Sums of Powers of Integers: A Little of the History

Author(s): A. W. F. Edwards
Source: *The Mathematical Gazette*, Vol. 66, No. 435 (Mar., 1982), pp. 22-28
Published by: The Mathematical Association
Stable URL: http://www.jstor.org/stable/3617302
Accessed: 06-02-2018 19:39 UTC

note that, by using pictorial ideas, it is possible to show that the blancmange function is nowhere differentiable, a fact that is considered 'too difficult' to explain in most undergraduate mathematics courses. What is more important is that these practical ideas translate into a correct formal proof, now invested with geometric insight sadly lacking in so much formal mathematics.

"Intuition" is not a low-level phenomenon to be excluded from higher mathematics, it is a highly personal mental activity produced by experience. If we give the right experiences and enhance intuition then it can result in a much more profound understanding.

DAVID TALL

Mathematics Education Research Centre, University of Warwick, Coventry CV4 7AL

Sums of powers of integers: a little of the history

A. W. F. EDWARDS

The lack of any obvious pattern amongst the Bernoulli numbers $(1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, -\frac{1}{30}, \ldots)$ is one of the shocks of analysis which subsequent familiarity with the many beautiful and simple means of deriving them does not altogether assuage. Historically, they first arose in connection with the sums of the rth powers of the first n integers

$$\sum_{i=1}^{n} i^r = 1^r + 2^r + 3^r + \cdots + n^r \tag{1}$$

which it is convenient to write as $\sum n^r$. The Greeks, Hindus, and Arabs all had rules amounting to

$$\left.\begin{aligned} \sum n &= \tfrac{1}{2}n(n+1) = \tfrac{1}{2}n^2 + \tfrac{1}{2}n \\ \sum n^2 &= \tfrac{1}{6}n(n+1)(2n+1) = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n \\ \sum n^3 &= [\tfrac{1}{2}n(n+1)]^2 = \tfrac{1}{4}n^4 + \tfrac{1}{2}n^3 + \tfrac{1}{4}n^2 \end{aligned}\right\} \tag{2}$$

whilst a fifteenth-century Arab rule for the fourth powers was equivalent to

$$\begin{aligned} \sum n^4 &= \tfrac{1}{30}n(n+1)(2n+1)(3n^2+3n-1) \\ &= \tfrac{1}{5}n^5 + \tfrac{1}{2}n^4 + \tfrac{1}{3}n^3 - \tfrac{1}{30}n, \end{aligned} \tag{3}$$

there being no n^2 term in the second form.

With these four formulae to go on (prefaced if necessary by the trivial $\sum 1 = n$), mathematicians in the first half of the seventeenth century attempted to find general forms for $\sum n^r$, $r = 5, 6, 7 \ldots$. In 1636 Fermat discovered a recurrence relation, based on the figurate numbers, which gave $\sum n^r$ in terms of $\sum n^{r-1}, \sum n^{r-2}, \ldots$, but the successive algebraic substitutions soon become intractable, whilst in 1654 Pascal derived a more practicable formula from the binomial expansion as a sequel to his investigations on the Arithmetical Triangle. It still gives $\sum n^r$ in terms of the lower-order sums, but this time the coefficients are readily computed, as we show below. But both these famous French authors had overlooked the results of the weaver of Ulm, Johann Faulhaber, who by 1631 had completed the publication of all the sums up to $\sum n^{17}$. Nor did he find them by brute force alone, even though by 1615 he was aware of the result in the figurate numbers that makes Fermat's method possible; rather, he generalised the first of the forms given above, using his discovery that

$$\left.\begin{aligned} &\textstyle\sum n^r \ (r \text{ odd}) = \text{a polynomial in } n(n+1) \\ &\textstyle\sum n^r \ (r \text{ even}) = (2n+1) \times (\text{a polynomial in } n(n+1)) \end{aligned}\right\} \qquad (4)$$

Having obtained the rules by which these polynomials could be found for successive r's, in each case he expanded his result into the second of the forms given above in (2) and (3), that is, as a polynomial in n of degree $r + 1$. It is a little surprising that the French authors were ignorant of these results because the fame of Faulhaber was such that Descartes had spent some time with him in Ulm in 1620, and he was a prolific author.

Then, in his posthumous *Ars conjectandi* of 1713, James Bernoulli, mentioning Faulhaber, gave the second forms up to $r = 10$ in a table (Table 1) from which it is easy to see the pattern of the coefficients of the powers of n, for each coefficient derives from the one above it by a simple rule. It is only the numbers at the heads of the columns which are mysterious, and Bernoulli saw (as Faulhaber had done) that they could easily be determined since the coefficients in each row must sum to 1 (putting $n = 1$, $\sum 1^r = 1$). Bernoulli actually made a mistake in the

$$\begin{aligned}
\textstyle\sum n &= \tfrac{1}{2}n^2 + \tfrac{1}{2}n \\
\textstyle\sum n^2 &= \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n \\
\textstyle\sum n^3 &= \tfrac{1}{4}n^4 + \tfrac{1}{2}n^3 + \tfrac{1}{4}n^2 \\
\textstyle\sum n^4 &= \tfrac{1}{5}n^5 + \tfrac{1}{2}n^4 + \tfrac{1}{3}n^3 - \tfrac{1}{30}n \\
\textstyle\sum n^5 &= \tfrac{1}{6}n^6 + \tfrac{1}{2}n^5 + \tfrac{5}{12}n^4 - \tfrac{1}{12}n^2 \\
\textstyle\sum n^6 &= \tfrac{1}{7}n^7 + \tfrac{1}{2}n^6 + \tfrac{1}{2}n^5 - \tfrac{1}{6}n^3 + \tfrac{1}{42}n \\
\textstyle\sum n^7 &= \tfrac{1}{8}n^8 + \tfrac{1}{2}n^7 + \tfrac{7}{12}n^6 - \tfrac{7}{24}n^4 + \tfrac{1}{12}n^2 \\
\textstyle\sum n^8 &= \tfrac{1}{9}n^9 + \tfrac{1}{2}n^8 + \tfrac{2}{3}n^7 - \tfrac{7}{15}n^5 + \tfrac{2}{9}n^3 - \tfrac{1}{30}n \\
\textstyle\sum n^9 &= \tfrac{1}{10}n^{10} + \tfrac{1}{2}n^9 + \tfrac{3}{4}n^8 - \tfrac{7}{10}n^6 + \tfrac{1}{2}n^4 - \tfrac{3}{20}n^2 \\
\textstyle\sum n^{10} &= \tfrac{1}{11}n^{11} + \tfrac{1}{2}n^{10} + \tfrac{5}{6}n^9 - n^7 + n^5 - \tfrac{1}{2}n^3 + \tfrac{5}{66}n
\end{aligned}$$

TABLE 1. James Bernoulli's *Summae Potestatum* (corrected)

coefficient of n^2 in $\sum n^9$, which he gave as $-1/12$ and which has been faithfully repeated by commentators (though not, of course, by the constructors of new tables) ever since. He did not attempt to prove the general form, and made no use of the relations (4) above.

Euler, in turn, tackled the problem of summing the powers and in 1755 published a proof of the Bernoulli forms based on the calculus of finite differences, christening the coefficients of n taken positively for $r = 2, 4, 6 \ldots$ the *Bernoulli numbers* in honour of James. Poor Faulhaber might reasonably have felt a little aggrieved at this, since he had not only published the 'Bernoulli' numbers up to $r = 16$ a hundred and twenty-four years previously, but he had given a method, albeit needlessly laborious, by which further numbers could be computed. There is little we can do to repair the nomenclature now, but we could call (4) *Faulhaber polynomials*. They were not rediscovered until Jacobi applied the Euler–Maclaurin summation formula to the sums of the powers in 1834; since that formula came out of a knowledge of the sums of the powers in the first place, and Euler mentioned Bernoulli who mentioned Faulhaber, the trail of two centuries would not have been too difficult to follow. Even the usually well-informed J. W. L. Glaisher, writing about the sums of the powers in 1899, said he knew of nothing on the subject beside Jacobi's paper and a brief and inconsequential note by Cayley in 1858. Since then the Faulhaber forms have been rediscovered on more than one occasion, but never, to my knowledge, correctly attributed.

So much for history; I now show how, with elementary matrix theory applied to Pascal's method, the sums of the powers may be derived with consummate ease, and in particular how one of the standard algorithms for calculating the Bernoulli numbers may be made self-evident. The name 'Pascal matrices' may be given to a family of infinite triangular matrices whose elements are derived, more-or-less directly, from the coefficients of Pascal's Arithmetical Triangle. Pascal, of course, knew nothing of matrices, nor even of determinants, but the peculiar appropriateness of attaching his name to this particular kind of matrix will become clear in due course.

The first Pascal matrix can only be:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & . & . \\ 1 & 1 & 0 & 0 & 0 & . & . \\ 1 & 2 & 1 & 0 & 0 & . & . \\ 1 & 3 & 3 & 1 & 0 & . & . \\ 1 & 4 & 6 & 4 & 1 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix}$$

and our first theorem is:

Theorem 1

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & . & . \\ -1 & 1 & 0 & 0 & 0 & . & . \\ 1 & -2 & 1 & 0 & 0 & . & . \\ -1 & 3 & -3 & 1 & 0 & . & . \\ 1 & -4 & 6 & -4 & 1 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix}$$

Proof A hint will be sufficient: consider the expansions of $(x+1)^r$ and $(x-1)^r$.

Our second theorem is simply the matrix formulation of the identities found by Euler and Vandermonde in the 1770s, which may themselves be thought of as generalisations of the addition relation for finding binomial coefficients:

Theorem 2

$$PP' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & . & . \\ 1 & 2 & 3 & 4 & 5 & . & . \\ 1 & 3 & 6 & 10 & 15 & . & . \\ 1 & 4 & 10 & 20 & 35 & . & . \\ 1 & 5 & 15 & 35 & 70 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix}$$

Proof Again, a hint will suffice: just as the identity $(x+1)^{r+1} = (x+1)(x+1)^r$ establishes the addition relation, identities of the form $(x+1)^{r+s} = (x+1)^s(x+1)^r$, $s = 2, 3, 4, \ldots$ establish its generalizations.

In 1654 Pascal gave what has become the standard school method for summing the powers:

$$(n+1)^r = (n+1) + r\sum n + \binom{r}{2}\sum n^2 + \binom{r}{3}\sum n^3 + \cdots + \binom{r}{r-1}\sum n^{r-1}. \quad (5)$$

Writing it successively for $r = 1, 2, 3, \ldots$ and applying matrix notation to the resultant simultaneous linear equations we find:

$$\begin{pmatrix} (n+1) \\ (n+1)^2 \\ (n+1)^3 \\ (n+1)^4 \\ (n+1)^5 \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & . & . \\ 1 & 2 & 0 & 0 & 0 & . & . \\ 1 & 3 & 3 & 0 & 0 & . & . \\ 1 & 4 & 6 & 4 & 0 & . & . \\ 1 & 5 & 10 & 10 & 5 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix} \begin{pmatrix} n+1 \\ \sum n \\ \sum n^2 \\ \sum n^3 \\ \sum n^4 \\ \cdot \\ \cdot \end{pmatrix}$$

The matrix may be called the second Pascal matrix, Q, derived from the first by the loss of its main diagonal. Inverting Q and replacing n by $n-1$ we have

Theorem 3

$$\begin{pmatrix} n \\ \sum(n-1) \\ \sum(n-1)^2 \\ \sum(n-1)^3 \\ \sum(n-1)^4 \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & . & . \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & . & . \\ \frac{1}{6} & -\frac{1}{2} & \frac{1}{3} & 0 & 0 & . & . \\ 0 & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} & 0 & . & . \\ -\frac{1}{30} & 0 & \frac{1}{3} & -\frac{1}{2} & \frac{1}{5} & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix} \begin{pmatrix} n \\ n^2 \\ n^3 \\ n^4 \\ n^5 \\ \cdot \\ \cdot \end{pmatrix}$$

This theorem gives the sums of the powers of the integers, in each case up to the term $(n-1)^r$, as polynomials in n. To obtain Bernoulli's *Summae Potestatum* exactly as given in our Table 1 it is only necessary to add n^r to each sum, thereby changing all the coefficients in the diagonal below the main diagonal of the matrix Q^{-1} from $-\frac{1}{2}$ to $+\frac{1}{2}$.

In his recent *Mathematical Gazette* article (December 1980) S. H. Scott arrived at Q^{-1}, but without realising its precise significance. He had obtained (5) by Pascal's method, but instead of carrying straight on as we have done he found 'the repeated use of this method ... formidable' and diverted down a rougher road. His second method for computing the coefficients is, incidentally, the one actually advocated by James Bernoulli in *Ars conjectandi* (1713).

How delighted Pascal would have been to learn that his own method for finding the sums of the powers could be completed by inverting a matrix of coefficients from the Arithmetical Triangle! (Hence the appropriateness of attaching his name to such matrices.)

The first column of Q^{-1} gives the Bernoulli numbers B_0, B_1, B_2, ... as they are often nowadays defined, even though B_1 is then $-\frac{1}{2}$ rather than the $+\frac{1}{2}$ of Bernoulli's table. When Euler first christened the Bernoulli numbers he avoided this problem by starting with $B_2 = 1/6$.

Theorem 4

$$B_0 = 1 \qquad \sum_{k=0}^{j} \binom{j}{k} B_k = B_j \quad (j \geqslant 2)$$

Proof An immediate consequence of the B's being the first column of Q^{-1}.

This well-known relation is sometimes described symbolically in the form '$B_j = (B + 1)^j$, where the exponents are to be degraded to subscripts after the binomial has been expanded'. Exactly the same principle can be used to find the elements in any other column of Q^{-1}, starting always with the number in the main diagonal as the 'seed'.

If one is upset by the fact that $B_1 = -\frac{1}{2}$ in this approach, matters may be set right by deriving:

Theorem 5

$$\begin{pmatrix} n \\ \sum n \\ \sum n^2 \\ \sum n^3 \\ \sum n^4 \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdot & \cdot \\ -1 & 2 & 0 & 0 & 0 & \cdot & \cdot \\ 1 & -3 & 3 & 0 & 0 & \cdot & \cdot \\ -1 & 4 & -6 & 4 & 0 & \cdot & \cdot \\ 1 & -5 & 10 & -10 & 5 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}^{-1} \begin{pmatrix} n \\ n^2 \\ n^3 \\ n^4 \\ n^5 \\ \cdot \\ \cdot \end{pmatrix}$$

Proof Recall the standard school proof of Pascal's formula (5), which relies on the expansion of $(x + 1)^r - x^r$; repeat the argument on $(x - 1)^r - x^r$, and then continue in a way analogous to the derivation of Theorem 3.

Theorem 5 gives the coefficients in the expressions for the sums of the powers exactly as they occur in Bernoulli's table; moreoever, the corresponding theorem to Theorem 4 leads to the modified Bernoulli numbers B_j' where $B_j' = B_j$, $j \geqslant 2$, $B_0' = B_0$, $B_1' = \frac{1}{2} = B_1 + 1$, constructed symbolically from '$B_j' = (B' - 1)^j$, $j \geqslant 2$, $B_0' = 1$, where the exponents are to be degraded to subscripts after the binomial has been expanded.'

All the properties of the Bernoulli numbers can, of course, be derived from these matrix formulations, in particular their known expressions in terms of determinants, given originally by F. Siacci in 1865 and then by J. Hammond in 1875 and E. Lucas in 1876.

A. W. F. EDWARDS

Gonville and Caius College, Cambridge

Parallels

P. J. GIBLIN

Given a straight line L in the plane, how can we find all lines parallel to L? Here is one way: at each point P of L draw a straight line perpendicular to L. Now choose a number d and measure a distance d down these lines, always in the same direction. Joining up the points P' so obtained gives the parallel to L at distance d.

This definition of parallel works for any curve C, provided C has a tangent line at each point. For we can measure a distance d along the "normal" to C at each point P. The normal is the line through P perpendicular to the tangent at P. Joining up the points P' gives the parallel

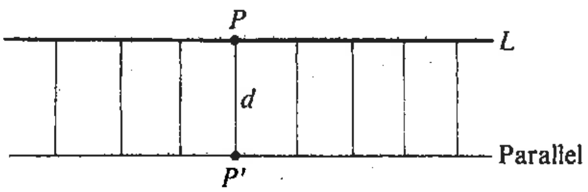

FIGURE 1.

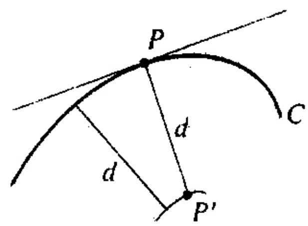

FIGURE 2.

to C at distance d. (Measuring d the other side of C may now give a completely different looking parallel.) It is reasonably clear that the parallel to a circle, centre A, radius r, at distance d, will be another circle centre A. The radius will be $r+d$ or $r-d$ or $d-r$: can you see why it might be any of these?

Given C and a reasonably accurate way of drawing normals it is possible to get a good idea of what the parallels to C look like. An alternative method, given C, is to draw lots of circles, centred on C, of radius d. These will all touch the two parallels at distance d on the two sides of C. The parallels appear as the "envelope" of the circles. In Fig. 3, C is the curve $y^2 = x - x^3$ (for $0 \leqslant x \leqslant 1$) and $d = 0{\cdot}2$. By tracing C and using different values of d you can obtain other parallels.

Parallels are physically interesting because we can imagine light, or other radiation, propagated from each point of C along the normals to C: the