POSITIVE SEQUENCES IN THE OEIS

Philipp Nuspl

July 19, 2023

JOHANNES KEPLER
UNIVERSITY LINZ

- Neil Sloane collects integer sequences since 1964.
- Since 1996 the database was first published online as the On-Line Encyclopedia of Integer Sequences (OEIS).
- At the moment it contains around 360 000 integer sequences.
- How many of them satisfy a linear recurrence?

Figure: Neil Sloane at the e-party to celebrate the 100,000-th sequence in 2004

Definition: C-finite

A sequence $c(n) \in \mathbb{Q}^{\mathbb{N}}$ is called C-finite if there are constants $\gamma_0, \ldots, \gamma_r \in \mathbb{Q}$, not all zero, with

$$\gamma_0\, c(n) + \cdots + \gamma_r\, c(n+r) = 0 \quad \text{for all } n \in \mathbb{N}.$$

Definition: C-finite

A sequence $c(n) \in \mathbb{Q}^{\mathbb{N}}$ is called C-finite if there are constants $\gamma_0, \ldots, \gamma_r \in \mathbb{Q}$, not all zero, with

$$\gamma_0 \, c(n) + \cdots + \gamma_r \, c(n+r) = 0 \quad \text{for all } n \in \mathbb{N}.$$

- Examples: Fibonacci sequence, Perrin numbers, geometric sequences, etc.
- The minimal r is called the order of c.

Definition: P-recursive

A sequence $c(n) \in \mathbb{Q}^{\mathbb{N}}$ is called *P-recursive* if there are polynomials $p_0, \ldots, p_r \in \mathbb{Q}[x]$, not all zero, such that

$$p_0(n)\, c(n) + \cdots + p_r(n)\, c(n+r) = 0 \quad \text{for all } n \in \mathbb{N}.$$

Definition: P-recursive

A sequence $c(n) \in \mathbb{Q}^{\mathbb{N}}$ is called P-recursive if there are polynomials $p_0, \ldots, p_r \in \mathbb{Q}[x]$, not all zero, such that

$$p_0(n)\, c(n) + \cdots + p_r(n)\, c(n+r) = 0 \quad \text{for all } n \in \mathbb{N}.$$

- Examples: C-finite sequences, Harmonic numbers, factorials, etc.
- We call r the order of the recurrence.
- We call $\max_i \deg p_i$ the degree of the recurrence.
- A P-recursive sequence satisfies many recurrences (order/degree curve)...

Does the sequence with initial values

$$1, 1, 2, 3, 5, 8, \ldots$$

satisfy a linear recurrence?

Does the sequence with initial values

$$1, 1, 2, 3, 5, 8, \ldots$$

satisfy a linear recurrence? Suppose it satisfies

$$\gamma_0\, c(n) + \gamma_1\, c(n+1) + \gamma_2\, c(n+2) = 0 \quad \text{for all } n \in \mathbb{N}.$$

Does the sequence with initial values

$$1, 1, 2, 3, 5, 8, \ldots$$

satisfy a linear recurrence? Suppose it satisfies

$$\gamma_0\, c(n) + \gamma_1\, c(n+1) + \gamma_2\, c(n+2) = 0 \quad \text{for all } n \in \mathbb{N}.$$

Hence, in particular

$$\gamma_0\, c(0) + \gamma_1\, c(1) + \gamma_2\, c(2) = 0$$
$$\gamma_0\, c(1) + \gamma_1\, c(2) + \gamma_2\, c(3) = 0$$
$$\gamma_0\, c(2) + \gamma_1\, c(3) + \gamma_2\, c(4) = 0$$
$$\gamma_0\, c(3) + \gamma_1\, c(4) + \gamma_2\, c(5) = 0$$

Does the sequence with initial values

$$1, 1, 2, 3, 5, 8, \ldots$$

satisfy a linear recurrence? Suppose it satisfies

$$\gamma_0\, c(n) + \gamma_1\, c(n+1) + \gamma_2\, c(n+2) = 0 \qquad \text{for all } n \in \mathbb{N}.$$

Hence, in particular

$$\gamma_0\, 1 + \gamma_1\, 1 + \gamma_2\, 2 = 0$$

$$\gamma_0\, 1 + \gamma_1\, 2 + \gamma_2\, 3 = 0$$

$$\gamma_0\, 2 + \gamma_1\, 3 + \gamma_2\, 5 = 0$$

$$\gamma_0\, 3 + \gamma_1\, 5 + \gamma_2\, 8 = 0$$

Does the sequence with initial values

$$1, 1, 2, 3, 5, 8, \ldots$$

satisfy a linear recurrence? Suppose it satisfies

$$\gamma_0\, c(n) + \gamma_1\, c(n+1) + \gamma_2\, c(n+2) = 0 \qquad \text{for all } n \in \mathbb{N}.$$

Hence, in particular

$$\gamma_0\, 1 + \gamma_1\, 1 + \gamma_2\, 2 = 0$$
$$\gamma_0\, 1 + \gamma_1\, 2 + \gamma_2\, 3 = 0$$
$$\gamma_0\, 2 + \gamma_1\, 3 + \gamma_2\, 5 = 0$$
$$\gamma_0\, 3 + \gamma_1\, 5 + \gamma_2\, 8 = 0$$

This linear equation has the nontrivial solution

$$\gamma_0 = 1, \gamma_1 = 1, \gamma_2 = -1$$

which gives rise to the recurrence

$$c(n) + c(n+1) - c(n+2) = 0.$$

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?
- If c is P-recursive of order r and degree d, we get a linear system

$$\sum_{i=0}^{r} \sum_{j=0}^{d} p_{i,j} n^j c(n+i) = 0 \quad \text{for } n = 0, \ldots, D-r$$

of size

$$D - r + 1 \times (r+1)(d+1)$$

for the unknown coefficients $p_{i,j} \in \mathbb{Q}$ of the recurrence.

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?
- If c is P-recursive of order r and degree d, we get a linear system

$$\sum_{i=0}^{r} \sum_{j=0}^{d} p_{i,j} n^j c(n+i) = 0 \quad \text{for } n = 0, \ldots, D-r$$

  of size

$$D - r + 1 \times (r+1)(d+1)$$

  for the unknown coefficients $p_{i,j} \in \mathbb{Q}$ of the recurrence.
- If the system is overdetermined and we get a non-trivial solution, it is probably there for a reason...

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?
- If c is P-recursive of order r and degree d, we get a linear system

$$\sum_{i=0}^{r} \sum_{j=0}^{d} p_{i,j} n^j c(n+i) = 0 \quad \text{for } n = 0, \ldots, D-r$$

  of size

$$D - r + 1 \times (r+1)(d+1)$$

  for the unknown coefficients $p_{i,j} \in \mathbb{Q}$ of the recurrence.
- If the system is overdetermined and we get a non-trivial solution, it is probably there for a reason...
- The more equations compared to variables we have, the more confident we can be in our guess.

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?
- If c is P-recursive of order r and degree d, we get a linear system

$$\sum_{i=0}^{r} \sum_{j=0}^{d} p_{i,j} n^j c(n+i) = 0 \quad \text{for } n = 0, \ldots, D-r$$

of size

$$D - r + 1 \times (r+1)(d+1)$$

for the unknown coefficients $p_{i,j} \in \mathbb{Q}$ of the recurrence.
- If the system is overdetermined and we get a non-trivial solution, it is probably there for a reason...
- The more equations compared to variables we have, the more confident we can be in our guess.
- Here, $e \in \mathbb{N}$ denotes the confidence we have in our guess.

- Given $c(0), c(1), \ldots, c(D)$. Does c satisfy a recurrence?
- If c is P-recursive of order r and degree d, we get a linear system

$$\sum_{i=0}^{r} \sum_{j=0}^{d} p_{i,j} n^j c(n+i) = 0 \quad \text{for } n = 0, \ldots, D-r$$

  of size

$$D - r + 1 \times (r+1)(d+1)$$

  for the unknown coefficients $p_{i,j} \in \mathbb{Q}$ of the recurrence.
- If the system is overdetermined and we get a non-trivial solution, it is probably there for a reason...
- The more equations compared to variables we have, the more confident we can be in our guess.
- Here, $e \in \mathbb{N}$ denotes the confidence we have in our guess.
- Idea: use guessing to determine how many sequences in the OEIS are C-finite or P-recursive.

- If we have more terms of a sequence, we can guess recurrences better.

- If we have more terms of a sequence, we can guess recurrences better.
- For about 6% of the sequences at most 10 terms are given, about 50% have at most 100 terms, about 13% have at least 10 000 terms given.

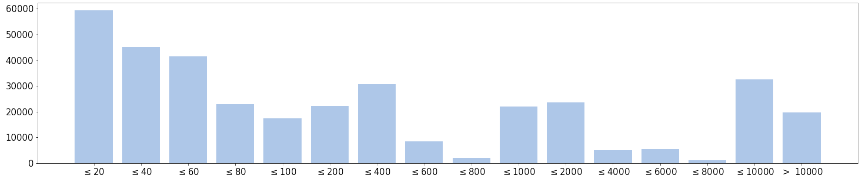


Figure: Number of sequences with number of terms specified in the OEIS

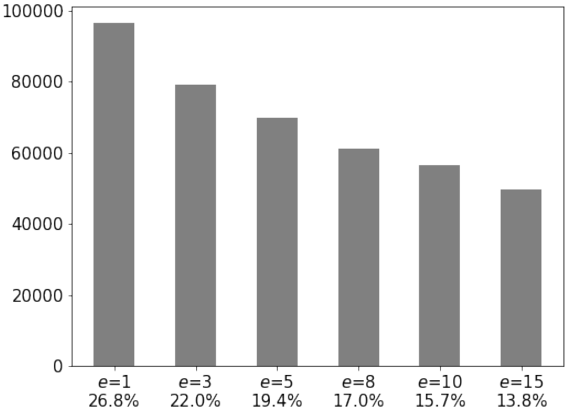


Figure: Number of C-finite sequences with order ≤ 100

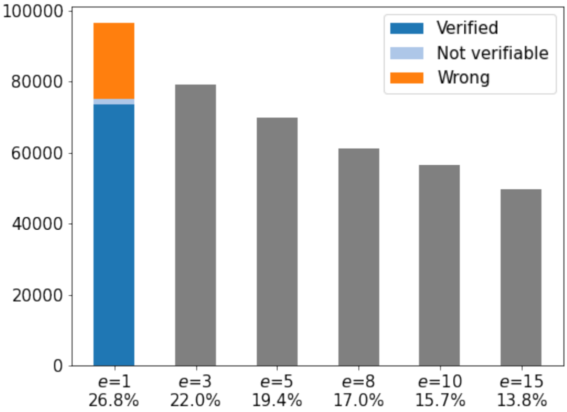


Figure: Number of C-finite sequences with order ≤ 100

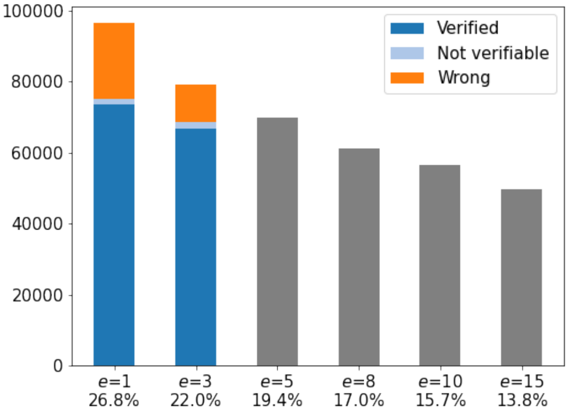


Figure: Number of C-finite sequences with order ≤ 100

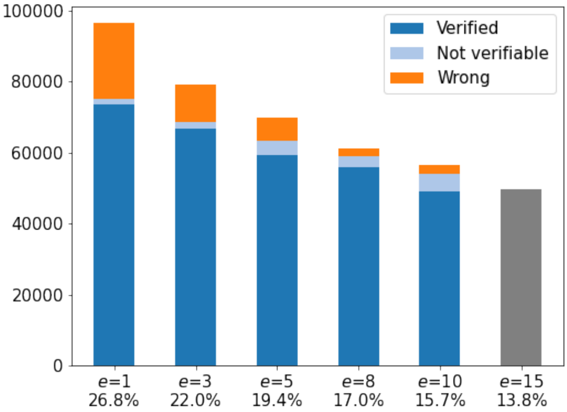


Figure: Number of C-finite sequences with order ≤ 100

C-finite sequences orders

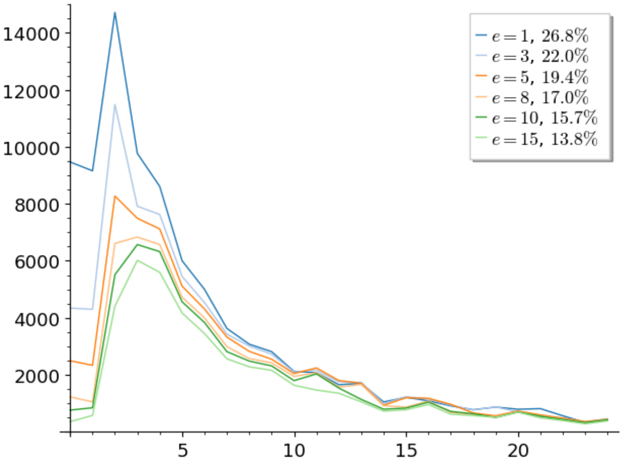


Figure: Number of C-finite sequences of given order

Let $a(n)$ be the sequence A000122 with initial values

$$1, 2, 0, 0, 2, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, \ldots$$

Guessing yields the recurrence (if $e = 1$)

$$n(n-1)\,a(n) = 0.$$

Let $a(n)$ be the sequence A000122 with initial values

$$1, 2, 0, 0, 2, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, \ldots$$

Guessing yields the recurrence (if $e = 5$)

$$n(n-1)(n-4)\, a(n) = 0.$$

Let $a(n)$ be the sequence A000122 with initial values

$$1, 2, 0, 0, 2, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, \ldots$$

Guessing yields the recurrence (if $e = 50$)

$$n(n-1)(n-4)(n-9)(n-16)(n-25)(n-36)(n-49)\,a(n) = 0.$$

Let $a(n)$ be the sequence A000122 with initial values

$$1, 2, 0, 0, 2, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, \ldots$$

Guessing yields the recurrence (if $e = 50$)

$$n(n-1)(n-4)(n-9)(n-16)(n-25)(n-36)(n-49)\,a(n) = 0.$$

Idea: If sequence has many zeros, make sure that the system is even more overdetermined (Kauers and Verron 2019).

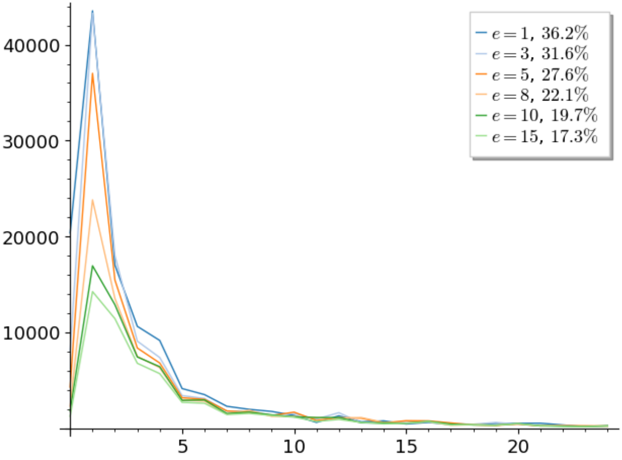


Figure: Number of P-recursive sequences of given order

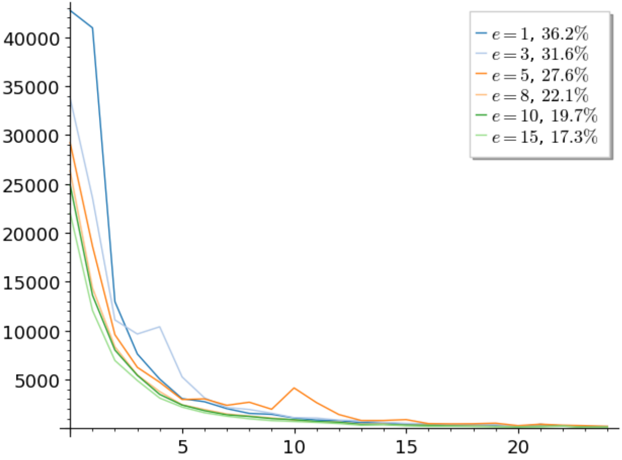


Figure: Number of P-recursive sequences of given degree

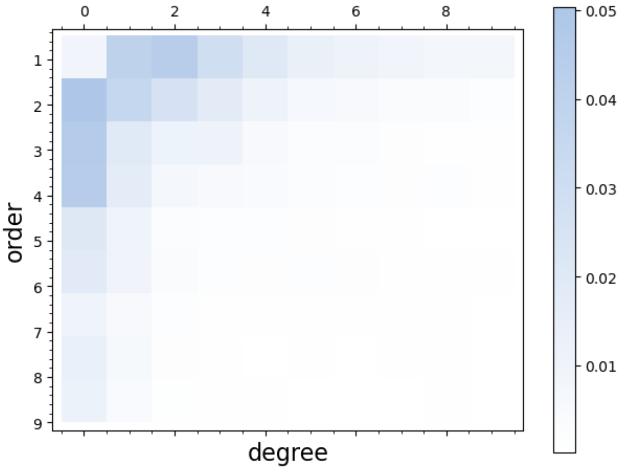


Figure: Number of P-recursive sequences of given order/degree

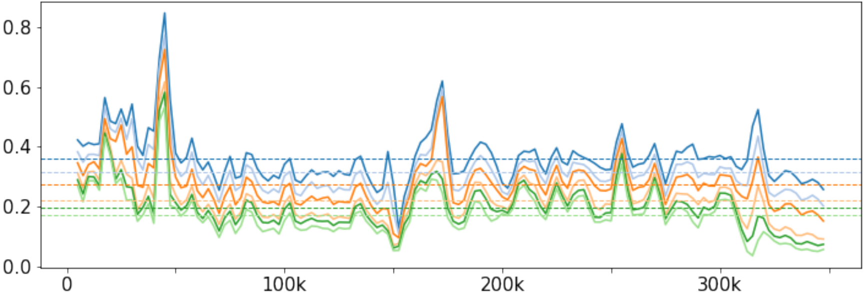


Figure: Ratio of P-recursive sequences in the OEIS

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N}\colon c(n) > 0$$

holds. We write $c > 0$, if this is the case.

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N} \colon c(n) > 0$$

holds. We write $c > 0$, if this is the case.

- The sequence c has no zero terms if and only if c^2 is a positive sequence (Skolem Problem).

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N}\colon c(n) > 0$$

holds. We write $c > 0$, if this is the case.

- The sequence c has no zero terms if and only if c^2 is a positive sequence (Skolem Problem).
- Exact same ideas apply for non-negativity: $\forall n \in \mathbb{N}\colon c(n) \geq 0$.

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N} \colon c(n) > 0$$

holds. We write $c > 0$, if this is the case.

- The sequence c has no zero terms if and only if c^2 is a positive sequence (Skolem Problem).
- Exact same ideas apply for non-negativity: $\forall n \in \mathbb{N} \colon c(n) \geq 0$.
- The problem $c > d$ is equivalent to $c - d > 0$ as C-finite sequences form a ring.

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N}\colon c(n) > 0$$

holds. We write $c > 0$, if this is the case.

- The sequence c has no zero terms if and only if c^2 is a positive sequence (Skolem Problem).
- Exact same ideas apply for non-negativity: $\forall n \in \mathbb{N}\colon c(n) \geq 0$.
- The problem $c > d$ is equivalent to $c - d > 0$ as C-finite sequences form a ring.
- Only known to be decidable for sequences of small order:
  - Order 2 (Halava, Harju, and Hirvensalo 2006),
  - Order 3 (Laohakosol and Tangsupphathawat 2009),
  - Order 4 and 5 (Ouaknine and Worrell 2014).

Suppose we are given a C-finite sequence c. We want to decide whether

$$\forall n \in \mathbb{N}\colon\, c(n) > 0$$

holds. We write $c > 0$, if this is the case.

- The sequence c has no zero terms if and only if c^2 is a positive sequence (Skolem Problem).
- Exact same ideas apply for non-negativity: $\forall n \in \mathbb{N}\colon\, c(n) \geq 0$.
- The problem $c > d$ is equivalent to $c - d > 0$ as C-finite sequences form a ring.
- Only known to be decidable for sequences of small order:
  - Order 2 (Halava, Harju, and Hirvensalo 2006),
  - Order 3 (Laohakosol and Tangsupphathawat 2009),
  - Order 4 and 5 (Ouaknine and Worrell 2014).
- Can we prove positivity of sequences in practice?

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(1) = 1 \quad \text{because} \quad 1 = 1.$$

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(2) = 2 \quad \text{because} \quad 2 = 1 + 1,\ 2 = 2.$$

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(3) = 2 \quad \text{because} \quad 3 = 1 + 1 + 1,\ 3 = 1 + 2.$$

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(4) = 3 \quad \text{because} \quad 4 = 1 + 1 + 1 + 1,\ 4 = 1 + 1 + 2,\ 4 = 2 + 2.$$

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(4) = 3 \quad \text{because} \quad 4 = 1+1+1+1,\ 4 = 1+1+2,\ 4 = 2+2.$$

In fact, $c(n)$ is C-finite of order 8 satisfying

$$\begin{aligned} c(n) - c(n+1) - c(n+2) + c(n+3) - c(n+5) \\ + c(n+6) + c(n+7) - c(n+8) = 0 \end{aligned}$$

with initial values $c = \langle 1, 1, 2, 2, 3, 4, 5, 6, \ldots \rangle$.

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(4) = 3 \quad \text{because} \quad 4 = 1+1+1+1,\ 4 = 1+1+2,\ 4 = 2+2.$$

In fact, $c(n)$ is C-finite of order 8 satisfying

$$\begin{aligned} c(n) - c(n+1) - c(n+2) + c(n+3) - c(n+5) \\ + c(n+6) + c(n+7) - c(n+8) = 0 \end{aligned}$$

with initial values $c = \langle 1, 1, 2, 2, 3, 4, 5, 6, \dots \rangle$.

We can compute

$$c(10\,n + k) = 5\,n^2 + (k+4)\,n + c(k) > 0 \quad \text{for all } k = 0, \dots, 9.$$

Let $c(n)$ denote the number of partitions of n into parts $1, 2$ or 5.

$$c(4) = 3 \quad \text{because} \quad 4 = 1 + 1 + 1 + 1,\ 4 = 1 + 1 + 2,\ 4 = 2 + 2.$$

In fact, $c(n)$ is C-finite of order 8 satisfying

$$\begin{aligned} c(n) - c(n+1) - c(n+2) + c(n+3) - c(n+5) \\ + c(n+6) + c(n+7) - c(n+8) = 0 \end{aligned}$$

with initial values $c = \langle 1, 1, 2, 2, 3, 4, 5, 6, \dots \rangle$.

We can compute

$$c(10\,n + k) = 5\,n^2 + (k+4)\,n + c(k) > 0 \qquad \text{for all } k = 0, \dots, 9.$$

Hence, c is the interlacing of positive sequences, so $c(n) > 0$ for all $n \in \mathbb{N}$.

- We took 1000 C-finite sequences from the OEIS which look positive.

- We took 1000 C-finite sequences from the OEIS which look positive.
- The sequences have the following orders:

order	1	2	3	4	5	6	7	8	9	10	> 10
	73	134	117	139	120	80	87	36	47	27	140

- We took 1000 C-finite sequences from the OEIS which look positive.
- The sequences have the following orders:

order	1	2	3	4	5	6	7	8	9	10	> 10
	73	134	117	139	120	80	87	36	47	27	140

- We implemented several known and new algorithms which can be used to show positivity of a C-finite sequence:
  - The `SageMath` implementation is part of the `rec_sequences` package.
  - The `Mathematica` implementation is part of the `RISCErgoSum` package.

- We took 1000 C-finite sequences from the OEIS which look positive.
- The sequences have the following orders:

order	1	2	3	4	5	6	7	8	9	10	> 10
	73	134	117	139	120	80	87	36	47	27	140

- We implemented several known and new algorithms which can be used to show positivity of a C-finite sequence:
  - The `SageMath` implementation is part of the `rec_sequences` package.
  - The `Mathematica` implementation is part of the `RISCErgoSum` package.
- The implementations can prove positivity of all 1000 sequences (N. and Pillwein 2022).
  - Decompose sequence into subsequences which have simple asymptotic behavior.

How can we prove positivity of C-finite sequences with SageMath? First, C-finite sequences can be defined in different ways:

How can we prove positivity of C-finite sequences with SageMath? First, C-finite sequences can be defined in different ways:

```
sage: C = CFiniteSequenceRing(QQ)
```

How can we prove positivity of C-finite sequences with SageMath? First, C-finite sequences can be defined in different ways:

```
sage: C = CFiniteSequenceRing(QQ)
sage: fib = C([1,1,-1], [1,1]) # use recurrence and initial values
```

How can we prove positivity of C-finite sequences with SageMath? First, C-finite sequences can be defined in different ways:

```
sage: C = CFiniteSequenceRing(QQ)
sage: fib = C([1,1,-1], [1,1]) # use recurrence and initial values
sage: fib[:10]
```

How can we prove positivity of C-finite sequences with SageMath? First, C-finite sequences can be defined in different ways:

```
sage: C = CFiniteSequenceRing(QQ)
sage: fib = C([1,1,-1], [1,1]) # use recurrence and initial values
sage: fib[:10]
[1, 1, 2, 3, 5, 8, 13, 21, 34, 55]
```

Several methods for proving C-finite (termwise) inequalities are implemented:

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1,1,2,2,3,4,5,6])
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1,1,2,2,3,4,5,6])
sage: 0 in c
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1,1,2,2,3,4,5,6])
sage: 0 in c
False
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1,1,2,2,3,4,5,6])
sage: 0 in c
False
sage: c > 0
```

Several methods for proving C-finite (termwise) inequalities are implemented:

```
sage: fib >= 0
True
sage: fib < 100
False
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1,1,2,2,3,4,5,6])
sage: 0 in c
False
sage: c > 0
True
```

What have we done?

- We estimated ratio of C-finite (around 15%) and P-recursive (around 20%) sequences in the OEIS.
- The results are available online. Hence, these sequences can be used for testing algorithms.
- We showed that proving positivity of C-finite sequences can usually be done in practice.

What is left?

- Can we prove positivity of P-recursive sequences in practice?

References I

Halava, Vesa, Tero Harju, and Mika Hirvensalo (2006). "Positivity of second order linear recurrent sequences". In: Discrete Applied Mathematics 154.3, pp. 447–451. ISSN: 0166-218X.

Kauers, Manuel and Thibaut Verron (2019). "Why You Should Remove Zeros from Data before Guessing". In: ACM Commun. Comput. Algebra 53.3, 126fffdfffdfffd129.

Laohakosol, Vichian and Pinthira Tangsupphathawat (2009). "Positivity of third order linear recurrence sequences". In: Discrete Applied Mathematics 157.15.

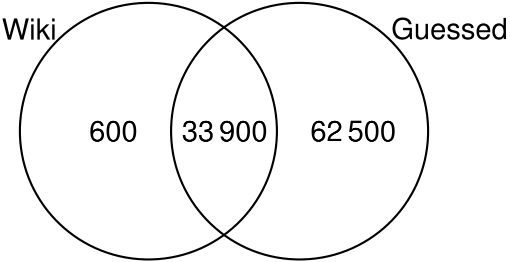
N., Philipp and Veronika Pillwein (2022). "A Comparison of Algorithms for Proving Positivity of Linearly Recurrent Sequences". In: Computer Algebra in Scientific Computing. Vol. 13366. LNCS. Springer International Publishing, pp. 268–287.

References II

Ouaknine, Joël and James Worrell (2014). "Positivity problems for low-order linear recurrence sequences". In: SODA '14: Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms.

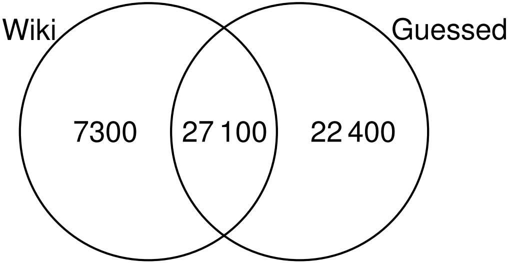
Comparison

- There is also a Wiki page with a list of C-finite sequences in the OEIS.
- Some that are clearly C-finite are not in there (e.g., A272636).
- Some look very C-finite but the recurrence is only conjectured and therefore the sequence not on the Wiki page (e.g., A281605).
- It contains about 34 000 sequences. Compare with guessed sequences with ensure=1.



Comparison

- There is also a Wiki page with a list of C-finite sequences in the OEIS.
- Some that are clearly C-finite are not in there (e.g., A272636).
- Some look very C-finite but the recurrence is only conjectured and therefore the sequence not on the Wiki page (e.g., A281605).
- It contains about 34 000 sequences. Compare with guessed sequences with ensure=10.



Example

The Berstel sequence (A007420) is C-finite of order 3 satisfying

$$4c(n) - 4c(n+1) + 2c(n+2) - c(n+3) = 0$$

with $c(0) = c(1) = 0, c(2) = 1$.

Define $d(n) = c(n+53)^2$. Then, d looks positive. However, it does not have a unique dominant root and cannot be decomposed to obtain one.