Motzkin Numbers

ROBERT DONAGHEY AND LOUIS W. SHAPIRO

Baruch College, City University of New York, New York 10010 and Howard University, Washington, D. C. 20001

Communicated by John Riordan

Two equations relate the well-known Catalan numbers with the relatively unknown Motzkin numbers which suggest that the combinatorial settings of the Catalan numbers should also yield Motzkin numbers. In this paper we provide a representative selection of 14 situations where the Motzkin numbers occur along with the Catalan numbers.

1. INTRODUCTION

The Catalan numbers have an extensive history and some 450 papers have appeared involving them. The bibliographies of Alter [1], Brown [2], and Gould [7] give an excellent account of this literature. In contrast, only a handful of papers [3, 5, 8–11] have appeared on the closely related sequence of Motzkin numbers m_n defined by

$$m(x) = 1 + xm(x) + x^2 m^2(x) = \sum_{n=0}^{\infty} m_n x^n. \tag{1}$$

In Section 2 we examine the algebraic relations between the Motzkin and Catalan numbers which suggest that situations in which Catalan numbers appear should also give rise to Motzkin numbers.

In Section 3 we examine a representative selection of 14 situations where the Motzkin numbers occur, establishing briefly in each that the enumerating sequence is indeed the Motzkin numbers. Many more situations (≈ 40) are known in which the Motzkin numbers appear, and are available from the authors.[1]

---

[1] D. G. Rogers has recently come up with a situation involving quasi-similarity relations not on our list.

291

2. RELATIONS BETWEEN MOTZKIN AND CATALAN NUMBERS

The Motzkin numbers first appear in [9] in a circle chording setting. There the numbers $\{m_n\}_{n=0}^{\infty} = 1, 1, 2, 4, 9, 21, 51, 127,\ldots$ are shown to enumerate the number of ways of connecting a subset of n points on a circle by non-intersecting chords. For $n = 4$ these are:

(2)

Restricting attention to chordings which meet every point on the circle yields a well-known Catalan family (there are $c_n = (2n)!/n!\,(n + 1)!$ complete chordings of $2n$ points on a circle), from which it follows by summing over the number of points which meet a chord, that:

$$m_n = \sum_{k=0}^{} \binom{n}{2k} c_k .$$
(3)

However, the Motzkin and Catalan numbers also satisfy the equation

$$b_n = c_{n+1} = \sum_{k=0}^{} \binom{n}{k} m_k .$$
(4)

In fact, combining Eqs. (3) and (4) leads to the identity

$$c_{n+1} = \sum_{k=0}^{} \binom{n}{k} \sum_{j=0}^{} \binom{k}{2j} c_j = \sum_{j=0}^{} \binom{n}{2j} c_j 2^{n-2j}$$
(5)

which appears in Touchard [13, 1924] (cf. [5] for combinatorial proofs of (4) and (5)).

We define a third related sequence by $a_{2n} = c_n$, $a_{2n+1} = 0$ (the aerated Catalan numbers). The three sequences have as generating functions

$$a(x) = (1 - (1 - 4x^2)^{1/2})/2x^2 \qquad (=c(x^2)),$$
(6)

$$m(x) = ((1 - x) - ((1 - x)^2 - 4x^2)^{1/2})/2x^2,$$
(7)

$$b(x) = ((1 - 2x) - ((1 - 2x)^2 - 4x^2)^{1/2})/2x^2 \qquad (=(c(x) - 1)/x).$$
(8)

The Euler transformation, defined by $T(f(x)) = (1/(1 + x)) f(x/(1 + x))$, maps $b(x)$ to $m(x)$, and $m(x)$ to $a(x)$. Heuristically, it might now be reasonable to expect one occurrence of the Motzkin numbers for each two occurrences of the Catalan numbers.

3. Settings for the Motzkin Numbers

(M1) Consider a random walk, beginning at the origin, on the non-negative integers. Requiring that at each step we move from p to either $p + 1$ or $p - 1$ yields a Catalan family, but if we also allow moves in place we get Motzkin numbers. The following table, $m(n, k)$, enumerates the number of paths from 0 to k in n steps. We will refer to this as the Motzkin triangle.

		Position k						
		0	1	2	3	4	5	6
Number of moves	0	1						
n	1	1	1					
	2	2	2	1				
	3	4	5	3	1			
	4	9	12	9	4	1		
	5	21	30	25	14	5	1	
	6	51	76	69	44	20	6	1

$$(9)$$

Note that $m(n + 1, k) = m(n, k - 1) + m(n, k) + m(n, k + 1)$, the three terms on the right corresponding to the three points from which the point k can be reached in a single step.

An alternate interpretation is that these are the ballot numbers modified to allow abstentions. Hence setting $n = l + 2k$, where l is the number of abstentions, and summing over l,

$$m(n, 0) = \sum_{l=0} \binom{n}{l} c_k = \sum_{k=0} \binom{n}{2k} c_k$$
$$= m_n \text{ by (3).} \tag{10}$$

Note that the second column has $xm^2(x)$ as its generating function and the kth column has $x^{k-1}m^k(x)$ as its generating function.

Note also that the diagonal sums $\sum_k m(n - k, k) = 1, 1, 2 + 1, 4 + 2, 9 + 5 + 1, \ldots$ are essentially the sequence $\{\gamma_n\}_{n=0}^{\infty} = 1, 0, 1, 1, 3, 6, 15, 36, 91, \ldots$ defined by $m_n = \gamma_{n+1} + \gamma_n$ which appear in [10]. These numbers γ_n, which are also the alternating row sums $\sum_k (-1)^k m(n, k)$, further satisfy $c_n = \sum_{k=0}^{n} \binom{n}{k} \gamma_k$, and will reappear as we progress.

The row sums $M_n = \sum_k m(n, k)$ clearly staisfy $M_{n+1} + m_n = 3M_n$ (since each entry save $m(n, 0)$ contributes to three terms in the next row). The sums M_n also satisfy the identity

$$M_n = \tfrac{1}{2}(\beta_n + \beta_{n+1}),$$

where β_n [12, Seq. 1070] is the coefficient of t^n in $(1 + t + t^2)^n$, with g.f. $\beta(x) = 1/((1 - x)^2 - 4x^2)^{1/2}$ (compare with Eq. (7)). These numbers are close cousins of the Motzkin numbers in that $\beta_n = \sum_{k=0} \binom{n}{2k}(k + 1) c_k$ and $(n + 1) c_n = \sum_{k=0} \binom{n}{k} \beta_k$ (compare with Eqs. (3) and (4)). It follows that the row sums M_n satisfy $(2n + 1) c_n = \sum_{k=0} \binom{n}{k} M_k$ (the details are omitted).

(M2)   The ballot problem is equivalent to a series of competitions in which each game yields one point to the victor, and in which player A is never behind player B in points. However, if we allow ties, with the point split $\frac{1}{2} - \frac{1}{2}$ in this event, then a Motzkin family arises.

On the grid of half-integers in the first quadrant up to the main diagonal, let a win by A be represented by an edge from (p, q) to $(p + 1, q)$, a win by B be represented by an edge from (p, q) to $(p, q + 1)$, and a tie be represented by an edge from (p, q) to $(p + \frac{1}{2}, q + \frac{1}{2})$. Then for a four game series ending in a tie we have nine possible situations:


| TTTT | ABTT | TABT | TTAB | ATBT | TATB | ATTB | ABAB | AABB |
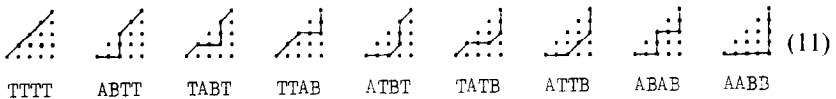
By tracking A's lead over B, a tie yields no change, so A can be k points ahead of B after n games in $m(n, k)$ ways, returning us to situation M1 above.

(M3)   The rooted plane trees are a classical Catalan family. Adding half-edges, or loops, yields a Motzkin family. We illustrate the trees with $2k + l = 4$, k counting the edges and l counting the loops (i.e., with each loop weighted as one-half of an edge):

(12)

To enumerate these trees with loops we modify the coding principle described in [4]:

$$\to \quad ULDL \qquad (13)$$
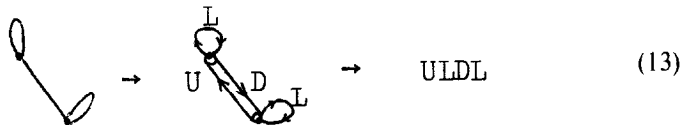
Proceeding from left to right and starting at the root, we use a U for "up," an L for a loop, and a D for "down."

Then map this sequence of letters to family M1 by replacing U by a step to the right, L by a step in place, and D by a step to the left.

(M4)   This setting is the family of increasing bipartite graphs whose edges are a subset, S, of $\{1,..., n\} \times \{1,..., n\}$ satisfying the restraints:

(1) if $(a_1, b_1),..., (a_k, b_k)$ are the ordered pairs of S and are ordered so that $a_1 < a_2 < \cdots < a_k$, then $b_1 < b_2 < \cdots < b_k$;

(2) $a_i \leqslant b_i$, $i = 1(1)k$;

(3) $\{a_1, b_1,..., a_k, b_k\} = \{1,..., n\}$.

Conditions 1 and 2 alone specify a new Catalan family, with c_{n+1} of these graphs satisfying the first two conditions. Condition 3, requiring that the edges meet every row of the graph, restricts us to a Motzkin subfamily. We illustrate for $n = 1(1)4$:

$$(14)$$

For each number m in $\{1,..., n\}$, if m is only in $\{a_1,..., a_k\}$ then step to the right, and if m is only in $\{b_1,..., b_k\}$ then step to the left. If m is in both sets, then step in place (it must be in at least one by condition 3). Let m go from 1 to n and we obtain a random walk as specified in M1.

Alternatively, this family can be mapped to the Catalan family specified by conditions 1 and 2 by relaxing condition 3. This is done by adding an arbitrary number of extra rows of dots around the n rows present. This can be done in $\binom{N}{n}$ ways, where N is the number of rows in the resulting graph. Summing over all $n \leqslant N$, with N fixed, yields Eq. (4).

(M5) The numbers presented here were called the total information numbers in [11] but turn out to be the Motzkin numbers. Suppose we have n distinct weights $w_1 < w_2 < \cdots < w_n$ and a crude balance that perhaps can not well order them. The balance is faithful, though, in that if it can not distinguish between two weights, then it cannot distinguish between these two weights and any intermediate weight.

It is not necessary, however, that the balance distinguish between every two weights for it to be possible to well order the n weights. (For example, with $n = 3$, it is enough to have w_1 distinguished as less than w_3, and w_2 not distinguishable from w_1 or w_3.)

As was shown in [11], the number of situations in which the well ordering can be deciphered yields the Motzkin numbers, while the total number of distinct partial orderings that might be induced by these kinds of scales yields the Catalan numbers. Calling $\{w_i, w_{i+1},..., w_j\}$ a run if it is a maximal set of indistinguishable weights (with possibly $i = j$), the necessary and sufficient condition for the well ordering to be deciphered is that, for each i, either w_i ends a run or w_{i+1} starts a run (or both). This family can be converted

quickly into family M4 above by mapping each run $\{w_i, w_{i+1}, ..., w_j\}$ to the ordered pair (i, j).

(M6) Modifying the ballot problem by requiring that B never receives three votes in a row yields a Motzkin family. (Here we do not allow abstentions or draws.) We tabulate the number of ways, $M(n, k)$, of A being k votes ahead of B after the nth vote cast:

		A's lead over B k									
		0	1	2	3	4	5	6	7	8	
Number of votes	0	1									
n	1	0	1								
	2	1	0	1							
	3	0	2	0	1						
	4	2	0	3	0	1					(15)
	5	0	5	0	4	0	1				
	6	4	0	9	0	5	0	1			
	7	0	12	0	14	0	6	0	1		
	8	9	0	25	0	20	0	7	0	1	

The entry $M(n, k)$ is the sum $M(n - 1, k - 1) + M(n - 2, k) + M(n - 3, k + 1)$, the three terms corresponding to the three sequences of votes leading up to vote n:

$\cdots A$, with $k - 1$ more A's than B's in the "$\cdots$";

$\cdots AB$, with k more A's than B's in the "$\cdots$";

$\cdots ABB$, with $k + 1$ more A's than B's in the "$\cdots$"

(the sequence $\cdots BBB$ is impossible because B cannot get three votes in a row). After rearranging, both the numbers and the law of formation are the same as in the Motzkin triangle in family M1.

(M7) Another family of bipartite graphs like those in family M4 can be formed by requiring:

(1) $a_1 < \cdots < a_k \Rightarrow b_1 < \cdots < b_k$;

(2) $a_i \leqslant b_i$;

the first two conditions given in family M4, and requiring further:

(3') For every $m \in \{1, ..., n - 1\}$, $\{m, m + 1\} \cap \{a_1, ..., a_k\} \neq \varnothing$.

Hence at least every other point in the left column meets an edge. For example, when $n = 3$ the graphs are:

$$\text{(16)}$$

Note that the number of distinct sets $\{a_1, ..., a_k\}$ is given by the Fibonacci numbers.

This family can be mapped to the ballot family M6 above by mapping each ordered pair (a_i, b_i) to the point (b_i, a_i) in the first quadrant, and drawing a ballot path from $(0, 0)$ to (n, n) passing through these points but otherwise remaining as far from the main diagonal as possible. For example, the graphs in (15) map to the ballot paths:

$$\text{(17)}$$

The condition that either m or $m + 1 \in \{a_1, ..., a_k\}$ is equivalent to B never getting three votes in a row.

(M8) Consider the rooted plane trees in which no vertex, the root excepted, has degree 2. These trees have been called the branch reduced trees in [5], but are called bushes here, for short. For example, the nine bushes with five edges are:

$$\text{(18)}$$

If we let $\tilde{m}_k$ enumerate the k edge bushes, we can relate the $\tilde{m}_k$ to the Catalan numbers as follows.

There are c_n rooted plane trees with n edges, and eliminating all vertices, save the root, of degree 2 by merging edges maps these trees to the bushes. Reversing this process requires starting with the bushes of up to n edges and adding, for a k edge bush, $n - k$ extra vertices midedge (with possibly more than one extra vertex per edge) in any of $\binom{k+(n-k)-1}{(n-k)}$ ways. Hence, summing over k,

$$c_n = \sum_{k=1}^{} \binom{k + (n - k) - 1}{(n - k)} \tilde{m}_k = \sum_{k=1}^{} \binom{n - 1}{k - 1} \tilde{m}_k. \tag{19}$$

Comparing Eqs. (18) and (4) leads to the result that $\tilde{m}_n = m_{n-1}$.

This brief argument and result is from [5], but is repeated here for completeness. Note that the same numbers γ_n which appeared in family M1 reappear here. There are γ_{n+1} such bushes where the root has degree greater than 1, and thus γ_n bushes where the root has degree 1.

(M9)   The bushes of family M8 can immediately be translated into a general bracketing (g.b.) subfamily. Consider all g.b.'s with no occurrence of $((*))$, where $*$ is itself any (possibly vacuous) g.b.

We use the same translation employed in family M3 above, so that for example,


$$\text{UDUUDUDD} \;\rightarrow\; (\,)((\,)(\,)) \tag{20}$$
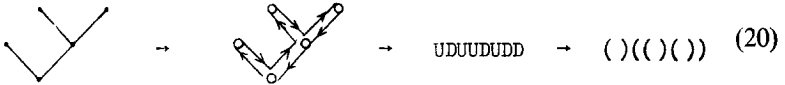
where for the last step we replace each U by "(" and each D by ")." A vertex other than the root of degree 2 would correspond to $((*))$. The nine g.b.'s corresponding to (17) are:

$$()()()()(),\quad (()()()()(),\quad ()(()()(),\quad ()()(()()()),\quad (()()()())(),$$

$$()(()()())),\quad (()()()()()),\quad ((()()()())),\quad (()(()()())). \tag{21}$$

Note that the g.b.'s contained within a single outermost pair of parentheses are enumerated by γ_n, those not so contained by γ_{n+1}.

(M10)   Consider those plane trees in which the edge leading down from each tip vertex of degree 1 is the left-most edge above the vertex below. For five edges, these trees are:
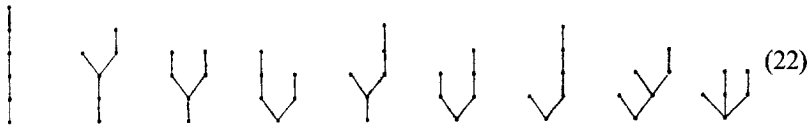
These trees have in common that there is never a terminating edge to the right of any other edge, so to map these trees to the bushes of family (M8) we modify the procedure of [4] as follows:
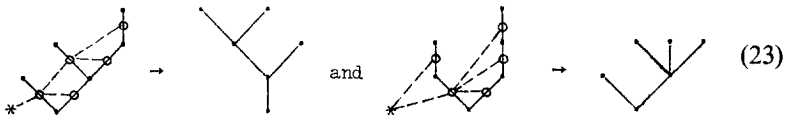

Note that
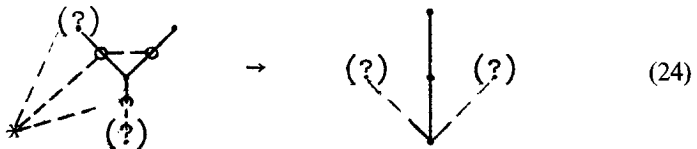
so the lack of any terminating edge to the right of another edge guarantees that no vertex in the image of the map, the root excepted, can have degree 2. The trees of (22) are mapped to the corresponding trees of (18).

Note that the trees in which a terminating edge adjoins the root are enumerated by γ_n, and there are γ_{n+1} trees without such a terminal edge adjoining the root.

(M11)   The ballot problem can be modified by requiring that A always receives votes in blocks of two or more. This restricted family is enumerated by γ_n.

For example, the 10 vote paths are:

 $\qquad$ (25)
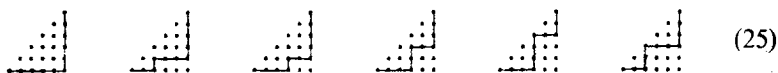
There are $\gamma_5 = 6$ such paths.

Allowing the voting sequence to *begin* $AB \cdots$ as well as $AA \cdots$ yields the extra γ_{n-1} paths, and since $\gamma_n + \gamma_{n+1} = m_n$ this yields a Motzkin family. We can map this family to family M10 by changing all A's to U's and all B's to D's and letting the sequences of U's and D's be up–down codes for the rooted trees, reversing the process described in families M3 and M9.

(M12)   Related to the ballots of family M11 above are the increasing bipartite graphs satisfying:
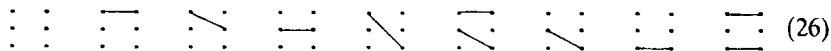
(1) $a_1 < \cdots < a_k \Rightarrow b_1 < \cdots < b_k$;

(2) $a_i \leqslant b_i$;

and the extra restriction (see families M4 and M7):

(3″) $m \in \{b_1,...,b_k\} \Rightarrow \{m-1, m+1\} \cap \{b_1,...,b_k\} = \varnothing$
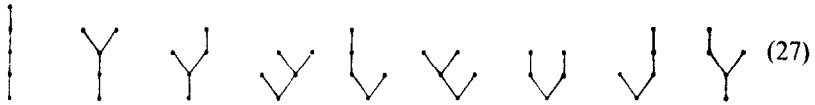   [or: $b_i + 1 < b_{i+1}$, $i = 1(1)k - 1$].

Hence no two consecutive points in the right column both meet edges. For example, when $n = 3$ the graphs are:
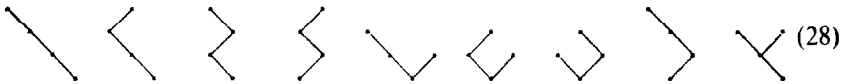
 (26)

These graphs are mapped to family M11 above in exactly the same way that family M7 was mapped to family M6.

(M13)   The rooted trees in which every vertex has degree at most 3, and in which the root has degree at most 2, have appeared in both [5, 8] and are enumerated by the Motzkin numbers. For example, the trees with four edges are:

$$\Big| \quad \Upsilon \quad \dot{\Upsilon} \quad \dot{\sqrt{}} \quad \dot{\mathsf{L}} \quad \dot{\vee} \quad \mathsf{U} \quad \mathsf{J} \quad \dot{\mathsf{V}} \quad (27)$$

Klarner [8] and Motzkin [9] derived different, but equivalent generating function equations for the Motzkin numbers. Klarner also obtained, in this setting, a closed form for $m(x)$ trivially equivalent to (7). A second enumeration of this family is provided in [5] using Eq. (4).

(M14)   Binary trees, called zig-zag trees in [5], in which every edge slants up to the left or right, are a Catalan family (they are actually the interiors of the trivalent plane trees). Considering only those binary trees in which no two consecutive edges slant to the right, we obtain our last Motzkin family. With three edges these are:

$$\diagdown \quad \langle \quad \rangle \quad \zeta \quad \diagdown\!\!\!\vee \quad \diagup\!\!\!\diagdown \quad \diagdown\!\!\!\rangle \quad \diagdown \quad \diagdown\!\!\!\diagup \quad (28)$$

By applying another map given in [4] we find that these trees map to the trees of family M13 above, with the trees of (28), in particular, mapping to the corresponding trees of (27). We illustrate this process with the middle three of the above nine trees, which we first tilt 45° to the right:

$$(29)$$
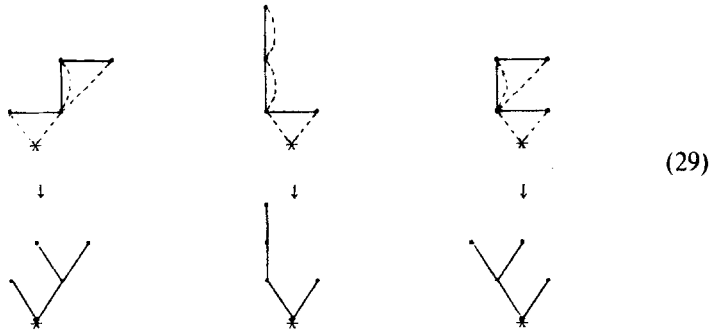
The condition of no consecutive right (horizontal) edges implies the condition of maximum degree 3.

Note.   The following bibliography is, at the time of this writing, believed to be complete with regards to the Motzkin numbers. The authors would appreciate any further applicable references.

REFERENCES

1. R. ALTER, Some remarks and results on Catalan numbers, *in* "Proceedings of the Second Louisiana Conference on Combinatorics, Graph Theory, and Computing, Baton Rouge, 1971, pp. 109–132.
2. W. G. BROWN, Historical note on a recurrent combinatorial problem, *Amer. Math. Monthly* **72** (1965), 973–977.
3. L. CARLITZ, Solutions of certain recurrences, *SIAM J. Appl. Math.* **17** (1969), 251–259.
4. N. G. DEBRUIJN AND B. J. M. MORSELT, A note on plane trees, *J. Combinatorial Theory* **2** (1967), 27–34.
5. R. DONAGHEY, Restricted plane tree representations of four Motzkin–Catalan equations, *J. Combinatorial Theory Ser. B* **22** (1977), 114–121.
6. R. DONAGHEY, Binomial self-inverse sequences and tangent coefficients, *J. Combinatorial Theory Ser. A* **21** (1976), 155–163.
7. H. W. GOULD, Research bibliography of two special number sequences, *Math. Monongaliae* **12** (1971).
8. D. KLARNER, Correspondence between plane trees and binary sequences, *J. Combinatorial Theory* **9** (1970), 401–411.
9. T. MOTZKIN, Relations between hypersurface cross ratios, and a combinatorial formula for partitions of a polygon, for permanent preponderance, and for non-associative products, *Bull. Amer. Math. Soc.* **54** (1948), 352–360.
10. J. RIORDAN, Enumeration of plane trees by branches and endpoints, *J. Combinatorial Theory Ser. A* **19** (1975), 214–222.
11. L. SHAPIRO, Catalan numbers and total information numbers, *in* "Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory, and Computing, Boca Raton, 1975," pp. 531–539.
12. N. J. A. SLOANE, "A Handbook of Integer Sequences," Academic Press, New York, 1973.
13. J. TOUCHARD, Sur certaines equations fonctionelles, *in* "Proc. Inter. Congr. Math.," pp. 465–472, Univ. of Toronto Press, Toronto, 1928.