Teaching the Negatives, 1870–1970: A Medley of Models

GAVIN HITCHCOCK

Although there are as many ways of teaching any subject as there are teachers, there is much to be learned from even a transcript of a good teacher at work. This is one reason for giving here a series of snapshots of important representative moments in the hundred-year period considered. But there is the added dimension of a comparative study, both of trends in time and of varieties of style; and beyond these there is the suggestion that it might be good to expose a class of children, or older students, to a variety of approaches to this or any particular topic, and encourage them to be active critics instead of passive recipients. There is a depth of insight and breadth of perspective to be gained by such interactive experience with multiple approaches, for, after all, that is how most early learning proceeds.

The six extracts (slightly edited) are presented here in historical sequence as classroom *scenes*, in order to emphasize the context, the person of the teacher, the time and the place, and to enliven the experience of reading or enacting. There is inevitable loss in the base prose of a written text, but much of the original impact could be regained in imagination, or by a sensitive teacher interacting with a lively class, using the script as a basis for improvisation. The sequence as a whole is structured as a play: *Embracing the Strangers*, opening with a Prologue (as from Augustus de Morgan) and closing with an Epilogue from Felix Klein. Both De Morgan and Klein were great mathematicians with a deep concern for (and involvement with) mathematical education at all levels. The aim of the Prologue is to sketch briefly some of the human and intellectual background to the mathematics, setting the scene as it would have been perceived by De Morgan in 1871 (the year of his death). This draws from De Morgan's writings and aims at reflecting his personality, but is not to be quoted as from him. The Epilogue is an almost direct extract from Klein, reviewing both the mathematical story and the art of teaching.

At the end, in addition to notes and references for the play, there is a set of questions and exercises intended to encourage reflection on the play. These (like the scenes themselves) vary in depth and difficulty, but are mainly at a level suited to pre-service or in-service teacher training workshops. Many of the issues arising have relevance far beyond the topic of negative numbers, or even the teaching of number systems and algebra. I would like to acknowledge the inspiration of Abraham Arcavi and Maxim Bruckheimer, whose work in developing historical sourcework collections for use in teacher training in Israel was described in FLM Vol. 3, no. 1; the context and motivation behind my own work in dramatising the history of mathematics was described in FLM Vol. 12, no. 1.

The play

EMBRACING THE STRANGERS

One Hundred Years of Didactics

PROLOGUE: *Reminiscences*

AUGUSTUS DE MORGAN[1], in 1871, the last year of his life, aged 65, musing in an armchair.

DE MORGAN

I think the most striking change in mathematics over my life-time has been the joyous assertion of logical freedom! Our laws — whether of number, algebra or even geometry — are not absolute, not logically necessary after all. There are new Geometries, new Algebras, to explore; new entities, such as Sir William Hamilton's quaternions and Professor Arthur Cayley's matrices, obeying quite remarkable laws. And the way to all this was opened, I think, by the gradual acceptance of the negative numbers, closely followed by the imaginary numbers, as mathematicians began to realize the relative meaning of the terms "possible" and "impossible" or, indeed, the terms "real" and "imaginary"! It is human tradition, drawing upon the resources of human imagination, which sets the limits on the field of operation, which erects the fences and draws the horizons. We must, of course, ensure that any proposed law is logically *permissible* — that is, consistent with its fellows; but our structures are otherwise agreeably arbitrary, free creations of the human spirit, regulated by considerations of convenience and expediency (such as the principle of permanence), or by considerations of elegance or the desired applicability of the resulting theory.

This quality of freedom would have shocked the mathematicians of the last century. But they were nevertheless unconsciously preparing the way, as they were won over by the negative numbers and the imaginary numbers, and swept along by the exhilarating currents of symbolic Algebra and Analysis.

It is truly astounding, when one comes to reflect upon it, how great a degree of unanimity we mathematicians have achieved over previously contentious issues. Not that we don't still have our petty differences, but I believe it would have been generally admitted, by about the middle of this century, that the only subject yet remaining (of an elementary character), on which a serious schism existed among mathematicians, as to the absolute correctness or incorrectness of results, was the question of divergent series.[2] And even those monstrous creatures are rapidly becoming domesticated and their somewhat embarrassing uses regarded as legitimate.

What I said back in the forties about the way we should react to anomalies and embarrassments has been proved true in striking ways. The history of algebra shows us that nothing is more unsound than rejection of any method which naturally arises, merely because of one or more valid cases in which such a method leads to erroneous results. Such cases should indeed teach *caution*, but not *rejection*. For if the latter had been preferred to the former, negative quantities, and still more, their square roots, would have been an effectual bar to the progress of algebra, and those immense fields over which even the rejectors of divergent series now roam without fear, would have been not so much as discovered, much less cultivated and settled.[3]

How singular, in retrospect, that the burning issue of the *reality* of negative numbers should have appeared a *logical* one, and turned out in the end as a victory, not for logic, but for the imagination! Babbage saw it earlier than most of us in England, I think; and poor Peacock fought for the *logical* status of his Principle of Permanence, only to see it become a handmaid to the imagination! The realisation has dawned slowly, but is now clear to all of us: the moving power of mathematical invention is not reasoning but imagination![4]

The great difficulty of the opponents of Algebra — the so-called "pure arithmeticians" — lay in a lack of ability or will to see *extension* of terms; to admit any use of symbols which outstrips the limits of absolute number. They would forbid all extensions of language[5], and so cut themselves off from one of the great creative forces of the imagination, which is operative in all poetry and great literature: to allow the words, the symbols, to carry one beyond oneself!

Perhaps this sect is extinct now. During the last century, its chief writers were Robert Simson, Francis Masères and William Frend. So far as these opponents [of negative numbers and symbolic algebra] set out their objections, it is seen that there is much force in them against the mode of elementary writing then in vogue. Having been casually brought into contact with Mr Frend in early life (he later became my father-in-law), at a time when I was engaged in examination of first principles, and having had many discussions with him, and been led thereby to an attentive examination of Masères, Simson, and others, I long ago came to the conclusion that in those minds which are irresistibly led to a sweeping condemnation of almost everyone else, on matters of subjective nature, the bias is a craving for simplicity — a craving which will, in the end, find a way of rejecting whatever cannot be immediately reduced to earliest axioms. A sadly common state of mind ... But I suspect that even those opponents played a useful part in the strange story of algebra — by goading others to defend and analyse their own principles.[6]

A strange story indeed! — where will algebra go in the future, I wonder? How will students of the twentieth century be taught these ideas that have been forged in such creative fires of the human imagination? Will they simply take them for granted, unquestioning, unmoved by the triumphs of previous generations of mathematicians?

[*EXIT with aid of walking stick*]

SCENE 1

CHARLES SMITH, English mathematician, reading/lecturing from his "A Treatise on Algebra," of 1888.[7]

SMITH [*writes where necessary*]
The following is taken as the definition of multiplication: "To multiply one number by a second, is to do to the first what is done to unity to obtain the second". Thus:

$$(-5) \times 4 = (-5) + (-5) + (-5) + (-5) = -20$$

With the above definition, multiplication by a negative quantity presents no difficulty.

For example, to multiply 4 by –5. Since to subtract 5 by one subtraction is the same as to subtract five units successively,

$$-5 = -1 -1 -1 -1 -1,$$
$$4 \times (-5) = -4 -4 -4 -4 -4 = -20$$

Again, to multiply –5 by –4:

$$-4 = -1 -1 -1 -1; \quad (-5) \times (-4) =$$
$$-(-5) - (-5) - (-5) - (-5) = +5 +5 +5 +5 = +20$$

The *Law of Signs* may be enunciated briefly as follows: like signs give plus, and unlike signs give minus.

[*CURTAIN*]

SCENE 2

ALFRED NORTH WHITEHEAD[8]*, English mathematician and philosopher, delivering a "popular" lecture from his book "An Introduction to Mathematics", in about the year 1918.*

WHITEHEAD [*writes all displayed formulae*]
An extension of the concept of number comes from the introduction of the idea of an operation or a step. We will start with a particular case. Consider the statement 2`345. We add 3 to 2 and obtain 5. Think of the operation of adding 3: let this be denoted by +3.

$$2 + 3 = 5, \qquad +3$$

Again consider the statement 41341. Think of the operation of subtracting 3: let this be denoted by –3.

$$4 - 3 = 1, \qquad -3$$

Thus instead of considering the real numbers in themselves, we consider the *operations* of adding, and of subtracting; instead of the number $\sqrt{2}$, for instance, we consider the operations of adding $\sqrt{2}$ and subtracting $\sqrt{2}$:

$$+\sqrt{2}, \ -\sqrt{2}$$

Then we can add these operations, of course in a different sense of addition to that in which we add numbers. The sum of two operations is the single operation which has the same effect as the two operations applied successively. In what order are the two operations to be applied? The answer is that it is indifferent, since for example:

$$2 + 3 + 1 = 2 + 1 + 3$$

so that the addition of the steps +3 and +1 is commutative.

Mathematicians have a habit, which is puzzling to those engaged in tracing out meanings, but is very convenient in practice, of using the same symbol in different though allied senses. The one essential requisite for a symbol in their eyes is that, whatever its possible varieties of meaning, the formal laws for its use shall always be the same. In accordance with this habit the addition of operations is denoted by ` as well as the addition of numbers. Accordingly we can write.

$$(+3) + (+1) = +4$$

where the middle + on the left-hand side denotes the addition of the operations +3 and +1. But, furthermore, we need not be so very pedantic in our symbolism, except in the rare instances when we are directly tracing meanings; thus we always drop the first + of a line and the brackets, and never write + signs running. So the above equation becomes

$$3 + 1 = 4$$

which we interpret as simple numerical addition, or as the more elaborate addition of operations which is fully expressed in the previous way of writing the equation, or lastly as expressing the result of applying the operation +1 to the number 3 and obtaining the number 4. Any interpretation which is possible is always correct. But the only interpretation which is always possible, under certain conditions, is that of operations. The other interpretations often give nonsensical results.

This leads us at once to a question, which must have been rising insistently in your mind: What is the use of all this elaboration? At this point our friend, the practical man, will surely step in and insist on sweeping away all these silly cobwebs of the brain. The answer is that what the mathematician is seeking is *Generality.* This is an idea worthy to be placed beside the notions of the *Variable* and of *Form* so far as concerns its importance in governing mathematical procedure.

[Variable, Form, Generality]

Any limitation whatsoever upon the generality of theorems, or of proofs, or of interpretation is abhorrent to the mathematical instinct. These three notions, of the variable, of form, and of generality, compose a sort of mathematical trinity which preside over the whole subject. They all really spring from the same root, namely from the abstract nature of the science.

Let us see how generality is gained by the introduction of this idea of operations. Take the equation $x + 1 = 3$; the solution is $x = 2$. Here we can interpret our symbols as mere numbers, and the recourse to "operations" is entirely unnecessary. But if x is a mere number, the equation $x + 3 = 1$ is nonsense. For x should be the number of things which remain when you have taken three things away from one thing; and no such procedure is possible. At this point our idea of algebraic form steps in, itself only generalization under another aspect. We consider, therefore, the general equation of the same form as $x + 1 = 3$. This equation is $x + a = b$, and its solution is $x = b - a$. Here our difficulties become acute; for this form can only be used for the numerical interpretation so long as b is greater than a, and we cannot say without qualification that a and b may be any constants. In other words we have introduced a limitation on the variability of the "constants" a and b, which we must drag like a chain throughout all our reasoning. Really prolonged mathematical investigations would be impossible under such conditions. Every equation would at last be buried under a pile of limitations. But if we now interpret our symbols as "operations", all limitation vanishes like magic. The equation $x + 1 = 3$ gives $x = +2$, the equation $x + 3 = 1$ gives $x = -2$, the equation $x + a = b$ gives $x = b - a$, which is an operation of addition or subtraction as the case may be.

$$\begin{array}{ll} x + 1 = 3, & x = +2 \\ x + 3 = 1, & x = -2 \\ x + a = b, & x = b - a \end{array}$$

We need never decide whether $b - a$ represents the operation of addition or of subtraction, for the rules of procedure with the symbols are the same in either case. We shall not further explain the detailed rules by which the "positive and negative numbers" are multiplied and otherwise combined. We have explained above that positive and negative numbers are operations. They have also been called "steps". Thus +3 is the step by which we go from 2 to 5, and −3 is the step backwards by which we go from 5 to 2. Consider the line OX divided up, so that its points represent numbers.

$$\begin{array}{ccccccccc} X' & C' & B' & A' & O & A & B & C & X \\ & -3 & -2 & -1 & & 1 & 2 & 3 & \end{array}$$

Then +2 is the step from O to B, or from A to C, or (if the divisions are taken backwards along OX') from C' to A', or from D' to B', and so on. Similarly −2 is the step from O to B' or from B' to D', or from B to O, or from C to A.

We may consider the point which is reached by a step from O, as representative of that step. Thus A represents +1, B represents +2, A' represents −1, B' represents −2, and so on. It will be noted that, whereas previously with the mere "unsigned" real numbers the points on one side of O only, namely along OX, were representative of numbers, now with steps every point on the whole line stretching on both sides of O is representative of a step. This is a pictorial representation of the superior generality introduced by the positive and negative numbers, namely the operations or steps. These "signed" numbers are also particular cases of what have been called vectors (from the Latin *veho*, I draw or carry). For we may think of a particle as carried from O to A, or from A to B.

In suggesting just now that the practical man would object to the subtlety involved by the introduction of the positive and negative numbers, we were libelling that excellent individual. For in truth we are on the scene of one of his greatest triumphs. If the truth must be confessed, it was the practical man himself who first employed the actual symbols + and −. Their origin is not very certain, but it seems most probable that they arose from the marks chalked on chests of goods in German warehouses, to denote excess or defect from standard weight. The earliest notice of them occurs in a book published at Leipzig, in A.D. 1489. They seem first to have been employed in mathematics by a German mathematician, Stifel, in a book published in

Nūremburg in 1544 A.D. But then it is only recently that the Germans have come to be looked on as emphatically a practical nation. There is an old epigram which assigns the empire of the sea to the English, of the land to the French, and of the clouds to the Germans. Surely it was from the clouds that the Germans fetched + and –; the ideas which these symbols have generated are much too important for the welfare of humanity to have come from the sea or from the land!

The possibilities of application of the positive and negative numbers are very obvious. If lengths in one direction are represented by positive numbers, those in the opposite direction are represented by negative numbers. If a velocity in one direction is positive, that in the opposite direction is negative. If a rotation round a dial in the opposite direction to the hands of a clock (anti-clockwise) is positive, that in the clockwise direction is negative. If a balance at the bank is positive, an overdraft is negative. If vitreous electrification is positive, resinous electrification is negative. Indeed, in this latter case, the term positive electrification and negative electrification, considered as mere names, have practically driven out the other terms. An endless series of examples could be given. The idea of positive and negative numbers has been practically the most successful of mathematical subtleties.

[CURTAIN]

SCENE 3

EDMUND LANDAU,[9] German mathematician, lecturing to a class of students at the University of Göttingen, from his book "Foundations of Analysis", in about the year 1930.

LANDAU [*writes where necessary*]
We are now going to define, in this lecture, what we will call "real numbers", including positive numbers, zero and negative numbers. But first I will remind you of the foundation we have been laying so far in the previous three lectures.

Recall that I set out, asking of you only the ability to read English and to think logically — no high school mathematics, and certainly no higher mathematics.

We assumed, in the beginning, that we were given a set, or totality, or objects (which we called "natural numbers"), possessing five properties (which we named Axions 1 to 5), and we proceeded to define and prove theorems about two *operations* called addition and multiplication, and a *relation* called ordering.

Next we defined "fractions" as ordered pairs x_1, x_2, of natural numbers, written:

$$x_1/x_2$$

and we defined an ordering relation and the two operations of addition and multiplication on these new objects, and proved theorems about them. Then we defined "rational numbers" as sets of equivalent fractions, under the obvious equivalence relation: the pair x_1, x_2 is equivalent to the pair y_1, y_2 if x_1y_2 is equal to y_1x_2.

$$x_1/x_2 \sim y_1/y_2 \quad \text{if} \quad x_1y_2 = y_1/x_2$$

We extended our ordering relation and our two operations to these rational numbers, and proved theorems about them.

We called a rational number an "integer" (or "whole number") if its equivalence class of fractions contains one of the form: x, 1.

$$x/1$$

We proved that these integers satisfy the five Axioms of the natural numbers and have all their properties relating to equality, ordering, sum and product. Therefore we threw out the natural numbers and replaced them by the corresponding integers. Since the fractions also became superfluous, we agreed to speak only of rational numbers.

The next new entity we created was called a "cut" defined as a certain kind of subset of the rational numbers, and we went on to define and prove theorems about the ordering, the addition and the multiplication of cuts.

Certain cuts were singled out and called "rational cuts", and certain of these were singled out and called "integral cuts". And these were shown to have all the properties which we proved for rational numbers and integers, respectively. Therefore, we agreed to throw out the rational numbers, so that in all that follows we will only have to speak in terms of cuts whenever any of the foregoing material is involved.

Now, at last, we arrive at the creation of the whole system of "real numbers"!

Definition 43: The cuts will henceforth be called "positive numbers"; similarly, what we have been calling "rational numbers" and "integers" will henceforth be called "positive rational numbers" and "positive integers", respectively. We create a new number "zero", written: 0, distinct from the positive numbers. We also create numbers which are distinct from the positive numbers as well as distinct from zero, and which we will call "negative numbers", in such a way that to each positive number x we assign a negative number called minus x, and denoted by

$$-x.$$

In this assignment, minus x and minus y will be considered as the same number (as equivalent, that is) if and only if x and y are the same numbers. The totality, consisting of all positive numbers, of zero, and of all negative numbers will be called the real numbers.

Now we set out to define a concept of "absolute value" of a real number, and a concept of ordering of real numbers; and we shall prove (in Theorem 169) that the positive numbers are the numbers which are greater than zero; and the negative numbers are those which are less than zero. Thereafter we will define operations of addition and multiplication on our real numbers, and prove theorems about their properties. For example, here are some of the results we will prove to hold for all real numbers x, y, z:

Theorem 179: $x + (-x) = 0$.
Theorem 197: $(-x)y = y(-x)$.
Theorem 198: $(-x)(-y) = xy$.

So — let us get on with the task …

[CURTAIN]

SCENE 4

TOM M. APOSTOL,[10] *American mathematician, lecturing to a class of students at the California Institute of Technology, from his book "Mathematical Analysis: A Modern Approach to Advanced Calculus", in about the year 1957.*

APOSTOL

Section 1.1. Introduction. The real number system is one of the fundamental concepts of mathematics. A thorough and exhaustive study of mathematical analysis would have to include a careful definition of what is meant by a real number, a discussion of how real numbers are constructed (starting, for example, with the integers), and a derivation of the principal properties of real numbers. Although these elements form a very interesting part of the foundations of mathematics, they will not be treated in detail here. As a matter of fact, in most phases of analysis it is only the properties of real numbers that concern us, rather than the methods used to construct the real number system. Therefore I shall simply list a set of axioms from which all the properties of real numbers can be derived. For discussions of the methods used to construct real numbers you should consult the references, for example, Landau's *Foundations of Analysis.*

I shall assume that you are familiar with most of the properties of real numbers, as well as some of their elementary consequences.

Section 1.2. Arithmetical properties of real numbers. Given any two real numbers, x and y, we can form their sum $x + y$ and their product xy, and these satisfy the following axioms:

AXIOM 1. $x + y = y + x, \quad xy = yx$ (commutative laws)

AXIOM 2. $x + (y + z) = (x + y) + z, \quad x(yz) = (xy)z$ (associative laws)

AXIOM 3. $x(y + z) = xy + xz$ (distributive law)

AXIOM 4. Given any two real numbers x and y, there exists a real number z such that $x + z = y$. This z is denoted by $y - x$; the number $x - x$ is denoted by 0. (It can be proved that 0 is independent of x.) We write $-x$ for $0 - x$.

AXIOM 5: There exists at least one real number $x \neq 0$. If x and y are two real numbers with $x \neq 0$, then there exists a real number z such that $xz=y$. This z is denoted by y/x; the number x/x is denoted by 1 and can be shown to be independent of x. We write x^{-1} for $1/x$ if $x \neq 0$.

From these axioms we can derive all the usual laws of arithmetic; for example:

$$-(-x) = x, \ (x^{-1})^{-1}=x,$$
$$-(x-y) = y - x, \ x - y = x + (-y), \text{ etc.}$$

(For a more detailed explanation see Landau's book.)

Section 1.3. Order properties of real numbers. We also have a relation "less than", written <, which establishes an ordering among the real numbers and which satisfies the following axioms:

AXIOM 6: Exactly one of the following relations holds: $x = y, x < y, x > y.$
(Note: $x > y$ means the same thing as $y < x$.)

AXIOM 7: If $x < y$ then for every z we have $x + z < y + z$.

AXIOM 8: If $x > 0$ and $y > 0$, then $xy > 0$.

AXIOM 9: If $x > y$ and $y > z$, then $x > z$

From these axioms we can derive the usual rules for operating with inequalities. For example, if we have $x < y$, then $xz < yz$ if $z > 0$, whereas $xz > yz$ if $z < 0$. Also, if $x > y > 0$ and $z > w > 0$, then $xz > yw$.

$$x < y, z > 0 \rightarrow xz < yz$$
$$x < y, z < 0 \rightarrow xz < yz$$
$$x > y > 0, z > w > 0 \rightarrow xz > yw.$$

Section 1.4. Geometrical representation of real numbers. The real numbers can be represented geometrically as points on a line (the real axis). A point is selected to represent 0 and another point to represent 1, and these points determine the scale. Then each point on the real axis corresponds to one and only real number and, conversely, each real number is represented by a single point.

[CURTAIN]

SCENE 5

AMERICAN SCHOOLTEACHER, teaching from the School Mathematics Study Group (SMSG)[11] *textbook, "First Course in Algebra, Part I", in about the year 1961.*

SCHOOLTEACHER *[writes all displayed formulae]*

Today we are going to talk about another operation, called "multiplication". We have already defined and studied the properties of the operation called "addition". Now let us decide how we should multiply two real numbers to obtain another real number. All that we can say at present is that we know how to multiply two non-negative numbers.

Of primary importance here, as in the definition of addition, is that we maintain the "structure" of the number system. We know that if a, b, c are any numbers of arithmetic, then

$$ab = ba$$
$$(ab)c = a(bc)$$
$$a.1 = a$$
$$a.0 = 0$$
$$a(b + c) = ab + ac.$$

(What names did we give to these properties of multiplication?) Whatever meaning we give to the product of two real numbers, we must be sure that it agrees with the products which we already have for non-negative real numbers and that the above *properties of multiplication still hold for all real numbers.*

Consider some possible products:

$$(2)(3), \ (3)(0), \ (0)(0), \ (-3)(0), \ (3)(-2), \ (-2)(-3).$$

(Do these include examples of every case of multiplication of positive and negative numbers and zero?) Notice that the first three products involve only non-negative numbers and are therefore already determined:

$$(2)(3) = 6, \quad (3)(0) = 0, \quad (0)(0) = 0.$$

Now let us try to see what the remaining three products will have to be in order to preserve the basic properties of multiplication listed above. In the first place, if we want the multiplication property of zero to hold for all real numbers, then we have

$$(-3)(0) = 0.$$

The other two products can be obtained as follows:

$$0 = 3.0 = 3(2 + (-2)) \quad \text{by writing } 0 = 2 + (-2);$$

(Notice how this introduces a negative number into the discussion.) Hence:

$$\begin{aligned} 0 &= (3)(2) + (3)(-2), \\ &\quad \text{if the distributive property is to hold;} \\ &= 6 + (3)(-2), \quad \text{since} (3)(2) = 6. \end{aligned}$$

We know from uniqueness of the additive inverse that the only real number which yields zero when added to six is the number minus six. Therefore, if the properties of numbers are expected to hold, the only possible value for "three times minus two" which we can accept is minus six:

$$(3)(-2) = -6.$$

Now, finally, what about the product of minus two and minus three? The method is quite similar, so try to find and justify the answer yourselves?

[*CURTAIN*]

SCENE 6

ENGLISH SCHOOLTEACHER, teaching from the School Mathematics Project (SMP)[12] *"Additional Mathematics Book, Part 1", in about the year 1966.*

SCHOOLTEACHER

We have been looking at the way various number systems have developed, and at their main features. We first looked at "counting numbers", also known as natural numbers; then we looked at the practical everyday considerations which led to the invention of a new kind of number, the fraction or "sharing number". Once this new kind of number had been invented, it was necessary to specify rules for combining such numbers, suggested by the practical situations to which they were applied. So we examined these rules. Next we saw how the notion of *measuring*, rather than counting, which made the invention of fractions inevitable, also led naturally to another step forward in generalising the idea of number. This was the extension to include numbers corresponding to the lengths of all lines. We gave such numbers the name of "size numbers", and saw that they were divided into two categories, rational and irrational numbers.

Historically, the next kind of number to be invented was the directed number together with the number "zero". These numbers were first used by Hindu mathematicians, and they reached the western world through the Arab scholars of the Middle Ages.

It is easy to see how such numbers are useful in setting up a correspondence between numbers and physical observations. If we have only the size numbers, then we can use these to register a temperature above freezing point, or a height above sea level; but this correspondence breaks down as soon as there is frost or we wish to map the Dead Sea. The invention of positive and negative numbers enables us to apply arithmetic to displacements which may be made in either direction.

In discussing the fundamental ideas about directed numbers the usual notation is confusing, since the symbol [*writes*: –] (the minus sign) is used in more than one sense: for the operation of subtraction, and also as the prefix denoting a negative number. It will be convenient for the present purpose to distinguish these, which we do by saying that with each size number a we associate two directed numbers, known as "positive a" and "negative a", and written:

$$^{+}a, \quad ^{-}a.$$

We also invent a number zero with the property that, for all a, the sum of positive a and negative a is zero.

$$(^{+}a) + (^{-}a) = 0.$$

It is then necessary to devise rules for adding and multiplying these numbers. We begin by laying down that the rules for handling the positive numbers must correspond exactly to those for the unsigned size numbers. For example, we require that positive three times positive four should equal positive twelve, and that positive three plus positive four should equal positive seven. In general:

$$(^{+}a) \cdot (^{+}b) = {}^{+}(ab),$$

and

$$^{+}a + {}^{+}b = {}^{+}(a + b).$$

For negative numbers we make rules such as:

$$\begin{aligned} ^{+}a \,.\, ^{-}b &= {}^{-}(ab), \\ ^{-}a \,.\, ^{-}b &= {}^{+}(ab), \\ ^{-}a + {}^{-}b &= {}^{-}(a+b), \end{aligned}$$

and

$$^{+}a + {}^{-}b = \begin{cases} ^{+}(a + b) & \text{if } a > b, \\ ^{-}\ b - a) & \text{if } b > a, \\ 0 & \text{if } a = b. \end{cases}$$

Thus

$$\begin{aligned} ^{+}6 + {}^{-}2 &= {}^{+}(6 - 2) = {}^{+}4, \\ ^{+}5 + {}^{-}8 &= {}^{-}(8 - 5) = {}^{-}3, \end{aligned}$$

For zero there are two important rules:

(i) Addition of zero leaves a number unchanged; that is,

$$x + 0 +)x \quad \text{for all directed numbers } x.$$

(ii) Multiplication by zero gives zero; that is,

$$x.0 = 0 \quad \text{for all directed numbers } x.$$

An important consequence of (ii) is that division by zero is impossible. For if the result of dividing a number a by zero were denoted by x, this would mean that

$$a = 0.x,$$

so that

$$a = 0.$$

If a itself is not zero, this is clearly absurd; and if a were zero, the statement would be true for all values of x, so that x would be indeterminate. Thus in either case no useful meaning can be attached to the symbol $a \div 0$.

These rules for manipulating directed numbers cannot be proved; they are definitions. But it can be shown (although the details are long) from the definitions that all the usual associative, commutative, and distributive properties hold, so that directed numbers can be handled in much the same way as other kinds of numbers.

The important feature of directed numbers from the point of view of arithmetic is that with these numbers subtraction is always possible. In fact, if p and q are directed numbers, the equation

$$p + x = q$$

can always be solved by adding to q the "opposite" of p; by which we mean the number formed from p by changing its sign. For example, the solution of

$$^{+}5 + x = {}^{+}2$$

is given by

$$x = {}^{+}2 + {}^{-}5 = {}^{-}(5 - 2) = {}^{-}3;$$

and that of

$$^{-}5 + x = {}^{-}1$$

is given by

$$x = {}^{-}1 + {}^{+}5 = {}^{+}(5 - 1) = {}^{+}4$$

We therefore write

$$^{+}2 - {}^{+}5 = {}^{-}3 \text{ and } {}^{-}1 - {}^{-}5 = {}^{+}4.$$

In practice, of course, we have become accustomed to using the symbol 2 as shorthand for $^{+}2$, and -3 rather than $^{-}3$, so that we write:

$$2 - 5 = -3.$$

But one should realize that in this statement the two minus signs are used with different meanings:

[CURTAIN]

EPILOGUE: Retrospection on the Long Journey

FELIX KLEIN, German mathematician (1849–1925), Professor of Mathematics at the University of Göttingen, addresses an audience of German schoolteachers from his "Elementary Mathematics from an Advanced Standpoint: Arithmetic, Algebra, Analysis", in about the year 1908.

KLEIN[13]

One must not think that the negative numbers are the invention of some clever person who manufactured them, together with their consistency perhaps, out of the geometric representation. Rather, during a long period of development, the use of negative numbers forced itself, so to speak, upon mathematicians. Only in the nineteenth century, after people had been operating with them for centuries, was the consideration of their consistency taken up. The ancient Greeks certainly had no negative numbers, so that one cannot yield to them the first place, in this case, as so many people are otherwise prone to do. One must attribute this invention to the Hindus, who also created our system of digits and in particular our zero. In Europe, negative numbers came gradually into use at the time of the Renaissance, just as the transition to operating with letters had been completed. I must not omit to mention here that this completion of operations with letters is said to have been accomplished by Vièta in his book *In Artem Analyticam Isagoge.*[…]

Now, upon this basis of operations with letters, the most important psychological moment associated with the introduction of negative numbers is the involuntary inclination to employ rules such as the known formulae.

$$(a - b)(c - d) = ac - ad - bc + bd,$$
$$c - (a - b) = c - a + b,$$

under circumstances more general than are warranted by the original derivations of these rules. This general peculiarity of human nature shows itself, in the particular case before us, as a desire to forget the original assumptions as to the relative magnitude of a and b. Thus, for example, if one applies that first formula to:

$$a = c = 0,$$

for which the formulae were not proved at all, one obtains

$$(-b) \cdot (-d) = +bd,$$

that is, the *sign rule for multiplication of negative numbers.* In this manner we may derive, in fact almost unconsciously, all the rules for these numbers, which we must designate as *almost necessary assumptions — necessary insofar as one would have validity of the old rules for the new concepts.* To be sure, the old mathematicians were not happy with this abstraction, and their uneasy consciences found expression in names like *invented numbers, false numbers,* etc., which they gave to the negative numbers on occasion. But in spite of all scruples, the negative numbers found more and more general recognition in the sixteenth and seventeenth centuries, because they justified themselves by their usefulness. To this end, the development of analytic geometry without doubt contributed materially.

Nevertheless the doubts persisted, and were bound to persist, so long as one continued to seek for a representation in the concept of a number of *things*, and had not recognised the leading role of formal laws when new concepts are set up. In connection with this stood the continually recurring attempts to prove the rule of signs (above). The simple explanation, which was brought out in the nineteenth century, is that it is *idle to talk of the logical necessity of the theorem*; in other words, *the rule of signs is not susceptible of proof, one can only be concerned with recognising the logical permissibility of the rule,* and, at the same time, that it is arbitrary, and regulated by considerations of expedience, such as the principle of permanence.

In this connection one cannot repress that oft-recurring thought that *things* sometimes seem to be more sensible than human beings. Think of it: one of the greatest advances in mathematics, the introduction of negative numbers and of operations with them, was not created by the conscious logical reflection of an individual. On the contrary, its slow organic growth developed as a result of intensive occupation with *things*, so that it almost seems as though men had learned from the letters! The rational reflection that one devised here something correct, compatible with strict logic, came at a much later time. And, after all, the function of pure logic, when it comes to setting up new concepts, is only to *regulate* and *never to act as the sole guiding principle*; for there will always be, of course, many other conceptual systems which satisfy the single demand of logic, namely, freedom from contradiction.

If we now look critically at the way in which negative numbers are presented in the schools, we find frequently the error of trying to prove the logical necessity of the rule of signs, corresponding to the above noted efforts of the older mathematicians. One is to derive

$$(-b)(-d) = +bd$$

heuristically, from the formula

$$(a - b)(c - d)$$

and to think that one has a proof, completely ignoring the fact that the validity of this formula depends on the inequalities: a greater than b, c greater than d:

$$a > b, c > d.$$

Thus the proof is fraudulent, and the psychological consideration which would lead us to the rule by way of the principle of permanence is lost in favor of quasi-logical considerations. Of course the pupil, to whom it is thus presented for the first time, cannot possibly comprehend it, but in the end he/she must nevertheless believe it; and if, as it often happens, the repetition in a higher class does not supply the corrective, the conviction may become lodged with some students that the whole thing is mysterious, incomprehensible.

In opposition to this practice, I should like to urge you, in general, never to attempt to make impossible proofs appear valid. One should convince the pupil by simple examples, or, if possible, let him find out for himself (or her for herself) that, in view of the actual situation, *precisely these conventions, suggested by the principle of permanence, are appropriate in that they yield a uniformly convenient algorithm, whereas every other convention would always compel the consideration of numerous special cases.* To be sure, one must not be precipitate, but must allow the pupil time for the revolution in his thinking which this knowledge will provoke. And while it is easy to understand that other conventions are not advantageous, one must emphasize to the pupil how really wonderful the fact is that a general useful convention really *exists*; it should become clear to him/her that this is by no means self-evident.

From the standpoint of mathematical pedagogy, we must protect against putting abstract and difficult things before the pupils too early. In order to give precise expression to my own view on this point, I should like to bring forward the biogenetic fundamental law, according to which the individual in his/her development goes though, in an abridged series, all the stages in the development of the species. Now, I think that instruction in mathematics, as well as in everything else, should follow this law, at least in general. Taking into account the native ability of youth, instruction should guide it slowly to higher things, and finally to abstract formulations; and in doing this it should follow the same road along which the human race has striven from its naïve original state to higher forms of knowledge. It is necessary to formulate this principle frequently, for there are always people who, after the fashion of the mediaeval scholastics, begin their instruction with the most general ideas, defending this method as the "only scientific one". And yet this justification is based on anything but truth. To instruct scientifically can only mean to induce the person to think scientifically, but by no means to confront him (or her), from the beginning, with cold, scientifically polished systematics.

An essential obstacle to the spreading of such a natural and truly scientific method of instruction is the lack of historical knowledge which so often makes itself felt. In order to combat this, I have made a point of introducing historical remarks into my presentation. By doing this, I trust that I have made it clear to you how slowly all mathematical ideas have come into being; how they have nearly always appeared first in rather prophetic form, and only after long development have crystallized into the rigid form so familiar in systematic presentation. It is my earnest hope that this knowledge may exert a lasting influence upon the character of your own teaching.

[*CURTAIN*]

THE END

Questions and exercises

1. In the Prologue, Augustus De Morgan's reminiscences evoke something of the intellectual adventure and the human story behind the mathematics. How useful would this (or similar personal perspectives on the mathematical scene at moment in time — authentic or fictional) be, as a classroom activity or a reading assignment, in arousing the interest and curiosity of students of mathematics at various levels?
2. In Scene 1, Charles Smith proceeds by giving, without motivation or explanation, a definition of *product* which, however, may have some attractions. For Smith the application to negative numbers "presents no difficulty". Do you agree? In the seventeenth century, Antoine Arnauld (one of Coolidge's "great amateurs") expressed acute uneasiness at the extension to negative numbers of what he referred to as a basic principle of multiplication of two factors: the ratio of unity to the first factor is equal to the ratio of the second factor to the whole product. His example was

   $$(-4)(-5) = +20, \quad \text{leading to } +1 :-4 \quad \text{as} \quad -5 :+20$$

   so that bigger is to smaller as smaller is to bigger, as Arnauld saw it, and cried foul!

   How are Smith's definition and Arnauld's principal related? Observe that Smith's is the more general, involving as he does the interaction between the two operations $+, \cdot$, such that $(S,+)$ forms an Abelian group with identity 0, and $(S\backslash\{0\},\cdot)$ forms an Abelian group with identity 1. Show that Smith's product is well-defined if and only if the operation $\cdot$ is distributive over the operation $+$ (that is, $(S,+,\cdot)$ is a field, or satisfies Apostol's Axioms 1-5).

   Is Smith's definition an evasive ruse? Is it a too-clever definition contrived in the service of authoritarian deductivism? Is it an attempt to provide a sure rule-of-thumb for keeping on track in perilous un-mapped regions, in contrast to the later emphasis on *structure* of the numbers perceived as homeground? Is it a good (transparent, natural, easy to use) definition?

Is the response to such questions an indicator of psychological citizenship — of the worlds before or after the great transition described in the prologue? Is a similar transition experienced in the classroom, or in the maturing consciousness of children?

3. In the eighteenth century, Nicholas Saunderson, the blind Lucasian Professor at Cambridge, in his *The Elements of Algebra* (Cambridge, 1741), assumed that multiplying an arithmetic progression by any number (even a negative number) would yield another arithmetic progression. He then used this to prove the law of signs: $(-x)(-y) = xy$. Leonhard Euler, and others of his time, assumed implicitly that the *cancellation law* extended to negative numbers, and used this to justify the law of signs.

   In the nineteenth century, George Peacock and Hermann Hankel found logical justification for arriving at and extending such laws in a *Principle of Permanence.*

   Now, in the twentieth century, the two Schoolteachers and Tom Apostol all lay down, more or less explicitly, certain rules or *axioms* as a logical basis for all results. These may be taken as (i) *granted* for the sake of argument (by whom?); (ii) *self-evident* to any right-thinking person; (iii) a *set of rules* for the game we are going to play; (iv) *freely creating* a world of things we are pleased to call numbers; (v) *carefully selected* on the basis of some limited permanence principle for carrying over from the restricted old habitat to the new — and therefore defining the limits of the new habitat; (vi) *absolutely necessary minimum requirements* for the survival and success for our particular (specified?) purposes in the (specified?) wider environment; (vii) a *sufficient constitution*, or legal framework for settling all disputes — hence a basis for condemnation of outlaws and a recourse lest we are goaded to defend and analyse the legality of our own principles (in De Morgan's phrase).

   Which of these is predominant in each of Scenes 4, 5 and 6? Carry out a proof, in each of the three cases, of Saunderson's and Euler's assumptions.

4. What do the approaches of Whitehead and Landau have in common, and what distinguishes them from the others? At what levels of maturity, talent or experience would you expect each of these expositions to be helpful?

5. The negative numbers may variously be thought of and described as: created, invented, discovered, explored, imagined, contrived, accepted, recognised, admitted, summoned, introduced, given, imposed, exposed, taken for granted, demanded, smuggled in ... How do you feel about their treatment in each Scene?

6. The American Schoolteacher uses the *distributive law* to justify the rules $(-a)b = -(ab)$, and $(-a)(-b) = ab$. That is, old laws are decreed to hold still for the new numbers; or "maintaining the structure" is held to be of "primary importance". The English Schoolteacher points out that if we devise certain rules for operating with these new "directed numbers" then "all the usual associative commutative and distributive properties hold. "That is, the devised rules are used to justify the persistence of the old laws.

   (i) Prove those sign rules in the style of the American Schoolteacher.

   (ii) Prove the *distributive law* in the style of the English Schoolteacher:

   $^{+}a \cdot (^{+}b + {}^{-}c) = {}^{+}a \cdot {}^{+}b + {}^{+}a \cdot {}^{-}c$, when $b < c$.

   Which of these two directions of deduction do you think is more natural/elegant/appropriate for teaching at various levels?

7. The English Schoolteacher asserts that "the usual notation is confusing," for the same reasons that Whitehead describes a certain habit of mathematics as "puzzling [...] but [...] very convenient." What is this habit? Can you think of other instances of its use in mathematics? Does it cause confusion and puzzlement? Do you think the English Schoolteacher's notation helps to clear up the confusion? How much are we engaged in "directly tracing meanings" in the classroom? What do you think of Whitehead's clam that only rarely need we "be so very pedantic in our symbolism"? At what moments is it useful, in teaching, to explicitly acknowledge (and even highlight with notational devices) the distinct meanings of certain symbols? How can we teach the *convenience* of symbols as finely-honed labour-savers and burden-bearers — as liberators (like all good habits)?

8. Use Apostol's axioms to prove:
   (a) $-(-x) = x$, (b) $(-x)(-y) = xy$, (c) $x < y, z < 0 \rightarrow xz > yz$.

9. Which of Apostol's axioms makes the same point as the English Schoolteacher's final paragraph, that negative numbers are precisely those things which provide solutions to ALL equations of the form $a + x = \mathrm{b}$, where a, b can be any (positive) numbers?

   In the light of this, how may we "construct" the set of integers Z out of the set of natural numbers N? (Hint: Landau constructs the rational numbers as classes of "equivalent" ordered pairs (x_1, x_2); Whitehead constructs the integers Z by means of *operations* on N; think of a system of *ordered pairs* of natural numbers (a,b), grouped together according to whether they determine the same operation in getting from a to b.)

10. Whitehead describes the three great ideas: [*Variable, Form, Generality*], as a "Mathematical Trinity" which is to "preside over the whole subject." If we accept his noble orthodoxy, how, in practice, do we give honour to the Three in our teaching of the subject? Are there tendencies to heresy in any of the Scenes?

11. On the basis of his remarks in the Prologue, how do you think De Morgan would respond to each of the Scenes? Are our students "unmoved"? Where did "algebra go in the future"? What are some of the surprises that would still spring from the "inexhaustible store" of the "mathematical menagerie"?

12. How do you think Klein's critique in the Epilogue applies to each of the Scenes?

References

1. More about De Morgan and his writings can be found in his biography, written by his widow; Sophia Elizabeth De Morgan, *Memoir of Augustus De Morgan*, [London: Longmans, Green, 1882]. See also H. Pycior, "Augustus De Morgan's algebraic work: the three stages," *Isis*, 74 [1983], pp. 211-226; Ivor Grattan-Guinness, "An eye for method: Augustus De Morgan and mathematical education," *Paradigm*, no. 9 [December 1992], pp. 1-7; there are articles on him in the *Dictionary of National Biography* and the *Encyclopaedia Britannica.* For a chronological list of De Morgan's papers and books, see: G. Smith (ed.), *The Boole-De Morgan correspondence 1842–1864*, [Oxford: Clarendon Press, 1982].
2. Augustus De Morgan, *Trans. Camb. Phil. Soc.* 8, Part II [1844], pp. 182-203; pub. 1849.
3. Augustus de Morgan, *The Differential and Integral Calculus* [London: Society for the Diffusion of Useful Knowledge, 1842], p. 566.
4. The last assertion is quoted by Morris Kline in *Mathematics in Western Culture* [New York: Oxford University Press, 1953]. The quote appears on page 170 of the Pelican edition.
5. Up to this point the paragraph is drawn from Augustus De Morgan, "On Infinity and On the Sign of Equality," *Trans. Camb. Phil. Soc.* XI, Part 1 [1864], footnote on p. 38.
6. This paragraph is drawn almost verbatim from De Morgan, "On Infinity."

Continued on page 42.