A Combinatorial View of Sums of Powers

Michael Z. Spivey
University of Puget Sound
Tacoma, Washington 98416-1043 USA
mspivey@pugetsound.edu

May 20, 2021

The sum of powers of the first n positive integers,

$$1^m + 2^m + 3^m + \cdots + n^m,$$

has been a topic of mathematical study for centuries. For example, the formulas for this sum when m equals 1, 2, and 3 are widely known:

$$\begin{aligned} 1 + 2 + 3 + \cdots + n &= \frac{n(n+1)}{2}, \\ 1^2 + 2^2 + 3^2 + \cdots + n^2 &= \frac{n(n+1)(2n+1)}{6}, \\ 1^3 + 2^3 + 3^3 + \cdots + n^3 &= \frac{n^2(n+1)^2}{4}. \end{aligned}$$

In this note we describe a simple combinatorial way to view the power sum. Then we use this perspective to give combinatorial proofs of two known formulas for the power sum: one involving Stirling numbers of the second kind and one involving Eulerian numbers. Finally, we make a minor modification to the latter proof to produce a combinatorial proof of Worpitzky's identity.

A convention and some notation: Throughout this note we assume that m, n, and x are nonnegative integers, and we use the symbol $[n]$ to denote the set $\{1, 2, \ldots, n\}$.

1 The power sum, combinatorially

Here is our combinatorial interpretation of the power sum.

Theorem 1. *The power sum $\sum_{k=1}^n k^m$ is the number of functions $f : [m+1] \mapsto [n+1]$ such that, for all $i \in [m]$,*

$$f(i) < f(m+1).$$

Before we prove Theorem 1, let's take a look at a couple of its special cases in order to understand a little better what it says.

- With $m = 1$, Theorem 1 says that the sum $\sum_{k=1}^{n} k$ is the number of functions f from $\{1, 2\}$ to $\{1, 2, \ldots, n + 1\}$ with $f(1) < f(2)$.
- With $m = 2$, Theorem 1 says that the sum $\sum_{k=1}^{n} k^2$ is the number of functions f from $\{1, 2, 3\}$ to $\{1, 2, \ldots, n + 1\}$ with $f(1) < f(3)$ and $f(2) < f(3)$.

In other words, the power sum can be thought of as the number of functions $f : [m + 1] \mapsto [n + 1]$ for which $f(m + 1)$ is larger than any of the other function values.

Proof. For such a function f, let $k \in [n]$, and suppose $f(m+1) = k+1$. Since we must have $f(1) < f(m+1)$, there are k choices for $f(1)$; i.e., we must have $f(1) \in \{1, 2, \ldots, k\}$. Similarly, there are k choices for $f(2)$, k choices for $f(3)$, and so forth, up to $f(m)$, so that we have k^m total choices for the values of $f(1), f(2), \ldots, f(m)$. Summing over all possible values of k yields $\sum_{k=1}^{n} k^m$ as the total number of functions $f : [m+1] \mapsto [n+1]$ satisfying $f(i) < f(m + 1)$ for all $i \in [m]$. □

2 The power sum via Stirling numbers

Theorem 1 leads to a quick combinatorial proof of a formula for the power sum featuring the Stirling numbers of the second kind. Let's talk a little about these numbers before we discuss the formula.

The Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$ can be generated by the following recursion:

$$\left\{ {n \atop k} \right\} = k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\},$$

valid for $n + 1 \geq k \geq 1$. For the boundary cases, we have $\left\{ {n \atop n} \right\} = 1$, and, for $n > 0$, $\left\{ {n \atop 0} \right\} = 0$. The first several rows of the triangle of Stirling numbers of the second kind are given in Figure 1.

The Stirling numbers of the second kind have a combinatorial interpretation as well: They count the number of ways to partition n objects into k nonempty subsets. For example, suppose we want to partition the set $\{1, 2, 3, 4\}$ into two nonempty subsets. Here are the possibilities:

$$\begin{array}{llll} \{1\} \cup \{2,3,4\}, & \{2\} \cup \{1,3,4\}, & \{3\} \cup \{1,2,4\}, & \{4\} \cup \{1,2,3\}, \\ \{1,2\} \cup \{3,4\}, & \{1,3\} \cup \{2,4\}, & \{1,4\} \cup \{2,3\}. & \end{array}$$

Thus $\left\{ {4 \atop 2} \right\} = 7$, which we also see in Figure 1.

Let's now look at our power sum formula featuring the Stirling numbers of the second kind.

n	$\left\{ {n \atop 0} \right\}$	$\left\{ {n \atop 1} \right\}$	$\left\{ {n \atop 2} \right\}$	$\left\{ {n \atop 3} \right\}$	$\left\{ {n \atop 4} \right\}$	$\left\{ {n \atop 5} \right\}$	$\left\{ {n \atop 6} \right\}$	$\left\{ {n \atop 7} \right\}$
0	1							
1	0	1						
2	0	1	1					
3	0	1	3	1				
4	0	1	7	6	1			
5	0	1	15	25	10	1		
6	0	1	31	90	65	15	1	
7	0	1	63	301	350	140	21	1

Figure 1: Rows 0 through 7 of the triangle of Stirling numbers of the second kind

Theorem 2.

$$\sum_{k=1}^{n} k^m = \sum_{k=0}^{m} \binom{n+1}{k+1} \left\{ {m \atop k} \right\} k!.$$

Proof. By Theorem 1, the left side is the number of functions $f : [m+1] \mapsto [n+1]$ such that, for all $i \in [m]$, $f(i) < f(m+1)$.

For the right side, let $k \in \{0, 1, \ldots, m\}$, and suppose $k+1$ is the number of elements in the image of f. Then there are $\binom{n+1}{k+1}$ ways to choose exactly which elements these will be. The largest element must be $f(m+1)$, and the remaining m elements in the domain must be mapped to the other k elements in the image. There are $\left\{ {m \atop k} \right\}$ ways to assign the remaining m elements to k nonempty groups, and then $k!$ ways to assign these groups to the k remaining elements in the image. (More generally, the number of onto functions from an m-set to a k-set is $\left\{ {m \atop k} \right\} k!$.) Finally, sum over all possible values of k. □

As an example, take $m = 3$. Then Theorem 2 becomes

$$\sum_{k=1}^{n} k^m = \binom{n+1}{2} + 6\binom{n+1}{3} + 6\binom{n+1}{4}.$$

We see, then, that Theorem 2 gives us a formula for the power sum in terms of the binomial coefficients in row $n+1$ of Pascal's triangle.

3 The power sum via Eulerian numbers

Our second formula for the power sum involves the Eulerian numbers. Let's take a look at those numbers now.

As with the Stirling numbers of the second kind, the Eulerian numbers $\left\langle {n \atop k} \right\rangle$ satisfy a simple two-term recurrence relation. We have

$$\left\langle {n \atop k} \right\rangle = (k+1)\left\langle {n-1 \atop k} \right\rangle + (n-k)\left\langle {n-1 \atop k-1} \right\rangle,$$

valid for $n+1 \geq k \geq 1$. For the boundary cases, we have $\left\langle {n \atop 0} \right\rangle = 1$ and, for $n > 0$, $\left\langle {n \atop n} \right\rangle = 0$. The first several rows of the triangle of Eulerian numbers are given in Figure 2.

n	$\left\langle {n \atop 0} \right\rangle$	$\left\langle {n \atop 1} \right\rangle$	$\left\langle {n \atop 2} \right\rangle$	$\left\langle {n \atop 3} \right\rangle$	$\left\langle {n \atop 4} \right\rangle$	$\left\langle {n \atop 5} \right\rangle$	$\left\langle {n \atop 6} \right\rangle$	$\left\langle {n \atop 7} \right\rangle$
0	1							
1	1	0						
2	1	1	0					
3	1	4	1	0				
4	1	11	11	1	0			
5	1	26	66	26	1	0		
6	1	57	302	302	57	1	0	
7	1	120	1191	2416	1191	120	1	0

Figure 2: Rows 0 through 7 of the triangle of Eulerian numbers

Also like the Stirling numbers of the second kind, the Eulerian numbers have a combinatorial interpretation. To explain this combinatorial interpretation we need to provide some background.

Given a permutation $\pi : [m] \mapsto [m]$, we say that i, $1 \leq i \leq m-1$, is an *ascent* of π if $\pi(i) < \pi(i+1)$. For example, suppose we have the permutation

$$\pi = (5, 7, 2, 1, 6, 4, 9, 3, 8)$$

of $\{1, 2, \ldots, 9\}$ (so that $\pi(1) = 5$, $\pi(2) = 7$, and so forth). There are four locations in this permutation for which $\pi(i) < \pi(i+1)$: from 5 to 7, from 1 to 6, from 4 to 9, and from 3 to 8. Thus π has four ascents; they occur at positions 1, 4, 6, and 8. (For emphasis, these are underlined here: $(\underline{5}, 7, 2, \underline{1}, 6, \underline{4}, 9, \underline{3}, 8)$.)

The Eulerian number $\left\langle {m \atop k} \right\rangle$ counts the number of permutations on $[m]$ containing exactly k ascents. For example, $\left\langle {m \atop m} \right\rangle = 0$ for $m \geq 1$; this is because a permutation on $[m]$ can have at most $m-1$ ascents. Moreover, $\left\langle {m \atop m-1} \right\rangle = 1$ because there is only one permutation on $[m]$ with exactly $m-1$ ascents: the identity permutation, which outputs values in increasing order. In addition, $\left\langle {m \atop 0} \right\rangle = 1$ because there is only one permutation on $[m]$ with zero ascents: the permutation that outputs values in decreasing order.

We need a bit more setup in order to prove our second formula for the power sum. Suppose we have a function $f : [m] \mapsto [n]$. Sort the function values so that

$$f(\pi(1)) \leq f(\pi(2)) \leq \cdots \leq f(\pi(m)),$$

where π is the resulting permutation of the domain set $[m]$ and ties are broken by requiring

$$f(\pi(i)) = f(\pi(j)) \implies \pi(i) < \pi(j).$$

For example, suppose we have the function f given by

$$f = (2, 7, 1, 8, 2, 8, 1, 8, 2)$$

(i.e., $f(1) = 2, f(2) = 7$, etc.). Sorting these by function value and using our rule for breaking ties, we have

$$f(3) \leq f(7) \leq f(1) \leq f(5) \leq f(9) \leq f(2) \leq f(4) \leq f(6) \leq f(8),$$

as well as the permutation

$$\pi = (3, 7, 1, 5, 9, 2, 4, 6, 8).$$

Given a nondecreasing function $g : [m] \mapsto [n]$, we say that g features a *duplication* at $i \in [m-1]$ if $g(i) = g(i+1)$. For example,

$$g = (1, 1, 2, 2, 2, 7, 8, 8, 8)$$

features duplications at positions 1, 3, 4, 7, and 8.

We also need the following lemma.

Lemma 1. *Given a function $f : [m] \to [n]$ and the implied permutation π of its domain set obtained by sorting the values of f, let $g = f \circ \pi$. If g features a duplication at i then π has an ascent at i.*

Proof. If g features a duplication at i, then $g(i) = g(i+1)$. This means that $f(\pi(i)) = f(\pi(i+1))$. By the rule for breaking ties in the definition of π, we have $\pi(i) < \pi(i+1)$, which means π has an ascent at i. $\square$

We're now ready to give a combinatorial proof of the following power sum formula involving the Eulerian numbers.

Theorem 3.

$$\sum_{k=1}^{n} k^m = \sum_{k=0}^{m-1} \left\langle {m \atop k} \right\rangle \binom{n+k+1}{m+1}.$$

Proof. As we know, Theorem 1 says that the left side is the number of functions $f : [m+1] \mapsto [n+1]$ such that, for all $i \in [m]$, $f(i) < f(m+1)$.

For the right side, let $k \in \{0, 1, \ldots, m-1\}$. Suppose $k+1$ is the number of ascents in π, where π is the implied permutation for a function f obtaining by sorting the values of f. Since we're counting functions $f : [m+1] \mapsto [n+1]$ such that, for all $i \in [m]$, $f(i) < f(m+1)$, the rule for constructing π requires

$$\pi(m+1) = m+1.$$

In addition, $m+1$ is the largest possible value of $[m+1]$, so π must have at least one ascent; namely, at m. Because of these restrictions, choosing π can be accomplished by choosing a permutation σ on $[m]$ with k ascents and extending it via $\pi(i) = \sigma(i)$ for $i \in [m]$ and $\pi(m+1) = m+1$. By the definition of the Eulerian numbers, the selection of σ (and thus of π) can be done in $\left\langle {m \atop k} \right\rangle$ ways.

Denote the ascents in σ by $A_1, A_2, \ldots, A_k$, where $0 \leq k \leq m-1$. Then choose $m+1$ distinct elements from the set

$$\{1, 2, \ldots, n+1, A_1, A_2, \ldots, A_k\}.$$

There are $\binom{n+k+1}{m+1}$ ways to make this selection.

Now, construct a nondecreasing function $g : [m+1] \mapsto [n+1]$ by placing the chosen elements in a list of length $m+1$ according to the following scheme:

1. For each ascent A_i chosen, place a D (for duplication) in position A_i.
2. Place the elements chosen from $[n+1]$ in order in the remaining spaces.
3. Replace each D with the next numbered element that appears after it.

For example, suppose we have the permutation

$$\sigma = (2, 4, 1, 3, 5).$$

It has ascents $A_1 = 1, A_2 = 3$, and $A_3 = 4$. (Remember that we have defined an ascent to be a position, so the ascents occur at positions 1, 3, and 4.) Suppose we choose from the set

$$\{1, \ldots, 9, A_1, \ldots, A_3\}$$

the elements $\{1, 3, 4, 5, 9, A_1\}$. Since $A_1 = 1$, the scheme above gives us the following:

After Step 1: $(D, \_, \_, \_, \_, \_)$.
After Step 2: $(D, 1, 3, 4, 5, 9)$.
After Step 3: $(1, 1, 3, 4, 5, 9)$, which is g.

(Note that this scheme follows Lemma 1 in that duplications in g can only occur at ascents in σ. In particular, there will never be a duplication at position m, as the largest possible ascent in σ occurs at position $m-1$.)

Since π is a permutation it is invertible. Let $f = g \circ \pi^{-1}$, so that $f \circ \pi = g$. Since there are $\left\langle {m \atop k} \right\rangle$ ways to select π and $\binom{n+k+1}{m+1}$ ways to select g given π, there are

$$\left\langle {m \atop k} \right\rangle \binom{n+k+1}{m+1}$$

possible ways to construct f if its implied permutation π has $k+1$ ascents. Summing over all possible values of k gives the right side. $\square$

In the example discussed in the proof of Theorem 3, requiring $f \circ \pi = g$ yields $f = (3, 1, 4, 1, 5, 9)$.

As we did with Theorem 2, let's take a look at the special case $m = 3$ of Theorem 3. We have

$$\sum_{k=1}^{n} k^3 = \binom{n+1}{4} + 4\binom{n+2}{4} + \binom{n+3}{4}.$$

In contrast to Theorem 2, Theorem 3 gives us a formula for the power sum in terms of the binomial coefficients in *column* $m+1$ of Pascal's triangle.

4 Worpitzky's identity

As a side note, a slight simplification of the proof of Theorem 3 gives us what is known as *Worpitzky's identity.*

Corollary 1 (Worpitzky's Identity)**.**

$$x^m = \sum_{k=0}^{m-1} \left\langle {m \atop k} \right\rangle \binom{x+k}{m}.$$

Proof. The left side is the number of functions f from $[m]$ to $[x]$.

For the right side, let π be a permutation of $[m]$. (Here is where the simplification occurs, as we are choosing π directly rather than choosing σ and then constructing π from σ.) Fix $k \in \{0, 1, \ldots, m-1\}$, and suppose k is the number of of ascents in π. Then the selection of π can be done in $\left\langle {m \atop k} \right\rangle$ ways.

Then, as in the proof of Theorem 3, select m elements from the set

$$\{1, \ldots, x, A_1, \ldots, A_k\},$$

and construct a non-decreasing function $g : [m] \mapsto [x]$ from those elements. Given a π with k ascents, there are $\binom{x+k}{m}$ different non-decreasing functions g that can be constructed this way.

Finally, let $f = g \circ \pi^{-1}$, so that $f \circ \pi = g$. With $\left\langle {m \atop k} \right\rangle$ ways to construct π and $\binom{x+k}{m}$ ways to construct g, there are

$$\left\langle {m \atop k} \right\rangle \binom{x+k}{m}$$

ways to construct f. Summing over all possible values of k completes the proof. □

The case $m = 3$ of Worpitzky's Identity is the following:

$$x^3 = \binom{x}{3} + 4\binom{x+1}{3} + \binom{x+2}{3}.$$

This, of course, looks quite similar to the $m = 3$ case of Theorem 3.

5 Additional reading

As we have seen, the interpretation of the power sum given in Theorem 1 leads to combinatorial proofs of two different formulas for the power sum, as well as a combinatorial proof of Worpitzky's identity. Several others have done similar or related work as well.

For more on the Stirling numbers of the second kind, see Graham et al. [6, p. 258]. Treviño [13] gives an argument similar to our proof of Theorem 2. See also Mackiw [8].

Petersen [10] is a great resource for more information on the Eulerian numbers. Stanley [12] is as well. (Both authors define $\left\langle {m \atop k} \right\rangle$ as the number of permutations on $[m]$ with k *descents*, but the values of $\left\langle {m \atop k} \right\rangle$ are the same under this definition and ours.)

Engbers and Stocker [3] contains a combinatorial proof of a generalization of Theorem 3 extended to multisets.

Foata and Schützenberger [4], Knuth [7, pp. 36–37], and Rawlings [11] each contain proofs of Worpitzky's identity similar to ours. Dzhumadil'daev [2] uses *barred permutations* to prove a multipermutation generalization of Worpitzky's identity. An even further generalization of Worpitzky is due to MacMahon (see Equation (3.3), with $q = 1$, in Gessel [5] for a more modern formulation). In addition, Petersen [9] uses barred permutations to derive a variation of Worpitzky's identity for two-sided Eulerian numbers.

References

[1] Benjamin, Arthur T. and Jennifer J. Quinn (2003). *Proofs that Really Count*. Washington, DC: Mathematical Association of America.

[2] Dzhumadil'daev, A. S. (2011). Worpitzky identity for multipermutations. *Math. Notes* 90(3): 448-450.

[3] Engbers, John, and Christopher Stocker (2016). Two combinatorial proofs of identities involving sums of powers of binomial coefficients. *Integers* 16: Article A58.

[4] Foata, Dominique, and Marcel-Paul Schützenberger (1970). Théorie géométrique des polynômes Eulériens. *Lecture Notes in Mathematics*. 138. Berlin: Springer-Verlag.

[5] Ira M. Gessel (2016). A historical survey of P-partitions, in *The Mathematical Legacy of Richard P. Stanley*. Providence, RI: American Mathematical Society, 169–188.

[6] Graham, Ronald L., Donald E. Knuth, and Oren Patashnik (1994). *Concrete Mathematics*, 2nd ed. Reading, MA: Addison-Wesley.

[7] Knuth, Donald E. (1998). *The Art of Computer Programming, Vol. III: Sorting and Searching*, 2nd ed. Reading, MA: Addison-Wesley.

[8] Mackiw, George (2000). A combinatorial approach to sums of integer powers. *Math. Mag.* 73(1): 44–46.

[9] Petersen, T. Kyle (2013). Two-sided Eulerian numbers via balls in boxes. *Math. Mag.* 86(3): 159–176.

[10] Petersen, T. Kyle (2015). *Eulerian Numbers*. New York: Birkhäuser.

[11] Rawlings, Don (1981). Generalized Worpitzky identities with applications to permutation enumeration. *Europ. J. Combinatorics*. 2: 67–78.

[12] Stanley, Richard P. (2012). *Enumerative Combinatorics, Vol. 1*, 2nd ed. Cambridge, UK: Cambridge University Press.

[13] Treviño, Enrique (2018). A short proof of a sum of powers formula. *Am. Math. Monthly*. 125(7): 659.

Mike Spivey is a professor of mathematics at the University of Puget Sound. His research interests include assignment problems and the binomial coefficients. He has written a book on proving binomial identities, as well as three mathematically-inclined works of interactive fiction.