Parabola Volume 50, Issue 1 (2014)

History of Mathematics: Mathematical Induction and the Foundations of Arithmetic

Michael A B Deakin[1]

In my last column, I introduced a technique of proof known as "mathematical induction". It is a powerful technique, aimed at proving general formulas, by showing that each instance of the formula implies the truth of the next. For completeness, and for the sake of new readers, I reiterate the underlying principle, before going on to describe the application of the technique to the proof of some of the most basic properties of the number system.

Here is the basis of the method as summarized in my last column. If we write $S(n)$ as a shorthand for the statement "Statement S about the number n is true", then all that we need do to prove $S(n)$ for all values of n is to show two simpler things:

(a) $S(1)$,

(b) $S(n) \Rightarrow S(n+1)$, for all n.

The reason is as follows: if, by (a), $S(1)$ holds, then, by (b), $S(2)$ holds, and then, again by (2b), $S(3)$ holds, and so on for all values of n. So the method consists of assuming $S(n)$, the *induction hypothesis* and from this *deducing* $S(n+1)$.

Late in the nineteenth century, two separate researchers both had the idea of applying this approach to the basis of arithmetic. The two researchers were Richard Dedekind (1831-1916) in Germany and Giuseppe Peano (1858-1932) in Italy.

Dedekind's work appeared in a short 1888 pamphlet, entitled *Was sind und was sollen die Zahlen?* (What are numbers and what should they be?) This was translated (rather clumsily) into English in 1901 as *The Nature and Meaning of Numbers* and put together with another of Dedekind's works in a booklet, *Essays on the Theory of Numbers*.[2] Peano's independent advancement of exactly the same ideas came the following year in a booklet called *Arithmetices principia, nova methodo exposita* (*Arithmetical principles newly explained*), written for a still unknown reason in Latin. It is, however, Peano's name rather than Dedekind's that has become attached to the mathematical advance in question. Peano's biographer sees a certain justice in the attribution because he judges his account to be clearer; certainly it avoids the somewhat unusual notation that Dedekind used.

[1] Dr Michael Deakin is an Adjunct Senior Research Fellow in the School of Mathematical Sciences at Monash University.

[2] An aspect of this same work was the subject of a (rather critical) analysis in my column in Vol. **44**, No. **1**.

What concerned both mathematicians was to devise a set of axioms that would suffice to prove the properties of the natural numbers (which, for brevity, I shall simply call *numbers*).

There are many versions of the Dedekind-Peano approach, but all are equivalent apart from minor differences. The one I shall follow here is due to the American textbook writer C C MacDuffee (Wiley: 1940). In MacDuffee's version, there are five *Peano Axioms*:

1. There exists a number 1.
2. Every number a has a successor a^+.
3. The number 1 is the successor of no number.
4. If $a^+ = b^+$, then $a = b$.
5. If a set of numbers contains 1 and also contains the successor of every number, then it contains all numbers.

There are slightly different versions of these as given by different authors. For example, some treat the first two as **definitions** rather than as **axioms**, and others give the different axioms in somewhat different order, etc. Here I deal with them as I have just listed them and comment sequentially on each.

In starting the count at 1, MacDuffee is following Dedekind rather than Peano. Peano himself started the count at 0 rather than 1, and this is the more modern way of looking at things. However, the psychological relation to the counting process makes MacDuffee's choice more natural. I discussed this question at some length in my column in Vol. **45**, No. **2**.

The second axiom introduces the other key element in the process. There is to be a *succession* of numbers beginning with 1. The first two axioms introduce these concepts. I referred above to these axioms as being "definitions", but strictly speaking this is a little imprecise. Rather 1 and succession are *undefined* notions. This means that they have no strict definition in terms of the system being developed. [*Psychologically*, however, we know that we are talking about our familiar counting numbers: how we start at 1 and continue a succession of subsequent numbers. As MacDuffee puts it: "They are not, however, truly undefined, for they are implicitly defined by the postulates they are assumed to satisfy."]

The third axiom makes the number 1 special; it is the *start* of the succession. No other number has this status.

The fourth makes the successor of each number unique; two different numbers cannot share the same successor.

It is the fifth of these axioms that incorporates the inductive idea. If you compare the statement of this fifth axiom with the description of the inductive principle given above, you will see that the two ways of looking at things are the same: the set of numbers that exhibit a property is complete if it contains 1 and the successor of every number.

With this behind us, we are now in a position to develop the basic rules of arithmetic. The first thing we need to do is to develop the laws of addition. So we have to start with a definition of what we mean by addition. The definition itself is inductive.[3] It comes in two parts. We begin by saying what we mean by $a + 1$ for some number a. We have:

$$a + 1 = a^+. \tag{0.1}$$

This means that if we add 1 to a number, the result is the next number in the sequence. The second half of the definition says what happens if (having got some way with our definition) we wish to proceed further.

$$a + b^+ = (a + b)^+ \,. \tag{0.2}$$

The idea is that we start by adding 1 according to Equation (1) and then we can use Equation (2) to add 1^+ (or more familiarly 2). And we just keep going in this way.

Now we have used mathematical induction to say what we *mean* by addition, we can also use it to prove properties of the addition process. First, we show that it is *associative*, that is to say:

$$(a + b) + c = a + (b + c)\,. \tag{0.3}$$

The understanding of this law is that it doesn't matter in which order we add terms. If we first add a and b and *then* add c, we get the result as if we had first added the b and the c and then added the result onto the number a. The proof of this property is (surprise, surprise!) inductive. Indeed, it uses a double induction. First we show that

$$a + (b + 1) = (a + b) + 1. \tag{0.4}$$

But now $a+(b+1) = a+b^+$ by Equation (1), and this equals $(a + b)^+$ by Equation (2). Now apply Equation (1) to the number $a + b$ and find that we have indeed $(a + b) + 1$ as claimed.

In order to complete the proof, we need to show that

$$(a + b) + c = a + (b + c) \Longrightarrow (a + b) + c^+ = a + (b + c^+).$$

$$\begin{array}{lll} \text{But} \quad (a + b) + c^+ & = ((a + b) + c)^+ & \text{by Equation (2)} \\ & = (a + (b + c))^+ & \text{by the induction hypothesis} \\ & = a + (b + c^+) & \text{again by Equation (2).} \end{array}$$

The proof is now complete.

Next we show that addition is *commutative*. That is to say:

$$a + b = b + a. \tag{0.5}$$

Again the proof proceeds via a double induction. First it is demonstrated that:

$$a + 1 = 1 + a. \tag{0.6}$$

[3] It is somewhat of a novelty to use induction to *define* a process. It may be that Dedekind and Peano were the first to do this. Certainly I know of no earlier example.

This is clearly true if $a = 1$, so to complete the first part of the induction we need to show that if Equation (6) holds, then $a^+ + 1 = 1 + a^+$.

$$\begin{aligned} \text{But now} \quad a^+ + 1 &= (a+1)+1 && \text{by Equation (1)} \\ &= (1+a)+1 && \text{by the induction hypothesis} \\ &= 1+(a+1) && \text{by the associative property just proved} \\ &= 1+a^+ && \text{again by Equation (1).} \end{aligned}$$

So the first half of the inductive argument is complete.

It remains to show that Equation (5) implies that

$$a + b^+ = b^+ + a. \tag{0.7}$$

$$\begin{aligned} \text{But} \quad a+b^+ &= a+(b+1) && \text{by Equation (1)} \\ &= (a+b)+1 && \text{by the associative property} \\ &= 1+(a+b) && \text{by the first part of the proof} \\ &= 1+(b+a) && \text{by the induction hypothesis} \\ &= (1+b)+a && \text{again by the associative property} \\ &= (b+1)+a && \text{again by the first part} \\ &= b^+ + a && \text{by Equation (1).} \end{aligned}$$

The proof is now complete.

These two properties (associativity and commutativity) assure us that in adding any set of numbers, we may take them in any order we please without affecting the result. When I was in Year 4 of my schooling, there was a fashion for addition problems known colloquially as “long and cross tots”. Such a problem presented the student with a rectangular array of numbers. Each column was to be added to give a partial sum of the entire array, and likewise each row. The partial sums from the columns were added to deliver the grand total, as were the partial sums for the rows. The two processes were to agree, giving the same number for the grand total. Of course there were kids who thought there ought to be space for two different answers, but such concerns met short shrift from the teachers!

After addition comes multiplication, and this too is defined inductively. We have:

$$a * 1 = a. \tag{0.8}$$

$$a * b^+ = a * b + a. \tag{0.9}$$

The first property we need to prove has a long and somewhat involved title: *Multiplication is left-distributed with respect to addition*. This means that

$$a * (b + c) = a * b + a * c. \tag{0.10}$$

When $c = 1$, $a * (b + c) = a * (b + 1) = a * b^+ = a * b + a = a * b + a * 1$, as a result of Equations (8) and (9). So Equation (10) holds for $c = 1$. It remains to show that

$$a * (b + c) = a * b + a * c \Longrightarrow a * (b + c^+) = a * b + (a * c)^+ \ .$$

$$\begin{aligned}\text{But}\quad a*(b+c^{+}) &= a*(b+c)+a && \text{by Equation (9)}\\ &= (a*b+a*c)+a && \text{by the induction hypothesis}\\ &= a*b+(a*c+a) && \text{by the associative property}\\ &= a*b+a*c^{+} && \text{again by Equation (9).}\end{aligned}$$

This completes the proof.

Like addition, multiplication is associative. That is to say:

$$(a*b)*c = a*(b*c). \tag{0.11}$$

Again, the proof is inductive. In the case $c = 1$, $(a*b)*1 = (a*b) = a*(b*1)$, as a result of Equation (8).

$$\begin{aligned}\text{But}\quad (a*b)*c^{+} &= (a*b)*c+a*b && \text{by Equation (9)}\\ &= (a*b)*c+a*b && \text{by the inductive hypothesis}\\ &= a*(b*c+b) && \text{by the distributive property}\\ &= a*(b*c+b) && \text{by Equation (9).}\end{aligned}$$

The proof is now complete.

We are leading to the proof of the most difficult of the properties, the so-called *commutative law of multiplication*. This states that

$$a*b = b*a. \tag{0.12}$$

We begin with a special case, $1*a = a*1 = a$, and again the proof is inductive. It is clearly true for $a = 1$. We now need to establish that

$$1*a = a*1 \Longrightarrow 1*a^{+} = a^{+}*1$$

$$\begin{aligned}\text{We have}\quad 1*a^{+} &= 1*(a+1) && \\ &= 1*a+1*1 && \text{by the distributive law}\\ &= 1*a+1 && \\ &= a+1 && \text{by the induction hypothesis}\\ &= a^{+} = a^{+}*1 && \text{as required.}\end{aligned}$$

So now Equation (12) holds for $b = 1$.

It is now required to complete the proof by showing that Equation (12) implies that $a*b^{+} = b^{+}*a$. This is a little tricky. We first need a subsidiary result

$$b^{+}*a = b*a+a. \tag{0.13}$$

The proof is once again inductive. It holds for $a = 1$, because, by Equation (8), $b^{+}*1 = b^{+} = b+1 = b*1+1$.

And now we need to show that Equation (13) implies that

$$b^{+}*a^{+} = b*a^{+}+a^{+}.$$

$$\begin{aligned}\text{We have}\quad b^{+}*a^{+} &= b^{+}*a+b^{+} && \text{by Equation (9)}\\ &= b*a+a+b+1 && \text{by the induction hypothesis}\\ &= b*a+b+a+1 && \text{by Equation (5)}\\ &= b*a^{+}+a^{+} && \text{again by Equation (9).}\end{aligned}$$

This completes the proof of the subsidiary result.

With this behind us, the proof of commutativity, Equation (12), is straightforward. We have seen that it is true for $a = 1$, and now we need to show that $a * b^+ = b^+ * a$ follows from the inductive hypothesis. But

$$a * b^+ = a * b + a = b * a + a = b^+ * a.$$

I leave the reader to fill in the details.

There remains one final property to be proved: multiplication is right-distributed with respect to addition, i.e.

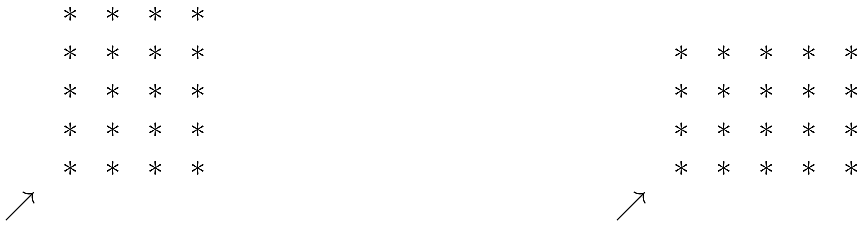
$$(b + c) * a = b * a + c * a. \tag{0.14}$$

And this time we need not use an inductive argument, because the result is a ready corollary of Equations (10) and (12).

That the proof of Equation (12) (commutativity) is the most difficult and intricate is fitting. Different sorts of numbers build on the simple natural numbers we have been discussing here. There are signed (directed) numbers, rational numbers, real numbers, complex numbers and more exotic ones beyond that. The commutative law holds for all these up to and including the complex numbers, but no longer applies beyond those. As we move to ever more complicated number systems, our familiar laws of algebra start to break down. Equation (12) is the first to go.

I incline to the view also that the commutative law of multiplication is *psychologically* the most difficult of the familiar laws for us to grasp. Certainly this was my own experience. When I was in Year 5, the most challenging task set for us was the memorization of the multiplication table. It was my mother who pointed out to me that $3 \times 4 = 4 \times 3$ and so on. Certainly this piece of information greatly reduced the magnitude of the task involved, and so I found it a great help. Nonetheless I did not accept that it was universally true. It seemed to me eminently possible that it only applied to the small numbers encountered in the table itself. For several years, I would multiply what to me were large numbers in the hope of discovering a counterexample. Needless to say I had no success!

Eventually, I came up with an argument that made the law apparent to me. It went as follows: the number $n \times m$ can be represented an a rectangular array of n rows, each containing m objects, while the number $m \times n$ could be represented an a rectangular array of m rows, each containing n objects. The diagram below illustrates the situation.



Viewed from the bottom left, an observer would count n (here 4) columns to the right and m (here 5) rows to the right. But if the array were rotated, the position would be reversed. In fact, we need not even rotate the array: merely view it from a different direction. Once I realized this, I was immediately convinced, because it seemed obvious that neither the rotation nor the different viewpoint could possibly affect the actual number of objects. This argument is, to my mind, utterly convincing. However, it probably falls short of being a strict proof.

In a vague way, I must have realized this because, when I later discovered MacDuffee's account and the Peano axioms, I immediately became a convert!