A CONCRETE APPROACH TO ABSTRACT ALGEBRA

W. W. Sawyer

A Concrete Approach to Abstract Algebra

W. W. Sawyer

NEW YORK

The writing of this book, which was prepared while the author was teaching at the University of Illinois, as a member of the Academic Year Institute, 1957–1958, was supported in part by a grant from the National Science Foundation.

Published in Canada by General Publishing Company, Ltd., 30 Lesmill Road, Don Mills, Toronto, Ontario.
Published in the United Kingdom by Constable and Company, Ltd., 10 Orange Street, London WC2H 7EG.

This Dover edition, first published in 1978, is an unabridged republication of the work originally published in 1959. It is reprinted by special arrangement with the original publisher, W. H. Freeman and Company, 660 Market Street, San Francisco, California 94104.

International Standard Book Number: 0-486-63647-X
Library of Congress Catalog Card Number: 78-52146

Manufactured in the United States of America
Dover Publications, Inc.
180 Varick Street
New York, N.Y. 10014

Contents

Introduction

The Aim of This Book and How to Read It

AT THE PRESENT time there is a widespread desire, particularly among high school teachers and engineers, to know more about "modern mathematics." Institutes are provided to meet this desire, and this book was originally written for, and used by, such an institute. The chapters of this book were handed out as mimeographed notes to the students. There were no "lectures"; I did not in the classroom try to expound the same material again. These chapters *were* the "lectures." In the classroom we simply argued about this material. Questions were asked, obscure points were clarified.

In planning such a course, a professor must make a choice. His aim may be to produce a perfect mathematical work of art, having every axiom stated, every conclusion drawn with flawless logic, the whole syllabus covered. This sounds excellent, but in practice the result is often that the class does not have the faintest idea of what is going on. Certain axioms are stated. How are these axioms chosen? Why do we consider these axioms rather than others? What is the subject about? What is its purpose? If these questions are left unanswered, students feel frustrated. Even though they follow every individual deduction, they cannot think effectively about the subject. The framework is lacking; students do not know where the subject fits in, and this has a paralyzing effect on the mind.

On the other hand, the professor may choose familiar topics as a starting point. The students collect material, work problems, observe regularities, frame hypotheses, discover and prove theorems for themselves. The work may not proceed so quickly; all topics may not be covered; the final outline may be jagged. But the student knows what he is doing and where he is going; he is secure in his mastery of the subject, strengthened in confidence of himself. He has had the experience of discovering mathematics. He no longer thinks of mathematics as static dogma learned by rote. He sees mathematics as something growing and developing, mathematical concepts as something continually revised and enriched in the light of new knowledge. The course may have covered a very limited region, but it should leave the student ready to explore further on his own.

This second approach, proceeding from the familiar to the unfamiliar, is the method used in this book. Wherever possible, I have tried to show how modern higher algebra grows out of traditional elementary algebra. Even so, you may for a time experience some feeling of strangeness. This sense of strangeness will pass; there is nothing you can do about it; we all experience such feelings whenever we begin a new branch of mathematics. Nor is it surprising that such strangeness should be felt. The traditional high school syllabus—algebra, geometry, trigonometry—contains little or nothing discovered since the year 1650 A.D. Even if we bring in calculus and differential equations, the date 1750 A.D. covers most of that. Modern higher algebra was developed round about the years 1900 to 1930 A.D. Anyone who tries to learn modern algebra on the basis of traditional algebra faces some of the difficulties that Rip Van Winkle would have experienced, had his awakening been delayed until the twentieth century. Rip would only overcome that sense of strangeness by riding around in airplanes until he was quite blasé about the whole business.

Some comments on the plan of the book may be helpful. Chapter 1 is introductory and will not, I hope, prove difficult reading. Chapter 2 is rather a long one. In a book for professional mathematicians, the whole content of this chapter would fill only a few lines. I tried to spell out in detail just what those few lines would convey to a mathematician. Chapter 2 was the result. The chapter contains a solid block of rather formal calculations (pages 50–56). Psychologically, it seemed a pity to have such a block early in the book, but logically I did not see where else I could put it. I would advise you not to take these calculations too seriously at a first reading. The ideas are explained before the calculations begin. The calculations are there simply to show that the program can be carried through. At a first reading, you may like to take my word for this and skip pages 50–56. Later, when you have seen the trend of the whole book, you may return to these formal proofs. I would particularly emphasize that the later chapters do not in any way depend on the *details* of these calculations—only on the *results*.

The middle of the book is fairly plain sailing. You should be able to read these chapters fairly easily.

I am indebted to Professor Joseph Landin of the University of Illinois for the suggestion that the book should culminate with the proof that angles cannot be trisected by Euclidean means. This proof, in chapter 11, shows how modern algebraic concepts can be used to solve an ancient problem. This proof is a goal toward which the earlier chapters work.

I assume, if you are a reader of this book, that you are reasonably familiar with elementary algebra. One important result of elementary algebra seems not to be widely known. This is the remainder theorem. It states that when a polynomial $f(x)$ is divided by $x - a$, the remainder is $f(a)$. If you are not familiar with this theorem and its simple proof, it would be wise to review these, with the help of a text in traditional algebra.

Chapter 1

The Viewpoint of Abstract Algebra

THERE ARE two ways in which children do arithmetic—by understanding and by rote. A good teacher, certainly in the earlier stages, aims at getting children to understand what $5 - 2$ and 6×8 mean. Later, he may drill them so that they will answer "48" to the question "Eight sixes?" without having to draw eight sets of six dots and count them.

Suppose a foreign child enters the class. This child knows no arithmetic, and no English, but has a most retentive memory. He listens to what goes on. He notices that some questions are different from others. For instance, when the teacher makes the noise "What day is it today?" the children may make the noise "Monday" or "Tuesday" or "Wednesday" or "Thursday" or "Friday." This question, he notices, has five different answers. There are also questions with two possible answers, "Yes" and "No." For example, to the question "Have you finished this sum?" sometimes one, sometimes the other answer is given.

However, there are questions that always receive the same answer. "Hi" receives the answer "Hi." "Twelve twelves?" receives the answer "A hundred and forty-four"—or, at least, the teacher seems more satisfied when this response is given. Soon the foreign child might learn to make these responses, without realizing that "Hi" and "144" are in rather different categories.

Suppose that the foreign child comes to school after the children in his class have finished working with blocks and beads. He sees $12 \times 12 = 144$ written and hears it spoken, but is never present when 12 is related to the counting of twelve objects.

One cannot say that he understands arithmetic, but he may be top of the class when it comes to reciting the multiplication table. With an excellent memory, he may have complete mastery of formal, mechanical arithmetic.

We may thus separate two elements in arithmetic. (i) The formal element—this covers everything the foreign child can observe and learn. Formal arithmetic is arithmetic seen from the outside. (ii) The intuitive element—the understanding of arithmetic, its meaning, its connection with the actual world. This understanding we derive by being part of the actual universe, by experiencing life and seeing it from the inside.

For teaching, both elements of arithmetic are necessary. But there are certain activities for which the formal approach is helpful. In the formalist philosophy of mathematics, a kind of behaviorist view is taken. Instead of asking "How do mathematicians think?" the formalist philosophers ask "What do mathematicians do?" They look at mathematics from the outside: they see mathematicians writing on paper, and they seek rules or laws to describe how the mathematicians behave.

Formalist philosophy is hardly likely to provide a full picture of mathematics, but it does illuminate certain aspects of mathematics.

A practical application of formalism is the design of all kinds of calculating machines and automatic appliances. A calculating machine is not expected to understand what 71×493 means, but it is expected to give the right answer. A fire alarm is not expected to understand the danger to life and the damage to property involved in a fire. It is expected to ring bells, to turn on sprinklers, and so forth. There may even be some connection between the way these mechanisms operate and the behavior of certain parts of the brain.

One might say that the abstract approach studies what a machine *is*, without bothering about what it is *for*.

Naturally, you may feel it is a waste of time to study a mechanism that has no purpose. But the abstract approach does not imply that a system has no meaning and no use; it merely implies that, for the moment, we are studying the structure of the system, rather than its purpose.

Structure and purpose are in fact two ways of classifying things. In comparing a car and an airplane, you would say that the propeller of an airplane corresponds to the driving wheels of a car if you are thinking in terms of purpose; you would however say that the propeller corresponds to the cooling fan if you are thinking in terms of structure.

Needless to say, a person familiar with all kinds of mechanical structures—wheels, levers, pulleys, and so on—can make use of that knowledge in inventing a mechanism. In a really original invention, a structure might be put to a purpose it had never served before.

Arithmetic Regarded as a Structure

Accordingly, we are going to look at arithmetic from the viewpoint of the foreign student. We shall forget that 12 is a number used for counting, and that + and × have definite meanings. We shall see these things purely as signs written on the keys of a machine.

Stimulus: 12×12.
Response: 144.

Our calculating machine would have the following visible parts:

(i) A space where the first number is recorded.
(ii) A space for the operations $+$, $\times$, $-$, $\div$.
(iii) A space for the second number.

These constitute the input.

The output is the answer, a single number.

Playing around with our machine, we would soon observe certain things. Order is important with $\div$ and $-$. Thus $6 \div 2$ gives the answer 3, while $2 \div 6$ gives the answer 1/3. But order is not important with $+$ and $\times$. Thus $3 + 4$ and $4 + 3$ both give 7; 3×4 and 4×3 both give 12.

We have the *commutative laws:* $a + b = b + a$, $a \times b = b \times a$. (Or $ab = ba$, with the usual convention of leaving out the multiplication sign.)

Commutativity is not something that could have been predicted in advance. Since $6 \div 2$ is not the same as $2 \div 6$, we could not say, for any sign S, that

$$a \text{ S } b = b \text{ S } a.$$

Some comment may be made here on the symbol S. In school algebra, letters usually stand for numbers. In what we are doing, letters stand for *things written on the keys of machines*. The form a S b covers, for example,

a "times" b,
a "plus" b,
a "minus" b,
a "over" b,
a "to the power" b,
a "-th root of" b,
a "-'s log to base" b,

as well as many more complicated ways of combining *a* and *b* that one could devise.

Commutativity, then, is something we may notice about a machine. It is one example of the kind of remark that can be made about a machine.

Ordinary arithmetic has one property that is inconvenient for machine purposes: it is infinite. If we make a calculating machine that goes up to 999,999 we are unable to work out, say, 999,999 + 999,999 or 999,999 × 999,999 by following the ordinary rules for operating the machine.

We can consider a particular calculating machine that is very much simpler, and that avoids the trouble of infinity. This machine will answer any question appropriate to its system. It deals with a particular part or aspect of arithmetic.

If two even numbers are added together, the result is an even number. If an even number is added to an odd number, the result is odd. We may, in fact, write

$$\begin{aligned} \text{Even} + \text{Even} &= \text{Even} \\ \text{Even} + \text{Odd} &= \text{Odd} \\ \text{Odd} + \text{Odd} &= \text{Even.} \end{aligned}$$

Similarly, there are multiplication facts,

$$\begin{aligned} \text{Even} \times \text{Even} &= \text{Even} \\ \text{Even} \times \text{Odd} &= \text{Even} \\ \text{Odd} \times \text{Odd} &= \text{Odd.} \end{aligned}$$

Here we have a miniature arithmetic. There are only two elements in it, Even and Odd. Let us abbreviate, writing A for Even, B for Odd. Then

$$\begin{array}{ll} A + A = A & \qquad A \times A = A \\ A + B = B & \qquad A \times B = A \\ B + A = B & \qquad B \times A = A \\ B + B = A & \qquad B \times B = B \end{array}$$

which may be written more compactly as

$+$	A	B
A	A	B
B	B	A

$\times$	A	B
A	A	A
B	A	B

Our foreign student would have no reason for regarding A and B in the tables above as any different from 0, 1, 2, 3, $\cdots$, in the ordinary addition and multiplication tables. He might think of it as "another arithmetic." He does not know anything about its meaning. What can he observe about its structure? Does it behave at all like ordinary arithmetic? In actual fact, the similarities are very great. I shall only mention a few of them at this stage.

Both addition and multiplication are commutative in the A, B arithmetic. For instance, $A + B = B + A$ and $A \times B = B \times A$.

In ordinary arithmetic the number zero occurs. We know the meaning of zero. But how could zero be identified by someone who only saw the structure of arithmetic? Quite easily, for there are two properties of zero that single it out. First, when zero is added to a number, it makes no difference. Second, whatever number zero is multiplied by, the result is always zero.

Thus

$$x + 0 = x,$$
$$x \cdot 0 = 0.$$

Is there a symbol in the A, B arithmetic that plays the role of zero? It makes a difference when B is added: $A + B$ is not A, nor is $B + B$ the same as B. A is the only possible candidate, and in fact A passes all the tests. When you add A, it makes no difference; when anything is multiplied by A, you get A.

Is there anything that corresponds to 1? The only distinguishing property I can think of for 1 is that multiplication by 1 has no visible effect:

$$x \cdot 1 = x.$$

In the *A*, *B* arithmetic, multiplication by *B* leaves any symbol unchanged. So *B* plays the part of 1.

This suggests that we might have done better to choose O (capital o) as a symbol instead of *A* and I as a symbol instead of *B*, because O looks like zero, and I looks rather like 1.

Our tables would then read

O + O = O	O × O = O
O + I = I	O × I = O
I + O = I	I × O = O
I + I = O	I × I = I

Now this looks very much like ordinary arithmetic. In fact, the only question that would be raised by somebody who thought I stood for 1 and O for zero would be, "Haven't you made a mistake in writing I + I = O?" All the other statements are exactly what you would expect from ordinary arithmetic.

The tables of this "arithmetic" are

+	O	I
O	O	I
I	I	O

×	O	I
O	O	O
I	O	I

We arrived at the tables above by considering even and odd numbers. But we could arrive at the same pattern without any mention of numbers.

Imagine the following situation. There is a narrow bridge with automatic signals. If a car approaches from either end, a signal "All clear—Proceed" is flashed on. But if cars approach from both ends, a warning signal is flashed, and the car at, say, the north end is instructed to withdraw.

In effect, the mechanism asks two questions: "Is a car approaching from the south? Is one approaching from the north?" The answers to these questions are the input, the stimulus. The output, the response of the mechanism, is to switch on an appropriate signal.

For the all-clear signal the scheme is as follows.

Should all-clear signal be flashed?

		Car from north? No	Car from north? Yes
Car from south?	No	No	Yes
	Yes	Yes	No

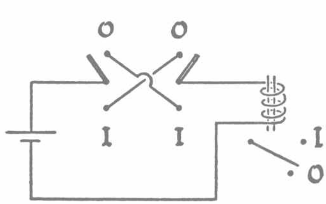
For the warning signal the scheme is as follows.

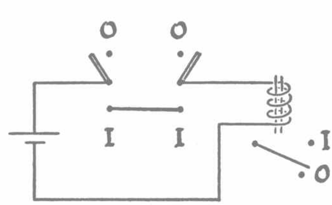
Should warning signal be flashed?

		Car from north? No	Car from north? Yes
Car from south?	No	No	No
	Yes	No	Yes

If you compare these tables with the earlier ones, you will see that they are exactly the same in structure. "No" replaces "O," "Yes" replaces "I"; "all clear" is related to +, "warning signal" to ×.

One could also realize this pattern by simple electrical circuits.





If you had this machine in front of you, you would not know whether it was intended for calculations with even and odd numbers, or for traffic control, or for some other purpose.

When the same pattern is embodied in two different systems, the systems are called *isomorphic*. In our example above, the traffic control system is isomorphic with the arithmetic of Even and Odd. The same machine does for both.

Isomorphism does not simply mean that there is some general resemblance between the two systems. It means that they have exactly the same pattern. Our example above shows this exact correspondence. Wherever "O" occurs in one system, "No" occurs in the other; wherever "I" occurs in one system, "Yes" occurs in the other.

The statements, "these two systems are isomorphic" and "there is an isomorphism between them," are two different ways of saying the same thing. To prove two systems isomorphic, you must demonstrate a correspondence between them, like the one in our example.

The study of structures has two things to offer us. First, the same structure may have many different realizations. By studying the single structure, we are simultaneously learning several different subjects.

Second, even though we have only one realization of our structure in mind, we may be able to simplify our proofs and clarify our understanding of the subject by treating it abstractly—that is to say, by leaving out details that merely complicate the picture and are not relevant to our purpose.

Our Results Considered Abstractly

So far we have been *concrete* in our approach. That is, we have been talking of things whose meaning we under-