Éric Angelini’s Comma Sequence

Experimental Math Seminar, Rutgers, January 18 2024

Neil J. A. Sloane, Visiting Scholar, Math. Dept., Rutgers University;
and The OEIS Foundation, Highland Park, NJ
(njasloane@gmail.com)

Start with video:

Lorenzo Angelini,
Happy
birthday Éric!!, Youtube video.

The Comma Sequence (Eric Angelini, 2006) is defined by:

Comma Numbers = First Differences

a(n): 1 , 12 , 35 , 94 , 135 , 186 , 248 , 331 , 344 , … A121805

cn(n): 11 , 23 , 59 , 41 , 51 , 62 , 83 , 13 , … A366487

comma numbers

Edwin Clark: a(2137453) = 99999945

and the next term does not exist!

a landmine

But if we start with 3 we get 3 , 36

(ends)

33 61 fails etc

The manuscript:

The Comma Sequence: A Simple Sequence With Bizarre Properties

Éric Angelini, Michael S. Branicky, Giovanni Resta, N. J. A. Sloane, and David W. Wilson

Draft, January 18 2024. Will be on my homepage NeilSloane.com

Summary of talk

- Base-dependent sequences with remarkable properties
- Will give good model of comma sequence:
  kangaroo hopping along infinite road with occasional small clumps of landmines.
- Know where landmines are (Th 1)
- Sequence is always finite for base $B>2$ (conj)
- Expected length $\exp(2B)$ (conj)

- n-th term (if it exists) is about $n B^2/2$
- Very fast program
- Occasionally 2 choices for comma-number, definition says take smaller
- If allowed to take other path, infinite comma sequences exist (theorem)

- In base $B=3$, know how to get to oo
- In base 10, know oo path exists, know 10^{84} terms, don’t know explicit path to oo

Outline

- Summary of talk
- The comma sequence
- The Landmines, Theorem 1
- The Successor Graph G_s and the Child Graph G_c
- The child graph contains infinite paths
- Periodicity of comma numbers, the algorithm
- Chance of hitting landmine, aveage life of comma sequence
- Open problems

If we start at n, the sequence has X terms and the last term is Y:

n	1	2	3	4	5	6	7	OEIS
X	2137453	194697747222394	2	199900	19700	209534289952018960	15	A330128
Y	99999945	9999999999999918	36	9999945	999945	9999999999999999936	936	A330129

What on earth is going on?

Basic facts about comma sequences

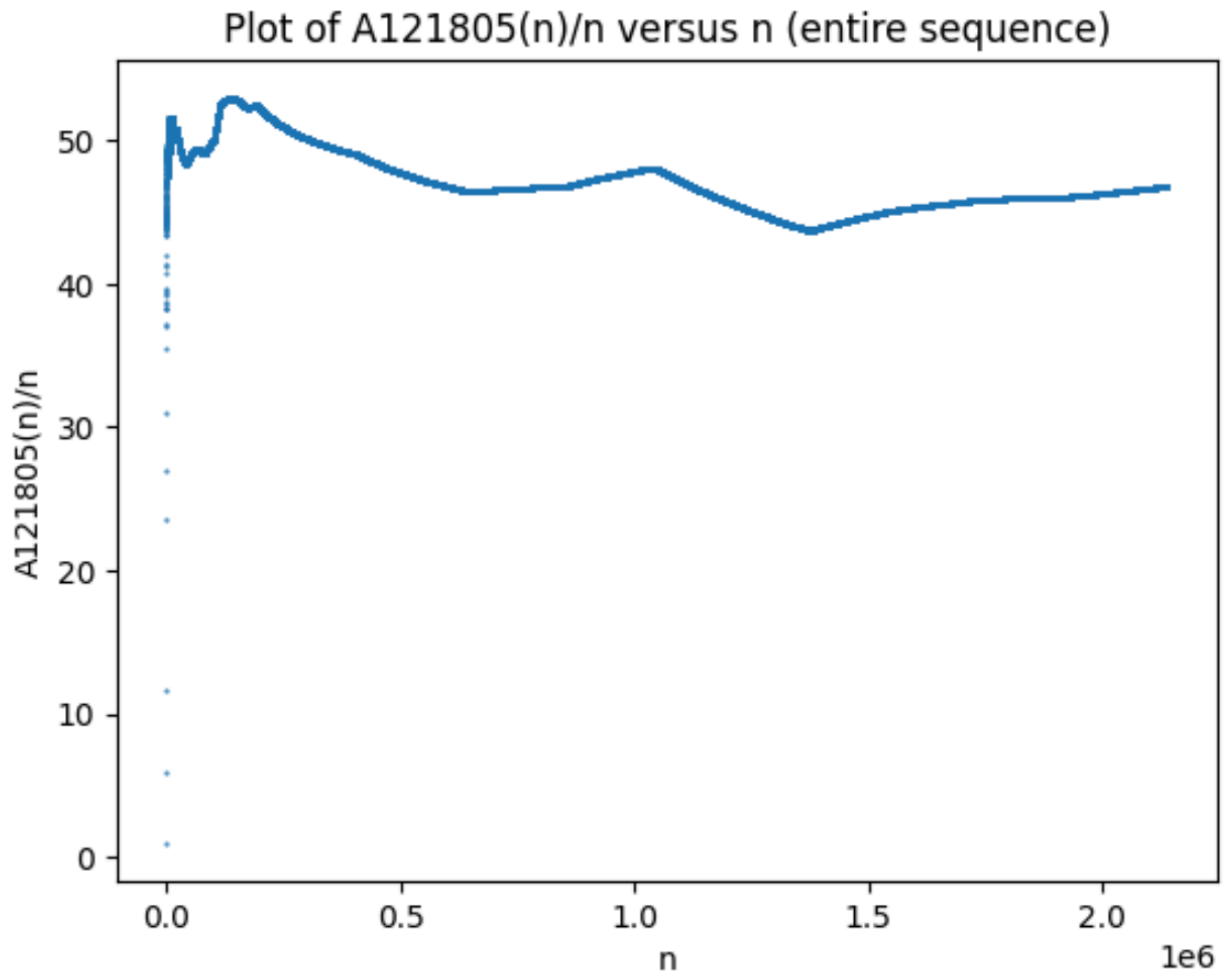
Comma-successor to a(n)   a(n+1) = a(n) + cn(n)   (add the comma number)

If leading digit didn't change, and a(n) = bcdef then cn(n) = fb

In base B, 1 <= cn(n) <= B^2 - 1, so small steps

Average step is B^2/2, so good rule of thumb is

a(n)   approx=   n*B^2/2 if it exists



47

A121805(n)/n

The Landmines

Theorem 1. In base B >= 2, the landmines (numbers with no successor) are

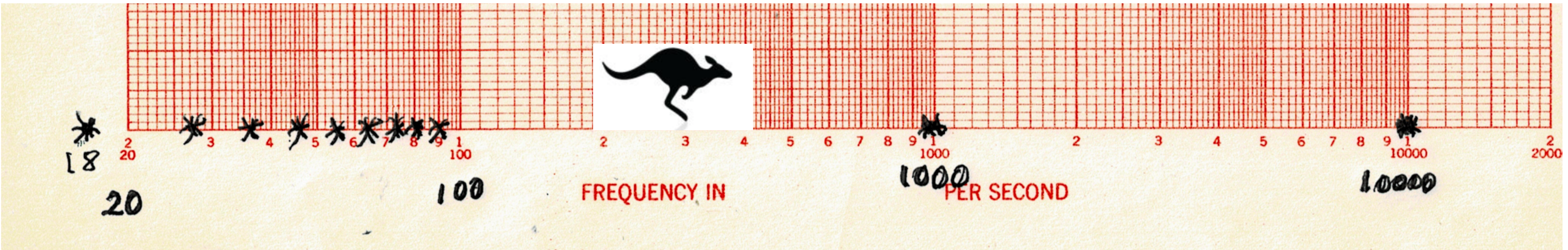
the numbers B-1 B-1 …B-1 x y with i >= 0 copies of B-1, 1 <= x <= B-2, y=B-1-x.

Base B=10: 18 27 36 45 54 63 72 81, 918 927 936 945 954 963 972 981,

9918 9927 … 9981, 99918 99927 … 99981, 999918 999927 … 999981, … A367341

Note: There are 8 with 2 digits, 8 with 3 digits, 8 with 4 digits, … Landmines are

rare and only occur near the end of each interval from 10^k to 10^(k+1)-1.



Model: Kangaroo making bounded jumps on infinite road with clusters of 8 mines at every power of 10 cells

Proof of Theorem 1

LD = Leading Digit

(1) The listed k are landmines. If k has 2 digits, easy. WLOG assume base B = 10.
Suppose k = 999xy, with at least one 9, and x + y = 9.

Suppose k has successor k'. If k' = 9…, k' = k + y9 >= 100…0, contradiction.
If k' = 1…, k' = k + y1 < 100…0, contradiction.

(2a) Conversely, suppose k has no successor, and at least 3 digits, say k = f s x y, where f, x, y have 1 digit and s has i >= 0 digits.
If fs is not 999…9, then LD(k) = LD(k') and can take k' = yf, so k' exists, contradiction.
Therefore fs = 999…9.
Suppose k = 999..9xy has successor k' = g….
if g = 9 there was no overflow from last two digits, so xy + y9 < 100, and k' = k + yg works, so need xy + yg >= 100,
that is 10(x+y) + 10y + 9 >= 100 …..(1)
On the other hand, if g = 1, then we can take k' = 99…9xy + y1 unless xy + y1 < 100, so need
10(x+y) + y + 1 < 100 ….(2)
(1) and (2) imply xy is one of 18, 27, 36, 45, 54, 63, 72, or 81. Done.

(2b) If k has no successor and two digits, then …. [see the manuscript!]

If more than one comma-number works:

Rules say take smallest successor. In rare cases there are two choices for the comma-number, never more. Call them comma-children.

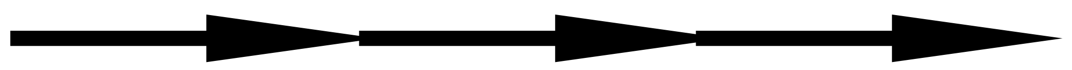
14 has two comma-children, 59 and 60	14 , 59	14 , 60
comma-number	45	46

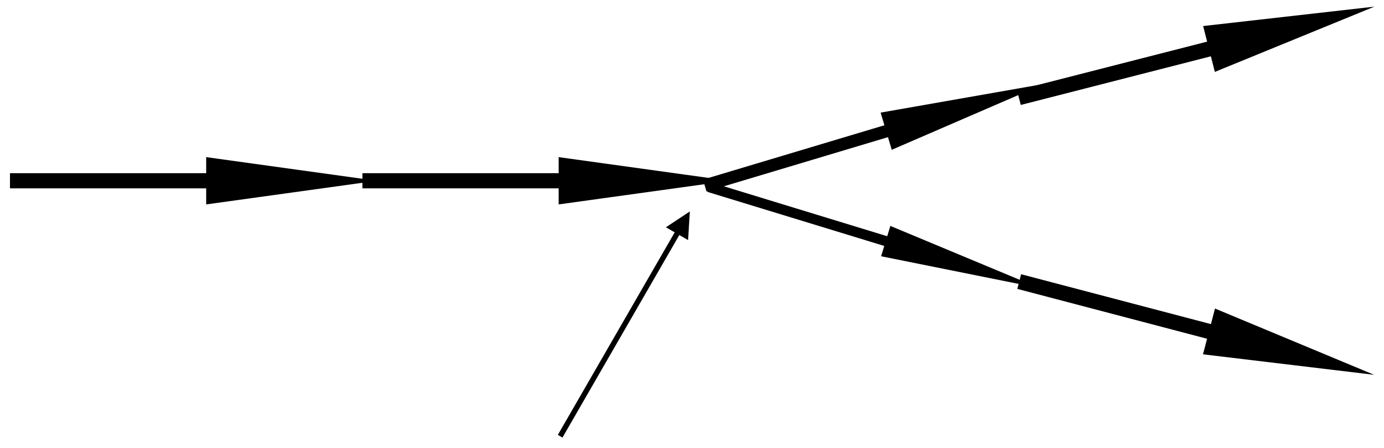
Theorem 2. Numbers with 2 children in base B are wx with 1 <= x <= [(B-1)/2] and w = B-1-2x, and the numbers d B-1 B-1 … B-1 d B-1-d with i >= 0 copies of B-1 and 1 <= d <= B-2.

In base 10: **14, 33, 52, 71, 118, 227, 336, 445, 554, 663, 772, 881, 1918, 2927… A367346**

By the Law of Primogeniture, the first-born child (the smallest) is the successor.

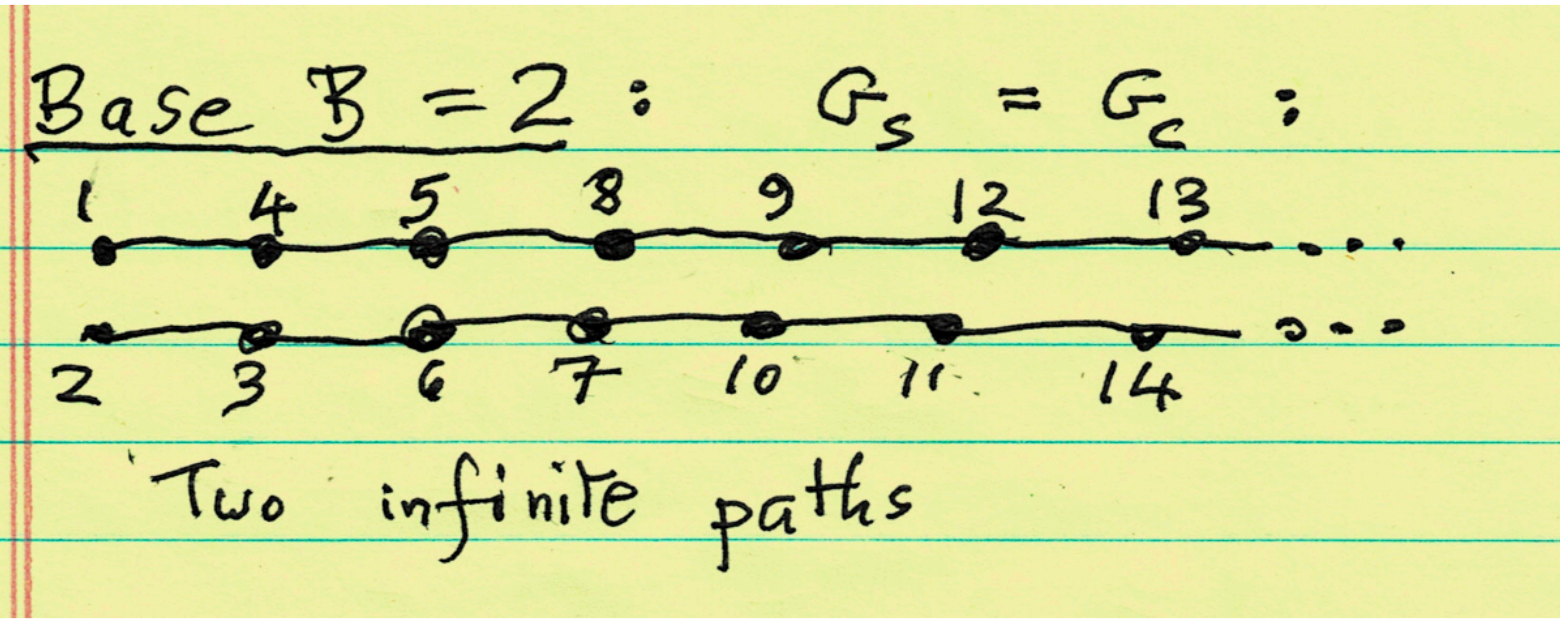
The successor graph G_s and the child graph G_c

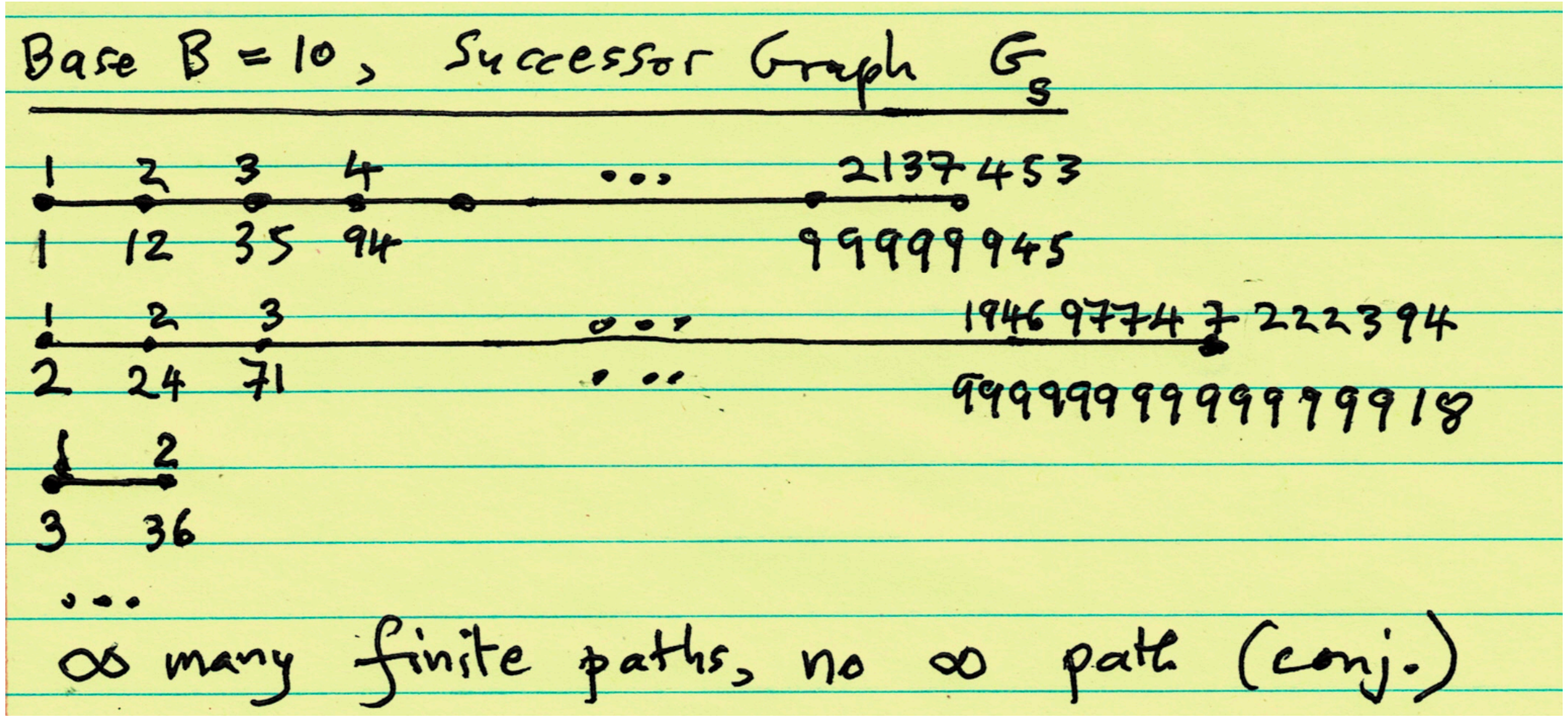
G_s 



G_c

Nodes are all positive integers, in both cases.





Predecessors

Lemma: If n is a child of k, then
$k = n - x*B - f$, f = first digit of n, $x = (n-f) \bmod B$

k ⟶ n

Notes: If n IS a child, k is unique. If $k < 0$, n is not a child.
If n is a child of k, n may not be the successor to k.

Theorem 3: In base B >= 3, all numbers n >= B^2-1 are comma-successors
except c*B^i, where 2 <= c <= B-1, i >= 2.

In base B = 10, the numbers that are not successors are (infinitely many)

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25, 26, 27, 28, 29, 30, 31, 32, 37, 38, 39, 40, 41, 42, 43, 49, 50, 51, 52, 53, 54, 60, 62, 63, 64, 65, 70, 74, 75, 76, 80, 86, 87, 90, 98, 200, 300, 400, 500, 600,

Theorem 4: In base B >= 2, the only numbers that are not children are in the range 1 to B^2 -1.

[Follows from Lemma and Theorem 4]

In base B = 10, there are exactly 50 numbers that are not children (A367611):
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25, 26, 27, 28, 29, 30, 31, 32, 37, 38, 39, 40, 41, 42, 43, 49, 50, 51, 52, 53, 54, 62, 63, 64, 65, 74, 75, 76, 86, 87, 98

Child graphs G_c contain infinite paths

Theorem 5: For any base B >= 2, the child graph G_c contains an infinite path.

Proof: By Theorem 4, every node in G_c belongs to a tree with root <= B^2 - 1.
By the Infinite Pigeonhole Principle, one tree contains infinitely many nodes.
By Kõnig's Infinity Lemma(*), that tree contains an infinite path starting at the root.

Reference: L. J. Halbeisen, Combinatorial Set Theory, Springer 2012, page 1.

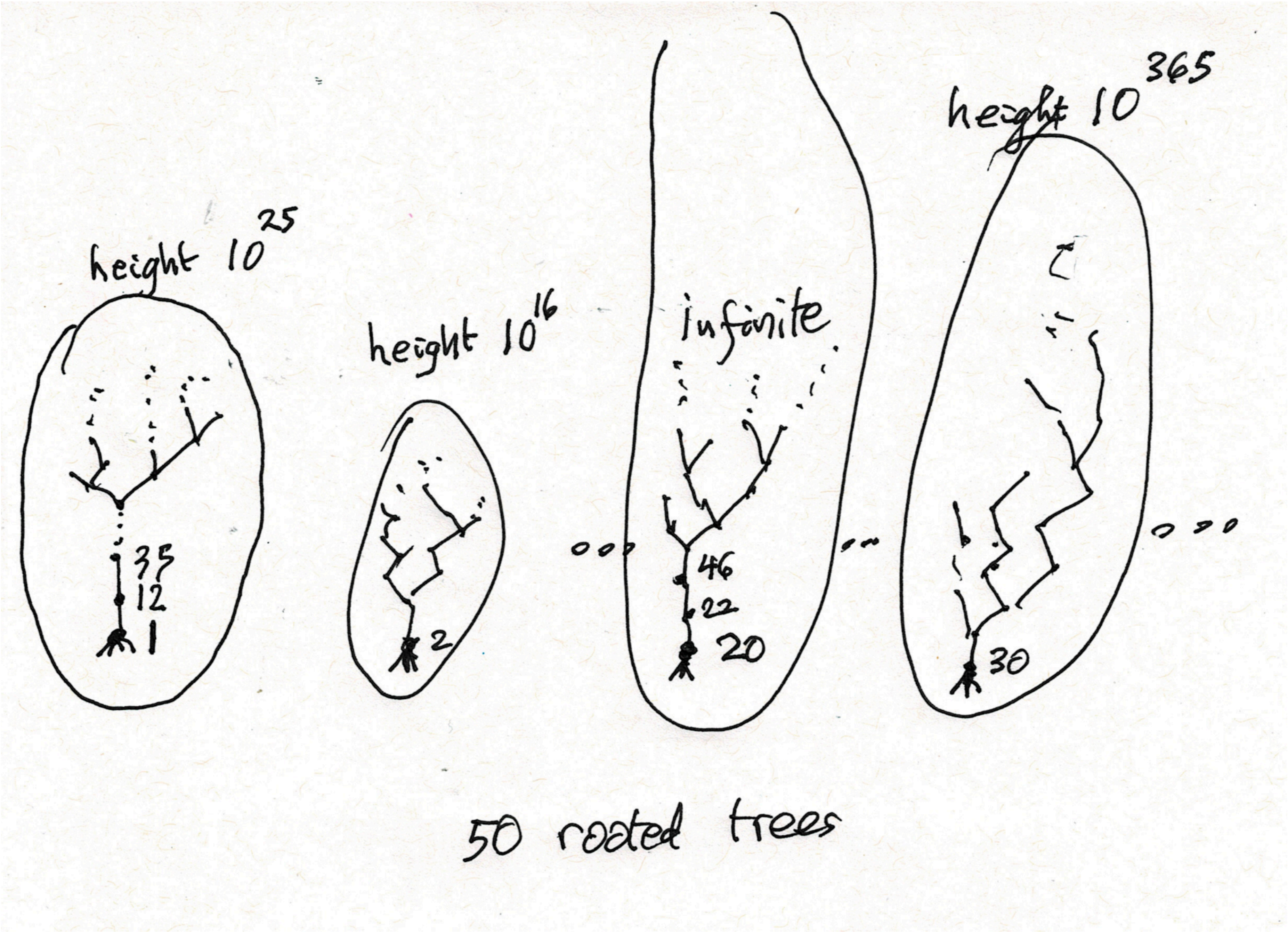
So if we are allowed to choose which child to follow at the branch-points, we can avoid all the landmines.

(*) An infinite rooted tree with bounded degree contains an infinite path starting at the root.

Proof uses Zermelo-Fraenkel-Axiom-of-Choice (ZFC) axioms

The Base 10 Child Graph G_c

It is made up of 50 rooted trees, all but one of which is finite



So we know there is at least one infinite path rooted at 20, and we can define A367620 to be the lexicographically earliest infinite tree.
It begins 20, 22, 46, 107, 178, 260, 262, 284, 327, 401,…
and the first branch point is at
a(412987860) = 19999999918.

We have followed this tree out through the first 69 branch points.
At the 30th branch point there were three candidates left,
but only one of them survived. So we know the first 30 choices
to make: 0 0 1 1 1 0 0 1 1 1 0 1 1 0 0 0 0 1 1 0 0 1 0 1 1 1 1 1 1 0 ← What is this sequence?

We have not identified that binary sequence yet!
By 69 branches there are still multiple candidates left.
We do know the first 10^84 terms of A367620 for certain.

For base 3 we know exactly what the infinite path is. It is unique,
and the choice sequence is simply 01010101…

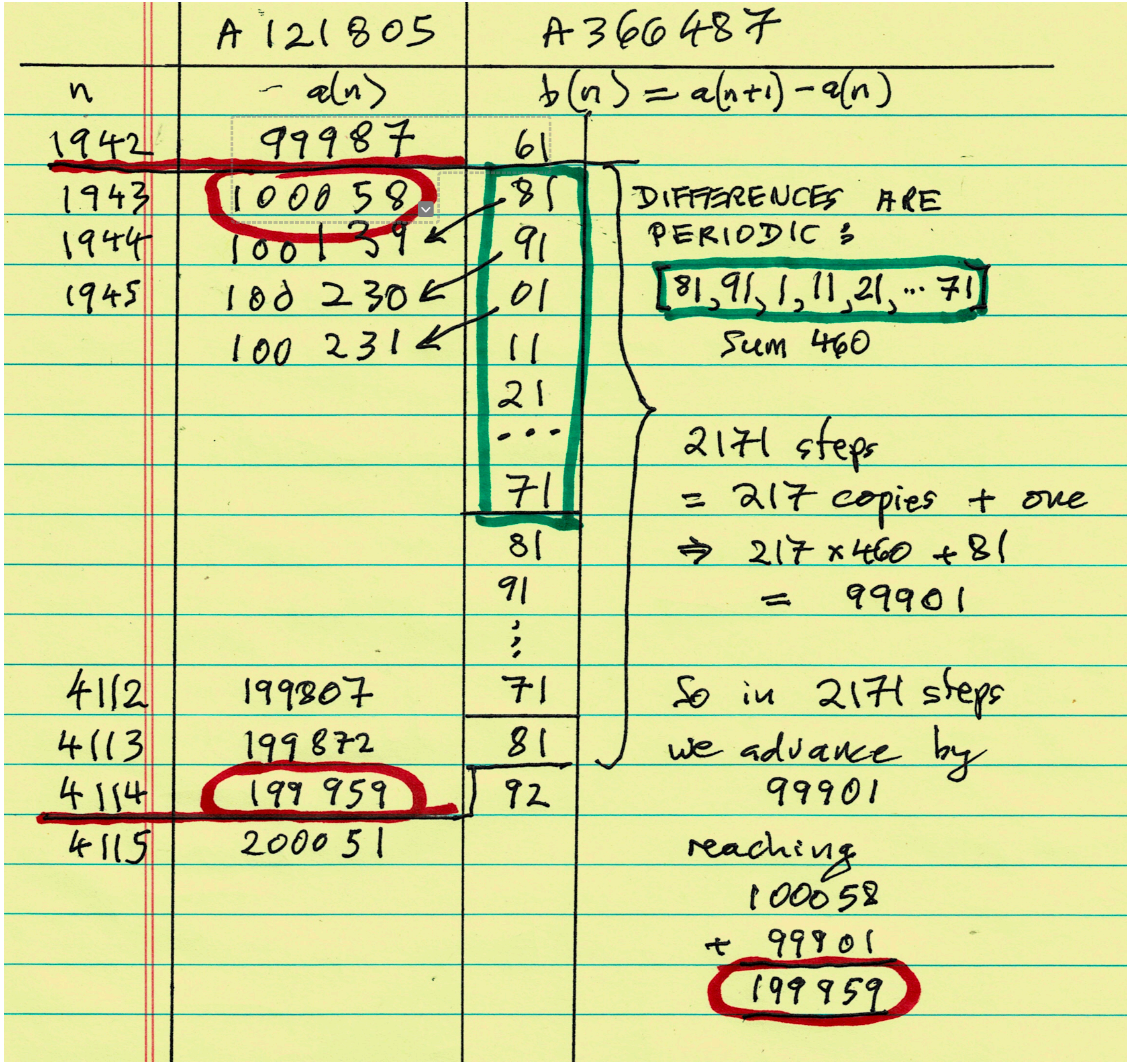
Periodicy of Comma Numbers - The Algorithm

1. As long as the leading digit of the comma sequence a(n) doesn’t change,
the comma number (= first difference!) is periodic,
and we cross that region in a small finite number of steps
(see next slide)
2. All landmines are within the last 100 steps between two powers of 10.
These two facts are the basis for our fast algorithm.
See Mathematica and Python implementations in A121805.

Speeding up the computation

A121805 and comma numbers A366487

Can move forward 2171 steps in one step



n	A121805 a(n)	A366487 b(n) = a(n+1) − a(n)
1942	99987	61
1943	100058	81
1944	100139	91
1945	100230	01
	100231	11
		21
		...
		71
		81
		91
		⋮
4112	199807	71
4113	199872	81
4114	199959	92
4115	200051	

DIFFERENCES ARE PERIODIC:
[81, 91, 1, 11, 21, ... 71]
Sum 460

2171 steps
= 217 copies + one
⇒ 217 × 460 + 81
= 99901

So in 2171 steps we advance by 99901
reaching
100058
+ 99901
199959

Chance of hitting landmine, average length of comma sequence

What is chance of kangaroo getting from 10^k to 10^(k+1)? There are 8 landmines in last 100 cells.
Say K. jumps in at square s, 999900 <= s <= 999999, and follows comma rule until hits mine
or reaches safety at 1000000. Answer: 88 succeed, 12 die.
Chance of death for each k is 12/100 = 1/8.33. Expected life = 10^8.33.

In base B>=2, not 12 but this sequence:

(Conjecture!)

0, 1, 2, 4, 5, 7, 8, 11, 12, 14, 16, 18, 20, 23, 24, 26, 29, … (100 terms) …. (*)

It was not in OEIS(**), but first differences match A136107 (Thank you, OEIS) = no. of ways of
writing B as difference of two triangular numbers = A001227 - A010054.
Asked Václav Kotěšovec for help. Answer (see next slide):

(**) It is now: see A368364.

Chance of hitting landmine, average length of comma sequence (2)

Want growth of (*)

$\overset{\text{conj.}}{=}$ Partial sums of A136107

= Partial sums of A001227 − ~~A010054~~ negligable

From Václav Kotěšovec:

A001227 = no of odd divisors of B

Dirichlet g.f. $= \zeta(s)\left(1-\frac{1}{2^s}\right)$

By Perron's theorem,

Partial sums of A1227

$\sim$ Residue at $s=1$ of

$$\zeta(s)\left(1-\frac{1}{2^s}\right)\cdot\frac{B^s}{s}$$

$$\sim B\left(\frac{\log(2B)}{2}+\gamma-\frac{1}{2}\right)$$

Chance of dying from B^k to B^{k+1}

$$\sim B\,\frac{\log B}{2}\Big/B^2 = \frac{\log B}{2B}$$

Expected life $= \frac{2B}{\log B}$ "centuries"

$$= B^{\frac{2B}{\log B}} \text{ steps}$$

$$= e^{2B} \text{ steps (conj.)}$$

Thank you, Václav Kotěšovec!

The Comma Transform of a sequence of nonnegative numbers

Replace each comma ijk , pqr by the one- or two-digit number kp,

then throw away the original sequence

Example:

0 , 1 ,   2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 ,10 ,11,12 ,13 , …

1 ,12 , 23 ,34 ,45 ,56 ,67,78,89,91, 1 ,11, 21 , …   A367362, Nov. 2023

The main unsolved problems

- Given the starting term and the base B, how long is the comma sequence?

  (Expected value is conjectured to be exp(2B))

- Show all comma sequences are finite in base B > 3.

  (Have proof for B=3, and it is false for B=2.)

- In base B = 10, what is the infinite path?

  (Know first 10^85 terms, first 30 decisions.)