### Catalan Numbers: From EGS to BEG

Drew Armstrong

University of Miami www.math.miami.edu/~armstrong

MIT Combinatorics Seminar May 15, 2015 This talk was inspired by an article of Igor Pak on the history of Catalan numbers, (http://arxiv.org/abs/1408.5711) which now appears as an appendix in Richard Stanley's monograph *Catalan Numbers*. Igor also maintains a webpage with an extensive bibliography and links to original sources:

http://www.math.ucla.edu/~pak/lectures/Cat/pakcat.htm



Pak



Stanley

# Goal of the Talk

The goal of the current talk is to connect the history of Catalan numbers with recent trends in geometric representation theory. To make the story coherent I'll have to skip some things (sorry).



Hello!

The talk will follow Catalan numbers through three levels of generality:

Amount of Talk	Level of Generality
41.94%	Catalan
20.97%	Fuss-Catalan
24.19%	Rational Catalan

On September 4, 1751, Leonhard Euler wrote a letter to his friend and mentor Christian Goldbach.



Euler



Goldbach

In this letter Euler considered the problem of counting the triangulations of a convex polygon. He gave a couple of examples.

61 Disgenter 1 ag; 11. 50 ; 11 50 ; 11 4 maraly, Quel

The pentagon *abcde* has five triangulations:

Here's a bigger example that Euler computed but didn't put in the letter.



A convex heptagon has 42 triangulations.

He gave the following table of numbers and he **conjectured** a formula.

 $1 = \frac{3}{2}, 2 = 1.\frac{6}{2}, 5 = 2.\frac{19}{2}, 14 = 5.\frac{19}{4}, 42 = 14.\frac{18}{4}, 102 = 42\frac{13}{4}, 102$ n = 3, 4, 5, 6, 7, 8, 9, 10x = 1, 2, 5, 14, 42, 152, 429, 1430

$$x = \frac{2 \cdot 6 \cdot 10 \cdot 14 \cdot 18 \cdot 22 \cdots (4n - 10)}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdots (n - 1)}$$

This was the first<sup>\*</sup> appearance of what we now call **Catalan numbers**. Following modern terminology we will define  $C_n$  as the number of triangulations of a convex (n + 2)-gon. The most commonly quoted formula is

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

\*See Ming Antu.

Does this agree with Euler's formula?

$$\frac{2 \cdot 6 \cdot 10 \cdot 14 \cdot 18 \cdot 22 \cdots (4n - 10)}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdots (n - 1)}$$

$$= \frac{2^{n-2}}{(n-1)!} \cdot 1 \cdot 3 \cdot 5 \cdots (2n - 5)$$

$$= \frac{2^{n-2}}{(n-1)!} \cdot \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdots (2n - 5)(2n - 4)}{2 \cdot 4 \cdot 6 \cdots (2n - 4)}$$

$$= \frac{2^{n-2}}{(n-1)!} \cdot \frac{(2n - 4)!}{2^{n-2}} \cdot (n - 2)!$$

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

$$= C_{n-2}.$$

Yes it does.

At the end of the letter Euler even guessed the generating function for this sequence of numbers.

$$1 + 2a + 5a^{2} + 14a^{3} + 42a^{4} + 132a^{5} + \text{etc.} = \frac{1 - 2a - \sqrt{(1 - 4a)}}{2aa}$$

He knew that this generating function agrees with the closed formula.

But he also knew that something was missing.

La all and file for the file of the Arming for the second for the

Die Induction aber, so ich gebraucht, war ziemlich mühsam, doch zweifle ich nicht, dass diese Sach nicht sollte weit leichter entwickelt werden können. [However, the induction that I employed was pretty tedious, and I do not doubt that this result can be reached much more easily.] Goldbach replied a month later, observing that the generating function

$$A = \frac{1 - 2a - \sqrt{(1 - 4a)}}{2aa}$$

satisfies the functional equation

$$1+aA=A^{\frac{1}{2}},$$

which is equivalent to infinitely many equations in the coefficients. He suggested that these equations in the coefficients might be proved directly. Euler replied, giving more details on the derivation of the generating function, but he did not finish the proof.

Somewhat later (date unknown), Euler communicated the problem of triangulations to Johann Andreas von Segner.







Segner

Segner published a paper in 1758 with a combinatorial proof of the following recurrence:

$$C_{n+1} = C_0 C_n + C_1 C_{n-1} + \cdots + C_{n-1} C_1 + C_n C_0.$$

factis, quia terminum intermedium vnum superesse necesse est, qui sit d, huius quadratum factis adde, vt fiat x = 2aq + 2bp + 2co + etc. + dd

Ad indicem o eff  $a \equiv x$ . Si enim linea recta

The proof is not hard.



Segner's recurrence was exactly the piece that Euler and Goldbach were missing. Thus, by 1758, Euler, Goldbach, and Segner had pieced together a proof that the number of triangulations of a convex (n + 2)-gon is  $\frac{1}{n+1} \binom{2n}{n}$ . However, no one bothered to publish it.

The last word on the matter was an unsigned summary of Segner's work published in the same volume. (It was written by Euler.)

All the pieces were available, but a self-contained proof was not published until 80 years later, after a French mathematician named Orly Terquem communicated the problem to Joseph Liouville in 1838. Specifically, he asked whether Segner's recurrence can be used to prove Euler's product formula for the Catalan numbers. Liouville, in turn, proposed this problem to "various geometers".



Terquem



Liouville

At this point, several French mathematicians published papers improving the methods of Euler-Goldbach-Segner. One of these mathematicians was a student of Liouville named *Eugène Charles Catalan*.



Catalan

Catalan's contributions were modest: He was the first to observe that  $C_n = \binom{2n}{n} - \binom{2n}{n-1}$  and he was also the first to interpret triangulations as "bracketed sequences".

Like so.



 $(((x_1((x_2x_3)((x_4x_5)x_6)))x_7)x_8)$ 

Despite this modest contribution, Catalan's name eventually stuck to the problem. Maybe this is because he published several papers throughout his life that popularized the subject.

The first occurence of the term "Catalan's numbers" appears in a 1938 paper by Eric Temple Bell, but seen in context he was not suggesting this as a name.

reterence. It is an immediate consequence of (1.10), and there are similar (but more complicated) summations involving  $\xi_n^{(m)}$ , m > 1. Catalan's problem of 1338 (loc. cit. (1), p. 508-16, vol. 3) of finding the number of different products of *n* factors is similar to Epstein's (1) of finding the total number of decompositions of an integer into coprime factors where, instead of Catalan's sequence 1, 2, 5, 14, 42, 132, ..., the sequence where 1, 2, 3, ... distinct primes are concerned is 1, 2, 5, 15, 52, 203, ..., that is,  $\xi_{1}^{(m)}$ . Catalan's numbers first appeared in Euler's and Segner's problem of finding the number of partitions of a rectilinear *n*-gon into triangles by diagonals. In my previous papers (1) the connection between the  $\xi_{1}^{(D)}$  and partitions was noted. A combinatorial inter-

According to Pak, the term "Catalan numbers" became standard only after John Riordan's book *Combinatorial Identities* was published in 1968. (Catalan himself had called them "Segner numbers".)

Now let me skip forward a bit. Eventually interest in triangulations passed beyond their enumeration to their structure. We will say that two triangulations are **adjacent** if they differ in exactly one diagonal.



The first person to consider this structure was born in Germany in 1911 as **Bernhard Teitler**. In 1942 he officially changed his name to **Dov Tamari**, after spending time in a Jerusalem prison for being a member of the militant underground organization IZL and for being caught with explosives in his room.



Tamari

In 1942 Bernhard Teitler officially changed his name to Dov Tamari [20]. Unofficially Teitler had already used the new name earlier [21]. 'Bernhard' has the meaning 'strong as a bear', and 'Dov' is Hebrew for 'bear'. The name 'Teitler' may have its origin in the Yiddish word 'teytl' for 'date', the fruit, and 'tamar' is the name of a palm tree carrying this fruit.

After several more adventures, Tamari finally completed his thesis at the Sorbonne in 1951, entitled *Monoides préordeonnés et chaînes de Malcev*. In this work he considered the set of bracketed sequences as a partially ordered set, and he also thought of this poset as the skeleton of a convex polytope. This partially ordered set is now called the **Tamari lattice**.



Picture from Tamari's Thesis

Ten years later, in 1961, the topologist Jim Stasheff independently considered the same structure in his thesis at Princeton, entitled *Homotopy associativity of H-spaces, etc.* John Milnor verified that it is a polytope and it becamse known as the **Stasheff polytope** in the topology literature.



Stasheff



Picture from Stasheff's Thesis

But there is one more name for this structure. In 1984, apparently independently of Tamari and Stasheff, Micha Perles asked Carl Lee whether the collection of mutually non-crossing sets of diagonals in a convex polygon is isomorphic to the boundary complex of some simplicial polytope. Mark Haiman verified that this is the case and he called this polytope the **associahedron**. Here is a quote from his unpublished manuscript of 1984, *Constructing the associahedron*:



The associated is a nuthical polytope whose face structure represents the lattice of partial parenthesizations of a sequence of variables, in a way to be made precise below. The purpose of these notes is to give an explicit construction of such a polytope.

Haiman

Here is a picture.



Thus the (simplicial) **associahedron** is merely the dual polytope of the (simple) **Stasheff polytope**, and the names are used interchangeably. Study of the associahedron has exploded in recent years and I will confine myself to just one aspect: computation of the *f*-vector.

Triangulations correspond to maximal faces of the associahedron, so they are counted by the Catalan numbers. As Haiman mentions, the lower-dimensional faces correspond to "partial triangulations". The problem of counting these partial triangulations had been solved over 100 years earlier.

In the year 1857, unaware of the previous work on triangulations, the Reverend Thomas Kirkman considered the problem of placing k non-crossing diagonals in a convex r-gon:

This summation I must leave to the learned and industrious reader; but, meanwhile, I shall venture to enunciate with the best demonstration, such as it is, that occurs to me, the following

THEOREM T. The number of (1+k)-divisions of an r-gon, i. e. of all the ways in which k diagonals can be drawn in it, none crossing another, is

$$D(r,k) = \frac{r^{k|1}}{(k+2)} \times \frac{(r-k-2)^{k|1}}{(k+1)}.$$

I don't know what-the-heck kind of terminology that is. In modern language the answer is

$$D(r,k) = \frac{1}{k+1} \binom{r+k-1}{k} \binom{r-3}{k}$$

Kirkman freely admitted that his proof was incomplete. In 1891, Arthur Cayley used generating function methods to completely solve Kirkman's problem.



Kirkman



Cayley

Thus, we now call D(r, k) the Kirkman-Cayley numbers.

This man is Nicolas Fuss.



Fuss

He was born in 1723 in Switzerland and moved to St. Petersbug in 1773 to become Euler's assistant. He later married Euler's granddaughter Albertina Philippine Louise (daughter of Johann Albrecht Euler).

In the year 1793 it seems that Johann Friedrich Pfaff sent a letter to Fuss.









He asked whether anything was known about the problem of dissecting a convex *n*-gon into convex *m*-gons. Fuss admitted that he knew nothing about this and he was tempted to investigate the problem. He wrote a paper about it one month later.



riofi, abftinere non potui, quominus ipfe, quomodo inueftigatio ifta in genere fuscipienda fit, tentarem. Methodum

a me adhibitam, conatuumque meorum fucceffus hic exhi-

bere conftitui.

#### SOLUTION OF THE QUESTION

... the 15th of last month...

... in how many ways can an n-gon be resolved along diagonals into m-gons?...

... I was tempted ...

Clearly this is only possible for certain m. Following modern notation we let  $C_n^{(m)}$  be the number of ways a convex (mn + 2)-gon can be dissected by diagonals into (m + 2)-gons. Note that  $C_n^{(1)} = C_n$ .

I don't read Latin but I believe that Fuss proved the following recurrence, generalizing Segner's recurrence:

$$C_{n+1}^{(m)} = \sum_{n_1+n_2+\dots+n_{m+1}=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_{m+1}}^{(m)}.$$

Today we call these  $C_n^{(m)}$  the **Fuss-Catalan numbers**.

Again, the proof is not hard.



Actually, I can't find the recurrence stated explicitly in Fuss' paper. But I do see a functional equation that is equivalent.

```
primum autem coëfficientem A vnitati effe aequalem ipfa
pofitio (r + A x + B x^2 + C x^3 + etc.)^{n-1}
= A + B x + C x^2 + D x^3 + etc.,
```

Fuss did not find a product formula for the coefficients and it seems that this remained open an open question until Liouville solved the problem using Lagrange inversion in 1843.



```
De là résulte \varphi(\mathbf{z}) = m - 1, puis, en général,
```

```
\varphi(i) = \frac{(im-i)(im-i-1)...(im-2i+2)}{1.2...i}
```

Telle est la valeur de  $\varphi(i)$ , que Fuss aurait dù trouver. On voit par les Nova acta que Pfaff s'était aussi occupé de ce problème, mais j'ignore entièrement quelle solution il avait obtenue.

Let's examine Liouville's formula.

$$\varphi(i) = \frac{(im-i)(im-i-1)\cdots(im-2i+2)}{1\cdot 2\cdots i}$$
  
=  $\frac{(im-i)!}{i!(im-2i+1)!}$   
=  $\frac{1}{i(m-2)+1} \binom{i(m-1)}{i}.$ 

In our language we would write

$$C_n^{(m)} = \frac{1}{nm+1} \binom{n(m+1)}{n}.$$

Observe again that  $C_n^{(1)} = C_n$ .

And where does the name "Fuss-Catalan" come from? It must certainly be more recent than Riordan's 1968 book. The earliest occurrence I can find is in the 1989 textbook *Concrete Mathematics* by Graham, Knuth, and Patashnik.

mn + i = n occurrences of +1, and Raney's lemma tells us that the number of such sequences with all partial sums positive is exactly

$$\binom{mn+1}{n}\frac{1}{mn+1} = \binom{mn}{n}\frac{1}{(m-1)n+1}.$$
 (7.67)

So this is the number of m-Raney sequences. Let's call this a Fuss-Catalan number  $C_n^{(m)}$ , because the sequence  $\langle C_n^{(m)} \rangle$  was first investigated by N. I. Fuss [135] in 1791 (many years before Catalan himself got into the act). The ordinary Catalan numbers are  $C_n = C_n^{(2)}$ .

Now that we know the answer (7 67) let's play "leopardy" and figure

The earliest occurrence I can find in a research paper is:

 Algebras associated to intermediate subfactors (1997), by Dietmar Bisch and Vaughan Jones.

Bisch and Jones refer to [GKP] for the terminology.

Back in Cayley's 1891 paper he had listed and solved the following three problems:

1. The partitions are made by non-intersecting diagonals; the problems which have been successively considered are (1) to find the number of partitions of an r-gon into triangles, (2) to find the number of partitions of an r-gon into k parts, and (3) to find the number of partitions of an r-gon into p-gons, r of the form n(p-2)+2.

(1) is Euler's problem, (2) is Kirkman's problem, and (3) is the problem solved by Fuss. However, Cayley did **not** consider the natural problem of finding a common generalization of (2) and (3), i.e., to count partial dissections of an (sn + 2)-gon by *i* non-crossing diagonals into (sj + 2)-gons for various *j*.

This problem was posed and solved in 1998 by Józef Przytycki and Adam Sikora, motivated directly by the 1997 paper of Bisch and Jones.

Our work on knot theory motivated an elementary and short "bijective" proof of the common generalization of (2) and (3). Let  $Q_i(s, n)$  denote the set of dissections of a convex (sn+2)-gon by *i* non-crossing diagonals into sj + 2-gons  $(1 \le j \le n-1)$ , i.e. we allow dissections of a convex (sn+2)-gon by  $j \ge 1$ .

tions which can be subdivided to dissections into (s+2)-gons. Let

They proved that

$$\#Q_i(s,n)=\frac{1}{i+1}\binom{sn+i+1}{i}\binom{n-1}{i}.$$

And here are their pictures.



Przytycki



Sikora

The Przytycki-Sikora numbers can be viewed as the *f*-vector of a certain "generalized associahedron" that was studied by Sergey Fomin, Nathan Reading, and Eleni Tzanaki in 2005.



Fomin



Reading



Tzanaki

Here is a generalized associahedron.



Now I will describe some joint work with Brendon Rhoades and Nathan Williams.



Rhoades



Williams

Let a < b be **coprime positive integers** and consider the set of lattice paths from (0,0) to (b, a) staying above the line of slope a/b. We will call these (a, b)-Dyck paths.



The history of lattice path enumeration is very murky, but it was known to Howard Grossman in 1950 and proved by M. T. L. Bizley in 1954 (in the *Journal of the Institute of Actuaries*) that the number of (a, b)-Dyck paths is

$$\frac{1}{a+b}\binom{a+b}{a}$$

Grossman(4) announced without proof in 1950 a formula for the number of paths from (0, 0) to (km, kn) which may touch but never rise above the line my = nx, where k is a positive integer and m and n are coprime positive integers; thus (km, kn) is any point having positive integral coefficients. Grossman's formula is  $\sum F^{k_1}F^{k_2}_{n_1}...,k_1 \mid k_n \mid ....$ 

			1.	-	in \
F	1		11	m + j	"
$r_j =$	i(m-	$\vdash n$		im	
	)(""	· ~,	1	<i>J</i>	/

the sum extending over all positive integral  $k_i$  such that  $k_i \ge 0$  and  $\sum i k_i = k$ . If k = i this takes the simple form  $\frac{I}{m+n} {m \choose m}$ .

The object of the present note is to supply a proof of Grossman's formula and

Call these the rational Catalan numbers:

$$Cat(a,b) := \frac{1}{a+b} \binom{a+b}{a}.$$

Note that the classical Catalan numbers and Fuss-Catalan numbers occur as special cases:

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

$$C_n^{(m)} = \frac{1}{mn+1} \binom{n(m+1)}{n} = \operatorname{Cat}(n, mn+1).$$

In the spirit of this talk, we wish to define a **rational associahedron** generalizing the earlier examples.

Start with a Dyck path. Here (a, b) = (5, 8).



Label the columns by  $\{1, 2, \ldots, b+1\}$ .



Shoot lasers from the bottom left with slope a/b.



Lift the lasers up.



This defines a partial dissection of the convex b + 1-gon. We will call it an (a, b)-dissection.



**Definition:** Given coprime positive integers a < b, let Ass(a, b) be the abstract simplicial complex whose vertices are certain diagonals of a convex (b + 1)-gon and whose maximal faces are the (a, b)-dissections.

#### Theorems (Armstrong-Rhoades-Williams, 2013):

- Ass(n, n + 1) is the classical associahedron.
- ► Ass(n, mn + 1) is the generalized associahedron of Fomin-Reading and Tzanaki.
- ▶ The *f*-vector of Ass(*a*, *b*) is given by the numbers

$$\frac{1}{a}\binom{a}{k}\binom{b+k-1}{k-1},$$

generalizing the Kirkman-Caylay-Przytycki-Sikora numbers. We will just call them the rational Kirkman numbers.

• The complexes Ass(a, b) and Ass(b - a, b) are Alexander duals.

Here is a picture of Ass(2,5) and Ass(3,5) as subcomplexes of Ass(4,5).



Finally, I will give a highbrow reason why rational Catalan numbers may be interesting.

Let the symmetric group  $S_n$  act on affine space  $\mathfrak{h} := \mathbb{C}^n$  by permuting coordinates, and consider the quotient variety  $(\mathfrak{h} \times \mathfrak{h}^*)/S_n$ . This is a singular variety with coordinate ring given by the **diagonal symmetric polynomials** in two sets of variables:

$$\mathbb{C}[(\mathfrak{h}\times\mathfrak{h}^*)/S_n]=\mathbb{C}[x_1,\ldots,x_n,y_1,\ldots,y_n]^{S_n}.$$

The singularities of  $(\mathfrak{h} \times \mathfrak{h}^*)/S_n$  can be resolved by the Hilbert scheme  $\operatorname{Hilb}^n(\mathbb{C}^2)$  of *n* points in the plane. Furthermore, the smooth locus of  $(\mathfrak{h} \times \mathfrak{h}^*)/S_n$  has a natural symplectic structure that lifts to the resolution  $\operatorname{Hilb}^n(\mathbb{C}^2)$ . And where there is a symplectic manifold, there are always 'quantum' deformations of its coordinate ring.

In this case, there is a natural (or so I'm told) quantum deformation  $U_c(S_n)$  of the coordinate ring  $\mathbb{C}[(\mathfrak{h} \times \mathfrak{h}^*)/S_n]$  depending on a complex parameter  $c \in \mathbb{C}$  such that

$$\lim_{c\to 0} U_c(S_n) = \mathbb{C}[(\mathfrak{h} \times \mathfrak{h}^*)/S_n].$$

This defomation is called the **spherical rational Cherednik algebra**. It was introduced in 2001 by Pavel Etingof and Victor Ginzburg. The idea (or so I'm told) is that the representation theory of  $U_c(S_n)$  should tell us about the geometry of the resolution

$$\operatorname{Hilb}^{n}(\mathbb{C}^{2}) \twoheadrightarrow (\mathfrak{h} \times \mathfrak{h}^{*})/S_{n}.$$

In this context, the following theorem of Yuri Berest, Pavel Etingof, and Victor Ginzburg is probably significant. First see their pictures.



Berest



Etingof



Ginzburg

#### Theorem (Berest-Etingof-Ginzburg, 2002):

Let *a* be a positive integer. The only values of  $c \in \mathbb{C}$  such that the algebra  $U_c(S_a)$  has a nonzero finite dimensional representation are c = b/a, where *b* is a positive integer coprime to *a*. In this case, there exists a **unique finite dimensional representation**, of dimension

$$\frac{1}{a+b}\binom{a+b}{a},$$

i.e., a rational Catalan number.



# Thanks!