1.3 Fermat and Pascal Use Figurate Numbers, Binomials, and the Arithmetical Triangle to Calculate Sums of Powers

In the autumn of 1636, Pierre de Fermat (1601–1665) wrote to Marin Mersenne (1588–1648) in Paris [133, pp. 436, 481f], implying that he had solved “what is perhaps the most beautiful problem of all arithmetic,” finding the precise sum of powers in an arithmetic progression, no matter what the power [19], [73, volume II, pp. 70, 84–85]. Fermat, who did mathematics in his spare time, was one of the truly great figures in the history of mathematics, often stating his results without proof in his letters, leaving it to others to try to work out the details and verify his claims. Some of his assertions about numbers have driven the development of mathematics for hundreds of years. More about Fermat’s life can be found in our chapter on the quadratic reciprocity law between primes.

In his letter to Mersenne, Fermat wrote that he had reached his results on sums of powers by using the following theorem on “natural progressions”:

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 [19], [163, p. 230f], [73, vol. II, pp. 70, 84–85].

Although Fermat characteristically does not say any more, we will be able to understand what he is claiming here, and see how it produces formulas for sums of powers. This analysis will also prepare us for reading Pascal’s and Bernoulli’s works on sums of powers.

First, exactly what does Fermat mean by the collateral triangle, pyramid, and triangulo-triangle, examples of what we call figurate numbers? We saw earlier that a triangular number counts the dots in a triangular figure composed of rows starting with one dot and increasing in length by one (Figure 1.2). Thus by a “collateral triangle” Fermat means the triangular number with a specified number of rows, i.e., a certain length to its side, such as the triangular number $1 + 2 + 3 = 6$ with side length three. Likewise, a pyramidal (or
1.3 Fermat and Pascal Sum Powers Using Figurates and Binomials

tetrahedral) number counts the dots in a three-dimensional pyramidal figure formed by piling shrinking triangles on top of each other as in Figure 1.5. So the collateral pyramid with side length three is the sum $1 + 3 + 6 = 10$ of the first three triangular numbers. Likewise, Fermat’s four-dimensional triangulotriangle with side length three is the sum $1 + 4 + 10 = 15$ of the first three pyramidal numbers; and so on.

Fermat’s claims are not merely geometric interpretations of relationships between certain whole numbers; they are actually discrete analogues of continuous results we already know about 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 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 1.2. 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 continuous 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. The reader may enjoy seeing the discrete relationship geometrically as well, by packing six discrete triangular pyramidal figures into a box figure, half of which is the discrete prism. 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 1.5). 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

\[
F_{n,0} = 1 \quad (n \geq 1), \\
F_{1,j} = 1 \quad (j \geq 0), \\
F_{n+1,j+1} = F_{n+1,j} + F_{n,j+1} \quad (n \geq 1, j \geq 0).
\]
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

$$nF_{n+1,1} = 2F_{n,2},$$
$$nF_{n+1,2} = 3F_{n,3},$$
$$nF_{n+1,3} = 4F_{n,4},$$

and

Taking these temporarily for granted, let us see how sums of powers can be calculated, as Fermat claimed. As an example, we will derive the formula for a sum of squares. We know that any pyramidal number is obtained by piling up triangular numbers, i.e.,

$$F_{n,3} = \sum_{i=1}^{n} F_{i,2}.$$

We can also obtain closed formulas for the figurate numbers on both sides of this equality by iteratively using Fermat’s formulas:

$$F_{n,3} = \frac{n}{3} F_{n+1,2} = \left( \frac{n}{3} \right) \left( \frac{n+1}{2} \right) F_{n+2,1} = \left( \frac{n}{3} \right) \left( \frac{n+1}{2} \right) \left( \frac{n+2}{1} \right),$$

and

$$F_{i,2} = \frac{i}{2} F_{i+1,1} = \left( \frac{i}{2} \right) \left( \frac{i+1}{1} \right).$$

So the two sides of the piling-up equation above expand to produce

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

Now we simply substitute $n(n+1)/2$ for $\sum_{i=1}^{n} i$, which we already know inductively, and solve for $\sum_{i=1}^{n} i^2$, clearly yielding a polynomial with leading term $n^3/3$, as expected. The reader may verify that we get the correct formula, and may continue this method to sums of higher powers (Exercise 1.27). In fact, this very method is advanced by Bernoulli in the text we will read in the next section. 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.

We shall now study the figurate numbers further to see why Fermat’s claim about them is true, and simultaneously prepare the groundwork for reading Pascal. The reader may have noticed that the individual figurate numbers
seem familiar, from looking back to Pascal’s arithmetical triangle displayed in the introduction (Figure 1.6). The numbers shown there match the figurate numbers, i.e., $F_{n,j}$ appears in Pascal’s “parallel row” $n$ and “perpendicular row” $j + 1$. Indeed, in his treatise on the arithmetical triangle, Pascal defines the numbers in the triangle by starting the process off with a 1 in the corner, and defines 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 by which we formally defined the figurate numbers.

Perhaps the reader also already recognizes the numbers in the arithmetical triangle as binomial coefficients or combination numbers. The numbers occurring along Pascal’s ruled diagonals in Figure 1.6 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, then the entry in diagonal $m$ at column $j$ will be the binomial coefficient $\binom{m}{j}$. This is easy to prove (Exercise 1.28).

Since we have now identified both the figurate numbers and the binomial coefficients as the numbers in the triangle generated by the basic recursion relation, their precise relationship follows just by comparing their indexing:

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

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

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

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 1.29).

This ubiquitous triangle of numbers 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 [133]. A key fact, which Pascal called the Twelfth Consequence, is that neighboring numbers along a diagonal in the triangle are always in a simple ratio:

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

which is easily obtained from the factorial formula above (Exercise 1.30).
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

$$n \binom{n + 2}{2} = 3 \binom{n + 2}{3},$$

or

$$nF_{n+1,2} = 3F_{n,3},$$

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 1.31), and we are ready to move on to Pascal’s work on sums of powers.

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, as we described in the introduction. 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. 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 [92, 63]. Pascal was an extremely complex person, and one of the outstanding scientists of the mid-seventeenth century, but we will never know how much 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 1.32). For instance, there are \( \binom{11}{5} = 462 \) different teams of 5 players possible from a pool of 11 people available to play.

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 (Exercises 1.33, 1.34). The reader will enjoy reading Pascal’s actual treatise [100, v. 30].

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), which analyzes sums of powers in arithmetic progressions 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

---

6 Later in the seventeenth century, Leibniz, 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.
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*. Below we will see the transition from a geometric to algebraic approach almost complete, since Pascal, unlike Archimedes, Al-Karaji, al-Haytham, 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 present a translation from the French (*Sommation des puissances numériques*), published in his collected works with the Latin on facing pages [182, v. III, pp. 341–367]. 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,.... 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,..., 2, 4, 6, 8,..., or differing by three units like 1, 4, 7, 10, 13,... 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,...: 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,... 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

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

\[
14^4 + 1214^3 + 5414^2 + 10814 + 81.
\]

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

\[
14^4 + 1214^3 + 5414^2 + 10814 + 81,
\]

is \( 1214^3 + 5414^2 + 10814 + 81 \); 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

\[
1211^3 + 5411^2 + 10811 + 81,
\]

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

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

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 $5^2 + 8^2 + 11^2 + 14^2$, or 25 + 64 + 121 + 196, or again 406, which, multiplied by 54, gives 21924;

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
1.3 Fermat and Pascal Sum Powers Using Figurates and Binomials

\[17^4 - 14^4 + 14^4 - 11^4 + 11^4 - 8^4 + 8^4 - 5^4 + 5^4.\]

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 \(17^4\):

\[
\begin{align*}
12.14^3 + 54.14^2 + 108.14 + 81 \\
+ 12.11^3 + 54.11^2 + 108.11 + 81 \\
+ 12.8^3 + 54.8^2 + 108.8 + 81 \\
+ 12.5^3 + 54.5^2 + 108.5 + 81 \\
+ 5^4,
\end{align*}
\]

or, on interchanging the order of the terms:

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

If therefore one subtracts on both sides the sum:

\[
\begin{align*}
5 + 8 + 11 + 14 & \text{ multiplied by } 108, \\
+ 5^2 + 8^2 + 11^2 + 14^2 & \text{ multiplied by } 54, \\
+ 81 + 81 + 81 + 81 & \text{ multiplied by } 108, \\
+ 5^4; \\
\end{align*}
\]

There remains \(17^4\) diminished by the previously known quantities:

\[
\begin{align*}
- 5 - 8 - 11 - 14 & \text{ multiplied by } 108, \\
- 5^2 - 8^2 - 11^2 - 14^2 & \text{ multiplied by } 54, \\
- 81 - 81 - 81 - 81 & \text{ multiplied by } 54, \\
- 5^4; \\
\end{align*}
\]
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 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. Like

---

7 The French version mistakenly says “higher” here.
al-Karajī for exponent three, and al-Haytham for exponents four and higher, Pascal presents a rule obtained by generalizable example, but expanded in scope on two fronts: to arithmetic progressions with arbitrary differences and to those beginning with any number. The reader is urged to consider whether his example convinces one of the general rule, and then try applying it to obtain the sum of fourth powers in his progression 5, 8, 11, 14 (Exercise 1.35), and finally use it to obtain a formula for the sum of the first \(n\) fifth powers (Exercise 1.36). Doing this displays a clear advantage over al-Haytham’s equation. Pascal’s only requires us to substitute directly our known formulas for sums with previous exponents, and then solve immediately, without having first to expand and rearrange al-Haytham’s double summation and then apply previous formulas a second time before solving. In this sense we can say that Pascal’s prescription represents the first explicit recipe for sums of powers.

Although it requires knowledge of the formulas for all previous exponents, Pascal’s is an attractive formulation. 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 \textit{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 1.37).

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

\[
\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2} n^k + \frac{2}{4} n^{k-1} + \cdots + \frac{k}{2} n + 0 \text{ for } k \geq 1.
\]

We leave it to the reader to confirm these features and even to push one step further to discover and confirm a simple pattern for the coefficients of \(n^{k-1}\) in the polynomial formulas (Exercise 1.38). We can also use Pascal’s equation to prove that sums of powers satisfy Roberval’s inequalities discussed in the introduction (Exercise 1.39).

The patterns in the polynomial coefficients 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 term \(\frac{1}{2} n^k\) amounts to improving the right-endpoint approximation to a trapezoidal approximation, and the next term also has interpretation as a further correction (Exercise 1.14). We can speculate that the other coefficients in these polynomials continue to follow an interesting pattern. We pursue this in the next
episode 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

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

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 (Exercises 1.40, 1.41, 1.42).

**Exercise 1.27.** Finish deducing and checking the sum of squares formula derived from Fermat’s claims about figurate numbers, and carry the procedure forward to obtain the formula for a sum of cubes from Fermat’s claims. Discuss what would be involved in carrying this to higher powers.

**Exercise 1.28.** 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 $(a+b)^{m+1} = (a + b)(a + b)^m$.)

**Exercise 1.29.** 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.)

**Exercise 1.30.** Prove Pascal’s Twelfth Consequence from the factorial formula for binomial coefficients.
Exercise 1.31. State Fermat’s general claim about figurate numbers, and prove it from Pascal’s Twelfth Consequence.

Exercise 1.32. 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).

Exercise 1.33. Prove that the sum of the numbers in each row of the arithmetical triangle is a power of two. Hint: binomial theorem.

Exercise 1.34. Find a pattern of your own in the arithmetical triangle, and then prove that it holds.

Exercise 1.35. 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 1.36. 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., \( \sum_{i=1}^{n} i^5 \).

Exercise 1.37. 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.

Exercise 1.38. In the text and exercises we have obtained explicit polynomial formulas for sums of powers up to exponent five. From these we conjecture

\[
\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2} n^k + \binom{k}{2} n^{k-1} + \cdots + \binom{k}{k-1} n + 0.
\]

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.

Exercise 1.39. 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?
Exercise 1.40. Compare and contrast the work of Archimedes and Pascal toward finding the area of a bounded curvilinear region in the plane. Cite specific contributions of each author toward the solution of this problem. Explain what specific mathematical constructions from the era of Fermat and Pascal facilitate a discussion of “area under a curve.”

Exercise 1.41. Use \[ \sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + \frac{1}{2} n^k + \cdots \] as obtained in Exercise 1.38 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 (discussed further in the next section). 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.

Exercise 1.42. Read about the work of Fermat and Pascal, and discuss how close each came to the modern idea of integration and the fundamental theorem of calculus? What prevented them from developing calculus?

1.4 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.

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