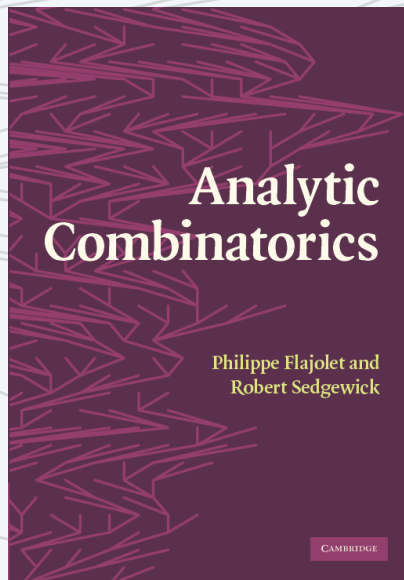


ANALYTIC COMBINATORICS

PART TWO



<http://ac.cs.princeton.edu>

# 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.

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

## 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**.

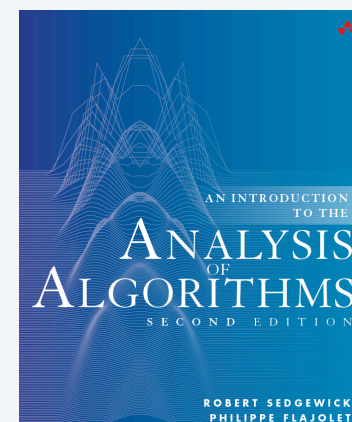
- Define a *class* of combinatorial objects.
- Define a notion of *size* (and associated generating function).
- Use standard operations to develop a *specification* of the structure.

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.



<http://aofa.cs.princeton.edu>

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.

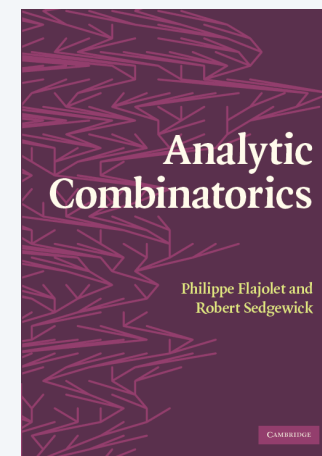
Result: A direct derivation of a **GF equation** (implicit or explicit).

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.



<http://ac.cs.princeton.edu>

# Analytic combinatorics overview

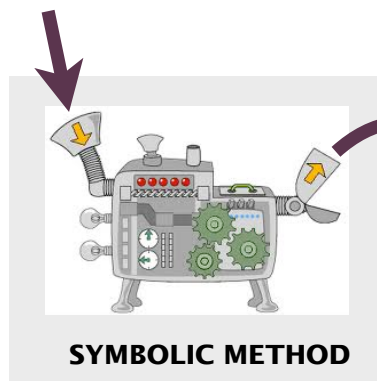
## A. SYMBOLIC METHOD

- ➔ 1. OGFs
- 2. EGFs
- 3. MGFs

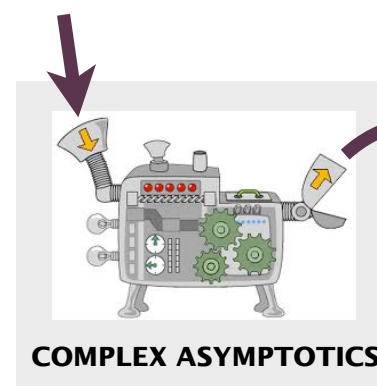
## B. COMPLEX ASYMPTOTICS

- 4. Rational & Meromorphic
- 5. Applications of R&M
- 6. Singularity Analysis
- 7. Applications of SA
- 8. Saddle point

specification

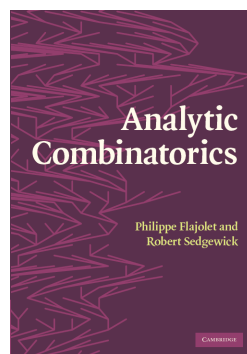


GF  
equation



asymptotic  
estimate

desired  
result !



## The symbolic method

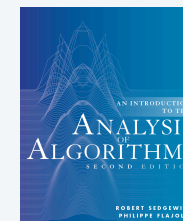
---

An approach for *directly* deriving GF equations.

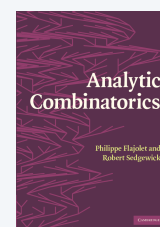
- Define a *class* of combinatorial objects.
- Define a notion of *size* (and associated generating function).
- Define *operations* suitable for constructive definitions of objects.
- Prove *correspondences* between operations and GFs.

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  $A(z) = \sum_{a \in A} z^{|a|}$  ← size function

object name ↗
↖ class name

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.**  $A_N = [z^N]A(z)$

*Fantasy:*  
Different letter for each class

*Reality:*  
Only 26 letters!

*Usual conventions*

class name	roman	$A$
OGF name	roman with arg	$A(z)$
object variable	lowercase	$a$
coefficient	subscripted	$A_N$
size	$N$ or $n$	

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  $B(z)$ .

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

Stay tuned for other constructs.

## Proofs of correspondences

---

$A + B$

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

$A \times B$

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

$SEQ(A)$	construction	OGF
	$SEQ_k(A) \equiv A^k$	$A(z)^k$
	$SEQ_T(A) \equiv A^{t_1} + A^{t_2} + A^{t_3} + \dots$ where $T \equiv t_1, t_2, t_3, \dots$ is a subset of the integers	$A(z)^{t_1} + A(z)^{t_2} + A(z)^{t_3} + \dots$
	$SEQ(A) \equiv \epsilon + A + A^2 + A^3 + \dots$	$1 + A(z) + A(z)^2 + A(z)^3 + \dots = \frac{1}{1 - A(z)}$

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- Substitution



Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

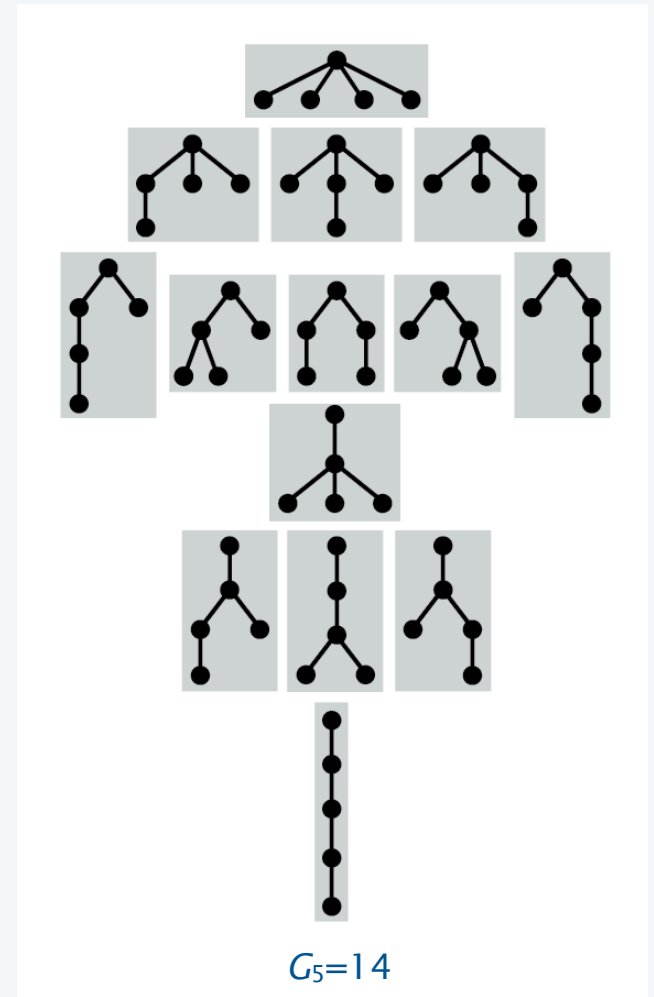
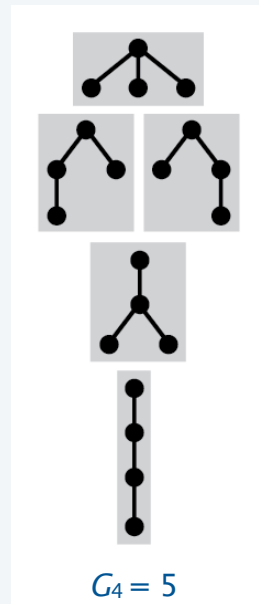
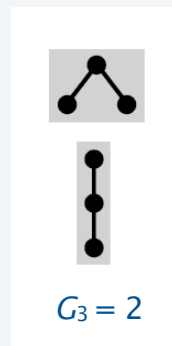
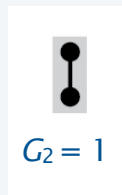
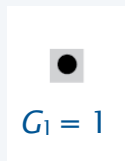
<http://ac.cs.princeton.edu>

# 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 ordered **trees** with  $N$  nodes?



# Analytic combinatorics: How many ordered trees with $N$ nodes?

## Symbolic method

Combinatorial class

Construction

OGF equation

Quadratic equation

Classic next steps

Binomial theorem

Extract coefficients

Stirling's approximation

Simplify

$G$ , the class of all trees

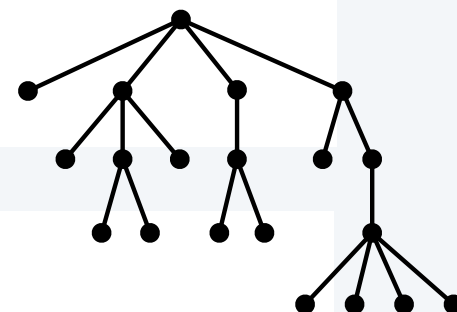
$$G = Z \times \text{SEQ}(G)$$

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

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

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

$$G(z) = \frac{1 - \sqrt{1 - 4z}}{2}$$



$$G(z) = -\frac{1}{2} \sum_{N \geq 1} \binom{\frac{1}{2}}{N} (-4z)^N$$

$$G_N = -\frac{1}{2} \binom{\frac{1}{2}}{N} (-4)^N = \frac{1}{N} \binom{2N-2}{N-1} = \frac{1}{4N-2} \binom{2N}{N}$$

detailed calculations omitted

$$\sim \frac{1}{4N} \exp(2N \ln(2N) - 2N + \ln \sqrt{4\pi N} - 2(N \ln(N) - N + \ln \sqrt{2\pi N}))$$

$$G_N \sim \frac{4^{N-1}}{\sqrt{\pi N^3}}$$

# Analytic combinatorics: How many ordered trees with $N$ nodes?

## Symbolic method

Combinatorial class

$G$ , the class of all trees

Construction

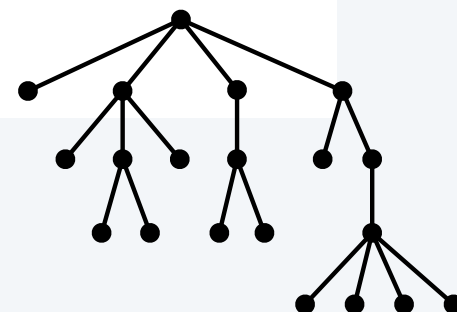
$$G = Z \times \text{SEQ}(G)$$

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

OGF equation

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

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



## 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

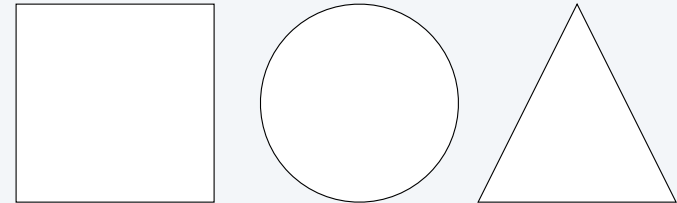
**This lecture:** Focus on symbolic method for deriving OGF equations (stay tuned for asymptotics).

## A standard paradigm for the symbolic method

---

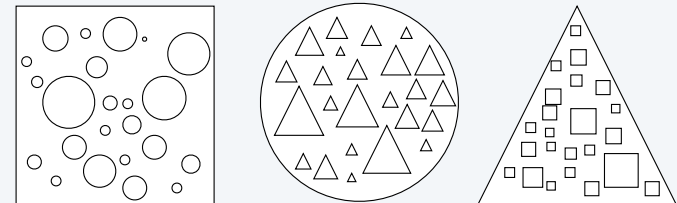
### Fundamental constructs

- elementary or trivial
- confirm intuition



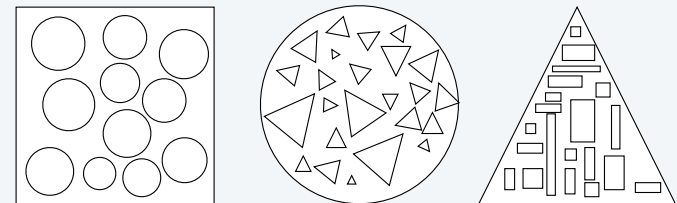
### Compound constructs

- many possibilities
- classical combinatorial objects
- expose underlying structure
- one of many paths to known results



### Variations

- unlimited possibilities
- *not* easily analyzed otherwise



## Variations on a theme 1: Trees

### Fundamental construct

Combinatorial class

$G$ , the class of all trees

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

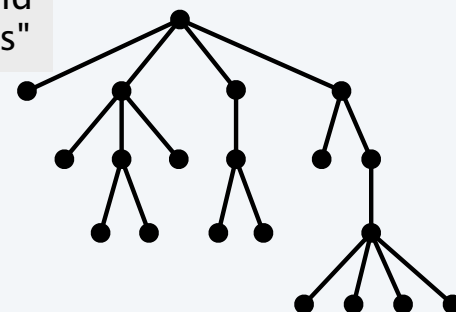
Construction

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

OGF equation

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

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



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

Combinatorial class

$T$ , the class of binary trees

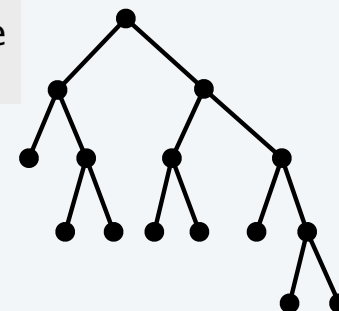
"a binary tree is a node and a sequence of 0 or 2 binary trees"

Construction

$$T = Z \times \text{SEQ}_{0,2}(T)$$

OGF equation

$$T(z) = z(1 + T(z)^2)$$



## Variations on a theme 1: Trees (continued)

Variation on the theme: multiple node types

Combinatorial class

$T^\bullet$ , binary trees, *enumerated by internal nodes*

Atoms

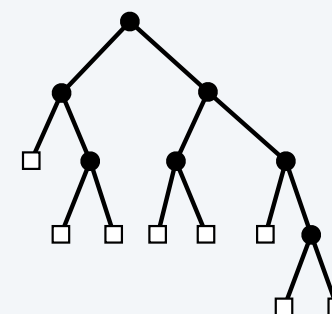
type	class	size	GF
external node	$\varepsilon$	0	1
internal node	$Z$	1	$z$

Construction

$$T = \varepsilon + T \times Z \times T$$

OGF equation

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



Combinatorial class

$T^\square$ , binary trees, *enumerated by external nodes*

OGF equation

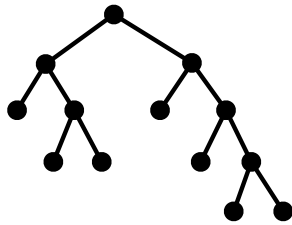
$$T^\square(z) = z + T^\square(z)^2$$

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

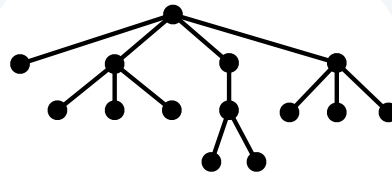
Binary



$$T = Z \times \text{SEQ}_{0,2}(T)$$

$$T(z) = z(1 + T(z)^2)$$

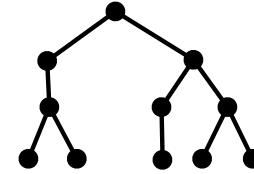
Ordered



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

$$G(z) = \frac{z}{1 - G(z)}$$

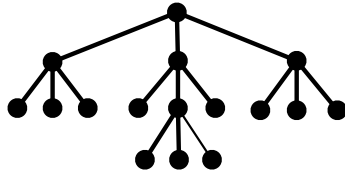
Unary-binary



$$M = Z \times \text{SEQ}_{\leq 2}(M)$$

$$M(z) = z(1 + M(z) + M(z)^2)$$

Ternary



$$T = Z \times \text{SEQ}_{0,3}(T)$$

$$T(z) = z(1 + T(z)^3)$$

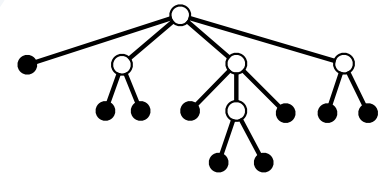
Arbitrary restrictions

$$T^{\Omega} = Z \times \text{SEQ}_{\Omega}(T^{\Omega})$$

$$T^{\Omega}(z) = z\phi(T^{\Omega}(z))$$

$$\phi(u) \equiv \sum_{\omega \in \Omega} u^{\omega}$$

Bracketings



$$S = Z + \text{SEQ}_{\geq 2}(S)$$

$$S(z) = z + \frac{S(z)^2}{1 - S(z)}$$



## Variation on a theme 2: Strings

---

### Fundamental construct

**Combinatorial class**  $B$ , the class of all binary strings

**Construction**  $B = \varepsilon + (Z_0 + Z_1) \times B$

“a binary string is empty or a bit followed by a binary string”

**OGF equation**  $B(z) = 1 + 2zB(z)$

### Variation on the theme: *disallow sequences of $P$ or more 0s*

**Combinatorial class**  $B_P$ , the class of all binary strings with no  $0^P$

**Construction**  $B_P = Z_{<P}(\varepsilon + Z_1 B_P)$

“a string with no  $0^P$  is a string of 0s of length  $<P$  followed by an empty string or a 1 followed by a string with no  $0^P$ ”

**OGF equation**  $B_P(z) = (1 + z + \dots + z^{P-1})(1 + zB_P(z))$

More variations: disallow any pattern (autocorrelation), REs, CFGs ...

## Some variations on strings

---

*M*-ary

$$B = \text{SEQ}(Z_0 + \dots + Z_{M-1})$$

$$B(z) = \frac{1}{1 - Mz}$$

Binary

$$B = \epsilon + (Z_0 + Z_1) \times B$$

$$B = \text{SEQ}(Z_0 + Z_1)$$

$$B(z) = \frac{1}{1 - 2z}$$

Exclude  $0^P$

$$B_P = Z_{<P}(\epsilon + Z_1 \times B_P)$$

$$B_P(z) = \frac{1 - z^P}{1 - 2z + z^{P+1}}$$

Regular languages

[Rational OGFs]

Context-free languages

[Algebraic OGFs]

Exclude pattern  $p$

$$S_p(z) = \frac{c_p(z)}{z^P + (1 - 2z)c_p(z)}$$

[See Part I, Lecture 8]

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- **Trees and strings**
- Powersets and multisets
- Compositions and partitions
- Substitution

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 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  $A(z)$ .

operation	notation	semantics	OGF
<i>powerset</i>	$PSET(A)$	finite sets of objects from $A$ (no repetitions)	[stay tuned]
<i>multiset</i>	$MSET(A)$	finite sets of objects from $A$ (with repetitions)	[stay tuned]

## Powersets

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

PSET  $\{a\}$

$\{\}$   
 $\{a\}$

$$P_1 = 2$$

PSET  $\{a, b\}$

$\{\}$   
 $\{a\}$   
 $\{b\}$   
 $\{a, b\}$

$$P_2 = 4$$

PSET  $\{a, b, c\}$

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

$$P_3 = 8$$

PSET  $\{a, b, c, d\}$

$\{\}$	$\{d\}$
$\{a\}$	$\{a, d\}$
$\{b\}$	$\{b, d\}$
$\{a, b\}$	$\{a, b, d\}$
$\{c\}$	$\{c, d\}$
$\{a, c\}$	$\{a, c, d\}$
$\{b, c\}$	$\{b, c, d\}$
$\{a, b, c\}$	$\{a, b, c, d\}$

↑  
subsets  
without  $d$

↑  
same subsets  
with  $d$

$$P_4 = 16$$

Lemma:  $\text{PSET } \{a_1, a_2, \dots, a_M\} = \text{PSET } \{a_1, a_2, \dots, a_{M-1}\} \times (\{\} + \{a_M\})$

# Powersets

Combinatorial class

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

Atoms

notation	size	GF
$a_k$	1	$z$

Example

$\{a, c, f, g, h\}$

OGF

$$P_M(z) = \sum_{p \in P_M} z^{|p|} = \sum_{N \geq 0} P_{MN} z^N$$

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

Construction

$$P_M = ( \{ \} + \{a_1\} ) \times ( \{ \} + \{a_2\} ) \times \dots \times ( \{ \} + \{a_M\} )$$

OGF equation

$$P_M(z) = (1 + z)^M$$

Expansion

$$P_{MN} = \binom{M}{N} \quad \checkmark$$

$$P_M(1) = 2^M \quad \checkmark$$

↑  
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*.

	MSET {a, b}	MSET {a, b, c}		
MSET {a}	{} {a} {a, a} {a, a, a}	{} {a} {a, a} {a, a, a}	{c} {a, c} {a, a, c} {a, a, a, c}	{c, c} {a, c, c} {a, a, c, c} {a, a, a, c, c}
	{b} {a, b} {a, a, b} {a, a, a, b}	{b} {a, b} {a, a, b} {a, a, a, b}	{b, c} {a, b, c} {a, a, b, c} {a, a, a, b, c}	{b, c, c} {a, b, c, c} {a, a, b, c, c} {a, a, a, b, c, c}
	{b, b} {a, b, b} {a, a, b, b} {a, a, a, b, b}	{b, b} {a, b, b} {a, a, b, b} {a, a, a, b, b}	{b, b, b, c} {a, b, b, b, c} {a, a, b, b, b, c} {a, a, a, b, b, b, c}	{b, b, c, c} {a, b, b, c, c} {a, a, b, b, c, c} {a, a, a, b, b, c, c}

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



# Multisets

## Atoms

notation	size	GF
$a_k$	1	$z$

Combinatorial class

$S_M$ , the multiset class for  $M$  atoms

Example

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

OGF

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

$S_{MN}$  is the # of subsets of size  $N$  (with repetitions)

Construction

$$S_M = SEQ(a_1) \times SEQ(a_2) \times \dots \times SEQ(a_M)$$

OGF equation

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

Expansion

$$S_{MN} = \binom{N + M - 1}{M - 1} \checkmark$$

## The symbolic method (two additional constructs)

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

operation	notation	semantics	OGF
<i>powerset</i>	$PSET(A)$	finite sets of objects from $A$ (no repetitions)	$\prod_{n \geq 0} (1 + z^n)^{A_n} = \exp\left(-\sum_{k \geq 1} \frac{(-1)^k A(z^k)}{k}\right)$
<i>multiset</i>	$MSET(A)$	finite sets of objects from $A$ (with repetitions)	$\prod_{n \geq 0} \frac{1}{(1 - z^n)^{A_n}} = \exp\left(\sum_{k \geq 1} \frac{A(z^k)}{k}\right)$

Note: For multisets,  $A_0$  must be 0; then  $n$  can start from 1 instead of 0.

## Proof of correspondences for powersets

$PSET(A)$	construction	OGF
	$PSET(\{a, b\}) = (\{\} + \{a\}) \times (\{\} + \{b\})$	$(1 + z^{ a })(1 + z^{ b })$
	$PSET(A) \equiv \prod_{a \in A} (\{\} + \{a\})$	$\prod_{a \in \mathcal{A}} (1 + z^{ a }) = \prod_{N \geq 0} (1 + z^N)^{A_N}$

**exp-log version**

$$\begin{aligned}
 \prod_{N \geq 0} (1 + z^N)^{A_N} &= \exp\left(\sum_{N \geq 0} A_N \ln(1 + z^N)\right) \\
 &= \exp\left(-\sum_{N \geq 0} A_N \sum_{k \geq 1} (-1)^k \frac{z^{Nk}}{k}\right) \\
 &= \exp\left(-\sum_{k \geq 1} (-1)^k \frac{A(z^k)}{k}\right) \\
 &= \exp\left(A(z) - \frac{A(z^2)}{2} + \frac{A(z^3)}{3} - \dots\right)
 \end{aligned}$$

## Proof of correspondences for multisets

$MSET(A)$

construction

OGF

$$MSET(\{a, b\}) = SEQ(\{a\}) \times SEQ(\{b\})$$

$$\frac{1}{(1 - z^{|a|})(1 - z^{|b|})}$$

$$MSET(A) \equiv \prod_{a \in A} SEQ(\{a\})$$

$$\prod_{a \in A} \frac{1}{(1 - z^{|a|})} = \prod_{N \geq 0} \frac{1}{(1 - z^N)^{A_N}}$$

exp-log version

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

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

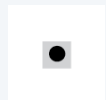
$$= \exp\left(\sum_{k \geq 1} \frac{A(z^k)}{k}\right)$$

$$= \exp\left(A(z) + \frac{A(z^2)}{2} + \frac{A(z^3)}{3} + \dots\right)$$

Note:  $A_0$  must be zero.

## Multiset application example

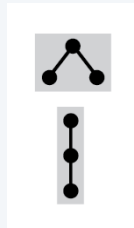
Q. How many **unordered** trees with  $N$  nodes?



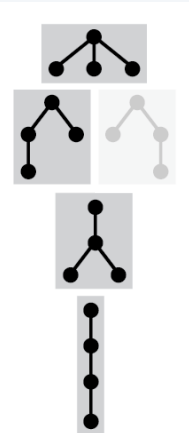
$H_1 = 1$



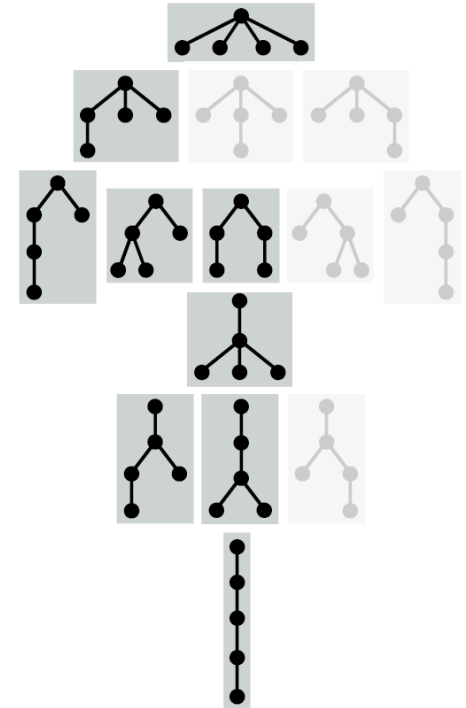
$H_2 = 1$



$H_3 = 2$



$H_4 = 4$



$H_5 = 9$

**Combinatorial class**  $H$ , the class of all unordered trees

**Construction**

$$H = Z \times MSET(H)$$

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

**OGF equation**

$$H(z) = z \exp(H(z) + H(z^2)/2 + H(z^3)/3 + \dots)$$

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- **Powersets and multisets**
- Compositions and partitions
- Substitution

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- **Compositions and partitions**
- Substitution

## Compositions

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

$$\begin{array}{c} 1 \\ l_1 = 1 \end{array}$$

$$\begin{array}{c} 1 + 1 \\ 2 \\ l_2 = 2 \end{array}$$

$$\begin{array}{c} 1 + 1 + 1 \\ 1 + 2 \\ 2 + 1 \\ 3 \\ l_3 = 4 \end{array}$$

$$\begin{array}{c} 1 + 1 + 1 + 1 \\ 1 + 1 + 2 \\ 1 + 2 + 1 \\ 1 + 3 \\ 2 + 1 + 1 \\ 2 + 2 \\ 3 + 1 \\ 4 \\ l_4 = 8 \end{array}$$

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

$$\begin{array}{c} 1 + 1 + 1 + 1 + 1 \\ 1 + 1 + 1 + 2 \\ 1 + 1 + 2 + 1 \\ 1 + 1 + 3 \\ 1 + 2 + 1 + 1 \\ 1 + 2 + 2 \\ 1 + 3 + 1 \\ 1 + 4 \\ 2 + 1 + 1 + 1 \\ 2 + 1 + 2 \\ 2 + 2 + 1 \\ 2 + 3 \\ 3 + 1 + 1 \\ 3 + 2 \\ 4 + 1 \\ 5 \\ l_5 = 16 \end{array}$$



## Integers as a combinatorial class

Atom	notation	size	GF
	$z$	1	$z$

Combinatorial class  $I$ , the class of all positive integers

Example

• • • • • ← unary notation for 7

OGF

$$I(z) = \sum_{i \in I} z^{|i|} = \sum_{N \geq 0} I_N z^N$$

Construction

$$I = \text{SEQ}_{>0}(Z)$$

OGF equation

$$I(z) = \frac{z}{1 - z}$$

Expansion

$$I_N = 1 \text{ for } N > 0 \quad \checkmark$$

# Compositions

Combinatorial class

$\mathcal{C}$ , the class of all compositions

Example

$\bullet \mid \bullet \bullet \bullet \mid \bullet \mid \bullet \bullet \bullet \bullet \mid \bullet \bullet = \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet$

← unary notation for  
 $1 + 3 + 1 + 5 + 2 = 12$

OGF

$$C(z) = \sum_{c \in \mathcal{C}} z^{|c|} = \sum_{N \geq 0} C_N z^N$$

Construction

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

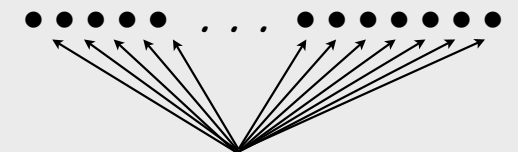
← "a composition is a sequence  
of positive integers"

OGF equation

$$\begin{aligned} C(z) &= \frac{1}{1 - I(z)} \\ &= \frac{1}{1 - \frac{z}{1-z}} = \frac{1-z}{1-2z} \end{aligned}$$

Expansion

$$C_N = 2^N - 2^{N-1} = \underbrace{2^{N-1}}_{\checkmark} \text{ for } N > 0$$



$N-1$  spaces between dots  
each could have a bar or not  
 $= 2^{N-1}$  possibilities ✓

## Partitions

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

$$\begin{array}{c} 1 \\ P_1 = 1 \end{array}$$

$$\begin{array}{c} 1 + 1 \\ 2 \\ P_2 = 2 \end{array}$$

$$\begin{array}{c} 1 + 1 + 1 \\ 1 + 2 \\ 2 + 1 \\ 3 \\ P_3 = 3 \end{array}$$

$$\begin{array}{c} 1 + 1 + 1 + 1 \\ 1 + 1 + 2 \\ 1 + 2 + 1 \\ 1 + 3 \\ 2 + 1 + 1 \\ 2 + 2 \\ 3 + 1 \\ 4 \\ P_4 = 5 \end{array}$$

$$\begin{array}{c} 1 + 1 + 1 + 1 + 1 \\ 1 + 1 + 1 + 2 \\ 1 + 1 + 2 + 1 \\ 1 + 1 + 3 \\ 1 + 2 + 1 + 1 \\ 1 + 2 + 2 \\ 1 + 3 + 1 \\ 1 + 4 \\ 2 + 1 + 1 + 1 \\ 2 + 1 + 2 \\ 2 + 2 + 1 \\ 2 + 3 \\ 3 + 1 + 1 \\ 3 + 2 \\ 4 + 1 \\ 5 \\ P_5 = 7 \end{array}$$

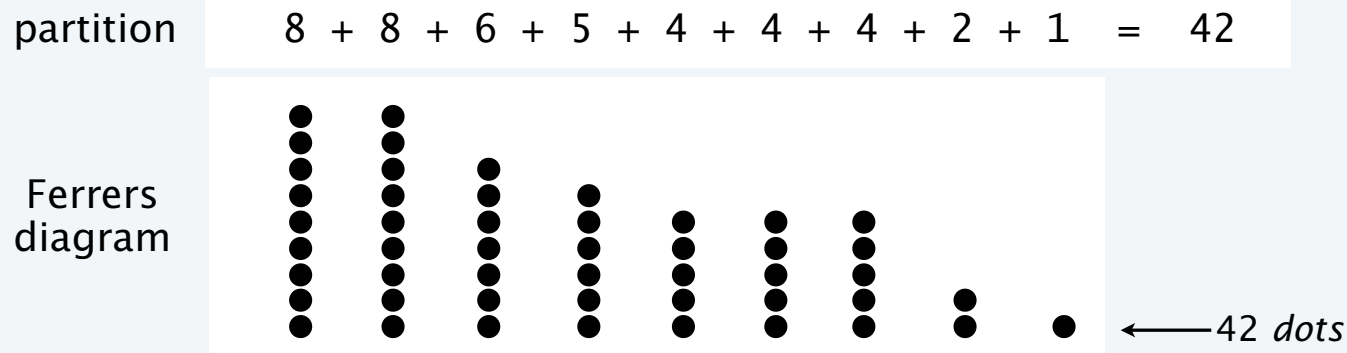
representations  
of the same  
partition

keep the one  
whose parts  
are nonincreasing

A. Not so obvious !

## Ferrers diagrams

**Def.** A *Ferrers diagram* is a 2D representation of a partition: one column of dots per part.



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

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

Applications: analysis of algorithms, representation theory, Lie algebras, particle physics.

# Partitions

Combinatorial class  $\mathcal{P}$ , the class of all partitions

Example



Ferrers diagram for  
 $5+3+2+1+1=12$

OGF

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

Construction

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

"a partition is a *multiset* of positive integers"

OGF equation

$$P(z) = \frac{1}{(1-z)(1-z^2)(1-z^3)\dots}$$

$$\begin{aligned} \text{MSET}(A) &\equiv \prod_{a \in A} \text{SEQ}(\{a\}) \\ \prod_{a \in A} \frac{1}{(1-z^{|a|})} &= \prod_{N \geq 0} \frac{1}{(1-z^N)^{A_N}} \end{aligned}$$

Expansion

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

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

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

$$Q(z) = (1 + z)(1 + z^2)(1 + z^3) \dots$$

### 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. 1

$$\begin{aligned}\prod_{j \geq 0} (1 + z^{2^j}) &= (1 + z)(1 + z^2)(1 + z^4)(1 + z^8) \dots \\ &= (1 + z + z^2 + z^3)(1 + z^4)(1 + z^8) \dots \\ &= (1 + z + z^2 + z^3 + z^4 + z^5 + z^6 + z^7)(1 + z^8) \dots \\ &= 1 + z + z^2 + z^3 + z^4 + z^5 + z^6 + z^7 + z^8 + z^9 + z^{10} + \dots\end{aligned}$$

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

A. 1

$$\prod_{j \geq 0} (1 + z^{2^j}) = \frac{1}{1 - z}$$

## How many ways to change a dollar?

---

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

A. **1**

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



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

A. **3**

$$\begin{aligned} [z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} &= [z^{100}] (1 + z^{25} + z^{50} + \dots)(1 + z^{10} + z^{20} + \dots) \\ &= [z^{100}] (1 + z^{50} + z^{100})(1 + z^{50} + z^{100}) \end{aligned}$$





## How many ways to change a dollar?

---

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

A. **1**  $[z^{100}] \frac{1}{1 - z^{25}} = [z^{100}] (1 + z^{25} + z^{50} + \dots) = 1$

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

A. **3**  $[z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} = [z^{100}] (1 + z^{25} + z^{50} + \dots)(1 + z^{10} + z^{20} + \dots)$

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

A. ?  $[z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} \frac{1}{1 - z^5}$  ← need a computer?

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

A. ?  $[z^{100}] \frac{1}{1 - z^{25}} \frac{1}{1 - z^{10}} \frac{1}{1 - z^5} \frac{1}{1 - z}$  ← 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.

	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
$[z^n] \frac{1}{1-z}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$[z^n] \frac{1}{1-z} \frac{1}{1-z^5}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$[z^n] \frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}}$	1	2	4	6	9	12	16	20	25	30	36	42	49	56	64	72	81	90	100	110	121
$[z^n] \frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}} \frac{1}{1-z^{25}}$	1					13					49					121					242

## In-class exercise

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

How many ways to change a dollar?

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

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- **Compositions and partitions**
- Substitution

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- **Substitution**

## The symbolic method for unlabeled objects (summary)

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

Additional constructs are available (and still being invented)—one example to follow

## Another construct for the symbolic method: substitution

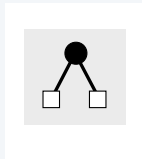
---

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

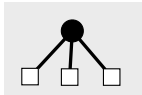
operation	notation	semantics	OGF
<i>substitution</i>	$A \circ [B]$	replace each object in an instance of $A$ with an object from $B$	$A(B(z))$

## Substitution application example

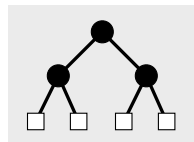
Q. How many 2-3 trees with  $N$  nodes?



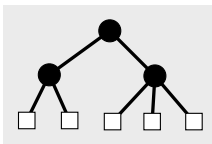
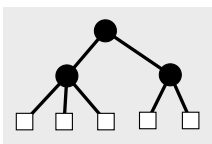
$W_2 = 1$



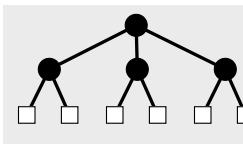
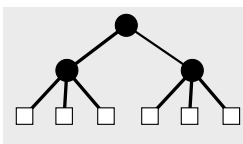
$W_3 = 1$



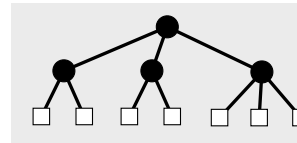
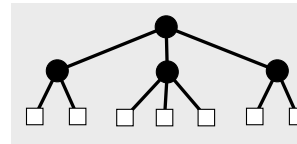
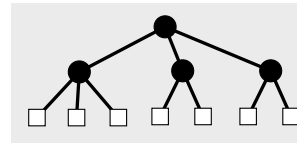
$W_4 = 1$



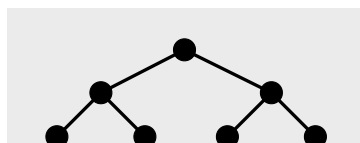
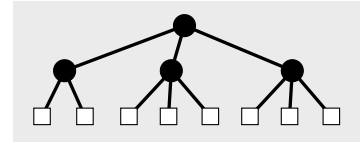
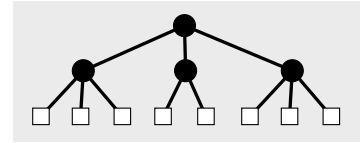
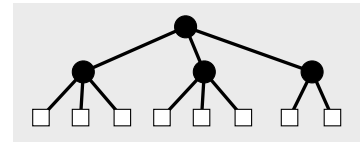
$W_5 = 2$



$W_6 = 2$



$W_7 = 3$



$W_8 = 4$



## Substitution application example

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

Combinatorial class  $W$ , the class of all 2-3 trees

Construction  $W = Z + W \circ [ (Z \times Z) + (Z \times Z \times Z) ]$

← “a 2-3 tree is a 2-3 tree with each external node replaced by a 2-node or a 3-node”

OGF equation  $W(z) = z + W(z^2 + z^3)$

$$\begin{aligned} W(z) &= z + z^2 + z^3 + z^4 + 2z^5 + 2z^6 + 3z^7 + 4z^8 \\ W(z^2 + z^3) &= z^2 + z^3 + (z^2 + z^3)^2 + (z^2 + z^3)^3 + (z^2 + z^3)^4 + \dots \\ &= z^2 + z^3 + (z^4 + 2z^5 + z^6) + (z^6 + 3z^7 + 3z^8 + z^9) + z^8 + \dots \checkmark \end{aligned}$$

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

See A. Odlyzko, *Periodic oscillations of coefficients of power series that satisfy functional equations*, Adv. in Mathematics (1982).

## Two French mathematicians on the utility of GFs

---



*"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*



*"**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*

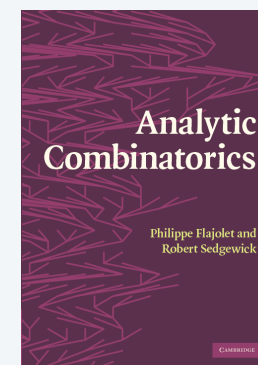
## Analytic combinatorics overview

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).



*Important note: GF equations vary widely in nature*

$$\begin{aligned} P(z) &= \frac{1}{(1-z)(1-z^2)(1-z^3)\dots} & C(z) &= \frac{1}{1-I(z)} & T(z) &= z + T(z^2 + z^3) & B(z) &= \frac{1}{1-2z} \\ S_M(z) &= \frac{1}{(1-z)^M} & H(z) &= z \exp(H(z) + H(z^2)/2 + H(z^3)/3 + \dots) \\ B_P(z) &= \frac{1-z^P}{1-2z+z^{P+1}} & G(z)^2 - G(z) + z &= 0 & Q(z) &= (1+z)(1+z^2)(1+z^3)\dots \end{aligned}$$

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

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs

- Symbolic method
- Trees and strings
- Powersets and multisets
- Compositions and partitions
- **Substitution**

Analytic  
Combinatorics

Philippe Flajolet and  
Robert Sedgewick

CAMBRIDGE

<http://ac.cs.princeton.edu>

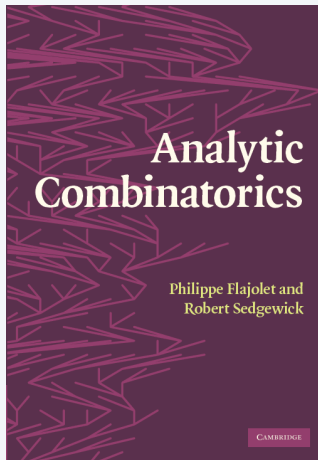
## 1. Combinatorial structures and OGFs

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

## Note 1.23

---

### Alice, Bob, and coding bounds



▷ **I.23.** *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  $\mathcal{C}$  be the class of allowed code words. For words of length  $n$ , a code of length  $L \equiv L(n)$  is achievable only if there exists a one-to-one mapping from  $\{0, 1\}^n$  into  $\bigcup_{j=0}^L \mathcal{C}_j$ , i.e.,  $2^n \leq \sum_{j=0}^L C_j$ . Working out the OGF of  $\mathcal{C}$ , one finds that necessarily

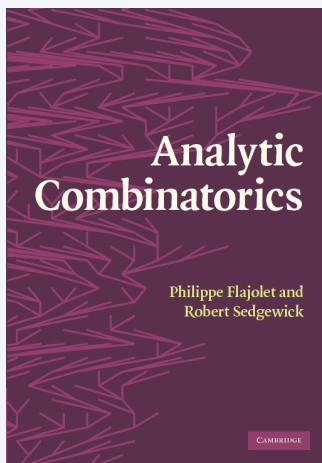
$$L(n) \geq \lambda n + O(1), \quad \lambda = \frac{1}{\log_2 \varphi} \doteq 1.440420, \quad \varphi = \frac{1 + \sqrt{5}}{2}.$$

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



▷ **I.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.



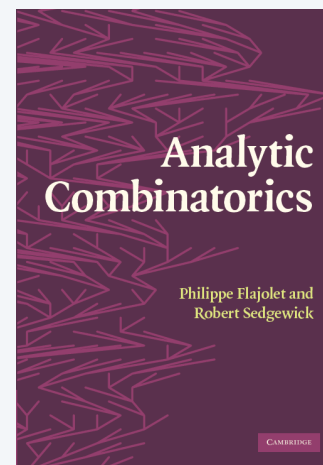
2. Write up solutions to Notes 1.23 and 1.43.

3. Programming exercises.



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

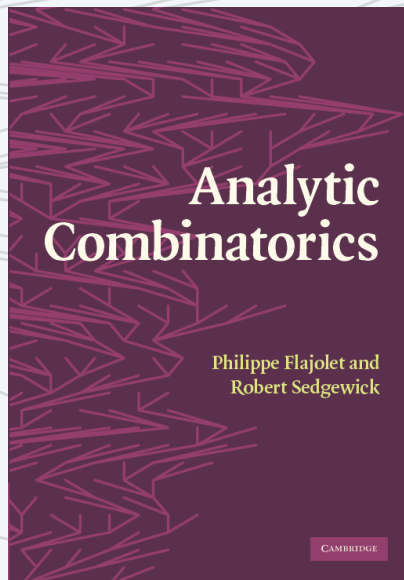
**Program I.2.** Write programs that estimate the rate of growth of the Cayley numbers and the partition numbers ( $H_n/H_{n-1}$  and  $P_n/P_{n-1}$ ). See Note I.43.





ANALYTIC COMBINATORICS

PART TWO



<http://ac.cs.princeton.edu>

# 1. Combinatorial structures and OGFs