

Brief Introduction to Bernoulli Numbers

Heewon Lee

December 22, 2025

1 Introduction: Sum of Powers

Our story begins with the following sum:

$$S_k(n) := 1^k + 2^k + \cdots + n^k = \sum_{i=0}^n i^k.$$

One, of course, could motivate with Gauss' story of adding 1 through 100. Let us take it as an axiom that the reader is interested in these sums. Sometime during high school, we learn the following formulae:

$$\begin{aligned} S_1(n) &= 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \\ S_2(n) &= 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \\ S_3(n) &= 1^3 + 2^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2 \end{aligned}$$

How do we go forward? If you're lucky, you may have learned an inductive algorithm for the successive formulae using the binomial formula. The following illustrates this for $S_4(n)$:

$$\begin{aligned} 1^5 &= 5 \cdot 1^4 - 10 \cdot 1^3 + 10 \cdot 1^2 - 5 \cdot 1^1 + 1 \cdot 1^0 \\ 2^5 - 1^5 &= 5 \cdot 2^4 - 10 \cdot 2^3 + 10 \cdot 2^2 - 5 \cdot 2^1 + 1 \cdot 2^0 \\ 3^5 - 2^5 &= 5 \cdot 3^4 - 10 \cdot 3^3 + 10 \cdot 3^2 - 5 \cdot 3^1 + 1 \cdot 3^0 \\ &\vdots \\ n^5 - (n-1)^5 &= 5 \cdot n^4 - 10 \cdot n^3 + 10 \cdot n^2 - 5 \cdot n^1 + 1 \cdot n^0 \end{aligned}$$

$$n^5 = 5S_4(n) - 10S_3(n) + 10S_2(n) - 5S_1(n) + S_0(n)$$

Then, we can solve for $S_4(n)$ (and recognizing $S_0(n) = n$):

$$\begin{aligned} S_4(n) &= \frac{n^5}{5} + 2S_3(n) - 2S_2(n) + S_1(n) - \frac{1}{5}S_0(n) \\ &= \frac{n^5}{5} + \frac{n^2(n+1)^2}{2} - \frac{n(n+1)(2n+1)}{3} + \frac{n(n+1)}{2} - \frac{n}{5} \\ &= \dots \\ &= \frac{n^5}{5} + \frac{n^4}{2} + \frac{n^3}{3} - \frac{n}{30}. \end{aligned}$$

This, of course, can be generalized to obtain $S_k(n)$ for arbitrary k .

Theorem 1.1. *For any $k \geq 1$, $S_k(n)$ can be expressed as a linear combination of $S_0(n), S_1(n), \dots, S_{k-1}(n)$ and n^{k+1} .*

Proof. Recall the binomial formula:

$$(x-1)^{k+1} = \sum_{j=0}^{k+1} \binom{k+1}{j} (-1)^{k+1-j} x^j$$

$$\Rightarrow x^{k+1} - (x-1)^{k+1} = \sum_{j=0}^k \binom{k+1}{j} (-1)^{k_j} x^j$$

By taking the sum of both sides from $x = 1$ to $x = n$, the left side telescopes to n^{k+1} and the right side can be summed termwise:

$$n^{k+1} = \sum_{j=0}^k \binom{k+1}{j} (-1)^{k-j} S_j(n)$$

$$\Rightarrow S_k(n) = \frac{n^{k+1}}{k+1} + \frac{1}{k+1} \sum_{j=0}^{k-1} \binom{k+1}{j} (-1)^{k+1-j} S_j(n).$$

Hence, $S_k(n)$ is a linear combination of $S_0(n), S_1(n), \dots, S_{k-1}(n)$ and n^{k+1} . \square

Corollary 1.1.1. $S_k(n)$ is a polynomial on n of degree $k+1$.

Proof. This is justified inductively. $S_0(n) = n$ is a degree-1 polynomial on n . If each $S_l(n)$ is a degree- $(l+1)$ polynomial on n for every $l < k$, then $S_k(n)$ is a linear combination of n^{k+1} and polynomials of degrees less than $k+1$, where the coefficient of n^{k+1} is nonzero. Therefore, $S_k(n)$ is a degree- $(k+1)$ polynomial. \square

Using this algorithm, however, is cumbersome: algebraically solving for $S_k(n)$ is tedious and error-prone. We need a better method, and fortunately, Corollary 1.1.1 provides one.

2 Going Linear

Instead of solving for the function $S_k(n)$ itself, let us shift the question to calculating their coefficients. Let

$$S_k(n) = a_{k,k} n^{k+1} + a_{k,k-1} n^k + \dots + a_{k,1} n^2 + a_{k,0} n.$$

(Notice the absence of the constant term. This can be justified either inductively or by the fact that $S_k(0) = 0$.) We can reformulate this using the following matrix equation:

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} a_{00} & 0 & 0 & 0 & \dots \\ a_{10} & a_{11} & 0 & 0 & \dots \\ a_{20} & a_{21} & a_{22} & 0 & \dots \\ a_{30} & a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} n^1 \\ n^2 \\ n^3 \\ n^4 \\ \vdots \end{pmatrix}$$

The inductive formula can also be formulated with matrices:

$$\begin{pmatrix} n^1 \\ n^2 \\ n^3 \\ n^4 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ -1 & 2 & 0 & 0 & \dots \\ 1 & -3 & 3 & 0 & \dots \\ -1 & 4 & -6 & 4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ \vdots \end{pmatrix}$$

If we chop off at a certain number of rows, then the resulting two coefficient matrices must be inverses of each other. Thus, we are left with the task of computing

$$\begin{pmatrix} a_{00} & 0 & 0 & 0 & \dots \\ a_{10} & a_{11} & 0 & 0 & \dots \\ a_{20} & a_{21} & a_{22} & 0 & \dots \\ a_{30} & a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ -1 & 2 & 0 & 0 & \dots \\ 1 & -3 & 3 & 0 & \dots \\ -1 & 4 & -6 & 4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}^{-1}$$

or

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ -1 & 2 & 0 & 0 & \cdots \\ 1 & -3 & 3 & 0 & \cdots \\ -1 & 4 & -6 & 4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_{00} & 0 & 0 & 0 & \cdots \\ a_{10} & a_{11} & 0 & 0 & \cdots \\ a_{20} & a_{21} & a_{22} & 0 & \cdots \\ a_{30} & a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Directly solving for all a_{ij} seems hard, so let's try this diagonal by diagonal.

0. Main diagonal:

$$1 = \left((-1)^k \binom{k+1}{0} \quad (-1)^{k-1} \binom{k+1}{1} \quad \cdots \quad (-1)^0 \binom{k+1}{k} \quad 0 \quad \cdots \right) \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_{k,k} \\ a_{k+1,k} \\ \vdots \end{pmatrix} = (k+1)a_{k,k}$$

$$\Rightarrow a_{k,k} = \frac{1}{k+1}$$

1. Subdiagonal:

$$0 = -\binom{k+1}{k-1}a_{k-1,k-1} + \binom{k+1}{k}a_{k,k-1} = -\frac{k(k+1)}{2}\frac{1}{k} + (k+1)a_{k,k-1}$$

$$\Rightarrow a_{k,k-1} = \frac{1}{2}$$

2. Subsubdiagonal:

$$0 = \binom{k+1}{k-2}a_{k-2,k-2} - \binom{k+1}{k-1}a_{k-1,k-2} + \binom{k+1}{k}a_{k,k-2} = \frac{(k+1)k(k-1)}{6}\frac{1}{k-1} - \frac{k(k+1)}{2}\frac{1}{2} + (k+1)a_{k,k-2}$$

$$\Rightarrow a_{k,k-2} = \frac{k}{12}$$

We proceed in this way to obtain:

3.

$$a_{k,k-3} = 0$$

4.

$$a_{k,k-4} = -\frac{k(k-1)(k-2)}{720}$$

5.

$$a_{k,k-5} = 0$$

6.

$$a_{k,k-6} = \frac{k(k-1)(k-2)(k-3)(k-4)}{30240}$$

...

The numerators, in particular, make us form the following ansatz:

Theorem 2.1. *We may write*

$$a_{k,k-l} = \frac{B_l}{l!} k(k-1)\cdots(k-l+2) \quad (0 \leq l \leq k),$$

where B_l are some coefficients and we define the product as $\frac{1}{k+1}$ for $l=0$ and 1 for $l=1$.

Proof. This statement has been proven by brute force for $l = 0, 1, \dots, 6$. We proceed inductively, using the l th subdiagonal in our matrix equation:

$$\begin{aligned}
0 &= (-1)^l \binom{k+1}{k-l} a_{k-l,k-l} + (-1)^{l-1} \binom{k+1}{k-l+1} a_{k-l+1,k-l} + \cdots + (-1)^0 \binom{k+1}{k} a_{k,k-l} \\
\Rightarrow a_{k,k-l} &= \frac{k}{2!} a_{k-1,k-l} - \frac{k(k-1)}{3!} a_{k-2,k-l} + \cdots + (-1)^{l+1} \frac{k(k-1) \cdots (k-l+1)}{(l+1)!} a_{k-l,k-l} \\
&= \sum_{j=1}^l (-1)^{j+1} \frac{\prod_{i=0}^{j-1} (k-i)}{(j+1)!} \cdot a_{k-j,k-l} \\
&= \sum_{j=1}^l (-1)^{j+1} \frac{\prod_{i=0}^{j-1} (k-i)}{(j+1)!} \cdot \frac{B_{l-j}}{(l-j)!} \prod_{i=0}^{l-j-2} (k-j-i) \\
&= \left(\frac{1}{l+1} \sum_{j=1}^l (-1)^{j+1} \binom{l+1}{j+1} B_{l-j} \right) \cdot \frac{k(k-1) \cdots (k-l+2)}{l!}
\end{aligned}$$

Therefore, the formula holds for all l . \square

Corollary 2.1.1. *For any l ,*

$$\sum_{j=0}^l (-1)^j \binom{l+1}{j+1} B_{l-j} = \delta_{l0},$$

where δ_{l0} is the Kronecker delta, equal to 1 if $l = 0$ and 0 otherwise.

Proof. For $l > 0$, the above formula

$$B_l = \frac{1}{l+1} \sum_{j=1}^l (-1)^{j+1} \binom{l+1}{j+1} B_{l-j}$$

is equivalent to this theorem. For $l = 0$, the left-hand side is B_0 , which is equal to 0. \square

We now have an algorithm for calculating arbitrary sums of powers!

1. Calculate sufficiently many B_n 's, using Corollary 2.1.1.
2. $a_{k,k-j} = \frac{B_j}{j!} k(k-1) \cdots (k-j+2)$
3. $S_k(n) = a_{k,k} n^{k+1} + a_{k,k-1} n^k + \cdots + a_{k,0} n^1$

As an example, let us calculate $S_{10}(10)$.

$$\{B_n\} = 1, \frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, -\frac{1}{30}, 0, \frac{5}{66}, \dots$$

(Why so many zeros? Stay tuned...)

$$\begin{aligned}
\Rightarrow a_{10,10} &= \frac{1}{11}, \quad a_{10,9} = \frac{1}{2}, \quad a_{10,8} = \frac{5}{6}, \quad a_{10,7} = 0, \quad a_{10,6} = -1, \quad a_{10,5} = 0, \quad a_{10,4} = 1, \\
a_{10,3} &= 0, \quad a_{10,2} = -\frac{1}{2}, \quad a_{10,1} = 0, \quad a_{10,0} = \frac{5}{66}
\end{aligned}$$

$$\begin{aligned}
1^{10} + 2^{10} + \cdots + 10^{10} &= \frac{10^{11}}{11} + \frac{10^{10}}{2} + \frac{5}{6} 10^9 - 10^7 + 10^5 - \frac{10^3}{2} + \frac{5}{66} 10^1 \\
&= 10 \left(\frac{5}{66} + 100 \left(-\frac{1}{2} + 100 \left(1 + 100 \left(\frac{5}{6} + 10 \left(\frac{1}{2} + \frac{10}{11} \right) \right) \right) \right) \right) \\
&= 14914341925
\end{aligned}$$

This was barely less effort than directly summing the expression!

3 Bernoulli Numbers

The numbers B_n we have defined are the infamous *Bernoulli numbers*, and the power sum formulae we found are called *Faulhaber's formula*. Explicitly, this is:

$$S_k(n) = 1^k + 2^k + \cdots + n^k = \frac{1}{k+1} \sum_{j=1}^{k+1} \binom{k+1}{j} B_{k+1-j} n^j.$$

When analyzing any sequence, one overpowered tool is its generating function. Consider the following:

$$f(x) := \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n.$$

(Let us postpone the question of convergence for later, which is standard procedure.) Let us use the inductive formula from Corollary 2.1.1 again:

$$\delta_{l0} = \sum_{j=0}^l (-1)^j \binom{l+1}{j+1} B_{l-j} = (l+1)! \sum_{j=0}^l \frac{(-1)^j}{(j+1)!} \frac{B_{l-j}}{(l-j)!}.$$

The last expression could very well arise from a product of $f(x)$ and

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} x^n &= -\frac{1}{x} \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(n+1)!} x^{n+1} = \frac{1-e^{-x}}{x}. \\ \Rightarrow \frac{1-e^{-x}}{x} f(x) &= \sum_{n=0}^{\infty} x^n \sum_{j=0}^n \frac{(-1)^j}{(j+1)!} \frac{B_{n-j}}{(n-j)!} = \sum_{n=0}^{\infty} x^n \delta_{n0} = 1 \\ \therefore f(x) &= \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n = \frac{x}{1-e^{-x}} \end{aligned}$$

We thus obtain the closed-form expression

$$B_n = \left. \frac{d^n}{dx^n} \left(\frac{x}{1-e^{-x}} \right) \right|_{x=0}$$

or using Cauchy's integral formula,

$$B_n = \frac{n!}{2\pi i} \oint_{|z|=1} \frac{dz}{z^n (1-e^{-z})}$$

with the contour oriented counterclockwise around the origin. Returning to the question of convergence, notice that the poles of the complex function $f(z)$ are precisely $z = \pm in\pi$ for positive integers n . The closest poles to the origin are $\pm i\pi$; therefore, the series converges when $|z| < \pi$. The convergence for when $|z| = \pi$ is left as an exercise.

Theorem 3.1. *For $n > 1$ odd, $B_n = 0$.*

Proof. Notice that

$$\begin{aligned} f(x) - f(-x) &= \sum_{n=0}^{\infty} \frac{B_n}{n!} (x^n - (-x)^n) = \sum_{n \text{ odd}} \frac{2B_n}{n!} x^n \\ &= \frac{x}{1-e^{-x}} - \frac{-x}{1-e^x} = \frac{x(1-e^{-x})}{1-e^{-x}} = x. \end{aligned}$$

Therefore, $B_n = 0$ for all odd $n > 1$. □

4 Euler-Maclaurin Formula

We now turn our attention to generalizing this summation to general functions, as in, consider the following sum:

$$S(a, b, h; f) := \sum_{j=1}^n f(a + jh) = f(a + h) + f(a + 2h) + f(a + 3h) + \dots,$$

where $n = \frac{b-a}{h}$ and f is a sufficiently good function. This whole thing is too complicated, though, so let's first look at the first term more closely.

$$\begin{aligned} f(a + h) &= \sum_{j=0}^{\infty} \frac{f^{(j)}(a)}{j!} h^j \\ &= \sum_{j=0}^{\infty} \frac{(hD)^j}{j!} f(a) \\ &= e^{hD} f(a), \end{aligned}$$

where D denotes the differentiation operator. Thus, shifting the argument by h is equivalent to the operator e^{hD} ! We can use this to simplify our sum:

$$\begin{aligned} S(a, b, h; f) &= \sum_{j=1}^n f(a + jh) \\ &= \sum_{j=1}^n \sum_{k=0}^{\infty} \frac{(jhD)^k}{k!} f(a) \\ &= \sum_{k=0}^{\infty} \frac{(hD)^k}{k!} \sum_{j=1}^n j^k f(a) \\ &= \sum_{k=0}^{\infty} \frac{(hD)^k}{k!} \cdot \frac{1}{k+1} \sum_{j=1}^{k+1} \binom{k+1}{j} B_{k+1-j} n^j f(a) \\ &= \sum_{k=1}^{\infty} (hD)^k \left(\sum_{j=0}^{k+1} \frac{B_{k+1-j}}{(k+1-j)!} \cdot \frac{n^j}{j!} - \frac{B_{k+1}}{(k+1)!} \right) f(a) \\ &= (hD)^{-1} \left(\left(\sum_{k=0}^{\infty} (hD)^k \frac{B_k}{k!} \right) \cdot \left(\sum_{k=0}^{\infty} (hD)^k \frac{n^k}{k!} \right) - \sum_{k=0}^{\infty} (hD)^k \frac{B_k}{k!} \right) f(a) \\ &= (hD)^{-1} (e^{nhD} - 1) \cdot \sum_{k=0}^{\infty} \frac{B_k h^k}{k!} D^k f(a) \\ &= (e^{nhD} - 1) \left((hD)^{-1} + \frac{1}{2} + \sum_{k=1}^{\infty} \frac{B_{2k} h^{2k-1}}{(2k)!} D^{2k-1} \right) f(a) \\ &= \frac{1}{h} \int_a^b dx f(x) + \frac{f(b) - f(a)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k} h^{2k-1}}{(2k)!} (f^{(2k-1)}(b) - f^{(2k-1)}(a)) \end{aligned}$$

This final expression is the famed *Euler-Maclaurin formula*! Or, adding $f(a)$ to both sides, we get the following two expressions:

$$\begin{aligned} f(a + h) + f(a + 2h) + \dots + f(b) &= \frac{1}{h} \int_a^b dx f(x) + \frac{f(b) - f(a)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k} h^{2k-1}}{(2k)!} (f^{(2k-1)}(b) - f^{(2k-1)}(a)), \\ f(a) + f(a + h) + \dots + f(b) &= \frac{1}{h} \int_a^b dx f(x) + \frac{f(b) + f(a)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k} h^{2k-1}}{(2k)!} (f^{(2k-1)}(b) - f^{(2k-1)}(a)). \end{aligned}$$

Thus revealed is the importance of the Bernoulli numbers: they allow the error correction between discrete sums and continuous integrals!

Example: Faulhaber's formulae (again). Let us perform a sanity check for this formula using $f(x) = x^k$.

$$\begin{aligned}
S_k(n) &= (0+0)^k + (0+1)^k + \cdots + (0+n)^k \\
&= f(0) + S(0, n, 1; f) \\
&= \int_0^n dx x^k + \frac{n^k + 0^k}{2} + \sum_{j=1}^{\infty} \frac{B_{2j}}{(2j)!} \left(f^{(2j-1)}(n) - f^{2j-1}(0) \right) \\
&= \frac{n^{k+1}}{k+1} + \frac{n^k}{2} + \sum_{j=1}^{\lfloor \frac{k+1}{2} \rfloor} \frac{B_{2j}}{(2j)!} \cdot \frac{k!}{(k-2j+1)!} n^{k-2j+1} \\
&= \frac{n^{k+1}}{k+1} + \frac{n^k}{2} + \frac{1}{k+1} \sum_{j=1}^{\lfloor \frac{k+1}{2} \rfloor} \binom{k+1}{2j} B_{2j} n^{k-2j+1}
\end{aligned}$$

We thus obtain the same expression as before.

Example: Riemann zeta function. Let

$$\zeta(s) := 1^{-s} + 2^{-2} + \cdots.$$

The importance of this function is beyond the scope of this introduction; you could win a million buck by studying this thing. For complex-valued inputs s , this series absolutely converges for $\Re\{s\} > 1$ (see p -test). Let us attempt to evaluate this sum with Euler-Maclaurin formula with $f(x) = x^{-s}$.

$$\begin{aligned}
\zeta(s) &= f(1) + S(1, \infty, 1; f) \\
&= \int_1^\infty \frac{dx}{x^s} + \frac{1}{2} + \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!} \left(-f^{(2n-1)}(1) \right) \\
&= \frac{1}{s-1} + \frac{1}{2} + \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!} \frac{\Gamma(s+2n-1)}{\Gamma(s)}
\end{aligned}$$

Here, the Γ fraction is a shorthand for $s(s+1)\cdots(s+2n-2)$.

Take, for example, $s = -1$:

$$\zeta(-1) = -\frac{1}{2} + \frac{1}{2} - \frac{B_2}{2} = -\frac{1}{12}.$$

Of course, the series does not converge for $s = -1$, the derivation above does not work as is, and the naïve interpretation yields the nonsensical

$$1 + 2 + 3 + \cdots \stackrel{?}{=} -\frac{1}{12}.$$

But we are doing physics, so it must *mean* something, right? Right? ... Let's not dwell on this for so long.¹

¹For the interested, the relevant search term is Ramanujan summation.