CHAPTER 30 UNIQUENESS OF THE REAL NUMBERS

We shall now revert to the usual notation for real numbers, reserving boldface symbols for other fields which may turn up. Moreover, we will regard integers and rational numbers as special kinds of real numbers, and forget about the specific way in which real numbers were defined. In this chapter we are interested in only one question: are there any complete ordered fields other than $\mathbf{R}$? The answer to this question, if taken literally, is "yes." For example, the field F_3 introduced in Chapter 28 is a complete ordered field, and it is certainly not $\mathbf{R}$. This field is a "silly" example because the pair (a, a) can be regarded as just another name for the real number a; the operations

$$(a, a) \mathbin{\boldsymbol{+}} (b, b) = (a + b, a + b),$$
$$(a, a) \mathbin{\boldsymbol{\cdot}} (b, b) = (a \cdot b, a \cdot b),$$

are consistent with this renaming. This sort of example shows that any intelligent consideration of the question requires some mathematical means of discussing such renaming procedures.

If the elements of a field F are going to be used to rename elements of $\mathbf{R}$, then for each a in $\mathbf{R}$ there should correspond a "name" $f(a)$ in F. The notation $f(a)$ suggests that renaming can be formulated in terms of functions. In order to do this we will need a concept of function much more general than any which has occurred until now; in fact, we will require the most general notion of "function" used in mathematics. A function, in this general sense, is simply a rule which assigns to some things, other things. To be formal, a **function** is a collection of ordered pairs (of objects of any sort) which does not contain two distinct pairs with the same first element. The **domain** of a function f is the set A of all objects a such that (a, b) is in f for some b; this (unique) b is denoted by $\boldsymbol{f(a)}$. If $f(a)$ is in the set B for all a in A, then f is called a function **from** A **to** B. For example,

> if $f(x) = \sin x$ for all x in $\mathbf{R}$ (and f is defined only for x in $\mathbf{R}$), then f is a function from $\mathbf{R}$ to $\mathbf{R}$; it is also a function from $\mathbf{R}$ to $[-1, 1]$;
>
> if $f(z) = \sin z$ for all z in $\mathbf{C}$, then f is a function from $\mathbf{C}$ to $\mathbf{C}$;
>
> if $f(z) = e^z$ for all z in $\mathbf{C}$, then f is a function from $\mathbf{C}$ to $\mathbf{C}$; it is also a function from $\mathbf{C}$ to $\{z \text{ in } \mathbf{C} : z \neq 0\}$;
>
> θ is a function from $\{z \text{ in } \mathbf{C} : z \neq 0\}$ to $\{x \text{ in } \mathbf{R} : 0 \leq x < 2\pi\}$;
>
> if f is the collection of all pairs $(a, (a, a))$ for a in $\mathbf{R}$, then f is a function from $\mathbf{R}$ to F_3.

THEOREM If F is a complete ordered field, then F is isomorphic to $\mathbf{R}$.

PROOF Since two fields are defined to be isomorphic if there is an isomorphism between them, we must actually construct a function f from $\mathbf{R}$ to F which is an isomorphism. We begin by defining f on the integers as follows:

$$\begin{aligned} f(0) &= \mathbf{0}, \\ f(n) &= \underbrace{\mathbf{1} \boldsymbol{+} \dots \boldsymbol{+} \mathbf{1}}_{n \text{ times}} \qquad \text{for } n > 0, \\ f(n) &= \boldsymbol{-}\underbrace{(\mathbf{1} \boldsymbol{+} \dots \boldsymbol{+} \mathbf{1})}_{|n| \text{ times}} \qquad \text{for } n < 0. \end{aligned}$$

It is easy to check that

$$\begin{aligned} f(m+n) &= f(m) \boldsymbol{+} f(n), \\ f(m \cdot n) &= f(m) \boldsymbol{\cdot} f(n), \end{aligned}$$

for all integers m and n, and it is convenient to denote $f(n)$ by $\boldsymbol{n}$. We then define f on the rational numbers by

$$f(m/n) = \boldsymbol{m/n} = \boldsymbol{m} \boldsymbol{\cdot} \boldsymbol{n}^{\boldsymbol{-1}}$$

(notice that the n-fold sum $\mathbf{1} \boldsymbol{+} \cdots \boldsymbol{+} \mathbf{1} \neq \mathbf{0}$ if $n > 0$, since F is an ordered field). This definition makes sense because if $m/n = k/l$, then $ml = nk$, so $\boldsymbol{m \cdot l} = \boldsymbol{k \cdot n}$, so $\boldsymbol{m} \boldsymbol{\cdot} \boldsymbol{n}^{\boldsymbol{-1}} = \boldsymbol{k} \boldsymbol{\cdot} \boldsymbol{l}^{\boldsymbol{-1}}$. It is easy to check that

$$\begin{aligned} f(r_1 + r_2) &= f(r_1) \boldsymbol{+} f(r_2), \\ f(r_1 \cdot r_2) &= f(r_1) \boldsymbol{\cdot} f(r_2), \end{aligned}$$

for all rational numbers r_1 and r_2, and that $f(r_1) \boldsymbol{<} f(r_2)$ if $r_1 < r_2$.

The definition of $f(x)$ for arbitrary x is based on the now familiar idea that any real number is determined by the rational numbers less than it. For any x in $\mathbf{R}$, let A_x be the subset of F consisting of all $f(r)$, for all rational numbers $r < x$. The set A_x is certainly not empty, and it is also bounded above, for if r_0 is a rational number with $r_0 > x$, then $f(r_0) \boldsymbol{>} f(r)$ for all $f(r)$ in A_x. Since F is a complete ordered field, the set A_x has a least upper bound; we define $f(x)$ as $\sup A_x$.

We now have $f(x)$ defined in two different ways, first for rational x, and then for any x. Before proceeding further, it is necessary to show that these two definitions agree for rational x. In other words, if x is a rational number, we want to show that

$$\sup A_x = f(x),$$

where $f(x)$ here denotes $\boldsymbol{m/n}$, for $x = m/n$. This is not automatic, but depends on the completeness of F; a slight digression is thus required.

Since F is complete, the elements

$$\underbrace{\mathbf{1} \boldsymbol{+} \dots \boldsymbol{+} \mathbf{1}}_{n \text{ times}} \qquad \text{for natural numbers } n$$

Suppose that F_1 and F_2 are two fields; we will denote the operations in F_1 by $\oplus$, $\odot$, etc., and the operations in F_2 by $+$, $\cdot$, etc. If F_2 is going to be considered as a collection of new names for elements of F_1, then there should be a function from F_1 to F_2 with the following properties:

(1) The function f should be one-one, that is, if $x \neq y$, then we should have $f(x) \neq f(y)$; this means that no two elements of F_1 have the same name.
(2) The function f should be "onto," that is, for every element z in F_2 there should be some x in F_1 such that $z = f(x)$; this means that every element of F_2 is used to name some element of F_1.
(3) For all x and y in F_1 we should have

$$f(x \oplus y) = f(x) + f(y),$$
$$f(x \odot y) = f(x) \cdot f(y);$$

this means that the renaming procedure is consistent with the operations of the field.

If we are also considering F_1 and F_2 as ordered fields, we add one more requirement:

(4) If $x \otimes y$, then $f(x) < f(y)$.

A function with these properties is called an *isomorphism* from F_1 to F_2. This definition is so important that we restate it formally.

DEFINITION

If F_1 and F_2 are two fields, an **isomorphism** from F_1 to F_2 is a function f from F_1 to F_2 with the following properties:

(1) If $x \neq y$, then $f(x) \neq f(y)$.
(2) If z is in F_2, then $z = f(x)$ for some x in F_1.
(3) If x and y are in F_1, then

$$f(x \oplus y) = f(x) + f(y),$$
$$f(x \odot y) = f(x) \cdot f(y).$$

If F_1 and F_2 are ordered fields we also require:

(4) If $x \otimes y$, then $f(x) < f(y)$.

The fields F_1 and F_2 are called **isomorphic** if there is an isomorphism between them. Isomorphic fields may be regarded as essentially the same—any important property of one will automatically hold for the other. Therefore, we can, and should, reformulate the question asked at the beginning of the chapter; if F is a complete ordered field it is silly to expect F to equal $\mathbf{R}$—rather, we would like to know if F is isomorphic to $\mathbf{R}$. In the following theorem, F will be a field, with operations $+$ and $\cdot$, and "positive elements" $\mathbf{P}$; we write $a < b$ to mean that $b - a$ is in $\mathbf{P}$, and so forth.

form a set which is not bounded above; the proof is exactly the same as the proof for $\mathbf{R}$ (Theorem 8-2). The consequences of this fact for $\mathbf{R}$ have exact analogues in F: in particular, if a and b are elements of F with $a \prec b$, then there is a rational number r such that

$$a \prec f(r) \prec b.$$

Having made this observation, we return to the proof that the two definitions of $f(x)$ agree for rational x. If y is a rational number with $y < x$, then we have already seen that $f(y) \prec f(x)$. Thus every element of A_x is $\prec f(x)$. Consequently,

$$\sup A_x \preceq f(x).$$

On the other hand, suppose that we had

$$\sup A_x \prec f(x).$$

Then there would be a rational number r such that

$$\sup A_x \prec f(r) \prec f(x).$$

But the condition $f(r) \prec f(x)$ means that $r < x$, which means that $f(r)$ is in the set A_x; this clearly contradicts the condition $\sup A_x \prec f(r)$. This shows that the original assumption is false, so

$$\sup A_x = f(x).$$

We thus have a certain well-defined function f from $\mathbf{R}$ to F. In order to show that f is an isomorphism we must verify conditions (1)–(4) of the definition. We will begin with (4).

If x and y are real numbers with $x < y$, then clearly A_x is contained in A_y. Thus

$$f(x) = \sup A_x \preceq \sup A_y = f(y).$$

To rule out the possibility of equality, notice that there are rational numbers r and s with

$$x < r < s < y.$$

We know that $f(r) \prec f(s)$. It follows that

$$f(x) \preceq f(r) \prec f(s) \preceq f(y).$$

This proves (4).

Condition (1) follows immediately from (4): If $x \neq y$, then either $x < y$ or $y < x$; in the first case $f(x) \prec f(y)$, and in the second case $f(y) \prec f(x)$; in either case $f(x) \neq f(y)$.

To prove (2), let a be an element of F, and let B be the set of all rational numbers r with $f(r) \prec a$. The set B is not empty, and it is also bounded above, because there is a rational number s with $f(s) \succ a$, so that $f(s) \succ f(r)$ for r in B, which implies that $s > r$. Let x be the least upper bound of B; we claim that $f(x) = a$. In order to prove this it suffices to eliminate the alternatives

$$f(x) \prec a,$$
$$a \prec f(x).$$

In the first case there would be a rational number r with

$$f(x) \prec f(r) \prec a.$$

But this means that $x < r$ and that r is in B, which contradicts the fact that $x = \sup B$. In the second case there would be a rational number r with

$$a \prec f(r) \prec f(x).$$

This implies that $r < x$. Since $x = \sup B$, this means that $r < s$ for some s in B. Hence

$$f(r) \prec f(s) \prec a,$$

again a contradiction. Thus $f(x) = a$, proving (2).

To check (3), let x and y be real numbers and suppose that $f(x + y) \neq f(x) \boldsymbol{+} f(y)$. Then either

$$f(x + y) \prec f(x) \boldsymbol{+} f(y) \qquad \text{or} \qquad f(x) \boldsymbol{+} f(y) \prec f(x + y).$$

In the first case there would be a rational number r such that

$$f(x + y) \prec f(r) \prec f(x) \boldsymbol{+} f(y).$$

But this would mean that

$$x + y < r.$$

Therefore r could be written as the sum of two rational numbers

$$r = r_1 + r_2, \qquad \text{where } x < r_1 \text{ and } y < r_2.$$

Then, using the facts checked about f for *rational* numbers, it would follow that

$$f(r) = f(r_1 + r_2) = f(r_1) \boldsymbol{+} f(r_2) \succ f(x) \boldsymbol{+} f(y),$$

a contradiction. The other case is handled similarly.

Finally, if x and y are positive real numbers, the same sort of reasoning shows that

$$f(x \cdot y) = f(x) \bullet f(y);$$

the general case is then a simple consequence. ▮

This theorem brings to an end our investigation of the real numbers, and resolves any doubts about them: There *is* a complete ordered field and, up to isomorphism, only one complete ordered field. It is an important part of a mathematical education to follow a construction of the real numbers in detail, but it is not necessary to refer ever again to this particular construction. It is utterly irrelevant that a real number happens to be a collection of rational numbers, and such a fact should never enter the proof of any important theorem about the real numbers. Reasonable proofs should use only the fact that the real numbers are a complete ordered field, because this property of the real numbers characterizes them up to isomorphism, and any significant mathematical property of the real numbers will be true for all isomorphic fields. To be candid I should admit that this last assertion is just a prejudice of the author, but it is one shared by almost all other mathematicians.

PROBLEMS

1. Let f be an isomorphism from F_1 to F_2.

(a) Show that $f(\mathbf{0}) = \mathbf{0}$ and $f(\mathbf{1}) = \mathbf{1}$. (Here $\mathbf{0}$ and $\mathbf{1}$ on the left denote elements in F_1, while $\mathbf{0}$ and $\mathbf{1}$ on the right denote elements of F_2.)

(b) Show that $f(\boldsymbol{-}a) = \boldsymbol{-}f(a)$ and $f(a^{\mathbf{-1}}) = f(a)^{\mathbf{-1}}$, for $a \neq \mathbf{0}$.

2. Here is an opportunity to convince yourself that any significant property of a field is shared by any field isomorphic to it. The point of this problem is to write out very formal proofs until you are certain that all statements of this sort are obvious. F_1 and F_2 will be two fields which are isomorphic; for simplicity we will denote the operations in both by $\boldsymbol{+}$ and $\boldsymbol{\cdot}$. Show that:

(a) If the equation $x^2 \boldsymbol{+} \mathbf{1} = \mathbf{0}$ has a solution in F_1, then it has a solution in F_2.

(b) If every polynomial equation $x^n \boldsymbol{+} a_{n-1} \boldsymbol{\cdot} x^{n-1} \boldsymbol{+} \cdots \boldsymbol{+} a_0 = \mathbf{0}$ with $a_0, \dots, a_{n-1}$ in F_1, has a root in F_1, then every polynomial equation $x^n \boldsymbol{+} b_{n-1} \boldsymbol{\cdot} x^{n-1} \boldsymbol{+} \cdots \boldsymbol{+} b_0 = \mathbf{0}$ with $b_0, \dots, b_{n-1}$ in F_2 has a root in F_2.

(c) If $1 \boldsymbol{+} \cdots \boldsymbol{+} 1$ (summed m times) $= \mathbf{0}$ in F_1, then the same is true in F_2.

(d) If F_1 and F_2 are ordered fields (and the isomorphism f satisfies $f(x) \boldsymbol{<} f(y)$ for $x \boldsymbol{<} y$) and F_1 is complete, then F_2 is complete.

3. Let f be an isomorphism from F_1 to F_2 and g an isomorphism from F_2 to F_3. Define the function $g \circ f$ from F_1 to F_3 by $(g \circ f)(x) = g(f(x))$. Show that $g \circ f$ is an isomorphism.

4. Suppose that F is a complete ordered field, so that there is an isomorphism f from $\mathbf{R}$ to F. Show that there is actually only *one* isomorphism from $\mathbf{R}$ to F. Hint: In case $F = \mathbf{R}$, this is Problem 3-17. Now if f and g are two isomorphisms from $\mathbf{R}$ to F consider $g^{-1} \circ f$.

5. Find an isomorphism from $\mathbf{C}$ to $\mathbf{C}$ other than the identity function.

SUGGESTED READING

A man ought to read
just as inclination leads him;
for what he reads as a task
will do him little good.

SAMUEL JOHNSON

One purpose of this bibliography is to guide the reader to other sources, but the most important function it can serve is to indicate the variety of mathematical reading available. Consequently, there is an attempt to achieve diversity, but no pretense of being complete. The present plethora of mathematics books would make such an undertaking almost hopeless in any case, and since I have tried to encourage independent reading, the more standard a text, the less likely it is to appear here. In some cases, this philosophy may seem to have been carried to extremes, as some entries in the list cannot be read by a student just finishing a first course of calculus until several years have elapsed. Nevertheless, there are many selections which can be read now, and I can't believe that it hurts to have some idea of what lies ahead.

For most references, only the title and author have been given, since so many of these books have gone through numerous editions and printings, often having gone out of print at some point only to be resurrected later on by a different publisher (often as an inexpensive paperback by the redoubtable Dover Publications or by the Mathematical Association of America). More exact information really isn't necessary, since it is now so easy to search for books on-line at Amazon.com and other sites.

† is used to indicate books whose availability, either new or used, is problematical. Author and title searches may turn up other intriguing books by the same author, or other books with similar titles. In addition, many of these books will still be found in well-stocked academic libraries, perhaps the best place of all to search; despite the convenience of the internet, nothing matches the experience of an actual (as opposed to a virtual) library, with books stacked according to subject, awaiting serendipitous discovery.

One of the most elementary unproved theorems mentioned in this book is the "Fundamental Theorem of Arithmetic", that every natural number can be written as a product of primes in only one way. This follows from the basic fact alluded to on page 444, a proof of which will be found near the beginning of almost any book on elementary number theory. Few books have won so enthusiastic an audience as

[1] *An Introduction to the Theory of Numbers*, by G. H. Hardy and E. M. Wright.

Two other recommended books are

† [2] *A Selection of Problems in the Theory of Numbers*, by W. Sierpinski.
[3] *Three Pearls of Number Theory*, by A. Khinchin.

The Fundamental Theorem also applies in more general algebraic settings, see references [33] and [34].

The subject of irrational numbers straddles the fields of number theory and analysis. An excellent introduction will be found in

[4] *Irrational Numbers*, by I. M. Niven.