An Analysis of OEIS Generating Functions

Michael Assis

School of Mathematical and Physical Sciences
The University of Newcastle

13 December, 2017

What is the OEIS?

- **O**nline **E**ncyclopedia of **I**nteger **S**equences
- Started in 1965 by Neil Sloane
- Originally a punched card database
- Books published in 1973, 1995
- E-mail service 1994, website 1996
- User contributed in Wiki from 2010

The OEIS

- A searchable database of integer sequences
- Two kinds of searches:
  - Simple look-up
  - Superseeker
- Results include:
  - Sequence name and definition
  - Known formulas
  - References
  - Code for use in Maple, Mathematica, PARI/GP, . . .
- Resources:
  - Graphing
  - Musical rendering

Generating functions

Given a sequence of numbers, what is their pattern?

Examples:

- 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, ...
  n-th term: $2n$
- 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, ...
  n-th term: $\frac{\phi_1^n - \phi_2^n}{\phi_1 - \phi_2}$, $\quad \phi_1 = \frac{1+\sqrt{5}}{2}$, $\phi_2 = \frac{1-\sqrt{5}}{2}$
- $1, -\frac{1}{2}, \frac{1}{24}, -\frac{1}{720}, \frac{1}{40320}, -\frac{1}{3628800}, \frac{1}{479001600}, \frac{1}{87178291200}, \ldots$
  n-th term: $\frac{(-1)^n}{(2n)!}$
- 1, 6, 29, 130, 561, 2368, 9855, 40622, 166303, 677420, ...
  n-th term: ?

Generating functions

Questions for a given sequence/series:

- Does the sequence converge or diverge?
- If it converges, what is the radius of convergence?
- How do you analytically continue beyond the radius of convergence?
- Is there an efficient algorithm to compute the terms?

Generating functions

Convergence tests:

- Ratio test
  $$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r, \quad (r<1):\text{converges}$$
- Root test
  $$\lim_{n\to\infty}\sup\sqrt[n]{|a_n|} = r, \quad (r<1):\text{converges}$$
- Radius of convergence: $\dfrac{1}{r}$, $\quad r$ from Ratio or Root tests

Generating functions

Analytic continuation:

- From a series expansion around x_0 with radius of converge r:
  1. Find point x_1 within $|x_1 - x_0| < r$ with radius outside original domain
  2. Repeat
- Find a functional equation for the function
- Use an integral definition of the function
- Use the ODE that the function satisfies

Generating functions

Efficient computation:

- Have an explicit n-th term expression
- Use a recurrence relation
- Find series solution to ODE
- Find a Puiseux series to an algebraic equation

Example:

Treat sequence as coefficients of a power series:

2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, ...

$$\Rightarrow 2x + 4x^2 + 6x^3 + 8x^4 + 10x^5 + \ldots = \frac{2x}{(x-1)^2}$$

- Sequence diverges
- Radius of convergence 1
- Easy to compute sequence terms

Generating functions

Example:

Treat sequence as coefficients of a power series:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, \ldots$$

$$\Rightarrow 1 + 1x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + \ldots = \frac{1}{1 - x - x^2}$$

- Sequence diverges
- Radius of convergence $\frac{1}{\phi_1} \approx 0.618033\ldots$
- Terms satisfy recursion relation:
  $a_n = a_{n-1} + a_{n-2}, \quad a_0 = 1, \ a_1 = 1$
- and satisfies n-th term expression:
  $\dfrac{\phi_1^n - \phi_2^n}{\phi_1 - \phi_2}$

Generating functions

Example:

Treat sequence as coefficients of a power series:

$$1, -\frac{1}{2}, \frac{1}{24}, -\frac{1}{720}, \frac{1}{40320}, -\frac{1}{3628800}, \frac{1}{479001600}, \frac{1}{87178291200}, \cdots$$

$$\Rightarrow 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \ldots = \cos(x)$$

- Sequence converges
- Radius of converge is ∞
- Easy to calculate sequence terms

Example:

Treat sequence as coefficients of a power series:

$$1, 6, 29, 130, 561, 2368, 9855, 40622, 166303, 677420, \ldots$$

$$\Rightarrow 1x + 6x^2 + 29x^3 + 130x^4 + 561x^5 + 2368x^6 + \ldots = \ ?$$

Generating functions

Practical problems:

- It can take a long time to compute sequence terms
- You may not know an n-th term expression
- Radius of convergence is only known to within some error

Finding generating functions

When you don't have an n-th term expression, what to do?

- You can use the Ratio and Root tests to estimate the radius of convergence
- You can search for an ODE the series satisfies

Finding generating functions

Example:

- $$1 + 1x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + \ldots = \frac{1}{1 - x - x^2}$$
- Know in advance it satisfies ODE:
  $$(1 - x - x^2)\,\frac{dy(x)}{dx} - (1 + 2x)\,y(x) = 0$$
- Assuming 1st order ODE, 2nd degree polynomial coefficients:
  1. Compute: $x\,y,\ x^2 y,\ y',\ x\,y',\ x^2\,y'$
  2. Find series: $a_0\,y + a_1 x\,y + a_2 x^2\,y + b_0\,y' + b_1 x\,y' + b_2 x^2\,y'$
  3. Solve for coefficients $a_j,\ b_j$ such that each term vanishes:
     $a_0 = 1,\ a_1 = -1,\ a_2 = -1,\ b_0 = -1,\ b_1 = -2,\ b_2 = 0$

Finding generating functions

Things to consider:

- The full pattern may not have revealed itself yet
  - $\Rightarrow$ The solution may fail after a certain point
- Save terms for use in checking any solution
- Not enough terms known for finding polynomial coefficients
  - $\Rightarrow$ No solution found until more terms are known
- The series satisfies no finite order ODE
  - $\Rightarrow$ At best an approximation can be found

Software

Software

Maple:

- **listtodiffeq** and **seriestodiffeq** in gfun package
- linear ODE search
- Default max order is 2, default polyn coeff degree is 3
- Can handle arbitrary precision integers, rationals
- No control over saving later terms for checking
- Very slow

Software

Mathematica:

- **FindGeneratingFunction** and **FindSequenceFunction**
- Search in function spaces
  - Polynomials
  - Rational functions
  - Hypergeometric functions
  - Holonomic functions (linear ODEs)
- Can take symbolic terms, arbitrary precision integers, rationals
- Control over saving later terms for checking
- Time limit, default 10s, (likely) slow
- Only yields functions, not ODEs (?)

Software

Iwan Jensen's code:

- Linear ODE search
- Written in Fortran, fast
- All calculations done mod a prime, reconstruction necessary
- Input integers must be less than $2^{16} - 1$
- Not openly distributed

Software

Writing my own software:

- Written in C, very fast
- Use of arbitrary precision linear algebra package
- Search for linear and non-linear ODEs, $p_k\, y^n \left(\frac{dy}{dx}\right)^m$
- Automated search
- Control over saving later terms for checking
- Open source, `https://github.com/mike352/serintode`

Analyzing the OEIS

- Downloadable file with all “a” sequence files:
  - File updated daily, $\sim$ 24MB
  - About 300,000 sequences
  - Typical sequence length of 40 terms, up to 100
- Using my program to analyze file:
  - Takes 2 hours
  - Finds linear ODE solutions to $\sim$ 15% sequences
  - Automated LaTeX generation
  - Output pdf file has 3235 pages

Analyzing the OEIS

The OEIS stores longer versions of sequences, “b-files”:

- Contains the most number of terms known for most sequences
- No single file available for easy download
  - Neil Sloane said "No" to me
- All are accessible from the website at individualized URLs
  - Each b-file can be downloaded from its URL
  - Beware of automated download methods . . .
- Total file size zipped 5GB, unzipped 29GB

Analyzing the OEIS

The b-files:

- Some are identical to the a-file version
- Largest sequence has 1,578,730 terms
- Breakdown by size:

length	$n \leq 20$	$20 < n \leq 50$	≤ 100	≤ 500	≤ 1000	≤ 5000	≤ 10000	> 10000
% of OEIS	19.9%	21.5%	19%	15.9%	5.6%	8.3%	5.3%	4.6%

Linear ODE solutions:

- Found total of 55,431 solutions so far
- 19% of OEIS
- Breakdown by size of sequence:

length	≤100	≤200	≤300	≤500	≤900	>900
total	21392	4667	12251	1566	1815	14468+

- $\sim$ 1% have very large integers

Non-linear algebraic ODE solutions:

- Found total of 897 solutions so far
- 0.3% of OEIS
- Breakdown by size of sequence:

length	≤100	≤200	≤300	≤500	≤900	>900
total	751	55	30	61		

Analyzing the OEIS

The OEIS:

- Most are simple 1st order linear ODEs
- Most have very small, simple integer coefficients
- $\sim$20% of OEIS has solutions
- Very few are solutions of non-linear algebraic ODEs
- Possible OEIS strategy:
  - Make a searchable database for solutions of low order ODEs with polynomials of low degree

Going further

- Software can be generalized for functional equations
  For example: $p_n\, y(a_1 x + a_2 x^2 + \ldots) = y(x)$
- Putting my C program online
- Incorporation my program into the OEIS itself
- Analysis of $> 50,000$ solutions ongoing ...

Thank you!