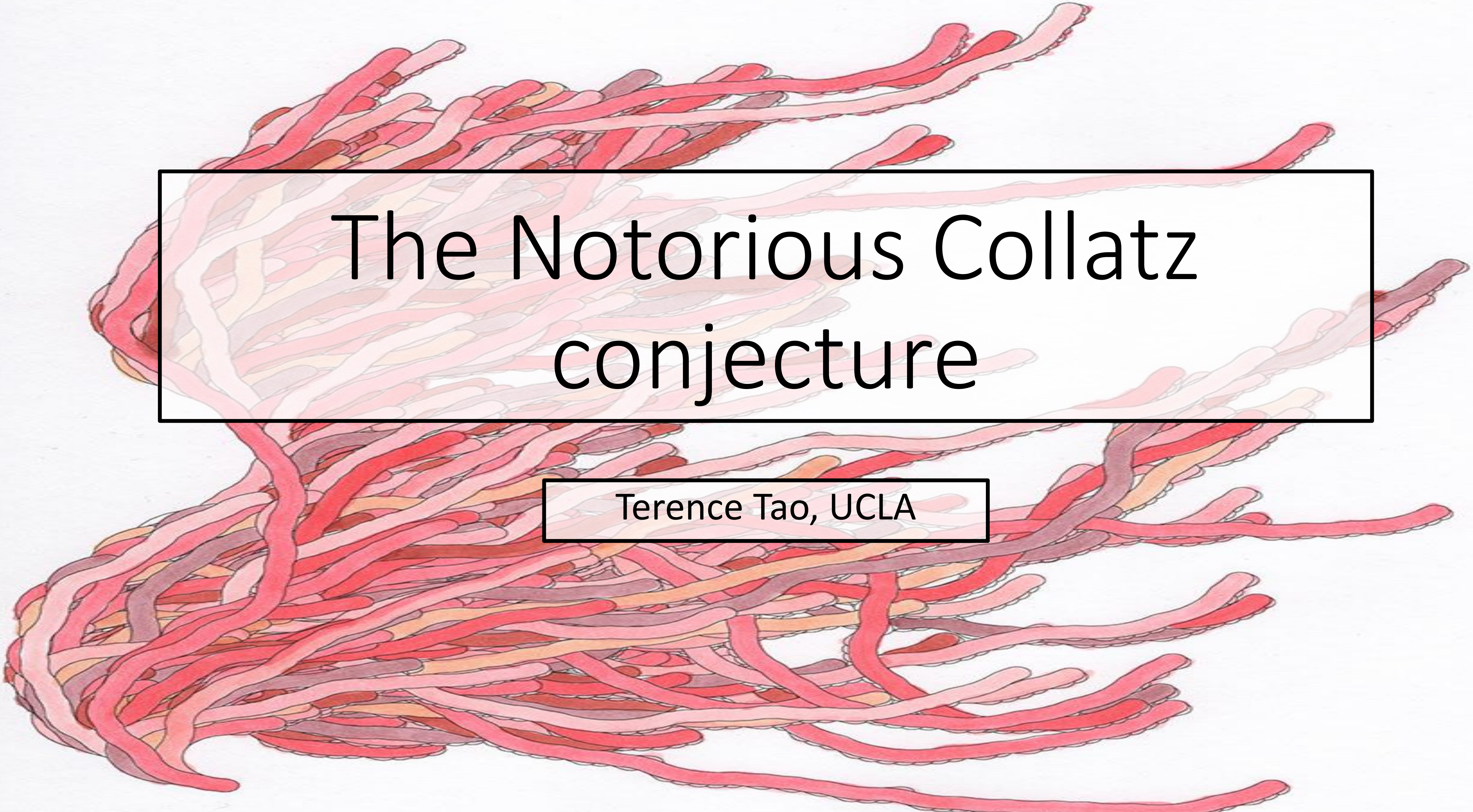
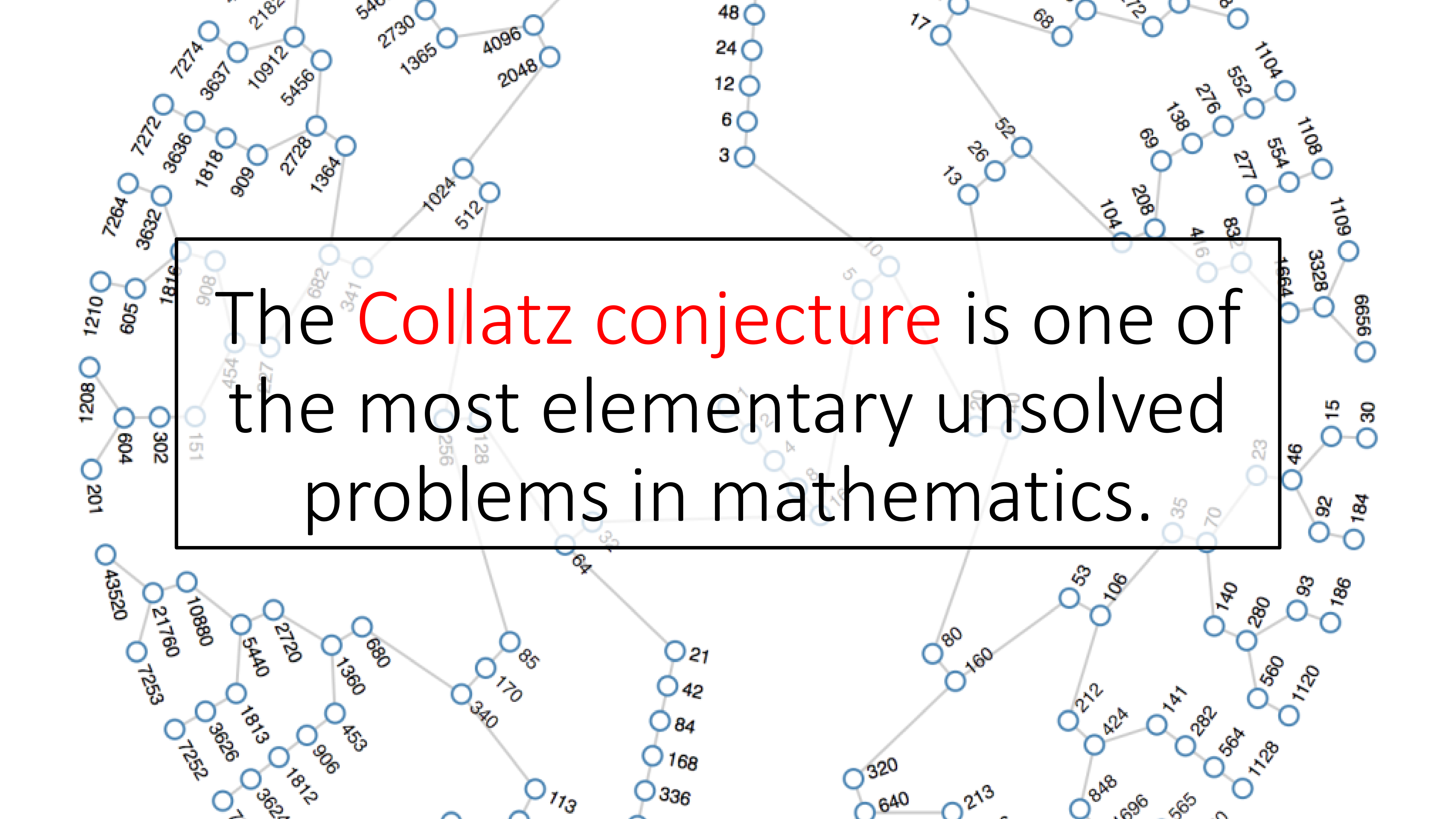
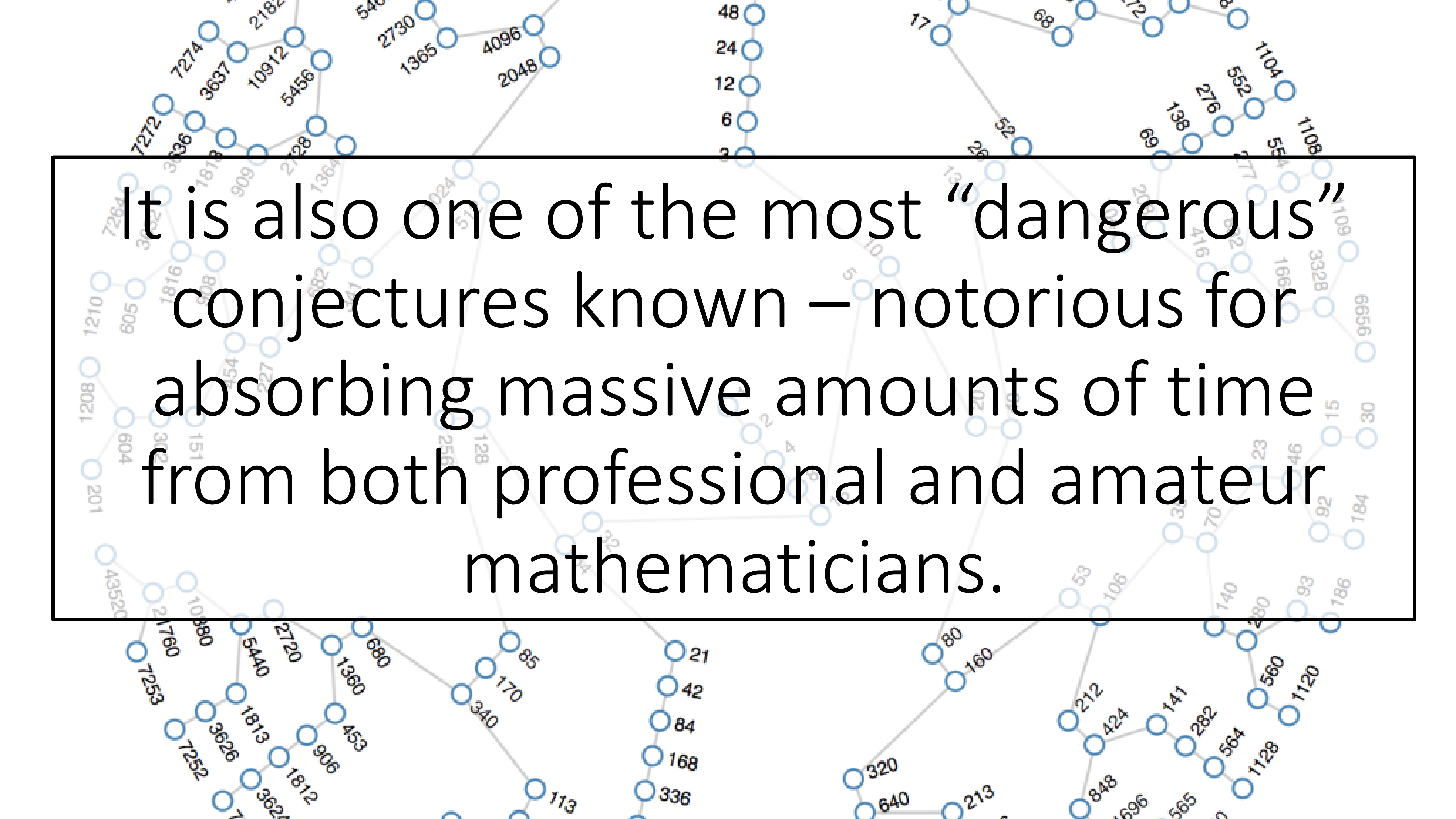


The Notorious Collatz conjecture

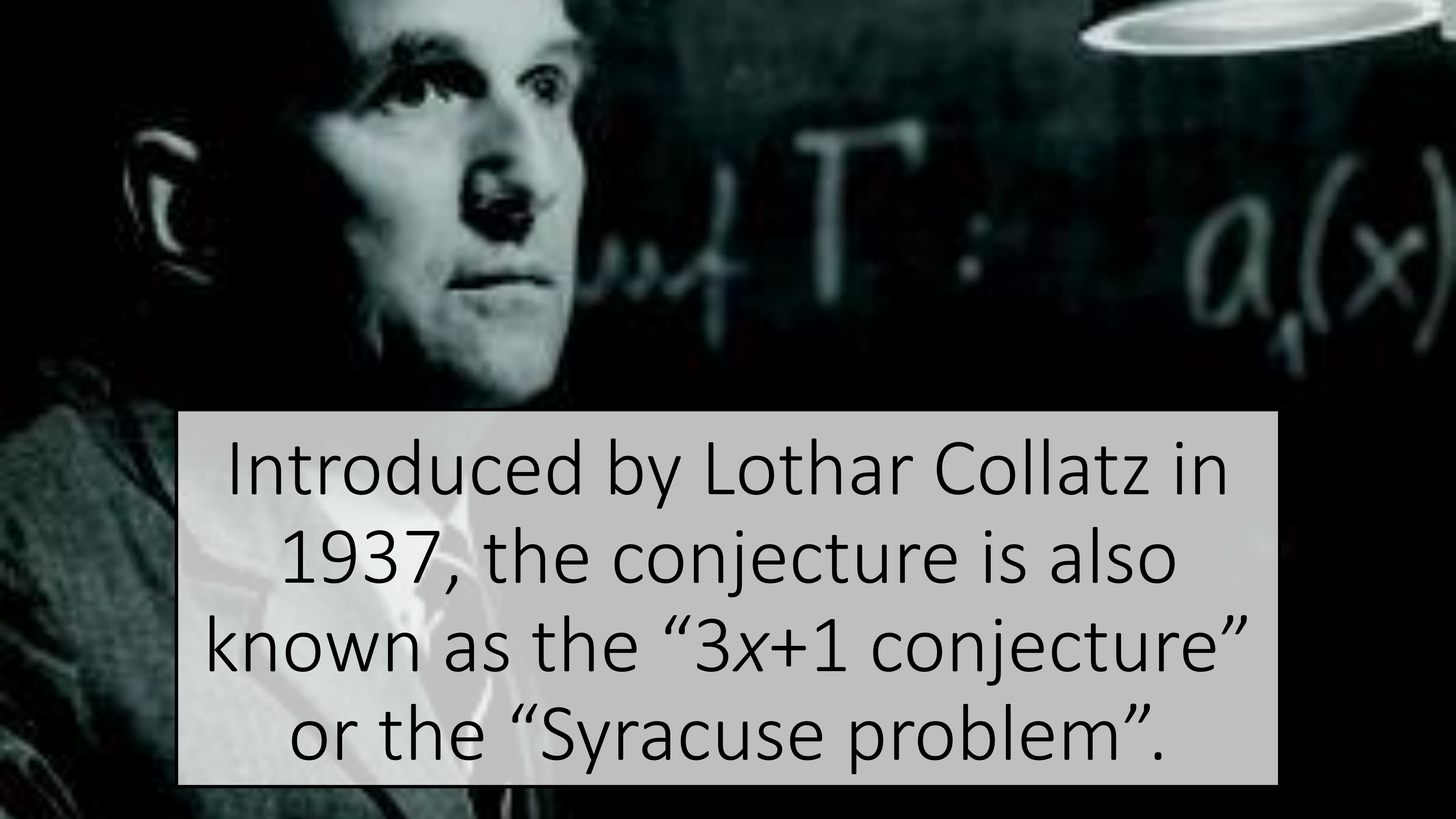
Terence Tao, UCLA



The Collatz conjecture is one of the most elementary unsolved problems in mathematics.



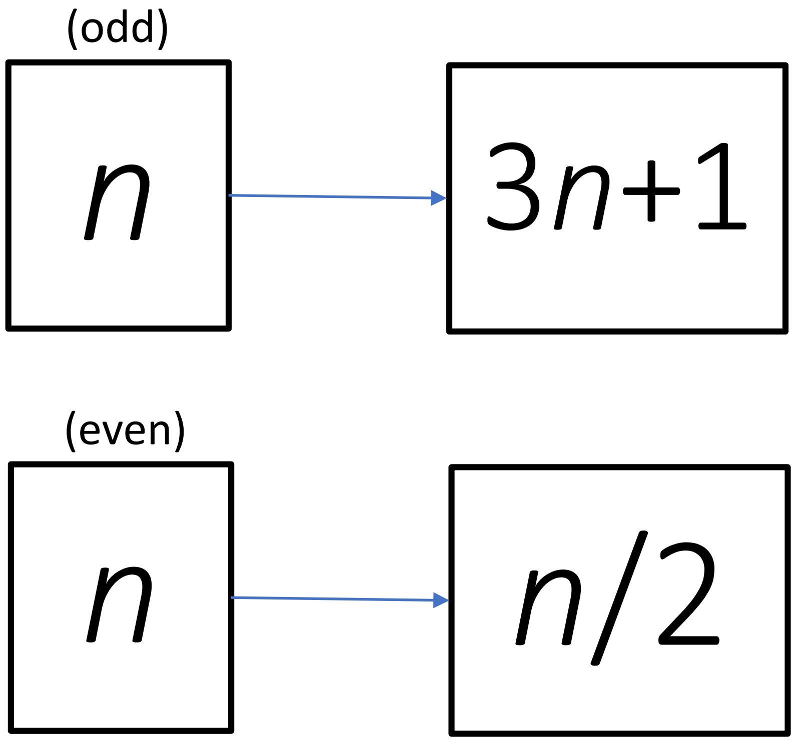
It is also one of the most “dangerous” conjectures known – notorious for absorbing massive amounts of time from both professional and amateur mathematicians.



Introduced by Lothar Collatz in 1937, the conjecture is also known as the “3*x*+1 conjecture” or the “Syracuse problem”.

The conjecture involves an innocuous function `Col` on the natural numbers {1,2,3,...} defined by the following rule:

- `Col`(n) equals $3n+1$ if n is odd.
- `Col`(n) equals $n/2$ if n is even.



n	1	2	3	4	5	6	7	8	9	10
Col(n)	4	1	10	2	16	3	22	4	28	5

n	11	12	13	14	15	16	17	18	19	20
Col(n)	34	6	40	7	46	8	52	9	58	10

n	21	22	23	24	25	26	27	28	29	30
Col(n)	64	11	70	12	76	13	82	14	88	15

Now consider iterates of the Collatz function `Col`, in which the output of the function is fed back into the input:

$$\texttt{Col}^2(n) = \texttt{Col}(\texttt{Col}(n))$$
$$\texttt{Col}^3(n) = \texttt{Col}(\texttt{Col}(\texttt{Col}(n)))$$

etc.

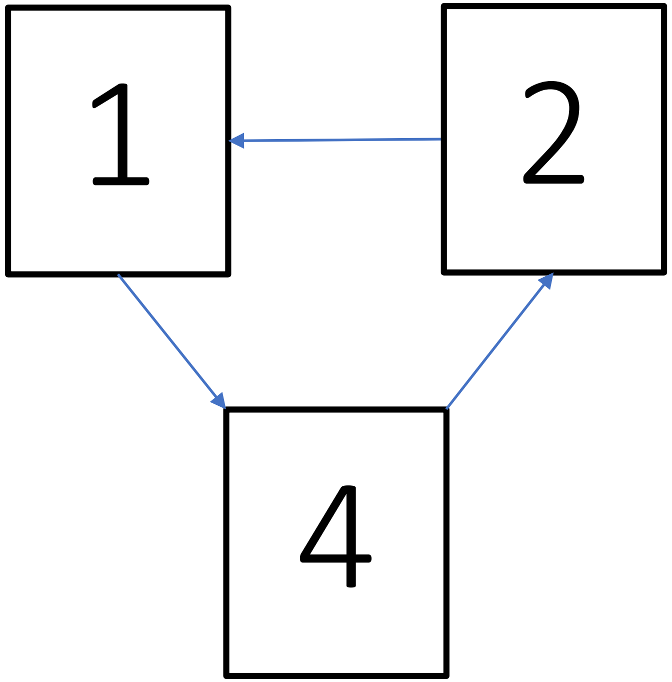
n	1	2	3	4	5	6	7	8	9	10
Col(n)	4	1	10	2	16	3	22	4	28	5
$Col^2(n)$	2	4	5	1	8	10	11	2	14	16
$Col^3(n)$	1	2	16	4	4	5	34	1	7	8
$Col^4(n)$	4	1	8	2	2	16	17	4	22	4
$Col^5(n)$	2	4	4	1	1	8	52	2	11	2
$Col^6(n)$	1	2	2	4	4	4	26	1	34	1
$Col^7(n)$	4	1	1	2	2	2	13	4	17	4

Every natural number n generates a Collatz sequence (or Collatz orbit)

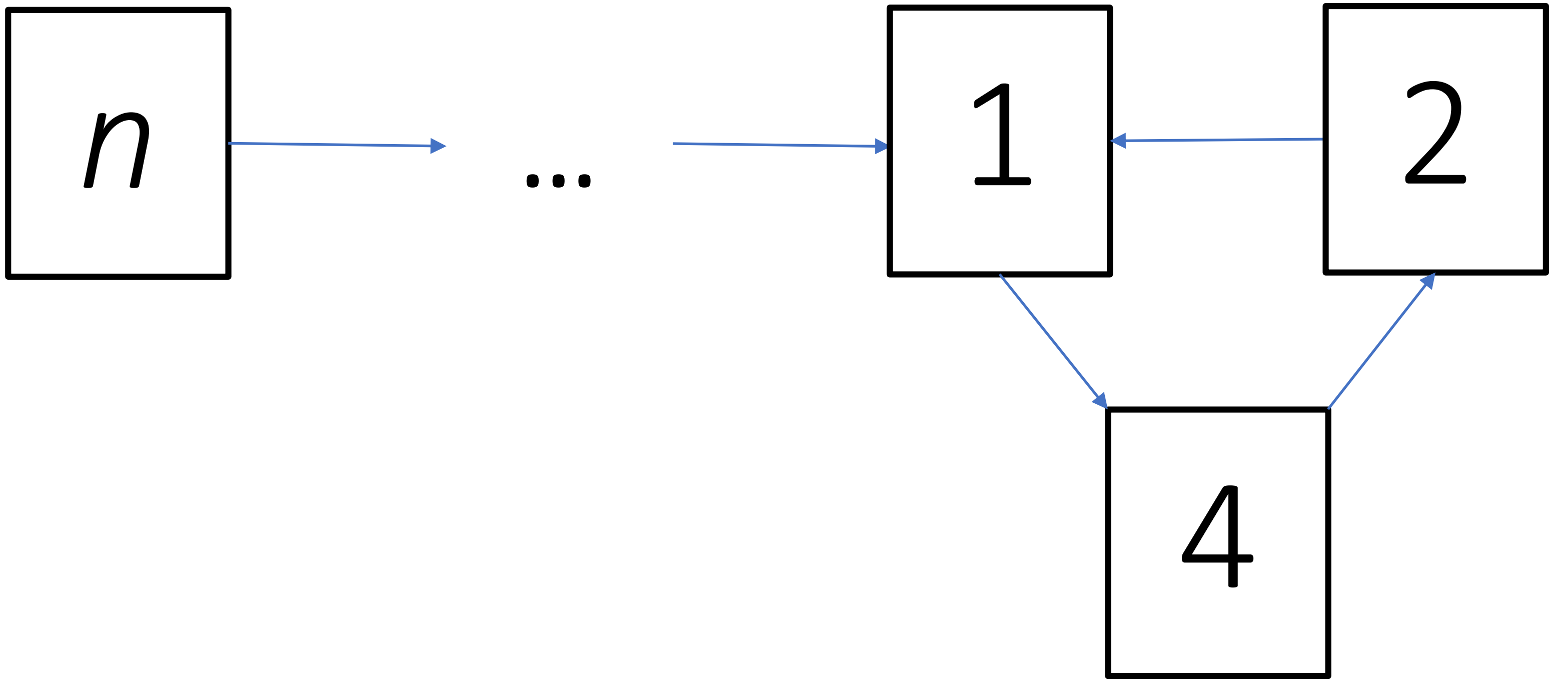
$n, \mathtt{Col}(n), \mathtt{Col}^2(n), \mathtt{Col}^3(n), \ldots$

$$\boxed{n} \to \boxed{\mathtt{Col}(n)} \to \boxed{\mathtt{Col}^2(n)} \to \ldots$$

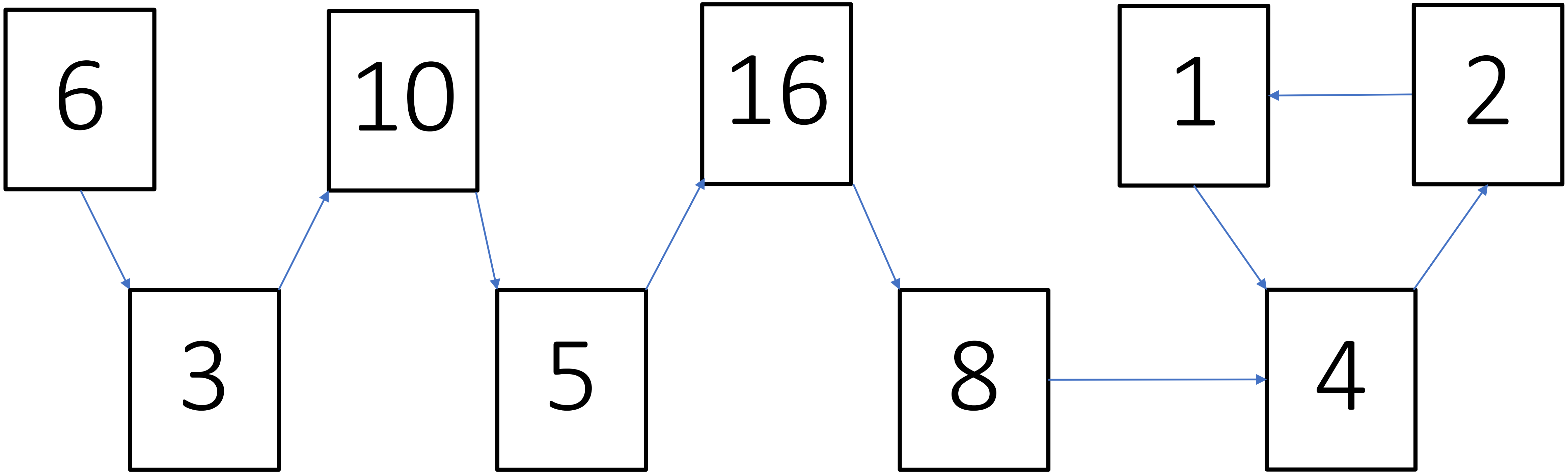
For instance, n=1 generates the periodic Collatz sequence
1, 4, 2, 1, 4, 2, 1, 4, 2, 1,…



If a Collatz sequence reaches the value 1, it will then cycle through the values 1, 4, 2 indefinitely.

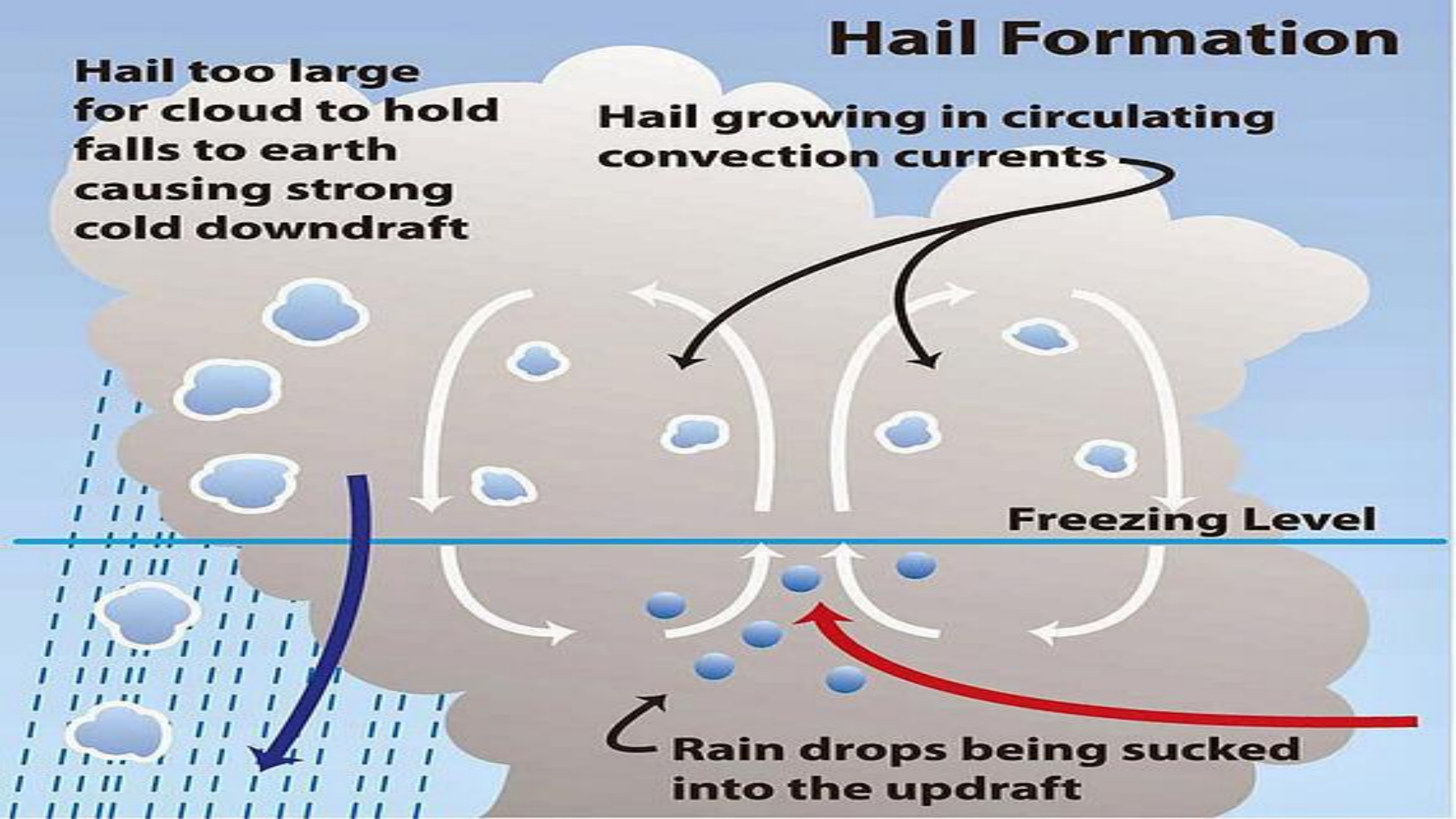


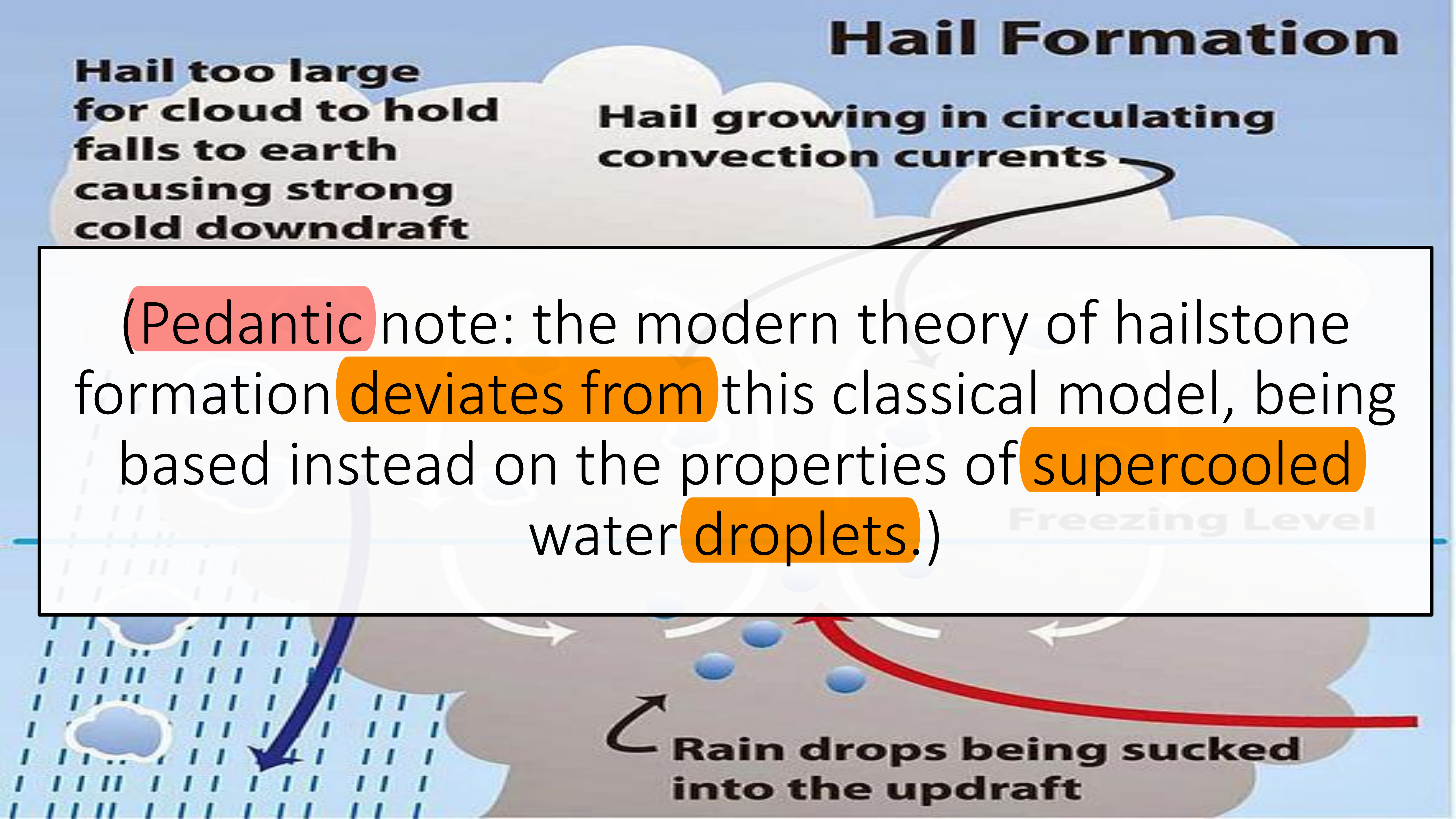
For instance, n=6 generates the Collatz sequence 6, 3, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, 4, 2, 1, 4, 2, 1,…



Collatz sequences are also known as hailstone sequences, as they can bounce up and down much like hailstones in a cloud were thought to.

For instance, n=27 generates the Collatz sequence
27, 82, 41, 124, 62, 31, 94, 47, 142, 71, 214, 107, 322, 161, 484, 242, 121, 364, 182, 91, 274, 137, 412, 206, 103, 310, 155, 466, 233, 700, 350, 175, 526, 263, 790, 395, 1186, 593, 1780, 890, 445, 1336, 668, 334, 167, 502, 251, 754, 377, 1132, 566, 283, 850, 425, 1276, 638, 319, 958, 479, 1438, 719, 2158, 1079, 3238, 1619, 4858, 2429, 7288, 3644, 1822, 911, 2734, 1367, 4102, 2051, 6154, 3077, 9232, 4616, 2308, 1154, 577, 1732, 866, 433, 1300, 650, 325, 976, 488, 244, 122, 61, 184, 92, 46, 23, 70, 35, 106, 53, 160, 80, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, …


Hail Formation
Hail too large for cloud to hold falls to earth causing strong cold downdraft
Hail growing in circulating convection currents
Freezing Level
Rain drops being sucked into the updraft



(Pedantic note: the modern theory of hailstone formation deviates from this classical model, being based instead on the properties of supercooled water droplets.)

But just as every hailstone eventually falls to the ground, we have the infamous

Collatz conjecture. Every Collatz sequence eventually attains the value 1.

The 3x+1 problem: An annotated bibliography (1963--1999) (sorted by author)

Jeffrey C. Lagarias

(Submitted on 13 Sep 2003 (v1), last revised 11 Jan 2011 (this version, v13))

The 3x+ 1 problem concerns iteration of the map on the integers given by T(n) = (3n+1)/2 if n is odd; T(n) = n/2 if n is even. The 3x+1 Conjecture asserts that for every positive integer n > 1 the forward orbit of n under iteration by T includes the integer 1. This paper is an annotated bibliography of work done on the 3x+1 problem and related problems from 1963 through 1999. At present the 3x+1 Conjecture remains unsolved.

Comments: 74 pages latex, 197 references, second title change to distinguish from 3x+1 book; first title change indicates abridgment of earlier versions to papers 1999 and earlier ; part II now covers papers 2000 and later, see arXiv:math.NT/0608208; v.11 cutoff date changed from 2000 to 1999, v.13 added Oulipo references

Despite hundreds of published papers on the Collatz conjecture, and many more unpublished works (including countless failed proofs), the conjecture remains unsolved today.

The 3x+1 Problem: An Annotated Bibliography, II (2000-2009)

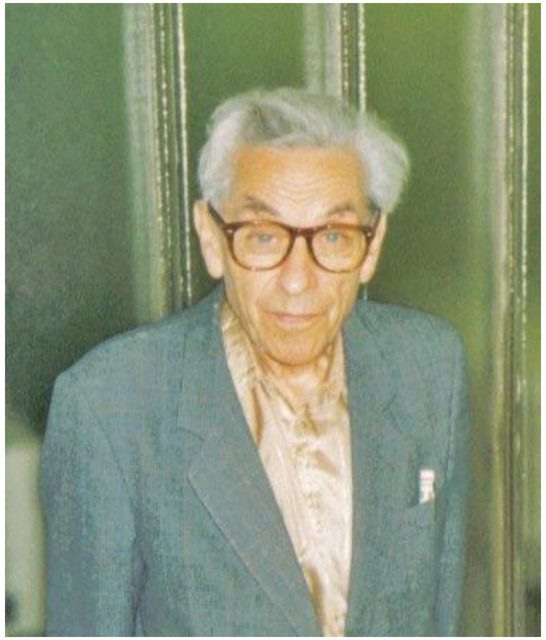
Jeffrey C. Lagarias

(Submitted on 9 Aug 2006 (v1), last revised 12 Feb 2012 (this version, v6))

The 3x+1 problem concerns iteration of the map T(n) =(3n+1)/2 if n odd; n/2 if n even. The 3x +1 Conjecture asserts that for every positive integer n>1 the forward orbit of n includes the integer 1. This paper is an annotated bibliography of work done on the 3x+1 problem published from 2000 through 2009, plus some later papers that were preprints by 2009. This is a sequel to an annotated bibliography on the 3x+1 problem covering 1963-1999.
At present the 3x+1 Conjecture remains unsolved.

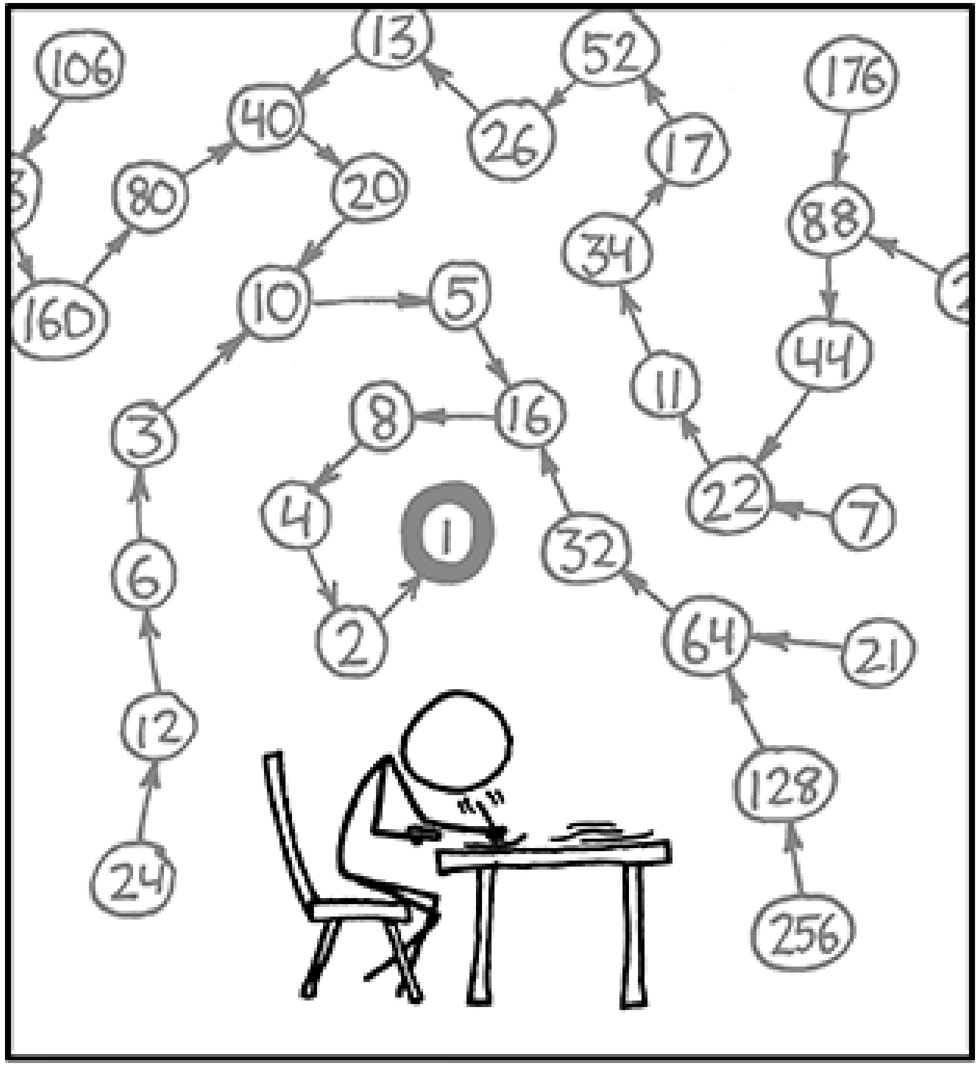
Comments: 42 pages; will be periodically updated, cover papers till 2009+ later papers preprints by 2009; sequel to arXiv:math.NT/0309224; v4 now runs from 2000 rather than 2001, 121 entries; v5 40 pages, 126 entries, v6 42 pages, 134 enries

“For about a month everyone at Yale worked on it, with no result. A similar phenomenon happened when I mentioned it at the University of Chicago. A joke was made that this problem was part of a conspiracy to slow down mathematical research in the U.S.” – Shizuo Kakutani, 1960



“Mathematics is not yet ripe enough for such questions.” – Paul Erdős, 1983

“This is an extraordinarily difficult problem, completely out of reach of present day mathematics.” – Jeff Lagarias, 2010



THE COLLATZ CONJECTURE STATES THAT IF YOU PICK A NUMBER, AND IF IT'S EVEN DIVIDE IT BY TWO AND IF IT'S ODD MULTIPLY IT BY THREE AND ADD ONE, AND YOU REPEAT THIS PROCEDURE LONG ENOUGH, EVENTUALLY YOUR FRIENDS WILL STOP CALLING TO SEE IF YOU WANT TO HANG OUT.

XKCD, Randall Monroe,
March 5, 2010

The Collatz conjecture appears to be a mere mathematical curiosity, with no obvious real-world applications. Why should we try to solve it?

- Pure intellectual challenge
- A benchmark for testing our understanding of number theory
- Proof attempts have linked the problem to other areas of mathematics
- It is a simple, but non-trivial, toy model of a dynamical system
- Modest cash prizes ($50, Harold Coxeter; $500, Paul Erdős; £1000, Sir Bryan Thwaites)
- Bragging rights

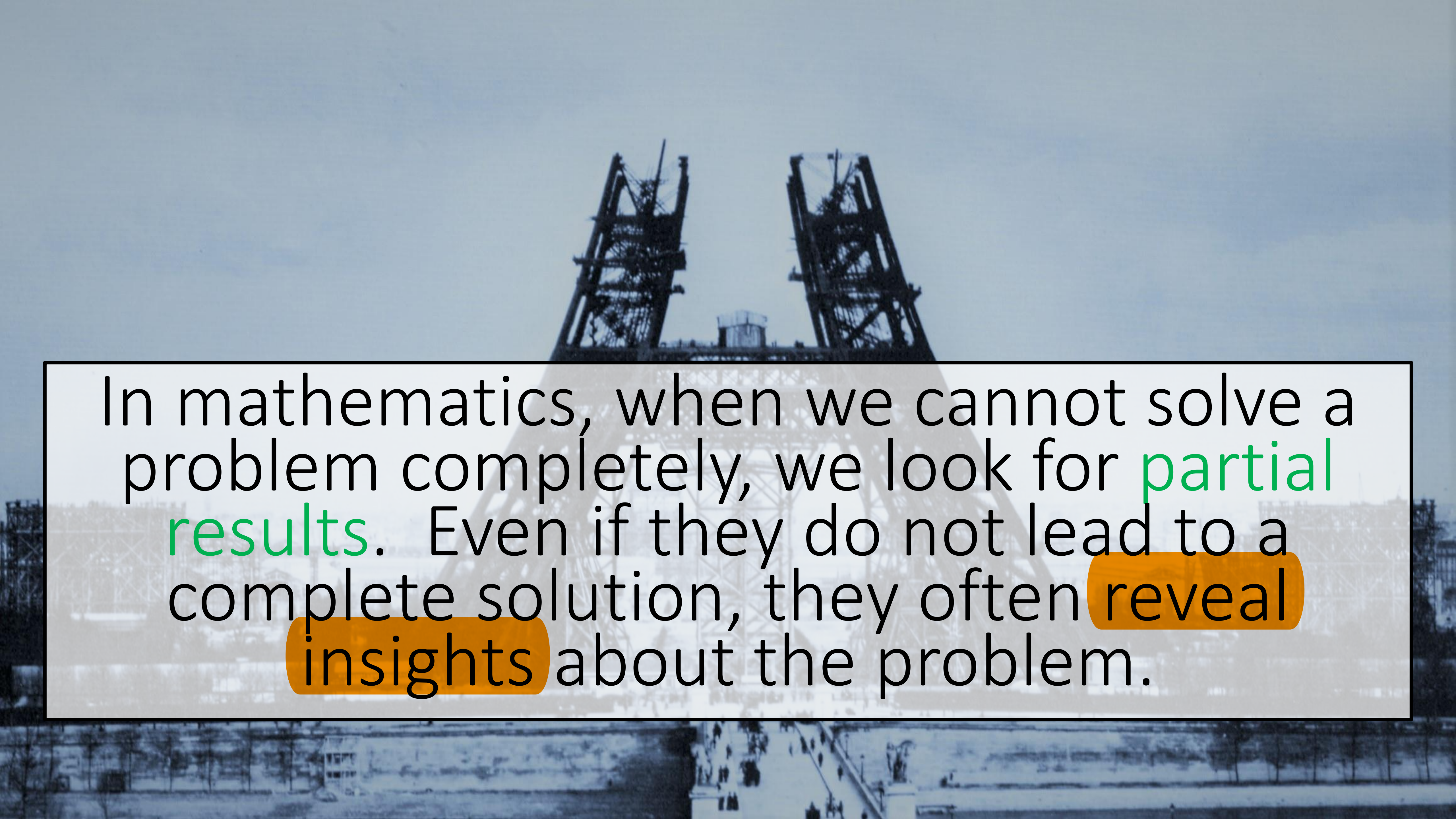
Mathematically speaking, a (discrete) dynamical system is a state space X, together with a shift map T from X to itself. The iterates $T, T^2, T^3, \ldots$ describe the dynamics of the system.

In the Collatz dynamical system, the state space is the natural numbers $\mathbf{N} = \{1,2,3,\ldots\}$ and the shift map is the Collatz map `Col`.

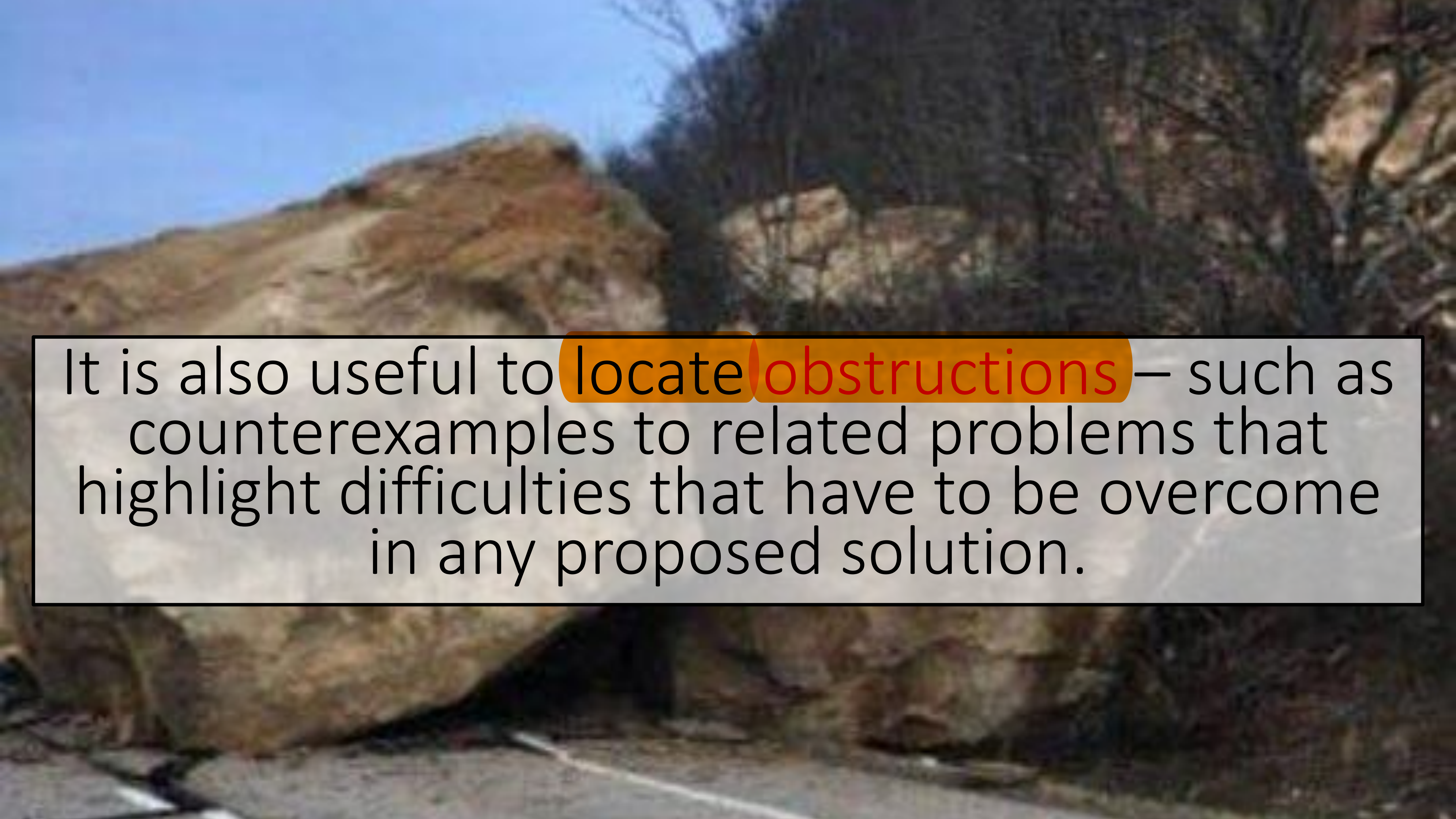
A sibling to the discrete dynamical systems are the continuous dynamical systems, where the dynamics are given by ordinary differential equations (ODE) or partial differential equations (PDE).

Many important real-world systems, such as fluids, ecosystems, and the climate, can be viewed as (continuous) dynamical systems.

The Collatz conjecture highlights the basic fact that even very simple equations can lead to amazingly complicated dynamics.



In mathematics, when we cannot solve a problem completely, we look for partial results. Even if they do not lead to a complete solution, they often reveal insights about the problem.



It is also useful to locate obstructions – such as counterexamples to related problems that highlight difficulties that have to be overcome in any proposed solution.

What partial results and obstructions do we have for the Collatz conjecture?

Partial result: in 2017, a distributed computing project verified the Collatz conjecture for all starting values n up to 10^{20}. So it is highly unlikely that a counterexample can be found just from pen and paper search.

One way the Collatz conjecture could fail is if there is a cycle – a Collatz sequence that repeats itself indefinitely – other than the known cycle 1, 2, 4, 1, 2, 4,… (or its shifts).

Partial result: it is known that any such cycle must have length at least 17,087,915. (Eliahou, 1993). So one cannot simply produce a short cycle to easily disprove the conjecture!

Obstruction: On the other hand, there are variants of the Collatz conjecture that have non-trivial cycles. For instance, if one modifies `Col` by sending an odd number n to $3n$-1 rather than $3n$+1, then two additional cycles appear:

- 5, 14, 7, 20, 10, 5,…
- 17, 50, 25, 74, 37, 110, 55, 164, 82, 41, 122, 61, 192, 91, 272, 136, 68, 34, 17,…

We don't know if there are any further cycles for this map.

This obstruction shows that any proof of the Collatz conjecture must at some point use a property of the $3n+1$ map that is not shared by the $3n-1$ map.

Obstruction: the absence of non-trivial Collatz cycles can be shown to imply a difficult result in number theory:

Theorem: The gap between powers of 2 and powers of 3 goes to infinity.

$3^2-2^3 = 9-8 = 1$; $2^5-3^3 = 32-27 = 5$; $2^8-3^5 = 256-243 = 13$; $3^7-2^{11} = 2187-2048 = 139$; …

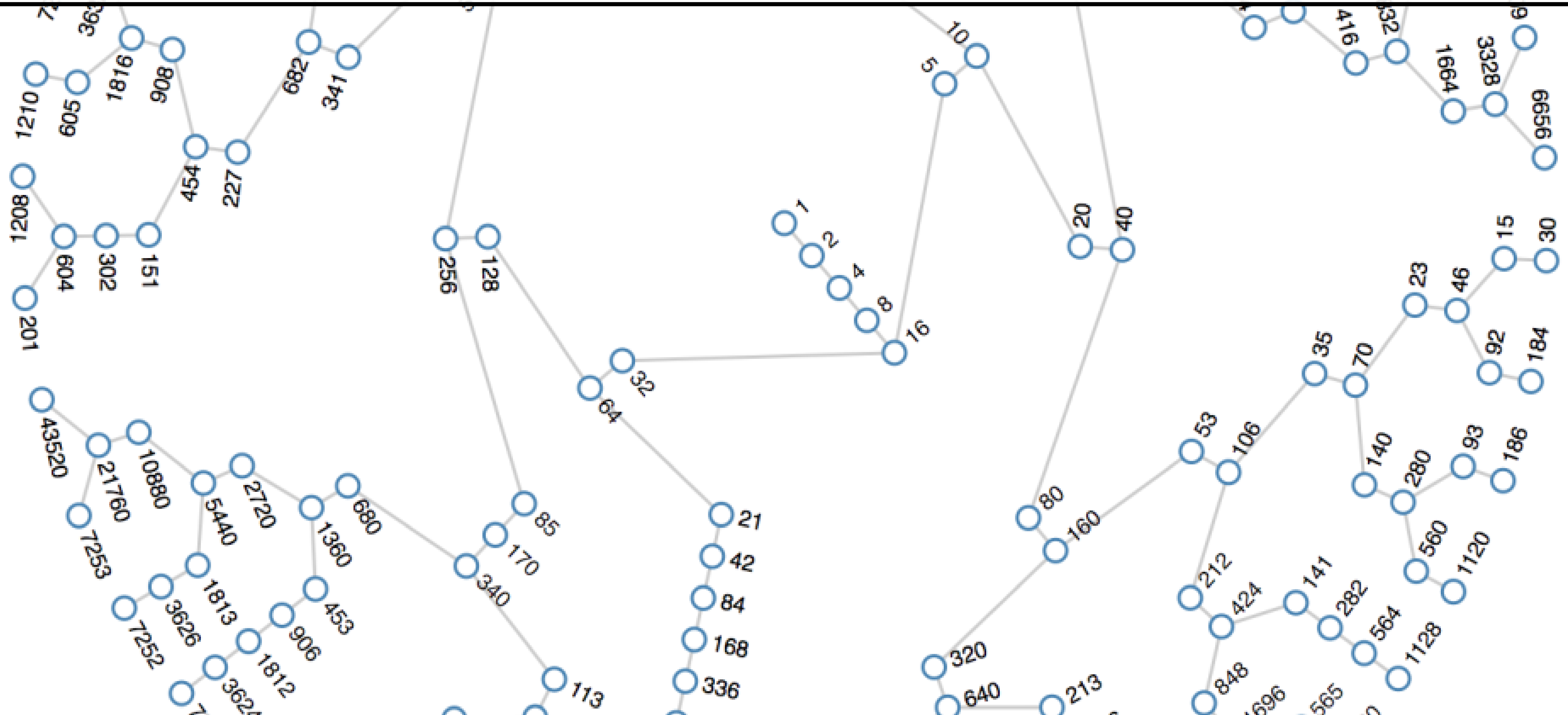
Basically, if a power of 2 and power of 3 are too close together, they can be used to create a Collatz cycle.

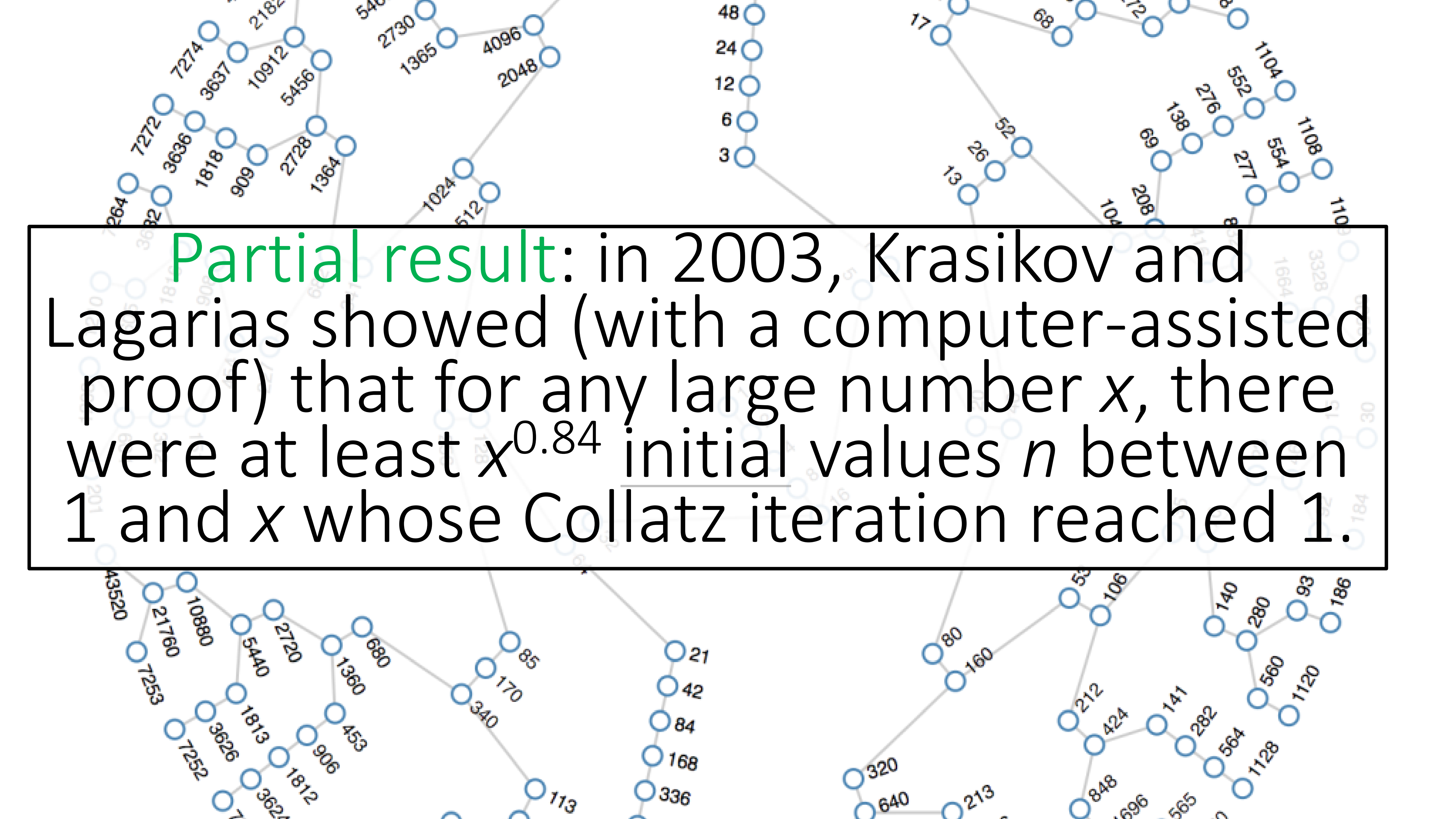
Theorem: The gap between powers of 2 and powers of 3 goes to infinity.

This theorem is known to be true, but its proof is difficult, requiring a deep result known as Baker's theorem (which earned Alan Baker the Fields medal in 1970).

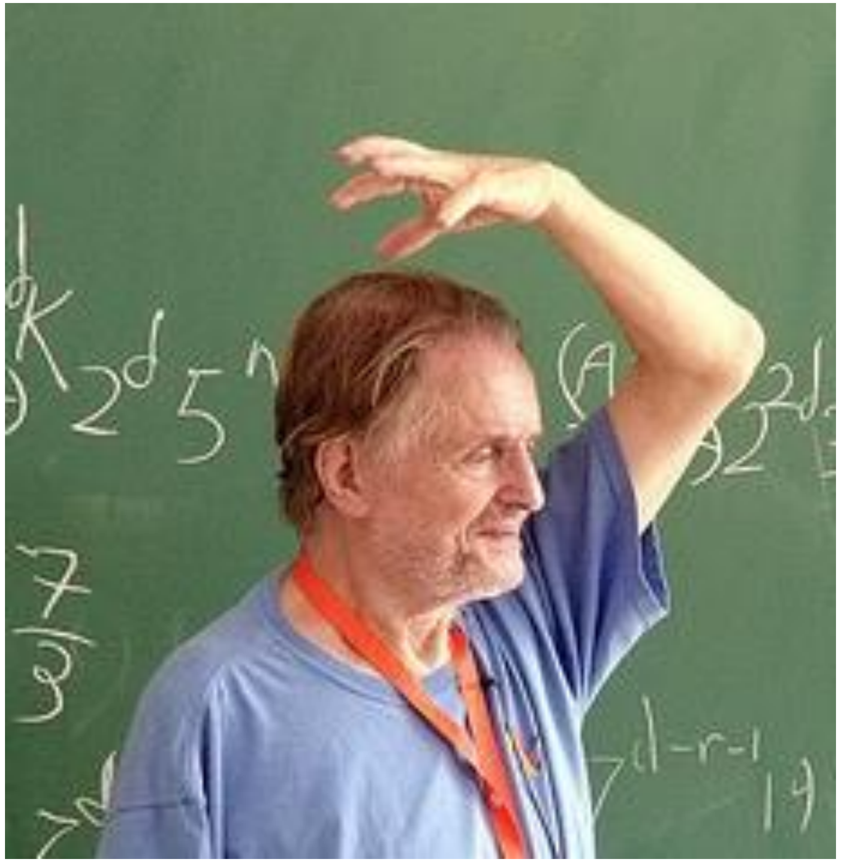
So solving the Collatz conjecture may be at least as hard as proving Baker's theorem!

One can try to work backwards and show that lots and lots of numbers get sent to 1 by the Collatz iteration.





Partial result: in 2003, Krasikov and Lagarias showed (with a computer-assisted proof) that for any large number x, there were at least $x^{0.84}$ initial values n between 1 and x whose Collatz iteration reached 1.



In 1987, John H. Conway invented a computer language called `FRACTRAN`, in which every program was a variant of the Collatz function `Col`. The output of sequences could be used to perform mathematical computations!

For instance, the `FRACTRAN` program `Prime` maps any natural number *n* to the number `Prime`(*n*), defined to equal

- 17n/91 if n is divisible by 91; else
- 78n/85 if n is divisible by 85; else
- 19n/51 if n is divisible by 51; else
- 23n/38 if n is divisible by 38; else
- 29n/33 if n is divisible by 33; else
- 77n/29 if n is divisible by 29; else
- 95n/23 if n is divisible by 23; else
- 77n/19 if n is divisible by 19; else
- n/17 if n is divisible by 17; else
- 11n/13 if n is divisible by 13; else
- 13n/11 if n is divisible by 11; else
- 15n/2 if n is divisible by 2; else
- n/7 if n is divisible by 7; else
- 55n.

Remarkable fact: the `Prime` orbit

$$2,\ \texttt{Prime}(2),\ \texttt{Prime}^2(2),\ \texttt{Prime}^3(2),\ \ldots$$

contains precisely the powers 2^p of 2 whose exponents are primes (together with many non-powers of two). This `FRACTRAN` program computes primes!

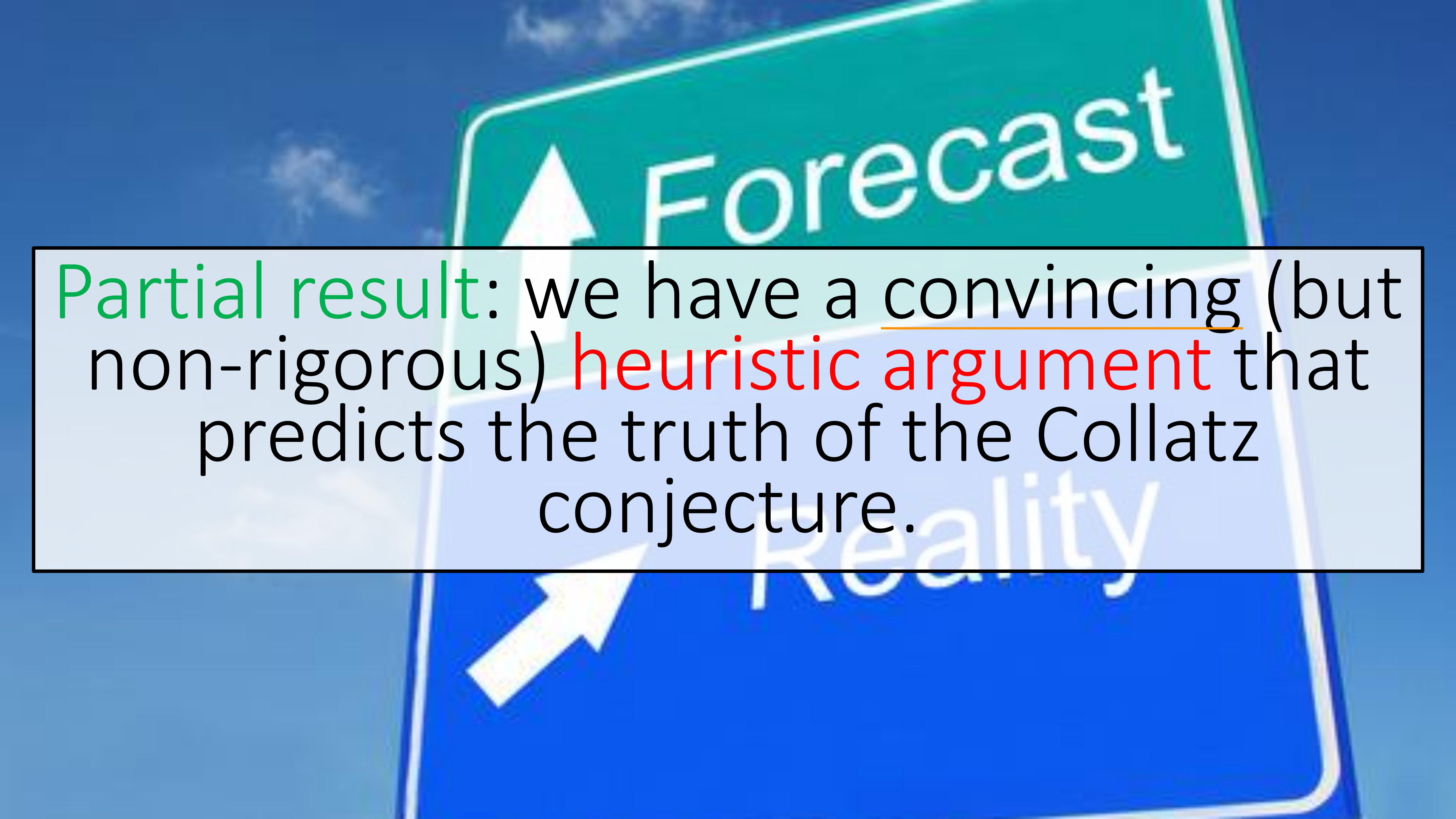
In fact, `FRACTRAN` is Turing Complete. Roughly speaking, this means that *any* computation that can be performed by an ordinary computer, can also be computed by a FRACTRAN program!

Obstruction: There are `FRACTRAN` program sequences for which it is **undecidable** whether they will ever reach a certain target value n_0.

This is related to the undecidability of the **halting problem** for Turing Machines.

This obstruction demonstrates that there is NO general algorithm that can definitively resolve all questions resembling the Collatz conjecture.

Any solution to that conjecture must use special properties of the Collatz map `Col` that are not shared by general `FRACTRAN` programs.



Partial result: we have a convincing (but non-rigorous) heuristic argument that predicts the truth of the Collatz conjecture.

The argument proceeds like this. The Collatz map `Col` can take an odd number n to a larger number $3n+1$. But this new number $3n+1$ is necessarily even, so the next application of `Col` will divide it by 2.

$$\boxed{n} \rightarrow \boxed{3n+1} \rightarrow \boxed{(3n+1)/2}$$

Heuristically, there is a fifty-fifty chance that the number $(3n+1)/2$ will also be even, leading to further divisions by 2. Indeed, a probability theory calculation reveals that the "expected number" of divisions by 2 one experiences before reaching an odd number again is equal to two.

$$\boxed{n} \rightarrow \boxed{3n+1} \rightarrow \boxed{(3n+1)/2} \rightarrow \boxed{(3n+1)/4}$$

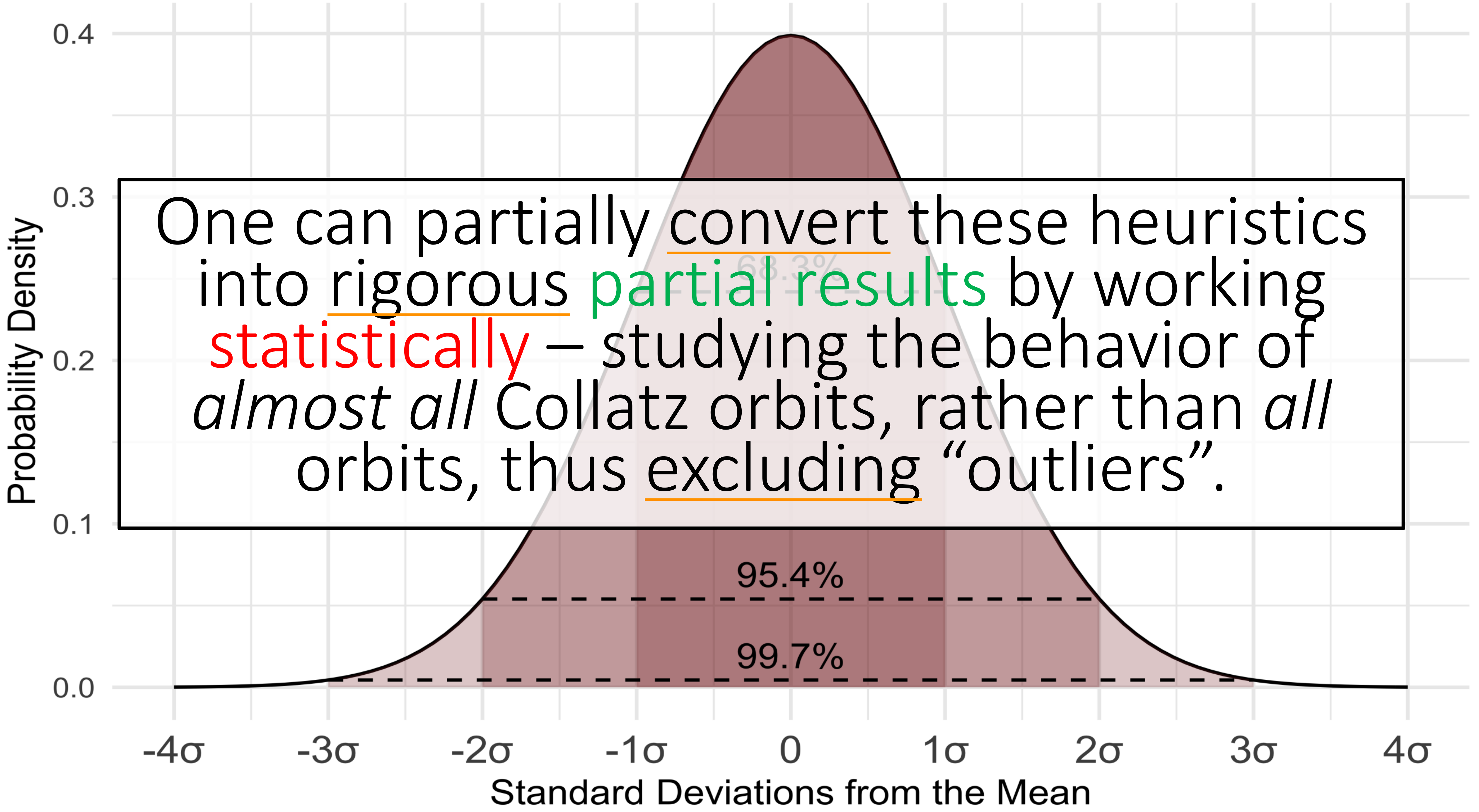
As a consequence, if one starts with an odd number *n*, the next odd number in the Collatz sequence would be expected to equal approximately $3n/4$ on the average. Thus the average size of the odd numbers in the sequence will decrease towards 1, which supports the validity of the Collatz conjecture.

$$\boxed{n} \rightarrow \boxed{3n+1} \rightarrow \boxed{(3n+1)/2} \rightarrow \boxed{(3n+1)/4}$$

This heuristic also predicts that some variants of the Collatz map, such as the 5n+1 map, will have orbits that go to infinity. This appears to be supported by numerics.

$$n \to 5n+1 \to (5n+1)/2 \to (5n+1)/4$$

7, 36, 18, 9, 46, 23, 116, 58, 29, 146, 73, 366, 183, 916, 458, 229, 1146, 573, 2866, 1433, 7166, 3583, 17916, …



One can partially convert these heuristics into rigorous partial results by working statistically – studying the behavior of *almost all* Collatz orbits, rather than *all* orbits, thus excluding "outliers".

Partial result: in 1976, Terras showed that almost all initial values *n* eventually iterated to a value less than *n*. (As a first approximation, think of “almost all” as meaning “at least 99.99% of all”.)

If one could show that all initial values *n* (other than 1) iterated to something less than *n*, this would imply the Collatz conjecture by further iteration.

(all except 1?)

$$\boxed{n} \longrightarrow \quad \ldots \quad \longrightarrow \boxed{<n}$$

Partial result: Terras's result was refined over the years. In 1979, Allouche showed that almost all initial values n eventually iterated to a value less than $n^{0.869}$.

(almost all)

$$\boxed{n} \longrightarrow \cdots \longrightarrow \boxed{< n^{0.869}}$$

Partial result: in 1994, Korec lowered this bound further to $n^{0.7925}$.

(almost all)

$$\boxed{n} \longrightarrow \quad \ldots \quad \longrightarrow \boxed{< n^{0.7925}}$$

Partial result: In 2019, I showed that almost all initial values n eventually iterated to a value less than $f(n)$, for any function $f()$ that grew to infinity, no matter how slowly. "Almost all Collatz orbits attain almost bounded values."

(almost all)

$$\boxed{n} \longrightarrow \quad \cdots \quad \longrightarrow \boxed{< f(n)}$$

For instance: almost all initial values n eventually iterate to a value less than $\log(\log(\log(\log(n))))$.

(almost all)

$$\boxed{n} \longrightarrow \quad \ldots \quad \longrightarrow \boxed{< \log(\log(\log(\log(n)))).}$$

This is about as close as one can get to the Collatz conjecture without actually solving it.

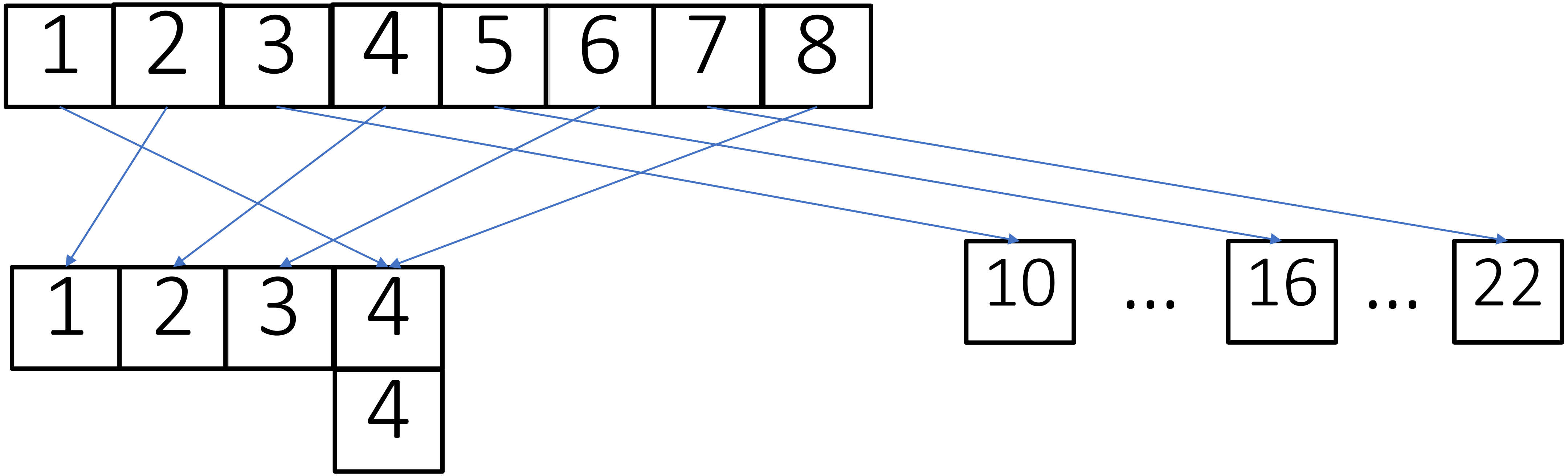
(all?)

$$\boxed{n} \rightarrow \cdots \rightarrow \boxed{1}$$

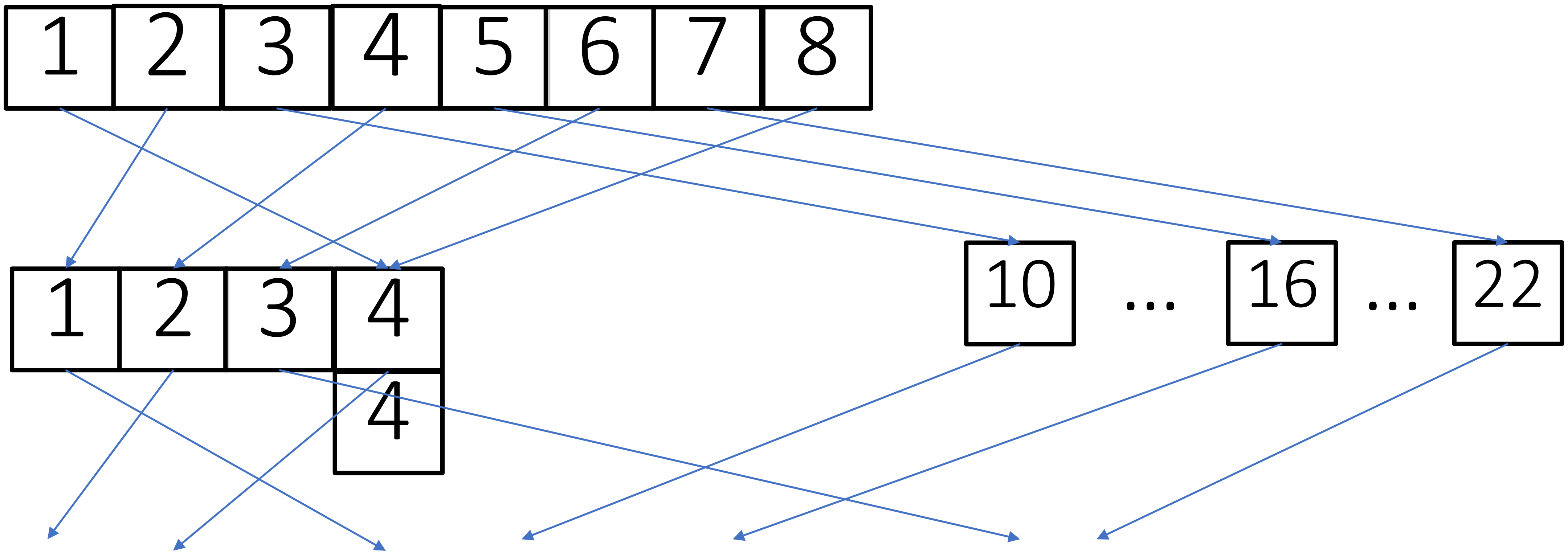
Unfortunately, the statistical methods used in the proof seem to be unable to fully resolve the conjecture, which remains out of reach for now.

The argument was inspired by other dynamical systems results, and in particular by a 1994 result of Bourgain on constructing an invariant measure for the nonlinear Schrödinger equation.

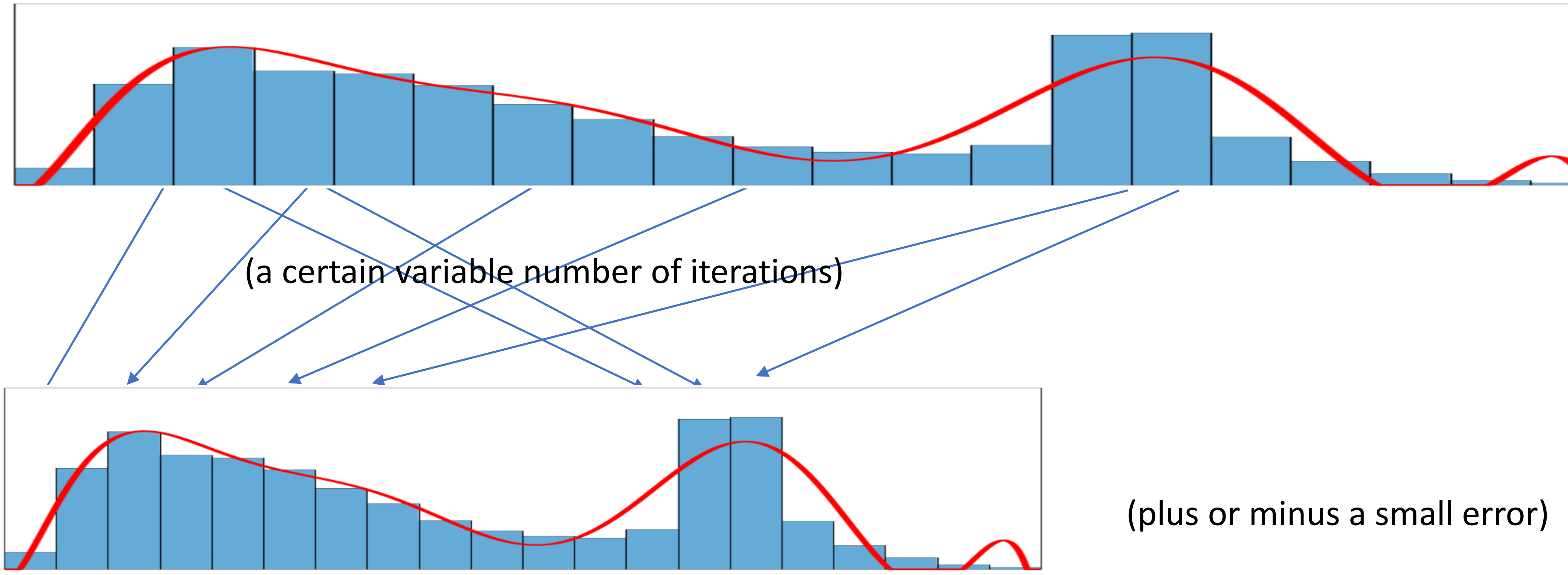
A key difficulty with the Collatz iteration is that it can greatly distort the distribution of a set of numbers – some numbers collide into each other, others get skipped entirely.



As a consequence, the statistical behavior of Collatz iteration quickly becomes intractable to study.

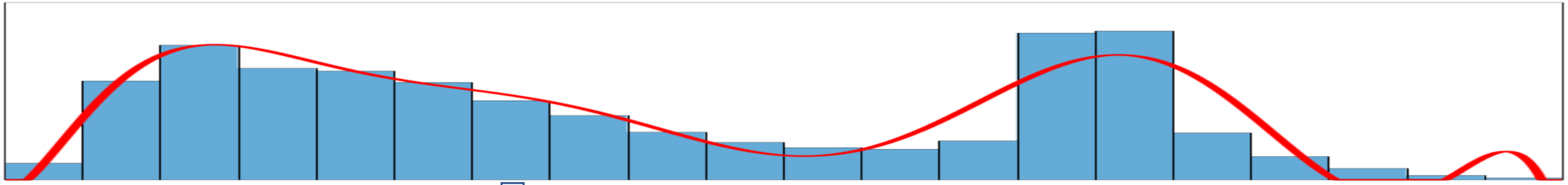


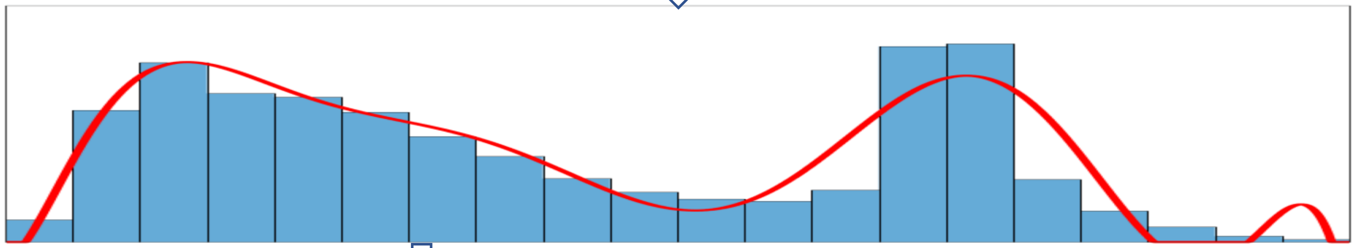
However, I was able to construct an (approximate) invariant measure – a distribution of numbers that iterates to something resembling a smaller version of itself.



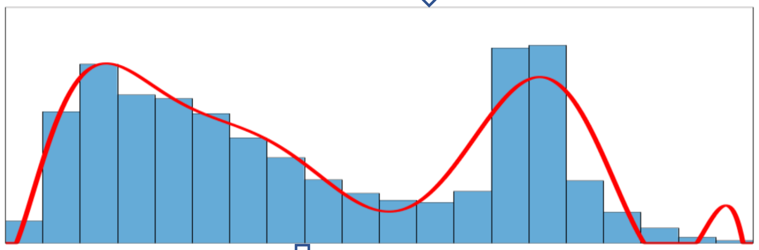
Iterating this fact gives the result.

(after 49 pages of argument)

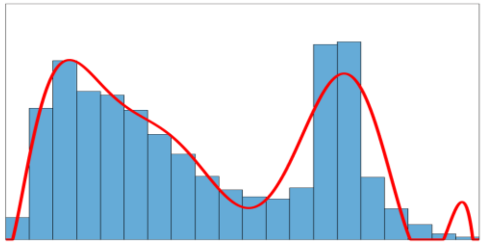




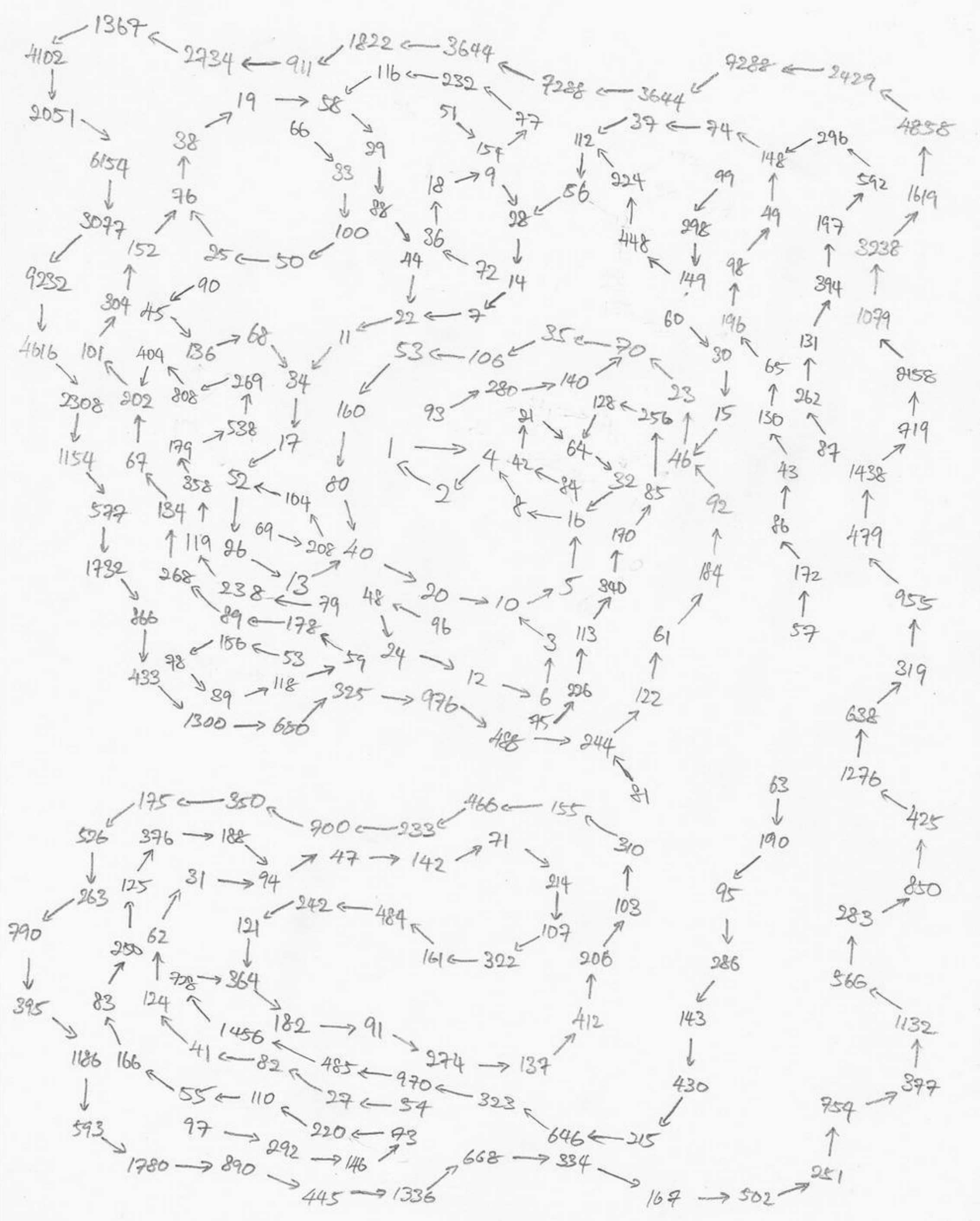
(plus or minus a small error)



(plus or minus a small error)



(plus or minus a small error)



Thanks for listening!