An Analysis of OEIS Generating Functions

Michael Assis

School of Mathematical and Physical Sciences
The University of Newcastle

13 December, 2017
Online Encyclopedia of Integer Sequences

- Started in 1965 by Neil Sloane
- Originally a punched card database
- E-mail service 1994, website 1996
- User contributed in Wiki from 2010
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?
Generating functions

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: \( \frac{\phi_1^n - \phi_2^n}{\phi_1 - \phi_2} \),  
  \( \phi_1 = \frac{1 + \sqrt{5}}{2} \),  
  \( \phi_2 = \frac{1 - \sqrt{5}}{2} \)

- 1, \( -\frac{1}{2} \), \( \frac{1}{24} \), \( -\frac{1}{720} \), \( \frac{1}{40320} \), \( -\frac{1}{3628800} \), \( \frac{1}{479001600} \), \( \frac{1}{87178291200} \), ...  
  \( n \)-th term: \( \frac{(-1)^n}{(2n)!} \)

- 1, 6, 29, 130, 561, 2368, 9855, 40622, 166303, 677420, ...  
  \( n \)-th term: ?
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?
Convergence tests:

- **Ratio test**
  \[
  \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = r, \quad (r < 1) : \text{converges}
  \]

- **Root test**
  \[
  \lim_{n \to \infty} \sup n \sqrt{|a_n|} = r, \quad (r < 1) : \text{converges}
  \]

- Radius of convergence: \( \frac{1}{r} \), \( r \) from Ratio or Root tests
Analytic continuation:

1. From a series expansion around $x_0$ with radius of converge $r$:
   - Find point $x_1$ within $|x_1 - x_0| < r$ with radius outside original domain
   - Repeat

2. Find a functional equation for the function

3. Use an integral definition of the function

4. Use the ODE that the function satisfies
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:

\[ 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, \ldots \]

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

- 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:
  \[a_n = a_{n-1} + a_{n-2}, \quad a_0 = 1, \quad a_1 = 1\]
- and satisfies \(n\)-th term expression:
  \[\frac{\phi_1^n - \phi_2^n}{\phi_1 - \phi_2}\]
Example:

Treat sequence as coefficients of a power series:

\[
1, \quad -\frac{1}{2}, \quad \frac{1}{24}, \quad -\frac{1}{720}, \quad \frac{1}{40320}, \quad -\frac{1}{3628800}, \quad \frac{1}{479001600}, \quad \frac{1}{87178291200}, \quad \cdots
\]

\[1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \cdots = \cos(x)\]

- Sequence converges
- Radius of converge is $\infty$
- Easy to calculate sequence terms
Example:

Treat sequence as coefficients of a power series:

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

⇒ \[ 1x + 6x^2 + 29x^3 + 130x^4 + 561x^5 + 2368x^6 + \ldots = ? \]
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: $xy, x^2y, y', xy', x^2y'$
2. Find series: $a_0y + a_1xy + a_2x^2y + b_0y' + b_1xy' + b_2x^2y'$
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
  ⇒ The solution may fail after a certain point

- Save terms for use in checking any solution

- Not enough terms known for finding polynomial coefficients
  ⇒ No solution found until more terms are known

- The series satisfies no finite order ODE
  ⇒ At best an approximation can be found
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

Michael Assis
An Analysis of OEIS Generating Functions
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
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 (?)
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
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, ~ 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 ~ 15% sequences
  - Automated LaTeX generation
  - Output pdf file has 3235 pages
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:

<table>
<thead>
<tr>
<th>length</th>
<th>$n \leq 20$</th>
<th>$20 &lt; n \leq 50$</th>
<th>$\leq 100$</th>
<th>$\leq 500$</th>
<th>$\leq 1000$</th>
<th>$\leq 5000$</th>
<th>$\leq 10000$</th>
<th>$&gt; 10000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of OEIS</td>
<td>19.9%</td>
<td>21.5%</td>
<td>19%</td>
<td>15.9%</td>
<td>5.6%</td>
<td>8.3%</td>
<td>5.3%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>
Analyzing the OEIS

Linear ODE solutions:

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

<table>
<thead>
<tr>
<th>length</th>
<th>≤100</th>
<th>≤200</th>
<th>≤300</th>
<th>≤500</th>
<th>≤900</th>
<th>&gt;900</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>21392</td>
<td>4667</td>
<td>12251</td>
<td>1566</td>
<td>1815</td>
<td>14468+</td>
</tr>
</tbody>
</table>

- ~ 1% have very large integers
Analyzing the OEIS

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

<table>
<thead>
<tr>
<th>length</th>
<th>≤100</th>
<th>≤200</th>
<th>≤300</th>
<th>≤500</th>
<th>≤900</th>
<th>&gt;900</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>751</td>
<td>55</td>
<td>30</td>
<td>61</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The OEIS:

- Most are simple 1st order linear ODEs
- Most have very small, simple integer coefficients
- ~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
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!