Normality and the Digits of π

David H. Bailey* Jonathan M. Borwein† Cristian S. Calude‡
Michael J. Dinneen§ Monica Dumitrescu¶ Alex Yee‖

February 15, 2012

1 Introduction

The question of whether (and why) the digits of well-known constants of mathematics are statistically random in some sense has long fascinated mathematicians. Indeed, one prime motivation in computing and analyzing digits of π is to explore the age-old question of whether and why these digits appear "random." The first computation on ENIAC in 1949 of π to 2037 decimal places was proposed by John von Neumann to shed some light on the distribution of π (and of e) [8, pp. 277–281].

Since then, numerous computer-based statistical checks of the digits of π, for instance, so far have failed to disclose any deviation from reasonable statistical norms. See, for instance, Table 1, which presents the counts of individual hexadecimal digits among the first trillion hex digits, as obtained by Yasumasa Kanada. By contrast,

*Lawrence Berkeley National Laboratory, Berkeley, CA 94720. dhbailey@lbl.gov. Supported in part by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC02-05CH11231.

†Centre for Computer Assisted Research Mathematics and its Applications (CARMA), University of Newcastle, Callaghan, NSW 2308, Australia. jonathan.borwein@newcastle.edu.au. Distinguished Professor King Abdul-Aziz Univ, Jeddah.

‡Department of Computer Science, University of Auckland, Private Bag 92019, Auckland, New Zealand. cristian@cs.auckland.ac.nz.

§Department of Computer Science, University of Auckland, Private Bag 92019, Auckland, New Zealand. mjd@cs.auckland.ac.nz.

¶Faculty of Mathematics and Computer Science, University of Bucharest, Bucharest, Romania. mdumi@fmi.unibuc.ro.

‖University of Illinois Urbana-Champaign, Urbana, IL, USA. a-yee@u.northwestern.edu.

Table 1: Digit counts in the first trillion hexadecimal (base-16) digits of π. Note that deviations from the average value 62,500,000,000 occur only after the first six digits, as expected from the central limit theorem.

Hex Digits	Occurrences	Hex Digits	Occurrences
0	62499881108	8	62500216752
1	62500212206	9	62500120671
2	62499924780	A	62500266095
3	62500188844	B	62499955595
4	62499807368	C	62500188610
5	62500007205	D	62499613666
6	62499925426	E	62499875079
7	62499878794	F	62499937801
Total	1000000000000		

the early computations did reveal provable abnormalities in the behavior of e [10, §11.2]. Figure 2 show π as a random walk drawn as we describe below.

In the first part of this paper we look at various classical numbers — such as $\sqrt{2}$ and π — and discuss current knowledge regarding normality of such irrational numbers. In the second part of the paper we analyze roughly four trillion hexadecimal digits of π. This is only possible because of several extremely large recent computations [23]. We then propose a Poisson process model of normality of the digits of a number and deduce that in this model, it is extraordinarily unlikely that π is not asymptotically normal base 16, given the behaviour of its initial segment.

1.1 Normality of real numbers

In each of the pictures in Figures 2 through 5, a digit string for a given number is used to determine the angle of unit steps (multiples of 120 degrees base 3, 90 degrees base four, etc), while the color is shifted up the spectrum after a fixed number of steps (red-orange-yellow-green-cyan-blue-purple-red). In Figure 2 we show a walk on the first billion base 4 digits of π. This may be viewed in more detail online at http://gigapan.org/gigapans/e76a680ea683a233677109fddd36304a. In Figures 3 through 5, we similarly illustrate Theorems 2 through 8 below. We note that the behavior or random walks of various numbers constructed to be provably normal looks entirely different from the expected behavior of a genuine pseudorandom walk as in Figure 1.

In the following, given some positive integer base b, we will say that a real number

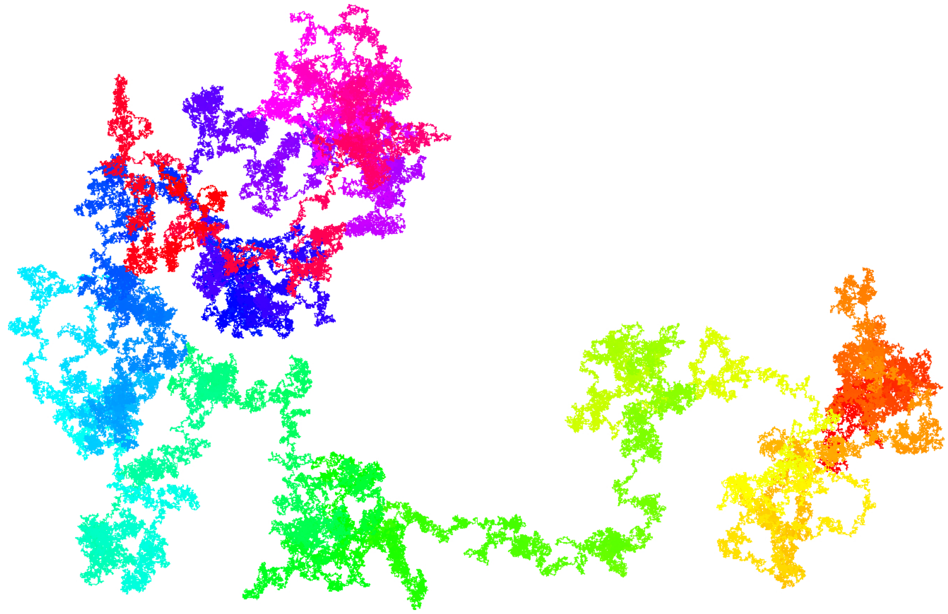


Figure 1: A uniform pseudo-random walk.

α is b-*normal* if every m-long string of base-b digits appears in the base-b expansion of α with precisely the expected limiting frequency $1/b^m$. It follows, from basic measure theory, that almost all real numbers are b-normal for any specific base b and even for all bases simultaneously. But proving normality for specific constants of interest in mathematics has proven remarkably difficult.

Borel was the first to conjecture that *all* irrational algebraic numbers are b-normal for *every* integer $b \geq 2$. Yet not a single instance of this conjecture has ever been proven. We do not even *know* for certain whether or not the limiting frequency of zeroes in the binary expansion of $\sqrt{2}$ is one-half, although numerous large statistical analyses have failed to show any significant deviation from statistical normals.

Recently two of the present authors, together with Richard Crandall and Carl Pomerance, proved the following: If a real y has algebraic degree $D > 1$, then the number $\#(|y|, N)$ of 1-bits in the binary expansion of $|y|$ through bit position N satisfies

$$\#(|y|, N) \quad > \quad CN^{1/D} \tag{1}$$

for a positive number C (depending on y) and all sufficiently large N [2]. For example, there must be at least $\sqrt{N}$ 1-bits in the first N bits in the binary expansion of $\sqrt{2}$, in the limit. A related and more refined result has been obtained by Hajime Kaneko of Kyoto University in Japan. He obtained the bound in $C(\log N)^{3/2}/[(\log(6D))^{1/2}(\log\log N)^{1/2}]$ and extended his results to a very general class of bases and algebraic irrationals [17]. However, each of these results falls far short of establishing b-normality for any irrational algebraic in any base b, even in the single-digit sense.

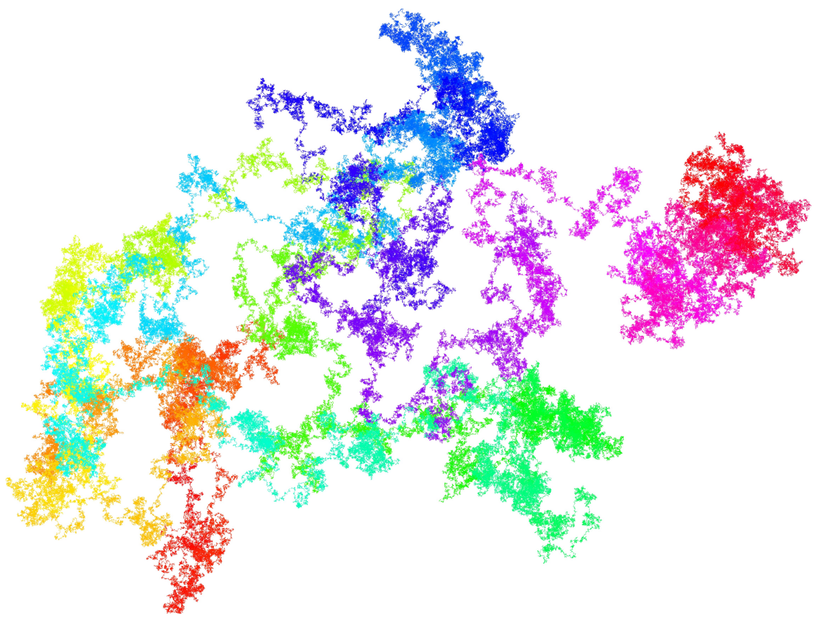


Figure 2: A random walk on the first two billion bits of π (normal?).

The same can be said for π and other basic constants, such as $e, \log 2$ and $\zeta(3)$. Clearly any result (one way or the other) for one of these constants would be a mathematical development of the first magnitude.

We do record the following known stability result [9, pp. 165–166]:

Theorem 1 *If α is normal in base b and r, s are positive rational numbers then $r\alpha + s$ is also normal in base b.*

1.2 The Champernowne number and relatives

The first mathematical constant proven to be 10-normal is the *Champernowne number*, which is defined as the concatenation of the decimal values of the positive integers, i.e., $C_{10} = 0.12345678910111213141516\ldots$, which can also be written as

$$C_{10} = \sum_{n=1}^{\infty} \sum_{k=10^{n-1}}^{10^n-1} \frac{k}{10^{kn-9\sum_{k=0}^{n-1} 10^k (n-k)}}. \tag{2}$$

Champernowne proved that C_{10} is 10-normal in 1933 [14]. This, was later extended to base-b normality (for base-b versions of the Champernowne constant).

In 1946, Copeland and Erdös established that the corresponding concatenation of primes $0.23571113171923\ldots$ and also the concatenation of composites $0.46891012141516\ldots$, among others, are also 10-normal [15]. In general they proved:

Theorem 2 ([15]) *If $a_1, a_2, \cdots$ is an increasing sequence of integers such that for every $\theta < 1$ the number of a_i's up to N exceeds N^θ provided N is sufficiently large, then the infinite decimal*

$$0.a_1 a_2 a_3 \cdots$$

is normal with respect to the base β in which these integers are expressed.

This clearly applies the Champernowne numbers (Figure 5(d)) and to the primes of the form $ak + c$ with a and c relatively prime in any given base (Figure 5(c)) and to the integers which are the sum of two squares (since every prime of the form $4k+1$ is included). In further illustration, using the primes in binary lead to normality in base two of the number

$$0.10111011111011110110001100111011111101111111001011010011010111011111\ldots,$$

as shown as a random walk in Figure 3.

Some related results were established by Schmidt, including the following [20]. Write $p \sim q$ if there are positive integers r and s such that $p^r = q^s$. Then

Theorem 3 *If $p \sim q$, then any real number that is p-normal is also q-normal. However, if $p \not\sim q$, then there are uncountably many p-normal reals that are not q-normal.*

Queffelec [19] described the above result in a recent survey which also presented the following theorem:

Theorem 4 (Korobov) *Numbers of the form $\sum_k p^{-2^k} q^{-p^{2^k}}$, where $p > 1$ and $q > 1$ are relatively prime, are q-normal.*

We will return to such numbers in Theorem 5 of Section 4. Nonetheless, we are still completely in the dark as to the b-normality of "natural" constants of mathematics.

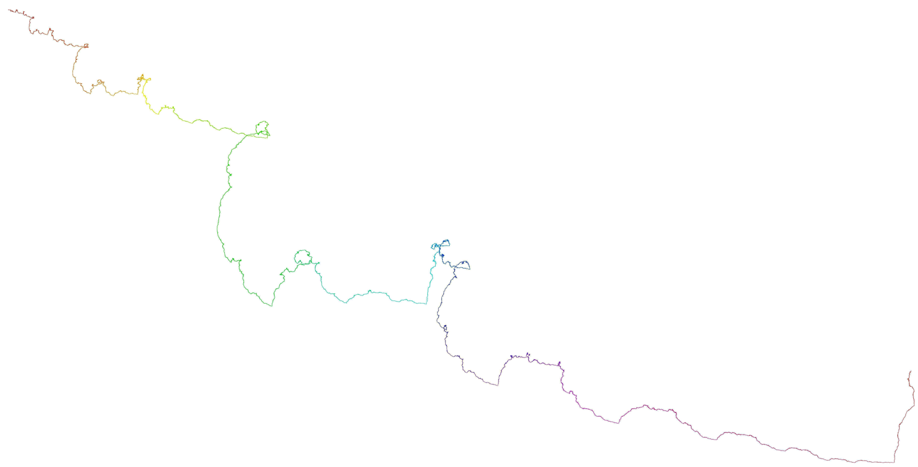


Figure 3: A random walk on the first 100,000 bits of the primes base two (normal).

2 The BBP formula for π

In 1996, one of the present authors (Bailey), together with Peter Borwein (brother of Jonathan Borwein) and Simon Plouffe, published what is now known as the BBP formula for π [3], [9, Ch. 3]:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right). \tag{3}$$

This formula has the remarkable property that it permits one to directly calculate binary or hexadecimal digits of π beginning at an arbitrary starting position, without needing to calculate any of the preceding digits. The resulting simple algorithm requires only minimal memory, does not require multiple-precision arithmetic, and is very well suited to highly parallel computation. The cost of this scheme increases only slightly faster than the index of the starting position.

The BBP formula (3) was discovered via a computer-based search using the PSLQ *integer relation detection algorithm* of mathematician-sculptor Helaman Ferguson [4], in a process that some have described as an exercise in "reverse mathematical engineering." The motivation for this search was the earlier observation by the authors of [3] that $\log 2$ also has this arbitrary position digit calculating property. This can be seen by analyzing the classical formula

$$\log 2 = \sum_{k=1}^{\infty} \frac{1}{k 2^k}, \tag{4}$$

which has been known at least since the time of Euler, and which is closely related to the functional equation for the dilogarithm.

The digit-calculating scheme can be demonstrated for the $\log 2$ formula (4) as follows. Let $r \bmod 1$ denote the fractional part of a nonnegative real number r, and let d be a nonnegative integer. Then the binary fraction of $\log 2$ after the "decimal" point has been shifted to the right d places can be written as

$$\begin{aligned}(2^d \log 2) \bmod 1 &= \left(\sum_{k=1}^{d} \frac{2^{d-k}}{k} \bmod 1 + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k} \bmod 1\right) \bmod 1 \\ &= \left(\sum_{k=1}^{d} \frac{2^{d-k} \bmod k}{k} \bmod 1 + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k} \bmod 1\right) \bmod 1,\end{aligned} \tag{5}$$

where "mod k" has been inserted in the numerator of first term since we are only interested in the fractional part of the result after division. The operation $2^{d-k} \bmod k$ can be performed very rapidly by means of the *binary algorithm for exponentiation*, where all of the intermediate multiplication results are reduced modulo k at each step. This algorithm, together with the division and summation operations indicated in the first term, can be performed in ordinary double-precision floating-point arithmetic, or, for very large calculations, by using quad- or oct-precision arithmetic. Expressing the final fractional value in binary notation yields a string of digits corresponding to the binary digits of $\log 2$ beginning immediately after the first d digits. Applying this same scheme to the four terms of (3) yields binary digits of π at an arbitrary starting position.

3 BBP formulas and normality

Interest in BBP-type formulas was heightened by the 2001 observation, by one of the present authors (Bailey) and Richard Crandall, that the normality of BBP-type constants such as $\pi, \pi^2, \log 2$ and Catalan's constant can be reduced to a certain hypothesis regarding the behavior of a class of chaotic iterations [5], [9, pg. 141–173]. No proof is known for this general hypothesis, but even proofs in specific instances would be quite interesting. For example, if it could be established that the iteration given by $w_0 = 0$, and

$$w_n = \left(2w_{n-1} + \frac{1}{n}\right) \bmod 1 \tag{6}$$

is equi-distributed in $[0,1)$ (i.e., is a "good" pseudorandom number generator), then, according to the Bailey-Crandall result, it would follow that $\log 2$ is 2-normal. In a similar vein, if it could be established that the iteration given by $x_0 = 0$ and

$$x_n = \left(16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21}\right) \bmod 1 \qquad (7)$$

is equi-distributed in $[0,1)$, then it would follow that π is 2-normal.

3.1 The Erdös-Borwein constants

In a base $b \geq 2$, we define the *Erdös-(Peter) Borwein constant* $EB(b)$ by the Lambert series [10]:

$$EB(b) := \sum_{n \geq 1} \frac{1}{b^n - 1} = \sum_{n \geq 1} \frac{\sigma(n)}{b^n}, \qquad (8)$$

where σ_k the sum of the k-th power of the divisors and $\sigma = \sigma_0 = d$ counts the number of divisors. It is known that the numbers $\sum 1/(q^n - r)$ are irrational for r rational and $q = 1/b, b = 2, 3, ...$ [11]. Whence, it is interesting to consider their normality.

Crandall has used the BBP-like structure made obvious in (8), and some non-trivial knowledge of the arithmetic properties of σ to establish results such as that the googol-th bit–i.e., the bit in position 10^{100} to the right of the floating point—is a 1.

Crandall and Mitchell also computed the full first 2^{43} bits of $EB(2)$ (a Terabyte in about a day), and find that there are **435**9105565638 zeroes and **443**6987456570 ones. There is corresponding variation in the second and third place in the single digit hex distributions. This certainly leaves some doubt as to its normality. See also Figure 5(e).

Our own more modest computations of $EB(10)$ base-ten again leave it far from clear that $EB(10)$ is 10-normal. For instance, Crandall finds that in the first 1000 decimal positions after the quintillionth digit 10^{18}, the respective digit counts for digits $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ are $104, 82, 87, 100, 73, 126, 87, 123, 114, 104$. See also Figure 5(f).

The identity

$$\sum_{n \geq 1} \frac{q^n}{1 - q^n} = \sum_{n \geq 1} q^{n^2} \frac{1 + q^n}{1 - q^n},$$

for $|q| < 1$, due to Clausen, is what we used for computational purposes, as did Crandall [16].

4 A class of provably normal constants

In 2002, Bailey and Crandall showed that given a real number r in $[0, 1)$, with r_k denoting the k-th binary digit of r, the real number

$$\alpha_{2,3}(r) := \sum_{k=0}^{\infty} \frac{1}{3^k 2^{3^k + r_k}} \tag{9}$$

is 2-normal. It can be seen that if $r \neq s$, then $\alpha_{2,3}(r) \neq \alpha_{2,3}(s)$, so that these constants are all distinct. Since r can range over the unit interval, this class of constants is uncountably infinite. A similar result applies if 2 and 3 in this formula are replaced by any pair of coprime integers (b, c) with $b \geq 2$ and $c \geq 2$ [6], [9, pg. 141–173]. So, for example, the constant $\alpha_{2,3}(0) = \sum_{k \geq 0} 1/(3^k 2^{3^k}) = 0.5418836808315030\ldots$ is provably 2-normal. This special case was proven by Stoneham in 1973 [21].

More recently, one of the present authors (Bailey) and Michal Misieurwicz were able to establish this normality result by a simpler argument, by utilizing techniques of ergodic theory [7] and a "hot spot" lemma. In the Appendix, we present proofs of the following two results, illustrating the usage of the "hot spot" lemma:

Theorem 5 *The constant*

$$\alpha_{2,3}(0) = \sum_{k=0}^{\infty} \frac{1}{3^k 2^{3^k}} \tag{10}$$

is 2-normal.

Theorem 6 *Given coprime integers $b \geq 2$ and $c \geq 2$, the constant*

$$\alpha_{b,c}(0) = \sum_{k=0}^{\infty} \frac{1}{c^k b^{c^k}} \tag{11}$$

is b-normal.

The first proof is adapted from [7], but the second proof has not previously been published.

5 A non-normality result

Almost as interesting is the following fact:

Theorem 7 (Non-normality) *$\alpha_{2,3}(0)$ is* not *6-normal.*

Discussion: In the following, we will use α to mean $\alpha_{2,3}(0)$. Note that the base-6 digits immediately following position n in the base-6 expansion of α can be obtained by computing $6^n\alpha \bmod 1$, which can be written as follows:

$$\begin{aligned}(6^n\alpha) \bmod 1 &= \left(\sum_{m=0}^{\lfloor \log_3 n \rfloor} 3^{n-m}2^{n-3^m}\right) \bmod 1 \\ &+ \left(\sum_{m=\lfloor \log_3 n \rfloor+1}^{\infty} 3^{n-m}2^{n-3^m}\right) \bmod 1.\end{aligned}$$

Now note that the first portion of this expression is *zero*, since all terms of the summation are integers. That leaves the second expression.

Consider first the special case when $n = 3^m$, where $m \geq 1$ is an integer. The first three terms of the second summation are:

$$\begin{aligned}3^{3^m-(m+1)}2^{3^m-3^{m+1}} &= 3^{3^m-m-1}2^{-2\cdot 3^m} = (3/4)^{3^m}/3^{m+1}, \\ 3^{3^m-(m+2)}2^{3^m-3^{m+2}} &= 3^{3^m-m-2}2^{-8\cdot 3^m} = (3/8)^{3^m}/3^{m+2}, \\ 3^{3^m-(m+3)}2^{3^m-3^{m+3}} &= 3^{3^m-m-3}2^{-26\cdot 3^m} = (3/26)^{3^m}/3^{m+3}. \qquad (12)\end{aligned}$$

Note that for $m \geq 3$, we can generously bound the sum of all these terms by $1 + 10^{-6}$ times the first term, and by ratios arbitrarily closer to one for larger m. Thus we have $(6^{3^m}\alpha) \bmod 1 \approx (3/4)^{3^m}/3^{m+1}$, and this approximation is as accurate as desired (in ratio) for all sufficiently large m.

Given the very small size of the expression $(3/4)^{3^m}/3^{m+1}$ for even moderate-sized m, it is clear the base-6 expansion will have very long stretches of zeroes beginning at positions $3^m + 1$. For example, by explicitly computing α to high precision, one can obtain the counts shown in Table 2 of consecutive zeroes Z_m that immediately follow position 3^m in the base-6 expansion of α.

Examining Table 2, there 14256 zeroes in these ten segments, which, including the last segment, span the first $59049 + 9487 = 68536$ base-6 digits of α. In this tabulation we have of course ignored the many zeroes in the large "random" segments of the expansion. Thus the fraction of the first 68536 digits that are zero is at least $14256/68536 = 0.20800747\ldots$. A more careful analysis shows that the limiting ratio

$$\begin{aligned}\lim_{m\to\infty} \frac{\sum_{m\geq 1} Z_m}{3^m + Z_m} &= \frac{3}{2} \cdot \frac{\log_6(4/3)}{1 + \log_6(4/3)} \\ &= \frac{1}{2}\log_2(4/3) = 0.2075187496\ldots, \qquad (13)\end{aligned}$$

Table 2: Counts of consecutive zeroes.

m	3^m	Z_m
1	3	1
2	9	3
3	27	6
4	81	16
5	243	42
6	729	121
7	2187	356
8	6561	1058
9	19683	3166
10	59049	9487

which is significantly more than the expected value $1/6 = 0.166666\ldots$, thus establishing the non-normality base 6 of α. A completely detailed proof is presented in the Appendix. The Appendix also gives a full proof of this generalization of Theorem 7:

Theorem 8 (Non-normality) *Given coprime integers $b \geq 2$ and $c \geq 2$, the constant*

$$\alpha_{b,c}(0) : \;=\; \sum_{k \geq 1} \frac{1}{c^k b^{c^k}} \tag{14}$$

is b-normal but is not *bc-normal.*

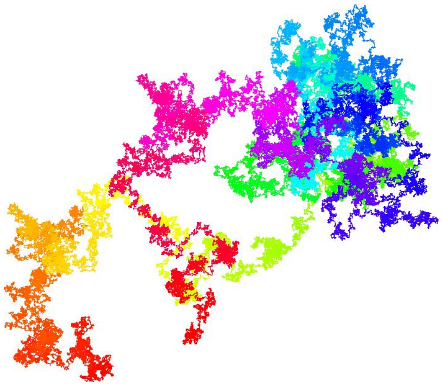
We further conjecture that this result is true not just for the class $\alpha_{b,c}(0)$, but also for all $\alpha_{b,c}(r)$ where r is the argument in (9) above.

The proofs of Theorems 7 and 8 given in the Appendix are adapted from [1].

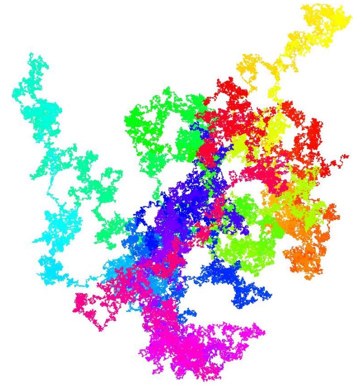
In Figure 4 the non-normality of $\alpha_{2,3}$ base 6 is very visible.

Figure 4: A random walk on the first 50,000 bits of $\alpha_{2,3}(0)$ base six (abnormal).

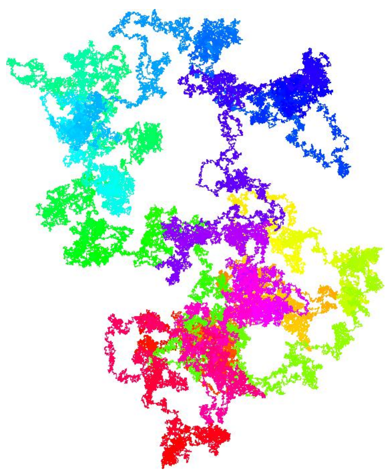
We now turn to an analysis of the normality of finite strings.



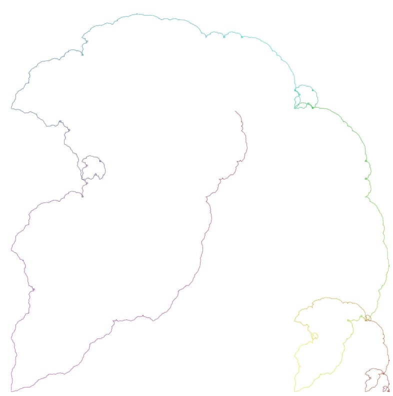
(a) A half-million step walk on $\alpha_{2,3}(0)$ base 2 (normal)



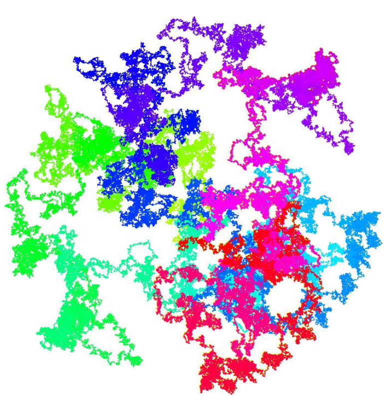
(b) A million step walk on $\alpha_{2,3}(0)$ base 3 (normal?)



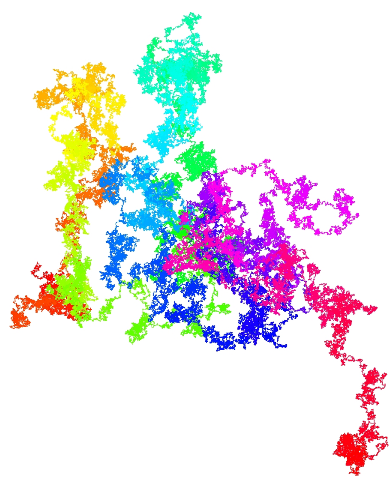
(c) A million step walk on 23571113... base 2 (normal?)



(d) A 600,000 step walk on Champernowne's number base 4 (normal)



(e) A million step walk on $EB(2)$ base 4 (normal?)



(f) A million step walk on $EB(10)$ base 4 (normal?)

Figure 5: Random walks on prefixes of various numbers.

6 Normality for words

Let x be a (finite) binary word. We denote by $N_i^m(x)$ the number of occurrences of the ith word of length m ($1 \leq i \leq 2^m$), ordered lexicographically, where $|x|_m = \lfloor |x|/m \rfloor$ is the number of (contiguous, non-overlapping) of length m words in x. The *prefix* of length n of the infinite (binary) sequence $\mathbf{x} = x_1 x_2 \ldots x_m \ldots$ is denoted by $\mathbf{x} \restriction n = x_1 x_2 \ldots x_n$.

Definition 1 ([12, 13]) *Let $\varepsilon > 0$ and m be a positive integer. We say:*

1. *x is $(\varepsilon, m)-$normal if, for every $1 \leq i \leq 2^m$,*

$$\left| \frac{N_i^m(x)}{|x|_m} - \frac{1}{2^m} \right| \leq \varepsilon.$$

2. *x is $m-$normal if, for every $1 \leq i \leq 2^m$,*

$$\left| \frac{N_i^m(x)}{|x|_m} - \frac{1}{2^m} \right| \leq \sqrt{\frac{\log_2 |x|}{|x|}}. \tag{15}$$

3. *x is normal if it is $m-$normal for every $1 \leq m \leq \log_2(\log_2 |x|)$.*

If for every positive integer n, the word $\mathbf{x} \restriction n$ is normal, then $\mathbf{x}$ is normal, but the converse is not necessarily true (because $\mathbf{x}$ can be normal but with a different "speed").

7 Testing normality of prefixes of π

We had access to an extremely large data-set, thanks to recent record computations by Kondo and Yee, of π initially to five trillion hexadecimal (base 16) places in August 2010 and then to ten trillion in October 2011 [23]. We first converted these bits — which Kondo and Yee had confirmed by a computation with (3) — to a true binary string of bits using the Python module `binascii`.

All input lines contained an even number of characters so it was easy to convert pairs of hexadecimal digits to bytes.

```
import sys, binascii
for line in sys.stdin.readlines():
    sys.stdout.write(binascii.unhexlify(line.strip()))
```

For our normality test we needed to split a *big* binary string of length n into $\lfloor n/k \rfloor$ pieces (non-overlapping words) of length $k = 1, 2, \ldots, \log\log n$. We use the term *word* to denote a binary string of length k. We then proceeded to calculate the minimum and maximum occurrences of such words.

This calculation is done by running the following Algorithm 1 once for each different value of k.

Algorithm 1: Frequency range of words of a given length.

Input: Binary string X, word length k
Output: Minimum and maximum counts over all possible 2^k words of length k in string X
integer array $counts[0, \ldots, 2^k - 1] = [0, 0, \ldots, 0]$;
for $i = 0$ **to** $|X| - k$ ***step*** k **do**
  $w = integer(X[i, \ldots, i + k - 1])$;
  increment $counts[w]$;
return min($counts$), max($counts$);

It is essential to do an efficient streaming implementation of Algorithm 1 so that the actual bits of input X are only read into main memory as needed.

Finally to check that these minimum and maximum frequencies satisfy the expected range for the normality test we used the following Python code snippet to generate a table using our earlier formula (15):

```
import math, sys
n=int(sys.argv[1]) # n = |X|
r = int(math.floor(math.log(math.log(n,2),2))) # r = lg lg n
m1,m2=[0]*(r+1),[0]*(r+1)
sqrtV = math.sqrt(math.log(n,2)/n)
for k in range(1,r+1):
    floorNk = math.floor(n/k)
    m1[k] = int(math.floor(((1.0/2.0**i)-sqrtV)*floorNk))
    m2[k] = int(math.ceil((sqrtV+(1.0/2.0**k))*floorNk))
    print "expected range k=",k, "[",m1[k],"...",m2[k],"]"
```

We tested normality for the prefix of $N = 15,925,868,541,400$ bits of π—nearly 16 trillion bits—calculated with the *y-cruncher-multi-threaded pi* program [22] and we have found it to be within the normality range as described above. The frequency counts passed our `expectedCheck.py` test script.

Table 3: Frequency summary for $N = 15,925,868,541,400$ bits of π.

m	min frequency found	max frequency found	expected range
1	**79629**33149184	**79629**35392216	7962907842460, ..., 7962960698940
2	**19907**32495242	**19907**35357049	1990720353555, ..., 1990746781795
3	**6635**76589836	**6635**79050172	663569046478, ..., 663586665305
4	**2488**41171873	**2488**42651924	248835088899, ..., 248848303020
5	**9953**5989611	**995**37473460	99531392735, ..., 99541964032

8 Normality of π

We have tested the prefix of $N = 15,925,868,541,400$ bits of π—nearly 16 trillion bits—and we have found it to be normal as described above.

Does this "information" tell us anything about the classical normality of π? In the next subsection, we will use a Poisson process model to provide an affirmative answer to this question.

8.1 A Poisson process model

We denote by

$$\mathbf{b} = b(1)\, b(2) \ldots b(n) \ldots$$

the (infinite) binary expansion of π (b is a computable function) and by

$$\mathbf{b} \restriction n = b(1)\, b(2) \ldots b(n)$$

the finite prefix of $\mathbf{b}$ of length n.

We base our model on the distribution on 1s' and 0's only, i.e., we work with $N_1^1(\mathbf{b} \restriction n)$, the number of occurrences of 1's in $\mathbf{b} \restriction n$, so $N_0^1(\mathbf{b} \restriction n) = n - N_1^1(\mathbf{b} \restriction n)$. A similar, slightly more elaborate model, can be developed for words of any length.

The number $N_1^1(\mathbf{b} \restriction n)$ can be connected with π by means of a *counting (Poisson) process* [18]:

$$\begin{aligned} &Y_n = \#\{j \mid 1 \leq j \leq n,\ b(j) = 1\},\ n = 1, 2, \ldots \\ &Y_0 = 0, \end{aligned}$$

where $Y_n = N_1^1(\mathbf{b} \restriction n)$, $n = 1, 2, \ldots$

Theorem 9 *If π is normal, then $\{Y_n, n = 0, 1, 2...\}$ can be approximated by a homogenous Poisson process with intensity $\lambda = 0.5$.*

Proof. By construction, $\{Y_n, n = 0, 1, 2...\}$ is a Poisson process with an unspecified parameter λ. Hence Y_n is a random variable with parameter $n\lambda$ with the following properties: $E(Y_n) = V(Y_n) = n\lambda$, $\lim_{n\to\infty} Y_n = \infty$ almost sure.

We apply Chebysev's inequality, so for every $c > 0$,

$$P(|Y_n - E(Y_n)| < c) \geq 1 - \frac{V(Y_n)}{c^2},$$

we have

$$P(|Y_n - n\lambda| < c) \geq 1 - \frac{n\lambda}{c^2},$$

hence

$$P\left(\left|\frac{Y_n}{n} - \lambda\right| < \frac{c}{n}\right) \geq 1 - \frac{n\lambda}{c^2}\cdot$$

In view of (15) we take

$$\frac{c}{n} = \varepsilon = \sqrt{\frac{\log_2 n}{n}},$$

so we obtain

$$P\left(\left|\frac{Y_n}{n} - \lambda\right| < \varepsilon\right) \geq 1 - \frac{n\lambda}{(n\varepsilon)^2} = 1 - \frac{\lambda}{\log_2 n}\cdot \tag{16}$$

If π is normal, then

$$\left|\frac{N_1^1\left(x_{(n)}\right)}{n} - \frac{1}{2}\right| \leq \varepsilon = \sqrt{\frac{\log_2 n}{n}}$$

or

$$\left|\frac{Y_n}{n} - \frac{1}{2}\right| \leq \varepsilon = \sqrt{\frac{\log_2 n}{n}}\cdot \tag{17}$$

If we identify the random event in relation (16) and the certain event in relation (17) we get $\lambda = 1/2$ and

$$P\left(\left|\frac{Y_n}{n} - \frac{1}{2}\right| < \sqrt{\frac{\log_2 n}{n}}\right) \geq 1 - \frac{1}{2\log_2 n}\cdot$$

QED

A Poisson process with intensity λ has the following properties [21]:

- The Poisson process $\{Y_n, n = 0, 1, 2, ...\}$ has independent increments.
- For $n > r$, $Y_n - Y_r$ has a Poisson distribution with parameter $\lambda(n-r)$, and $Y_n - Y_r$ is independent of $\{Y_t, t \leq r\}$.

Let us denote the positions where 1*s* occur (*jump moments*) by

$$\tau_r = \inf\{n \mid Y_n = r\}, \ r = 1, 2, ...$$

Then

$$Y_n = 0, \ n < \tau_1,$$
$$Y_n = r, \ \tau_r \leq n < \tau_{r+1}.$$

With the convention $\tau_0 = 0$, we can introduce the *sojourn times,* or *inter-arrival times*

$$T_r = \tau_r - \tau_{r-1}, \ r = 1, 2, ...$$

Note that the sojourn times represent the distances between two successive 1s. Thus, for the word 10^s1 the sojourn time is $s + 1$.

- $\{T_r, r = 1, 2, ...\}$ is a sequence of independent, identical distributed random variables, with the Exponential distribution $Expo(\lambda)$.Then

$$E(T_r) = \frac{1}{\lambda}, \ V(T_r) = \frac{1}{\lambda^2}.$$

Note that the jump moments $\tau_r = T_1 + ... + T_r$ have an Erlang distribution with parameters $(r; \lambda)$, hence

$$E(T_r) = \frac{r}{\lambda}, \ V(T_r) = \frac{r}{\lambda^2}.$$

Corollary 1 *If π is normal, then the sojourn times $\{T_r, r = 1, 2, ...\}$ form a sequence of independent, identical distributed random variables, with the Exponential distribution $Expo(1/2)$. Hence*

$$P(T_r > t_r, \ r = 1, ..., k) = \prod_{r=1}^{k}\left(\exp\left(-\frac{t_r}{2}\right)\right) = \exp\left(-\frac{1}{2}\sum_{r=1}^{k} t_r\right).$$

8.2 Testing the hypothesis "π is normal"

We test the hypothesis H: "π is normal" against the alternative H_A: "π is not normal". If H is true, then for every d there exists K_d such that the sojourn tine exceeds the value d if we wait long enough, up to the rank $(K_d + 1)$:

$$P\left(T_1 \leq d, ..., T_{K_d} \leq d, T_{K_d+1} > d \mid H \text{ true}\right) = \prod_{r=1}^{K_d} \left(1 - \exp\left(-\frac{d}{2}\right)\right) \cdot \exp\left(-\frac{d}{2}\right)$$
$$= \exp\left(-\frac{d}{2}\right)\left(1 - \exp\left(-\frac{d}{2}\right)\right)^{K_d} > 0.$$

We can base our decision of accepting/rejecting normality (hypothesis H) on the following implication: "π is a normal sequence" *implies* "for every d there exists K_d such that $P\left(T_1 \leq d, ..., T_{K_d} \leq d, T_{K_d+1} > d\right) > 0$"),

As one cannot explore the whole sequence π, we deal with an *evidence body* represented by a prefix of π, of length N. In this evidence body, we look for the largest value $d_{\max}$ for which a rank $K_{d\max}$ can be identified or, equivalently, we look for the first value $(d+1)$ which is not reached by the sojourn time T. Accordingly, the decision of accepting/rejecting the hypothesis H : "π is normal" is taken according to the following algorithm:

- If there is no such $d_{\max}$ in the evidence body, we conclude that the sequence π is normal.
- If $d_{\max}$ and the corresponding $K_{d_{\max}}$ exist, we can decide that the sequence π is not normal. The decision is based on the event

  $$\left\{T_1 \leq d_{\max}, ..., T_{K_{d_{\max}}} \leq d_{\max}, T_{K_{d_{\max}}+1} > d_{\max}\right\}$$

  whose probability is

  $$P\left(T_1 \leq d_{\max}, ..., T_{K_{d_{\max}}} \leq d_{\max}, T_{K_{d_{\max}}+1} > d_{\max}\right)$$
  $$= \exp\left(-\frac{d_{\max}}{2}\right)\left(1 - \exp\left(-\frac{d_{\max}}{2}\right)\right)^{K_{d_{\max}}}.$$

We interpret the above probability as the decision "π is normal" has credibility equal to

$$1 - \exp\left(-\frac{d_{\max}}{2}\right)\left(1 - \exp\left(-\frac{d_{\max}}{2}\right)\right)^{K_{d_{\max}}}.$$

Table 4: d and K_d values for 400 million bits of π.

d	1	2	3	4	5	6	7
K_d	9	1	14	3	46	56	41
d	8	9	10	11	12	13	14
K_d	78	1276	446	2090	18082	8633	4175
d	15	16	17	18	19	20	21
K_d	239183	5856	56453	218007	643030	363117	2787207
d	22	23	24	25	26	27	28
K_d	13733056	1003213	21127913	100317701	not found	85745944	not found
d	29						
K_d	not found						

8.3 Results

Suppose first that the evidence body is represented by a prefix of 400 million bits of π. The $d-$values and their corresponding ranks K_d are given in Table 4; $\max K_d$=100317701.

The value $d = 28$ has the property that for every K, the event

$$\{T_1 \le 28, ..., T_K \le 28, T_{K+1} > 28\}$$

has not been identified in the the evidence body, so, based on the algorithm in Section 8.2, the decision "π is not normal" has credibility

$$P\left(T_s \le 27,\ s = 1, ..., 100317701, T_{100317702} > 27\right)$$
$$= \left(1 - \exp\left(-\frac{27}{2}\right)\right)^{100317701} \cdot \exp\left(-\frac{27}{2}\right) = 2.557\,6 \times 10^{-66}.$$

Suppose now that the evidence body has increased to the prefix of π of N = 15925868541400 bits. The $d-$values and their corresponding ranks K_d are given in following Table 5; $\max K_d = 9274770297096$.

The value $d = 43$ has the property that for every K, the event

$$\{T_1 \le 43, ..., T_K \le 43, T_{K+1} > 43\}$$

Table 5: d and K_d values for 15925868541400 bits of π.

d	1	2	3	4	5
K_d	9	1	14	3	46
d	6	7	8	9	10
K_d	56	41	78	1276	446
d	11	12	13	14	15
K_d	2090	18082	8633	4175	239183
d	16	17	18	19	20
K_d	5856	56453	218007	643030	363117
d	21	22	23	24	25
K_d	2787207	13733056	1003213	21127913	100317701
d	26	27	28	29	30
K_d	273575848	85745944	234725219	611367301	1075713943
d	31	32	33	34	35
K_d	703644000	10621041176	27019219636	15063287853	10887127703
d	36	37	38	39	40
K_d	48115888750	19128531469	1218723032299	1334087352175	792460189481
d	41	42	43	44	45
K_d	9274770297096	4368224447710	not found	not found	not found

has not been identified in the evidence body, so, based on the algorithm in Section 8.2, the decision "π is not normal" has credibility

$$P\left(T_s \leq 42,\ s = 1, ..., 9274770297096, T_{9274770297097} > 42\right)$$
$$= \left(1 - \exp\left(-\frac{42}{2}\right)\right)^{9274770297096} \cdot \exp\left(-\frac{42}{2}\right) = 4.349\,7 \times 10^{-3064}.$$

This is perhaps 'incredible'?

9 Conclusion

A prime motivation in computing and analyzing digits of π is to explore the age-old question of whether and why these digits appear "random." Numerous computer-based statistical checks of the digits of π have failed to disclose any deviation from reasonable statistical norms. A new avenue for studying the normality of π was explored: we proved that the prefix of length 15, 925, 868, 541, 400 bits of π is normal when viewed as a binary word [12].

This result was used in a Poisson process model to show that the probability that π is not normal is extraordinarily small, reinforcing the empirical evidence we have presented evidence for the normality of π. In future work we intend to look methodically at other numerical constants.

Acknowledgement Thanks are due to Dr. Francisco Aragon for his generous assistance with the pictures of random walks.

10 Appendix

In the following, we will utilize the following result from [7]. Let $A(\alpha, y, n, m)$ denote the count of occurrences where the m-long binary string y is found to start at position p in the binary expansion of α, where $1 \le p \le n$.

Lemma 1 ("Hot Spot" Lemma): *If x is not b-normal, then there is some $y \in [0, 1)$ with the property*

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{b^m A(x, y, n, m)}{n} \quad = \quad \infty. \tag{18}$$

Conversely, if for all $y \in [0, 1)$,

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{b^m A(x, y, n, m)}{n} \quad < \quad \infty, \tag{19}$$

then x is b-normal.

Note that Lemma 1 implies that if a real constant α is not b-normal, then there must exist some interval $[r_1, s_1)$ with the property that successive shifts of the base-b expansion of α visit $[r_1, s_1)$ ten times more frequently, in the limit, relative to its length $s_1 - r_1$; there must be another interval $[r_2, s_2)$ that is visited 100 times more often relative to its length; there must be a third interval $[r_3, s_3)$ that is visited 1,000 times more often relative to its length; etc. Furthermore, there exists at least one real number y (a "hot spot") such that sufficiently small neighborhoods of y are visited too often by an arbitrarily large factor, relative to the lengths of these neighborhoods. On the other hand, if it can be established that no subinterval of the unit interval is visited 1,000 times (for instance) more often in the limit relative to its length, then this suffices to prove that the constant in question is b-normal (and thus that each subinterval is visited with precisely the correct frequency, in the limit, relative to the size of the subinterval).

Here are two simple examples of the "hot spots" of non-normal decimal numbers. First, consider the fraction $1/28$. Obviously this is not a 10-normal number since its decimal expansion repeats. It is easy to see by examining its decimal expansion that it possesses six base-10 hot spots, namely $1/7, 2/7, \cdots, 6/7$. As a second example, consider $\sum_{n\ge1} 10^{-n^2}$, which is irrational but also clearly not 10-normal. This has zero as a base-10 hot spot.

We will first establish 2-normality for the constant $\alpha_{2,3}$. This result was established first by Stoneham [21] and more recently in [6]. Here we reprise a simpler proof that was first presented in [7].

Proof of Theorem 5: $\alpha_{2,3} = \sum_{m=0}^{\infty} 1/(3^m 2^{3^m})$ is 2-normal.

First we note that the successive shifted binary fractions of $\alpha = \alpha_{2,3}$ can be written as

$$2^n \alpha \bmod 1 \quad = \quad \left(\sum_{m=0}^{\lfloor \log_3 n \rfloor} \frac{2^{n-3^m} \bmod 3^m}{3^m} \right) \bmod 1 + \sum_{m=\lfloor \log_3 n \rfloor + 1}^{\infty} \frac{2^{n-3^m}}{3^m} \cdot \qquad (20)$$

As it turns out, the first term of this expression can be generated by means of the recursion $z_0 = 0$ and, for $n \geq 1$, $z_n = (2z_{n-1} + r_n) \bmod 1$, where $r_n = 1/n$ if $n = 3^k$ for some integer k, and zero otherwise. The first few members of the z sequence are given as follows:

$$\begin{array}{l}
0,\ 0, \\
\frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \\
\frac{4}{9}, \frac{8}{9}, \frac{7}{9}, \frac{5}{9}, \frac{1}{9}, \frac{2}{9}, \quad \text{(repeated 3 times)}, \\
\frac{13}{27}, \frac{26}{27}, \frac{25}{27}, \frac{23}{27}, \frac{19}{27}, \frac{11}{27}, \frac{22}{27}, \frac{17}{27}, \frac{7}{27}, \frac{14}{27}, \frac{1}{27}, \frac{2}{27}, \frac{4}{27}, \frac{8}{27}, \frac{16}{27}, \frac{5}{27}, \\
\frac{10}{27}, \frac{20}{27}, \quad \text{(repeated 3 times), etc.}
\end{array} \qquad (21)$$

It can be shown, via straightforward combinatorial arguments [6], that indeed this sequence has the pattern evident here: It is a concatenation of triply repeated segments, where each individual segment consists of fractions with numerators, at stage m, that range over all integers relatively prime to the denominator 3^m. From this pattern it follows that if $n < 3^{p+1}$, then z_n is a multiple of $1/3^p$, and furthermore that the set $(z_k, 1 \leq k \leq n)$ contains at most three repetitions of any particular value.

These fractions (z_k) constitute an accurate set of approximations to the sequence $2^n \alpha \bmod 1$ of shifted fractions of α. In fact, by examining (20) it can be readily seen that for all $n \geq 3$,

$$|2^n \alpha \bmod 1 - z_n| < \frac{1}{96n} \cdot \qquad (22)$$

To establish that α is 2-normal via Lemma 1, we seek an upper bound for $2^m A(\alpha, y, n, m)/n$. A binary sequence y out to some length m, translated to a subset of the real unit interval, is $[r, s)$, where $r = 0.y_1 y_2 y_3 \ldots y_m$, and s is the next largest binary fraction of length m, so that $s - r = 2^{-m}$. Observe that $A(\alpha, y, n, m)$ is equal to the number of those j between 0 and $n-1$ for which $2^j\alpha \bmod 1 \in [r,s)$. Also observe, in view of (22), that if $2^j\alpha \bmod 1 \in [r,s)$, then $z_j \in [r-1/(96j), s+1/(96j))$.

Let n be any integer greater than 2^{2m}, and let 3^p denote the largest power of 3 less than or equal to n, so that $3^p \le n < 3^{p+1}$. Now note that for $j \ge 2^m$, we have $[r-1/(96j), s+1/(96j)) \subset [r-2^{-m-1}, s+2^{-m-1})$. Since the length of this latter interval is 2^{-m+1}, the number of multiples of $1/3^p$ that it contains cannot exceed $\lfloor 3^p 2^{-m+1}\rfloor + 1$. Thus there can be at most three times this many j's less than n for which $z_j \in [r-2^{-m-1}, s+2^{-m-1})$. Therefore we can write

$$\begin{aligned}
\frac{2^m A(\alpha,y,n,m)}{n} &= \frac{2^m \#_{0\le j<n}\big(2^j\alpha \bmod 1 \in [r,s)\big)}{n} \\
&\le \frac{2^m\big[2^m + \#_{2^m\le j<n}\big(z_j \in [r-2^{-m-1}, s+2^{-m-1})\big)\big]}{n} \\
&\le \frac{2^m\big[2^m + 3(3^p 2^{-m+1}+1)\big]}{n} < 8.
\end{aligned}$$

We have shown that for all $y \in [0,1)$ and all $m > 0$,

$$\limsup_{n\to\infty} \frac{2^m A(\alpha,y,n,m)}{n} \le 8, \tag{23}$$

so by Lemma 1, α is 2-normal. **QED**

Proof of Theorem 6: For every coprime pair of integers (b,c) with $b \ge 1$ and $c \ge 1$, the constant $\alpha_{b,c} = \sum_{m=0}^{\infty} 1/(c^m b^{c^m})$ is b-normal.

In this more general case, we can write

$$b^n\alpha_{b,c} \bmod 1 = \left(\sum_{m=0}^{\lfloor \log_c n\rfloor} \frac{b^{n-c^m} \bmod c^m}{c^m}\right) \bmod 1 + \sum_{m=\lfloor \log_c n\rfloor+1}^{\infty} \frac{b^{n-c^m}}{c^m}\cdot \tag{24}$$

As before, note that the first expression can be generated by means of the recursion $z_0 = 0$ and, for $n \ge 1$, $z_n = (bz_{n-1} + r_n) \bmod 1$, where $r_n = 1/n$ if $n = c^k$ for some integer k, and zero otherwise. When c is prime, one always obtains a sequence of fractions quite similar to the sequence above (21). When c is composite, though, there is slight complication in that some powers of $1/c^p$ do not appear. For example, consider the case $b = 3$ and $c = 4$. Then the first few members of the z sequence are given as follows:

$$\begin{aligned}&0,\ 0,\ 0,\\&\frac{1}{4},\ \frac{3}{4},\quad \text{(repeated 6 times)}\\&\frac{5}{16},\ \frac{15}{16},\ \frac{13}{16},\ \frac{7}{16},\quad \text{(repeated 12 times)},\\&\frac{21}{64},\frac{63}{64},\frac{61}{64},\frac{55}{64},\frac{37}{64},\frac{47}{64},\frac{13}{64},\frac{39}{64},\frac{53}{64},\frac{31}{64},\frac{29}{64},\frac{23}{64},\frac{5}{64},\frac{15}{64},\frac{45}{64},\frac{7}{64},\\&\text{(repeated 12 times), etc.}\end{aligned} \tag{25}$$

Note here that the fraction $1/2$ is omitted in the first set, the fractions $1/8, 3/8, 5/8, 7/8$ are omitted in the second set, and the fractions with 32 in the denominators are omitted in the third set. Nonetheless, this critical property holds, both in this particular case and in general, so long as $b \ge 2$ and $c \ge 2$ are coprime: if $n < c^{p+1}$ then z_n is a multiple of $1/c^p$, and furthermore the set $(z_k, 1 \le k \le n)$ contains at most t repetitions of any particular value, where the integer t depends only on (b, c). For the case $(2, 3)$, the repetition factor $t = 3$. For the case $(3, 4)$, $t = 12$.

As before, these fractions (z_k) constitute an accurate set of approximations to the sequence $b^n \alpha_{b,c} \bmod 1$ of shifted fractions of $\alpha_{b,c}$. In fact, by examining (24) it can be readily seen that for all (b, c) as above and all $n \ge c$,

$$|b^n \alpha_{b,c} \bmod 1 - z_n| < \frac{1}{9n} \tag{26}$$

(and in most cases is much smaller than this).

To establish that $\alpha_{b,c}$ is b-normal via Lemma 1, we seek an upper bound for $b^m A(\alpha_{b,c}, y, n, m)/n$. A base-$b$ sequence y out to some length m, translated to a subset of the real unit interval, is $[r, s)$, where $r = 0.y_1y_2y_3\ldots y_m$, and s is the next largest base-b fraction of length m, so that $s - r = b^{-m}$. Observe that $A(\alpha_{b,c}, y, n, m)$ is equal to the number of those j between 0 and $n-1$ for which $b^j \alpha_{b,c} \bmod 1 \in [r, s)$. Also observe, in view of (26), that if $b^j \alpha \bmod 1 \in [r, s)$, then $z_j \in [r - 1/(9j), s + 1/(9j))$.

Let n be any integer greater than b^{2m}, and let c^p denote the largest power of c less than or equal to n, so that $c^p \le n < c^{p+1}$. Now note that for $j \ge b^m$, we have $[r - 1/(9j), s + 1/(9j)) \subset [r - b^{-m-1}, s + b^{-m-1})$. Since the length of this latter interval is no greater than $2b^{-m}$, the number of multiples of $1/c^p$ that it contains cannot exceed $\lfloor 2c^p b^{-m} \rfloor + 1$. Thus there can be at most t times this many j's less than n for which $z_j \in [r - b^{-m-1}, s + b^{-m-1})$. Therefore we can write

$$\begin{aligned}
\frac{b^m A(\alpha_{b,c}, y, n, m)}{n} &= \frac{b^m \#_{0 \le j < n}\big(b^j \alpha_{b,c} \bmod 1 \in [r,s)\big)}{n} \\
&\le \frac{b^m \big[b^m + \#_{b^m \le j < n}\big(z_j \in [r - b^{-m-1}, s + b^{-m-1})\big)\big]}{n} \\
&\le \frac{b^m \big[b^m + t(2c^p b^{-m} + 1)\big]}{n} < 2t + 2,
\end{aligned}$$

where t is the repetition factor for (b, c), mentioned above. For a fixed pair of integers (b, c), we have shown that for all $y \in [0, 1)$ and all $m > 0$,

$$\limsup_{n\to\infty} \frac{b^m A(\alpha_{b,c}, y, n, m)}{n} \le 2t + 2, \tag{27}$$

so by Lemma 1, $\alpha_{b,c}$ is b-normal. **QED**

Proof of Theorem 7: $\alpha_{2,3}$ is not 6-normal.

Let Q_m be the base-6 expansion of $\alpha_{2,3}$ immediately following position 3^m (i.e., after the "decimal" point has been shifted to the right 3^m digits). We can write

$$\begin{aligned}
Q_m &= 6^{3^m} \alpha_{2,3} \bmod 1 \\
&= \left(\sum_{k=0}^{m} 3^{3^m - k} 2^{3^m - 3^k}\right) \bmod 1 + \sum_{k=m+1}^{\infty} 3^{3^m-k} 2^{3^m - 3^k}.
\end{aligned} \tag{28}$$

The first portion of this expression is zero, since all terms in the summation are integers. The small second portion is very accurately approximated by the first term of the series, namely $(3/4)^{3^m}/3^{m+1}$. In fact, for all $m \ge 1$,

$$\frac{(3/4)^{3^m}}{3^{m+1}} \ < \ Q_m \ < \ \frac{(3/4)^{3^m}}{3^{m+1}}(1 + 2 \cdot 10^{-6}). \tag{29}$$

Let $Z_m = \lfloor \log_6 1/Q_m \rfloor$ be the number of zeroes in the base-6 expansion of α that immediately follow position 3^m. Then for all $m \ge 1$, (29) can be rewritten

$$\begin{aligned}
& 3^m \log_6 \left(\frac{4}{3}\right) + (m+1)\log_6 3 - 2 \\
< \ Z_m \ < \ & 3^m \log_6 \left(\frac{4}{3}\right) + (m+1) \log_6 3.
\end{aligned} \tag{30}$$

Now let F_m be the fraction of zeroes in the base-6 expansion of α up to position $3^m + Z_m$ (i.e., up to the end of the block of zeroes that immediately follows position 3^m). Clearly

$$F_m \quad > \quad \frac{\sum_{k=1}^m Z_k}{3^m + Z_m}, \tag{31}$$

since the numerator only counts zeroes in the long stretches. The summation in the numerator satisfies, for all sufficiently large m,

$$\begin{aligned}\sum_{k=1}^m Z_k \quad &> \quad \frac{3}{2}\left(3^m - \frac{1}{3}\right)\log_6\left(\frac{4}{3}\right) + \frac{m(m+3)}{2}\log_6 3 - 2m \\ &> \quad \frac{3}{2}\cdot 3^m \log_6\left(\frac{4}{3}\right) - \frac{1}{2}\log_6\left(\frac{4}{3}\right) - 2m. \end{aligned}\tag{32}$$

Now given any $\varepsilon > 0$, we can write, for all sufficiently large m,

$$\begin{aligned} F_m \quad &> \quad \frac{\frac{3}{2}\cdot 3^m \log_6\left(\frac{4}{3}\right) - \frac{1}{2}\log_6\left(\frac{4}{3}\right) - 2m}{3^m + 3^m\log_6\left(\frac{4}{3}\right) + (m+1)\log_6 3} \\ &= \quad \frac{\frac{3}{2}\log_6\left(\frac{4}{3}\right) - \frac{1}{3^m}\left(\frac{1}{2}\log_6\left(\frac{4}{3}\right) + 2m\right)}{1 + \log_6\left(\frac{4}{3}\right) + \frac{(m+1)\log_6 3}{3^m}} \\ &\geq \quad \frac{\frac{3}{2}\log_6\left(\frac{4}{3}\right) - \varepsilon}{1 + \log_6\left(\frac{4}{3}\right) + \varepsilon} \;\geq\; \frac{1}{2}\log_2\left(\frac{4}{3}\right) - 2\varepsilon. \end{aligned}\tag{33}$$

But $\beta = \frac{1}{2}\log_2(4/3)$ (which has numerical value $0.2075187496\ldots$) is clearly greater than $1/6$, since $(4/3)^3 = 64/27 > 2$. This means that infinitely often (namely, whenever $n = 3^m + Z_m$) the fraction of zeroes in the base-6 expansion of α up to position n exceeds $\frac{1}{2}(1/6 + \beta) > 1/6$. Thus α is not 6-normal. **QED**

Proof of Theorem 8: Given co-prime integers $b \geq 2$ and $c \geq 2$, the constant $\alpha_{b,c} = \sum_{k\geq 0} 1/(c^k b^{c^k})$ is not bc-normal.

Let $Q_m(b,c)$ be the base-bc expansion of $\alpha_{b,c}$ immediately following position c^m. Then

$$\begin{aligned} Q_m(b,c) \quad &= \quad (bc)^{c^m}\alpha_{b,c} \bmod 1 \\ &= \quad \left(\sum_{k=0}^m c^{c^m - k} b^{c^m - c^k}\right) \bmod 1 + \sum_{k=m+1}^\infty c^{c^m-k} b^{c^m - c^k}. \end{aligned}\tag{34}$$

As above, the first portion of this expression is zero, since all terms in the summation are integers, and the second portion is very accurately approximated by the first term of the series, namely $[\frac{c}{b(c-1)}]^{c^m}/c^{m+1}$. In fact, for any choice of b and c as above, and for all $m \geq 1$,

$$\frac{1}{c^{m+1}}\left[\frac{c}{b(c-1)}\right]^{c^m} \;<\; Q_m(b,c) \;<\; \frac{1}{c^{m+1}}\left[\frac{c}{b(c-1)}\right]^{c^m} \cdot (1+1/10). \tag{35}$$

Let $Z_m(b,c) = \lfloor \log_{bc} 1/Q_m(b,c) \rfloor$ be the number of zeroes that immediately follow position c^m. Then for all $m \geq 1$, (35) can be rewritten as

$$\begin{aligned} & c^m \log_{bc}\left[\frac{b(c-1)}{c}\right] + (m+1)\log_{bc} c - 2 \\ < \;\; & Z_m(b,c) \;<\; c^m \log_{bc}\left[\frac{b(c-1)}{c}\right] + (m+1)\log_{bc} c. \end{aligned} \tag{36}$$

Now let $F_m(b,c)$ be the fraction of zeroes up to position $c^m + Z_m(b,c)$. Clearly

$$F_m(b,c) \;>\; \frac{\sum_{k=1}^m Z_k(b,c)}{c^m + Z_m(b,c)}, \tag{37}$$

since the numerator only counts zeroes in the long stretches. The summation in the numerator of $F_m(b,c)$ satisfies

$$\begin{aligned} \sum_{k=1}^m Z_k(b,c) \;&>\; \frac{c}{c-1}\left(c^m - \frac{1}{c}\right)\log_{bc}\left[\frac{b(c-1)}{c}\right] + \frac{m(m+3)}{2}\log_{bc} c - 2m \\ &>\; \frac{c^{m+1}}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - \frac{1}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - 2m. \end{aligned} \tag{38}$$

Thus given any $\varepsilon > 0$, we can write, for all sufficiently large m,

$$\begin{aligned}
F_m(b,c) &> \frac{\frac{c^{m+1}}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - \frac{1}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - 2m}{c^m + c^m\log_{bc}\left(\frac{b(c-1)}{c}\right) + (m+1)\log_{bc} c} && (39)\\
&= \frac{\frac{c}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - \frac{1}{c^m}\left(\frac{1}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] + 2m\right)}{1 + \log_{bc}\left[\frac{b(c-1)}{c}\right] + \frac{(m+1)\log_{bc} c}{c^m}}\\
&\geq \frac{\frac{c}{c-1}\log_{bc}\left[\frac{b(c-1)}{c}\right] - \varepsilon}{1 + \log_{bc}\left[\frac{b(c-1)}{c}\right] + \varepsilon}\\
&\geq \frac{c}{c-1}\cdot\frac{\log_{bc}\left[\frac{b(c-1)}{c}\right]}{1 + \log_{bc}\left[\frac{b(c-1)}{c}\right]} - 2\varepsilon.\\
&= T(b,c) - 2\varepsilon, && (40)
\end{aligned}$$

where

$$T(b,c) = \frac{c}{c-1}\cdot\frac{\log_{bc}\left[\frac{b(c-1)}{c}\right]}{1 + \log_{bc}\left[\frac{b(c-1)}{c}\right]}. \qquad (41)$$

To establish the desired result that $T(b,c) > 1/(bc)$, first note that

$$T(b,c) > \frac{1}{2}\log_{bc}\left[\frac{b(c-1)}{c}\right] \geq \frac{1}{2}\log_{bc}\left(\frac{b}{2}\right). \qquad (42)$$

Raise bc to the power of the right-hand side, and also to the power $1/(bc)$. Then it suffices to demonstrate that

$$\frac{b}{2} > \left[(bc)^{1/(bc)}\right]^2. \qquad (43)$$

The right-hand side is bounded above by $(e^{1/e})^2 = 2.0870652286\ldots$. Thus this inequality is clearly satisfied whenever $b \geq 5$.

If we also presume that $c \geq 5$, then by examining the middle of (42) it suffices to demonstrate that

$$\frac{1}{2}\log_{bc}\frac{4b}{5} > \frac{1}{bc} \qquad (44)$$

or

$$\frac{4b}{5} \quad > \quad \left(e^{1/e}\right)^2 . \tag{45}$$

But this is clearly satisfied whenever $b \geq 3$. For the case $b = 2$ and $c \geq 5$, we can write

$$T(b,c) \quad = \quad \frac{c}{c-1} \cdot \frac{\log_{2c}\left[\frac{2(c-1)}{c}\right]}{1 + \log_{2c}\left[\frac{2(c-1)}{c}\right]} \;\geq\; \frac{\log_{2c}\left[\frac{2(c-1)}{c}\right]}{1 + \log_{10} 2}, \tag{46}$$

so by similar reasoning it suffices to demonstrate that

$$\frac{2(c-1)}{c} \quad > \quad \left(e^{1/e}\right)^{1+\log_{10} 2} \quad = \quad 1.61384928833\ldots . \tag{47}$$

But this is clearly satisfied whenever $c \geq 6$.

The five remaining cases, namely $(2,3), (2,5), (3,2), (3,4), (4,3)$, are easily verified by explicitly computing numerical values of $T(b,c)$ using (41). As it turns out, the simple case that we worked out in detail above, namely $b = 2$ and $c = 3$, is the worst case, in the sense that for all other (b,c), the fraction $T(b,c)$ exceeds the natural frequency $1/(bc)$ by greater margins. **QED**

The proofs of Theorems 7 and 8 above are adapted from [1].

References

[1] David Bailey and Jonathan Borwein, "Normal numbers and pseudorandom generators," submitted to *Proceedings of the Workshop on Computational and Analytical Mathematics in Honour of Jonathan Borwein's 60th Birthday*, Springer, September 2011, available at http://crd.lbl.gov/~dhbailey/dhbpapers/normal-pseudo.pdf.

[2] David H. Bailey, Jonathan M. Borwein, Richard E. Crandall and Carl Pomerance, "On the binary expansions of algebraic numbers," *Journal of Number Theory Bordeaux*, vol. 16 (2004), pg. 487–518.

[3] David H. Bailey, Peter B. Borwein and Simon Plouffe, "On the rapid computation of various polylogarithmic constants," *Mathematics of Computation*, vol. 66, no. 218 (Apr 1997), pg. 903–913.

[4] David H. Bailey and David J. Broadhurst, "Parallel integer relation detection: Techniques and applications," *Mathematics of Computation*, vol. 70, no. 236 (Oct 2000), pg. 1719–1736.

[5] David H. Bailey and Richard E. Crandall, "On the random character of fundamental constant expansions," *Experimental Mathematics*, vol. 10, no. 2 (Jun 2001), pg. 175–190.

[6] David H. Bailey and Richard E. Crandall, "Random generators and normal numbers," *Experimental Mathematics*, vol. 11 (2002), no. 4, pg. 527–546.

[7] David H. Bailey and Michal Misiurewicz, "A strong hot spot theorem," *Proceedings of the American Mathematical Society*, vol. 134 (2006), no. 9, pg. 2495–2501.

[8] L. Berggren, J. M. Borwein and P. B. Borwein, *Pi: A Source Book,* Springer-Verlag, Third Edition, 2004.

[9] J. M. Borwein and D. H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A K Peters, Natick, MA, second edition, 2008.

[10] J. M. Borwein and P. B. Borwein, Pi and the AGM: *A Study in Analytic Number Theory and Computational Complexity*, John Wiley, New York, 1987, paperback 1998.

[11] P. B. Borwein, "On the irrationality of certain series," *Mathematical Proceedings of the Cambridge Philosophical Society* **112** (1992), pg. 141–146.

[12] C. S. Calude, "Borel normality and algorithmic randomness," in G. Rozenberg, A. Salomaa (eds.), *Developments in Language Theory*, World Scientific, Singapore, 1994, pg. 113–129.

[13] C. S. Calude, *Information and Randomness: An Algorithmic Perspective*, 2nd Edition, Revised and Extended, Springer-Verlag, Berlin, 2002.

[14] D. G. Champernowne, "The construction of decimals normal in the scale of ten," *Journal of the London Mathematical Society*, vol. 8 (1933), pg. 254–260.

[15] A. H. Copeland and P. Erdös, "Note on normal numbers," *Bulletin of the American Mathematical Society*, vol. 52 (1946), pg. 857–860.

[16] R. E. Crandall, "The googol-th bit of the Erdös-Borwein constant," preprint, November 8, 2011.

[17] Hajime Kaneko, "On normal numbers and powers of algebraic numbers," *Integers*, vol. 10 (2010), pg. 31–64.

[18] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, 1983.

[19] Martine Queffelec, "Old and new results on normality," *Lecture Notes – Monograph Series*, vol. 48, *Dynamics and Stochastics*, 2006, Institute of Mathematical Statistics, pg. 225–236.

[20] W. Schmidt, "On normal numbers," *Pacific Journal of Mathematics*, vol. 10 (1960), pg. 661–672.

[21] R. Stoneham, "On absolute (j, ε)-normality in the rational fractions with applications to normal numbers," *Acta Arithmetica*, vol. 22 (1973), pg. 277–286.

[22] A. J. Yee, "Y-cruncher-multi-threaded pi program," http://www.numberworld.org/y-cruncher, 2010.

[23] A. J. Yee and S. Kondo, "10 trillion digits of pi: A case study of summing hypergeometric series to high precision on multicore systems," preprint, 2011, available at http://hdl.handle.net/2142/28348.