# Figurate numbers and sums of numerical powers: Fermat, Pascal, Bernoulli

David Pengelley*

In the year 1636, one of the greatest mathematicians of the early seventeenth century, the Frenchman Pierre de Fermat (1601–1665), wrote to his correspondent Gilles Persone de Roberval (1602–1675) that he could solve "what is perhaps the most beautiful problem of all arithmetic" [3], and stated that he could do this using the following theorem on the *figurate numbers* derived from *natural progressions*, as he first stated in a letter to another correspondent, Marin Mersenne (1588–1648) in Paris:

Pierre de Fermat, from

*Letter to Mersenne. September/October, 1636, and again to Roberval, November 4, 1636*

The last number multiplied by the next larger number is double the collateral triangle;
the last number multiplied by the triangle of the next larger is three times the collateral pyramid;
the last number multiplied by the pyramid of the next larger is four times the collateral triangulo-triangle;
and so on indefinitely in this same manner [3], [12, p. 230f], [6, vol. II, pp. 70, 84–85].

Fermat, most famous for his *last theorem,*[1] worked on many problems, some of which had ancient origins. Fermat had a law degree and spent most of his life as a government official in Toulouse. There are many indications that he did mathematics partly as a diversion from his professional duties, solely for personal gratification. That was not unusual in his day, since a mathematical profession comparable to today's did not exist. Very few scholars in Europe made a living through their research accomplishments. Fermat had one especially unusual trait: characteristically he did not divulge proofs for the discoveries he wrote of to others; rather, he challenged them to find proofs of their own. While he enjoyed the attention and esteem he received from his correspondents, he never showed interest in publishing a book with his results. He never traveled to the centers of mathematical activity, not even Paris, preferring to communicate with the scientific community through an exchange of letters, facilitated by the Parisian theologian Marin Mersenne, who served as a clearinghouse for scientific correspondence from all over Europe, in the absence as yet of scientific research journals. While Fermat made very important contributions to the development

*Mathematical Sciences; Dept. 3MB, Box 30001; New Mexico State University; Las Cruces, NM 88003; davidp@nmsu.edu.

[1] Fermat claimed that the equation $x^n + y^n = z^n$ has no solution in positive whole numbers $x, y, z$ when $n > 2$. One of the greatest triumphs of twentieth-century mathematics was the proof in the 1990s of his famous long-standing claim.

Figure 1: Fermat

of the differential and integral calculus and to analytic geometry [12, Chapters III, IV], his life-long passion belonged to the study of properties of the integers, now known as number theory, and it is here that Fermat has had the most lasting influence on the subsequent course of mathematics. In hindsight, Fermat was one of the great mathematical pioneers, who built a whole new paradigm for number theory on the accomplishments of his predecessors, and laid the foundations for a mathematical theory that would later be referred to as the "queen of mathematics."

What was this "most beautiful problem of all arithmetic" to which Fermat refers in his letter to Roberval? And what are the figurate numbers about which he makes his geometric sounding claims? He writes as if these numbers are themselves geometric figures like triangles and pyramids. In fact they are the numbers that "count" how many equally spaced dots occur in forming discrete versions of these geometric figures, and such figurate numbers had first been studied millennia earlier. Fermat's "most beautiful problem of all arithmetic", which he claimed he could solve using figurate numbers, also had a long history. He was referring to finding formulas for *sums of powers* in a natural progression, i.e., in an arithmetic progression $1, 2, \ldots, n$ whose "last number" is $n$. We shall see shortly what is meant by a sum of powers, and see why formulas for sums of powers became central in mathematics as the ideas of calculus emerged through the seventeenth century.

Both figurate numbers and formulas for sums of powers were important in mathematics long before and after Fermat. Previous and subsequent episodes of our projects on this theme address other eras, and the whole story is told in full in [11]. This project focuses on the intertwined study of the figurate numbers and formulas for sums of powers through the work of three of the greatest seventeenth century mathematicians, Pierre de Fermat, Blaise Pascal (1623–1662), and Jakob Bernoulli (1654–1705). We will see that while figurate numbers are not generally well known by that name today, they are the same numbers that count combination of choices, that appear as coefficients in binomial expansions, and that appear in *Pascal's Triangle*, and are thus still the most important numbers in combinatorial mathematics today.

### Fermat's figurate numbers and sums of powers

Already in the classical Greek era, figurate numbers were of great interest, and formulas for sums of powers were sought for solving area and volume problems. The Pythagoreans, a mysterious group led by Pythagoras in ancient Greece around the sixth century B.C.E., believed that number was the substance of all things, and geometric patterns seen in dots or pebbles in the sand showed them relationships between numbers. For instance, in Figure 2 we see dot pictures for three types of numbers the Pythagoreans probably considered, and we can obtain some formulas for discrete sums from them.

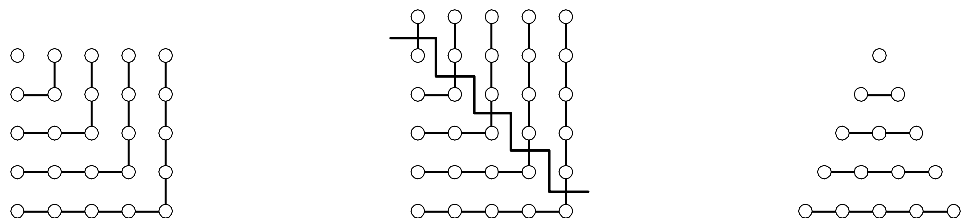


Figure 2: Square, rectangular, and triangular numbers

**Exercise** 1. Using the dot pictures of Figure 2, explain how to obtain summation formulas like

$$\sum_{i=1}^{n}(2i-1)=n^2, \quad \sum_{i=1}^{n}2i=n\,(n+1)\,, \quad \sum_{i=1}^{n}i=\frac{n\,(n+1)}{2},$$

and then prove these formulas by mathematical induction.

We can now also understand Fermat's first claim, if we realize that for the progression $1, 2, \ldots, n$, what Fermat means by the collateral triangle is the triangle with rows of dots whose quantity ranges from 1 through $n$, as in Figure 2. We will call the total number of dots in such a triangle a triangular number.

**Exercise** 2. Interpret and justify Fermat's first claim about the collateral triangle using the results of the previous exercise.

The further figurate numbers, such as Fermat's collateral pyramid, will generalize this by counting dots in analogous higher-dimensional figures. We shall study these shortly, but first we begin the parallel story of sums of powers.

In classical Greek mathematics we find an interest not just in knowing how to sum the terms in a progression, as in the closed formula appearing on the right side of $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. We also find sums like $\sum_{i=1}^{n} i^2$ and $\sum_{i=1}^{n} i^3$, which today we call sums of powers, i.e., the sum of $n$ terms in the progression each raised to a fixed power. These arose in specific determinations of areas and volumes of geometric objects with curved sides, by what is today called the Greek method of exhaustion. For instance, Archimedes of Syracuse (c. 287–212 B.C.E.), perhaps the greatest mathematician of antiquity, determined areas bounded by parabolas and spirals. To see the connection to sums of powers, the reader should recall and review from modern calculus that an area bounded by curves is given by an integral, defined as a limit of sums that approximate the area with sums of areas of rectangles. Recall, for instance, that for the important integral $\int_0^1 x^k dx$, with $k$ an arbitrary fixed natural number, the simplest Riemann sums will involve $\sum_{i=1}^n i^k$, and that to determine the integral as a limiting value of the Riemann sums, one will need some very good knowledge about the size of these sums as $n$ grows. While you may be tempted simply to evaluate this integral by antidifferentiation using the fundamental theorem of calculus, note that neither antidifferentiation nor the fundamental theorem of calculus were even dreamt of by anyone in the world until the second half of the seventeenth century, two millennia after Archimedes and several decades after Fermat's work. However, by Fermat's time in the first half of the seventeenth century, finding formulas for sums of powers $\sum_{i=1}^n i^k$ for the purpose of computing the integrals $\int_0^1 x^k dx$ was a major quest in mathematics. We should therefore not be surprised that Archimedes desired knowledge about sums of squares $\sum_{i=1}^n i^2$, since $\int_0^1 x^2 dx$ gives an area bounded by a parabola, and the reader may check that the same integral arises when finding the area bounded by an Archimedean spiral, a curve described today by $r = a\theta$ in polar coordinates. In Fermat's time the curves $y = x^k$ for $k > 2$ were called *higher parabolas* by analogy.

**Exercise** 3. Carry out an analysis with sums of rectangles to approximate the area $\int_0^1 x^k dx$ of the region under the curve $y = x^k$ by an expression involving $\sum_{i=1}^n i^k$, and state what limit needs to be computed to determine the area. Specialize to the case $k = 2$ to specify what limit will give the area under the parabola.

Archimedes actually solved his area problem for the parabola by two incredibly clever methods that did not require using a sum of squares, and the reader may see these in [10]. But he found the area bounded by a turn of his spiral by discovering, proving, and utilizing a result about the sum of squares, $\sum_{i=1}^n i^2$, within the Greek method of exhaustion. We will not show his geometric viewpoint or proof of his result on a sum of squares, but will state his result here in modern numerical terminology:

$$(n+1)\,n^2 + \sum_{i=1}^n i = 3\sum_{i=1}^n i^2.$$

**Exercise** 4. Use this equality of Archimedes, and the earlier formula for the sum of terms in the progression, to find an explicit polynomial formula for $\sum_{i=1}^n i^2$.

**Exercise** 5. Now use the explicit polynomial formula you have for $\sum_{i=1}^n i^2$ to compute the necessary limit to obtain the value of $\int_0^1 x^2 dx$.

**Exercise** 6. Show from the polynomial formula for $\sum_{i=1}^n i^2$ that

$$\frac{n^3}{3} < \sum_{i=1}^n i^2 < \frac{(n+1)^3}{3},$$

and determine the value of $\int_0^1 x^2 dx$ directly from these inequalities without needing the precise formula for $\sum_{i=1}^n i^2$.

**Exercise** 7. Roberval wrote to Fermat that he could find the areas under all the higher parabolas (which we would label as the integrals $\int_0^1 x^k dx$ for $k > 2$) using the inequalities

$$\frac{n^{k+1}}{k+1} < \sum_{i=1}^n i^k < \frac{(n+1)^{k+1}}{k+1},$$

which he claimed hold. Indeed, calculate $\int_0^1 x^k dx$ by considering lower and upper sums of rectangles based on left and right endpoints of equally spaced partitions of the interval, and by using Roberval's inequalities to compute the appropriate limit. Fermat went further with sums of powers, claiming he could find explicit formulas for them.

Let us now reconnect our exploration of sums of powers to figurate numbers by seeing how we can obtain the polynomial formula for $\sum_{i=1}^n i^2$ from Fermat's second claim about figurate numbers. This was exactly the viewpoint expressed by Fermat, that the sums of powers problem can be solved by understanding figurate numbers. To do this, consider Fermat's second claim, that "the last number multiplied by the triangle of the next larger is three times the collateral pyramid". We need to know what he meant by the pyramid collateral to the number $n$. Analogous to the collateral triangle, by the collateral pyramid Fermat means to count the dots in a three-dimensional triangular pyramid with sides each having $n$ dots (Figure 3).

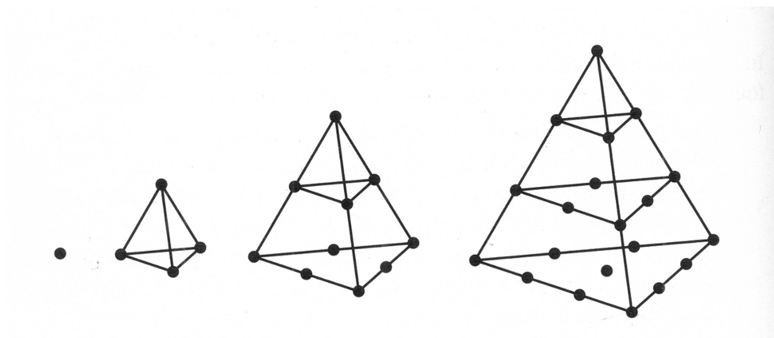


Figure 3: Pyramidal numbers

We can now apply Fermat's claim, by thinking first about the geometry of a triangular pyramid with $n$ dots on each side. We see that its vertical layers consist of $n$ triangular numbers stacked on top of each other, with side lengths from 1 to $n$. Since we know from before that the triangular number with side length $i$ has $\frac{i(i+1)}{2}$ dots, we see that altogether the pyramid has $\sum_{i=1}^n \frac{i(i+1)}{2}$ dots. Then since the "triangle of the next larger" (number than $n$) must have $\frac{(n+1)(n+2)}{2}$ dots, we see that Fermat's claim can be expressed as

$$n\frac{(n+1)(n+2)}{2} = 3\sum_{i=1}^{n} \frac{i(i+1)}{2}$$

Now notice that we can separate the right side via

$$\sum_{i=1}^{n} \frac{i(i+1)}{2} = \sum_{i=1}^{n} \frac{i^2}{2} + \sum_{i=1}^{n} \frac{i}{2},$$

and should now be able to solve for an explicit formula for $\sum_{i=1}^n i^2$, since we already have an explicit formula for $\sum_{i=1}^n i$.

**Exercise** 8. Solve for an explicit formula for $\sum_{i=1}^n i^2$ using Fermat's claim, and check to see if it is the same formula that you found in a previous exercise from the equality of Archimedes.

The previous exercise performs exactly what Fermat had in mind when he said he could solve the sums of powers problem using his claims about figurate numbers.

**Exercise** 9. We have seen how to obtain the formula for a sum of squares from the equality of Archimedes, and also from Fermat's second claim about figurate numbers. But so far we have proofs for neither of these precursors, so just for good measure, prove the formula for sums of squares by mathematical induction.

Moving towards higher powers and higher figurate numbers, we note here that formulas for sums of cubes and sums of fourth powers were also known centuries before Fermat, and some work towards general methods had been achieved. To give a flavor of the diversity of sources, knowledge of a formula for sums of cubes appears implicitly in the work of the neo-Pythagorean Nicomachus of Gerasa in the first century C.E., and explicitly in verse by Āryabhaṭa in India written in 499 C.E., along with a proof by the Islamic mathematician Abū Bakr al-Karajī (c. 1000 C.E.) of the House of Wisdom established in Baghdad in the ninth century. The Egyptian mathematician Abū 'Alī al-Ḥasan ibn al-Haytham (965–1039) gives us the first steps along a path toward understanding these formulas in general, in which he finds formulas for sums of fourth powers in order to find the volume of a general paraboloid of revolution (in contemporary terms this involves integrating $x^4$). His complicated method could in principle be generalized to higher powers. These developments are explored in a precursor project to this one. What we do not see in these earlier works is a connection between the figurate numbers and sums of powers. This is what Fermat brought to the scene, and to which we return.

**Exercise** 10. Guess a formula for sums of cubes: First calculate the first six sums and look for a pattern. Then prove by mathematical induction that your guess is correct.

**Exercise** 11. Above we showed how Fermat's second claim about figurate numbers could be used to derive a formula for a sum of squares. Generalize this approach to obtain the formula for a sum of cubes from Fermat's third claim. Of course a *collateral triangulo-triangle* is the generalization of a pyramid to a four-dimensional object consisting of $n$ stacked triangular pyramids with side lengths ranging from 1 to $n$. Discuss what would be involved in carrying this to even higher powers.

**Exercise** 12. From Fermat's claim above, derive a formula for the number of dots in a 3-dimensional triangular pyramid with $n$ dots on a side.

**Exercise** 13. From Fermat's final claim above, derive a formula for the number of dots in a 4-dimensional triangulo-triangle with $n$ dots on a side. Generalize to conjecture formulas for higher dimensional 'pyramids'.

Since Fermat did not reveal his methods, it is up to us to decipher what he may have been thinking when he made his claims about figurate numbers, and to see if they are true. We will be able to do this by studying the figurate numbers and discovering their agreement with the numbers in the "arithmetical triangle"[2] (Figure 4) of his contemporary and correspondent Blaise Pascal. The numbers in the arithmetical triangle have yet other extremely important properties in combinatorial mathematics, namely in their roles as combination numbers and binomial coefficients.

Fermat's claims are not merely geometric interpretations of relationships between certain whole numbers; they are actually discrete analogues of results we already know about continuous objects like areas and volumes. For instance, his first claim is a discrete version of the fact that for any line segment, the product of the segment with itself creates a rectangle with area twice that of the triangle with that side and height. Fermat's use of the "next larger number" for one of the discrete

[2] Today called Pascal's triangle.

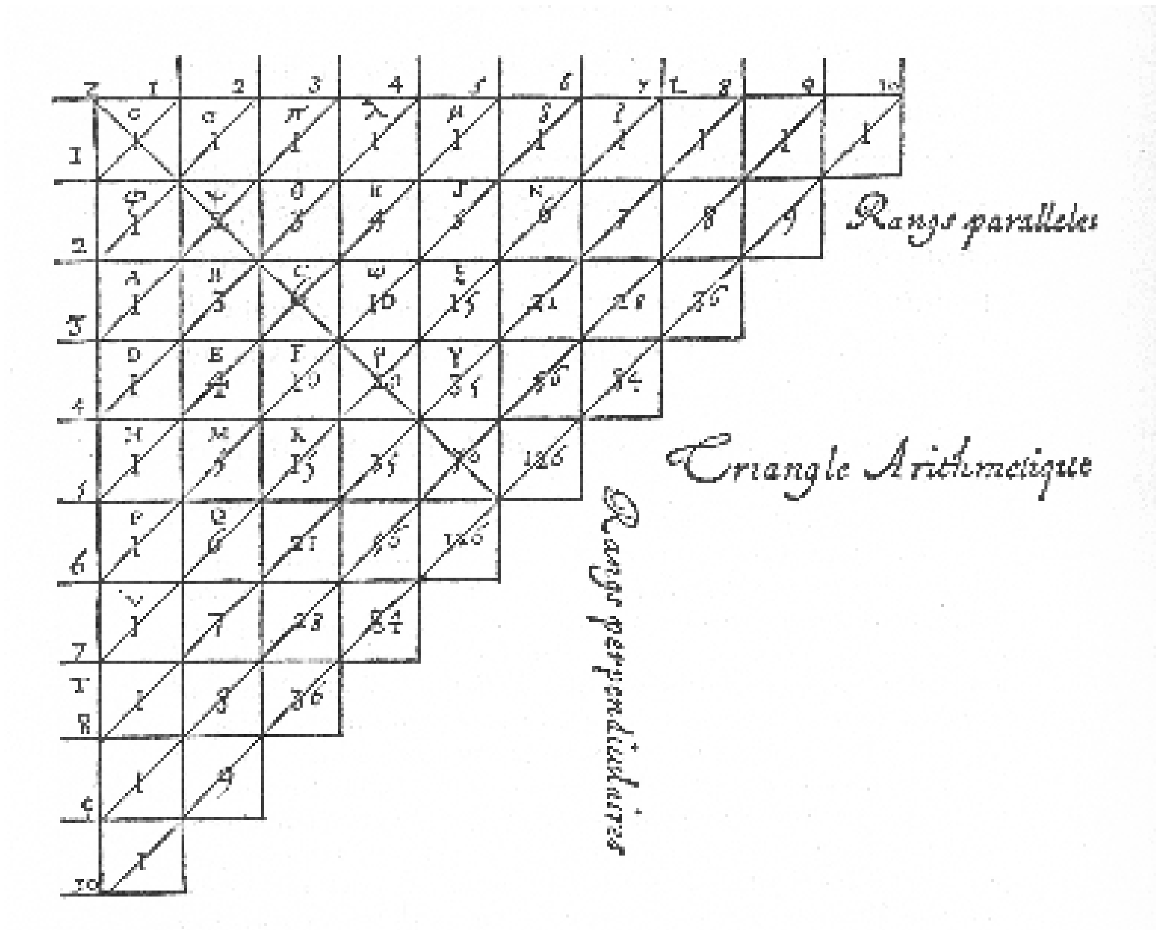



Figure 4: Pascal's *arithmetical triangle*

lengths is exactly what is necessary to ensure a precise two-to-one ratio of the discrete count of dots in a rectangular figure to the count of dots in an inscribed triangle at the corners, as we know happens for the continuous measurements of the analogous areas. This is visible geometrically in the packing of dots in the middle of Figure 2. Similarly, when he says "the last number multiplied by the triangle of the next larger is three times the collateral pyramid," he is expressing a delicate discrete version of the fact that a three-dimensional pyramid on a triangular base has volume one-third that of the prism created with the same base and height. Fermat claims that similar relationships hold in all dimensions.

Let us formally define the *figurate numbers* by their recursive *piling up* property. Here $F_{n,j}$ will denote the figurate number that is $j$-dimensional with $n$ dots along each side. Call $n$ its *side length.* For instance, $F_{3,2}$ is the planar triangular number with 3 dots per side, so by counting up its rows of dots, $F_{3,2} = 1 + 2 + 3 = 6$. The piling up property, by which we define larger figurate numbers from smaller ones, is encoded in the formula $F_{n+1,j+1} = F_{n,j+1} + F_{n+1,j}$. This formalizes the idea that to increase the side length of a $(j + 1)$-dimensional figurate number from $n$ to $n + 1$, we simply add on another layer at its base, consisting of a $j$-dimensional figurate number of side length $n + 1$ (Figure 3). This defining formula is called a *recursion relation*, since it defines each of the figurate numbers in terms of those of lower dimension and side length, provided we start correctly by specifying those numbers with the smallest dimension and those with the smallest side length. While we could start from dimension one, it is useful to begin with zero-dimensional figurate numbers, all of which we will define to have the value one. We thus define

$$\begin{aligned} F_{n,0} &= 1 \qquad (n \geq 1), \\ F_{1,j} &= 1 \qquad (j \geq 0), \\ F_{n+1,j+1} &= F_{n+1,j} + F_{n,j+1} \qquad (n \geq 1, j \geq 0). \end{aligned}$$

We must check that our starting data determine just what was intended for Fermat's figurate numbers. We have required that any figurate number with side length one will have exactly one dot, no matter what its dimension. And the recursion relation clearly produces the desired one-dimensional figurate numbers $F_{n,1} = n$ for all $n \geq 1$ from the starting data of zero-dimensional numbers.

Now Fermat's claimed formulas to Mersenne assume the form

$$\begin{aligned} nF_{n+1,1} &= 2F_{n,2}, \\ nF_{n+1,2} &= 3F_{n,3}, \\ nF_{n+1,3} &= 4F_{n,4}, \\ &\;\;\ldots \end{aligned}$$

**Exercise** 14. Explain why Fermat's claimed formulas to Mersenne assume the form

$$\begin{aligned} nF_{n+1,1} &= 2F_{n,2}, \\ nF_{n+1,2} &= 3F_{n,3}, \\ nF_{n+1,3} &= 4F_{n,4}, \\ &\;\;\ldots . \end{aligned}$$

We shall study the figurate numbers further to see why Fermat's claim about them is true, and simultaneously prepare the groundwork for reading Pascal's work on sums of powers. The reader may have noticed that the individual figurate numbers seem familiar, from looking at Pascal's arithmetical triangle (Figure 4). The numbers shown there seem to match the figurate numbers, i.e., $F_{n,j}$ appears in Pascal's "parallel row" (horizontal) labeled $n$ and "perpendicular row" (vertical) labeled $j+1$. Indeed, in his *Treatise on the Arithmetical Triangle*, Pascal defined the numbers in the triangle by starting the process off with a 1 in the corner, and defined the rest simply by saying that each number is the sum of the two numbers directly above and directly to the left of it, which corresponds precisely to the recursion relation and starting data by which we formally defined the figurate numbers. So they must agree.

Perhaps you already also recognize the numbers in the arithmetical triangle as binomial coefficients or combination numbers. The numbers occurring along Pascal's ruled diagonals in Figure 4 appear to be the coefficients in the expansion of a binomial; for instance, the coefficients of $(a+b)^4 = 1a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + 1b^4$ occur along the diagonal Pascal labels with 5. Our modern notation for these binomial coefficients is that $\binom{m}{j}$ denotes the coefficient of $a^{m-j}b^j$ in the expansion of $(a+b)^m$. Indeed, if in the arithmetical triangle we index both the diagonals and their individual entries beginning with zero (rather than one, as Pascal does), then the entry in diagonal $m$ at column $j$ will be the binomial coefficient $\binom{m}{j}$. This is easy to prove, by showing that they, too, satisfy the recursion relation and starting data of the arithmetical triangle.

**Exercise** 15. Show that the coefficients in the expansion of a binomial satisfy the starting data and the recursion relation of the arithmetical triangle. In other words, if for all $m \geq 0$ we write $(a+b)^m = \sum_{j=0}^{m} \binom{m}{j} a^{m-j}b^j$, show that these coefficients satisfy the starting data $\binom{m}{0} = \binom{m}{m} = 1$, and Pascal's recursion relation $\binom{m+1}{j} = \binom{m}{j} + \binom{m}{j-1}$. (Hint: write $(a+b)^{m+1} = (a+b)\,(a+b)^m$.)

Since we have now identified both the figurate numbers and the binomial coefficients as the numbers in the arithmetical triangle generated by the basic recursion relation, their precise relationship follows just by comparing their indexing: $F_{i+1,j}$ is the number in Pascal's arithmetical triangle in the row and column he labels $i+1$ and $j+1$, which we reindexed (starting with zero) to be labeled as $i$ and $j$, which is in diagonal $i+j$ and column $j$, so it is the binomial coefficient $\binom{i+j}{j}$. We have

$$F_{i+1,j} = \binom{i+j}{j} \quad \text{for } i, j \geq 0.$$

We can also verify a closed formula in terms of factorials for the numbers in the triangle:

$$\binom{m}{j} = \frac{m!}{j!(m-j)!} = \frac{m(m-1)\cdots(m-j+1)}{j!} \quad \text{for } 0 \leq j \leq m,$$

where the notation $i!$ (read "$i$ factorial") is defined to mean

$$i \cdot (i-1) \cdot (i-2) \cdots 3 \cdot 2 \cdot 1,$$

and $0!$ is defined to be 1.

**Exercise** 16. Show that

$$\binom{m}{j} = \frac{m!}{j!(m-j)!} \quad \text{for } 0 \leq j \leq m.$$

(Hint: Show that this factorial formula satisfies the starting data and the recursion relation of the arithmetical triangle.)

The ubiquitous numbers in the arithmetical triangle had already been in use for over 500 years, in places ranging from China to the Islamic world, before Pascal developed and applied its properties in his *Traité du Triangle Arithmétique* (Treatise on the Arithmetical Triangle), written by 1654 [9]. A key fact, which Pascal called his *Twelfth Consequence*, is that neighboring numbers along a diagonal in the triangle are always in a simple ratio:

$$(m-j)\binom{m}{j} = (j+1)\binom{m}{j+1} \quad \text{for } j < m,$$

which is easily obtained from the factorial formula above.

**Exercise** 17. Prove Pascal's Twelfth Consequence from the factorial formula for binomial coefficients.

Translated into figurate numbers, Pascal's Twelfth Consequence reemerges as precisely Fermat's claim in his letter to Mersenne! For instance, letting $j = 2$ and $m = n + 2$ yields

$$n\binom{n+2}{2} = 3\binom{n+2}{3},$$
$$\text{or} \quad nF_{n+1,2} = 3F_{n,3},$$

exactly Fermat's claim that "the last number multiplied by the triangle of the next larger is three times the collateral pyramid." The reader may now easily confirm Fermat's general claim.

**Exercise** 18. State Fermat's general claim about figurate numbers, and prove it from Pascal's Twelfth Consequence.

We have achieved our goal of understanding and verifying Fermat's claims about figurate numbers, although we cannot know that this is precisely how Fermat thought of things. In the process we have connected the figurate numbers to the arithmetical triangle and binomial coefficients, and found a factorial formula for them. All this will prove to be a powerful shift.

We have also seen how Fermat could use his knowledge of figurate numbers to find formulas for sums of powers, at least for squares and cubes. While it is clear that the process can be continued indefinitely, it quickly becomes impractically complicated, and it is also not clear that it yields any new general insight about sums of powers. However, Pascal too wrote about sums of powers, using the binomial coefficient meaning of the figurate numbers to great advantage, and this is where we continue.

### Pascal's Arithmetical Triangle, binomials, and sums of powers of arithmetic progressions

Pascal wrote two treatises of interest to us at around the same time. In his *Treatise on the Arithmetical Triangle* [8, v. 30], Pascal made a systematic study of the numbers in his triangle, simultaneously encompassing their figurate, combinatorial, and binomial roles. Although these numbers had emerged in the mathematics of several cultures over many centuries [9], Pascal was the first to connect binomial coefficients with combinatorial coefficients in probability.

A major motivation for Pascal's treatise was a question from the beginnings of probability theory, about the equitable division of stakes in an interrupted game of chance. The question had been posed to Pascal around 1652 by Antoine Gombaud, the Chevalier de Méré, who wanted to improve his chances at gambling: Suppose two players are playing a fair game, to continue until one player wins a certain number of rounds, but the game is interrupted before either player reaches the winning number. How should the stakes be divided equitably, based on the number of rounds each player has won [9, p. 431, 451ff]? The solution requires the combinatorial properties inherent in

Figure 5: Blaise Pascal

the numbers in the arithmetical triangle, as Pascal demonstrated in his *Treatise*, since they count the number of ways various occurrences can combine to produce a given result.

Blaise Pascal (1623–1662) was born in Clermont-Ferrand, in central France. Even as a teenager his father introduced him to meetings of Marin Mersenne's circle of mathematical discussion in Paris. He quickly became involved in the development of projective geometry, the first in a sequence of highly creative mathematical and scientific episodes in his life, punctuated by periods of religious fervor. Around age twenty-one he spent several years developing a mechanical addition and subtraction machine, in part to help his father in tax computations as a local administrator. It was the first of its kind ever to be marketed. Then for several years he was at the center of investigations of the problem of the vacuum, which led to an understanding of barometric pressure. In fact, the scientific unit of pressure is named the *pascal.* He is also known for Pascal's law on the behavior of fluid pressure.

Around 1654 Pascal conducted his studies on the arithmetical triangle and its relationship to probabilities. His correspondence with Fermat in that year marks the beginning of probability theory. Several years later, Pascal refined his ideas on area problems via the method of indivisibles already being developed by others, and solved various problems of areas, volumes, centers of gravity, and lengths of curves.[3] After only two years of work on the calculus of indivisibles, Pascal fell gravely ill, abandoned almost all intellectual work to devote himself to prayer and charitable work, and died three years later at age thirty-nine. In addition to his work in mathematics and physics, Pascal is prominent for his *Provincial Letters* defending Christianity, which gave rise to his posthumously published *Pensées* on religious philosophy [7, 5]. Pascal was an extremely complex person, and one of the outstanding scientists of the mid-seventeenth century, but we will never know how much

[3] Later in the seventeenth century, Gottfried Wilhelm Leibniz (1646–1716), one of the two inventors of the infinitesimal calculus, which supplanted the method of indivisibles, explicitly credited Pascal's approach as stimulating his own ideas on the so-called characteristic triangle of infinitesimals in his fundamental theorem of calculus.

more he might have accomplished with more sustained efforts and a longer life.

The relation of the arithmetical triangle to counting combinations, and thus their centrality in probability theory, follows easily from the factorial formula above for the triangle's numbers. The reader may verify that $\binom{m}{j}$ represents the number of different combinations of $j$ elements that can occur in a set of $m$ elements

**Exercise** 19. Prove that the number of distinct five-card hands possible from a standard deck of fifty-two playing cards is $\binom{52}{5}$. Then prove that $\binom{m}{j}$ is the number of different combinations of $j$ elements that can occur in a set of $m$ elements. Hint: First see why

$$m\,(m-1)\cdots(m-j+1)$$

is the number of different ways of selecting a sequence of $j$ elements from a sequence with $m$ elements (where different orderings of the same elements count as different sequences).

We have seen that the numbers in the arithmetical triangle have three interchangeable interpretations: as figurate numbers, combination numbers, and binomial coefficients. Given this multifaceted nature, it is no wonder that they arose early on, in various manners and parts of the world, and that they are ubiquitous today. The arithmetical triangle in fact overflows with fascinating patterns.

**Exercise** 20. (Optional) Prove that the sum of the numbers in each diagonal of the arithmetical triangle is a power of two. Hint: binomial theorem.

In about the same year as his *Treatise on the Arithmetical Triangle*, Pascal produced the text we will now study, *Potestatum Numericarum Summa* (Sums of Numerical Powers) [13, v. III, pp. 341–367], which analyzes sums of powers of an arithmetic progression (a sequence of numbers with a fixed difference between each term and its successor) in terms of the numbers in the arithmetical triangle, interpreted as binomial coefficients. Pascal also makes the connection between these results and area problems via the method of indivisibles.

Fermat's great enthusiasm in 1636 for the problem of calculating sums of powers was not immediately embraced by others, and Pascal, although a direct correspondent of Fermat's, was apparently unaware of Fermat's work when he made his own analysis about eighteen years later in *Sums of Numerical Powers*. Here we will see the transition from a geometric to algebraic approach almost complete, since Pascal, unlike others before him, and even Fermat, is bent on presenting a generalized arithmetic solution for the problem, albeit still using mostly verbal descriptions of formulas, instead of modern algebraic notation. We will find that Pascal obtains a compact formula directly relating sums of powers for various exponents, using binomial coefficients as the intermediary.

## Blaise Pascal, from
## *Sums of Numerical Powers*

### **Remark**.

Given, starting with the unit, some consecutive numbers, for example $1, 2, 3, 4$, one knows, by the methods the Ancients made known to us, how to find the sum of their squares, and also the sum of their cubes; but these methods, applicable only to the second and third degrees, do not extend to higher degrees. In this treatise, I will teach how to calculate not only the sum of squares and of cubes, but also the sum of the fourth powers and those of higher powers up to infinity: and that, not only for a sequence of consecutive numbers beginning with the unit, but for a sequence beginning with any number, such as the sequence $8, 9, 10, \ldots$. And I will not restrict myself to the natural sequence of numbers: my method will apply also to a progression having as ratio [difference] $2, 3, 4$, or any other number,—that is to say to a sequence of numbers different by two units, like $1, 3, 5, 7, \ldots$, $2, 4, 6, 8, \ldots$, or differing by three units like $1, 4, 7, 10, 13, \ldots$. And what is more, whatever the first term in the sequence may be: if the first term is $1$, as in the sequence with ratio three, $1, 4, 7, 10, \ldots$: or if it is another term in the progression, as in the sequence $7, 10, 13, 16, 19$; or even if it is alien to the progression, as in the sequence with ratio three, $5, 8, 11, 14, \ldots$ beginning with $5$. It is remarkable that a single general method will suffice to treat all these different cases. This method is so simple that it will be explained along several lines, and without the preparation of algebraic notations to which difficult demonstrations must have recourse. One can judge this after having read the following problem.

### **Definition**.

Consider a binomial $A + 3$, whose first term is the letter $A$, and the second a number: raise this binomial to any power, the fourth for example, which gives

$$A^4 + 12.A^3 + 54.A^2 + 108.A + 81;$$

the numbers $12, 54, 108$, which are multiplying the different powers of $A$, and are formed by the combination of the figurate numbers with the second term, $3$, of the binomial, will be called the *coefficients* of $A$.

Thus, *in the cited example*, $12$ *will be the coefficient of the cube* $A$; $54$, *that of the square; and* $108$, *that of the first power*.

As for the number $81$, it will be called the *pure number*.

### **Lemma**

Suppose any number, like $14$, be given, and a binomial $14+3$, whose first term is $14$ and the second any number $3$, in such a manner that the difference of the numbers $14$ and $14 + 3$ will equal $3$. Let us raise these numbers to a same power, the fourth for example: the fourth power of $14$ is $14^4$, that of the binomial, $14 + 3$, is

$$14^4 + 12.14^3 + 54.14^2 + 108.14 + 81.$$

*In this expression, the powers of the first term,* $14$, *of the binomial are obviously affected by the same coefficients as the powers of* $A$ *in the expansion of* $(A + 3)^4$. This put down, the difference of the two fourth powers, $14^4$ and

$$14^4 + 12.14^3 + 54.14^2 + 108.14 + 81,$$

is $12.14^3 + 54.14^2 + 108.14 + 81$; this difference comprises: on the one hand, the powers of $14$ whose degree is less than the proposed degree $4$, these powers being affected by the coefficients which the same powers of $A$ have in the expansion of $(A + 3)^4$; on the other hand, the number $3$ (*the difference of the proposed numbers*) raised to the *fourth* power [because the *absolute number* $81$ is the *fourth* power of the number 3]. From this we deduce the *following Rule*:

**The difference of like powers of two numbers comprises: the difference of these numbers raised to the proposed power; plus the sum of all the powers of lower degree of the smaller of the two numbers, these powers being respectively multiplied by the coefficients which the**

**same powers of $A$ have in the expansion of a binomial raised to the proposed power and having as first term $A$ and as second term the difference of the given numbers**.

Thus, the difference of $14^4$ and $11^4$ will be

$$12.11^3 + 54.11^2 + 108.11 + 81,$$

since the difference of the first powers is $3$. And so forth.

**A single general method for finding the sum of like powers of the terms of any progression**.

**Given, beginning with any term, any sequence of terms of an arbitrary progression, find the sum of like powers of these terms raised to any degree**.

*Suppose an arbitrary number* $5$ *is chosen as the first term of a progression whose ratio [difference], arbitrarily chosen, is for example three; consider, in this progression, as many of the terms as one wishes, for instance the terms* $5, 8, 11, 14$, *and raise these terms to any power, suppose to the cube. The question is to find the sum of the cubes* $5^3 + 8^3 + 11^3 + 14^3$.

**These cubes are** $125, 512, 1331, 2744$; **and their sum is** $4712$. **Here is how one finds this sum**.

**Let us consider the binomial $A + 3$ having as first term $A$ and as second term the difference of the progression**.

**Raise this binomial to the fourth power, the power immediately higher than the proposed degree three; we obtain the expression**

$$A^4 + 12.A^3 + 54.A^2 + 108.A + 81.$$

*This admitted, we consider the number* $17$, *which, in the proposed progression, immediately follows the last term considered,* $14$. *We take the* fourth *power of* $17$, *known as* $83521$, *and subtract from it:*

First: *the sum* $38$ *of the terms considered,* $5 + 8 + 11 + 14$, *multiplied by the number* $108$ *which is the coefficient of* $A$;

Second: *the sum of the squares of the same terms* $5, 8, 11, 14$, *multiplied by the number* $54$, *which is the coefficient of* $A^2$.

*And so on, in case one still has the powers of $A$ of lesser degree than the proposed degree* three.

*With these subtractions made, one subtracts also the* fourth *power of the first term proposed,* $5$.

*Finally one subtracts the number* $3$ *(ratio [difference] of the progression) itself raised to the* fourth *power and taken as many times as one considers terms in the progression, here* four *times.*

*The remainder of the subtraction will be a multiple of the sum sought; it will be the product of this sum with the number* $12$, *which is the coefficient of* $A^3$, *that is to say the coefficient of the term* $A$ *raised to the proposed power* three.

Thus, in practice, one must form the fourth power of $17$, being $83521$, then subtracting from it successively:

*First*, the sum of the terms proposed, $5 + 8 + 11 + 14$, being $38$, multiplied by $108$,—that is, the product $4104$;

*Then* the sum of the squares of the same terms, $5^2 + 8^2 + 11^2 + 14^2$, or $25 + 64 + 121 + 196$, or again $406$, which, multiplied by $54$, gives $21924$;

*Then* the number $5$ to the *fourth* power, which is $625$;

*Finally* the number $3$ to the *fourth* power, being $81$, multiplied by *four*, which gives $324$. In summary one must subtract the numbers $4104$, $21924$, $625$, $324$, whose sum is $26977$. Taking this sum away from $83521$, there remains $56544$.

The remainder thus obtained is equal to the sum sought, $4712$, multiplied by $12$; and, in fact, $4712$ multiplied by $12$ equals $56544$.

The rule is, as one sees, easy to apply. Here now is how one proves it.

The number 17 raised to the *fourth* power, which one writes $17^4$, is equal to

$$17^4 - 14^4 + 14^4 - 11^4 + 11^4 - 8^4 + 8^4 - 5^4 + 5^4.$$

In this expression, *only the term $17^4$ appears with the single sign $+$; the other terms are in turns added and subtracted.*

But the difference of the terms 17 and 14 is 3; likewise the difference of the terms 14 and 11, and of the terms 11 and 8, and of the terms 8 and 5. Thenceforth, according to our preliminary lemma: $17^4 - 14^4$ equals $12.14^3 + 54.14^2 + 108.14 + 81$.

Likewise $14^4 - 11^4$ equals $12.11^3 + 54.11^2 + 108.11 + 81$.

Likewise $11^4 - 8^4$ equals $12.8^3 + 54.8^2 + 108.8 + 81$.

Likewise $8^4 - 5^4$ equals $12.5^3 + 54.5^2 + 108.5 + 81$.

The term $5^4$ does not need to be transformed.

One then finds as the value of $17^4$:

$$\begin{aligned}
&12.14^3 + 54.14^2 + 108.14 + 81\\
&+ 12.11^3 + 54.11^2 + 108.11 + 81\\
&+ 12.8^3 + 54.8^2 + 108.8 + 81\\
&+ 12.5^3 + 54.5^2 + 108.5 + 81\\
&+ 5^4,
\end{aligned}$$

or, *on interchanging the order of the terms*:

$$\begin{aligned}
&5 + 8 + 11 + 14 \text{ multiplied by } 108,\\
&+ 5^2 + 8^2 + 11^2 + 14^2 \text{ multiplied by } 54,\\
&+ 5^3 + 8^3 + 11^3 + 14^3 \text{ multiplied by } 12,\\
&+ 81 + 81 + 81 + 81\\
&+ 5^4.
\end{aligned}$$

If therefore one subtracts on both sides the sum:

$$\begin{aligned}
&5 + 8 + 11 + 14 \text{ multiplied by } 108,\\
&+ 5^2 + 8^2 + 11^2 + 14^2 \text{ multiplied by } 54,\\
&+ 81 + 81 + 81 + 81\\
&+ 5^4;
\end{aligned}$$

There remains $17^4$ diminished by the previously known quantities:

$$\begin{aligned}
&- 5 - 8 - 11 - 14 \text{ multiplied by } 108,\\
&- 5^2 - 8^2 - 11^2 - 14^2 \text{ multiplied by } 54,\\
&- 81 - 81 - 81 - 81\\
&- 5^4;
\end{aligned}$$

which will be found equal to the sum $5^3 + 8^3 + 11^3 + 14^3$ multiplied by $12$. Q.E.D.

One may thus present as follows the statement and the general solution of the proposed problem.

### The sum of powers

**Given, beginning with any term, any sequence of terms of an arbitrary progression, find the sum of like powers of these terms raised to any degree**.

*We form a binomial having $A$ as its first term, and for its second term the difference of the given progression; we raise this binomial to the degree immediately higher than the proposed degree, and we consider the coefficients of the various powers of $A$ in the expansion obtained.*

*Now we raise to the same degree the term that, in the given progression, immediately follows the last term considered. Then we subtract from the number obtained the following quantities*:

First: *The first term given in the progression,—that is, the smallest of the given terms,—itself raised to the same power (immediately higher than the proposed degree).*

Second: *The difference of the progression, raised to the same power, and taken as many times as of the terms considered in the progression.*

Third: *The sums of the given terms, raised to the various degrees less than the proposed degree, these sums being respectively multiplied by the coefficients of the same powers of $A$ in the expansion of the binomial formed above.*

*The* remainder *of the subtraction thus accomplished is a multiple of the sum sought: it contains it as many times as unity is contained in the coefficient of the power of $A$ whose degree is equal to the proposed degree.*

### NOTE

The reader himself will deduce practical rules that apply in each particular case. Suppose, for example, that one wishes to find the sum of a certain number of terms in the natural sequence [i.e., of natural numbers] beginning with an arbitrary number: here is the rule that one deduces from our general method:

**In a natural progression beginning with any number, the square of the number immediately above the last term, diminished by the square of the first term and the number of terms given, is equal to double the sum of the stated terms**.

Suppose given a sequence of any consecutive numbers whose first term is arbitrary, for example the *four* numbers $5, 6, 7, 8$: I say that $9^2 - 5^2 - 4$ equals the double of $5 + 6 + 7 + 8$.

One will easily obtain analogous rules giving the sums of powers of higher degrees and which apply to all progressions.

### Conclusion.

Any who are a little acquainted with the doctrine of *indivisibles* will not fail to see what profit one may make from the preceding results for the determination of curvilinear areas. These results permit the immediate squaring of all types of parabolas and an infinity of other curves.

If then we extend to continuous quantities the results found for numbers, by the method expounded above, we will be able to state the following rules:

### Rules relating to the natural progression beginning with unity.

**The sum of a certain number of lines is to the square of the largest as** $1$ **is to** $2$.
**The sum of the squares of the same lines is to the cube of the largest as** $1$ **is to** $3$.
**The sum of their cubes is to the fourth power of the largest as** $1$ **is to** $4$.

**General rule relating to the natural progression beginning with unity.**

**The sum of like powers of a certain number of lines is to the immediately greater power of the largest among them as unity is to the exponent of this same power**.

I will not pause here for the other cases, because this is not the place to study them. It will be enough for me to have cursorily stated the preceding rules. One can discover the others without difficulty by relying on the principle that *one does not increase a continuous magnitude when one adds to it, in any number one wishes, magnitudes of a lower*[4] *order of infinitude*. Thus points add nothing to lines, lines to surfaces, surfaces to solids; or—to speak in numbers as is proper in an arithmetical treatise,—roots do not count in regard to squares, squares in regard to cubes, and cubes in regard to square-squares. In such a way one must disregard, as nil, quantities of smaller order.

I have insisted on adding these few remarks, familiar to those who practise indivisibles, in order to bring out the always wonderful connection that nature, in love with unity, establishes between objects distant in appearance. It appears in this example, where we see the calculation of the *dimensions of continuous magnitudes* joined with the *summation of numerical powers*.

Pascal's approach to sums of powers is rich with detail, and ends with his view on how this topic displays the connection between the continuous and the discrete. His idea of using a sum of equations in which one side "telescopes" via cancellations is masterful, and is a tool widely used in mathematics today. Pascal presents a rule obtained by generalizable example. The idea of a generalizable example is to prove the claim for a particular number, but in a way that clearly shows that it works for any number. This was a common method of proof for centuries, in part because there was no notation adequate to handle the general case, and in particular no way of using indexing as we do today to deal with sums of arbitrarily many terms.

Pascal's study of sums of powers is expanded over Fermat's in scope on two fronts: to arithmetic progressions with arbitrary differences and to those beginning with any number. The reader should analyze whether his example convinces one of the general rule, and then apply it to obtain sums of fourth and fifth powers.

**Exercise** 21. Apply Pascal's method to obtain the sum of the fourth powers in the progression $5, 8, 11, 14$, and then check your answer by direct calculation.

**Exercise** 22. Apply Pascal's method to obtain the polynomial formula for the sum of the fourth powers in a natural progression beginning with one, i.e., $\sum_{i=1}^{n} i^4$.

**Exercise** 23. Apply Pascal's method to obtain the polynomial formula for the sum of the fifth powers in a natural progression beginning with one, i.e., $\sum_{i=1}^{n} i^5$.

Pascal's prescription does requires us to substitute known formulas for sums with previous exponents, and then solve for the sum with desired exponent. But we can at least say that Pascal's prescription represents the first general recipe for sums of powers, and it is an attractive formulation,

[4]The French version mistakenly says "higher" here.

connecting sums of powers to each other very directly using binomial coefficients. Let us transcribe his verbal prescription into modern notation for how to find the sum of the first $n$ $k$th powers:

$$(k+1)\sum_{i=1}^{n} i^k = (n+1)^{k+1} - 1^{k+1} - n \cdot 1^{k+1} - \sum_{j=1}^{k-1} \binom{k+1}{j} \sum_{i=1}^{n} i^j \ .$$

We call this *Pascal's equation.* The reader may verify that his method generalizes to produce what he claims in his verbal prescription for more general progressions.

**Exercise** 24. (Optional) Write out Pascal's general result, in modern notation, and provide a proof (based on the method of his example) to justify his general prescription for a sum of powers of any arithmetic progression, i.e., with arbitrary difference and beginning with any number. Include a modern formulation of his algorithm, and apply it to compute some examples.

We can use Pascal's equation to confirm patterns that have slowly been emerging, namely that sums-of-powers polynomials have a particular degree and predictable leading and trailing coefficients:

$$\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2}n^k + \boxed{?}\,n^{k-1} + \cdots + \boxed{?}\,n + 0 \text{ for } k \geq 1.$$

We can confirm these features and even push one step further to discover and confirm a simple pattern for the coefficients of $n^{k-1}$ in the polynomial formulas.

**Exercise** 25. In the text and exercises we have obtained explicit polynomial formulas for sums of powers up to exponent five. From these we conjecture

$$\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2}n^k + \boxed{?}\,n^{k-1} + \cdots + \boxed{?}\,n + 0.$$

Use Pascal's equation to prove these observed patterns for all $k$. (Hint: mathematical induction. You will need a strong form of induction, in which you assume the truth of all preceding statements, not just the one prior to the one you are trying to verify. Why is this stronger form of mathematical induction a valid method of proof?) Then push one step further to conjecture and prove a pattern for the coefficient of $n^{k-1}$ in the formulas.

At this point we can also use Pascal's equation to prove that sums of powers satisfy Roberval's inequalities discussed earlier, which he used to find the areas under all the higher parabola curves.

**Exercise** 26. Prove the inequality $\sum_{i=1}^{n} i^k < (n+1)^{k+1}/\,(k+1)$ of Roberval using Pascal's equation. Can you also use it to prove his second inequality $n^{k+1}/\,(k+1) < \sum_{i=1}^{n} i^k$, which is harder to show?

The patterns in the coefficients of the polynomial formula for sums of powers above begin to reveal more of the connection between the continuous and the discrete. The term $\frac{n^{k+1}}{k+1}$ is the area $\int_0^n x^k dx$ under the curve $y = x^k$ between 0 and $n$. The left side of the above equation is the area of a right-endpoint approximating sum of rectangles for this area, and thus the rest constitutes "correcting terms" interpolating between the area under the curve and the sum of rectangles. The second term, $\frac{1}{2}n^k$, amounts to improving the right-endpoint approximation to a trapezoidal approximation, and the next term also has geometric interpretation as a further correction.

**Exercise** 27. There is a simple geometric interpretation of the second term in the polynomial formula above. Draw a picture illustrating the difference between the region under the curve $y = x^k$ for $0 \le x \le n$ and the region of circumscribing rectangles with ends at integer values. Interpreting their areas as $\int_0^n x^k dx = \frac{n^{k+1}}{k+1}$ and $\sum_{i=1}^n i^k$, find an interpretation in the picture of how the term $\frac{1}{2}n^k$ above represents part of the region between these two, and explain what its connection is to the trapezoid rule from calculus as a numerical approximation for definite integrals. This should suggest to you the sign of the next term in the formula above. What should it be and why?

We can speculate that the other coefficients in these polynomials continue to follow an interesting pattern. We pursue this in the next part of our story, emerging at the turn of the eighteenth century in the work of Jakob Bernoulli (1654–1705).

We end this section by remarking on Pascal's *Conclusion* about indivisibles and the squaring (area) of higher parabolas. Clearly for him the connection between "dimensions of continuous magnitudes" and "summation of numerical powers" is striking and subtle, and was probably a prime motivation for his investigations on sums of powers in an era when many were vying to square higher parabolas and other curves. His view is that for continuous quantities, terms of "lower order of infinitude" (i.e., lesser dimension) add nothing, and one must "disregard [them] as nil," so that the sums of powers formulas above become

$$\sum_{i=1}^{n} i^k \approx \frac{n^{k+1}}{k+1},$$

which is his statement about summing continuous quantities. Today we recognize this as analogous to our integration formula

$$\int_0^n t^k dt = \frac{n^{k+1}}{k+1}$$

for the area under a higher parabola. Turning these analogies into a tight logical connection between discrete summation formulas and continuous area results was part of the long struggle to define and rigorize calculus, which began with the classical Greek mathematics exemplified by Archimedes, and lasted until well into the nineteenth century.

**Exercise** 28. Use

$$\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2}n^k + \cdots$$

as obtained in an exercise above to prove that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n i^k}{n^{k+1}} = \frac{1}{k+1},$$

a result known as Wallis's theorem. Utilize Wallis's theorem and the modern definition of the integral as a limit of approximating sums to calculate

$$\int_0^x t^k dt = \frac{x^{k+1}}{k+1} \text{ for any real } x > 0.$$

Discuss how this supports what Pascal is arguing in his *Conclusion*.

## Jakob Bernoulli finds a pattern

Jakob (Jacques, James) Bernoulli (1654–1705) was one of two spectacular mathematical brothers in a large family of mathematicians spanning several generations. The Bernoulli family had settled in Basel, Switzerland, when fleeing the persecution of Protestants by Catholics in the Netherlands in the sixteenth century.

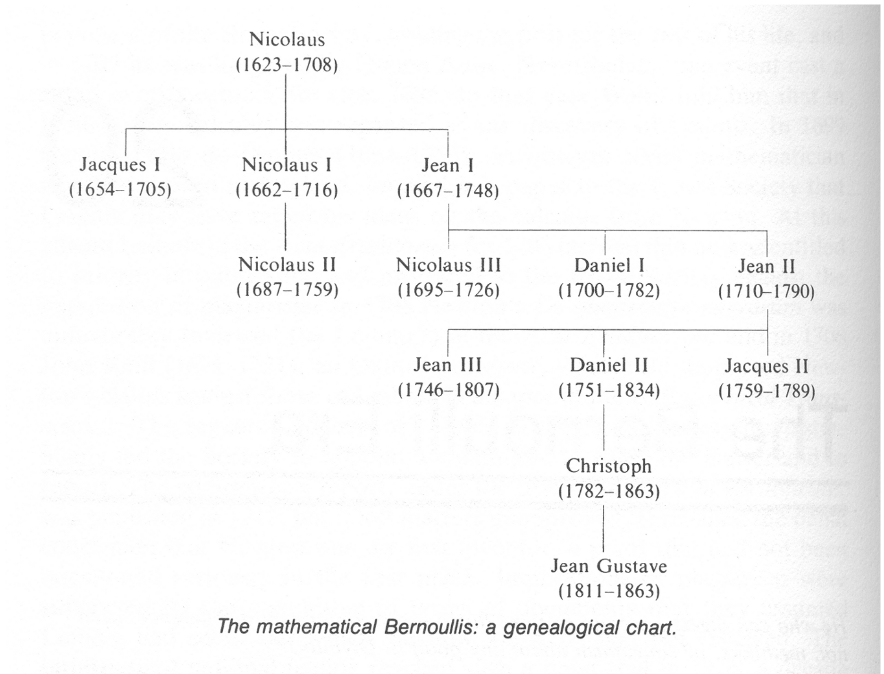



Figure 6: The mathematical Bernoulli family [4, p. 416]

Jakob at first studied mathematics against the will of his father, who wanted him to become a minister, and then traveled widely to learn from prominent mathematicians and scientists in France, the Netherlands, and England. He was appointed professor of mathematics in Basel, and he and his younger brother Johann (Jean, 1667–1748) were among the first to fully absorb Gottfried Leibniz's (1646–1716) newly invented methods of calculus, and to apply them to solve many fascinating mathematical questions. For instance, in 1697 Jakob used a differential equation to solve the brachistochrone problem, i.e., to find the curve down which a frictionless bead will slide from one point to another in the least time. His method began a new mathematical field, the calculus of variations, in which one seeks among all curves the one that maximizes or minimizes some property [9, pp. 547–549]. Bernoulli also used the calculus to discover numerous wonderful properties of the logarithmic spiral, leading him to request that this "spira mirabilis" be engraved on his tombstone [4, p. 417f], [17, pp. 148–153]. And he worked and published much on infinite series.

Jakob wrote the earliest substantial book on probability theory, *Ars Conjectandi* (The Art of Conjecturing). Its posthumous publication in 1713 contained much original work, including the pattern we have been seeking in the formulas for sums of powers. It should not surprise us that Jakob connects probability theory and sums of powers, for we have learned that the figurate numbers, binomial coefficients, and combination numbers are simply different interpretations of the numbers in the arithmetical triangle of Pascal.

While Pascal's equation displayed a compact connection between sums of powers formulas for

Figure 7: Jakob Bernoulli

various exponents, its recursive nature still prevents quick and easy calculation of the polynomial formulas representing $\sum_{i=1}^{n} i^k$ for various values of $k$. Nor did it reveal any general pattern in all the coefficients of these polynomials, even though we suspect there is one. Bernoulli addresses both issues in his short addendum[5] on sums of powers (*Summae Potestatum*) in a chapter of *Ars Conjectandi* on permutations and combinations [1, v. 3, pp. 164–167], [2], [15, pp. 85–90]. Here the *Bernoulli numbers* first appear, and our knowledge takes a tremendous leap forward. We will comment in detail after our source, but we note as an aid beforehand that Bernoulli uses the integral sign to represent finite summations!

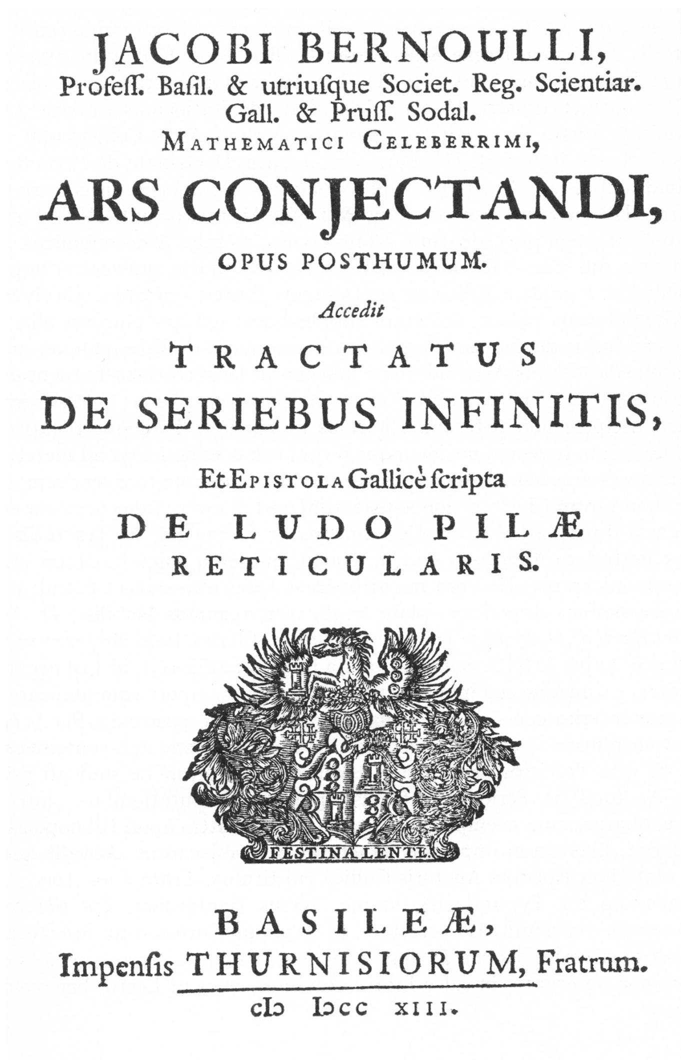
Jakob Bernoulli, from
*The Art of Conjecturing*
*Part Two*

A THEORY OF PERMUTATIONS AND COMBINATIONS
*On combinations of particular numbers of things; which leads to figurate numbers and their properties*

[...] *Scholium*. We note in passing that many (among others, Faulhaber and Remmelin from Ulm, Wallis, Mercator in his Logarithmotechnia, Prestet) have engaged themselves in the study of figurate numbers. But I have found no one who has given a universal and scientific demonstration of this property. Wallis put forward his fundamental methods in the *Arithm. Infinitorum,* where he investigates inductively[6] the ratios of the series of Squares, Cubes, or other powers of the natural numbers to the

[5] We are indebted to Daniel E. Otero for this translation from Latin.

[6] That is, by inductive, as opposed to deductive, reasoning. Inductive reasoning argues from particular cases (specific examples) towards a general conclusion. Deductive reasoning argues from a known principle to an unknown, from the general to the specific, or from a premise to a logical conclusion. When Bernoulli refers here to inductive reasoning, he is not referring to the method of deductive proof called mathematical induction, which is entirely different.


JACOBI BERNOULLI,
Profeſſ. Baſil. & utriuſque Societ. Reg. Scientiar.
Gall. & Pruſſ. Sodal.
MATHEMATICI CELEBERRIMI,

ARS CONJECTANDI,

OPUS POSTHUMUM.

*Accedit*

TRACTATUS
DE SERIEBUS INFINITIS,

Et EPISTOLA Gallicè ſcripta
DE LUDO PILÆ
RETICULARIS.



BASILEÆ,
Impenſis THURNISIORUM, Fratrum.
cIↄ Iↄcc xiii.

Figure 8: *Ars Conjectandi*

series, having as many terms, of the largest of these powers. From this he moves [...] to the study of Triangular, Pyramidal, and the remaining figurate numbers. But it would have been more convenient and appropriate in the nature of things had he instead first prepared a treatise on the figurate numbers, with universal and accurate demonstrations, and then later continued the investigation of sums of powers. For after all, the method of demonstration by induction is not particularly scientific, and besides, each series requires its own special methods. Those series which should be considered first, by general estimation, and whose natures are most fundamental and simple, are seen to be the figurate numbers, which are generated by addition, while the powers are generated by multiplication. Moreover, the series of figurates, beginning with their respective zeros, have exact fractional ratios with the series having the same number of constant terms equal to the largest of these,[7] which is not necessarily so for the powers (at least not in a finite number of terms, regardless how many zeros, by excess or defect, are prefixed to it). Furthermore, from the knowledge of the sum of figurates, it is no more difficult to determine the sums of powers, and so the author has concluded from these first ideas, as I will now do most briefly.

Let there be given the series of natural numbers from unity: $1, 2, 3, 4, 5$, etc., up to $n$, and suppose that we ask for the sums of these, or of their squares, their cubes, etc.: In the Table of Combinations[8] the indefinite term in the [...][9] third column is found to be

$$\frac{n-1 \cdot n-2}{1 \cdot 2} = \frac{nn-3n+2}{2},$$

and the sum of all the terms (that is, all $\frac{nn-3n+2}{2}$) is

$$\frac{n \cdot n-1 \cdot n-2}{1 \cdot 2 \cdot 3} = \frac{n^3-3nn+2n}{6};$$

$$\text{this gives}^{10} \int \overline{\frac{nn-3n+2}{2}} \text{ or } \int \frac{1}{2}nn - \int \frac{3}{2}n + \int 1 = \frac{n^3-3nn+2n}{6}.$$

$$\text{So } \int \frac{1}{2}nn = \frac{n^3-3nn+2n}{6} + \int \frac{3}{2}n - \int 1.$$

$$\text{But } \int \frac{3}{2}n = \frac{3}{2}\int n = \text{(by what was shown above) } \frac{3}{4}nn + \frac{3}{4}n,$$

and $\int 1 = n$; substituting these above gives

$$\int \frac{1}{2}nn = \frac{n^3-3nn+2n}{6} + \frac{3nn+3n}{4} - n = \frac{1}{6}n^3 + \frac{1}{4}nn + \frac{1}{12}n,$$

and by doubling, $\int nn$ (the sum of the squares of all n)

$$= \frac{1}{3}n^3 + \frac{1}{2}nn + \frac{1}{6}n.$$

[...][11] And by proceeding to higher powers in turn, we easily build up the following formulas:

[7] The reader may wish to decipher this claim, and see that it is actually equivalent to Fermat's claims in his letter to Mersenne. Hint: Bernoulli means that to obtain an unchanging fractional ratio $1/(r+1)$ using the sum of any series of $r$-dimensional figurate numbers in the numerator, one should prefix the sequence with $r$ zeros for determining the number of constant terms in the denominator of the ratio. He then contrasts this with sums of powers, and compares with the infinite case; can you see what he is getting at?

[8] That is, the arithmetical triangle.

[9] We omit Bernoulli's derivation of the sum of first powers, moving directly to a sum of squares.

[11] Bernoulli continues on to a sum of cubes.

$$\begin{aligned}
\int n &= \frac{1}{2}nn + \frac{1}{2}n.\\
\int nn &= \frac{1}{3}n^3 + \frac{1}{2}nn + \frac{1}{6}n.\\
\int n^3 &= \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}nn.\\
\int n^4 &= \frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 * -\frac{1}{30}n.\\
\int n^5 &= \frac{1}{6}n^6 + \frac{1}{2}n^5 + \frac{5}{12}n^4 * -\frac{1}{12}nn.\\
\int n^6 &= \frac{1}{7}n^7 + \frac{1}{2}n^6 + \frac{1}{2}n^5 * -\frac{1}{6}n^3 * +\frac{1}{42}n.\\
\int n^7 &= \frac{1}{8}n^8 + \frac{1}{2}n^7 + \frac{7}{12}n^6 * -\frac{7}{24}n^4 * +\frac{1}{12}nn.\\
\int n^8 &= \frac{1}{9}n^9 + \frac{1}{2}n^8 + \frac{2}{3}n^7 * -\frac{7}{15}n^5 * +\frac{2}{9}n^3 * -\frac{1}{30}n.\\
\int n^9 &= \frac{1}{10}n^{10} + \frac{1}{2}n^9 + \frac{3}{4}n^8 * -\frac{7}{10}n^6 * +\frac{1}{2}n^4 * -\frac{3}{20}nn.\\
\int n^{10} &= \frac{1}{11}n^{11} + \frac{1}{2}n^{10} + \frac{5}{6}n^9 * -1n^7 * +1n^5 * -\frac{1}{2}n^3 * +\frac{5}{66}n.
\end{aligned}$$

Indeed, a pattern can be seen in the progressions herein, which can be continued by means of this rule: Suppose that $c$ is the value of any power; then the sum of all $n^c$ or

$$\begin{aligned}
\int n^c = \frac{1}{c+1}n^{c+1} + \frac{1}{2}n^c + \frac{c}{2}An^{c-1} + \frac{c\cdot c-1\cdot c-2}{2\cdot 3\cdot 4}Bn^{c-3}\\
+ \frac{c\cdot c-1\cdot c-2\cdot c-3\cdot c-4}{2\cdot 3\cdot 4\cdot 5\cdot 6}Cn^{c-5}\\
+ \frac{c\cdot c-1\cdot c-2\cdot c-3\cdot c-4\cdot c-5\cdot c-6}{2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8}Dn^{c-7}\ldots \text{ \& so on,}
\end{aligned}$$

where the value of the power $n$ continues to decrease by two until it reaches $n$ or $nn$. The uppercase letters $A$, $B$, $C$, $D$, etc., in order, denote the coefficients of the final term of $\int nn$, $\int n^4$, $\int n^6$, $\int n^8$, etc., namely

$$A = \frac{1}{6}, \quad B = -\frac{1}{30}, \quad C = \frac{1}{42}, \quad D = -\frac{1}{30}.$$

These coefficients are such that, when arranged with the other coefficients of the same order, they add up to unity: so, for $D$, which we said signified $-\frac{1}{30}$, we have

$$\frac{1}{9} + \frac{1}{2} + \frac{2}{3} - \frac{7}{15} + \frac{2}{9}(+D) - \frac{1}{30} = 1.$$

[12] The symbol $*$ in Bernoulli's table means that a term with a particular power of $n$ is not shown because the coefficient is zero.

[13] There is an error in the original published Latin table of sums of powers formulas. The last coefficient in the formula for $\int n^9$ should be $-\frac{3}{20}$, not $-\frac{1}{12}$; we have corrected this here.

By means of these formulas, I discovered in under a quarter hour's work that the tenth (or quadrato-sursolid) powers of the first thousand numbers from unity, when collected into a sum, yield

$$91409924241424243424241924242500.$$

Clearly this renders obsolete the work of Ismael Bulliald, who wrote so as to thicken the volumes of his *Arithmeticae Infinitorum* with demonstrations involving immense labor, unexcelled by anyone else, of the sums of up to the first six powers (which is only a part of what we have superseded in a single page).

**Exercise** 29. See whether you can duplicate Bernoulli's claim that he calculated (by hand, of course) the sum of the tenth powers of the first thousand numbers in less than a quarter of an hour.

Interestingly, while Bernoulli indicates his familiarity with the work of several others on sums of powers, he mentions neither Fermat nor Pascal. John Wallis (1616–1703) had studied sums of powers in his *Arithmetica Infinitorum* of 1655, with the same motivation as Fermat and Pascal, to find the areas under higher parabolas. Bernoulli contrasts Wallis's work with his own, including comparing sums of figurate numbers with sums of powers. While finite sums of powers do not behave as nicely as sums of figurate numbers, Bernoulli's subsequent formulas shed light on the nature of the difference between them by providing a precise expression for $\sum_{i=1}^{n} i^k$ as a polynomial in $n$.

Notice Bernoulli's summation notation as he proceeds to analyze sums of powers. The expression after the integral indicates both the general term and the ending index, i.e., he writes $\int n^7$ for $\sum_{i=1}^{n} i^7$. He also uses an asterisk to indicate "missing" terms, i.e., monomials with zero coefficient.

**Exercise** 30. Today we use the notation $\sum_{i=1}^{n} i^7$ instead of Bernoulli's $\int n^7$. What are the advantages and disadvantages of the two notations?

Bernoulli first shows how to derive sum formulas for the first few exponents, using his knowledge of the arithmetical triangle, by exactly the same method we presented when considering Fermat's claim to have solved the problem. He presents the results of calculation in a table of polynomials for sums up to the tenth powers. And now suddenly he claims:

Indeed, a pattern can be seen in the progressions herein which can be continued by means of this rule:

Perhaps readers will discover this pattern for themselves before reading closely Bernoulli's description of it.

**Exercise** 31. (Optional) Guess, as did Bernoulli, the complete pattern of coefficients for sums of powers formulas just from the examples in Bernoulli's table. Clearly the pattern is to be sought down each column of Bernoulli's table. The key is to multiply each column of numbers by a common denominator, and then compare with the arithmetical triangle (computing the sequence of successive differences in a column, and the successive differences in that sequence, etc., may also help). Can you also express the general rule for calculating the special numbers $A, B, C, D, \ldots$, which Bernoulli introduces? Hint: What happens when $n = 1$?

The reader should check that in modern notation, Bernoulli is claiming

$$\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{n^k}{2} + \sum_{j=2}^{k} \frac{1}{k+1}\binom{k+1}{j} B_j n^{k+1-j} \text{ for } k \geq 1,$$

where we have represented the special sequence of numbers that Bernoulli calls $A, B, C, D, \ldots$ by $B_{2m}$ for $l \geq 1$, and $B_{2m+1} = 0$. These numbers have been important in mathematics ever since their introduction here by Bernoulli[14]. The great eighteenth century mathematician Leonhard Euler (1707–1783) christened them the *Bernoulli numbers*.

Bernoulli also claims that he can compute his special sequence of numbers $A, B, C, D, \ldots$. First he notes that they

in order, denote the coefficients of the final term of $\int nn$, $\int n^4$, $\int n^6$, $\int n^8$.

Indeed, we notice that the coefficient of $n$ in the general formula he gives is always the first occurrence of a new Bernoulli number in the process. And he says:

These coefficients are such that, when arranged with the other coefficients of the same order, they add up to unity.

Here he is simply evaluating both sides of his general formula at $n = 1$. Since the left side is then 1, the $k$th formula simplifies to

$$1 = \frac{1}{k+1} + \frac{1}{2} + \sum_{j=2}^{k} \frac{1}{k+1}\binom{k+1}{j} B_j.$$

Since the last term in the sum is the newest Bernoulli number $B_k$, one can solve for it in terms of the previous ones. Thus the Bernoulli numbers are recursively defined by these formulas. He gives as an example the computation of $D = B_8 = -\frac{1}{30}$ from the formula for $k = 8$ and the previous numbers. While this still leaves a step-by-step aspect to the determination of sums of powers formulas, the process is now greatly simplified. Moreover, we see a general pattern in the relationship between the coefficients for different values of $k$, since the Bernoulli numbers are the same in the formulas for all $k$. These are the great steps forward that Bernoulli provided beyond the work of Fermat and Pascal.

How might we attempt to verify the general validity of the pattern Bernoulli guessed? Since Pascal gave us an equation relating the sums of $k$th powers to those of lower powers, we should be able to proceed by strong mathematical induction on $k$, by simply substituting all the formulas of Bernoulli's into Pascal's equation to verify the inductive claim at each stage. All but one of Bernoulli's formulas substituted in Pascal's equation are assumed true inductively, and the $k$th is thus shown true by verifying the equality itself.

[14]The evidence suggests that around the same time, Takakazu Seki (1642?–1708) in Japan also discovered the same numbers [14, 16].

**Exercise** 32. Prove Bernoulli's claimed formulas by strong mathematical induction, in the manner suggested in the text, using Pascal's equation, Bernoulli's claims, and the Bernoulli numbers as defined recursively. At some point in your calculations you may need to prove and use the identity $\binom{a}{b}\binom{b}{c} = \binom{a}{c}\binom{a-c}{b-c}$. Hint: When substituting Bernoulli's claims into Pascal's equation, verify equality by calculating and comparing the coefficients for an arbitrary power of $n$ on each side of the equation.

The reader may notice a pattern in the Bernoulli numbers not even mentioned by Bernoulli.

**Exercise** 33. What do you conjecture about the signs of the Bernoulli numbers? Compute several more Bernoulli numbers to see whether your conjecture has promise.

Leonhard Euler in the eighteenth century was the first to prove Bernoulli's claimed patterns in the coefficients of the sums of powers polynomials, as part of his development of spectacular new mathematics for finding sums of infinite series, a discovery called the Euler-MacLaurin summation formula. For instance, he was able to guess, and then prove, that the sum of the reciprocal squares, $\sum_{i=1}^{\infty} \frac{1}{i^2}$, is exactly $\pi^2/6$, one of the most astonishing discoveries of the era. Before this time, almost the only infinite series whose sums were known were geometric series, and their sums never involved totally unexpected numbers like $\pi$. This next episode of our story continues in another project. Euler's work led to modern methods of studying the distribution of prime numbers, one of the most active research areas in mathematics today [11].

## Notes to the instructor

This project is the third episode of four connected projects following the epic story of formulas for sums of powers from the Pythagoreans to Euler; it is a central theme in the development of discrete mathematics and combinatorics. The audience for this episode is students of intermediate discrete mathematics or combinatorics, and the episode connects sums of powers to figurate numbers, binomial coefficients, Pascal's triangle, and Bernoulli numbers. The project is quite flexible, and the instructor should be able to pick and choose, if desired, from the various activities offered. For a shorter project the instructor can choose selectively from this module; some exercises are marked as optional if they are not critical to later work. Students may need substantial guidance with some parts, and the instructor should be sure to work through all the details before assigning any student work.

The goal is for students to learn many basic notations, techniques, and skills in the context of an historically and mathematically authentic big motivating problem with multiple connections to other mathematics. Hopefully this will be much more effective and rewarding than simply being asked to learn various skills for no immediately apparent application. Many of the techniques and phenomena introduced in a discrete mathematics or combinatorics course simply arise naturally in this project, like recursive definitions, delicate work with summations and inequalities, counting and geometry, binomial coefficients and combination numbers, and proofs by mathematical induction. Instead of separately covering various such topics and techniques, that class time could simply be spent on the project, and students will learn those things in the process.

The project also asks students to conjecture from patterns they generate, develop their mathematical intuition and judgement, and try proving their conjectures, i.e., putting students in the creative mathematical driver seat in an authentic context. The setting of sums of powers in the context of primary sources allows a richness of questions and interpretations, especially includes deep connections to geometry and the two-way interplay with calculus, as well as basic algebra and linear algebra, and a richness of proof techniques, including natural comparison of the efficacy of various proof methods.

## References

[1] J. Bernoulli, *Die Werke von Jakob Bernoulli*, Naturforschende Gesellschaft in Basel, Birkhäuser Verlag, Basel, 1975.

[2] J. Bernoulli *The Art of Conjecturing: Together with "Letter to a Friend on Sets in Court Tennis"* (transl. E.D. Sylla), Johns Hopkins University Press, Baltimore, 2006.

[3] C. Boyer, *Pascal's Formula For the Sums of Powers of the Integers*, Scripta Mathematica **9** (1943), 237–244.

[4] C.B. Boyer, U.C. Merzbach, *A History of Mathematics* (second ed.), John Wiley & Sons, New York, 1989.

[5] *Encyclopædia Britannica*, Chicago, 1986.

[6] P. de Fermat, *Oeuvres*, P. Tannery (ed.), Paris, 1891–1922.

[7] C.C. Gillispie, F.L. Holmes (eds.), *Dictionary of Scientific Biography*, Scribner, New York, 1970.

[8] *Great Books of the Western World*, Mortimer Adler (ed.), Encyclopædia Britannica, Inc., Chicago, 1991.

[9] V.J. Katz, *A History of Mathematics: An Introduction* (second ed.), Addison-Wesley, New York, 1998.

[10] R. Laubenbacher, D. Pengelley, *Mathematical Expeditions: Chronicles by the Explorers*, Springer Verlag, New York, 1999.

[11] A. Knoebel, R. Laubenbacher, J. Lodder, D. Pengelley, *Mathematical Masterpieces: Further Chronicles by the Explorers,* Springer Verlag, New York, 2007.

[12] M. Mahoney, *The Mathematical Career of Pierre de Fermat*, 2nd edition, Princeton University Press, Princeton, New Jersey, 1994.

[13] B. Pascal, *Oeuvres*, L. Brunschvieg (ed.), Paris, 1908–14; Kraus Reprint, Vaduz, Liechtenstein, 1976.

[14] K. Shen, *Seki Takakazu and Li Shanlan's Formulae on the Sum of Powers and Factorials of Natural Numbers. (Japanese)* Sugakushi Kenkyu **115** (1987), 21–36.

[15] D.J. Struik, *A Source Book in Mathematics, 1200–1800*, Cambridge, Harvard University Press, 1969; Princeton University Press, Princeton, 1986.

[16] K. Yosida, *A Brief Biography on Takakazu Seki (1642?–1708)*, Math. Intelligencer **3** (1980/81), no. 3, 121–122.

[17] R.M. Young, *Excursions in Calculus: An Interplay of the Continuous and the Discrete*, Mathematical Association of America, Washington, D.C., 1992.