

Regularity versus Complexity in the Binary Representation of 3^n

Eric S. Rowland

*Mathematics Department
Tulane University
New Orleans, LA 70118, USA*

We use the grid consisting of bits of 3^n to motivate the definition of 2-adic numbers. Specifically, we exhibit diagonal stripes in the bits of 3^{2^n} , which turn out to be the first in an infinite sequence of such structures. Our observations are explained by a 2-adic power series, providing some regularity among the disorder in the bits of powers of 3. Generally, the base- p representation of k^{p^n} has these features.

1. Several Mysteries

The binary representation of a number m can be thought of as encoding the unique set of distinct powers of 2 that sum to m . For example,

$$81 = 1010001_2 = 1 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 2^6 + 2^4 + 2^0.$$

We will display binary representations graphically by rendering 0 and 1 respectively as \square and \blacksquare . For reasons that will become clear, the convention in this paper when displaying binary representations graphically is to reverse the order of the digits relative to the standard ordering, so that higher indices are to the right. For example, we write $81 = \blacksquare\square\square\square\square\square\blacksquare = 2^0 + 2^4 + 2^6$, where the dot identifies the 2^0 position (somewhat like a decimal point).

The binary representations of the first several powers of 3 grow steadily in length:

$$\begin{aligned} 3^0 &= \blacksquare \\ 3^1 &= \blacksquare\blacksquare \\ 3^2 &= \blacksquare\square\square\blacksquare \\ 3^3 &= \blacksquare\square\square\blacksquare\blacksquare \\ 3^4 &= \blacksquare\square\square\square\blacksquare\blacksquare\blacksquare \end{aligned}$$

Figure 1 displays a grid in which the n^{th} row contains the binary digits of 3^n . Pictures like this were considered by Stephen Wolfram [1,

page 119]. Small triangles and other local features can be seen, but overall we get an impression of uniform disorder. There is no global structure evident, aside from the right boundary of the pattern, which has slope $\log_2 3$.

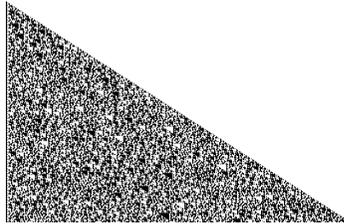


Figure 1. Powers of 3 in base 2. The n^{th} row consists of the binary digits of 3^n , in order of increasing exponents.

There are some global regularities, however. In particular, every column is eventually periodic. This is because there are only finitely many (in fact, 2^a) states that can be assumed by the first a columns taken together, so eventually the first a columns return to a state that they have reached before, at which point they become periodic.

In fact, each column is not just eventually periodic but periodic from the start. This is because each row has a unique predecessor, namely the integer obtained by dividing by 3. Put algebraically, 3 is invertible modulo 2^a for every $a \geq 1$, so from a given row we may compute the previous row to as many bits as we want.

What if we try to compute a bits of “row -1 ”, that is, the predecessor to the initial condition? Certainly we can do this, and the result simply maintains the periodicity of the columns. We can iteratively compute predecessors and thereby uniquely continue the picture up the page. Figure 2 shows the end of the unique infinite “history” leading up to the initial condition.

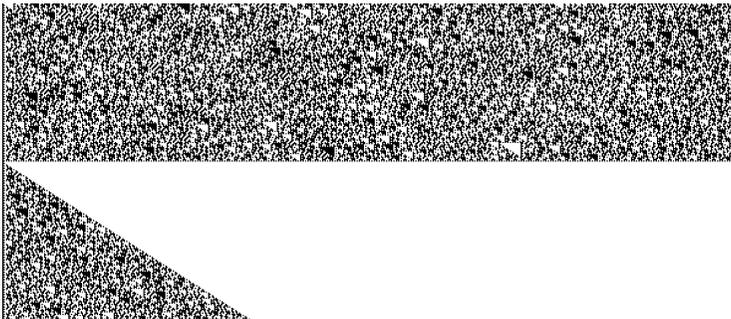


Figure 2. Part of the history obtained by periodically continuing each column up the page.

For those readers familiar with cellular automata, we mention that the ability to evolve the system backward in time is analogous to the same ability in a class of cellular automata whose local rules are bijective functions in the rightmost position. As with multiplying each row by 3 to form the next, information only propagates to the right and is not lost in these automata, and consequently they are reversible under the condition that the left half of each row is determined (say, all white) [2]. Indeed, several themes of the present paper can be carried over to such cellular automata.

But what do the bits in such rows mean? Rows $n < 0$ in the history illustrated in Figure 2 do not represent integers, since they contain 1s in positions arbitrarily far to the right: The sum used to compute the value of such a row diverges. For example, row -1 represents the “infinite integer”

$$\dots \square \square \square \square \blacksquare \blacksquare \blacksquare \blacksquare \blacksquare \blacksquare \square \dots = 2^0 + 2^1 + 2^3 + 2^5 + 2^7 + 2^9 + \dots = 1 + \sum_{i=0}^{\infty} 2^{2i+1}.$$

However, formally applying the geometric series formula to this divergent series produces

$$1 + 2 \sum_{i=0}^{\infty} 4^i \stackrel{?}{=} 1 + \frac{2}{1-4} = \frac{1}{3} = 3^{-1},$$

which is certainly a natural object for row $n = -1$ to correspond to. Similarly, row -2 represents

$$\dots \square \square \square \square \blacksquare \square \blacksquare \blacksquare \square \blacksquare \blacksquare \square \square \square \dots = 2^0 + \sum_{i=0}^{\infty} (2^{6i+3} + 2^{6i+4} + 2^{6i+5}) \stackrel{?}{=} 1 + \frac{2^3}{1-2^6} + \frac{2^4}{1-2^6} + \frac{2^5}{1-2^6} = \frac{1}{9} = 3^{-2}.$$

This is our first mystery, and in fact for each $n < 0$ there is a divergent series which produces 3^n under invalid applications of the geometric series formula.

Mystery 1. Each power 3^n for $n < 0$ is the “sum” of a divergent series.

In order to resolve this mystery we must first encounter several additional mysteries, all related to the first, regarding the binary representation of 3^n .

Since there are only two cell values (\square and \blacksquare) in Figure 1, the period length of each column is a power of 2. (In fact for $a \geq 3$, the period length of $3^n \bmod 2^a$ is 2^{a-2} .) Therefore, another consequence

satisfies

$$\lim_{n \rightarrow \infty} \frac{\frac{3^{2^n} - c_0}{2^n} - c_1}{2^n} - c_2 = 0.$$

Let us take this expression and unravel it to see the structure better. We find

$$3^{2^n} - (c_0 + c_1 2^n + c_2 2^{2^n} + c_3 2^{3^n}) \rightarrow 0$$

as $n \rightarrow \infty$. Replacing 2^n with x reveals that this is a power series:

$$3^x = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + O(x^4).$$

Of course, we know a power series for 3^x , namely

$$3^x = e^{x \log 3} = 1 + x \log 3 + \frac{1}{2!} (x \log 3)^2 + \frac{1}{3!} (x \log 3)^3 + \dots,$$

so we might conjecture that $c_i = (\log 3)^i / i!$.

For $i = 0$ we indeed have $(\log 3)^0 / 0! = 1 = c_0$. But for $i = 1$ the conjecture seems to fail, because c_1 is not a real number but an “infinite integer”. (In any case, the bits of c_1 do not resemble the binary representation of the real number $\log 3 = 1.00011001001111101010 \dots$.)

Mystery 3. “ $\log 3$ ” is not the real number $\log 3$.

A final observation we make is that the direction of convergence is opposite that of real numbers. All the sequences we have seen approach their limits by filling in bits from low indices to high indices, which is toward the right in our graphical convention of reversing the digits of integers. A convergent sequence of real numbers, on the other hand, fills in bits from high indices to low indices. Take the sequence $(1 + 1/n)^n$, for example. Some terms of this sequence (as real numbers) are shown in Figure 6. The convergence proceeds from left to right, which is the same graphical direction but opposite numerical direction as the convergence of the sequence 3^{2^n} in Figure 3.

In the setting of bits of 3^{2^n} , then, the low indices of a number are somehow stronger than the higher indices. Therefore, we should really think of the “tail” of numbers as being backward from the normal sense: In this mode of convergence, two numbers are close to each other if their leftmost bits agree, that is, if their difference is divisible by a large power of 2. This is why we have chosen the convention that higher indices are to the right.

Mystery 4. Two numbers are close if their difference is highly divisible by 2.

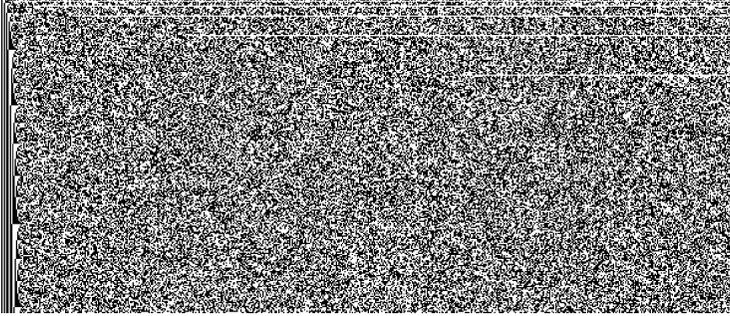


Figure 6. Binary representations of $(1 + 1/n)^n$ as real numbers, with most significant bits on the left. The terms are slowly converging to $e = 10.10110111111000010101 \dots_2$.

2. 2-adic Numbers

Our four mysteries suggest that there is a notion of number presenting itself through the binary representation of 3^{2^n} that is quite different from the real numbers. From Mystery 4 we must conclude that in some sense 2^i gets small as i gets large, and the other mysteries support this conclusion. Let us therefore make this a *definition* instead of a mystery and introduce a new notion of “size” to make this precise.

Every rational number $r \neq 0$ has a representation $r = 2^\alpha n/d$ for integers α , n , and d , where n and d are not divisible by 2. Moreover, α is unique. We want $|r|_2$ to be large when α is small, and small (but positive) when α is large. A natural choice is to let $|r|_2 = 2^{-\alpha}$; this is called the *2-adic norm* of r . For example, $|64|_2 = 1/64$ and $|-691/2730|_2 = 2$. Since 0 is very highly divisible by 2, let us define $|0|_2 = 0$.

Since large powers 2^i are small in the 2-adic norm, a rational number can be a sum of arbitrarily large powers of 2 when thought of 2-adically, just as it can be a sum of arbitrarily large powers of $1/2$ when thought of as a real number. For example,

$$\frac{1}{3} = 1 + \sum_{i=0}^{\infty} 2^{2i+1}.$$

In fact, every rational number has the representation $\sum_{i=N}^{\infty} c_i 2^i$ for some integer N and $c_i \in \{0, 1\}$. For example, -1 is rendered 2-adically as

$$-1 = \frac{1}{1-2} = \sum_{i=0}^{\infty} 2^i = 2^0 + 2^1 + 2^2 + 2^3 + \dots = \blacksquare \dots,$$

which illustrates that the 2-adic representation of a number is really a “limit” of its representations modulo 2^a as $a \rightarrow \infty$. For finite a , the representation modulo 2^a of course coincides with its two’s complement representation.

In general, a bits of the 2-adic representation of a rational number $r = 2^\alpha n / d$ can be found by computing the inverse $d^{-1} \pmod{2^a}$, which is an integer, and multiplying by $2^\alpha n$.

One can check that the 2-adic norm induces a metric $d(x, y) = |x - y|_2$ on the rational numbers, akin to the usual metric induced by the absolute value. In particular, it satisfies the triangle inequality $|x - y|_2 + |y - z|_2 \geq |x - z|_2$.

There are some strange properties of this metric, however. Perhaps the most immediate is that every triangle is isosceles: If $|x - y|_2 = |y - z|_2$, then the triangle is isosceles by definition. On the other hand, if $|x - y|_2 \neq |y - z|_2$, then $|x - z|_2 = \max(|x - y|_2, |y - z|_2)$. For example, if $x - y = 20$ and $y - z = 6$ then $|x - y|_2 = 1/4 \neq 1/2 = |y - z|_2$ and $|x - z|_2 = |26|_2 = 1/2$.

Naturally, we write $x_n \rightarrow x$ if the sequence of 2-adic norms $|x - x_n|_2$ approaches 0 as $n \rightarrow \infty$. So indeed $3^{2^n} \rightarrow 1$ in the 2-adic metric.

Of course, when we take a limit of rational numbers we may not get another rational number. Traditionally, the real numbers can be constructed by taking limits of rationals with respect to the real metric; each real number has an expansion $\sum_{i=N}^{\infty} c_i 2^{-i}$ for $c_i \in \{0, 1\}$. Similarly, we can take limits of rationals with respect to the 2-adic metric and get a different completion of the rationals. This completion is called the *set of 2-adic numbers*, and each 2-adic number has a representation $\sum_{i=N}^{\infty} c_i 2^i$, where again $c_i \in \{0, 1\}$. Like the real numbers, this set is complete, meaning that it contains all its limit points.

It turns out that our power series $3^x = \sum_{i=0}^{\infty} (\log 3)^i x^i / i!$ is correct, but it must be interpreted not as a real power series but as a 2-adic power series. This means that “ $\log 3$ ” is not the real number $\log 3$ but the 2-adic number $\log 3$. How do we compute it? The function $\log(1 - x)$ has a 2-adic power series that coincides with its real power series:

$$\log(1 - x) = - \sum_{i=1}^{\infty} \frac{x^i}{i}.$$

Of course, in the real metric this power series diverges at $x = -2$, so it cannot be used to compute the real $\log 3$. But 2-adically this series converges at $x = -2$ to the 2-adic $c_1 = \log 3$. Similarly, $c_2 = (\log 3)^2 / 2$, $c_3 = (\log 3)^3 / 3$, and so on.

To be precise, the standard objects of calculus over the 2-adic numbers such as derivatives, power series, tests for convergence, and so on, must be established. We do not undertake this task here but refer the reader to texts on the subject. Gouvêa [3] serves as a solid introduction, and Koblitz [4] provides a more advanced treatment.

3. Generalizations

The results of Section 2 can be generalized in several directions, and we discover that the power series structure we have seen is quite common.

Euler's theorem tells us that there is nothing particularly special about 3^{2^n} , and in fact the binary representations of 5^{2^n} and 7^{2^n} have exactly analogous structures, as shown in the first row of Figure 7. In general, if k is odd then k^x has a 2-adic power series $1 + x \log k + \dots$. Gouvêa [3, Section 4.5] discusses the region of convergence of such power series.

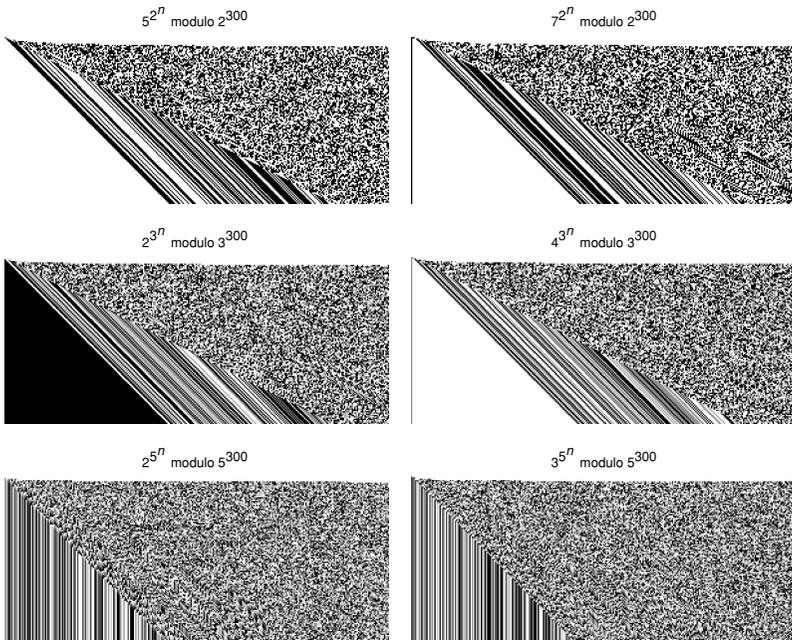


Figure 7. Powers k^{p^n} in base p .

What about other bases $b > 2$? The second and third rows of Figure 7 show several examples. To address these cases we briefly generalize the discussion to p -adic numbers for prime p .

Of course, we may define $|x|_b$ for general b (prime or composite) in the analogous way. For primes p , $|x|_p$ is a norm on the set of rational numbers. For composite b it is not since in general $|x \cdot y|_b \neq |x|_b \cdot |y|_b$; for example, $|4|_4 = 1/4 \neq 1 = |2|_4 \cdot |2|_4$. In fact, it is a theorem of Ostrowski that the p -adic norms and the real norm are (up to equivalence) the only nontrivial norms on the set of rational numbers.

Evidently $4^{3^n} \rightarrow 1$ in the 3-adic metric. However, the 3-adic limit of 2^{3^n} is not 1 but $2 + 2 \cdot 3^1 + 2 \cdot 3^2 + \dots = -1$. For general k relatively prime to p , Euler's theorem provides that $k^{(p-1)p^n} \equiv 1 \pmod{p^{n+1}}$. In the limit, then, k^{p^n} approaches a $(p-1)^{\text{th}}$ root of unity $1^{1/(p-1)}$. This root of unity is congruent to k modulo p and is called the *Teichmüller representative* of k . This accounts for the vertical stripes in the base-5 digits of 2^{5^n} and 3^{5^n} ; the 5-adic fourth roots of unity congruent to 2 and 3 modulo 5 are irrational. Note also that $2^{5^n} + 3^{5^n} \rightarrow 0$.

The p -adic power series of functions $f(x)$ other than k^x are also evident in the base- p digits of $f(p^n)$. Let F_n , C_n , M_n , and B_n be the sequences of Fibonacci, Catalan, Motzkin, and Bell numbers. The sequences C_{2^n} and M_{2^n} have 2-adic limits. The sequences F_{2^n} and B_{2^n} do not have 2-adic limits, but $F_{2^{2n}}$, $F_{2^{2n+1}}$, $B_{2^{2n}}$, and $B_{2^{2n+1}}$ do, giving some indication of the ubiquity of p -adic convergence in combinatorial sequences.

Finally, consider the factorial function $x!$. The terms of the sequence $2^n!$, of course, become highly divisible by 2, so $2^n! \rightarrow 0$ in the 2-adic norm. However, a theorem of Legendre implies that $|2^n!|_2 = 1/2^{2^n-1}$, and it turns out that $2^n!/2^{2^n-1}$ has a (nonzero) 2-adic limit.

References

- [1] S. Wolfram, *A New Kind of Science*, Champaign, IL: Wolfram Media, Inc., 2002.
- [2] E. Rowland, "Local Nested Structure in Rule 30," *Complex Systems*, 16(3), 2006 pp. 239-258.
- [3] F. Gouvêa, *p-adic Numbers: An Introduction*, 2nd ed., New York: Springer-Verlag, 1997.
- [4] N. Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, 2nd ed., New York: Springer-Verlag, 1984.