Catalan Numbers

Richard P. Stanley

March 20, 2018

An OEIS entry

OEIS: *Online Encyclopedia of Integer Sequences* (**Neil Sloane**). See http://oeis.org. A database of over 270,000 sequences of integers.

An OEIS entry

OEIS: *Online Encyclopedia of Integer Sequences* (**Neil Sloane**). See `http://oeis.org`. A database of over 270,000 sequences of integers.

A000108: $1, 1, 2, 5, 14, 42, 132, 429, \ldots$

$C_0 = 1,\ C_1 = 2,\ C_2 = 3,\ C_3 = 5,\ C_4 = 14, \ldots$

C_n is a **Catalan number**.

An OEIS entry

OEIS: *Online Encyclopedia of Integer Sequences* (**Neil Sloane**). See http://oeis.org. A database of over 270,000 sequences of integers.

A000108: $1, 1, 2, 5, 14, 42, 132, 429, \ldots$

$C_0 = 1,\ C_1 = 2,\ C_2 = 3,\ C_3 = 5,\ C_4 = 14, \ldots$

$\boldsymbol{C_n}$ is a **Catalan number**.

Comments. ... This is probably the longest entry in OEIS, and rightly so.

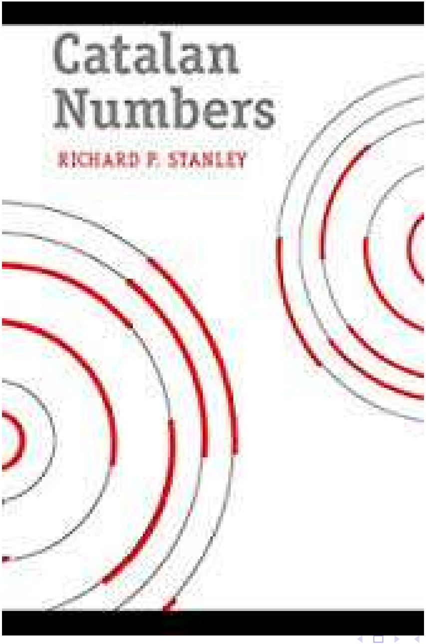
Catalan monograph

R. Stanley, *Catalan Numbers*, Cambridge University Press, 2015.

Catalan monograph

R. Stanley, *Catalan Numbers*, Cambridge University Press, 2015.

Includes 214 combinatorial interpretations of C_n and 68 additional problems.


Catalan
Numbers
RICHARD P. STANLEY

History

Sharabiin Myangat, also known as **Minggatu**, **Ming'antu** (明安图), and **Jing An** (c. 1692–c. 1763): a Mongolian astronomer, mathematician, and topographic scientist who worked at the Qing court in China.

History

Sharabiin Myangat, also known as **Minggatu**, **Ming'antu** (明安图), and **Jing An** (c. 1692–c. 1763): a Mongolian astronomer, mathematician, and topographic scientist who worked at the Qing court in China.

Typical result (1730's):

$$\sin(2\alpha) = 2\sin\alpha - \sum_{n=1}^{\infty} \frac{C_{n-1}}{4^{n-1}} \sin^{2n+1}\alpha$$

History

Sharabiin Myangat, also known as **Minggatu**, **Ming'antu** (明安图), and **Jing An** (c. 1692–c. 1763): a Mongolian astronomer, mathematician, and topographic scientist who worked at the Qing court in China.

Typical result (1730's):

$$\sin(2\alpha) = 2\sin\alpha - \sum_{n=1}^{\infty} \frac{C_{n-1}}{4^{n-1}} \sin^{2n+1}\alpha$$

First example of an infinite trigonometric series.

History

Sharabiin Myangat, also known as **Minggatu**, **Ming'antu** (明安图), and **Jing An** (c. 1692–c. 1763): a Mongolian astronomer, mathematician, and topographic scientist who worked at the Qing court in China.

Typical result (1730's):

$$\sin(2\alpha) = 2\sin\alpha - \sum_{n=1}^{\infty} \frac{C_{n-1}}{4^{n-1}} \sin^{2n+1}\alpha$$

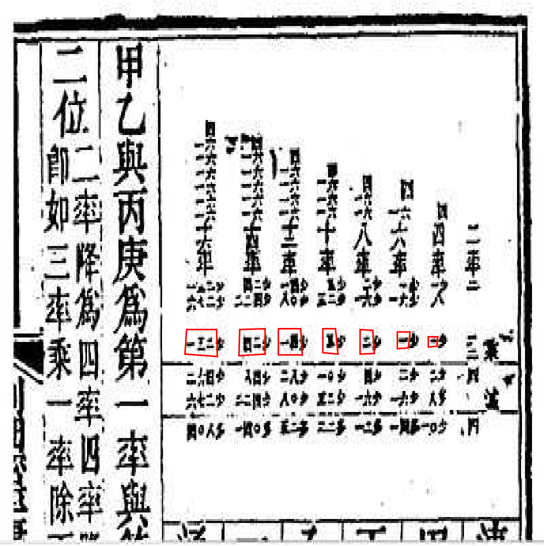
First example of an infinite trigonometric series.

No combinatorics, no further work in China.

Ming'antu

Manuscript of Ming'antu

Manuscript of Ming'antu



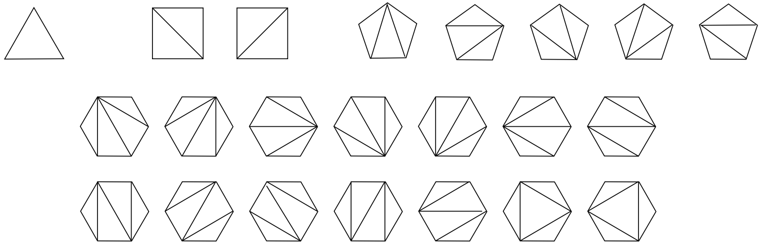
Manuscript of Ming'antu


甲乙與丙庚為第一率與
二位
二率降為四率四率
即如三率乘一率除一

少

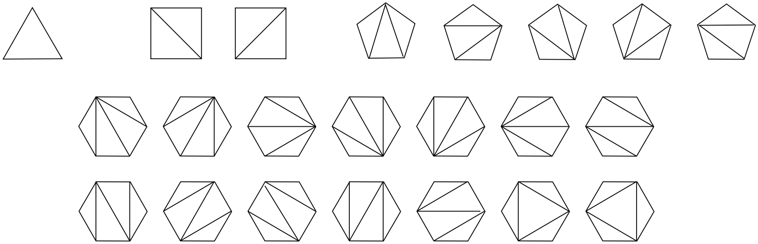
More history, via Igor Pak

- **Euler** (1751): conjectured formula for the number of triangulations of a convex $(n+2)$-gon. In other words, draw $n-1$ noncrossing diagonals of a convex polygon with $n+2$ sides.



More history, via Igor Pak

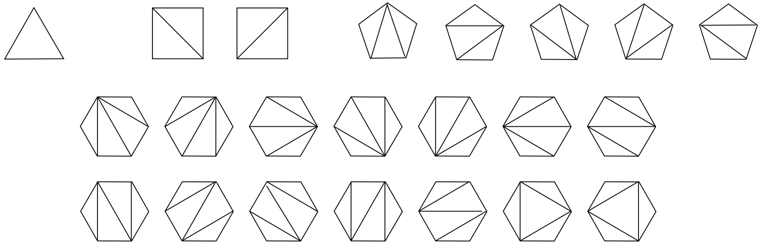
- **Euler** (1751): conjectured formula for the number of triangulations of a convex $(n+2)$-gon. In other words, draw $n-1$ noncrossing diagonals of a convex polygon with $n+2$ sides.



$$1, \quad 2, \quad 5, \quad 14, \quad \ldots$$

More history, via Igor Pak

- **Euler** (1751): conjectured formula for the number of triangulations of a convex $(n+2)$-gon. In other words, draw $n-1$ noncrossing diagonals of a convex polygon with $n+2$ sides.



$$1, \quad 2, \quad 5, \quad 14, \quad \ldots$$

We **define** these numbers to be the Catalan numbers C_n.

Completion of proof

- **Goldbach and Segner** (1758–1759): helped Euler complete the proof, in pieces.
- **Lamé** (1838): first self-contained, complete proof.

Catalan

- **Eugène Charles Catalan** (1838): wrote C_n in the form $\frac{(2n)!}{n!\,(n+1)!}$ and showed it counted (nonassociative) **bracketings** (or **parenthesizations**) of a string of $n+1$ letters.

Catalan

- **Eugène Charles Catalan** (1838): wrote C_n in the form $\frac{(2n)!}{n!\,(n+1)!}$ and showed it counted (nonassociative) **bracketings** (or **parenthesizations**) of a string of $n+1$ letters.

Born in 1814 in Bruges (now in Belgium, then under Dutch rule). Studied in France and worked in France and Liège, Belgium. Died in Liège in 1894.

Why "Catalan numbers"?

- **John Riordan** (1948): introduced the term "Catalan number" in *Math Reviews*.

Why "Catalan numbers"?

- **John Riordan** (1948): introduced the term "Catalan number" in *Math Reviews*.
- **Riordan** (1964): used the term again in *Math. Reviews*.

Why "Catalan numbers"?

- **John Riordan** (1948): introduced the term "Catalan number" in *Math Reviews*.
- **Riordan** (1964): used the term again in *Math. Reviews*.
- **Riordan** (1968): used the term in his book *Combinatorial Identities*. Finally caught on.

Why “Catalan numbers”?

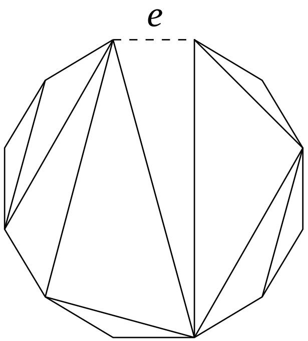
- **John Riordan** (1948): introduced the term “Catalan number” in *Math Reviews*.
- **Riordan** (1964): used the term again in *Math. Reviews*.
- **Riordan** (1968): used the term in his book *Combinatorial Identities*. Finally caught on.
- **Martin Gardner** (1976): used the term in his Mathematical Games column in *Scientific American*. Real popularity began.

The primary recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1$$

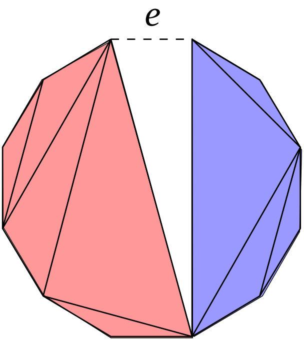
The primary recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1$$



The primary recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1$$



Solving the recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1$$

Let $\boldsymbol{y} = \sum_{n\geq 0} C_n x^n$ (**generating function**).

$$\Rightarrow xy^2 - y + 1 = 0$$

Solving the recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1$$

Let $y = \sum_{n\geq 0} C_n x^n$ (**generating function**).

$$\Rightarrow xy^2 - y + 1 = 0$$

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!\,(n+1)!}$$

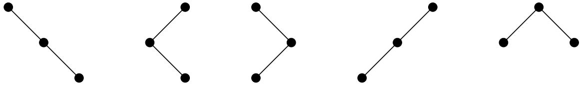
Other combinatorial interpretations

$$\begin{aligned} \mathcal{P}_n &:= \{\text{triangulations of convex } (n+2)\text{-gon}\} \\ \Rightarrow \#\mathcal{P}_n &= C_n \text{ (where } \#S = \text{number of elements of } S) \end{aligned}$$

We want other combinatorial interpretations of C_n, i.e., other sets $\mathcal{S}_n$ for which $C_n = \#\mathcal{S}_n$.

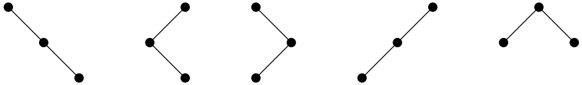
“Transparent” interpretations

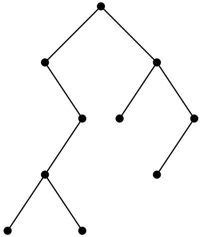
4. **Binary trees** with *n* vertices (each vertex has a left subtree and a right subtree, which may be empty)



“Transparent” interpretations

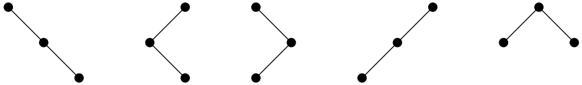
4. Binary trees with n vertices (each vertex has a left subtree and a right subtree, which may be empty)

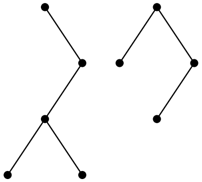




“Transparent” interpretations

4. Binary trees with n vertices (each vertex has a left subtree and a right subtree, which may be empty)





Binary parenthesizations

3. Binary **parenthesizations** or **bracketings** of a string of $n+1$ letters

$$(xx \cdot x)x \quad x(xx \cdot x) \quad (x \cdot xx)x \quad x(x \cdot xx) \quad xx \cdot xx$$

Binary parenthesizations

3. Binary **parenthesizations** or **bracketings** of a string of $n+1$ letters

$$(xx\cdot x)x \quad x(xx\cdot x) \quad (x\cdot xx)x \quad x(x\cdot xx) \quad xx\cdot xx$$

$$((x(xx))x)(x((xx)(xx)))$$

Binary parenthesizations

3. Binary **parenthesizations** or **bracketings** of a string of $n+1$ letters

$$(xx \cdot x)x \quad x(xx \cdot x) \quad (x \cdot xx)x \quad x(x \cdot xx) \quad xx \cdot xx$$

$$\mathbf{((x(xx))x)(x((xx)(xx)))}$$

The ballot problem

Bertrand's ballot problem: first published by **W. A. Whitworth** in 1878 but named after **Joseph Louis François Bertrand** who rediscovered it in 1887 (one of the first results in probability theory).

The ballot problem

Bertrand's ballot problem: first published by **W. A. Whitworth** in 1878 but named after **Joseph Louis François Bertrand** who rediscovered it in 1887 (one of the first results in probability theory).

Special case: there are two candidates A and B in an election. Each receives n votes. What is the probability that A will never trail B during the count of votes?

Example. $AABABBBAAB$ is bad, since after seven votes, A receives 3 while B receives 4.

Definition of ballot sequence

Encode a vote for A by 1, and a vote for B by -1 (abbreviated $-$). Clearly a sequence $a_1 a_2 \cdots a_{2n}$ of n each of 1 and -1 is allowed if and only if $\sum_{i=1}^{k} a_i \geq 0$ for all $1 \leq k \leq 2n$. Such a sequence is called a **ballot sequence**.

Ballot sequences

77. Ballot sequences, i.e., sequences of n 1's and n -1's such that every partial sum is nonnegative (with -1 denoted simply as $-$ below)

$$111---\quad 11-1--\quad 11--1-\quad 1-11--\quad 1-1-1-$$

Ballot sequences

77. Ballot sequences, i.e., sequences of n 1's and n -1's such that every partial sum is nonnegative (with -1 denoted simply as $-$ below)

$$111---\quad 11-1--\quad 11--1-\quad 1-11--\quad 1-1-1-$$

Note. Answer to original problem (probability that a sequence of n each of 1's and -1's is a ballot sequence) is therefore

$$\frac{C_n}{\binom{2n}{n}} = \frac{\frac{1}{n+1}\binom{2n}{n}}{\binom{2n}{n}} = \frac{1}{n+1}.$$

The ballot recurrence

$$1\,1 - 1\,1 - 1 - - - 1 - 1\,1 - 1 - -$$

The ballot recurrence

$$1\,1 - 1\,1 - 1 - - - \; 1 - 1\,1 - 1 - -$$

$$\mathbf{1}\,1 - 1\,1 - 1 - - - \,|\, 1 - 1\,1 - 1 - -$$

The ballot recurrence

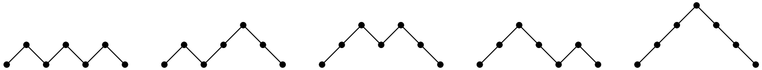
$$1\,1 - 1\,1 - 1 - - - \; 1 - 1\,1 - 1 - -$$

$$\color{blue}{1}\,1 - 1\,1 - 1 - - \color{blue}{-}\color{red}{|}1 - 1\,1 - 1 - -$$

$$1 - 1\,1 - 1 - - \quad \color{red}{|}1 - 1\,1 - 1 - -$$

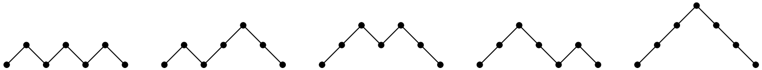
Dyck paths

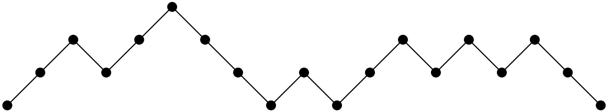
25. Dyck paths of length $2n$, i.e., lattice paths from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$, never falling below the x-axis



Dyck paths

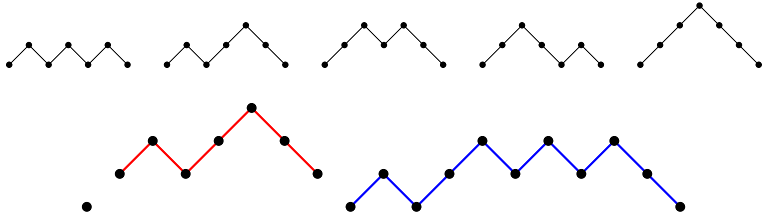
25. Dyck paths of length $2n$, i.e., lattice paths from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$, never falling below the x-axis





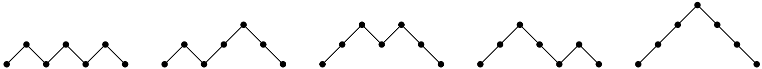
Dyck paths

25. Dyck paths of length $2n$, i.e., lattice paths from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$, never falling below the x-axis

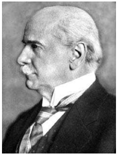


Dyck paths

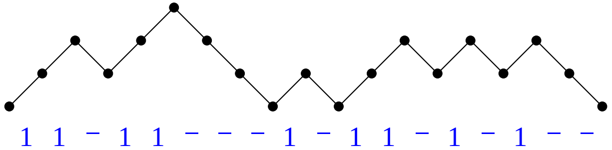
25. **Dyck paths** of length $2n$, i.e., lattice paths from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$, never falling below the x-axis



Walther von Dyck (1856–1934)



Bijection with ballot sequences



For each upstep, record 1.

For each downstep, record -1.

312-avoiding permutations

116. Permutations $a_1 a_2 \cdots a_n$ of $1, 2, \ldots, n$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called **312-avoiding** permutations)

$$123 \quad 132 \quad 213 \quad 231 \quad 321$$

312-avoiding permutations

116. Permutations $a_1a_2\cdots a_n$ of $1, 2, \ldots, n$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called **312-avoiding** permutations)

123 132 213 231 321

34251768

312-avoiding permutations

116. Permutations $a_1 a_2 \cdots a_n$ of $1, 2, \ldots, n$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called **312-avoiding** permutations)

123 132 213 231 321

3425 768

312-avoiding permutations

116. Permutations $a_1a_2\cdots a_n$ of $1, 2, \ldots, n$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called **312-avoiding**) permutations)

123 132 213 231 321

3425 768 (note **red**<**blue**)

312-avoiding permutations

116. Permutations $a_1a_2\cdots a_n$ of $1, 2, \ldots, n$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called **312-avoiding** permutations)

123 132 213 231 321

3425 768 (note **red**<**blue**)

part of the subject of **pattern avoidance**

321-avoiding permutations

Another example of pattern avoidance:

115. Permutations $a_1a_2\cdots a_n$ of $1, 2, \dots, n$ with longest decreasing subsequence of length at most two (i.e., there does not exist $i < j < k$, $a_i > a_j > a_k$), called **321-avoiding** permutations

123 213 132 312 231

321-avoiding permutations

Another example of pattern avoidance:

115. Permutations $a_1 a_2 \cdots a_n$ of $1, 2, \ldots, n$ with longest decreasing subsequence of length at most two (i.e., there does not exist $i < j < k$, $a_i > a_j > a_k$), called **321-avoiding** permutations

123 213 132 312 231

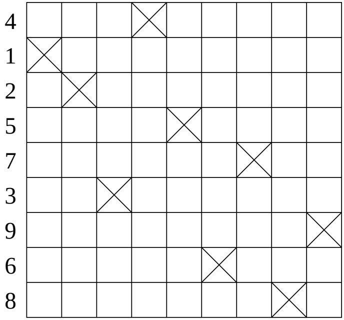
more subtle: no obvious decomposition into two pieces

Bijection with ballot sequences

$$w = 412573968$$

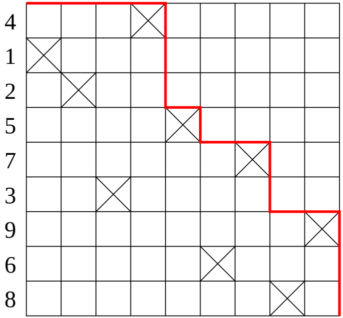
Bijection with ballot sequences

$$w = 412573968$$



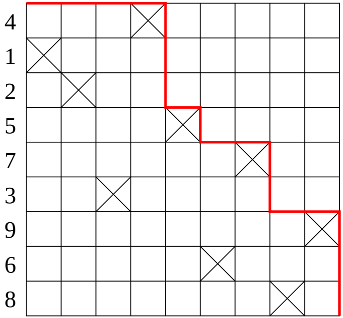
Bijection with ballot sequences

$$w = 412573968$$



Bijection with ballot sequences

$$w = 412573968$$



$$1\,1\,1\,1\,-\,-\,-\,1\,-\,1\,1\,-\,-\,1\,1\,-\,-\,-$$

An unexpected interpretation

92. n-tuples $(a_1, a_2, \ldots, a_n)$ of integers $a_i \geq 2$ such that in the sequence $1a_1a_2\cdots a_n1$, each a_i divides the sum of its two neighbors

14321 13521 13231 12531 12341

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$1 \quad 2 \quad 5 \quad 3 \quad 4 \quad 1$$

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$1 \; | 2 \; \mathbf{5} \; 3 \; 4 \; 1$$

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$1\ |2\ \ \mathbf{5}\ |3\ \ \mathbf{4}\ \ 1$$

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$1 \,|\, |2 \quad \mathbf{5} \; |\mathbf{3} \quad \mathbf{4} \quad 1$$

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$|1||\mathbf{2}\ \ \mathbf{5}\ |\mathbf{3}\ \ \mathbf{4}\ \ 1$$

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$|1||\mathbf{2}\ \ \mathbf{5}\ |\mathbf{3}\ \ \mathbf{4}\ \ 1$$

|	1	|	|	2	5	|	3	4	1
1	–	1	1	–	–	1	–		

Bijection with ballot sequences

remove largest; insert bar before the element to its left; continue until only 1's remain; then replace bar with 1 and an original number with -1, except last two

$$|1||\mathbf{2}\ \ \mathbf{5}\ |\mathbf{3}\ \ \mathbf{4}\ \ 1$$

$$\begin{array}{ccccccccccc} | & 1 & | & | & 2 & 5 & | & 3 & 4 & 1 \\ 1 & - & 1 & 1 & - & - & 1 & - & & \end{array}$$

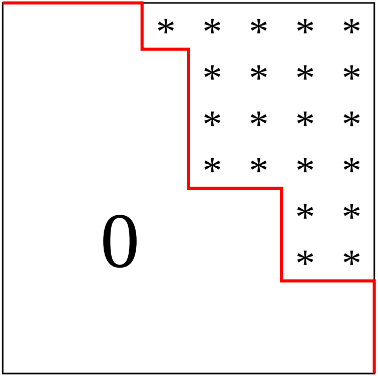
tricky to prove

A8. Algebraic interpretations

(a) Number of two-sided ideals of the algebra of all $(n-1) \times (n-1)$ upper triangular matrices over a field

A8. Algebraic interpretations

(a) Number of two-sided ideals of the algebra of all $(n-1) \times (n-1)$ upper triangular matrices over a field



Diagonal harmonics

(i) Let the symmetric group $\mathfrak{S}_n$ act on the polynomial ring $A = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ by
$w \cdot f(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_{w(1)}, \ldots, x_{w(n)}, y_{w(1)}, \ldots, y_{w(n)})$
for all $w \in \mathfrak{S}_n$. Let I be the ideal generated by all invariants of positive degree, i.e.,

$$I = \langle f \in A : w \cdot f = f \text{ for all } w \in \mathfrak{S}_n, \text{ and } f(0) = 0 \rangle.$$

Diagonal harmonics (cont.)

Then C_n is the dimension of the subspace of A/I affording the sign representation, i.e.,

$$C_n = \dim\{f \in A/I : w \cdot f = (\operatorname{sgn} w)f \text{ for all } w \in \mathfrak{S}_n\}.$$

Diagonal harmonics (cont.)

Then C_n is the dimension of the subspace of A/I affording the sign representation, i.e.,

$$C_n = \dim\{f \in A/I : w \cdot f = (\operatorname{sgn} w)f \text{ for all } w \in \mathfrak{S}_n\}.$$

Very deep proof by **Mark Haiman**, 1994.

Generalizations & refinements

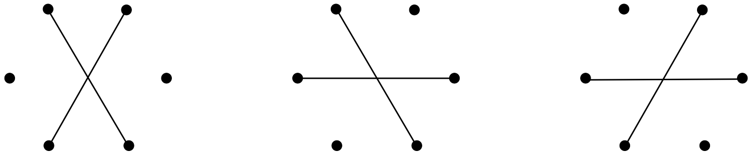
A12. ***k*-triangulation** of n-gon: maximal collections of diagonals such that no $k+1$ of them pairwise intersect in their interiors

$k = 1$: an ordinary triangulation

superfluous edge: an edge between vertices at most k steps apart (along the boundary of the n-gon). They appear in all k-triangulations and are irrelevant.

An example

Example. 2-triangulations of a hexagon (superfluous edges omitted):



Some theorems

Theorem (**Nakamigawa**, **Dress-Koolen-Moulton**). *All k-triangulations of an n-gon have $k(n - 2k - 1)$ nonsuperfluous edges.*

Some theorems

Theorem (**Nakamigawa**, **Dress-Koolen-Moulton**). *All k-triangulations of an n-gon have $k(n-2k-1)$ nonsuperfluous edges.*

Theorem (**Jonsson**, **Serrano-Stump**). *The number $T_k(n)$ of k-triangulations of an n-gon is given by*

$$\begin{aligned} T_k(n) &= \det\left[C_{n-i-j}\right]_{i,j=1}^{k} \\ &= \prod_{1\leq i<j\leq n-2k} \frac{2k+i+j-1}{i+j-1}. \end{aligned}$$

Representation theory?

Note. The number $T_k(n)$ is the dimension of an irreducible representation of the symplectic group $\mathrm{Sp}(2n-4)$.

Representation theory?

Note. The number $T_k(n)$ is the dimension of an irreducible representation of the symplectic group $\mathrm{Sp}(2n-4)$.

Is there a direct connection?

A61. Let $b(n)$ denote the number of 1's in the binary expansion of n. Using Kummer's theorem on binomial coefficients modulo a prime power, show that the exponent of the largest power of 2 dividing C_n is equal to $b(n+1) - 1$.

Sums of three squares

Let $f(n)$ denote the number of integers $1 \leq k \leq n$ such that k is the sum of three squares (of nonnegative integers). Well-known:

$$\lim_{n\to\infty} \frac{f(n)}{n} = \frac{5}{6}.$$

Sums of three squares

Let $f(n)$ denote the number of integers $1 \le k \le n$ such that k is the sum of three squares (of nonnegative integers). Well-known:

$$\lim_{n\to\infty} \frac{f(n)}{n} = \frac{5}{6}.$$

A63. Let $g(n)$ denote the number of integers $1 \le k \le n$ such that C_k is the sum of three squares. Then

$$\lim_{n\to\infty} \frac{g(n)}{n} = ??.$$

Sums of three squares

Let $\boldsymbol{f(n)}$ denote the number of integers $1 \leq k \leq n$ such that k is the sum of three squares (of nonnegative integers). Well-known:

$$\lim_{n\to\infty} \frac{f(n)}{n} = \frac{5}{6}.$$

A63. Let $\boldsymbol{g(n)}$ denote the number of integers $1 \leq k \leq n$ such that C_k is the sum of three squares. Then

$$\lim_{n\to\infty} \frac{g(n)}{n} = \frac{7}{8}.$$

Analysis

A65.(b)

$$\sum_{n\geq 0} \frac{1}{C_n} = ??$$

Analysis

A65.(b)

$$\sum_{n\geq 0} \frac{1}{C_n} = ??$$

$$1 + 1 + \frac{1}{2} + \frac{1}{5} = 2.7$$

Analysis

A65.(b)

$$\sum_{n\geq 0} \frac{1}{C_n} = 2 + \frac{4\sqrt{3}\pi}{27}$$

$$1 + 1 + \frac{1}{2} + \frac{1}{5} = 2.7$$

Analysis

A65.(b)

$$\sum_{n\geq 0} \frac{1}{C_n} = 2 + \frac{4\sqrt{3}\pi}{27}$$

$$1 + 1 + \frac{1}{2} + \frac{1}{5} = 2.7$$

$$2 + \frac{4\sqrt{3}\pi}{27} = 2.806133\cdots$$

Why?

A65.(a)

$$\sum_{n\geq 0} \frac{x^n}{C_n} = \frac{2(x+8)}{(4-x)^2} + \frac{24\sqrt{x}\sin^{-1}\left(\frac{1}{2}\sqrt{x}\right)}{(4-x)^{5/2}}.$$

Why?

A65.(a)

$$\sum_{n\geq 0} \frac{x^n}{C_n} = \frac{2(x+8)}{(4-x)^2} + \frac{24\sqrt{x}\sin^{-1}\left(\frac{1}{2}\sqrt{x}\right)}{(4-x)^{5/2}}.$$

Based on a (difficult) calculus exercise: let

$$y = 2\left(\sin^{-1}\frac{1}{2}\sqrt{x}\right)^2.$$

Then $y = \sum_{n\geq 1} \frac{x^n}{n^2\binom{2n}{n}}$. Use $\sin^{-1} x = \sum_{n\geq 0} 4^{-n}\binom{2n}{n}\frac{x^{2n+1}}{2n+1}$.

The last slide

The last slide 

The last slide