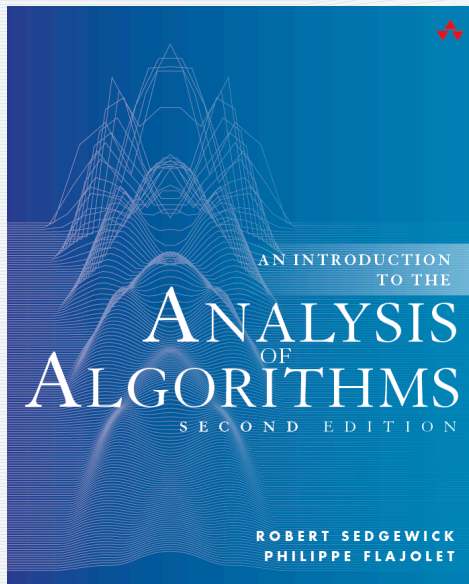


# ANALYTIC COMBINATORICS

## PART ONE

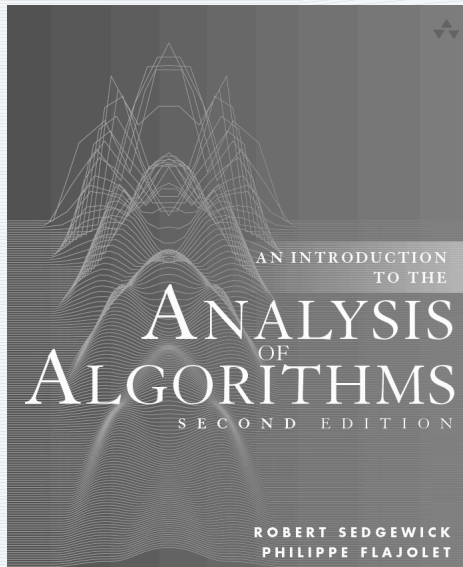


<http://aofa.cs.princeton.edu>

### 3. Generating Functions

# ANALYTIC COMBINATORICS

## PART ONE



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

### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- EGFs
- Counting with GFs

## Ordinary generating functions

Definition.

$$A(z) = \sum_{k \geq 0} a_k z^k \quad \text{is the ordinary generating function (OGF)}$$

of the sequence  $a_0, a_1, a_2, \dots, a_k, \dots$

Notation.  $[z^N]A(z)$  is “the coefficient of  $z^N$  in  $A(z)$ ”

| sequence               | OGF                                    |
|------------------------|----------------------------------------|
| 1, 1, 1, 1, 1, ...     | $\sum_{N \geq 0} z^N = \frac{1}{1-z}$  |
| 1, 1/2, 1/6, 1/24, ... | $\sum_{N \geq 0} \frac{z^N}{N!} = e^z$ |

←  $[z^N]e^z = 1/N!$

Significance. Can represent an entire sequence with a single function.

## Operations on OGFs: Scaling

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $A(cz) = \sum_{k \geq 0} a_k c^k z^k$  is the OGF of  $a_0, ca_1, c^2 a_2, c^3 a_3, \dots$

| sequence                | OGF                                        |
|-------------------------|--------------------------------------------|
| 1, 1, 1, 1, 1, ...      | $\sum_{N \geq 0} z^N = \frac{1}{1-z}$      |
| 1, 2, 4, 8, 16, 32, ... | $\sum_{N \geq 0} 2^N z^N = \frac{1}{1-2z}$ |

$$\leftarrow [z^N] \frac{1}{1-2z} = 2^N$$

## Operations on OGFs: Addition

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

and  $B(z) = \sum_{k \geq 0} b_k z^k$  is the OGF of  $b_0, b_1, b_2, \dots, b_k, \dots$

then  $A(z) + B(z)$  is the OGF of  $a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots, a_k + b_k, \dots$

Example:

| sequence                | OGF                                        |
|-------------------------|--------------------------------------------|
| 1, 1, 1, 1, 1, ...      | $\sum_{N \geq 0} z^N = \frac{1}{1-z}$      |
| 1, 2, 4, 8, 16, 32, ... | $\sum_{N \geq 0} 2^N z^N = \frac{1}{1-2z}$ |
| 0, 1, 3, 7, 15, 31, ... | $\frac{1}{1-2z} - \frac{1}{1-z}$           |

## Operations on OGFs: Differentiation

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $zA'(z) = \sum_{k \geq 1} k a_k z^k$  is the OGF of  $0, a_1, 2a_2, 3a_3, \dots, k a_k, \dots$

| OGF                                                          | sequence                          |
|--------------------------------------------------------------|-----------------------------------|
| $\frac{1}{1-z} = \sum_{N \geq 0} z^N$                        | 1, 1, 1, 1, 1, ...                |
| $\frac{z}{(1-z)^2} = \sum_{N \geq 1} N z^N$                  | 0, 1, 2, 3, 4, 5, ...             |
| $\frac{z^2}{(1-z)^3} = \sum_{N \geq 2} \binom{N}{2} z^N$     | 0, 0, 1, 3, 6, 10, ...            |
| $\frac{z^M}{(1-z)^{M+1}} = \sum_{N \geq M} \binom{N}{M} z^N$ | 0, ..., 1, M+1, (M+2)(M+1)/2, ... |
| $\frac{1}{(1-z)^{M+1}} = \sum_{N \geq 0} \binom{N+M}{M} z^N$ | 1, M+1, (M+2)(M+1)/2, ...         |

## Operations on OGFs: Integration

---

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $\int_0^z A(t) dt = \sum_{n \geq 1} \frac{a_{n-1}}{n} z^n$  is the OGF of  $0, a_0, \frac{a_1}{2}, \frac{a_2}{3}, \dots, \frac{a_{k-1}}{k}, \dots$

Example:

| OGF                                                 | sequence                      |
|-----------------------------------------------------|-------------------------------|
| $\frac{1}{1-z} = \sum_{N \geq 0} z^N$               | 1, 1, 1, 1, 1, ...            |
| $\ln \frac{1}{1-z} = \sum_{N \geq 1} \frac{z^N}{N}$ | 0, 1, 1/2, 1/3, 1/4, 1/5, ... |

## Operations on OGFs: Partial sum

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $\frac{1}{1-z} A(z)$  is the OGF of  $a_0, a_0 + a_1, a_0 + a_1 + a_2, \dots$

Proof.

$$\frac{1}{1-z} A(z) = \sum_{k \geq 0} z^k \sum_{n \geq 0} a_n z^n$$

Distribute

$$= \sum_{k \geq 0} \sum_{n \geq 0} a_n z^{n+k}$$

Change  $n$  to  $n-k$

$$= \sum_{k \geq 0} \sum_{n \geq k} a_{n-k} z^n$$

Switch order of summation.

$$= \sum_{n \geq 0} \left( \sum_{0 \leq k \leq n} a_{n-k} \right) z^n$$

Change  $k$  to  $n-k$

$$= \sum_{n \geq 0} \left( \sum_{0 \leq k \leq n} a_k \right) z^n$$

## Operations on OGFs: Partial sum

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $\frac{1}{1-z} A(z)$  is the OGF of  $a_0, a_0 + a_1, a_0 + a_1 + a_2, \dots$

Example:

| OGF                                                         | sequence                       |
|-------------------------------------------------------------|--------------------------------|
| $\frac{1}{1-z} = \sum_{N \geq 0} z^N$                       | 1, 1, 1, 1, 1, ...             |
| $\ln \frac{1}{1-z} = \sum_{N \geq 1} \frac{z^N}{N}$         | 0, 1, 1/2, 1/3, 1/4, 1/5, ...  |
| $\frac{1}{1-z} \ln \frac{1}{1-z} = \sum_{N \geq 1} H_N z^N$ | 1, 1 + 1/2, 1 + 1/2 + 1/3, ... |

## Operations on OGFs: Convolution

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

and  $B(z) = \sum_{k \geq 0} b_k z^k$  is the OGF of  $b_0, b_1, b_2, \dots, b_k, \dots$

then  $A(z)B(z)$  is the OGF of  $a_0 b_0, a_1 b_0 + a_0 b_1, \dots, \sum_{0 \leq k \leq n} a_k b_{n-k}, \dots$

Proof.

$$A(z)B(z) = \sum_{k \geq 0} a_k z^k \sum_{n \geq 0} b_n z^n$$

Distribute

$$= \sum_{k \geq 0} \sum_{n \geq 0} a_k b_n z^{n+k}$$

Change  $n$  to  $n-k$

$$= \sum_{k \geq 0} \sum_{n \geq k} a_k b_{n-k} z^n$$

Switch order of summation.

$$= \sum_{n \geq 0} \sum_{0 \leq k \leq n} a_k b_{n-k} z^n$$

## Operations on OGFs: Convolution

---

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

and  $B(z) = \sum_{k \geq 0} b_k z^k$  is the OGF of  $b_0, b_1, b_2, \dots, b_k, \dots$

then  $A(z)B(z)$  is the OGF of  $a_0 b_0, a_1 b_0 + a_0 b_1, \dots, \sum_{0 \leq k \leq n} a_k b_{n-k}, \dots$

Example:

| OGF                                            | sequence           |
|------------------------------------------------|--------------------|
| $\frac{1}{1-z} = \sum_{N \geq 0} z^N$          | 1, 1, 1, 1, 1, ... |
| $\frac{1}{(1-z)^2} = \sum_{N \geq 0} (N+1)z^N$ | 1, 2, 3, 4, 5, ... |

## Expanding a GF (summary)

---

The process of expressing an unknown GF as a power series (finding the coefficients) is known as **expanding** the GF.

Techniques we have been using:

1. Taylor theorem: 
$$f(z) = f(0) + f'(0)z + \frac{f''(0)}{2!}z^2 + \frac{f'''(0)}{3!}z^3 + \frac{f''''(0)}{4!}z^4 + \dots$$

Example.

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots$$

2. Reduce to known GFs.

Example.

$$[z^N] \frac{1}{(1-z)} \ln \frac{1}{1-z} = \sum_{1 \leq k \leq N} \frac{1}{k} = H_N$$

Integrate  $\frac{1}{1-z}$  to get  $\ln \frac{1}{1-z}$   
then convolve  $\frac{1}{1-z}$  with  $\ln \frac{1}{1-z}$

## In-class exercise

**Exercise 3.4** Prove that  $\sum_{1 \leq k \leq N} H_k = (N+1)(H_{N+1} - 1)$

1. Find GF for LHS (convolve  $\frac{1}{1-z}$  with  $\frac{1}{1-z} \ln \frac{1}{1-z}$  )

$$\frac{1}{(1-z)^2} \ln \frac{1}{1-z}$$

2. Expand GF to find RHS coefficients (convolve  $\ln \frac{1}{1-z}$  with  $\frac{1}{(1-z)^2}$  )

$$[z^N] \frac{1}{(1-z)^2} \ln \frac{1}{1-z} = \sum_{1 \leq k \leq N} \frac{1}{k} (N+1-k)$$

### Operations on OGFs: Partial sum

If  $A(z) = \sum_{k \geq 0} a_k z^k$  is the OGF of  $a_0, a_1, a_2, \dots, a_k, \dots$

then  $\frac{1}{1-z} A(z)$  is the OGF of  $a_0, a_0 + a_1, a_0 + a_1 + a_2, \dots$

| Example: | OGF                                                         | sequence                       |
|----------|-------------------------------------------------------------|--------------------------------|
|          | $\frac{1}{1-z} = \sum_{N \geq 0} z^N$                       | 1, 1, 1, 1, 1, ...             |
|          | $\ln \frac{1}{1-z} = \sum_{N \geq 1} \frac{z^N}{N}$         | 0, 1, 1/2, 1/3, 1/4, 1/5, ...  |
|          | $\frac{1}{1-z} \ln \frac{1}{1-z} = \sum_{N \geq 1} H_N z^N$ | 1, 1 + 1/2, 1 + 1/2 + 1/3, ... |

3. Do some math

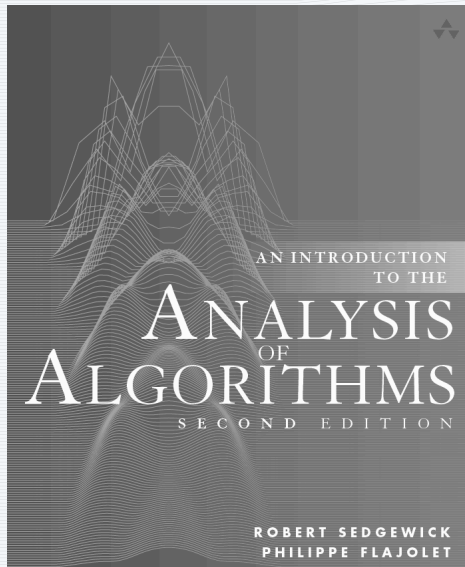
$$= (N+1)H_N \cdot N$$

$$= (N+1)(H_{N+1} \cdot \frac{1}{N+1}) \cdot N$$

$$= (N+1)(H_{N+1} \cdot 1)$$

# ANALYTIC COMBINATORICS

## PART ONE



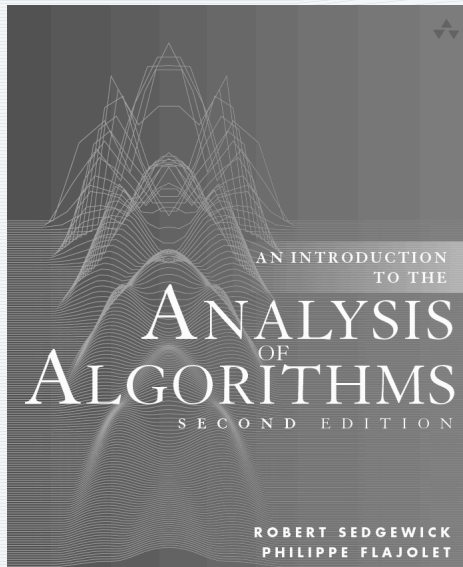
<http://aofa.cs.princeton.edu>

### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- EGFs
- Counting with GFs

# ANALYTIC COMBINATORICS

## PART ONE



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

### 3. Generating Functions

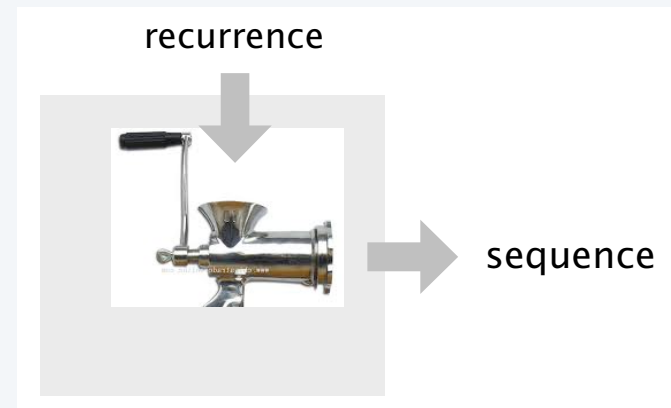
- OGFs
- **Solving recurrences**
- Catalan numbers
- EGFs
- Counting with GFs

## Solving recurrences with OGFs

---

General procedure:

- Make recurrence valid for all  $n$ .
- Multiply both sides of the recurrence by  $z^n$  and sum on  $n$ .
- Evaluate the sums to derive an equation satisfied by the OGF.
- Solve the equation to derive an explicit formula for the OGF.  
(Use the initial conditions!)
- Expand the OGF to find coefficients.



## Solving recurrences with GFs

For linear recurrences with constant coefficients, the GF equation is a polynomial, so the general procedure is a **algorithm**.

Example 4 from previous lecture.

$$a_n = 5a_{n-1} - 6a_{n-2} \quad \text{for } n \geq 2 \text{ with } a_0 = 0 \text{ and } a_1 = 1$$

Make recurrence valid for all  $n$ .

$$a_n = 5a_{n-1} - 6a_{n-2} + \delta_{n1}$$

Multiply by  $z^n$  and sum on  $n$ .

$$A(z) = 5zA(z) - 6z^2A(z) + z$$

Solve.

$$A(z) = \frac{z}{1 - 5z + 6z^2}$$

Use partial fractions:  
solution must be of the form

$$A(z) = \frac{c_0}{1 - 3z} + \frac{c_1}{1 - 2z}$$

Solve for coefficients.

$$c_0 + c_1 = 0$$

$$2c_0 + 3c_1 = -1$$

Solution is  $c_0 = 1$  and  $c_1 = -1$

$$A(z) = \frac{1}{1 - 3z} - \frac{1}{1 - 2z}$$

Expand.

$$a_n = 3^n - 2^n$$

## Solving linear recurrences with GFs

For linear recurrences with constant coefficients, the GF equation is a polynomial, so the general procedure is a **algorithm**.

Example with multiple roots.

$$a_n = 5a_{n-1} - 8a_{n-2} + 4a_{n-3} \quad \text{for } n \geq 3 \text{ with } a_0 = 0, a_1 = 1 \text{ and } a_2 = 4$$

Make recurrence valid for all  $n$ .

$$a_n = 5a_{n-1} - 8a_{n-2} + 4a_{n-3} + \delta_{n1} - \delta_{n2}$$

Multiply by  $z^n$  and sum on  $n$ .

$$A(z) = 5zA(z) - 8z^2A(z) + 4z^3A(z) + z - z^2$$

Solve.

$$A(z) = \frac{z - z^2}{1 - 5z + 8z^2 - 4z^3}$$

Simplify.

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

Expand.

$$a_n = n2^{n-1}$$

← multiplicity 3 gives terms of the form  $n^2\beta^n$ , etc.

## Solving linear recurrences with GFs

For linear recurrences with constant coefficients, the GF equation is a polynomial, so the general procedure is a **algorithm**.

Example with complex roots.

$$a_n = 2a_{n-1} - a_{n-2} + 2a_{n-3} \quad \text{for } n \geq 3 \text{ with } a_0 = 1, a_1 = 0 \text{ and } a_2 = -1$$

Make recurrence valid for all  $n$ .

$$a_n = 2a_{n-1} - a_{n-2} + 2a_{n-3} + \delta_{n0} - 2\delta_{n1}$$

Multiply by  $z^n$  and sum on  $n$ .

$$A(z) = 2zA(z) - z^2A(z) + 2z^3A(z) + 1 - 2z$$

Solve.

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

Simplify.

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

Use partial fractions.

$$A(z) = \frac{1}{2} \left( \frac{1}{1 - iz} + \frac{1}{1 + iz} \right)$$

Expand.

$$a_n = \frac{1}{2} (i^n + (-i)^n) = \frac{1}{2} i^n (1 + (-1)^n)$$

$$1, 0, -1, 0, 1, 0, -1, 0, 1, \dots$$

## Solving linear recurrences with GFs (summary)

---

Solution to  $a_n = x_1 a_{n-1} + x_2 a_{n-2} + \dots + x_t a_{n-t}$

is a linear combination of  $t$  terms.

Suppose the roots of the polynomial  $1 - x_1 z + x_2 z^2 + \dots + x_t z^t$

are  $\beta_1, \beta_2, \dots, \beta_r$  where the multiplicity of  $\beta_i$  is  $m_i$  so  $m_1 + m_2 + \dots + m_r = t$

Solution is 
$$\sum_{0 \leq j < m_1} c_{1j} n^j \beta_1^n + \sum_{0 \leq j < m_2} c_{2j} n^j \beta_2^n + \dots + \sum_{0 \leq j < m_r} c_{rj} n^j \beta_r^n \quad \leftarrow t \text{ terms}$$

The  $t$  constants  $c_{ij}$  are determined from the initial conditions.

**Note:** complex roots (and  $-1$ ) introduce periodic behavior.

## Solving the Quicksort recurrence with OGFs

$$C_N = N + 1 + \frac{2}{N} \sum_{1 \leq k \leq N} C_{k-1}$$

Multiply both sides by  $N$ .

$$NC_N = N(N + 1) + 2 \sum_{1 \leq k \leq N} C_{k-1}$$

Multiply by  $z^N$  and sum.

$$\sum_{N \geq 1} NC_N z^N = \sum_{N \geq 1} N(N + 1) z^N + 2 \sum_{N \geq 1} \sum_{1 \leq k \leq N} C_{k-1} z^N$$

Evaluate sums to get an ordinary differential equation

$$C'(z) = \frac{2}{(1 - z)^3} + 2 \frac{C(z)}{1 - z}$$

homogeneous equation  
 $\rho'(z) = 2\rho(z)/(1 - z)$   
 solution (integration factor)  
 $\rho(z) = 1/(1 - z)^2$

Solve the ODE.

$$\begin{aligned} ((1 - z)^2 C(z))' &= (1 - z)^2 C'(z) - 2(1 - z)C(z) \\ &= (1 - z)^2 \left( C'(z) - 2 \frac{C(z)}{1 - z} \right) = \frac{2}{1 - z} \end{aligned}$$

Integrate.

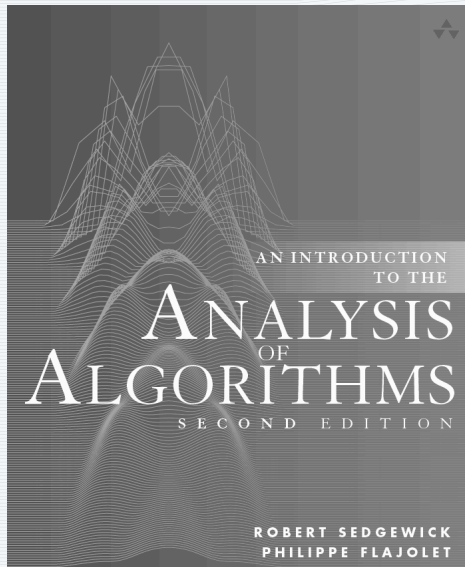
$$C(z) = \frac{2}{(1 - z)^2} \ln \frac{1}{1 - z}$$

Expand.

$$C_N = [z^N] \frac{2}{(1 - z)^2} \ln \frac{1}{1 - z} = 2(N + 1)(H_{N+1} - 1)$$

# ANALYTIC COMBINATORICS

## PART ONE



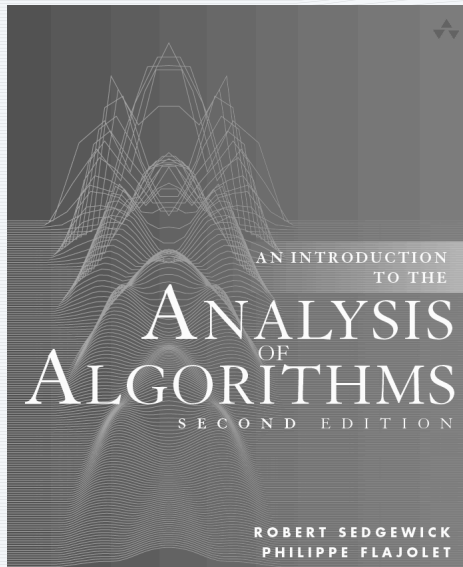
<http://aofa.cs.princeton.edu>

### 3. Generating Functions

- OGFs
- **Solving recurrences**
- Catalan numbers
- EGFs
- Counting with GFs

# ANALYTIC COMBINATORICS

## PART ONE



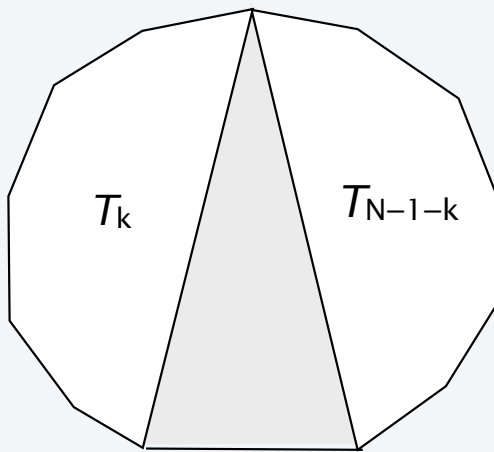
<http://aofa.cs.princeton.edu>

### 3. Generating Functions

- OGFs
- Solving recurrences
- **Catalan numbers**
- EGFs
- Counting with GFs

# Catalan numbers

How many **triangulations** of an  $(N+2)$ -gon?



$$T_N = \sum_{0 \leq k < N} T_k T_{N-1-k} + \delta_{N0}$$

$$T_0 = 1$$

$$T_1 = 1$$

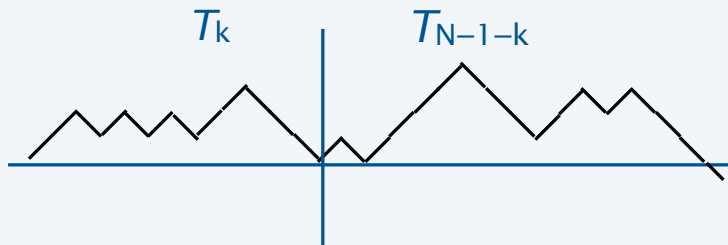
$$T_2 = 2$$

$$T_3 = 5$$

$$T_4 = 14$$

# Catalan numbers

How many **gambler's ruin sequences** with  $N$  wins?



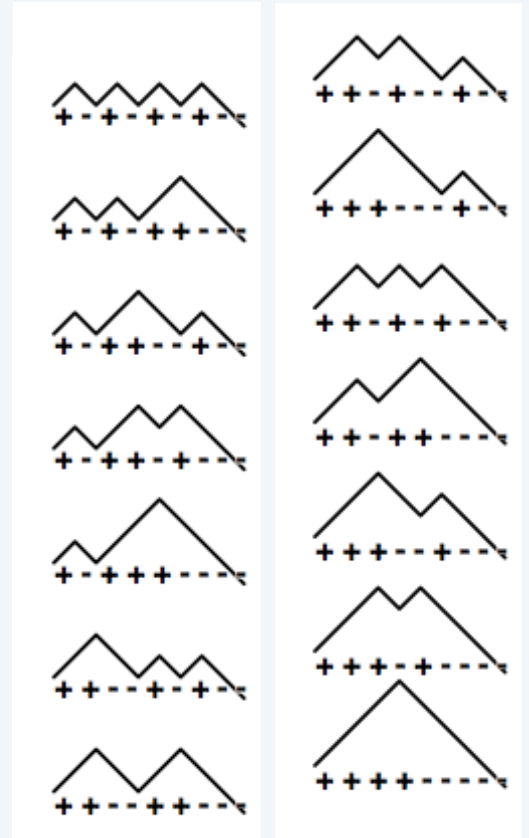
$$T_N = \sum_{0 \leq k < N} T_k T_{N-1-k} + \delta_{N0}$$

$$\nearrow$$
  
 $T_0 = 1$

$$\nearrow$$
  
 $T_1 = 1$

$$\nearrow$$
  
 $T_2 = 2$

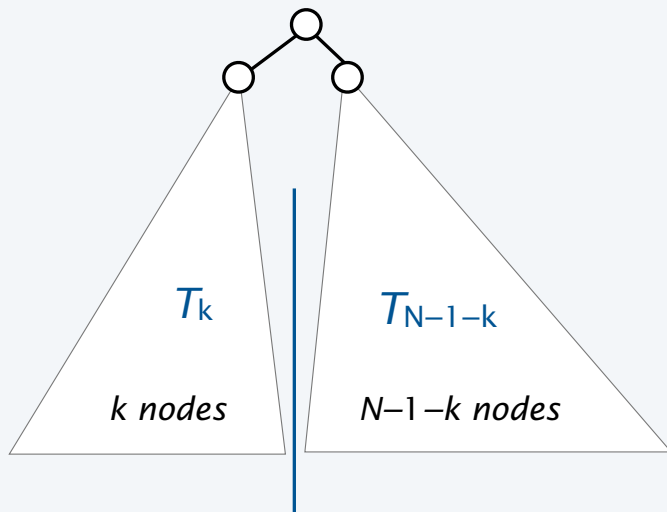
$$\nearrow$$
  
 $T_3 = 5$



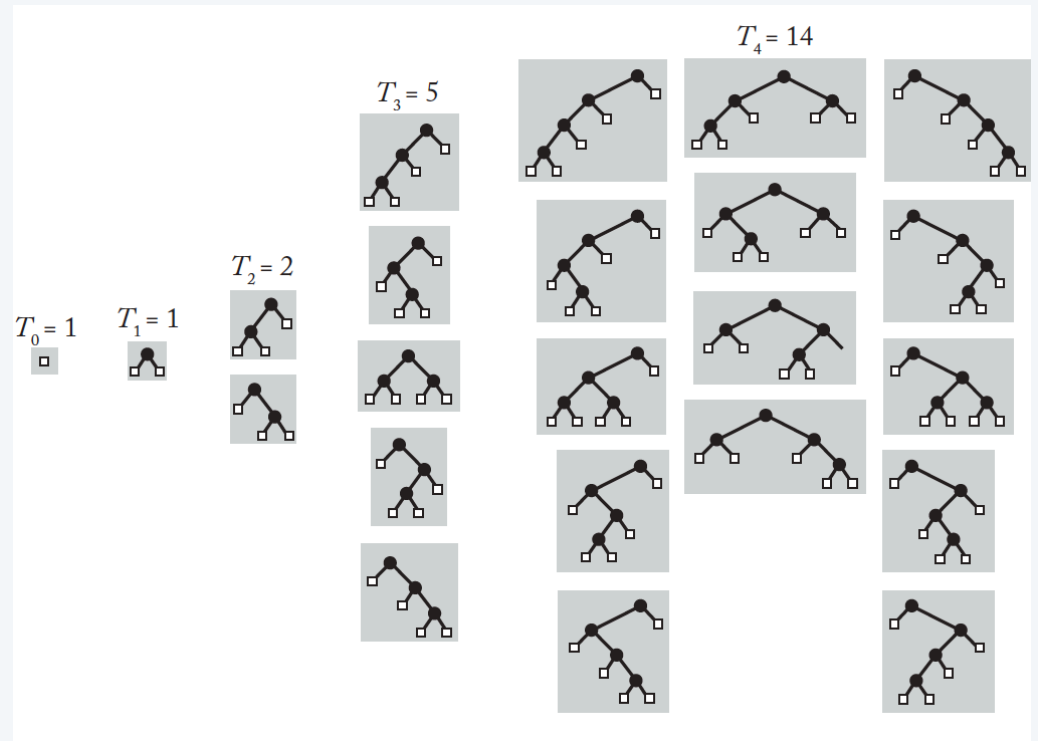
$T_4 = 14$

## Catalan numbers

How many **binary trees** with  $N$  nodes?

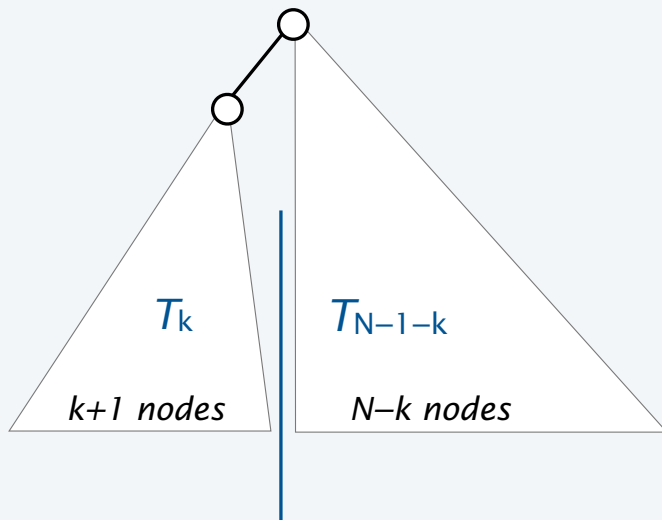


$$T_N = \sum_{0 \leq k < N} T_k T_{N-1-k} + \delta_{N0}$$

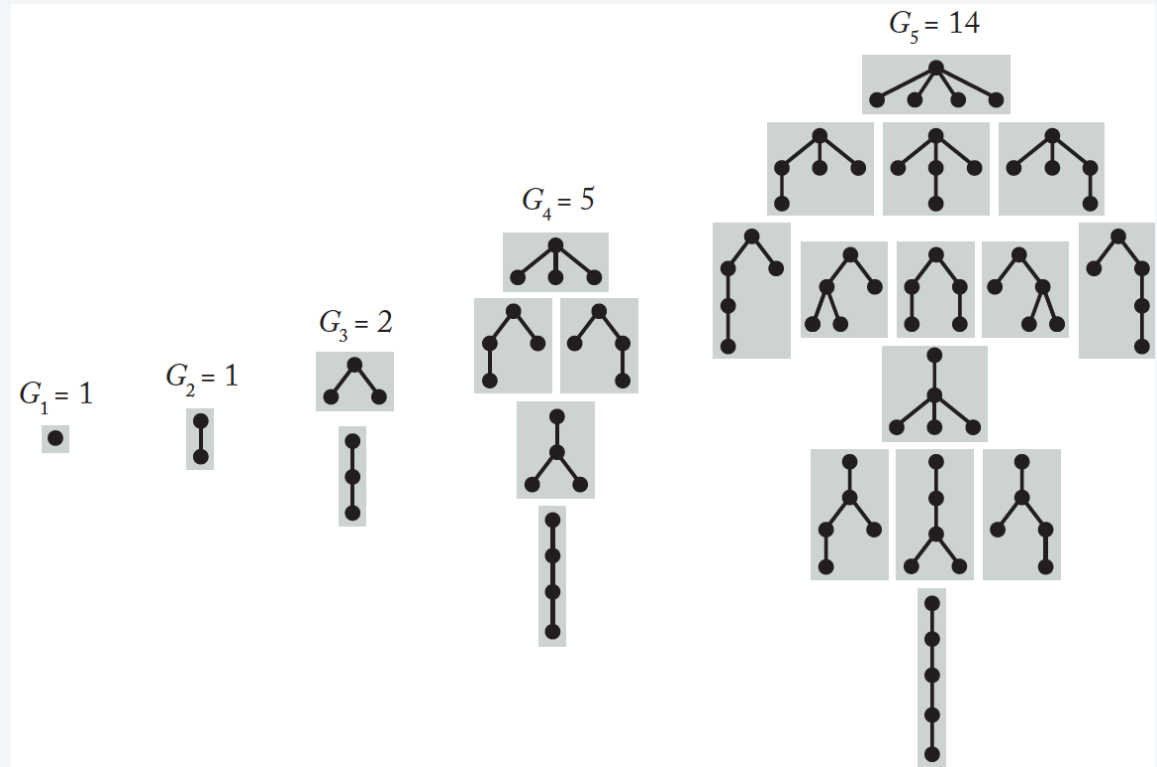


## Catalan numbers

How many **trees** with  $N+1$  nodes?



$$T_N = \sum_{0 \leq k < N} T_k T_{N-1-k} + \delta_{N0}$$



## Solving the Catalan recurrence with GFs

Recurrence that holds for all  $N$ .

$$T_N = \sum_{0 \leq k < N} T_k T_{N-1-k} + \delta_{N0}$$

Multiply by  $z^N$  and sum.

$$T(z) \equiv \sum_{N \geq 0} T_N z^N = \sum_{N \geq 0} \sum_{0 \leq k < N} T_k T_{N-1-k} z^N + 1$$

Switch order of summation

$$T(z) = 1 + \sum_{k \geq 0} \sum_{N > k} T_k T_{N-1-k} z^N$$

Change  $N$  to  $N+k+1$

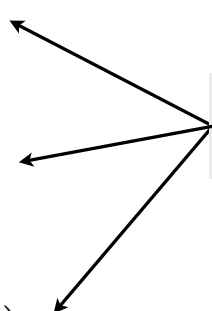
$$T(z) = 1 + \sum_{k \geq 0} \sum_{N \geq 0} T_k T_N z^{N+k+1}$$

Distribute.

$$T(z) = 1 + z \left( \sum_{k \geq 0} T_k z^k \right) \left( \sum_{N \geq 0} T_N z^N \right)$$

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

convolution  
(backwards)



## Common-sense rule for working with GFs

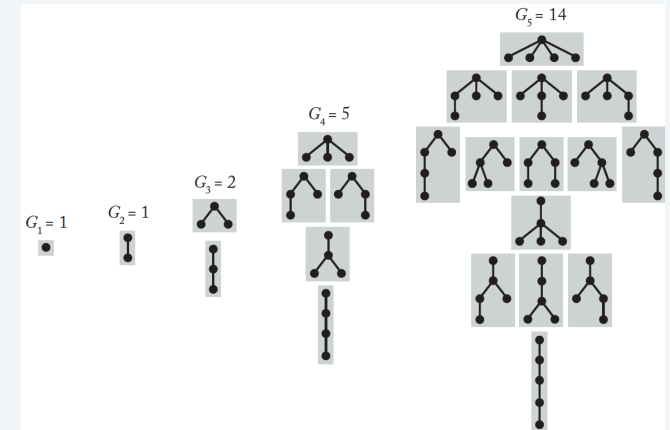
It is **always** worthwhile to check your math with your computer.

Known from initial values:

$$T(z) = 1 + z + 2z^2 + 5z^3 + 14z^4 + \dots$$

Check:

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



```
sage: ZP.<z> = ZZ[]
```

```
sage: 1 + z*(1+z+2*z^2+5*z^3+14*z^4)*(1+z+2*z^2+5*z^3+14*z^4)
```

```
196*z^9 + 140*z^8 + 81*z^7 + 48*z^6 + 42*z^5 + 14*z^4 + 5*z^3 + 2*z^2 + z + 1
```

not valid because  
 $z^5$  and beyond  
missing in factors



## Solving the Catalan recurrence with GFs (continued)

Functional GF equation.

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

Solve with quadratic formula.

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

Expand via binomial theorem.

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

Set coefficients equal

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

Expand via definition.

$$= -\frac{1}{2} \frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)\dots(\frac{1}{2}-N)(-4)^{N+1}}{(N+1)!}$$

Distribute  $(-2)^N$  among factors.

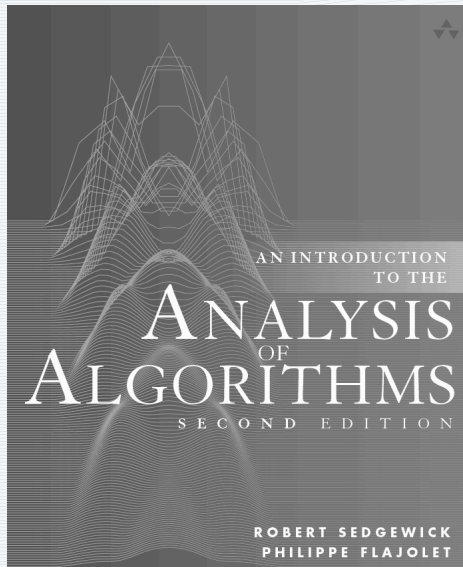
$$= \frac{1 \cdot 3 \cdot 5 \dots (2N-1) \cdot 2^N}{(N+1)!}$$

Substitute  $(2/1)(4/2)(6/3)\dots$  for  $2^N$ .

$$\begin{aligned} &= \frac{1}{N+1} \frac{1 \cdot 3 \cdot 5 \dots (2N-1)}{N!} \frac{2 \cdot 4 \cdot 6 \dots 2N}{1 \cdot 2 \cdot 3 \dots N} \\ &= \frac{1}{N+1} \binom{2N}{N} \end{aligned}$$

# ANALYTIC COMBINATORICS

## PART ONE



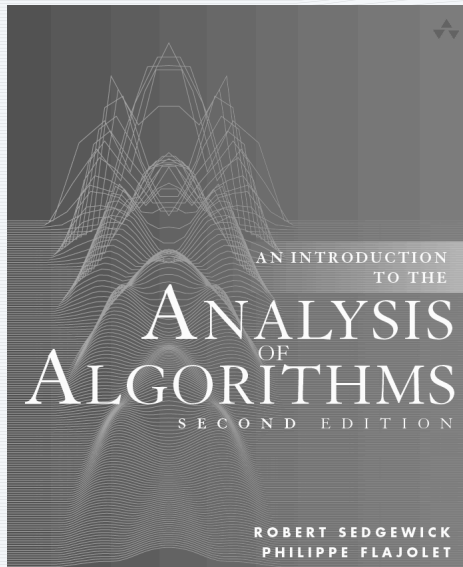
<http://aofa.cs.princeton.edu>

### 3. Generating Functions

- OGFs
- Solving recurrences
- **Catalan numbers**
- EGFs
- Counting with GFs

# ANALYTIC COMBINATORICS

## PART ONE



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

### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- **EGFs**
- Counting with GFs

## Exponential generating functions (EGFs)

Definition.

$A(z) = \sum_{k \geq 0} a_k \frac{z^k}{k!}$  is the **exponential generating function (EGF)**

of the sequence  $a_0, a_1, a_2, \dots, a_k, \dots$

| sequence                | EGF                                                 |
|-------------------------|-----------------------------------------------------|
| 1, 1, 1, 1, 1, ...      | $\sum_{N \geq 0} \frac{z^N}{N!} = e^z$              |
| 1, 2, 4, 8, 16, 32, ... | $\sum_{N \geq 0} 2^N \frac{z^N}{N!} = e^{2z}$       |
| 1, 1, 2, 6, 24, 120 ... | $\sum_{N \geq 0} N! \frac{z^N}{N!} = \frac{1}{1-z}$ |

## Operations on EGFs: Binomial convolution

If  $A(z) = \sum_{k \geq 0} a_k \frac{z^k}{k!}$  is the EGF of  $a_0, a_1, a_2, \dots, a_k, \dots$   
and  $B(z) = \sum_{k \geq 0} b_k \frac{z^k}{k!}$  is the EGF of  $b_0, b_1, b_2, \dots, b_k, \dots$   
then  $A(z)B(z)$  is the EGF of  $a_0b_0, a_0b_1 + a_1b_0, \dots, \binom{n}{k} a_k b_{n-k}, \dots$

Proof.

$$A(z)B(z) = \sum_{k \geq 0} a_k \frac{z^k}{k!} \sum_{n \geq 0} b_n \frac{z^n}{n!}$$

Distribute.

$$= \sum_{k \geq 0} \sum_{n \geq 0} \frac{a_k}{k!} \frac{b_n}{n!} z^{n+k}$$

Change  $n$  to  $n-k$

$$= \sum_{k \geq 0} \sum_{n \geq k} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} z^n$$

Multiply and divide by  $n!$

$$= \sum_{k \geq 0} \sum_{n \geq k} \binom{n}{k} a_k b_{n-k} \frac{z^n}{n!}$$

Switch order of summation.

$$= \sum_{n \geq 0} \left( \sum_{0 \leq k \leq n} \binom{n}{k} a_k b_{n-k} \right) \frac{z^n}{n!}$$

## Solving recurrences with EGFs

Choice of EGF vs. OGF is typically dictated naturally from the problem.

Example. 
$$f_n = \sum_k \binom{n}{k} \frac{f_k}{2^k}$$

Multiply by  $z^n/n!$  and sum on  $n$ .

Switch order of summation.

Change  $n$  to  $n+k$ .

Simplify.

Distribute.

Evaluate and telescope.

Expand.

$$f(z) = \sum_{n \geq 0} \sum_k \binom{n}{k} \frac{f_k}{2^k} \frac{z^n}{n!}$$

$$f(z) = \sum_{k \geq 0} \sum_{n \geq k} \binom{n}{k} \frac{f_k}{2^k} \frac{z^n}{n!}$$

$$f(z) = \sum_{k \geq 0} \sum_{n \geq 0} \binom{n+k}{k} \frac{f_k}{2^k} \frac{z^{n+k}}{(n+k)!}$$

$$f(z) = \sum_{k \geq 0} \sum_{n \geq 0} f_k \frac{(z/2)^k}{k!} \frac{z^n}{n!}$$

$$f(z) = \left( \sum_{k \geq 0} f_k \frac{(z/2)^k}{k!} \right) \left( \sum_{n \geq 0} \frac{z^n}{n!} \right)$$

$$f(z) = e^z f(z/2) = e^{z+z/2+z/4+z/8+\dots} = e^{2z}$$

$$f_n = 2^n$$

binomial  
convolution  
(backwards)

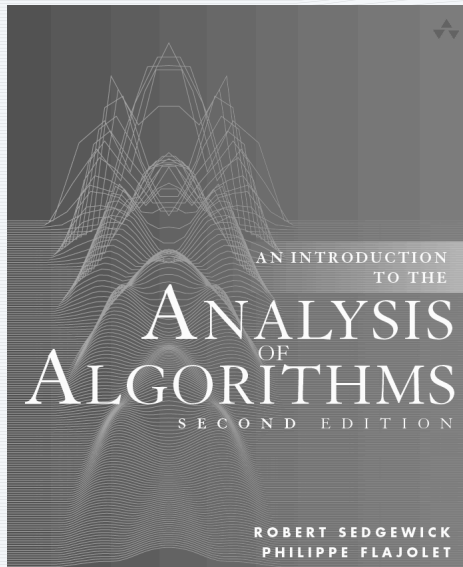
convergence  
not assured

Check.

$$2^n = \sum_k \binom{n}{k} \frac{2^k}{2^k}$$

# ANALYTIC COMBINATORICS

## PART ONE



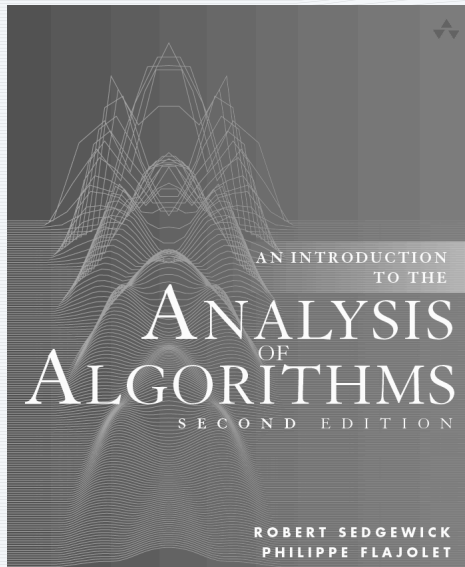
<http://aofa.cs.princeton.edu>

### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- **EGFs**
- Counting with GFs

# ANALYTIC COMBINATORICS

## PART ONE



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

### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- EGFs
- Counting with GFs

## Counting with generating functions

An alternative (combinatorial) view of GFs

- Define a *class* of combinatorial objects with associated *size* function.
- GF is sum over all members of the class.

Example.

$\mathcal{T} \equiv$  set of all binary trees

$|t| \equiv$  number of *internal* nodes in  $t \in \mathcal{T}$

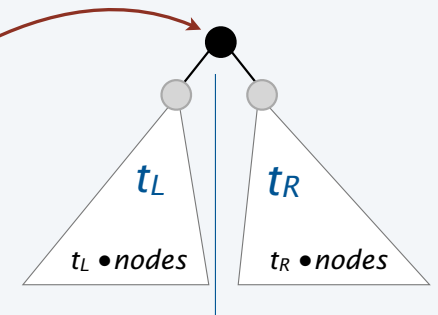
$T_N \equiv$  number of  $t \in \mathcal{T}$  with  $|t| = N$

$$T(z) \equiv \sum_{t \in \mathcal{T}} z^{|t|} = \sum_{N \geq 0} T_N z^N$$

Decompose from definition

Distribute

$$\begin{aligned} T(z) &= 1 + \sum_{t_L \in \mathcal{T}} \sum_{t_R \in \mathcal{T}} z^{|t_L| + |t_R| + 1} \\ &= 1 + z \left( \sum_{t_L \in \mathcal{T}} z^{|t_L|} \right) \left( \sum_{t_R \in \mathcal{T}} z^{|t_R|} \right) \\ &= 1 + zT(z)^2 \end{aligned}$$



## Combinatorial view of Catalan GF

Each term  $z^N$  in the GF corresponds to an object of size  $N$ .

*Collect all the terms with the same exponent to expose counts.*

Each term  $z^i z^j$  in a product corresponds to an object of size  $i + j$ .

$$\begin{aligned}
 T(z) &= 1 + z + z^2 + z^2 + z^3 + z^3 + z^3 + z^3 + z^3 + \dots \\
 &\quad \square \quad \begin{array}{|c|} \hline \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \dots \\
 &= 1 + z + 2z^2 + 5z^3 + \dots
 \end{aligned}$$

$$\begin{aligned}
 T(z) &= 1 + zT(z)^2 \\
 &= 1 + z(1 + z + z^2 + z^2 + \dots)(1 + z + z^2 + z^2 + \dots)
 \end{aligned}$$

$$\begin{aligned}
 &= 1 + z + z^2 + z^2 + z^3 + z^3 + z^3 + z^3 + z^3 + \dots \\
 &\quad \square \quad \begin{array}{|c|} \hline \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \begin{array}{|c|} \hline \bullet \bullet \\ \hline \square \square \\ \hline \end{array} \quad \dots
 \end{aligned}$$

$z \times 1 \times 1$      $z \times 1 \times z$      $z \times z \times 1$      $z \times 1 \times z^2$      $z \times 1 \times z^2$      $z \times z \times z$      $z \times z^2 \times 1$      $z \times z^2 \times 1$

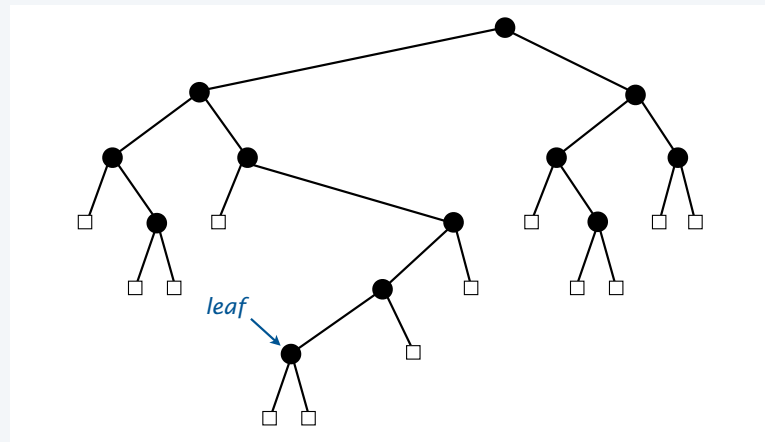
## Values of parameters ("costs")

are often the object of study in the analysis of algorithms.

## How many 1 bits in a random bitstring? (Easy)

01110100100010001110101000001010000

## How many leaves in a random binary tree? (Not so easy)



## Computing expected costs by counting

An alternative (combinatorial) view of probability

- Define a *class* of combinatorial objects.
- Model: All objects of size  $N$  are equally likely

$\mathcal{P} \equiv$  set of all objects in the class

$|p| \equiv$  size of  $p \in \mathcal{P}$

$P_N \equiv$  number of  $p \in \mathcal{P}$  with  $|p| = N$

$\text{cost}(p) \equiv$  cost associated with  $p$

$P_{Nk} \equiv$  number of  $p \in \mathcal{P}$  with  $|p| = N$  and  $\text{cost}(p) = k$

Expected cost of an object of size  $N$

$$C_N \equiv \sum_{k \geq 0} k \frac{P_{Nk}}{P_N} \quad \leftarrow \frac{P_{Nk}}{P_N} \text{ is the probability that the cost of an object of size } N \text{ is } k$$
$$= \frac{\sum_{k \geq 0} k P_{Nk}}{P_N} \quad \leftarrow \text{"cumulated cost"}$$

Def. *Cumulated cost* is total cost of all objects of a given size.

*Expected cost* is cumulated cost divided by number of objects.

## Counting with generating functions: cumulative costs

---

An alternative (combinatorial) view of GFs

- Define a *class* of combinatorial objects.
- Model: All objects of size  $N$  are equally likely
- GF is sum over all members of the class.

$\mathcal{P} \equiv$  set of all objects in the class

$|p| \equiv$  size of  $p \in \mathcal{P}$

$P_N \equiv$  number of  $p \in \mathcal{P}$  with  $|p| = N$

$\text{cost}(p) \equiv$  cost associated with  $p$

Counting GF

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

Cumulative cost GF

$$C(z) \equiv \sum_{p \in \mathcal{P}} \text{cost}(p) z^{|p|} = \sum_{N \geq 0} \sum_{k \geq 0} k P_{Nk} z^N$$

Average cost

$$[z^N]C(z)/[z^N]P(z)$$

Bottom line: Reduces computing expectation to GF counting

## Warmup: How many 1 bits in a random bitstring?

$B$  is the set of all bitstrings.

$|b|$  is the number of bits in  $b$ .

$\text{ones}(b)$  is the number of 1 bits in  $b$ .

$B_N$  is the # of bitstrings of size  $N$  ( $2^N$ ).

$C_N$  is the total number of 1 bits in all bitstrings of size  $N$ .

Counting GF.

$$B(z) = \sum_{b \in B} z^{|b|} = \sum_{N \geq 0} 2^N z^N = \frac{1}{1 - 2z}$$

Cumulative cost GF.

$$\begin{aligned} C(z) &= \sum_{b \in B} \text{ones}(b) z^{|b|} \\ &= \sum_{b' \in B} (1 + 2 \cdot \text{ones}(b')) z^{|b'|+1} \\ &= zB(z) + 2zC(z) \\ &= \frac{z}{(1 - 2z)^2} \end{aligned}$$

|   |      |
|---|------|
| 0 | $b'$ |
| 1 | $b'$ |

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

Average # 1 bits in a random bitstring of length  $N$ .

$$\frac{[z^N]C(z)}{[z^N]B(z)} = \frac{N2^{N-1}}{2^N} = \frac{N}{2} \quad \checkmark$$

## Leaves in binary trees

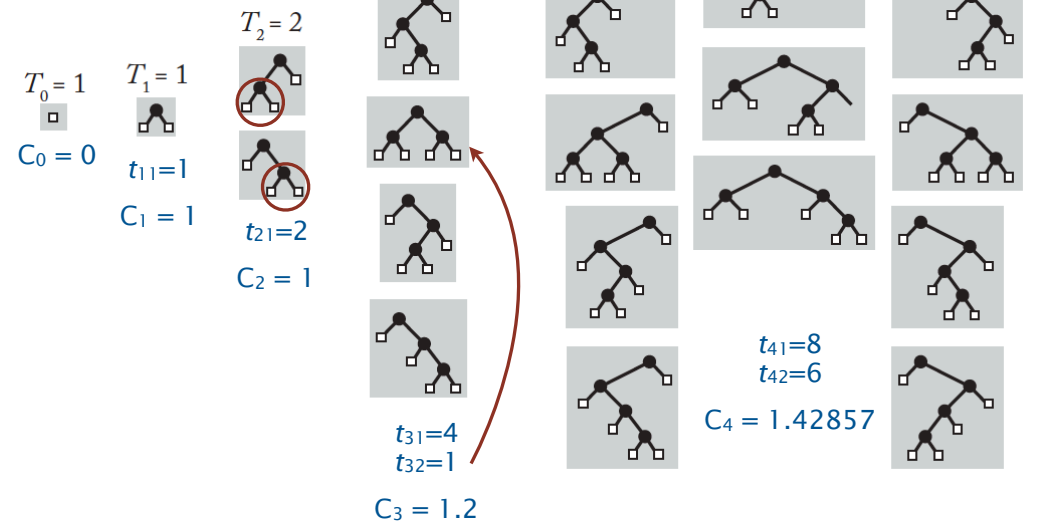
are internal nodes whose children are both external.

Definitions:

$T_N$  is the # of binary trees with  $N$  nodes.

$t_{Nk}$  is the # of  $N$ -node binary trees with  $k$  leaves

$C_N$  is the average # of leaves in a random  $N$ -node binary tree



Q. How many leaves in a random binary tree?

## How many leaves in a random binary tree?

$T$  is the set of all binary trees.

$|t|$  is the number of internal nodes in  $t$ .

$\text{leaves}(t)$  is the number of leaves in  $t$ .

$T_N$  is the # of binary trees of size  $N$  (Catalan).

$C_N$  is the total number of leaves in all binary trees of size  $N$ .

Counting GF.

$$T(z) = \sum_{t \in T} z^{|t|} = \sum_{N \geq 0} T_N z^N = \sum_{N \geq 0} \frac{1}{N+1} \binom{2N}{N} z^N$$

Cumulative cost GF.

$$C(z) = \sum_{t \in T} \text{leaves}(t) z^{|t|}$$

Average # leaves in a random  
 $N$ -node binary tree.

$$\frac{[z^N]C(z)}{[z^N]T(z)} = \frac{[z^N]C(z)}{T_N}$$

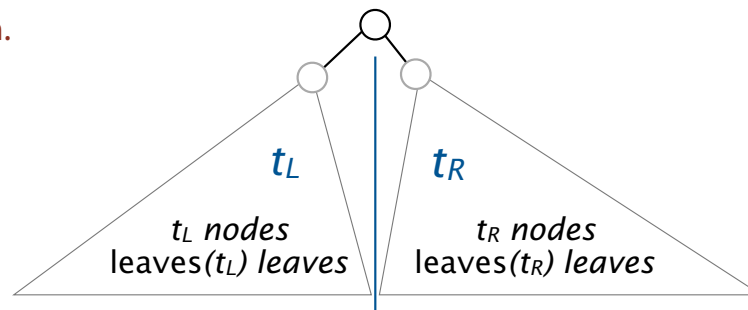
Next: Derive a functional equation for the CGF.

## CGF functional equation for leaves in binary trees

CGF.

$$C(z) = \sum_{t \in T} \text{leaves}(t) z^{|t|}$$

Decompose from definition.



$$\begin{aligned}
 C(z) &= z + \sum_{t_L \in T} \sum_{t_R \in T} (\text{leaves}(t_L) + \text{leaves}(t_R)) z^{|t_L| + |t_R| + 1} \\
 &= z + z \sum_{t_L \in T} \text{leaves}(t_L) z^{|t_L|} \sum_{t_R \in T} z^{|t_R|} + z \sum_{t_L \in T} z^{|t_L|} \sum_{t_R \in T} \text{leaves}(t_R) z^{|t_R|} \\
 &= z + 2zC(z)T(z)
 \end{aligned}$$

## How many leaves in a random binary tree?

CGF.

$$C(z) = \sum_{t \in T} \text{leaves}(t) z^{|t|}$$

Decompose from definition.

$$\begin{aligned} C(z) &= z + \sum_{t_L \in T} \sum_{t_R \in T} (\text{leaves}(t_L) + \text{leaves}(t_R)) z^{|t_L| + |t_R| + 1} \\ &= z + 2zC(z)T(z) \end{aligned}$$

Compute number of trees  $T_N$ .  
*Catalan numbers*

$$\begin{aligned} T(z) &= zT(z)^2 - z \\ &= \frac{1}{2z} (1 - \sqrt{1 - 4z}) \end{aligned}$$

$$\begin{aligned} T_N &= [z^N] \frac{1}{2z} (1 - \sqrt{1 - 4z}) \\ &= \frac{1}{N+1} \binom{2N}{N} \end{aligned}$$

Compute cumulated cost  $C_N$ .

$$\begin{aligned} C(z) &= z + 2zT(z)C(z) \\ &= \frac{z}{1 - 2zT(z)} = \frac{z}{\sqrt{1 - 4z}} \end{aligned}$$

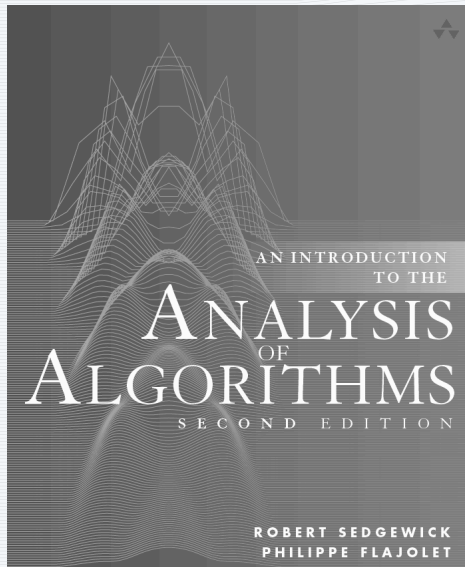
$$\begin{aligned} C_N &= [z^N] \frac{z}{\sqrt{1 - 4z}} \\ &= \binom{2N-2}{N-1} \end{aligned}$$

Compute average number of leaves.

$$C_N/T_N = \frac{\binom{2N-2}{N-1}}{\frac{1}{N+1} \binom{2N}{N}} = \frac{(N+1) \cdot N \cdot N}{2N(2N-1)} \sim \textcircled{N/4}$$

# ANALYTIC COMBINATORICS

## PART ONE



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

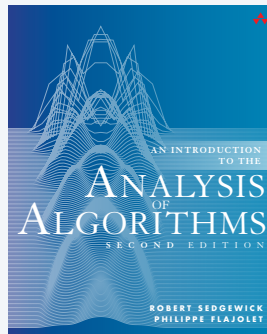
### 3. Generating Functions

- OGFs
- Solving recurrences
- Catalan numbers
- EGFs
- Counting with GFs

## Exercise 3.20

---

Solve a linear recurrence. Initial conditions matter.



**Exercise 3.20** Solve the recurrence

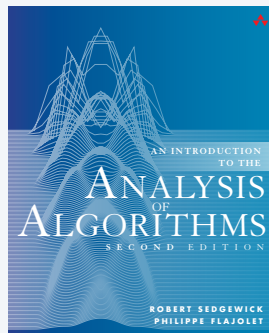
$$a_n = 3a_{n-1} - 3a_{n-2} + a_{n-3} \quad \text{for } n > 2 \text{ with } a_0 = a_1 = 0 \text{ and } a_2 = 1.$$

Solve the same recurrence with the initial condition on  $a_1$  changed to  $a_1 = 1$ .

## Exercise 3.28

---

The art of expanding GFs.



**Exercise 3.28** Find an expression for

$$[z^n] \frac{1}{\sqrt{1-z}} \ln \frac{1}{1-z}.$$

(*Hint:* Expand  $(1-z)^{-\alpha}$  and differentiate with respect to  $\alpha$ .)

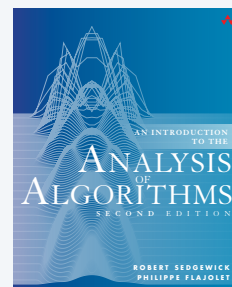
## Assignments for next lecture

---

1. Use a symbolic mathematics system  
to check initial values for  $C(z) = z + 2C(z)T(z)$ .



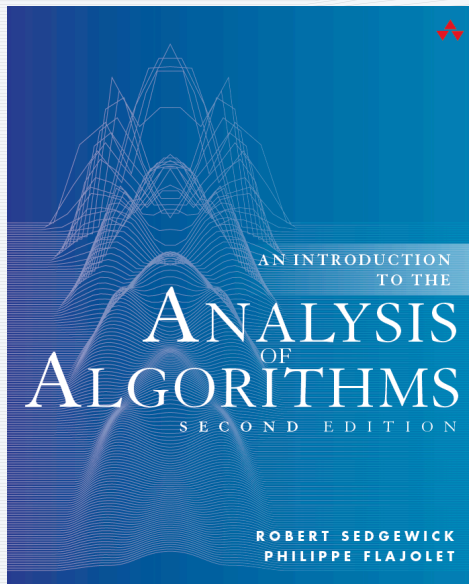
2. Read pages 89-147 in text.



3. Write up solutions to Exercises 3.20 and 3.28.

# ANALYTIC COMBINATORICS

## PART ONE



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

### 3. Generating Functions