ANALYTIC COMBINATORICS

Philippe FLAJOLET

Bologna, June 2010
Counting...

Figure 3.1  All binary trees with 1, 2, 3, 4, and 5 external nodes
Counting (and asymptotics)

Binary trees $\Rightarrow$ \textit{Catalan numbers}

Formula is

\[
C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{n! (n+1)!}
\]

Growth rate is (asymptotics)

\[
C_n \sim \frac{4^n}{\sqrt{\pi n}}
\]
Counting (and probabilities)

Plane tree

Increasing tree
Counting (methods)

E.g. binary trees: $1, 1, 2, 5, 14, 42, \ldots$

Bijective combinatorics = first principles

Generating function methods ... 

Algebraic methods (e.g., symmetric fns, operator)
Generating Functions (GFs)

Combinatorial class \( \mathcal{C} \); counting sequence \((C_n)\):

\[
\mathcal{C} \quad \rightarrow \quad \begin{cases} 
C(z) &= \sum C_n z^n \quad \text{(OGF)} \\
\hat{C}(z) &= \sum C_n \frac{z^n}{n!} \quad \text{(EGF)}
\end{cases}
\]

- Get GFs
  - combinatorics \( \rightsquigarrow \) algebra of special fns
- Look at GFs as mappings of complex plane, \( z \in \mathbb{C} \)
  - algebra of special fns \( \rightsquigarrow \) complex analysis
- For parameters, add extra variables
  - complex analysis \( \rightsquigarrow \) perturbation theory
A Calculus of Discrete Structures

Discrete

Continuous

(a digital tree aka trie of size 500)
(a generating function in the complex plane)
Analytic Combinatorics

Philippe Flajolet and Robert Sedgewick

algo.inria.fr/flajolet
Analytic Combinatorics

~ A. Combinatorial structures
~ B. Analytic structures
~ C. Randomness properties

for objects given by constructions
Laplace discovered the remarkable correspondence between set theoretic operations and operations on formal power series and put it to great use to solve a variety of combinatorial problems. — G.–C. ROTA

La méthode des fonctions génératrices, qui a exercé ses ravages pendant un siècle, est tombée en désuétude... — Claude BERGE
Quotations (2)

Despite all appearances they [generating functions] belong to algebra and not to analysis.

Combinatorialists use recurrence, generating functions, and such transformations as the Vandermonde convolution; others to my horror, use contour integrals, differential equations, and other resources of mathematical analysis. — John RIORDAN
PART I
Symbolic Methods

*1. Unlabelled structures & OGFs
*2. Labelled structures and EGFs
*3. Parameters and multivariate GFs

Embed a fragment of set theory into a language of constructions; map to algebra(s) of special functions.
Chapter I
Unlabelled structures and OGFs
Symbolic Methods

Embed a fragment of set theory into a language of constructions;

map combinatorics to algebra(s) of special functions.
1 UNLABELLED STRUCTURES AND OGFS

Ordinary Generating Function (OGF)

\[(f_n) \rightarrow f(z) := \sum_{n=0}^{\infty} f_n z^n.\]

\((f_n)\) is number sequence, e.g., counting sequence.

Later: Exponential Generating function (EGF): \((f_n) \rightarrow f(z) := \sum_{n=0}^{\infty} f_n \frac{z^n}{n!}.\)
Symbolic approach

- An object of size $n$ is viewed as composed of $n$ atoms (with additional structure): words, trees, graphs, permutations, etc.
- Replace each atom by symbolic weight $z$:
  - Class: $\sum$ objects. Object: $\gamma \rightsquigarrow z^{|\gamma|}$.

Gives the Ordinary Generating Function (OGF):

\[
C \rightsquigarrow C(z) := \sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_{n} C_n z^n.
\]
E.g.: a class of graphs enumerated by \( \# \) vertices

\[
\begin{align*}
C & = \quad \begin{tikzpicture}[scale=0.5] 
\foreach \i in {1,...,4} { 
\node[shape=circle,fill,inner sep=2pt] (n\i) at ({\i-0.5},0) {};
\node[shape=circle,fill,inner sep=2pt] (n5) at (4,0) {};
\draw (n\i) -- (n\i+1);
}\end{tikzpicture} & + & \begin{tikzpicture}[scale=0.5] 
\foreach \i in {1,...,3} { 
\node[shape=circle,fill,inner sep=2pt] (n\i) at ({\i-0.5},0) {};
}\node[shape=circle,fill,inner sep=2pt] (n4) at (3,0) {};
\draw (n1) -- (n2) -- (n3) -- (n4) -- (n1);
\end{tikzpicture} & + & \begin{tikzpicture}[scale=0.5] 
\foreach \i in {1,...,3} { 
\node[shape=circle,fill,inner sep=2pt] (n\i) at ({\i-0.5},0) {};
}\node[shape=circle,fill,inner sep=2pt] (n4) at (3,0) {};
\draw (n1) -- (n2) -- (n3) -- (n4) -- (n1);
\draw (n2) -- (n4);
\end{tikzpicture} & + & \begin{tikzpicture}[scale=0.5] 
\foreach \i in {1,...,3} { 
\node[shape=circle,fill,inner sep=2pt] (n\i) at ({\i-0.5},0) {};
}\node[shape=circle,fill,inner sep=2pt] (n4) at (3,0) {};
\draw (n1) -- (n2) -- (n3) -- (n4) -- (n1);
\draw (n2) -- (n3);
\draw (n4) -- (n2);
\end{tikzpicture} & + & \begin{tikzpicture}[scale=0.5] 
\foreach \i in {1,...,4} { 
\node[shape=circle,fill,inner sep=2pt] (n\i) at ({\i-0.5},0) {};
}\node[shape=circle,fill,inner sep=2pt] (n5) at (4,0) {};
\draw (n1) -- (n2) -- (n3) -- (n4) -- (n1);
\draw (n2) -- (n3);
\draw (n4) -- (n5);
\end{tikzpicture} & + & \begin{tikzpicture}[scale=0.5] 
\node[shape=circle,fill,inner sep=2pt] (n1) at (0,0) {};
\node[shape=circle,fill,inner sep=2pt] (n2) at (1,0) {};
\node[shape=circle,fill,inner sep=2pt] (n3) at (2,0) {};
\node[shape=circle,fill,inner sep=2pt] (n4) at (3,0) {};
\node[shape=circle,fill,inner sep=2pt] (n5) at (4,0) {};
\draw (n1) -- (n2) -- (n3) -- (n4) -- (n1);
\draw (n2) -- (n3);
\draw (n4) -- (n5);
\end{tikzpicture} \\
C(z) &= z z z z + z z z + z z z + z z z z z + z \\
&= 1 \cdot z + 2 \cdot z^3 + 2 \cdot z^4 \\
(C_n) &= (0, 1, 0, 2, 2).
\end{align*}
\]

Principle (Symbolic method)

The OGF of a class: (i) encodes the counting sequence; (ii) is nothing but a reduced form of the class itself.
How many binary trees $B_n$ with $n$ external nodes?

$B = \square + \bullet$, $(B \times B)$.

Euler-Segner (1743): Recurrence

$B_n = \sum_{k=1}^{n-1} B_k B_{n-k}$.

Form OGF: $B(z) = z + (B(z) \times B(z))$.

Solve equation (quadratic):

$B(z) = \frac{1}{2} \left( 1 - \sqrt{1 - 4z} \right) = \frac{1}{2} - \frac{1}{2} (1 - 4z)^{1/2}$.

Expand:

$B_n = \frac{1}{n} \binom{2n-2}{n-1}$ (Catalan numbers)

Analogy: $B = \square + (\bullet B \times B) \sim B(z) = z + (B(z) \times B(z))$
Define a collection of constructions

union, product, sequence, set, cycle, …

allowing for recursive definitions.

meta-THM1: OGFs are automatically computable (equations!)

meta-THM2: Counting sequences are automatically computable in time $O(n^2)$, and even $O(n^{1+\varepsilon})$.

meta-THM3: Random generation is fast in $O(n \log n)$ arithmetic op´ns.
Several set-theoretic constructions translate into GFs.

\begin{align*}
\sum_{A \oplus B} &= \sum_{A} + \sum_{B} \\
\sum_{A \times B} &= \sum_{A} \cdot \sum_{B}
\end{align*}

There is a micro-dictionary:

\begin{align*}
\text{disjoint union} & \quad C = A \cup B \quad \implies \quad C(z) = A(z) + B(z) \\
\text{cartesian product} & \quad C = A \times B \quad \implies \quad C(z) = A(z) \cdot B(z)
\end{align*}
**Theorem.** There exists a dictionary:

<table>
<thead>
<tr>
<th>Construction</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C = \mathcal{A} + \mathcal{B}$</td>
<td>$C(z) = A(z) + B(z)$</td>
</tr>
<tr>
<td>$C = \mathcal{A} \times \mathcal{B}$</td>
<td>$C(z) = A(z) \cdot B(z)$</td>
</tr>
<tr>
<td>$C = \text{SEQ}(\mathcal{A})$</td>
<td>$C(z) = \frac{1}{1 - A(z)}$</td>
</tr>
<tr>
<td>$C = \text{MSET}(\mathcal{A})$</td>
<td>$C(z) = \text{Exp}(A(z))$</td>
</tr>
<tr>
<td>$C = \text{PSET}(\mathcal{A})$</td>
<td>$C(z) = \widehat{\text{Exp}}(A(z))$</td>
</tr>
<tr>
<td>$C = \text{Cyc}(\mathcal{A})$</td>
<td>$C(z) = \text{Log} \frac{1}{1 - A(z)}$</td>
</tr>
</tbody>
</table>

$\mathcal{E}$ or 1: “neutral class” formed with element of size 0 $\leftrightarrow E(z) = 1$.

$\mathcal{Z}$: “atomic class” formed with element of size 1 $\leftrightarrow E(z) = 1$.

\[
\text{Exp}(g(z)) = \exp \left( \sum_{k \geq 1} \frac{1}{k} g(z^k) \right); \quad \widehat{\text{Exp}}(g(z)) = \exp \left( \sum_{k \geq 1} \frac{(-1)^k}{k} g(z^k) \right);
\]

\[
\text{Log}(g(z)) = \sum_{k \geq 1} \frac{\varphi(k)}{k} g(z^k) \quad \text{with} \quad \varphi(k) = \text{Euler totient}.
\]
\text{SEQUENCE}
\begin{align*}
C &= \text{Seq}(A) \\
(i) \quad C &= 1 + (A \times C) \\
(ii) \quad \frac{1}{1-f} &= 1 + f + f^2 + \ldots
\end{align*}

\text{SET}
\begin{align*}
C &= \text{Set}(A) \\
&= \prod_{\alpha \in A} (1 + z^\alpha) \\
C(z) &= \prod_{n=1}^{\infty} (1 + z^n)^{A_n} = \exp \left( \sum_n A_n \log (1 + z^n) \right) \ldots
\end{align*}

\text{MULTISET}
\begin{align*}
C &= \text{MSet}(A) \equiv \prod_{\alpha \in A} \text{Seq}(\{\alpha\}) \\
C(z) &= \prod_{n=1}^{\infty} (1 - z^n)^{-A_n} = \exp(A(z) + \frac{1}{2} A(z^2) + \ldots).
\end{align*}
Theorem (Symbolic method)

A dictionary translates constructions into generating functions:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Union</strong></td>
<td>+</td>
</tr>
<tr>
<td><strong>Product</strong></td>
<td>×</td>
</tr>
<tr>
<td><strong>Sequence</strong></td>
<td>( \frac{1}{1 - \cdots} )</td>
</tr>
<tr>
<td><strong>Set</strong></td>
<td>Exp</td>
</tr>
<tr>
<td><strong>Cycle</strong></td>
<td>Log</td>
</tr>
</tbody>
</table>
This theorem permits us to write \( B \) for binary trees automatically.

\[
B = z + \frac{B(z)^2}{2}
\]
Roots...

- A modicum of Pólya theory (1937)
- Schützenberger: languages and GFs (~1960)
- Rota-Stanley = MIT School (1970s)
- Goulden-Jackson = constructions (~1980)
- Joyal’s theory of species +BLL (1980s)
Example 1. Binary words

\[ \mathcal{W} = \text{SEQ}(\{a, b\}) \implies W(z) = \frac{1}{1 - 2z}. \]

Get \( W_n = 2^n \) (!?). Words starting with \( b \) and < 4 consecutive \( a \)'s:

\[ \mathcal{W}^\bullet \cong \text{SEQ}(b \times (1+a+aa+aaa)) \implies W^\bullet(z) = \frac{1}{1 - (z + z^2 + z^3 + z^4)}. \]

Longest run statistics lead to rational functions (Feller).
Example 2. Plane trees ("general" = all degrees allowed)

\[ P = \mathbb{Z} \times \text{Seq}(P) \]

\[
\begin{align*}
P_1 &= 1 \\
P_2 &= 1 \\
P_3 &= 2 \\
P_4 &= 5
\end{align*}
\]

\[
P(z) = \frac{z}{1 - P(z)} \implies P(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z}
\]

\[
P_n = \frac{1}{n} \binom{2n-2}{n-1}
\]

Catalan numbers again!
Example 3. Nonplane trees (all degrees allowed) \( \mathcal{U} = \mathbb{Z} \times \text{MSET}(\mathcal{U}) \). \( U_1 = 1, \ U_2 = 1, \ U_3 = 2, \ U_4 = 5. \)

\[
U(z) = z \exp \left( \frac{1}{1} U(z) + \frac{1}{2} U(z^2) + \frac{1}{3} U(z^3) + \cdots \right).
\]

Cayley: recurrences; Pólya: asymptotics of this infinite functional equation.

Exercise: computable in polynomial time \( O(n^2) \).
Example 4. Words containing a pattern \((a^2b)\)

\[
\begin{array}{c}
0 \rightarrow a \rightarrow 1 \rightarrow b \rightarrow 2 \rightarrow b \rightarrow 3 \rightarrow a, b \\
\end{array}
\]

\[\mathcal{L}_j := \text{language accepted from state } j.\]

\[\{\mathcal{L}_0 = a\mathcal{L}_1 + b\mathcal{L}_0, \mathcal{L}_1 = a\mathcal{L}_1 + b\mathcal{L}_2, \mathcal{L}_2 = a\mathcal{L}_1 + b\mathcal{L}_3, \ldots\}\]

**Theorem.** Regular language (finite automaton) has rational GF.

\[
\text{Reg} \quad \leftrightarrow \quad \mathbb{Q}(z).
\]

Patterns of all sorts in words. Applications in pattern matching algorithms and computational biology.
Example 5. Walks and excursions.

$$\text{gains} \quad \Rightarrow \text{time}$$

$$\text{Excursion} = \text{Seq} (\uparrow \text{Excursion} \downarrow)$$

Positive path = Excursion $\times$ Seq($\uparrow$ Excursion $\downarrow$)

Draw game = Seq($\uparrow$ Excursion $\downarrow$ + $\uparrow$ Excursion $\downarrow$) etc.
Simple families of plane trees.

Let $\Omega \subseteq \mathbb{Z}_{\geq 0}$ be the set of allowed (out)degrees. Define

$$\phi(y) := \sum_{w \in \Omega} y^w.$$  

Then the simple family $\mathcal{Y}$ has OGF:

$$Y(z) = z \phi(Y(z)).$$

If $\phi$ is finite, get an algebraic function.

Lagrange Inversion Theorem.

$$[z^n] Y(z) = \frac{1}{n} \text{coeff}[w^n] \phi(w)^n.$$  

If $\phi$ is finite, get multinomial sums.
A variety of classes => a variety of “special functions”

Some constructible families and generating functions

- Regular languages, FA, paths in graphs: \( \sim \)
  - rational fns
- Unambiguous context-free languages \( \sim \)
  - algebraic functions
- Terms trees \( \sim \) [+Pólya operators]
  - implicit functions

\[
\text{Tree} = \begin{array}{c} \text{Tree} \\
\end{array} \quad \Rightarrow \quad T = z \Phi(T)
\]
Algebraic functions (1)

~ Arise from specifications (CF grammars), with +, x, Seq

~ *Elimination*: system \( \rightarrow \) single equation
\[ P(x,y)=0 \]

~ Coefficients are “combinatorial sums”

[e.g., Sokal, *SLC* 2009]
MAPS: Tutte’s **quadratic method**;
cf Cori, Bousquet-Mélou et al., Bordeaux School...

**EXCURSIONS**: the **kernel method**;
cf Lalley 1993, Banderier-F 2001, MBM

\[
F(z, u) = 1 + z(u^{-2} + u^{0} + u^{3})F(z, u) - \text{coeff } [u^{<0}]u^{-2}F(z, u)
\]

\[\implies \text{solve } 1 - z(u^{-2} + u^{0} + u^{3}) = 0.\]
Chapter 2
Labelled structures and EGFs
2 LABELLED STRUCTURES AND EGFS

EGF = exponential generating function

\[(f_n) \rightarrow f(z) = \sum_{n \geq 0} f_n \frac{z^n}{n!} .\]

A labelled object has atoms that bear distinct integer labels (canonically numbered on \([1 \ldots n]\)).

Example. How many (undirected) graphs on \(n\) (distinguishable) vertices? \(G^n = 2^{n(n-1)/2}\).

Graphs: unlabelled problem is harder (Pólya theory). In general, can get unlabelled by identification of labelled.
PERMUTATIONS = typical labelled objects: write \( \sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma_1 & \sigma_2 & \cdots & \sigma_n \end{pmatrix} \)
as \( \sigma_1 \sigma_2 \cdots \sigma_n \) and view as linear digraph that is labelled:

\[
\begin{array}{cccc}
& 1 & 1 & 2 \\
\text{E}, & & & \\
\text{O}, & & & \\
\text{O}, & & & \\
\text{O}, & & & \\
\end{array}
\]

EGF is \( \frac{1}{1 - z} \) since \( P(z) = \sum_n n! \frac{z^n}{n!} \).

DISCONNECTED GRAPHS (labelled) = no edges aka “Urns”.

\[
\begin{array}{cccc}
& 1 & 1 & 1 \\
1 & & & \\
1 & & & \\
\end{array}
\]

EGF is \( U(z) = \exp(z) = e^z \).

CYCLIC GRAPHS (directed)

EGF \( K(z) = \log \frac{1}{1 - z} \).
ROOTED TREES (graphs) nonplane and labelled

\[ T_n = ?? \]

\[ T_1 = 1, \quad T_2 = 2, \quad T_3 = 9, \quad (T_4 = 64...) \]

» Unlabelled:

\[ U_1 = 1, \quad U_2 = 1, \quad U_3 = 2, \quad U_4 = 4, ... \]
Labelled product. Let $A$ and $B$ be labelled classes. Then the cartesian product $A \times B$ is not well-labelled (why?).

Given $(\beta, \gamma)$ form all possible relabellings that preserve the order structure within $\beta, \gamma$, while giving rise to well-labelled objects.

• **Labelled product of two objects.**

\[
(a \star \beta) := \{ \gamma \mid \gamma = (\alpha', \beta') \},
\]

where $\gamma$ is well-labelled and $\alpha' \equiv_{\text{order}} \alpha$ and $\beta' \equiv_{\text{order}} \beta$.

• **Labelled product of two classes.**

\[
C := \bigcup_{\alpha \in A, \beta \in B} (a \star \beta).
\]
\[ C = A \ast B \]

\[ C_n = \sum_{k=0}^{m} \binom{m}{k} A_k B_{n-k} \]

\[ \frac{C_n}{n!} = \sum_{k} \frac{A_k}{k!} \cdot \frac{B_{n-k}}{(n-k)!} \]

\[ \Rightarrow \hat{C}(z) = \hat{A}(z) \cdot \hat{B}(z) \leftarrow \text{EGF} \]
Sequences, Sets, Cycles

- $\mathcal{E}$ (or 1): neutral class.
- $\mathcal{Z}$: atomic class $\equiv [1]$.
- Define $\text{SEQ}(\mathcal{A})$, $\text{SET}(\mathcal{A})$, $\text{Cyc}(\mathcal{A})$ by relabellings:

$$\text{SEQ}(\mathcal{A}) = 1 + \mathcal{A} + (\mathcal{A} \star \mathcal{A}) + \cdots.$$ 

Sets: quotient up to perms. Cyc: up to cyclic perms.
**Theorem.** There exists a dictionary:

<table>
<thead>
<tr>
<th>Construction</th>
<th>EGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C = A + B )</td>
<td>( C(z) = A(z) + B(z) )</td>
</tr>
<tr>
<td>( C = A \cdot B )</td>
<td>( C(z) = A(z) \cdot B(z) )</td>
</tr>
<tr>
<td>( C = SEQ(A) )</td>
<td>( C(z) = \frac{1}{1 - A(z)} )</td>
</tr>
<tr>
<td>( C = SET(A) )</td>
<td>( C(z) = \exp(A(z)) )</td>
</tr>
<tr>
<td>( C = CYC(A) )</td>
<td>( C(z) = \log \frac{1}{1 - A(z)} )</td>
</tr>
</tbody>
</table>

\( E \) or 1: “neutral class” formed with element of size 0 \( \leftrightarrow E(z) = 1 \).

\( Z \): “atomic class” formed with element of size 1 \( \leftrightarrow E(z) = 1 \).
Seq: $1 + A + A^2 + \cdots = \frac{1}{1 - A}$.

Set: $1 + \frac{A}{1!} + \frac{A^2}{2!} + \cdots = \exp(A)$.

Cyc: $1 + \frac{A}{1} + \frac{A^2}{2} + \cdots = \log \frac{1}{1 - A}$.

(End of proof of Theorem)
— Perms \( \mathcal{P} \cong \text{Seg}(\mathcal{Z}) \)
— Urn \( \mathcal{U} \cong \text{Set}(\mathcal{Z}) \)
— Circulars graphs \( \mathcal{K} \cong \text{Cyc}(\mathcal{Z}) \)

\[ m \text{ times} \]

— \( m \)-functions: \( \mathcal{F}^m \cong \mathcal{U} \ast \cdots \ast \mathcal{U} \cong \text{Seg}_m(\mathcal{U}) \)
— \( m \)-surjections: \( \text{Seg}(\mathcal{V}) \), \( \mathcal{V} = \text{Set}_{\geq 1}(\mathcal{Z}) \)
— Set partitions: \( \text{Set}(\text{Set}_{\geq 1}(\mathcal{Z})) \)
— Lab. trees: \( T = \mathcal{Z} \ast \text{Set}(T) \).
**Example 1.** Permutations and cycles:

\[ \mathcal{P} = \text{SET}(\text{CYC}(\mathcal{Z})) \implies P(z) = \exp \left( \log \frac{1}{1 - z} \right) = \frac{1}{1 - z}. \]

Derangements (no fixed point)

\[ \mathcal{D} = \text{SET}(\text{CYC}(\mathcal{Z}) \setminus \mathcal{Z}) \implies D(z) = \exp \left( \log \frac{1}{1 - z} - z \right) \equiv \frac{e^{-z}}{1 - z}. \]

Thus

\[ \frac{D_n}{n!} = 1 - \frac{1}{1!} + \frac{2}{2!} - \cdots + \frac{(-1)^n}{n!} \sim e^{-1}. \]

**Example 2.** Labelled (Cayley) trees:

\[ T = \mathcal{Z} \star \text{SET}(T) \implies T(z) = ze^{T(z)}. \]

Thus \[ T_n = n^{n-1} \] by Lagrange Inversion Th. 
Example 3. Set partitions:

\[ B = \text{SET}(\text{SET}_{\geq 1}(\mathcal{Z})) \implies B(z) = e^{e^z} - 1. \]

Bell numbers:

\[ B_n = e^{-1} \sum_{k \geq 0} \frac{k^n}{k!}. \]
**Example 4. Allocations to \([1 \ldots m]\):**

- all: \(e^{mz} \sim F_n = m^n\).
- surjective: \((e^z - 1)^m \sim \text{Stirling numbers}, m!{m \choose n} = \sum (m \choose k)(-1)^{m-k} k^n.\)
- injective: \((1 + z)^m \sim {m \choose n} n! \) (arrangement #).

**Exercise: Birthday Problem and Coupon Collector.**

\[
\mathbb{E}(B) = \int_0^\infty \left(1 + \frac{t}{m}\right)^m e^{-t} \, dt, \quad \mathbb{E}(C) = \int_0^\infty \left(e^t - (e^{t/m} - 1)^m\right) e^{-t} \, dt.
\]

Multiple birthdays, multiple collections. (Cf Poissonization.)
**Example 5.** Mappings aka functional graphs = endofunctions of finite set.

\[ T = z e^T, \quad K = \log(1-T)^{-1}, \quad M = e^K : \quad M_n = n^n. \quad P(\text{connected}) = O \left( \frac{1}{\sqrt{n}} \right). \]

**Exercise:** A binary functional graph is such that each \( x \) has either 0 or 2 preimages (cf \( x^2 + a \mod p \)). **q1.** Construct; **q2.** enumerate.

**Exercise:** All graphs \( G(z) = 1 + \sum_{n=1}^{\infty} 2^{n(n-1)/2} z^n / n! \). **q1.** EGF \( K(z) \) of connected graphs? **q2.** Probability of connectedness. **q3**. Prove not constructible.
Additional constructions: *substitution, pointing, order constraints*:

\[ f \circ g, \quad \partial f, \quad \int f. \]
Linear probing hashing: From Knuth’s original derivation (rec.):

\[
M(N, k) = \sum_{n=1}^{N^2} m P(\pi_k, n) = \frac{N(N+1)}{2} \left( 1 - \frac{k-1}{N^2} \right) - \frac{N-k}{N^2} \sum_{x=1}^{N} \left( \frac{1}{N^2} - \frac{1}{x^2} \right) x^2.
\]

\[
= \frac{1}{2} \left[ N(N+1)(1 - \frac{k-1}{N}) - (1 - \frac{k}{N}) \sum_{x=1}^{N} \left( x N^2 \right) \left( 1 - \frac{2}{N} \right) - N \left( 1 - \frac{4}{N} \right)^2 \right] R(N, k, N).
\]

**to symbolic GFs:**

\[
\text{Island} = \mathbb{F} +
\]

\[
I(z) = 1 + \int \frac{\partial}{\partial z} (z I(z)) \times I(z)
\]

Get nonempty island by joining two islands by means of a gluing element.

\[\rightsquigarrow\] wide encompassing extensions of original analyses [F-Poblete-Viola, Pittel, Knuth 1998, Janson, Chassaing-Marckert, ...].
Some constructible families and generating functions

labelled

\[
\text{Tree} = \quad \implies \quad T = z \Phi(T)
\]

- Increasing trees \(\rightsquigarrow Y = \int \Phi(Y)\)
- Mappings \(\rightsquigarrow\)

\[
\begin{align*}
    m &= \text{Set}(K) \\
    K &= \text{Cycle}(T) \\
    T &= Z \ast \text{Set}(T)
\end{align*}
\]

\[
\begin{align*}
    M &= \exp(K) \\
    K &= \log(1 - T)^{-1} \\
    T &= z \exp(T)
\end{align*}
\]
Chapter 3. Parameters and Multivariate GFs

runs in_perms
Discrete Probabilistic Model

\[ E_N \text{ with } k \]

\( \{E_N, k\} \)

\( \chi = 6 \)
\( \chi = 5 \)
\( \chi = 4 \)
\( \chi = 3 \)
\( \chi = 2 \)
\( \chi = 1 \)

\( \text{eg. Trees}_N \text{ with path length} \)

\[ \Pr \{ X = k \} = \frac{E_{N,k}}{E_N} \]

\[ E[X] = \sum k \cdot \frac{E_{N,k}}{E_N} \]

\[ + \text{variance, etc.} \]
Bivariate GF (ordinary) \((E_{n,k})\) \(\leadsto E(z, u) = \sum_{n,k} E_{n,k} u^k z^n.\)

Bivariate GF (exponential) \((E_{n,k})\) \(\leadsto E(z, u) = \sum_{n,k} E_{n,k} u^k \frac{z^n}{n!}.\)

• BGF encodes exact distributions. hence, moments.

\[
\mathbb{E}_{\mathcal{E}_n}[\chi] = \sum_k k \cdot \frac{E_{n,k}}{E_n} = \frac{1}{E_n} \text{coeff}[z^n] \frac{\partial}{\partial u} E(z, u) \bigg|_{u=1}.
\]

Variance & moment of order 2: second derivative, etc.

Chebyshev inequalities: \(\sigma_n/\mu_n \to 0\) implies convergence in probability.
Parameters: via multivariate GFs.

\[
C = \begin{array}{c}
\text{Diagram 1} \\
\text{Diagram 2} \\
\text{Diagram 3} \\
\text{Diagram 4}
\end{array}
\]

\[
C(z, u) = \frac{zzzz}{uuuu} + \frac{zzz}{uuu} + \frac{zzz}{uuu} + \frac{zzzz}{uuuuuuuu} + \frac{z}{u^0}
\]

Also: combinational forms. E.g. for ordinary GFs

\[
C(z, u) = \sum_{i=0}^{\infty} \frac{X[i]}{i!} z^i.
\]
PRINCIPLE: Add variables marking parameters at appropriate places and recycle:

Theorem (Symbolic method)
A dictionary translates constructions into generating functions:

<table>
<thead>
<tr>
<th>Union</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>×</td>
</tr>
<tr>
<td>Sequence</td>
<td>(\frac{1}{1 - \cdots})</td>
</tr>
<tr>
<td>Set</td>
<td>Exp</td>
</tr>
<tr>
<td>Cycle</td>
<td>Log</td>
</tr>
</tbody>
</table>
Conclusions (Part I)

~ [Chapter 3]: Multivariate GFs give access to parameters; those that can be obtained by “marking” in combinatorial constructions.

~ [Chapters 1-2-3]: Exploit all this asymptotically? counting; mean, variance, distribution?