1

Basic Properties

1.1. The Definition of Catalan Numbers

There are many equivalent ways to define Catalan numbers. In fact, the main focus of this monograph is the myriad combinatorial interpretations of Catalan numbers. We also discuss some algebraic interpretations and additional aspects of Catalan numbers. We choose as our basic definition their historically first combinatorial interpretation, whose history is discussed in Appendix B along with further interesting historical information on Catalan numbers. Let $\mathcal{P}_{n+2}$ denote a convex polygon in the plane with $n+2$ vertices (or *convex* $(n+2)$*-gon* for short). A *triangulation* of $\mathcal{P}_{n+2}$ is a set of $d-1$ diagonals of $\mathcal{P}_{n+2}$ which do not cross in their interiors. It follows easily that these diagonals partition the interior of $\mathcal{P}_{n+2}$ into n triangles. Define the nth *Catalan number* C_n to be the number of triangulations of $\mathcal{P}_{n+2}$. Set $C_0 = 1$. Figure 1.1 shows that $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, and $C_4 = 14$. Some further values are $C_5 = 42$, $C_6 = 132$, $C_7 = 429$, $C_8 = 1430$, $C_9 = 4862$, and $C_{10} = 16796$.

In this chapter we deal with the following basic properties of Catalan numbers: (1) the fundamental recurrence relation, (2) the generating function, (3) an explicit formula, and (4) the primary combinatorial interpretations of Catalan numbers. Throughout this monograph we use the following notation:

$\mathbb{C}$	complex numbers
$\mathbb{R}$	real numbers
$\mathbb{Q}$	rational numbers
$\mathbb{Z}$	integers
$\mathbb{N}$	nonnegative integers $\{0,1,2,\dots\}$
$\mathbb{P}$	positive integers $\{1,2,3,\dots\}$
$[n]$	the set $\{1,2,\dots,n\}$, where $n \in \mathbb{N}$
$\#S$	number of elements of the (finite) set S

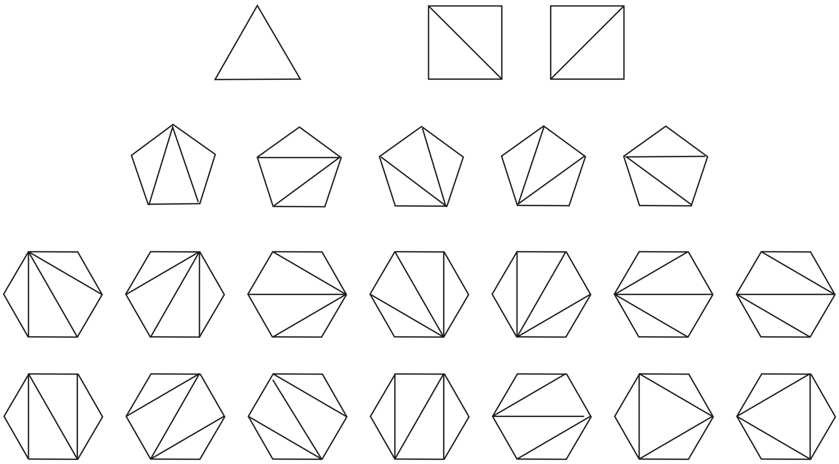

Figure 1.1. Triangulated polygons.

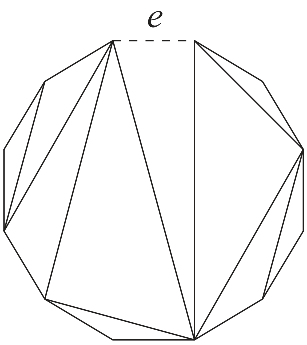

Figure 1.2. The recursive structure of a triangulated polygon.

1.2. The Fundamental Recurrence

To obtain a recurrence relation for Catalan numbers, let $\mathcal{P}_{n+3}$ be a convex $(n+3)$-gon. Fix an edge e of $\mathcal{P}_{n+3}$, and let T be a triangulation of $\mathcal{P}_{n+3}$. When we remove the edge e from T we obtain two triangulated polygons, say $\mathcal{Q}_1$ and $\mathcal{Q}_2$, in counterclockwise order from e, with one common vertex. If $\mathcal{Q}_i$ has a_i+2 vertices, then $a_1+a_2=n$. See Figure 1.2, where $n=9$, $a_1=5$, and $a_2=4$. It is possible that one of $\mathcal{Q}_1$ or $\mathcal{Q}_2$ is just a single edge, which occurs when the triangle of T containing e has an additional edge on $\mathcal{P}_{n+3}$, necessarily adjacent to e. In this case we consider the edge as a 2-gon, which has $C_0=1$ triangulations.

Conversely, given two triangulated polygons with a_1+2 and a_2+2 vertices, we can put them together to form a triangulated $(n+3)$-gon by reversing the above procedure. Since there are C_{a_i} triangulations of Q_i, we obtain the recurrence and initial condition

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}, \quad C_0 = 1. \tag{1.1}$$

This is the most important and most transparent recurrence satisfied by C_n. It easily explains many of the combinatorial interpretations of Catalan numbers, where the objects being counted have a decomposition into two parts, analogous to what we have just done for triangulated polygons. For other "Catalan objects," however, it can be quite difficult, if not almost impossible, to see directly why the recurrence (1.1) holds.

1.3. A Generating Function

Given the recurrence (1.1), it is a routine matter for those familiar with generating functions to obtain the next result. For some background information on generating functions and the rigorous justification for our manipulations, see [64], especially Chapter 1. Let us just mention now one aspect of generating functions, namely, the binomial theorem for arbitrary exponents. When a is any complex number, or even an indeterminate, and $k \in \mathbb{N}$, then we define the *binomial coefficient*

$$\binom{a}{k} = \frac{a(a-1)\cdots(a-k+1)}{k!}.$$

The "generalized binomial theorem" due to Isaac Newton asserts that

$$(1+x)^a = \sum_{n\geq 0} \binom{a}{n} x^n. \tag{1.2}$$

This formula is just the formula for the Taylor series of $(1+x)^a$ at $x=0$. For our purposes we consider generating function formulas such as Equation (1.2) to be "formal" identities. Questions of convergence are ignored.

1.3.1 Proposition. *Let*

$$\begin{aligned} C(x) &= \sum_{n\geq 0} C_n x^n \\ &= 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + 132x^6 + 429x^7 + 1430x^8 + \cdots . \end{aligned}$$

Then

$$C(x) = \frac{1-\sqrt{1-4x}}{2x}. \tag{1.3}$$

Proof. Multiply the recurrence (1.1) by x^n and sum on $n \geq 0$. On the left-hand side we get

$$\sum_{n\geq 0} C_{n+1}x^n = \frac{C(x)-1}{x}.$$

Since the coefficient of x^n in $C(x)^2$ is $\sum_{k=0}^{n} C_k C_{n-k}$, on the right-hand side we get $C(x)^2$. Thus

$$\frac{C(x)-1}{x} = C(x)^2,$$

or

$$xC(x)^2 - C(x) + 1 = 0. \tag{1.4}$$

Solving this quadratic equation for $C(x)$ gives

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}. \tag{1.5}$$

We have to determine the correct sign. Now, by the binomial theorem for the exponent 1/2 (or by other methods),

$$\sqrt{1-4x} = 1 - 2x + \cdots.$$

If we take the plus sign in Equation (1.5) we get

$$\frac{1+(1-2x+\cdots)}{2x} = \frac{1}{x} - 1 + \cdots,$$

which is not correct. Hence we must take the minus sign. As a check,

$$\frac{1-(1-2x+\cdots)}{2x} = 1 + \cdots,$$

as desired. □

1.4. An Explicit Formula

From the generating function it is easy to obtain a formula for C_n.

1.4.1 Theorem. *We have*

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!\,(n+1)!}. \tag{1.6}$$

Cambridge University Press & Assessment
978-1-107-42774-7 — Catalan Numbers
Richard P. Stanley
Excerpt
More Information

Proof. We have

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{n\geq 0} \binom{1/2}{n} x^n.$$

Hence by Proposition 1.3.1,

$$\begin{aligned} C(x) &= \frac{1}{2x}\left(1 - \sum_{n\geq 0}\binom{1/2}{n}(-4x)^n\right) \\ &= -\frac{1}{2}\sum_{n\geq 0}\binom{1/2}{n+1}(-4)^{n+1}x^n. \end{aligned}$$

Equating coefficients of x^n on both sides gives

$$C_n = -\frac{1}{2}\binom{1/2}{n+1}(-4)^{n+1}. \tag{1.7}$$

It is a routine matter to expand the right-hand side of Equation (1.7) and verify that it is equal to $\frac{1}{n+1}\binom{2n}{n}$. □

The above proof of Theorem 1.4.1 is essentially the same as the proof of Bernoulli-Euler-Segner discussed in Appendix B. In Section 1.6 we will present a more elegant proof.

The expression $\frac{1}{n+1}\binom{2n}{n}$ is the standard way to write C_n explicitly. There is an equivalent expression that is sometimes more convenient:

$$C_n = \frac{1}{2n+1}\binom{2n+1}{n}. \tag{1.8}$$

Note also that

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

1.5. Fundamental Combinatorial Interpretations

Among the myriad of combinatorial interpretations of Catalan numbers, a few stand out as being the most fundamental, namely, polygon triangulations (already considered), binary trees, plane trees, ballot sequences, parenthesizations, and Dyck paths. They will be the subject of the current section. For all of them the recurrence (1.1) is easy to see, or, what amounts to the same thing, there are simple bijections among them. We begin with the relevant definitions.

A *binary tree* is defined recursively as follows. The empty set Ø is a binary tree. Otherwise a binary tree has a *root vertex* v, a *left subtree* T_1, and a *right*

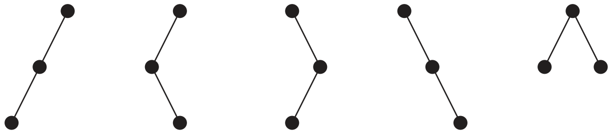

Figure 1.3. The five binary trees with three vertices.

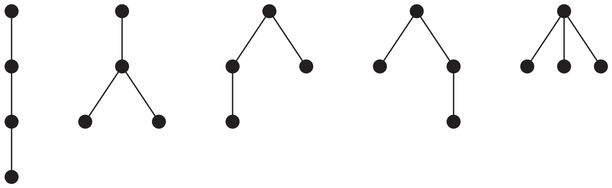

Figure 1.4. The five plane trees with four vertices.

subtree T_2, both of which are binary trees. We also call the root of T_1 (if T_1 is nonempty) the *left child* and the root of T_2 (if T_2 is nonempty) the *right child* of the vertex v. We draw a binary tree by putting the root vertex v at the top, the left subtree T_1 below and to the left of v, and the right subtree T_2 below and to the right of v, with an edge drawn from v to the root of each nonempty T_i. Figure 1.3 shows the five binary trees with three vertices.

A *plane tree* (also called an *ordered tree* or *Catalan tree*) P may be defined recursively as follows. One specially designated vertex v is called the *root* of P. Thus plane trees, unlike binary trees, cannot be empty. Then either P consists of the single vertex v, or else it has a sequence $(P_1,\dots,P_m)$ of *subtrees* P_i, $1 \le i \le m$, each of which is a plane tree. Thus the subtrees of each vertex are linearly ordered. When drawing such trees, these subtrees are written in the order left-to-right. The root v is written on the top, with an edge drawn from v to the root of each of its subtrees. Figure 1.4 shows the five plane trees with four vertices.

A *ballot sequence* of length $2n$ is a sequence with n each of 1's and -1's such that every partial sum is nonnegative. The five ballot sequences of length six (abbreviating -1 by just $-$) are given by

$$111{-}{-}{-} \qquad 11{-}1{-}{-} \qquad 11{-}{-}1{-} \qquad 1{-}11{-}{-} \qquad 1{-}1{-}1{-}\,.$$

The term "ballot sequence" arises from the following scenario. Two candidates A and B are running in an election. There are $2n$ voters who vote sequentially for one of the two candidates. At the end each candidate receives n votes. What is the probability p_n that A never trails B in the voting and both candidates receive n votes? If we denote a vote for A by 1 and a vote for B by -1, then clearly the sequence of votes forms a ballot sequence if and only if A

never trails B. Moreover, the total number of ways in which the $2n$ voters can cast n votes for each of A and B is $\binom{2n}{n}$. Hence if f_n denotes the number of ballot sequences of length $2n$, then $p_n = f_n/\binom{2n}{n}$. We will see in Theorem 1.5.1(iv) that $f_n = C_n$, so $p_n = 1/(n+1)$. For the generalization where A receives m votes and B receives $n \le m$ votes, see Problem A2.[1] For the history behind this result, see Section B.7.

A *parenthesization* or *bracketing* of a string of $n+1$ x's consists of the insertion of n left parentheses and n right parentheses that define n binary operations on the string. An example for $n = 6$ is $(((xx)x)((xx)(xx)))$. In general we can omit the leftmost and rightmost parentheses without loss of information. Thus our example denotes the product of $(xx)x$ with $(xx)(xx)$, where $(xx)x$ denotes the product of xx and x, and $(xx)(xx)$ denotes the product of xx and xx. There are five ways to parenthesize a string of four x's, namely,

$$x(x(xx)) \quad x((xx)x) \quad (xx)(xx) \quad (x(xx))x \quad ((xx)x)x.$$

Let S be a subset of $\mathbb{Z}^d$. A *lattice path* in $\mathbb{Z}^d$ of length k with steps in S is a sequence $v_0, v_1, \ldots, v_k \in \mathbb{Z}^d$ such that each consecutive difference $v_i - v_{i-1}$ lies in S. We say that L *starts at* v_0 and *ends at* v_k, or more simply that L *goes from* v_0 *to* v_k. A *Dyck path* of length $2n$ is a lattice path in $\mathbb{Z}^2$ from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$, with the additional condition that the path never passes below the x-axis. Figure 1.5(a) shows the five Dyck paths of length six (so $n = 3$). A trivial but useful variant of Dyck paths (sometimes also called a Dyck path) is obtained by replacing the step $(1,1)$ with $(1,0)$ and $(1,-1)$ with $(0,1)$. In this case we obtain lattice paths from $(0,0)$ to (n,n) with steps $(1,0)$ and $(0,1)$, such that the path never rises above the line $y = x$. See Figure 1.5(b) for the case $n = 3$.

It will come as no surprise that the objects we have just defined are counted by Catalan numbers. We will give simple bijective proofs of this fact. By a *bijective proof* we mean a proof that two finite sets S and T have the same cardinality (number of elements) by exhibiting an explicit bijection $\varphi : S \to T$. To *prove* that φ is a bijection, we need to show that it is injective (one-to-one) and surjective (onto). This can either be shown directly or by defining an *inverse* $\psi : T \to S$ such that $\psi\varphi$ is the identity map on S, i.e., $\psi\varphi(x) = x$ for all $x \in S$, and $\varphi\psi$ is the identity map on T. (Note that we are composing functions *right-to-left*.) Often we will simply define φ without proving that it is a bijection when such a proof is straightforward. As we accumulate more and more objects counted by Catalan numbers, we have more and more choices

[1] A reference to a problem whose number is preceded by A refers to a problem in Chapter 4.

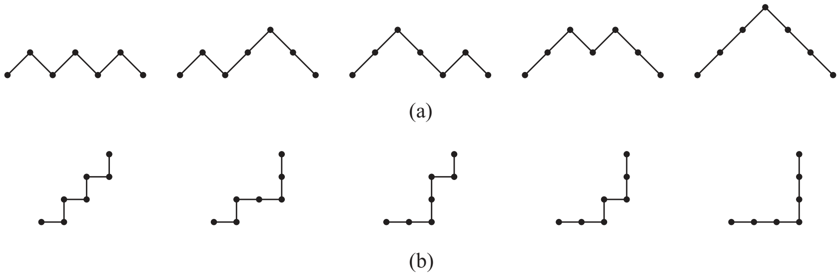
Figure 1.5. The five Dyck paths of length six.

for which of these objects we can biject to when trying to show that some new objects are also counted by Catalan numbers.

1.5.1 Theorem. *The Catalan number C_n counts the following:*

(i) *Triangulations T of a convex polygon with $n+2$ vertices.*
(ii) *Binary trees B with n vertices.*
(iii) *Plane trees P with $n+1$ vertices.*
(iv) *Ballot sequences of length $2n$.*
(v) *Parenthesizations (or bracketings) of a string of $n+1$ x's subject to a nonassociative binary operation.*
(vi) *Dyck paths of length $2n$.*

Proof. (i)→(ii) (that is, the construction of a bijection from triangulations T of polygons to binary trees B). Fix an edge e of the polygon as in Figure 1.2. Put a vertex in the interior of each triangle of T. Let the root vertex v correspond to the triangle for which e is an edge. Draw an edge f' between any two vertices that are separated by a single edge f of T. As we move along edges from the root to reach some vertex v after crossing an edge f of T, we can traverse the edges of the triangle containing v in counterclockwise order beginning with the edge f. Denote by f_1 the first edge after f and by f_2 the second edge. Then we can define the (possible) edge f_1' crossing f_2 to be the *left edge* of the vertex v of B, and similarly f_2' is the *right edge*. Thus we obtain a binary tree B, and this correspondence is easily seen to be a bijection. Our rather long-winded description should become clear by considering the example of Figure 1.6(a), where the edges of the tree B are denoted by dashed lines. In Figure 1.6(b) we redraw B in "standard" form.

(iii)→(ii) Given a plane tree P with $n+1$ vertices, first remove the root vertex and all incident edges. Then remove every edge that is not the leftmost edge from a vertex. The remaining edges are the left edges in a binary tree B

Cambridge University Press & Assessment
978-1-107-42774-7 — Catalan Numbers
Richard P. Stanley
Excerpt
More Information

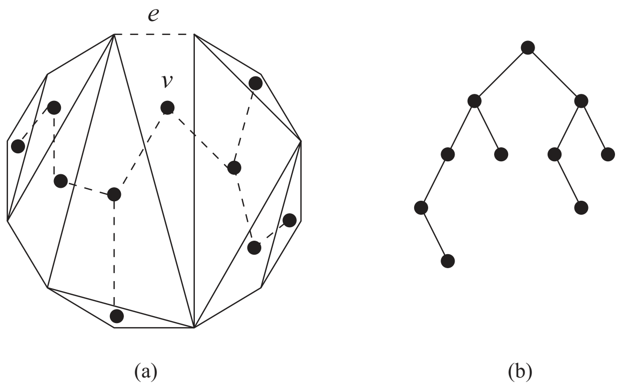

Figure 1.6. A binary tree associated with a triangulated polygon.

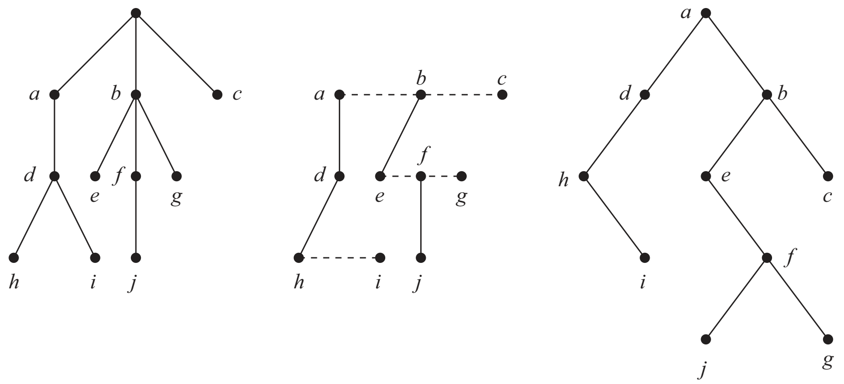

Figure 1.7. A binary tree constructed from a plane tree.

whose root is the leftmost child of the root of P. Now draw edges from each child w of a vertex v of P to the next child (the one immediately to the right of w) of v (if such a child exists). These horizontal edges are the right edges of B. The steps can be reversed to recover P from B, so the map $P \mapsto B$ gives the desired bijection. See Figure 1.7 for an example. On the left is the plane tree P. In the middle is the binary tree B with left edges shown by solid lines and right (horizontal) edges by dashed lines. On the right is B rotated 45° clockwise and "straightened out" so it appears in standard form.

This elegant bijection is due to de Bruijn and Morselt [11]. Knuth [33, §2.3.2] calls it the "natural correspondence." For some extensions, see Klarner [32].

(iii)→(iv) We wish to associate a ballot sequence $a_1a_2\cdots a_{2n}$ of length n with a plane tree P with $n+1$ vertices. To do this, we first need to define a certain canonical linear ordering on the vertices of P, called *depth first order* or *preorder*, and denoted $\text{ord}(P)$. It is defined recursively as follows.

Let $P_1,\ldots,P_m$ be the subtrees of the root v (listed in the order defining P as a plane tree). Set

$$\text{ord}(P) = v,\ \text{ord}(P_1),\ \ldots,\ \text{ord}(P_m)\ \text{(concatenation of words)}.$$

The preorder on a plane tree has an alternative informal description as follows. Imagine that the edges of the tree are wooden sticks, and that a worm begins facing left just above the root and crawls along the outside of the sticks, until (s)he (or it) returns to the starting point. Then the order in which vertices are seen for the first time is preorder. Figure 1.8 shows the path of the worm on a plane tree P, with the vertices labeled 1 to 11 in preorder.

We can now easily define the bijection between plane trees P and ballot sequences. Traverse P in preorder. Whenever we go down an edge (away from the root), record a 1. Whenever we go up an edge (toward the root), record a -1. For instance, for the plane tree P of Figure 1.8, the ballot sequence is (writing as usual $-$ for -1)

$$111-1---1-11-11-1---.$$

It is also instructive to see directly how ballot sequences are related to the recurrence (1.1). Given a ballot sequence $\beta = a_1a_2\cdots a_{2n+2}$ of length $2(n+1)$,

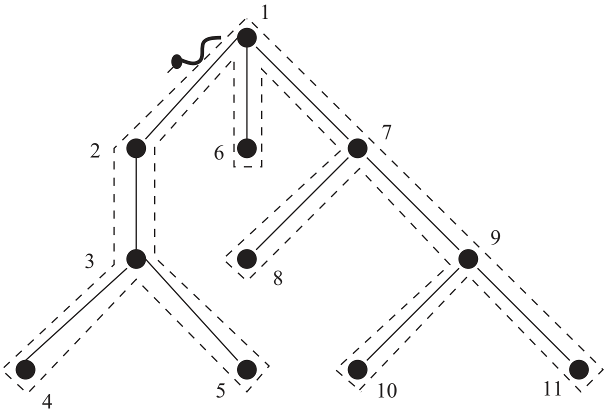

Figure 1.8. A plane tree traversed in preorder.