Infinity in Combinatorics: Asymptotic Enumeration

Helmut Prodinger

Stellenbosch

November 4, 2007

I am a discrete mathematician and deal with finite structures.
Why do I talk at a conference where the subject is INFINITY?
OK: finite, but large: what happens?

I am a discrete mathematician and deal with finite structures.
Why do I talk at a conference where the subject is INFINITY?
OK: finite, but large: what happens?

I am a discrete mathematician and deal with finite structures.
Why do I talk at a conference where the subject is INFINITY?
OK: finite, but large: what happens?

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600, 6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula $n!$

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600, 6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula n!

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600, 6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula $n!$

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600,$
$6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula $n!$

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600, 6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula $n!$

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, 479001600, 6227020800, 87178291200, 1307674368000$

What can we say about these numbers?

They grow very fast.

Aha! The numbers are given by the formula $n!$

How fast do they grow?

We want to compare with some "standard quantities"

Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$10! = 3628800$
Stirling: $3598695.61\ldots$

$30! = 265252859812191058636308480000000$
Stirling: $264517095922964306151924784917114$

In what sense is this an approximation?

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$10! = 3628800$
Stirling: $3598695.61\ldots$

$30! = 265252859812191058636308480000000$
Stirling: $264517095922964306151924784917114$

In what sense is this an approximation?

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$10! = 3628800$
Stirling: $3598695.61\ldots$

$30! = 265252859812191058636308480000000$
Stirling: $264517095922964306151924784917114$

In what sense is this an approximation?

Definition:

$$a_n \sim b_n$$

as $n \to \infty$ if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1$$

(asymptotic equivalence)

More precision:

$$a_n = b_n + O(c_n)$$

Meaning:

$$|a_n - b_n| \leq K \cdot c_n$$

This is Bachmann's celebrated big-Oh notation, popularised by Landau.

Why are we interested?

Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!

Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

Why are we interested?
Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!
Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

Why are we interested?
Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!
Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

Why are we interested?
Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!
Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

Why are we interested?
Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!
Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

Why are we interested?
Given

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{k}.$$

First idea: Simplify!
Indeed

$$S_n = H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$$

(harmonic number)

OK, but what about

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{2^k - 1}?$$

1., 1.6666667, 2.1428571, 2.5047619, 2.7941628, 3.0348182, 3.2408489, 3.4210777, 3.5813102, 3.7255593, 3.8567220, 3.9769661, 4.0879615, 4.1910240, 4.2872095, 4.3773779, 4.4622386, 4.5423828, 4.6183084, 4.6904384

????

We need some methods!

OK, but what about

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{2^k - 1}?$$

1., 1.6666667, 2.1428571, 2.5047619, 2.7941628, 3.0348182, 3.2408489, 3.4210777, 3.5813102, 3.7255593, 3.8567220, 3.9769661, 4.0879615, 4.1910240, 4.2872095, 4.3773779, 4.4622386, 4.5423828, 4.6183084, 4.6904384

????

We need some methods!

OK, but what about

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{2^k - 1}?$$

1., 1.6666667, 2.1428571, 2.5047619, 2.7941628, 3.0348182, 3.2408489, 3.4210777, 3.5813102, 3.7255593, 3.8567220, 3.9769661, 4.0879615, 4.1910240, 4.2872095, 4.3773779, 4.4622386, 4.5423828, 4.6183084, 4.6904384

????

We need some methods!

OK, but what about

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{2^k - 1}?$$

$1., 1.6666667, 2.1428571, 2.5047619, 2.7941628, 3.0348182, 3.2408489, 3.4210777, 3.5813102, 3.7255593, 3.8567220, 3.9769661, 4.0879615, 4.1910240, 4.2872095, 4.3773779, 4.4622386, 4.5423828, 4.6183084, 4.6904384$

????

We need some methods!

Why would a combinatorialist be interested?

She encounters combinatorial structures, like permutations, compositions, partitions, trees, lattice paths, words, . . .

She wants to know how many objects *of size n* there are. And sometimes, an exact formula does not exist, or, if it exists, it is not conclusive.

Why would a combinatorialist be interested?
She encounters combinatorial structures, like permutations, compositions, partitions, trees, lattice paths, words, ...
She wants to know how many objects *of size* n there are. And sometimes, an exact formula does not exist, or, if it exists, it is not conclusive.

Why would a combinatorialist be interested?
She encounters combinatorial structures, like permutations, compositions, partitions, trees, lattice paths, words, ...
She wants to know how many objects *of size n* there are. And sometimes, an exact formula does not exist, or, if it exists, it is not conclusive.

Why would a computer scientist be interested?

Standard answer: It is in Knuth's book, and Knuth is interested ...

Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).

Some people: Algorithms performs in $O(n^3)$.

What does that mean? Worst-case behaviour.

Even less precise: "polynomial performance"

People who analyse algorithms:

How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested . . .
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,. . .).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested ...
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested ...
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested ...
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?

Standard answer: It is in Knuth's book, and Knuth is interested ...

Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).

Some people: Algorithms performs in $O(n^3)$.

What does that mean? Worst-case behaviour.

Even less precise: "polynomial performance"

People who analyse algorithms:

How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested ...
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,...).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

Why would a computer scientist be interested?
Standard answer: It is in Knuth's book, and Knuth is interested . . .
Seriously: Her algorithms depends on an input of size n, and she wants to know the performance. (Sorting, searching,. . .).
Some people: Algorithms performs in $O(n^3)$.
What does that mean? Worst-case behaviour.
Even less precise: "polynomial performance"
People who analyse algorithms:
How does the algorithm perform *on average*, with a precisely defined (probabilistic) input model?

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

What shall I do when I encounter an asymptotic problem?

- Maple knows a little bit: asympt(f(n),n)
- Concrete Mathematics (Graham/Knuth/Patashnik) has a section on it.
- Bender wrote a survey in SIAM Review (1974); it is a bit dated.
- Odlyzko wrote a long chapter in the handbook of combinatorics.
- The long-awaited book "Analytic Combinatorics" by Flajolet/Sedgewick will answer many questions.
- The old book by de Bruijn on asymptotics is still useful.
- Textbooks on analytic number theory will perhaps help.

Use generating functions!

Consider

$$\sum_{n\geq 0} a_n z^n$$

and treat it as a function of a complex variable z.

The shortest path between two truths in the real domain passes through the complex domain.

(Hadamard)

Sometimes it is better to consider

$$\sum_{n\geq 0} a_n \frac{z^n}{n!},$$

which is called *exponential generating function.*

Use generating functions!
Consider

$$\sum_{n\geq 0} a_n z^n$$

and treat it as a function of a complex variable z.

The shortest path between two truths in the real domain passes through the complex domain.
(Hadamard)

Sometimes it is better to consider

$$\sum_{n\geq 0} a_n \frac{z^n}{n!},$$

which is called *exponential generating function.*

Use generating functions!
Consider

$$\sum_{n\geq 0} a_n z^n$$

and treat it as a function of a complex variable z.

The shortest path between two truths in the real domain passes through the complex domain.
(Hadamard)

Sometimes it is better to consider

$$\sum_{n\geq 0} a_n \frac{z^n}{n!},$$

which is called *exponential generating function.*

Use generating functions!
Consider

$$\sum_{n\geq 0} a_n z^n$$

and treat it as a function of a complex variable z.
The shortest path between two truths in the real domain passes through the complex domain.
(Hadamard)
Sometimes it is better to consider

$$\sum_{n\geq 0} a_n \frac{z^n}{n!},$$

which is called *exponential generating function.*

Simple example: Fibonacci numbers.
$F_n = F_{n-1} + F_{n-2}$, $F_0 = 0$, $F_1 = 1$.

$$F(z) = \sum_{n \geq 0} F_n z^n = \frac{z}{1 - z - z^2}.$$

Partial fraction decomposition:

$$F(z) = \frac{1}{\sqrt{5}} \frac{1}{1 - z\alpha} - \frac{1}{\sqrt{5}} \frac{1}{1 - z\beta},$$

with

$$\alpha = \frac{1 + \sqrt{5}}{2}, \qquad \beta = \frac{1 - \sqrt{5}}{2}$$

Simple example: Fibonacci numbers.
$F_n = F_{n-1} + F_{n-2}$, $F_0 = 0$, $F_1 = 1$.

$$F(z) = \sum_{n \geq 0} F_n z^n = \frac{z}{1 - z - z^2}.$$

Partial fraction decomposition:

$$F(z) = \frac{1}{\sqrt{5}} \frac{1}{1 - z\alpha} - \frac{1}{\sqrt{5}} \frac{1}{1 - z\beta},$$

with

$$\alpha = \frac{1 + \sqrt{5}}{2}, \qquad \beta = \frac{1 - \sqrt{5}}{2}$$

Simple example: Fibonacci numbers.
$F_n = F_{n-1} + F_{n-2}$, $F_0 = 0$, $F_1 = 1$.

$$F(z) = \sum_{n \geq 0} F_n z^n = \frac{z}{1 - z - z^2}.$$

Partial fraction decomposition:

$$F(z) = \frac{1}{\sqrt{5}} \frac{1}{1 - z\alpha} - \frac{1}{\sqrt{5}} \frac{1}{1 - z\beta},$$

with

$$\alpha = \frac{1 + \sqrt{5}}{2}, \qquad \beta = \frac{1 - \sqrt{5}}{2}$$

$$F_n = [z^n]F(z) = \frac{1}{\sqrt{5}}\alpha^n - \frac{1}{\sqrt{5}}\beta^n.$$

(explicit)

$$F_n \sim \frac{1}{\sqrt{5}}\alpha^n.$$

Why? Since $|\beta| < |\alpha|$, or the pole $1/\alpha$ is closer to the origin than $1/\beta$.

$$F_n = [z^n]F(z) = \frac{1}{\sqrt{5}}\alpha^n - \frac{1}{\sqrt{5}}\beta^n.$$

(explicit)

$$F_n \sim \frac{1}{\sqrt{5}}\alpha^n.$$

Why? Since $|\beta| < |\alpha|$, or the pole $1/\alpha$ is closer to the origin than $1/\beta$.

$$F_n = [z^n]F(z) = \frac{1}{\sqrt{5}}\alpha^n - \frac{1}{\sqrt{5}}\beta^n.$$

(explicit)

$$F_n \sim \frac{1}{\sqrt{5}}\alpha^n.$$

Why? Since $|\beta| < |\alpha|$, or the pole $1/\alpha$ is closer to the origin than $1/\beta$.

Rational functions are easy!

For enumerative generating functions, 0 isn't a pole. We can (in principle) perform partial fraction decomposition, and so we only need to know what

$$[z^n]\frac{1}{(1-z\alpha)^{m+1}}$$

is. A simple shift gives

$$\alpha^n[z^n]\frac{1}{(1-z)^{m+1}},$$

so we only need to deal with the pole 1.

Rational functions are easy!

For enumerative generating functions, 0 isn't a pole. We can (in principle) perform partial fraction decomposition, and so we only need to know what

$$[z^n]\frac{1}{(1-z\alpha)^{m+1}}$$

is. A simple shift gives

$$\alpha^n[z^n]\frac{1}{(1-z)^{m+1}},$$

so we only need to deal with the pole 1.

Rational functions are easy!

For enumerative generating functions, 0 isn't a pole. We can (in principle) perform partial fraction decomposition, and so we only need to know what

$$[z^n]\frac{1}{(1-z\alpha)^{m+1}}$$

is. A simple shift gives

$$\alpha^n[z^n]\frac{1}{(1-z)^{m+1}},$$

so we only need to deal with the pole 1.

$$[z^n]\frac{1}{(1-z)^{m+1}} = \binom{n+m}{m} \sim \frac{n^m}{m!}$$

Even if you cannot locate all poles exactly, you win, if you know that they are far away:

The closer the pole is to the origin, the larger the contribution.

$$[z^n]\frac{1}{(1-z)^{m+1}} = \binom{n+m}{m} \sim \frac{n^m}{m!}$$

Even if you cannot locate all poles exactly, you win, if you know that they are far away:

The closer the pole is to the origin, the larger the contribution.

$$[z^n]\frac{1}{(1-z)^{m+1}} = \binom{n+m}{m} \sim \frac{n^m}{m!}$$

Even if you cannot locate all poles exactly, you win, if you know that they are far away:
The closer the pole is to the origin, the larger the contribution.

Compositions:

$$1 + 2 + 2$$

is a composition of 5;

$$2 + 1 + 2$$

is another one.

How many compositions of n are there?

Compositions:

$$1 + 2 + 2$$

is a composition of 5;

$$2 + 1 + 2$$

is another one.
How many compositions of n are there?

From a composition of $n = c_1 + \cdots + c_k$ we can create 2 compositions of $n + 1$:

$$c_1 + \cdots + (c_k + 1)$$

and

$$c_1 + \cdots + c_k + 1.$$

Hence, since for $n = 1$, there is 1 composition, the answer is 2^{n-1}. This was too easy, no asymptotics needed.

Carlitz Compositions:

Additional restriction: Always $c_i \neq c_{i+1}$. Now it is not so easy to find the generating function $f(z)$. Trick: Use $K(z,u)$, where the coefficient of $z^n u^k$ is the number of Carlitz compositions of n with last summand k. One derives:

$$K(z,u) = \frac{uz}{1-uz} + \frac{uz}{1-uz}K(z,1) - K(z,zu).$$

This functional equation can be solved by iteration; specialising $u=1$ eventually leads to

$$f(z) = K(z,1) = \left(1 + \sum_{j\geq 1} \frac{(-z)^j}{1-z^j}\right)^{-1}$$

This is not a rational function, but "almost"

Carlitz Compositions:
Additional restriction: Always $c_i \neq c_{i+1}$. Now it is not so easy to find the generating function $f(z)$. Trick: Use $K(z,u)$, where the coefficient of $z^n u^k$ is the number of Carlitz compositions of n with last summand k. One derives:

$$K(z,u) = \frac{uz}{1-uz} + \frac{uz}{1-uz}K(z,1) - K(z,zu).$$

This functional equation can be solved by iteration; specialising $u = 1$ eventually leads to

$$f(z) = K(z,1) = \left(1 + \sum_{j\geq 1} \frac{(-z)^j}{1-z^j}\right)^{-1}$$

This is not a rational function, but "almost"

Carlitz Compositions:

Additional restriction: Always $c_i \neq c_{i+1}$. Now it is not so easy to find the generating function $f(z)$. Trick: Use $K(z,u)$, where the coefficient of $z^n u^k$ is the number of Carlitz compositions of n with last summand k. One derives:

$$K(z,u) = \frac{uz}{1-uz} + \frac{uz}{1-uz}K(z,1) - K(z,zu).$$

This functional equation can be solved by iteration; specialising $u = 1$ eventually leads to

$$f(z) = K(z,1) = \left(1 + \sum_{j\geq 1} \frac{(-z)^j}{1-z^j}\right)^{-1}$$

This is not a rational function, but "almost"

Carlitz Compositions:

Additional restriction: Always $c_i \neq c_{i+1}$. Now it is not so easy to find the generating function $f(z)$. Trick: Use $K(z,u)$, where the coefficient of $z^n u^k$ is the number of Carlitz compositions of n with last summand k. One derives:

$$K(z,u) = \frac{uz}{1-uz} + \frac{uz}{1-uz} K(z,1) - K(z,zu).$$

This functional equation can be solved by iteration; specialising $u = 1$ eventually leads to

$$f(z) = K(z,1) = \left(1 + \sum_{j \geq 1} \frac{(-z)^j}{1-z^j}\right)^{-1}$$

This is not a rational function, but "almost"

Carlitz Compositions:
Additional restriction: Always $c_i \neq c_{i+1}$. Now it is not so easy to find the generating function $f(z)$. Trick: Use $K(z,u)$, where the coefficient of $z^n u^k$ is the number of Carlitz compositions of n with last summand k. One derives:

$$K(z,u) = \frac{uz}{1-uz} + \frac{uz}{1-uz} K(z,1) - K(z,zu).$$

This functional equation can be solved by iteration; specialising $u = 1$ eventually leads to

$$f(z) = K(z,1) = \left(1 + \sum_{j \geq 1} \frac{(-z)^j}{1-z^j}\right)^{-1}$$

This is not a rational function, but "almost"

There is a dominant pole ρ: The smallest positive root of the equation

$$\sum_{j\geq 1} \frac{\rho^j}{1+\rho^j} = 1.$$

This can be computed numerically. Consequently

$$[z^n]f(z) \sim C \cdot \beta^n,$$

with 0.45636, $\beta = 1.75024$. (Unrestricted case: 2)

Derangements:

D_n is the number of fix-point free permutations.

In elementary courses, one derives by inclusion-exclusion:

$$D_n = n!\Big(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\Big).$$

One notices the truncated series for $1/e$, and thus $D_n/n! \sim \frac{1}{e}$.

More elegantly, one uses (exponential) generating functions:

Permutations are a collection of cycles, and from basic principles of exponential generating functions:

$$\exp\Big(\log\frac{1}{1-z}\Big) = \frac{1}{1-z}.$$

This is of course trivial and just tells us that there are $n!$ permutations.

Derangements:
D_n is the number of fix-point free permutations.
In elementary courses, one derives by inclusion-exclusion:

$$D_n = n!\Big(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\Big).$$

One notices the truncated series for $1/e$, and thus $D_n/n! \sim \frac{1}{e}$.
More elegantly, one uses (exponential) generating functions:
Permutations are a collection of cycles, and from basic principles of exponential generating functions:

$$\exp\Big(\log\frac{1}{1-z}\Big) = \frac{1}{1-z}.$$

This is of course trivial and just tells us that there are $n!$ permutations.

But: We can handle restrictions: Forbidding fixpoints, means:

$$f(z) := \exp\left(\log\frac{1}{1-z} - z\right) = \frac{e^{-z}}{1-z}.$$

Again: This is not a rational function, but almost: There is a dominant pole at 1.

$$f(z) \sim \frac{e^{-1}}{1-z},$$

and

$$[z^n]f(z) \sim [z^n]\frac{e^{-1}}{1-z} = e^{-1}.$$

There are no other poles. It does NOT mean that our formula is an identity, but that the error is smaller than any M^{-n}.

But: We can handle restrictions: Forbidding fixpoints, means:

$$f(z) := \exp\left(\log \frac{1}{1-z} - z\right) = \frac{e^{-z}}{1-z}.$$

Again: This is not a rational function, but almost: There is a dominant pole at 1.

$$f(z) \sim \frac{e^{-1}}{1-z},$$

and

$$[z^n]f(z) \sim [z^n]\frac{e^{-1}}{1-z} = e^{-1}.$$

There are no other poles. It does NOT mean that our formula is an identity, but that the error is smaller than any M^{-n}.

Binary trees.

Recursion definition: A root and a left resp. right subtree. Consequently

$$B(z) = 1 + zB^2(z)$$

(this is almost a context-free grammar).

Solving

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

This is not rational, but has a square-root singularity at $z = \frac{1}{4}$.

Local expansion: (Puiseaux)

$$B(z) \sim 2 - 2\sqrt{1-4z}.$$

Coefficients:

$$[z^n]\sqrt{1-z} \sim \frac{n^{-3/2}}{\Gamma(-\frac{1}{2})}.$$

Binary trees.

Recursion definition: A root and a left resp. right subtree.

Consequently

$$B(z) = 1 + zB^2(z)$$

(this is almost a context-free grammar).

Solving

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

This is not rational, but has a square-root singularity at $z = \frac{1}{4}$.

Local expansion: (Puiseaux)

$$B(z) \sim 2 - 2\sqrt{1-4z}.$$

Coefficients:

$$[z^n]\sqrt{1-z} \sim \frac{n^{-3/2}}{\Gamma(-\frac{1}{2})}.$$

Binary trees.
Recursion definition: A root and a left resp. right subtree.
Consequently

$$B(z) = 1 + zB^2(z)$$

(this is almost a context-free grammar).
Solving

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

This is not rational, but has a square-root singularity at $z = \frac{1}{4}$.
Local expansion: (Puiseaux)

$$B(z) \sim 2 - 2\sqrt{1 - 4z}.$$

Coefficients:

$$[z^n]\sqrt{1 - z} \sim \frac{n^{-3/2}}{\Gamma(-\frac{1}{2})}.$$

More generally:

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}.$$

For us (binary trees)

$$\sim \frac{1}{\sqrt{\pi} n^{3/2}} 4^n$$

(Also from the explicit formula $\frac{1}{n+1}\binom{2n}{n}$).

Local expansions can be *translated* into asymptotic expansions of the coefficients. (Under mild technical conditions.)

SINGULARITY ANALYSIS OF GENERATING FUNCTIONS.

(Flajolet and Odlyzko)

More generally:

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}.$$

For us (binary trees)

$$\sim \frac{1}{\sqrt{\pi} n^{3/2}} 4^n$$

(Also from the explicit formula $\frac{1}{n+1}\binom{2n}{n}$).
Local expansions can be *translated* into asymptotic expansions of the coefficients. (Under mild technical conditions.)
SINGULARITY ANALYSIS OF GENERATING FUNCTIONS. (Flajolet and Odlyzko)

More generally:

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}.$$

For us (binary trees)

$$\sim \frac{1}{\sqrt{\pi} n^{3/2}} 4^n$$

(Also from the explicit formula $\frac{1}{n+1}\binom{2n}{n}$).
Local expansions can be *translated* into asymptotic expansions of the coefficients. (Under mild technical conditions.)
SINGULARITY ANALYSIS OF GENERATING FUNCTIONS.
(Flajolet and Odlyzko)

So one needs a catalogue of known expansions:

$$\frac{1}{(1-z)^\alpha} \log^\beta \frac{1}{1-z}$$

and perhaps other ones.

Another occurrence of a square-root singularity: the tree function.

Implicitly defined: $y = ze^y$.

$$y(z) = \sum_{n\geq 1} a_n \frac{z^n}{n!}$$

Where is a singularity?

Where the implicit function theorem ceases to apply!

Differentiate w.r.t y:

Another occurrence of a square-root singularity: the tree function.
Implicitly defined: $y = ze^y$.

$$y(z) = \sum_{n \geq 1} a_n \frac{z^n}{n!}$$

Where is a singularity?
Where the implicit function theorem ceases to apply!
Differentiate w.r.t y:

Another occurrence of a square-root singularity: the tree function.
Implicitly defined: $y = ze^y$.

$$y(z) = \sum_{n \geq 1} a_n \frac{z^n}{n!}$$

Where is a singularity?
Where the implicit function theorem ceases to apply!
Differentiate w.r.t y:

Another occurrence of a square-root singularity: the tree function.
Implicitly defined: $y = ze^y$.

$$y(z) = \sum_{n\geq 1} a_n \frac{z^n}{n!}$$

Where is a singularity?
Where the implicit function theorem ceases to apply!
Differentiate w.r.t y:

$$y = ze^y$$
$$1 = ze^y$$

Consequently $y = \tau = 1$, and $z = \rho = \frac{1}{e}$. The singularity is at $z = 1/e$, and there is a local expansion

$$y(z) \sim 1 - \sqrt{2}\sqrt{1 - ez}$$

It is a fact that the precise knowledge of the behaviour of an analytic function in the vicinity of its singular points is a source of arithmetic properties. (Erich Hecke)

Basic principle:

- location and type of singular points
- local expansion
- translate ("Transfer theorems")

It is a fact that the precise knowledge of the behaviour of an analytic function in the vicinity of its singular points is a source of arithmetic properties. (Erich Hecke)

Basic principle:

- location and type of singular points
- local expansion
- translate ("Transfer theorems")

It is a fact that the precise knowledge of the behaviour of an analytic function in the vicinity of its singular points is a source of arithmetic properties. (Erich Hecke)

Basic principle:

- location and type of singular points
- local expansion
- translate ("Transfer theorems")

It is a fact that the precise knowledge of the behaviour of an analytic function in the vicinity of its singular points is a source of arithmetic properties. (Erich Hecke)

Basic principle:

- location and type of singular points
- local expansion
- translate ("Transfer theorems")

Longest runs in binary words.

We are interested in binary words (letters a, b), such that no run of k consecutive a's can be found.

$$a^{<k} = \varepsilon + a + a^2 + \cdots + a^{k-1}$$

Decomposition:

$$L = a^{<k}\left(ba^{<k}\right)^*$$

Translate into generating functions:

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

Longest runs in binary words.

We are interested in binary words (letters a, b), such that no run of k consecutive a's can be found.

$$a^{<k} = \varepsilon + a + a^2 + \cdots + a^{k-1}$$

Decomposition:

$$L = a^{<k}\left(ba^{<k}\right)^*$$

Translate into generating functions:

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

Longest runs in binary words.

We are interested in binary words (letters a, b), such that no run of k consecutive a's can be found.

$$a^{<k} = \varepsilon + a + a^2 + \cdots + a^{k-1}$$

Decomposition:

$$L = a^{<k}\left(ba^{<k}\right)^*$$

Translate into generating functions:

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

This is a rational function, but where are the poles?
Equation:

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

“bootstrapping”:

$$\rho_k \sim \frac{1}{2}$$

Now plug in $\rho_k = \frac{1}{2} + \varepsilon_k$

$$\frac{1}{2} + \varepsilon_k \sim \frac{1}{2}(1 + \frac{1}{2^{k+1}}) \quad \text{or} \quad \varepsilon_k \sim \frac{1}{2^{k+2}}$$

This is a rational function, but where are the poles?
Equation:

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

"bootstrapping":

$$\rho_k \sim \frac{1}{2}$$

Now plug in $\rho_k = \frac{1}{2} + \varepsilon_k$

$$\frac{1}{2} + \varepsilon_k \sim \frac{1}{2}(1 + \frac{1}{2^{k+1}}) \quad \text{or} \quad \varepsilon_k \sim \frac{1}{2^{k+2}}$$

This is a rational function, but where are the poles?
Equation:

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

"bootstrapping":

$$\rho_k \sim \frac{1}{2}$$

Now plug in $\rho_k = \frac{1}{2} + \varepsilon_k$

$$\frac{1}{2} + \varepsilon_k \sim \frac{1}{2}(1 + \frac{1}{2^{k+1}}) \quad \text{or} \quad \varepsilon_k \sim \frac{1}{2^{k+2}}$$

This is a rational function, but where are the poles?
Equation:

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

“bootstrapping”:

$$\rho_k \sim \frac{1}{2}$$

Now plug in $\rho_k = \frac{1}{2} + \varepsilon_k$

$$\frac{1}{2} + \varepsilon_k \sim \frac{1}{2}(1 + \frac{1}{2^{k+1}}) \quad \text{or} \quad \varepsilon_k \sim \frac{1}{2^{k+2}}$$

One can show that the other poles are "far away" (Rouché's theorem!)

$$\frac{1 - z^k}{1 - 2z + z^{k+1}} \sim \frac{\text{constant}}{1 - z/\rho_k}$$

The coefficient of z^n:

$$\sim \text{constant}\rho_k^{-n}$$

Probabilities: divide by 2^n

$$\sim \text{constant}(2\rho_k)^{-n} \sim \text{constant}(1+\frac{1}{2^{k+1}})^{-n} \sim \text{constant}\exp(-\frac{n}{2^{k+1}})$$

One can show that the other poles are "far away" (Rouché's theorem!)

$$\frac{1 - z^k}{1 - 2z + z^{k+1}} \sim \frac{\text{constant}}{1 - z/\rho_k}$$

The coefficient of z^n:

$$\sim \text{constant}\rho_k^{-n}$$

Probabilities: divide by 2^n

$$\sim \text{constant}(2\rho_k)^{-n} \sim \text{constant}(1 + \frac{1}{2^{k+1}})^{-n} \sim \text{constant}\exp(-\frac{n}{2^{k+1}})$$

One can show that the other poles are "far away" (Rouché's theorem!)

$$\frac{1-z^k}{1-2z+z^{k+1}} \sim \frac{\text{constant}}{1-z/\rho_k}$$

The coefficient of z^n:

$$\sim \text{constant}\rho_k^{-n}$$

Probabilities: divide by 2^n

$$\sim \text{constant}(2\rho_k)^{-n} \sim \text{constant}(1+\frac{1}{2^{k+1}})^{-n} \sim \text{constant}\exp(-\frac{n}{2^{k+1}})$$

One can show that the other poles are "far away" (Rouché's theorem!)

$$\frac{1 - z^k}{1 - 2z + z^{k+1}} \sim \frac{\text{constant}}{1 - z/\rho_k}$$

The coefficient of z^n:

$$\sim \text{constant}\rho_k^{-n}$$

Probabilities: divide by 2^n

$$\sim \text{constant}(2\rho_k)^{-n} \sim \text{constant}(1+\frac{1}{2^{k+1}})^{-n} \sim \text{constant}\exp(-\frac{n}{2^{k+1}})$$

One can show that the other poles are "far away" (Rouché's theorem!)

$$\frac{1 - z^k}{1 - 2z + z^{k+1}} \sim \frac{\text{constant}}{1 - z/\rho_k}$$

The coefficient of z^n:

$$\sim \text{constant}\rho_k^{-n}$$

Probabilities: divide by 2^n

$$\sim \text{constant}(2\rho_k)^{-n} \sim \text{constant}(1+\frac{1}{2^{k+1}})^{-n} \sim \text{constant}\exp(-\frac{n}{2^{k+1}})$$

If one works out the details, in order to compute the average of the longest run of a's:

$$E_n \sim \sum_{k\geq 1}\left(1 - e^{-n/2^k}\right).$$

The asymptotics of this is harder than what we have seen so far!

Something similar is of relevance in von Neumann's addition scheme.

Two binary numbers x and y of length n are added as follows: $s = s_n \dots s_0$, where $s_i = x_i + y_i \mod 2$. The carries will be collected into a separate string (number) c. The process is repeated by adding s and c. It stops when no more carries are there.

The number of iterations is the parameter of interest. If one adds x and y bitwise, it is the longest sequence of 1's, which has a 2 at the right end.

Something similar is of relevance in von Neumann's addition scheme.

Two binary numbers x and y of length n are added as follows: $s = s_n \dots s_0$, where $s_i = x_i + y_i \mod 2$. The carries will be collected into a separate string (number) c. The process is repeated by adding s and c. It stops when no more carries are there.

The number of iterations is the parameter of interest. If one adds x and y bitwise, it is the longest sequence of 1's, which has a 2 at the right end.

Something similar is of relevance in von Neumann's addition scheme.
Two binary numbers x and y of length n are added as follows: $s = s_n \dots s_0$, where $s_i = x_i + y_i \mod 2$. The carries will be collected into a separate string (number) c. The process is repeated by adding s and c. It stops when no more carries are there.
The number of iterations is the parameter of interest. If one adds x and y bitwise, it is the longest sequence of 1's, which has a 2 at the right end.

The height of planar trees.

The average (planar trees with $n+1$ nodes):

$$E_n = (n+1)\sum_{k\geq 1} d(k)\frac{\binom{2n}{n+1-k} - 2\binom{2n}{n-k} + \binom{2n}{n-1-k}}{\binom{2n}{n}}$$

$d(k)$ is the number of (positive) divisors of k.

Approximation:

$$E_n \sim \sum_{k\geq 1} d(k)\Big(-2 + 4\frac{k^2}{n}\Big)e^{-k^2/n}$$

The technique for this and the average length of the longest run of ones is the Mellin transform!

The height of planar trees.

The average (planar trees with $n+1$ nodes):

$$E_n = (n+1) \sum_{k \geq 1} d(k) \frac{\binom{2n}{n+1-k} - 2\binom{2n}{n-k} + \binom{2n}{n-1-k}}{\binom{2n}{n}}$$

$d(k)$ is the number of (positive) divisors of k.

Approximation:

$$E_n \sim \sum_{k \geq 1} d(k) \Big(-2 + 4\frac{k^2}{n} \Big) e^{-k^2/n}$$

The technique for this and the average length of the longest run of ones is the Mellin transform!

The height of planar trees.

The average (planar trees with $n+1$ nodes):

$$E_n = (n+1)\sum_{k\geq 1} d(k)\frac{\binom{2n}{n+1-k} - 2\binom{2n}{n-k} + \binom{2n}{n-1-k}}{\binom{2n}{n}}$$

$d(k)$ is the number of (positive) divisors of k.

Approximation:

$$E_n \sim \sum_{k\geq 1} d(k)\Big(-2 + 4\frac{k^2}{n}\Big)e^{-k^2/n}$$

The technique for this and the average length of the longest run of ones is the Mellin transform!

Ramanujan's Q-function.

$$Q(n) = 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots$$

$$Q(n) \sim ?$$

Not so easy. It is also related to a famous open question by Ramanujan (that I solved together with some coauthors).

The function appears in many contexts, e.g., in the birthday paradox.

It can be attacked by the Laplace method for integrals, or otherwise.

$$Q(n) \sim \sqrt{\frac{\pi n}{2}}$$

Ramanujan's Q-function.

$$Q(n) = 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots$$

$$Q(n) \sim ?$$

Not so easy. It is also related to a famous open question by Ramanujan (that I solved together with some coauthors).
The function appears in many contexts, e.g., in the birthday paradox.
It can be attacked by the Laplace method for integrals, or otherwise.

$$Q(n) \sim \sqrt{\frac{\pi n}{2}}$$

Ramanujan's Q-function.

$$Q(n) = 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots$$

$$Q(n) \sim ?$$

Not so easy. It is also related to a famous open question by Ramanujan (that I solved together with some coauthors).
The function appears in many contexts, e.g., in the birthday paradox.
It can be attacked by the Laplace method for integrals, or otherwise.

$$Q(n) \sim \sqrt{\frac{\pi n}{2}}$$

In discrete mathematics, we deal more with sums than with integrals.

However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i}\int_{\mathcal{C}} f(z)\frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.

However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i}\int_{\mathcal{C}} f(z)\frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.

However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.
However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.
However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.
However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

In discrete mathematics, we deal more with sums than with integrals.
However, Cauchy's integral formula extracts coefficients from power series:

$$[z^n]f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

A carefully crafted path of integration (going through saddle-points via steepest descents) can be used to estimate the integral.

Mellin analysis

$$E_n \sim \sum_{k\geq 1} d(k)\Big(-2+4\frac{k^2}{n}\Big)e^{-k^2/n}$$

$$E_n \sim \sum_{k\geq 1}\left(1-e^{-n/2^k}\right).$$

Another one:

$$E_n \sim \sum_{k\geq 1} v_2(k)\Big(-2 + 4\frac{k^2}{n}\Big) e^{-k^2/n}$$

where $v_2(k)$ is the number of trailing zeroes of k when written in binary. Equivalently, the highest j such that 2^j divides k. The last expression occurs in the analysis of the *register function*. This is the minimal number of extra registers (optimal strategy) to evaluate a given binary (expression) tree. This was invented by Ershov, but appeared earlier in the natural sciences (Horton, Strahler).

Another one:

$$E_n \sim \sum_{k\geq 1} v_2(k)\Big(-2+4\frac{k^2}{n}\Big)e^{-k^2/n}$$

where $v_2(k)$ is the number of trailing zeroes of k when written in binary. Equivalently, the highest j such that 2^j divides k. The last expression occurs in the analysis of the *register function*. This is the minimal number of extra registers (optimal strategy) to evaluate a given binary (expression) tree. This was invented by Ershov, but appeared earlier in the natural sciences (Horton, Strahler).

After a substitution:

$$E_n \sim \sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$

$$E_n \sim \sum_{k\geq 1} \Big(1-e^{-x/2^k}\Big)$$

$$E_n \sim \sum_{k\geq 1} v_2(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$

Shape:

$$F(x) = \sum_{k\geq 1} a_k f(b_k x)$$

("harmonic sum")

After a substitution:

$$E_n \sim \sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$

$$E_n \sim \sum_{k\geq 1} \left(1-e^{-x/2^k}\right)$$

$$E_n \sim \sum_{k\geq 1} v_2(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$

Shape:

$$F(x) = \sum_{k\geq 1} a_k f(b_k x)$$

("harmonic sum")

Mellin transform

$$f^*(s) = \int_0^\infty f(x)x^{s-1}dx$$

Inversion formula

$$f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f^*(s)x^{-s}ds,$$

c in a suitable interval ("critical strip").

Mellin transform

$$f^*(s) = \int_0^\infty f(x) x^{s-1} dx$$

Inversion formula

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s} ds,$$

c in a suitable interval ("critical strip").

A simple change of variable:

$$f(ax) \longrightarrow a^{-s} f^*(s)$$

and therefore

$$F(x) = \sum_{k \geq 1} a_k f(b_k x) \longrightarrow \sum_{k \geq 1} a_k b_k^{-s} \cdot f^*(s).$$

$$\sum_{k \geq 1} a_k b_k^{-s}$$

is a (generalised) *Dirichlet* series and often admits a nice form in terms of classical functions.

A simple change of variable:

$$f(ax) \longrightarrow a^{-s} f^*(s)$$

and therefore

$$F(x) = \sum_{k \geq 1} a_k f(b_k x) \longrightarrow \sum_{k \geq 1} a_k b_k^{-s} \cdot f^*(s).$$

$$\sum_{k \geq 1} a_k b_k^{-s}$$

is a (generalised) *Dirichlet* series and often admits a nice form in terms of classical functions.

A simple change of variable:

$$f(ax) \longrightarrow a^{-s} f^*(s)$$

and therefore

$$F(x) = \sum_{k \geq 1} a_k f(b_k x) \longrightarrow \sum_{k \geq 1} a_k b_k^{-s} \cdot f^*(s).$$

$$\sum_{k \geq 1} a_k b_k^{-s}$$

is a (generalised) *Dirichlet* series and often admits a nice form in terms of classical functions.

Example:

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2} \longrightarrow$$

$$\sum_{k\geq 1} d(k)k^{-s}\left(\Big(-2+4x^2\Big)e^{-x^2}\right)^*(s)$$

Both parts are readily computed:

$$\sum_{k\geq 1} d(k)k^{-s} = \zeta^2(s)$$

and

$$\left(\Big(-2+4x^2\Big)e^{-x^2}\right)^*(s) = (s-1)\Gamma\Big(\frac{s}{2}\Big)$$

Example:

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2} \longrightarrow$$

$$\sum_{k\geq 1} d(k)k^{-s}\bigg(\Big(-2+4x^2\Big)e^{-x^2}\bigg)^*(s)$$

Both parts are readily computed:

$$\sum_{k\geq 1} d(k)k^{-s} = \zeta^2(s)$$

and

$$\bigg(\Big(-2+4x^2\Big)e^{-x^2}\bigg)^*(s) = (s-1)\Gamma\Big(\frac{s}{2}\Big)$$

Example:

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2} \longrightarrow$$

$$\sum_{k\geq 1} d(k)k^{-s}\bigg(\Big(-2+4x^2\Big)e^{-x^2}\bigg)^{*}(s)$$

Both parts are readily computed:

$$\sum_{k\geq 1} d(k)k^{-s} = \zeta^2(s)$$

and

$$\bigg(\Big(-2+4x^2\Big)e^{-x^2}\bigg)^{*}(s) = (s-1)\Gamma\Big(\frac{s}{2}\Big)$$

Example:

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2} \longrightarrow$$

$$\sum_{k\geq 1} d(k)k^{-s}\left(\Big(-2+4x^2\Big)e^{-x^2}\right)^*(s)$$

Both parts are readily computed:

$$\sum_{k\geq 1} d(k)k^{-s} = \zeta^2(s)$$

and

$$\left(\Big(-2+4x^2\Big)e^{-x^2}\right)^*(s) = (s-1)\Gamma\Big(\frac{s}{2}\Big)$$

So

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

The integral is computed (with an error) by *residues*.
(Technical remark. The line of integration is either shifted to the left or right, depending on $x \to 0$ or $x \to \infty$.)
The average height of a random planar tree with n nodes is $\sim \sqrt{\pi n}$.

So

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

The integral is computed (with an error) by *residues*.
(Technical remark. The line of integration is either shifted to the left or right, depending on $x\to 0$ or $x\to\infty$.)
The average height of a random planar tree with n nodes is $\sim\sqrt{\pi n}$.

So

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$= \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

The integral is computed (with an error) by *residues*.
(Technical remark. The line of integration is either shifted to the left or right, depending on $x \to 0$ or $x \to \infty$.)
The average height of a random planar tree with n nodes is $\sim \sqrt{\pi n}$.

So

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$= \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

The integral is computed (with an error) by *residues*.
(Technical remark. The line of integration is either shifted to the left or right, depending on $x \to 0$ or $x \to \infty$.)
The average height of a random planar tree with n nodes is $\sim \sqrt{\pi n}$.

So

$$\sum_{k\geq 1} d(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

The integral is computed (with an error) by *residues*.
(Technical remark. The line of integration is either shifted to the left or right, depending on $x \to 0$ or $x \to \infty$.)
The average height of a random planar tree with n nodes is $\sim \sqrt{\pi n}$.

If one does the same for the register instance, one gets

$$\sum_{k\geq 1} v_2(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{\zeta(s)}{2^s-1}(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

And now, there are poles at $\frac{2\pi ik}{\log 2}$ (solutions of $2^s=1$).
The corresponding residues altogether can be collected into a periodic function.

If one does the same for the register instance, one gets

$$\sum_{k\geq 1} v_2(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{\zeta(s)}{2^s-1}(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

And now, there are poles at $\frac{2\pi ik}{\log 2}$ (solutions of $2^s=1$).
The corresponding residues altogether can be collected into a periodic function.

If one does the same for the register instance, one gets

$$\sum_{k\geq 1} v_2(k)\Big(-2+4k^2x^2\Big)e^{-k^2x^2}$$
$$=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{\zeta(s)}{2^s-1}(s-1)\Gamma\Big(\frac{s}{2}\Big)x^{-s}ds$$

And now, there are poles at $\frac{2\pi ik}{\log 2}$ (solutions of $2^s=1$).
The corresponding residues altogether can be collected into a periodic function.

Final answer: The average number of registers is asymptotic to

$$\log_4 n + \delta(\log_4 n),$$

with $\delta(x)$ a 1-periodic function.

Such a periodic term appears also in the average of the maximal runs.

One of my favourites (and of M Archibald, the next speaker):

Rice's method:

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{n!(-1)^n}{z(z-1)\dots(z-n)} f(z) dz$$

"n-th difference of the sequence $f(n)$"

Looks like a curiosity, but appears surprisingly often.

Example: study of tries and patricia tries, $f(k) = \frac{k-1}{2^{k-1}-1}$, etc.

Integral is again evaluated by changing the contour and evaluating residues.

One of my favourites (and of M Archibald, the next speaker): Rice's method:

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{n!(-1)^n}{z(z-1)\dots(z-n)} f(z) dz$$

"n-th difference of the sequence $f(n)$"

Looks like a curiosity, but appears surprisingly often.

Example: study of tries and patricia tries, $f(k) = \frac{k-1}{2^{k-1}-1}$, etc.

Integral is again evaluated by changing the contour and evaluating residues.

One of my favourites (and of M Archibald, the next speaker):
Rice's method:

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{n!(-1)^n}{z(z-1)\dots(z-n)} f(z) dz$$

"n-th difference of the sequence $f(n)$"
Looks like a curiosity, but appears surprisingly often.
Example: study of tries and patricia tries, $f(k) = \frac{k-1}{2^{k-1}-1}$, etc.
Integral is again evaluated by changing the contour and evaluating residues.

One of my favourites (and of M Archibald, the next speaker):
Rice's method:

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{n!(-1)^n}{z(z-1)\dots(z-n)} f(z) dz$$

"n-th difference of the sequence $f(n)$"
Looks like a curiosity, but appears surprisingly often.
Example: study of tries and patricia tries, $f(k) = \frac{k-1}{2^{k-1}-1}$, etc.
Integral is again evaluated by changing the contour and evaluating residues.

One of my favourites (and of M Archibald, the next speaker): Rice's method:

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{n!(-1)^n}{z(z-1)\dots(z-n)} f(z) dz$$

"n-th difference of the sequence $f(n)$"
Looks like a curiosity, but appears surprisingly often.
Example: study of tries and patricia tries, $f(k) = \frac{k-1}{2^{k-1}-1}$, etc.
Integral is again evaluated by changing the contour and evaluating residues.

Thank you!