A computable absolutely normal Liouville number

Verónica Becher   Pablo Ariel Heiber   Theodore A. Slaman
vbecher@dc.uba.ar   pheiber@dc.uba.ar   slaman@math.berkeley.edu

January 30, 2014

Abstract

We give an algorithm that computes an absolutely normal Liouville number.

1   The main result

The set of *Liouville numbers* is

$$\{x \in \mathbb{R}\backslash\mathbb{Q} : \forall k \in \mathbb{N},\ \exists q \in \mathbb{N},\ q > 1 \text{ and } ||qx|| < q^{-k}\}$$

where $||x|| = \min\{|x - m| : m \in \mathbb{Z}\}$ is the distance of a real number x to the nearest integer and other notation is as usual. Liouville's constant, $\sum_{k \geqslant 1} 10^{-k!}$, is the standard example of a Liouville number. Though uncountable, the set of Liouville numbers is small, in fact, it is null, both in Lebesgue measure and in Hausdorff dimension (see [6]).

We say that a *base* is an integer s greater than or equal to 2. A real number x is *normal to base s* if the sequence $(s^j x : j \geqslant 0)$ is uniformly distributed in the unit interval modulo one. By Weyl's Criterion [11], x is normal to base s if and only if certain harmonic sums associated with $(s^j x : j \geqslant 0)$ grow slowly. *Absolute normality* is normality to every base.

Bugeaud [6] established the existence of absolutely normal Liouville numbers by means of an almost-all argument for an appropriate measure due to Bluhm [3, 4]. The support of this measure is a perfect subset of the Liouville numbers, Bluhm's fractal. The Fourier transform of this measure decays quickly enough to ensure that those harmonic sums grow slowly on a set of measure one. Thus, Bugeaud's proof exhibits a nonempty set but does not provide a construction of an absolutely normal Liouville number.

A real number x is *computable* if there is a base s and an algorithm to output the digits for the base-s expansion of x, one after the other. In this note we show the following:

Theorem. *There is a computable absolutely normal Liouville number.*

We regard this result as a step into the ancient problem posed by Émile Borel [5] on exhibiting a natural instance of an absolutely normal number. Borel's understanding of "natural" may have been towards numbers that can be described geometrically (as π), analytically (as e), or algebraically (as $\sqrt{2}$). To our mind, algorithmic descriptions are also explicit, immediate and worthy of investigation.

We give an algorithm that determines a real number in the unit interval by recursively constructing a nested sequence of dyadic intervals. At each step the algorithm obtains a new subinterval containing sufficiently many points that satisfy, simultaneously, a better

approximation to the Liouville condition, and a better approximation to absolute normality. The number obtained by the algorithm is the unique point in the intersection of these intervals.

The discrepancy of a finite sequence of real numbers is a quantitative indicator of whether its elements are uniformly distributed modulo one in the unit interval. We translate between bounds on harmonic sums and bounds on discrepancy using a discrete version of LeVeque's Inequality (Theorem 2.4 in [9]), proved in Lemma 7.

Like Bugeaud, we use the ingredients of Bluhm's measure. However, we combine those ingredients differently so as to work within subintervals of the unit interval and make explicit the Liouville exponent and the level in Bluhm's fractal. By adapting an argument of Davenport, Erdös and LeVeque [8], we prove that the set of points in a given interval having small harmonic sums has large measure. See Lemma 10.

Our algorithm relies on the fact that the Fourier transform of Bluhm's measure decays not only quickly but also uniformly quickly over all intervals. For any given positive ϵ there is an extension length L with the following property. Consider any interval the form $[p2^{-a}, (p+1)2^{-a})$, for some non-negative integer p. So, the endpoints have a finite expansion in base 2, requiring at most a digits. Let b be the counterpart number of digits in the expansion of the left endpoint in base s (precisely, $b = \lceil a/\log_2 s \rceil$). Then, there is a level in Bluhm's fractal such that for the corresponding measure and for every ℓ as large as L, the set of reals x in this interval whose harmonic sum associated with $(s^j x : b \leqslant j < b + \ell)$ is below ϵ, has large measure. We prove this in Lemma 11.

In addition, we exploit another feature of discrepancy: as a function of finite sequences $(s^j x : a \leqslant j < b)$, it is continuous in two ways. One is with respect to the real variable x. That is, for any real numbers such that $|x - y|$ is small, if $(s^j x : a \leqslant j < b)$ has small discrepancy then $(s^j y : a \leqslant j < b)$ also has small discrepancy. Lemma 12 formalizes this idea giving quantitative estimates. The second way is with respect to the length of the sequence, given by the variables a and b. That is, for any c such that $c - a$ is non-negative and $c/(b - a)$ is small, if $(s^j x : a \leqslant j < b)$ has small discrepancy then both $(s^j x : a \leqslant j < b + c)$ and $(s^j x : a - c \leqslant j < b)$ also have small discrepancy. Lemma 13 formalizes this feature in a way that is conveniently applicable.

The algorithm constructs a real number x as the point in the intersection of a nested sequence of dyadic intervals. At each step, the algorithm determines one such dyadic interval, ensuring that the set of real numbers in it has small discrepancy and meets a designated Liouville exponent. However, at each step we do not consider the discrepancy of the entire sequence but the discrepancy of the current extension. Using the mentioned continuity of discrepancy, we conclude that the discrepancy of the limit point x output by the algorithm converges to zero.

We do not provide bounds on the time complexity for our algorithm. Without the Liouville condition, it is possible to compute absolutely normal numbers efficiently. Specifically, there are algorithms that output the first n digits of an absolutely normal number in time polynomial in n. The most efficient of these algorithms requires time just above quadratic [1], where speed is achieved by controlling, at each step, the size of the subinterval and how much progress is done towards absolute normality. The algorithm we present here is not consistent with such a strategy because it does not control the size of the subinterval at each step. So, the estimations of harmonic sums, which are inherently costly, are associated with necessarily long sequences. Constructing Liouville numbers that are normal to a given base, but not necessarily absolutely normal, as done in [10], admits a much simpler approach and can be done in linear time.

2 Bluhm's measure for computing Liouville numbers

We write $e(z)$ to denote e^z. We write the Fourier transform of a real function f as

$$\widehat{f}(x) = \int_{\mathbb{R}} f(t)e(-2\pi ixt) \, dt.$$

Recall that the Fourier transform of a positive bounded measure ν is defined, for $x \in \mathbb{R}$, by

$$\widehat{\nu}(x) = \int_{\mathbb{R}} e(-2\pi ixt) \, d\nu(t).$$

We write log without subscript for the logarithm in base e, and add a subscript for other bases.

2.1 Continuous replacements for step functions

We make use of measures which are supported by subintervals I of $[0, 1]$ and have Fourier transforms which decay quickly. Bluhm [3] gives examples of such and we employ them here.

Definition 1. Let R be a real number less than $1/2$. Define the function F_R on $[-1/2, 1/2]$ by

$$F_R(x) = \frac{15}{16}R^{-5}(R^2 - x^2)^2 \text{ when } |x| \leqslant R, \text{ and } F_R(x) = 0 \text{ otherwise.}$$

Let the Fourier series for $F_R(x)$ be denoted by

$$\sum_{n \in \mathbb{Z}} c_n^{(R)} e(2\pi inx).$$

Notice that the definition is such that

$$\int_{\mathbb{R}} F_R(x)dx = 1.$$

As Bluhm points out, the Fourier coefficients

$$c_n^{(R)} = \int_{-1/2}^{1/2} F_R(t) \, e(-2\pi int) \, dt$$

satisfy

$$c_0^{(R)} = 1, \qquad |c_n^{(R)}| \leqslant 1, \qquad \text{and} \qquad |c_n^{(R)}| \leqslant n^{-2}R^{-2}.$$

Definition 2. For a subinterval I of $[0, 1]$, let R_I be such that $4R_I$ is equal to the length of I. Let b be the center point of the interval I and let F_I be the translation of F_{R_I} by b, defined as

$$F_I(x) = \frac{15}{16}R_I^{-5}(R_I^2 - (x - b)^2)^2 \text{ when } |x - b| \leqslant R_I, \text{ and } F_I(x) = 0, \text{ otherwise.}$$

The support of F_I is contained in I and the analogous inequalities hold for the coefficients of the Fourier series for F_I,

$$c_0^{(I)} = 1, \qquad |c_n^{(I)}| \leqslant 1, \qquad \text{and} \qquad |c_n^{(I)}| \leqslant n^{-2}R_I^{-2}.$$

2.2 Bluhm-style measures

Bluhm [3, 4] showed that the set of Liouville numbers supports a Rajchman measure, that is, a positive measure whose Fourier transform vanishes at infinity. Bluhm's measure is the limit of a sequence of measures μ_k, for $k \in \mathbb{N}$. The measure μ_k is supported by a set of real numbers x such that there is at least one rational number p/q such that $0 < |x - p/q| < 1/q^k$. Here, rather than taking a limit of measures, we perform a sequence of finite steps to compute a real x in the limit and argue that for each step there is an appropriate action to take by appealing to an appropriate Bluhm-style measure.

Definition 3 ([3]). For every pair of integers m and k such that $k \geqslant 1$, let

$$E(m, k) = \bigcup_{\substack{m \leqslant q < 2m \\ \text{prime } q}} \{x \in \mathbb{R} : ||qx|| \leqslant q^{-1-k}\}.$$

As usual, we write C^2 for the class of functions whose first and second derivative both exist and are continuous.

Lemma 4 ([3, Lemma 3.2]). *There is a family of C^2 functions $g_{m,k}$, parametrized by the pairs of positive integers m and k, such that the support of $g_{m,k}$ is included in $E(m,k)$, $\widehat{g_{m,k}}(0) = 1$, and such that for every function Ψ in C^2 of compact support, for every positive integer k and for every positive real δ, there is an integer $M = M(\Psi, k, \delta)$ such that for every $m \geqslant M$ and for every $x \in \mathbb{R}$,*

$$|(\widehat{\Psi g_{m,k}})(x) - \widehat{\Psi}(x)| \leqslant \delta(1 + |x|)^{-1/(2+k)} \log(e + |x|) \ \log\log(e + |x|).$$

Bluhm defines $g_{m,k}$ by taking the sum of functions F_I for appropriate subintervals of those comprising $E(m, k)$ and then normalizing so that $\widehat{g_{m,k}}(0) = 1$.

Definition 5. We let ν_I be the measure on $[0, 1]$ obtained by integrating F_I. For m and k positive integers, we let $\nu_{I,m,k}$ be the measure on $[0, 1]$ obtained by integrating $F_I g_{m,k}$.

Lemma 6. *For every subinterval I of $[0, 1]$ and every positive integer k, there is a positive integer M such that for all $m \geqslant M$, $\nu_{I,m,k}(I) = 1$.*

Proof. $\nu_{I,m,k}(I)$ is equal to $\widehat{F_I g_{m,k}}(0)$, which by Lemma 4 can be made equal to $\widehat{F_I}(0) = 1$ by choosing m sufficiently large. $\qquad \square$

Observe that the support of $\nu_{I,m,k}$ is included in the support of $g_{m,k}$, which in turn is included in $E(m, k)$.

3 Lemmas

We say an interval is s-adic if it is of the form $(ps^{-a}, (p + 1)s^{-a})$ for non-negative integers a and p. We use $\langle a; s \rangle$ to denote $\lceil a/\log_2 s \rceil$. We write $\{x\}$ to denote the non-integral part of a real x. The cardinality of a set S is denoted by $\#S$.

The discrepancy of a finite sequence $(x_1, \ldots, x_n)$ of reals in the unit interval with respect to a fixed interval $[u, v]$ is

$$D([u, v], (x_1, \ldots, x_n)) = \left| \frac{\#\{j : 1 \leqslant j \leqslant n \text{ and } u \leqslant x_j < v\}}{n} - (v - u). \right|.$$

4

If we consider its discrepancy with respect to every subinterval in the unit interval, we have

$$D(x_1, \ldots, x_n) = \sup_{0 \leqslant u < v \leqslant 1} D([u, v], (x_1, \ldots, x_n))$$

Thus, a real number x is normal to base s exactly when the sequence $(\{s^j x\} : j \geqslant 0)$ is uniformly distributed in the unit interval, that is,

$$\lim_{n \to \infty} D(\{s^j x\} : 0 \leqslant j < n) = 0.$$

Lemma 7 ([2, Lemma 3.8]). *For any positive real ϵ there is a finite set of positive integers T and a positive real δ such that for any sequence of reals in the unit interval $(x_1, \ldots, x_n)$,*

$$\text{if for all } t \in T, \quad \frac{1}{n^2} \left| \sum_{j=1}^{n} e(t x_j) \right|^2 < \delta \text{ then } D(x_1, \ldots, x_n) < \epsilon.$$

Furthermore, such T and δ can be computed from ϵ.

Proof. By LeVeque's Inequality (see Theorem 2.4 [9])

$$D(x_1, \ldots, x_n) \leqslant \left(\frac{6}{\pi^2} \sum_{t=1}^{\infty} \frac{1}{t^2} \left| \frac{1}{n} \sum_{j=1}^{n} e(t x_j) \right|^2 \right)^{\frac{1}{3}}.$$

Since

$$\left| \frac{1}{n} \sum_{j=1}^{n} e(t x_j) \right|^2 \leqslant 1.$$

we get for each positive integer h,

$$\sum_{t=h+1}^{\infty} \frac{1}{t^2} \left| \frac{1}{n} \sum_{j=1}^{n} e(t x_j) \right|^2 \leqslant \sum_{t=h+1}^{\infty} \frac{1}{t^2} \leqslant \int_{h+1}^{\infty} x^{-2} dx \leqslant \frac{1}{h+1}.$$

Assume $\dfrac{1}{n^2} \left| \displaystyle\sum_{j=1}^{n} e(t x_j) \right|^2 < \dfrac{\epsilon^3}{2}$ for all positive integers t less than or equal to h. Thus,

$$\sum_{t=1}^{h} \frac{1}{t^2} \left| \frac{1}{n} \sum_{j=1}^{n} e(t x_j) \right|^2 + \sum_{t=h+1}^{\infty} \frac{1}{t^2} \left| \frac{1}{n} \sum_{j=1}^{n} e(t x_j) \right|^2 \leqslant \sum_{t=1}^{h} \frac{1}{t^2} \frac{\epsilon^3}{2} + \frac{1}{h+1} \leqslant \frac{\epsilon^3 \pi^2}{12} + \frac{1}{h+1}.$$

We can computably choose h so that $(6/\pi^2)\left(\epsilon^3 \pi^2/12 + 1/(h+1)\right)^{\frac{1}{3}} < \epsilon$. Then, $D(x_1, \ldots, x_n) < \epsilon$. This proves the lemma with $T = \{1, \ldots, h\}$ and $\delta = \epsilon^3/2$. $\qquad \square$

Lemma 8. *Let I be a dyadic interval with length 2^{-a} and let k be a positive integer. Let s be a base and let t be a positive integer. Then, there is an integer M such that for every $m \geqslant M$ and every positive integer ℓ,*

$$\int \frac{1}{\ell^2} \left| \sum_{n=\langle a;s \rangle + 1}^{\langle a;s \rangle + \ell} e(2\pi i t s^n x) \right|^2 d\nu_{I,m,k}(x) < \frac{100}{\ell}.$$

Moreover, M is uniformly computable from I, k, s and t.

Proof. Let

$$Z = \int \frac{1}{\ell^2} \left| \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n y) \right|^2 d\nu_{I,m,k}(y).$$

Then,

$$Z = \int \frac{1}{\ell^2} \sum_{n,q=1}^{\ell} e(2\pi i t s^{\langle a;s\rangle+n} y - 2\pi i t s^{\langle a;s\rangle+q} y) d\nu_{I,m,k}(y)$$

$$= \frac{1}{\ell^2} \left(\int \sum_{n=1}^{\ell} 1 d\nu_{I,m,k}(y) + \int \sum_{n\neq q} e(2\pi i t s^{\langle a;s\rangle+n} y - 2\pi i t s^{\langle a;s\rangle+q} y) d\nu_{I,m,k}(y) \right).$$

$$= \frac{1}{\ell} + \frac{1}{\ell^2} \sum_{n\neq q} \widehat{\nu_{I,m,k}}(t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q}).$$

We now give upper bounds for the values of $\widehat{\nu_{I,m,k}}$. Since R is one fourth of the length of I, which is 2^{-a}, by the above

$$\widehat{F_I}(x) \leqslant x^{-2} R^{-2} = \frac{2^{2a+4}}{x^2}.$$

According to Lemma 4, for sufficiently large m,

$$\left| \widehat{g_{m,k} F_I}(x) - \widehat{F_I}(x) \right| \leqslant \delta \theta_k(x),$$

where δ is a constant and can be made arbitrarily small by choice of m, and

$$\theta_k(x) = (1 + |x|)^{-1/(2+k)} \log(e + |x|) \, \log\log(e + |x|).$$

Let c be such that

$$\theta_k(x) \leqslant c(1 + |x|)^{-1/k}.$$

Notice c is uniformly computable from k. Thus,

$$|\widehat{\nu_{I,m,k}}(t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q})| = |\widehat{g_m F_I}(t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q})|$$

$$\leqslant \widehat{F_I}(t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q}) + \delta c(1 + |t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q}|)^{-1/k}$$

$$\leqslant \frac{2^{2a+4}}{|t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q}|^2} + \delta c \frac{1}{|t s^{\langle a;s\rangle+n} - t s^{\langle a;s\rangle+q}|^{1/k}}.$$

$$\leqslant \frac{2^4}{t^2 |s^n - s^q|^2} + \frac{\delta c}{t^{1/k} s^{\langle a;s\rangle/k}} \frac{1}{(s^n - s^q)^{1/k}}.$$

From which it follows,

$$Z \leqslant \frac{1}{\ell} + \frac{2}{\ell^2} \sum_{q=1}^{\ell-1} \sum_{n=q+1}^{\ell} \left(\frac{16}{t^2 |s^n - s^q|^2} + \frac{\delta c}{t^{1/k} s^{\langle a;s\rangle/k}} \frac{1}{(s^n - s^q)^{1/k}} \right).$$

6

Now, for positive p, analyze the sum

$$\sum_{q=1}^{\ell-1}\sum_{n=q+1}^{\ell}\frac{1}{(s^n-s^q)^{1/p}} \leqslant \sum_{q=1}^{\ell-1}\sum_{n=q+1}^{\ell}\frac{1}{(s^n/2)^{1/p}}$$

$$\leqslant 2^{1/p}\sum_{q=1}^{\ell-1}\sum_{n=q+1}^{\ell}s^{-n/p}$$

$$\leqslant 2^{1/p}\sum_{q=1}^{\ell-1}\int_q^\ell s^{-x/p}dx$$

$$\leqslant 2^{1/p}\sum_{q=1}^{\ell-1}\frac{p}{\log(s)}s^{-q/p}$$

$$\leqslant 2^{1/p}\frac{p}{\log(s)}\int_0^\ell s^{-x/p}dx$$

$$\leqslant 2^{1/p}\left(\frac{p}{\log(s)}\right)^2.$$

Hence,

$$Z \leqslant \frac{1}{\ell} + \frac{2}{\ell^2}\frac{16}{t^2}\,2^2\left(\frac{1/2}{\log(s)}\right)^2 + \frac{2}{\ell^2}\frac{\delta c}{t^{1/k}s^{\langle a;s\rangle/k}}\,2^{1/k}\left(\frac{k}{\log(s)}\right)^2.$$

By noting that t is a positive integer and choosing M so that for every $m \geqslant M$, δ is sufficiently small,

$$Z \leqslant \frac{1}{\ell} + \frac{32}{\ell^2\log(s)^2} + \frac{1}{\ell^2} < \frac{100}{\ell}.$$

By, [3, page 314], M is uniformly computable from the other parameters of the construction.
$\square$

Definition 9. Let s be a base. Let I be a dyadic interval of length 2^{-a}. Let ℓ and t be positive integers, and ϵ be a positive real. We define

$$A(I,\ell,s,t,\epsilon) = \left\{x \in I : \left|\frac{1}{\ell}\sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell}e(2\pi its^nx)\right| \geqslant \epsilon\right\}.$$

The following is an adaptation of an argument due to Davenport, Erdős and LeVeque in [8] reproduced as Lemma 1.8 in [7]).

Lemma 10. *Let s be a base, I be a dyadic interval of length 2^{-a} and k a positive integer. Let t be a positive integer. Let ϵ and δ be positive reals with ϵ less than $1/2$. Then there are positive integers M and L such that for all m greater than or equal to M*

$$\nu_{I,m,k}\left(\bigcup_{\ell\geqslant L}A(I,\ell,s,t,\epsilon)\right) < \delta.$$

Furthermore, L is uniformly computable from k, ϵ, δ, s and t, and does not depend upon I. On the other hand, M is uniformly computable from I, k, s and t.

Proof. By the definition of $A(I, \ell, s, t, \gamma)$,

$$\gamma^2 \, \nu_{I,m,k}(A(I,\ell,s,t,\gamma)) \leqslant \int \frac{1}{\ell^2} \left| \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n x) \right|^2 d\nu_{I,m,k}(x).$$

By Lemma 8, there is an M such that for all $m \geqslant M$ and all positive integers ℓ,

$$\int \frac{1}{\ell^2} \left| \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n x) \right|^2 d\nu_{I,m,k}(x) < \frac{100}{\ell}.$$

We fix this value of M and let m be greater than or equal to it. Then,

$$\nu_{I,m,k}(A(I,\ell,s,t,\gamma)) < \frac{100}{\gamma^2 \, \ell}.$$

Note that the above inequality holds for any real number γ. For the rest of the proof we fix parameters $I, s, t, \ \gamma = \epsilon/7$ and abbreviate $A(I, \ell, s, t, \epsilon/7)$ by writing A_ℓ, for varying values of ℓ. Set $\ell_1 = 1$ and $\ell_{j+1} = \lceil \ell_j/(1-\epsilon/7) \rceil + 1$ for $j \geqslant 1$ such that

$$\frac{\ell_{j+1} - \ell_j}{\ell_{j+1}} > \frac{\epsilon}{7} \quad \text{and} \quad \frac{\ell_{j+1} - \ell_j}{\ell_j} < 2\frac{\epsilon}{7} + \frac{2}{\ell_j}.$$

For each j, let h_j be the integer in $[\ell_j, \ell_{j+1})$ that minimizes $\nu_{I,m,k}(A_{h_j})/h_j$. Hence,

$$\sum_{\ell=\ell_j}^{\ell_{j+1}-1} \frac{\nu_{I,m,k}(A_\ell)}{\ell} \geqslant (\ell_{j+1} - \ell_j)\frac{\nu_{I,m,k}(A_{h_j})}{h_j} \geqslant \frac{\epsilon}{7}\nu_{I,m,k}(A_{h_j}).$$

Thus,

$$100\frac{7^3}{\epsilon^3} \sum_{\ell=\ell_j}^{\ell_{j+1}-1} \frac{1}{\ell^2} > \nu_{I,m,k}(A_{h_j}),$$

and so for every j,

$$100\frac{7^3}{\epsilon^3} \frac{1}{\ell_j - 1} > \sum_{p=j}^{\infty} \nu_{I,m,k}(A_{h_p}).$$

Let j_0 be minimal such that

$$\ell_j - 1 \geqslant \frac{100}{\delta}\frac{7^3}{\epsilon^3} \quad \text{and} \quad \frac{2}{\ell_{j_0}} < \frac{\epsilon}{7}.$$

Note that j_o does not depend on I. Then, by the first inequality,

$$\delta > \sum_{j=j_0}^{\infty} \nu_{I,m,k}(A_{h_j}).$$

Now, consider an $x \in I$ such that for all $j \geqslant j_0$, $x \notin A_{h_j}$, or equivalently,

$$\left| \frac{1}{h_j} \sum_{\ell=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^\ell x) \right| < \frac{\epsilon}{7},$$

and let ℓ be a positive integer greater than or equal to ℓ_{j_0}. Let j be such that $\ell_j \leqslant \ell < \ell_{j+1}$. Note that,

$$\frac{1}{\ell} \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n x) - \frac{1}{h_j} \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^n x) =$$

$$\frac{1}{\ell} \left(\sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n x) - \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^n x) \right) + \left(\frac{1}{\ell} - \frac{1}{h_j} \right) \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^n x).$$

It follows that for such x,

$$\left| \frac{1}{\ell} \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^n x) - \frac{1}{h_j} \sum_{n=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^n x) \right| < \left(\frac{\ell_{j+1} - \ell_j}{\ell_j} \right) + \left(\frac{\ell_{j+1} - \ell_j}{\ell_j} \right)$$

$$< \left(2\frac{\epsilon}{7} + \frac{2}{\ell_j} \right) + \left(2\frac{\epsilon}{7} + \frac{2}{\ell_j} \right)$$

$$< 6\frac{\epsilon}{7}.$$

and by the triangle inequality, for this x,

$$\left| \frac{1}{\ell} \sum_{\ell=\langle a;s\rangle+1}^{\langle a;s\rangle+\ell} e(2\pi i t s^\ell x) \right| < \left| \frac{1}{h_j} \sum_{\ell=\langle a;s\rangle+1}^{\langle a;s\rangle+h_j} e(2\pi i t s^\ell x) \right| + 6\frac{\epsilon}{7} < \frac{\epsilon}{7} + 6\frac{\epsilon}{7} = \epsilon.$$

Then, $x \notin A(I, \ell, s, t, \epsilon)$. So, $\bigcup_{\ell \geqslant \ell_{j_0}} A(I, \ell, s, t, \epsilon)$ is contained in $\bigcup_{j \geqslant j_0} A_{h_j}$ and thereby $\nu_{I,m,k}\left(\bigcup_{\ell \geqslant \ell_{j_0}} A(I, \ell, s, t, \epsilon) \right) < \delta$. This proves the lemma for M as above and L equal to ℓ_{j_0}, with the observation that since j_0 did not depend on I neither does L. $\qquad\square$

Lemma 11. *Let S be a set of bases, I be a dyadic interval of length 2^{-a} and k a positive integer. Let ϵ be a positive real number. Then, there are positive integers M and L such that for every $m \geqslant M$,*

$$\nu_{I,m,k}(I) = 1, \text{ and}$$

$$\nu_{I,m,k}\left(\{y \in I : \forall s \in S, \forall \ell \geqslant L, \ D(\{s^j y\} : \langle a;s\rangle \leqslant j < \langle a + \ell; s\rangle) < \epsilon\} \right) \geqslant \tfrac{1}{2}.$$

Furthermore, L is uniformly computable from k, S and ϵ, and does not depend on I; M is uniformly computable from I, k, S and ϵ.

Proof. Let T and γ be, respectively, the set of positive integers and the real number determined by Lemma 7 with input ϵ.

For each $s \in S$ and $t \in T$ apply Lemma 10 with ϵ equal to γ, and with δ equal to $1/(2\#T\#S)$. Let M_0 and L_0 be, respectively, the maximum of the values of M and L output by Lemma 10 in each case. Then, for every $m \geqslant M_0$,

$$\sum_{t \in T} \sum_{s \in S} \nu_{I,m,k} \left(\bigcup_{\ell \geqslant L_0} A(I, \ell, s, t, \gamma) \right) < \sum_{t \in T} \sum_{s \in S} \frac{1}{2\#T\#S} = \frac{1}{2}.$$

The proof is completed by taking L to be the minimum integer such that for every $s \in S$, $\langle a + L; s \rangle \geqslant \langle a; s \rangle + L_0$, and by taking M to be the maximum between M_0 and the value M determined by Lemma 6. $\qquad \square$

Lemma 12. *For every base s and for every positive real ϵ there is an integer L with the following property. Let $[u, v]$ be any subinterval of $[0, 1]$ such that*

$$2\epsilon < v - u, \quad 2\epsilon < u \text{ in case } u \neq 0, \text{ and } 2\epsilon < 1 - v \text{ in case } v \neq 1.$$

For all $\ell \geqslant L$ and for all real numbers x and y satisfying $|x - y| < 2^{-\ell}$, if for each interval $J \subseteq [0, 1]$ with endpoints in $\{0, u \pm \epsilon, v \pm \epsilon, 1\}$, $D(J, (\{s^j x\} : j < \langle \ell; s \rangle)) < \epsilon$, then

$$D([u, v], (\{s^j y\} : j < \langle \ell; s \rangle)) < 8\epsilon.$$

Proof. Let s, $[u, v]$ and ϵ be fixed as above. Let L be an integer such that $\lceil |\log_s(\epsilon)| \rceil$ is less than $\epsilon \langle L; s \rangle$. Suppose that ℓ is greater than or equal to L and that x and y are real numbers such that $|x - y| < 1/2^\ell$ and for each interval J with endpoints contained in $\{0, u \pm \epsilon, v \pm \epsilon, 1\}$ $D(J, (\{s^j x\} : j < \langle \ell; s \rangle)) < \epsilon$.

We first show that $\#\{j : \{s^j y\} \in [u, v] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\}$ is greater than $(|v - u| - 8\epsilon)\langle \ell; s \rangle$. Consider the interval $[u + \epsilon, v - \epsilon]$. By assumption,

$$\#\{j : \{s^j x\} \in [u + \epsilon, v - \epsilon] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\} > \big(((v - \epsilon) - (u + \epsilon)) - \epsilon\big)\langle \ell; s \rangle$$
$$> ((v - u) - 3\epsilon)\langle \ell; s \rangle.$$

If $s^j 2^{-\ell}$ is less than ϵ and $x \in [u + \epsilon, v - \epsilon]$ then $y \in [u, v]$. Further, $s^j 2^{-\ell} > \epsilon$ only when $j > \langle \ell; s \rangle - |\log_s(\epsilon)|$ and, by choice of L, there are at most $\epsilon \langle \ell; s \rangle$ many such integers j less than or equal to $\langle \ell; s \rangle$. Then,

$$\#\{j : \{s^j y\} \in [u, v] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\} > ((v - u) - 3\epsilon)\langle \ell; s \rangle - \epsilon \langle \ell; s \rangle$$
$$> ((v - u) - 4\epsilon)\langle \ell; s \rangle.$$

Similarly, regarding only the fully non-trivial case in which $0 < u$ and $v < 1$,

$$\#\{j : \{s^j y\} \in [0, u] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\} > (u - 4\epsilon)\langle \ell; s \rangle$$
$$\#\{j : \{s^j y\} \in [v, 1] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\} > ((1 - v) - 4\epsilon)\langle \ell; s \rangle$$

and so

$$((v - u) + 8\epsilon)\langle \ell; s \rangle > \#\{j : \{s^j y\} \in [u, v] \text{ and } 0 \leqslant j < \langle \ell; s \rangle\} > ((v - u) - 8\epsilon)\langle \ell; s \rangle,$$

as required. $\qquad \square$

Lemma 13. *Let ϵ be a positive real. Let any sequence of reals in the unit interval $(x_1, \ldots, x_n)$ of length n. Let $(y_1, \ldots, y_\ell)$ is a sequence of reals of length ℓ such that $D(y_1, \ldots, y_\ell) < \epsilon$. If $n < \epsilon \ell$ then for all $k \leqslant \ell$, $D(x_1, \ldots, x_n, y_1, \ldots, y_k) < 2\epsilon$ and $D(y_1, \ldots, y_k, x_1 \ldots, x_n) < 2\epsilon$.*

Proof. Immediate from the definition. $\qquad \square$

4 Proof of the Theorem

For n an integer greater than or equal to 2 and ϵ a positive real number, let $L_{11}(n, \epsilon)$ and $L_{12}(n, \epsilon)$ be the supremum of the output numbers L in Lemmas 11 and 12, respectively, for inputs s, a base less than or equal to n, Liouville exponent k equal to n, and ϵ, a positive real number. Without loss of generality, we assume that L_{11} and L_{12} increase as the first argument increases and as the second argument decreases.

Definition 14. An interval $I \subseteq [0, 1]$ *meets the Liouville condition for exponent k* if for any real $x \in I$ there is an integer $q > 1$ such that $||qx|| < q^{-k}$.

Then, a real number is Liouville when for each exponent k there is an interval that contains x and meets the Liouville for exponent k.

4.1 Algorithm

We proceed by recursion to define a sequence of dyadic intervals $[x_n, x_n + 2^{-a_n})$, that is to say that a_n is a non-negative integer and x_n is of the form $p/2^{a_n}$ with $0 \leqslant p < 2^{a_n}$. To simplify notation, let $\epsilon_n = 1/8^n$.

Let $x_0 = 0$ and $a_0 = 0$. Given $[x_n, x_n + 2^{-a_n})$ from the previous step, let $[x_{n+1}, x_{n+1} + 2^{-a_{n+1}})$ be the dyadic interval minimizing a_{n+1} and breaking ties by minimizing x_n, with the following conditions:

- $[x_{n+1}, x_{n+1} + 2^{-a_{n+1}}) \subseteq [x_n, x_n + 2^{-a_n})$.

- $[x_{n+1}, x_{n+1} + 2^{-a_{n+1}}]$ meets the Liouville condition for exponent $n + 1$.

- $a_{n+1} > L_{12}(n + 1, \epsilon_{n+1}/16)$.

- For every base s less than or equal to $n + 1$,

$$\langle a_{n+1}; s \rangle \epsilon_{n+1}/16 > \langle L_{11}(n + 2, \epsilon_{n+2}/16); s \rangle.$$

- For every base s less than or equal to n, for all nontrivial intervals $J \subseteq [0, 1]$ with rational endpoints of the form $[p_1/8^{n-2}, p_2/8^{n-2}]$ and for all integers $\ell \in [a_n, a_{n+1})$,

$$D(J, (\{s^j x_{n+1}\} : j < \langle \ell; s \rangle)) < 128\epsilon_n.$$

- For every base s less than or equal to $n + 1$, for all nontrivial intervals $J \subseteq [0, 1]$ with rational endpoints of the form $[p_1/8^{n+1}, p_2/8^{n+1}]$,

$$D(J, (\{s^j x_{n+1}\} : j < \langle a_{n+1}; s \rangle)) < \epsilon_{n+1}.$$

4.2 Verification

We first check by induction that the sequence $(x_n : n \geqslant 0)$ is well defined. We have specified x_0 and a_0 explicitly. It is immediate that for step $n = 1$ there is a suitable choice for x_1 and a_1. Assume that the sequence is defined up to and including $[x_n, x_n + 2^{-a_n})$, where $n \geqslant 1$. Let I be the interval $[x_n, x_n + 2^{-a_n}]$ and let S be the set of bases less than or equal to $n + 1$. Apply Lemma 11 for $\epsilon = \epsilon_{n+1}/16, k = n + 1, I, a = a_n$ and S. Obtaining $L_{11}(n + 1, \epsilon_{n+1}/16) = L$ and $m = M$.

Let a be a positive integer with the following properties:

- $a > (n+3)\log_2 m$,

- $a > L_{12}(n+1, \epsilon_{n+1}/16)$,

- For every base s less than or equal to $n+1$,

$$\langle a; s \rangle \epsilon_{n+1}/16 > \langle L_{11}(n+2, \epsilon_{n+2}/16); s \rangle.$$

- For every base s less than or equal to $n+1$,

$$\big(\langle a; s \rangle - \langle a_n; s \rangle\big)\epsilon_{n+1}/16 > \langle a_n; s \rangle.$$

Let Y be the set of reals $y \in I$ such that

$$\forall s \in S, \forall \ell \geqslant L_{11}(n+1, \epsilon_{n+1}/16),\ D(\{s^j y\} : \langle a_n; s \rangle \leqslant j < \langle a_n + \ell; s \rangle) < \epsilon_{n+1}/16.$$

By definition, Y satisfies

$$\nu_{I,m,n+1}(Y) \geqslant \frac{1}{2}.$$

Fix a real number $y \in Y$, which implies $y \in E(m, n+1) \cap I$ and for every $s \in S$ and every $\ell \geqslant L_{11}(n+1, \epsilon_{n+1}/16)$,

$$D(\{s^j y\} : \langle a_n; s \rangle \leqslant j < \langle a_n + \ell; s \rangle) < \epsilon_{n+1}/16.$$

By inductive hypothesis, $a_n > L_{12}(n, \epsilon_n/16) \geqslant L_{12}(n, \epsilon_n)$ and for every base $s \leqslant n$ and all nontrivial intervals $J \subseteq [0,1]$ with rational endpoints of the form $[p_1/8^n, p_2/8^n]$,

$$D(J, (\{s^j x_n\} : j < \langle a_n; s \rangle)) < \epsilon_n.$$

Let J^* be a subinterval of $[0,1]$ of the form $[p_1/8^{n-1}, p_2/8^{n-1}]$. Then, Lemma 12 applies to the pair x_n and y and the interval J^* to conclude that

$$D(J^*, (\{s^j y\} : j < \langle a_n; s \rangle)) < 8\epsilon_n.$$

By inductive hypothesis again, for all $s \leqslant n$, we have $\langle a_n; s \rangle \epsilon_n/16 > \langle L_{11}(n+1, \epsilon_{n+1}/16); s \rangle$, and so by Lemma 13, for all $\ell \leqslant L_{11}(n+1, \epsilon_{n+1}/16)$ and all $s \leqslant n$,

$$D(J^*, (\{s^j y\} : j < \langle a_n + \ell; s \rangle)) < 16\epsilon_n.$$

If ℓ is such that $L_{11}(n+1, \epsilon_{n+1}/16) < \ell \leqslant a - a_n$, then $(\{s^j y\} : j < \langle a_n + \ell; s \rangle)$ is the concatenation of the two sequences $(\{s^j y\} : j < \langle a_n; s \rangle)$ and $(\{s^j y\} : \langle a_n; s \rangle \leqslant j < \langle a_n + \ell; s \rangle)$, both of which have discrepancy, with respect to interval J^*, less than $8\epsilon_n$, and so it has discrepancy less than $8\epsilon_n$. Thus, for every ℓ between a_n and a,

$$D(J^*, (\{s^j y\} : j < \langle \ell; s \rangle)) < 16\epsilon_n.$$

Let $[\overline{y}, \overline{y} + 2^{-a})$ be a dyadic interval such that $y - \overline{y} < 2^{-a}$. As above for x_n and y, Lemma 12 applies to the pair y and $\overline{y}$ to conclude that for every ℓ between a_n and a and every subinterval J of $[0,1]$ of the form $[p_1/8^{n-2}, p_2/8^{n-2}]$,

$$D(J, (\{s^j \overline{y}\} : j < \langle \ell; s \rangle)) < 8 \cdot 16\epsilon_n = 128\epsilon_n.$$

Further, for every $s \leqslant n + 1$, since $D(\{s^j y\} : \langle a_n; s \rangle \leqslant j < \langle a; s \rangle) \leqslant \epsilon_{n+1}/16$ (obtained above) and $(\langle a; s \rangle - \langle a_n; s \rangle)\epsilon_{n+1}/16 \geqslant \langle a_n; s \rangle$, for any interval $J \subseteq [0, 1]$,

$$D(J, (\{s^j y\} : j < \langle a; s \rangle)) < 2 \cdot \epsilon_{n+1}/16 = \epsilon_{n+1}/8,$$

and so, by Lemma 13,

$$D(J, (\{s^j \overline{y}\} : j < \langle a; s \rangle)) < 8 \cdot \epsilon_{n+1}/8 = \epsilon_{n+1}.$$

Lastly, since $y \in E(m, n + 1)$, let q be a positive prime such that $m \leqslant q$ and $||qy|| < q^{-n-2}$. Let $y^* \in [\overline{y}, \overline{y} + 2^{-a})$. We verify that $||qy^*|| < q^{-n-1}$.

Since $y - \overline{y} < 2^{-a}$, then, $y - y^* < 2^{-a}$.

Since $a > (n + 3)\log_2 m$, then, $2^{-a} < m^{-n-3} < q^{-n-3}(q - 1)$.

Thus,

$$q \, 2^{-a} < q^{-n-2}(q - 1)$$
$$q^{-n-2} + q \, 2^{-a} < q^{-n-1}$$
$$||qy|| + |qy - qy^*| < q^{-n-1}$$
$$||qy^*|| < q^{-n-1}, \text{ as required.}$$

Hence, $[\overline{y}, \overline{y} + 2^{-a})$ satisfies all requirements to be a dyadic interval for step $n + 1$.

Now, let x be the limit of the sequence $(x_n : n \geqslant 0)$. By virtue of the second condition in the specification of x_{n+1} from x_n, x is a Liouville number. To check that x is absolutely normal, let s be a base, ϵ a real number and J an interval. By a continuity argument, we may fix positive integer m and assume that the endpoints of J are dyadic rational numbers of the form $p/8^m$. Consider N so large that $s \leqslant N$, $m < N - 3$ and $\epsilon > 8\epsilon_N$. Let ℓ be a positive integer greater than or equal to a_N. Let n be such that $a_n \leqslant \ell < a_{n+1}$. By choice of x_{n+1}, Lemma 12 applies to x_{n+1} and x to conclude that

$$D(J, (\{s^j x\} : j < \langle \ell; s \rangle)) < 8 \cdot 128\epsilon_n.$$

Since the sequence ϵ_n goes to zero as n goes to infinity, the discrepancy of $(\{s^j x\} : j < \langle \ell; s \rangle)$ goes to zero as ℓ goes to infinity. Hence, x is absolutely normal. This completes the proof.

References

[1] Verónica Becher, Pablo Ariel Heiber, and Theodore A. Slaman. A polynomial-time algorithm for computing absolutely normal numbers. *Information and Computation*, 232:1–9, 2013.

[2] Verónica Becher and Theodore A. Slaman. On the normality of numbers to different bases. preprint, arXiv:1311.0333, 2013.

[3] Christian Bluhm. On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets. *Arkiv för Matematik*, 36(2):307–316, 1998.

[4] Christian Bluhm. Liouville numbers, Rajchman measures, and small Cantor sets. *Proceedings American Mathematical Society*, 128(9):2637–2640, 2000.

[5] Émile Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Supplemento di Rendiconti del circolo matematico di Palermo*, 27:247–271, 1909.

[6] Yann Bugeaud. Nombres de Liouville et nombres normaux. *Comptes Rendus de l'Académie des Sciences Paris*, 335(2):117–120, 2002.

[7] Yann Bugeaud. *Distribution Modulo One and Diophantine Approximation*. Number 193 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, UK, 2012.

[8] H. Davenport, P. Erdös, and W. J. LeVeque. On Weyl's criterion for uniform distribution. *Michigan Mathematical Journal*, 10:311–314, 1963.

[9] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Dover, 2006.

[10] Satyadev Nandakumar and Santhosh Kumar Vangapelli. Normality and finite-state dimension of liouville numbers. preprint, arXiv:1204.4104, 2012.

[11] Hermann Weyl. Ueber die gleichverteilung von zahlen mod. eins. *Mathematische Annalen*, 77(3):313–352, 1916.

Verónica Becher
Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Pabellón I, Ciudad Universitaria, (1428) Buenos Aires, Argentina. Also, CONICET Argentina.
Email: `vbecher@dc.uba.ar`

Pablo Ariel Heiber
Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Pabellón I, Ciudad Universitaria, (1428) Buenos Aires, Argentina. Also, CONICET Argentina.
Email: `pheiber@dc.uba.ar`

Theodore A. Slaman
The University of California, Berkeley, Department of Mathematics
719 Evans Hall #3840, Berkeley, CA 94720-3840 USA
Email: `slaman@math.berkeley.edu`