Mathematical induction

The **mathematical induction** (in Latin: *inductio matematica*, *inductio plena*) is a method of mathematical proof (and, in spite of its name, it is a deductive reasoning). The proof realized with the mathematical induction is called an **inductive proof**.

Typically the mathematical induction is used to establish a claimed statement, T(n), for every natural number n (bigger than a certain one, n_0; in many cases this initial value is 0 or 1), and we here outline its idea in this case only. It is done in two steps.
The first step (called a **base case**, *inceptum inductionis*) is the check that the statement T(n) holds true for the initial value $n = n_0$.
The second step (known as an **inductive step**, *gradus inductionis*) is to follow the schema: we assume T(n) is true
and we conclude the truth of the statement T(n+1).
Usually it is spoken as follows:
we do an **inductive assumption** , or inductive hypothesis,
and we derive the **inductive thesis**.

The inductive proofs are based on the **principle of mathematical induction**, i.e., on the following schema:

$$\left\{ \mathrm{T}(n_0) \ \& \ \left\{ \forall_{m > n_0} \ \mathrm{T}(m) \Rightarrow \mathrm{T}(m+1) \right\} \right\} \Rightarrow \forall_{n \geq n_0} \ \mathrm{T}(n)$$

(and commonly, instead of m it is used n).

The first inductive proofs, although not in the perfect form, were presented, implicitly, by Plato in his philosophical dialogues *Parmenides* (around 370 BC) and by Euclid in his Στοιχεῖα (*Elements*, c.300 BC), who used it to show that there are infinitely many prime numbers. Elements of the mathematical induction appeared in India, where Jayadeva (c.1000 AD) and Bhaskara II (c.1150) worked out *ćarkwala* (cyclic algorithm) to solve indetermined quadratic equation [1]. The mathematical induction was also independently discovered by a Persian mathematician and engineer al-Karaji, in his lost work (written c.1000 AD) he used it when writing on the binomial theorem and Pascal triangle (both these notions are presented later on). The first explicit formulation of the mathematical induction was given by Blaise Pascal in his *Traité du*

[1]) More precisely, in 628 Brahmagupta found integer solutions, i.e., pairs (x, y) of integer numbers, of the equation of the form $x^2 = N y^2 + 1$ for some values of integer N. His approach was generalized by Jayadeva and Bhaskara II and it provided solutions for arbitrary integer N. Their result was unknown in Europe, where this equation with $N = 61$ was solved in 1658 by William Brouncker, and a general method was discovered by J.L.Lagrange in 1766; nota bene, Lagrange technique is more laborious that ćarkwala.

triangle arithmétique (published in 1665). The term ‘mathematical induction’ and a modern rigorous treatment of its principle was provided by Augustus de Morgan in his Penny Cyclopedia article *Induction* of 1838. This principle was investigated in detail by George Boole (in *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*, 1835) and by such researches in foundations of mathematics as Gottlob Frege, Giuseppe Peano and Richard Dedekind.

Beside the already mentioned proof on the infinite number of primes the most known inductive proofs are that on

- the sum of consecutive natural numbers,
- the number of permutations,

and we below present them (and some other ones, too). Let’s start with the first of them.

Triangular numbers

Maybe the most popularized proof by the mathematical induction is that showing that the sum of n consecutive natural numbers starting at 1 is equal to $n \cdot (n+1)/2$; for example,

for $n = 4$: $1 + 2 + 3 + 4 = 10 = 2 \cdot 5 = 4/2 \cdot (4+1) = 4 \cdot (4+1)/2$,

for $n = 100$: $1 + 2 + 3 + \ldots + 100 = 5500 = 100 \cdot (100+1)/2$.

The first of above sums, valued 10 and called a tetraktys, was particularly admired by Pythagoreans. They acted in the 6th century BC (then so-called golden age of Greek mathematics started) and they knew that, for every concrete n, instead of evaluating the sum $1 + 2 + 3 + \ldots + n$ one can produce the correct result by one addition, one division by 2 and one multiplication.

In the mathematical notation [2)] the claim on the considered sum can be written down in any of following forms:

1°: if $a_k = k$, then for every natural n there holds $\sum_{k=1}^{n} a_k = \frac{n\cdot(n+1)}{2}$,

2°: $\forall_{n \in \boldsymbol{N}} \sum_{k=1}^{n} k = \frac{n\cdot(n+1)}{2}$,

3°: T(n), where T(n) denotes the statement: $\forall_{n \in \boldsymbol{N}} \sum_{k=1}^{n} k = \frac{n\cdot(n+1)}{2}$.

Below let's present the (inductive) proof of the above, that is the proof that

$$\forall_{n \in \boldsymbol{N}} \; L_n = P_n ,$$

where $L_n := \sum_{k=1}^{n} k$, $P_n := \frac{n\cdot(n+1)}{2}$.

The base case is nothing else than the check that the thesis holds true for $n = 1$. Since $L_1 = \sum_{k=1}^{1} k = 1$ and $P_1 = \frac{1\cdot(1+1)}{2} = 1$, so $L_1 = P_1$.

The base case completed, we go to do the inductive step. We assume that there holds true the equality $L_n = P_n$ for some n (this is the inductive assumption) and we will show that $L_{n+1} = P_{n+1}$ (this is the inductive thesis).

[2)] It is used the summation symbol (capital Greek letter Σ, read as 'the sum'),

$$\sum_{k=m}^{n} a_k := a_m + a_{m+1} + a_{m+2} + \ldots \; a_n,$$

k is called a summation index, it runs from the initial value m and it increases by 1 till the final value n; by definition, if $m < n$, then the sum equals 0,

a_k is called a k-th addend, or a k-th component, of the sum at hand;

the summation is said to be realized wrt k

We will show it by transforming the left side, L_{n+1}, of the thesis to be proven. Step by step there is

$$\begin{aligned} L_{n+1} = \sum_{k=1}^{n+1} k &= \sum_{k=1}^{n} k + (n+1) = \\ &= L_n + (n+1) = \\ &= P_n + (n+1) = \\ &= \frac{n\cdot(n+1)}{2} + (n+1) = (n+1)\cdot\left(\frac{n}{2}+1\right) = \frac{(n+1)\cdot(n+2)}{2} = P_{n+1}, \end{aligned}$$

where consecutive equality signs are because of

1) the definition of what L_n stands for,
2) the decomposition of the sum into the sum of two components,
3) the meaning of the symbol L_n,
4) the inductive assumption,
5) the use of the denotation P_n,
6) the pulling out the common factor $(n+1)$ in front of the parentheses,
7) the writing out both factor on the same solidus,
8) the recognizing that the previous expression is P_n with the index increased by 1.

The above proves that the implication

$$\{ L_n = P_n \} \Rightarrow \{ L_{n+1} = P_{n+1} \}$$

holds true. This way the inductive step is completed, and – in virtue of the principle of mathematical induction – it proves that $\forall_{n\in \boldsymbol{N}}\ L_n = P_n$ Q.E.D. [3)].

It is not difficult to notice that 1, 3, 6, 10, 15 and 21 circles can be arranged in the triangle of height 1, 2, 3, 4, 5 and 6, respectively (see figure).

We ask how many circles are needed to arrange them into the triangle of height n (this triangle is made of n layers, and k-th layer is composed of k triangles, when layers are counted from the top as the 1st, the 2nd, the 3rd, the 4th a.s.o.). Let's call this triangle as the n-th triangle and let's denote the number of circles completing the n-th triangle by T_n.

[3)] Q.E.D. (*quod erat demonstrandum*) means 'that which was to be demonstrated' and marks the end of the mathematical proof or philosophical argument. This is Latin analogous to Greek O.E.Δ. (ὅπερ ἔδει δεῖξαι) used, a.o., by Euclid and Archimedes. In some texts it is replaced by W[5] (which was what was wanted), in recent decades it is common to mark the end of the proof by the solid black square ■ (the symbol phoned as a mathematical tombstone, introduced by Paul Halmos and implemented in TeX by Donald Knuth).

An immediate observation that a $(n+1)$-th triangle is formed when to the n-th triangle there is joined the next layer and it contains $n+1$ circles, is mathematically memorized as the recursion [4)]

$$T_{n+1} = T_n + (n+1) \text{ for } n = 1, 2, 3, \ldots$$

and with $T_1 = 1$ (because 1 circle is a 1-st triangle).

It is very easy to see that, for arbitrary natural $n \geq 2$ there is

$$T_n = T_{n-1} + n = T_{n-2} + (n-1) + n = T_{n-3} + (n-2) + (n-1) + n = \ldots =$$
$$T_1 + 2 + \ldots + (n-2) + (n-1) + n = 1 + 2 + \ldots + (n-2) + (n-1) + n = \sum_{k=1}^{n} k,$$

so, by the formula proven above,

$$T_n = \frac{n \cdot (n+1)}{2}.$$

[4)] Here it is enough to say that a recursion wrt the quantity Q_n is the algebraic relation involving quantity Q_n and at least one Q_m with the index $m < n$. The distance between the most distant indexes is called the order of the recursion, so the formula $T_{n+1} = T_n + (n+1)$ is the recursion of the 1st order, and the equation $F_{n+1} = F_n + F_{n-1}$ is the recursion, or the recurrence, or the recursive equation, of order 2 (this is known as **Fibonacci recurrence**). As far as we know, first recursive formulas appeared in India in works of Panini (6th century BC) and Aryabhata (5th century BC).

Handshake problem

(m–1)-th triangular number solves the **handshake problem** which is defined as follows: how many handshakes, H_m, are made between m persons if every one of them shakes hands once with each other. Really, $H_2 = 1 = T_1$ (and, although it is not necessary, one can direct examine that there are $H_3 = 3 = T_2$ handshakes between 3 persons, $H_4 = 6 = T_3$ handshakes between 4 persons a.s.o.). Assuming that m persons do H_m handshakes, we see that with the arrival of the (m+1)-st person the number of handshakes rise up by m. It validates the recursive equation

$$H_{m+1} = H_m + m \text{ for } m = 1, 2, 3, \ldots$$

with $H_1 = 0$ (no handshake if there is only one person).

Therefore

$$H_m = H_{m-1} + m-1 = H_{m-2} + (m-2) + m-1 = H_1 + 1 + 3 + \ldots + (m-2) + m-1 =$$

$$= 1 + 2 + 3 + \ldots + (m-2) + m-1 = \sum_{k=1}^{m-1} k = T_{m-1} = \frac{m\cdot(m-1)}{2}.$$

A similar question to that forming the handshake problem concerns the **number of diagonals**, D_n, in a convex n-gon (and it does not restrict the generality to consider a regular n-gon). Obviously, in a triangle there is no diagonal, so $D_3 = 0$, in the square there are $D_4 = 2$ diagonals, in the pentagon there are $D_5 = 5$ diagonals, in a hexagon there are $D_6 = 9$ diagonals. The general formula is

$$D_n = \frac{n\cdot(n-3)}{2}.$$

It's easy to produce it when noticing that the problem is identical as that of handshakes:

a) n persons sitting at a table form n-gon,
b) there are H_n handshakes the persons do, so there are H_n linear segments between all n vertices of a n-gon,
c) between all segments starting at a vertex no.v there are two segments which are sides of considered n-gon,
d) $D_n = H_n - n$.

Notice that the formula $D_n = n\cdot(n-3)/2$ can be obtained without referring to H_n; indeed, from very vertex of n-gon there goes out n–3 diagonals, so the total number of diagonals is $n\cdot(n-3)/2$, and the division by 2 corresponds to the fact that in the presented approach every diagonal is taken into account twice (as it joins two vertices).

Binomial theorem and binomial coefficients

The formula

$$(a + b)^2 = a^2 + 2ab + b^2,$$

for the square of the sum of two scalars, a and b, is known (let's say) from ever [5]. It is easy to produce analogous formulas for $(a + b)^n$ with $n = 3, 4$ etc., e.g.

$$(a + b)^3 = (a + b)^2 \cdot (a + b) = a^3 + 3a^2b + 3ab^2\, b^3,$$
$$(a + b)^4 = (a + b)^3 \cdot (a + b) = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4,$$
$$(a + b)^5 = \ldots = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5.$$

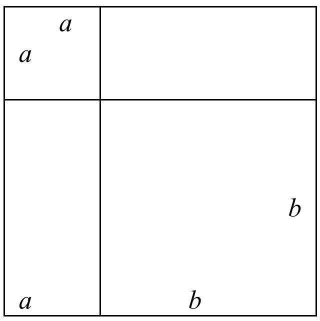
Let's denote the coefficient in the term a^kb^{n-k} by $C_{n,k}$ (at this moment let's say that we use the letter C as coefficient; later on we will see that it stands rather because of an other relation) it means that there is

$$(a + b)^n = \sum_{k=0}^{n} C_{n,k} \cdot a^k b^{n-k} .$$

This equality is stating what is known as a **binomial theorem**, made famous via *Traité du triangle arithmétique* (1665) written by Blaise Pascal, and the number $C_{n,k}$ is called a (n, k)-**binomial coefficient** (and it is English adaptation of Latin *binomium coefficent*) [6].

---

[5] Its geometrical proof is to draw the square with its side of length a+b and to divide it by two lines (to obtain two squares, of area a2 and b2, and two rectangles, each of area ab



[6] The binomial theorem is given, for several values of the exponent n, in Euclid's *Elements* (c. 300 BC). It, as well and its extension to higher exponents, is discussed in Chandaḥśāstra. This 8 chapter work was authored in India around 200 BC by Pingala. The commentary to Pingala's work, made by Halayudha in 10th century CE, gives a graphical presentation of Pingala recurrence (it is named as *Meru prastara* and is equivalent to the Pascal triangle). About 1150 a clear exposition of binomial coefficients was given by Bhaskaracharya in his book *Lilavati*. Independently of Halayudha similar results gave his contemporary Persian mathematician al-Karaji. Later binomial coefficients appeared in papers by Omar Khayyam (11th century), in Apianus', Stiefel's, Tartaglia's, Cardano's and Viéte's papers (16th century). Leibniz in his *Dissertatio de arte combinatoria* (1666) derived what is here named Pingall's recurrence.

Immediately from this definition it follows that

a) $C_{n,0} = C_{n,n} = 1$ for arbitrary $n \in \boldsymbol{N}_0$,

b) $C_{n,k} = C_{n,n-k} = 1$ for every $k \in \{0, 1, 2, \ldots, n\}$

c) $\sum_{k=0}^{n} C_{n,k} = 1$,

and by mathematical induction it can be stated that

d) $C_{n,k} = \binom{n}{k}$

– this is contemporary denotation of the equality stated by Blaise Pascal, it involves so-called combinatorial symbol and we will familiarize it later on, when discussing combinations.

Binomial coefficients $C_{n,k}$ can be arranged in a table (see Table 1) where every line corresponds to the value of the first index n and each column corresponds to the value of the second index k.

Table 1. First binomial coefficients

$C_{n,k}$	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	…
$n = 0$	1							
$n = 1$	1	1						
$n = 2$	1	2	1					
$n = 3$	1	3	3	1				
$n = 4$	1	4	6	4	1			
$n = 5$	1	5	10	10	5	1		
$n = 6$	1	6	15	20	15	6	1	
…	…	…	…	…	…	…	…	

In the table of numbers $C_{n,k}$ one can find such regularities as:

a) the column no.1 contains consecutive natural numbers,

b) the column no.2 contains consecutive triangular numbers,

c) a number staying in line no. n and column no. k equals the sum of two numbers staying in preceding line: that located directly above and its left neighbor.

A mathematical notation of the last above property is

$$C_{n,k} = C_{n-1,k-1} + C_{n-1,k}$$

for $n = 2, 3, 4, \ldots$ and $k = 1, 2, 3, \ldots, n-1$.

This is the recursion (but of other type than recursion for triangular numbers or Fibonacci recurrence; the recurrence at hand is spread over two indexes, n and k, and therefore it is qualified as a double recursion) and it is referred to as a **Pingala recurrence**, or a **Pingala recursion** [7)].

Table 2. Pascal triangle

										0		1		2		3	
n=0:									1								
n=1:								1		1							
n=2:							1		2		1						
n=3:						1		3		3		1					
n=4:					1		4		6		4		1				
n=5:				1		5		10		10		5		1			
n=6:			1		6		15		20		15		6		1		
n=7:		1		7		21		35		35		21		7		1	
n=8:	1		8		28		46		70		46		28		8		1
… …		…		…		…		…		…		…		…		…	

The Pingala recurrence is very well noticed when binomial coefficients are arranged to form the triangle (see Fig.2): n-th level of this triangle lists the binomial coefficients with the first index equal to n,

$$C_{n,0},\ C_{n,1},\ C_{n,2},\ \ldots,\ C_{n,n},$$

and for every fixed index k the numbers

$$C_{j,k},\ C_{j+1,k},\ C_{j+2,k},\ C_{j+3,k},\ \ldots$$

are written along k-th line which is parallel to the left side of this triangle (this left side itself is 0-th line). This configuration of binomial coefficients is commonly referred to as a **Pascal triangle** [8)].

[7)] We call it so to mention that it was written down by Pingala around 200 BC. In Europe this formula was derived by G.W.Leibniz in *Dissertatio de arte combinatoria* (1666).

[8)] In Europe this triangle appeared for the first time in *Traité du triangle arithmétique* written by Blaise Pascal in 1654 (edited posthumously in 1665). Its name, the Pascal triangle, was proposed by Montmort (*Essay d'analysis sur les jeux de hazard*, 1708), and it was accepted in most European countries (but in Italy it is referred to as Tartaglia triangle). In India it was composed by Halayudha (as the visualization, called *Meru prastara*, of Pingala recurrence). It seems that in Persia and China the binomial theorem was discovered around 1100, in China the Pascal triangle is known as Yang Hui's triangle after the author of the work *Jiuzhang suan fa zuan lei* (*Reclassification of the mathematical procedures in nine chapters*, c.1275), where he presents first seven rows of Pascal triangle (and informs that it is copied form Jia Xian's *Huangdi jiuzhang suanfa* (*Yellow Emperor's nine chapters on mathematical methods*, c.1050).

Now

the line no.0 (i.e., the left side of the triangle) is formed from 1s only,
the 1st line lists all naturals,
consecutive elements in the 2nd line are triangular numbers.

A deeper insight into Pascal triangle reveals many interesting dependencies. As an example notice that by putting

$n = C_{4,1} = 4$ points on a circle we can join them to form $C_{4,2} = 6$ segments (4 sides and 2 diagonals), $C_{4,3} = 4$ triangles and $C_{4,4} = 1$ quadrilateral (which is a square if our 4 points are distributed equidistantly),

$n = C_{5,1} = 5$ points on a circle we can join them to form $C_{5,2} = 10$ segments (5 sides and 5 diagonals), $C_{5,3} = 10$ triangles, $C_{5,4} = 4$ quadrilaterals and $C_{5,5} = 1$ pentagon.

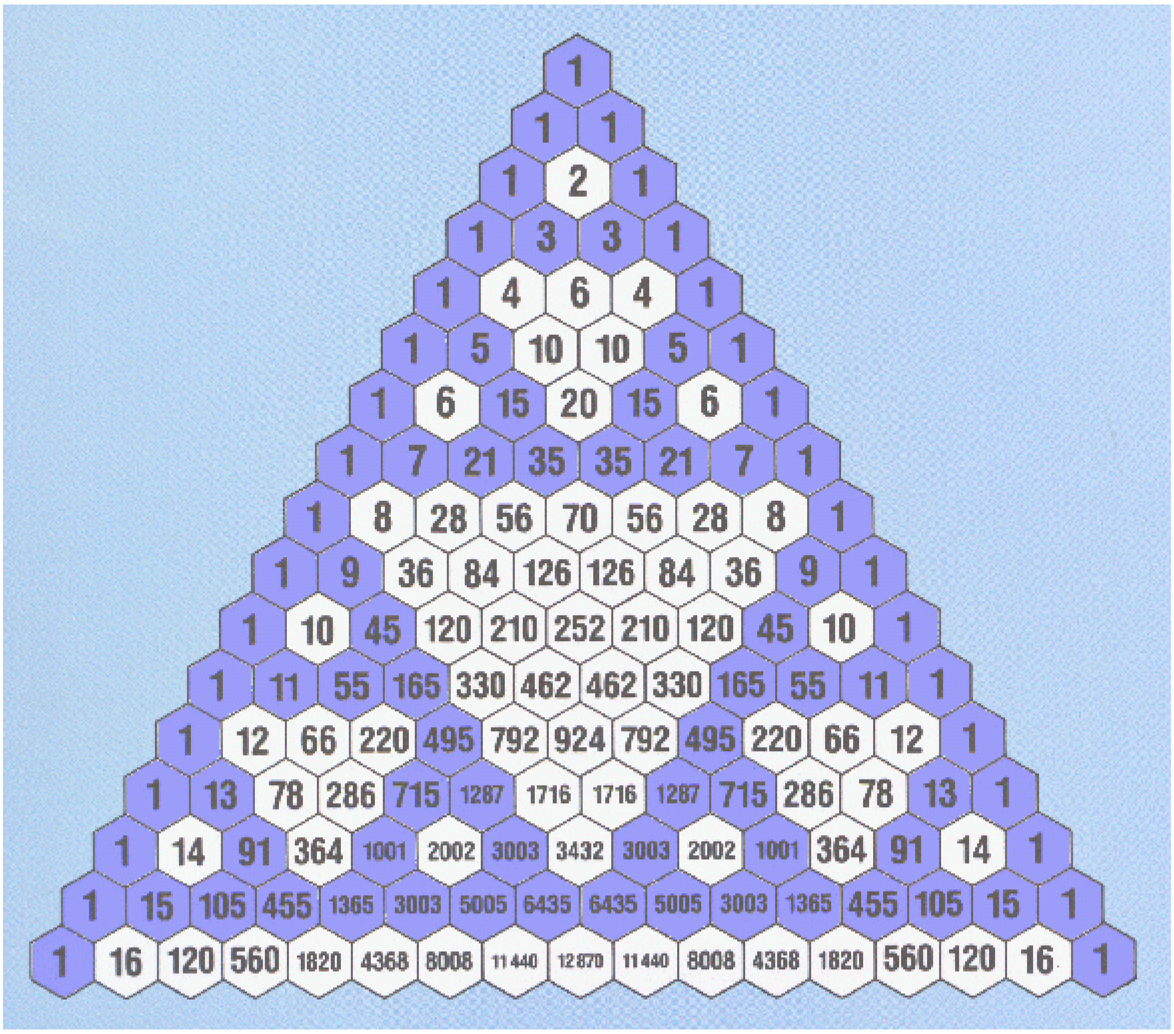


Fig.1. Pascal triangle (more precisely: its 17 top rows) built of hexagons, which are colored blue if contain odd binomial coefficients, and white otherwise
http://malsmath.blogspot.com/2006/12/pascals-triangle-and-prime-numbers.html

Pascal triangle can be presented via appropriately arranged isosceles elementary triangles, in every triangle there is sitting a binomial coefficient, or (see Fig.1.) as an infinite triangle composed of an infinite honeycomb, in every hexagon there is sitting a binomial coefficient. When we replace every odd number by the same color (i.e., when we color all numbers which are divisible by 2) and we white every hexagon where there is an even coefficient $C_{n,k}$, we find a structure similar to Sierpiński triangle, aka Sierpiński gasket (created by Wacław Sierpiński and desribed in the paper *Sur une courbe dont tout point est un point de ramification*, 1915, which was also published in Polish: *O krzywej, której każdy punkt jest punktem rozgałęzienia* 1916), although this pattern appeared in 13th century mosaic in the cathedral of Agnani (in the province Latium) and in Rome, in the *Basilica di Santa Maria in Cosmedin*, aka *S.Maria de Schola Graeca*. This is a self-similar structure, a fractal set. Maybe the most known such set is the Mandelbrot set, its name recall Benoit Mandelbrot who investigated it in 1980. It was Mandelbrot who advocated to use the name 'Sierpiński triangle'

Recall that the binomial coefficients satisfy the Pingala recurrence,

$$C_{n,k} = C_{n-1,k-1} + C_{n-1,k},$$

with boundary values $C_{n,0} = C_{n,0} = 1$. Both the Pingala recurrence and boundary conditions are also satisfied by numbers called as combinatorial symbols: a (n, k)-th combinatorial symbol is denoted (after Andreas von Ettinghausen's proposal presented in 1826) as

$$\binom{n}{k} := \frac{n\,!}{k\,! \cdot (n-k)\,!}$$

(and it can be read 'n over k').

Another way in which binomial coefficients can be arranged is presented in Table 3. Here we have an infinite array, its left column (said to be the column no.0) and upper row are filled with numbers $C_{n,0} = 1$ and $C_{n,n} = 1$, resp., and the numbers $C_{n,0}$, $C_{n,1}$, $C_{n,2}$, …, $C_{n,n}$ are listed along the diagonals joining values $C_{n,0}$ and $C_{n,n}$. The table of numbers $C_{n,2}$ arranged in this way can be called a **binomial rectangle**, or a **Pascal rectangle**.

It has a direct interpretation which is easily illustrated on two orthogonal families: one containing street and the other containing avenues, just as a little idealized Manhattan transport network, where there are streets and avenues, street runs latitudinally (east-west direction) [9], while avenues run longitudinally (south-north direction), so cross streets perpendicularly. Staying at the corner of

[9] to be precise: in Manhattan the most of streets and avenues cross perpendicularly (and Broadway, probably the most famous Manhattan street, in its southern part, close to Central Park, does not observe this perpendicularity), the directions are not exactly east-west and south-north, they are rotated by about 30°

an avenue and a street, we can ask how many shortest pathways are to reach another corner. Let's say that we are on the corner where the avenue no.0 meets the street no.0, and we aim to pass to the corner of the avenue no.4 and the street no.2. The shortest pathway is only if we move downwards and/or to the east. To arrive to the intersection of 4th avenue and 2nd street we have to move only The number of pathways is exactly $C_{4+2,2} = 15$.

Table.3. Rectangular arrangement of binomial coefficients.
There is distinguished 6th diagonal (n-th diagonal is the line joining cells described as n: and :n ; the index n runs 0, 1, 2, 3, …).
Along n-th diagonal there are placed coefficients $C_{n-k,k}$ (with k running from 0 up to n).
For example, $C_{6,1} = 15$ stays on the 6th diagonal (it joins 6: and :6) and in the column no.2 (it is marked as :2, because columns are counted from 0)

$C_{n,k}$		:0	:1	:2	:3	:4	:5	:6	:7	: ...
	1	1	1	1	1	1	1	1	1	…
0:	1	2	3	4	5	6	7	8	9	…
1:	1	3	6	10	15	21	28	36	45	…
2:	1	4	10	20	35	56	84	120	165	…
3:	1	5	15	35	70	126	210	330	495	…
4:	1	6	21	56	126	252	462	792	1287	…
5:	1	7	28	84	210	462	924	1716	3003	…
6:	1	8	36	120	330	792	1716	3432	6435	…
7:	1	9	45	165	495	1287	3003	6435	12870	…
8:	1	10	55	220	715	2002	5005	11440	24310	…
9:	1	11	66	286	1001	3003	8008	19448	43758	…
10:	1	12	78	364	1365	4368	12376	31824	75582	…
11:	1	13	91	455	1820	6188	18564	50388	125970	…
... :	…	…	…	…	…	…	…	…	…	…

Newton coefficients

Let's finish with the information that the binomial theorem written with $(a, b) = (1, x)$ and with the use of combinatorial symbols reads

$$(1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k .$$

It says that binomial coefficients are coefficients in the representation of the function $x \to (1 + x)^n$ in the Stevin basis (recall: Stevin basis is the system of functions $x \to x^k$, where the exponent k assumes nonnegative integer values). This is equivalent to the statement that binomial coefficients are coefficients in the Maclaurin expansion of the function $x \to (1 + x)^n$.

Probably in 1665, while sojouring in Woolsthorpe, Isaac Newton, highly influenced by calculations performed by John Wallis [10)], realized that the binomial theorem writes down almost identically when instead of n there sits a real number r, i.e., he produced Maclaurin series and wrote (the equality which will be later on referred to as a **Newtonian expansion**)

$$(1 + x)^r = \sum_{k=0}^{\infty} \binom{r}{k} \cdot x^k ,$$

where $\displaystyle\binom{r}{k} := \prod_{j=0}^{k-1} \frac{r-j}{k-j} = \frac{r}{k} \cdot \frac{r-1}{k-1} \cdot \frac{r-2}{k-2} \cdot \ldots \cdot \frac{r-k+2}{2} \cdot \frac{r-k+1}{1}$,

e.g. $$\binom{\frac{7}{4}}{3} = \frac{\frac{7}{4} \cdot \left(\frac{7}{4} - 1\right) \cdot \left(\frac{7}{4} - 2\right)}{3 \cdot 2 \cdot 1} = \frac{-7}{128},$$

$$\binom{\frac{7}{5}}{4} = \frac{\frac{7}{5} \cdot \left(\frac{7}{5} - 1\right) \cdot \left(\frac{7}{5} - 2\right) \cdot \left(\frac{7}{5} - 3\right)}{4 \cdot 3 \cdot 2 \cdot 1} = \frac{14}{625}.$$

---

10) John Wallis (1616-1703) was made famous by his *Treatise on the conic sections* (1655), *Arithmetica infinitorum* (1656) and *Opera Mathematica* (1695). In the first book he presented analytical definitions of conics (and popularized the use of symbol ∞), in the next he generalized facts observed by Archimedes, by Ibn al-Haytham (aka Alhazen) reported in his 7-volume *Kitab al-manazir* (*Book of optics*, c.1000 AD) and by Bonaventura Cavalieri (discussed in *Geometria indivisibilibus continuorum nova quadam ratione promota*, 1635) and gave the formula which in today symbolism is $\int_0^1 x^m \, dx = \frac{1}{m+1}$ with $m \neq -1$. Starting with the sequence $\int_0^1 (1-x^2)^0 \, dx = 1$, $\int_0^1 (1-x^2)^1 \, dx = \frac{2}{3}$, $\int_0^1 (1-x^2)^2 \, dx = \frac{8}{15}$, … and applying the interpolation he found that $\frac{\pi}{2} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdot \ldots}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdot \ldots}$, which is now known as a **Wallis product**.

The Newtonian expansion was produced by Newton [11)] and this fact is remembered in another name given to coefficients $C_{n,k}$ – they are aka **Newton coefficients**. Newton himself did not occupy with the convergence of the series generated by the function

$$x \to (1+x)^r .$$

The convergence was discussed by Euler and Gauss, and (for a complex argument x) by Abel who ultimately stated that it holds if $|x| < 1$.

Plato is my friend, Aristotle is my friend, but my best friend is truth." --Head of Newton's *Quaestiones Quaedam Philosophicae* (Certain Philosophical Questions), ca. 1664.

11) Newton found this expansion when staying in family's country manor in Woolstrophe, in 1665 (so only 2 years after he started to be interested in mathematics he started to learn in Cambridge). In the same time he presented the expansion for

$$(P+PQ)^{m/n} = P^{m/n} + \frac{m}{n}AQ + \frac{m-n}{2n}BQ + \frac{m-2n}{3n}CQ + + \frac{m-3n}{4n}DQ + \ldots,$$

where each of A, B, C, D, … represents the previous term, so that

$$(1+Q)^{m/n} = 1 + \frac{m}{n}Q + \frac{\frac{m}{n}\cdot\left(\frac{m}{n}-1\right)}{2\cdot 1}Q^2 + \frac{\frac{m}{n}\cdot\left(\frac{m}{n}-1\right)\cdot\left(\frac{m}{n}-2\right)}{3\cdot 2\cdot 1}Q^3 + \ldots$$

Newtonian, falling and rising powers

Considered as functions of r, the binomial coefficients $\binom{r}{k} = \prod_{j=0}^{k} \frac{r-j}{k-j}$ define the function called a k-th **Newtonian power** of r. Replacing r by x (in fact, x is the standard letter used to denote the argument of the function) we have

$$\binom{x}{k} = \prod_{j=0}^{k} \frac{x-j}{k-j} = \frac{\prod_{j=0}^{k}(x-j)}{\prod_{j=0}^{k}(k-j)}.$$

In the last quotient the denominator does not depend on x, while the nominator is the polynomial, of degree $k+1$ in variable x. This polynomial is the product of factors $x–j$ and each next factor is smaller by 1 than the previous one (equal to x); we can tell that in consecutive factors x falls down by 1. The polynomial defined in this way is called a falling power; more precisely, for any natural n a n-th **falling power** of x, or a n-th falling factorial power, or a falling power of degree n, or a **descending power**, is the polynomial in x defined by the formula

$$x^{\underline{n}} := \prod_{j=0}^{n-1} (x-j)$$

(read: 'x to the n falling'); additionally, a falling power of degree 0 is set to be 1,

$$x^{\underline{0}} := 1.$$

Analogously, for any natural n a n-th **rising power** of x, or a n-th raising factorial power, or a rising power of degree n, or an **ascending power**, is the polynomial in x defined by the formula

$$x^{\overline{n}} := \prod_{j=0}^{n} (x+j)$$

(read: 'x to the n rising'); additionally, a rising power of degree 0 is set to be 1,

$$x^{\overline{0}} := 1.$$

n-th rising power is aka a **Pochhammer symbol** [12)] and denoted by $(x)_n$.

[12)] Leo August Pochhammer (1841-1920) famous for his works on vibrations of circular cylinders and bending of beams, as well as for the introduction, in 1870, of generalized hypergeometric functions, intensively used rising powers $(x)_n$. The term 'Pochhammer symbol' was proposed by P.E.Appell. Rising powers were first considered by A.L.Crelle in 1831.

With K denoting a field (in practice, K is the set $\boldsymbol{Q}$ of rational numbers, the set $\boldsymbol{R}$ of reals, or the set $\boldsymbol{C}$ of complex numbers), $K_n[x]$ denotes the space completed by polynomials whose coefficients are taken from K and have degree up to n. Later on we deal with $K = \boldsymbol{R}$, so we have the space $\boldsymbol{R}_n[x]$ of real polynomials in x of degree up to n. This is a $(n+1)$-dimensional linear space, so every its element is uniquely represented, via the sequence of $n+1$ real numbers, in arbitrary basis of the space $\boldsymbol{R}_n[x]$. Exemplary bases are:

a) a **standard basis**, aka a **natural basis** or a **Stevin basis**;
   it is composed of natural powers: $1, x, x^2, \ldots, x^n$;

b) a **Newtonian basis**; its elements are: $\binom{x}{0} = 1$, $\binom{x}{1} = x$, $\binom{x}{2}, \ldots, \binom{x}{n}$;

c) a **descending power basis**, aka a **falling power basis**;
   its elements are $x^{\underline{0}} = 1$, $x^{\underline{1}} = x, \ldots, x^{\underline{n}} = x \cdot (x-1) \cdot \ldots \cdot (x-n+1)$;

d) an **ascending power basis**, aka a **rising power basis** or a **Pochhammer basis**; it consists of $x^{\overline{0}} = 1$, $x^{\overline{1}} = x, \ldots., x^{\overline{n}} = x \cdot (x+1) \cdot \ldots \cdot (x+n-1)$.

It is easy to notice that falling and rising powers are related to Newtonian powers via the equalities

$$x^{\overline{n}} = n! \cdot \binom{x+n-1}{n}, \quad x^{\underline{n}} = n! \cdot \binom{x}{n},$$

and the binomial theorem fully conserves its form when written with the use of falling powers and of rising powers,

$$(a+b)^{\overline{n}} = \sum_{j=0}^{n} \binom{n}{j} \cdot a^{\overline{n-j}} \cdot b^{\overline{j}}, \quad (a+b)^{\underline{n}} = \sum_{j=0}^{n} \binom{n}{j} \cdot a^{\underline{n-j}} \cdot b^{\underline{j}}.$$

Both above relations, with a and b replaced by 1 and x, resp., are representations
of the $1+x$ to the n falling in the falling power basis
and of the $1+x$ to the n rising in the rising power basis,

namely
$$(1+x)^{\overline{n}} = \sum_{j=0}^{n} \binom{n}{j} \cdot x^{\overline{j}}, \quad (1+x)^{\underline{n}} = \sum_{j=0}^{n} \binom{n}{j} \cdot x^{\underline{j}},$$

and it exhibits that binomial coefficients are coefficients of these representations. In the next we recognize the coefficients of rising powers in Stevin basis and, inversely, the coefficients of standard powers in the rising power basis.

Stirling numbers

Coefficients in the representation of the n-th rising power in the Stevin basis are denoted by $s_{n,k}$, or by $\begin{bmatrix} n \\ k \end{bmatrix}$, and are called **Stirling numbers of the first kind** [13]. Thus

$$x^{\overline{n}} = \sum_{k=0}^{n} s_{n,k} \cdot x^k = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} \cdot x^k$$

and, for instance, with $n = 5$ we have

$$x^{\overline{6}} = \sum_{k=0}^{6} s_{6,k} \cdot x^k ,$$

or

$x \cdot (x+1) \cdot (x+2) \cdot (x+3) \cdot (x+4) \cdot (x+5) = x^6 + 15x^5 + 85x^4 + 225x^3 + 274x^2 + 120x$,

so

$$s_{6,0} = 0,\ s_{6,1} = 120,\ \ s_{6,2} = 274,\ \ s_{6,3} = 225,\ \ s_{6,4} = 85,\ \ s_{6,5} = 15,\ \ s_{6,6} = 1,$$

Coeffcicients in the representation of the n-th falling power in the Stevin basis are denoted by $S_{n,k}$ or by $\begin{Bmatrix} n \\ k \end{Bmatrix}$, and are called **Stirling numbers of the second kind**. Thus

$$x^n = \sum_{k=0}^{n} s_{n,k} \cdot x^{\underline{k}} = \sum_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} \cdot x^{\underline{k}}$$

and, for instance, with $n = 4$ we have

$x^4 = s_{4,0} \cdot 1 + s_{4,1} \cdot x + s_{4,2} \cdot x \cdot (x–1) + s_{4,3} \cdot x \cdot (x–1) \cdot (x–2) + s_{4,4} \cdot x \cdot (x–1) \cdot (x–2) \cdot (x–3)$

and for $x = 0, 1, 2, 3$ and 4 it gives

$$\begin{aligned}
0^4 &= s_{4,0} \cdot 1, \\
1^4 &= s_{4,0} \cdot 1 + s_{4,1} \cdot 1, \\
2^4 &= s_{4,0} \cdot 1 + s_{4,1} \cdot 2 + s_{4,2} \cdot 2 \cdot 1, \\
3^4 &= s_{4,0} \cdot 1 + s_{4,1} \cdot 3 + s_{4,2} \cdot 3 \cdot 2 + s_{4,3} \cdot 3 \cdot 2 \cdot 1, \\
4^4 &= s_{4,0} \cdot 1 + s_{4,1} \cdot 4 + s_{4,2} \cdot 4 \cdot 3 + s_{4,3} \cdot 4 \cdot 3 \cdot 2 + s_{4,4} \cdot 4 \cdot 3 \cdot 2 \cdot 1.
\end{aligned}$$

[13] James Stirling presented these numbers, as well as that now called Stirling numbers of the second kind (and later on in this chapter denoted as $S_{n,k}$), in his treatise *Methodus differentialis* (1749), which we mentioned of when talking about the approximation of n !. Before Stirling these numbers appeared, a.o., in Thomas Harriot manuscript dated about 1600. The name 'Stirling numbers' was coined by Niels Nielsen in his paper *Recherches sur les polynomes et les nombres de Stirling* (1904) nad disseminated by his *Handbuch der Theorie der Gammafunktion* (1906)

This system of 5 equations is solved by

$$s_{4,0} = 0,\ s_{4,1} = 1,\ s_{4,2} = 7,\ s_{4,3} = 6,\ s_{4,4} = 1.$$

From the definitions of Stirling numbers of the first kind, $s_{n,k}$, and that of the second kind, $S_{n,k}$, it follows that they satisfy, for every natural n and with $k \in \{1, 2, \ldots, n-1\}$, following recursive relations:

$$s_{n,k} = s_{n-1,k-1} + (n-1) \cdot s_{n-1,k},$$
$$S_{n,k} = S_{n-1,k-1} + k \cdot S_{n-1,k},$$

with (so-called boundary values)

$$s_{n,0} = S_{n,0} = \delta_{n,0},\ \ s_{n,n} = S_{n,n} = 1$$

valid for $n \in \boldsymbol{N}_0$.

Both recurrences are similar to the recursion

$$b_{n,k} = b_{n-1,k-1} + b_{n-1,k}$$

which, with $b_{n,0} = b_{n,n} = 1$, holds true for binomial coefficients, $b_{n,k} = \binom{n}{k}$.

Similarly, Stirling numbers $s_{n,k}$ and $S_{n,k}$ can be arranged in triangles. These triangles are presented in Table 2 and Table 3 in the next page.

There are numerous formulas which are analogous to the above three ones and define families of numbers. For example, the family of numbers $A_{n,k}$ is defined by relations

$$A_{n,0} = A_{n,n} = 1 \text{ for } n = 0, 1, 2, \ldots,$$
$$A_{n,k} = (n-k) \cdot A_{n-1,k-1} + (k+1) \cdot A_{n-1,k} \text{ for } k = 1, 2, \ldots, n-1.$$

Numbers An,k are called Euler ascent numbers, we will discuss them later on.

Table 2. Triangle listing Stirling numbers of the first kind, $s_{n,k}$ for $n = 0, 1, 2, 3, 4, 5, 6$

								0		1		2		3
n=0:							1							
n=1:						0		1						
n=2:					0		1		1					
n=3:				0		2		3		1				
n=4:			0		6		11		6		1			
n=5:		0		24		50		35		10		1		
n=6:	0		120		274		225		85		15		1	
…		…		…		…		…		…		…		…

for instance, $s_{5,2} = s_{4,1} + 4 \cdot s_{4,2} =$
$= 3! + 4 \cdot (s_{3,1} + 3 \cdot s_{3,2}) =$
$= 6 + 4 \cdot (2! + 3 \cdot (s_{2,1} + 2 \cdot s_{1,1})) =$
$= 6 + 4 \cdot (2 + 3 \cdot (1! + 2 \cdot 1)) = 50.$

Table 2. Triangle listing Stirling numbers of the second kind, $S_{n,k}$ for $n = 0, 1, 2, 3, 4, 5, 6$

								0		1		2		3
n=0:							1							
n=1:						0		1						
n=2:					0		1		1					
n=3:				0		1		3		1				
n=4:			0		1		7		6		1			
n=5:		0		1		15		25		10		1		
n=6:	0		1		31		90		65		15		1	
…		…		…		…		…		…		…		…

for instance, $S_{5,2} = S_{4,1} + 2 \cdot S_{4,2} =$
$= 1 + 2 \cdot (S_{3,1} + 2 \cdot S_{3,2}) =$
$= 1 + 2 \cdot (1 + 2 \cdot (S_{2,1} + 2 \cdot S_{1,1})) =$
$= 1 + 2 \cdot (1 + 2 \cdot (1 + 2 \cdot 1)) = 15.$