When normality and distribution normality coincide for nice classes of Cantor series

University of Texas at Austin Dynamical systems seminar

Sohail Farhangi (Joint work with Bill Mance)
Slides available on sohailfarhangi.com

April 21, 2025

Table of Contents

Base-b normality 1/2

Definition

For $b \in \mathbb{N}_{\geq 2}$, a number $x \in [0,1]$ is **normal base-b** if for $\ell \in \mathbb{N}$ and any word $w \in \{0,1,\cdots,b-1\}^{\ell}$, the word w appears in the base-b expansion of x with the correct frequency, and we denote the set of such numbers by $\mathcal{N}_b$. More explicitly, if $x = 0.x_1x_2\cdots x_n\cdots_b$, then

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N \mid w = x_nx_{n+1}\cdots x_{n+\ell-1}\} = b^{-\ell}. \quad (1)$$

Equivalently, x is **normal base-b** if the sequence $(b^n x)_{n=1}^{\infty}$ is uniformly distributed in $[0,1]$. More explicitly, if for any $0 < a < c < 1$ we have

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N \mid b^n x \pmod 1 \in (a,c)\} = c - a. \quad (2)$$

Base-b normality 2/2

We observe that for $x \in [0,1]$ with a base-2 expansion of $x = 0.x_1x_2\cdots x_n\cdots$, we have

$$2^n x \pmod 1 \in \begin{cases} [0, \frac{1}{2}) & \text{iff } x_{n+1} = 0 \\ [\frac{1}{2}, 1) & \text{iff } x_{n+1} = 1 \\ [0, \frac{1}{4}) & \text{iff } (x_{n+1}, x_{n+2}) = (0,0) \\ [\frac{1}{4}, \frac{2}{4}) & \text{iff } (x_{n+1}, x_{n+2}) = (0,1) \\ [\frac{2}{4}, \frac{3}{4}) & \text{iff } (x_{n+1}, x_{n+2}) = (1,0) \\ [\frac{3}{4}, \frac{4}{4}) & \text{iff } (x_{n+1}, x_{n+2}) = (1,1). \end{cases}$$

More generally, if $x = 0.x_1x_2\cdots x_n\cdots_b$ and $w = (w_1, \cdots, w_\ell) \in \{0,1,\cdots,b-1\}^\ell$, then

$$b^n x \pmod 1 \in [\sum_{j=1}^{\ell} \frac{w_j}{b^j}, \sum_{j=1}^{\ell} \frac{w_j}{b^j} + \frac{1}{b^\ell}) \text{ iff } (x_{n+1}, \cdots, x_{n+\ell}) = w.$$

Cantor series

Given a basic sequence (sequence of bases) $Q = (q_n)_{n=1}^{\infty} \in \mathbb{N}_{\geq 2}^{\mathbb{N}}$ and some $x \in [0,1]$, the base Q expansion $x = 0.x_1 x_2 \cdots x_n \cdots_Q$ with $0 \leq x_i < q_i$ is defined by

$$x = \sum_{n=1}^{\infty} x_i \left(\prod_{j=1}^{n} q_j \right)^{-1} = \frac{x_1}{q_1} + \frac{x_2}{q_1 q_2} + \frac{x_3}{q_1 q_2 q_3} + \cdots \qquad (3)$$

It is unique for all but countably many points in $[0,1]$.

Normality for a Cantor series

Given a basic sequence $Q = (q_n)_{n=1}^{\infty}$ and an $x = 0.x_1x_2\cdots x_n\cdots_Q \in [0,1]$, we say that x is **Q-normal** if for any block of digits $D = (d_1, \cdots, d_\ell) \in \mathbb{Z}_{\geq 0}^{\ell}$ appears in the base Q expansion of x with the correct frequency, and we denote the set of such x by $\mathcal{N}(Q)$. More precisely, if D satisfies

$$M_D(N) := \sum_{n=1}^{N} \left(\prod_{j=1}^{\ell} \frac{1}{q_{n+j}} \mathbb{1}_{[0,q_{n+j})}(d_j) \right) \underset{N\to\infty}{\longrightarrow} \infty, \text{ then}$$

$$\lim_{N\to\infty} \#\{1 \leq n \leq N \mid x_n x_{n+1} \cdots x_{n+\ell-1} = D\}/M_D(N) = 1.$$

Distribution normality for a Cantor series

Given a basic sequence $Q = (q_n)_{n=1}^{\infty}$ and an $x \in [0,1]$, we say that x is **Q-distribution normal** if the sequence $(x, q_1x, q_2q_1x, \cdots, q_nq_{n-1}\cdots q_2q_1x, \cdots)$ is uniformly distributed, and the set of such x is denoted by $\mathcal{DN}(Q)$. For a general basic sequence Q, the notions of Q-normality and Q-distribution normality don't need to be the same.

Theorem (Airey, Jackson, and Mance [1])

If $Q = (q_n)_{n=1}^{\infty}$ is such that $\lim_{n\to\infty} q_n = \infty$ and $\sum_{n=1}^{\infty} q_n^{-1} = \infty$, then $\mathcal{DN}(Q) \setminus \mathcal{N}(Q)$ and $\mathcal{N}(Q) \setminus \mathcal{DN}(Q)$ are $D_2(\Pi_3^0)$-complete.

See also [3] for results regarding the Hausdorff dimension of the difference sets, which are usually 1.

Remark: For any Q, the set $\mathcal{DN}(Q)$ has Lebesgue measure 1, and if Q is such that $\sum_{n=1}^{\infty} q_n^{-1} < \infty$, then $\mathcal{N}(Q) = [0,1]$.

Table of Contents

Dynamically generated basic sequences

Definition

A basic sequence $Q = (q_n)_{n=1}^{\infty}$ is **dynamically generated** if there exists a (continuous) measure preserving system $(X, \mathscr{B}, \mu, T)$ on a (not necessarily compact) polish space X, a continuous function $f : X \to \mathbb{N}_{\geq 2}$, and a μ-generic point $y \in X$ for which $q_n = f(T^n y)$.

1. If $X = \{1\}$ and T is (necessarily) the identity, then we recover base-b, i.e., $q_n = b$ for all n.
2. If $X = \{0, 1\}$, $Tx = x + 1 \pmod 2$, and $f(i) = b_i$, then we get $q_n = b_0$ if n is even and $q_n = b_1$ if n is odd. (See [2])
3. If $X = [0, 1]$, $Tx = x + \alpha \pmod 1$ with $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and $f = 2\mathbb{1}_{[0,\frac{1}{2})} + 3\mathbb{1}_{[\frac{1}{2},1)}$, then $(q_n)_{n=1}^{\infty}$ will be almost periodic sequence of 2s and 3s. (There is a model in which f is continuous)
4. We may also consider $f(x) = \left\lfloor \frac{1}{x} \right\rfloor$ in the previous example.

Uniform normality of a Cantor series

Let $Q = (q_n)_{n=1}^{\infty}$ be a dynamically generated basic sequences. For each block $B = (b_1, \cdots, b_\ell) \in \mathbb{N}_{\geq 2}^{\mathbb{N}}$, we have that

$$Q_B := \lim_{N\to\infty} \frac{1}{N} \underbrace{\#\{1 \leq n \leq N \mid (q_n, q_{n+1}, \cdots, q_{n+\ell-1}) = B\}}_{Q_B(N)}$$

exists, and it is 0 if and only if the block B never appears in Q. $x = 0.x_1x_2\cdots x_n\cdots_Q \in [0,1]$ is Q-**uniformly normal** if for any block of digits $D = (d_1, \cdots, d_\ell)$, and any block of bases $B = (b_1, \cdots, b_\ell)$ with $Q_B > 0$ and $b_j > d_j$ for all j, we have

$$\lim_{N\to\infty} \frac{1}{Q_B(N)} \#\{1 \leq n \leq N \mid (x_n, \cdots, x_{n+\ell-1}) = D \;\& \; (q_n, \cdots, q_{n+\ell-1}) = B\} = \prod_{j=1}^{\ell} \frac{1}{b_j},$$

and we denote the set of such x by $\mathcal{UN}(Q)$. I.e., $x \in \mathcal{UN}(Q)$ iff the pairs (D, B) of digits and bases occur with the correct frequency.

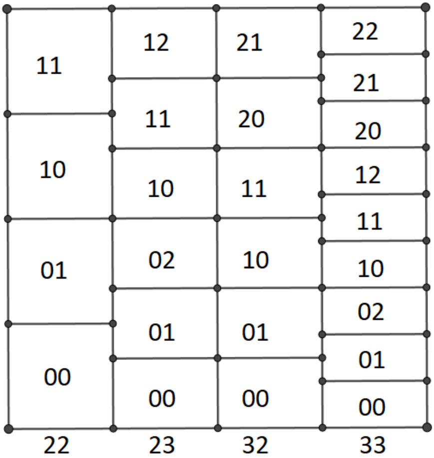
Uniform distribution normality of a Cantor series

Let $Q = (q_n)_{n=1}^{\infty}$ be a dynamically generated basic sequence, generated by the (continuous) m.p.s. $(X, \mathscr{B}, \mu, T)$ and the continuous function $f : X \to \mathbb{N}_{\geq 2}$. In particular, we have $q_n = f(T^n y)$ for some $y \in X$. Furthermore, let us assume that this representation is *minimal* in the sense that X is a polish space, f is continuous, and the topology generated by f and T is that of X. A number $x \in [0, 1]$ is Q-**uniformly distribution normal** if $(S^n(y, x))_{n=1}^{\infty}$ is uniformly distributed in $X \times [0, 1]$, where $S(y, x) = (Ty, f(y)x)$, and we denote the set of such x by $\mathcal{UDN}(Q)$.

A nice equivalence 1/2

Theorem

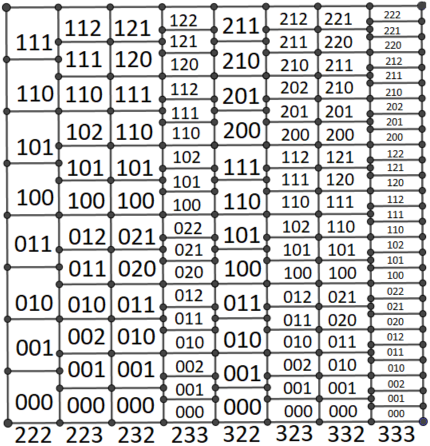
If $Q = (q_n)_{n=1}^{\infty}$ is a dynamically generated basic sequence, then $\mathcal{UN}(Q) = \mathcal{UDN}(Q)$.

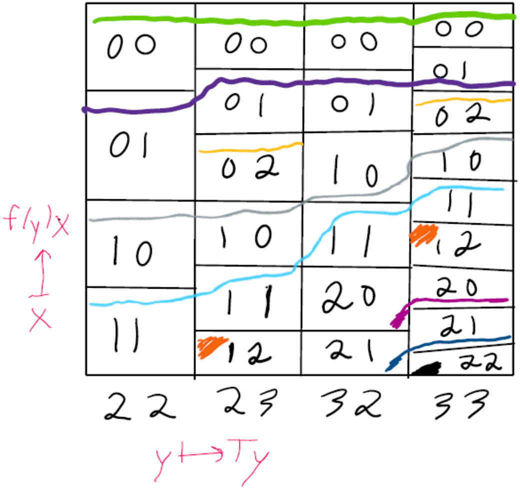


A nice equivalence 2/2

Theorem

If $Q = (q_n)_{n=1}^{\infty}$ is a dynamically generated basic sequence, then $\mathcal{UN}(Q) = \mathcal{UDN}(Q)$.





Q-normality means $S^n(y,x)$ visits the colored regions with the correct frequency, where $S(y,x) = (Ty, f(y)x)$

When $\mathcal{N}(Q) = \mathcal{DN}(Q)$

Theorem

Suppose that $Q = (q_n)_{n=1}^{\infty}$ is a basic sequence (dynamically) generated by $(X, \mathscr{B}, \mu, T, y, f)$ with $(X, \mathscr{B}, \mu, T)$ being ergodic and having zero (measurable) entropy, and $\int_X \ln(f) d\mu < \infty$. Then $\mathcal{N}(Q) = \mathcal{UN}(Q) = \mathcal{UDN}(Q) = \mathcal{DN}(Q)$. Furthermore, if $f : X \to \{b^n\}_{n=1}^{\infty}$ for some $b \geq 2$, then $\mathcal{N}(Q) = \mathcal{N}_b$.

Examples of zero entropy systems to consider.

1. $Tx = x + \alpha \pmod 1$ with $\alpha \in \mathbb{R}$.
2. T is any finite interval exchange transformation.
3. T is the Horocycle flow.
4. T is a rank 1 transformation.
5. $(q_n)_{n=1}^{\infty}$ "is" the Thue-Morse sequence with 2s and 4s.

When $\mathcal{DN}(Q) \not\to \mathcal{N}(Q)$

Theorem

There exists a dynamically generated sequences $Q = (q_n)_{n=1}^{\infty}$ and a sequence of digits $(E_n)_{n=1}^{\infty}$ for which $x = x_1 x_2 \cdots x_n \cdots_Q$ is distribution normal but not normal.

Sketch: Let $x \in [0, 1]$ be normal base-4 (which is the same as normal base-2). We will now construct a sequence $(q_n)_{n=1}^{\infty} \in \{2, 4\}^{\mathbb{N}}$ in which the 2s always appear in blocks of even size (groups of 2,4,6,...). We let $q_1 = q_2 = 2$ if $x \in [\frac{1}{2}, 1)$ and $q_1 = 4$ otherwise. We now replace x with $4x$ and repeat this procedure inductively to construct the rest of the q_n. The number x is Q-distribution normal by construction, but it is not Q-normal since the digits 2 and 3 never appear.

When $\mathcal{N}(Q) \not\to \mathcal{DN}(Q)$

Theorem

There exists a dynamically generated sequences $Q = (q_n)_{n=1}^{\infty}$ and a sequence of digits $(E_n)_{n=1}^{\infty}$ for which $x = E_1 E_2 \cdots E_n \cdots_Q$ is normal but not distribution normal.

Proof: Let Ω be a probability space and $(q_n(\omega_1))_{n=1}^{\infty}$ a sequences of i.i.d. random variables taking values in $\{2, 4\}$ (can also be done for {2,6}) with $\mathbb{P}(X_n = 2) = \mathbb{P}(X_n = 4) = \frac{1}{2}$. Consider

$$E_n(\omega_1)(\omega_2) = \begin{cases} 0 & \text{with probability } \frac{1}{2} + \epsilon \text{ if } q_n(\omega_1) = 2 \\ 1 & \text{with probability } \frac{1}{2} - \epsilon \text{ if } q_n(\omega_1) = 2 \\ 0 & \text{with probability } \frac{1}{4} - \epsilon \text{ if } q_n(\omega_1) = 4 \\ 1 & \text{with probability } \frac{1}{4} + \epsilon \text{ if } q_n(\omega_1) = 4 \\ 2 & \text{with probability } \frac{1}{4} \qquad \text{if } q_n(\omega_1) = 4 \\ 3 & \text{with probability } \frac{1}{4} \qquad \text{if } q_n(\omega_1) = 4. \end{cases}$$

Hot Spot Theorems

Pyatetskii-Shapiro [11] introduced what is now commonly known as the Hot Spot Theorem.

Theorem

Fix $b \in \mathbb{N}_{\geq 2}$ and $x \in [0,1]$. If there exists $C > 0$ such that for all $0 \leq a < c \leq 1$ we have

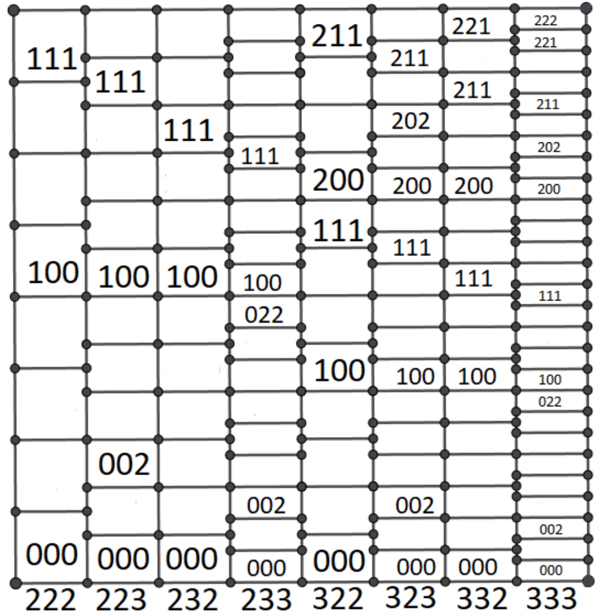
$$\limsup_{N\to\infty} \frac{1}{N}|\{1 \leq n \leq N \mid b^n x \pmod 1 \in (a,c)\}| \leq C(c-a), \quad (4)$$

then x is normal base b.

This result was generalized in [4, 9, 8, 5]. We have analogues of the strongest version of the Hot Spot Theorem proven in [5] for a large class of deterministic dynamically genereated basic sequences.

Table of Contents

Dynamics and normality?



We saw on slide 14 that $x \in \mathcal{N}(Q)$ if and only if $(S^n(y,x))_{n=1}^{\infty}$ equidistributes in a particular S-invariant collection of open sets in $X \times [0,1]$.

Question

What is the factor of $(X, \mathscr{B}, \mu, T)$ that is generated by this collection of open sets?

Selection rules

Kamae and Weiss [7, 6] proved the following selection rule:

Let $(n_k)_{k=1}^{\infty} \subseteq \mathbb{N}$ be an increasing sequence with positive lower density.

(i) If $(n_k)_{k=1}^{\infty}$ is deterministic, then for any $x = 0.x_1x_2\cdots x_n\cdots_b$ that is normal base b, $x' := 0.x_{n_1}x_{n_2}\cdots x_{n_k}\cdots_b$ will be normal base b.

(ii) If $(n_k)_{k=1}^{\infty}$ is not deterministic, then there exists a $x = 0.x_1x_2\cdots x_n\cdots_b$ that is normal base b for which $x' := 0.x_{n_1}x_{n_2}\cdots x_{n_k}\cdots_b$ is not normal base b.

Question: What are the selection rules for dynamically generated basic sequences?

We have some partial results, but not a complete classification.

Normality preservation through addition

Rauzy [10] characterized those $y \in [0,1]$ for which $y + \mathcal{N}_b \pmod 1 = \mathcal{N}_b$. Given a dynamically generated basic sequence Q, can we characterize those $y \in [0,1]$ for which $y + \mathcal{N}(Q) \pmod 1 = \mathcal{N}(Q)$? How about the same question for $\mathcal{DN}(Q)$ or $\mathcal{UN}(Q)$?

Difference Sets

Question: Suppose that Q is a dynamically generated basic sequence for which $\mathcal{N}(Q) \neq \mathcal{DN}(Q)$. What can be said about the descriptive complexity and the Hausdorff dimensions of the difference sets $\mathcal{DN}(Q) \setminus \mathcal{N}(Q), \mathcal{N}(Q) \setminus \mathcal{DN}(Q), \mathcal{N}(Q) \setminus \mathcal{UN}(Q)$, and $\mathcal{DN}(Q) \setminus \mathcal{UN}(Q)$, and $(\mathcal{N}(Q) \cap \mathcal{DN}(Q)) \setminus \mathcal{UN}(Q)$?

References I

[1] D. Airey, S. Jackson, and B. Mance.
Descriptive complexity in Cantor series.
J. Symb. Log., 87(3):1023–1045, 2022.

[2] D. Airey and B. Mance.
Normal equivalencies for eventually periodic basic sequences.
Indag. Math. (N.S.), 26(3):476–484, 2015.

[3] D. Airey and B. Mance.
On the Hausdorff dimension of some sets of numbers defined through the digits of their Q-Cantor series expansions.
J. Fractal Geom., 3(2):163–186, 2016.
MR3501345.

References II

[4] D. H. Bailey and M. Misiurewicz.
A strong hot spot theorem.
Proc. Amer. Math. Soc., 134(9):2495–2501, 2006.

[5] V. Bergelson and J. Vandehey.
A hot spot proof of the generalized Wall theorem.
Amer. Math. Monthly, 126(10):876–890, 2019.

[6] T. Kamae.
Subsequences of normal sequences.
Israel J. Math., 16:121–149, 1973.

[7] T. Kamae and B. Weiss.
Normal numbers and selection rules.
Israel J. Math., 21(2-3):101–110, 1975.

References III

[8] I. I. Pjateckiĭ-Šapiro.
On the distribution of the fractional parts of the exponential function.
Moskov. Gos. Ped. Inst. Uč. Zap., 108:317–322, 1957.

[9] A. G. Postnikov.
On the distribution of the fractional parts of the exponential function.
Doklady Akad. Nauk SSSR (N.S.), 86:473–476, 1952.

[10] G. Rauzy.
Nombres normaux et processus déterministes.
Acta Arith., 29(3):211–225, 1976.

References IV

[11] I. I. Šapiro Pyateckiĭ.
On the laws of distribution of the fractional parts of an exponential function.
Izv. Akad. Nauk SSSR Ser. Mat., 15:47–52, 1951.