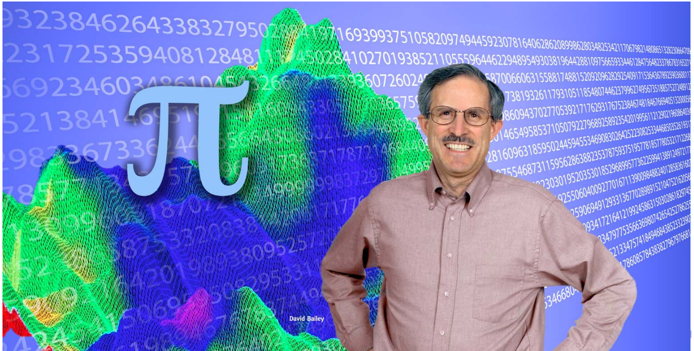
Are the Digits of Pi Random?

David H. Bailey
http://www.davidhbailey.com
Lawrence Berkeley National Laboratory (retired)
Computer Science Department, University of California, Davis



"Pi" by Kate Bush

Sweet and gentle sensitive man
With an obsessive nature and deep fascination
For numbers
And a complete infatuation with the calculation
Of PI

Oh he love, he love, he love
He does love his numbers
And they run, they run, they run him
In a great big circle
In a circle of infinity

3.1415926535 897932
3846 264 338 3279

Oh he love, he love, he love
He does love his numbers
And they run, they run, they run him
In a great big circle
In a circle of infinity
But he must, he must, he must
Put a number to it

50288419 716939937510
582319749 44 59230781
6406286208 821 4808651 32

Oh he love, he love, he love
He does love his numbers
And they run, they run, they run him
In a great big circle
In a circle of infinity

82306647 0938446095 505 8223...

New York Times PiDay (March 14, 2007) crossword puzzle

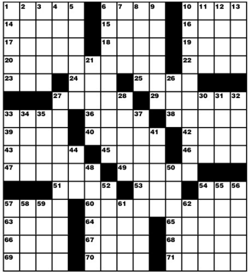
The New York Times Crossword | Edited by Will Shortz | No. 0314

Across

1 Enlighten
6 A couple CBS spinoffs
10 1972 Broadway musical
14 Metal giant
15 Evict
16 Area
17 Surface again, as a road
18 Pirate or Padre, briefly
19 Camera feature
20 Barracks artwork, perhaps
22 River to the Ligurian Sea
23 Keg necessity
24 "… ___ he drove out of sight"
25 ___ St. Louis, Ill.
27 Preen
29 Greek peak
33 Vice president after Hubert
36 Patient wife of Sir Geraint
38 Action to an ante
39 Gain ___
40 French artist Odilon ___
42 Grape for winemaking
43 Single-dish meal
45 Broad valley
46 See 21-Down
47 Artery inserts
49 Offspring
51 Mexican mouse catcher
53 Medical procedure, in brief
54 "Wheel of Fortune" option
57 Animal with striped legs
60 Editorial
63 It gets bigger at night
64 "Hold your horses!"
65 Idiots
66 Europe/Asia border river
67 Suffix with launder
68 Leaning
69 Brownback and Obama, e.g.: Abbr.
70 Rick with the 1976 #1 hit "Disco Duck"
71 Yegg's targets

Down

1 Mastodon trap
2 "Mefistofele" soprano
3 Misbehave
4 Pen
5 More pleased
6 Treated with disdain
7 Enterprise crewman
8 Rhone feeder
9 Many a webcast
10 Mushroom, for one
11 Unfortunate
12 Nevada's state tree
13 Disney fish
21 Colonial figure with 46-Across
26 Poker champion Ungar
27 Self-medicating excessively
28 March 14, to mathematicians
30 Book part
31 Powder, e.g.
32 007 and others: Abbr.
33 Drains
34 Stove feature
35 Feet per second, e.g.
37 Italian range
41 Prefix with surgery
44 Captain's announcement, for short
48 Tucked away
50 Stealthy fighters
52 Sedative
54 Letter feature
55 Jam
56 Settles in
57 Symphony or sonata
58 Japanese city bombed in W.W. II
59 Beelike
61 Evening, in ads
62 Religious artwork



Puzzle by Peter A. Collins

ANSWER TO PREVIOUS PUZZLE

A	R	F	S		A	C	H	E		O	R	G	A	N
C	O	R	K		T	R	E	X		K	E	R	R	Y
O	D	A	Y		L	A	I	T		S	T	A	T	S
L	E	N		R	A	N	D	O	M		I	P	S	E
D	O	C	T	O	R	K	I	L	D	A	R	E		
			R	A	G				S	I	E	S	T	A
T	Y	R	O	N	E	P	O	W	E	R		O	H	M
R	U	E	D			A	L	I			I	D	E	S
I	M	P		H	O	L	D	T	H	E	M	A	Y	O
B	A	A	B	A	A				O	R	E			
		I	R	I	S	H	C	O	U	N	T	I	E	S
P	E	R	I		T	I	R	A	D	E		D	X	C
A	R	M	E	D		T	A	T	I		Y	I	P	E
C	L	A	R	E		I	T	E	N		D	O	W	N

Solution to crossword puzzle

ANSWER TO PREVIOUS PUZZLE

T	E	A	C	H		C	S	I	S		π	P	π	N
A	L	C	O	A		O	U	S	T		Z	O	N	E
R	E	T	O	P		N	L	E	R		Z	O	O	M
π	N	U	P	π	C	T	U	R	E		A	R	N	O
T	A	P		E	R	E		E	A	S	T			
			P	R	I	M	P		M	T	O	S	S	A
S	π	R	O		E	N	I	D		U	P	π	N	G
A	L	A	P		R	E	D	O	N		π	N	O	T
P	O	T	π	E		D	A	L	E		N	E	W	S
S	T	E	N	T	S		Y	O	U	N	G			
			G	A	T	O		M	R	I		S	π	N
O	K	A	π		O	π	N	I	O	N	π	E	C	E
P	U	π	L		W	A	I	T		J	E	R	K	S
U	R	A	L		E	T	T	E		A	T	I	L	T
S	E	N	S		D	E	E	S		S	A	F	E	S

Pre-computer history of π

Name	Year	Digits
Babylonians	2000? BCE	1
Hebrews (1 Kings 7:23)	550? BCE	1
Archimedes	250? BCE	3
Ptolemy	150	3
Liu Hui	263	5
Tsu Ch'ung Chi	480?	7
Madhava	1400?	13
Al-Kashi	1429	14
Romanus	1593	15
Van Ceulen	1615	35
Sharp (and Halley)	1699	71
Machin	1706	100
Strassnitzky and Dase	1844	200
Rutherford	1853	440
W. Shanks	1853	(607) 527
W. Shanks	1874	(707) 527
Ferguson (Calculator)	1947	808

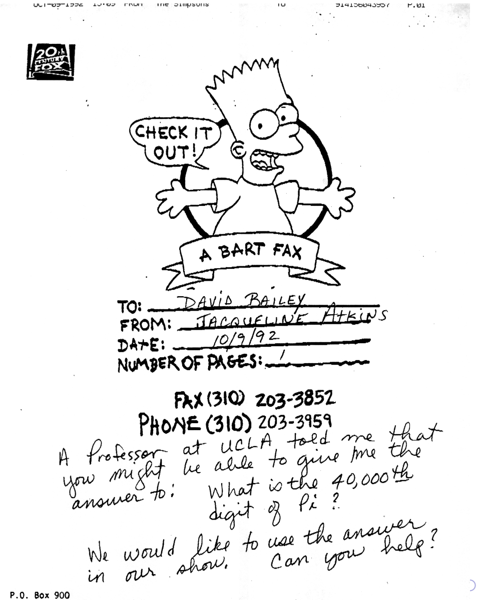
Computer-era π calculations

Name	Year	Correct Digits
Miyoshi and Kanada	1981	2,000,036
Kanada-Yoshino-Tamura	1982	16,777,206
Gosper	1985	17,526,200
Bailey	Jan. 1986	29,360,111
Kanada et. al	Jan. 1987	134,217,700
Kanada and Tamura	Jan. 1988	201,326,551
Kanada and Tamura	Nov. 1989	1,073,741,799
Chudnovskys	Aug. 1991	2,260,000,000
Chudnovskys	May 1994	4,044,000,000
Kanada and Takahashi	Jul. 1997	51,539,600,000
Kanada and Takahashi	Sep. 1999	206,158,430,000
Kanada-Ushiro-Kuroda	Dec. 2002	1,241,100,000,000
Takahashi	Jan. 2009	1,649,000,000,000
Takahashi	Apr. 2009	2,576,980,377,524
Bellard	Dec. 2009	2,699,999,990,000
Kondo and Yee	Aug. 2010	5,000,000,000,000
Kondo and Yee	Dec. 2013	12,000,000,000,000

DHB receives a fax from the Simpsons

In October 1992, I received this fax from the Simpsons TV crew, requesting the 40,000th digit of π. I computed the first 40,000 digits, and faxed the result back (noting that the 40,000th digit is a 1).

In the episode airing May 6, 1993, Apu, the manager of a convenience store, was challenged by Marge's attorney in a courtroom. He replied that he has a perfect memory. For example, he said, he can recite 40,000 digits of π, and the last digit is a 1.



TO: DAVID BAILEY
FROM: JACQUELINE ATKINS
DATE: 10/9/92
NUMBER OF PAGES: 1

FAX (310) 203-3852
PHONE (310) 203-3959

A Professor at UCLA told me that you might be able to give me the answer to: What is the 40,000th digit of Pi?

We would like to use the answer in our show. Can you help?

P.O. Box 900

Normal numbers

Given integer $b \geq 2$, a real number x is b-normal (or "normal base b") if every m-long string of digits appears in the base-b expansion of x with limiting frequency $1/b^m$.

Using measure theory, it is can be shown that almost all real numbers are b-normal for a given integer base b (in fact, for all b simultaneously).

These are widely believed to be b-normal, for all integer bases $b \geq 2$:

- $\pi = 3.14159265358979323846\ldots$
- $e = 2.7182818284590452354\ldots$
- $\sqrt{2} = 1.41421356237309504880\ldots$
- $\log(2) = 0.69314718055994530942\ldots$
- *Every* irrational algebraic number (this conjecture is due to Borel).

But there are *no proofs of normality* for any of these constants in any base.

Until recently, normality proofs were known only for a few contrived examples such as Champernowne's constant $= 0.12345678910111213141516\ldots$ (which is 10-normal).

One (very weak) result for algebraic numbers

If x is algebraic of degree $d > 1$, then its binary expansion through position n must have at least $Cn^{1/d}$ 1-bits, for all sufficiently large n and for some C that depends on x.

Simple case: The first n binary digits of $\sqrt{2}$ must have at least $\sqrt{n}$ one bits.

In this case, the result follows by noting that the one-bit count of the product of two integers is less than or equal to the product of the one-bit counts of the two integers. The more general result above requires a more sophisticated approach.

$$
\begin{array}{rrrrrrrrrr}
 & & & & & 1 & 0 & 1 & 1 & 1 \\
 & & & & \times & 1 & 1 & 0 & 1 & 1 \\
\hline
1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1
\end{array}
$$

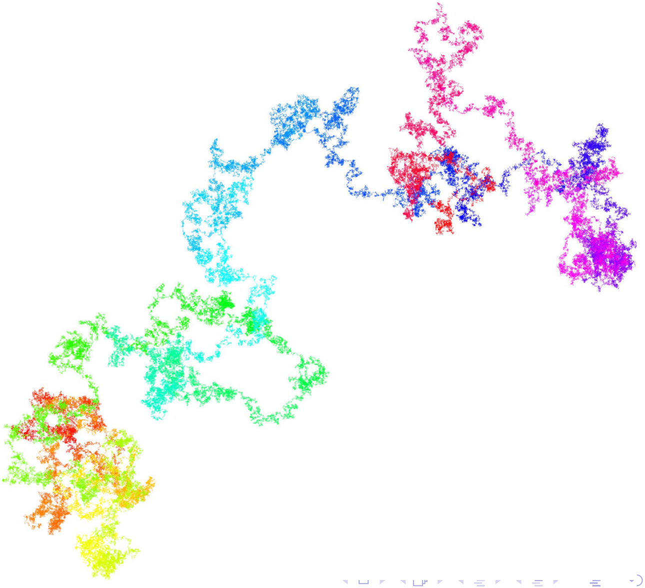
However, note that these results are still a far cry from even single-digit normality.

- D. H. Bailey, J. M. Borwein, R. E. Crandall and C. Pomerance, "On the Binary Expansions of Algebraic Numbers," *Journal of Number Theory Bordeaux*, vol. 16 (2004), pg. 487–518.

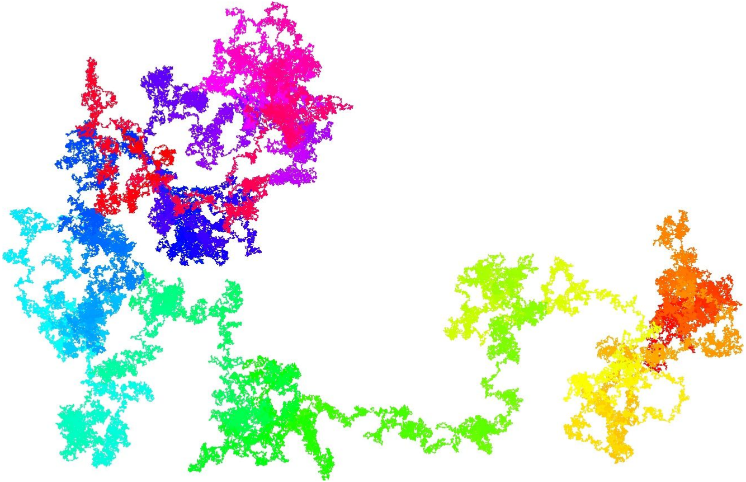
A random walk on π

This is a plot of a "random walk" on the first 10 billion base-4 digits of π. At each point, the plot moves to right one unit for a 0, up one unit for 1, left one unit for 2, or down one unit for 3. The colors indicate position in the sequence.

A user-searchable random walk on the first 100 billion base-4 digits of π (courtesy F. Aragon-Artacho) is available at `http://gigapan.com/gigapans/106803`.



A random walk on pseudorandom data



Experimental mathematics: Discovering new mathematical results by computer

Methodology:

1. Compute various mathematical entities (limits, infinite series sums, definite integrals, etc.) to high precision, typically 100–10,000 digits.
2. Use algorithms such as PSLQ to recognize these numerical values in terms of well-known mathematical constants.
3. When results are found experimentally, seek formal mathematical proofs of the discovered relations.

Many results have recently been found using this methodology, both in pure mathematics and in mathematical physics.

> *"If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied in mathematics just the same as in physics." – Kurt Godel*

The PSLQ integer relation algorithm

Let (x_n) be a given vector of real numbers. An integer relation algorithm either finds integers (a_n) such that

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \quad = \quad 0$$

(to within the "epsilon" of the arithmetic being used), or else finds bounds within which no relation can exist.

The "PSLQ" algorithm of mathematician-sculptor Helaman Ferguson is the most widely used integer relation algorithm.

Integer relation detection requires very high precision (at least $n \times d$ digits, where d is the size in digits of the largest a_k), both in the input data and in the operation of the algorithm.

1. H.R.P. Ferguson, D.H. Bailey and S. Arno, "Analysis of PSLQ, An Integer Relation Finding Algorithm," *Mathematics of Computation*, vol. 68, no. 225 (Jan 1999), pg. 351–369.
2. D.H. Bailey and D.J. Broadhurst, "Parallel Integer Relation Detection: Techniques and Applications," *Mathematics of Computation*, vol. 70, no. 236 (Oct 2000), pg. 1719–1736.

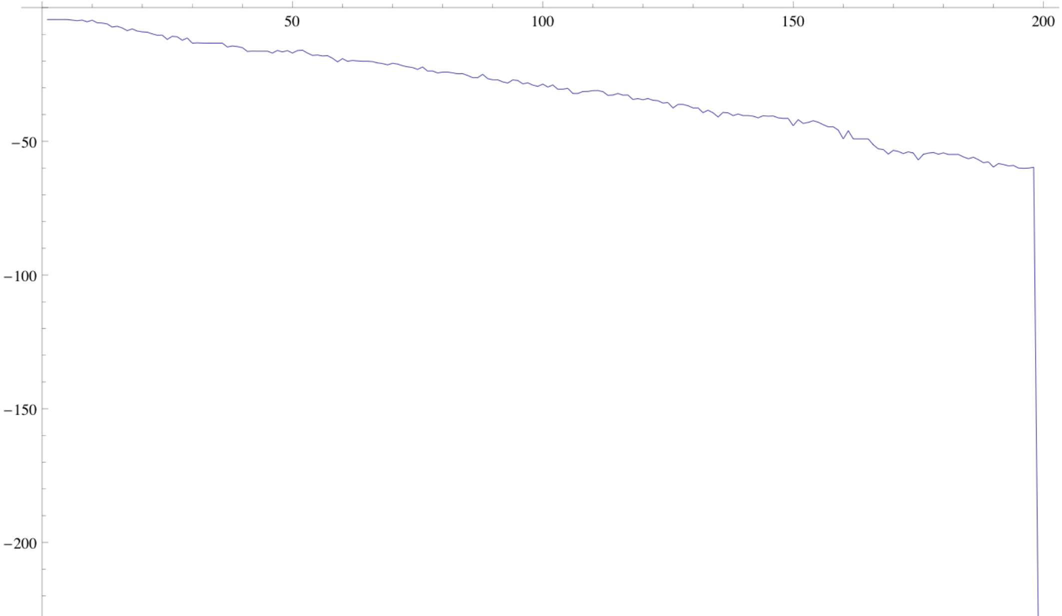
PSLQ, continued

- PSLQ constructs a sequence of integer-valued matrices B_n that reduce the vector $y = x \cdot B_n$, until either the relation is found (as one of the columns of matrix B_n), or else precision is exhausted.
- A relation is detected when the size of smallest entry of the y vector suddenly drops to roughly "epsilon" (i.e. 10^{-p}, where p is the number of digits of precision).
- The size of this drop can be viewed as a "confidence level" that the relation is not a numerical artifact: a drop of 20+ orders of magnitude almost always indicates a real relation.

Efficient variants of PSLQ:

- 2-level and 3-level PSLQ perform almost all iterations with only double precision, updating full-precision arrays as needed. They are hundreds of times faster than the original PSLQ.
- Multi-pair PSLQ dramatically reduces the number of iterations required. It was designed for parallel systems, but runs faster even on 1 CPU.

Decrease of $\log 10(\min |y_i|)$ in multipair PSLQ run



The first major PSLQ discovery: The BBP formula for π

In 1996, a PSLQ program discovered this new formula for π:

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

This formula permits one to compute binary (or hexadecimal) digits of π beginning at an arbitrary starting position, using a very simple scheme that requires only standard 64-bit or 128-bit arithmetic.

In 2004, Borwein, Galway and Borwein proved that no base-n formulas of this type exist for π, except when $n = 2^m$.

BBP-type formulas (discovered with PSLQ) are now known for numerous other mathematical constants.

1. D.H. Bailey, P.B. Borwein and S. Plouffe, "On the rapid computation of various polylogarithmic constants," *Mathematics of Computation*, vol. 66, no. 218 (Apr 1997), pg. 903–913.
2. J.M. Borwein, W.F. Galway and D. Borwein, "Finding and excluding b-ary Machin-type BBP formulae," *Canadian Journal of Mathematics*, vol. 56 (2004), pg. 1339–1342.

Some other BBP-type formulas found using PSLQ

$$\pi^2 = \frac{1}{8}\sum_{k=0}^{\infty}\frac{1}{64^k}\left(\frac{144}{(6k+1)^2}-\frac{216}{(6k+2)^2}-\frac{72}{(6k+3)^2}-\frac{54}{(6k+4)^2}+\frac{9}{(6k+5)^2}\right)$$

$$\begin{aligned}\pi^2 = \frac{2}{27}\sum_{k=0}^{\infty}\frac{1}{729^k}&\left(\frac{243}{(12k+1)^2}-\frac{405}{(12k+2)^2}-\frac{81}{(12k+4)^2}-\frac{27}{(27k+5)^2}\right.\\&\left.-\frac{72}{(12k+6)^2}-\frac{9}{(12k+7)^2}-\frac{9}{(12k+8)^2}-\frac{5}{(12k+10)^2}+\frac{1}{(12k+11)^2}\right)\end{aligned}$$

$$\begin{aligned}\zeta(3) = \frac{1}{1792}\sum_{k=0}^{\infty}\frac{1}{2^{12k}}&\left(\frac{6144}{(24k+1)^3}-\frac{43008}{(24k+2)^3}+\frac{24576}{(24k+3)^3}+\frac{30720}{(24k+4)^3}-\frac{1536}{(24k+5)^3}\right.\\&+\frac{3072}{(24k+6)^3}+\frac{768}{(24k+7)^3}-\frac{3072}{(24k+9)^3}-\frac{2688}{(24k+10)^3}-\frac{192}{(24k+11)^3}-\frac{1536}{(24k+12)^3}\\&-\frac{96}{(24k+13)^3}-\frac{672}{(24k+14)^3}-\frac{384}{(24k+15)^3}+\frac{24}{(24k+17)^3}+\frac{48}{(24k+18)^3}-\frac{12}{(24k+19)^3}\\&\left.+\frac{120}{(24k+20)^3}+\frac{48}{(24k+21)^3}-\frac{42}{(24k+22)^3}+\frac{3}{(24k+23)^3}\right)\end{aligned}$$

- D. H. Bailey, J. M. Borwein, A. Mattingly and G. Wightwick, "The computation of previously inaccessible digits of Pi^2 and Catalan's constant," *Notices of the AMS*, vol. 60 (2013), no. 7, pg. 844–854.

How to compute arbitrary digits of $\log 2$

The BBP algorithm for computing binary digits of π starting at an arbitrary position can be shown more easily for $\log 2$:

$$\log 2 = \sum_{n=1}^{\infty} \frac{1}{n2^n} = 0.10110001011100100001011111110111110100011100111101 1\ldots_2$$

Note that the binary digits of $\log 2$ beginning after position d can be written as $\{2^d \log 2\}$, where $\{\cdot\}$ denotes fractional part. Thus we can write:

$$\begin{aligned}
\{2^d \log 2\} &= \left\{ \sum_{n=1}^{d} \frac{2^{d-n}}{n} \right\} + \sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n} \\
&= \left\{ \sum_{n=1}^{d} \frac{2^{d-n} \mod n}{n} \right\} + \sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n}
\end{aligned}$$

where we have inserted $\bmod n$ since were are only interested in the fractional part when divided by n. Now note that the numerator $2^{d-n} \mod n$ can be calculated very rapidly by the binary algorithm for exponentiation.

The binary algorithm for exponentiation

Problem: What is $3^{17} \bmod 10$?

Algorithm A:
$3^{17} = 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times \times 3 = 129140163$,
so answer $= 3$.

Algorithm B (faster): $3^{17} = ((((3^2)^2)^2)^2) \times 3 = 129140163$, so answer $= 3$.

Algorithm C (fastest):
$3^{17} = ((((3^2 \bmod 10)^2 \bmod 10)^2 \bmod 10)^2 \bmod 10) \times 3 \bmod 10 = 3$.

Note that in Algorithm C, we never have to deal with integers larger than $9 \times 9 = 81$.
In general, if reducing mod n, we never have to deal with integers larger than $(n-1)^2$.

BBP formulas and normality

Consider a general BBP-type constant (i.e., a formula that permits the BBP "trick"):

$$\alpha = \sum_{n=0}^{\infty} \frac{p(n)}{b^n q(n)},$$

where p and q are integer polynomials, $\deg p < \deg q$, and q has no zeroes for nonnegative arguments.

In 2001, myself and Richard Crandall proved that α is b-normal iff the sequence $x_0 = 0$, and

$$x_n = \left\{ b x_{n-1} + \frac{p(n)}{q(n)} \right\},$$

where $\{\cdot\}$ again denote fractional part, is equidistributed in the unit interval. Here "equidistributed" means that the sequence visits each subinterval $[c, d)$ with limiting frequency $d - c$.

- D. H. Bailey and R. E. Crandall, "On the Random Character of Fundamental Constant Expansions," *Experimental Mathematics*, vol. 10, no. 2 (Jun 2001), pg. 175–190.

Two specific examples

Consider the sequence $x_0 = 0$ and

$$x_n = \left\{ 2x_{n-1} + \frac{1}{n} \right\}$$

Then $\log 2$ is 2-normal iff this sequence is equidistributed in the unit interval.

Similarly, consider the sequence $x_0 = 0$ and

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

Then π is 16-normal (and hence 2-normal) iff this sequence is equidistributed in the unit interval.

A class of provably normal constants

Myself and Crandall also proved that an infinite class of constants are 2-normal, e.g.

$$
\begin{aligned}
\alpha_{2,3} &= \sum_{n=1}^{\infty} \frac{1}{3^n 2^{3^n}} \\
&= 0.041883680831502985071252898624571682426096\ldots_{10} \\
&= 0.0ab8e38f684bda12f684bf35ba781948b0fcd6e9e0\ldots_{16}
\end{aligned}
$$

This constant was proven 2-normal by Stoneham in 1971, but we have extended this to the case where $(2,3)$ are any pair (p,q) of relatively prime integers ≥ 2. We also extended this result to an uncountable class: For any real r in $[0,1)$, the constant

$$
\alpha_{2,3}(r) = \sum_{n=1}^{\infty} \frac{1}{3^n 2^{3^n + r_n}}
$$

is 2-normal, where r_n is the n-th bit in the binary expansion of r in $[0,1)$. These constants are all distinct, so the class is uncountable.

- D. H. Bailey and R. E. Crandall, "Random Generators and Normal Numbers," *Experimental Mathematics*, vol. 11, no. 4 (2002), pg. 527–546.

The constant $\alpha_{2,3}$ is *not* 6-normal

First 900 base-6 digits of $\alpha_{2,3}$:

3130140430003334251130502130000001243555045432233011500243525320551352
3435410104300000000000000005141130054040555455303144250433435101241345
2351125142125134505503545015053522052044340452151505102411552500425130
0511244540010441311500324203032130000000000000000000000000000000000000
0000014212034311121452013525445342113412240220525301054204423552411055
4150155204350414555400310145303033532002534340401301240104453254343502
1420204324150255551010040433000455441145010313314511510144514123443342
3412400551313335045423530553151153501533452435450250055521453054234342
1530350125024205404135451231323245353031534552304115020154242121145201
5422225343403404505301233255344404431033324453321414150142334545424124
3203125340050134150245514404300000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000
0000000000031335054244443111055534141052014540213412313001424333133115

Note the long stretches of zeroes after power-of-three positions, e.g., after 3, 9, 27, 81, 243, 727, etc. This observation can be fashioned into a rigorous proof of non-normality.

Two new normality results on Stoneham constants

1. Given co-prime integers $b \geq 2$ and $c \geq 2$, and integers $p, q, r \geq 1$, with neither b nor c dividing r, let $B = b^p c^q r$, and assume $D = c^{q/p} r^{1/p} / b^{c-1} < 1$. Then the constant

$$\alpha_{b,c} = \sum_{n=1}^{\infty} \frac{1}{c^n b^{c^n}}$$

is B-nonnormal. Thus, for example, $\alpha_{b,c}$ is b-normal but bc-nonnormal. It is not known whether or not this is a complete catalog of the bases for which a Stoneham constant is probably nonnormal.

2. Let α_{b_1,c_1} and α_{b_2,c_2} be two Stoneham constants satisfying the conditions above to be B-nonnormal. Assume further there are no integers s and t such that $c_1^s = c_2^t$. Then $\alpha_{b_1,c_1} + \alpha_{b_2,c_2}$ is B-nonnormal.

- D. H. Bailey and J. M. Borwein, "Nonnormality of Stoneham constants," *Ramanujan Journal*, vol. 29 (2012), pg. 409–422.

Summary

This talk is available at `http://www.davidhbailey.com/dhbtalks/dhb-normality.pdf`.

References:

- David H. Bailey and Jonathan M. Borwein, "Pi day is upon us again, and we still do not know if pi is normal," *American Mathematical Monthly*, March 2014, pg. 191–206, available at urlhttp://www.davidhbailey.com/dhbpapers/pi-monthly.pdf.
- David H. Bailey and Jonathan M. Borwein, "Nonnormality of Stoneham constants," *Ramanujan Journal*, vol. 29 (2012), pg. 409–422, available at `http://www.davidhbailey.com/dhbpapers/nonnormality.pdf`.