Combinatorial Proof of an Abel-type Identity*

P. Mark Kayll
Department of Mathematical Sciences, University of Montana
Missoula MT 59812-0864, USA
mark.kayll@umontana.edu

David Perkins
Department of Mathematics and Computer Science
Houghton College, Houghton NY 14744, USA
david.perkins@houghton.edu

Identity (1) below resulted from our investigation in [21] of chip-firing games on complete graphs K_n, for $n \geq 1$; see, e.g., [2] for antecedents. The left side expresses the sum of the probabilities of a game experiencing firing sequences of each possible length $\ell = 0, 1, \ldots, n$. This note gives a combinatorial proof that these probabilities sum to unity.

We first manipulate

$$\frac{n-1}{n+1} + \sum_{\ell=1}^{n} \binom{n}{\ell} \frac{\ell^{\ell-1}(n+1-\ell)^{n-1-\ell}}{n(n+1)^{n-1}} = 1 \tag{1}$$

into a form amenable to combinatorial proof. Multiplying by $n(n+1)^n$ and using the relation $\binom{n}{\ell} = \binom{n+1}{\ell}(n+1-\ell)/(n+1)$, we transform (1) to the equivalent form

$$\sum_{\ell=1}^{n} \binom{n+1}{\ell} \ell^{\ell-2}(n+1-\ell)^{n-1-\ell} \ell(n+1-\ell) = 2n(n+1)^{n-1}. \tag{2}$$

To see that (2) holds, first observe that the right side enumerates the pairs $(T, \vec{e}\,)$, where T is a spanning tree of K_{n+1} for which one edge e (of

2000 MSC: Primary 05A19; Secondary 05C30, 60C05.
*Preprint to appear in *J. Combin. Math. Combin. Comput.*

its n edges) has been distinguished and oriented (in one of two possible directions). The left side also enumerates these pairs. Given $(T, \vec{e}\,)$, notice that deleting the oriented edge $\vec{e}$ from T leaves behind a spanning forest of K_{n+1} with two components L, R (that we may consider ordered from left to right). If $|V(L)| = \ell$, for an integer ℓ with $1 \le \ell \le n$, then $|V(R)| = n+1-\ell$. Conversely, given such a spanning forest, we can recover $(T, \vec{e}\,)$ by selecting a node x of L and a node y of R and letting $\vec{e} = (x, y)$. On the left side of (2), the factor $\binom{n+1}{\ell}$ accounts for the selection of $V(L)$ (hence for the selection of $V(R)$). Since L, R are, respectively, spanning trees of the induced (complete) subgraphs $K_{n+1}[V(L)]$, $K_{n+1}[V(R)]$, the factors $\ell^{\ell-2}$ and $(n+1-\ell)^{n-1-\ell}$ are delivered by Cayley's Formula [5]. Finally, the factors ℓ, $(n+1-\ell)$ count the number of ways to select the vertices $x \in V(L)$ and $y \in V(R)$ determining $\vec{e}$.

Epilogue

Having attempted to prove (2) before realizing that it follows from established identities, we fortuitously discovered the proof above. Here, we outline a multitude of connections between (2) and the work of authors who preceded us. Our presentation order is mainly chronological.

For integers r, n with $1 \le r \le n$, we denote by T_n the number of spanning trees of K_n and by $F_{n,r}$ the number of spanning forests of K_n that consist of r disjoint trees such that r specified nodes (*roots*) belong to distinct trees. We thrice invoked Cayley's Formula, $T_n = n^{n-2}$, in our proof above and shall connect one of its generalizations,

$$F_{n,r} = rn^{n-r-1}, \tag{3}$$

to (2) in our discussion below.

As it turns out, (2) goes back (at least) to 1917, when Dziobek [10] used it in the form

$$\sum_{\ell=1}^{n-1} \binom{n}{\ell} T_\ell T_{n-\ell}\, \ell(n-\ell) = 2(n-1)T_n \tag{4}$$

to give, among other results, an inductive proof of Cayley's Formula. Bol [3] also proved and applied (4) in deriving a generating-function identity for the numbers T_n. Of course, once Cayley's Formula is known, then (4) yields a proof of (2).

Clarke [6] determined the number $T_{n,d}$ of spanning trees of K_n with a specified node of valency $d \le n-1$: $T_{n,d} = \binom{n-2}{d-1}(n-1)^{n-d-1}$. Thus, the number of spanning trees of K_{n+2}, with, say, the node $(n+2)$ of valency

2, is

$$T_{n+2,2} = n(n+1)^{n-1}. \tag{5}$$

One obtains a combinatorial proof of (5)—reminiscent of our proof of (2)—by reverse-subdividing at the bivalent node $(n+2)$ and noticing that then $T_{n+2,2}$ enumerates the spanning trees of K_{n+1} with a distinguished (but not oriented) edge. In due course, we indicate a more direct connection between (5) and (2).

The year 1959 saw two more appearances of (2) in the literature. Dénes [9] proved (2) in an article where he showed that T_n gives the number of ways of representing a cyclic permutation $(1, 2, \ldots, n)$ as a product of $(n-1)$ transpositions. Then Rényi [23], in proving (3), established a generalization of (2) and cited [10, 9] as two earlier sources of the identity. Although (3) was stated by Cayley [5], Rényi's paper probably contains the first published proof of this result. Close on its heels, however, was an argument by Göbel [15], who showed that

$$F_{n,r} = \sum_{k=1}^{n-r} \binom{n-r}{k} r^k F_{n-r,k} \tag{6}$$

and then used this recurrence to prove (3) inductively.

Busacker and Saaty [4, p. 137] possibly marked the first (discrete mathematical) textbook appearance of (2), that formed the starting point for the authors' treatment of Dziobek's proof [10] of Cayley's Formula. Because no proof of (2) was supplied, the authors omitted some essential details, and, curiously, they referenced Rényi [23] but not Dziobek.

An exercise in Riordan's book [24, p. 116] directs the reader to derive (algebraically) the relation

$$\sum_{\ell=1}^{n} \binom{n}{\ell} \ell^{\ell-1}(n+1-\ell)^{n-\ell} = n(n+1)^{n-1}. \tag{7}$$

Using (7) and Pascal's Identity, a routine calculation leads to (2). The agreement between the right sides of (7) and (5) reveals the more direct connection between Clarke's expression for $T_{n+2,2}$ and (2), to which we alluded earlier.

The relation (7) specializes an identity belonging to a theory initiated by Abel. In 1826, he published a generalization [1] of the Binomial Theorem that spawned many identities and led eventually to the theory of polynomials of 'binomial-type'; see [20] or, e.g., [26, Exercise 5.37] for a more recent reference. In deriving one of these identities (specifically (8) below), Riordan [24, pp. 18–23] studied the class of 'Abel sums' of the form

$$A_m^{p,q}(x,y) := \sum_{k=0}^{m} \binom{m}{k} (x+k)^{k+p}(y+m-k)^{m-k+q}.$$

He gave an elementary proof that

$$A_m^{-1,0}(x,y) = x^{-1}(x+y+m)^m, \tag{8}$$

which, when combined with the 'Pascal-like' recurrence

$$A_m^{p,q}(x,y) = A_{m-1}^{p,q+1}(x,y+1) + A_{m-1}^{p+1,q}(x+1,y)$$

and the observation that

$$A_m^{p,q}(x,y) = A_m^{q,p}(y,x),$$

leads to

$$A_m^{-1,-1}(x,y) = (x^{-1}+y^{-1})(x+y+m)^{m-1}. \tag{9}$$

Clearing the denominator on the right side of (9) gives the polynomial identity

$$\sum_{k=0}^{m} \binom{m}{k} x(x+k)^{k-1} y(y+(m-k))^{m-k-1} = (x+y)(x+y+m)^{m-1}, \tag{10}$$

which shows that the sequence $\{z(z+m)^{m-1}\}_{m\geq 0}$ is of binomial type (*cf.* [27]).

It is not difficult to see that (2), in which $n \geq 1$, is equivalent to

$$\sum_{k=0}^{m} \binom{m}{k} (k+1)^{k-1}(m+1-k)^{m-1-k} = 2(m+2)^{m-1}, \tag{11}$$

in which $m := n-1 \geq 0$. Putting $x = y = 1$ in (10) yields (11) and thus delivers a sixth proof of (2). Notice that (3) gives

$$F_{m+2,2} = 2(m+2)^{m-1}; \tag{12}$$

this enumerative interpretation of the right side of (11) leads to a direct combinatorial proof of (11) along the lines of our proof of (2).

Evidently, forest and tree enumeration played a central role in many of these earlier results. Moon's monograph [18, especially Chapters 3, 4] presents an exceptionally well-organized account of these results to 1970, including more detailed discussions of the references [5, 10, 3, 6, 9, 23, 15, 24] than is feasible to present here. Since its appearance, at least seven more research articles have addressed results related to (2) and to identities we introduced in connection with (2). These articles present a surprising variety of contexts where (2) and its relatives have arisen.

Cooper [8] derived (2) by analyzing the busy period for a single-server queue with Poisson input and constant service times. Because he arrived at (2) by first deriving a version of (9), the final step of his proof is equivalent to the "sixth proof" mentioned above. Françon [13] proved combinatorially several identities of binomial-type, including (10). Phrased in terms of labelled rooted forests, this proof was perhaps the first within our scope to exploit the function-counting technique introduced by Foata and Fuchs [12]; see [7, p. 129] for a textbook treatment. Getu and Shapiro [14] presented another combinatorial proof of (10)—also using the Foata-Fuchs encoding—that invoked (3) as their main lemma. This is a step up in the level of generalization from the direct combinatorial proof of (11) that we mentioned above: for (11) is a specialization of (10), and (12) is a specialization of (3). Another decade later, Shapiro [25] produced a combinatorial proof of a generalization of (10). A seven-line computer-generated proof of (8) appeared as [11]. Finally, Pitman [22, p. 178] reincarnated both (3) and (6) (resp., as (10) and (11) in [22]).

Beyond the handful of textbook-length works [4, 24, 18, 7, 26] already cited, we must also mention [17]. Exercise 1.44 [*ibid.*, p. 22] comprises our (8), (9), and (2), that Lovász obtained via formal manipulations and induction. Exercise 4.6 [*ibid.*, p. 34] establishes the identity

$$\sum_{\ell=1}^{n-1} \ell \binom{n-2}{\ell-1} T_\ell T_{n-\ell} = T_n \tag{13}$$

as a basis for deriving Cayley's Formula. Lovász deduced (13) using an argument similar to our proof of (2) (but involving over-counting) and then used (13) to obtain (4). From there, he presented Dziobek's proof [10] of Cayley's Formula and along the way derived and applied (2).

Despite the glut of earlier occurrences of (2), our proof is, as far as we can tell, new. Incidentally, our approach also gives a combinatorial proof of the American Mathematical Monthly Problem 4984, for which the published solution [16] relied directly on Abel's Identity.

An early draft of the present article closed with a problem. Riordan [24, p. 97] connected a generalized version of the right side of (11) to an enumeration problem for forests of labelled rooted trees. As a special case, he derived (p. 117) the formula

$$\sum_{k=0}^{m} \binom{m}{k} k\, 2^k m^{m-1-k} = 2(m+2)^{m-1}, \tag{14}$$

which led to our problem: find a combinatorial proof that the left sides of (11) and (14) coincide.

Moon [19] supplied the following solution. The paragraph containing (11) alludes to a combinatorial proof of this identity, whose left side can be seen to count the number of labelled forests on $\{1, 2, \ldots, m+2\}$ consisting of two trees rooted at two specified nodes and having an additional m labelled non-root nodes. The left side of (14) enumerates these forests according to the number k of nodes attached directly to the two roots. The binomial coefficient counts the number of ways of selecting these k nodes; (3) shows that there are km^{m-1-k} ways of forming k trees, on m labelled nodes, that are rooted at the k selected nodes, and, finally, there are 2^k ways of attaching the k selected nodes to one or the other of the two nodes originally specified as the roots of the final forest.

Moon [19] noted that his solution is a special case of the approach Göbel [15] used to derive (3), an account of which also appears in [18, p. 17]; this explains why the instance of (6) with $n = m+2$, $r = 2$ coincides with (14).

Let us close with one more combinatorial proof—also due to Moon [19]—this time for (7). Notice that the right side counts the spanning trees of K_{n+1} that are rooted at a node other than the node $(n+1)$. The left side counts the same thing by classifying these trees according to the number ℓ of nodes in the subtree obtained by discarding all the nodes in the branch, attached at the root, that contains the node $(n+1)$.

Acknowledgements

We are grateful to Karel Stroethoff for pointing us to Abel's Identity and deeply indebted to J.W. Moon, who generously supplied us with the references [3, 4, 8, 9, 10, 11, 14, 15, 18, 22, 23, 25] and the two proofs attributed in the narrative. This article was revised during the first author's sabbatical at Université de Montréal; many thanks to Geňa Hahn for making this possible.

References

[1] N.H. Abel, Beweis eines Ausdruckes, von welchem die Binomial-Formel ein einzelner Fall ist. *J. Reine Angew. Math.* 1(1826), 159–160.

[2] A. Björner, L. Lovász and P.W. Shor, Chip-firing games on graphs. *European J. Combin.* 12(1991), 283–291.

[3] G. Bol, Über eine kombinatorische Frage. *Abh. Math. Semin. Hansische Univ.* 12(1938), 242–245.

[4] R.G. Busacker and T.L. Saaty, *Finite Graphs and Networks: An Introduction with Applications* (McGraw-Hill, New York, 1965).

[5] A. Cayley, A theorem on trees. *Quart. J. Math.* 23(1889), 376–378; pp. 26–28 in: *The Collected Mathematical Papers of Arthur Cayley*, vol. 13 (Cambridge University Press, Cambridge, 1897).

[6] L.E. Clarke, On Cayley's formula for counting trees. *J. London Math. Soc.* 33(1958), 471–474.

[7] L. Comtet, *Advanced Combinatorics* (Reidel, Boston, 1974).

[8] R.B. Cooper, Derivation from queueing theory of an identity related to Abel's generalization of the binomial theorem, which is useful in graph theory. *Management Sci.* 19(1973), 582–584.

[9] J. Dénes, The representation of a permutation as the product of a minimal number of transpositions, and its connection with the theory of graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 4(1959), 63–71.

[10] O. Dziobek, Eine Formel der Substitutionstheorie. *Sitzungsber. Berl. Math. Ges.* 16(1917), 64–67.

[11] S. Ekhad and J.E. Majewicz, A short WZ-style proof of Abel's identity. *Electron. J. Combin.* 3(1996), #R16, approx. 1 p.

[12] D. Foata and A. Fuchs, Réarrangements de fonctions et dénombrement. *J. Combinatorial Theory* 8(1970), 361–375.

[13] J. Françon, Preuves combinatoires des identités d'Abel. *Discrete Math.* 8(1974), 331–343; corrigendum, *ibid.* 10(1974), 193.

[14] S. Getu and L.W. Shapiro, A natural proof of Abel's identity. *Congr. Numer.* 28(1980), 447–452.

[15] F. Göbel, *Mutual Choices*, Technical Report S 317 (VP 22) (Math. Centrum Amsterdam, Afd. Math. Statistik, 1963).

[16] H.W. Gould, Solution of advanced problem 4984. *Amer. Math. Monthly* 69(1962), 572.

[17] L. Lovász, *Combinatorial Problems and Exercises, Second edition* (North-Holland, Amsterdam, 1993).

[18] J.W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, No. 1 (Canadian Mathematical Congress, Montréal, 1970).

[19] J.W. Moon, private communication, June 2005.

[20] R. Mullin and G.-C. Rota, On the foundations of combinatorial theory III: Theory of binomial enumeration. pp. 167–213 in: B.E. Harris, ed., *Graph Theory and its Applications*, Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, WI, 1969 (Academic Press, New York, 1970).

[21] D. Perkins, *Investigations of a Chip-firing Game*, Dissertation (University of Montana, Missoula MT, 2005).

[22] J. Pitman, Forest volume decompositions and Abel-Cayley-Hurwitz multinomial expansions. *J. Combin. Theory Ser. A* 98(2002), 175–191.

[23] A. Rényi, Some remarks on the theory of trees. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 4(1959), 73–85.

[24] J. Riordan, *Combinatorial Identities* (Wiley, New York, 1968).

[25] L.W. Shapiro, Voting blocks, reluctant functions, and a formula of Hurwitz. *Discrete Math.* 87(1991), 319–322.

[26] R.P. Stanley, *Enumerative Combinatorics, Vol. 2* (Cambridge University Press, New York, 1999).

[27] V. Strehl, Identities of Rothe-Abel-Schläfli-Hurwitz-type. *Discrete Math.* 99(1992), 321–340.