Elem. Math. 65 (2010) 68 – 77
0013-6018/10/020068-10
DOI 10.4171/EM/142

© Swiss Mathematical Society, 2010

Elemente der Mathematik

An elementary proof of a generalization of Bernoulli's formula

Kevin J. McGown and Harold R. Parks

Harold R. Parks received his Ph.D. from Princeton University in 1974. After a postdoctoral instructorship at Brown University, he joined the mathematics faculty at Oregon State University. His main field of interest is geometric analysis.

Kevin J. McGown received his Master of Science degree from Oregon State University in 2005. He is presently at the University of California, San Diego, working on his Ph.D. thesis. His main field of interest is number theory.

1 Introduction

The familiar formulas

$$\sum_{n=1}^{N} n = \frac{N(N+1)}{2} \quad \text{and} \quad \sum_{n=1}^{N} n^2 = \frac{N(N+1)(2N+1)}{6}$$

are special cases of a general formula discovered by Jakob Bernoulli (1654–1705) and published posthumously in *Ars conjectandi* (1713). To describe Bernoulli's formula, let a and N be positive integers, then the sum of the ath powers of the first N positive integers

Jakob Bernoulli (1654–1705) verdanken wir eine geschlossene Formel für die Summe der a-ten Potenzen ($a \in \mathbb{N}$) der ersten N natürlichen Zahlen unter Verwendung der nach ihm benannten Bernoullischen Zahlen. Offenbar hat Johann Faulhaber (1580–1635) bereits vorher für kleine Werte von a entsprechende Formeln gekannt. In der vorliegenden Arbeit verallgemeinern die beiden Autoren die Bernoullische Formel auf den Fall beliebiger reeller Exponenten $a > -1$. Die von ihnen verwendeten Methoden sind elementar und benutzen als wesentliches Werkzeug die Taylorentwicklung der Potenzfunktion. Die resultierende Summe wird durch die klassische Bernoullische Formel und einen speziellen Wert der Riemannschen Zetafunktion $\zeta(s)$ angenähert. Ist a eine gerade natürliche Zahl, so liegt eine triviale Nullstelle von $\zeta(s)$ vor, und die Annäherung wird zur Gleichheit.

is given by

$$(a+1)\sum_{n=1}^{N} n^a = \sum_{k=0}^{a}(-1)^k\binom{a+1}{k}B_k\,N^{a-k+1}\,. \tag{1}$$

The constants B_k in (1) are the Bernoulli numbers. Here the Bernoulli numbers will be defined by the recurrence

$$B_0 = 1, \qquad \sum_{i=0}^{k}\binom{k+1}{i}B_i = 0, \quad k > 0\,, \tag{2}$$

or equivalently, by the power series

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty}\frac{B_k}{k!}\,x^k\,. \tag{3}$$

(A different convention regarding the Bernoulli numbers is used in [9].) The first five Bernoulli numbers are

$$B_0 = 1, \qquad B_1 = -\tfrac{1}{2}, \qquad B_2 = \tfrac{1}{6}, \qquad B_3 = 0, \qquad B_4 = -\tfrac{1}{30}.$$

Also we have

$$B_{2k+1} = 0\,, \quad \text{for } k = 1, 2, \ldots\,, \tag{4}$$

as is seen by verifying that $x/(e^x - 1) + x/2$ is an even function and using (3). One reference for (1) (from among many possible references) is [2]. Bernoulli's original derivation of (1) can be found on pages 214–216 of [1], the modern translation of *Ars conjectandi*. See [4] for a proof of the equivalence of (2) and (3).

Some authors (e.g. [2], [6]) refer to (1) as Faulhaber's formula in honor of Johann Faulhaber (1580–1635) who studied power sums extensively, publishing his results in his *Academia Algebrae* (1631). The cases of $a = 1, \ldots, 23$ of this formula were known to Faulhaber, but he did not recognize the pattern in the coefficients that eventually gave rise to the Bernoulli numbers (see [3], [5], [8]). Although the general formula (1) is due to Bernoulli, we will nonetheless denote the polynomial on the right-hand side of (1) by $F_a(N)$ in Faulhaber's honor.

Because of its antiquity, it is believed that no original volume of *Academia Algebrae* has ever existed in the Americas. Indeed, the book is rare in Europe as well. Fortunately, A.W.F. Edwards provided Donald Knuth with a photocopy of the volume (once owned by Jacobi) that is in the Cambridge University Library. Knuth has deposited the annotated photocopy in the Mathematical Sciences Library at Stanford University.

While the title of *Academia Algebrae* is in Latin, the text is in German. The style is alien, and the notation, terminology, and typography are unfamiliar. Remarkably, Knuth managed to penetrate these difficulties and to describe Faulhaber's work in [5]. Also in [5], Knuth described the cryptogram in *Academia Algebrae* that was solved to demonstrate that Faulhaber knew the correct formulas for $a = 22$ and $a = 23$. Additionally, this work of Knuth showed that Faulhaber's calculations beyond $a = 23$ were not reliable.

It is natural to ask whether a formula similar to Bernoulli's formula (1) is valid for non-integral powers of a. While instances of approximating formulas for summing non-integral powers of integers have certainly been published (for example, see [5] and [7]), it is perhaps surprising that only recently has it been realized that, with but modest adjustment, Bernoulli's formula (1) generalizes to complex powers with real part greater than -1. In [6] we proved the following result:

Theorem 1 *Suppose $a = \sigma + i\tau$ with $\sigma \geq -1$. If*

$$\begin{aligned} m &= \lfloor \sigma + 1 \rfloor = \min\{k \in \mathbb{Z} : \sigma < k\}\,, \\ \widetilde{\gamma} &= -(m - a)\,, \\ F_a(N) &= \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k \; N^{m-k+1}\,, \end{aligned}$$

then

$$\lim_{N\to\infty} \left[(a+1) \sum_{n=1}^{N} n^a - N^{\widetilde{\gamma}} \, F_a(N) \right] = (a+1)\zeta(-a)\,, \tag{5}$$

where $\zeta(\cdot)$ is the Riemann zeta function. Moreover, if a is a non-negative integer, then the sequence on the left-hand side of (5) *is constant.*

A few words of explanation are needed for the case $a = -1$ which is special: Since the summation on the left-hand side of (5) is multiplied by $a + 1 = 0$, the left-hand side of (5) equals $-N^{-1}N = -1$. The only meaningful interpretation of the right-hand side of (5) is $\lim_{a\to -1}(a+1)\,\zeta(-a)$ and that limit equals -1, because of the fact $\zeta(s)$ has a simple pole with residue 1 at $s = 1$.

Notice that if a is a non-negative integer, then the theorem says

$$(a+1)\sum_{n=1}^{N} n^a = N^{-1} F_a(N) + (a+1)\zeta(-a)\,,$$

which is the classical result since $(a+1)\zeta(-a) = (-1)^a \, B_{a+1}$, for all integers $a \geq 0$, and hence

$$N^{-1} F_a(N) + (a+1)\zeta(-a) = \sum_{k=0}^{a} (-1)^k \binom{a+1}{k} B_k \; N^{a-k+1}\,.$$

On the other hand, when a is non-integral, the polynomial $F_a(N)$ has the same formal appearance as the polynomial on the right-hand side of Bernoulli's formula (1), but it is not equal to any of those polynomials because the binomial coefficients, defined by

$$\binom{a+1}{k} = \frac{(a+1)a\ldots(a+2-k)}{k!}\,,$$

are non-integral.

Also, when a is non-integral, a further refinement of line (5) is possible; namely,

$$(a+1)\sum_{n=1}^{N} n^a - N^{\widetilde{\gamma}}\, F_a(N) = (a+1)\zeta(-a) + O(N^{-\beta})\,, \tag{6}$$

where $\beta = m - \sigma$.

In [6], Theorem 1, and the refinement (6), were proved using Euler-Maclaurin summation and an identity for the Riemann zeta function. For the case in which a is real, one might hope that complex analysis could be avoided and an elementary proof be given. In this paper, we give such a proof.

To facilitate the elementary proof for real a it is convenient to make a slight change of notation and state the result in the following form.

Theorem 2 *Suppose $a > -1$ is real and $a = m + \gamma$ with $m \in \mathbb{Z}$, $-1 < \gamma \le 0$. In terms of the ceiling function, $\lceil\cdot\rceil$, we have $m = \lceil a \rceil$. If*

$$F_a(N) = \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k\, N^{m-k+1}\,,$$

then there exists a real number C_a so that

$$\lim_{N\to\infty}\left[(a+1)\sum_{n=1}^{N} n^a - N^{\gamma}\, F_a(N)\right] = C_a\,. \tag{7}$$

Moreover, if a is a non-negative integer, then the sequence on the left-hand side of (7) *is identically zero.*

When $a > -1$ is non-integral, the value γ from Theorem 2 equals the value $\widetilde{\gamma}$ used in Theorem 1; but, when a is a non-negative integer, we have $\widetilde{\gamma} = \gamma - 1$. While it would be nice to avoid the differing notations in the two theorems, we feel that Theorem 1 has a cleaner statement using $\widetilde{\gamma}$ and Theorem 2 yields a cleaner statement using γ.

Notice that, when a is a non-negative integer, the polynomial $F_a(N)$ in Theorem 2 is identical to the polynomial on the right-hand side of Bernoulli's formula (1) and the factor N^{γ} appearing in (7) is identically equal to 1; in this case Theorem 2 gives the classical result of Bernoulli.

In the final section of this paper, we give an application to summing the pth roots of the first N positive integers with an example computation.

2 Proof of the theorem

Our aim here is to give an elementary proof of Theorem 2. The idea is simple and is as follows: When a is a positive integer, one can prove Bernoulli's formula inductively. In fact, doing so for small integral a is often used to illustrate proof by induction. We will do almost the same inductive process even when a is non-integral. Of course, the inductive

process cannot work for non-integral a exactly as it does for integral a. Instead, when a is non-integral there is an "error" made at each step of the induction. We keep track of the accumulated errors, and their sum equals the constant C_a. Carrying out this process requires calculation and estimation, but it is nonetheless elementary.

The following lemma expresses $N^\gamma F_a(N)$ in terms of $(N+1)^\gamma F_a(N+1)$ and will be the crucial part of our argument.

Lemma 3 *If $a > -1$ is real and m, γ, and $F_a(N)$ are as above, then*

$$(a+1)(N+1)^a - (N+1)^\gamma F_a(N+1) + N^\gamma F_a(N) = \begin{cases} O(N^{\gamma-1}), & \text{if } a \text{ is not an integer,} \\ 0, & \text{if } a \text{ is an integer.} \end{cases}$$

In order to further motivate the lemma, we will first demonstrate how Theorem 2 easily follows.

Proof of Theorem 2. Suppose $a > -1$ is real. First consider the case in which a is not an integer. Define the sequence

$$A_N = (a+1)\sum_{n=1}^{N} n^a - N^\gamma F_a(N),$$

and observe that

$$A_{N+1} - A_N = (a+1)(N+1)^a - (N+1)^\gamma F_a(N+1) + N^\gamma F_a(N). \tag{8}$$

By Lemma 3, the right-hand side of (8) is $O(N^{\gamma-1})$ and therefore the series

$$\sum_{N=1}^{\infty} (A_{N+1} - A_N) \tag{9}$$

is convergent. Since the series (9) telescopes, we see that the sequence A_N is convergent as well. This proves the theorem when a is non-integral.

In case a is an integer, the argument proceeds as above except that now the right-hand side of (8) equals 0. Thus we have

$$A_M - A_1 = \sum_{N=1}^{M-1} (A_{N+1} - A_N) = 0,$$

so

$$(a+1)\sum_{n=1}^{M} n^a - F_a(M) = A_M = A_1.$$

To complete the proof, we must verify that $A_1 = 0$. This equation is trivial when $a = 0$. Otherwise, it follows readily from (4) (i.e., the fact that $B_k = 0$, whenever $k \geq 3$ is odd) as follows: Using $m = a$, (2), and $B_1 = -\frac{1}{2}$, we have

$$\begin{aligned} A_1 &= (a+1) - F_a(1) \\ &= (a+1) - \left[\sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k \right] \\ &= (a+1) - \left[\sum_{k=0}^{m} \binom{a+1}{k} B_k - 2(a+1) B_1 \right] \\ &= 0\,. \end{aligned}$$

□

Proof of Lemma 3. First we apply Taylor's theorem to $f(t) = t^p$. The values of p that we need will be specified later; they are all positive.

Since the hth derivative of f is given by

$$f^{(h)}(t) = p\,(p-1)\,\ldots\,(p-h+1)\,t^{p-h}\,,$$

we have

$$f(x_0 + y) = \sum_{h=0}^{q} \binom{p}{h} x_0^{p-h}\, y^h + \binom{p}{q+1} (x_0 + \xi)^{p-q-1}\, y^{q+1}\,,$$

where ξ is between 0 and y. In particular, if we set $x_0 = N + 1$ and $y = -1$, we obtain

$$N^p = \sum_{h=0}^{q} (-1)^h \binom{p}{h} (N+1)^{p-h} + (-1)^{q+1} \binom{p}{q+1} (N + \xi_{N,p,q})^{p-q-1}\,, \tag{10}$$

where $\xi_{N,p,q}$ is between 0 and 1. For convenience, we will simplify (10) by writing

$$R_{p,q}(N) = (-1)^{q+1} \binom{p}{q+1} (N + \xi_{N,p,q})^{p-q-1}\,,$$

so that

$$N^p = \sum_{h=0}^{q} (-1)^h \binom{p}{h} (N+1)^{p-h} + R_{p,q}(N)\,. \tag{11}$$

Note that

$$R_{p,q}(N) = \begin{cases} 0, & \text{if } p \text{ is a non-negative integer and } p \leq q, \\ O\left(N^{p-q-1}\right), & \text{otherwise.} \end{cases} \tag{12}$$

Next we use the Taylor expansion (11) with $q = m - k + 1$, $p = a - k + 1$, for $k = 0, 1, \ldots, m$, to express $N^\gamma\ F_a(N)$ in terms of $N + 1$:

$$\begin{aligned}
N^\gamma\ F_a(N) &= \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k\ N^{a-k+1} \\
&= \sum_{k=0}^{m} \sum_{h=0}^{m-k+1} (-1)^{k+h} \binom{a+1}{k} B_k \binom{a-k+1}{h} (N+1)^{a-k+1-h} \\
&\quad + \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k\ R_{a-k+1,m-k+1}(N) \\
&= \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k\ (N+1)^{a-k+1} \qquad (13) \\
&\quad + \sum_{k=0}^{m} \sum_{h=1}^{m-k+1} (-1)^{k+h} \binom{a+1}{k} B_k \binom{a-k+1}{h} (N+1)^{a-k+1-h} \qquad (14) \\
&\quad + \sum_{k=0}^{m} (-1)^k \binom{a+1}{k} B_k\ R_{a-k+1,m-k+1}(N)\,. \qquad (15)
\end{aligned}$$

Note that line (13) equals $(N + 1)^\gamma F_a(N + 1)$.

In case a is an integer, we see that line (15) equals 0. In case a is not an integer, we have $R_{p,q}(N) = O(N^{p-q-1})$, so that

$$R_{a-k+1,m-k+1}(N) = O(N^{a-m-1}) = O(N^{\gamma-1})$$

holds, for $k = 0, 1, 2, \ldots, m$, and hence line (15) is $O(N^{\gamma-1})$.

To complete the proof, it suffices to show that line (14) equals

$$-(a+1)(N+1)^a\,.$$

By direct calculation we observe that

$$\binom{a+1}{k}\binom{a-k+1}{h} = \binom{a+1}{k+h}\binom{k+h}{k}.$$

By setting $\ell = k + h$, we may rewrite the summation $\sum_{k=0}^{m} \sum_{h=1}^{m-k+1}$, which appears below, in the form $\sum_{k=0}^{m} \sum_{\ell=k+1}^{m+1}$ and then, by reversing the order of summation, we reorganize it into the form $\sum_{\ell=1}^{m+1} \sum_{k=0}^{\ell-1}$.

This yields:

$$\begin{aligned}
&\sum_{k=0}^{m}\sum_{h=1}^{m-k+1}(-1)^{k+h}\binom{a+1}{k}B_k\binom{a-k+1}{h}(N+1)^{a-k+1-h}\\
&\quad=\sum_{k=0}^{m}\sum_{h=1}^{m-k+1}(-1)^{k+h}\binom{a+1}{k+h}\binom{k+h}{k}B_k\,(N+1)^{a-k+1-h}\\
&\quad=\sum_{\ell=1}^{m+1}\sum_{k=0}^{\ell-1}(-1)^{\ell}\binom{\ell}{k}\binom{a+1}{\ell}B_k\,(N+1)^{a-\ell+1}\\
&\quad=\sum_{\ell=1}^{m+1}(-1)^{\ell}\binom{a+1}{\ell}(N+1)^{a-\ell+1}\left[\sum_{k=0}^{\ell-1}\binom{\ell}{k}B_k\right].
\end{aligned}\tag{16}$$

Directly from definition (2) we have

$$\sum_{k=0}^{\ell-1}\binom{\ell}{k}B_k=\begin{cases}1, & \ell=1\,,\\ 0, & \ell>1\,,\end{cases}$$

and hence line (16) collapses to $-(a+1)(N+1)^a$, as desired. □

3 An application

Let $p \geq 2$ be an integer. As an application of Theorem 1, we consider the problem of summing the pth roots of the first N positive integers. Substituting $a = 1/p$ in (6) and simplifying, we obtain the identity

$$\sum_{n=1}^{N} n^{1/p} = N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)+\zeta(-1/p)+O(N^{(1-p)/p})\,,$$

which we write informally as

$$\sum_{n=1}^{N} n^{1/p} \approx N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)+\zeta(-1/p)\,. \tag{17}$$

Using Euler-Maclaurin summation to approximate ζ and ζ', one can verify that $\zeta(x) \in [-1/2, 0]$ holds for $x \in [-1/2, 0]$ (alternatively, one might reasonably believe the graph that any mathematics software produces). Thus, we are justified in writing

$$\sum_{n=1}^{N} n^{1/p} \approx N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)-\frac{1}{2}\,. \tag{18}$$

Indeed, since $\zeta(0) = -1/2$, the difference between the approximations given by (17) and (18) tends to zero as $p \to \infty$. Using $p = 7$ and $N = 10^6$, we perform some

example computations using the mathematics software SAGE (http://www.sagemath.org). We compute the sum directly using 100 bits of precision (about 30 decimal places), as well as the approximations given by (17) and (18).

$$\sum_{n=1}^{N} n^{1/p} \approx 6\,297\,252\,.\,8505457990916984049647$$
$$348 \text{ seconds}, \qquad \text{error} < 10^{-23},$$

$$N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)+\zeta(-1/p) \approx 6\,297\,252\,.\,8505457134148325714971$$
$$0.61 \text{ seconds}, \qquad \text{error} < 8.5676\cdot 10^{-8},$$

$$N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)-\frac{1}{2} \approx 6\,297\,252\,.\,7371884451801369808624$$
$$0.00 \text{ seconds}, \qquad \text{error} < 0.11336.$$

The direct computation of the sum takes almost 6 minutes, while the computation of (18) is instantaneous and comes within 0.1134 of the true value; the computation of (17) takes 0.61 seconds and comes within $8.5676\cdot 10^{-8}$ of the true value!

As a second demonstration of the power of this method, we use the approximation (17) to perform a computation that is infeasible using the sum directly; namely, we set $p = 7$ and $N = 10^{100}$. Using 500 bits of precision, we find:

$$\begin{aligned}\sum_{n=1}^{N} n^{1/p} &\approx N^{1/p}\Big(\frac{pN}{p+1}+\frac{1}{2}\Big)+\zeta(-1/p)\\ &\approx 1.68936051277284389613119044986016540582133494697373 0\\ &\quad 8154372640299936006500361980618669600742989582681748\\ &\quad 3957153077976238315890719780919249375218697683\cdot 10^{114}\,.\end{aligned} \tag{19}$$

The whole computation takes 0.71 seconds! The error is $O(N^{-6/7})$, so we expect the error in this approximation to be about $8.5676\cdot 10^{-8}\cdot(10^{100-6})^{-6/7} \approx 2.2984\cdot 10^{-88}$. As a result, it is very likely that the approximation (19) is correct to all displayed decimal places! All computations were done on a MacBook with a 2GHz Intel Core Duo processor and 2 GB of memory, running Mac OS 10.5.

References

[1] Bernoulli, J.: *The Art of Conjecturing*. Translated with an introduction and notes by E. Dudley Sylla. The Johns Hopkins University Press, Baltimore 2006.

[2] Conway, J.H.; Guy, R.K.: *The Book of Numbers*. Springer-Verlag, New York 1996, 106.

[3] Faulhaber, J.: *Academia Algebrae. Darinnien die miraculosische Inventiones zu den höchsten Cossen weiters continuirt und profitiert werden*. Johann Ulrich Schönigs, Augsburg 1631.

[4] Ireland, K.F.; Rosen, M.I.: *A Classical Introduction to Modern Number Theory*. Springer-Verlag, New York 1990, 228–230.

[5] Knuth, D.E.: Johann Faulhaber and Sums of Powers. *Math. Comp.* 203 (1993), 277–294.

[6] McGown, K.J.; Parks, H.R.: The Generalization of Faulhaber's Formula to Sums of Non-Integral Powers. *J. Math. Anal. Appl.* 330 (2007), 571–575.

[7] Parks, H.R.: Sums of Non-Integral Powers. *J. Math. Anal. Appl.* 297 (2004), 343–349.

[8] Schneider, I.: *Johannes Faulhaber 1580–1635*. Birkhäuser Verlag, Basel–Boston–Berlin 1993, 131–159.

[9] Whittaker, E.T.; Watson, G.N.: *A Course of Modern Analysis*. 4th ed., Cambridge University Press, 1969.

Kevin J. McGown
Department of Mathematics
University of California at San Diego
La Jolla, CA 92093-0112, USA
e-mail: `kmcgown@math.ucsd.edu`

Harold R. Parks
Department of Mathematics
Oregon State University
Corvallis, OR 97331-4605, USA
e-mail: `parks@math.oregonstate.edu`