
AN EXPLORATION OF FAULHABER'S FORMULA AND THE BERNOULLI NUMBERS

MDM 4U7

JUNE 14, 2017

By

ANDREW WANG

Contents

1	Introduction and Initial Exposure	3
1.1	Gauss	3
1.2	Sum of Squares	4
1.3	Sum of Cubes and then some	6
2	Generalizations	8
2.1	The Bernoulli Numbers	9
2.2	Faulhaber's Formula	10

1 Introduction and Initial Exposure

1.1 Gauss

The story begins with a boy whose teacher was discontent with the class. As a result, the teacher punished his students by having them perform rote calculation and compelled them to calculate the sum of the first 100 positive integers. The activity was meant to take up the good part of an hour, but the boy only thought for a moment, after which he scribbled down his answer on the slate. He tossed the slate on his teacher's desk and rested as he waited for the others. The teacher said nothing, but was evidently skeptical, and so he decided that he would humiliate the boy later. Slowly, the students finished one by one, and after everyone had completed their work, the teacher turned the pile over. He was to look at the boy's first. Lo and behold, the boy's answer was 5050. The teacher's face flushed red in disbelief. Supposedly, that boy's name was Carl Friedrich Gauss.

The technique he used was relatively straight forward: he paired the first and last number, the second and second last, and so on, then summed the pairs. In fact, for a general arithmetic sequence with n terms, starting term a , and common difference d , the technique still works. Namely, for

$$\begin{aligned} S &= a + (a + d) + (a + 2d) + \dots + (a + (n - 2)d) + (a + (n - 1)d) \\ &= (a + (n - 1)d) + (a + (n - 2)d) + \dots + (a + 2d) + (a + d) + a \end{aligned}$$

summing corresponding terms pairwise, vertically, yields n terms, all equal to $2a + (n - 1)d$. Thus, after performing the addition this way,

$$\begin{aligned} 2S &= n(2a + (n - 1)d) \\ \Rightarrow S &= \frac{n}{2}(2a + (n - 1)d) \end{aligned}$$

for Gauss' instance of the problem, $a = 1$, $d = 1$, and

$$S = \frac{n}{2}(2 + n - 1) = \frac{n}{2}(n + 1) = \frac{1}{2}n^2 + \frac{1}{2}n. \tag{1}$$

Indeed, substituting $n = 100$ yields 5050.

My initial reaction to the story was one of awe. Gauss' clever exploitation of the commutative property of addition here was both ingenious and elegant, and this story is really

what made me realize the importance of computational efficiency. As compared to the boy's classmates who had to sum the, say, n terms, the simple formula would produce an answer in pretty well the same amount of time with $n = 100$ as $n = 10000$.

But yet, I still was not satisfied. Why was the problem limited to the sum of the first n numbers? What about their reciprocals? Their squares? Their cubes? Their k^{th} powers? And so, I sought after a generalization.

1.2 Sum of Squares

At first, I looked into the sum of squares. I sought after a function $f(n)$, defined on the natural numbers, which would return the sum $1^2 + 2^2 + \dots + n^2$ in some closed form. I took the most direct approach, which involved first constructing a table of values:

n	$\sum_{i=1}^n i^2$	1 st difference	2 nd difference	3 rd difference
1	1			
2	5	4		
3	14	9	5	
4	30	16	7	2
5	55	25	9	2
6	91	36	11	2

Table 1. the first six sum of the squares of naturals and their differences.

The first observation I made was that it took three differences to reach a constant rate of change. A polynomial of degree three would also exhibit this property since differentiating three times would leave a constant. In fact, integrating three times the constant, 2, here yields $\frac{1}{3}x^3 + C$ where C is the constant of integration. The issue here is that all terms of degree less than 3 in the original polynomial would have contributed 0 to the third derivative of $f(n)$, and so those terms would be unknown when reversing the process.

At this point I did not see another way to approach the problem, so I took to Microsoft Excel's graphing software with the data in table 1. Sure enough, as can be seen in figure 1, after setting the equation to a polynomial of degree three, Excel produced the result $y = \frac{1}{3}x^3 + \frac{1}{2}x^2 + \frac{1}{6}x$. Of course, there was a constant here, but for the purpose of simplicity, I will say that the $3E - 12$ term is negligible.

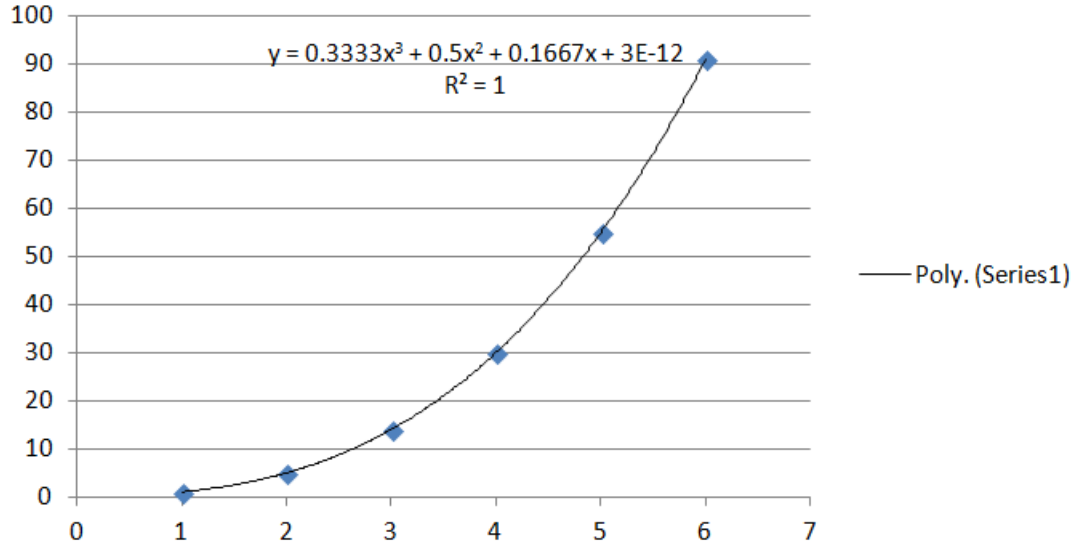


Figure 1. $y = f(n)$.

But of course, this method feels like cheating. After looking through a bit of mathematical literature, sure enough, there was a proof of the above result. Notice its use of clever telescoping technique that reduces so that the desired sum is expressed in terms of other known sum(s) and, of course, n .

Since the desired sum is a cubic polynomial, start with the expansion,

$$(n+1)^3 = n^3 + 3n^2 + 3n + 1$$

which, when rearranged to facilitate heavy cancellation, becomes,

$$(n+1)^3 - n^3 = 3n^2 + 3n + 1. \quad (2)$$

Rewriting the above identity with $n-m$, $m \in \mathbb{N}$ in place of n holds true as well. Namely, with $m \in \{1, 2, 3, \dots, n-2\}$,

$$\begin{aligned} n^3 - (n-1)^3 &= 3(n-1)^2 + 3(n-1) + 1 \\ (n-1)^3 - (n-2)^3 &= 3(n-2)^2 + 3(n-2) + 1 \\ &\vdots \\ 3^3 - 2^3 &= 3(2)^2 + 3(2) + 1 \\ 2^3 - 1^3 &= 3(1)^2 + 3(1) + 1 \end{aligned}$$

Adding all n expressions on the left side and on the right, including (2), yields

$$(n+1)^3 - 1 = 3 \sum_{i=1}^n i^2 + 3 \sum_{i=1}^n i + \sum_{i=1}^n 1$$

Substituting (1),

$$\begin{aligned} 3 \sum_{i=1}^n i^2 &= n^3 + 3n^2 + 3n - 3 \left(\frac{(n)(n+1)}{2} \right) + n \\ 6 \sum_{i=1}^n i^2 &= 2n^3 + 3n^2 + n \\ &= n(n+1)(2n+1) \end{aligned}$$

$$\therefore \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6} = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n, \text{ as desired.} \quad (3)$$

1.3 Sum of Cubes and then some

I proceeded to work on the sum of cubes. This is one of the more interesting results. It turns out that

$$\sum_{i=1}^n i^3 = \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}n^2 = \frac{n^2(n+1)^2}{4} = \left(\sum_{i=1}^n i \right)^2. \quad (4)$$

Incidentally, this result is often attributed to the Pythagorean, Nicomachus of Gerasa, who lived as early as c. 100 CE (Midonick, 1965, p. 15-6).

After (4), it is possible to make two observations and possibly even a conjecture. The first is that for every sum of the k^{th} powers of the natural numbers, the formula is a polynomial of degree $k+1$. Next, the lack of a constant term in any of these formulae makes sense. When the upper bound, n , is 0, the sum of the k^{th} powers of the first 0 numbers had better be 0. But here, in (4), there is not only no constant term but no linear term. Thus, it was proving to be difficult to spot an insightful pattern in the coefficient in these polynomials.

Perhaps, with more cases, key insights might reveal themselves. From this point on, for the sake of convenience, define $S_k(n) = \sum_{i=1}^n i^k$. Computing polynomials $S_k(n)$, $k \in \{0, 1, 2, 3, 4, 5, 6\}$ helps develop the following table of coefficients (Ezra, 2016):

k	n	n^2	n^3	n^4	n^5	n^6	n^7
0	1						
1	$\frac{1}{2}$	$\frac{1}{2}$					
2	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$				
3	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$			
4	$-\frac{1}{30}$	0	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{5}$		
5	0	$-\frac{1}{12}$	0	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{6}$	
6	$\frac{1}{42}$	0	$-\frac{1}{6}$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{7}$

Table 2. Coefficients of $S_k(n)$, $k \in \{0, 1, 2, 3, 4, 5, 6\}$.

Now, there are four observations that can be made. First, the leading coefficient (in red) in each instance of S_k is simply the reciprocal of $k + 1$. Second, for $k \geq 1$, the coefficient of the second highest degree term (in blue) is consistently $\frac{1}{2}$. Third, It appears that for odd $k \geq 3$, there is no linear term. All the orange 0's form a diagonal, too. Finally, the coefficients in the green diagonal are all negative.

At this point in my exploration, I am finding it increasingly tedious to compute the polynomials with the method from before. Furthermore, the method requires that all sum formulas below the desired power be known. Thus, the previous issue of computational inefficiency Gauss so cleverly avoided has once again presented itself in this extension of the problem. Finally, then, we inquire whether there is a closed form of $S_m(n)$, $m \in \mathbb{N}$.

But first, we observe the telescoping method's generalization, so to speak, for $S_m(n)$:

Begin with the expansion of $(n + 1)^{m+1}$,

$$\begin{aligned}
(n + 1)^{m+1} &= n^{m+1} + (m + 1)n^m + \binom{m + 1}{2}n^{m-1} + \binom{m + 1}{3}n^{m-2} + \dots + 1 \\
(n + 1)^{m+1} - n^{m+1} &= (m + 1)n^m + \binom{m + 1}{2}n^{m-1} + \binom{m + 1}{3}n^{m-2} + \dots + 1
\end{aligned}$$

replacing n with $n - 1$, $n - 2$, \dots , 2 , 1 will also satisfy the identity:

$$\begin{aligned}
n^{m+1} - (n-1)^{m+1} &= (m+1)(n-1)^m + \binom{m+1}{2}(n-1)^{m-1} + \\
&\quad \binom{m+1}{3}(n-1)^{m-2} + \dots + 1 \\
(n-1)^{m+1} - (n-2)^{m+1} &= (m+1)(n-2)^m + \binom{m+1}{2}(n-2)^{m-1} + \\
&\quad \binom{m+1}{3}(n-2)^{m-2} + \dots + 1 \\
&\vdots \\
3^{m+1} - 2^{m+1} &= 2^m(m+1) + \binom{m+1}{2}2^{m-1} + \binom{m+1}{3}2^{m-2} + \dots + 1 \\
2^{m+1} - 1^{m+1} &= (m+1)1^m + \binom{m+1}{2}1^{m-1} + \binom{m+1}{3}1^{m-2} + \dots + 1
\end{aligned}$$

Again, adding all n expressions on the left sides and those on the right sides causes heavy cancellation:

$$\begin{aligned}
(n+1)^{m+1} - 1 &= (m+1)S_m(n) + \binom{m+1}{2}S_{m-1}(n) + \binom{m+1}{3}S_{m-2}(n) \dots + S_0(n) \\
&= (m+1)S_m(n) + \sum_{i=1}^m \binom{m+1}{i+1}S_{m-i}(n) \\
&= \sum_{j=1}^{m+1} \binom{m+1}{j}n^{m+1-j} - 1 \\
S_m(n) &= \left(\frac{1}{m+1}\right) \left(\sum_{j=1}^{m+1} \binom{m+1}{j}n^{m+1-j} - \sum_{i=1}^m \binom{m+1}{i+1}S_{m-i}(n) - 1\right). \tag{5}
\end{aligned}$$

A rather monstrous result, really.

2 Generalizations

Johann Faulhaber of Ulm (1580-1635) devoted a substantial part of his life to this very problem—he notably calculated the sum formulas of odd powers up to 17 (Knuth, 1993, p.277-8). In particular, he defined $N = \frac{n^2+n}{2}$ and expressed all the sum formulas as polynomials in N .

It was after Johann Faulhaber that the Faulhaber formula was named, though the result is more accurately ascribed to Jacob Bernoulli and, simultaneously, Takakazu Seki (Beery, 2010). Indeed, the closed form expression given by Bernoulli features the Bernoulli numbers: a sequence of which Faulhaber likely had no knowledge seeing as Bernoulli discovered it after Faulhaber had past away.

Therefore, since the closed form formula involves the Bernoulli numbers, there ought to be a brief discussion on the sequence.

2.1 The Bernoulli Numbers

The Bernoulli numbers, B_n are a sequence of rational numbers whose appearances in the world of number theory are plentiful, for instance, the Basel problem (Sullivan, 2013).

By definition,

$$\sum_{i=0}^m \binom{m+1}{i} B_i = 0$$

and $B_0 = 1$. m is chosen as needed. Indeed, B_1 can be computed by letting $m = 1$:

$$\binom{2}{0} B_0 + \binom{2}{1} B_1 = 0 = 1 + 2B_1 \Rightarrow B_1 = -\frac{1}{2}.$$

By repeating this process, we can produce the first 7 Bernoulli numbers:

n	B_n
0	1
1	$-\frac{1}{2}$
2	$\frac{1}{6}$
3	0
4	$-\frac{1}{30}$
5	0
6	$\frac{1}{42}$

Table 3. the first 7 Bernoulli numbers.

These numbers look familiar. In fact, they are the coefficients of the linear terms in table 2! The one difference is that B_1 is negative. This is a discrepancy that is addressed by the "first" and "second" Bernoulli numbers. The difference between the two is only B_n , where

the first Bernoulli numbers, B_n^- , are such that $B_1 = -\frac{1}{2}$. The second Bernoulli numbers have $B_1 = +\frac{1}{2}$. Both work, but by convention the first Bernoulli numbers are used. The difference in the end is simply that the formula would contain $(-1)^j$ so that all odd Bernoulli numbers have their signs switched. However, the odd Bernoulli numbers, B_k $k \geq 3$, are all zero, and so that the -1 effectively only changes B_1 . I will examine the version with the second Bernoulli numbers.

2.2 Faulhaber's Formula

Faulhaber's formula states that:

$$S_m(n) = \frac{1}{m+1} \sum_{i=1}^m \binom{m+1}{i} B_i n^{m+1-i}$$

.

The derivation I considered is from Aigner's "A Course in Enumeration" (Springer, 2007). It is very involved, and so only the overview will be provided.

Lemma 1 Let a_n and b_n be two sequences not necessarily unique. Define $\hat{A}(z) = \sum_{n \geq 0} a_n \frac{z^n}{n!}$ and $\hat{B}(z) = \sum_{n \geq 0} b_n \frac{z^n}{n!}$. $\exists \hat{D}(z) = \sum_{n \geq 0} d_n \frac{z^n}{n!} = (\hat{A}(z))(\hat{B}(z))$, where $d_n = \sum_{i=0}^n \binom{n}{i} a_i b_{n-i}$.

This result makes more sense when the sums are written out in full:

$$\left(\frac{a_0}{0!}z^0 + \frac{a_1}{1!}z^1 + \frac{a_2}{2!}z^2 + \dots\right) \left(\frac{b_0}{0!}z^0 + \frac{b_1}{1!}z^1 + \frac{b_2}{2!}z^2 + \dots\right) = \frac{a_0b_0}{0!0!}z^0 + \left(\frac{a_0b_1}{0!1!} + \frac{a_1b_0}{1!0!}\right)z^1 + \left(\frac{a_0b_2}{0!2!} + \dots + \frac{a_2b_0}{2!0!}\right)z^2 + \dots$$

By matching each coefficient, c_i , by powers of z , we see that

$$c_i = \sum_{i=0}^k \frac{a_i b_{k-i}}{i!(k-i)!}$$

.

After defining $d_i = (k!)c_i$ and simplifying, lemma 1 is proved.

The derivation then looks at $S_m(n)$, though it is defined slightly differently from what I had, previously. Here, we redefine $S_m(n)$ with upper bound $n-1$ instead of n .

After defining $\hat{S}(z) = \sum_{m \geq 0} S_m(n) \frac{z^m}{m!}$ and manipulating this generating function by observing the presence of the Maclaurin series of e^x (Weisstein, result 14), we arrive at:

$$\hat{S}(z) = \frac{e^{nz} - 1}{e^z - 1}.$$

Next, we observe that the numerator is very close to the exponential generating function for the nz^{th} powers of e , differing only by the -1 . Thus, we define

$$\begin{aligned}\hat{P}_n(z) &= \sum_{m \geq 0} n^m \frac{z^m}{m!} \\ &= e^{nz}.\end{aligned}$$

A part of the proof that is omitted for now is the result: that the generating function of the Bernoulli numbers is

$$\hat{B}(z) = \sum_{n \geq 0} B_n \frac{z^n}{n!} = \frac{z}{e^z - 1}.$$

This result is shown with lemma 1 and a manipulation of the definition of the Bernoulli numbers provided in **2.1**.

Nonetheless, we arrive at

$$(\hat{P}_n(z) - 1)(\hat{B}(z)) = z\hat{S}(z).$$

Now that there are two unique expressions of the same entity, we may compare them. Namely, by matching the coefficients of the z^{m+1} terms and isolating S_m , the formula is proven.

References

- [1] Beery, Janet. (University of Redlands), "Sums of Powers of Positive Integers," Convergence (July 2010), DOI:10.4169/loci003284
- [2] Getzler, Ezra. (2016). The Bernoulli Numbers [PowerPoint slides]. Retrived from <https://www.youtube.com/watch?v=UALayaK5mMU>
- [3] Knuth, D. (1993). Johann Faulhaber and Sums of Powers. Mathematics of Computation, 61(203), 277-294. doi:10.2307/2152953
- [4] Midonick, Henrietta (1965), The Treasury of Mathematics, Volume 2, pages 15-16. Penguin Books.
- [5] Kolbert, E. (2011, August 15). Sleeping with the Enemy. Retrieved May 03, 2016, from <http://www.newyorker.com/magazine/2011/08/15/sleeping-with-the-enemy>
- [6] Weisstein, Eric W. "Maclaurin Series." From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/MaclaurinSeries.html>