A DIDACTIC UNIT ON MATHEMATICS AND SCIENCE EDUCATION: THE PRINCIPLE OF MATHEMATICAL INDUCTION

Paolo Bussotti
University of Udine, Italy

Among the mathematical methods which are taught in the last years of almost every high school, the mathematical induction deserves particular attention. It can be used both to define mathematical entities and to prove theorems. The second use is more common at high school level and is easier. Thus, I will basically focus on it, though analysing in depth two definitions by induction. The aim of this contribution is to offer the basic elements for a didact unit which could be developed in six/seven hours of lesson.

As a method of proof, the principle of mathematical induction is formulated as follows. *Base*: if a property *P* holds for the number 1 and if, *step of the induction*, $P(n) \rightarrow P(n+1)$, then *P* holds for every natural number. The reasons of interest behind this method are:

1) its meaning and its connection to infinity;

2) its link with empirical induction and with science education;

3) its use;

4) equivalent formulations of the principle.

All these four items are of considerable educational importance: 1) and 2) for general conceptual reasons; 3) because of specific mathematical reasons; 4) for the connection between logic and mathematics.

Meaning of the Principle of Mathematical Induction and Its Connection with Infinity

The intuitive idea upon which the principle of mathematical induction relies can be explained in this way: suppose that a theorem *T* concerning natural numbers has to be proved and also suppose that, by means of practical calculations, we know that this theorem is true for the number 1 or for a set of initial numbers. Under these conditions, why is one allowed to deduce that it is sufficient to prove the inference from *T*(*n*) to *T*(*n*+1) to be sure that *T* holds for all the natural numbers? A student is legitimate to ask this question. Suppose that one has proved that $T(n) \rightarrow T(n+1)$. Name n' the number *n*+1. Then, with the same reasoning used to prove the inference from *T*(*n*) to *T*(*n*+1), one infers $T(n'+1)$ from $T(n')$. But $n'+1 = n+2$. It is now obvious how to continue: one poses $n'' = n'+1$ and applies the same reasoning so to reach $n''+1 = n'+2 = n+3$. In this way every number is reached, so that the theorem holds for every natural number. This is the intuitive meaning of the principle of mathematical induction.

The connection with infinity is a fundamental trait of this method. As a matter of fact, if we were interested in proving a theorem only for a finite set of numbers, we could proceed by verifying its validity for every single number. A long time is necessary if the set of numbers has a great dimension, or if the property we want to prove

https://doi.org/10.33225/jbse/23.22.04

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

implies complicate calculations, but, in principle, nobody could prevent us from using such technique. This is not true anymore when the set of numbers in which we are interested is infinite. In this situation, to prove our theorem for single cases, however numerous these cases might be, is useless; we need a general method of proof, as mathematical induction. This is the reason why such a procedure is strictly connected to the potential infinity of the series of whole numbers. This aspect has to be highlighted in an educational context.

Link between Mathematical Induction and Empirical Induction

This issue concerns the relations between the term "induction" in mathematics and in sciences. Therefore, its development is significant for the learners to understand the analogies and the differences between the logic behind mathematics and empirical science. Thus, it is crucial not only for the reflection on maths education, but for science education, as well. Empirical induction is based on the attempt to obtain a general law from a series of single cases which take place without any exception. This perfectly makes sense in a physical context (with regard to the relation between conceptual and empirical aspects of physics education, see, e.g., Bussotti, 2021; Chen, et al., 2020; Truyol, et al., 2014). For example, given some well-known conditions, mechanic energy is conserved in all the numerous cases we have experienced. Therefore, we are allowed to induce the conservation of energy to be a general law. Of course, in mathematical physics many propositions are deduced as theorems like in abstract mathematics, but the truth of the basic principles is inferred by induction. It is paramount that the learners understand that such an inductive principle has no value in mathematics. Specifically, in the theory of numbers, it is useless to know that a proposition holds without exception for all the single cases we can calculate. A formal and general proof is necessary. This clarified, it is spontaneous to wonder why a mathematical method includes the name "induction". The answer is that our procedure can be applied only if the proposition *P* we are going to prove is empirically valid for a base, which is in general, but not necessarily, the number 1, namely *P*(1) holds. This is the inductive aspect: a concrete case in which *P* is valid has to be found. For let us suppose that one is able to infer the truth of $P(n+1)$ from the truth of $P(n)$, but is unable to verify that a *k* exists such that $P(k)$ holds. What could he infer? He has proved only a hypothetical proposition: if $P(n)$ is true, then $P(n+1)$ is true, but no categorical one. In contrast to this, if one is able to show that *P*(*k*) is true and that, for *n*>*k*, $P(n) \rightarrow P(n+1)$, then *P* is true for every number bigger than *k*, given the arbitrariness of *n* and the reasoning expounded in the previous section. If *k*=1, *P* is true for every integer. Therefore, the term "induction" is justified by the fact that the proposition to prove has to be empirically valid for a number (generally 1). The conjunction of the inductive basis and the inferential step from *n* to *n*+1 makes mathematical induction a true demonstrative method, which has only an analogical connection with the empirical induction, but whose logic is completely different.

Use of the Principle of Mathematical Induction

Let us see three, among numerous, examples which show how to use the principle of induction. it is advisable for each teacher to present examples with an increasing degree of difficulty in order to train students to reason by induction and to make them understand the wide use of this method.

1) Prove that for every natural number *n*≥1, the sum of the first *n* squares is equal to

$$\frac{n(n+1)(2n+1)}{6}$$

Base: $1^2 = \frac{1 \cdot 2 \cdot 3}{6}$ with *n*=1

Step: Suppose that the identity $\sum_{k=1}^{n-1}(k-1)^2 = \frac{(n-1)n(2n-1)}{6}$ holds. Add n^2 to each member. On the left one obtains $\sum_{k=1}^{n} k^2$. On the right, it is

$$\frac{(n-1)n(2n-1)+6n^2}{6} = \frac{n(2n^2-3n+1)+6n^2}{6} = \frac{2n^3+3n^2+n}{6} = \frac{n(n+1)(2n+1)}{6}.$$

Therefore, the theorem is proved.

2) Prove that for every natural number *n*≥1, $2^{n-1} \leq n!$

Base: $2^0 = 1 = 1!$, with *n*=1

Step: For inductive hypothesis, it is $2^{n-2} \leq (n-1)!$ Multiply each member by 2. One obtains $2^{n-1} \leq 2(n-1)!$. Therefore, a fortiori, $2^{n-1} \leq n(n-1)! = n!$ This proves the proposition.

https://doi.org/10.33225/jbse/23.22.04

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

3) Prove that, for every natural number *n*>1, it is $\left\{ r_n := \underbrace{\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}}_{n \text{ times}} \notin \mathbb{Q} \right.$

Base: if *n*=2, one obtains $\sqrt{2}$, which is not a rational number. Note that in this case, because of the conditions of the problem, the inductive basis cannot be referred to *n*=1.

Step: Suppose the proposition to be true for *n* and let us prove that it is also true for *n*+1.

Take into account that the following identity holds:

$$\left\{ r_{n+1} := \underbrace{\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}}_{(n+1) \text{ times}} = \sqrt{r_n + 1} \right.$$

Suppose *ad absurdum* that r_{n+1} is a rational number and set $r_{n+1} = \frac{p}{q}$, so that

$$\frac{p^2}{q^2} = r_{n+1}^2 = 1 + r_n$$

Therefore, it would be $\frac{p^2}{q^2} - 1 = r_n$. But this is absurd because, for inductive hypothesis, r_n is an irrational number. Thus, the proposition is true.

This example is interesting because it shows that the induction basis is not necessarily referred to *n*=1 and that mathematical induction can also be associated with an *ad absurdum* deduction.

In order to show the richness of the concept of mathematical induction, let us see two easy examples of series defined by induction or by recurrence. In this case our principle is used to define an object rather than to prove a theorem.

1) In the 12th chapter of his celebrated *Liber Abaci* (1202, second edition 1228) Leonardo Pisano told Fibonacci (about 1175-1245) posed the following problem: "How many couples of rabbits are produced in one year starting from a single couple, if every month each couple gives birth to a couple which becomes fertile from the second month of her life"?

In January there is an only couple. Since she is not fertile, in February there is a single couple, as well. In March she becomes fertile, so that we have two couples. In April, the first couple is fertile, but the second is not, so that we have three couples. In May the first two couples are fertile, but the third is not. Thence, we have five couples: the three existing in April, that generated by the first couple and that generated by the second couple. The logic of the reasoning is now clear: the number of couples existing at the month *m* is given by the sum of the number of couples existing at the month *m*-1 plus the number existing at the month *m*-2. Let us indicate the general term of this progression by f_i, where *i* indicates the number of months. Such a progression can be defined by induction or recurrence as follows:

$$\begin{cases} f_1 = f_2 = 1 \\ f_n = f_{n-1} + f_{n-2} \end{cases}$$

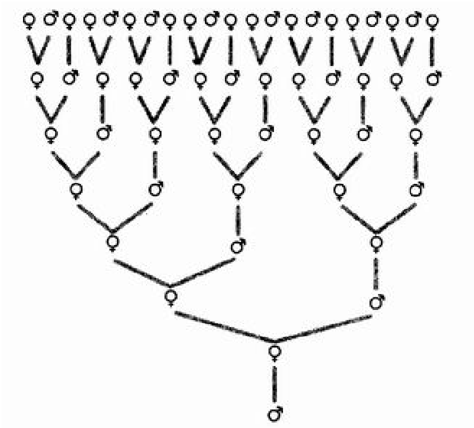
Thus, the numbers of the sequence are 1,1,2,3,5,8,13,21,34,55, …

This is a classical case of a definition by recurrence because $f_1 = f_2 = 1$ is the inductive base of the definition and $f_n = f_{n-1} + f_{n-2}$ is the inductive step.

The series of Fibonacci has many interesting applications which are of educational interest because they connect mathematics with other sections of science. Let us consider the following example: among the bees, the drones are generated for parthenogenesis by the queen, while the females are generated by the queen mating with a drone. Thus, the drones have one parent, the females have two parents. A drone has two grandparents – a female and a male –, three great-grandparents – two females and one male –. Therefore, a drone has f_n predecessors of *n*-th generation of which f_{n-1} are females and f_{n-2} are males, if f_1 is considered the drone itself. Thus, the genetic tree of a drone is represented in each branch by a number belonging to the Fibonacci sequence, as also clarified in Figure 1.

https://doi.org/10.33225/jbse/23.22.04

Figure 1
The Genetic Tree of a Drone



From a mathematical point of view, the Fibonacci numbers have interesting properties. For example, two consecutive of them are mutually prime. They are also connected with the golden ratio: a segment is defined to be divided in extreme and means ratio if the ratio between the bigger and the smaller part of the segment is equal to that between the whole segment and its bigger part. If we indicate by *x* the bigger part of the segment, by 1 the smaller, the whole segment is *x*+1, thus the required proportion is

$$\frac{x}{1} = \frac{x+1}{x}$$

which is the equation $x^2 - x - 1 = 0$, whose positive solution is $x = \frac{1+\sqrt{5}}{2}$. If this number is developed into a continuous fraction, one can check that the numerator of each convergent is the *i*-th Fibonacci number and its denominator the (*i*-1)-th Fibonacci number.

Thus, this progression defined by recurrence has many applications in mathematics and in scientific branches different from mathematics, which is an important educational aspect that is appropriate to develop.

2) Let us consider the series defined as follows:

$$\begin{cases} x_0 = 0 \\ x_{n+1} = \sqrt{2 + x_n} \end{cases}$$

Many properties can be proved on this series relying upon its definition by recurrence.

First of all, it can be proved that the series is well defined, namely that $2 + x_n$ is not negative. This is true thanks to a reasoning by induction because $x_0 \geq 0$ and if $x_n \geq 0$, then $x_{n+1} = \sqrt{2 + x_n}$ is well defined as $2 + x_n > 0$.

A further information we can infer by induction is that each term of this series is less than 2. For, $x_0 < 2$ and if $x_n < 2$, then $x_{n+1} = \sqrt{2 + x_n} < \sqrt{2 + 2} = 4$.

Let us now prove that our series is increasing: indeed $x_n < x_{n+1}$ means that $x_n < \sqrt{2 + x_n}$, that is $x_n^2 - x_n - 2 < 0$, which is the case if $-1 < x_n < 2$, and this is true because every term of the series is less than 2.

Thus, this series is positive, increasing and is superiorly limited by 2. Therefore, its limit is a number $L \leq 2$. The only possible values of *L* are obtained by solving the equation $L = \sqrt{2 + L}$, that is $L^2 - L - 2 = 0$, whose solutions are $L = 2$ and $L = -1$. Since the sequence is positive, its limit is *L*=2 (example drawn from Finzi-Vita, 2009).

This example is important from an educational point of view because it shows the profound link between the definitions by recurrence and the proofs by mathematical induction. For the features of this series, which is defined by recurrence, are obtained through mathematical induction.

https://doi.org/10.33225/jbse/23.22.04

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Equivalent Formulations of the Principle

The concept of logical equivalence is significant from a conceptual and educational standpoint. Two demonstrative methods are logically equivalent if, given a proof through a method, it is possible to transcribe it in a proof by the other method and reciprocally. The relation between logical equivalence and the real use of two logically equivalent methods in mathematics is a complicated topic, which is not appropriate to face in this context. Therefore, I only mention three equivalent formulations of mathematical induction: 1) infinite descent; 2) strong induction; 3) least number principle.

Infinite descent. It is an *ad absurdum* reasoning: a theorem *T* concerning natural numbers has to be proved. Suppose *T* to be false and that, hence, $\neg T$ is true. If this assumption implies that between two natural numbers an infinite quantity of numbers exists, this is absurd. Hence, $\neg T$ is false, and *T* is true.

The easiest example of a theorem proved by infinite descent is also the first historical one: Euclid in his *Elements*, Book VII, Proposition 31 proves that every composite number is divided by a prime number. Be given a composite number *p*. Suppose that *p* is not divided by any prime number. Be p_1 and p_2 two divisors of *p*. Given the *ad absurdum* hypothesis, none of them is prime or divided by a prime number. Consider p_1. It cannot either be prime or be divided by a prime number. Thence, it is divided by two composite numbers p_3 and p_4. None of them can be prime, and so on. In this way, an infinite descent of numbers between *p* and 1 should exists. But this is absurd because only *p*-1 numbers exist between *p* and 1. Thus, the theorem is proved.

The method of infinite descent was rediscovered by Fermat (1601-1665), who applied it to several difficult problems in number theory. For example: a) the area of no Pythagorean triangle can be the square of an integer; b) every prime of the form 4*n*+1 is the sum of two squares; c) the equation $x^3 + y^3 = z^3$ has no integral solution apart from the trivial ones; d) every integer is the sum of four squares. Fermat left an explicit proof only of the theorem a). After Fermat, Euler (1707-1783) and Lagrange (1736-1813) used extensively this method and some variants of which Fermat had spoken (see, e.g., Bussotti 2006). Today this method is used especially in number theory and abstract algebra.

The infinite descent is more easily applicable to problems concerning prime numbers than mathematical induction, because it does not imply the step from *n* to *n*+1, which is generally useless if one deals with prime numbers. Rather, the assumption that the theorem to prove is false implies the existence of a form which maintains its mathematical structure during the descent but represents progressively decreasing numbers. This form can be repeated at infinity, but between two numbers only a finite quantity of numbers can exist. Such logical structure is often more malleable than mathematical induction. This is the reason why it is inappropriate to see the infinite descent as a reversed induction.

Strong mathematical induction. Like in mathematical induction, in this method, there is a *base* for which the proposition to prove is true. The difference is that for the inductive step from *n* to *n*+1 one has to assume the proposition to be true not only for *n*, but for *all the integers less than n*+1.

Example: Every positive integer *n* can be written as a sum of distinct nonnegative integer powers of 2. *Base*: since $1 = 2^0$, the theorem is true for *n* = 1. *Step*. Suppose that every number less than or equal to *n* be written as sum of distinct powers of 2. Let us prove that also *n*+1 can be written in this form. Let *l* be the largest integer such that $2^l \le n + 1$. Set $m = n + 1 - 2^l$. Since $2^l \geq 1$, it is $m < n + 1$. Then by the strong inductive hypothesis, there are distinct integers $r_1, r_2, \ldots r_s$ such that $m = 2^{r_1} + 2^{r_2} + \cdots + 2^{r_s}$. Therefore, it is $m = 2^{r_1} + 2^{r_2} + \cdots + 2^{r_s} + 2^l$. Now, it is only necessary to verify that *l* is different from each r_j. Suppose *ad absurdum* that an index *j* exists such that $l = r_j$. Under this hypothesis $m \geq 2^l$. Hence $n + 1 = m + 2^l \geq 2^l + 2^l = 2^{l+1}$. In this case 2^l would not be the largest power of 2 less than *n*+1. But this is in contradiction with our assumption. Thence, the proposition to prove is true (this example is drawn from Radcliffe, without date, pp. 8-9).

This proof is of extreme interest because in its final part where the *ad absurdum* reasoning is developed, a further principle logically equivalent to mathematical induction is used: the *least integer principle*, which states that if a property is true in integers, there exists the smallest number for which it is true. This principle, which was for the first time employed by Lagrange, is used in an *ad absurdum* reasoning. Suppose one has to prove a theorem *T*. Suppose $\neg T$ to be true. Then the smallest number for which $\neg T$ is true must exists. Set this value equal to *m*. During the reasoning it is shown that another value less than *m* exists for which $\neg T$ is true. This contradicts the initial hypothesis, thence *T* is true. As a matter of fact, in the proof of the analysed example a principle which might be called of *the biggest integer principle* is applied because 2^l should be the biggest integer less than *n*+1, but we discover that 2^{l+1} should also be less than *n*+1, which contradicts our hypothesis. Obviously, the logic behind

https://doi.org/10.33225/jbse/23.22.04

the least and biggest integer principle is the same: a number should have and at the same time should not have a certain property, but this is absurd.

The principles of strong induction and of the least integer are fundamental in many branches of mathematics, as in abstract algebra where many propositions concerning group, ring and field theories are proved through these methods. For example, it is possible to prove by strong induction the fundamental Cauchy theorem of group theory that given a group *G* and its order *O*(*G*), if a prime *p* divides *O*(*G*), then *G* contains an element of order *p*. Among the numerous proofs of Sylow theorem that if *p* is a prime and p^{α} divides *O*(*G*), then a subgroup of order p^{α} exists in *G*, one is given by strong induction. The least integer principle is, for example, used to prove that two Abelian groups of order p^{n} are isomorphic if and only if they have the same invariants. In ring theory this principle is used to prove that if *R* is a Euclidean ring and *A* is an ideal of *R*, an element $a \in A$ exists such that *A* is composed of the elements whose form is *ax*, when *x* varies in *R*. In the theory of the invertible elements belonging to a Euclidean ring, strong induction and least integer principle are also extensively used. Further examples could be offered, but these are sufficient.

Summing-up

In this editorial a didactic unit has been proposed. Here only the basic elements have been given. The unit might be enriched by tracing the history of the inductive principle, which is interesting and formative from an educational standpoint (see, e.g., Palladino-Bussotti 2002). From an epistemological-methodological perspective, it should be pointed out that mathematical induction is a method of proof, but not a method of discovery. It is not a heuristic procedure. The presentation of an organised set of lessons on the principle of mathematical induction is significant and appropriate in the last years of high school because: 1) it introduces the learners within a method typical of whole numbers; 2) it connects mathematical issues with logical ones; 3) it is useful in order to clarify the difference between the way of reasoning connoting mathematic and that typical of empirical sciences. Hence, it is also useful in a science education context; 4) it can be related to the history of mathematics, so as to show that mathematics is also a humanistic discipline, born from conceptual problems, and not only a technical one; 5) it has also connection with epistemological themes linked to the heuristic of mathematics.

References

Bussotti, P. (2006). *From Fermat to Gauss. Indefinite descent and methods of reduction in number theory*. Rauner, Augsburg.

Bussotti, P. (2021). The concept of inertia: An interdisciplinary approach. *Journal of Baltic Science Education*, *20*(1), 4-9. https://doi.org/10.33225/jbse/21.20.04

Chen, Y. C., Yang, F. Y., & Chang, C.-C. (2020). Conceptualizing spatial abilities and their relation to science learning from a cognitive perspective. *Journal of Baltic Science Education*, *19*(1), 50-63. https://doi.org/10.33225/jbse/20.19.50

Finzi-Vita, S. (2009). *Le successioni definite per ricorrenza* [The series defined by recurrence]. https://www.mat.uniroma1.it/mat_cms/materiali/2009-2010/finzivita-15035-ricorrenza_LPC.pdf

Palladino, D., & Bussotti, P. (2002). Il principio di induzione. Storia e ruolo [The principle of induction. Its history and role]. *Nuova Secondaria*, *2*, 41-51.

Radcliffe, M. w.d. Math 127. Induction. https://www.math.cmu.edu/~mradclif/teaching/127S19/Notes/Induction.pdf

Truyol, M. E., Sanjosé, V., & Gangoso, Z. (2014). Obstacles modelling reality: Two exploratory studies on physics defined and undefined problems. *Journal of Baltic Science Education*, *13*(6), 883-895. https://doi.org/10.33225/jbse/14.13.883

Received: *January 22, 2023* Revised: *February 02, 2023* Accepted: *February 10, 2023*

Cite as: Bussotti, P. (2023). A didactic unit on mathematics and science education: The principle of mathematical induction. *Journal of Baltic Science Education*, *22*(1), 4-9. https://doi.org/10.33225/jbse/23.22.04

Paolo Bussotti PhD, Associate Professor, University of Udine, via Palladio 8, 33100 Udine, Italy.
E-mail: paolo.bussotti@uniud.it
ORCID: https://orcid.org/0000-0002-7883-8618