1. Combinatorial structures and OGFs
Attention: Much of this lecture is a quick review of material in Analytic Combinatorics, Part I.

One consequence: it is a bit longer than usual.

To: Students who took Analytic Combinatorics, Part I

Bored because you understand it all?

GREAT! Skip to the section on labelled trees and do the exercises.

To: Students starting with Analytic Combinatorics, Part II

Moving too fast? Want to see details and motivating applications?

No problem, watch Lectures 5, 6, and 8 in Part I.
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution
Analytic combinatorics overview

To analyze properties of a large combinatorial structure:

1. Use the **symbolic method**
   - Define a **class** of combinatorial objects
   - Define a notion of **size** (and associated generating function)
   - Use standard operations to develop a **specification** of the structure

Result: A direct derivation of a **GF equation** (implicit or explicit)

Classic next steps:
   - Extract coefficients
   - Use classic asymptotics to estimate coefficients

Result: **Asymptotic estimates** that quantify the desired properties

See *An Introduction to the Analysis of Algorithms* for a gentle introduction
Analytic combinatorics overview

To analyze properties of a large combinatorial structure:

1. Use the **symbolic method**
   - Define a *class* of combinatorial objects.
   - Define a notion of *size* (and associated generating function)
   - Use standard operations to develop a *specification* of the structure.

Result: A direct derivation of a **GF equation** (implicit or explicit).

2. Use **complex asymptotics** to estimate growth of coefficients.
   - [no need for explicit solution]
   - [stay tuned for details]

Result: **Asymptotic estimates** that quantify the desired properties

See *Analytic Combinatorics* for a rigorous treatment
Analytic combinatorics overview

A. SYMBOLIC METHOD
1. OGFs
2. EGFs
3. MGFs

B. COMPLEX ASYMPTOTICS
4. Rational & Meromorphic
5. Applications of R&M
6. Singularity Analysis
7. Applications of SA
8. Saddle point
The symbolic method

An approach for *directly* deriving GF equations.

- Define a *class* of combinatorial objects.
- Define a notion of *size* (and associated generating function).
- Define *operations* suitable for constructive definitions of objects.
- Prove *correspondences* between operations and GFs.

Result: A *GF equation* (implicit or explicit).

See *An Introduction to the Analysis of Algorithms* for a gentle introduction.

See *Analytic Combinatorics* for a rigorous treatment.

**This lecture:** An overview that assumes *some* familiarity.

Ex: Part I of this course
Basic definitions

Def. A **combinatorial class** is a set of combinatorial objects and an associated **size** function.

Def. The **ordinary generating function** (OGF) associated
with a class is the formal power series

\[ A(z) = \sum_{a \in A} z^{|a|} \]

Fundamental (elementary) identity

\[ A(z) \equiv \sum_{a \in A} z^{|a|} = \sum_{N \geq 0} A_N z^N \]

Q. How many objects of size \( N \)?
A. \( A_N = [z^N]A(z) \)

With the symbolic method, we specify the class **and at the same time** characterize the OGF.
Unlabeled classes: cast of characters

**TREES**
Recursive structures
\[ T_N = [\text{Catalan #s}] \]

**STRINGS**
Sequences of characters
\[ S_N = N^M \]

**INTEGERS**
N objects
\[ I_N = 1 \]

**COMPOSITIONS**
Positive integers sum to N
\[ C_N = 2^{N-1} \]

**LANGUAGES**
Sets of strings
[REs and CFGs]

**PARTITIONS**
Unordered compositions
[enumeration not elementary]
The symbolic method (basic constructs)

Suppose that $A$ and $B$ are classes of unlabeled objects with enumerating OGFs $A(z)$ and $B(z)$.

<table>
<thead>
<tr>
<th>operation</th>
<th>notation</th>
<th>semantics</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>disjoint union</td>
<td>$A + B$</td>
<td>disjoint copies of objects from $A$ and $B$</td>
<td>$A(z) + B(z)$</td>
</tr>
<tr>
<td>Cartesian product</td>
<td>$A \times B$</td>
<td>ordered pairs of copies of objects, one from $A$ and one from $B$</td>
<td>$A(z)B(z)$</td>
</tr>
<tr>
<td>sequence</td>
<td>$SEQ(A)$</td>
<td>sequences of objects from $A$</td>
<td>$\frac{1}{1 - A(z)}$</td>
</tr>
</tbody>
</table>

Stay tuned for other constructs
Proofs of correspondences

\[ A + B \]
\[ \sum_{c \in A+B} z^{|c|} = \sum_{a \in A} z^{|a|} + \sum_{b \in B} z^{|b|} = A(z) + B(z) \]

\[ A \times B \]
\[ \sum_{c \in a \times b} z^{|c|} = \sum_{a \in A} \sum_{b \in B} z^{|a|+|b|} = \left( \sum_{a \in A} z^{|a|} \right) \left( \sum_{b \in B} z^{|b|} \right) = A(z)B(z) \]

<table>
<thead>
<tr>
<th>SEQ(A)</th>
<th>construction</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEQ_k(A)</td>
<td>( A^k )</td>
<td>( A(z)^k )</td>
</tr>
<tr>
<td>SEQ_T(A)</td>
<td>( A^{t_1} + A^{t_2} + A^{t_3} + \ldots )</td>
<td>( A(z)^{t_1} + A(z)^{t_2} + A(z)^{t_3} + \ldots )</td>
</tr>
<tr>
<td>where</td>
<td>T \equiv t_1, t_2, t_3, \ldots is a subset of the integers</td>
<td></td>
</tr>
<tr>
<td>SEQ(A)</td>
<td>( \epsilon + A + A^2 + A^3 + \ldots )</td>
<td>( 1 + A(z) + A(z)^2 + A(z)^3 + \ldots = \frac{1}{1 - A(z)} )</td>
</tr>
</tbody>
</table>
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution
Classic example of the symbolic method

**Q.** How many trees with $N$ nodes?

- $G_1 = 1$
- $G_2 = 1$
- $G_3 = 2$
- $G_4 = 5$
- $G_5 = 14$
Analytic combinatorics: How many trees with $N$ nodes?

Symbolic method

**Combinatorial class**  
$G$, the class of all trees  
$G = ullet \times \text{SEQ}(G)$  
"a tree is a node and a sequence of trees"

**Construction**  
$G(z) = z(1 + G(z) + G(z)^2 + G(z)^3 + \ldots) = \frac{z}{1 - G(z)}$

**OGF equation**  
$G(z) - G(z)^2 = z$

**Quadratic equation**  
$G(z) = \frac{1 + \sqrt{1 - 4z}}{2}$

Classic next steps

**Binomial theorem**  
$G(z) = -\frac{1}{2} \sum_{N \geq 1} \left( \frac{1}{2} \right)^N (-4z)^N$

**Extract coefficients**  
$G_N = -\frac{1}{2} \left( \frac{1}{N} \right) (-4)^N = \frac{1}{N} \left( \frac{2N - 2}{N - 1} \right) = \frac{1}{4N - 2} \left( \frac{2N}{N} \right)$

**Stirling’s approximation**  
$\sim \frac{1}{4N} \exp(2N \ln(2N) - 2N + \ln \sqrt{4\pi N} - 2(N \ln(N) - N + \ln \sqrt{2\pi N}))$

**Simplify**  
$G_N \sim \frac{4^{N-1}}{\sqrt{\pi N^3}}$
Analytic combinatorics: How many trees with $N$ nodes?

**Symbolic method**

**Combinatorial class**

$G$, the class of all trees

$G = \bullet \times \text{SEQ}(G)$

"a tree is a node and a sequence of trees"

**Construction**

$G(z) = z(1 + G(z) + G(z)^2 + G(z)^3 + \ldots) = \frac{z}{1 - G(z)}$

$G(z) - G(z)^2 = z$

**OGF equation**

**Complex asymptotics**

**Singularity analysis**

$G_N = [z^N]G(z) \sim \frac{4^N}{\Gamma(1/2)\sqrt{N}} = \frac{4^N}{\sqrt{\pi N}}$

GF equation directly implies asymptotics

**This lecture:** Focus on symbolic method for deriving OGF equations (stay tuned for asymptotics).
A standard paradigm for the symbolic method

<table>
<thead>
<tr>
<th>Fundamental constructs</th>
<th>![square]</th>
<th>![circle]</th>
<th>![triangle]</th>
</tr>
</thead>
<tbody>
<tr>
<td>• elementary or trivial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• confirm intuition</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound constructs</th>
<th>![compound_square]</th>
<th>![compound_circle]</th>
<th>![compound_triangle]</th>
</tr>
</thead>
<tbody>
<tr>
<td>• many possibilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• classical combinatorial objects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• expose underlying structure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• one of many paths to known results</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variations</th>
<th>![variations_square]</th>
<th>![variations_circle]</th>
<th>![variations_triangle]</th>
</tr>
</thead>
<tbody>
<tr>
<td>• unlimited possibilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <em>not</em> easily analyzed otherwise</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Variations on a theme 1: Trees

Fundamental construct

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>$G$, the class of all trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>$G = \bullet \times \text{SEQ}(G)$</td>
</tr>
<tr>
<td>OGF equation</td>
<td>$G(z) = z(1 + G(z) + G(z)^2 + G(z)^3 + \ldots) = \frac{z}{1 - G(z)}$</td>
</tr>
<tr>
<td></td>
<td>$G(z) - G(z)^2 = z$</td>
</tr>
</tbody>
</table>

"a tree is a node and a sequence of trees"

Variation on the theme: restrict each node to 0 or 2 children

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>$T$, the class of binary trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>$T = \bullet \times \text{SEQ}_{0,2}(T)$</td>
</tr>
<tr>
<td>OGF equation</td>
<td>$T(z) = z(1 + T(z)^2)$</td>
</tr>
</tbody>
</table>

"a binary tree is a node and a sequence of 0 or 2 binary trees"
Variations on a theme 1: Trees (continued)

Variation on the theme: multiple node types

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>$T^\bullet$, binary trees, enumerated by internal nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Atoms</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>external node</td>
</tr>
<tr>
<td></td>
<td>$\square$</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$GF$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>internal node</td>
</tr>
<tr>
<td></td>
<td>$\bullet$</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$z$</td>
</tr>
</tbody>
</table>

| Construction        | $T = \square + T \times \bullet \times T$             |

| OGF equation        | $T^\bullet(z) = 1 + zT^\bullet(z)^2$                  |

More variations: unary-binary trees, ternary trees, ...

Still more variations: gambler’s ruin sequences, context-free languages, triangulations, ...
Some variations on ordered (rooted plane) trees

**Binary**

\[ T = \bullet \times SEQ_{0,2}(T) \]

\[ T(z) = z(1 + T(z)^2) \]

**Ordered**

\[ G = \bullet \times SEQ(G) \]

\[ G(z) = \frac{z}{1 - G(z)} \]

**Unary-binary**

\[ M = \bullet \times SEQ_{\leq 2}(M) \]

\[ M(z) = z(1 + M(z) + M(z)^2) \]

**Ternary**

\[ T = \bullet \times SEQ_{0,3}(T) \]

\[ T(z) = z(1 + T(z)^3) \]

**Arbitrary restrictions**

\[ T = \bullet \times SEQ_\Omega(T) \]

\[ T^\Omega(z) = z\phi(T^\Omega(z)) \]

\[ \phi(u) = \sum_{\omega \in \Omega} u^\omega \]

\[ S = \bullet + SEQ_{\leq 2}(S) \]

\[ S(z) = z + \frac{S(z)^2}{1 - S(z)} \]
### Variation on a theme 2: Strings

#### Fundamental construct

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>B, the class of all binary strings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>( B = E + (Z_0 + Z_1) \times B )</td>
</tr>
<tr>
<td>OGF equation</td>
<td>( B(z) = 1 + 2zB(z) )</td>
</tr>
</tbody>
</table>

#### Variation on the theme: disallow sequences of \( P \) or more 0s

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>( B_P ), the class of all binary strings with no 0(^P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>( B_P = Z_{&lt;P}(E + Z_1 B_P) )</td>
</tr>
<tr>
<td>OGF equation</td>
<td>( B_P(z) = (1 + z + \ldots + z^P)(1 + zB_P(z)) )</td>
</tr>
</tbody>
</table>

#### More variations: disallow any pattern (autocorrelation), REs, CFGs ...
### Some variations on strings

<table>
<thead>
<tr>
<th>M-ary</th>
<th>Binary</th>
<th>Exclude 0&lt;sup&gt;p&lt;/sup&gt;</th>
<th>Exclude pattern p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B = SEQ(Z_0 + \ldots + Z_{M-1})$</td>
<td>$B = E + (Z_0 + Z_1) \times B$</td>
<td>$B_P = Z_{&lt;P}(E + Z_1 \times B_P)$</td>
<td>$S_P(z) = \frac{c_P(z)}{z^P + (1 - 2z)c_P(z)}$</td>
</tr>
<tr>
<td>$B(z) = \frac{1}{1 - Mz}$</td>
<td>$B(z) = \frac{1}{1 - 2z}$</td>
<td>$B_P(z) = \frac{1 - z^P}{1 - 2z + z^{P+1}}$</td>
<td>[See Part I, Lecture 8]</td>
</tr>
</tbody>
</table>

**Regular languages**

[Rational OGFs]

**Context-free languages**

[Algebraic OGFs]

**Exclude pattern**

$P$
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Power sets and multisets
- Compositions and partitions
- Substitution
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- **Powersets and multisets**
- Compositions and partitions
- Substitution
The symbolic method (two additional constructs)

Suppose that $A$ is a class of unlabeled objects with enumerating OGF $A(z)$.

<table>
<thead>
<tr>
<th>operation</th>
<th>notation</th>
<th>semantics</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>powerset</td>
<td>$PSET(A)$</td>
<td>finite sets of objects from $A$ (no repetitions)</td>
<td>[stay tuned]</td>
</tr>
<tr>
<td>multiset</td>
<td>$MSET(A)$</td>
<td>finite sets of objects from $A$ (with repetitions)</td>
<td>[stay tuned]</td>
</tr>
</tbody>
</table>
Powersets

**Def.** The *powerset* of a class A is the class consisting of all subsets of A.

- **PSET** \{a\} = \{
  \{
  \}
  \{a\}
\}

- **PSET** \{a, b\} = \{
  \{
  \}
  \{a\}
  \{b\}
  \{a, b\}
\}

- **PSET** \{a, b, c\} = \{
  \{
  \}
  \{a\}
  \{b\}
  \{c\}
  \{a, b\}
  \{a, c\}
  \{b, c\}
  \{a, b, c\}
\}

- **PSET** \{a, b, c, d\} = \{
  \{
  \}
  \{a\}
  \{b\}
  \{c\}
  \{d\}
  \{a, d\}
  \{b, d\}
  \{c, d\}
  \{a, b, d\}
  \{a, c, d\}
  \{b, c, d\}
  \{a, b, c, d\}
\}

**Lemma:** PSET \{a_1, a_2, \ldots, a_M\} = PSET \{a_1, a_2, \ldots, a_{M-1}\} \times (\{\} + \{a_M\})

**Note:**
- \(P_1 = 2\)
- \(P_2 = 4\)
- \(P_3 = 8\)
- \(P_4 = 16\)
## Powersets

**Combinatorial class**  
$P_M$, the powerset class for $M$ atoms

**Example**  
\{a, c, f, g, h\}

**OGF**  
\[ P_M(z) = \sum_{p \in P_M} z^{|p|} = \sum_{N \geq 0} P_{MN} z^N \]

**Construction**  
\[ P_M = (\{\} + \{a_1\}) \times (\{\} + \{a_2\}) \times \ldots \times (\{\} + \{a_M\}) \]

**OGF equation**  
\[ P_M(z) = (1 + z)^M \]

**Expansion**  
\[ P_{MN} = \binom{M}{N} \quad \checkmark \]

**Atoms**  
<table>
<thead>
<tr>
<th>notation</th>
<th>size</th>
<th>GF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_k$</td>
<td>1</td>
<td>$z$</td>
</tr>
</tbody>
</table>

$P_{MN}$ is the # of subsets of size $N$ (no repetitions)

$P_M(1) = 2^M$  

**total # subsets of $M$ atoms**
**Multisets**

**Def.** The *multiset* of a class $A$ is the class consisting of all subsets of $A$ *with repetitions allowed*.

**Lemma:** $\text{MSET} \{a_1, a_2, \ldots, a_M\} = \text{MSET} \{a_1, a_2, \ldots, a_{M-1}\} \times \text{SEQ} \{a_M\}$
**Multisets**

**Combinatorial class**

\(S_M\), the multiset class for \(M\) atoms

**Example**

\{a, a, a, b, b, b, c\}

**OGF**

\[S_M(z) = \sum_{s \in S_M} z^{|s|} = \sum_{N \geq 0} S_{MN} z^N\]

**Construction**

\[S_M = \text{SEQ}(a_1) \times \text{SEQ}(a_2) \times \ldots \times \text{SEQ}(a_M)\]

**OGF equation**

\[S_M(z) = \frac{1}{(1 - z)^M}\]

**Expansion**

\[S_{MN} = \binom{N + M - 1}{M - 1}\]

**Atoms**

<table>
<thead>
<tr>
<th>notation</th>
<th>size</th>
<th>GF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_k)</td>
<td>1</td>
<td>(z)</td>
</tr>
</tbody>
</table>

\(S_{MN}\) is the # of subsets of size \(N\) (with repetitions)
### The symbolic method (two additional constructs)

Suppose that $A$ is a class of unlabeled objects with enumerating OGF $A(z)$.

<table>
<thead>
<tr>
<th>operation</th>
<th>notation</th>
<th>semantics</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>powerset</strong></td>
<td>$PSET(A)$</td>
<td>finite sets of objects from $A$ (no repetitions)</td>
<td>$\prod_{n \geq 1} (1 + z^n)^{A_n} = \exp\left(- \sum_{k \geq 1} \frac{(-1)^k A(z^k)}{k}\right)$</td>
</tr>
<tr>
<td><strong>multiset</strong></td>
<td>$MSET(A)$</td>
<td>finite sets of objects from $A$ (with repetitions)</td>
<td>$\prod_{n \geq 1} \frac{1}{(1 - z^n)^{A_n}} = \exp\left(\sum_{k \geq 1} \frac{A(z^k)}{k}\right)$</td>
</tr>
</tbody>
</table>
Proof of correspondences for powersets

<table>
<thead>
<tr>
<th>$PSET(A)$</th>
<th>construction</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PSET{a, b} = ({} + {a}) \times ({} + {b})$</td>
<td>$(1 + z^{</td>
<td>a</td>
</tr>
<tr>
<td>$PSET(A) \equiv \prod_{a \in A} ({} + {a})$</td>
<td>$\prod_{a \in A} (1 + z^{</td>
<td>a</td>
</tr>
</tbody>
</table>

**exp-log version**

$$\prod_{N \geq 0} (1 + z^N)^{A_N} = \exp\left(\sum_{N \geq 0} A_N \ln(1 + z^N)\right)$$

$$= \exp\left(-\sum_{N \geq 0} A_N \sum_{k \geq 1} (-1)^k \frac{z^{Nk}}{k}\right)$$

$$= \exp\left(-\sum_{k \geq 1} (-1)^k A(z^k) \frac{k}{k}\right)$$

$$= \exp\left(A(z) - \frac{A(z^2)}{2} + \frac{A(z^3)}{3} - \ldots\right)$$
**Proof of correspondences for multisets**

<table>
<thead>
<tr>
<th>$MSET(A)$</th>
<th>construction</th>
<th>$MSET({a, b}) = SEQ({a}) \times SEQ({b})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$MSET(A) \equiv \prod_{a \in A} SEQ({a})$</td>
<td>$\prod_{a \in A} \frac{1}{z^{</td>
<td>a</td>
</tr>
</tbody>
</table>

**exp-log version**

$$\prod_{N \geq 0} \frac{1}{(1 - z^N)^{A_N}} = \exp\left(\sum_{N \geq 0} A_N \ln \frac{1}{1 - z^N}\right)$$

$$= \exp\left(\sum_{N \geq 0} A_N \sum_{k \geq 1} \frac{z^{Nk}}{k}\right)$$

$$= \exp\left(\sum_{k \geq 1} \frac{A(z^k)}{k}\right)$$

$$= \exp\left(A(z) + \frac{A(z^2)}{2} + \frac{A(z^3)}{3} + \ldots\right)$$
**Multiset application example**

**Q.** How many unordered trees with \( N \) nodes?

\[
\begin{align*}
H_1 &= 1 \\
H_2 &= 1 \\
H_3 &= 2 \\
H_4 &= 4 \\
H_5 &= 9
\end{align*}
\]

**Combinatorial class**  
\( H \), the class of all unordered trees

**Construction**  
\( H = \bullet \times MSET(H) \)

"a tree is a node and a multiset of trees"

**OGF equation**  
\( H(z) = z \exp\left(H(z) + H(z^2)/2 + H(z^3)/3 + \ldots\right) \)
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- **Powersets and multisets**
- Compositions and partitions
- Substitution
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Power sets and multisets
- Compositions and partitions
- Substitution
Compositions

Q. How many ways to express $N$ as a sum of positive integers?

A. $I_N = 2^{N-1}$

- $l_1 = 1$
- $l_2 = 2$
- $l_3 = 4$
- $l_4 = 8$
- $l_5 = 16$
## Integers as a combinatorial class

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>$I$, the class of all positive integers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>![Unary notation for 7] (●●●●●●●)</td>
</tr>
<tr>
<td>OGF</td>
<td>$I(z) = \sum_{i \in I} z^{</td>
</tr>
<tr>
<td>Construction</td>
<td>$I = SEQ_{&gt;0}(\bullet)$</td>
</tr>
<tr>
<td>OGF equation</td>
<td>$I(z) = \frac{z}{1 - z}$</td>
</tr>
<tr>
<td>Expansion</td>
<td>$I_N = 1$ for $N &gt; 0$  ✓</td>
</tr>
</tbody>
</table>

### Atom

<table>
<thead>
<tr>
<th>notation</th>
<th>size</th>
<th>GF</th>
</tr>
</thead>
<tbody>
<tr>
<td>●</td>
<td>1</td>
<td>z</td>
</tr>
</tbody>
</table>
### Compositions

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>( C ), the class of all compositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>[ \bullet</td>
</tr>
<tr>
<td>OGF</td>
<td>( C(z) = \sum_{c \in C} z^{</td>
</tr>
<tr>
<td>Construction</td>
<td>( C = SEQ(I) )</td>
</tr>
<tr>
<td>OGF equation</td>
<td>( C(z) = \frac{1}{1 - I(z)} = \frac{1}{1 - \frac{z}{1-z}} = \frac{1 - z}{1 - 2z} )</td>
</tr>
<tr>
<td>Expansion</td>
<td>( C_N = 2^N - 2^{N-1} = 2^{N-1} ) for ( N &gt; 0 )</td>
</tr>
</tbody>
</table>
Q. How many ways to express \( N \) as a sum of unordered positive integers?

A. Not so obvious!
Def. A Ferrers diagram is a 2D representation of a partition: one column of dots per part.

partition: 8 + 8 + 6 + 5 + 4 + 4 + 4 + 2 + 1 = 42

Ferrers diagram:

Q. How many Ferrers diagrams with $N$ dots?

A. Not so obvious [need symbolic method plus saddle-point asymptotics—stay tuned]

Partitions

Combinatorial class

\( P, \) the class of all partitions

Example

\[
\begin{array}{cccccccc}
& & & & & & & \\
& & & & & \bullet & & \\
& & & & \bullet & & & \\
& & & \bullet & & & & \\
& & \bullet & & & & & \\
& \bullet & & & & & & \\
\end{array}
\]

Ferrers diagram for \( 5 + 3 + 2 + 1 + 1 = 12 \)

OGF

\[
P(z) = \sum_{p \in P} z^{|p|} = \sum_{N \geq 0} P_N z^N
\]

Construction

\( P = MSET(I) \)

"a partition is a \textit{multiset} of positive integers"

OGF equation

\[
P(z) = \frac{1}{(1 - z)(1 - z^2)(1 - z^3) \ldots}
\]

Expansion

\[
P_N \sim \frac{e^{\pi \sqrt{2N/3}}}{4N^{3/2}}
\]

Classic result of Hardy and Ramanujan (need saddle-point asymptotics)
Some variations on compositions and partitions

Restricted compositions

\( T = \{ \text{any subset of } I \} \)
\( C^T = \text{SEQ} ( \text{SEQ}_T(Z) ) \)

\[
C^T(z) = \frac{1}{1 - T(z)}
\]

Compositions

\( C = \text{SEQ}(I) \)

\[
C(z) = \frac{1 - z}{1 - 2z}
\]

Compositions into \( M \) parts

\( C_M = \text{SEQ}_M(I) \)

\[
C_M(z) = \frac{z^M}{1 - z^M}
\]

Partitions into distinct parts

\( Q = \text{PSET}(I) \)

\[
Q(z) = (1 + z)(1 + z^2)(1 + z^3) \ldots
\]

Partitions

\( P = \text{MSET}(I) \)

\[
P_N \sim \frac{e^{\pi \sqrt{2N/3}}}{4N\sqrt{3}}
\]

Restricted partitions

\( T = \{ \text{any subset of } I \} \)
\( P^T = \text{MSET}(\text{SEQ}_T(Z)) \)

\[
p^T(z) = \prod_{N \in T} \frac{1}{1 - z^N}
\]
In-class exercises

Q. OGF for compositions into parts less than or equal to R?

Q. How many partitions into parts that are powers of 2?

A. \( \prod_{j \geq 0} (1 + z^{2^j}) = (1 + z)(1 + z^2)(1 + z^4)(1 + z^8) \ldots \)

A. \( \prod_{j \geq 0} (1 + z^{2^j}) = (1 + z + z^2 + z^3)(1 + z^4)(1 + z^8) \ldots \)

A. \( = (1 + z + z^2 + z^3 + z^4 + z^5 + z^6 + z^7)(1 + z^8) \ldots \)

A. \( = 1 + z + z^2 + z^3 + z^4 + z^5 + z^6 + z^7 + z^8 + z^9 + z^{10} + \ldots \)

Q. How many ways to represent an integer as a sum of powers of 2?

A. \( \prod_{j \geq 0} (1 + z^{2^j}) = \frac{1}{1 - z} \)
How many ways to change a dollar?

Q. How many ways to change a dollar with quarters?

A. 1

\[
\left[ z^{100} \right] \frac{1}{1 - z^{25}} = \left[ z^{100} \right] (1 + z^{25} + z^{50} + \ldots) = 1
\]

Q. How many ways to change a dollar with quarters and dimes?

A. 3

\[
\left[ z^{100} \right] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} = \left[ z^{100} \right] (1 + z^{25} + z^{50} + \ldots) (1 + z^{10} + z^{20} + \ldots) = \left[ z^{100} \right] (1 + z^{50} + z^{100}) (1 + z^{50} + z^{100})
\]
How many ways to change a dollar?

**Q.** How many ways to change a dollar with quarters?

**A.** 1

\[
[Z^{100}] \frac{1}{1 - z^{25}} = [Z^{100}](1 + z^{25} + z^{50} + \ldots) = 1
\]

**Q.** How many ways to change a dollar with quarters and dimes?

**A.** 3

\[
[Z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} = [Z^{100}](1 + z^{25} + z^{50} + \ldots)(1 + z^{10} + z^{20} + \ldots)
\]

**Q.** How many ways to change a dollar with quarters, dimes and nickels?

**A.**?

\[
[Z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} \frac{1}{1 - z^{5}} = [Z^{100}](1 + z^{25} + z^{50} + \ldots)(1 + z^{10} + z^{20} + \ldots)(1 + z^{5} + \ldots)
\]

**Q.** How many ways to change a dollar with quarters, dimes, nickels and pennies?

**A.**?

\[
[Z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} \frac{1}{1 - z^{5}} \frac{1}{1 - z} = [Z^{100}](1 + z^{25} + z^{50} + \ldots)(1 + z^{10} + z^{20} + \ldots)(1 + z^{5} + \ldots)(1 + z + \ldots)
\]

Do you need a computer?
How many ways to change a dollar?

Key insight (Pólya): If \( b(z) = a(z) \frac{1}{1 - z^M} \) then \( b(z)(1 - z^M) = a(z) \) and therefore \( b_n = b_{n-M} + a_n \)

Gives an easy way to compute small values by hand.
In-class exercise

For whatever reason, the government switches to 20-cent pieces instead of dimes.

How many ways to change a dollar?

\[
\begin{array}{cccccccccccccccc}
0 & 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 & 45 & 50 & 55 & 60 & 65 & 70 & 75 & 80 & 85 & 90 & 95 & 100 \\
\frac{1}{1-z} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\frac{1}{1-z} \frac{1}{1-z^5} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 \\
\frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}} & 1 & 2 & 3 & 4 & 6 & 8 & 10 & 12 & 15 & 18 & 21 & 24 & 28 & 32 & 36 & 40 & 45 & 50 & 55 & 60 & 66 \\
\frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}} \frac{1}{1-z^{25}} & 1 & 9 & 30 & 70 & 136 \\
\end{array}
\]
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Power sets and multisets
- Compositions and partitions
- Substitution
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution
The symbolic method for unlabeled objects (summary)

<table>
<thead>
<tr>
<th>operation</th>
<th>notation</th>
<th>semantics</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>disjoint union</td>
<td>$A + B$</td>
<td>disjoint copies of objects from $A$ and $B$</td>
<td>$A(z) + B(z)$</td>
</tr>
<tr>
<td>Cartesian product</td>
<td>$A \times B$</td>
<td>ordered pairs of copies of objects, one from $A$ and one from $B$</td>
<td>$A(z)B(z)$</td>
</tr>
<tr>
<td>sequence</td>
<td>$SEQ(A)$</td>
<td>sequences of objects from $A$</td>
<td>$\frac{1}{1 - A(z)}$</td>
</tr>
<tr>
<td>powerset</td>
<td>$PSET(A)$</td>
<td>finite sets of objects from $A$ (no repetitions)</td>
<td>$\prod_{n \geq 1} (1 + z^n)^{A_n} = \exp \left( - \sum_{k \geq 1} \frac{(-1)^k A(z^k)}{k} \right)$</td>
</tr>
<tr>
<td>multiset</td>
<td>$MSET(A)$</td>
<td>finite sets of objects from $A$ (with repetitions)</td>
<td>$\prod_{n \geq 1} \frac{1}{(1 - z^n)^{A_n}} = \exp \left( \sum_{k \geq 1} \frac{A(z^k)}{k} \right)$</td>
</tr>
</tbody>
</table>

Additional constructs are available (and still being invented)—one example to follow
Another construct for the symbolic method: substitution

Suppose that $A$ and $B$ are classes of unlabeled objects with enumerating OGFs $A(z)$ and $B(z)$.

<table>
<thead>
<tr>
<th>operation</th>
<th>notation</th>
<th>semantics</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>substitution</td>
<td>$A \circ [B]$</td>
<td>replace each object in an instance of $A$ with an object from $B$</td>
<td>$A(B(z))$</td>
</tr>
</tbody>
</table>
Q. How many 2-3 trees with $N$ nodes?

- $W_2 = 1$
- $W_3 = 1$
- $W_4 = 1$
- $W_5 = 2$
- $W_6 = 2$
- $W_7 = 3$
- $W_8 = 4$
Substitution application example

Q. How many 2-3 trees with $N$ nodes?

<table>
<thead>
<tr>
<th>Combinatorial class</th>
<th>$W$, the class of all 2-3 trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>$W = Z + W \circ [ (Z \times Z) + (Z \times Z \times Z) ]$ “a 2-3 tree is a 2-3 tree with each external node replaced by a 2-node or a 3-node”</td>
</tr>
<tr>
<td>OGF equation</td>
<td>$W(z) = z + W(z^2 + z^3)$</td>
</tr>
</tbody>
</table>

$W(z) = z^2 + z^3 + z^4 + 2z^5 + 2z^6 + 3z^7 + 4z^8 + \ldots$

$W(z^2 + z^3) = z^2 + z^3 + (z^2 + z^3)^2 + (z^2 + z^3)^3 + (z^2 + z^3)^4 + \ldots$

$= z^2 + z^3 + (z^4 + 2z^5 + z^6) + (z^6 + 3z^7 + 3z^8 + z^9) + z^8 + \ldots \checkmark$

Coefficient asymptotics are complicated (oscillations in the leading term).

Two French mathematicians on the utility of GFs

“Generating functions are the central objects of the theory, rather than a mere artifact to solve recurrences, as it is still often believed.”

— Philippe Flajolet, 2007

“A property... is understood better, when one constructs a bijection... than when one calculates the coefficients of a polynomial whose variables have no particular meaning. The method of generating functions, which has had devastating effects for a century, has fallen into obsolescence, for this reason.”

— Claude Bergé, 1968
Analytic combinatorics overview

To analyze properties of a large combinatorial structure:

1. Use the **symbolic method**
   - Define a *class* of combinatorial objects.
   - Define a notion of *size* (and associated generating function)
   - Use standard operations to develop a *specification* of the structure.
   Result: A direct derivation of a *GF equation* (implicit or explicit).

2. Use **complex asymptotics** to estimate growth of coefficients (stay tuned).

*Important note: GF equations vary widely in nature*

\[
P(z) = \frac{1}{(1 - z)(1 - z^2)(1 - z^3) \cdots} \quad C(z) = \frac{1}{1 - I(z)} \quad T(z) = z + T(z^2 + z^3) \quad B(z) = \frac{1}{1 - 2z}
\]

\[
S_M(z) = \frac{1}{(1 - z)^M} \quad H(z) = z \exp\left( H(z) + H(z^2)/2 + H(z^3/3 + \cdots) \right)
\]

\[
B_P(z) = \frac{1 - z^p}{1 - 2z + z^{p+1}} \quad G(z)^2 - G(z) + z = 0 \quad Q(z) = (1 + z)(1 + z^2)(1 + z^3) \cdots
\]
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Power sets and multisets
- Compositions and partitions
- Substitution
1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution
- Exercises
Note 1.23

Alice, Bob, and coding bounds

\[ I.23. \text{Alice, Bob, and coding bounds.} \] Alice wants to communicate \( n \) bits of information to Bob over a channel (a wire, an optic fibre) that transmits 0, 1-bits but is such that any occurrence of 11 terminates the transmission. Thus, she can only send on the channel an encoded version of her message (where the code is of some length \( \ell \geq n \)) that does not contain the pattern 11.

Here is a first coding scheme: given the message \( m = m_1 m_2 \cdots m_n \), where \( m_j \in \{0, 1\} \), apply the substitution: \( 0 \mapsto 00 \) and \( 1 \mapsto 10 \); terminate the transmission by sending 11. This scheme has \( \ell = 2n + O(1) \), and we say that its rate is 2. Can one design codes with better rates? with rates arbitrarily close to 1, asymptotically?

Let \( C \) be the class of allowed code words. For words of length \( n \), a code of length \( L = L(n) \) is achievable only if there exists a one-to-one mapping from \( \{0, 1\}^n \) into \( \bigcup_{j=0}^{L} C_j \), i.e.,
\[ 2^n \leq \sum_{j=0}^{L} C_j. \]
Working out the OGF of \( C \), one finds that necessarily
\[ L(n) \geq \lambda n + O(1), \quad \lambda = \frac{1}{\log_2 \varphi} \approx 1.440420, \quad \varphi = \frac{1 + \sqrt{5}}{2}. \]
Thus no code can achieve a rate better than 1.44; i.e., a loss of at least 44% is unavoidable.
Note 1.43

Calculating Cayley numbers and partition numbers

> 1.43. *Fast determination of the Cayley–Pólya numbers.* Logarithmic differentiation of $H(z)$ provides for the $H_n$ a recurrence by which one computes $H_n$ in time polynomial in $n$. (Note: a similar technique applies to the partition numbers $P_n$; see p. 42.)
Assignments

1. Read pages 15-94 in text.

2. Write up solutions to Notes 1.23 and 1.43.


**Program I.1.** Determine the choice of four coins that maximizes the number of ways to change a dollar.

**Program I.2.** Write programs that estimate the rate of growth of the Cayley numbers and the partition numbers \(H_n/H_{n-1}\) and \(P_n/P_{n-1}\). See Note I.43.
1. Combinatorial structures and OGFs