Bernoulli numbers

Bill Casselman
University of British Columbia
cass@math.ubc.ca

The sequence of Bernoulli numbers is among the most fascinating things in all of mathematics. On the one hand, these numbers appear in the solutions to very natural problems, but on the other they exhibit an apparently random quality that defies simple characterization.

Contents

1. Introduction

The Bernoulli numbers are so exotic that one might well wonder how they came to be discovered. We are fortunate that this has been explained quite clearly in the posthumously published [Bernoulli:1713].

The story begins with the question, *what is the formula for the sum of k-powers $\sum_1^n n^k$?* The case $k = 0$ is trivial and the case $k = 1$ is well known:

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2} \,.$$

At the time Bernoulli took up the problem, many more cases were known, but it seems to have been he who first attacked it systematically. The key to his method is Pascal's triangle, which lays out the coefficients in the expansion

$$(x+y)^n = \sum_{i=0}^{n} c_{n,i}\, x^{n-i} y^i \,.$$

Here is a table of the first few:

n	$c_{n,0}$	$c_{n,1}$	$c_{n,2}$	$c_{n,3}$	$c_{n,4}$	$c_{n,5}$
0	1	0	0	0	0	0
1	1	1	0	0	0	0
2	1	2	1	0	0	0
3	1	3	3	1	0	0
4	1	4	6	4	1	0
5	1	5	10	10	5	1

The figure illustrates one feature of the table—the sum of the numbers in column i up through row n is the same as the number in column $i+1$ and row $n+1$. On the other hand, there is a simple formula for the entries. Therefore

(1.1) $$\sum_{m=0}^{n} \binom{m}{k} = \binom{n+1}{k+1} \,.$$

This allows us to find inductively a formula for $\sum_0^n m^k$. For example, suppose $k = 2$. Then applying (1.1)

$$\begin{aligned}\sum_{m=0}^n \frac{m(m-1)}{2} &= \frac{(n+1)n(n-1)}{6}\\ \sum_{m=0}^n \frac{m^2}{2} &= \frac{n^3-n}{6} - n(n+1)/4\\ &= \frac{n^3}{6} + \frac{n^2}{4} + \frac{n}{12}\\ \sum_{m=0}^n m^2 &= \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\end{aligned}$$

In this way, Bernoulli arrived at the following table:

$$\begin{array}{lccccccccccccccc}
\sum_{m=0}^n 1 & = & n & + & 1 \\
\sum_{m=0}^n m & = & \frac{n^2}{2} & + & \frac{n}{2} \\
\sum_{m=0}^n m^2 & = & \frac{n^3}{3} & + & \frac{n^2}{2} & + & \frac{n}{6} \\
\sum_{m=0}^n m^3 & = & \frac{n^4}{4} & + & \frac{n^3}{2} & + & \frac{n^2}{4} \\
\sum_{m=0}^n m^4 & = & \frac{n^5}{5} & + & \frac{n^4}{2} & + & \frac{n^3}{3} & - & \frac{n}{30} \\
\sum_{m=0}^n m^5 & = & \frac{n^6}{6} & + & \frac{n^5}{2} & + & \frac{5n^4}{12} & - & \frac{n^2}{12} \\
\sum_{m=0}^n m^6 & = & \frac{n^7}{7} & + & \frac{n^6}{2} & + & \frac{n^5}{2} & - & \frac{n^3}{6} & + & \frac{n}{42} \\
\sum_{m=0}^n m^7 & = & \frac{n^8}{8} & + & \frac{n^7}{2} & + & \frac{7n^6}{12} & - & \frac{7n^4}{24} & + & \frac{n^2}{12} \\
\sum_{m=0}^n m^8 & = & \frac{n^9}{9} & + & \frac{n^8}{2} & + & \frac{2n^7}{3} & - & \frac{7n^5}{15} & + & \frac{2n^3}{9} & - & \frac{n}{30} \\
\sum_{m=0}^n m^9 & = & \frac{n^{10}}{10} & + & \frac{n^9}{2} & + & \frac{3n^8}{4} & - & \frac{7n^6}{10} & + & \frac{n^4}{2} & - & \frac{n^2}{12} \\
\sum_{m=0}^n m^{10} & = & \frac{n^{11}}{11} & + & \frac{n^{10}}{2} & + & \frac{5n^9}{6} & - & n^7 & + & n^5 & - & \frac{n^3}{2} & + & \frac{5n}{66}
\end{array}$$

At first, some patterns in this table appear easily, but others do not. One hint is given by expressing the coefficients in the third column as $2/12, 3/12, \ldots, 10/12$. Another trick that might make things more obvious is to multiply the k-th row by $k+1$. In any event, what Bernoulli came up with is the following observation. Let β_m be the sequence of coefficients of n occurring in rows after the first. Thus

$$\begin{aligned}\beta_0 &= \ \ 1\\ \beta_1 &= \ \ 1/2\\ \beta_2 &= \ \ 1/6\\ \beta_3 &= \ \ 0\\ \beta_4 &= -1/30\\ \beta_5 &= \ \ 0\\ \beta_6 &= \ \ 1/42\,.\end{aligned}$$

Then what the table suggests is that all other numbers in the column m are simple integer multiples of β_m. This leads to the remarkable formula

$$(k+1)\Big(\sum_0^n m^k\Big) = n^{k+1} + (k+1)\cdot\beta_1\cdot n^k + \frac{(k+1)k}{2}\cdot\beta_2\cdot n^{k-1} + \frac{(k+1)k(k-1)(k-2)}{4!}\cdot\beta_4\cdot n^{k-3} + \cdots$$

In other words, the β_m ($m \geq 1$) are a constant factor in the coefficients of n^{k+1-m} in the formula for sums of k-th powers.

It is not at all clear that Bernoulli knew how to prove this formula. This was done a bit later by Euler, as we shall see in a moment. What Bernoulli did point out was that if you set $n = 2$ for any value of k, you would come up with a recursive formula for β_m, since then $\sum_{m=0}^{2} m^k = 1$.

$$\beta_n = (k+1) - \sum\nolimits_{m=0}^{n-1} \binom{k+1}{m} \cdot \beta_m \,.$$

Here is an expanded list of a few non-zero Bernoulli numbers:

$$\begin{aligned}
\beta_0 &= 1 \\
\beta_1 &= -1/2 \\
\beta_2 &= 1/6 \\
\beta_4 &= -1/30 \\
\beta_6 &= 1/42 \\
\beta_8 &= -1/30 \\
\beta_{10} &= 5/66 \\
\beta_{12} &= -691/2730 \\
\beta_{14} &= 7/6 \\
\beta_{16} &= -3617/510 \\
\beta_{18} &= 43867/798 \\
\beta_{20} &= -174611/330 \\
\beta_{22} &= 854513/138 \\
\beta_{24} &= -236364091/2730 \\
\beta_{26} &= 8553103/6 \\
\beta_{28} &= -23749461029/870 \\
\beta_{28} &= 8615841276005/14322 \\
\beta_{30} &= -7709321041217/510 \\
\beta_{32} &= 2577687858367/6 \\
\beta_{34} &= -26315271553053477373/1919190 \\
\beta_{36} &= 2929993913841559/6
\end{aligned}$$

Bernoulli also demonstrated implicitly his remarkable computational abilities by asserting that it took him only about 15 minutes to see that

$$1 + 2^{10} + \cdots + 1000^{10} = 91,409,924,241,424,243,424,241,924,242,400\,.$$

A little after Jacob Bernoulli died, Euler took up the investigation of the Bernoulli numbers. One of his first discoveries seems to have been that Bernoulli's recursive formula is equivalent to:

1.2. Proposition. *We have*

$$\frac{x}{e^x - 1} = \sum\nolimits_{m=0} \frac{\beta_m}{m!}\, x^m \,.$$

Estimates of the size of β_m will imply that this series converges for $|x| < 2\pi$, but of course modern analysis has this as a consequence of the location of zeroe sof $e^x = 1$.

2. The Euler-Maclaurin formula

Euler and the Scottish mathematician Colin Maclaurin discovered independently a very general formula from which Bernoulli's formula for the sum of powers follows easily. It is already hinted at by the dominant term in that formula, since

$$\frac{n^{k+1}}{k+1} = \int_0^n x^k \, dx \, .$$

In other words, one might often expect a discrete sum to be approximated by an integral.

I'll lead up to the Euler-Maclaurin formula with an initial computation. Suppose f to be a smooth function on the interval $[k, \ell]$. Then

$$\int_k^\ell f(x)\, dx = \sum\nolimits_{m=k}^{m=\ell-1} \int_m^{m+1} f(x)\, dx$$

and

$$\int_m^{m+1} f(x)\, dx = [f(x)\Psi_1(x)]_m^{m+1} - \int_m^{m+1} f'(x)\Psi_1(x)\, dx$$

if $\Psi_1'(x) = 1$ in the open interval $(m, m+1)$. Therefore

$$\begin{aligned}\int_k^\ell f(x)\, dx = &\big(f(k+1)\Psi_1^-(k+1) - f(k)\Psi_1^+(k)\big) \\ &+ \big(f(k+2)\Psi_1^-(k+2) - f(k+1)\Psi_1^+(k+1)\big) \\ &\quad + \cdots \\ &+ \big(f(\ell)\Psi_1^-(\ell) - f(\ell-1)\Psi_1^+(\ell-1)\big) \\ &\quad - \int_k^\ell f'(x)\Psi_1(x)\, dx \, .\end{aligned}$$

in which $\Psi_1^\pm(x)$ is the limit of Ψ_1 at x, taken from above (or below). I now specify

$$\Psi_1(x) = x - 1/2$$

in the interval $(0, 1)$, and of period 1. Since $\Psi^+(m) = -1/2$, $\Psi^-(m) = 1/2$, this gives us

$$\int_k^\ell f(x)\, dx = (1/2)(f(k) + f(\ell)) + f(k+1) + \cdots f(\ell-1) - \int_k^\ell f'(x)\Psi_1(x)\, dx \, .$$

or

$$f(k) + f(k+1) + \cdots + f(\ell-1) = \int_k^\ell f(x)\, dx - (1/2)\big(f(\ell) - f(k)\big) + \int_k^\ell f'(x)\Psi_1(x)\, dx \, .$$

This is in effect an approximation of a sum by an integral, together with an estimate of the approximation error.

We can now continue. Suppose $\Psi_2(x)$ to be a function such that $\Psi_2'(x) = \Psi_1(x)$ on the interior of intervals $[m, m+1]$. Since

$$\Psi_2(m+1) - \Psi_2(m) = \int_m^{m+1} \Psi_1(x)\, dx = 0$$

the function Ψ_2 is also periodic. Again integrating by parts, this gives us

$$\begin{aligned} & f(k) + f(k+1) + \ldots + f(\ell-1) \\ & \quad = \int_k^\ell f(x)\,dx - \frac{1}{2}\big(f(\ell) - f(k)\big) + \int_k^\ell f'(x)\Psi_1(x)\,dx \\ & \quad = \int_k^\ell f(x)\,dx - \frac{1}{2}\big(f(\ell) - f(k)\big) + \Psi_2(0)\big(f'(\ell) - f'(k)\big) - \int_k^\ell f^{(2)}(x)\Psi_2(x)\,dx\,. \end{aligned}$$

Specifying that $\Psi_2' = \Psi_1$ determines it only up to a constant. This constant is fixed by the additional condition that

$$\int_0^1 \Psi_2(x)\,dx = 0\,.$$

Now we have

$$\Psi_2(x) = \frac{x^2}{2} - \frac{x}{2} + \frac{1}{12}\,.$$

If we define a sequence of periodic functions Ψ_n by the inductive conditions

$$\begin{aligned} \Psi_{n+1}' &= \Psi_n \quad \text{on } (0,1) \\ \int_0^1 \Psi_n(x)\,dx &= 0 \end{aligned}$$

Then we arrive at:

2.1. Proposition. *Suppose f to be a function on the interval $[k,\ell]$ which has continuous derivatives up to order n. Then*

$$f(k) + f(k+1) + \ldots f(\ell-1) = \int_k^\ell f(x)\,dx - {}^1\!/_2\big(f(\ell) - f(k)\big) + \sum_{m=2}^{n} \Psi_m\big(f^{(m-1)}(\ell) - f^{m-1)}(k)\big) + R_n$$

where

$$R_n = \pm \int f^{(n)}(x)\Psi_n(x)\,dx\,.$$

We shall see later an estimate for the size of the remainder R_n. If f is a polynomial, then eventually the remainder vanishes, and this proves Bernoulli's formula for the sums of powers.

3. Bernoulli polynomials

Let $B_m(x)$ be the polynomials whose restriction to $[0,1]$ is the function $m!\Psi_m(x)$. These are called the **Bernoulli polynomials**, although I believe they were first recognized by Euler.

3.1. Proposition. *We have*

$$\frac{se^{sx}}{e^s - 1} = 1 + \sum\nolimits_{m=0} B_m(x) \cdot \frac{s^m}{m!}\,.$$

Proof. It suffices to verify that the polynomials defined by the equation satisfy the conditions defining $B_m(x)$. To see the first, take the derivative of the eqaution in the Proposition. To see the second, integrate both sides with respect to x, from 0 to 1. The miracle is that the left hand side gives

$$\frac{s}{e^s - 1} \cdot \left[\frac{e^{sx}}{s}\right]_0^1 = 1\,.$$

3.2. Corollary. *The Bernoulli numbers are the constant terms of the Bernoulli polynomials:*

$$\beta_m = B_m(0)\,.$$

One consequence is that the $B_n(x)$ are monic. Explicitly, if we set

$$B_n(x) = \sum_0^n b_{n,m} x^m$$

then these formulae lead to

$$b_{n,m} = \frac{n}{m} b_{n-1,m-1} \quad (m = 1 \dots n)$$

$$b_{n,0} = -\sum_{m=1}^{n} \frac{b_{n,m}}{m+1}$$

Here are some additional properties of these polynomials:

3.3. Proposition. *The Bernoulli polynomials possess the following properties:*

(a) *for all* n

$$(-1)^n B_n(1-x) = B_n(x)\,;$$

(b) *for* $n \geq 2$

$$B_{2n+1}(0) = B_{2n+1}(1/2) = B_{2n+1}(1) = 0\,.$$

Proof. Property (a) because the left hand side satisfies the same defining condition as $B_n(x)$ (proof by induction). As for (b), it follows from

$$B_{2n+1}(1) = -B_{2n+1}(0), \quad B_{2n+1}(1/2) = -B_{2n+1}(1/2)\,.$$

3.4. Proposition. *For* $n \geq 1$, $(-1)^{n-1} B_{2n+1}(x)$ *is positive in* $(0, 1/2)$, *negative in* $(1/2, 0)$. *The constant term of* $(-1)^{n-1} B_{2n}(x)$ *is positive.*

More explicitly, $(-1)^{n-1} B_{2n}(0) > 0$, decreases to a minimum at $x = 1/2$, crossing the x-axis exactly once in $(0, 1/2)$, then behaves symmetrically in $(1/2, 1)$.

Proof. These inequalities can be proven by induction, noting that it is true for $B_2(x)$ and using the formula

$$(-1)^{n-1} B_{2n+1}(x) = \frac{(-1)^{n-1}}{2n+1} \int_0^x B_{2n}(x)\,dx\,.$$

Therefore for $n > 1$

$$(-1)^{n-1}\beta_{2n} > 0, \quad \beta_{2n+1} = 0$$

The polynomials can be constructed directly from the β_n, and the β_n themselves inductively:

$$B_n(x) = \sum_{i=0}^{n} \binom{n}{i} \beta_{n-i} x^i\,.$$

There is apparently no simple way to calculate the Bernoulli polynomials (or numbers). They have interesting number-theoretic properties (see the book **Number Theory** by Borevitch and Shafarevitch). In spite of the small size of the first few, their magnitude grows rapidly, and it fact it will follow from things I prove later that for large m we have

$$\beta_{2m} \sim 4\sqrt{m\pi} \left(\frac{m}{\pi e}\right)^{2m}.$$

4. Fourier series

Define $\psi_m(x)$ to be the periodic extension of the restriction of $B_m(x)$ to $[0,1]$. For $m > 1$ these are continuous functions. The Fourier series of $\psi_1(x)$ can be calculated explicitly. Its n-th Fourier coefficient is

$$\int_0^1 (x - {}^1\!/_2)\, e^{-2\pi i n x}\, dx = -\frac{1}{2\pi n i}$$

so that

$$\psi_1(x) = -\left(\frac{\sin 2\pi x}{\pi} + \frac{\sin 4\pi x}{2\pi} + \frac{\sin 6\pi x}{3\pi} + \ldots\right).$$

For $n > 0$ the constant terms of the Fourier series of $\psi_n(x)$ must be null. If the n-th Fourier coefficient of f is c_n then that of $f'(x)$ is $2\pi n i c_n$. Therefore the n-th coefficient of $\psi_k(x)$ must be $-k!/(2\pi n i)^k$. In other words:

$$\psi_{2n}(x) = (-1)^{n-1}\, 2\,(2n)! \sum_1^\infty \frac{\cos 2\pi k x}{(2\pi k)^{2n}}$$

$$\psi_{2n+1}(x) = (-1)^{n-1}\, 2\,(2n+1)! \sum_1^\infty \frac{\sin 2\pi k x}{(2\pi k)^{2n+1}}$$

From the first:

4.1. Proposition. *For any positive integer* m

$$\zeta(2n) = (-1)^{n-1}(2\pi)^{2n}\psi_{2n}(0)/2(2n)! = (2\pi)^{2n}|\beta_{2n}|/2(2n)!$$

So for example

$$\zeta(2) = \pi^2/6, \quad \zeta(4) = \pi^4/90.$$

We have

$$|\psi_{2n}(x)| \le |\beta_{2n}|$$

$$|\psi_{2n+1}(x)| \le (n + {}^1\!/_2)\,|\beta_{2n}|$$

The first is immediate from the Fourier series. The second follows from the first, the fact that $\psi_{2n+1}(x) = 0$ for $x = 0$ or $x = 1/2$, and the Mean Value Theorem of elementary calculus.

5. Calculating Euler's constant

Suppose that f satisfies (a) $f(\ell) \to 0$ as $\ell \to \infty$, and (b) each $f^{(k)}(x)$ is integrable near ∞ for all $k \ge 1$. Then we can write an earlier formula as

$$f(1) + f(2) + \ldots + f(\ell - 1) - \int_1^\ell f(x)\, dx = \frac{1}{2}\big(f(1) - f(\ell)\big) + \int_1^\ell f'(x)\psi_1(x)\, dx.$$

As $\ell \to \infty$ the right-hand side has as limit the constant

$$C_f = \frac{1}{2} f(1) + \int_1^\infty f'(x)\psi_1(x)\, dx$$

so that we can also write

$$C_f = f(1) + f(2) + \ldots + f(\ell-1) - \int_1^\ell f(x)\,dx + \frac{1}{2}f(\ell) + \int_\ell^\infty f'(x)\psi_1(x)\,dx$$

and again apply successive integration by parts to get

$$C_f = f(1)+f(2)+\ldots+f(\ell-1)-\int_1^\ell f(x)\,dx+\frac{1}{2}f(\ell)-\frac{\beta_2}{2!}f'(\ell)-\frac{\beta_4}{4!}f^{(3)}(\ell)-\ldots-\frac{\beta_{2n}}{(2n)!}f^{(2n-1)}(\ell)+R_{2n+1}(\ell)$$

where

$$R_{2n+1}(\ell) = \frac{1}{(2n+1)!}\int_\ell^\infty f^{(2n+1)}(x)\psi_{2n+1}(x)\,dx.$$

To apply this to calculate γ we set

$$f(x) = \frac{1}{x},\ f'(x) = -\frac{1}{x^2},\ \ldots,\ f^{(n)}(x) = (-1)^n\frac{n!}{x^{n+1}}.$$

5.1. Proposition. *We have the asymptotic expansion*

$$\gamma_n \sim 1 + \frac{1}{2} + \frac{1}{3} + \ldots \frac{1}{\ell-1} - \log\ell + \frac{1}{2\ell} + \frac{\beta_2}{2\ell^2} + \frac{\beta_4}{4\ell^4} + \frac{\beta_6}{6\ell^6} + \frac{\beta_8}{8\ell^8}\cdots$$

where the series is asymptotic in the strong sense that γ lies between any two successive partial sums.

Note that the signs of successive terms alternate.

The last claim follows from the general fact that if $f > 0$ is decreasing on $[0,1]$ then

$$(-1)^{n-1}\int_0^1 f(x)B_{2n+1}(x)\,dx > 0$$

which follows from an earlier proposition. This guarantees that the sign of the remainder changes with every increment of n.

This series does not converge, since the β_{2n} increase in magnitude so rapidly as $n \to \infty$. But it can be used to calculate γ with arbitrary accuracy if only ℓ is chosen large. It is what the English call just a **divergent series**, and Carl Ludwig Siegel with more accuracy a **semi-convergent series**. In fact, because the first few terms β_{2n} are relatively small, even a moderate-sized ℓ can give astounding accuracy. Here is how the calculation goes with 20 decimal accuracy for $\ell = 10$:

Sum	Next term
0.5 2638316097420828423	0.05000000000000000000
0.57 638316097420828423	0.00083333333333333333
0.57721 649430754161756	-0.00000083333333333333
0.57721566 097420828423	0.00000000396825396825
0.5772156649 4246225248	-0.00000000004166666666
0.57721566490 079558582	0.00000000000075757575
0.5772156649015 5316157	-0.00000000000002109279
0.577215664901532 06878	0.00000000000000083333
0.577215664901532 90211	-0.00000000000000004432
0.5772156649015328 5779	0.00000000000000000305
0.577215664901532860 84	-0.00000000000000000026
0.577215664901532860 58	0.00000000000000000002
0.57721566490153286060	0

One might expect some rounding errors in this calculation, but in fact

$$\gamma = 0.57721\,56649\,01532\,86060\,65120\,90082$$

correct to 30 decimals. We shall see later that very similar calculations can be used to evaluate $\Gamma(s)$ and $\zeta(s)$ with reasonable efficiency.

6. References

1. Jacob Bernoulli, **Ars Conjectandi**, Basel, 1713.

An English translation of the discussion we are interested in appeared in the anthology [Smith:1929]. It is very readable. A translation of the entire book, by Edith Sylla, was published only in 2005 by the Johns Hopkins University Press.

2. David Eugene Smith, **Source Book in Mathematics**, McGraw-Hill, 1929.