The Bernoulli Numbers

John C. Baez, December 23, 2003

The numbers B_k *are defined by the equation*

$$\frac{x}{e^x - 1} = \sum_{n \geq 0} B_k \frac{x^k}{k!}.$$

They are called the **Bernoulli numbers** *because they were first studied by Johann Faulhaber in a book published in 1631, and mathematical discoveries are never named after the people who made them. For example, the 'Čech compactification' was invented by Tychonoff, while 'Tychonoff's Theorem' is due to Čech.*

It had been known since antiquity that

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

but Faulhaber, the 'Arithmetician of Ulm', seems to have been the first to seek a general formula for the sum of pth powers of the numbers from 1 *to n. This general formula requires what are now called the Bernoulli numbers.*

Faulhaber's work was cited and further developed by Bernoulli in his famous Ars Conjectandi, *published posthumously in 1713. In this book, Bernoulli wrote that he used the method described below to compute "in less than half a quarter an hour" that*

$$1^{10} + \cdots + 1000^{10} = 91409924241424243424241924242500.$$

What a showoff! Let's see how he did it. I'll give a modern explanation of the trick.

Sums are to integrals as the **difference operator**

$$(\Delta f)(z) := f(z+1) - f(z)$$

is to derivatives. It would be crazy to tackle fancy integrals before mastering derivatives. So, before trying to do fancy sums, we should master the difference operator! We will think of this as an operator

$$\Delta \colon \mathcal{E} \to \mathcal{E}$$

where $\mathcal{E}$ *is the space of entire functions on the complex plane. Note that* $\mathcal{E}$ *is contained in the space of formal power series* $\mathbb{C}[[z]]$*, so it gives us a slight variation on a theme we know and love: the Fock representation. In particular, we have* **annihilation** *and* **creation** *operators*

$$a \colon \mathcal{E} \to \mathcal{E}, \qquad a^* \colon \mathcal{E} \to \mathcal{E}$$

given by the usual formulas:

$$\begin{aligned} (af)(z) &= \tfrac{d}{dz} f(z) \\ (a^* f)(z) &= z f(z) \end{aligned}$$

The advantage of $\mathcal{E}$ *is that we can evaluate its elements at any point of the complex plane, while formal power series may not converge. We need this to make* Δ *well-defined.*

Now: since the difference operator is a discretized version of the derivative, let us find a formula for Δ *in terms of* a*!*

1. For any $t \in \mathbb{C}$, define the operator

$$e^{ta} \colon \mathcal{E} \to \mathcal{E}$$

by

$$e^{ta} = \sum_{k \geq 0} \frac{(ta)^k}{k!}.$$

Show that

$$(e^{ta} f)(z) = f(z+t).$$

2. Show that

$$\Delta = e^a - 1.$$

Next: just as integration is the inverse of differentiation, summation should be the inverse of the difference operator. So, to do sums we just need an inverse for the difference operator Δ*. To see what I mean, prove the following discrete version of the Fundamental Theorem of Calculus:*

3. Show that if $F \in \mathcal{E}$ and

$$\Delta F = f$$

then

$$\sum_{i=0}^{n-1} f(i) = F(n) - F(0).$$

So, if we could compute F *from* f *via* $F = \Delta^{-1} f$ *we'd have a nifty formula for summing the values of* f*. There's just one slight catch: the operator* $\Delta \colon \mathcal{E} \to \mathcal{E}$ *isn't invertible! The reason is that* $\Delta F = 0$ *whenever* F *is a constant function. Since* Δ *has a nontrivial kernel, it can't have an inverse.*

In fact, this problem is already familiar from ordinary calculus. I was lying slightly: integration isn't really *the inverse of differentiation as an operator on functions. The derivative of any constant is zero, so the integral of a function is only well-defined up to a constant.*

However, the operators a *and* Δ *do have* one-sided *inverses, and this all we really need!*

4. Define the operator $a^{-1} \colon \mathcal{E} \to \mathcal{E}$ by

$$(a^{-1}) f(z) = \int_0^z f(u) du.$$

Show that

$$a a^{-1} f = f$$

but

$$\text{not necessarily } a^{-1} a f = f$$

for all $f \in \mathcal{E}$.

We now use a^{-1} *to concoct a one-sided inverse for the difference operator* Δ *as follows. We saw how the operator* $\Delta = e^a - 1$ *can be written as a power series in* a*. We'd like to do the same for*

$$\Delta^{-1} = \frac{1}{e^a - 1},$$

but we can't, since the function

$$\frac{1}{e^x - 1}$$

has a simple pole at $x = 0$. To get rid of this pole, we can work instead with

$$\frac{x}{e^x - 1}.$$

This is an entire function if we define its value at $x = 0$ to be 1, so we can write it as a power series:

$$\frac{x}{e^x - 1} = \sum_{k \geq 0} B_k \frac{x^k}{k!}$$

where B_k are the **Bernoulli numbers**. *We can then define an operator*

$$\frac{a}{e^a - 1} \colon \mathcal{E} \to \mathcal{E}$$

by the same power series:

$$\frac{a}{e^a - 1} := \sum_{k \geq 0} B_k \frac{a^k}{k!}$$

Next, we define the operator Δ^{-1} by

$$\Delta^{-1} = \frac{a}{e^a - 1} \, a^{-1}.$$

where a^{-1} is the one-sided inverse of a. And believe it or not, Δ^{-1} is a one-sided inverse of Δ:

5. Starting from the definition of the Bernoulli numbers, show that

$$x = \sum_{j \geq 1} \frac{x^j}{j!} \sum_{k \geq 0} B_k \frac{x^k}{k!}. \tag{1}$$

Use this to show

$$a = \sum_{j \geq 1} \frac{a^j}{j!} \sum_{k \geq 0} B_k \frac{a^k}{k!}$$

and therefore

$$a = \Delta \, \frac{a}{e^a - 1}. \tag{2}$$

6. Using parts 4 and 5 together with the definition of Δ^{-1}, show that

$$\Delta \Delta^{-1} f = f$$

for all $f \in \mathcal{E}$, but

not necessarily $\Delta^{-1} \Delta f = f$.

Warning: here's how not *to do the first part:*

$$\Delta \Delta^{-1} = (e^a - 1) \frac{a}{e^a - 1} a^{-1} = a a^{-1} = 1$$

You can't simply cancel the factors of $e^a - 1$, because $\frac{a}{e^a - 1}$ is not defined as the product of a and the (nonexistent) operator $(e^a - 1)^{-1}$ — it's defined as a power series involving the Bernoulli numbers! This is why we need the rigamarole in part 5.

Now, here's how we use this abstract nonsense to calculate a sum! Let's do the example Faulhaber did, where the function to be summed is

$$f(z) = z^p.$$

Then we have

$$\Delta^{-1} f = \frac{a}{e^a - 1} a^{-1} f$$

but by definition,

$$(a^{-1} f)(z) = \int_0^z f(u)du = \frac{z^{p+1}}{p+1}$$

so

$$\begin{aligned}
(\Delta^{-1} f)(z) &= \frac{a}{e^a - 1} \frac{z^{p+1}}{p+1} \\
&= \sum_{k \geq 0} \frac{B_k}{k!} \frac{d^k}{dz^k} \frac{z^{p+1}}{p+1} \\
&= \sum_{k=0}^{p+1} \frac{B_k}{k!} (p+1)p \cdots (p+1-k) \frac{z^{p+1-k}}{p+1} \\
&= \frac{1}{p+1} \sum_{k=0}^{p+1} B_k \binom{p+1}{k} z^{p+1-k}
\end{aligned}$$

and since Δ^{-1} is a one-sided inverse of Δ, part 3 gives

$$\begin{aligned}
\sum_{i=0}^{n-1} i^p &= (\Delta^{-1} f)(n) - (\Delta^{-1} f)(0) \\
&= \frac{1}{p+1} \sum_{k=0}^{p+1} B_k \binom{p+1}{k} n^{p+1-k}
\end{aligned}$$

or if you prefer,

$$\sum_{i=1}^{n} i^p = \frac{1}{p+1} \sum_{k=0}^{p+1} B_k \binom{p+1}{k} (n+1)^{p+1-k}.$$

Since the right-hand sum has just $p+1$ terms, regardless of n, it's actually much easier to calculate than the left-hand sum when n gets big.

... or at least it would be easy if we knew the Bernoulli numbers! We could calculate these by grinding out the Taylor series of $x/(e^x - 1)$, but here's a much quicker way:

7. By taking the coefficient of x^{i+1} on both sides of equation (1), show that

$$0 = \frac{B_0}{0!} \cdot \frac{1}{(i+1)!} + \frac{B_1}{1!} \cdot \frac{1}{i!} + \cdots + \frac{B_i}{i!} \cdot \frac{1}{1!}$$

for $i \geq 1$.

8. Using part 7, show that

$$\begin{aligned}
0 &= \mathbf{1}B_0 + \mathbf{2}B_1 \\
0 &= \mathbf{1}B_0 + \mathbf{3}B_1 + \mathbf{3}B_2 \\
0 &= \mathbf{1}B_0 + \mathbf{4}B_1 + \mathbf{6}B_2 + \mathbf{4}B_3 \\
0 &= \mathbf{1}B_0 + \mathbf{5}B_1 + \mathbf{10}B_2 + \mathbf{10}B_3 + \mathbf{5}B_4
\end{aligned}$$

and so on, where the numbers in boldface are taken from Pascal's triangle in an obvious way.

9. Show that $B_0 = 1$, and by using the recurrence in part 8 work out B_1, B_2, B_3 and B_4.

10. Use everything we've done to work out an explicit formula for

$$1^4 + \cdots + n^4.$$

11. Show that

$$1^p + \cdots + n^p = \frac{(B + n + 1)^{p+1} - B^{p+1}}{p + 1}$$

where the rules of the game are that we expand the right-hand side using the binomial theorem and then set B^i equal to the Bernoulli number B_i. What formula do you get if you set $n = 0$ here? Where have you seen that formula before?

12. *Extra Credit:* categorify this homework assignment as much as possible, replacing $\mathcal{E}$ by the category of **entire structure types**: that is, structure types whose generating functions are entire functions on the complex plane. Why is this hard?