On the distribution of mantissae in nonautonomous difference equations

A. Berger[a,*], S. Siegmund[b]

[a] *Dept. of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand*
[b] *Fachbereich Mathematik, Johann Wolfgang Goethe–Universität, Frankfurt am Main, Germany*

Received 22 December 2006, revised 28 February 2007

Dedicated to the memory of Bernd Aulbach

Abstract

Mantissa distributions generated by dynamical processes continue to attract much interest. In this article, it is demonstrated that one-dimensional projections of (at least) almost all orbits of many multi-dimensional nonautonomous dynamical systems exhibit a mantissa distribution that is a convex combination of a trivial point mass and Benford's Law, i.e., the mantissa distribution of the non-trivial part of the orbit is asymptotically logarithmic, typically for all bases. Both linear and power-like systems are considered, and Benford behaviour is found to be ubiquitous for either class. The results unify previously known facts and extend them to the nonautonomous setting, with many of the conclusions being best possible in general.

Key words and phrases: Nonautonomous dynamical system, Benford's Law, uniform distribution mod 1 in $\mathbb{R}^d$, exponential dichotomy, shadowing.

AMS 2000 subject classification: 11K06, 37A45, 37B55.

1 Introduction

Studying the distribution of digits and mantissae of numerical data generated by some dynamical process or other is an intriguing task that continues to attract interest from a wide spectrum of disciplines, among which are for instance number theory [15], statistics [20], and accounting [21]. One recurring theme in this context is the astonishing ubiquity of a logarithmic distribution commonly referred to as *Benford's Law* (BL). The latter denotes the probability distribution for the mantissa with respect to the base $b \in \mathbb{N} \setminus \{1\}$ given by

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

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

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

[*] Corresponding author.
E-mail addresses: `arno.berger@canterbury.ac.nz`, `siegmund@math.uni-frankfurt.de`

Examples of empirical data sets following (1) have been discussed extensively, for instance in real-life data (e.g., physical constants, stock market indices, tax returns [12, 17, 21, 23, 25]), in stochastic processes (e.g., sums and products of random variables [11, 23]), and in deterministic sequences (e.g., $(n!)$ and Fibonacci numbers [3, 8, 10]). Recently a thorough mathematical analysis of BL for dynamical systems has been initiated [4, 5, 6, 11, 24, 26]. Following physical experiments and numerical simulations, it has been shown that orbits of many dynamical systems follow BL surprisingly often; for an application to Newton's method and other root-finding algorithms see [7].

Dynamical systems, be they deterministic or stochastic, autonomous or nonautonomous, are widely used as models for real-world phenomena. If the latter exhibit, on an empirical level, a striking statistical property like (1) — as often they do — then it is natural to ask for a rigorous manifestation of this property in the underlying mathematical model. For *autonomous* systems this task has to a certain extent been completed in [4, 6] where it was shown that BL will typically emerge in a very strong sense, provided that the dynamics under consideration shows uniform growth or decay of orbits. For *nonautonomous* dynamical systems, on the other hand, the situation is less clear, and only the one-dimensional case (which of course is somewhat special) has been analysed in some depth [5, 11]. With the overall appreciation for and understanding of nonautonomous dynamics increasing, it seems timely to address the problem of mantissa distributions for these systems as well.

The purpose of the present article is to initiate the study of mantissa and digit distributions for nonautonomous dynamical systems. Given the vast variety of nonautonomous behaviour, this can naturally be but a first step. Already, however, and not completely unexpectedly, it becomes apparent that the task ahead may pose a considerable challenge, as the information provided by standard tools (e.g. Sacker–Sell spectrum, Lyapunov exponents) may not be accurate enough for this purpose. This article uses a version of *shadowing*, an important technique in both the autonomous and the nonautonomous setting, to overcome this problem. In tribute to Bernd Aulbach, the sole focus is on difference equations (see e.g. [1, 2] and references therein). For the class of systems considered here, this is not a substantial restriction at all, as the results and counterexample carry over to the continuous time case in a straightforward way.

From one of the main results of this article, Theorem 9, one can for instance deduce that, unless $x_0 = 0$, each component of the sequence (x_n) in $\mathbb{R}^2$, defined iteratively as

$$x_n = \begin{pmatrix} a_n & 1+b_n \\ 1+c_n & 2+d_n \end{pmatrix} x_{n-1}\,, \qquad n = 1, 2, \ldots\,,$$

exhibits the distribution (1), for every $b \in \mathbb{N}\backslash\{1\}$ — provided that

$$\sum\nolimits_{n=1}^{\infty} (3+2\sqrt{2})^n \max\{|a_n|, |b_n|, |c_n|, |d_n|\} < \infty\,.$$

(While best possible in general, the latter condition can be weakened considerably in some special cases, see Example 10 and Remark 11(ii).) Similarly, each component of Lebesgue almost every (nonautonomous) orbit generated by *any successive combination* of the two bi-variate polynomial maps

$$T_a : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} 3u^2v^2+4 \\ 5u^2v^4 - 6v^2 + 7 \end{pmatrix}, \qquad T_b : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} v \\ 2uv \end{pmatrix},$$

exhibits (1), provided the orbit starts (and hence forever stays) sufficiently far away from the two coordinate axes $u = 0$ and $v = 0$, see Theorem 16 and Example 19. Besides giving complete proofs for these and other new results, this article, by demonstrating how the ad-hoc approach chosen here may fail, also highlights some of the difficulties which a future, more systematic treatment will have to address.

2 Basic notations and definitions

The sets of natural, non-negative integer, integer, rational, positive real, real, and complex numbers are symbolised by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}^+$, $\mathbb{R}$, and $\mathbb{C}$, respectively. The real part, imaginary part, complex conjugate and absolute value (modulus) of $z \in \mathbb{C}$ is denoted by $\Re z$, $\Im z$, $\overline{z}$ and $|z|$, respectively. For $z \neq 0$, the argument $\arg z$ is the unique number in $]-\pi, \pi]$ which makes $z = |z|e^{i \arg z}$ hold; for ease of notation $\arg 0 := 0$. The unit circle S^1 is interpreted as $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. If A is a complex $d \times d$-matrix, then $A^\top$ symbolises its transpose and $\sigma(A) \subseteq \mathbb{C}$ its spectrum, i.e. the *set* of all eigenvalues of A. For any $A \in \mathbb{C}^{d\times d}$ let $A_\mathbb{R}$ be its realification, i.e. the real matrix

$$A_\mathbb{R} = \begin{pmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{pmatrix} \in \mathbb{R}^{2d\times 2d}.$$

Accordingly, the set $\sigma^+(A) = \{\lambda \in \sigma(A_\mathbb{R}) : \Im\lambda \geq 0\}$ is the "upper half" of the symmetrised spectrum $\sigma(A_\mathbb{R}) = \sigma(A) \cup \overline{\sigma(A)}$; obviously, $\sigma(A_\mathbb{R}) = \sigma(A)$ whenever A is real. The standard inner product on $\mathbb{C}^d$ is $\langle x, y\rangle = \sum_{j=1}^d x^{(j)}\overline{y^{(j)}}$; it induces the Euclidean norm $\|x\| = \sqrt{\langle x, x\rangle}$. The symbols $r_\sigma(A)$ and $\|A\|$ denote, respectively, the spectral radius and the matrix norm of A as induced by $\|\cdot\|$, that is, $r_\sigma(A) = \max\{|\lambda| : \lambda \in \sigma^+(A)\}$ and $\|A\| = \sqrt{|\mu|}$ where μ is the largest eigenvalue of $\overline{A}^\top A$. The specific choice of a norm on $\mathbb{C}^d$ (and, correspondingly, of an induced norm on $\mathbb{C}^{d\times d}$) is largely irrelevant for this article, and the Euclidean norm is chosen for convenience only; see however Remark 11(i).

Throughout, b denotes a natural number larger than one (called a *base*). Every $x \in \mathbb{R}^+$ can be written uniquely as $x = M_b(x)b^l$ with $M_b(x) \in [1, b[$ and the appropriate $l \in \mathbb{Z}$. The function $M_b : \mathbb{R}^+ \to [1, b[$ is called the (*base* b) *mantissa function*; for convenience let $M_b(0) := 0$ for all b. For every $x \in \mathbb{R}$, the numbers $\lfloor x\rfloor$ and $\lceil x\rceil$ denote the largest integer not larger, and the smallest integer not smaller than x, respectively. The number $\lfloor M_b(x)\rfloor \in \{1, \ldots, b-1\}$ is called the *first significant digit* of x (with respect to base b). For a given base b, $\log_b$ will denote the logarithm with respect to b, where, to avoid cumbersome formulations, $\log_b 0 := 0$ for all b; if used without a subscript, the log symbol denotes the natural logarithm. The cardinality of the finite set S is $\#S$, and λ^d symbolises the d-dimensional Lebesgue measure on $\mathbb{R}^d$ (or parts thereof).

Definition 1. A sequence $(x_n)_{n\in\mathbb{N}_0}$ of real numbers is called b-*Benford* if

$$\lim_{n\to\infty} \frac{\#\{l < n : M_b(|x_l|) \leq t\}}{n} = \log_b t\,, \qquad \forall t \in [1, b[\,,$$

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

The Borel probability measure on $[1, b]$ with distribution function $\log_b t$ will be denoted by $\mathbb{B}_b$. Thus $(x_n)_{n\in\mathbb{N}_0}$ is b-Benford precisely if $\frac{1}{n}\sum_{l=0}^{n-1}\delta_{M_b(|x_l|)}$ converges weakly to $\mathbb{B}_b$; here δ_x is, for any $x \in \mathbb{R}$, the Dirac measure at x. Generally, given any (real or complex) sequence $(x_n)_{n\in\mathbb{N}_0}$ the (possibly finite or even empty) subsequence of its non-zero elements will be denoted by $(x_n)^*$, that is, $(x_n)^* = (x_{n_k})_{k\in\mathbb{N}_0}$ where $0 \leq n_0 < n_1 < \ldots$ and $\{n_k : k \in \mathbb{N}_0\} = \{n \in \mathbb{N}_0 : x_n \neq 0\}$.

The following correspondence between Benford sequences and uniform distribution modulo one is well known [10]. The term *uniformly distributed modulo one* (*in* $\mathbb{R}^d$) will henceforth be abbreviated as *u.d.* mod 1 (in $\mathbb{R}^d$).

Proposition 2 ([10])**.** *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.

This article studies Benford properties of recursively defined sequences,

$$x_n = T_n(x_{n-1})\,, \qquad n \in \mathbb{N}\,, \tag{2}$$

where, for each n, T_n denotes a map from $\mathbb{R}^d$ (or $\mathbb{C}^d$) or a part thereof into itself. For ease of notation, no distinction will be made between row and column vectors, that is, $x \in \mathbb{C}^d$ should be thought of as a column but will nevertheless be written as $x = \big(x^{(1)}, \dots, x^{(d)}\big)$. For $n \in \mathbb{N}$ the n-fold composition of maps, $T_n \circ T_{n-1} \circ \dots \circ T_1$, is denoted by T^n, and $T^0 := id$. The sequence generated by (2) starting from the initial state x_0 is thus $\big(T^n(x_0)\big)_{n\in\mathbb{N}_0}$; this sequence is denoted by $O_T(x_0)$ and referred to as the (*nonautonomous*) *orbit* of x_0 under T. Note that this interpretation of the orbit as a *sequence* differs from terminology in dynamical systems theory (e.g. [14]) according to which the orbit of x_0 is the mere *set* $\{x_n : n \in \mathbb{N}_0\}$. For any function φ defined on $\mathbb{C}^d$, the symbol $\varphi\big(O_T(x_0)\big)$ stands for the sequence $\big(\varphi \circ T^n(x_0)\big)_{n\in\mathbb{N}_0}$; for example, $\|O_T(x_0)\|$ denotes the sequence of Euclidean norms $(\|x_n\|)_{n\in\mathbb{N}_0}$, and $\langle c, O_T(x_0)\rangle$ symbolises the sequence of inner products $\big(\langle c, x_n\rangle\big)_{n\in\mathbb{N}_0}$. As a special case of the latter, with $c = e_j$ representing the j-th vector of the canonical basis, $\langle e_j, O_T(x_0)\rangle = O_T^{(j)}(x_0) = (x_n^{(j)})_{n\in\mathbb{N}_0}$ is the sequence of j-th components of $O_T(x_0)$.

3 Linear Systems

This section studies, under the perspective of BL, nonautonomous *linear* equations

$$x_n = A_n x_{n-1}\,, \qquad n \in \mathbb{N}\,, \tag{3}$$

where, for each n, A_n is a real $d \times d$-matrix. Thus each map T_n in (2) is linear, $T_n(x) = A_n x$. To analyse the mantissa distribution of $O_T(x_0)$ generated by (3) first recall the definition of b-resonance from [4].

Definition 3. (i) A set $\Lambda \subset \mathbb{C}$ is *b-resonant* if there exists a finite non-empty subset $\Lambda_0 = \{\lambda_1, \dots, \lambda_l\} \subset \Lambda$ with $|\lambda_1| = \dots = |\lambda_l|$ such that either $\#(\Lambda_0 \cap \mathbb{R}) = 2$ or the numbers 1, $\log_b|\lambda_1|$ as well as the elements of $\big\{\frac{1}{2\pi}\arg\lambda_1, \dots, \frac{1}{2\pi}\arg\lambda_l\big\}\backslash\big\{0, \frac{1}{2}\big\}$ are $\mathbb{Q}$-dependent.

(ii) A (real or complex) matrix A has *b-resonant spectrum* if the set $\sigma^+(A)$ is b-resonant.

For the autonomous case, that is for A_n in (3) not depending on n, [4, Thm.3.3] shows that, for every $c \in \mathbb{R}^d$, the sequence $\langle c, O_T(x_0)\rangle^*$ is either finite or b-Benford — provided that $A \equiv A_n$ does not have b-resonant spectrum. Thus, for every $c, x_0 \in \mathbb{R}^d$, and if A does not have b-resonant spectrum, the asymptotic mantissa distribution base b of $\big(\langle c, A^n x_0\rangle\big)_{n\in\mathbb{N}_0}$ is of the form $\rho\mathbb{B}_b + (1-\rho)\delta_0$ with $\rho \in \{0, 1\}$. This result generalises easily to the case of a p-periodic sequence (A_n), i.e., $A_{n+p} = A_n$ for some $p \geq 1$ and all $n \in \mathbb{N}$.

Theorem 4. *Let (A_n) in (3) be p-periodic for some $p \geq 1$ and assume that the matrix $A_p \cdot \ldots \cdot A_1$ does not have b-resonant spectrum. Then, for every $c, x_0 \in \mathbb{R}^d$, the sequence $\langle c, O_T(x_0)\rangle^*$ is either finite or b-Benford.*

Proof. Let $B = A_p \cdot \ldots \cdot A_1$, and observe that, for every $l = 0, 1, \dots, p-1$, the sequence

$$\big(\log_b|\langle c, x_{np+l}\rangle|\big)^* = \big(\log_b|\langle c, A_l \cdot \ldots \cdot A_1 \cdot B^n x_0\rangle|\big)^* = \big(\log_b|\langle A_1^\top \cdot \ldots \cdot A_l^\top c, B^n x_0\rangle|\big)^*$$

is either finite or u.d. mod 1, by [4, Thm.3.3]. With [16, Exc.2.12] the same is true for the entire sequence $\big(\log_b\langle c, x_n\rangle\big)^*$. □

Example 5. Let (A_n) be the 2-periodic sequence of matrices with

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & 3 + (-1)^n \end{pmatrix}, \qquad n \in \mathbb{N}\,.$$

The set $\sigma^+(A_2A_1) = \{5 \pm 2\sqrt{6}\}$ is not b-resonant for any b. Also, for every non-zero c and x_0, $\langle c, O_T(x_0)\rangle^*$ is infinite. Thus $\langle c, O_T(x_0)\rangle^*$ is Benford whenever $c, x_0 \in \mathbb{R}^d\backslash\{0\}$.

Corollary 6. *Under the assumptions of Theorem 4 the asymptotic mantissa distribution base b of $\langle c, O_T(x_0)\rangle$ is given by $\rho\mathbb{B}_b + (1-\rho)\delta_0$, with $\rho \in \left\{0, \frac{1}{p}, \frac{2}{p}, \ldots, \frac{p-1}{p}, 1\right\}$ depending on c and x_0.*

Remarks 7. (i) The crucial non-resonance condition in Theorem 4 could be rephrased as (A_n) not having b-resonant spectrum *on average*. For the latter to be the case, it is neither necessary nor sufficient that each of the p matrices $A_1, \ldots, A_p$ does not have b-resonant spectrum. This can be seen from very simple examples already.

(ii) If A does not have b-resonant spectrum then neither does A^l for any $l \geq 2$. Since the converse is not true in general, Theorem 4 constitutes a (slight) generalisation of [4, Thm.3.3] even in the autonomous case.

Requiring that (A_n) in (3) be periodic clearly is restrictive. Beyond periodicity, however, a non-resonance condition *on average* will typically not suffice to guarantee the generation of Benford sequences, even if the nonautonomy of (A_n) is quite mild.

Example 8. Pick $0 < \varepsilon < 1$, let

$$a_n = 2^{\frac{1}{2\pi} + \varepsilon\left(\sin n - \sin(n-1)\right)}, \qquad n \in \mathbb{N},$$

and consider (3) with $d = 1$ and $A_n = (a_n)$. Obviously, $n \mapsto A_n$ is *almost periodic* and, uniformly in k,

$$\frac{1}{n}\sum\nolimits_{l=k}^{k+n-1} \log_b a_l = \left(\frac{1}{2\pi} + \frac{\varepsilon}{n}\sin(k+n-1) - \frac{\varepsilon}{n}\sin(k-1)\right)\log_b 2 \to \frac{1}{2\pi}\log_b 2 \qquad \text{as } n \to \infty .$$

Thus, if $b = 2^m$ for some $m \in \mathbb{N}$, then (A_n) does — in a rather strong average sense — not have b-resonant spectrum, yet from

$$\log_b |x_n| = \log_b |a_n \cdot \ldots \cdot a_1 x_0| = \frac{n}{2m\pi} + \frac{\varepsilon}{m}\sin n + \log_b |x_0| , \qquad n \in \mathbb{N},$$

it is easy to deduce (see for instance [16, Exc.2.7]) that $(\log_b |x_n|)$ is not u.d. mod 1, and hence $(x_n)^*$ is neither finite nor b-Benford.

Without further assumptions, therefore, Theorem 4 does not even generalise to almost periodic sequences (A_n). Formulating appropriate assumptions is comparatively easy in the one-dimensional case, and various results guaranteeing the generation of Benford sequences for (at least) almost all x_0 have been discussed in [6, 11]. In the higher-dimensional setting, i.e. for $d \geq 2$, the situation is more complicated as the information provided by standard tools like e.g. exponential dichotomies and Lyapunov exponents will often be inconclusive as to whether (3) generically generates Benford sequences, as by Theorem 4 it does in the periodic case. The following theorem utilises a shadowing technique to derive conditions under which one-dimensional projections of sequences (x_n) generated by (3) are b-Benford (cf. [5, 6, 22]).

Theorem 9. *Assume that $A_n = A + B_n$ holds in (3) for all $n \in \mathbb{N}$, where A does not have b-resonant spectrum and, with $\beta_A = \|A\| \max\left\{1, \|A^{-1}\|\right\}$,*

$$\sum\nolimits_{n=1}^{\infty} \beta_A^n \|B_n\| < \infty . \tag{4}$$

Then, for every $c \in \mathbb{R}^d$, the sequence $\left(\langle c, x_n\rangle\right)$ is b-Benford provided that it is unbounded.

Proof. Note first that the matrix A, not having b-resonant spectrum, is invertible. Hence $\beta_A \geq 1$ is finite, and (4) implies $\lim_{n\to\infty} \|A\|^n \|B_n\| = 0$. From

$$x_n = (A + B_n) \cdot \ldots \cdot (A + B_1)x_0 = A^n(I + A^{-n}B_n A^{n-1}) \cdot \ldots \cdot (I + A^{-1}B_1)x_0 , \qquad n \in \mathbb{N},$$

it follows that

$$\|x_n\| \le \|A\|^n \prod_{l=1}^n \left(1 + \|A^{-1}\|^l \|B_l\| \|A\|^{l-1}\right) \|x_0\| \le D_1 \|A\|^n , \qquad n \in \mathbb{N} ,$$

with an appropriate constant $D_1 > 0$. Since there is nothing to prove otherwise, assume w.l.o.g. $\|A\| > 1$. Let P denote the (linear) projection onto the stable subspace $E_s = \oplus_{|\lambda|<1} \ker(A - \lambda I)^d$ of A along the complementary unstable subspace. With $\delta = \min\left\{|\lambda - z| : \lambda \in \sigma^+(A), z \in S^1\right\} > 0$ and any κ satisfying $(1+\delta)^{-1} < \kappa < 1$ there exists a constant $K > 0$ such that, for all $n \in \mathbb{N}_0$,

$$\max\left\{\|A^n P\|, \|A^{-n}(I-P)\|\right\} \le K\kappa^n .$$

Since $\|B_n x_{n-1}\| \to 0$ as $n \to \infty$, the point

$$\xi = x_0 + \sum_{l=1}^\infty A^{-l}(I-P)B_l x_{l-1}$$

is well-defined. Moreover,

$$\begin{aligned} x_n - A^n \xi &= A^n x_0 + \sum_{l=1}^n A^{n-l} B_l x_{l-1} - A^n x_0 - \sum_{l=1}^\infty A^{n-l}(I-P)B_l x_{l-1} \\ &= \sum_{l=1}^n A^{n-l} P B_l x_{l-1} - \sum_{l=1}^\infty A^{-l}(I-P) B_{l+n} x_{l+n-1} , \end{aligned}$$

and therefore, for all $1 \le m \le n$,

$$\begin{aligned} \|x_n - A^n\xi\| &\le \sum_{l=1}^m K\kappa^{n-l}\|B_l x_{l-1}\| + \sum_{l=m+1}^n K\kappa^{n-l}\|B_l x_{l-1}\| + \sum_{l=1}^\infty K\kappa^l \|B_{l+n}x_{l+n-1}\| \\ &\le K\frac{\kappa^{n-m}}{1-\kappa} \sup_{l\in\mathbb{N}} \|B_l x_{l-1}\| + K\frac{1+\kappa}{1-\kappa} \sup_{l \ge m+1} \|B_l x_{l-1}\| , \end{aligned}$$

showing that $\|x_n - A^n\xi\| \to 0$ as $n \to \infty$.

Assume now that $\big(\langle c, x_n\rangle\big)$ is unbounded. Then $\big(\langle c, A^n\xi\rangle\big)$ is unbounded as well and, since A does not have b-resonant spectrum, even b-Benford [4, Thm.3.3]. As a consequence, with $D = 2\|c\| \max\left(\frac{1}{\log b}, \sup_{n\in\mathbb{N}} \|x_n - A^n\xi\|\right)$ the set $N_D = \{n \in \mathbb{N}_0 : |\langle c, A^n\xi\rangle| \le D\}$ has density zero, that is, $\lim_{n\to\infty} \frac{1}{n}\#\big(N_D \cap \{0, \ldots, n-1\}\big) = 0$. For all $n \notin N_D$ the elementary estimate $|\log(1+x)| \le 2|x|$, valid whenever $|x| \le \frac{1}{2}$, implies that

$$\begin{aligned} \big|\log_b |\langle c, x_n\rangle| - \log_b |\langle c, A^n\xi\rangle|\big| &= \left|\log_b \left|1 + \frac{\langle c, x_n\rangle - \langle c, A^n\xi\rangle}{\langle c, A^n\xi\rangle}\right|\right| \le \frac{2}{D\log b}|\langle c, x_n\rangle - \langle c, A^n\xi\rangle| \\ &\le \|x_n - A^n\xi\| \ \to\ 0 , \end{aligned}$$

and hence shows that $\big(\langle c, x_n\rangle\big)$ is also b-Benford. □.

Example 10. (i) Let (A_n) be given as

$$A_n = \begin{pmatrix} \rho_n & 1+\rho_n \\ 1+\rho_n & 2+3\rho_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} + \rho_n \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = A + \rho_n B , \qquad n \in \mathbb{N} . \tag{5}$$

The matrix A does not have b-resonant spectrum for any b, and $\beta_A = 3 + 2\sqrt{2}$. Consequently, Theorem 9 guarantees that every unbounded sequence $\big(\langle c, x_n\rangle\big)$ with (x_n) generated by (3) and (5) is Benford provided that $\sum_{n=1}^\infty \beta_A^n |\rho_n| < \infty$. This condition can be relaxed considerably. Indeed, a straightforward calculation shows that, for all $c, x_0 \in \mathbb{R}^2$,

$$\langle c, x_n\rangle = D_1\big(1+\sqrt{2}\big)^n \prod_{l=1}^n \big(1 + \rho_l\sqrt{2}\big) + D_2\big(1-\sqrt{2}\big)^n \prod_{l=1}^n \big(1 - \rho_l\sqrt{2}\big) ,$$

where the constants D_1, D_2 depend linearly on c, x_0. The much weaker condition $\lim_{n\to\infty} \rho_n = 0$, therefore, is sufficient to ensure that for *every* $c, x_0 \in \mathbb{R}^2$ the sequence $\big(\langle c, x_n\rangle\big)^*$ is either finite or Benford.

(ii) Consider the sequence (A_n) defined according to

$$A_n = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix} + \rho_n \begin{pmatrix} 2 & 1-\sqrt{5} \\ 1+\sqrt{5} & -2 \end{pmatrix} = A + B_n \, , \qquad n \in \mathbb{N} \, , \tag{6}$$

so that $\beta_A = \sqrt{5} + 1 \approx 3.24$. Again, Theorem 9 guarantees that every unbounded sequence $\big(\langle c, x_n \rangle\big)$ is Benford whenever $\sum_{n=1}^{\infty} \beta_A^n |\rho_n| < \infty$. As before, this condition is not optimal. An explicit calculation shows that

$$\langle c, x_n \rangle = D_1 \big(1+\sqrt{5}\big)^n + D_2 \big(1+\sqrt{5}\big)^n \sum\nolimits_{l=1}^{n} \left(\frac{\sqrt{5}-3}{2} \right)^l \rho_l + D_3 \big(1-\sqrt{5}\big)^n \, ,$$

with the constants D_1, D_2, D_3 being linear functionals of c and x_0. With this it is easy to check that for every unbounded sequence $\big(\langle c, x_n \rangle\big)$ to be Benford it is enough to require that $\sum_{n=1}^{\infty} \beta^n |\rho_n| < \infty$ where $\beta = \sqrt{5} - 1 = \beta_A - 2 \approx 1.24$.

Remarks 11. (i) Condition (4) depends on the chosen norm, but the generation of Benford sequences via (3) is of course independent of the latter. To reformulate (4) without referring to any norm, for every regular matrix A define $\widehat{\beta}_A = r_\sigma(A) \max\{1, r_\sigma(A^{-1})\}$. Obviously, $\widehat{\beta}_A \le \beta_A$ and, as in [14, Prop.1.2.2], it is easy to see that, given $\varepsilon > 0$, there exists a norm on $\mathbb{R}^d$ (or $\mathbb{C}^d$) for which $\widehat{\beta}_A \le \beta_A \le \widehat{\beta}_A + \varepsilon$. Thus (4) can be replaced by

$$\sum\nolimits_{n=1}^{\infty} \beta^n \|B_n\| < \infty \quad \text{for some } \beta > \widehat{\beta}_A \, . \tag{7}$$

In view of (7) it is natural to ask whether Theorem 9 remains true with β_A in (4) replaced by $\widehat{\beta}_A$. While this is obviously the case for diagonalisable matrices A (and hence generically in $\mathbb{C}^{d \times d}$), no overall answer to this question is yet known to the authors.

(ii) Conditions (4) and (7) which force $(\|B_n\|)$ to decay exponentially* are obviously quite restrictive. Contrary to what the above examples may suggest, however, the assertion in Theorem 9 will not hold in general if for instance (4) is replaced by

$$\sum\nolimits_{n=1}^{\infty} \beta^n \|B_n\| < \infty \tag{8}$$

for some $1 < \beta < \beta_A$. Without further assumptions, therefore, condition (4) is best possible. For a concrete example consider

$$A_n = \begin{pmatrix} \varphi & \rho_n \\ 0 & \varphi \end{pmatrix} = \begin{pmatrix} \varphi & 0 \\ 0 & \varphi \end{pmatrix} + \rho_n \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = A + B_n \, , \qquad n \in \mathbb{N} \, ,$$

where φ is the golden mean (i.e. $\varphi > 1$ and $\varphi^2 = \varphi + 1$). Clearly, $\beta_A = \varphi$, and with $\|B_n\| = \rho_n = n\varphi^{-n}$ condition (8) holds for all $|\beta| < \varphi$. Nevertheless, it follows from

$$A_n \cdot \ldots \cdot A_1 = \begin{pmatrix} \varphi^n & \varphi^{n-1} \sum_{l=1}^{n} \rho_l \\ 0 & \varphi^n \end{pmatrix} , \qquad n \in \mathbb{N} \, ,$$

that choosing $c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $x_0 = \begin{pmatrix} \sum_{l=1}^{\infty} \rho_l \\ -\varphi \end{pmatrix}$ yields

$$\langle c, x_n \rangle = \varphi^n \sum\nolimits_{l=n+1}^{\infty} \rho_l = \sum\nolimits_{l=1}^{\infty} (l+n) \varphi^{-l} = \varphi^3 + n\varphi \, , \qquad n \in \mathbb{N} \, .$$

*When computed using the Euclidean norm, $\beta_A > 1$ unless αA is unitary for some number α with $|\alpha| \ge 1$; the latter case clearly is irrelevant here.

Hence the sequence $\big(\langle c, x_n\rangle\big)$ is unbounded yet not b-Benford for any b.

(iii) The fact that condition (4) is in some sense best possible in general does of course not rule out the possibility of substantial improvement in special cases. For instance, if $AB_n = B_nA$ for all n then β_A in (4) can be replaced by $\max\{\|A\|, 1\}$, and the resulting condition is best possible in the sense of (ii). (Note that $AB_n \equiv B_nA$ holds only for the first of the two systems in Example 10.) Similarly, the conditions $\lim_{n\to\infty} \rho_n = 0$ and $\sum_{n=1}^\infty (\sqrt{5}-1)^n|\rho_n| < \infty$ in Example 10(i) and (ii), respectively, are optimal in that for every $\varepsilon > 0$ the weaker conditions $\overline{\lim}_{n\to\infty}|\rho_n| < \varepsilon$ and $\sum_{n=1}^\infty (\sqrt{5}-1-\varepsilon)^n|\rho_n| < \infty$ are in general not sufficient for Theorem 9 to remain correct. Finally, if $d = 1$ then $\lim_{n\to\infty} b_n = 0$, together with $\log_b |a| \notin \mathbb{Q}$, is enough to guarantee that $(x_n)^*$ is either finite or b-Benford.

(iv) If, in the setting of Theorem 9, the sequence $\big(\langle c, x_n\rangle\big)^*$ is bounded it may or may not be b-Benford. In Example 10(ii) for instance, choosing $\rho_n = (-2)^{-n}$ as well as

$$c = \begin{pmatrix} \sqrt{5}-1 \\ 2 \end{pmatrix}, \qquad x_0 = \begin{pmatrix} 6-2\sqrt{5} \\ \sqrt{5}-2 \end{pmatrix},$$

yields

$$\langle c, x_n\rangle = 10(\sqrt{5}-2)\left(\frac{\sqrt{5}-1}{2}\right)^n \to 0,$$

a bounded sequence which is strict Benford. On the other hand, choosing $\rho_n = (-\rho)^{-n}$ with $\rho > \sqrt{5}-1$ such that $\log_b \frac{\sqrt{5}-1}{\rho}$ is rational will, for appropriate c, x_0, yield a sequence $\big(\langle c, x_n\rangle\big)^*$ which is bounded yet not b-Benford.

(v) The systems (3) covered by Theorem 9 have the property that (A_n) converges rapidly and therefore are *asymptotically autonomous* in a fairly strong sense (see [19] for details on continuous-time asymptotically autonomous systems). It is straightforward to formulate and prove an analogous result for asymptotically *periodic* systems (3) in the spirit of Theorem 4; details are left to the reader. Such a result shows for instance that for

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & 3+(-1)^n \end{pmatrix} + B_n, \qquad n \in \mathbb{N},$$

every unbounded sequence $\big(\langle c, x_n\rangle\big)^*$ is Benford provided that $\sum_{n=1}^\infty \beta_A^n \|B_n\| < \infty$, where $A = A_2A_1$.

(vi) The results of this section can easily be extended to *complex* matrices A_n, so as to yield an analogue of Theorem 9 which gives conditions ensuring that the sequences $\big(\Re\langle c, x_n\rangle\big)^*$ and $\big(\Im\langle c, x_n\rangle\big)^*$ are Benford whenever unbounded. For the (elementary) details of this extension, the reader may wish to consult [4, Rem.3.8].

(vii) The argument in the proof of Theorem 9 may be considered a variant of the basic shadowing lemma [6, Thm.2.5] in that properties of (x_n) are established by means of the corresponding properties of $(A^n\xi)$. It is well-known that for an invertible hyperbolic matrix A many signatures of hyperbolicity, e.g. the shadowing property and the existence of an exponential dichotomy, carry over from $x_n = Ax_{n-1}$ to (3) with $A_n = A + B_n$ under *much* weaker assumptions on $(\|B_n\|)$; in many cases it is enough to assume that $\sup_{n\in\mathbb{N}} \|B_n\|$ is sufficiently small, see [22] for details. As evidenced by (ii), however, these weaker assumptions are generally not sufficient to ensure the persistence of finer statistical properties such as BL.

4 Some non-linear examples

Two classes of non-linear nonautonomous systems on $\mathbb{R}^d$ will be discussed in this section. Both classes are fairly specific, and the results presented should be considered mainly as an invitation to further study the distribution of nonautonomous orbits.

First consider *linearly dominated* systems

$$x_n = A_n x_{n-1} + f_n(x_n)\,, \qquad n \in \mathbb{N}\,, \tag{9}$$

where the functions $f_n : \mathbb{R}^d \to \mathbb{R}^d$ are uniformly bounded, that is, $\sup_{x\in\mathbb{R}^d} \|f_n(x)\| \le C$ for all $n \in \mathbb{N}$ and some constant $C \ge 0$. The following result extends Theorem 9; it also significantly generalises [4, Thm.4.1].

Theorem 12. *Assume that $A_n = A + B_n$ holds in* (9) *for all $n \in \mathbb{N}$, where A does not have b-resonant spectrum, and (B_n) satisfies* (4). *Then, for every $c \in \mathbb{R}^d$ the sequence $\bigl(\langle c, x_n\rangle\bigr)^*$ is b-Benford provided that it is unbounded.*

Proof. Since discarding finitely many terms of $\bigl(\langle c, x_n\rangle\bigr)$ does not affect the statement, and since (4) entails $\|B_n\| \to 0$, it can be assumed that $\sup_{n\in\mathbb{N}} \|B_n\|$ is so small that $A_n = A + B_n$ is invertible for all n, and also that (A_n) has an exponential dichotomy (see e.g. [22, Lem.2.8]). Let (P_n), $K > 0$, and $0 < \kappa < 1$ denote, respectively, the sequence of projections, constant, and exponent associated with this exponential dichotomy. Thus with

$$\Phi(l,m) = \begin{cases} A_{l-1}\cdot \ldots \cdot A_m & \text{if } l > m\,, \\ I & \text{if } l = m\,, \\ A_l^{-1}\cdot \ldots \cdot A_{m-1}^{-1} & \text{if } l < m\,, \end{cases} \qquad \forall l, m \in \mathbb{N}\,,$$

the invariance condition $\Phi(l,m)P_m = P_l\Phi(l,m)$ holds for all $l, m \in \mathbb{N}$, as do the estimates

$$\|\Phi(l,m)P_m\| \le K\kappa^{l-m} \quad \text{and} \quad \|\Phi(l,m)(I-P_m)\| \le K\kappa^{m-l} \tag{10}$$

whenever $l \ge m$ and $l \le m$, respectively. Given $x_0 \in \mathbb{R}^d$, consider its orbit (x_n) as generated by (9) and let

$$\xi = x_0 + \sum\nolimits_{l=1}^{\infty} \Phi(1, l+1)(I - P_{l+1}) f_l(x_{l-1})\,. \tag{11}$$

It is obvious from (10) and the uniform boundedness of (f_n) that the series in (11) converges absolutely. Moreover,

$$\Phi(n+1,1)\xi = x_n - \sum\nolimits_{l=1}^{n}\Phi(n+1,l+1)P_{l+1}f_l(x_{l-1}) + \sum\nolimits_{l=n+1}^{\infty}\Phi(n+1,l+1)(I-P_{l+1})f_l(x_{l-1})\,,$$

and consequently, for all $n \in \mathbb{N}$,

$$\|\Phi(n+1,1)\xi - x_n\| \le \sum\nolimits_{l=1}^{\infty} CK\kappa^{|n-l|} \le CK\frac{1+\kappa}{1-\kappa}\,.$$

As detailed in the proof of Theorem 9 there exists a point $\widetilde{\xi}$ such that $\|A^n\widetilde{\xi} - \Phi(n+1,1)\xi\| \to 0$ as $n \to \infty$. Thus, if the sequence $\bigl(\langle c, x_n\rangle\bigr)^*$ is unbounded then so is $\bigl(\langle c, A^n\widetilde{\xi}\rangle\bigr)$. The latter is actually b-Benford, and $N_D = \{n \in \mathbb{N}_0 : |\langle c, A^n\widetilde{\xi}\rangle| \le D\}$ has density zero for every $D > 0$. Given $0 < \varepsilon < 1$, choose

$$D \ge 2\|c\|\left(1 + \frac{1}{\varepsilon \log b}\right)\left(CK\frac{1+\kappa}{1-\kappa} + \sup\nolimits_{n\in\mathbb{N}} \|A^n\widetilde{\xi} - \Phi(n+1,1)\xi\|\right)\,.$$

For all $n \notin N_D$ it follows that

$$\bigl|\log_b|\langle c, x_n\rangle| - \log_b|\langle c, A^n\widetilde{\xi}\rangle|\bigr| = \left|\log_b\left|1 + \frac{\langle c, x_n\rangle - \langle c, A^n\widetilde{\xi}\rangle}{\langle c, A^n\widetilde{\xi}\rangle}\right|\right| \le \varepsilon\,,$$

which, together with [4, Lem.2.3], shows that $\bigl(\langle c, x_n\rangle\bigr)^*$ is b-Benford as well. $\square$

Remarks 13. (i) Just as Theorem 9, the above result is best possible in general but can be strengthened considerably in special cases. For instance, the sequence (A_n) given by (5) has an exponential dichotomy whenever $\overline{\lim}_{n\to\infty}|\rho_n| < \sqrt{2}-1$. For this particular (A_n) it follows from Example 10(i) that the condition given there, that is, $\lim_{n\to\infty}\rho_n = 0$ is sufficient to guarantee that for every orbit generated by (9) and every $c \in \mathbb{R}^d$ the sequence $\big(\langle c, x_n\rangle\big)^*$ is strict Benford provided that it is unbounded. Similarly, in Example 10(ii) convergence of $\sum_{n=1}^{\infty}(\sqrt{5}-1)^n|\rho_n|$ is enough for (A_n) as given by (6) to have an exponential dichotomy, and Theorem 12 can be strengthened accordingly.

(ii) Under the conditions of Theorem 12, and similarly to Remark 11(iv), a *bounded* sequence $\big(\langle c, x_n\rangle\big)^*$ may or may not be b-Benford, see [4, Rem.4.2] for simple examples.

The second class of non-linear nonautonomous systems (2) considered here consists of *power-like* maps $T_n : \mathbb{R}^d \to \mathbb{R}^d$, for which each component $T_n^{(j)}$ has a dominant monomial term and can be written as

$$T_n^{(j)}(x) = (x^{(1)})^{(A_n)_{j1}} \cdots (x^{(d)})^{(A_n)_{jd}}\big(1 + f_n^{(j)}(x)\big)\,, \qquad \forall j = 1,\ldots,d; n \in \mathbb{N}\,; \tag{12}$$

here $(A_n)_{jl} \in \mathbb{N}_0$ for all $j, l \in \{1,\ldots,d\}$, and f_n is a C^1 function with $\|f_n(x)\| \to 0$ as $\|x\| \to \infty$. To ensure that T_n is actually power-like, it will be assumed throughout that, for every $n \in \mathbb{N}$,

$$\sum\nolimits_{l=1}^{d}(A_n)_{jl} \ge 1\,, \qquad \forall j = 1,\ldots d\,, \tag{13}$$

that is, each row of A_n contains at least one non-zero entry. Maps of this type have been used to model aspects of economic growth and socio-spatial dynamics (see [9, 18] and the references therein). To analyse the distribution of orbits generated by (2) with maps T_n given by (12) the following definition and lemma will be used.

Definition 14. (i) A sequence (Q_n) of (linear) projections in $\mathbb{R}^d$ is *Diophantine* if for every $h \in \mathbb{Z}^d$ there exists a number $\gamma_h > 0$ such that for each $n \in \mathbb{N}$ either $Q_n h = 0$ or else $\|Q_n h\| \ge \gamma_h$.

(ii) Let the sequence (A_n) of invertible matrices have an exponential dichotomy with associated projections (P_n). Then (A_n) is said to have a *Diophantine exponential dichotomy* if $(I - P_n^\top)$ is Diophantine.

Trivially, (Q_n) is Diophantine whenever $\{Q_n : n \in \mathbb{N}\}$ is finite. Thus for $d = 1$ every sequence of projections is Diophantine, as is every exponential dichotomy. If (Q_n) is a sequence of projections in $\mathbb{R}^2$ let $N = \{n \in \mathbb{N} : \text{rank } Q_n = 1\}$ and, for each $n \in N$, let $\ker Q_n$ be spanned by $(\cos\varphi_n)e_1 + (\sin\varphi_n)e_2$ with a uniquely defined $0 \le \varphi_n < \pi$. Then (Q_n) is Diophantine if and only if $\mathrm{cl}\{\varphi_n : n \in N\}\backslash\{\varphi_n : n \in N\}$ does neither contain $\frac{\pi}{2}$ nor any φ for which $\tan\varphi$ is rational.

Lemma 15. *Let (A_n) be a sequence of invertible matrices all of whose entries are integers. Assume that (A_n) has a Diophantine exponential dichotomy. Then, for (Lebesgue) almost all $y_0 \in \mathbb{R}^d$, the sequence $(A_n \cdot \ldots \cdot A_1 y_0)$ is u.d.* mod 1 *in $\mathbb{R}^d$.*

Proof. Fix $h \in \mathbb{Z}^d\backslash\{0\}$ as well as $a \in \mathbb{R}^d$ and define, for every $n \in \mathbb{N}$,

$$\begin{aligned} I_h(n) &= \frac{1}{n^2}\sum\nolimits_{l,m=1}^{n}\int_{y\in\prod_{j=1}^{d}[a^{(j)},a^{(j)}+1]} e^{2\pi i\langle h,\,(A_l\cdot\ldots\cdot A_1 - A_m\cdot\ldots\cdot A_1)y\rangle}dy \\ &= \frac{1}{n} + \frac{2}{n^2}\#\big\{1 \le l < m \le n : A_{l+1}^\top\cdot\ldots\cdot A_m^\top h = h\big\}\,. \end{aligned}$$

To find an upper bound on $I_h(n)$, assume that $m > l$ and $A_{l+1}^\top\cdot\ldots\cdot A_m^\top h = \Phi(m+1,l+1)^\top h = h$. If $P_{m+1}^\top h = h$ then, with the notation of the proof of Theorem 12,

$$\|h\| = \|\Phi(m+1,l+1)^\top P_{m+1}^\top h\| = \|(\Phi(m+1,l+1)P_{l+1})^\top h\| \le K\kappa^{m-l}\|h\|\,,$$

so that $m - l \le D_1$ with an appropriate positive constant D_1 not depending on l. If, on the other hand, $P_{m+1}^\top h \neq h$ then, since (A_n) has a Diophantine exponential dichotomy,

$$\begin{aligned} 0 < \gamma_h \le \|h - P_{m+1}^\top h\| &= \|(I - P_{m+1}^\top)\Phi(l+1, m+1)^\top \Phi(m+1, l+1)^\top h\| \\ &= \left\| \bigl(\Phi(l+1, m+1)(I - P_{m+1})\bigr)^\top \Phi(m+1, l+1)^\top h \right\| \le K\kappa^{m-l}\|h\| , \end{aligned}$$

so $m - l \le D_2$ with another constant $D_2 > 0$. Overall, $nI_h(n) \le 1 + 2\max(D_1, D_2)$, implying in turn that $\sum_{n=1}^\infty I_h(n)/n < \infty$. By (the d-dimensional version of) [16, Thm.4.2] the sequence $(\Phi(n+1, 1)y_0)$ is u.d. mod 1 in $\mathbb{R}^d$ for almost all y_0 in the unit cube $\prod_{j=1}^d [a^{(j)}, a^{(j)} + 1]$. Since $\mathbb{R}^d$ is a countable union of unit cubes, the proof is complete. □

To apply Lemma 15 to nonautonomous orbits generated by (12) observe that (13) forces each component of T_n to vanish on some coordinate hyperplane $x^{(j)} \equiv 0$, and Benford sequences may be generated only from initial points sufficiently far away from these hyperplanes. Therefore, for every $\alpha > 0$, define the cone $C_\alpha = \{x \in \mathbb{R}^d : \min_{j=1}^d |x^{(j)}| \ge \alpha\}$. Also, let $D_x f_n$ denote the Jacobian of f_n at x, that is $(D_x f_n)_{jk} = \frac{\partial}{\partial x^{(k)}} f_n^{(j)}(x)$. The following is a natural generalisation of [4, Thm.4.5].

Theorem 16. *Let (A_n) be a sequence of invertible matrices all of whose entries are non-negative integers satisfying* (13). *Assume that (A_n) has a Diophantine exponential dichotomy and that $(\frac{1}{n}\log\|A_n\|)$ is bounded. Furthermore assume that the sequence of C^1 functions (f_n) satisfies $\sup_{x\in C_\alpha}\sup_{n\in\mathbb{N}}\|f_n(x)\| \to 0$ as $\alpha \to \infty$, as well as $\sup_{x\in C_\alpha}\sup_{n\in\mathbb{N}}\|x\|^{1+\varepsilon}\|D_x f_n\| < \infty$ for some $\varepsilon > 0$, $\alpha > 0$. Then, for α sufficiently large and almost all $x_0 \in C_\alpha$, each component of $O_T(x_0)$ is Benford.*

Proof. Let $\mathbf{1} \in \mathbb{R}^d$ denote the vector all of whose components equal 1. Throughout the subsequent argument, scalar functions and inequalities applied to elements of $\mathbb{R}^d$ are to be read coordinate-wise, e.g., $\log|x|$ is understood to be $(\log|x^{(1)}|, \dots, \log|x^{(d)}|) \in \mathbb{R}^d$, and $x \ge \alpha\mathbf{1}$ means $x^{(j)} \ge \alpha$ for all $j = 1, \dots, d$. Also, the notation of the proofs of Theorem 12 and Lemma 15 will be used.

Observe first that under the stated assumptions on (A_n) there exists a number $0 < D \le 1$ such that

$$A_n \cdot \ldots \cdot A_1\mathbf{1} \ge D\kappa^{-n}\mathbf{1} , \qquad \forall n \in \mathbb{N}_0 .$$

Using this, it will now be shown that $T^n(C_{\beta^{2/D}}) \subset C_\beta$ holds for all $n \in \mathbb{N}_0$, provided that β is sufficiently large. To this end let $y_n = \log|x_n|$ so that

$$y_n = A_n y_{n-1} + \log|1 + f_n(x_{n-1})| , \qquad n \in \mathbb{N} .$$

Let $\gamma = \frac{D}{2K\sqrt{d}}\frac{1-\kappa}{1+\kappa}$ and choose $\beta > e^{D(D+1)}$ so large that $\|f_n(x)\| \le \min\{\frac{1}{2}, 1 - e^{-\gamma}\}$ holds for all $x \in C_\beta$ and $n \in \mathbb{N}$. Assume that $x_0 \in C_{\beta^{2/D}}$ as well as $x_l \in C_\beta$ for all $l = 1, \dots, n$. Then $\left|\log\bigl(1 + f_l(x_{l-1})\bigr)\right| \le \gamma\mathbf{1}$ for all $l = 1, \dots, n+1$, and consequently

$$\begin{aligned} y_{n+1} &= \Phi(n+2, 1)y_0 + \sum\nolimits_{l=1}^{n+1} \Phi(n+2, l+1)\log|1 + f_l(x_{l-1})| \\ &\ge \frac{2}{D}(\log\beta)\Phi(n+2, 1)\mathbf{1} - \gamma\sum\nolimits_{l=1}^{n+1} \Phi(n+2, l+1)\mathbf{1} . \end{aligned}$$

Setting $\eta_0 = (\frac{2}{D}\log\beta)\mathbf{1}$ and, inductively for $l \in \mathbb{N}$, $\eta_l = A_l\eta_{l-1} - \gamma\mathbf{1}$ therefore yields $y_{n+1} \ge \eta_{n+1}$. To find a lower bound for η_{n+1}, let

$$\widetilde{\eta} = \eta_0 - \gamma\sum\nolimits_{l=1}^\infty \Phi(1, l+1)(I - P_{l+1})\mathbf{1} ,$$

and observe that

$$\|\Phi(n+2,1)\widetilde{\eta}-\eta_{n+1}\| = \gamma\left\|\sum\nolimits_{l=1}^{n+1}\Phi(n+2,l+1)P_{l+1}\mathbf{1}-\sum\nolimits_{l=n+2}^{\infty}\Phi(n+2,l+1)(I-P_{l+1})\mathbf{1}\right\|$$
$$\leq \gamma K\sqrt{d}\sum\nolimits_{l=1}^{\infty}\kappa^{|n+1-l|} \leq \gamma K\sqrt{d}\frac{1+\kappa}{1-\kappa}=\frac{D}{2}<D\,,$$

which in turn implies that

$$\begin{aligned} y_{n+1} \geq \eta_{n+1} > \Phi(n+2,1)\widetilde{\eta}-D\mathbf{1} > \Phi(n+2,1)(\eta_0-D\mathbf{1})-D\mathbf{1} \\ \geq \kappa^{-(n+1)}(2\log\beta-D^2)\mathbf{1}-D\mathbf{1} \geq (2\log\beta-D^2-D) > (\log\beta)\mathbf{1}\,. \end{aligned} \tag{14}$$

Hence $|x_{n+1}|>\beta\mathbf{1}$, that is, $x_{n+1}\in C_\beta$. Thus, setting $\alpha=\beta^{2/D}$ yields $x_n\in C_{\alpha^{D/2}}$ for all $n\in\mathbb{N}$ whenever $x_0\in C_\alpha$.

Fix now a base b and, for every $\beta>1$, let $C_\beta^+=C_\beta\cap(\mathbb{R}^+)^d$. Obviously, C_β has 2^d connected components, and the map $\psi_b: x\mapsto \log_b|x|$ maps each component diffeomorphically onto $C^+_{\log_b\beta}$. Therefore, the inverse of ψ_b has 2^d branches, and the symbol ψ_b^{-1} will be used to denote any one of them. With α sufficiently large define the map

$$H_b: z\mapsto z+\sum\nolimits_{l=1}^{\infty}\Phi(1,l+1)(I-P_{l+1})\log_b\bigl(1+f_l\circ T^{l-1}\circ\psi_b^{-1}(z)\bigr)\,, \tag{15}$$

which is well-defined for all $z\in C^+_{\log_b\alpha}$. Given $x_0\in C_\alpha$ let $z_l=\psi_b(x_l)$ for all $l\in\mathbb{N}_0$ and choose the (unique) branch ψ_b^{-1} satisfying $x_0=\psi_b^{-1}(z_0)$. Note that (14) in particular implies

$$x_n\in C_{e^{\kappa^{-(n+1)}(D\log\alpha-D^2)-D}}\,,\qquad \forall n\in\mathbb{N}_0\,,$$

which, when combined with (15), yields

$$\|\Phi(n+1,1)H_b(z_0)-z_n\|\leq K\sum\nolimits_{l=1}^{\infty}\kappa^{|n-l|}\bigl\|\log_b\bigl(1+f_l(x_{l-1})\bigr)\bigr\| \;\to\; 0 \quad \text{as } n\to\infty\,.$$

Therefore (z_n) is u.d. mod 1 in $\mathbb{R}^d$ if and only if $\bigl(\Phi(n+1,1)H_b(z_0)\bigr)$ is. By Lemma 15, the latter is the case whenever $H_b(z_0)\in\mathbb{R}^d\backslash N_b$ with N_b denoting an appropriate set of measure zero. A straightforward albeit lengthy calculation, using termwise differentiation of (15) and the fact that $(\frac{1}{n}\log\|A_n\|)$ is bounded, shows that H_b is a local diffeomorphism, and $\sup_{z\in C_\beta^+}\|H_b(z)-z\|\to 0$ as $\beta\to\infty$. Thus $H_b^{-1}\bigl(N_b\cap H_b(C^+_{\log_b\alpha})\bigr)$ has measure zero as well, and as a consequence (z_n) is u.d. mod 1 in $\mathbb{R}^d$ for almost all $z_0\in C^+_{\log_b\alpha}$.

Overall, the above argument shows that, for every base b, there exists a set $M_b\subset C_\alpha$ such that each component of $O_T(x_0)$ is b-Benford provided that $x_0\in C_\alpha\backslash M_b$. Setting $M=\bigcup_{b\geq 2}M_b\subset C_\alpha$ yields $\lambda^d(M)=0$, and $O_T^{(j)}(x_0)$ is Benford for all $j=1,\ldots,d$ whenever $x_0\in C_\alpha\backslash M$. □

Corollary 17. *Let the maps (T_n) according to (12) satisfy all assumptions of Theorem 16. Then, for every sufficient large α, there exists an uncountable dense set $E\subset C_\alpha$ such that for every $x_0\in E$ no component of $O_T(x_0)$ is Benford.*

Proof. Let $(k_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{Z}^d$. With the notation used in the proof of Theorem 16, pick a natural number $N\geq 8$ so large that $\kappa^N<(1+8K\max_{n\in\mathbb{N}}\|k_n\|)^{-1}$, and consider the point

$$\zeta=\sum\nolimits_{l=1}^{\infty}\Phi(1,n_lN^2+1)(I-P_{n_lN^2+1})k_l\,, \tag{16}$$

where (n_l) is any strictly increasing sequence of natural numbers. For every $n=m_1+m_2N^2$ with $N\leq m_1\leq N^2-N$ and $m_2\in\mathbb{N}_0$ a simple calculation shows that

$$\min\nolimits_{h\in\mathbb{Z}^d}\|\Phi(n+1,1)\zeta-h\|\leq 2K\max\nolimits_{l\in\mathbb{N}}\|k_l\|\frac{\kappa^N}{1-\kappa^N}<\frac{1}{4}\,,$$

and consequently

$$\overline{\lim}_{n\to\infty}\frac{\#\{l<n:\min_{h\in\mathbb{Z}^d}\|\Phi(l+1,1)\zeta-h\|<\frac{1}{4}\}}{n}\geq\frac{N^2-2N+1}{N^2}>1-\frac{2}{N}>\frac{3}{4}\,.$$

Obviously this implies that no component of $(\Phi(n+1,1)\zeta)$ is u.d. mod 1. To produce uncountably many *different* points ζ by means of (16) the sequences (n_l) and (k_n) are chosen inductively as follows: Assume that numbers $1=n_1<n_2<\ldots<n_L$ and vectors $\widehat{k}_1,\ldots,\widehat{k}_L\in\{e_1,\ldots,e_d\}$ have already been chosen such that

$$\widehat{k}_i\notin \operatorname{range} P_{n_iN^2+1}\,,\qquad \forall i=1,\ldots,L\,.$$

(Such a choice is possible because $\operatorname{rank} P_n\leq d-1$ for all n.) Choose now $n_{L+1}>n_L$ so large that

$$2K\frac{\kappa^{n_{L+1}N^2}}{1-\kappa^{N^2}}<\|\Phi(1,n_LN^2+1)(I-P_{n_LN^2+1})\widehat{k}_L\|\,;$$

also choose $\widehat{k}_{L+1}\in\{e_1,\ldots,e_d\}\backslash\operatorname{range} P_{n_{L+1}N^2+1}$. Finally, for each n let k_n be either $\widehat{k}_n$ or 0. Any two among the uncountably many different sequences (k_n) thus constructed give rise to different points ζ in (16).

Fix now a base b and, for sufficiently large α, take $h\in\mathbb{Z}^d\cap C^+_{2+\log_b\alpha}$. The above construction provides uncountably many different points $h+\zeta\in C^+_{1+\log_b\alpha}$ for which no component of the sequence $\big(\Phi(n+1,1)(h+\zeta)\big)$ is u.d. mod 1. The latter is also true if $h+\zeta$ is replaced by any element of $\mathbb{Q}^d$, a dense subset of $\mathbb{R}^d$. Overall, there exists an uncountable dense subset $F\subset H_b(C^+_{\log_b\alpha})$ such that no component of $(\Phi(n+1,1)z)$ is u.d. mod 1 whenever $z\in F$. Setting

$$E=\{x\in C_\alpha: H_b(\log_b|x|)\in F\cap C^+_{\log_b\alpha}\}$$

therefore completes the proof. □

Remarks 18. (i) For concrete examples it may be desirable to multiply each component $T_n^{(j)}$ in (12) by some non-zero number $c_n^{(j)}$. As long as the additional nonautonomy thus introduced does not dominate the overall dynamics, Theorem 16 will still hold. More concretely, under the assumption that $(\log|c_n^{(j)}|)_{n\in\mathbb{N}}$ is bounded for all $j=1,\ldots,d$, Theorem 16 remains unchanged while its proof requires only one fairly obvious minor modification; details are left to the reader, see also [5].

(ii) Even in the autonomous case, that is for T_n not depending on n, and even when slightly generalised according to (i), Theorem 16 does not cover *all* d-variate polynomial maps unless $d=1$, see [4].

Example 19. Let $A_1,\ldots,A_L$ be a finite family of invertible hyperbolic matrices with non-negative integer entries, that is, $A_l\in\mathbb{N}_0^{d\times d}$ and $\sigma^+(A_l)\cap(S^1\cup\{0\})=\emptyset$ for all $l=1,\ldots,L$. Assume that $A_lA_m=A_mA_l$ for all l,m and also that (A_{ω_n}) has a Diophantine exponential dichotomy for every $\omega=(\omega_n)\in\{1,\ldots,L\}^{\mathbb{N}}$. (A simple condition guaranteeing the latter is that the stable and unstable generalised eigenspaces associated with the matrix A_l,

$$E_{s,l}=\oplus_{|\lambda|<1}\ker\,(A_l-\lambda I)^d\,,\quad E_{u,l}=\oplus_{|\lambda|>1}\ker\,(A_l-\lambda I)^d\,,\qquad l=1,\ldots,L\,,$$

do actually not depend on l.) For each ω then (A_{ω_n}) satisfies the assumptions of Theorem 16.

For a concrete example, consider the two bi-variate polynomial maps

$$T_a:\begin{pmatrix}u\\v\end{pmatrix}\mapsto\begin{pmatrix}3u^2v^2+4\\5u^2v^4-6v^2+7\end{pmatrix}\quad\text{and}\quad T_b:\begin{pmatrix}u\\v\end{pmatrix}\mapsto\begin{pmatrix}v\\2uv\end{pmatrix}\,,$$

which are of the form (12) and whose associated matrices

$$A_a = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} \quad \text{and} \quad A_b = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

satisfy all the conditions mentioned above. Consequently, for *every* sequence $\omega = (\omega_n)$ on the two symbols a, b, there exists $\alpha > 0$ such that for almost all $x \in C_\alpha$ both components of the nonautonomous orbit $\big(T_{\omega_n} \circ \ldots \circ T_{\omega_1}(x)\big)$ are Benford. There is, however, also an uncountable dense set $E \subset C_\alpha$ such that for every $x \in E$ no component of $O_T(x)$ is Benford.

Acknowledgements

The first author was partially supported by a Humboldt Research Fellowship; he wishes to thank B. Martin for a helpful comment. The second author was supported by the Emmy Noether Program funded by the DFG.

References

[1] B. Aulbach, S. Elaydi, and K. Ziegler, Asymptotic solutions of a discrete Schrödinger equation arising from a Dirac equation with random mass. In: B. Aulbach & S. Elaydi & G. Ladas (eds.), *New Progress in Difference Equations*, Taylor and Francis, London, 349-358 (2004).

[2] B. Aulbach and T. Wanner, Invariant foliations and decoupling of nonautonomous difference equations, *J. Difference Equ. Appl.* **9** (2003), 459-472.

[3] F. Benford, The law of anomalous numbers, *Proceedings of the American Philosophical Society* **78** (1938), 551–572.

[4] A. Berger, Multi-dimensional Dynamical Systems and Benford's Law, *Discrete Contin. Dyn. Syst.* **13** (2005), 219–237.

[5] A. Berger, Benford's Law in power-like dynamical systems, *Stoch. Dyn.* **5** (2006), 587–607.

[6] 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.

[7] A. Berger and T.P. Hill, Newton's Method obeys Benford's Law, to appear in *Amer. Math. Monthly.*

[8] J. Brown and R. Duncan, Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences, *Fibonacci Quarterly* **8** (1970), 482–486.

[9] D.S. Dendrinos and M. Sonis, *Chaos and Socio-Spatial Dynamics*, Springer, New York Berlin Heidelberg (1990).

[10] P. Diaconis, The distribution of leading digits and uniform distribution mod 1, *Ann. Probab.* **5** (1979), 72–81.

[11] G. Gottwald and M. Nicol, On the nature of Benford's Law, *Phys. A* **303** (2001), 387–396.

[12] T.P. Hill, A statistical derivation of the significant-digit law, *Statistical Science* **10** (1996), 354–363.

[13] Z. Huang, and S. Solomon, Stochastic multiplicative processes for financial markets, *Phys. A* **306** (2002), 412–422.

[14] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge (1995).

[15] A. Kontorovich and S. Miller, Benford's Law, values of L-functions and the 3x+1 problem, *Acta Arithmetica* **120** (2005), 269-297.

[16] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York (1974).

[17] E. Ley, On the Peculiar Distribution of the U.S. Stock Indices Digits, *Amer. Stat.* **50** (1996), 311–313.

[18] H.-W. Lorenz, *Nonlinear Dynamical Economics and Chaotic Motion* (Second Edition), Springer, Berlin Heidelberg New York (1993).

[19] L. Markus, Asymptotically autonomous differential systems, *Annals of Mathematics Studies* **3** (1956), 17–29.

[20] S. Miller and M.J. Nigrini, Order Statistics and Shifted Almost Benford Behavior, preprint (2006).

[21] M.J. Nigrini, *Digital Analysis Using Benford's Law: Tests and Statistics for Auditors* (Second Edition), Global Audit Publications, Vancouver (2000).

[22] K. Palmer, *Shadowing in Dynamical Systems. Theory and Applications*, Kluwer, Dordrecht (2000).

[23] R. Raimi, The first digit problem, *Amer. Math. Monthly* **102** (1976), 322–327.

[24] 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.

[25] P. Schatte, On mantissa distributions in computing and Benford's law, *J. Information Processing and Cybernetics* **24** (1998), 443–455.

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