Benford solutions of linear difference equations

Arno Berger and Gideon Eshun

Abstract Benford's Law (BL), a notorious gem of mathematics folklore, asserts that leading digits of numerical data are usually not equidistributed, as might be expected, but rather follow one particular logarithmic distribution. Since first recorded by Newcomb in 1881, this apparently counter-intuitive phenomenon has attracted much interest from scientists and mathematicians alike.

This article presents a comprehensive overview of the theory of BL for autonomous linear difference equations. Necessary and sufficient conditions are given for solutions of such equations to conform to BL in its strongest form. The results extend and unify previous results in the literature. Their scope and limitations are illustrated by numerous instructive examples.

1 Introduction

The study of digits generated by dynamical processes is a classical and rather wide subject that continues to attract interest from disciplines as diverse as ergodic and number theory [1, 14, 15, 27, 30], statistics [18, 21, 32], political science [16, 31, 40], and accounting [12, 13, 33, 37, 38]. Across these disciplines, one recurring theme is the surprising ubiquity of a logarithmic distribution of digits often referred to as *Benford's Law* (BL). The most well-known special case of BL is the so-called *first-digit law* which asserts that

$$\mathbb{P}(\text{leading digit} = d) = \log\left(1 + d^{-1}\right), \quad \forall d = 1, 2, \ldots, 9, \tag{1.1}$$

where *leading digit* refers to the first significant (decimal) digit (see Section 2 for rigorous definitions) and log is the base-10 logarithm; for example, the leading digit of $e = 2.718$ is 2, whereas the leading digit of $-e^e = -15.15$ is 1. Note that (1.1) is

Arno Berger, Gideon Eshun
University of Alberta, Edmonton AB, Canada, e-mail: {berger, geshun}@ualberta.ca

heavily skewed towards the smaller digits: For instance, the leading digit is almost seven times as likely to equal 1 (probability $\log 2 = 30.10\%$) as it is to equal 9 (probability $\log \frac{10}{9} = 4.57\%$).

Ever since first recorded by S. Newcomb [36] in 1881 and re-discovered by F. Benford [3] in 1938, examples of data and systems conforming to (1.1) in one form or another have been discussed extensively, for instance in real-life data (e.g. [19, 41]), stochastic processes (e.g. [44]), and in deterministic sequences (e.g. $(n!)$ and the prime numbers [17]). There now exists a large body of literature devoted to the mechanisms whereby mathematical objects, such as e.g. sequences or random variables, do or do not satisfy (1.1) or variants thereof, see also Fig. 1. Beyond mathematics, BL has found diverse applications throughout the sciences. Given that the ubiquity of BL in these fields is still somewhat of a mystery [8], some BL-based tools (e.g. for fraud detection in tax, census, election or image processing data) have proved remarkably successful in practice. This in turn has triggered further research on the many unique features of BL [22, 23, 45]. It still rings true that, as R. Raimi [39] observed almost 40 years ago,

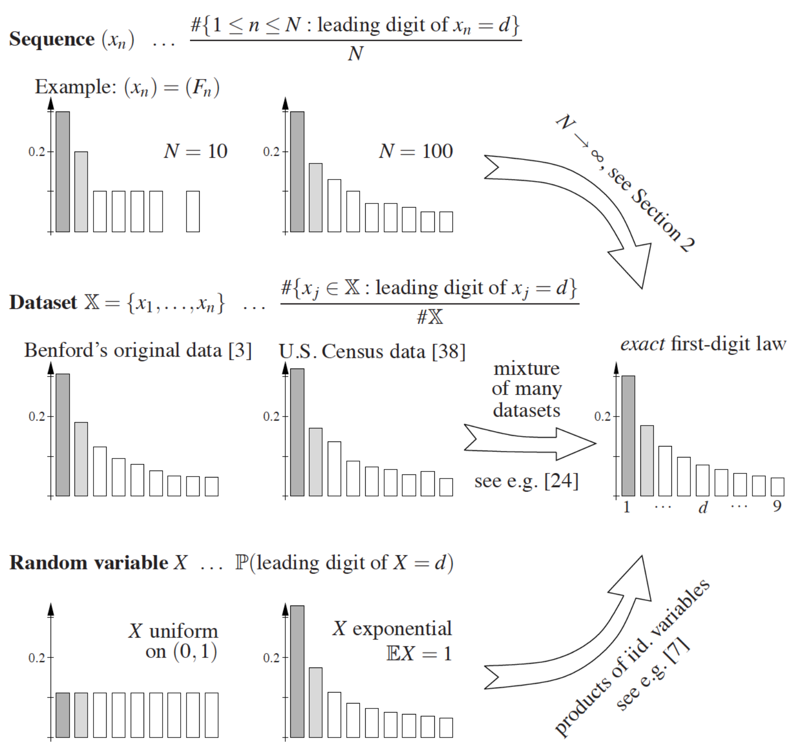

Fig. 1 Different interpretations of (1.1) for sequences, datasets, and random variables, respectively, and scenarios that may lead to conformance to the first-digit law.

> [t]his particular logarithmic distribution of the first digits, while not universal, is so common and yet so surprising at first glance that it has given rise to a varied literature, among the authors of which are mathematicians, statisticians, economists, engineers, physicists and amateurs.

As of this writing, an online database [4] devoted exclusively to BL lists more than 750 references.

Due to their important role as elementary models throughout science, linear difference and differential equations have, from very early on, been studied for their conformance to (1.1). A simple early example [11, 20, 28, 48] is the sequence $(x_n) = (F_n) = (1,1,2,3,5,\ldots)$ of Fibonacci numbers, i.e. $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for all $n \geq 3$, which satisfies (1.1) in the sense that

$$\lim_{N\to\infty} \frac{\#\{1 \leq n \leq N : \text{leading digit of } x_n = d\}}{N} = \log(1 + d^{-1}), \quad \forall d = 1,2,\ldots 9, \tag{1.2}$$

see also Fig. 2. Another simple case in point is $(x_n) = (2^n)$ for which (1.2) also holds [2, §24.4]. On the other hand, the sequence of primes $(x_n) = (2,3,5,7,11,\ldots)$ does not satisfy (1.2), as was in essence observed already by [47], yet may conform to BL in some weaker sense [14, 42].

1	121393	20365011074	3416454622906707
1	196418	32951280099	5527939700884757
2	317811	53316291173	8944394323791464
3	514229	86267571272	14472334024676221
5	832040	139583862445	23416728348467685
8	1346269	225851433717	37889062373143906
13	2178309	365435296162	61305790721611591
21	3524578	591286729879	99194853094755497
34	5702887	956722026041	160500643816367088
55	9227465	1548008755920	259695496911122585
89	14930352	2504730781961	420196140727489673
144	24157817	4052739537881	679891637638612258
233	39088169	6557470319842	1100087778366101931
377	63245986	10610209857723	1779979416004714189
610	102334155	17167680177565	2880067194370816120
987	165580141	27777890035288	4660046610375530309
1597	267914296	44945570212853	7540113804746346429
2584	433494437	72723460248141	12200160415121876738
4181	701408733	117669030460994	19740274219868223167
6765	1134903170	190392490709135	31940434634990099905
10946	1836311903	308061521170129	51680708854858323072
17711	2971215073	498454011879264	83621143489848422977
28657	4807526976	806515533049393	135301852344706746049
46368	7778742049	1304969544928657	218922995834555169026
75025	12586269025	2111485077978050	354224848179261915075

	#	*exact* BL
1	30	30.10
2	18	17.60
3	13	12.49
4	9	9.69
5	8	7.91
6	6	6.69
7	5	5.79
8	7	5.11
9	4	4.57

Fig. 2 Already the first one-hundred Fibonacci numbers conform to BL quite well.

Both positive examples mentioned above, i.e. the sequences (F_n) and (2^n), are obviously solutions of (very simple) *autonomous linear difference equations*. Building on earlier work, notably [5, 26, 35, 43], it is the purpose of this article to provide a comprehensive overview of the theory of BL for such equations. Thus the central question throughout is as follows: Given $d \in \mathbb{N}$ and real numbers $a_1, a_2, \ldots, a_{d-1}, a_d$ with $a_d \neq 0$, consider the (autonomous, d-th order) linear difference equation

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_{d-1} x_{n-d+1} + a_d x_{n-d}\,, \quad n \geq d+1\,. \tag{1.3}$$

Under which conditions on $a_1, a_2, \ldots, a_{d-1}, a_d$, and presumably also on the initial values $x_1, x_2, \ldots, x_d$, does the solution (x_n) of (1.3) satisfy (1.2)? Early work in this regard seems to have led merely to *sufficient* conditions that are either restrictive or difficult to state. By contrast, two of the main results presented here (Corollary 3.7 and Theorem 4.11) provide easy-to-state, *necessary and sufficient* conditions for every non-trivial solution of (1.3) to conform to (1.1) in a sense much stronger than (1.2). The classical results in the literature are then but simple special cases.

The organisation of this article is as follows. Section 2 introduces the formal definitions and analytic tools required for the analysis. In Section 3, difference equations (1.3), as well as the matrices associated with them are studied under the additional assumption of *positivity*. Though restrictive, this assumption holds for some important applications, and it yields a particularly simple answer to the central question raised earlier. Dropping the positivity assumption, Section 4 studies the case of general equations and matrices. The emergence of *resonances*, the key problem in the general case, is dealt with by means of a taylor-made definition (Definition 4.2). While the main results (Theorems 4.5 and 4.11) are stated in full generality, proofs are given here only under an additional non-degeneracy condition (and the interested reader is referred to the authors' forthcoming work [6] for complete proofs). Finally, Section 5 demonstrates how the presented results can be used to explain the "cancellation of resonance" phenomenon first observed in the context of finite-state Markov chains [10].

2 Basic definitions and tools

Throughout, the following, mostly standard notation is adhered to. 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 cardinality of any finite set $Z \subset \mathbb{C}$ is $\#Z$. The real part, imaginary part, complex conjugate, and absolute value (modulus) of $z \in \mathbb{C}$ is denoted by $\Re z$, $\Im z$, $\bar{z}$, and $|z|$, respectively. Let $\mathbb{S} := \{z \in \mathbb{C} : |z| = 1\}$. The argument $\arg z$ of $z \neq 0$ is understood to be the unique number in $(-\pi, \pi]$ for which $z = |z| e^{\imath \arg z}$; for convenience, let $\arg 0 := 0$. For any set $Z \subset \mathbb{C}$ and number $w \in \mathbb{C}$, define $wZ := \{wz : z \in Z\}$. Thus for instance $w\mathbb{S} = \{z \in \mathbb{C} : |z| = |w|\}$ for every $w \in \mathbb{C}$. Given $Z \subset \mathbb{C}$, denote by $\mathrm{span}_{\mathbb{Q}} Z$ the smallest subspace of $\mathbb{C}$ (over $\mathbb{Q}$) containing Z; equivalently, if $Z \neq \varnothing$ then $\mathrm{span}_{\mathbb{Q}} Z$ is the set of all *finite* rational linear combinations of elements of Z, i.e.

$$\mathrm{span}_{\mathbb{Q}} Z = \big\{\rho_1 z_1 + \rho_2 z_2 + \ldots + \rho_n z_n : n \in \mathbb{N}, \rho_1, \rho_2, \ldots, \rho_n \in \mathbb{Q}, z_1, z_2, \ldots, z_n \in Z\big\}\,;$$

note that $\mathrm{span}_{\mathbb{Q}}\varnothing = \{0\}$. For every $x \in \mathbb{R}^+$, $\log x$ and $\ln x$ are, respectively, the base-10 and the natural (base-e) logarithm of x; for convenience, set $\log 0 := \ln 0 := 0$. For every $x \in \mathbb{R}$, denote by $\lfloor x \rfloor$ the largest integer not larger than x, hence $\langle x \rangle := x - \lfloor x \rfloor$ is the non-integer (or fractional) part of x.

Given $x \in \mathbb{R} \setminus \{0\}$, there exists a unique $S(x) \in [1,10)$ such that $|x| = S(x)10^k$ for some (necessarily unique) integer k. The number $S(x)$ is the (*decimal*) *significand* of x. Note that

$$S(x) = 10^{\langle \log |x| \rangle}, \quad \forall x \in \mathbb{R} \setminus \{0\};$$

for convenience let $S(0) := 0$. For $x \neq 0$, the integer $\lfloor S(x) \rfloor \in \{1, 2, \ldots, 9\}$ is the *first significant* (*decimal*) *digit* of x. More generally, for every $m \in \mathbb{N}$, the integer $\lfloor 10^{m-1} S(x) \rfloor - 10 \lfloor 10^{m-2} S(x) \rfloor \in \{0, 1, \ldots, 9\}$ is the *m-th significant* (*decimal*) *digit* of x, see e.g. [7, Prop.2.5].

Throughout this article, conformance to (1.1) for solutions of difference equations is studied using the following terminology.

Definition 2.1. A sequence (x_n) of real numbers is a *Benford sequence*, or simply *Benford*, if

$$\lim_{N \to \infty} \frac{\#\{1 \le n \le N : S(x_n) \le t\}}{N} = \log t, \quad \forall t \in [1, 10). \tag{2.1}$$

Note that every Benford sequence (x_n) satisfies (1.2). For the purpose of this work, the following well-known characterization of the Benford property is indispensable.

Proposition 2.2. [17, Thm.1] *A sequence (x_n) is Benford if and only if the sequence $(\log |x_n|)$ is uniformly distributed modulo one.*

The term *uniformly distributed modulo one* is henceforth abbreviated *u.d.* mod 1. In view of Proposition 2.2, a few basic facts regarding uniform distribution of sequences are used throughout; for an authoritative overall account on the subject, the reader is referred to [29].

Proposition 2.3. [29, Sec.I.2] *The following statements are equivalent for any sequence (y_n) in $\mathbb{R}$:*

(i) (y_n) *is u.d.* mod 1*;*
(ii) *For every $\varepsilon > 0$ there exists a sequence (z_n) that is u.d.* mod 1*, and*

$$\limsup_{N \to \infty} \frac{\#\{1 \le n \le N : |y_n - z_n| > \varepsilon\}}{N} < \varepsilon;$$

(iii) *Whenever (z_n) is convergent then $(y_n + z_n)$ is u.d.* mod 1*;*
(iv) (py_n) *is u.d.* mod 1 *for every non-zero integer p;*
(v) $(y_n + \alpha \log n)$ *is u.d.* mod 1 *for every $\alpha \in \mathbb{R}$.*

One of the simplest yet also most fundamental examples of a sequence u.d. mod 1 is $(n\vartheta)$ with $\vartheta \in \mathbb{R} \setminus \mathbb{Q}$. The following, therefore, is an immediate consequence of Propositions 2.2 and 2.3.

Proposition 2.4. *Let (x_n) be a sequence in $\mathbb{R}$, and $\alpha \in \mathbb{R}\setminus\{0\}$. If $\lim_{n\to\infty} x_n/\alpha^n$ exists (in $\mathbb{R}$) and is non-zero, then (x_n) is Benford if and only if $\log|\alpha|$ is irrational.*

Example 2.5. Since $\log 2$ is irrational (even transcendental), (2^n) is Benford, and so is the sequence (F_n) of Fibonacci numbers because, with $\varphi = \frac{1}{2}(1+\sqrt{5})$, $\lim_{n\to\infty} F_n/\varphi^n = 1/\sqrt{5} \neq 0$, and $\log\varphi$ is irrational as well. $\triangleleft$

Remark. The Benford property can be studied w.r.t. any integer base $b \ge 2$, simply by replacing the decimal significand $S(x)$ in (2.1) with the base-b significand $S_b(x) = b^{\langle \log_b |x|\rangle}$, where $\log_b$ denotes the base-b logarithm. With the obvious modifications, the results in this work carry over to arbitrary base $b \in \mathbb{N}\setminus\{1\}$, cf. [5, 6, 7]. For the sake of clarity, however, only the familiar case $b = 10$ is considered from now on.

When studying solutions of linear difference equations, sequences of a particular form are often encountered, and the following lemma clarifies their properties.

Lemma 2.6. *Let $\alpha \in \mathbb{R}$, $z \in \mathbb{C}\setminus\{0\}$, and (z_n) a sequence in $\mathbb{C}$ with $\lim_{n\to\infty} z_n = 0$. If $\vartheta_1, \vartheta_2 \in \mathbb{R}$ are irrational then the following statements are equivalent:*

(i) $\vartheta_1 \notin \mathrm{span}_{\mathbb{Q}}\{1, \vartheta_2\}$*;*
(ii) *The sequence (y_n) with*

$$y_n = n\vartheta_1 + \alpha\log n + \log\left|\Re\left(z e^{\imath\pi n\vartheta_2} + z_n\right)\right|, \quad n \in \mathbb{N},$$

is u.d. mod 1.

Proof. If $\vartheta_1 \notin \mathrm{span}_{\mathbb{Q}}\{1, \vartheta_2\}$ then $1, \vartheta_1, \vartheta_2$ are rationally independent, and [5, Lem.2.9] shows that (y_n) is u.d. mod 1. On the other hand, if $\vartheta_1 \in \mathrm{span}_{\mathbb{Q}}\{1, \vartheta_2\}$ then $k_1\vartheta_1 = k_0 + k_2\vartheta_2$, where k_0, k_1, k_2 are appropriate integers with $k_1 k_2 \neq 0$; assume w.l.o.g. that $k_1 > 0$. Consider now the sequence (η_n) with

$$\eta_n = n\vartheta_1 + \log\left|\Re(z e^{\imath\pi n\vartheta_2})\right| + \tfrac{k_2}{k_1}\left(\tfrac{1}{2} + \tfrac{\arg z}{\pi}\right) - \log|z|, \quad n \in \mathbb{N}.$$

If (y_n) was u.d. mod 1, then so would be (η_n), and hence also $(k_1\eta_n)$, by Proposition 2.3. Moreover, for every $n \in \mathbb{N}$,

$$\begin{aligned}\langle k_1\eta_n\rangle &= \left\langle k_2\left(n\vartheta_2 + \tfrac{1}{2} + \tfrac{\arg z}{\pi}\right) + k_1\log\left|\sin\left(\pi\left(n\vartheta_2 + \tfrac{1}{2} + \tfrac{\arg z}{\pi}\right)\right)\right|\right\rangle \\ &= \left\langle f\left(n\vartheta_2 + \tfrac{1}{2} + \tfrac{\arg z}{\pi}\right)\right\rangle,\end{aligned}$$

with the measurable function $f : \mathbb{R} \to \mathbb{R}$ given by $f(t) = k_2 t + k_1\log|\sin(\pi t)|$. Note that $f(t+1) - f(t) \in \mathbb{Z}$ for all $t \in \mathbb{R}$, and so f induces the measurable map $T := \langle f\rangle$ on $[0,1)$. Recall now that the sequence $(n\vartheta_2 + \frac{1}{2} + \frac{\arg z}{\pi})$ is u.d. mod 1 because ϑ_2 is irrational. For every continuous, 1-periodic function $g : \mathbb{R} \to \mathbb{R}$, therefore,

$$\begin{aligned}\frac{1}{N}\sum\nolimits_{n=1}^{N} g(\langle k_1\eta_n\rangle) &= \frac{1}{N}\sum\nolimits_{n=1}^{N} g\circ f\left(n\vartheta_2 + \tfrac{1}{2} + \tfrac{\arg z}{\pi}\right) \\ &\stackrel{N\to\infty}{\longrightarrow} \int_0^1 g\circ f(t)\,\mathrm{d}t = \int_{[0,1]} g\circ T\,\mathrm{d}\lambda_{0,1} = \int_{[0,1]} g\,\mathrm{d}(\lambda_{0,1}\circ T^{-1}),\end{aligned}$$

because $g \circ f$ is Riemann integrable on $[0,1]$. On the other hand, if $(k_1\eta_n)$ was u.d. mod 1 then $\lim_{N\to\infty} N^{-1}\sum_{n=1}^N g(\langle k_1\eta_n\rangle) = \int_{[0,1]} g\,\mathrm{d}\lambda_{0,1}$ for every g, and hence $\lambda_{0,1}\circ T^{-1} = \lambda_{0,1}$. However, it is intuitively clear that the latter equality of measures does not hold. To see this formally, note that f is smooth on $(0,1)$ and has a (unique) non-degenerate maximum at some $0 < t_0 < 1$. Thus if $\lambda_{0,1}\circ T^{-1} = \lambda_{0,1}$ then, for all $\varepsilon > 0$ sufficiently small,

$$\begin{aligned}\frac{f(t_0-\varepsilon)-f(t_0-2\varepsilon)}{\varepsilon} &= \frac{\lambda_{0,1}\big([T(t_0-2\varepsilon), T(t_0-\varepsilon)]\big)}{\varepsilon} \\ &= \frac{\lambda_{0,1}\circ T^{-1}\big([T(t_0-2\varepsilon), T(t_0-\varepsilon)]\big)}{\varepsilon} \\ &\geq \frac{\lambda_{0,1}([t_0-2\varepsilon, t_0-\varepsilon])}{\varepsilon} = 1\,,\end{aligned}$$

which is impossible since $f'(t_0)=0$, see also Fig. 3 which depicts the special case $k_1 = k_2 = 1$. Hence $(k_1\eta_n)$ is not u.d. mod 1, and neither are (η_n) and (y_n). $\square$

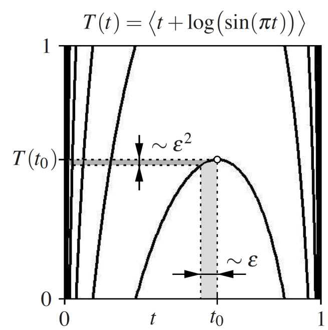

Fig. 3 The map T does not preserve $\lambda_{0,1}$, see the proof of Lemma 2.6.

Although it would be possible to study the Benford property of solutions of (1.3) directly, the analysis in subsequent sections becomes more transparent by means of a standard matrix-vector approach. To this end, associate with (1.3) the matrix

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_{d-1} & a_d \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{d\times d}\,, \tag{2.2}$$

which is invertible as $a_d \neq 0$, and recall that, given initial values $x_1, x_2, \ldots, x_d \in \mathbb{R}$, the solution of (1.3) can be expressed neatly in the form

$$x_n = e_d^\top A^n y\,, \quad \text{where } y = A^{-1}\begin{bmatrix} x_d \\ \vdots \\ x_2 \\ x_1 \end{bmatrix} \in \mathbb{R}^d\,; \tag{2.3}$$

here $e_1, e_2, \ldots, e_d$ represent the standard basis of $\mathbb{R}^d$, and $x^\top$ denotes the transpose of $x \in \mathbb{R}^d$, with $x^\top y$ being understood simply as the real number $\sum_{j=1}^d x_j y_j$. In what follows, therefore, the following, more general question suggested by (2.3) will be addressed: Under which conditions is $(x^\top A^n y)$ Benford, where A is any fixed real $d \times d$-matrix and $x, y \in \mathbb{R}^d$ are given vectors? Note that specifically choosing $x = e_j$ and $y = e_k$, with $j, k \in \{1, 2, \ldots, d\}$, simply yields $e_j^\top A^n e_k = [A^n]_{jk}$, i.e. the entry of A^n at the position (j, k). Also, if $A \in \mathbb{R}^{d\times d}$ is given by (2.2) then *every* sequence $(x^\top A^n y)$ solves (1.3), and (2.3) establishes a one-to-one correspondence between all sequences of the form $(e_d^\top A^n y)$ and *all* solutions of (1.3).

In the analysis of powers of matrices in the subsequent sections, d always is a fixed but usually unspecified positive integer. For every $x \in \mathbb{R}^d$, the number $|x| \ge 0$ is the *Euclidean norm* of x, i.e. $|x| = \sqrt{x^\top x} = \sqrt{\sum_{j=1}^d x_j^2}$. A vector $x \in \mathbb{R}^d$ is a *unit* vector if $|x| = 1$. The $d \times d$-identity matrix is I_d. For every matrix $A \in \mathbb{R}^{d\times d}$, its spectrum, i.e. the set of its eigenvalues, is denoted by $\sigma(A)$. Thus $\sigma(A) \subset \mathbb{C}$ is non-empty, contains at most d numbers and is symmetric w.r.t. the real axis, i.e., all non-real elements of $\sigma(A)$ come in complex-conjugate pairs. The number $r_\sigma(A) := \max\{|\lambda| : \lambda \in \sigma(A)\} \ge 0$ is the *spectral radius* of A. Note that $r_\sigma(A) > 0$ unless A is *nilpotent*, i.e. unless $A^N = 0$ for some $N \in \mathbb{N}$. For every $A \in \mathbb{R}^{d\times d}$, the number $|A|$ is the (*spectral*) *norm* of A, as induced by $|\cdot|$, i.e. $|A| = \max\{|Ax| : |x| = 1\}$. It is well-known that $|A| = \sqrt{r_\sigma(A^\top A)} \ge r_\sigma(A)$.

3 A simple special case: Positive matrices

The analysis of sequences $(x^\top A^n y)$ is especially simple if the matrix A or one of its powers happens to be positive. Recall that $A \in \mathbb{R}^{d\times d}$ is *positive*, in symbols $A > 0$, if $[A]_{jk} > 0$ for every $j, k \in \{1, 2, \ldots, d\}$. The following classical result, due to O. Perron, lists some of the remarkable properties of positive matrices, as they pertain to the present section. For a concise formulation, call $x \in \mathbb{R}^d$ *positive* (*non-negative*), in symbols $x > 0$ ($x \ge 0$), if $x_j > 0$ ($x_j \ge 0$) for every $j \in \{1, 2, \ldots, d\}$.

Proposition 3.1. [25, Sec.8.2] *Assume that $A \in \mathbb{R}^{d\times d}$ is positive. Then:*

(i) *The number $r_\sigma(A) > 0$ is an (algebraically) simple eigenvalue of A, i.e. a simple root of the characteristic polynomial of A;*
(ii) $|\lambda| < r_\sigma(A)$ *for every eigenvalue $\lambda \ne r_\sigma(A)$ of A;*
(iii) *There exists a positive eigenvector q, unique up to multiplication by a positive number, corresponding to the eigenvalue $r_\sigma(A)$;*

(iv) *The limit* $Q := \lim_{n\to\infty} A^n / r_\sigma(A)^n$ *exists, and* $Q > 0$ *satisfies* $Q^2 = Q$ *as well as* $AQ = QA = r_\sigma(A)Q$. *(In fact,* Q *is a rank-one projection with* $Qq = q$*.)*

Recall that (α^n) with $\alpha > 0$ is Benford if and only if $\log\alpha$ is irrational. The following is a generalization of this simple fact to arbitrary dimension. Informally put, it asserts that as far as the Benford property is concerned, matrices with some positive power behave just like the one-dimensional sequence $\big(r_\sigma(A)^n\big)$.

Theorem 3.2. *Let* A *be a real* $d\times d$*-matrix, and assume that* $A^N > 0$ *for some* $N \in \mathbb{N}$. *Then the following four statements are equivalent:*

(i) *The number* $\log r_\sigma(A)$ *is irrational;*
(ii) *The sequence* $(x^\top A^n y)$ *is Benford for every* $x, y \neq 0$ *with* $x \geq 0$ *and* $y \geq 0$*;*
(iii) *The sequence* $(|A^n x|)$ *is Benford for every* $x \neq 0$ *with* $x \geq 0$*;*
(iv) *The sequence* $(|A^n|)$ *is Benford.*

Proof. Since $A^N > 0$, the number $r_\sigma(A^N) = r_\sigma(A)^N > 0$ is an algebraically simple eigenvalue of A^N, by Proposition 3.1. It follows that exactly one of the two numbers $r_\sigma(A) > 0$ and $-r_\sigma(A) < 0$ is an algebraically simple eigenvalue of A. Denote this eigenvalue by λ_0, and let P be the spectral projection associated with it, that is, $P = bc^\top / b^\top c$ where b, c are eigenvectors of, respectively, A and $A^\top$ corresponding to the eigenvalue λ_0, i.e. $Ab = \lambda_0 b$ and $A^\top c = \lambda_0 c$. Thus $P^2 = P$ and $AP = PA = \lambda_0 P$. Moreover, the matrix $R := A - \lambda_0 P$ clearly satisfies $AR = RA$ and $PR = RP = 0$, and hence

$$A^n = \lambda_0^n P + R^n, \quad \forall n \in \mathbb{N}. \tag{3.1}$$

Since $|\lambda| < |\lambda_0| = r_\sigma(A)$ for every eigenvalue λ of R, $\lim_{n\to\infty} R^n / r_\sigma(A)^n = 0$, and an evaluation of (3.1) along even multiples of N yields

$$\lim_{n\to\infty} \frac{(A^N)^{2n}}{r_\sigma(A^N)^{2n}} = \lim_{n\to\infty} \left(\frac{\lambda_0^{2nN}}{r_\sigma(A)^{2nN}} P + \frac{R^{2nN}}{r_\sigma(A)^{2nN}} \right) = P\,.$$

This shows that $P = Q > 0$, with Q according to Proposition 3.1(iv) applied to A^N.

With these preparations, the asserted equivalences are now easily established. Indeed, given any $x, y \neq 0$ with $x \geq 0$ and $y \geq 0$, the vector Qy is positive, and

$$\frac{|x^\top A^n y|}{r_\sigma(A)^n} = \frac{|\lambda_0^n x^\top Q y + x^\top R^n y|}{r_\sigma(A)^n} = \left| x^\top Q y + \frac{x^\top R^n y}{\lambda_0^n} \right| \stackrel{n\to\infty}{\longrightarrow} x^\top Q y > 0\,, \tag{3.2}$$

together with Proposition 2.4, shows that $(x^\top A^n y)$ is Benford if and only if $\log r_\sigma(A)$ is irrational. A similar argument applies to $(|A^n x|)$, as

$$\frac{|A^n x|}{r_\sigma(A)^n} = \frac{|\lambda_0^n Q x + R^n x|}{r_\sigma(A)^n} = \left| Q x + \frac{R^n x}{\lambda_0^n} \right| \stackrel{n\to\infty}{\longrightarrow} |Qx| > 0\,,$$

whenever $x \geq 0$, $x \neq 0$, and also to $(|A^n|)$, as

$$\frac{|A^n|}{r_\sigma(A)^n} = \frac{|\lambda_0^n Q + R^n|}{r_\sigma(A)^n} = \left| Q + \frac{R^n}{\lambda_0^n} \right| \stackrel{n\to\infty}{\longrightarrow} |Q| > 0\,. \quad \square$$

Remark. The proof of Theorem 3.2 shows that if $\log r_\sigma(A)$ is rational then $(x^\top A^n y)$ and $(|A^n x|)$ are not Benford for *any* $x, y \geq 0$, and neither is $(|A^n|)$ Benford. Also, in (iii) and (iv), the Euclidean norm $|\cdot|$ can be replaced by any norm on, respectively, $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$.

Corollary 3.3. *Let $A \in \mathbb{R}^{d\times d}$, and assume that $A^N > 0$ for some $N \in \mathbb{N}$. Then, for every $j,k \in \{1,2,\ldots,d\}$, the sequence $([A^n]_{jk})$ is Benford if and only if $\log r_\sigma(A)$ is irrational.*

Example 3.4. The matrix associated with the Fibonacci recursion

$$x_n = x_{n-1} + x_{n-2}\,, \quad n \geq 3\,, \tag{3.3}$$

is $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, with $r_\sigma(A) = \varphi = \frac{1}{2}(1+\sqrt{5})$. While A is non-negative, i.e. $[A]_{jk} \geq 0$ for all j,k, but fails to be positive, the matrix A^2 is positive, and so is A^n for every $n \geq 2$. Since $\log r_\sigma(A)$ is irrational (even transcendental), every entry of (A^n) is Benford. This is consistent with the fact that

$$A^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix}, \quad n \geq 2\,,$$

and the sequence (F_n) is Benford.

Consider now the sequence (x_n) with $x_n = e_1^\top A^n(3e_2 - e_1)$. Recall that (x_n) thus defined also solves (3.3). However, since $3e_2 - e_1$ is not non-negative, Theorem 3.2 does not allow to decide whether $(x_n) = (2,1,3,4,7,\ldots)$, traditionally referred to as the sequence of *Lucas numbers* and denoted (L_n), is Benford. Corollary 3.7 below shows very easily that this is indeed the case. $\triangleleft$

Example 3.5. Consider the (symmetric) matrix

$$B = \begin{bmatrix} -3 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 6 \end{bmatrix},$$

the characteristic polynomial of which is

$$p_B(\lambda) = \det(B - \lambda I_3) = -\lambda^3 + 3\lambda^2 + 20\lambda - 3\,.$$

It is readily confirmed that p_B has three different real roots. If $\lambda = \pm 10^{m/n}$ was a root of p_B with any relatively prime $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, then $n \leq 3$, and 10^m would divide $|\det B^n| = |\det B|^n = 3^n$, hence $m = 0$, that is, $\lambda = \pm 1$. But $p_B(\pm 1) = \pm 19 \neq 0$. It follows that $r_\sigma(B)$, albeit algebraic, is not a rational power of 10, and so $\log r_\sigma(B)$ is irrational (even transcendental). Moreover, B^n contains both positive and negative entries for $n = 1,2,\ldots,7$, yet

$$B^8 = \begin{bmatrix} 13841 & 1929 & 37034 \\ 1929 & 56662 & 335235 \\ 37034 & 335235 & 2031038 \end{bmatrix} > 0\,,$$

hence Theorem 3.2 and Corollary 3.3 apply. In particular, every entry of (B^n) is Benford. Note that the actual value of $r_\sigma(B)$,

$$r_\sigma(B) = 1 + \tfrac{2\sqrt{69}}{3}\cos\left(\tfrac{1}{3}\arccos\tfrac{57\sqrt{69}}{1058}\right) = 6.165\,,$$

is not needed at all to draw this conclusion. $\triangleleft$

Example 3.6. When $A > 0$ and $\log r_\sigma(A)$ is irrational, the sequence $(x^\top A^n y)$ may nevertheless *not* be Benford for some non-zero $x, y \in \mathbb{R}^d$. By Theorem 3.2, such x, y cannot both be non-negative. For instance, the matrix $A = \begin{bmatrix} 5 & 15 \\ 15 & 5 \end{bmatrix}$ is positive, and $\log r_\sigma(A) = 1 + \log 2$ is irrational, yet $\left(e_1^\top A^n(e_1 - e_2)\right) = \left((-10)^n\right)$ is not Benford. On the other hand, even if $r_\sigma(B)$ is rational, $(x^\top B^n y)$ may be Benford for *some* $x, y \in \mathbb{R}^d$. Again, x, y cannot both be non-negative, by virtue of Theorem 3.2. Concretely, $B = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix} > 0$ has $\log r_\sigma(B) = 1$ rational, yet $\left(e_1^\top B^n(e_1 - e_2)\right) = (2^n)$ is Benford. $\triangleleft$

Corollary 3.7. *Let (x_n) be a solution of the linear difference equation*

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_{d-1} x_{n-d+1} + a_d x_{n-d}\,, \quad n \ge d+1\,,$$

with $a_1, a_2, \ldots, a_{d-1}, a_d > 0$. Assume that the numbers $x_1, x_2, \ldots, x_d$ are non-negative, and at least one is positive. Then (x_n) is Benford if and only if $\log \zeta$ is irrational, where $z = \zeta$ is the right-most root of $z^d = a_1 z^{d-1} + a_2 z^{d-2} + \ldots + a_{d-1} z + a_d$.

Proof. The associated matrix A according to (2.2) is non-negative, and $A^n > 0$ for $n \ge d$. Moreover, A has the characteristic polynomial

$$p_A(\lambda) = (-1)^d(\lambda^d - a_1\lambda^{d-1} - a_2\lambda^{d-2} - \ldots - a_{d-1}\lambda - a_d)\,.$$

Since $x_n = x^\top A^{n-1} y$ with $x = e_d \ge 0$ and $y = \sum_{j=1}^d x_{d+1-j} e_j \ge 0$, the claim follows directly from Theorem 3.2. $\square$

Example 3.8. Every solution of (3.3) with $x_1 x_2 > 0$ is Benford. (For the case $x_1 < 0$ simply note that $(-x_n)$ is a solution of (3.3) as well.) Evidently, this includes the Fibonacci sequence, where $x_1 = x_2 = 1$, but also the Lucas numbers, where $x_1 = 2$, $x_2 = 1$. As they stand, however, Theorem 3.2 and Corollary 3.7 do not allow to decide whether the solution of (3.3) with, say, $x_1 = 2, x_2 = -3$ is Benford.

More generally, every solution (x_n) with $x_1 x_2 > 0$ of

$$x_n = a_1 x_{n-1} + a_2 x_{n-2}\,, \quad n \ge 3\,, \tag{3.4}$$

where a_1, a_2 are positive *integers*, is Benford if and only if $10^{2m} - a_2 \ne a_1 \cdot 10^m$ for every $m = 0, 1, \ldots, \lfloor \log(a_1 + a_2) \rfloor$. Again, this leaves open the question regarding the Benford property of solutions of (3.4) with $x_1 x_2 < 0$. The results of the next section allow to settle this question without any further calculation: Except for the trivial case $x_1 = x_2 = 0$, *every* solution of (3.4) is Benford if and only if

$$|10^{2m} - a_2| \neq a_1 \cdot 10^m \,, \quad \forall m = 0, 1, \ldots, \lfloor \log(a_1 + a_2) \rfloor \,. \tag{3.5}$$

For the Fibonacci recursion (3.3), for instance, (3.5) reduces to $|1-1| \neq 1$, which is obviously true. Thus, apart from $x_n \equiv 0$, *every* solution of (3.3) is Benford. $\triangleleft$

The following examples aim at illustrating the scope and limitations of Theorem 3.2. Although the latter is easy to state and prove, and quite useful in a variety of situations, its overall applicability is somewhat limited because

- it does not apply in general if the matrix in question fails to have a positive power, see Example 3.9;
- even if it applies, the Benford property of individual solutions of a linear difference equation (1.3), or equivalently of sequences $(x^\top A^n y)$ with A according to (2.2) and *arbitrary* $x, y \in \mathbb{R}^d$, is generally unrelated to the Benford property of $(x^\top A^n y)$ with *non-negative* x, y, see Example 3.10;
- it does not apply to various sequences that are closely related to (A^n) and often of interest in their own right, for instance $(A^{n+1} - r_\sigma(A)A^n)$, see Example 3.11.

In view of these limitations, in the next section the Benford property is studied more generally for sequences $(x^\top A^n y)$ with *arbitrary* $A \in \mathbb{R}^{d\times d}$ and $x, y \in \mathbb{R}^d$.

Example 3.9. Theorem 3.2 may fail if $A \in \mathbb{R}^{d\times d}$ does not have a positive power. Simply consider the (non-negative) matrix $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, for which $\log r_\sigma(A) = \log 2$ is irrational, yet $([A^n]_{jk})$ is constant and hence not Benford except for $j = k = 1$. Neither is $(|A^n e_2|)$ Benford. Thus the implications (i)$\Rightarrow$(ii) and (i)$\Rightarrow$(iii) in Theorem 3.2 do not even hold for non-negative matrices. As will be seen in the next section, however, (ii)$\Rightarrow$(i) and (iii)$\Rightarrow$(i) remain true for arbitrary (non-nilpotent) matrices in that if $(x^\top A^n y)$ or $(|A^n x|)$ is, for every $x, y \in \mathbb{R}^d$, either Benford or vanishes for all $n \geq d$ then $\log r_\sigma(A)$ is irrational. Similarly, if A does not have a positive power then $(|A^n|)$ may not be Benford even when $\log r_\sigma(A)$ is irrational, see Example 4.10.

Note also that even if B does not have any positive power, all entries of (B^n), or in fact all non-trivial sequences $(x^\top B^n y)$, may nevertheless be Benford, as the example $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ shows, for which

$$B^n = 2^{n-1} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \,, \quad n \in \mathbb{N} \,. \quad \triangleleft$$

Example 3.10. Consider the difference equation

$$x_n = \tfrac{1}{2}(x_{n-1} + x_{n-2}) \,, \quad n \geq 3 \,, \tag{3.6}$$

and the associated matrix

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \,. \tag{3.7}$$

Similarly to Example 3.4, $A \geq 0$ and $A^2 > 0$. In addition, A evidently has the property that the entries in each of its rows add up to 1. Thus A is a (*row-*) *stochastic* matrix. It is well known (and easy to see) that $r_\sigma(A) = 1$ for every (row- or column-) stochastic matrix. According to Theorem 3.2, none of the sequences $(x^\top A^n y)$ with $x, y \ge 0$ is Benford. In fact, a short calculation yields

$$A^n = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} + \frac{(-\frac{1}{2})^n}{3}\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \quad n \in \mathbb{N}_0, \tag{3.8}$$

showing that each sequence $(x^\top A^n y)$ converges to a finite limit (which is positive unless $x = 0$ or $y = 0$) and hence cannot be Benford. Recall that each such sequence is a solution of (3.6). On the other hand, the solution of (3.6) with $x_1 = -2$, $x_2 = 1$ is $(x_n) = \big((-\frac{1}{2})^{n-2}\big)$ and clearly Benford. Thus a solution of a linear difference equation may be Benford even if the associated matrix A has a positive power but does not satisfy (i)–(iv) in Theorem 3.2.

To see that the reverse situation — some solution of a difference equation is not Benford despite the associated matrix having a positive power and satisfying (i)–(iv) in Theorem 3.2 — can also occur, consider

$$x_n = 19x_{n-1} + 20x_{n-2}, \quad n \ge 3. \tag{3.9}$$

The solution of (3.9) with $x_1 = -1, x_2 = 1$ is $\big((-1)^n\big)$ and hence not Benford. On the other hand, the associated matrix $B = \begin{bmatrix} 19 & 20 \\ 1 & 0 \end{bmatrix}$ has a positive power as $B^2 > 0$, and $\log r_\sigma(B) = 1 + \log 2$ is irrational. $\lhd$

Example 3.11. If $A^N > 0$ for some $N \in \mathbb{N}$ then exactly one of the two numbers $\lambda_0 = r_\sigma(A) > 0$ or $\lambda_0 = -r_\sigma(A) < 0$ is an eigenvalue of A, and $Q := \lim_{n\to\infty} A^n/\lambda_0^n$ exists and is a positive matrix. This fact, which has been instrumental in the proof of Theorem 3.2, is of particular interest in the case of A being a stochastic matrix, i.e. for $A \ge 0$ and each row (or column) of A summing up to 1. In this case, $\lambda_0 = r_\sigma(A) = 1$, and hence $Q = \lim_{n\to\infty} A^n$. Often, one is interested in the (Benford) properties of $(A^n - Q)$ and $(A^{n+1} - A^n)$. Entries of these sequences may well be Benford, notwithstanding the fact that Theorem 3.2 does not apply and $\log r_\sigma(A) = 0$ is rational. For instance, with A from (3.7), it follows from (3.8) that

$$A^n - Q = \frac{(-\frac{1}{2})^n}{3}\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \quad n \in \mathbb{N}_0,$$

but also

$$A^{n+1} - A^n = (-\tfrac{1}{2})^{n+1}\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \quad n \in \mathbb{N}_0,$$

and hence every entry of both $(A^n - Q)$ and $(A^{n+1} - A^n)$ is Benford. In general, note that $AQ = QA = \lambda_0 Q$, and consequently the sequences

$$\big([A^n - \lambda_0^n Q]_{jk}\big) = \big(e_j^\top (A^n - \lambda_0^n Q)e_k\big) = \big(e_j^\top A^n (I_d - Q)e_k\big)$$

as well as

$$\left([A^{n+1}-\lambda_0 A^n]_{jk}\right)=\left(e_j^\top A^n(A-\lambda_0 I_d)e_k\right)$$

are all of the form $(x^\top A^n y)$ with $x=e_j$ and the appropriate $y\in\mathbb{R}^d$ where, however, $y\geq 0$ may not hold and consequently Theorem 3.2 may not apply. $\lhd$

With a view towards Theorem 3.2, how does one decide in practice whether a given $d\times d$-matrix A has a positive power? Comprehensive answers to this question appear to be documented in the literature only for $A\geq 0$, that is, for non-negative matrices. In this case, Wielandt's Theorem [25, Cor.8.5.9] asserts that $A^N>0$ for some $N\in\mathbb{N}$ (if and) only if $A^{d^2-2d+2}>0$. The number d^2-2d+2 is smallest possible in general, but can be reduced in many special cases, see [25, Sec.8.5]. An equivalent condition is that A be *irreducible* and *aperiodic*, i.e., for any two indices $j,k\in\{1,2,\ldots,d\}$ there exists a positive integer $N(j,k)$ such that $[A^n]_{jk}>0$ for every $n\geq N(j,k)$; see e.g. [25, Sec.8.4]. Note that if $A\geq 0$ but the matrix A^n is not positive for *any* $n\in\mathbb{N}$ then there exists $j,k\in\{1,2,\ldots,d\}$ such that eventually the sequence $([A^n]_{jk})$ vanishes periodically and hence cannot be Benford. Overall, by combining these known facts, Theorem 3.2 can be re-stated specifically for non-negative matrices.

Theorem 3.12. *Let $A\in\mathbb{R}^{d\times d}$ be non-negative. Then the following three statements are equivalent:*

(i) *A is irreducible and aperiodic, and $\log r_\sigma(A)$ is irrational;*
(ii) *$A^{d^2-2d+2}>0$ and $\log r_\sigma(A)$ is irrational;*
(iii) *The sequence $(x^\top A^n y)$ is Benford for every $x,y\neq 0$ with $x\geq 0$ and $y\geq 0$.*

Moreover, if (i)–(iii) *hold then, for every $x\neq 0$ with $x\geq 0$, the sequence $(|A^n x|)$ is Benford, and so is $(|A^n|)$.*

Proof. If A^{d^2-2d+2} is not positive then neither is A^n for any n, by Wielandt's Theorem, and hence A cannot be irreducible and aperiodic. Thus (i)$\Rightarrow$(ii). According to Theorem 3.2, (iii) follows from (ii). Assume in turn that (i) does not hold. Then either A is not irreducible and aperiodic, or $\log r_\sigma(A)$ is rational. In the former case, $([A^n]_{jk})=(e_j^\top A^n e_k)$ vanishes periodically for some $j,k\in\{1,2,\ldots,d\}$, hence (iii) fails with $x=e_j\geq 0$ and $y=e_k\geq 0$. In the latter case, assume w.l.o.g. that A is irreducible and aperiodic. Then (iii) fails again, by virtue of Theorem 3.2. Overall, (iii)$\Rightarrow$(i). Finally, the assertions regarding $(|A^n x|)$ and $(|A^n|)$ are obvious from Theorem 3.2. $\square$

Remark. The non-negative matrix $A=\begin{bmatrix}0&2\\2&0\end{bmatrix}$ has neither of the properties (i)–(iii) in Theorem 3.12, and yet $(|A^n x|)$ is Benford for every $x\neq 0$, and so is $(|A^n|)$.

In general, i.e. without the assumption that A be non-negative, the clear-cut situation of the non-negative case persists only for the special cases $d=1$ (trivial) and $d=2$ (a simple exercise), where $A^N>0$ for some $N\in\mathbb{N}$ (if and) only if $A^2>0$. In stark contrast, if $d\geq 3$ then the minimal positive integer N with $A^N>0$ can be arbitrarily large. For example, for every $\alpha\in\mathbb{R}$ the (symmetric) 3×3-matrix

$$A_\alpha := \begin{bmatrix} 10-10^{4\alpha} & 10^{\alpha+1}\sqrt{2}(10^{2\alpha-1}+1) & 9\cdot 10^{2\alpha} \\ 10^{\alpha+1}\sqrt{2}(10^{2\alpha-1}+1) & 18\cdot 10^{2\alpha} & 10^{\alpha}\sqrt{2}(10^{2\alpha+1}+1) \\ 9\cdot 10^{2\alpha} & 10^{\alpha}\sqrt{2}(10^{2\alpha+1}+1) & 10^{4\alpha+1}-1 \end{bmatrix}$$

is positive precisely if $|\alpha| < \frac{1}{4}$, and for $|\alpha| \geq \frac{1}{4}$ a short calculation shows that

$$\min\{n \in \mathbb{N} : A_\alpha^n > 0\} = 2\lfloor|\alpha|\rfloor + 2 > 2|\alpha|\,.$$

Note that, for every $\alpha \in \mathbb{R}$, the matrix A_α has at most *one* negative entry (and is, in the terminology of [25, Exc.8.3.9], *essentially non-negative*).

The example of A_α demonstrates that unlike in the non-negative case, for $d \geq 3$ the minmal exponent N with $A^N > 0$ does not admit an upper bound independent of A. Still, the property that $A^N > 0$ for *some* $N \in \mathbb{N}$ can be characterized rather neatly.

Proposition 3.13. *The following properties are equivalent for every $A \in \mathbb{R}^{d\times d}$:*

(i) $A^N > 0$ *for some* $N \in \mathbb{N}$*;*
(ii) $A^{2n} > 0$ *for all sufficiently large* $n \in \mathbb{N}$*;*
(iii) *Either $\lambda_0 = r_\sigma(A) > 0$ or $\lambda_0 = -r_\sigma(A) < 0$ is an algebraically simple eigenvalue of A with $|\lambda| < r_\sigma(A)$ for every $\lambda \in \sigma(A)\setminus\{\lambda_0\}$, and the spectral projection Q associated with λ_0 is positive, i.e.*

$$Q = \frac{bc^\top}{b^\top c} > 0\,, \tag{3.10}$$

where b and c are eigenvectors of, respectively, A and $A^\top$ corresponding to the eigenvalue λ_0, that is, $Ab = \lambda_0 b$ and $A^\top c = \lambda_0 c$.

Applying this result for instance to the 3×3-matrix B of Example 3.5 yields, with $\lambda_0 = r_\sigma(B) = 6.165$,

$$Q = 10^{-4}\begin{bmatrix} 3.158 & 28.95 & 175.3 \\ 28.95 & 265.3 & 1606 \\ 175.3 & 1606 & 9731 \end{bmatrix} > 0\,,$$

and hence immediately shows that $B^N > 0$ for *some* $N \in \mathbb{N}$. (In Example 3.5, the minimal such N was seen to be $N = 8$.) On the other hand, for the matrix B considered in Example 3.9, the spectral projection associated with $\lambda_0 = r_\sigma(B) = 2$,

$$Q = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\,,$$

is not positive, and neither is $B^n = 2^{n-1}B = 2^n Q$ positive for any $n \in \mathbb{N}$.

Example 3.14. Theorems 3.2 and 3.12 are especially easy to apply if A is an *integer* matrix, i.e., if $[A]_{jk} \in \mathbb{Z}$ for every j,k. In this case, an explicit calculation of $r_\sigma(A)$

is not required. In fact, if $A \in \mathbb{Z}^{d\times d}$ with $d \ge 2$ and $A^N > 0$ for some $N \in \mathbb{N}$ then $\log r_\sigma(A)$ is irrational (even transcendental) provided that

$$\begin{array}{l}\text{none of the numbers } \pm 10^m, \text{ with } m = 1,2,\ldots,\lfloor d \log \|A\|_\infty \rfloor \\ \text{and } \|A\|_\infty := \max_j \sum_{k=1}^{d} \big|[A]_{jk}\big|, \text{ is an eigenvalue of any of} \\ \text{the } d \text{ matrices } A, A^2, \ldots, A^d\,.\end{array} \tag{3.11}$$

Even simpler to check is the condition that

$$\det A \ \text{ is not divisible by } 10\,, \tag{3.12}$$

which implies (3.11) and hence also guarantees the irrationality of $\log r_\sigma(A)$.

For example, the matrix A associated with (3.3) is an integer matrix with $A^2 > 0$, and $\det A = -1$ obviously satisfies (3.12). Hence, as already seen in Example 3.4, Theorem 3.2 applies, and $(x^\top A^n y)$ is Benford for all $x, y \ge 0$ with $x \ne 0$ and $y \ne 0$. Similarly, for the matrix B discussed in Example 3.5, $B^8 > 0$, and $\det B = -3$ is not divisible by 10, hence $\log r_\sigma(B)$ is irrational, and again Theorem 3.2 can be applied without determining the actual value of $r_\sigma(B)$.

For another example, consider the matrix

$$C = \begin{bmatrix} -3 & -1 & -1 \\ -2 & 1 & -3 \\ 1 & -3 & -1 \end{bmatrix},$$

for which $\|C\|_\infty = 6$. As before, $C^8 > 0$, hence Theorem 3.2 applies. Note that $\det C = 30$, and so (3.12) fails. However, (3.11) holds, as $\lfloor 3 \log \|C\|_\infty \rfloor = 2$ and none of the four integers $\pm 10, \pm 10^2$ is an eigenvalue of any of the three matrices

$$C, \quad C^2 = \begin{bmatrix} 10 & 5 & 7 \\ 1 & 12 & 2 \\ 2 & -1 & 9 \end{bmatrix}, \quad C^3 = \begin{bmatrix} -33 & -26 & -32 \\ -25 & 5 & -39 \\ 5 & -30 & -8 \end{bmatrix},$$

as can easily be checked e.g. by means of row-reductions. Again, therefore, $\log r_\sigma(C)$ is irrational. $\lhd$

4 The case of arbitrary matrices

Given an arbitrary real $d \times d$-matrix A, this section presents a necessary and sufficient condition for the sequence $(x^\top A^n y)$ to be, for any vectors $x, y \in \mathbb{R}^d$, either Benford or identically zero for $n \ge d$. As explained earlier, the result also allows to characterize the Benford property for solutions of any linear difference equation. To provide the reader with some intuition as to which properties of such equations, or the matrices associated with them, may affect the Benford property, first a few simple examples are discussed.

Example 4.1. **(i)** Let the sequence (x_n) be defined recursively as

$$x_n = x_{n-1} - x_{n-2}\,, \quad n \ge 3\,, \tag{4.1}$$

with given $x_1, x_2 \in \mathbb{R}$. From the explicit representation for (x_n),

$$x_n = (x_1 - x_2)\cos\left(\tfrac{1}{3}\pi n\right) + \tfrac{1}{\sqrt{3}}(x_1 + x_2)\sin\left(\tfrac{1}{3}\pi n\right)\,, \quad n \in \mathbb{N}\,,$$

it is clear that $x_{n+6} = x_n$ for all n, i.e., (x_n) is 6-periodic. This oscillatory behaviour of (x_n) corresponds to the fact that the eigenvalues of (4.1), i.e. of the matrix associated with it, are $\lambda = e^{\pm \imath\pi/3}$ and hence lie on the unit circle $\mathbb{S}$. For no choice of x_1, x_2, therefore, is (x_n) Benford.

(ii) Consider the linear 3-step recursion

$$x_n = 2x_{n-1} + 10x_{n-2} - 20x_{n-3}\,, \quad n \ge 4\,. \tag{4.2}$$

For any $x_1, x_2, x_3 \in \mathbb{R}$, the value of x_n is given explicitly by

$$x_n = \alpha_1 2^n + \alpha_2 10^{n/2} + \alpha_3(-1)^n 10^{n/2}\,,$$

with the constants $\alpha_1, \alpha_2, \alpha_3$ according to

$$\alpha_1 = \tfrac{1}{12}(10x_1 - x_3)\,, \quad \alpha_{2,3} = \tfrac{1}{60}(x_3 + 3x_2 - 10x_1) \pm \tfrac{1}{12\sqrt{10}}(x_3 - 4x_1)\,.$$

Clearly, $\limsup_{n\to\infty}|x_n| = +\infty$ unless $x_1 = x_2 = x_3 = 0$, so unlike in (i) the sequence (x_n) is not bounded. However, if $|\alpha_2| \ne |\alpha_3|$ then

$$\log|x_n| = \frac{n}{2} + \log\left|\alpha_1 10^{-n(\frac{1}{2}-\log 2)} + \alpha_2 + (-1)^n\alpha_3\right| \approx \frac{n}{2} + \log|\alpha_2 + (-1)^n\alpha_3|\,,$$

showing that $\big(S(x_n)\big)$ is asymptotically 2-periodic and hence (x_n) is not Benford. Similarly, if $|\alpha_2| = |\alpha_3| \ne 0$ then $\big(S(x_n)\big)$ is convergent along even (if $\alpha_2 = \alpha_3$) or odd (if $\alpha_2 = -\alpha_3$) indices n, and again (x_n) is not Benford. Only if $\alpha_2 = \alpha_3 = 0$ yet $\alpha_1 \ne 0$, or equivalently if $x_3 = 2x_2 = 4x_1 \ne 0$ is (x_n) Benford. Obviously, the oscillatory behaviour of $\big(S(x_n)\big)$ in this example is due to the characteristic equation $\lambda^3 = 2\lambda^2 + 10\lambda - 20$ associated with (4.2) having two roots with the same modulus but opposite signs, namely $\lambda = \pm\sqrt{10}$.

(iii) Let $\gamma = \cos(\pi \log 2) = 0.5851$ and define (x_n) recursively as

$$x_n = 4\gamma x_{n-1} - 4x_{n-2}\,, \quad n \ge 3\,, \tag{4.3}$$

with given $x_1, x_2 \in \mathbb{R}$. As before, an explicit formula for x_n is easily derived as

$$\begin{aligned} x_n &= 2^{n-2}(4\gamma x_1 - x_2)\cos(\pi n \log 2) + 2^{n-2}\frac{\gamma x_2 - 2x_1(2\gamma^2 - 1)}{\sqrt{1-\gamma^2}}\sin(\pi n \log 2) \\ &= 2^n \beta \cos(\pi n \log 2 + \xi)\,, \end{aligned}$$

with the appropriate $\beta \ge 0$ and $\xi \in \mathbb{R}$. Although somewhat oscillatory, the sequence (x_n) is clearly unbounded. However, if $(x_1, x_2) \neq (0,0)$ then $\beta > 0$, and

$$\log|x_n| = n\log 2 + \log\beta + \log|\cos(\pi n \log 2 + \xi)|, \quad n \in \mathbb{N},$$

together with Lemma 2.6, where $\vartheta_1 = \vartheta_2 = \log 2$, $\alpha = 0$, $z = e^{\imath\xi}$, and $z_n \equiv 0$, shows that (x_n) is not Benford. The reason for this can be seen in the fact that, while $\log|\lambda| = \log 2$ is irrational for the roots $\lambda = 2e^{\pm\imath\pi\log 2}$ of the characteristic equation associated with (4.3), there clearly is a rational dependence between the two real numbers $\log|\lambda|$ and $\frac{1}{2\pi}\arg\lambda$, namely $\log|\lambda| - 2(\frac{1}{2\pi}\arg\lambda) = 0$. $\lhd$

The above examples indicate that, under the perspective of BL, the main difficulty when dealing with multi-dimensional systems is their potential for more or less cyclic behaviour, either of the orbits themselves or of their significands. (In the case of *positive* matrices, as seen in the previous section, cyclicality does not occur or, more correctly, remains hidden.) To precisely denominate this difficulty, the following terminology will prove useful. Recall that, given any set $Z \subset \mathbb{C}$, $\mathrm{span}_{\mathbb{Q}} Z$ denotes the smallest linear subspace of $\mathbb{C}$ (over $\mathbb{Q}$) containing Z.

Definition 4.2. A non-empty set $Z \subset \mathbb{C}$ with $|z| = r$ for some $r > 0$ and all $z \in Z$, i.e. $Z \subset r\mathbb{S}$, is *non-resonant* if its associated set $\Delta_Z \subset \mathbb{R}$, defined as

$$\Delta_Z := \left\{ 1 + \frac{\arg z - \arg w}{2\pi} : z, w \in Z \right\}$$

satisfies the following two conditions:

(i) $\Delta_Z \cap \mathbb{Q} = \{1\}$;
(ii) $\log r \notin \mathrm{span}_{\mathbb{Q}} \Delta_Z$.

An arbitrary set $Z \subset \mathbb{C}$ is *non-resonant* if, for every $r > 0$, the set $Z \cap r\mathbb{S}$ is either non-resonant or empty; otherwise Z is *resonant*.

Note that by its very definition the set Δ_Z always satisfies $1 \in \Delta_Z \subset (0,2)$ and is symmetric w.r.t. the point 1. The empty set $\varnothing$ and the singleton $\{0\}$ are non-resonant. On the other hand, $Z \subset \mathbb{C}$ is certainly resonant if either $\{-r, r\} \subset Z$ for some $r > 0$, in which case (i) is violated, or $Z \cap \mathbb{S} \neq \varnothing$, which causes (ii) to fail.

Example 4.3. The singleton $\{z\}$ with $z \in \mathbb{C}$ is non-resonant if and only if either $z = 0$ or $\log|z| \notin \mathbb{Q}$. Similarly, the set $\{z, \bar{z}\}$ with $z \in \mathbb{C} \setminus \mathbb{R}$ is non-resonant if and only if the three numbers 1, $\log|z|$ and $\frac{1}{2\pi}\arg z$ are rationally independent, i.e. linearly independent over $\mathbb{Q}$. $\lhd$

Remark. If $Z \subset r\mathbb{S}$ then, for every $z \in Z$,

$$\mathrm{span}_{\mathbb{Q}} \Delta_Z = \mathrm{span}_{\mathbb{Q}} \left(\{1\} \cup \left\{ \frac{\arg z - \arg w}{2\pi} : w \in Z \right\} \right),$$

which shows that the dimension of $\mathrm{span}_{\mathbb{Q}}\Delta_Z$, as a linear space over $\mathbb{Q}$, is at most $\#Z$. Also, if $Z \subset r\mathbb{S}$ is symmetric w.r.t. the real axis, then the condition (ii) in Definition 4.2 is equivalent to $\log r \notin \mathrm{span}_{\mathbb{Q}}\big(\{1\} \cup \{\frac{1}{2\pi} \arg z : z \in Z\}\big)$, cf. [5, Def.3.1].

Recall that the behaviour of (A^n) is completely determined by the eigenvalues of A, together with the corresponding (generalized) eigenvectors. As far as BL is concerned, the key question turns out to be whether or not $\sigma(A)$ is non-resonant. Clearly $\log r_\sigma(A)$ is irrational whenever $\sigma(A)$ is non-resonant (and A is not nilpotent), but the converse is not true in general.

Example 4.4. The spectrum of the matrix A associated with the Fibonacci recursion (3.3), $\sigma(A) = \{-\varphi^{-1}, \varphi\}$, is non-resonant. On the other hand, the matrices

$$B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 10 & -20 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 4\gamma & -4 \\ 1 & 0 \end{bmatrix},$$

associated with the difference equations (4.1), (4.2), and (4.3), respectively, all have a resonant spectrum. Indeed, $\sigma(B) = \{e^{\pm\imath\pi/3}\}$, and hence $\Delta_{\sigma(B)} = \{\frac{2}{3}, 1, \frac{4}{3}\}$ contains rational numbers other than 1, which violates (i) in Definition 4.2. Also, $\log|e^{\pm\imath\pi/3}| = 0$, and so (ii) is violated, too. Similarly, $\sigma(C) = \{2, \pm\sqrt{10}\}$, and with $Z = \sigma(C) \cap \sqrt{10}\mathbb{S} = \{\pm\sqrt{10}\}$ again both (i) and (ii) in Definition 4.2 fail. Finally, $\sigma(D) = \{2e^{\pm\imath\pi\log 2}\}$, and so $\Delta_{\sigma(D)} = \{1, 1 \pm \log 2\}$ satisfies (i), yet (ii) is violated as $\log 2 \in \mathrm{span}_{\mathbb{Q}}\Delta_{\sigma(D)} = \mathrm{span}_{\mathbb{Q}}\{1, \log 2\}$. $\triangleleft$

The following theorem is the main result of the present section. Like Theorems 3.2 and 3.12, but without any assumptions on A, it extends to arbitrary dimensions the simple fact that for the sequence $(x\alpha^n y)$ with $\alpha \in \mathbb{R} \setminus \{0\}$ to be either Benford (if $xy \neq 0$) or trivial (if $xy = 0$) it is necessary and sufficient that $\log|\alpha|$ be irrational. To concisely formulate the result, call $(x^\top A^n y)$ and $(|A^n x|)$ with $A \in \mathbb{R}^{d\times d}$ and $x, y \in \mathbb{R}^d$ *terminating* if, respectively, $x^\top A^n y = 0$ or $A^n x = 0$ for all $n \geq d$; similarly, $(|A^n|)$ is terminating if $A^n = 0$ for all $n \geq d$.

Theorem 4.5. *Let A be a real $d \times d$-matrix. Then the following statements are equivalent:*

(i) *The set $\sigma(A)$ is non-resonant;*
(ii) *For every $x, y \in \mathbb{R}^d$, the sequence $(x^\top A^n y)$ is Benford or terminating.*

Moreover, if (i) *and* (ii) *hold then, for every $x \in \mathbb{R}^d$, the sequence $(|A^n x|)$ is Benford or terminating, and so is $(|A^n|)$.*

For a full proof of Theorem 4.5, the reader is referred to [6]. A simplified variant that applies to *most* matrices is given at the end of this section. From the argument, it will transpire that "terminating" can be replaced by "zero" (meaning "identically zero") whenever A is invertible, i.e. whenever $0 \notin \sigma(A)$. Before that, however, a few examples, corollaries, and remarks are presented.

Example 4.6. **(i)** As seen in Example 4.4, the (invertible) matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ associated with (3.3) has non-resonant spectrum. For every $x, y \in \mathbb{R}^2$, therefore, $(x^\top A^n y)$ is either Benford or zero. The latter happens precisely if x and y are proportional to, respectively, the eigenvector $\varphi e_1 + e_2$, corresponding to the eigenvalue φ of A, and to the eigenvector $\varphi e_2 - e_1$, corresponding to $-\varphi^{-1}$, or vice versa. In particular, the sequences $(F_n) = (e_1^\top A^n e_1)$ and $(L_n) = \big(e_1^\top A^n (3e_2 - e_1)\big)$ are Benford, as has already been observed in Examples 3.4 and 3.8.

Note that (F_n^2), for instance, is also Benford. This follows from Proposition 2.3 but can be seen directly as well by noticing that $\big(F_n^2 + \frac{2}{5}(-1)^n\big)$ is a solution of

$$x_n = 3x_{n-1} - x_{n-2}\,, \quad n \geq 3\,,$$

and the associated matrix $\begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}$ has non-resonant spectrum $\{\varphi^2, \varphi^{-2}\}$.

(ii) The 3×3-matrix B considered in Example 3.5 has non-resonant spectrum, as it has three real eigenvalues of different absolute value, none of which is of the form $\pm 10^{m/n}$ with $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. As in (i), every sequence $(x^\top B^n y)$ is either Benford or zero, with the latter being the case precisely if x and y are proportional to eigenvectors of B corresponding to two *different* eigenvalues. Note that even for this conclusion, which is stronger than the one reached in Example 3.5, it is not necessary to know $\sigma(B)$ explicitly. In fact, unlike in Example 3.5 it is not even necessary to know that $B^N > 0$ for some N. $\lhd$

Example 4.7. For the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ one finds $\sigma(A) = \{\sqrt{2} e^{\pm \imath \pi/4}\}$ which is resonant, as $\Delta_{\sigma(A)} = \{\frac{3}{4}, 1, \frac{5}{4}\}$. By Theorem 4.5, there must be $x, y \in \mathbb{R}^2$ for which $(x^\top A^n y)$ is neither Benford nor zero. Indeed, observe for instance that

$$e_1^\top A^n e_1 = e_1^\top 2^{n/2} \begin{bmatrix} \cos(\frac{1}{4}\pi n) & -\sin(\frac{1}{4}\pi n) \\ \sin(\frac{1}{4}\pi n) & \cos(\frac{1}{4}\pi n) \end{bmatrix} e_1 = 2^{n/2} \cos(\tfrac{1}{4}\pi n)\,, \quad n \in \mathbb{N}_0\,,$$

and hence $(e_1^\top A^n e_1)$ is neither Benford (because $e_1^\top A^{4n-2} e_1 = 0$ for all n) nor zero (because $e_1^\top A^{8n} e_1 = 2^{4n} \neq 0$ for all n). Note, however, that this of course does not rule out the possibility that *some* sequences derived from (A^n) may be Benford nevertheless. For instance, $(|A^n|) = (2^{n/2})$ is Benford. For another concrete example, fix any $x \neq 0$ and, for each $n \in \mathbb{N}$, denote by E_n the area of the triangle with vertices at $A^n x$, $A^{n-1} x$, and the origin. Then

$$E_n = \tfrac{1}{2} \left| \det(A^n x, A^{n-1} x) \right| = 2^{n-2} |x|^2\,, \quad n \in \mathbb{N}\,,$$

so (E_n) is Benford, see Fig. 4. Note also that while $\sigma(A)$ is resonant, the set $\sigma(A^4) = \{-4\}$ is not. (The reverse implication is easily seen to hold for all $d \in \mathbb{N}$ and $A \in \mathbb{R}^{d \times d}$: If $\sigma(A)$ is non-resonant then so is $\sigma(A^n)$ for every $n \in \mathbb{N}$.) $\lhd$

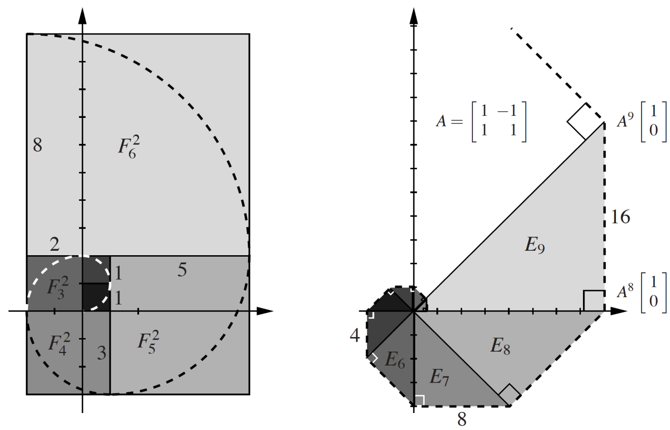

Fig. 4 Two Benford sequences, (F_n^2) and (E_n), derived from linear 2-dimensional systems, see Examples 4.6 and 4.7; note that $\sigma(A)$ is resonant for the matrix A associated with (E_n).

Example 4.8. For the matrix $B = \begin{bmatrix} 19 & 20 \\ 1 & 0 \end{bmatrix}$, first encountered in Example 3.10, $\sigma(B) = \{-1, 20\}$ is resonant. Consequently, there must be $x, y \in \mathbb{R}^2$ for which $(x^\top B^n y)$ is neither Benford nor zero. In essence, this has already been observed in Example 3.10, with $x = e_1$ and $y = e_1 - e_2$, for which $(x^\top B^n y) = \big((-1)^n\big)$. Note that failure of $(x^\top B^n y)$ to be Benford can occur only if $(20x_1 + x_2)(y_1 + y_2) = 0$. For *most* $x, y \in \mathbb{R}^2$, therefore, $(x^\top B^n y)$ is either Benford or zero. $\triangleleft$

Example 4.9. This example briefly reviews matrices and difference equations from earlier examples in the light of Theorem 4.5

(i) The matrices $A = \begin{bmatrix} 5 & 15 \\ 15 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$ both have resonant spectrum, $\sigma(A) = \{-10, 20\}$ and $\sigma(B) = \{2, 10\}$, which corroborates the observation, made in Example 3.6, that for some $x, y \in \mathbb{R}^2$, $(x^\top A^n y)$ is neither Benford nor zero, and similarly for B. Note, however, that $(x^\top A^n y)$ is Benford whenever $x, y \in \mathbb{R}^2$ are not multiples of $e_1 - e_2$, and hence for *most* $x, y \in \mathbb{R}^2$, whereas $(x^\top B^n y)$ can be Benford *only* if x or y is a multiple of $e_1 - e_2$.

(ii) While Theorem 3.2 did not apply to $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ in Example 3.9, every sequence $(x^\top B^n y)$ was seen to be Benford or terminating. This observation is consistent with $\sigma(B) = \{0, 2\}$ being non-resonant.

(iii) As is the case for every stochastic matrix, the matrix $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$ in Examples 3.10 and 3.11, has resonant spectrum $\sigma(A) = \{-\frac{1}{2}, 1\}$, and for most $x, y \in \mathbb{R}^2$, $(x^\top A^n y)$ is not Benford. The question, already raised in Example 3.11, whether, say, entries of $(A^{n+1} - A^n)$ can be Benford nevertheless is addressed in Section 5. $\triangleleft$

Example 4.10. Unlike in Theorem 3.2, within the wider scope of Theorem 4.5 the sequence $(|A^n x|)$ may, for every $x \in \mathbb{R}^d$, be Benford or terminating even if (i) and (ii) do not hold. Similarly, $(|A^n|)$ may be Benford. For an example, consider the 3×3-matrix

$$A = 10^{\varphi^2} \begin{bmatrix} \cos(2\pi\varphi) & -\sin(2\pi\varphi) & 0 \\ \sin(2\pi\varphi) & \cos(2\pi\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $\varphi = \frac{1}{2}(1+\sqrt{5})$, as usual, and hence $\varphi^2 = \varphi + 1$. The spectrum $\sigma(A) = \{10^{\varphi^2}, 10^{\varphi^2} e^{\pm \imath 2\pi\varphi}\}$ is resonant because

$$\tfrac{1}{2}(3+\sqrt{5}) = \varphi^2 = \log 10^{\varphi^2} \in \mathrm{span}_{\mathbb{Q}} \Delta_{\sigma(A)} = \mathrm{span}_{\mathbb{Q}}\{1, \sqrt{5}\}.$$

Nevertheless, for every $x \in \mathbb{R}^3$,

$$|A^n x| = 10^{n\varphi^2} \left| \begin{bmatrix} x_1 \cos(2\pi n\varphi) - x_2 \sin(2\pi n\varphi) \\ x_1 \sin(2\pi n\varphi) + x_2 \cos(2\pi n\varphi) \\ x_3 \end{bmatrix} \right| = 10^{n\varphi^2} |x|,$$

and since φ^2 is irrational, $(|A^n x|)$ is Benford whenever $x \neq 0$. Similarly, note that $10^{-n\varphi^2} A$ is an isometry for every n, and $(|A^n|) = (10^{n\varphi^2})$ is Benford. However, by Theorem 4.5, not *every* sequence $(x^\top A^n y)$ can be Benford or zero. That $(e_2^\top A^n e_1) = \big(10^{n\varphi^2} \sin(2\pi n\varphi)\big)$, for instance, is neither can be seen easily using Lemma 2.6.

Consider now also the matrix

$$B = 10^{\varphi^2} \begin{bmatrix} \cos(2\pi\varphi) & -\sin(2\pi\varphi) & \sin(\pi\varphi)\cos(\pi\varphi) \\ \sin(2\pi\varphi) & \cos(2\pi\varphi) & \sin(\pi\varphi)^2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Clearly, $\sigma(B) = \sigma(A)$, so the spectrum of B is resonant as well. A short calculation confirms that

$$B^n = 10^{n\varphi^2} \begin{bmatrix} \cos(2\pi n\varphi) & -\sin(2\pi n\varphi) & \sin(\pi n\varphi)\cos(\pi n\varphi) \\ \sin(2\pi n\varphi) & \cos(2\pi n\varphi) & \sin(\pi n\varphi)^2 \\ 0 & 0 & 1 \end{bmatrix}, \quad n \in \mathbb{N}_0,$$

from which it follows for instance that

$$|B^n \sqrt{2} e_3| = 10^{n\varphi^2} \sqrt{3 - \cos(2\pi n\varphi)}, \quad n \in \mathbb{N}_0,$$

and consequently, using $\varphi^2 = \varphi + 1$,

$$\big\langle \log |B^n \sqrt{2} e_3| \big\rangle = \big\langle n\varphi + \tfrac{1}{2} \log\big(3 - \cos(2\pi n\varphi)\big) \big\rangle = \langle f(n\varphi) \rangle,$$

with the smooth function $f : \mathbb{R} \to \mathbb{R}$ given by $f(s) = s + \frac{1}{2}\log\big(3 - \cos(2\pi s)\big)$. Recall that $(n\varphi)$ is u.d. mod 1. As in the proof of Lemma 2.6, consider the piecewise smooth map $T = \langle f \rangle$ on $[0,1)$ induced by f. Since T is a bijection of $[0,1)$ with non-constant slope, $\lambda_{0,1} \circ T^{-1} \neq \lambda_{0,1}$. This in turn means that $(|B^n e_3|)$, and in fact $(|B^n x|)$ for *most* $x \in \mathbb{R}^3$, is neither Benford nor zero. Similarly,

$$|B^n| = 10^{n\varphi^2} \sqrt{1 + \tfrac{1}{2}\sin(\pi n\varphi)^2 + \tfrac{1}{2}|\sin(\pi n\varphi)|\sqrt{4 + \sin(\pi n\varphi)^2}}\,, \quad n \in \mathbb{N}_0\,,$$

and a completely analogous argument shows that $(|B^n|)$ is not Benford either.

As evidenced by this example, the property of a matrix $A \in \mathbb{R}^{d\times d}$ that $(|A^n x|)$ is, for every $x \in \mathbb{R}^d$, either Benford or terminating is *not* a spectral property, i.e., it cannot be decided upon, at least for $d \geq 3$, by using $\sigma(A)$ alone. Similarly, $(|A^n|)$ being Benford is not a spectral property of A. $\lhd$

Remark. According to Theorem 4.5, non-resonance of $\sigma(A)$ is, for any invertible $A \in \mathbb{R}^{d\times d}$, equivalent to the widespread generation of Benford sequences of the form $(x^\top A^n y)$. Most $d\times d$-matrices are invertible with non-resonant spectrum, under a topological as well as a measure-theoretic perspective. To put this more formally, let

$$\mathscr{G}_d := \{A \in \mathbb{R}^{d\times d} : A \text{ is invertible and } \sigma(A) \text{ is non-resonant}\}\,.$$

Thus for example $\mathscr{G}_1 = \{[\alpha] : \alpha \in \mathbb{R}\setminus\{0\}, |\alpha| \neq 10^\rho \text{ for every } \rho \in \mathbb{Q}\}$. While the complement of $\mathscr{G}_d$ is dense in $\mathbb{R}^{d\times d}$, it is a topologically small set: $\mathbb{R}^{d\times d}\setminus\mathscr{G}_d$ is *of first category*, i.e. a countable union of nowhere dense sets. A (topologically) typical ("generic") $d\times d$-matrix therefore belongs to $\mathscr{G}_d$. Similarly, if A is an $\mathbb{R}^{d\times d}$-valued random variable, that is, a random matrix, whose distribution is a.c. with respect to the d^2-dimensional Lebesgue measure on $\mathbb{R}^{d\times d}$, then $\mathbb{P}(A \in \mathscr{G}_d) = 1$, i.e., with probability one A is invertible and $\sigma(A)$ non-resonant.

The next result is a corollary of Theorem 4.5 for difference equations and analogous to Corollary 3.7 but without any positivity assumptions on coefficients or initial values.

Theorem 4.11. *The following statements are equivalent for the difference equation*

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_{d-1}x_{n-d+1} + a_d x_{n-d}\,, \quad n \geq d+1\,, \tag{4.4}$$

where $a_1, a_2, \ldots, a_{d-1}, a_d \in \mathbb{R}$ *with* $a_d \neq 0$*:*

(i) *The set* $\{z \in \mathbb{C} : z^d = a_1 z^{d-1} + a_2 z^{d-2} + \ldots + a_{d-1}z + a_d\}$ *is non-resonant;*
(ii) *Every solution* (x_n) *of* (4.4) *is Benford, unless* $x_n \equiv 0$.

While the reader is again referred to [6] for a full proof of Theorem 4.11, a simplified argument applicable to most $a_1, a_2, \ldots, a_{d-1}, a_d$ is given at the end of the present section, following the very similar proof of Theorem 4.5.

Example 4.12. Since $\{z \in \mathbb{C} : z^2 = z+1\} = \{-\varphi^{-1}, \varphi\}$ is non-resonant, every solution of (3.3) except for $x_n \equiv 0$ is Benford, as was already seen in Example 4.6.

More generally, consider the second-order difference equation

$$x_n = a_1 x_{n-1} + a_2 x_{n-2}\,, \quad n \geq 3\,, \tag{4.5}$$

where a_1, a_2 are non-zero *integers*, and $a_2 > 0$. The set $\{z \in \mathbb{C} : z^2 = a_1 z + a_2\}$ consists of two real numbers with different absolute value, and is resonant if and only if one of them is of the form $\pm 10^N$ for some $N \in \mathbb{N}_0$. It follows that every solution (x_n) of (4.5), except for the trivial $x_n \equiv 0$, is Benford if and only if

$$|10^{2m} - a_2| \neq |a_1| 10^m, \quad \forall m = 0, 1, \ldots, \lfloor \log(|a_1| + a_2) \rfloor . \tag{4.6}$$

For example, for $a_1 = 2$, $a_2 = 5$, condition (4.6) reduces to $|1 - 5| \neq 2$. As the latter is obviously correct, every solution of

$$x_n = 2x_{n-1} + 5x_{n-2}, \quad n \geq 3,$$

is Benford unless $x_1 = x_2 = 0$. On the other hand, for

$$x_n = 19x_{n-1} + 20x_{n-2}, \quad n \geq 3,$$

(4.6) reads $|10^{2m} - 20| \neq 19 \cdot 10^m$ for $m = 0, 1$, which is violated for $m = 0$. This corroborates the observation, already made in Example 3.10, that $\big((-1)^n\big)$ is a solution that is neither Benford nor zero. $\lhd$

Remark. Earlier, weaker forms and variants of the implication (i)⇒(ii) in Theorems 4.5 and 4.11, or special cases thereof, can be traced back at least to [42] and may also be found in [5, 7, 9, 26, 35, 43]. The reverse implication (ii)⇒(i) seems to have been addressed only for $d < 4$, see [7, Thm.5.37]. A case in point is the 4×4-matrix

$$A = 10^{\sqrt{2}} \begin{bmatrix} \cos(2\pi\sqrt{3}) & -\sin(2\pi\sqrt{3}) & 0 & 0 \\ \sin(2\pi\sqrt{3}) & \cos(2\pi\sqrt{3}) & 0 & 0 \\ 0 & 0 & \cos(4\pi\sqrt{3}) & -\sin(4\pi\sqrt{3}) \\ 0 & 0 & \sin(4\pi\sqrt{3}) & \cos(4\pi\sqrt{3}) \end{bmatrix}.$$

In [7, Ex.5.36], it was observed that $(x^\top A^n y)$ is Benford or zero for every $x, y \in \mathbb{R}^4$ — despite the fact that A fails to be *Benford regular*, a property introduced there that is more restrictive than the non-resonance of $\sigma(A)$. This mismatch is resolved by Theorem 4.5, simply by noticing that $\sigma(A) = \{10^{\sqrt{2}} e^{\pm \imath 2\pi\sqrt{3}}, 10^{\sqrt{2}} e^{\pm \imath 4\pi\sqrt{3}}\}$ is indeed non-resonant.

Example 4.13. While satisfying theoretically, Theorems 4.5 and 4.11 may be difficult to use in practice, even if A is an integer 2×2-matrix (in Theorem 4.5), or $d = 2$ and a_1, a_2 are integers (in Theorem 4.11). To illustrate the basic difficulty, consider the innocent-looking difference equation

$$x_n = 2x_{n-1} - 5x_{n-2}, \quad n \geq 3. \tag{4.7}$$

For the set $Z = \{z \in \mathbb{C} : z^2 = 2z - 5\} = \{1 \pm 2\imath\} = \{\sqrt{5} e^{\pm \imath \arctan 2}\}$ it is not hard to see that $\Delta_Z = \{1, 1 \pm \frac{1}{\pi} \arctan 2\}$ satisfies (i) in Definition 4.2. Thus the non-resonance of Z is equivalent to $\log 5 \notin \operatorname{span}_{\mathbb{Q}} \Delta_Z = \operatorname{span}_{\mathbb{Q}}\{1, \frac{1}{\pi} \arctan 2\}$. While $\log 5$ and $\frac{1}{\pi} \arctan 2$ can both be shown to be transcendental, it seems to be unknown whether or not $1, \log 5, \frac{1}{\pi}\arctan 2$ are rationally independent [46]. In other words, it is not known whether the set Z is non-resonant. In the likely case that it is, every solution of (4.7), except for $x_n \equiv 0$, would be Benford; otherwise, none would. Experimental evidence strongly supports the former alternative, see Fig. 5. ◁

	1	2	3	4	5	6	7	8	9
$N = 10$	50.00	10.00	10.00	10.00	10.00	0.00	10.00	0.00	0.00
$N = 100$	20.00	11.00	13.00	7.00	10.00	12.00	12.00	9.00	6.00
$N = 1000$	29.80	17.20	13.80	8.90	7.80	6.80	6.50	5.20	4.00
$N = 10000$	29.99	17.23	12.78	9.51	7.92	6.61	6.01	5.19	4.76
exact BL	30.10	17.60	12.49	9.69	7.91	6.69	5.79	5.11	4.57

	0	1	2	3	4	5	6	7	8	9
$N = 10$	40.00	40.00	10.00	0.00	0.00	0.00	0.00	0.00	0.00	10.00
$N = 100$	15.00	11.00	8.00	10.00	15.00	7.00	10.00	9.00	7.00	8.00
$N = 1000$	10.80	10.70	11.10	9.50	11.60	10.10	10.70	9.30	8.40	7.80
$N = 10000$	11.44	11.48	11.18	9.98	10.22	10.03	9.27	9.04	8.85	8.51
exact BL	11.96	11.38	10.88	10.43	10.03	9.66	9.33	9.03	8.75	8.49

Fig. 5 Relative frequencies of the first (top) and second significant digits for the first N terms of the solution (x_n) of (4.7) with $x_1 = x_2 = 1$, see Example 4.13; the data suggests that (x_n) is Benford.

The practical difficulty alluded to in Example 4.13 can be avoided altogether only if all eigenvalues of A, or all roots of $z^d = a_1 z^{d-1} + a_2 z^{d-2} + \ldots + a_{d-1} z + a_d$ are *real*. In this situation, the following simple observation may be helpful.

Proposition 4.14. *A set $Z \subset \mathbb{R}$ is non-resonant if and only if every $z \in Z \setminus \{0\}$ satisfies*

$$\log|z| \notin \mathbb{Q} \quad \text{and} \quad |w| \neq |z| \quad \text{for every } w \in Z \setminus \{z\}\,.$$

The remainder of this section is devoted to presenting proofs of Theorems 4.5 and 4.11. Both proofs are given here only under the additional assumption that,

$$\text{for every } r > 0, \text{ the set } \sigma(A) \cap r\mathbb{S} \text{ contains at most two elements,} \tag{4.8}$$

i.e., the matrix A has at most two eigenvalues of modulus r, which may take the form of the real pair $-r, r$, or a non-real pair $\lambda, \overline{\lambda}$ with $|\lambda| = r$. (For complete proofs without this assumption, the reader is referred to [6]. Note that the matrices *not* satisfying (4.8) form a nowhere dense nullset in $\mathbb{R}^{d\times d}$.) For convenience, let

$$\sigma^+(A) := \{\lambda \in \sigma(A) : \Im\lambda \geq 0\} \setminus \{0\}\,.$$

Proof of Theorem 4.5: If $\sigma^+(A) = \varnothing$, then A is nilpotent, $\sigma(A) = \{0\}$ is non-resonant, every sequence $(x^\top A^n y)$ is identically zero for $n \ge d$, and the claimed equivalence trivially holds. Thus, from now on assume that $\sigma^+(A)$ is not empty.

Recall that, given any $x, y \in \mathbb{R}^d$, the value of $x^\top A^n y$ can be written in the form

$$x^\top A^n y = \Re\left(\sum\nolimits_{\lambda\in\sigma^+(A)} p_\lambda(n)\lambda^n\right), \quad n \ge d, \tag{4.9}$$

where p_λ is, for every $\lambda \in \sigma^+(A)$, a (possibly non-real) polynomial of degree at most $d-1$; moreover, p_λ is real whenever $\lambda \in \mathbb{R}$. The representation (4.9) follows for instance from the Jordan Normal Form Theorem. Note that p_λ also depends on x, y, but for the sake of notational clarity this dependence is not displayed explicitly.

To establish the asserted equivalence, assume first that $\sigma(A)$ is non-resonant and, given any $x, y \in \mathbb{R}^d$, that $p_\lambda \neq 0$ for some $\lambda \in \sigma^+(A)$. (Otherwise $x^\top A^n y = 0$ for all $n \ge d$.) Let

$$r := \max\{|\lambda| : \lambda \in \sigma^+(A), p_\lambda \neq 0\} > 0.$$

Recall that $\sigma(A) \cap r\mathbb{S}$ contains at most two elements. Note also that r and $-r$ cannot both be eigenvalues of A, as otherwise $\sigma(A)$ would be resonant. Hence either exactly one of the two numbers $r, -r$ is an eigenvalue of A, and $\log r$ is irrational, or else $\sigma(A) \cap r\mathbb{S} = \{re^{\pm\imath\pi\vartheta}\}$ with the appropriate irrational $0 < \vartheta < 1$, and

$$\log r \notin \operatorname{span}_{\mathbb{Q}}\{1, 1 \pm \vartheta\} = \operatorname{span}_{\mathbb{Q}}\{1, \vartheta\}.$$

In the former case, assume w.l.o.g. that r is an eigenvalue. (The case of $-r$ being an eigenvalue is completely analogous.) Recall that $|\lambda| < r$ for every other eigenvalue λ of A with $p_\lambda \neq 0$. Denote by $k \in \{0, 1, \ldots, d-1\}$ the degree of the polynomial p_r, and let $\gamma := \lim_{n\to\infty} p_r(n)/n^k$. Note that γ is non-zero and real. From (4.9), it follows that

$$|x^\top A^n y| = \left| p_r(n) r^n + \sum\nolimits_{\lambda\in\sigma^+(A):|\lambda|<r} p_\lambda(n)\lambda^n \right| = r^n n^k |\gamma + z_n|, \quad n \ge d,$$

with the (real) sequence (z_n) given by

$$z_n = \frac{p_r(n)}{n^k} - \gamma + \frac{1}{r^n n^k} \sum\nolimits_{\lambda\in\sigma^+(A):|\lambda|<r} p_\lambda(n)\lambda^n, \quad n \ge d.$$

Clearly, $\lim_{n\to\infty} z_n = 0$. Since $\log r$ is irrational and

$$\log|x^\top A^n y| = n\log r + k\log n + \log|\gamma + z_n|, \quad n \ge d,$$

Proposition 2.3 implies that $(x^\top A^n y)$ is Benford.

In the other case, the matrix A has $\lambda_0 = re^{\imath\pi\vartheta}$ and its conjugate $\overline{\lambda_0} = re^{-\imath\pi\vartheta}$ as eigenvalues, and $|\lambda| < r$ for every other eigenvalue λ of A with $p_\lambda \neq 0$. With k denoting the degree of p_{λ_0}, let again $\gamma := \lim_{n\to\infty} p_{\lambda_0}(n)/n^k$, and note that γ may now be non-real, yet is non-zero as before. Deduce from (4.9) that

$$|x^\top A^n y| = \left|\Re\left(p_{\lambda_0}(n) r^n e^{\imath \pi n \vartheta} + \sum\nolimits_{\lambda \in \sigma^+(A): |\lambda| < r} p_\lambda(n) \lambda^n\right)\right| = r^n n^k |\Re(\gamma e^{\imath \pi n \vartheta} + z_n)|,$$

with the (possibly non-real) sequence (z_n), given by

$$z_n = \left(\frac{p_{\lambda_0}(n)}{n^k} - \gamma\right) e^{\imath \pi n \vartheta} + \frac{1}{r^n n^k} \sum\nolimits_{\lambda \in \sigma^+(A): |\lambda| < r} p_\lambda(n) \lambda^n,$$

again satisfying $\lim_{n\to\infty} z_n = 0$. Since ϑ is irrational due to the non-resonance of $\sigma(A)$, the set $I := \{n \in \mathbb{N} : \Re(\gamma e^{\imath \pi n \vartheta} + z_n) = 0\}$ has density zero, that is, $\lim_{n\to\infty} \frac{1}{n} \#(I \cap \{1, 2, \ldots, n\}) = 0$, and

$$\log |x^\top A^n y| = n \log r + k \log n + \log |\Re(\gamma e^{\imath \pi n \vartheta} + z_n)|, \quad \forall n \in \mathbb{N} \setminus I. \tag{4.10}$$

(The reader familiar with the Skolem–Mahler–Lech Theorem [34, Thm.A] will notice that I is actually *finite*, though this much stronger property is not needed here.) Lemma 2.6 with $\vartheta_1 = \log r$, $\vartheta_2 = \vartheta$, $\alpha = k$, $z = \gamma$ shows that the sequence on the right in (4.10) is u.d. mod 1, and so is $(\log |x^\top A^n y|)$, by Proposition 2.3. Thus $(x^\top A^n y)$ is Benford, and the proof of (i)$\Rightarrow$(ii) is complete.

To establish the reverse implication, assume that $\sigma(A)$ is resonant. Then, for some $r_0 > 0$ and with $Z := \sigma(A) \cap r_0 \mathbb{S}$, the set Δ_Z contains rational numbers other than 1, or $\log r_0 \in \mathrm{span}_{\mathbb{Q}} \Delta_Z$, or both. Assume first that $1 + \rho \in \Delta_Z$ for some rational number $\rho > 0$. This implies that Z contains exactly two elements, either $r_0, -r_0$ or else $r_0 e^{\pm \imath \pi \rho}$. In the former case, let b, c be unit eigenvectors corresponding to, respectively, the eigenvalues r_0 and $-r_0$ of A, and let $x := y := b + c$. Then

$$x^\top A^n y = (b + c)^\top \left(r_0^n b + (-r_0)^n c\right) = (1 + b^\top c)\left(r_0^n + (-r_0)^n\right).$$

By the Cauchy–Schwarz inequality, $1 + b^\top c > 0$. Hence $x^\top A^n y = 0$ for all odd n but $x^\top A^n y > 0$ for all even n, and $(x^\top A^n y)$ is neither Benford nor terminating. In the case of non-real eigenvalues, there exist linearly independent unit vectors $b, c \in \mathbb{R}^d$ such that, for every $n \in \mathbb{N}_0$,

$$A^n b = r_0^n \cos(\pi n \rho) b - r_0^n \sin(\pi n \rho) c, \quad A^n c = r_0^n \sin(\pi n \rho) b + r_0^n \cos(\pi n \rho) c. \tag{4.11}$$

Hence with $x := y := b + c$,

$$\begin{aligned} x^\top A^n y &= r_0^n (b + c)^\top \big((\cos(\pi n \rho) + \sin(\pi n \rho)) b + (\cos(\pi n \rho) - \sin(\pi n \rho)) c\big) \\ &= 2(1 + b^\top c) r_0^n \cos(\pi n \rho), \end{aligned}$$

and again $x^\top A^n y = 0$ periodically but no identically. Thus $(x^\top A^n y)$ is neither Benford nor terminating.

It remains to consider the case where $\#(\Delta_Z \cap \mathbb{Q}) \leq 1$ for every $Z = \sigma(A) \cap r\mathbb{S}$ and $r > 0$, yet $\log r_0 \in \mathrm{span}_{\mathbb{Q}} \Delta_Z$ for some $r_0 > 0$. Again it is helpful to distinguish two cases: either $\sigma(A) \cap r_0 \mathbb{S} \subset \mathbb{R}$ or $\sigma(A) \cap r_0 \mathbb{S} \subset \mathbb{C} \setminus \mathbb{R}$. In the former case, exactly one of the two numbers r_0 and $-r_0$ is an eigenvalue of A. The argument for $-r_0$ being analogous, assume w.l.o.g. that $\sigma(A) \cap r_0 \mathbb{S} = \{r_0\}$. Then $\Delta_Z = \{1\}$ and hence $\log r_0$

is rational. Taking $x := y := b$, where b is any eigenvector of A corresponding to the eigenvalue r_0, yields $x^\top A^n y = r_0^n |b|^2$, and $(x^\top A^n y)$ is neither Benford nor terminating. In the other case, i.e. for $Z = \sigma(A) \cap r_0\mathbb{S} = \{r_0 e^{\pm \imath\pi\vartheta}\}$ with some irrational $0 < \vartheta < 1$, pick again linearly independent unit vectors $b, c \in \mathbb{R}^d$ such that (4.11) holds for all n, with ρ replaced by ϑ. With $x := y := b + c$, it follows that

$$\log|x^\top A^n y| = n \log r_0 + \log|2(1 + b^\top c)\cos(\pi n\vartheta)|\,.$$

Recall that $\log r_0 \in \mathrm{span}_{\mathbb{Q}}\{1, \vartheta\}$, by assumption. An application of Lemma 2.6 with $\vartheta_1 = \log r_0$, $\vartheta_2 = \vartheta$, $\alpha = 0$, $z = 2(1 + b^\top c) > 0$ and $z_n \equiv 0$ shows that the sequence $(x^\top A^n y)$ is not Benford. Clearly, it is not terminating either. Thus (ii)⇒(i), as claimed.

To complete the proof, the assertions regarding $(|A^n x|)$ and $(|A^n|)$ have to be verified. The above argument establishing (i)⇒(ii) can be used to verify the former assertion because, for every $x \in \mathbb{R}^d$, $|A^n x|^2 = \sum_{j=1}^d (e_j^\top A^n x)^2$, and so $(\log|A^n x|)$ is easily seen to be u.d. mod 1 using Lemma 2.6. Finally, if A is not nilpotent (otherwise $(|A^n|)$ obviously is terminating) assume first that $\sigma(A) \cap r_\sigma(A)\mathbb{S} = \{r_\sigma(A)e^{\pm\imath 2\pi\vartheta}\}$ for some irrational $0 < \vartheta < \frac{1}{2}$. Then, with the appropriate integer $k \geq 0$,

$$(A^n)^\top A^n = r_\sigma(A)^{2n} n^{2k} \big(B(n\vartheta)^\top B(n\vartheta) + C_n\big)\,, \quad n \in \mathbb{N}\,,$$

where the function $B : \mathbb{R} \to \mathbb{R}^{d\times d}$ is 1-periodic, real-analytic and does not vanish identically, and (C_n) is a sequence in $\mathbb{R}^{d\times d}$ with $|C_n| \to 0$. It follows that

$$\log|A^n| = n \log r_\sigma(A) + k \log n + \tfrac{1}{2} \log r_\sigma\big(B(n\vartheta)^\top B(n\vartheta) + C_n\big)\,.$$

Note that $r_\sigma\big(B(t)^\top B(t)\big) > 0$ for all but finitely many $t \in [0,1)$. By the assumed non-resonance of $\sigma(A)$, $\log r_\sigma(A) \notin \mathrm{span}_{\mathbb{Q}}\{1, \vartheta\}$, and hence [5, Lem.2.9] shows that $(|A^n|)$ is Benford. As the argument for the case $\sigma(A) \cap r_\sigma(A)\mathbb{S} \subset \mathbb{R}$ is completely analogous, the proof is complete. □

Proof of Theorem 4.11: Note first that the matrix A associated with (4.4) via (2.2) is invertible because $a_d \neq 0$. Hence (4.9) is valid for all $n \in \mathbb{N}$, and the sequence $(x^\top A^n y)$ vanishes identically unless $p_\lambda \neq 0$ for some $\lambda \in \sigma^+(A)$. Also note that $\sigma(A)$ simply equals $\{z \in \mathbb{C} : z^d = a_1 z^{d-1} + a_2 z^{d-2} + \ldots + a_{d-1}z + a_d\}$. Thus, if the latter set is non-resonant then $(x_n) = (e_d^\top A^n y)$ with $y = A^{-1}\sum_{j=1}^d x_{d+1-j} e_j$ either is Benford or else vanishes identically. This shows that (i)⇒(ii).

To establish the reverse implication, assume that $\sigma(A)$ is resonant, and distinguish cases just as in the above proof of Theorem 4.5. If $\sigma(A)$ is resonant due to failure of (i) in Definition 4.2 then, for some $r_0 > 0$ and rational number $\rho \in (0,1)$, either $\{-r_0, r_0\} \subset \sigma(A)$ or $\{r_0 e^{\pm\imath\pi\rho}\} \subset \sigma(A)$. In the former case, $(x_n) = \big(r_0^n + (-r_0)^n\big)$ solves (4.4) and is neither Benford nor zero. In the latter case, the same is true for $(x_n) = \big(r_0^n \cos(\pi n\rho)\big)$. If, on the other hand, $\sigma(A)$ is resonant due to failure of (ii) then, for some $r_0 > 0$ and irrational $\vartheta \in (0,1)$, either $r_0 \in \sigma(A)$ and $\log r_0 \in \mathbb{Q}$, or else $\{r_0 e^{\pm\imath\pi\vartheta}\} \subset \sigma(A)$ and $\log r_0 \in \mathrm{span}_{\mathbb{Q}}\{1, \vartheta\}$. In the first case, $(x_n) = (r_0^n)$ solves (4.4) and is neither Benford nor zero. In the second case,

$(x_n) = \big(r_0^n \cos(n\pi\vartheta)\big)$ is a non-zero solution of (4.4) that is not Benford since

$$\big(\log|r_0^n \cos(n\pi\vartheta)|\big) = \big(n\log r_0 + \log|\cos(n\pi\vartheta)|\big)$$

is not u.d. mod 1, by Lemma 2.6. Overall, (ii)⇒(i), and the proof is complete. □

5 An application to Markov chains

If A is a real $d\times d$-matrix and $\log r_\sigma(A)$ is rational then, as an immediate consequence of Theorem 4.5, the sequence $(x^\top A^n y)$ is, for most $x, y \in \mathbb{R}^d$, *not* Benford. If, in addition, A happens to have a positive power then, for instance, none of the entries $([A^n]_{jk})$ is Benford, according to Corollary 3.3. Even in this situation, however, it is quite possible that all entries of $(A^{n+1} - r_\sigma(A)A^n)$, and in fact most sequences $\big(x^\top(A^{n+1} - r_\sigma(A)A^n)y\big)$, are Benford. This phenomenon has already been observed in Examples 3.10 and 3.11 for the matrix $A = \begin{bmatrix} \frac12 & \frac12 \\ 1 & 0 \end{bmatrix}$, for which $\log r_\sigma(A) = \log 1 = 0$, and yet

$$A^{n+1} - A^n = \left(-\tfrac12\right)^{n+1} \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \quad n \in \mathbb{N}_0,$$

hence most sequences $(x^\top(A^{n+1} - A^n)y)$ are Benford. The purpose of the present section is to study this "cancellation of resonance" scenario and to demonstrate how it can be understood easily by utilizing the results from previous sections. The scenario is of particular interest in the case of *stochastic* matrices which often arise as transition probability matrices of finite-state Markov chains. (As observed in Example 4.9(iii), the spectrum of every stochastic matrix is resonant.) However, "cancellation of resonance" may occur whenever A has a dominant simple eigenvalue, and it is in this more general and transparent setting that the main result, Theorem 5.1 below, is formulated. The specific result for Markov chains is then a simple corollary (Corollary 5.4).

Assume, therefore, that the real $d\times d$-matrix A has a dominant eigenvalue λ_0 that is algebraically simple, i.e., $|\lambda| < |\lambda_0|$ for every $\lambda \in \sigma(A)\setminus\{\lambda_0\}$, and λ_0 is a simple root of the characteristic polynomial of A. Note that λ_0 is necessarily a real number, and $r_\sigma(A) = |\lambda_0|$. It is not hard to see that under these assumptions the limit

$$Q_A := \lim_{n\to\infty} \frac{A^n}{\lambda_0^n} \tag{5.1}$$

exists. Moreover, it is clear from (5.1) that $Q_A A = A Q_A = \lambda_0 Q_A$, but also $Q_A^2 = Q_A$. In fact, Q_A is nothing but the spectral projection associated with λ_0 and can also be represented in the form

$$Q_A = \frac{bc^\top}{b^\top c}\,, \tag{5.2}$$

where b, c are eigenvectors of, respectively, A and $A^\top$ corresponding to the eigenvalue λ_0. A dominant, algebraically simple eigenvalue is often observed in practice. For instance, it occurs whenever $A^N > 0$ for some $N \in \mathbb{N}$, see Proposition 3.1. (In this case even $Q_A > 0$.) But it also occurs for matrices such as e.g.

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},$$

of which no power is positive.

Consider now the sequences $(A^{n+1} - \lambda_0 A^n)$ and $(A^n - \lambda_0^n Q_A)$, both of which in a sense measure the speed of convergence in (5.1) and therefore are often of interest in their own right. Using the results of Section 4, the Benford behaviour of these sequences is easily analysed.

Theorem 5.1. *Assume $A \in \mathbb{R}^{d\times d}$ has a dominant eigenvalue λ_0 that is algebraically simple, and let Q_A be the associated projection according to* (5.2). *Then the following three statements are equivalent:*

(i) *The set $\sigma(A) \setminus \{\lambda_0\}$ is non-resonant;*
(ii) *The sequence $(x^\top (A^{n+1} - \lambda_0 A^n) y)$ is Benford or terminating for every $x, y \in \mathbb{R}^d$;*
(iii) *The sequence $(x^\top (A^n - \lambda_0^n Q_A) y)$ is Benford or terminating for every $x, y \in \mathbb{R}^d$.*

Proof. Since all assertions are trivially correct for $d = 1$, assume $d \ge 2$ from now on, and hence $\lambda_0 \ne 0$. As in the proof of Theorem 3.2, let $R := A - \lambda_0 Q_A$ and observe that $AR = RA$ as well as $Q_A R = 0 = R Q_A$, and hence

$$A^n = \lambda_0^n Q_A + R^n\,, \quad \forall n \ge 1\,. \tag{5.3}$$

Note that, for every $\lambda \in \sigma(A) \setminus \{\lambda_0\}$ and $x \in \mathbb{R}^d$ with $(A - \lambda I_d)^m x = 0$ (i.e., x is a generalized eigenvector of A corresponding to the eigenvalue $\lambda \ne \lambda_0$),

$$0 = c^\top (A - \lambda I_d)^m x = \big((A^\top - \lambda I_d)^m c \big)^\top x = (\lambda_0 - \lambda)^m c^\top x\,,$$

and so $c^\top x = 0$, which in turn implies $Q_A x = 0$, and hence $A^n x = R^n x$ for all n, by (5.3), and $\lambda \in \sigma(R)$. On the other hand, $Ab = \lambda_0 b = \lambda_0 Q_A b$ and therefore $Rb = 0$. Thus $0 \in \sigma(R)$. Also, if $Rx = \lambda_0 x$ for some $x \in \mathbb{R}^d$ then (5.3) yields

$$Q_A x = \lim\nolimits_{n\to\infty} \frac{A^n x}{\lambda_0^n} = Q_A x + x\,,$$

hence $x = 0$. In other words, $\lambda_0 \notin \sigma(R)$, and overall $\sigma(R) = (\sigma(A) \setminus \{\lambda_0\}) \cup \{0\}$, showing that $\sigma(R)$ is non-resonant if and only if $\sigma(A) \setminus \{\lambda_0\}$ is non-resonant. Moreover, deduce from (5.1) and (5.3) that

$$A^{n+1} - \lambda_0 A^n = R^n (R - \lambda_0 I_d)\,, \quad A^n - \lambda_0^n Q_A = R^n\,, \quad \forall n \ge 1\,.$$

Since $R - \lambda_0 I_d$ is invertible, the asserted equivalences are now obvious from Theorem 4.5. □

Example 5.2. **(i)** The (positive) matrix $B = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$ first encountered in Example 3.6 has the dominant simple eigenvalue $\lambda_0 = 10$. Thus Theorem 5.1 applies, with

$$Q_B = \lim_{n\to\infty} \frac{B^n}{\lambda_0^n} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since $\sigma(B) \setminus \{10\} = \{2\}$ is non-resonant, every sequence $(x^\top(B^{n+1} - 10B^n)y)$ and $(x^\top(B^n - 10^n Q_B)y)$ is Benford or terminating. This can also be seen directly from

$$B^{n+1} - 10B^n = -2^{n+2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad B^n - 10^n Q_B = 2^{n-1}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad n \in \mathbb{N}_0.$$

(ii) For the (non-negative) matrix $B = \begin{bmatrix} 19 & 20 \\ 1 & 0 \end{bmatrix}$ from Example 3.10, $\lambda_0 = 20$ is a dominant simple eigenvalue, and

$$Q_B = \frac{1}{21}\begin{bmatrix} 20 & 20 \\ 1 & 1 \end{bmatrix}.$$

However, $\sigma(B) \setminus \{20\} = \{-1\}$ is resonant, and hence some (in fact, most) sequences $(x^\top(B^{n+1} - 20B^n)y)$ and $(x^\top(B^n - 20^n Q_B)y)$ are neither Benford nor terminating. Again, this can be confirmed by an explicit calculation as well.

(iii) The matrix $C = \begin{bmatrix} 6 & -8 \\ 4 & -6 \end{bmatrix}$ does not have a dominant eigenvalue, as $\sigma(C) = \{\pm 2\}$, and hence Theorem 5.1 does not apply. Correspondingly, the limit $\lim_{n\to\infty} C^n/2^n$ does not exist. Note, however, that every entry of $(C^{n+1} - 2C^n)$, for instance, is Benford, as

$$C^{n+1} - 2C^n = 2(-2)^{n+1}\begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}, \quad n \in \mathbb{N}_0. \quad \triangleleft$$

Example 5.3. **(i)** The spectrum of

$$A = \begin{bmatrix} -10 & 15 & 15 \\ -24 & 29 & 27 \\ 24 & -24 & -22 \end{bmatrix}$$

equals $\sigma(A) = \{-10, 2, 5\}$, and hence is resonant, yet $\lambda_0 = -10$ is a dominant simple eigenvalue, and $\sigma(A) \setminus \{-10\} = \{2, 5\}$ is non-resonant. By Theorem 5.1, every entry of $(A^{n+1} + 10A^n)$ and $(A^n - (-10)^n Q_A)$, in particular, is Benford or terminating, where

$$Q_A = \begin{bmatrix} 1 & -1 & -1 \\ 2 & -2 & -2 \\ -2 & 2 & 2 \end{bmatrix}.$$

(ii) Consider the (non-negative) 3×3-matrix

$$B = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Clearly, $\lambda_0 = 3$ is a dominant eigenvalue, and $\sigma(B)\setminus\{3\} = \{2\}$ is non-resonant. However,

$$B^n = \begin{bmatrix} 3^n & n3^{n-1} & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & 2^n \end{bmatrix}, \quad n \in \mathbb{N}_0,$$

and so $\lim_{n\to\infty} B^n/3^n$ does not exist. The reason for this is that the eigenvalue λ_0, although dominant, is not simple. Thus Theorem 5.1 does not apply. Nevertheless, $(x^\top(B^{n+1} - 3B^n)y)$ is Benford or terminating for every $x, y \in \mathbb{R}^3$. $\lhd$

Remark. A close inspection of the proof of Theorem 5.1 shows that the assumption of algebraic simplicity for λ_0 can be relaxed somewhat. As a matter of fact, Theorem 5.1 remains unchanged if the dominant eigenvalue λ_0 is merely assumed to be *semi-simple*, meaning that its algebraic and geometric multiplicities coincide or, equivalently, that $A - \lambda_0 I_d$ and $(A - \lambda_0 I_d)^2$ have the same rank.

Arguably the most important application of Theorem 5.1 is to stochastic matrices. Recall that $A \in \mathbb{R}^{d\times d}$ is row-stochastic (column-stochastic) if $A \geq 0$ and the entries in each row (column) add up to 1; recall also that $r_\sigma(A) = 1 \in \sigma(A)$ for every (row- or column-) stochastic matrix. In probability textbooks, the letters P, Q etc. are traditionally used to denote stochastic matrices, a tradition adhered to for the remainder of this section. If $P \in \mathbb{R}^{d\times d}$ is a (row-) stochastic matrix, then it can naturally be interpreted as the matrix of 1-step transition probabilities of a time-homogeneous d-state Markov chain (X_n), i.e., (X_n) is a discrete-time Markov process on d symbols, $s_1, s_2, \ldots, s_d$, and, for every $n \in \mathbb{N}$,

$$\mathbb{P}(X_{n+1} = s_k | X_n = s_j) = [P]_{jk} \quad \forall j,k \in \{1,2,\ldots,d\}. \tag{5.4}$$

As a consequence of (5.4), the N-step transition probabilities are simply given by the entries of P^N, that is, for every $n \in \mathbb{N}$,

$$\mathbb{P}(X_{n+N} = s_k | X_n = s_j) = [P^N]_{jk} \quad \forall j,k \in \{1,2,\ldots,d\}.$$

Thus the long-term behaviour of the stochastic process (X_n) is governed by the sequence of (stochastic) matrices (P^n). Moreover, if $|\lambda| < 1$ for every eigenvalue $\lambda \neq 1$ of P then $Q_P := \lim_{n\to\infty} P^n$ exists and is itself a stochastic matrix. A common problem in many Markov chain models is to estimate Q_P through numerical simulation. In this context, the sequences $(P^{n+1} - P^n)$ and $(P^n - Q_P)$ are of special interest, as they both in a sense measure the speed of convergence of $P^n \to Q_P$. They are also rich sources of Benford sequences.

Corollary 5.4. [10, Thm.12] *Assume that the stochastic matrix $P \in \mathbb{R}^{d\times d}$ is irreducible and aperiodic, and let $Q_P := \lim_{n\to\infty} P^n$. If $\sigma(P)\setminus\{1\}$ is non-resonant then, for every $j,k \in \{1,2,\dots,d\}$, the sequences $([P^{n+1}-P^n]_{jk})$ and $([P^n - Q_P]_{jk})$ are Benford or terminating.*

Proof. Since P is irreducible and aperiodic, $P^N > 0$ for some $N \in \mathbb{N}$, and hence $\lambda_0 = 1$ is a dominant, algebraically simple eigenvalue of P. The claim then follows from Theorem 5.1. □

Example 5.5. For the stochastic matrix

$$P = \frac{1}{10}\begin{bmatrix} 9 & 0 & 1 \\ 6 & 3 & 1 \\ 1 & 0 & 9 \end{bmatrix},$$

$\sigma(P)\setminus\{1\} = \{\frac{3}{10}, \frac{4}{5}\}$ is non-resonant. Note that P fails to be irreducible, and hence Corollary 5.4 does not apply directly. However, $\lambda_0 = 1$ obviously is dominant and simple, and so Theorem 5.1 can be used to deduce that every entry of $(P^{n+1}-P^n)$ and $(P^n - Q_P)$ is Benford or terminating, with

$$Q_P = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}. \quad \triangleleft$$

Example 5.6. Consider the (irreducible and aperiodic) stochastic matrix

$$P = \frac{1}{30}\begin{bmatrix} 14 & 11 & 5 \\ 11 & 14 & 5 \\ 5 & 5 & 20 \end{bmatrix},$$

for which $\sigma(P)\setminus\{1\} = \{\frac{1}{10}, \frac{1}{2}\}$ is resonant. While Corollary 5.4 does not apply, Theorem 5.1 shows that there exist $x,y \in \mathbb{R}^3$ for which $(x^\top(P^{n+1}-P^n)y)$, for instance, is neither Benford nor terminating. For a concrete example that is neither, simply take $x = e_1$, $y = e_1 - e_2$, which yields $(x^\top(P^{n+1}-P^n)y) = (10^{-n})$. On the other hand, with

$$Q_P = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

it is straightforward to check that all entries of $(P^{n+1}-P^n)$ and $(P^n - Q_p)$ are Benford. Thus the non-resonance of $\sigma(P)\setminus\{1\}$ is not necessary for the latter property. In other words, the implication in Corollary 5.4 cannot in general be reversed. Moreover, the property asserted by Corollary 5.4, i.e. the property that all entries of $(P^{n+1}-P^n)$ and $(P^n - Q_P)$ are Benford or terminating, is not a spectral property of P. To see this, consider for example

$$\widetilde{P} = \frac{1}{10}\begin{bmatrix} 6 & 3 & 1 \\ 3 & 4 & 3 \\ 1 & 3 & 6 \end{bmatrix},$$

and note that $\sigma(\widetilde{P}) = \sigma(P)$ and $Q_{\widetilde{P}} = Q_P$. Again, it is readily confirmed that, for instance $([\widetilde{P}^{n+1} - \widetilde{P}^n]_{22}) = (-\frac{3}{5}10^{-n})$ and $([\widetilde{P}^n - Q_{\widetilde{P}}]_{22}) = (\frac{2}{3}10^{-n})$, and both sequences are neither Benford nor terminating. $\triangleleft$

The situation described in Corollary 5.4 is very common among stochastic matrices. To put this more formally, denote by $\mathscr{P}_d$ the family of all (row-) stochastic $d \times d$-matrices, that is

$$\mathscr{P}_d = \left\{ P \in \mathbb{R}^{d\times d} : P \geq 0, \sum\nolimits_{k=1}^{d} [P]_{jk} = 1 \ \ \forall j = 1,2,\ldots,d \right\}.$$

The set $\mathscr{P}_d$ is a compact and convex subset of $\mathbb{R}^{d\times d}$. For example,

$$\mathscr{P}_1 = \{[1]\} \quad \text{and} \quad \mathscr{P}_2 = \left\{ \begin{bmatrix} s & 1-s \\ 1-t & t \end{bmatrix} : 0 \leq s,t \leq 1 \right\}.$$

Note that $\mathscr{P}_d$ can be identified with a d-fold copy of the standard $(d-1)$-simplex, that is, $\mathscr{P}_d \simeq \{x \in \mathbb{R}^d : x \geq 0, \sum_{j=1}^{d} x_j = 1\}^d$, and hence carries the (normalized) $d(d-1)$-dimensional Lebesgue measure Leb. Consider now

$$\mathscr{H}_d := \left\{ P \in \mathscr{P}_d : P \ \text{ is irreducible and aperiodic, and } \sigma(P)\setminus\{1\} \ \text{ is non-resonant} \right\}.$$

Thus $\mathscr{H}_d$ is exactly the family of stochastic matrices covered by Corollary 5.4. For instance, $\mathscr{H}_1 = \{[1]\} = \mathscr{P}_1$,

$$\mathscr{H}_2 = \left\{ \begin{bmatrix} s & 1-s \\ 1-t & t \end{bmatrix} : 0 \leq s,t < 1,\, s+t = 1 \ \text{ or } \log|s+t-1| \notin \mathbb{Q} \right\},$$

and in both cases $\mathscr{H}_d$ constitues *most* of $\mathscr{P}_d$. The latter can be shown to be true in general: For every $d \in \mathbb{N}$, the complement of $\mathscr{H}_d$ in $\mathscr{P}_d$ is a set of first category and has Leb-measure zero. Thus if P is a $\mathscr{P}_d$-valued random variable, i.e. a random stochastic matrix, whose distribution is absolutely continuous (w.r.t. Leb, which means that $\mathbb{P}(P \in C) = 0$ whenever $C \subset \mathscr{P}_d$ and $\mathrm{Leb}(C) = 0$), then $\mathbb{P}(P \in \mathscr{H}_d) = 1$. Together with Corollary 5.4, this implies

Corollary 5.7. [10, Thm.17] *If the random stochastic matrix P has an absolutely continuous distribution then with probability one, P is irreducible and aperiodic, and every sequence $([P^{n+1} - P^n]_{jk})$ and $([P^n - Q_P]_{jk})$ is Benford or terminating.*

Note that for example the random stochastic matrix P has an absolutely continuous distribution whenever its d rows are chosen independently according to the same density on the standard $(d-1)$-simplex.

While the above genericity properties are very similar to the corresponding results for arbitrary matrices (see the *Remark* on p. 23), they do not follow directly from the latter. In fact, they are somewhat harder to prove, as they assert (topological as well as measure-theoretic) prevalence of $\mathscr{H}_d$ within the space $\mathscr{P}_d$ which, as a subset of $\mathbb{R}^{d\times d}$, is itself a nowhere dense nullset. The interested reader may want to consult [10] for details.

Acknowledgements This work was supported by an NSERC Discovery Grant. The authors are grateful to T. Hill, A. Weiss, and R. Zweimüller for helpful conversations.

References

1. Anderson, T.C., Rolen, L., Stoehr, R.: Benford's law for coefficients of modular forms and partition functions. Proc. Amer. Math. Soc. **139**, 1533–1541 (2011)
2. Arnold, V.I.: Ordinary Differential Equations. Springer, Berlin–Heidelberg–New York (1992)
3. Benford, F.: The law of anomalous numbers. Proc. Amer. Philos. Soc. **78**, 551–572 (1938)
4. Benford Online Bibliography, `http://www.benfordonline.net`
5. Berger, A.: Multi-dimensional dynamical systems and Benford's law. Discrete Contin. Dyn. Syst. **13**, 219–237 (2005)
6. Berger, A., Eshun, G.: A characterization of Benford's Law in discrete-time linear systems. In preparation (2014)
7. Berger, A., Hill, T.P.: A basic theory of Benford's Law. Probab. Surv. **8**, 1–126 (2011)
8. Berger, A., Hill, T.P.: Benfords law strikes back: no simple explanation in sight for mathematical gem. Math. Intelligencer **33**, 8591 (2011)
9. Berger, A., Siegmund, S.: On the distribution of mantissae in nonautonomous difference equations. J. Difference Equ. Appl. **13**, 829–845 (2007)
10. Berger, A., Hill, T.P., Kaynar, B., Ridder, A.: Finite-state Markov chains obey Benford's law. SIAM J. Matrix Anal. Appl. **32**, 665–684 (2011)
11. Brown, J.L., Duncan, R.L.: Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences. Fibonacci Quart. **8**, 482–486 (1970)
12. Carslaw, C.A.P.N.: Anomalies in Income Numbers: Evidence of Goal Oriented Behavior. The Accounting Review **63**, 321-327 (1988)
13. Cho, W.K.T., Gaines, B.J.: Breaking the (Benford) law: statistical fraud detection in campaign finance. Amer. Statist. **61**, 218–223 (2007)
14. Cohen, D.I.A., Katz, T.M.: Prime numbers and the first digit phenomenon. J. Number Theory **18**, 261–268 (1984)
15. Dajani, K., Kraikamp, C.: Ergodic theory of numbers. Carus Mathematical Monographs **29**. Mathematical Association of America, Washington DC (2002)
16. Deckert, J., Myagkov, M., Ordeshook, P.C.: Benford's Law and the Detection of Election Fraud. Political Analysis **19**, 245–268 (2011)
17. Diaconis, P.: The distribution of leading digits and uniform distribution mod 1. Ann. Probability **5**, 72–81 (1977)
18. Diekmann, A.: Not the first digit! Using Benford's law to detect fraudulent scientific data. J. Appl. Stat. **34**, 321–329 (2007)
19. Docampo, S., del Mar Trigo, M., Aira, M.J., Cabezudo, B., Flores-Moya, A.: Benfords law applied to aerobiological data and its potential as a quality control tool. Aerobiologia **25**, 275–283 (2009)
20. Duncan, R.L.: An application of uniform distributions to the Fibonacci numbers. Fibonacci Quart. **5**, 137–140 (1967)
21. Geyer, C.L., Williamson, P.P.: Detecting Fraud in Data Sets Using Benford's Law. Communications in Statistics: Simulation and Computation **33**, 229–246 (2004)
22. Hill, T.P.: Base-invariance implies Benford's law. Proc. Amer. Math. Soc. **123**, 887–895 (1995)

23. Hill, T.P.: The Significant-Digit Phenomenon. Amer. Math. Monthly **102**, 322–327 (1995)
24. Hill, T.P.: A Statistical Derivation of the Significant-Digit Law. Statistical Science **10**, 354–363 (1995)
25. Horn, R.A., Johnson, C.R.: Matrix analysis. Cambridge University Press, Cambridge (1985)
26. Kanemitsu, S., Nagasaka, K., Rauzy, G., Shiue, J-S.: On Benford's law: the first digit problem. In: Probability theory and mathematical statistics (Kyoto, 1986). Lecture Notes in Math. **1299**, 158–169 (1988)
27. Kontorovich, A.V., Miller, S.J.: Benford's law, values of L-functions and the $3x+1$ problem. Acta Arith. **120**, 269–297 (2005)
28. Kuipers, L.: Remark on a paper by R.L. Duncan concerning the uniform distribution mod 1 of the sequence of the logarithms of the Fibonacci numbers. Fibonacci Quart. **7**, 465–466, 473 (1969)
29. Kuipers, L., Niederreiter, H.: Uniform Distribution of Sequences. Wiley, New York (1974)
30. Lagarias, J.C., Soundararajan, K.: Benford's law for the $3x+1$ function. J. London Math. Soc. **74**, 289–303 (2006)
31. Mebane, W.R.: Fraud in the 2009 presidential election in Iran? Chance **23**, 6–15 (2010)
32. Miller, S.J., Nigrini, M.J.: Order statistics and Benford's law. Int. J. Math. Math. Sci. Art. ID 382948 (2008)
33. Moore, G.B., Benjamin, C.O.: Using Benford's Law for fraud detection. Internal Auditing **19**, 4–9 (2004)
34. Myerson, G., van der Poorten, A.J.: Some Problems Concerning Recurrence Sequences. Amer. Math. Monthly **102**, 698–705 (1995)
35. Nagasaka, K., Shiue, J-S.: Benford's law for linear recurrence sequences. Tsukuba J. Math. **11**, 341–351 (1987)
36. Newcomb, S.: Note on the frequency of use of the different digits in natural numbers. Amer. J. Math. **4**, 39–40 (1881)
37. Nigrini, M.J.: A taxpayer compliance application of Benford's law. The Journal of the American Taxation Association **18**, 72–91 (1996)
38. Nigrini, M.J.: Digital Analysis Using Benford's Law: Tests Statistics for Auditors. Global Audit Publications, Vancouver (2000)
39. Raimi, R.A.: The first digit problem. Amer. Math. Monthly **83**, 521–538 (1976)
40. Roukema, B.F.: A first-digit anomaly in the 2009 Iranian presidential election. Preprint arXiv:0906.2789 (2009).
41. Sambridge, M., Tkalčić, H., Jackson, A.: Benford's law in the natural sciences. Geophysical research letters **37**, L22301 (2010)
42. Schatte, P.: On H_∞-summability and the uniform distribution of sequences. Math. Nachr. **113**, 237–243 (1983)
43. Schatte, P.: On the uniform distribution of certain sequences and Benford's law. Math. Nachr. **136**, 271–273 (1988)
44. Schürger, K.: Extensions of Black–Scholes processes and Benford's law. Stochastic Process. Appl. **118**, 1219–1243 (2008)
45. Strzałka, D.: On some properties of Benford's law. J. Korean Math. Soc. **47**, 1055-1075 (2010)
46. Waldschmidt, M.: Diophantine Approximation on Linear Algebraic Groups. Springer, Berlin-Heidelberg-New York (2000)
47. Wintner, A.: On the cyclical distribution of the logarithms of the prime numbers. Q. J. Math. **6**, 65-68 (1935)
48. Wlodarski, J.: Fibonacci and Lucas Numbers tend to obey Benfords law. Fibonacci Quart. **9**, 87–88 (1971)