 1)^2 + \sum\nolimits_{l=1}^{n_N - 1} (j_{l+1} - j_l)^2 + (j_1 - 1)^2 ,$$

and, by virtue of a rough estimate,

$$e_N = \mathcal{O}\left(N^{2-\lambda^{-1}}\right) \quad \text{as } N \to \infty .$$

Therefore $\sum_N N^{-3} e_N$ converges, and $(B_n y)$ is u.d. mod 1 for almost all y by [14,Exp.4.1]. Even though the individual maps T_j generate Benford sequences only if $j \in J_\lambda$, an event which has vanishing asymptotic relative frequency, $O_T(x)$ is nevertheless Benford for almost all $x \in \mathbb{R}$. This is yet another manifestation of the striking robustness of BL.

Example 3: Contraction on average ensures BL

In view of the previous example it is tempting to predominantly choose maps like $T_j(x) = \sqrt{x}$ whose dynamics is even farther from the generation of Benford sequences in that the fixed points at 0 and infinity are not even attracting any more. If this non-Benford behavior is compensated for on average, typical orbits may well be Benford sequences.

Assume that the family (T_j), as given by (3.1), only consists of a *finite* number of different maps, hence

$$\alpha_j = \alpha_{J_i} \, , \quad \beta_j = \beta_{J_i} \, , \quad \forall i = 1, \ldots, d \, ; \, j \in J_i \, ,$$

with the sets J_i forming a finite partition of $\mathbb{N}$, i.e. $J_1 \cup \ldots \cup J_d = \mathbb{N}$ and $J_i \cap J_k = \emptyset$ whenever $i \ne k$. Suppose that the densities $\rho(J_i)$ exist for all i and that the relative frequencies

$$r_i(n) = \frac{\#(J_i \cap \{1, \ldots, n\})}{n} \, , \quad i = 1, \ldots, d \, ,$$

converge to their respective limit $\rho(J_i)$ at a reasonable speed; more formally, in terms of the error $\delta(n) := \sum_{i=1}^{d} |r_i(n) - \rho(J_i)|$ assume that

$$\sum\nolimits_{n=1}^{\infty} n^{-1}\delta(n) < \infty \,. \tag{3.8}$$

If the (finitely many) functions f_j are C^1, then Theorem 3.1 can be applied, once (3.2) and (3.3) are satisfied, which in turn is the case if

$$\langle \log\beta\rangle := \sum\nolimits_{i=1}^{d} \rho(J_i)\log\beta_{J_i} > 0 \,. \tag{3.9}$$

Thus, if (3.9) holds, then $O_T(x)$ is a b-Benford sequence for almost all x close to 0 if and only if $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all large y. From $\widetilde{S}_{j,b}(y) = -\log_b\alpha_j + \beta_j y$ one readily obtains the explicit formula

$$\widetilde{S}_b^n(y) = B_n\big(y - \sum\nolimits_{j=1}^{n} B_j^{-1}\log_b\alpha_j\big) \,,$$

where

$$\log B_n = n\sum\nolimits_{i=1}^{d} r_i(n)\log\beta_{J_i} \,. \tag{3.10}$$

For every integer h, define the quantity $I_h = I_h(N)$ by

$$I_h(N) := N^{-2}\sum\nolimits_{m,n=1}^{N}\int_k^{k+1} e^{2\pi i h(\widetilde{S}_b^m(y)-\widetilde{S}_b^n(y))}dy \,. \tag{3.11}$$

According to [14,Thm.4.2], the sequence $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all $y \in [k, k+1]$ provided that $\sum N^{-1}I_h(N)$ converges for all $h \neq 0$. As will be explained now, a reasonably good bound for $|I_h(N)|$ can be found if (3.8) holds; the abbreviation $\gamma = \langle\log\beta\rangle > 0$ will be used throughout.

By means of the quantity $\Delta(n) = \sum_{i=1}^{d}\big(r_i(n) - \rho(J_i)\big)\log\beta_{J_i}$ formula (3.10) may be rewritten as $B_n = e^{n(\gamma+\Delta(n))}$; clearly, $|\Delta(n)| \le C\delta(n)$ with an appropriate constant $C > 0$. Defining sets

$$\begin{aligned} E_1 &:= \{(m,n) \,:\, 1 \le m \le n \le N,\, n \ge m + 2C\gamma^{-1}m\delta(m) + \gamma^{-1}\log 2\} \,,\\ E_2 &:= \{(m,n) \,:\, 1 \le m \le n \le N,\, m \le n - 2C\gamma^{-1}n\delta(n) - \gamma^{-1}\log 2\} \,, \end{aligned} \tag{3.12}$$

it is easily checked that for every pair $(m,n) \in E_1 \cap E_2$,

$$B_n - B_m \ge e^{m(\gamma + C\delta(m))} \,.$$

Thus for $(m,n) \in E_1 \cap E_2$ the estimate

$$\left|\int_k^{k+1} e^{2\pi i h(B_n - B_m)y}dy\right| \le 2e^{-\gamma m} \tag{3.13}$$

follows. On the other hand, the cardinality of the set

$$E_N' := \{(m,n) \,:\, 1 \le m \le n \le N\}\backslash(E_1 \cup E_2)$$

obviously is bounded by

$$\#E_N' \le 4C\gamma^{-1}\sum\nolimits_{m=1}^{N} m\delta(m) + 2N\gamma^{-1}\log 2 \,, \tag{3.14}$$

which follows from (3.12). For $(m, n) \in E'_N$ the trivial estimate

$$\left| \int_k^{k+1} e^{2\pi i h (B_n - B_m) y} dy \right| \leq 1 \tag{3.15}$$

holds. Combining estimates (3.13) and (3.15), and taking into account the symmetry of the integral, it follws from (3.11) that

$$\begin{aligned} N^2 |I_h(N)| &\leq 8C\gamma^{-1} \sum\nolimits_{m=1}^{N} m\delta(m) + 4N\gamma^{-1} \log 2 + 4 \sum\nolimits_{m,n=1}^{N} e^{-m\gamma} \\ &\leq d_1 \sum\nolimits_{m=1}^{N} m\delta(m) + d_2 N \end{aligned} \tag{3.16}$$

with positive constants d_1, d_2. Since $\sum_N N^{-3} \sum_{m \leq N} m\delta(m)$ converges if and only if $\sum_m m^{-1}\delta(m)$ converges, estimate (3.16) implies that $\sum_{N=1}^{\infty} N^{-1} |I_h(N)| < \infty$ for every integer h. Therefore $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all y, and Theorem 3.1 yields the following

Corollary 3.2. *Let the family of maps (T_j), as given by (3.1), be finite such that for the finite partition $(J_i)_{i=1}^d$ of $\mathbb{N}$*

$$\alpha_j = \alpha_{J_i} > 0\,, \quad \beta_j = \beta_{J_i} > 0\,, \quad \forall i = 1, \ldots, d\,;\, j \in J_i\,,$$

and $f_j \in C^1$, $f_j(0) = 0$. Assume that the densities $\rho(J_i)$ exist for all i and that

$$\sum\nolimits_{n=1}^{\infty} n^{-1} \sum\nolimits_{i=1}^{d} \left| \frac{\#(J_i \cap \{1, \ldots, n\})}{n} - \rho(J_i) \right| < \infty\,. \tag{3.17}$$

If (T_j) is contracting on average, i.e., if

$$\sum\nolimits_{i=1}^{d} \rho(J_i) \log \beta_{J_i} > 0\,, \tag{3.18}$$

then $O_T(x)$ is a strict Benford sequence for almost all x sufficiently close to 0.

Remark 3.2. (i) Condition (3.18) is equivalent to $\lim_{n\to\infty} n^{-1} \log B_n > 0$ and may be interpreted as (T_j) having a *positive* Lyapunov exponent on the logarithmic scale [10,13].

(ii) It is not difficult to find examples for which the conclusion of Corollary 3.2 is true even though (3.17) does not hold. Consider for instance the sequence $j_n := Cn \log_{(2)} n$, where C is chosen so large that $j_{n+1} - j_n \geq 2$ for all $n \geq 3$. For the set $J := \{j_n : n \in \mathbb{N}, n \geq 3\}$ clearly $\rho(J) = 0$, while $\sum n^{-1}\delta(n)$ diverges. Choosing

$$T_j(x) = \begin{cases} \sqrt{x} & \text{if } j \in J\,, \\ x^4 & \text{if } j \notin J\,, \end{cases}$$

condition (3.18) is satisfied. Moreover, $O_T(x)$ is a strict Benford sequence for almost all $x > 0$ since $n \mapsto B_n$ is one-to-one (see [14,Thm.4.1]).

Thus while condition (3.17) is needed in the above proof to ensure $\sum N^{-1} |I_h(N)| < \infty$, a condition known to be best possible in general [7], one may wonder whether (3.17) may not be disposed of in the present context, i.e., whether Corollary 3.2 remains valid without the assumption that $\sum n^{-1}\delta(n) < \infty$; no proof or counter-example is yet known to the author.

(iii) If (3.18) is not satisfied, then $O_T(x)$ may fail to be a b-Benford sequence for any $x > 0$ and any base b. For a trivial example consider

$$T_j(x) = \begin{cases} x^4 & \text{if 3 divides } j\,, \\ \sqrt{x} & \text{otherwise}\,, \end{cases}$$

for which equality holds in (3.18), and every orbit is periodic since $T_{j+2} \circ T_{j+1} \circ T_j = \mathrm{id}_{\mathbb{R}^+}$ for all j.

(iv) The map $T(x) = \sqrt{x}$ has been mentioned above as an example of $O_T(x)$ not being a b-Benford sequence for any b and $x > 0$. However, the reader may have noticed that, in a different sense, T also shows some Benford behaviour as $\bigl(T^n(x) - 1\bigr)$ is a b-Benford sequence for all $x > 0$, $x \neq 1$ unless $\log_2 b$ is an integer. Even though T is power-like near its fixed point, this effect lies beyond the scope of the present article because $\beta_j \equiv 1$ in local coordinates. A different shadowing argument is required for this *linearly dominated* system, see [4,5].

Example 4: Choosing maps at random

The ideas underlying the last example allow for an obvious stochastic interpretation: among the (finite) family $(T_{J_i})_{i=1}^d$ of possible maps, the map

$$T_{J_i}(x) = \alpha_{J_i} x^{\beta_{J_i}}\bigl(1 - f_{J_i}(x)\bigr)$$

is chosen at random, with probability $\rho(J_i)$. In this probabilistic terminology, condition (3.18) simply reads

$$\mathbb{E}\log\beta > 0\,. \tag{3.20}$$

Accordingly, in this section the parameters α, β in $T(x) = \alpha x^{\beta}\bigl(1 - f(x)\bigr)$ will be real-valued random variables (not necessarily attaining only finitely many values). When combined with mild technical assumptions, condition (3.20) does in fact guarantee almost sure generation of Benford sequences for such stochastic power-like dynamical systems. For the statement of the following result, recall that a measurable map T is *non-singular* if $T^{-1}(N)$ has measure zero for every set N of measure zero.

Theorem 3.3. *Let the family of maps $(T_{\alpha,\beta})_{\alpha,\beta>0}$ be given by*

$$T_{\alpha,\beta}(x) = \alpha x^{\beta}\bigl(1 - f_{\alpha,\beta}(x)\bigr)\,,$$

and assume that for all α, β, the map $T_{\alpha,\beta}$ is non-singular on $[0,1]$, and the non-negative function $f_{\alpha,\beta}$ is C^1 with $f_{\alpha,\beta}(0) = 0$ and $|f'_{\alpha,\beta}(x)| \le C$ for all $0 \le x \le 1$

with some constant $C > 0$. Let (α_j), (β_j) be i.i.d. sequences of positive random variables which satisfy

$$\sup_{n\in\mathbb{N}} \left(\sum_{j=1}^{n} B_j^{-1} \log \alpha_j\right) < D < \infty \quad \text{with probability one}\,, \tag{3.21}$$

for some constant $D > 0$, as well as $\mathbb{E}|\log \alpha_j| < \infty$ and $\mathbb{E}|\log \beta_j|^2 < \infty$. If T_{α_j,β_j} is contracting on average, i.e., if

$$\mathbb{E}\log \beta = \mathbb{E}\log \beta_j > 0\,, \tag{3.22}$$

then, for almost all x sufficiently close to 0, the stochastic orbit $\big(T_{\alpha_n,\beta_n} \circ \ldots \circ T_{\alpha_1,\beta_1}(x)\big)_{n\in\mathbb{N}}$ is a Benford sequence with probability one.

Proof. Denote the underlying (abstract) probability space for $(\alpha_j, \beta_j)_{j\in\mathbb{N}}$ by $(\Omega, \mathcal{A}, \mathbb{P})$. By the law of the iterated logarithm,

$$\left| n^{-1} \sum_{j=1}^{n} \log \beta_j(\omega) - \mathbb{E}\log \beta \right| = \mathcal{O}\big(n^{-\frac{1}{2}} (\log_{(2)} n)^{\frac{1}{2}}\big) \quad \text{as } n \to \infty\,, \tag{3.23}$$

for all $\omega \in \Omega_1 \subset \Omega$ where $\mathbb{P}(\Omega_1) = 1$. Conditions (3.2) and (3.3) in Theorem 3.1 are thus satisfied on some set $\Omega_2 \subset \Omega_1$ with $\mathbb{P}(\Omega_2) = 1$. This in turn implies the existence of a threshold $\xi_1(\omega) > 0$ such that $O_T(x)$ is a b-Benford sequence for almost all $x \in [0, \xi_1(\omega)]$ precisely if $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all large y. Since $B_n(\omega) = e^{n(\gamma+\Delta(n;\omega))}$ with $\Delta(n;\omega) = n^{-1}\sum_{j=1}^{n} \log \beta_j(\omega) - \mathbb{E}\log \beta$ and $\gamma = \mathbb{E}\log \beta > 0$, the same argument as in Example 3 shows that for all $\omega \in \Omega_2$ and an appropriate $\xi_2(\omega) \le \xi_1(\omega)$, the stochastic orbit $\big(T_{\alpha_n(\omega),\beta_n(\omega)} \circ \ldots \circ T_{\alpha_1(\omega),\beta_1(\omega)}(x)\big)$ is a strict Benford sequence for all $x \in [0, \xi_2(\omega)]$ except for a set $N(\omega)$ of Lebesgue measure zero. (Note that (3.23) implies that $\sum n^{-\frac{1}{2}-\varepsilon}|\Delta(n;\omega)|$ converges for every $\varepsilon > 0$; thus $\Delta(n;\omega)$ a.s. decays much faster than required by condition (3.8).) From the assumptions on (α_j), for $\omega \in \Omega_3 \subset \Omega_2$ with $\mathbb{P}(\Omega_3) = 1$,

$$T^n(x) \le \left(e^{\sum_{j=1}^{n} B_j(\omega)^{-1}\log \alpha_j(\omega)} x\right)^{B_n(\omega)} \le \big(e^D x\big)^{B_n(\omega)} \to 0 \quad \text{as } n \to \infty\,,$$

whenever $x < e^{-D}$. This latter bound is *independent of* ω. Therefore, the orbit of any $x \in [0, e^{-D}]$ is *not* a Benford sequence only if $x \in \bigcup_{n=0}^{\infty} T^{-n}\big(N(\omega)\big)$; by the (uniform) non-singularity of $T_{\alpha,\beta}$ the latter set has measure zero. Consequently, for every $\omega \in \Omega_3$ there exists a set $\overline{N}(\omega)$ of measure zero such that $O_T(x)$ is Benford for all $x \in [0, e^{-D}]\backslash\overline{N}(\omega)$. Let G denote the set of all pairs (x, ω) in $[0, e^{-D}] \times \Omega$ such that $O_T(x)$ is a Benford sequence. For every $x \in [0, e^{-D}]$ define $G_x := \{\omega \in \Omega : (x, \omega) \in G\}$; analogously, for every $\omega \in \Omega$ define $G^\omega := \{x \in [0, e^{-D}] : (x, \omega) \in G\}$. For every x and ω, G_x and G^ω are measurable subsets of Ω and $[0, e^{-D}]$, respectively. With these notations, $\lambda(G^\omega) = e^{-D}$ for every $\omega \in \Omega_3$, and $\mathbb{P}(\Omega_3) = 1$. By Fubini's theorem

$$e^{-D} = \int_\Omega \lambda(G^\omega)\, d\mathbb{P}(\omega) = \int_{[0,e^{-D}]\times\Omega} \mathbf{1}_G\, d(\lambda \times \mathbb{P}) = \int_{[0,e^{-D}]} \mathbb{P}(G_x)\, d\lambda(x)\,,$$

and $\mathbb{P}(G_x) = 1$ for all $x \in [0, e^{-D}]$ except for a set of Lebesgue measure zero. Thus, for almost every $x \in [0, e^{-D}]$ the stochastic orbit $O_T(x)$ is a Benford sequence with probability one. $\square$

For a very simple example consider the maps

$$T_j(x) = x^{2^{\xi_j}}(1 - f(x)) ,$$

where (ξ_j) is an i.i.d. sequence with $\mathbb{P}(\xi_j = 1) = p = 1 - \mathbb{P}(\xi_j = -1)$. Condition (3.22) is then equivalent to $p > \frac{1}{2}$. In the latter case, and for $f_j \equiv f$ denoting any non-negative C^1 function with $f(0) = 0$ and T_j non-singular, all requirements of Theorem 3.3 are met. Therefore, for almost all $x \in [0, 1]$ the stochastic orbit $O_T(x)$ is a Benford sequence with probability one. In the special case $f = 0$, the sufficient condition (3.22) is easily seen to be necessary also. Indeed, if $\mathbb{E}\log 2^{\xi_j} = (2p - 1) \log 2 < 0$, then for all $x > 0$

$$\lim_{n\to\infty} T^n(x) = 1 \quad \text{with probability one} .$$

On the other hand, if $p = \frac{1}{2}$ then $T^n(x) = x^{2^{S_n}}$, where $S_n = \sum_{j=1}^n \xi_j$ is the simple symmetric random walk on $\mathbb{Z}$. Thus, for any $k \in \mathbb{Z}$

$$\limsup_{n\to\infty} \frac{\#\{1 \leq j \leq n : S_j \leq k\}}{n} = 1 \quad \text{with probability one} , \tag{3.24}$$

which shows that for any $x > 0$, almost surely $T^n(x)$ is close to 1 sufficiently often in order to prevent $O_T(x)$ from being a b-Benford sequence for any b. (A nice justification of (3.24) in terms of dynamical systems can be derived from [1].)

Remark 3.3. (i) The conditions on (α_j) in Theorem 3.3 are automatically satisfied if $\mathbb{E}|\log \alpha| < \infty$ and if a.s. either $0 < \alpha \leq 1$, or if $\alpha \leq A$ and $\beta \geq B > 1$ with some constants A, B.

(ii) The assumptions in Theorem 3.3, in particular the non-negativity of $f_{\alpha,\beta}$, condition (3.21), and the uniform non-singularity requirement on $T_{\alpha,\beta}$, seem artificial on first sight. However, in general it may not be possible to weaken these assumptions in a straightforward manner, as the following simple examples show.

Fix ε with $0 < \varepsilon < 1$ and define two monotone power-like maps $T_2 \, (= T_{2^{+1}})$ and $T_{1/2} \, (= T_{2^{-1}})$ on $\mathbb{R}^+$ according to

$$T_{2^{\pm 1}}(x) = x^{2^{\pm 1}}\left(1 - f_\pm(x)\right) ,$$

where $f_\pm$ are C^∞ functions with

$$f_\pm(x) \equiv 0 \ \text{ if } x \leq 1 - \tfrac{3}{8}\varepsilon \, , \quad f_\pm(x) = 1 - \tfrac{1}{4}\varepsilon x^{-2^{\pm 1}} e^{-(1-x)^{-2}} \ \text{ if } x \geq 1 - \tfrac{1}{4}\varepsilon \, ,$$

such that $T_{2^{\pm 1}}$ is monotone and non-singular, and $f_\pm(x) \leq 0$ for $1 - \frac{3}{8}\varepsilon \leq x \leq 1 - \frac{1}{4}\varepsilon$ as well as $f_\pm(x) \geq -\varepsilon$ for all x. (Here and throughout, equations containing $\pm$ or $\mp$ are to be read as two individual equations containing only upper or lower signs, respectively.) As in the example above, let (β_j) be an i.i.d. sequence with $\mathbb{P}(\beta_j = 2) = 1 - \mathbb{P}(\beta_j = 1/2) = p$. If $p > \frac{1}{2}$, then all assumptions of Theorem 3.3

are satisfied by (T_{β_j}) except for the seemingly harmless fact that $f_\pm(\mathbb{R}) \subset [-\varepsilon, 1]$ rather than $f_\pm(\mathbb{R}) \subset [0,1]$. Nevertheless, it is easy to see that for any $0 < x < 1$,

$$\mathbb{P}(O_T(x) \text{ is not a Benford sequence }) \geq \mathbb{P}\big(\lim_{n\to\infty} T^n(x) = 1\big) \geq \Pi_{k(x)}(p) > 0\,;$$

here $k(x) = \lfloor \log_2 |\log_2(1-\frac{1}{4}\varepsilon)| - \log_2 |\log_2 x| \rfloor$, and $\Pi_k(p) = \min\big((p/q)^k, 1\big)$ denotes the probability that a simple random walk on $\mathbb{Z}$ starting at the origin and moving with probabilities p and $q = 1-p$ one step to the right and left, respectively, ever hits $k \in \mathbb{Z}$. Consequently, points escape with positive probability from any interval $[0, \delta]$, and an almost sure statement in the spirit of Theorem 3.3 cannot possibly be made.

By virtually the same construction, it is possible to define two monotone, non-singular power-like maps $T_{2^{\pm 1}} = e^\varepsilon x^{2^{\pm 1}}\big(1 - g_\pm(x)\big)$ where $g_\pm$ are C^∞ non-negative monotone functions which vanish on $]-\infty, e^{-3\varepsilon}]$ and equal $1 - e^{-\varepsilon} x^{-2^{\pm 1}}(1 - \varepsilon e^{-(1-x)^{-2}})$ on $[e^{-\frac{1}{3}\varepsilon}, +\infty[$. Contrary to the functions $f_\pm$ from above, however, $g_\pm(\mathbb{R}) \subset [0,1]$. Now the only requirement in Theorem 3.3 not satisfied with $\alpha_j \equiv e^\varepsilon$ and (β_j) as above is that $\sum_{j=1}^\infty B_j^{-1} \log \alpha_j = \varepsilon \sum_{j=1}^\infty B_j^{-1}$ be bounded by a constant with probability one. (Although it is not bounded by a constant, the non-negative random quantity $\zeta := \sum_{j=1}^\infty B_j^{-1}$ is finite almost surely, and $\mathbb{E}\zeta^m < \infty$ if and only if $p = \mathbb{P}(\beta_j = 2) > (1 + 2^{-m})^{-1}$.) As before $\mathbb{P}\big(\lim_{n\to\infty} T^n(x) = 1\big) > 0$ for all $x > 0$, and no almost sure statement can be made for even the tiniest neighborhood of the origin.

Finally, to understand the importance of the uniform non-singularity requirement let $\rho_n := (4 + 2^{n+3})^{-1}$ and

$$a_n := (2^n + \rho_n)^{-1}\,, \quad c_n := \frac{2^n}{(2^n + \rho_n)^{1+2^{-n}}}\,, \quad n \in \mathbb{N}\,,$$

and, by means of these quantities, define C^1 functions

$$h_n(x) := \begin{cases} 0 & \text{if } x \leq (1-2\rho_n)a_n\,, \\ \frac{1}{4}\rho_n^{-1} a_n^{-1}\big(x - (1-2\rho_n)a_n\big)^2 & \text{if } (1-2\rho_n)a_n < x \leq a_n\,, \\ 1 - c_n x^{-2^{-n}} & \text{if } x > a_n\,. \end{cases}$$

It is easily checked that $h_n(\mathbb{R}) \subset [0,1]$, and $\sup_{x\in\mathbb{R}} |h_n'(x)| = 1$ for all $n \in \mathbb{N}$; obviously, $h_n(0) = 0$. Define a family of power-like C^1 maps on the positive reals as

$$T_{2^{-n}}(x) := x^{2^{-n}}\big(1 - h_n(x)\big)\,, \quad n \in \mathbb{N}\,. \tag{3.25}$$

Clearly, $T_{2^{-n}}(x) = x^{2^{-n}}$ for $x \leq (1-2\rho_n)a_n$, and $T_{2^{-n}}(x) = c_n$ for $x \geq a_n$. Also let h_{-1} be a non-negative C^1 function such that $h_{-1}(x) = 0$ for $x \leq \frac{1}{2}$ and $h_{-1}(x) = 1 - (2x)^{-2}$ for $x \geq \frac{2}{3}$; with this define $T_2(x) := x^2\big(1 - h_{-1}(x)\big)$. Next, assume that (β_j) is an i.i.d. sequence with

$$\mathbb{P}(\beta = 2) = p\,, \quad \mathbb{P}(\beta = 2^{-n}) = (1-p)2^{-n} \quad n \in \mathbb{N}\,.$$

Condition (3.22) is equivalent to $p > \frac{2}{3}$. In the latter case, all assumptions of Theorem 3.3 are satisfied except for the fact that each map $T_{2^{-n}}$ is nonsingular only on $[0, a_n]$, the length of which intervals shrinks to zero exponentially with n. As in the two previous examples, this apparently minor deviation causes Theorem 3.3 to fail for (3.25). Indeed, for each $\delta > 0$ the set

$$I_\delta := [0, \delta] \setminus \bigcup\nolimits_{n=1}^{\infty} [(1 - 2\rho_n) a_n, a_n]$$

has positive measure (in fact, $\lambda(I_\delta)/\delta \to 1$ as $\delta \to 0$), and for every $x \in I_\delta$ one has $T_{2^{-l}}(x) = c_l$, provided that $a_l \leq x$. Once x gets mapped to the interval $[\frac{2}{3}, 1]$, its orbit cannot be a Benford sequence. Thus for each $x \in I_\delta$

$$\mathbb{P}(O_T(x) \text{ is not a Benford sequence}) \geq \mathbb{P}(\beta \in \{2^{-n} : n \geq l\}) = (1-p)2^{-l+1} > 0 \,.$$

In stark contrast to the assertion of Theorem 3.3 most orbits starting arbitrarily close to zero therefore have a positive probability of *not* being Benford sequences. (The reader may have noticed that this pathology is due to the random nature of (β_j). If $(\beta_j)_{j\in\mathbb{N}}$ were a *given* sequence in $\{2\} \cup \{2^{-n} : n \in \mathbb{N}\}$ such that $n^{-1}\sum_{j=1}^{n} \log \beta_j$ converges sufficiently fast to a positive limit, then (an appropriately generalized version of) Corollary 3.2 would guarantee that almost every orbit close to the origin is a strict Benford sequence.)

In the light of the above examples, it is natural to expect that Theorem 3.3 can be strengthened considerably in the special case $f_{\alpha,\beta} \equiv 0$.

Corollary 3.3. *Let the family $(\widetilde{T}_{\alpha,\beta})_{\alpha,\beta>0}$ be given by $\widetilde{T}_{\alpha,\beta}(x) = \alpha x^\beta$, and let (α_j) and (β_j) be an arbitrary and an i.i.d. sequence of positive random variables, respectively. If $\mathbb{E}\log\beta > 0$ and $\mathbb{E}(\log\beta)^2 < \infty$ then for almost all $x > 0$, the stochastic orbit $\big(\widetilde{T}_{\alpha_n,\beta_n} \circ \ldots \circ \widetilde{T}_{\alpha_1,\beta_1}(x)\big)$ is a Benford sequence with probability one.*

Proof. Again, invoking the law of the iterated logarithm, and by means of calculations analogous to those following (3.11), it is easy to see that $\big(\widetilde{S}_b^n(y)\big)$ is u.d. mod 1 for almost all y, regardless of the specific form of the sequence (α_j). As before, an application of Fubini's theorem completes the proof. □

Remark 3.4. Although it simplifies the arguments, independence of the sequence (β_j) in not essential in the proof of Theorem 3.3 (and Corollary 3.3). Specifically, if $\alpha_j \in [\delta, 1]$ for some $\delta > 0$ and all j then all that is needed is that, with probability one, the family $(T_{\alpha,\beta})$ is contracting on average, i.e.

$$\langle \log \beta \rangle = \lim\nolimits_{n\to\infty} n^{-1} \sum\nolimits_{j=1}^{n} \log \beta_j > 0 \,,$$

and that, with probability one,

$$\sum n^{-1} \left| n^{-1} \sum\nolimits_{j=1}^{n} \log \beta_j - \langle \log \beta \rangle \right| < \infty \,.$$

For example, (β_j) could be any stationary stochastic process on $\mathbb{R}^+$ which satisfies a law of the iterated logarithm for $(\log \beta_j)$, and for which $\langle \log \beta \rangle > 0$. (See [10] for a related statement concerning affine dynamics.)

Example 5: Shadowing under stronger contraction

Maps showing contraction properties near zero (or expansion properties near infinity) that are stronger than power-like are not covered by the results in this article. For example, the following family of maps is not power-like near infinity:

$$T_j(x) = e^{x^2} P_j(x) , \quad j \in \mathbb{N} , \tag{3.26}$$

where $P_j(x) = a_{j,m_j} x^{m_j} + \ldots + a_{j,1}x + a_{j,0}$ are real polynomials satisfying (3.7), and $m_j \geq 1$ for all j. However, it follows from (the obvious non-autonomous analogue of) [5,Thm.4.4] that the orbit $O_T(x)$ is Benford for almost all sufficiently large $|x|$. Under the mild additional assumption that $\sup_{j\in\mathbb{N}} m_j < \infty$ it is illuminating to derive this result in an entirely different manner, namely from an analysis of (3.26) in the spirit of Theorem 3.1. As the argument resembles the proof of [5,Thm.4.1] only a sketch of the ideas will be given.

For $k, n \in \mathbb{N}$ define the function $\varphi_{k,n}$ as

$$\varphi_{k,n}(x) := \underbrace{\sqrt{\log \sqrt{\ldots \sqrt{\log}}}}_{n\text{-fold } \sqrt{\log(\cdot)}} \circ T_{k+n-1} \circ T_{k+n-2} \circ \ldots \circ T_k(x) ;$$

clearly, $\varphi_{k,n}$ is smooth for $x \geq x_0$ with x_0 sufficiently large but independent of both k and n. In analogy to (2.8) it is straightforward to see that

$$\varphi_k(x) := \lim_{n\to\infty} \varphi_{k,n}(x) \tag{3.27}$$

exists and defines a continuous function φ_k on $[x_0, +\infty[$. In fact the convergence in (3.27) is uniform in x and $k \in \mathbb{N}$, and φ_k is continuously differentiable. Moreover, it is easily verified that, uniformly in k,

$$|\varphi_k(x) - x| \to 0 , \quad |\varphi_k'(x) - 1| \to 0 \quad \text{as } x \to \infty . \tag{3.28}$$

By construction, $e^{\varphi_{j,n}^2} = \varphi_{j+1,n-1} \circ T_j$ and therefore also $e^{\varphi_j^2} = \varphi_{j+1} \circ T_j$ for all $j \in \mathbb{N}$. Thus, from (3.28) it follows that, for all $x \geq x_0$,

$$\left| \underbrace{e^{e^{2e^{\cdot^{\cdot^{\cdot 2e^{2\varphi_1^2(x)}}}}}}}_{n\text{-fold exponent}} - T^n(x) \right| = \bigl|\varphi_{n+1} \circ T^n(x) - T^n(x)\bigr| \to 0 \quad \text{as } n \to \infty .$$

Consequently, $O_T(x)$ is a Benford sequence for almost all $x \geq x_0$, precisely if $O_{\widetilde{T}}(x)$ is Benford for almost all sufficiently large x, with $\widetilde{T}(x) = e^{x^2}$. Theorem 4.4 in [5], for instance, shows that the latter is indeed the case.

Remark 3.5. (i) With T_j according to (3.26) the maps $\overline{T}_j : x \mapsto T_j(x^{-1})^{-1}$ can be written as $\overline{T}_j(x) = x^{m_j - x^{-2}(\log x)^{-1}} a_{j,m_j}^{-1}\bigl(1 - f_j(x)\bigr)$. Formally, therefore, $\overline{T}_j$ is power-like near 0, but $\beta_j = m_j - x^{-2}(\log x)^{-1}$ depends on x. With the appropriate uniformity requirements, Theorems 2.1 and 3.1 can be generalized so as to allow for the quantities β_j to depend on y and x, respectively. As a consequence, $O_{\overline{T}}(x)$ is a strict Benford sequence for almost all x near 0 if and only if the same is true for

$O_{\widetilde{T}}(x)$ where $\widetilde{T}_j(x) = a_{j,m_j}^{-1} e^{-x^{-2}} x^{m_j}$. Thus the simple shadowing idea discussed in this article extends naturally to more general maps; the details of any such generalization are left to the interested reader.

(ii) As in [4,5], uniform distribution theory plays a substantial role in the present article. In this context, it is interesting to note that an increasing sequence (x_n) may be both u.d. mod 1 and Benford. In fact, among the sequences studied here, this property is fairly widespread: take $x_n = x_0^{n+1}$ as an example with almost every $x_0 > 1$. However, if $\varphi > 1$ is an irrational Pisot number [14], then $x_n = \varphi^{n+1}$ is Benford but not u.d. mod 1, whereas $x_n = (n+1)\varphi$ is u.d. mod 1 yet fails to be Benford. Trivially, $x_n = \log(n+1)$ does not show either property. It may thus be interesting to study whether the simultaneous Benford property of (x_n) and $(\log_{(k)} x_n)$, where $k = 1, 2, \ldots$, implies further properties for the sequence (x_n).

Acknowledgments

The author is grateful for inspiring discussions with and helpful comments from L. Bunimovich, J. Elton, T. Hill and R. Zweimüller.

References

1. J. Aaronson, M. Thaler, M., and R. Zweimüller, Occupation times of sets of infinite measure for ergodic transformations, *preprint* (2002).
2. F. Benford, The law of anomalous numbers, *Proceedings of the American Philosophical Society* **78** (1938), 551–572.
3. A. Berger, *Chaos and Chance*, de Gruyter, Berlin-New York (2001).
4. A. Berger, Multi-dimensional Dynamical Systems and Benford's Law, *Discrete Contin. Dyn. Syst.* **13** (2005), 219–237.
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. J. Brown, and R. Duncan, Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences, *Fibonacci Quarterly* **8** (1970), 482–486.
7. H. Davenport, P. Erdös, and W.J. LeVeque, On Weyl's criterion for uniform distribution, *Michigan Math. J.* **10** (1963), 311–314.
8. D.S. Dendrinos, and M. Sonis, *Chaos and Socio-Spatial Dynamics*, Springer, New York Berlin Heidelberg (1990).
9. P. Diaconis, The distribution of leading digits and uniform distribution mod 1, *Ann. Probab.* **5** (1979), 72–81.
10. G. Gottwald, and M. Nicol, On the nature of Benford's Law, *Phys. A* **303** (2001), 387–396.
11. T.P. Hill, A statistical derivation of the significant-digit law, *Statistical Science* **10** (1996), 354–363.
12. Z. Huang, and S. Solomon, Stochastic multiplicative processes for financial markets, *Phys. A* **306** (2002), 412–422.
13. A. Katok, and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge (1995).
14. L. Kuipers, and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York (1974).

15. E. Ley, On the Peculiar Distribution of the U.S. Stock Indices Digits, *Amer. Stat.* **50** (1996), 311–313.
16. H.-W. Lorenz, *Nonlinear Dynamical Economics and Chaotic Motion* (Second Edition), Springer, Berlin Heidelberg New York (1993).
17. M.J. Nigrini, *Digital Analysis Using Benford's Law: Tests and Statistics for Auditors* (Second Edition), Global Audit Publications, Vancouver (2000).
18. K. Palmer, *Shadowing in Dynamical Systems. Theory and Applications*, Kluwer, Dordrecht (2000).
19. R. Raimi, The first digit problem, *Amer. Math. Monthly* **102** (1976), 322–327.
20. 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.
21. P. Schatte, On mantissa distributions in computing and Benford's law, *J. Information Processing and Cybernetics* **24** (1998), 443–455.
22. C. Tolle, J. Budzien, and R. LaViolette, Do dynamical systems follow Benford's law? *Chaos* **10** (2000), 331–337.