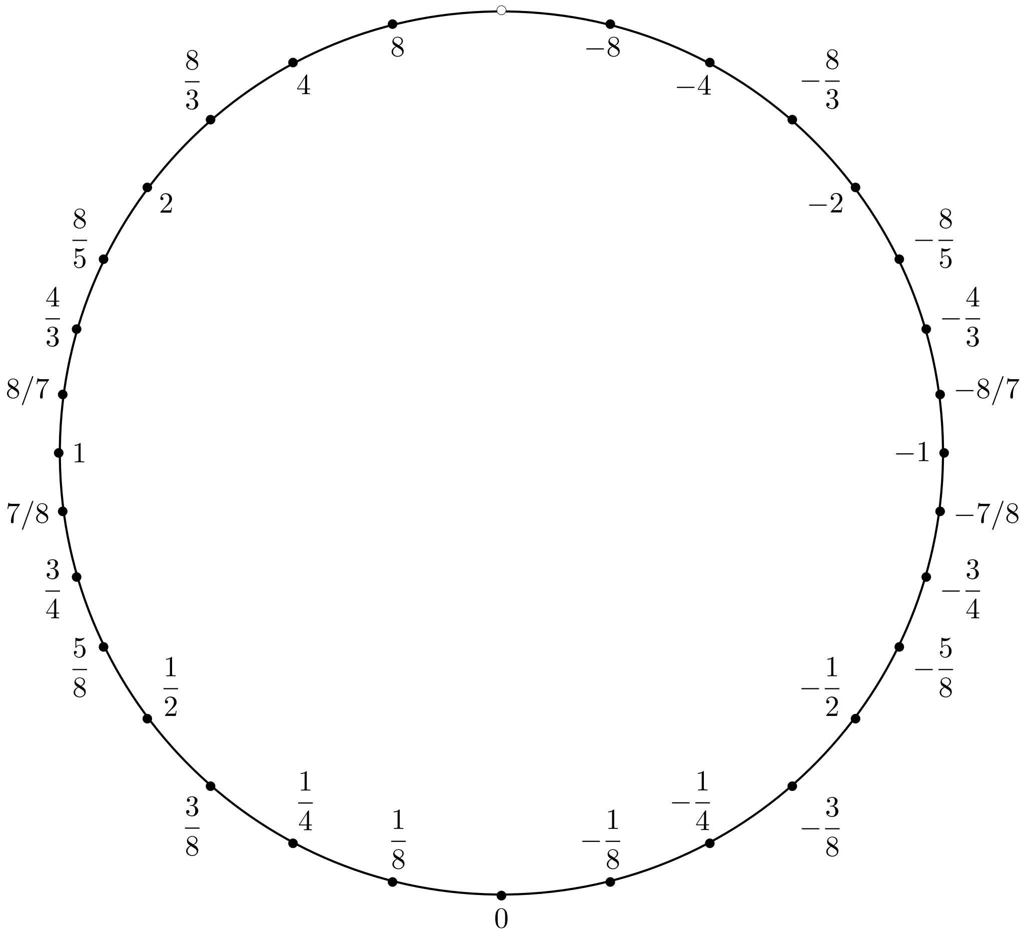
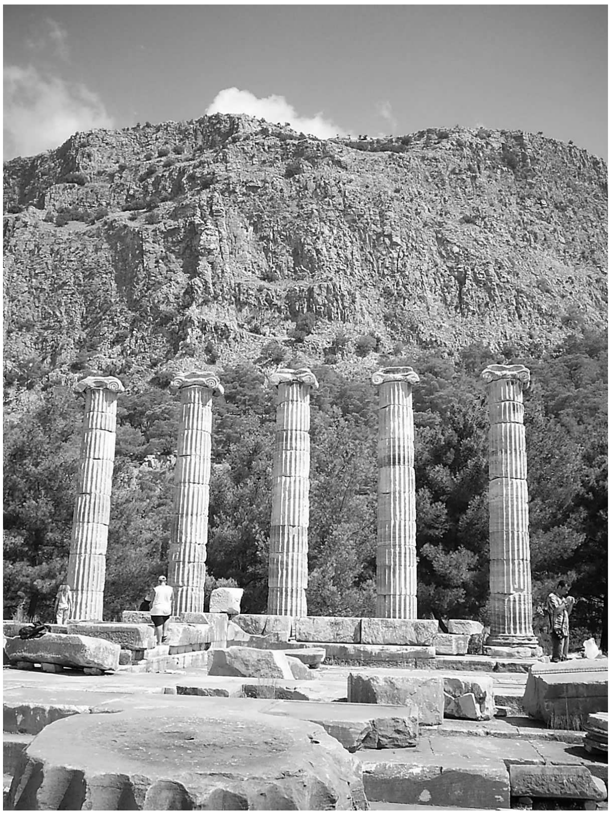
Non-standard analysis



David Pierce



The photograph on the previous page
was taken in Priene, Söke, Aydın, in 2008.
The columns are in the Ionic order; see p. 14.
Another version of the cover image is Fig. 4.4 of the text.

Non-standard analysis

David Pierce

June 29, 2010

Mathematics Department
Middle East Technical University
Ankara 06531 Turkey
http://metu.edu.tr/~dpierce/
dpierce@metu.edu.tr

Preface

This book is prepared for a course called Non-standard Analysis, given in July, 2010, at the Nesin Matematik Köyü (Nesin Mathematics Village), Şirince, Selçuk, İzmir, Turkey.

The book is a thorough revision of the notes prepared for the 2009 version of the same course. The course and the book could also be called Foundations of Analysis, like Landau's book [19]. Like Landau, I construct the real numbers, albeit with more attention to the *logic* of the construction. Unlike Landau, I construct also a field of so-called *hyper-real* or *non-standard real numbers;* then I show the use of this field in establishing some of the basic theorems of analysis.

The book is not a textbook, but just a record of what I have worked out so far for myself and for anybody else who may be interested. A week of lectures in Şirince can cover only lightly what is here. This book itself only lightly covers some topics.

The earlier version of the notes had mistakes and other infelicities; I have corrected some of these, but others remain. Mathematics is like art, as Collingwood [7, p. 2] describes it in recalling his childhood:

> . . . I was constantly watching the work of my father and mother, and the other professional painters who frequented the house, and constantly trying to imitate them; so that I learned to think of a picture not as a finished product exposed for the admiration of virtuosi, but as the visible record, lying about the house, of an attempt to solve a definite problem in painting, so far as the attempt has gone. I learned what some critics and aestheticians never know to the end of their lives, that no 'work of art' is ever finished, so that in that sense of the phrase there is no such thing as a 'work of art' at all.

It is time to print these notes. Then the next round of revision begins.

D. P.
Ankara
June 29, 2010

Contents

List of Figures

Introduction

Mathematical analysis is the theoretical side of calculus. Calculus consists of methods of solving certain sorts of problems; analysis studies those methods. The **standard** way of solving calculus problems is founded on the 'epsilon-delta' definition of *limit.* The **non-standard** approach uses *infinitesimals,* with rigorous logical justification. Abraham Robinson first gave this justification, which can be found in his book *Non-standard Analysis* [25].

An **infinitesimal** is a number whose absolute value is less than than every positive rational number. If two numbers a and b differ by an infinitesimal, we shall write

$$a \simeq b.$$

Zero is an infinitesimal, but there are no other infinitesimals among the so-called *real numbers.* Indeed, suppose ε is a positive real number. Then some integer n is greater than $1/\varepsilon$, so $0 < 1/n < \varepsilon$. Thus ε is not infinitesimal.

In standard analysis, a function f is said to be **continuous** at an element a of its domain if

$$\lim_{x \to a} f(x) = f(a);$$

this means that, for all positive numbers ε, there is a positive number δ such that, for all x in the domain of f, if $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$.

In non-standard analysis, there is an alternative formulation of continuity: f is continuous at a just in case, for all x in the domain of f, if $x \simeq a$, then $f(x) \simeq f(a)$.

The alternative formulation of continuity and many other things will be worked out in Chapter 7, the last chapter of these notes. The other chapters are meant to provide logical justication and motivation for the work of Chapter 7. Chapter 1 looks at Archimedes's solution of a calculus problem, and it also mentions the *Archimedean axiom,* which we used above to prove that no non-zero real numbers are infinitesimal. Today we think of calculus as involving the *complete ordered field* $\mathbb{R}$ of real numbers; this field is constructed in Chapters 2, 3, and 4. As we have seen, non-standard analysis requires a certain larger ordered field, which will be denoted by ${}^*\mathbb{R}$: this is an example of a *non-archimedean* ordered field. Non-archimedean ordered fields in general, and simple examples of these and related fields, are discussed in Chapter 5. The field ${}^*\mathbb{R}$ can be obtained as an *ultrapower* of $\mathbb{R}$; this construction is treated in Chapter 6. Readers can jump ahead to Chapter 7 at any time, provided they understand the meaning of Theorem 89 in § 6.3.

1. Archimedes's quadrature of the parabola

In the last chapter of *Non-standard Analysis* [25], Robinson treats the history of calculus in the light of non-standard analysis. Robinson begins with Leibniz; but I think it worthwhile to go back much further—about two thousand years further. In the work of Archimedes, both standard and non-standard approaches to calculus (in our terms) can be discerned. For example, Archimedes takes up the following

Problem 1. *Find a square equal to a given segment of a parabola.*

Parabolas (in the sense of this problem) will be defined below; meanwhile, a segment of a parabola can be seen in Figure 1.1, with an *inscribed* triangle. A solution to Problem 1 is a **quadrature:** a 'squaring' of the parabola. Archimedes's solution is given by the following

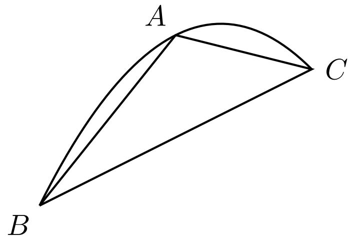


Figure 1.1. A parabolic segment with the inscribed triangle

Theorem 2 (Archimedes)**.** *A parabolic segment is a third again as large* as the inscribed triangle.*

In Figure 1.1, I say triangle ABC is **inscribed** in the parabolic segment cut off by the chord AB, because the tangent to the parabola at A is parallel to the chord BC. Once we have Archimedes's theorem, we can carry out the actual squaring of the parabolic segment by a standard procedure described in a series of propositions (I.42, I.45, II.14) in Euclid's *Elements* [10, 11]. Indeed, from Figure 1.1, we can obtain Figure 1.2 as follows. Bisect BC at D. Extend CB to E so that $BE = \frac{1}{3}DB$. Draw straight line AF parallel to BC, and construct

*ἐπίτριτος, one and a third times as much. See Appendix A for the Greek letters.

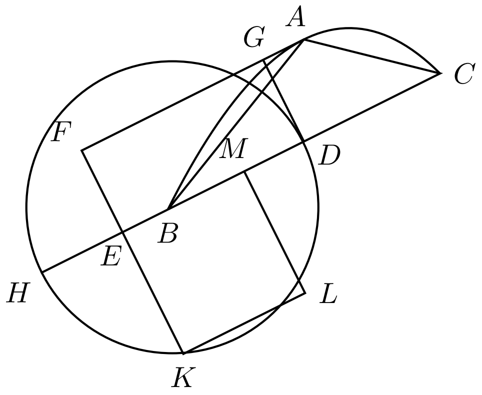


Figure 1.2. Quadrature of a parabolic segment

rectangle $DEFG$: this is a third again as large as triangle ABC, so it is equal to the parabolic segment on chord BC by Theorem 2. Extend BE to H so that $EH = EF$. Draw a circle with diameter DH, and extend FE to meet the circle at K. Then EK is a mean proportional to EF and ED (by *Elements* VI.13), so the square $EKLM$ constructed on EK is equal to the parabolic segment on chord BC (by *Elements* VI.17 or just II.14).

So Problem 1 can be solved by means of Theorem 2. Archimedes proves Theorem 2 in two ways in his *Quadrature of the parabola:* in Proposition 17 (and the propositions leading up to it), and in Proposition 24. Heath [3] provides an English version of this work; however, instead of simply translating, he rewrites Archimedes in a way intended to be more comprehensible to his modern readers. Selections from the Greek text of Archimedes, with more literal English translations, are provided by Thomas [29]. The first volume of a faithful translation of all of Archimedes's works by Netz [4] has appeared; but this does not contain the works that we are particularly interested in here.

Insight into the *discovery* of Theorem 2 is given in Archimedes's *Method.* This work was lost until 1906. Then, in İstanbul, the Danish scholar J. L. Heiberg discovered the *Archimedes Palimpsest:* a parchment codex of the works of Archimedes that had been washed and reused for writing prayers.

1.1. Parabolas

Archimedes does not use the word *parabola* [3, p. clxvii]; for him, a parabola is a *section of an orthogonal cone.*[*] Let me review what this means, sometimes

[*] ὀρθογωνίου κώνου τομή.

following also the account of Apollonius of Perga [1]. A cone is determined by a circle, called the **base,** and a point, called the **apex.** The apex is not in the plane of the base. The cone is traced out by a straight line that has one endpoint at the apex: the other endpoint is moved about the circumference of the base. The straight line drawn from the apex to the center of the base is the **axis** of the cone. A plane containing the axis intersects the cone in an **axial triangle.** See Figure 1.3. If the axis is perpendicular to the base, then the cone is a **right cone.** If an axial triangle of a right cone has a right angle at the apex, then the cone is an **orthogonal cone.*** Suppose a plane is perpendicular to one of the sides of an axial triangle of an orthogonal cone. Then the plane cuts the cone in a curve that—following Apollonius—we call a **parabola.** See Figure 1.4. The intersection of the cutting plane and the axial triangle is the straight line called the **axis** of the parabola (and this is different from the axis of the cone itself);

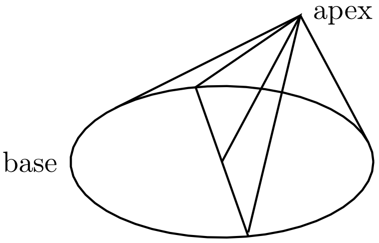


Figure 1.3. A cone, with an axial triangle

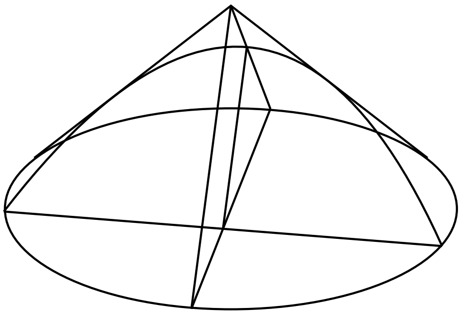


Figure 1.4. A parabola in a cone

*Or ‘right-angled cone’; but not every *right cone* is a *right-angled cone.*

the intersection of the axis of the parabola and the parabola itself is the point called the **vertex** of the parabola.

A straight line drawn from the parabola to the axis perpendicularly to the axis is an **ordinate;** the part of the axis between the ordinate and the vertex is the corresponding **abscissa.** See Figure 1.5. The word *abscissa* means *cut off* in Latin, while the word *ordinate* is related to *order.* The word *order* is used for any of the several classical styles of architecture. An order in this sense features columns standing parallel, like ordinates of a parabola. For an example, see p. 1.

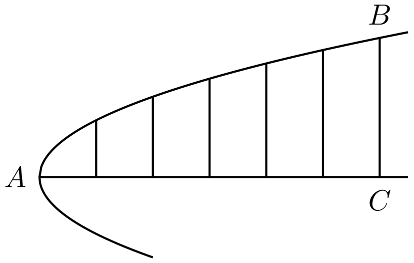


Figure 1.5. The ordinate BC cuts off from the axis the abscissa AC

Apollonius's reason for using the term *parabola* is shown in Proposition 11 of his *Conics* [1, 29]. Apollonius also shows that parabolas can be obtained from *all* cones, not necessarily orthogonal, not necessarily right. What is important for us are the following properties of a parabola, whose proofs can be found in Apollonius.

¶ 1. The squares on two ordinates are in the ratio of the corresponding abscissas [1, I.11]; in modern terms, there is a *parameter,* ℓ, such that, if x is an ordinate, and y the corresponding abscissa, then $\ell y = x^2$.

¶ 2. Suppose a parabola has the vertex A, and another point B is chosen on the parabola, and the ordinate BC is drawn, as in Figure 1.6. Extend the axis CA beyond A to a point D. The straight line BD is tangent to the parabola at B if and only if $AD = CA$ [1, I.33, 35].

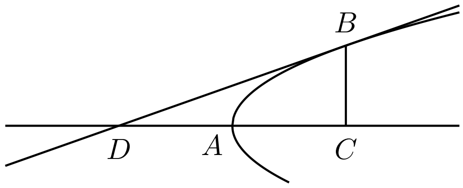


Figure 1.6. DB is tangent at B if and only if $CA = AD$

¶ 3. Every straight line parallel to the axis is a **diameter** in the following sense. Where such a straight line meets the parabola, a tangent can be drawn. If a chord of the parabola is drawn parallel to this tangent, then the given straight line bisects the chord [1, I.46], as in Figure 1.7. Half of such a chord can be called an ordinate with respect to the corresponding diameter, and the squares on two such ordinates are in the ratio of the corresponding abscissas, as in ¶ 1 [1, I.49].

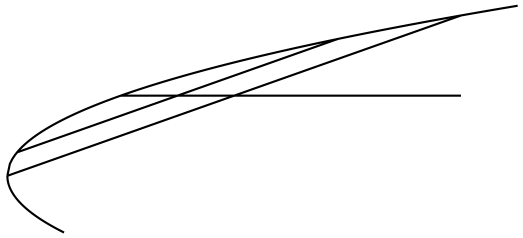


Figure 1.7. A new diameter for a parabola

1.2. The mechanical argument

In the *Method,* Archimedes solves Problem 1 as follows. We add some straight lines to Figure 1.1, getting Figure 1.8. Here D is the midpoint of BC, so that, since BC is parallel to the tangent at A, the straight line AD must be a diameter of the parabola by ¶ 3 above. The tangent to the parabola at C meets DA extended at E. Then A is the midpoint of DE, by ¶ 2 above. A straight line from B drawn parallel to DA meets CE extended at F. Extend CA to meet BF at G, then extend further to H so that $GH = CG$. The idea now is to consider CH as a *lever* with *fulcrum* at G. Conceiving our figures as having weights proportional to their sizes, we shall show that, if we place the weight of the parabolic segment ABC at H, then it will just balance triangle BCF where it is.

Since A is the midpoint of DE, also G is the midpoint of BF. Let DF be drawn, intersecting CG at K. Since D is the midpoint of BC, we can conclude that K is the center of gravity (or centroid) of triangle BCF. Then GK is a third of CG, hence a third of GH. Therefore triangle BCF is balanced by a third of its weight at H. If we can show that the parabolic segment balances the triangle, then the segment must be a third of the triangle. But triangle BCF is twice triangle BCG, which is twice triangle ABC. Then Theorem 2 follows.

To show that, when placed at H, the parabolic segment balances BCF, we

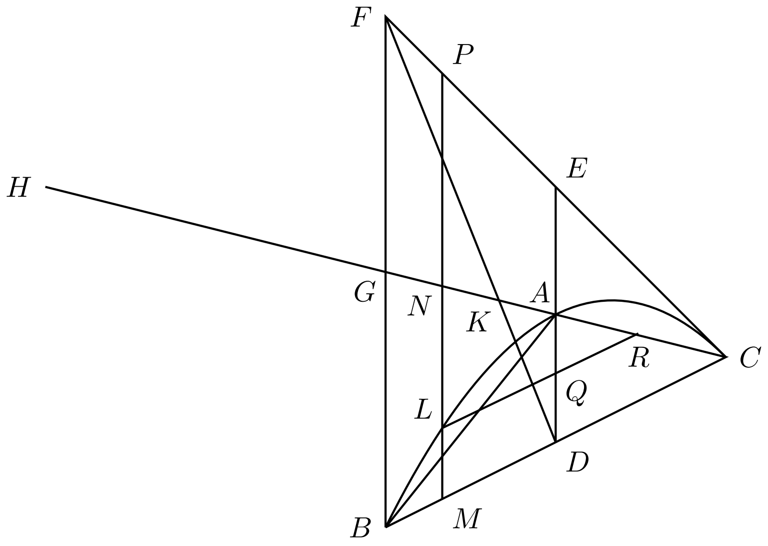


Figure 1.8. Quadrature by balancing

pick a point L at random on the parabola between B and C. Let the straight line drawn through L parallel to BF meet BC at M and CG at N and CF at P. We shall show that LM has to MP the same ratio that GN has to GH: more simply, LM is to MP as GN is to GH, or

$$LM : MP :: GN : GH. \tag{$*$}$$

This is the key point. If the proportion $(*)$ holds, then LM, if its midpoint is placed at H, will just balance MP. Since L was chosen arbitrarily, we conclude that, if all of the parabolic segment were placed at H, then it would balance BCF.

Although he does give this argument in the *Method,* Archimedes does not think it is rigorous. Indeed, in the preface to the *Method,* he writes,

> For some things first became clear to me by mechanics, though they had later to be proved geometrically owing to the fact that investigation by this method does not amount to actual proof; but it is, of course, easier to provide the proof when some knowledge of the things sought has been acquired by this method rather than to seek it with no prior knowledge. [29, p.223]

I want to look at the 'actual proof' of Archimedes later, in § 1.3. Meanwhile, let us establish $(*)$. Perhaps you think of it as an equation of fractions:

$LM/MP = GN/GH$. That is fine; but (∗) simply expresses a relation of *proportionality* among four *magnitudes.* Here are Definitions 3–6 from Book V of Euclid's *Elements.*

> 3. A **ratio** is a sort of relation in respect of size between two magnitudes of the same kind.
>
> 4. Magnitudes are said to **have a ratio** to one another which are capable, when multiplied, of exceeding one another.
>
> 5. Magnitudes are said to **be in the same ratio,** the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order.
>
> 6. Let magnitudes which have the same ratio be called **proportional.**

In view of Definition 6, we may refer to what is written in (∗) as a **proportion.** Since no number of lines are equal to an area, Definition 4 implies that there is no ratio between a line and an area. But we want to use (∗), a proportion of *lines,* to establish the proportion of *areas* that leads to Theorem 2. Archimedes calls this a 'mechanical' argument. The thinking behind Definition 4 may by why Archimedes does not find the 'mechanical' argument to be rigorous.

If a and b *are* magnitudes having a ratio in the sense of Definition 4, and c and d are also magnitudes with a ratio, then by Definition 5, we may say variously

(1) a has to b the same ratio that c has to d,

(2) a is to b as c is to d,

(3) $a : b :: c : d$,

provided that, whenever we take a multiple na of a, and the same multiple nc of c, and a multiple mb of b, and the same multiple md of d, then

$$\begin{aligned} na > mb &\text{ if and only if } nc > nd,\\ na = mb &\text{ if and only if } nc = nd,\\ na < mb &\text{ if and only if } nc < nd. \end{aligned}$$

The multiple na is a *number* of (copies of) a. To be more precise, we should say that na is a **counting number** or a **natural number** of a. We think today that there are other numbers besides the counting numbers 1, 2, 3, and so on; there is, for example, the square root of two, written $\sqrt{2}$. Justification for this way of thinking can be found in Euclid's definition of proportion. For example, $\sqrt{2}$ can be defined as the set of pairs (m, n) of counting numbers such that $m^2 < 2n^2$. These matters will be taken up later. See also Russo [26, § 2.5], who suggests that Weierstrass and his student Dedekind were able to develop the modern theory of real numbers precisely because they had studied Euclid's *Elements.* Meanwhile, let us just note that, in Books V and VI of the *Elements,* Euclid proves the properties of proportionality that we shall need.

We return to the problem of establishing $(*)$. From L draw a straight line parallel to BC, meeting AD at Q and AC at R. Then by property 1 of the parabola given above,

$$LQ^2 : CD^2 :: AQ : AD.$$

Since $LQ = MD$, and triangle ACD is similar to NCM, while ARQ is similar to ACD, we can rewrite the proportion as

$$NA^2 : AC^2 :: AR : AC.$$

Therefore NA is a mean proportional to AC and AR, that is,

$$NA : AC :: AR : NA,$$

so we have the following, where ± resolves to $+$ if L is to the left of A, and otherwise to $-$:

$$\begin{aligned} NA : AC &:: NA \pm AR : AC \pm NA \\ &:: NR : NC \\ &:: NL : NM. \end{aligned}$$

Since $AC = AG$, we obtain

$$\begin{gathered} NA : AG :: NL : NM, \\ AG \mp NA : AG :: NM \mp NL : NM, \\ NG : AG :: LM : NM, \\ NG : 2AG :: LM : 2NM. \end{gathered}$$

Since also $MN = NP$, we get finally

$$\begin{gathered} NG : GC :: LM : MP, \\ NG : GH :: LM : MP, \end{gathered}$$

which is $(*)$, as desired.

1.3. The proof

Archimedes works out the 'mechanical' argument for Theorem 2 also in the *Quadrature of the parabola;* but I prefer now to look at the alternative proof, which Archimedes gives for Proposition 24 of the *Quadrature.*

Start over from Figure 1.1, getting Figure 1.9. Here D is the midpoint of BC as before, and E is the midpoint of AC. From E a straight line is drawn parallel to DA, meeting the parabola at F. Draw straight lines AF and FC. Then triangle ACF has the same altitude as the parabolic segment in which it is inscribed. Similarly we can find G on the parabola between A and B so that the inscribed triangle ABG has the same height as its parabolic segment. We show that triangles ACF and ABG are together one fourth of triangle ABC.

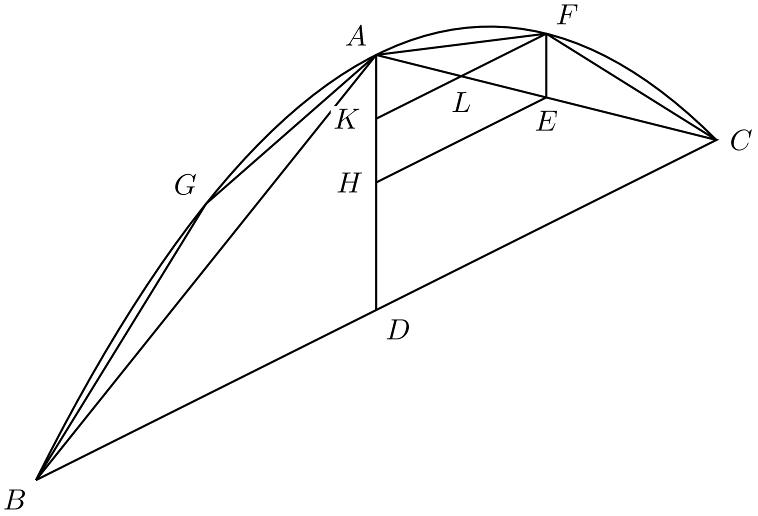


Figure 1.9. Quadrature by inscribing triangles

To this end, from E and F we draw parallels to BC, meeting AD at H and K respectively. Then

$$FK^2 : CD^2 :: AK : AD.$$

But $FK = EH$, and

$$EH : CD :: EA : CA :: 1 : 2.$$

Therefore

$$AK : AD :: EH^2 : CD^2 :: 1 : 4,$$

so AK is one fourth of AD. Consequently K is the midpoint of AH, and so, letting L be the intersection of FK and AC, we have that L is the midpoint of

AE. Hence triangle AKL is equal to triangle EFL. But EFL is one fourth of ACF, and AKL is one thirty-second of ABC. Therefore ACF is one eighth of ABC. Similarly, ABG is one eighth of ABC; so ABG and ACF together are one fourth of ABC.

We have started with the parabolic segment cut off by the chord BC, and we have removed from it the triangle ABC. Then we have removed triangles equal to a fourth of ABC. We can continue, removing triangles equal to a sixteenth of ABC, and so on. Moreover, at each step, we remove *more than half* the remainder of the original parabolic segment.

Therefore, if we continue long enough, we can make the remainder of the parabolic segment less than any pre-assigned area M. This is the conclusion of Proposition X.1 of Euclid's *Elements;* let us note the proof. The pre-assigned area M is assumed to *have a ratio* with the parabolic segment, so that, by Definition 4 quoted above from *Elements* V, some multiple nM of M exceeds the segment. Indeed, Archimedes himself makes this assumption explicit in the preface to the *Quadrature of the parabola;* it is what we may refer to as the **Archimedean axiom:**

> given [two] unequal areas, the exess by which the greater exceeds the less can, by being added to itself, be made to exceed any given finite area. [29, p.231]

If we take away at least half of the parabolic segment, and take M from nM, then in the latter case we are taking not more than half; so the remainder in the former case is still less than the remainder in the latter case, which is $(n-1)M$. If we repeat this process $n-2$ more times, then, in the end, the remainder of the parabolic segment will be less than M.

Suppose we have an area that is a third again as large as triangle ABC. If we remove an area equal to triangle ABC, then what is left is one third of this triangle. If we then remove an area equal to one fourth of triangle ABC, then what is left is one twelfth of that triangle, which is one fourth of the previous remainder. Continuing, if at each step we remove one fourth of what we last removed, then what remains is one fourth of the previous remainder. Therefore, continuing as far as necessary, we can make the remainder as small as we like. But this is the same process as we described in the original parabolic segment.

Now, suppose if possible that the original parabolic segment is *more than* a third again as large as ABC. Let it be greater by M. By the process described, we can inscribe in the parabolic segment a rectilinear figure which differs from the segment by less than M. But then the inscribed figure is more than a third greater than ABC, whereas the process of inscription always gives us a figure *less* then a third greater than ABC, which is absurd. There is a similar contradiction if the parabolic segment is less than a third again as large as ABC. Theorem 2 now follows.

2. The natural numbers

Today many number systems are recognized, and some of them form the chain

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}. \tag{$*$}$$

Here $\mathbb{N}$ is the set of **natural numbers** in the sense of p. 17;[*] $\mathbb{Z}$ comprises the **integers;** $\mathbb{Q}$, the **rational numbers;** $\mathbb{R}$, the **real numbers.** What, if anything, comes after $\mathbb{R}$? It depends on how we think of $\mathbb{R}$. If we think of it as a *field,* then we might think of $\mathbb{R}$ as included in $\mathbb{C}$, the field of **complex numbers.** But what if we think of $\mathbb{R}$ as an *ordered field?* I postpone an answer until Chapter 5. Meanwhile I want to look at how we obtain $(*)$ in the first place.

2.1. The numbers themselves

We can understand $\mathbb{N}$ axiomatically. To do this, we first have to name some features of $\mathbb{N}$.

1. It has an **initial element** called **one,** denoted by
$$1.$$
2. It has an operation of **succession,** denoted by
$$n \mapsto n+1;$$
here $n+1$ is the **successor** of n.

Let us denote succession also by S, so

$$\mathrm{S}(n) = n+1. \tag{$*$}$$

I propose to refer to the ordered triple $(\mathbb{N}, 1, \mathrm{S})$ as an *iterative structure.*[†] In general, by an **iterative structure,** I mean any set that has a distinuished element and a distinguished **singulary operation** (function from the set to itself). For

[*] That is, $\mathbb{N} = \{1, 2, 3, \dots\}$; but some writers consider 0 as a natural number—as indeed *we* shall in 6.1.1 and later.

[†] Stoll [28, §2.1, p. 58] uses the term ‘unary system’.

example, modular arithmetic involves the iterative structures $(\mathbb{Z}_n, 1, \mathrm{S})$.[*] The iterative structure $(\mathbb{N}, 1, \mathrm{S})$ is distinguished among iterative structures for satisfying the following axioms.

1. 1 is not a successor: $1 \neq n + 1$.
2. Succession is injective: if $m + 1 = n + 1$, then $m = n$.
3. **Proof by induction** is admitted, in the sense that a subset A is the whole set, provided we can establish

   (base step) $1 \in A$;

   (inductive step) for all n, if the **inductive hypothesis** $n \in A$ holds, then $n + 1 \in A$.

These axioms were published originally by Dedekind in 1888 [8, II, VI (71), p. 67]; then they were written down by Peano in a special notation in 1889 [24]. They are usually known as the **Peano axioms.** From these axioms, Landau develops the rational, real, and complex numbers rigorously, over the course of a book [19]. I want to do the same here, though more quickly and in a different style. In particular, Landau does not use the next theorem below. The proof is difficult, but the result is very useful. The function h found in the theorem is depicted in Figure 2.1.

Theorem 3 (Recursion). *For every iterative structure (A, b, f), there is a unique function h from $\mathbb{N}$ to A such that*

(1) $h(1) = b$,

(2) $h(n + 1) = f(h(n))$ for all n in $\mathbb{N}$.

$$\begin{tikzcd}
\{1\} \arrow[r, "\subseteq"] \arrow[d] & \mathbb{N} \arrow[r, "\mathrm{S}"] \arrow[d, "h"'] & \mathbb{N} \arrow[d, "h"] \\
\{b\} \arrow[r, "\subseteq"'] & A \arrow[r, "f"'] & A
\end{tikzcd}$$

Figure 2.1. Recursion

Proof. I use the set-theoretic conception whereby a function g is just the set of

[*] One can understand $\mathbb{Z}_n$ either as the set $\{1, \dots, n\}$ comprising the first n natural numbers, or as the set of congruence-classes $k + (n)$ of integers *modulo* n. In the latter case, the element of $\mathbb{Z}_n$ that is called one is really $1 + (n)$; but we can still denote it by 1. Likewise, though the operation of succession on $\mathbb{Z}_n$ is different from the operation on $\mathbb{N}$, we may still denote it by the same symbolism.

ordered pairs[*] (x, y) such that $g(x) = y$; so if (x, y) and (x, z) belong to g, then $y = z$. We now seek h as a particular subset of $\mathbb{N} \times A$.

Let $\mathscr{B}$ be the set whose elements are the subsets C of $\mathbb{N} \times A$ such that, if $(x, y) \in C$, then either

1. $(x, y) = (1, b)$ or else
2. C has an element (u, v) such that $(x, y) = (u + 1, f(v))$.

Let $R = \bigcup \mathscr{B}$; so R is a subset of $\mathbb{N} \times A$. We may say R is a *relation* from $\mathbb{N}$ to A. If $(x, y) \in R$, we may write also

$$x \mathrel{R} y.$$

Since $(1, b) \in \mathscr{B}$, we have $1 \mathrel{R} b$. If $n \mathrel{R} y$, then $(n, y) \in C$ for some C in $\mathscr{B}$, but then $C \cup \{(n + 1, f(y))\} \in \mathscr{B}$ by definition of $\mathscr{B}$, so $(n + 1) \mathrel{R} f(y)$. Therefore R is the desired function h, provided it is a *function* from $\mathbb{N}$ to A. Proving this has two stages.

1. For all n in $\mathbb{N}$, there is y in A such that $n \mathrel{R} y$. Indeed, let D be the set of such n. Then we have just seen that $1 \in D$, and if $n \in D$, then $n + 1 \in D$. By induction, $D = \mathbb{N}$.

2. For all n in $\mathbb{N}$, if $n \mathrel{R} y$ and $n \mathrel{R} z$, then $y = z$. Indeed, let E be the set of such n. Suppose $1 \mathrel{R} y$. Then $(1, y) \in C$ for some C in $\mathscr{B}$. Since 1 is not a successor, we must have $y = b$, by definition of $\mathscr{B}$. Therefore $1 \in E$. Suppose $n \in E$, and $(n + 1) \mathrel{R} y$. Then $(n + 1, y) \in C$ for some C in $\mathscr{B}$. Again since 1 is not a successor, we must have $(n + 1, y) = (m + 1, f(v))$ for some (m, v) in C. Since succession is injective, we must have $m = n$. Since $n \in E$, we know v is *unique* such that $n \mathrel{R} v$. Since $y = f(v)$, therefore y is unique such that $(n + 1) \mathrel{R} y$. Thus $n + 1 \in E$. By induction, $E = \mathbb{N}$.

So R is the desired function h. Finally, h is unique by induction. □

The function h in the statement of the Recursion Theorem (and in Figure 2.1) has three properties:

1. $h \colon \mathbb{N} \to A$ (that is, h is a function from $\mathbb{N}$ to A);
2. $h(1) = b$;
3. $h \circ \mathrm{S} = f \circ h$.

Because it has such properties, h is a **homomorphism** from $(\mathbb{N}, 1, \mathrm{S})$ to (A, b, f). The unique homomorphism guaranteed by the Recursion Theorem is said to be defined **recursively** or defined **by recursion.** Indeed, the word *recursion* has the etymological sense of *running back;* and the value of $h(n + 1)$ is obtained by running back to the value of $h(n)$ and applying f. The iterative structure $(\mathbb{N}, 1, \mathrm{S})$ itself can be said to **admit recursion,** because of the Recursion Theorem.

[*] A good definition of an ordered pair is Kuratowski's [18]: $(x, y) = \{\{x\}, \{x, y\}\}$.

Corollary 4. *For every set A with a distinguished element b, and for every function F from $\mathbb{N} \times A$ to A, there is a unique function H from $\mathbb{N}$ to A such that*

(1) $H(1) = b$,

(2) $H(n+1) = F(n, H(n))$ *for all n in $\mathbb{N}$.*

Proof. Let h be the unique homomorphism from $(\mathbb{N}, 1, \mathrm{S})$ to $(\mathbb{N} \times A, (1, b), f)$, where f is the operation $(n, x) \mapsto (n+1, F(n, x)))$ on $\mathbb{N} \times A$. In particular, $h(n)$ is always an ordered pair. By induction, the first entry of $h(n)$ is always n; so there is a function H from $\mathbb{N}$ to A such that $h(n) = (n, H(n))$. Then H is as desired. By induction, H is unique. □

The proof of the Recursion Theorem used each of the three Peano axioms; induction alone would not be enough. This is a consequence of the next two theorems.

An **isomorphism** of iterative structures is a bijective homomorphism. Note then that the inverse of an isomorphism is an isomorphism. Two iterative structures are **isomorphic** if there is an isomorphism from one to the other.

Theorem 5. *Every iterative structure that admits recursion is isomorphic to* $(\mathbb{N}, 1, \mathrm{S})$.

Proof. Suppose (A, b, f) is an iterative structure admitting recursion. Then there is a homomorphism h from this to $(\mathbb{N}, 1, \mathrm{S})$, and a homomorphism h' from the latter to the former. The composition $h \circ h'$ is a homomorphism from $(\mathbb{N}, 1, \mathrm{S})$ to itself; but so is the identity on $\mathbb{N}$. Such homomorphisms are unique, by definition of admitting recursion; therefore $h \circ h'$ is the identity. For the same reason, $h' \circ h$ is the identity on A. Therefore (A, b, f) and $(\mathbb{N}, 1, \mathrm{S})$ are isomorphic. □

Corollary 6. *The Recursion Theorem is logically equivalent to the Peano axioms.*

Theorem 7. *The Peano axioms are logically independent.*

Proof. $(\mathbb{Z}_n, 1, \mathrm{S})$ satisfies axioms 2 and 3, but not 1; and there are examples satisfying 1 and 3, but not 2; and satisfying 1 and 2, but not 3 (finding such examples is an exercise). □

Therefore no two of the Peano axioms are enough to prove the Recursion Theorem. In particular, the induction axiom by itself is not enough.

2.2. Their structure

2.2.1. Addition

Let $\mathscr{F}$ be the set of all singulary operations on $\mathbb{N}$. Then S is a particular element of $\mathscr{F}$; also, if $g \in \mathscr{F}$, then so is g', where g' is given by

$$g'(n) = \mathrm{S}(g(n)).$$

Thus we have an iterative structure $(\mathscr{F}, \mathrm{S}, ')$. For the moment, let h be the homomorphism from $(\mathbb{N}, 1, \mathrm{S})$ to this structure; but write $h(n)$ as h_n. Then

$$h_1(n) = \mathrm{S}(n), \qquad h_{m+1}(n) = \mathrm{S}(h_m(n)). \tag{$*$}$$

We can now define the binary operation of **addition** on $\mathbb{N}$ by

$$n + m = h_m(n).$$

Then $(*)$ can be written as

$$n + 1 = \mathrm{S}(n), \qquad n + \mathrm{S}(m) = \mathrm{S}(n + m). \tag{$\dagger$}$$

In short, these equations define addition recursively in its second argument. The former equation agrees with the convention established in $(*)$ of § 2.1; and the latter equation can be rewritten as

$$n + (m + 1) = (n + m) + 1.$$

We can now prove all of the standard properties of addition:

Lemma. *For all n and m in $\mathbb{N}$,*

$$1 + n = n + 1, \qquad (m + 1) + n = (m + n) + 1.$$

Proof. Use induction on n. Trivially $1 + n = n + 1$ when $n = 1$, and if it does when $n = k$, then it does when $n = k + 1$, since

$$\begin{aligned} 1 + (k + 1) &= (1 + k) + 1 && \text{[by definition of $+$]} \\ &= (k + 1) + 1 && \text{[by inductive hypothesis].} \end{aligned}$$

Also trivially $(m + 1) + n = (m + n) + 1$ when $n = 1$, and if it does when $n = k$, then it does when $n = k + 1$, since

$$\begin{aligned} (m + 1) + (k + 1) &= ((m + 1) + k) + 1 && \text{[by definition of $+$]} \\ &= ((m + k) + 1) + 1 && \text{[by inductive hypothesis]} \\ &= (m + (k + 1)) + 1 && \text{[by definition of $+$].} \qquad \square \end{aligned}$$

Theorem 8. *Addition on* $\mathbb{N}$ *is*

(1) ***commutative:*** $n + m = m + n$*; and*

(2) ***associative:*** $n + (m + \ell) = (n + m) + \ell$*.*

Proof. Use induction and the lemma. By the lemma, $n + m = m + n$ when $m = 1$, and if it does when $m = k$, then

$$\begin{aligned} n + (k + 1) &= (n + k) + 1 && \text{[by definition of $+$]} \\ &= (k + n) + 1 && \text{[by inductive hypothesis]} \\ &= (k + 1) + n && \text{[by the lemma].} \end{aligned}$$

By definition of addition, $n + (m + \ell) = (n + m) + \ell$ when $\ell = 1$, and if it does when $\ell = k$, then

$$\begin{aligned} n + (m + (k + 1)) &= n + ((m + k) + 1) && \text{[by definition of $+$]} \\ &= (n + (m + k)) + 1 && \text{[by definition of $+$]} \\ &= ((n + m) + k) + 1 && \text{[by inductive hypothesis]} \\ &= (n + m) + (k + 1) && \text{[by definition of $+$].} \end{aligned}$$

□

Like $(\mathbb{N}, 1, \mathrm{S})$, the pair $(\mathbb{N}, +)$ is an example of an **algebraic structure:** a set with one or more (or even no) operations on it. The set itself is the **universe** of the structure. In particular, by the theorem, $(\mathbb{N}, +)$ is a **commutative semigroup.**

2.2.2. Multiplication

We can define **multiplication** on $\mathbb{N}$ recursively in its second argument by

$$n \cdot 1 = n, \qquad n \cdot (m + 1) = n \cdot m + n. \tag{‡}$$

Note that $n \cdot m + n$ means $(n \cdot m) + n$.

Lemma. *For all* n *and* m *in* $\mathbb{N}$*,*

$$1 \cdot n = n, \qquad (m + 1) \cdot n = m \cdot n + n.$$

Proof. By definition, $1 \cdot n = n$ when $n = 1$, and if it does when $n = k$, then

$$\begin{aligned} 1 \cdot (k + 1) &= 1 \cdot k + 1 && \text{[by definition of $\cdot$]} \\ &= k + 1 && \text{[by inductive hypothesis].} \end{aligned}$$

Also by definition, $(m+1)\cdot n = m\cdot n + n$ when $n = 1$, and if when $n = k$, then

$$\begin{aligned}
(m+1)\cdot(k+1) &= (m+1)\cdot k + (m+1) && \text{[by definition of $\cdot$]}\\
&= (m\cdot k + k) + (m+1) && \text{[by inductive hypothesis]}\\
&= (m\cdot k + (k+m)) + 1 && \text{[by associativity of $+$]}\\
&= (m\cdot k + (m+k)) + 1 && \text{[by commutativity of $+$]}\\
&= (m\cdot k + m) + (k+1) && \text{[by associativity of $+$]}\\
&= m\cdot(k+1) + (k+1) && \text{[by definition of $\cdot$].}
\end{aligned}$$
□

Since $1\cdot n = n = n\cdot 1$, the element 1 is a **multiplicative identity** in $\mathbb{N}$. We may write $n\cdot m$ as nm.

Theorem 9. *Multiplication on $\mathbb{N}$ is*

(1) commutative: $nm = mn$;
*(2) **distributive** over addition: $n(m+\ell) = nm + n\ell$; and*
(3) associative: $n(m\ell) = (nm)\ell$.

Proof. By the lemma, $nm = mn$ when $m = 1$, and if when $m = k$, then

$$\begin{aligned}
n(k+1) &= nk + n && \text{[by definition of $\cdot$]}\\
&= kn + n && \text{[by inductive hypothesis]}\\
&= (k+1)n && \text{[by the lemma].}
\end{aligned}$$

By definition, $n(m+\ell) = nm + n\ell$ when $\ell = 1$, and if when $\ell = k$, then

$$\begin{aligned}
n(m+(k+1)) &= n((m+k)+1) && \text{[by definition of $+$]}\\
&= n(m+k) + n && \text{[by definition of $\cdot$]}\\
&= (nm + nk) + n && \text{[by inductive hypothesis]}\\
&= nm + (nk + n) && \text{[by associativity of $+$]}\\
&= nm + n(k+1) && \text{[by definition of $\cdot$].}
\end{aligned}$$

Finally, $n(m\ell) = (nm)\ell$ when $\ell = 1$ by definition of $\cdot$, and if when $\ell = k$, then

$$\begin{aligned}
n(m(k+1)) &= n(mk + m) && \text{[by definition of $\cdot$]}\\
&= n(mk) + nm && \text{[by distributivity]}\\
&= (nm)k + nm && \text{[by inductive hypothesis]}\\
&= (nm)(k+1) && \text{[by definition of $\cdot$].}
\end{aligned}$$
□

Now we know that the structure $(\mathbb{N}, \cdot)$ is a commutative semigroup. Because this is so, and $(\mathbb{N}, +)$ is a commutative semigroup, and multiplication distributes over addition, the structure $(\mathbb{N}, +, \cdot)$ can be called a **commutative semi-ring.** Since also 1 is a multiplicative identity, $(\mathbb{N}, 1, \cdot)$ is a commutative **monoid,** and $(\mathbb{N}, 1, +, \cdot)$ is a **unital** commutative semi-ring.

The algebraic terminology here is awkward. The structure $(\mathbb{N}, 1, +, \cdot)$ is fundamental: just about everything else of interest to us will be obtained from it. Unfortunately it has not got a simpler description (that I know of) than 'unital commutative semi-ring'. But it will be important that we can look at a set like $\mathbb{N}$ from different points of view: as the semigroup $(\mathbb{N}, +)$, as the semigroup $(\mathbb{N}, \cdot)$, and so forth.

All lemmas and theorems of this section have been proved by induction alone. But they have relied on the recursive definitions of addition and multiplication, and as we have shown with Corollary 6 and Theorem 7, the possibility of recursive definitions in general requires more than induction. Nonetheless, in his own development of arithmetic, Landau [19] proves *using induction alone* that addition and multiplication exist as given by the recursive definitions (†) and (‡) above. For the record, let us do the same.

Theorem 10. *On any iterative structure $(A, 1, \mathrm{S})$ that admits proof by induction, there are unique binary operations satisfying (†) and (‡); then $(A, 1, +, \cdot)$ is a unital commutative semi-ring.*

Proof. Let A' be the set of n in A for which there is an operation $x \mapsto n + x$ satisfying (†). Then $1 \in A'$, since if we define

$$1 + x = \mathrm{S}(x),$$

then indeed

$$1 + 1 = \mathrm{S}(1), \qquad\qquad 1 + \mathrm{S}(m) = \mathrm{S}(\mathrm{S}(m)) = \mathrm{S}(1 + m).$$

Suppose $k \in A'$. Then $\mathrm{S}(k) \in A'$, since if we define

$$\mathrm{S}(k) + x = \mathrm{S}(k + x),$$

then indeed

$$\mathrm{S}(k) + 1 = \mathrm{S}(k + 1) = \mathrm{S}(\mathrm{S}(k)),$$
$$\mathrm{S}(k) + \mathrm{S}(m) = \mathrm{S}(k + \mathrm{S}(m)) = \mathrm{S}(\mathrm{S}(k + m)) = \mathrm{S}(\mathrm{S}(k) + m).$$

By induction, $A' = A$. Also, for each n in A, by induction there is *at most* one operation $x \mapsto n+x$ satisfying (†). So addition on A exists as desired. Moreover, the proof of Theorem 8 goes through, so that $(A,+)$ is a commutative semigroup.

Now let A'' comprise those n in A for which there is an operation $x \mapsto n \cdot x$ satisfying (‡). Then $1 \in A''$, since if we define

$$1 \cdot x = x,$$

then indeed

$$1 \cdot 1 = 1, \qquad\qquad 1 \cdot (m+1) = m+1 = 1 \cdot m + 1.$$

Suppose $k \in A''$. Then $k+1 \in A''$, since if we define

$$(k+1) \cdot x = k \cdot x + x,$$

then

$$\begin{aligned}(k+1) \cdot 1 &= k \cdot 1 + 1 = k+1,\\ (k+1) \cdot (m+1) &= k \cdot (m+1) + (m+1)\\ &= (k \cdot m + k) + (m+1)\\ &= (k \cdot m + m) + (k+1)\\ &= (k+1) \cdot m + (k+1).\end{aligned}$$

By induction, $A'' = A$. Also by induction on the second argument, multiplication is unique. Then the proof of Theorem 9 goes through, and $(A,1,+,\cdot)$ is unital commutative semi-ring. □

In particular, $(\mathbb{Z}_n,1,+,\cdot)$ is a unital commutative semi-ring. Moreover, one can prove now as an exercise that, on $\mathbb{N}$ and on $\mathbb{Z}_n$, addition is **cancellative,** that is,

$$k = \ell \iff k+m = \ell+m;$$

but the proof requires more than induction. Multiplication also is cancellative on $\mathbb{N}$, but proving this takes more work (and it fails on $\mathbb{Z}_n$). We shall obtain both results about $\mathbb{N}$ as a corollary of Theorem 15 below.

2.2.3. Exponentiation

Exponentiation on $\mathbb{N}$ as a binary operation $(x,y) \mapsto x^y$ given by

$$n^1 = n, \qquad\qquad n^{m+1} = n^m \cdot n. \tag{§}$$

The existence of such an operation requires more than induction. Indeed, exponentiation cannot be defined in $(\mathbb{Z}_n, 1, \mathrm{S})$ by (§) unless n is 1, 2, 6, 42, or 1806: this results from an exercise in number theory worked out by Dyer-Bennet [9]. Note that exponentiation on $\mathbb{N}$ is not commutative either. We shall not make any particular use of exponentiation as a binary operation. We shall however obtain in 4.3.2, for each *real* number b that is greater than 1, the isomorphism $x \mapsto b^x$ from the additive *group* of real numbers to the multiplicative group of positive real numbers.

2.2.4. Ordering

Something else that distinguishes $\mathbb{N}$ from the sets $\mathbb{Z}_n$ is the possibility of *ordering* the former in a way compatible with addition. The usual ordering $<$ of $\mathbb{N}$ can be defined recursively as follows. First note that $m \leqslant n$ means simply $m < n$ or $m = n$. Then the definition of $<$ is:

(1) $m \nless 1$;
(2) $m < n + 1$ if and only if $m \leqslant n$.

In particular, $n < n + 1$. Really, it is the function $n \mapsto \{x \in \mathbb{N} \colon x < n\}$ that is defined by recursion:

(1) $\{x \in \mathbb{N} \colon x < 1\} = \varnothing$;
(2) $\{x \in \mathbb{N} \colon x < n + 1\} = \{x \in \mathbb{N} \colon x < n\} \cup \{n\}$.

We now have $<$ as a binary relation on $\mathbb{N}$. We prove first that it is a **partial ordering,** namely,

(1) it is **irreflexive:** $n \nless n$;
(2) it is **transitive:** if $k < m$ and $m < n$, then $k < n$.

Theorem 11. *The relation $<$ on $\mathbb{N}$ is transitive.*

Proof. It follows by induction on n that if $k < m$ and $m < n$, then $k < n$. □

Lemma. $m \neq m + 1$.

Proof. The claim is true when $m = 1$, since 1 is not a successor. Suppose the claim is true when $m = k$; that is, suppose $k \neq k + 1$. Then $k + 1 \neq (k + 1) + 1$, by injectivity of succession, so the claim is true when $m = k + 1$. By induction, the claim is true for all m. □

Theorem 12. *The relation $<$ on $\mathbb{N}$ is irreflexive.*

Proof. The claim is true when $m = 1$, since $m \nless 1$ by definition. Suppose the claim *fails* when $m = k + 1$. This means $k + 1 < k + 1$. Therefore $k + 1 \leqslant k$ by definition. By the lemma, $k+1 \neq k$, so $k+1<k$. But $k<k+1$ by definition. So $k<k+1$ and $k+1<k$; hence $k<k$ by Theorem 11, that is, the claim fails when $m=k$. Contrapositively, if $k \nless k$, then $k+1 \nless k+1$. By induction, the claim holds for all m. □

So we know the relation $<$ is a *partial* ordering of $\mathbb{N}$. The pair $(\mathbb{N},<)$ can therefore be called a **partial order.** It is an example of a new kind of structure: not an algebraic structure, but a **relational structure.**

Every partial ordering is **antisymmetric,** that is, if $x<y$, then $y \nless x$. Indeed, if $x<y$ and $y<x'$, then $x<x'$ by transitivity, so $x \neq x'$ by irreflexivity. We now want to show $<$ is simply an **ordering**—or more precisely a **linear ordering**[*]—of $\mathbb{N}$, that is, $k \leqslant m$ or $m \leqslant k$ (for all k and m in $\mathbb{N}$). Note that, once this linearity is known, the next two lemmas will be easy consequences of the definition of $<$.

Lemma. *$1 \leqslant m$.*

Proof. Induction. □

Lemma. *If $k<m$, then $k+1 \leqslant m$.*

Proof. The claim is vacuously true when $m=1$, since $k \nless 1$. Suppose the claim is true when $m=n$. Say $k<n+1$. Then $k \leqslant n$. If $k=n$, then $k+1=n+1$. If $k<n$, then $k+1 \leqslant n$ by inductive hypothesis, so $k+1<n+1$ by definition. In either case, $k+1 \leqslant n+1$. Thus the claim holds when $m=n+1$. By induction, the claim holds for all m. □

Theorem 13. *The partial ordering $<$ of $\mathbb{N}$ is linear.*

Proof. It follows by induction and the last two lemmas that either $\ell \leqslant m$ or $m \leqslant \ell$. Indeed, the claim holds by the next-to-last lemma when $\ell=1$. Suppose it holds when $\ell=k$, but $k+1 \nleqslant m$. Then $k \nless m$ by the last lemma, so either $k=m$ or $k \nleqslant m$, and then by inductive hypothesis $m \leqslant k$, so $m \leqslant k+1$. □

Therefore $(\mathbb{N},<)$ is simply an **order.** More is true:

Theorem 14. *$\mathbb{N}$ is **well ordered** by $<$: every nonempty set of natural numbers has a least element.*

[*] Or *total ordering;* but *linear* is more suggestive of arranging elements in one line. In any case, since all orderings in these notes will be total or linear, I shall generally call them just *orderings.* A set with an ordering is an **order.**

Proof. Suppose A is a set of natural numbers with no least element; we show A is empty. Let

$$B = \{x \in \mathbb{N} \colon \forall y\ (y < x \Rightarrow y \notin A)\}.$$

Then $1 \in B$ since nothing is less than 1. Suppose $m \in B$. Then $m + 1 \in B$, since otherwise m would be the least element of A. By induction, $B = \mathbb{N}$, so $A = \varnothing$. □

2.2.5. Interaction

Theorem 15. *For all m and n in $\mathbb{N}$, we have $m < n$ if and only if the equation*

$$m + x = n \tag{¶}$$

is soluble in $\mathbb{N}$.

Proof. By induction on k, if $m + k = n$, then $m < n$.

By induction on n, we prove conversely that if $m < n$, then (¶) is soluble. This is vacuously true when $n = 1$. Suppose the equation $m+x = r$ is soluble whenever $m < r$, but now $m < r+1$. Then $m \leqslant r$. If $m = r$, then $m+1 = r+1$. If $m < r$, then the equation $m + x = r$ has a solution k, and therefore $m + (k+1) = r + 1$. Thus the equation $m + x = r + 1$ is soluble whenever $m < r + 1$. □

The theorem means $<$ can be defined in terms of addition by an existential formula:

$$m < n \iff \exists x\ m + x = n.$$

Corollary 16. *On* $\mathbb{N}$,

$$\begin{aligned} k < n &\iff k + m < n + m, && (\|) \\ k = n &\iff k + m = n + m. && (**) \end{aligned}$$

Proof. The forward direction of $(**)$ is immediate. For the forward direction of $(\|)$, if $k < n$, then $k + \ell = n$ for some ℓ, so

$$(k + \ell) + m = n + m.$$

By commutativity and associativity,

$$(k + m) + \ell = n + m,$$

so $k + m < n + m$. The reverse directions of $(\|)$ and $(**)$ follow by the linearity of the ordering. □

Note that, also by the linearity of the ordering, (‖) implies (∗∗). It follows from (∗∗) that the equation (¶) has at most one solution: this can be denoted by

$$n - m \tag{††}$$

and is the **difference** of n and m. So, for now, $n - m$ exists if and only if $m < n$.

Corollary 17. *On* $\mathbb{N}$*,*

$$k < n \iff km < nm, \tag{‡‡}$$

$$k = n \iff km = nm. \tag{§§}$$

Proof. The forward direction of (§§) is immediate. For the forward direction of (‡‡), if $k < n$, then $k + \ell = n$ for some ℓ, so

$$(k + \ell)m = nm.$$

By distributivity,

$$km + \ell m = nm,$$

so $km < nm$. The reverse directions of (‡‡) and (§§) follow by the linearity of the ordering. □

We now have several ways of thinking of $\mathbb{N}$, in addition to those identified earlier in this section. For example, we can consider the triple $(\mathbb{N}, +, <)$. This is neither a (purely) algebraic structure, nor a (purely) relational structure; it is an example of a **structure,** simply. Because

(1) $(\mathbb{N}, +)$ is a commutative semigroup,
(2) $(\mathbb{N}, <)$ is an order, and
(3) Corollary 16 holds,

we may say that the structure $(\mathbb{N}, +, <)$ is an **ordered commutative semigroup.** Therefore, because $(\mathbb{N}, \cdot)$ is a commutative semigroup, and Corollary 17 holds, $(\mathbb{N}, \cdot, <)$ too is an ordered commutative semigroup; we may say further that $(\mathbb{N}, 1, \cdot, <)$ is an **ordered commutative monoid.**

There is stronger terminology. Because

(1) $(\mathbb{N}, +)$ is a commutative semigroup,
(2) $(\mathbb{N}, <)$ is an order, and
(3) Theorem 15 holds,

the ordering $<$ is a **natural ordering**[*] of the commutative semigroup $(\mathbb{N}, +)$, and $(\mathbb{N}, +, <)$ is a **naturally ordered commutative semigroup.** Then $<$ is

[*] This terminology is inspired the survey article [5, p. 308] of Clifford, where naturally ordered semigroups are defined.

not a natural ordering of $(\mathbb{N}, \cdot)$, since for example $2 < 3$, but $2x = 3$ is insoluble in $\mathbb{N}$. However, we may say that $(\mathbb{N}, 1, +, \cdot, <)$ is a **naturally ordered unital commutative semi-ring,** because

(1) $(\mathbb{N}, 1, +, \cdot)$ is a unital commutative semi-ring,
(2) $(\mathbb{N}, <)$ is an order, and
(3) Theorem 15 holds.

As such, $(\mathbb{N}, 1, +, \cdot, <)$ will turn out to embed in an *ordered commutative ring* by Porism 33. We do not define the general notion of an ordered unital commutative semi-ring.

3. Construction of the rational numbers

3.1. The positive rational numbers

The integers can be constructed from the natural numbers, and the rational numbers can be constructed from the integers. However, the *positive* rational numbers can also be constructed directly from the natural numbers, and indeed we are taught some aspects of this construction from an early age. If a and b are natural numbers, then there is a **fraction** denoted by

$$\frac{a}{b}$$

or a/b. Then there are definitions for adding and multiplying fractions:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}. \tag{$*$}$$

Also, an ordering of fractions can be defined by

$$\frac{a}{b} < \frac{c}{d} \iff ad < cb. \tag{$\dagger$}$$

We are taught to *reduce* fractions also: By $(*)$ we compute $1/3 + 1/6 = 9/18$, which reduces to $1/2$. In particular, $9/18$ and $1/2$ are *equal* fractions. Equality of fractions may be given by

$$\frac{a}{b} = \frac{c}{d} \iff ad = cb. \tag{$\ddagger$}$$

The fraction a/b need not uniquely determine the pair (a, b). This means that the validity of $(*)$ and $(\dagger)$ as definitions must be checked. Indeed, the validity of $(\ddagger)$ as a definition of equality must be checked. There are two ways to do this.

3.1.1. Equality by definition

In $(\ddagger)$, *some* binary relation on fractions is defined; for the moment, we can give it the arbitrary name R. So

$$\frac{a}{b} \mathrel{R} \frac{c}{d} \iff ad = cb.$$

We can replace R by any name we want; but if we want to refer to R as equality, it will be useful to know that, like the relation of identity, R is

(1) **reflexive** ($x \mathrel{R} x$),
(2) **symmetric** (if $x \mathrel{R} y$, then $y \mathrel{R} x$) and
(3) transitive (as on p. 31: if $x \mathrel{R} y$ and $y \mathrel{R} z$, then $x \mathrel{R} z$).
The relation R is indeed so: this is Theorem 18 below.

3.1.2. Equality by theorem

The foregoing approach to equality of fractions avoids the question of what a fraction *is.* That is fine. Indeed, we have already avoided the question of what a natural number is; we have said only that a natural number is a member of a set $\mathbb{N}$ with the properties expressed in the Peano axioms.

However, we shall give a set-theoretic definition of $\mathbb{N}$ in § 6.1. Meanwhile, assuming we do have $\mathbb{N}$, we can form the set $\mathbb{N} \times \mathbb{N}$ of ordered pairs of elements of $\mathbb{N}$; then we can define the binary relation $\sim$ on $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \sim (c, d) \iff ad = cb. \tag{§}$$

Now we can define

$$\frac{a}{b} = \{(x, y) \in \mathbb{N} \times \mathbb{N} \colon (a, b) \sim (x, y)\}. \tag{¶}$$

This definition is useful because of:

Theorem 18. *The relation $\sim$ on $\mathbb{N} \times \mathbb{N}$ is reflexive, symmetric, and transitive.*

Proof. Reflexivity and symmetry of $\sim$ follow immediately from the corresponding properties of equality; but transitivity needs more. Suppose $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then $ad = cb$ and $cf = ed$, so

$$(ad)f = (cb)f = c(bf) = c(fb) = (cf)b = (ed)b$$

by commutativity and associativity of multiplication. By these properties and also cancellation, we can go on to conclude

$$af = eb,$$

hence $(a, b) \sim (e, f)$. □

This means $\sim$ is an **equivalence relation,** and a/b is the **equivalence class** of (a, b) with respect to this relation. The set of all such classes can be denoted by

$$(\mathbb{N} \times \mathbb{N})/{\sim},$$

but we shall usually write simply

$$\mathbb{Q}^+.$$

This is the set of **positive rational numbers.** We can now characterize equality as a relation on $\mathbb{Q}^+$:

Corollary 19. *$a/b = x/y$ if and only if $(a, b) \sim (x, y)$.*

3.1.3. Structure

We are free to define operations $\oplus$ and $\otimes$ on $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \oplus (c, d) = (ad + cb, bd), \qquad (a, b) \otimes (c, d) = (ac, bd).$$

We are free to define a relation $\olessthan$ on $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \olessthan (c, d) \iff ad < cb.$$

What makes these useful is the following:

Theorem 20. *If $a/b = a'/b'$ and $c/d = c'/d'$, then*

$$\begin{aligned}(a, b) \oplus (c, d) &\sim (a', b') \oplus (c', d'),\\ (a, b) \otimes (c, d) &\sim (a', b') \otimes (c', d'),\\ (a, b) \olessthan (c, d) &\iff (a', b') \olessthan (c', d').\end{aligned} \tag*{$(\|)$}$$

Proof. We assume $ab' = a'b$ and $cd' = c'd$. To prove $(\|)$ we have

$$\begin{aligned}(a, b) \olessthan (c, d) &\iff ad < cb\\ &\iff adb'c' < cbb'c'\\ &\iff ab'c'd < cbb'c'\\ &\iff a'bcd' < cbb'c'\\ &\iff a'd' < c'b'\\ &\iff (a', b') \olessthan (c', d').\end{aligned}$$

□

Corollary 21. *On $\mathbb{Q}^+$, the equations $(*)$ define two binary operations, while the equivalence $(\dagger)$ defines a binary relation.*

In passing from $\mathbb{N}$ to $\mathbb{Q}^+$, we lose only the being well-ordered. Denoting $1/1$ in $\mathbb{Q}^+$ by 1, we have

Theorem 22. $(\mathbb{Q}^+, 1, +, \cdot, <)$ *is a naturally ordered unital commutative semiring.*

Proof. Easily from the definitions, $(\mathbb{Q}^+, 1, +, \cdot, <)$ is a unital commutative semiring. Also, we have $bda < bda + cb^2 = (ad + cb)b$, and therefore

$$\frac{a}{b} < \frac{ad + cb}{bd} = \frac{a}{b} + \frac{c}{d}.$$

Conversely, if $a/b < c/d$, then $ad < cb$, so $cb - ad \in \mathbb{N}$ and

$$\frac{a}{b} + \frac{cb - ad}{db} = \frac{adb + cb^2 - adb}{db^2} = \frac{c}{d}.$$

Thus $<$ naturally orders $(\mathbb{Q}^+, 1, +, \cdot, <)$. □

The whole point of defining $\mathbb{Q}^+$ is the following:

Theorem 23. *There is a well-defined operation $x \mapsto x^{-1}$ on $\mathbb{Q}^+$ given by*

$$\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}.$$

Then

$$x \cdot x^{-1} = 1.$$

Therefore, since $(\mathbb{Q}^+, 1, \cdot)$ is a commutative monoid, $(\mathbb{Q}^+, 1, {}^{-1}, \cdot)$ is an **abelian group;** since also $(\mathbb{Q}^+, \cdot, <)$ is an ordered commutative monoid, $(\mathbb{Q}^+, 1, {}^{-1}, \cdot, <)$ is an **ordered abelian group.** If r and s are in $\mathbb{Q}^+$, then the equation $r = s \cdot X$ has the unique solution $s^{-1}r$, which is written also as a fraction,

$$\frac{r}{s}.$$

If $a, b, c, d \in \mathbb{N}$, then

$$\frac{a/b}{c/d} = \frac{ad}{bc},$$

and in particular

$$\frac{a/1}{c/1} = \frac{a}{c}. \qquad (**)$$

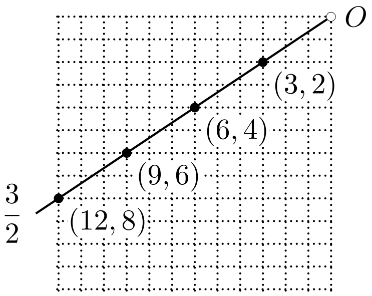


Figure 3.1. Fractions as straight lines

3.1.4. Numbers and fractions

By our construction, a natural number is not literally a positive rational number; a positive rational is a class of ordered pairs of natural numbers. One way to understand this is shown in Figure 3.1, where ordered pairs of natural numbers are depicted as points in a grid; then a fraction is the class of ordered pairs lying on a particular straight line through the point O.

A fraction may not literally be a natural number; but there are fractions that *behave* like natural numbers:

Theorem 24. *The function $x \mapsto x/1$ is an injection from $\mathbb{N}$ into $\mathbb{Q}^+$; moreover,*

$$\frac{x+y}{1} = \frac{x}{1} + \frac{y}{1}, \qquad \frac{x \cdot y}{1} = \frac{x}{1} \cdot \frac{y}{1}, \qquad x < y \iff \frac{x}{1} < \frac{y}{1}.$$

Proof. Immediate from the definitions. □

In a word, $x \mapsto x/1$ is an **embedding** of $(\mathbb{N}, 1, +, \cdot, <)$ in $(\mathbb{Q}^+, 1/1, +, \cdot, <)$. We may therefore forget about the distinction between natural numbers and positive rational numbers: we may *identify* a natural number n with its image $n/1$ in $\mathbb{Q}^+$. By $(**)$, there will be no ambiguity in writing fractions: a fraction of natural numbers as such will be the same as their fraction as positive rational numbers.

Using the idea in Figure 3.1, we can arrange the positive rational numbers along a semicircle, according to their ordering, as in Figure 3.2 (a). It is more usual to arrange the positive rational numbers along a straight line, as in Figure 3.2 (b); the point of using a semicircle is that here, if $k < m$, then m/k lies directly above k/m. Indeed, in Figure 3.3, since $BDCO$ is a semicircle, the angles AOB, OCB, and ODB are equal; if also AOB and COD are equal, then COD and ODB are equal, so the straight lines BD and OC are parallel.

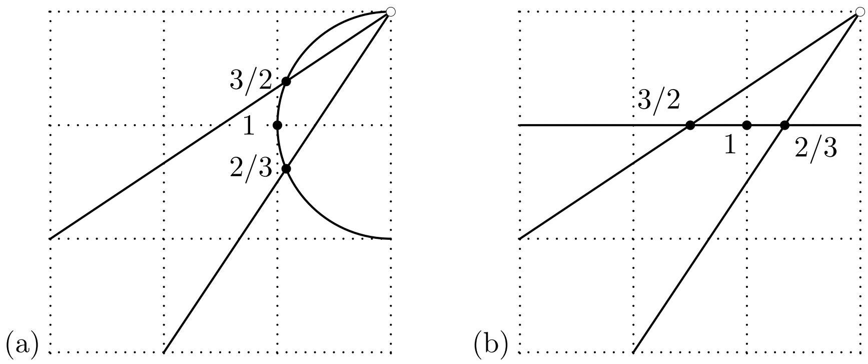


Figure 3.2. Positive rationals along a semicircle and a straight line

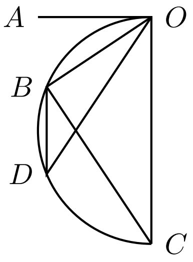


Figure 3.3. Fractions are below their reciprocals

3.2. The integers

We need only have $(\mathbb{N}, 1, \cdot, <)$ as an ordered commutative monoid in order to carry out the following activities in the last section:*

(1) to define $\sim$ on $\mathbb{N} \times \mathbb{N}$, so as to define $\mathbb{Q}^+$ as $(\mathbb{N} \times \mathbb{N})/{\sim}$;
(2) to define multiplication and an ordering on $\mathbb{Q}^+$;
(3) to embed $(\mathbb{N}, \cdot, <)$ in $(\mathbb{Q}^+, \cdot, <)$;
(4) to obtain $(\mathbb{Q}^+, 1, {}^{-1}, \cdot, <)$ as an ordered abelian group.

But $(\mathbb{N}, +, <)$ is also an ordered commutative semigroup, and it is 'almost' a monoid. With slight modifications to accommodate the lack of an additive inverse, we can carry out the listed activities using addition instead of multiplication, thus obtaining the integers.

*Clifford [5, p. 309] suggests that this observation is 'essentially well-known', though it did not appear in general form in the literature until the works he cites as Tamari 1949, Alimov 1950, and Nakada 1951.

In analogy with (§), let us define $\approx$ on $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \approx (c, d) \iff a + d = b + c. \tag{$*$}$$

Then we have a direct analogue of Theorem 18:

Theorem 25. *The relation $\approx$ on $\mathbb{N} \times \mathbb{N}$ is an equivalence-relation.*

In analogy with (¶), we can define

$$n - m = \{(x, y) \in \mathbb{N} \times \mathbb{N} \colon (n, m) \approx (x, y)\}. \tag{$\dagger$}$$

For the moment, this definition disagrees with the one given in § 2.2, at (††); ultimately, the two definitions will be in harmony. An equivalence-class as in (†) is just an **integer;** the set of all integers is

$$\mathbb{Z}.$$

As we have multiplication on $\mathbb{Q}^+$ given in the last section at ($*$), so we have:

Theorem 26. *On $\mathbb{Z}$, there are a well-defined operation of addition and a well-defined relation $<$ given by*

$$(a - b) + (c - d) = (a + c) - (b + d), \quad (a - b) < (c - d) \iff a + d < c + b.$$

The integer $1 - 1$ is **zero,** denoted by

$$0.$$

Then

$$0 < n - m \iff m < n. \tag{$\ddagger$}$$

In partial analogy with Theorem 22, we have

Theorem 27. *$(\mathbb{Z}, 0, +, <)$ is an ordered commutative monoid.*

In analogy with Theorem 23, we have

Theorem 28. *There is a well-defined operation $x \mapsto -x$ on $\mathbb{Z}$ given by*

$$-(k - n) = n - k,$$

and $(\mathbb{Z}, 0, -, +, <)$ is an ordered abelian group.

If a and b are in $\mathbb{Z}$, then the equation $a = b + X$ has the unique solution $-b+a$, which is also denoted by

$$a - b.$$

If $k, \ell, m, n \in \mathbb{N}$, then

$$(k - \ell) - (m - n) = (n + k) + (m + \ell).$$

The elements a of $\mathbb{Z}$ such that $0 < a$ are **positive;** $a < 0$, **negative.** So b is positive if and only if $-b$ is negative; also $n - m$ is positive if and only if $m < n$, by (‡). Similarly, when considering the ordered abelian group $(\mathbb{Q}^+, 1, {}^{-1}, \cdot)$, we may refer to those elements a of $\mathbb{Q}^+$ such that $1 < a$ as positive, or more precisely **multiplicatively positive.**

The embedding of $\mathbb{N}$ in $\mathbb{Q}^+$ has no immediate analogue for $\mathbb{Z}$, since 1 is a multiplicative identity in $\mathbb{N}$, but there is no additive identity. However, if k is an arbitrary element of $\mathbb{N}$, then the same embedding of $\mathbb{N}$ in $\mathbb{Q}^+$ can be defined as $n \mapsto nk/k$. In partial analogy with Theorem 24, we have

Theorem 29. *The function* $x \mapsto (x + 1) - 1$ *embeds* $(\mathbb{N}, +, <)$ *in* $(\mathbb{Z}, +, <)$.

We may identify a natural number n with its image $(n + 1) - 1$ in $\mathbb{Z}$. What makes $(\mathbb{Z}, +, <)$ different from $(\mathbb{Q}^+, \cdot, <)$ is the following.

Theorem 30. *The positive elements of* $(\mathbb{Z}, 0, -, +, <)$ *are just the elements of* $\mathbb{N}$.

Proof. As noted, every positive element of $\mathbb{Z}$ is $n - k$ for some n and k in $\mathbb{N}$ such that $k < n$. Because $<$ is a natural ordering of $\mathbb{N}$, we have $k + m = n$ for some m in $\mathbb{N}$, so $n - k = m$. Conversely, if $m \in \mathbb{N}$, then $m = (m + 1) - 1$, and $1 < m + 1$, so m is positive. □

In short, the ordered commutative semigroup $(\mathbb{N}, +, <)$ is the **positive part** of the ordered abelian group $(\mathbb{Z}, 0, -, +, <)$. The elements of $\mathbb{Z}$ are usually depicted on a straight line extending infinitely in both directions. Alternatively, we can arrange them in a circle, as in Figure 3.4, where, if $0 < n$, then $-n$ is directly to its right. The left half of the circle is the semicircle in Figure 3.2 (a).

Finally, we can define **multiplication** on $\mathbb{Z}$ as in school, by

$$a \cdot 0 = 0 \cdot a = 0, \qquad -m \cdot -n = m \cdot n, \qquad -m \cdot n = m \cdot -n = -(m \cdot n), \qquad (\S)$$

where $a \in \mathbb{Z}$ and m and n are in $\mathbb{N}$.

Theorem 31. $(\mathbb{Z}, 1, +, \cdot)$ *is a unital commutative semi-ring.*

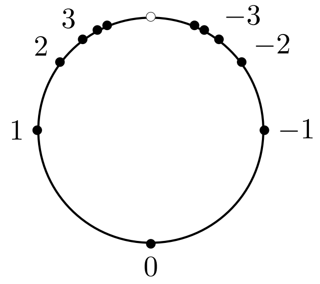


Figure 3.4. Integers on a circle

Proof. By Theorem 27, it is enough to show that multiplication on $\mathbb{Z}$ has the identity 1, is commutative and associative, and distributes over addition. Commutativity on $\mathbb{Z}$ with identity 1 follows immediately from commutativity on $\mathbb{N}$ with identity 1, along with the definitions (§). Associativity follows from considering the several cases, such as

$$(x \cdot -y) \cdot -z = -(x \cdot y) \cdot -z = (x \cdot y) \cdot z = x \cdot (y \cdot z) = x \cdot (-y \cdot -z).$$

For distributivity, we have for example, if $-y + z = w > 0$, then $z = w + y$, so $x \cdot z = x \cdot w + x \cdot y$, and therefore

$$x \cdot (-y + z) = -(x \cdot y) + x \cdot z = x \cdot -y + x \cdot z. \qquad \square$$

Since also $(\mathbb{Z}, 0, -, +)$ is an abelian group, $(\mathbb{Z}, 0, 1, -, +, \cdot)$ is a **commutative ring.** Since also $\mathbb{Z}$ is ordered so that the set of positive elements is closed under addition and multiplication, $(\mathbb{Z}, 0, 1, -, +, \cdot, <)$ is an **ordered commutative ring.** It follows then that

$$x \cdot y = 0 \Rightarrow x = 0 \lor y = 0,$$

so $(\mathbb{Z}, 0, 1, -, +, \cdot)$ is an **integral domain.**

3.3. The rational numbers

By mimicking the contructions of $\mathbb{Q}^+$ and $\mathbb{Z}$, we have two ways to define the *rational numbers* (which can be positive, negative, or 0):

3.3.1. The field of fractions of differences

Because the integers compose an integral domain, they embed in a **quotient field** or **fraction field,** which is constructed almost as $\mathbb{Q}^+$ is constructed from

$\mathbb{N}$ in § 3.1. The difference is that now the relation $\sim$ defined by (§) in that section must be understood as a relation on $\mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\})$. Still, $\sim$ is an equivalence-relation, and the equivalence-class of (m, n) is denoted by m/n. We define

$$\mathbb{Q} = (\mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\}))/{\sim};$$

the elements of this are the **rational numbers.** We obtain addition and multiplication as before, only now defined on $\mathbb{Q}$; and we obtain multiplicative inversion as before, now defined on $\mathbb{Q} \smallsetminus \{0\}$. There is additive inversion on $\mathbb{Q}$ also, namely $x/y \mapsto -x/y$. Then $(\mathbb{Q}, 0, 1, -, +, \cdot, <)$ is an **ordered field:** that is, it is an ordered commutative ring, and $(\mathbb{Q} \smallsetminus \{0\}, 1, {}^{-1}, \cdot)$ is an abelian group.

3.3.2. The field of differences of fractions

There is an alternative approach to defining $\mathbb{Q}$ that is better for our purposes and is suggested by Figures 3.2 (a) and 3.4. In obtaining the ordered abelian group $(\mathbb{Z}, 0, -, +, <)$ from $(\mathbb{N}, +, <)$, we need only that $(\mathbb{N}, +, <)$ is an ordered semigroup. Thus, to obtain $(\mathbb{Q}^+, 1, {}^{-1}, \cdot, <)$ as an ordered abelian group, we need only that $(\mathbb{N}, \cdot, <)$ is an ordered semigroup. The difference between the two ordered groups is that $(\mathbb{N}, +, <)$ is the positive part of $(\mathbb{Z}, 0, -, +, <)$ by Theorem 30, but $(\mathbb{N}, \cdot, <)$ is not the positive part of $(\mathbb{Q}^+, 1, {}^{-1}, \cdot, <)$. We have Theorem 30 simply because $(\mathbb{N}, +, <)$ is *naturally* ordered. In short, we have:[*]

Porism 32. *Every naturally ordered commutative semigroup consists of the positive elements of an ordered abelian group.*

Once we have $(\mathbb{Z}, 0, +, <)$ as an ordered abelian group whose positive part is $(\mathbb{N}, +, <)$, then in order to define multiplication and so obtain the ordered commutative ring $(\mathbb{Z}, 0, 1, -, +, \cdot, <)$, all we need is that $(\mathbb{N}, 1, +, \cdot)$ is a unital commutative semi-ring. That is, we have

Porism 33. *Every naturally ordered unital commutative semi-ring consists of the positive elements of an ordered commutative ring.*

Since $(\mathbb{Q}^+, 1, +, \cdot, <)$ is a naturally ordered unital commutative semi-ring by Theorem 22, it determines $(\mathbb{Q}, 0, 1, -, +, \cdot, <)$ as an ordered commutative ring. We now consider $\mathbb{Q}^+$ is a subset of $\mathbb{Q}$. Since $(\mathbb{Q}^+, 1, {}^{-1}, \cdot)$ is a group, the operation $x \mapsto x^{-1}$ can be extended to $\mathbb{Q} \smallsetminus \{0\}$ by defining

$$x^{-1} = -(-x)^{-1}$$

when $x < 0$.

[*] *Porism* (πόρισμα): I use the term to mean a corollary of an earlier *proof.*

Theorem 34. $(\mathbb{Q}, 0, 1, +, \cdot, <)$ *is an ordered field.*

Proof. It is an ordered commutative ring, and $(\mathbb{Q} \smallsetminus \{0\}, 1, {}^{-1}, \cdot)$ is an abelian group. □

3.3.3. Ordered fields

On any ordered field, there is an operation $x \mapsto |x|$, where

$$|a| = \begin{cases} a, & \text{if } a \geqslant 0; \\ -a, & \text{if } a < 0. \end{cases}$$

Here $|a|$ is the **absolute value** of a. An absolute value is always positive or zero.

Theorem 35. $\mathbb{Q}$ *embeds uniquely in every ordered field.*

Proof. Suppose K is an ordered field. Then K contains an element 1 and has the singulary operation $x \mapsto x + 1$ or S, so there is a unique homomorphism h from $(\mathbb{N}, 1, \mathrm{S})$ to $(K, 1, \mathrm{S})$. By induction on n, if $m < n$ in $\mathbb{N}$, then $h(m) < h(n)$. Therefore h is injective. Hence we can treat $\mathbb{N}$ as a subset of K, and then we can construct $\mathbb{Q}$ inside K. □

An order is **complete** if

(1) every nonempty subset with an upper bound has a *least* upper bound, and

(2) every nonempty subset with a lower bound has a greatest lower bound.

If a subset does have a least upper bound, then it is unique and is called the **supremum** of the subset. Likewise, a greatest lower bound is unique and is called an **infimum.** Only half of the definition is needed, because of:

Theorem 36. *If an order is such that every nonempty subset with an upper bound has a supremum, then the order is complete, and moreover the infimum of every nonempty subset with a lower bound is the supremum of the set of lower bounds.*

An order is **dense** if between any two distinct elements, there is a third. So every ordered field is dense, because

$$x < y \Rightarrow x < \frac{x+y}{2} < y.$$

However, we shall show that $\mathbb{Q}$ is not complete.

Theorem 37. *The equation*

$$x^2 = 2 \tag{$*$}$$

has no solution in $\mathbb{Q}$.

Proof. We use the method of **infinite descent.** Suppose there were a solution, n/m. We may assume m and n are positive integers. Then $n^2 = 2m^2$, so n must be *even:* say $n = 2k$. So $4k^2 = 2m^2$, hence $2k^2 = m^2$. Thus m/k is also a solution to $(*)$. But $0 < m < n$. Thus there is no *least* n in $\mathbb{N}$ such that, for some m in $\mathbb{N}$, n/m solves $(*)$. Therefore $(*)$ has no solution, by Theorem 14. □

Theorem 38. *The set* $\{x \in \mathbb{Q} \colon x^2 < 2\}$ *has an upper bound in* $\mathbb{Q}$*, but no supremum.*

Proof. Call the set A. It has 2 as an upper bound. Suppose b is an upper bound. We show:

(1) $2 < b^2$;

(2) A has upper bounds less than b.

For (1), suppose $c \in \mathbb{Q}$ and $c^2 \leqslant 2$. We show c is *not* an upper bound of A by finding some positive h in $\mathbb{Q}$ such that $(c+h)^2 < 2$. For all h, we have

$$(c+h)^2 = c^2 + 2ch + h^2 = c^2 + (2c+h)h.$$

We have $c^2 < 2$ by Theorem 37, and moreover $c < 2$. If also $0 < h < 1$, then $2c + h < 5$, so

$$(c+h)^2 < c^2 + 5h.$$

Thus, if we require also $h < (2 - c^2)/5$, then $(c+h)^2 < 2$. We can certainly find such h; just let h be the lesser of $1/2$ and $(2 - c^2)/6$. Therefore c is not an upper bound of A. This proves (1).

For (2), since 2 is an upper bound for A, we may assume $b \leqslant 2$. If $k > 0$, then

$$(b-k)^2 = b^2 - 2bk + k^2 > b^2 - 2bk \geqslant b^2 - 4k.$$

Since $b^2 > 2$, we can require $0 < k < (b^2 - 2)/4$; then $(b-k)^2 > 2$, so $b - k$ is an upper bound of A that is less than b. □

4. Construction of the real numbers

4.1. Dedekind cuts

As a consequence of Theorem 38, we can write $\mathbb{Q}$ as the union of two nonempty disjoint sets A and B, where

(1) each element of A is less than each element of B;
(2) A has no greatest element;
(3) B has no least element.

Indeed, just let $A = \{x \in \mathbb{Q} \colon x < 0 \vee x^2 < 2\}$, and $B = \{x \in \mathbb{Q} \colon x > 0 \mathbin{\&} x^2 > 2\}$. See Figure 4.1. Here the pair (A, B) is an example of a *cut* in the sense of Dedekind [8, I, IV., pp. 12 f.]. Since B can be obtained from A as $\mathbb{Q} \smallsetminus A$, we may just refer to A as a cut. To be precise then, we define a **cut** of $\mathbb{Q}$ to be a nonempty proper subset A of $\mathbb{Q}$ such that

Figure 4.1. A cut made by no point

(1) every element of A is less than every element of $\mathbb{Q} \smallsetminus A$,
(2) A has no greatest element.

So a cut may be as in Figure 4.1 *or* 4.2. Note that condition 2 is somewhat arbitrary; one might alternatively require a supremum of A, if it exists, to belong to A. We denote the set of cuts of $\mathbb{Q}$ by

Figure 4.2. A cut made by a point

$$\mathbb{R}.$$

That is, a cut of $\mathbb{Q}$ is precisely a **real number.**

Dedekind [8, I] observes that this construction of $\mathbb{R}$ results in the complete ordered field that we want. Details are worked out in Landau [19], and also in Spivak's *Calculus* [27, ch. 28]. Spivak writes,

> The mass of drudgery which this chapter necessarily contains is relieved by one truly first-rate idea

—namely, the idea of what Dedekind calls a cut. My own view is that, in mathematics, if you think something is drudgery, then perhaps you are not looking at it the right way.

4.2. Topology

4.2.1. Cuts

In the interest of finding some insight in the construction of $\mathbb{R}$, I note that the notion of a cut makes sense in any order. Suppose A is an order. More precisely, $(A, <)$ is the order; but I shall no longer follow the practice in previous chapters of writing out explicitly all of the operations and relations of a structure. If $b \in A$, let us define

$$\operatorname{pred}(b) = \{x \in A \colon x < b\};$$

this is the set of **predecessors** of b. If A is dense, then b is the supremum of $\operatorname{pred}(b)$. In any case, we have

$$\operatorname{pred}(x) \cap \operatorname{pred}(y) = \operatorname{pred}(z),$$

where $z = \min(x, y)$. We define a subset of A to be **open** if it is either A itself or a union

$$\bigcup\{\operatorname{pred}(x) \colon x \in Y\},$$

where $Y \subseteq A$. Note that $A = \bigcup\{\operatorname{pred}(x) \colon x \in A\}$ if and only if A has no greatest element; and this is the case we shall be interested in. In any case,

(1) the union of a family of open subsets of A is open;
(2) the intersection of two open sets is open;
(3) A itself is open.

In a word, the open subsets of A compose a **topology** on A. The subsets $\operatorname{pred}(b)$ of A, together with A itself, are **basic** open subsets of A, because every open subset of A is the union of a family of such subsets, and moreover we have:

Theorem 39. *If X and Y are open subsets of A, one of them is not A, and $b \in X \cap Y$, then there is z in A such that $b \in \operatorname{pred}(z)$ and $\operatorname{pred}(z) \subseteq X \cap Y$.*

Proof. We may assume that neither X nor Y is A. Then there are x and y in A such that $\operatorname{pred}(x) \subseteq X$ and $\operatorname{pred}(y) \subseteq Y$ and $b \in \operatorname{pred}(x) \cap \operatorname{pred}(y)$. Now let $z = \min(x, y)$. □

Theorem 40. *The family of open subsets of an order is itself linearly ordered by proper inclusion.*

Proof. Proper inclusion is automatically irreflexive and transitive. To establish linearity, suppose X and Y are distinct open subsets of A. We may assume $Y \smallsetminus X$ has an element b. Then

$$X \subseteq \operatorname{pred}(b) \subset Y.$$ □

An open subset X of A is a **cut** of A if, for some b and c in A,

$$\operatorname{pred}(b) \subseteq X \subseteq \operatorname{pred}(c).$$

That is,

(1) A is not a cut;
(2) $\varnothing$ is not a cut, unless A has a least element;
(3) every other open subset of A is a cut.

4.2.2. Completions

Let us denote the set of all cuts of A by

$$\overline{A}.$$

An **embedding** of A in an order B is an injective function f from A to B such that

$$x < y \iff f(x) < f(y). \tag{$*$}$$

This is consistent with the usage on p. 40 and later. Since our orderings are linear, the reverse direction of $(*)$ is implied by the forward.

Theorem 41. *An order A embeds in $\overline{A}$ under the map $x \mapsto \operatorname{pred}(x)$. The orders A and $\overline{A}$ alike*

(1) have a greatest element or not,
(2) have a least element or not,
(3) are dense or not.

Proof. If $x < y$, then $\operatorname{pred}(x) \subset \operatorname{pred}(y)$. If A has the greatest element b, then $\operatorname{pred}(b)$ is the greatest element of $\overline{A}$. If on the other hand b is not greatest, so that $b < c$ for some c, then $\operatorname{pred}(b) \subset \operatorname{pred}(c)$, so $\operatorname{pred}(b)$ is not greatest in $\overline{A}$. Similarly for least elements. Suppose $X \subset Y$ in $\overline{A}$. Then $X \subseteq \operatorname{pred}(b) \subset Y$ for some b as in the proof of Theorem 40. In particular, $\operatorname{pred}(b)$ is not the greatest basic open subset of A that is included in Y, so

$$X \subseteq \operatorname{pred}(b) \subset \operatorname{pred}(c) \subseteq Y \tag{†}$$

for some c, where $b < c$. If A is dense, then $b < d < c$ for some d, and then $X \subset \operatorname{pred}(d) \subset Y$. If, on the contrary, A is not dense, then A has elements e and f such that $e < f$, but nothing lies between; then no element of $\overline{A}$ lies between $\operatorname{pred}(e)$ and $\operatorname{pred}(f)$. □

Theorem 42. *If A is an order, then $\overline{A}$ is complete; indeed, if $\mathscr{B}$ is a non-empty subset of $\overline{A}$ with an upper bound, then*

$$\sup(\mathscr{B}) = \bigcup \mathscr{B}. \tag{‡}$$

Proof. Since $\bigcup \mathscr{B}$ is the union of a set of open subsets of A, it is open too. Let X be a member of $\mathscr{B}$, and Y, an upper bound of $\mathscr{B}$. Then $X \subseteq Y$, so

$$X \subseteq \bigcup \mathscr{B} \subseteq Y.$$

Thus (‡). Then $\overline{A}$ is complete by Theorem 36. □

Theorem 43. *Suppose f is an embedding of the ordered set A in a complete ordered set B. Then there is an embedding $\overline{f}$ of $\overline{A}$ in B such that*

$$\overline{f}(\operatorname{pred}(x)) = f(x) \tag{§}$$

for all x in A.

Proof. Define $\overline{f}$ by

$$\overline{f}(X) = \sup(\{f(y) \colon \operatorname{pred}(y) \subseteq X\}).$$

Then in particular $\overline{f}(\operatorname{pred}(x)) = \sup(\{f(y) \colon y \leqslant x\})$, so (§) holds since f is order-preserving. We also have

$$X \subseteq Y \implies \overline{f}(X) \leqslant \overline{f}(Y).$$

Suppose $X \subset Y$ in $\overline{A}$. As in the proof of Theorem 41, for some b and c we have (†) and therefore

$$\overline{f}(X) \leqslant \overline{f}(\operatorname{pred}(b)) = f(b) < f(c) = \overline{f}(\operatorname{pred}(c)) \leqslant \overline{f}(Y). \qquad \square$$

$$\begin{tikzcd} A \arrow[r, "\mathrm{pred}"] \arrow[d, "f"'] & \overline{A} \arrow[dl, "\overline{f}"] \\ B & \end{tikzcd}$$

Figure 4.3. Completion of an order

The situation of the theorem is shown in Figure 4.3. Because of the theorem, $\overline{A}$ can be called a **completion** of A with respect to the embedding $x \mapsto \mathrm{pred}(x)$. Note that the function $\overline{f}$ may not be unique; such a function could be defined also by (for example)

$$\overline{f}(X) = \inf\{f(y) \colon X \subseteq \mathrm{pred}(y)\}.$$

But $\overline{f}$ is unique if B is also a completion of A, by the next theorem. First note that an **isomorphism** is a surjective embedding, which implies that its inverse is also an embedding. (This is consistent with the usage on p. 25 and later.)

Theorem 44. *Suppose B is a completion of the ordered set A with respect to an embedding f. Then there is a unique embedding $\overline{f}$ of $\overline{A}$ in B such that* (§) *holds. Also, $\overline{f}$ is an isomorphism.*

Proof. By Theorem 43, there is *some* such embedding $\overline{f}$. By definition, there is an embedding g of B in $\overline{A}$ such that

$$g \circ f(x) = \mathrm{pred}(x).$$

Then by (§) we have

$$g \circ \overline{f}(\mathrm{pred}(x)) = \mathrm{pred}(x).$$

We shall show that $g \circ \overline{f}$ is the identity on $\overline{A}$. Since g is injective, it will follow that $\overline{f}$ is surjective; also $\overline{f} = g^{-1}$, so $\overline{f}$ is uniquely determined.

Say $X \in \overline{A}$. Then $g \circ \overline{f}(X)$ is an upper bound of $\{\mathrm{pred}(y) \colon \mathrm{pred}(y) \subseteq X\}$, so by Theorem 42,

$$X = \bigcup\{\mathrm{pred}(y) \colon \mathrm{pred}(y) \subseteq X\} = \sup(\{\mathrm{pred}(y) \colon \mathrm{pred}(y) \subseteq X\}) \subseteq g \circ \overline{f}(X).$$

To see the reverse inequality, say $X \subset Y$ in $\overline{A}$. Then $X \subseteq \mathrm{pred}(b) \subset Y$ for some b in A, so that

$$g \circ \overline{f}(X) \subseteq g \circ \overline{f}(\mathrm{pred}(b)) = \mathrm{pred}(b) \subset Y.$$

This shows $g \circ \overline{f}(X) \subseteq X$. Therefore $g \circ \overline{f}(X) = X$, as desired. □

We may now refer to $\overline{A}$ as **the completion** of A. Now we can define $\mathbb{R}$ as $\overline{\mathbb{Q}}$, the completion of $\mathbb{Q}$—that is, the completion of $\mathbb{Q}$ considered solely as an order.

4.3. Structure

We want to make $\mathbb{R}$ into an ordered field so that the embedding of $\mathbb{Q}$ in $\mathbb{R}$ is an embedding of ordered fields. Then $\mathbb{R}$ will be a **complete ordered field:** an ordered field that, as an order, is complete.

Let us denote by

$$\mathbb{Q}^{++}$$

the set of **improper fractions,**[*] namely, those elements x of $\mathbb{Q}^+$ such that $1 < x$. Then $\mathbb{Q}^{++}$ is an archimedean naturally ordered commutative multiplicative semigroup and is the positive part of the ordered multiplicative abelian group $\mathbb{Q}^+$. But $\mathbb{Q}^+$ is moreover a naturally ordered commutative semi-ring and is the positive part of the ordered field $\mathbb{Q}$. The procession from $\mathbb{Q}^{++}$ to $\mathbb{Q}^+$ to $\mathbb{Q}$ was suggested by Figures 3.2 (a) and 3.4; we can finally depict the whole thing as in Figure 4.4. We shall show now:

(1) $\overline{\mathbb{Q}^{++}}$ can be made into a naturally ordered commutative multiplicative semigroup;
(2) $\overline{\mathbb{Q}^{+}}$ can be made into an ordered multiplicative abelian group, of which $\overline{\mathbb{Q}^{++}}$ is the positive part;
(3) $\overline{\mathbb{Q}^{+}}$ can be made also into a naturally ordered additive semigroup, so that it is a naturally ordered unital commutative ring;
(4) $\overline{\mathbb{Q}}$ can be made into a field, of which $\overline{\mathbb{Q}^{+}}$ is the positive part.

We may relate our work to ancient mathematics by noting that a naturally ordered commutative semigroup has characteristics of a set of *magnitudes* that have a *ratio* in the sense of Euclid (see p. 17). Similarly, naturally ordered unital commutative semi-rings have characteristics of ratios themselves.

4.3.1. Ordered semigroups

We shall consider the completion of an arbitrary naturally ordered commutative semigroup; this could be $\mathbb{N}$ as well as $\mathbb{Q}^{++}$. A difference between these two

[*] The *Oxford English Dictionary* [23] defines a proper fraction as 'a fraction whose numerator is greater than (or equal to) its denominator', but quotes Robert Recorde from *The ground of artes, teachyng the worke and practise of arithetike,* 1575 edition: 'An Improper Fraction. . . a fraction in forme, which in dede is greater than an Unit.' The point is presumably that, etymologically speaking, a fraction is something *broken off* from something else; something greater than a unit cannot be broken off from a unit, so it is not properly a fraction.

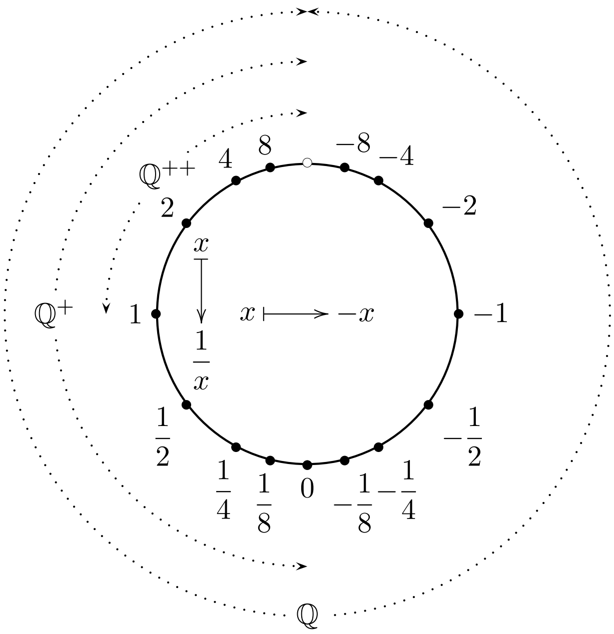


Figure 4.4. From $\mathbb{Q}^{++}$ to $\mathbb{Q}^{+}$ to $\mathbb{Q}$

structures as orders is that, while $\mathbb{Q}^{++}$ is dense, $\mathbb{N}$ is *discrete*. Discreteness is a topological notion. In addition to the topology used in § 4.2 for constructing the completion, an order has the **order topology,** defined by taking the **open intervals** as basic open sets. The open intervals are defined as in calculus: they have four possible forms,

$$(-\infty,\infty), \qquad (-\infty,b), \qquad (a,\infty), \qquad (a,b),$$

where $(-\infty,\infty)$ is the whole order, and

$$(-\infty,b)=\{x\colon x<b\}, \quad (a,\infty)=\{x\colon a<x\}, \quad (a,b)=(-\infty,b)\cap(a,\infty).$$

Here possibly $b \leqslant a$, in which case (a,b) is empty. In any case, the open interval (a,b) is completely different from the ordered pair (a,b). An element b of an order is **isolated** if $\{b\}$ is an open interval.

Theorem 45. *In a naturally ordered commutative semigroup, the following are equivalent:*

(1) there is an isolated element;

(2) there is a least element;
(3) every element is isolated.

Proof. Suppose b is isolated. Then b is the greatest element of some open interval $(-\infty, c)$. In particular, $b < c$; but then $c - b$ is the least element of the order.

Suppose d is a least element of the order. Then $(-\infty, d + d) = \{d\}$, and if $d < b$, then $(b - d, b + d) = \{b\}$. □

An order in which every element is isolated is called **discrete.**

Theorem 46. *$\mathbb{Z}$ is complete, and every complete discrete ordered abelian group is isomorphic to it.*

Proof. It is enough to prove the corresponding claim about $\mathbb{N}$. If X is a subset of $\mathbb{N}$ with an upper bound, then X has a least upper bound because $\mathbb{N}$ is well-ordered by Theorem 14; so $\mathbb{N}$ is complete. Suppose $(A, +, <)$ is a naturally ordered commutative semigroup with least element 1. By recursion, there is a homomorphism h from $(\mathbb{N}, 1, \mathrm{S})$ to $(A, 1, \mathrm{S})$. By induction, h is a *homomorphism* from $(\mathbb{N}, +)$ to $(A, +)$, that is,

$$h(m + n) = h(m) + h(n);$$

for, this equation holds immediately when $n = 1$, and if it holds when $n = k$, then $h(m+k+1) = h(m+k)+1 = h(m)+h(k)+1 = h(m)+h(k+1)$. Moreover, h preserves the ordering. Indeed, if $k < n$, then $k + m = n$ for some m, and therefore

$$h(k) < h(k) + h(m) = h(k + m) = h(n),$$

since $(A, +, <)$ is naturally ordered. In particular, h is injective; we may assume it is an inclusion.

Suppose $b \in A$ and b is less than some element of $\mathbb{N}$. If n is the *least* element of $\mathbb{N}$ that is greater than b, then $b = n - 1$; in particular, $b \in \mathbb{N}$. Suppose that there is some c in $A \smallsetminus \mathbb{N}$. Then c is an upper bound for $\mathbb{N}$, but not a least upper bound, since $c - 1$ is also an upper bound. Therefore $(A, <)$ is not complete. Contrapositively, if it is complete, then $(A, +, <)$ is isomorphic to $(\mathbb{N}, +, <)$. □

There are discrete ordered abelian groups that are not isomorphic to $\mathbb{Z}$. One example is $\mathbb{Z} \times \mathbb{Z}$ with the **left lexicographic ordering,** so that

$$(a, b) < (c, d) \iff (a < c \lor (a = c \;\&\; b < d)).$$

Written in order, the positive elements are

$$(0, 1), (0, 2), (0, 3), \dots, (1, -1), (1, 0), (1, 1), \dots;$$

these compose the naturally ordered semigroup $(\{0\}\times\mathbb{N})\cup(\mathbb{N}\times\mathbb{Z})$. This and $\mathbb{Z}\times\mathbb{Z}$ are not complete. Indeed, the subset $\{0\}\times\mathbb{N}$ has an upper bound, but no supremum. Moreover, by Theorem 46, the completion $\overline{\mathbb{Z}\times\mathbb{Z}}$ cannot be an ordered abelian group in such a way that $x\mapsto\operatorname{pred}(x)$ is an embedding of groups.

The problem with the example is that it is not *archimedean.* An ordered abelian group is **archimedean** if for all positive elements a and b, there is n in $\mathbb{N}$ such that $b\leqslant na$. The same definition holds for an ordered commutative semigroup, provided all elements are considered as positive. Compare this definition with the Archimedean Axiom on p. 20. From Theorem 46 we have:[*]

Porism 47. *No non-archimedean ordered abelian group is complete. Every discrete archimedean ordered abelian group is isomorphic to $\mathbb{Z}$ and is therefore complete.*

The semigroups $\mathbb{Q}^{++}$ and $\mathbb{Q}^{+}$ are archimedean, but not discrete. We have to show how to extend multiplication and addition respectively to the completions.

Suppose A is an arbitrary an ordered commutative additive semigroup. We want to define an addition on $\overline{A}$ this becomes a semigroup and $x\mapsto\operatorname{pred}(x)$ is an embedding of semigroups. The latter condition is just

$$\operatorname{pred}(x)+\operatorname{pred}(y)=\operatorname{pred}(x+y). \tag{$*$}$$

By the porism, we cannot hope that $\overline{A}$ will always be an *ordered* semigroup. However, we can aim for something weaker, namely that $\overline{A}$ is a semigroup, and

$$X\subseteq Y\iff X+Z\subseteq Y+Z. \tag{$\dagger$}$$

In particular, suppose $\operatorname{pred}(x)\subseteq X\subseteq\operatorname{pred}(x')$ and $\operatorname{pred}(y)\subseteq Y\subseteq\operatorname{pred}(y')$; then we want

$$\operatorname{pred}(x+y)\subseteq X+Y\subseteq\operatorname{pred}(x'+y').$$

One way to achieve this will be to define addition on $\overline{A}$ by

$$X+Y=\bigcup\{\operatorname{pred}(x+y)\colon\ \operatorname{pred}(x)\subseteq X\ \&\ \operatorname{pred}(y)\subseteq Y\}. \tag{$\ddagger$}$$

Another possibility is

$$X+'Y=\inf\{\operatorname{pred}(x'+y')\colon X\subseteq\operatorname{pred}(x')\ \&\ Y\subseteq\operatorname{pred}(y')\}; \tag{§}$$

however, we shall use (‡) as our official definition of addition on the completion of an ordered semigroup.

[*] Compare Huntington [16, 17], as quoted by Clifford [5]: 'Let S be a cancellative, naturally ordered semigroup without identity element and having a least element. If S is archimedean, then it is complete, and is isomorphic with Z [that is, $\mathbb{N}$].'

Theorem 48. *The completion of an ordered commutative semigroup is a commutative semigroup on which* $(*)$ *and* $(\dagger)$ *hold.*

Proof. Let the ordered commutative semigroup be A. Immediately from $(\ddagger)$, on $\overline{A}$ we have $(*)$ and

$$X + Y = Y + X.$$

Also, if $\operatorname{pred}(x) \subseteq X$ and $\operatorname{pred}(y) \subseteq Y$, then by definition,

$$\operatorname{pred}(x + y) \subseteq X + Y.$$

Conversely, suppose $\operatorname{pred}(u) \subseteq X + Y$. Then for some x and y such that $\operatorname{pred}(x) \subseteq X$ and $\operatorname{pred}(y) \subseteq Y$, we have $\operatorname{pred}(u) \subseteq \operatorname{pred}(x+y)$, that is, $u \leqslant x+y$. In this case, $u + z \leqslant (x + y) + z$, so $\operatorname{pred}(u + z) \subseteq \operatorname{pred}((x + y) + z)$. Therefore

$$\begin{aligned}
&(X + Y) + Z \\
&= \bigcup\{\operatorname{pred}(u + z)\colon \operatorname{pred}(u) \subseteq X + Y \mathbin{\&} \operatorname{pred}(z) \subseteq Z\} \\
&= \bigcup\{\operatorname{pred}((x + y) + z)\colon \operatorname{pred}(x) \subseteq X \mathbin{\&} \operatorname{pred}(y) \subseteq Y \mathbin{\&} \operatorname{pred}(z) \subseteq Z\} \\
&= \bigcup\{\operatorname{pred}(x + (y + z))\colon \operatorname{pred}(x) \subseteq X \mathbin{\&} \operatorname{pred}(y) \subseteq Y \mathbin{\&} \operatorname{pred}(z) \subseteq Z\} \\
&= \bigcup\{\operatorname{pred}(x + v)\colon \operatorname{pred}(x) \subseteq X \mathbin{\&} \operatorname{pred}(v) \subseteq Y + Z\} \\
&= X + (Y + Z).
\end{aligned}$$

Thus $\overline{A}$ is a commutative semigroup. The forward direction of $(\dagger)$ follows directly from $(\ddagger)$, and the converse is by linearity of the ordering. □

We can complete the left-lexicographically ordered group $\mathbb{Z} \times \mathbb{Z}$ by inserting elements (n, ∞), where $n \in \mathbb{Z}$ and

$$(n, x) < (n, \infty) < (n + 1, x).$$

Then

$$(m, x) + (n, \infty) = (m + n, \infty).$$

In particular, $(0, 1) < (0, 2)$, but $(0, 1) + (0, \infty) = (0, \infty) = (0, 2) + (0, \infty)$; so $\mathbb{Z} \times \mathbb{Z}$ is not an *ordered* semigroup. Also, using the definition (§), we have

$$(m, x) +' (n, \infty) = \begin{cases} (m + n, \infty), & \text{if } x \in \mathbb{Z}; \\ (m + n + 1, \infty), & \text{if } x = \infty; \end{cases}$$

so $+$ and $+'$ are different.

Theorem 49. *The completion of an archimedean naturally ordered commutative semigroup is an archimedean naturally ordered commutative semigroup.*

Proof. First, given arbitrary elements X and Y of the completion, we show $X \subset X + Y$. To this end, we can find y so that $\operatorname{pred}(y) \subseteq X \cap Y$, and z so that $X \subseteq \operatorname{pred}(z)$. Then there is n in $\mathbb{N}$ such that $z \leqslant ny$. Let k be the least n such that $\operatorname{pred}(ny) \subseteq X$. Then $\operatorname{pred}(ky + y) \nsubseteq X$, but $\operatorname{pred}(ky + y) \subseteq X + Y$.

Next, given elements X and Z of the completion such that $X \subset Z$, define

$$Y = \bigcup \{\operatorname{pred}(y) \colon \forall x\ (\operatorname{pred}(x) \subseteq X \Rightarrow \operatorname{pred}(x + y) \subseteq Z)\}.$$

Immediately $X + Y \subseteq Z$. To prove $Z \subseteq X + Y$, it is enough to show

$$\operatorname{pred}(z) \subset Z \implies \operatorname{pred}(z) \subset X + Y.$$

Indeed, if $X + Y \subset Z'$, then $Z' \smallsetminus (X + Y)$ contains some w, and then

$$X + Y \subseteq \operatorname{pred}(w) \subset Z'.$$

So we assume $\operatorname{pred}(z) \subset Z$. To show $\operatorname{pred}(z) \subset X + Y$, it is enough to find u such that

(1) $\operatorname{pred}(u) \subset X$;

(2) $\operatorname{pred}(z - u) \subseteq Y$.

The first condition requires X to be nonempty. However, suppose X is empty. Then the original naturally ordered semigroup has a least element: call this 1. Then $X = \operatorname{pred}(1)$ and

$$Y = \bigcup \{\operatorname{pred}(y) \colon \operatorname{pred}(1 + y) \subseteq Z\}.$$

Since Z is not all of A, there is v such that $Z \subseteq \operatorname{pred}(v)$, and then, by the archimedean property, there is k in $\mathbb{N}$ such that $v \leqslant k \cdot 1$. If n is the least k such that $Z \subseteq \operatorname{pred}(k \cdot 1)$, then $Z = \operatorname{pred}(n \cdot 1)$, so $Y = \operatorname{pred}((n-1) \cdot 1)$ and $X + Y = Z$. We may now assume that X is not empty.

The second condition requires $u < z$, so that $z - u$ is defined. This is impossible if z is the least element of A. But in this case, $z \in X$ (since this is nonempty), so $\operatorname{pred}(z) \subset X + Y$. We may now assume that z is not the least element of A.

By definition of Y, the second condition now means

$$\operatorname{pred}(x) \subseteq X \implies \operatorname{pred}(z + x - u) \subseteq Z.$$

Loosely, u must be closer to the top of X than z is to the top of Z. Using the archimedean property, we can achieve this.

To do so, note first that there is w such that $\operatorname{pred}(z) \subset \operatorname{pred}(w) \subseteq Z$; and there is v such that $X \subseteq \operatorname{pred}(v)$. Then there is k in $\mathbb{N}$ such that $v \leqslant k(w-z)$ and hence

$$X \subseteq \operatorname{pred}(k(w-z)).$$

Let n be the *least* k such that this inclusion holds.

If $n=1$, let u be an arbitrary element of X such that $u<z$. If $\operatorname{pred}(x) \subseteq X$, then

$$x \leqslant w-z, \qquad z+x \leqslant w, \qquad z+x-u \leqslant w,$$

and hence $\operatorname{pred}(z+x-u) \subseteq Z$.

Suppose now $n>1$. Let $u=(n-1)(w-z)$. If $\operatorname{pred}(x) \subseteq X$, then

$$x \leqslant n(w-z), \qquad z+x-u \leqslant z+(w-z)=w,$$

and again $\operatorname{pred}(z+x-u) \subseteq Z$. Therefore in all cases $X+Y=Z$. Thus the completion is naturally ordered as a commutative semigroup.

This ordered commutative semigroup is archimedean, because if it contains X and Y, then it contains some $\operatorname{pred}(x)$ and $\operatorname{pred}(y)$ such that $\operatorname{pred}(x) \subseteq X$ and $Y \subseteq \operatorname{pred}(y)$; then for some n in $\mathbb{N}$ we have $y \leqslant nx$, so

$$Y \subseteq \operatorname{pred}(y) \subseteq \operatorname{pred}(nx) \subseteq nX. \qquad \square$$

Theorem 50. *If A is an archimedean ordered abelian group, there is only one way to make $\overline{A}$ a semigroup so that (†) holds and $x \mapsto \operatorname{pred}(x)$ is an embedding of semigroups.*

Proof. The official definition (‡) of $X+Y$ in $\overline{A}$ is the least possible. Therefore, supposing $X+Y \subset Z$, we need only find x' and y' in A such that $X \subseteq \operatorname{pred}(x')$ and $Y \subseteq \operatorname{pred}(y')$, but

$$\operatorname{pred}(x'+y') \subset Z.$$

There are w and z in A such that

$$X+Y \subseteq \operatorname{pred}(w) \subset \operatorname{pred}(z) \subseteq Z.$$

We may assume A is dense. Then there are positive u and v in A such that $u+v<z-w$. By the archimedean property, there are m and n in $\mathbb{Z}$ such that

$$\operatorname{pred}(mu) \subset X \subseteq \operatorname{pred}((m+1)u), \qquad \operatorname{pred}(nv) \subset Y \subseteq \operatorname{pred}((n+1)v).$$

Let $x' = (m+1)u$ and $y' = (n+1)v$. Since

$$\operatorname{pred}(mu+nv) \subset X+Y \subseteq \operatorname{pred}(w),$$

we have

$$\operatorname{pred}(x'+y') \subseteq \operatorname{pred}(w+u+v) \subset \operatorname{pred}(z) \subseteq Z. \qquad \square$$

Suppose S is the positive part of an archimedean ordered abelian group G. Then $\overline{S}$ is the positive part of some ordered abelian group. We want to show that this ordered group is isomorphic to $\overline{G}$. To this end, we note that, for every order A, there is an **opposite order,** A^{op}: this means $x < y$ in A^{op} if and only if $y < x$ in A.

Lemma. *Let f be the embedding $x \mapsto \operatorname{pred}(x)$ of an order A in the completion $\overline{A}$. Then $\overline{A}^{\mathrm{op}}$ is a completion of A^{op} with respect to f.*

Proof. All of the relevant definitions are symmetric in the sense that they are equivalent if $<$ is replaced with $>$ throughout. In particular, the order $\overline{A}^{\mathrm{op}}$ is complete, and f embeds A^{op} in this, and then the claim follows. □

Lemma. *If A and B are orders, and every element of A is less than every element of B, then there is an isomorphism from $\overline{A \cup B}$ to $\overline{A} \cup \overline{B}$, where, in the latter, every element of $\overline{A}$ is less than every element of $\overline{B}$. The isomorphism is*

$$X \mapsto \begin{cases} X, & \textit{if } X \subset A, \\ B \cap X, & \textit{if } A \subseteq X. \end{cases}$$

Proof. The given function is well-defined, since A is an open subset of $A \cup B$, and also, for every element x of B, the predecessors of x in B are just the predecessors of x in $A \cup B$ that are actually in B. Then the function is order-preserving and bijective. □

Theorem 51. *Let G be an archimedean ordered abelian group with positive part S. Then $\overline{G}$ is an archimedean ordered abelian group, and the induced embedding of $\overline{S}$ in $\overline{G}$ is an isomorphism of $\overline{S}$ with the positive part of $\overline{G}$.*

Proof. We can understand G as the disjoint union $-S \cup \{0\} \cup S$, where

$$-S = \{-x \colon x \in S\}.$$

By the last lemma, $\overline{G}$ as an order can be considered as the disjoint union

$$\overline{-S} \cup \{\varnothing\} \cup \overline{S}.$$

Here both S and $-S$ are ordered commutative semigroups, so $\overline{S}$ and $\overline{-S}$ are commutative semigroups as in Theorem 48; moreover, $\overline{S}$ is a naturally ordered commutative semigroup, by Theorem 49. The map $x \mapsto -x$ is an isomorphism of the ordered semigroups S^{op} and $-S$. By this and the next-to-last lemma, there is an order isomorphism from $\overline{S}$ to $\overline{-S}^{\mathrm{op}}$ that takes $\{x\colon x < a\}$ to $\{y\colon y < -a\}$. Then we can make $\overline{G}$ into a group of which $\overline{S}$ is the positive part, and moreover $x \mapsto \operatorname{pred}(x)$ is an embedding of G in $\overline{G}$ as groups. By Theorem 50, the group-structure on G must be the semigroup-structure given by Theorem 48 □

4.3.2. Naturally ordered semi-rings

Theorem 52. *The completion of a naturally ordered commutative semi-ring is a commutative semi-ring.*

Proof. If A is a naturally ordered commutative semi-ring, then on $\overline{A}$ we have, as in the proof of Theorem 48,

$$\begin{aligned}
&X \cdot (Y + Z)\\
&= \bigcup\{\operatorname{pred}(x \cdot w)\colon \operatorname{pred}(x) \subseteq X \;\&\; \operatorname{pred}(w) \subseteq Y + Z\}\\
&= \bigcup\{\operatorname{pred}(x \cdot (y + z))\colon \operatorname{pred}(x) \subseteq X \;\&\; \operatorname{pred}(y) \subseteq Y \;\&\; \operatorname{pred}(z) \subseteq Z\}\\
&= \bigcup\{\operatorname{pred}(x \cdot y + x \cdot z)\colon \operatorname{pred}(x) \subseteq X \;\&\; \operatorname{pred}(y) \subseteq Y \;\&\; \operatorname{pred}(z) \subseteq Z\}\\
&= \bigcup\{\operatorname{pred}(x \cdot y + x' \cdot z)\colon \operatorname{pred}(x) \subseteq X \;\&\; \operatorname{pred}(y) \subseteq Y \qquad (\P)\\
&\qquad\qquad \;\&\; \operatorname{pred}(x') \subseteq X \;\&\; \operatorname{pred}(z) \subseteq Z\}\\
&= \bigcup\{\operatorname{pred}(u + v)\colon \operatorname{pred}(u) \subseteq X \cdot Y \;\&\; \operatorname{pred}(v) \subseteq X \cdot Z\}\\
&= X \cdot Y + X \cdot Z.
\end{aligned}$$

Indeed, (¶) is justified by noting

$$x \cdot y + x' \cdot z \leqslant \max(x, x') \cdot y + \max(x, x') \cdot z.$$ □

Now we have what was described at the beginning of the section:

Theorem 53. $\mathbb{R}$ *is a complete ordered field.*

Proof. Because $\mathbb{Q}^+$ is an archimedean naturally ordered commutative semigroup and, as such, is the positive part of $\mathbb{Q}$, also $\overline{\mathbb{Q}^+}$ is an archimedean naturally ordered commutative semigroup by Theorem 49; we can consider it as the positive part of $\overline{Q}$, by Theorem 51. Because $\mathbb{Q}^+$ is an archimedean ordered abelian multiplicative group, so is $\overline{\mathbb{Q}^+}$, by Theorem 51. Because $\mathbb{Q}^+$ is a naturally ordered commutative semi-ring, so is $\overline{\mathbb{Q}^+}$, by Theorem 52; it is unital, because it is a multiplicative group. Then the multiplication on this extends to $\overline{\mathbb{Q}}$, making this an ordered commutative ring, by Porism 33. Since $\overline{\mathbb{Q}^+}$ is a multiplicative group, $\overline{\mathbb{Q}}$ must be a field. □

We may denote by

$$\mathbb{R}^+, \qquad\qquad \mathbb{R}^{++},$$

the subsets $\{x \colon 0 < x\}$ and $\{x \colon 1 < x\}$ of $\mathbb{R}$, respectively; we can identify these subsets with $\overline{\mathbb{Q}^+}$ and $\overline{\mathbb{Q}^{++}}$. Say $b \in \mathbb{R}^{++}$. By recursion, we have a homomorphism $x \mapsto b^x$ of ordered semigroups from $\mathbb{N}$ to $\mathbb{R}^{++}$ given by

$$b^1 = b, \qquad\qquad b^{n+1} = b^n \cdot b.$$

This extends uniquely to a homomorphism of ordered abelian groups from $\mathbb{Z}$ to $\mathbb{R}^+$, where

$$b^{-x} = (b^x)^{-1}.$$

We can extend further, to an isomorphism of ordered abelian groups from $\mathbb{R}$ to $\mathbb{R}^+$, by Theorem 54 below.

A commutative group is **divisible** if for every element a and every n in $\mathbb{Z}$ there is a solution of the equation

$$nx = a$$

in the group.

Lemma. *Every complete dense ordered abelian group is divisible.*

Proof. Let A be such a group, $b \in A$, and $n \in \mathbb{Z}$. It is enough to suppose $b > 0$ and $n \in N$. By density, there are elements a_k of A such that

$$0 = a_0 < a_1 < \cdots < a_{n-1} < a_n = b.$$

Let a be the least of the differences $a_k - a_{k-1}$. Then $na \leqslant b$. Therefore the set $\{x \in A \colon nx \leqslant b\}$ contains a; it also has b as an upper bound; so it has a supremum, c. Then $nc = b$. Indeed, if $nc' < b$, then for some x_k we have

$$nc' = x_0 < x_1 < \cdots < x_{n-1} < x_n = b;$$

if d is the least of the $x_k - x_{k-1}$, then $n(c' + d) \leqslant b$, and therefore $c' < c' + d \leqslant c$. Similarly, if $b < nc''$, then $c < c''$. □

Theorem 54 (Hölder). *If A is a complete dense ordered abelian group, and b is a positive element of A, then there is a unique embedding of ordered abelian groups from $\mathbb{R}$ to A that takes 1 to b; and this embedding is an isomorphism.*

Proof. It is enough to find a unique embedding of $\mathbb{R}^+$ in the positive part A^+ of A, and then to show that this embedding is an isomorphism.* Let φ be the embedding of $\mathbb{N}$ in A^+ that takes 1 to b. If $m/n = m'/n'$ in $\mathbb{Q}^+$, then $mn' = m'n$, so in A^+,

$$nx = mb \iff mn'nx = m'nmb \iff n'x = m'b.$$

Therefore we can define φ on $\mathbb{Q}^+$ by letting $\varphi(m/n)$ be the unique solution of $nx = mb$. Then φ is the unique embedding of $\mathbb{Q}^+$ in A^+ that takes 1 to b.

We show now that A^+ is the completion of the image of $\mathbb{Q}^+$ under φ. If $c \in A^+$, let

$$c' = \sup(\{\varphi(x) \colon \varphi(x) \leqslant c\}).$$

Then $c' = c$. Indeed, immediately $c' \leqslant c$. Suppose $d < c$. Since A^+ is archimedean by Porism 47, for some n in $\mathbb{N}$ we have $b < n(c-d)$, so $nd+b < nc$. If m is the least k in $\mathbb{N}$ such that $nc < kb$, then

$$nd < (m-1)b \leqslant nc,$$
$$d < \varphi\Bigl(\frac{m-1}{n}\Bigr) \leqslant c,$$

so $d < c'$. Therefore $c' = c$. This shows that A^+ is the completion of $\varphi[\mathbb{Q}^+]$. Therefore φ extends to an isomorphism from $\mathbb{R}^+$ to A^+ by Theorem 43. □

Hence, for any choice of b in $\mathbb{R}^{++}$, we have the isomorphism $x \mapsto b^x$ of ordered abelian groups from $\mathbb{R}$ to $\mathbb{R}^+$.

Theorem 55. *The ordered field $\mathbb{R}$ embeds uniquely in every complete ordered field, and this embedding is an isomorphism.*

Proof. Let K be a complete ordered field. By Theorem 54, there is an embedding φ of ordered abelian groups from $\mathbb{R}$ to K that takes 1 to 1, and this is an isomorphism. But the ring-structure on $\mathbb{Q}$ is uniquely determined by the group-structure. Also, by Theorem 50, the group-structure on $\mathbb{R}^+$ is uniquely determined by the group-structure on $\mathbb{Q}^+$; and then the field-structure on $\mathbb{R}$ is determined by the additive homomorphism $x \mapsto -x$ on $\mathbb{R}^+$. Therefore φ must be a field-isomorphism. □

*Compare Hölder [13, 14, 15] as quoted by Clifford [5]: 'Let S be a cancellative, naturally ordered semigroup without identity element and without a least element. (A) S is isomorphic with P [namely $\mathbb{R}^+$] if and only if it is complete. (B) S can be embedded in P if and only if it is archimedean.'

Henceforth we may say $\mathbb{R}$ is **the complete ordered field.**

4.4. Sequences

Another way to construct $\mathbb{R}$ is by means of *Cauchy sequences.* First of all, a sequence $(a_n\colon n \in \mathbb{N})$ **converges** to the real number b if, for all positive real numbers ε, there is a natural number M such that, for all n in $\mathbb{N}$, if $n \geqslant M$, then

$$|a_n - b| < \varepsilon.$$

In this case, we write

$$\lim_{n\to\infty} a_n = b,$$

saying b is the **limit** of the sequence. If we denote the sequence $(a_n\colon n \in \mathbb{N})$ by the single letter a, then we may write simply

$$\lim(a) = b.$$

Theorem 56. *A bounded monotone sequence in $\mathbb{R}$ converges.*

Proof. Let a be a bounded increasing sequence, and let $b = \sup(a)$. Suppose $\varepsilon > 0$. Then $b - \varepsilon$ is not an upper bound of a, so for some M in $\mathbb{N}$, we have

$$b - \varepsilon < a_M \leqslant b.$$

Since the sequence is increasing, if $n \geqslant M$, we have

$$b - \varepsilon < a_M \leqslant a_n \leqslant b,$$

and therefore

$$|a_n - b| = b - a_n < \varepsilon.$$

Thus a converges to b. Similarly, bounded decreasing sequences converge. □

A sequence a of real numbers is a **Cauchy sequence** if, for every positive real number ε, there is a natural number M such that, for all m and n in $\mathbb{N}$, if $m \geqslant M$ and $n \geqslant M$, then

$$|a_m - a_n| < \varepsilon.$$

For example, let sequences p and q be defined recursively* by

$$\begin{aligned} p_1 &= 1, & p_{n+1} &= p_n + 2q_n,\\ q_1 &= 1, & q_{n+1} &= p_n + q_n. \end{aligned}$$

*Strictly, we are defining $((p_n, q_n)\colon n \in \mathbb{N})$ recursively.

Let $a_n = p_n/q_n$. Then

$$a = \left(1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \dots\right).$$

It is an exercise show that

$$p_n q_{n+1} - q_n p_{n+1} = (-1)^n,$$

and hence

$$a_{n+1} - a_n = \frac{(-1)^{n+1}}{q_{n+1}q_n}.$$

Since q is increasing, it follows that

$$a_1 < a_3 < a_5 < \dots < a_6 < a_4 < a_2,$$

and moreover a is a Cauchy sequence. The two subsequences $(a_{2n-1} \colon n \in \mathbb{N})$ and $(a_{2n} \colon n \in \mathbb{N})$ of a converge by Theorem 56. Moreover, since

$${p_n}^2 - 2{q_n}^2 = (-1)^n,$$

the two sequences must have the same limit, which is therefore $\lim(a)$. This limit is $\sqrt{2}$, which however is not in $\mathbb{Q}$, by Theorem 37.

Theorem 57. *Every Cauchy sequence in* $\mathbb{R}$ *is bounded.*

Proof. Let a be a Cauchy sequence. Let M be such that, if $m \geqslant M$ and $n \geqslant M$, then $|a_m - a_n| \leqslant 1$. In particular, if $m \geqslant M$, then

$$|a_m| \leqslant |a_m - a_M| + |a_M| \leqslant 1 + |a_M|.$$

Thus each $|a_n|$ is bounded by $\max(|a_0|, \dots, |a_{M-1}|, 1 + |a_M|)$. □

Theorem 58. *Every Cauchy sequence in* $\mathbb{R}$ *converges.*

Proof. Let a be a Cauchy sequence. By Theorem 57, the sequence is bounded below by some A. We can also define

$$b_k = \sup(\{a_n \colon n \geqslant k\}).$$

Then $(b_k \colon k \in \mathbb{N})$ is decreasing, but bounded below by A; so the sequence converges to some c by Theorem 56. We have

$$|a_m - c| \leqslant |a_m - a_n| + |a_n - b_k| + |b_k - c|. \qquad (*)$$

Let $\varepsilon > 0$. There is some M such that, if $k \geqslant M$, $m \geqslant M$, and $n \geqslant M$, then $|a_m - a_n| < \varepsilon/3$ and $|b_k - c| < \varepsilon/3$. For all k, there is n such that $n \geqslant k$ and $|a_n - b_k| < \varepsilon/3$. Therefore, if $m \geqslant M$, then there are some k and n such that the right-hand side of $(*)$ is less than ε; so $|a_m - c| < \varepsilon$. Thus a converges to c. □

For the alternative construction of $\mathbb{R}$, let us denote by

$$\mathbb{Q}^{\mathbb{N}}$$

the set of functions from $\mathbb{N}$ to $\mathbb{Q}$, that is, rational sequences. This becomes a commutative ring when we let 0 be $(0,0,0,\dots)$ and 1 be $(1,1,1,\dots)$ and, writing a for $(a_n \colon n \in \mathbb{N})$ as before, we define $a+b$, $-a$, and ab by

$$(a+b)_n = a_n + b_n, \qquad (-a)_n = -a_n, \qquad (ab)_n = a_n b_n.$$

Then the commutative ring $\mathbb{Q}$ embeds in $\mathbb{Q}^{\mathbb{N}}$ under the map that takes x to the sequence $(x,x,x,\dots)$. Let S be the set of Cauchy sequences in $\mathbb{Q}^{\mathbb{N}}$. In particular, the sequences $(x,x,x,\dots)$ belong to S. Moreover:

Theorem 59. *S is a sub-ring of $\mathbb{Q}^{\mathbb{N}}$.*

Proof. Since $\mathbb{Q}$ embeds in S, so that 0 and 1 are in S, we need only show in addition that S is closed under the ring-operations of addition, additive inversion, and multiplication. The most difficult part is multiplication. Let a and b be in S. By Theorem 57, there is M such that, for all n in $\mathbb{N}$, we have $|a_n| \leqslant M$ and $|b_n| \leqslant M$. Hence

$$\begin{aligned} |a_m b_m - a_n b_n| &= |a_m b_m - a_n b_m + a_n b_m - a_n b_n| \\ &\leqslant |a_m - a_n|\,|b_m| + |a_n|\,|b_m - b_n| \\ &\leqslant M(|a_m - a_n| + |b_m - b_n|). \end{aligned}$$

Then ab is Cauchy. □

By Theorem 58, we have a map $x \mapsto \lim(x)$ from S to $\mathbb{R}$.

Theorem 60. *The function $x \mapsto \lim(x)$ from S to $\mathbb{R}$ is a homomorphism.*

Let I be the set of sequences in S that converge to 0; so I is the kernel of the homomorphism $x \mapsto \lim(x)$. Then I is an **ideal** of S.

Theorem 61. *The homomorphism $x \mapsto \lim(x)$ from S to $\mathbb{R}$ is surjective, inducing an isomorphism from S/I onto $\mathbb{R}$. In particular, I is a maximal ideal of S.*

Proof. Let $A \in \mathbb{R}$. Then $A = \sup(\{x \in \mathbb{Q} \colon x < A\})$. By the *Axiom of Choice,* there is a sequence a in $\mathbb{Q}^{\mathbb{N}}$ such that

$$a_0 < A, \qquad a_n + \frac{A - a_n}{2} < a_{n+1} < A.$$

Then $\lim(a) = A$. □

Independently of the theorem—that is, without having previously defined $\mathbb{R}$—, one can show that I is a maximal ideal of S. Then S/I is a field, and one can show that it is a complete ordered field. Thus there is an alternative construction of $\mathbb{R}$. The construction is carried out more generally in Theorem 76 in the next chapter.

5. Non-archimedean fields and valuations

One purpose of this chapter, as suggested at the end of the last chapter, is to give an alternative construction of $\mathbb{R}$ from $\mathbb{Q}$. This construction can be understood as being generally applicable to *normed fields*—fields with a **norm function.** Ordered fields are examples of these, their norm function being the absolute value function; but there are other normed fields, namely the *valued fields.* Every normed field has a completion (5.2.3). In particular, non-archimedean ordered fields have completions as normed fields, and these completions are non-archimedean: thus $\mathbb{R}$ is not unique as a complete normed field or even as a complete field with absolute value function.

Essential to non-standard analysis is the non-archimedean ordered field ${}^*\mathbb{R}$, an instance of the general construction to be given in Ch. 6. A second purpose of this chapter is to give a simpler example of a non-archimedean ordered field. This example also leads to the general notion of a valued field.

5.1. Non-archimedean fields

Given an ordered field K, by Theorem 35 we may assume $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subseteq K$. An element x of K is:

(1) **infinite,** if $n < |x|$ for all n in $\mathbb{N}$;

(2) **finite,** if not infinite;

(3) **infinitesimal,** if $|x| < 1/n$ for all n in $\mathbb{N}$.

Theorem 62. *The following are equivalent conditions on an ordered field K.*

1. *K is non-archimedean.*
2. *K has infinite elements.*
3. *K has nonzero infinitesimal elements.*

Proof. Immediately x is infinite if and only if x^{-1} is a nonzero infinitesimal, and such x exist in ordered field only if it is non-archimedean. Conversely, in such a field, there are positive elements a and b such that $na < b$ for all n in $\mathbb{N}$. If a is infinitesimal, we are done. Suppose a is not infinitesimal. Then $1/n \leqslant a$ for some n in $\mathbb{N}$. Hence

$$k = \frac{kn}{n} \leqslant kna < b$$

for all k in $\mathbb{N}$, so b is infinite. □

Another way to prove the theorem is to suppose K is non-archimedean, so that there is a positive element a of $K \smallsetminus \mathbb{R}$. If a is not infinite, then $\{x \in \mathbb{R} \colon x < a\}$ is nonempty and bounded above, so it has a supremum, b; then $b - a$ is a nonzero infinitesimal.

5.1.1. Rational functions

An example of a non-archimedean ordered field can be obtained by ordering the field

$$\mathbb{R}(X)$$

of **rational functions** over $\mathbb{R}$ in one variable, X. Here $\mathbb{R}(X)$ is the quotient field of the commutative ring

$$\mathbb{R}[X]$$

of **polynomials** in X over $\mathbb{R}$. Such a polynomial, if it is not 0, can be written uniquely as

$$a_0 + a_1X + a_2X^2 + \cdots + a_nX^n \tag{$*$}$$

or

$$\sum_{k=0}^{n} a_kX^k,$$

where $n \in \mathbb{N} \cup \{0\}$, the coefficients a_k are in $\mathbb{R}$, and $a_n \neq 0$. Then $\mathbb{R}(X) \smallsetminus \{0\}$ consists of the fractions

$$\frac{a_0 + \cdots + a_nX^n}{b_0 + \cdots + b_mX^m}, \tag{$\dagger$}$$

where $a_nb_m \neq 0$.

Theorem 63. *The field $\mathbb{R}(X)$ can be ordered by defining the fraction in ($\dagger$) to be positive if and only if $a_nb_m > 0$. Then for all n in $\mathbb{N}$,*

$$0 < \cdots < \frac{1}{X^3} < \frac{1}{X^2} < \frac{1}{X} < \frac{1}{n} \leqslant n < X < X^2 < X^3 < \cdots, \tag{$\ddagger$}$$

so the positive powers of X are infinite, and the negative powers of X are infinitesimal.

5.1.2. Valuation rings

The **units** or multiplicatively invertible elements of a commutative ring R compose a multiplicative group denoted by

$$R^{\times}.$$

Theorem 64. *Let K be an arbitrary non-archimedean ordered field. The finite elements of K compose a sub-ring R, and the infinitesimal elements compose a maximal ideal I of R. An element a of $K^{\times}$ is infinite if and only if $a^{-1} \in I$. In particular, either a or a^{-1} is finite, so it belongs to R.*

An arbitrary commutative ring R is called a **valuation ring** if, for every nonzero element a of its quotient field, either a or a^{-1} is in R. The reason for the terminology will be seen below. Meanwhile, the finite elements of a non-archimedean field compose a valuation ring. Trivially, fields are valuation rings. A commutative ring R is a **local ring** if it has a unique maximal ideal; if this ideal is I, then R/I must be a field.

Theorem 65. *Every nontrivial valuation ring is a local ring whose unique maximal ideal consists of the elements of the ring that are not units.*

Proof. Suppose R is a valuation ring and $I = R \smallsetminus R^{\times}$. Let a and b be nonzero elements of R. Then we may assume $a/b \in R$. But $(a+b)/b = a/b + 1$, which is now also in R. Hence, if $a + b \in R^{\times}$, then also $b \in R^{\times}$. Contrapositively, if both a and b are in I, then $a+b \in I$; so I is closed under addition. Similarly, if $r \in R$, and $1/rb \in R$, then $b \in R^{\times}$; so I is closed under multiplication by members of R. If I is not all of R, then it is a maximal ideal, since every ideal containing an element of $R \smallsetminus I$ contains a unit and is therefore all of R. □

Theorem 66. *Let K be an ordered field that includes $\mathbb{R}$, and let R be the ring of finite elements of K, with maximal ideal I of infinitesimals. Then the quotient map $x \mapsto x + I$ determines an isomorphism from $\mathbb{R}$ onto R/I.*

Proof. Let h be $x \mapsto x + I$ on $\mathbb{R}$. Then $\ker(h) = I \cap \mathbb{R}$, which is $\{0\}$. Thus h is injective. It remains to show h is surjective onto R/I.

Let $a \in R$. Since a is finite, the set $\{x \in \mathbb{R} \colon x < a\}$ has an upper bound in $\mathbb{R}$, hence a supremum, a'. We shall show $h(a') = a + I$. To this end, suppose $b \in \mathbb{R}$, but $h(b) \neq a + I$. This means $b - a$ is not infinitesimal. In particular, for some real number δ, we have

$$0 < \delta < |b - a|\,.$$

If $b < a$, then $b < b + \delta < a$, so b is not an upper bound of $\{x \in \mathbb{R} \colon x < a\}$. If $a < b$, then $a < b - \delta$, so b is not the supremum of $\{x \in \mathbb{R} \colon x < a\}$. In either case, $b \neq a'$. □

In the notation of the theorem, if a and b are arbitrary elements of K such that $a - b \in I$, then a and b are **infinitely close,** and we write

$$a \simeq b. \qquad (\S)$$

By the theorem, if a is finite, then a is infinitesimally close to some *unique* real number; this number is called the **standard part** of a. In particular, the infinisimals are the elements whose standard part is 0.

For example, the finite elements of $\mathbb{R}(X)$ are those of the form

$$\frac{a_nX^n + \cdots + a_0}{b_nX^n + \cdots + b_0},$$

where $b_n \neq 0$. The standard part of this element is a_n/b_n, since

$$\frac{a_nX^n + \cdots + a_0}{b_nX^n + \cdots + b_0} - \frac{a_n}{b_n} = \frac{(a_{n-1} - a_nb_{n-1}/b_n)X^{n-1} + \cdots}{b_nX^n + \cdots + b_0}.$$

5.1.3. Power series

Using the division algorithm taught in school, we can formally compute the quotient of two nonzero elements of $\mathbb{R}[X]$, getting a possibly infinite series

$$c_0 + c_1X^{-1} + c_2X^{-2} + \cdots$$

or simply

$$\sum_{n=0}^{\infty} c_kX^{-k};$$

this is a **formal power series** in X^{-1} with coefficients from $\mathbb{R}$. For example, formally,

$$\frac{X}{X-1} = 1 + X^{-1} + X^{-2} + \cdots$$

The set of all formal power series in X^{-1} over $\mathbb{R}$ is denoted by $\mathbb{R}[[X^{-1}]]$ or rather

$$\mathbb{R}[[T]],$$

where $T = X^{-1}$. This set is an integral domain in the obvious way, and its quotient field is denoted by

$$\mathbb{R}((T));$$

this is the field of **formal Laurent series** in T with coefficients from $\mathbb{R}$, namely series

$$\sum_{n=k}^{\infty} a_nT^n, \tag{¶}$$

where $k \in \mathbb{Z}$, and each a_n is in $\mathbb{R}$. This field includes $\mathbb{R}(T)$, which is $\mathbb{R}(X)$.

The ordering of $\mathbb{R}(T)$, in which T is a positive infinitesimal, extends to $\mathbb{R}((T))$. Indeed, let a be the element in (¶), and assume $a_k \neq 0$. Then a is

1. positive if and only if $a_k > 0$,
2. finite if and only if $k \geqslant 0$,
3. infinitesimal if and only if $k > 0$.

If a is finite, then its standard part is a_0 (which is 0 if $k > 0$).

5.2. Valuations

5.2.1. Power series

The construction of $\mathbb{R}((T))$ as a field uses only that $\mathbb{R}$ is a field. Let K be an arbitrary field, not necessarily ordered; then we can form the field

$$K((T))$$

of formal Laurent series in T with coefficients from K. This has the sub-ring

$$K[[T]]$$

of formal power series in T with coefficients from K.

Theorem 67. *If K is a field, then the ring $K[[T]]$ is a valuation ring; its unique maximal ideal is (T), comprising the series $\sum_{n=1}^{\infty} a_n T^n$ with no constant term. Then K is isomorphic to*

$$K[[T]]/(T)$$

under $\xi \mapsto \xi + (T)$.

Proof. The given map is an isomorphism since $\sum_{n=0}^{\infty} a_n T^n + (T) = a_0 + (T)$, but $a + (T) \neq b + (T)$ if a and b are distinct elements of K. Hence (T) is a maximal ideal. It is unique as such since every non-element can be formally inverted. □

There is another quotient we can form, namely

$$K((T))^{\times}/K[[T]]^{\times}.$$

Here $K((T))^{\times}$ is just $K((T)) \smallsetminus \{0\}$, and $K[[T]]^{\times} = K[[T]] \smallsetminus (T)$ as in Theorem 65. If $a_k \neq 0$, then

$$\frac{1}{T^k} \sum_{n=k}^{\infty} a_n T^n = \sum_{n=0}^{\infty} a_{k+n} T^n,$$

which is in $K[[T]]^{\times}$; so we may write

$$\sum_{n=k}^{\infty} a_n T^n \equiv T^k \pmod{K[[T]]^{\times}}.$$

Thus the quotient map $\xi \mapsto \xi K[[T]]^{\times}$ on $K((T))^{\times}$ gives a bijection between $\langle T\rangle$ (that is, $\{T^n : n \in \mathbb{Z}\}$) and $K((T))^{\times}/K[[T]]^{\times}$. By defining

$$0 < \cdots < T^2 < T < 1 < T^{-1} < \cdots ,$$

we induce an ordering on $\{0\} \cup K((T))^{\times}/K[[T]]^{\times}$. The field $K((T))$ thus provides an example of the following general situation.

5.2.2. Valuations

Assume the following data:

(1) K is an arbitrary field;
(2) Γ is an ordered abelian group written multiplicatively;
(3) the ordering and multiplication are extended to $\{0\} \cup \Gamma$ so that

$$0 < v, \qquad 0 \cdot 0 = 0, \qquad 0 \cdot v = v \cdot 0 = 0,$$

for all v in Γ; and

(4) there is a function $x \mapsto |x|$ from K to $\{0\} \cup \Gamma$ so that

$$\begin{gathered} |xy| = |x|\,|y|\,, \\ |x| = 0 \iff x = 0, \\ |x+y| \leqslant \max(|x|\,, |y|). \end{gathered} \tag{$*$}$$

Then the function $x \mapsto |x|$ is a **valuation** on K.[*] The inequality $(*)$ is the **strong triangle inequality.**

There is a related notion. In the same situation, if Γ is the positive part of an ordered field, but the triangle inequality is assumed to hold only in its standard form

$$|x+y| \leqslant |x| + |y|\,,$$

then the function $x \mapsto |x|$ on K is an **absolute value function** or **norm.**[†] This generalizes the definition on p. 46: every ordered field has an absolute value function, but so does the non-orderable field $\mathbb{C}$. (For definiteness, let $\mathbb{C}$ be the field $\mathbb{R}(X)/(X^2+1)$, which, as a real vector space, has basis consisting of 1 and $X + (X^2+1)$; but we write the latter basis element as i. Then $|x + y\mathrm{i}| = \sqrt{x^2+y^2}$.)

[*] Sometimes Γ is written additively and the opposite ordering is used, while 0 is written as ∞; I follow the practice of Lang [20, § I.2].

[†] It is usually assumed that Γ here is the positive part of a subfield of $\mathbb{R}$; but I do not make this assumption.

Theorem 68. *Let $\mathfrak{O}$ be a valuation ring with unique maximal ideal $\mathfrak{p}$ and quotient field K. The multiplicative group $K^\times/\mathfrak{O}^\times$ can be ordered by the rule*

$$a\mathfrak{O}^\times \leqslant b\mathfrak{O}^\times \iff a/b \in \mathfrak{O}.$$

Say then that 0 is less than all elements of this group. Then the function $\xi \mapsto |\xi|_\mathfrak{p}$ from K to $\{0\} \cup K^\times/\mathfrak{O}^\times$ defined by

$$|\xi|_\mathfrak{p} = \begin{cases} \xi\mathfrak{O}^\times, & \text{if } \xi \neq 0, \\ 0, & \text{if } x = 0 \end{cases}$$

is a valuation.

Proof. If $b \neq 0$, then

$$|a+b|_\mathfrak{p} \leqslant |b|_\mathfrak{p} \iff \frac{a}{b} + 1 = \frac{a+b}{b} \in \mathfrak{O} \iff \frac{a}{b} \in \mathfrak{O} \iff |a|_\mathfrak{p} \leqslant |b|_\mathfrak{p}\,,$$

so the strong triangle inequality holds. □

The valuation in the theorem is the **$\mathfrak{p}$-adic valuation.** If R happens to be a **discrete valuation ring,** so that its maximal ideal is (π) for some element π of R, then the valuation is the **π-adic valuation,** denoted by $\xi \mapsto |\xi|_\pi$. In particular, by Theorem 67, $K((T))$ has the T-adic valuation.

Since a value group is written multiplicatively, we write 1 for the identity in a value group. Two valuations $x \mapsto |x|$ and $x \mapsto |x|'$ on a field are **equivalent** if

$$|x| < |y| \iff |x|' < |y|'\,,$$

which means the same thing as

$$|x| < 1 \iff |x|' < 1.$$

As a valuation can be obtained from a valuation ring, so a valuation ring can be recovered from a valuation:

Theorem 69. *Given a field K with valuation $x \mapsto |x|$, let*

$$\mathfrak{O} = \{x \in K\colon\ |x| \leqslant 1\}, \qquad \mathfrak{p} = \{x \in K\colon\ |x| < 1\}.$$

Then K is the quotient field of $\mathfrak{O}$, which is a valuation ring with maximal ideal $\mathfrak{p}$, and $x \mapsto |x|$ is equivalent to the $\mathfrak{p}$-adic valuation.

In the notation of the theorem, the pair $(K, \mathfrak{O})$ is a **valued field,** and the associated valuation on K is the $\mathfrak{p}$-adic valuation. Note that $\mathfrak{p} = \mathfrak{O} \smallsetminus \mathfrak{O}^{\times}$ by Theorem 65. The ordered group $K^{\times}/\mathfrak{O}^{\times}$ is the **value group,** and the field $\mathfrak{O}/\mathfrak{p}$ is the **residue field.** So $(K((T)), K[[T]])$ is a valued field with residue field (isomorphic to) K, by Theorem 67; the value group is infinite cyclic. In case K is an arbitrary ordered field extending $\mathbb{R}$, and $\mathfrak{O}$ is the ring of finite elements, then $(K, \mathfrak{O})$ is a valued field with residue field (isomorphic to) $\mathbb{R}$, by Theorem 66.

Theorem 70. *Given a valued field* $(K, \mathfrak{O})$ *with associated valuation be* $x \mapsto |x|$, *we have*

$$|1| = 1, \tag{†}$$

$$|-x| = |x|, \tag{‡}$$

$$|1 + \cdots + 1| \leqslant 1, \tag{§}$$

$$|x| < |y| \Rightarrow |x \pm y| = |y|. \tag{¶}$$

Proof. In the value group, 1 is $\mathfrak{O}^{\times}$. In the field, 1 and -1 belong to $\mathfrak{O}^{\times}$. This gives (†) and (‡). Then (§) is by definition of valuation. For (¶), since

$$|y| = |\pm y| = |x \pm y - x| \leqslant \max(|x \pm y|, |x|),$$

if $|x| < |y|$, we have $|y| \leqslant |x \pm y| \leqslant |y|$. □

Because of this theorem, we may refer to a non-constant valuation as a **non-archimedean** norm, even though the value group itself is usually archimedean. Then an absolute value function is an **archimedean** norm, at least when its codomain is the archimedean field $\mathbb{R}$. A **normed field** is then a field with a norm, archimedean or not.

5.2.3. Completions

In any normed field, there is the notion of **Cauchy sequence** and **convergent sequence:** the definitions are formally the same as in § 4.4 for sequences in $\mathbb{R}$. Convergent sequences are Cauchy sequences, but the converse may fail. For example, the sequence $(T^n \colon n \in N)$ in $K(T)$ converges to 0 with respect to the T-adic valuation, while the sequence $(\sum_{m=k}^{k+n} a_m T^m \colon n \in \mathbb{N})$ converges to $\sum_{m=k}^{\infty} a_m T^m$, which is in $K((T))$, but need not be in $K(T)$. Note that the values of the terms in the latter sequence are all the same, namely T^k (if $a_k \neq 0$). This illustrates a general rule:

Theorem 71. *In a valued field, if a Cauchy sequence* $(a_n \colon n \in \mathbb{N})$ *does not converge to* 0*, then the sequence* $(|a_n| : n \in \mathbb{N})$ *of values is eventually constant.*

Proof. Let $a = (a_n \colon n \in \mathbb{N})$. Since this does not converge to 0, for some non-zero value ε, for all positive integers M, there is an integer n such that $n > M$ and $|a_n| \geqslant \varepsilon$. Since a is Cauchy, for some positive integer N, if $m > N$ and $n > N$, then $|a_m - a_n| < \varepsilon$. But then there is k such that $k > N$ and $|a_k| \geqslant \varepsilon$. In this case, if $m > N$, then

$$|a_m - a_k| < \varepsilon \leqslant |a_k|,$$

so $|a_m| = |a_m - a_k + a_k| = |a_k|$ by Theorem 70. □

A normed field is **complete** with respect to its norm if every Cauchy sequence of its elements converges.

Theorem 72. *The following normed fields are complete:*
(1) $K((T))$ *with the* T*-adic valuation;*
(2) $\mathbb{R}$ *and* $\mathbb{C}$ *with the absolute value function;*
(3) $\mathbb{R}((T))$ *with the absolute value function induced by letting* T *be infinitesimal.*

A set is **countable** if it is the range of a function on $\mathbb{N}$. So $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are countable, but $\mathbb{R}$ is not.

Theorem 73. *Suppose* K *is a normed field such that every countable set of values has a positive lower bound. Then* K *is complete with respect to the norm.*

Proof. In such a field, every Cauchy sequence is eventually constant, so it already converges. Indeed, let a be a Cauchy sequence in such a field, and let A be the set of all *positive* values $|a_n - a_m|$. Then A has a positive lower bound ε. Let L be such that, when $m \geqslant L$ and $n \geqslant L$, then $|a_n - a_m| < \varepsilon$. Then, for such m and n, we must have $a_n = a_m$. □

An example of a normed field as in the theorem can be contrived by means of the theory of **ordinal numbers.** The ordinal numbers, or just **ordinals,** compose a well-ordered class, **ON**. Then **ON** is also an iterative structure: the least ordinal is denoted by

$$0,$$

and for every ordinal α, its successor, $\alpha + 1$, is the least ordinal that is greater than α. There are ordinals that are neither successors nor 0; these are **limit**

ordinals. The ordinals can be constructed so that each ordinal α is actually the set of ordinals that are less than α. Then

$$0 = \varnothing, \qquad \alpha + 1 = \alpha \cup \{\alpha\};$$

also, the least infinite ordinal is denoted by one of

$$\omega, \qquad \aleph_0.$$

We shall define ω in more detail in 6.1.1. Meanwhile, the least *uncountable* ordinal is denoted by

$$\aleph_1.$$

Then every countable subset of $\aleph_1$ has an upper bound in $\aleph_1$, since if A is such a subset, then $\bigcup A$ is a countable ordinal and an upper bound of A.

Given an ordered field K, we can define an extension $K(T_\alpha\colon \alpha < \aleph_1)$ by **transfinite recursion:**

1. $K(T_\alpha\colon \alpha < 0) = K$,
2. $K(T_\alpha\colon \alpha < \beta + 1) = K(T_\alpha\colon \alpha < \beta)(T_\beta)$,
3. $K(T_\alpha\colon \alpha < \gamma) = \bigcup_{\beta<\gamma} K(T_\alpha\colon \alpha < \beta)$, if β is a limit.

We order $K(T_\alpha\colon \alpha < \aleph_1)$ by letting each T_α be positive but less than each positive element of K, and also

$$T_\beta < T_\alpha \iff \alpha < \beta.$$

Then every countable set of the T_α has a positive lower bound, so $K(T_\alpha\colon \alpha < \aleph_1)$ is complete with respect to the absolute value function, by Theorem 73.

An **embedding** of a valued field $(K, \mathfrak{O})$ in a valued field $(K_1, \mathfrak{O}_1)$ is an embedding φ of K in K_1 such that $\varphi[\mathfrak{O}] = \varphi[K] \cap \mathfrak{O}_1$.

Theorem 74. *An embedding of valued fields induces embeddings of the value groups and the residue fields.*

Proof. Let φ be an embedding of the valued field $(K, \mathfrak{O})$ in the valued field $(K_1, \mathfrak{O}_1)$. We may assume that φ is just an inclusion, so

$$K \subseteq K_1, \qquad \mathfrak{O} = K \cap \mathfrak{O}_1.$$

Since $\mathfrak{O} \subseteq \mathfrak{O}_1$, we have also $\mathfrak{O}^\times \subseteq \mathfrak{O}_1^\times$, so the function $x\mathfrak{O}^\times \mapsto x\mathfrak{O}_1^\times$ is a well-defined homomorphism from the value group $K^\times/\mathfrak{O}^\times$ to the value group $K_1^\times/\mathfrak{O}_1^\times$. Suppose $x\mathfrak{O}^\times$ is in the kernel. Then $x \in K^\times \cap \mathfrak{O}_1^\times$, and also $x^{-1} \in$

$K^\times \cap \mathfrak{O}_1^\times$, so $x^{-1} \in \mathfrak{O}$ and hence $x \in \mathfrak{O}^\times$. Therefore the homomorphism of value groups is an embedding.

Let $\mathfrak{p} = \mathfrak{O} \smallsetminus \mathfrak{O}^\times$ and $\mathfrak{p}_1 = \mathfrak{O}_1 \smallsetminus \mathfrak{O}_1^\times$. If $x \in \mathfrak{p}$, then either $x = 0$ or $x^{-1} \in K \smallsetminus \mathfrak{O}$. In the latter case, $x^{-1} \notin \mathfrak{O}_1$; hence in either case, $x \notin \mathfrak{O}_1^\times$, so $x \in \mathfrak{p}_1$. Therefore the function $x + \mathfrak{p} \mapsto x + \mathfrak{p}_1$ is a well-defined homomorphism from the residue field $\mathfrak{O}/\mathfrak{p}$ to the residue field $\mathfrak{O}_1/\mathfrak{p}_1$. Since the homomorphism takes $1 + \mathfrak{p}$ to $1 + \mathfrak{p}_1$, which is not 0, the homomorphism must be an embedding. □

A **completion** of a valued field $(K, \mathfrak{O})$ is a complete valued field $(\overline{K}, \overline{\mathfrak{O}})$, together with an embedding ψ of the former in the latter, such that, whenever φ is an embedding of $(K, \mathfrak{O})$ in a complete valued field $(K_1, \mathfrak{O}_1)$, then there is an embedding $\overline{\varphi}$ of $(\overline{K}, \overline{\mathfrak{O}})$ in $(K_1, \mathfrak{O}_1)$ such that $\overline{\varphi} \circ \psi = \varphi$. The situation is as in Figure 5.1, which is analogous to Fig. 4.3.

$$\begin{tikzcd}
(K, \mathfrak{O}) \arrow[r, "\psi"] \arrow[d, "\varphi"'] & (\overline{K}, \overline{\mathfrak{O}}) \arrow[dl, "\overline{\varphi}"] \\
(K_1, \mathfrak{O}_1) &
\end{tikzcd}$$

Figure 5.1. Completion of a valued field

Theorem 75. *Every valued field has a completion, and the corresponding embeddings of value groups and residue fields are isomorphisms.*

Proof. Let $(K, \mathfrak{O})$ be a valued field, let R consist of its Cauchy sequences, and let I consist of those sequences that converge to 0. Then R is a ring with ideal I. Moreover, I is a *maximal* ideal of R: that is, if $a \in R \smallsetminus I$, then there are elements b of R and c of I such that

$$ab + c = 1. \tag{$\|$}$$

Indeed, suppose a is an element $(a_n \colon n \in \mathbb{N})$ of $R \smallsetminus I$. By Theorem 71, the sequence of values of the a_n is eventually some nonzero constant $|A|$. Hence, if m and n are large enough, then $a_m a_n \neq 0$ and

$$\left|\frac{1}{a_n} - \frac{1}{a_m}\right| = \left|\frac{a_m - a_n}{a_n a_m}\right| = \frac{|a_m - a_n|}{|A|^2}.$$

So we have (‖) as desired when

$$b_n = \begin{cases} {a_n}^{-1}, & \text{if } a_n \neq 0,\\ 0, & \text{if } a_n = 0, \end{cases} \qquad c_n = \begin{cases} 0, & \text{if } a_n \neq 0,\\ 1, & \text{if } a_n = 0. \end{cases}$$

Now R/I is a field $\overline{K}$, in which K embeds under the map $x \mapsto (x, x, \dots) + I$. Whenever a and b are in R, and $a + I = b + I$, then either $|a_n| = |b_n|$ when n is large enough, or else both a and b are in I. For, suppose $a \notin I$. Then again by Theorem 71, for some A in $K^\times \smallsetminus \mathfrak{O}^\times$, if n is large enough, then $|a_n| = |A|$. Since $b - a \in I$, if n is large enough, we have both $|b_n - a_n| < |A|$ and also

$$|b_n| = |a_n + b_n - a_n| = |a_n| = |A|\,.$$

Now we can define a valuation on $\overline{K}$ by letting $|a + I| = |A|$, if a eventually takes the value $|A|$, and $|I| = 0$. Letting $\overline{\mathfrak{O}}$ comprise those $a + I$ in $\overline{K}$ such that $|a + I| \leqslant 1$, we have a valued field $(\overline{K}, \overline{\mathfrak{O}})$ by Theorem 69, and, identifying K with its image in $\overline{K}$, we have $K \cap \overline{\mathfrak{O}} = \mathfrak{O}$, so $(K, \mathfrak{O}) \subseteq (\overline{K}, \overline{\mathfrak{O}})$. Immediately the embedding of value groups is surjective. Now let $\mathfrak{p} = \mathfrak{O} \smallsetminus \mathfrak{O}^\times$, and $\overline{\mathfrak{p}} = \overline{\mathfrak{O}} \smallsetminus \overline{\mathfrak{O}}^\times$. If $a + I \in \overline{K}^\times$, and m and n are large enough, we have $|a_m - a_n| < 1$, that is, $a_m - a_n \in \mathfrak{p}$, so $(a + I) + \overline{\mathfrak{p}} = (a_n + I) + \overline{\mathfrak{p}}$. Thus the embedding of residue fields is surjective.

By Theorem 73, we may assume there is a decreasing sequence $(\delta_n \colon n \in \mathbb{N})$ of values such that, for all values ε, there is n such that $\delta_n \leqslant \varepsilon$. Suppose $(a^n + I \colon n \in \mathbb{N})$ is a Cauchy sequence of elements of $\overline{K}$. Then there is a strictly increasing function f from $\mathbb{N}$ to itself such that, if $k \geqslant f(n)$ and $\ell \geqslant f(n)$, then

$$\left|a^k - a^\ell + I\right| < \delta_n.$$

Then there is a strictly increasing function g from $\mathbb{N}$ to itself such that, if

$$f(n) \leqslant j \leqslant f(n+1), \qquad f(n) \leqslant k \leqslant f(n+1), \qquad \ell \geqslant g(n), \qquad m \geqslant g(n),$$

then

$$|a^j_\ell - a^k_m| \leqslant \max(|a^j_\ell - a^k_\ell|, |a^k_\ell - a^k_m|) < \delta_n.$$

Now define a sequence b of elements of K by

$$b_m = \begin{cases} a^1_m, & \text{if } m < g(1);\\ a^{f(n)}_m, & \text{if } g(n) \leqslant m < g(n+1). \end{cases}$$

Then $b \in R$, since, if $g(n) \leqslant m < g(n+1)$ and $g(n+i-1) \leqslant m+j < g(n+i)$, then

$$\begin{aligned}
&|b_{m+j} - b_m| \\
&\leqslant \max(|b_{m+j} - b_{g(n+i)}|, |b_{g(n+i)} - b_{g(n+i-1)}|, \dots, |b_{g(n+1)} - b_{g(n)}|, |b_{g(n)} - b_m|) \\
&\leqslant \max(\delta_{n+i-1}, \delta_{n+i-1}, \dots, \delta_n, \delta_n) \\
&= \delta_n.
\end{aligned}$$

Moreover, $(a^n + I \colon n \in \mathbb{N})$ converges to $b + I$, since if $f(n) \leqslant k < f(n+1)$ and $g(n+i-1) \leqslant m < g(n+i)$, then

$$\begin{aligned}
|a^k_m - b_m| &= |a^k_m - a^{f(n+i-1)}_m| \\
&\leqslant \max(|a^k_m - a^{f(n)}_m|, |a^{f(n)}_m - a^{f(n-1)}_m|, \dots, |a^{f(n+i-2)}_m - a^{f(n+i-1)}_m|) \\
&\leqslant \max(\delta_n, \delta_n, \dots, \delta_{n+i-2}) \\
&= \delta_n.
\end{aligned}$$

Therefore $(\overline{K}, \overline{\mathfrak{O}})$ is complete. It is a completion of $(K, \mathfrak{O})$, because we can extend an inclusion of $(K, \mathfrak{O})$ in a complete valued field $(K_1, \mathfrak{O}_1)$ to an embedding of $(\overline{K}, \overline{\mathfrak{O}})$ by sending $(a_n \colon n \in \mathbb{N}) + I$ to the limit of $(a_n \colon n \in \mathbb{N})$ in K_1. □

For example, $K((T))$ is the completion of $K(T)$ with respect to the T-adic valuation.

If p is a prime number, let $\mathbb{Z}_{(p)}$ be the **localization** of $\mathbb{Z}$ at (p): this is the sub-ring of $\mathbb{Q}$ comprising those a/b such that b is not a multiple of p. Then $\mathbb{Z}_{(p)}$ is a valuation ring with maximal ideal generated by p, and $(\mathbb{Q}, \mathbb{Z}_{(p)})$ is a valued field. Usually the corresponding ***p*-adic valuation** on $\mathbb{Q}$ is given by

$$\left|p^n \cdot \frac{a}{b}\right|_p = \frac{1}{p^n},$$

where a and b are indivisible by p; in particular, the value group is taken as a subgroup of $\mathbb{Q}^+$ with the usual ordering. The completion of $(\mathbb{Q}, \mathbb{Z}_{(p)})$ is denoted by $(\mathbb{Q}_p, \mathbb{Z}_p)$. Here, each element of $\mathbb{Q}_p$ is called a ***p*-adic number** and is a formal sum

$$\sum_{n=k}^{\infty} a_n p^n,$$

where $k \in \mathbb{Z}$ and $a_n \in \{0, 1, \dots, p-1\}$; the p-adic value of this element is $1/p^k$, assuming $a_k \neq 0$. We have for example

$$-1 = \sum_{n=0}^{\infty} (p-1)p^k.$$

Elements of $\mathbb{Z}_p$ are **p-adic integers.** Even though $\mathbb{Q}_p$ is of characteristic 0, its residue field is finite, with p elements.

A **completion** of an ordered field K with respect to the absolute value function now has the obvious definition: it is an ordered field $\overline{K}$ such that

1. K embeds in $\overline{K}$,
2. $\overline{K}$ is complete with respect to its absolute value function,
3. every embedding of K in an ordered field that is complete with respect to its absolute value function extends to an embedding of $\overline{K}$ in that ordered field.

Theorem 76. *Every ordered field has a completion with respect to its absolute value function.*

Proof. Follow the proof of the last theorem. In the proof that I is a maximal ideal, the sequence of values of a need not be eventually constant; but for some nonzero value ε, if m and n are large enough, then $|a_n|, |a_n| \geqslant \varepsilon$, so

$$\left|\frac{1}{a_n} - \frac{1}{a_m}\right| \leqslant \frac{|a_m - a_n|}{\varepsilon^2}.$$

So I is still a maximal ideal, so R/I is a field $\overline{K}$. This can be ordered, since every element of $R \smallsetminus I$ is eventually either positive or negative. To prove completeness, we may assume that the sequence $(\delta_n \colon n \in \mathbb{N})$ satisfies also $2\delta_{n+1} \leqslant \delta_n$, so that

$$\delta_{n+1} + \delta_{n+2} + \cdots \leqslant \delta_n. \qquad \square$$

For example, $\mathbb{R}((T))$ is the completion of $\mathbb{R}(T)$ with respect to the absolute value function induced by the ordering in which T is infinitesimal.

In sum:

1. The field $\mathbb{R}$ is the unique complete ordered field.
2. The field $\mathbb{R}$ is complete with respect to the absolute value function, but so too is $\mathbb{C}$.
3. There are non-archimedean ordered fields. The completion of one of these with respect to the ordering is never a field. But there is a completion with respect to the absolute value function determined by the ordering, and this completion is always a field.

6. Ultrapowers and logic

6.1. Algebraic preliminaries

6.1.1. Natural numbers constructed

The simplest set is the empty set, $\varnothing$. One of the simplest things we can do with a set x (aside from doing nothing at all) is to make a new set, $\{x\}$, with only x as a member. Then we can form the set $x \cup \{x\}$, comprising the elements of x, along with x itself. Hence one of the simplest nontrivial functions defined recursively on $\mathbb{N}$ is the function f given by[*]

$$f(1) = \varnothing, \qquad\qquad f(n+1) = f(n) \cup \{f(n)\}. \qquad (*)$$

Theorem 77. *The function f on $\mathbb{N}$ defined as in $(*)$ is injective.*

Proof. By induction, each set $f(n)$ includes its elements, that is, if $x \in f(n)$, then $x \subseteq f(n)$.

Hence $f(n+1) \nsubseteq f(n)$. For, this is trivially true when $n = 1$. Suppose it is false when $n = m+1$, that is, $f(m+2) \subseteq f(m+1)$. But $f(m+2) = f(m+1) \cup \{f(m+1)\}$, so $f(m+1) \in f(m+1)$. Thus either $f(m+1) = f(m)$ or $f(m+1) \in f(m)$, and in either case, $f(m+1) \subseteq f(m)$; hence the claim fails when $n = m$. By induction, the claim is true for all n in $\mathbb{N}$.

Also by induction, if $k \leqslant n$, then $f(k) \subseteq f(n)$.

Suppose if possible that $f(k) = f(m)$, although $k < m$. But then $k \leqslant m - 1$, so $f(m) \subseteq f(m-1)$, contrary to what we have just shown. Therefore f is injective. □

The image $f[\mathbb{N}]$ of $\mathbb{N}$ under f is denoted by

$$\omega.$$

Writing x' for $x \cup \{x\}$, and 0 for $\varnothing$, we have that $(\omega, 0, ')$ is isomorphic to $(\mathbb{N}, 1, \mathrm{S})$. It will be convenient to treat ω as $\mathbb{N}$, although now 1 will denote $0'$,

[*] Here f is the unique homomorphism from $(\mathbb{N}, 1, \mathrm{S})$ into $(\mathbf{V}, \varnothing, x \mapsto x \cup \{x\})$, where $\mathbf{V}$ is the universe of all sets. Now, $\mathbf{V}$ is not a set itself, but is a *proper class,* since, by the Russell Paradox, the subclass $\{x \colon x \notin x\}$ of $\mathbf{V}$ cannot be a set. But the proof of the Recursion Theorem (Theorem 3) does not actually require A to be a set.

that is, $\{\varnothing\}$, and n' will be $n + 1$. With this understanding, we still define addition on ω by (†) in § 2.2. The ordering of ω is induced from $\mathbb{N}$ by f, that is, $f(k) < f(n) \iff k < n$.

Theorem 78. *If $n \in \omega$, then*

$$n = \{0, \dots, n - 1\};$$

that is, $k < n \iff k \in n$.

Proof. With f as in $(*)$, by induction, if $k < n$, then $f(k) \in f(n)$. Also by induction, if $x \in f(n)$, then $x = f(k)$ for some k such that $k < n$. Since f is injective by Theorem 77, if $f(k) \in f(n)$, then $k < n$. □

This, for us, is the point of defining ω. Given a natural number n, we shall want a set with n elements; now n itself is such a set.

6.1.2. Powers of sets

Let Ω be a set, and $n \in \omega$. We can now define

$$\Omega^n$$

as the set of functions from n to Ω; this is the nth **Cartesian power** of Ω. The definition is consistent with that of $\mathbb{Q}^{\mathbb{N}}$ in § 4.4. Since 0 is empty, Ω^0 consists of the empty set; but $\{\varnothing\} = \{0\} = 1$, so

$$\Omega^0 = 1.$$

A typical element of Ω^n might be denoted by

$$(x^0, \dots, x^{n-1})$$

or more simply

$$\boldsymbol{x};$$

then x^k is the kth **coordinate** of $\boldsymbol{x}$.

A subset of Ω^n is an n**-ary relation** on Ω. On any set, there are just two 0-ary relations, namely $\varnothing$ and $\{\varnothing\}$, that is, 0 and 1. A function from Ω^n to Ω is an n**-ary operation** on Ω. In particular, a 0-ary operation on Ω is a **constant** and can be identified with an element of Ω. Also, an n-ary operation can be identified with an $n + 1$-ary relation.

The set of functions from ω into Ω is denoted by

$$\Omega^{\omega}.$$

Given an indexed family $(\Omega_k \colon k \in \omega)$ of sets, we can form the product

$$\prod_{k\in\omega} \Omega_k;$$

this is the subset of $(\bigcup_{k\in\omega} \Omega_k)^{\omega}$ comprising those f such that $f(k) \in \Omega_k$ in each case. A **relation** on the family $(\Omega_k \colon k \in \omega)$ is a subset of some finite product $\prod_{k<n} \Omega_{g(k)}$, where $n \in \omega$ and $g \in \omega^n$. Similarly, an operation on the family is a function from some $\prod_{k<n} \Omega_{g(k)}$ into some $\Omega_{g(n)}$; this can be understood as a relation on $\prod_{k\leqslant n} \Omega_{g(k)}$.

Some *structures* were identified in § 2.2. In the most general sense, a **structure** is a set, or even an indexed family of sets, with some operations and relations on it. The most common structure based on more than one set is perhaps the *vector space,* which will come up in Ch. 7. We shall identify another example presently.

We have seen in Ch. 2 that, given the structure $(\mathbb{N}, 1, \mathrm{S})$, we can define new operations and relations, such as $+$, $\cdot$, and $<$. The **full structure** on the set Ω is just Ω together with *every* n-ary relation on Ω, for *every* n in ω. That is, the full structure on Ω is Ω together with the elements of all of the sets $\mathscr{P}(\Omega^n)$. We shall now consider some operations on the indexed family $(\mathscr{P}(\Omega^n) \colon n \in \omega)$.

The **diagonal** on Ω is the element Δ_Ω of $\mathscr{P}(\Omega^2)$ given by

$$\Delta_\Omega = \{(x, x) \colon x \in \Omega\}.$$

The set $\mathscr{P}(\Omega)$ (and hence each of the sets $\mathscr{P}(\Omega^n)$) can be equipped with the **Boolean operations,** namely

$$(X, Y) \mapsto X \cap Y, \qquad\qquad (X, Y) \mapsto X \smallsetminus Y$$

and their compositions, possibly involving the constants Ω and $\varnothing$. So for example we have

$$X^{\mathrm{c}} = \Omega \smallsetminus X, \qquad\qquad X \cup Y = (X^{\mathrm{c}} \cap Y^{\mathrm{c}})^{\mathrm{c}}.$$

Of particular interest is the operation $(X, Y) \mapsto X \vartriangle Y$, where

$$X \vartriangle Y = (X \smallsetminus Y) \cup (Y \smallsetminus X);$$

this is the **symmetric difference** of X and Y. The operation is commutative; also,

$$X \mathbin{\triangle} X = \varnothing.$$

In order to move between different powers of Ω, suppose m and n are in ω, and

$$f\colon m \to n.$$

If $n = 0$, then m must be 0. In any case, a function from Ω^n to Ω^m is induced, namely

$$\boldsymbol{x} \mapsto \boldsymbol{x} \circ f.$$

Let us denote this function by

$$f^*.$$

Then*

$$f^*(x^0, \dots, x^{n-1}) = (x^{f(0)}, \dots, x^{f(m-1)}).$$

An important example arises when f is the inclusion of m in $m+1$. In this case, $f^*(x^0, \dots, x^{m-1}, x^m) = (x^0, \dots, x^{m-1})$, or more simply

$$f^*(\boldsymbol{x}, y) = \boldsymbol{x};$$

that is, f^* is **projection** onto the first m coordinates.

In general, the function f from Ω^n to Ω^m induces two new functions. First we have the function $X \mapsto f^*[X]$ from $\mathscr{P}(\Omega^n)$ to $\mathscr{P}(\Omega^m)$, given by

$$f^*[A] = \{f^*(\boldsymbol{x})\colon \boldsymbol{x} \in A\}.$$

When f is the inclusion of m in $m+1$, then

$$f^*[A] = \{\boldsymbol{x}\colon \exists y\ (\boldsymbol{x}, y) \in A\}.$$

So applying f^* in this case corresponds to applying $\exists$, the **existential quantifier.** If f is just a permutation of m, then f^* is a permutation of coordinates.

In the general case, we also have a function $Y \mapsto f_*(Y)$ from $\mathscr{P}(\Omega^m)$ to $\mathscr{P}(\Omega^n)$, given by

$$f_*(B) = (f^*)^{-1}[B].$$

If again f is the inclusion of m in $m+1$, then

$$f_*(B) = \{(\boldsymbol{x}, y)\colon \boldsymbol{x} \in B \mathbin{\&} y \in \Omega\},$$

*In the language of category theory, the pair $(m \mapsto \Omega^m, f \mapsto f^*)$ is a *contravariant functor* from the category $(\omega, \{\text{functions}\})$ to the category $(\{\text{sets}\}, \{\text{functions}\})$.

which can be identified with $B \times \Omega$. Also, if f is a permutation of m, then

$$f^*[A] = (f^{-1})_*(A).$$

When we equip the family $(\mathscr{P}(\Omega^n)\colon n \in \omega)$ with the diagonal, the Boolean operations, and the various operations f_* and f^*, let us denote the resulting structure by

$$\mathscr{D}(\Omega).$$

The point of the notation is to simplify the statement of Theorem 86 on p. 91 below.

6.1.3. Boolean rings

So that the *proof* of Theorem 86 makes sense, suppose R is a commutative ring. As in § 4.4, we obtain the commutative ring

$$R^{\omega}$$

of sequences of elements of R (only now the sequences are $(x_0, x_1, x_2, \dots)$ rather than $(x_1, x_2, x_3, \dots)$). If a is an element $(a_n\colon n \in \omega)$ of R^{ω}, let

$$\operatorname{supp}(a) = \{n \in \omega\colon a_n \neq 0\};$$

this is the **support** of a. In one case of interest, R is $\mathbb{B}$, where

$$\mathbb{B} = \{0, 1\},$$

considered as a two-element field.

Theorem 79. *The map $x \mapsto \operatorname{supp}(x)$ on $\mathbb{B}^{\omega}$ is a bijection onto $\mathscr{P}(\omega)$. Also*

$$\begin{gathered}\operatorname{supp}(0) = \varnothing,\\ \operatorname{supp}(1) = \omega,\\ \operatorname{supp}(xy) = \operatorname{supp}(x) \cap \operatorname{supp}(y),\\ \operatorname{supp}(x + y) = \operatorname{supp}(x) \vartriangle \operatorname{supp}(y),\end{gathered}$$

Thus $(\mathscr{P}(\omega), \varnothing, \omega, \vartriangle, \cap)$ is a commutative unital ring in which each element is its own additive inverse.

A ring is called **Boolean** if in it

$$x^2 = x. \tag{†}$$

So $\mathbb{B}$, $\mathbb{B}^{\omega}$, and $\mathscr{P}(\omega)$ are Boolean rings.

Theorem 80. *Let R be a Boolean ring. In this ring,*

$$2x = 0, \tag{‡}$$

and hence

$$-x = x.$$

Also R is commutative, and R can be partially ordered by the rule

$$x \leqslant y \iff xy = x.$$

Then a nonempty subset I of R is an ideal of R if and only if

$$\begin{aligned} x \in I \mathbin{\&} y \in I &\implies x + y \in I, && (§)\\ x \in I \mathbin{\&} y \leqslant x &\implies y \in I. && (¶)\end{aligned}$$

All prime ideals of R are maximal, and and ideal I is maximal if and only if

$$x \in I \iff x + 1 \notin I.$$

Proof. For (‡), compute

$$2x = (2x)^2 = 4x^2 = 4x.$$

For commutativity then, compute

$$\begin{gathered} x + y = (x+y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y,\\ 0 = xy + yx. \end{gathered}$$

Immediately from the definitions, $x \leqslant x$. If $x \leqslant y$ and $y \leqslant x$, then $x = xy = yx = y$. If $x \leqslant y$ and $y \leqslant z$, then $xz = xyz = xy = x$, so $x \leqslant z$. Thus $\leqslant$ partially orders R.

For the characterization of ideals, note that (¶) is equivalent to $x \in I \implies xz \in I$.

From (†), we get

$$x(x-1) = 0,$$

so in every Boolean integral domain, the only elements are 0 and 1. In short, every Boolean integral domain is a field, so prime ideals of R are maximal. Moreover, an ideal I of R is maximal if and only if R/I is the disjoint union of two cosets, I and $1 + I$; this yields the characterization of maximal ideals. □

Theorem 81. *A subset I of $\mathscr{P}(\omega)$ is an ideal if and only if*

$$x \in I \mathbin{\&} y \in I \implies x \cup y \in I, \tag{$\|$}$$
$$x \in I \mathbin{\&} y \subseteq x \implies y \in I; \tag{$**$}$$

an ideal M of $\mathscr{P}(\omega)$ is maximal if and only if

$$x \in M \iff \omega \smallsetminus x \notin M.$$

A principal ideal (A) of $\mathscr{P}(\omega)$ is maximal if and only if $A = \omega \smallsetminus \{n\}$ for some n in ω. A maximal ideal of $\mathscr{P}(\omega)$ is non-principal if and only if it contains all finite subsets of ω.

Proof. The conditions ($\|$) and ($**$) are equivalent to (§) and (¶) since, in $\mathscr{P}(\omega)$,

$$\begin{gathered} y \subseteq x \iff y \cap x = y \iff y \leqslant x, \\ x \cup y = x + y + xy, \\ x + y \subseteq x \cup y. \end{gathered}$$ □

Theorem 82. *Let K be a field. The function $X \mapsto \operatorname{supp}[X]$ gives a one-to-one correspondence between the ideals of K^ω and the ideals of $\mathscr{P}(\omega)$.*

Proof. If $X \subseteq \omega$, let $u(X)$ be the element of K^ω defined by

$$u(X)_n = \begin{cases} 1, & \text{if } n \in X, \\ 0, & \text{if } n \notin X. \end{cases}$$

Then

$$X = \operatorname{supp}(u(X)).$$

Suppose I is an ideal of K^ω, and $a \in I$. Then $u(\operatorname{supp}(a)) \in I$. If $b \in K^\omega$, and $\operatorname{supp}(b) = \operatorname{supp}(a)$, then $b = bu(\operatorname{supp}(a))$, so $b \in I$. This shows $\operatorname{supp}^{-1}[\operatorname{supp}[I]] = I$. Therefore $X \mapsto \operatorname{supp}[X]$ is injective on the set of ideals of K^ω.

If $X \subseteq \operatorname{supp}(a)$, then $X = \operatorname{supp}(au(X))$. Also,

$$\operatorname{supp}(a) \mathbin{\triangle} \operatorname{supp}(b) \subseteq \operatorname{supp}(a+b).$$

This shows $\operatorname{supp}[I]$ is an ideal of $\mathscr{P}(\omega)$.

Finally, if J is an ideal of $\mathscr{P}(\omega)$, then $\operatorname{supp}^{-1}[J]$ is an ideal of K^ω by Theorem 81, since

$$\operatorname{supp}(a+b) \subseteq \operatorname{supp}(a) \cup \operatorname{supp}(b).$$ □

6.2. Ultrapowers

Throughout this section, $\mathfrak{m}$ is a maximal ideal of $\mathscr{P}(\omega)$. For now, K is a field.* Then by Theorem 82, $\operatorname{supp}^{-1}[\mathfrak{m}]$ is a maximal ideal of K^ω; for ease of writing and reading, let us denote this ideal also by $\mathfrak{m}$. Then we can form the quotient

$$K^\omega/\mathfrak{m},$$

which must be a field; it is called an **ultrapower** of K. If a and b are in K^ω, and $a+\mathfrak{m}=b+\mathfrak{m}$, then a and b are **congruent** *modulo* $\mathfrak{m}$, and we may write

$$a \equiv b$$

or more precisely $a \equiv b \pmod{\mathfrak{m}}$.

The elements of the ideal $\mathfrak{m}$ of $\mathscr{P}(\omega)$ should be considered as **small** subsets of ω; every other subset is **large.** This means, by Theorem 81:

(1) a subset of a small set is small,
(2) the union of two small sets is small, and
(3) a set is small if and only if its complement is large.

In particular, the empty set is small, but ω itself is large.

Theorem 83. *If $a, b \in K^\omega$, then*

$$\begin{aligned} a \equiv b &\iff \{n \in \omega\colon a_n \neq b_n\} \in \mathfrak{m} \\ &\iff \{n \in \omega\colon a_n = b_n\} \notin \mathfrak{m}. \end{aligned}$$

If a and b are constant as functions on ω, then $a \equiv b$ if and only if $a = b$.

In other words, two elements of K^ω are congruent if and only if the set of indices where they differ is small, that is, the set of indices where they agree is large; in particular, the **diagonal map**

$$x \mapsto (x, x, \dots) + \mathfrak{m}$$

is an embedding of K in $K^\omega/\mathfrak{m}$. Let us identify K with its image in $K^\omega/\mathfrak{m}$, so that we may write

$$K \subseteq K^\omega/\mathfrak{m}. \tag{$*$}$$

The theorem shows how congruence can be defined independently of the field structure of K. We shall usually be interested only in the case where K is a

[*] The approach here is inspired by a lecture of Angus Macintyre [22].

field, and in particular K is the complete ordered field $\mathbb{R}$. However, § 6.4 will consider a different structure, albeit one derived from a field. In any case, the question of whether the inclusion in $(*)$ is proper is settled by means of the next theorem below.

Lemma. *If $\omega = X_0 \cup \dots \cup X_n$, then one of the sets X_k is large.*

Proof. By induction, if each set X_i is small, then so is $X_0 \cup \dots \cup X_{n-1}$. □

Theorem 84. *The inclusion $(*)$ of K in $K^\omega/\mathfrak{m}$ is proper if and only if K is infinite and the maximal ideal $\mathfrak{m}$ of K^ω is not principal.*

Proof. If K is finite, and $a \in K^\omega$, then, by the lemma, the set $\{k\colon a_k = b\}$ is large for some b in K; but then $a \equiv (b, b, \dots)$.

In case K is infinite, there is (by the Axiom of Choice) an element $(a_0, a_1, \dots)$ of K^ω such that $a_j \neq a_k$ whenever $j < k$. If $(a_0, a_1, \dots) \equiv (b, b, \dots)$, then $a_n = b$ for some n, and $\{n\}$ is a large subset of ω, so $\omega \smallsetminus \{n\}$ is small; but in this case every subset of ω that does not contain n is small, so $\mathfrak{m}$ is generated by $\omega \smallsetminus \{n\}$. Conversely, by Theorem 81, if $\mathfrak{m}$ is principal, then $\mathfrak{m} = (\omega \smallsetminus \{n\})$ for some n in ω, and then $(a_0, a_1, \dots) \equiv (a_n, a_n, \dots)$. □

We assume henceforth that K is infinite and $\mathfrak{m}$ is not principal. We shall show that $K^\omega/\mathfrak{m}$ still has the features of K; to be precise, $\mathscr{D}(K)$ (as defined in 6.1.2) embeds in $\mathscr{D}(K^\omega/\mathfrak{m})$. Then an element of $K^\omega/\mathfrak{m}$ that is not in K will have 'generic' or 'ideal' properties that no one element of K has; for example, in case K is $\mathbb{R}$, and $a \in K$, then there will be elements b of $K^\omega/\mathfrak{m} \smallsetminus K$ that are 'absolutely' close to a in the sense that no elements of K lie between a and b.

To work this all out, some notational conventions will be useful. There is a bijection

$$\bigl((x_k^0\colon k \in \omega), \dots, (x_k^{n-1}\colon k \in \omega)\bigr) \mapsto \bigl((x_k^0, \dots, x_k^{n-1})\colon k \in \omega\bigr)$$

from $(K^\omega)^n$ onto $(K^n)^\omega$; we may write the bijection more simply as

$$(x^0, \dots, x^{n-1}) \mapsto (\boldsymbol{x}_k\colon k \in \omega).$$

So a plainface x^j may denote $(x_k^j\colon k \in \omega)$ in K^ω (and x may denote $(x_k\colon k \in \omega)$); while a boldface $\boldsymbol{x}_k$ denotes $(x_k^0, \dots, x_k^{n-1})$ in K^n for some n in ω (and $\boldsymbol{x}$ denotes $(x^0, \dots, x^{n-1})$). Now write

$$^*K = K^\omega/\mathfrak{m}. \tag{†}$$

There is a well-defined isomorphism

$$(x^0+\mathfrak{m},\dots,x^{n-1}+\mathfrak{m})\mapsto(x^0,\dots,x^{n-1})+\mathfrak{m}^n$$

from $({}^*K)^n$, that is, $(K^\omega/\mathfrak{m})^n$, onto $(K^\omega)^n/\mathfrak{m}^n$. We may treat this isomorphism as an identity and write $\mathfrak{m}$ for $\mathfrak{m}^n$, so that

$$(x^0+\mathfrak{m},\dots,x^{n-1}+\mathfrak{m})=(x^0,\dots,x^{n-1})+\mathfrak{m}.$$

Instead of $x^0\equiv y^0\ \&\ \cdots\ \&\ x^{n-1}\equiv y^{n-1}$, we may write

$$(x^0,\dots,x^{n-1})\equiv(y^0,\dots,y^{n-1}).$$

Then there is an analogue of Theorem 83:

Theorem 85. *If* $(a^0,\dots,a^{n-1})$ *and* $(b^0,\dots,b^{n-1})$ *are in* $(K^\omega)^n$*, then*

$$(a^0,\dots,a^{n-1})\equiv(b^0,\dots,b^{n-1})\iff\{n\in\omega\colon \boldsymbol{a}_n\neq\boldsymbol{b}_n\}\in\mathfrak{m}.$$

For each n in ω, there is a function $S\mapsto{}^*S$ from $\mathscr{P}(K^n)$ to $\mathscr{P}(({}^*K)^n)$ given by

$${}^*S=\{(x^0,\dots,x^{n-1})+\mathfrak{m}\colon(\boldsymbol{x}_k\colon k\in\omega)\in S^\omega\}.\tag{‡}$$

This function does indeed take K to the set ${}^*K$ defined by (†); also,

$${}^*(K^n)=({}^*K)^n,$$

so we may write simply ${}^*K^n$ for either member of this equation. The function $S\mapsto{}^*S$ will be a way to carry any structure on K over to ${}^*K$: if S is an n-ary relation on K, then ${}^*S$ is an n-ary relation on ${}^*K$.

If $(x^0,\dots,x^{n-1})+\mathfrak{m}\in{}^*S$, it need not be the case that $(x^0,\dots,x^{n-1})\in S^\omega$; but it is necessary and sufficient that $(x^0,\dots,x^{n-1})$ be congruent to a member of S^ω.

The following theorem is fundamental.* Notation introduced in 6.1.2 is used.

Theorem 86. *The function* $S\mapsto{}^*S$ *is an embedding of* $\mathscr{D}(K)$ *in* $\mathscr{D}({}^*K)$*, and*

$${}^*S\cap K^n=S\tag{§}$$

whenever $n\in\omega$ *and* $S\subseteq K^n$*.*

*An ultrapower is a special case of an *ultraproduct.* When generalized to ultraproducts and formulated in terms of logical symbolism, the next theorem is known as *Łoś's Theorem* and can be traced to Łoś's paper [21].

Proof. We have $S \subseteq {}^*S \cap K^n$ simply because, if $\boldsymbol{x} \in S$, then $(\boldsymbol{x}, \boldsymbol{x}, \dots) \in S^\omega$. Conversely, if $(\boldsymbol{x}, \boldsymbol{x}, \dots) \in {}^*S$, then $(\boldsymbol{x}, \boldsymbol{x}, \dots)$ is congruent to an element $(\boldsymbol{y}_0, \boldsymbol{y}_1, \dots)$ of S^ω; but then $\boldsymbol{x} = \boldsymbol{y}_k$ for some k, and therefore $\boldsymbol{x} \in S$. Thus ${}^*S \cap K^n \subseteq S$, and (§) holds.

To show that $S \mapsto {}^*S$ is an embedding, it is enough to show

$$^*(\Delta_K) = \Delta_{(^*K)}$$

(which is immediate from the definitions), and for all n in ω and all subsets S and T of K^n,

$$^*(S^{\mathrm{c}}) = (^*S)^{\mathrm{c}}, \tag{¶}$$

$$^*(S \cap T) = {}^*S \cap {}^*T, \tag{‖}$$

and for all m in ω and all subsets U of K^m, if $f : m \to n$ and $g : n \to m$, then

$$^*(g^*[U]) = g^*[^*U], \tag{**}$$

$$^*(f_*(S)) = f_*(^*S). \tag{††}$$

To prove these, let $(x^0, \dots, x^{n-1}) \in (K^\omega)^n$. For (¶) we have

$$\begin{aligned}
(x^0, \dots, x^{n-1}) + \mathfrak{m} \in {}^*(S^{\mathrm{c}}) &\iff \{k \colon \boldsymbol{x}_k \notin S^{\mathrm{c}}\} \in \mathfrak{m} \\
&\iff \{k \colon \boldsymbol{x}_k \notin S\} \notin \mathfrak{m} \\
&\iff (x^0, \dots, x^{n-1}) + \mathfrak{m} \notin {}^*S \\
&\iff (x^0, \dots, x^{n-1}) + \mathfrak{m} \in (^*S)^{\mathrm{c}},
\end{aligned}$$

and for (‖),

$$\begin{aligned}
&(x^0, \dots, x^{n-1}) + \mathfrak{m} \in {}^*(S \cap T) \\
&\iff \{k \colon \boldsymbol{x}_k \notin S \cap T\} \in \mathfrak{m} \\
&\iff \{k \colon \boldsymbol{x}_k \notin S\} \cup \{k \colon \boldsymbol{x}_k \notin T\} \in \mathfrak{m} \\
&\iff \{k \colon \boldsymbol{x}_k \notin S\} \in \mathfrak{m} \;\&\; \{k \colon \boldsymbol{x}_k \notin T\} \in \mathfrak{m} \\
&\iff (x^0, \dots, x^{n-1}) + \mathfrak{m} \in {}^*S \;\&\; (x^0, \dots, x^{n-1}) + \mathfrak{m} \in {}^*T \\
&\iff (x^0, \dots, x^{n-1}) + \mathfrak{m} \in {}^*S \cap {}^*T.
\end{aligned}$$

For $(**)$,

$$\begin{aligned}g^*[{}^*U] &= g^*(\{(y^0+\mathfrak{m},\dots,y^{m-1}+\mathfrak{m})\colon (\boldsymbol{y}_k\colon k\in\omega)\in U^\omega\})\\ &= \{g^*(y^0+\mathfrak{m},\dots,y^{m-1}+\mathfrak{m})\colon (\boldsymbol{y}_k\colon k\in\omega)\in U^\omega\})\\ &= \{g^*(y^0,\dots,y^{m-1})+\mathfrak{m}\colon (\boldsymbol{y}_k\colon k\in\omega)\in U^\omega\}\\ &= \{(t^0,\dots,t^{n-1})+\mathfrak{m}\colon (\boldsymbol{t}_k\colon k\in\omega)\in (g^*[U])^\omega\}\\ &= {}^*(g^*[U]),\end{aligned}$$

and for (††),

$$\begin{aligned}f_*({}^*S) &= \{(t^0,\dots,t^{n-1})+\mathfrak{m}\colon f^*(t^0,\dots,t^{n-1})+\mathfrak{m}\in{}^*T\}\\ &= \{(t^0,\dots,t^{n-1})+\mathfrak{m}\colon (f^*(\boldsymbol{t}_k)\colon k\in\omega)\in S^\omega\}\\ &= \{(t^0,\dots,t^{n-1})+\mathfrak{m}\colon (\boldsymbol{t}_k\colon k\in\omega)\in (t_*(S))^\omega\}\\ &= {}^*(f_*(S)).\end{aligned}$$

□

6.3. First order logic

Given some relations S_0, …, S_{n-1} on K, we may apply some of the operations defined in 6.1.2 repeatedly in order to get a new relation U. If we apply the same operations to ${}^*S_0$, …, ${}^*S_{n-1}$, then, by Theorem 86, we must get the relation ${}^*U$. For example, if $S\subseteq K^n$, and $T\subseteq K^{n+1}$, and f is the inclusion of n in $n+1$, then

$${}^*((f^*((S^{\mathrm{c}}\cup f_*(T))^{\mathrm{c}}))^{\mathrm{c}}) = (f^*((({}^*S)^{\mathrm{c}}\cup f_*({}^*T))^{\mathrm{c}}))^{\mathrm{c}}.$$

We now develop an alternative notation for such equations and their members.

Different people can have the same name. Supposing $S\subseteq K^n$, we want a symbol that can denote both S and ${}^*S$, depending on the context. A person is different from the name of the person; but we shall use S as a symbol for both itself and ${}^*S$. To distinguish which of these two relations is meant by S, we can write S^K for the relation S, and $S^{({}^*K)}$ for ${}^*S$. We may say S is **interpreted** in K as S, and in ${}^*K$ as ${}^*S$. (Of course we have already been using ${}^*S$ as a name for ${}^*S$; but the name does not show clearly that the relation is to be understood as being on ${}^*K$, rather than on ${}^*L$ for some field L that is different from K.)

Considered as a symbol for relations, S here is a **predicate.** Since $S^K\subseteq K^n$, we may say in particular that S is an **n-ary** predicate. Then we can write down the string

$$Sx^0\cdots x^{n-1}$$

of $n + 1$ symbols. This string is an example of a **formula,** and it too has interpretations: it is interpreted in K as S, and in $^*K$ as $^*S$. We may say also that the string **defines** its interpretations. The symbols x^k are **variables,** and the point of introducing them is to be able to write down new formulas that define new relations. So for example if $g\colon n \to m$, then the formula

$$Sx^{g(0)} \dots x^{g(n-1)}$$

defines in K the m-ary relation $g^*(S^K)$. We may also replace variables with **constants,** namely symbols for elements of K. Usually the symbol is the same as the element, so that if a^0, ..., a^{n-1} are elements of K, then we can write the formula

$$Sa^0 \cdots a^{n-1}.$$

Having no variables, this formula is a **sentence;** the sentence is **true** in K if

$$(a^0, \dots, a^{n-1}) \in S^K.$$

A set of predicates is a **signature.** The predicates we are considering come from $\bigcup_{n\in\omega} \mathscr{P}(K^n)$, which might be called the **full signature** signature of K. The formulas introduced so far are more precisely examples of *atomic* formulas in this signature. In general, an **atomic formula** in the full signature of K is a string

$$St^0 \cdots t^{n-1}$$

for some n in ω, where $S \subseteq K^n$, and each t^k is either a variable or a constant. In case $n = 2$, instead of Stu we customarily write

$$t\ S\ u.$$

All variables appearing in an atomic formula are called **free.** The **formulas** of our signature, along with their free variables, are defined recursively:

1. Atomic formulas are formulas, and all of their variables are free.
2. If φ is a formula, then so is its **negation,** $\neg\varphi$, and every free variable in φ is free in $\neg\varphi$.
3. Suppose φ and ψ are formulas, and no variable that occurs in φ, but is not free in φ, is free in ψ. Then the **conjunctions** $(\varphi\ \&\ \psi)$ and $(\psi\ \&\ \varphi)$ are formulas, and every variable that is free in φ or ψ is free in the conjunctions.
4. If φ is a formula, and x is a free variable of φ, then the **instantiation**[*] $\exists x\ \varphi$ is a formula.

[*] I don't know of a common term for formulas $\exists x\ \varphi$; *instantiation* seems to work, though, since the formula will be interpreted as saying that φ is true for some *instance* of x.

More precisely, the formulas so defined are the *good* formulas;[*] but the validity of their definition must be justified. Usually formulas are defined as above, but without any restrictions on their variables. This means the set F of formulas is $\bigcap\mathscr{A}$, where $\mathscr{A}$ consists of the sets B of strings such that B contains the atomic formulas and is closed under the operations of negation, conjunction, and instantiation. Then F admits induction, in the sense that no proper subset has the same closure properties. Also each formula has a set of free variables, and this is defined recursively. But an analogue of Theorem 7 shows that induction is not enough to ensure that such recursive definitions are valid: one needs in addition the following:

Theorem 87 (Unique Readability)**.** *Every formula is* uniquely *an atomic formula, a negation, a conjunction. or an instantiation. Every conjunction is* $(\varphi\mathbin{\&}\psi)$ *for some* unique *formulas* φ *and* ψ*.*

Then free variables of formulas can be defined; then one can go back and make the more restrictive definition of formulas as above.

We can introduce the other customary symbols as abbreviations:

$$\begin{gathered}(\varphi\Rightarrow\psi)\text{ means }\neg(\varphi\mathbin{\&}\neg\psi),\\(\varphi\lor\psi)\text{ means }(\neg\varphi\Rightarrow\psi),\\(\varphi\Leftrightarrow\psi)\text{ means }((\varphi\Rightarrow\psi)\mathbin{\&}(\psi\Rightarrow\varphi)),\\\forall x\ \varphi\text{ means }\neg\exists x\ \neg\varphi.\end{gathered}$$

If the free variables appearing in a formula are all on the list $(x^0,\dots,x^{n-1})$, then the formula can be called **n-ary.** In this case, if $n\leqslant r$, then the formula is also r-ary. If we want to understand a formula φ *as* n-ary, we may write it as $\varphi(x^0,\dots,x^{n-1})$.

Suppose t^k is in K or is a variable for each k in ω. For each n-ary formula θ, a formula $\theta(t^0,\dots,t^{n-1})$ is defined. The definition is recursive:

1. If θ is atomic, then $\theta(\boldsymbol{t})$ is the result of replacing each x^k with t^k.
2. If θ is $\neg\varphi$, then $\theta(\boldsymbol{t})$ is $\neg\psi$, where ψ is $\varphi(\boldsymbol{t})$.
3. If θ is $(\varphi\mathbin{\&}\psi)$, then $\theta(\boldsymbol{t})$ is $(\varphi(\boldsymbol{t})\mathbin{\&}\psi(\boldsymbol{t}))$.
4. If θ is $\exists x^\ell\ \varphi$, then we can understand φ as r-ary, where $r=\max(\ell+1,n)$. In this case, $\theta(\boldsymbol{t})$ is $\exists x^\ell\ \psi$, where ψ is $\varphi(\boldsymbol{u})$, where

$$u^k=\begin{cases}x, & \text{if } k=\ell,\\ t^k, & \text{if } k\neq\ell.\end{cases}$$

[*] Such terminology is used for example by Cohen in his treatment of logic in [6].

The case $n = 0$ is not excluded; in this case, $\theta(t^0, \dots, t^{n-1})$ is simply θ.

The **parameters** of a formula are the (symbols of) elements of K that appear in the formula. A **sentence** is a formula with no free variables, namely a 0-ary or **nullary** formula.

A sentence σ with parameters from K may be **true** in K, in which case we write

$$K \vDash \sigma;$$

otherwise σ is **false** in K, and we write

$$K \nvDash \sigma.$$

The definition is recursive:

1. $K \vDash Sa^0 \cdots a^{n-1}$ if and only if $(a^0, \dots, a^{n-1}) \in S$.
2. $K \vDash \neg\sigma$ if and only if $K \nvDash \sigma$.
3. $K \vDash (\sigma \mathbin{\&} \tau)$ if and only if $K \vDash \sigma$ and $K \vDash \tau$.
4. $K \vDash \exists x\ \varphi$ if and only if, assuming φ is n-ary, there is $\boldsymbol{a}$ in K^n such that $K \vDash \varphi(\boldsymbol{a})$.

All of the foregoing holds also with K replaced by ${}^*K$.

The definition of truth shows why formulas as we have defined them are more precisely called formulas of **first-order logic.** In our formulas, variables stand only for *elements* of K. If we allowed variables standing for *relations* on K, then our formulas would be *second order.* The third of the Peano axioms in § 2.1 is apparently second order; so is the definition of completeness of an ordered field. In Corollaries 95 and 97 of Chapter 7, we shall note that there is no first-order axiomatization of $\mathbb{N}$ or $\mathbb{R}$.

If $S = \{(x, x) \colon x \in K\}$, then, instead of $t\ S\ u$, we may write

$$t = u.$$

Then $K \vDash a = b$ if and only if $a = b$; and likewise in ${}^*K$, by Theorem 86. An n-ary formula φ **defines** an n-ary relation on K, namely $\{\boldsymbol{a} \in K^n \colon {}^*K \vDash \varphi(\boldsymbol{a})\}$; this relation can be denoted by

$$\varphi^K.$$

In case σ is nullary, we have $\sigma^K = \{x \in \{0\} \colon K \vDash \sigma\}$, so that

$$K \vDash \sigma \iff \sigma^K = 1.$$

In the same way, the formula with parameters from ${}^*K$ defines a relation on ${}^*K$, denoted by $\varphi^{({}^*K)}$.

Theorem 88. *Let θ be a formula with parameters from K. Then*

$$^*(\theta^K) = \theta^{(^*K)}.$$

Proof. Since formulas are defined recursively, we can argue inductively, using Theorem 86. Indeed, by this theorem, the claim is true when θ is atomic. If the claim is true when θ is φ, then

$$^*((\neg\varphi)^K) = {}^*((\varphi^K)^{\mathrm{c}}) = ({}^*(\varphi^K))^{\mathrm{c}} = (\varphi^{(^*K)})^{\mathrm{c}} = (\neg\varphi)^{(^*K)}.$$

so the claim is true when θ is $\neg\varphi$. Similarly, if the claim is true when θ is φ or ψ, then the claim is true when θ is $(\varphi \mathbin{\&} \psi)$.

For the final case, let us first note that, if the claim is true when θ is considered as m-ary, and $m \leqslant n$, then the claim is still true when θ is considered as n-ary. Indeed, let f be the inclusion of m in n. Then

$$\theta(x^0, \dots, x^{n-1})^K = \theta(x^0, \dots, x^{m-1})^K \times K^{n-m} = f_*(\theta(x^0, \dots, x^{m-1})^K),$$

and likewise with $^*K$ in place of K. To finish then, we suppose the claim is true when θ is φ, and we prove the claim when θ is $\exists x^\ell\ \varphi$.

We may assume φ and $\exists x^\ell\ \varphi$ are both n-ary, where $\ell < n$. Then we can understand $(x^0, \dots, x^{n-1})$ as $(\boldsymbol{x}, y, \boldsymbol{z})$, where $\boldsymbol{x}$ is $(x^0, \dots, x^{\ell-1})$, and y is x^ℓ, and $\boldsymbol{z}$ is $(x^{\ell+1}, \dots, x^{n-1})$. Let f be the function from $n-1$ to n given by

$$f(k) = \begin{cases} k, & \text{if } k < \ell, \\ k+1, & \text{if } \ell \leqslant k < n-1. \end{cases}$$

Then

$$(\exists x^\ell\ \varphi)^K = f_*(f^*[\varphi^K]), \tag{$*$}$$

which yield the claim when θ is $\exists x^\ell\ \varphi$. To prove $(*)$, we have

$$\begin{aligned} (\exists x^\ell\ \varphi)^K &= \{(\boldsymbol{a}, b, \boldsymbol{c}) \in K^n \colon K \vDash (\exists x^\ell\ \varphi)(\boldsymbol{a}, b, \boldsymbol{c})\} \\ &= \{(\boldsymbol{a}, b, \boldsymbol{c}) \in K^n \colon K \vDash \exists x^\ell\ \varphi(\boldsymbol{a}, x^\ell, \boldsymbol{c})\} \\ &= f_*(\{(\boldsymbol{a}, \boldsymbol{c}) \in K^{n-1} \colon K \vDash \exists x^\ell\ \varphi(\boldsymbol{a}, x^\ell, \boldsymbol{c})\}), \end{aligned}$$

We have also that $K \vDash \exists x^\ell\ \varphi(\boldsymbol{a}, x^\ell, \boldsymbol{c})$ if and only if $K \vDash \varphi(\boldsymbol{a}, b, \boldsymbol{c})$ for some b in K. Then

$$\{(\boldsymbol{a}, \boldsymbol{c}) \in K^{n-1} \colon K \vDash \exists x^\ell\ \varphi(\boldsymbol{a}, x^\ell, \boldsymbol{c})\} = f^*[\varphi^K].$$

Combining these results, we have $(*)$. □

Theorem 89. *Let σ be a sentence with parameters from K. Then*[*]

$$K \vDash \sigma \iff {}^*K \vDash \sigma.$$

Proof. When $n = 0$, then equation (§) in Theorem 86 is simply ${}^*S = S$. □

6.4. Mock higher-order logic

We may want our logic to be able to refer to relations on K, relations on sets of relations of K, and so forth. To achieve this, we can enlarge K to a set that contains, as *elements,* all of the relations just mentioned.

Each of the relations that we want to consider is of a certain **type.** Formally, a type is a string, and the set of types is defined recursively:

1. 0 is a type.
2. If $n \in \omega \smallsetminus \{0\}$, and $(t_0, \dots, \tau_{n-1})$ is a list of n types, then the string

$$n\tau_0 \cdots \tau_{n-1}$$

is a type.

These two conditions are really one, since the type 0 is the unique type of the form $n\tau_0 \cdots \tau_{n-1}$ where $n = 0$. So that we can define functions *on* the set of types by recursion, we observe:

Theorem 90 (Unique Readability)**.** *Every type has the form $n\tau_0 \cdots \tau_{n-1}$ for some* unique *n in ω and some* unique *list $(\tau_0, \dots, \tau_{n-1})$ of types.*

Given the set K, we can now define a function $\tau \mapsto K_\tau$ recursively by:

(1) $K_0 = K$;

(2) if τ is a type $n\tau_0 \cdots \tau_{n-1}$, where $n > 0$, then

$$K_\tau = \mathscr{P}(K_{\tau(0)} \times \cdots \times K_{\tau(n-1)}).$$

Here $\tau(j)$ is just τ_j when written as a subscript. The first part of the definition is not a special case of the second: if τ is not 0, then elements of K_τ are *relations;* but elements of K_0 are just elements of K. Letting T be the set of types, we define

$$\tilde{K} = \bigcup_{\tau \in T} K_\tau.$$

[*] In model-theoretic terms, the theorem is that the full structure on K is an *elementary substructure* of the structure induced on ${}^*K$ by $X \mapsto {}^*X$.

Letting $\mathfrak{m}$ be a non-principal maximal ideal of $\mathscr{P}(\omega)$ as in § 6.2, we have a special case of (†) there:

$$^*\tilde{K} = \tilde{K}^\omega/\mathfrak{m}.$$

Now we have an apparent ambiguity. In § 6.2, we first defined $^*K$ in (†), and then, if $S \subseteq K^n$, we defined $^*S$ in (‡). As we noted, when $S = K$, then the two definitions agree. However, the definition of $^*S$ in general depends on the prior choice of K, in the sense that $^*S$ is a set of congruence-classes of elements of $(K^\omega)^n$. But now there is another possibility: since $K \subseteq \tilde{K}$, we can understand $^*S$ also as a set of congruence-classes of elements of $(\tilde{K}^\omega)^n$. A congruence-class of elements of $(K^\omega)^n$ is never identical to a congruence-class of elements of $(\tilde{K}^\omega)^n$, since every element of $(K^\omega)^n$ is congruent to elements of $(\tilde{K}^\omega)^n$ that are not in $(K^\omega)^n$.

In the notation introduced in the last section, the two possibilities for $^*S$ can be distinguished as $S^{(^*K)}$ and $S^{(^*\tilde{K})}$. However, we need not worry about the distinction, once we observe the following.

Theorem 91. *$^*K$ embeds in $^*\tilde{K}$ under a map i given by*

$$i(\{x \in K^\omega \colon x \equiv a\}) = \{x \in \tilde{K}^\omega \colon x \equiv a\}.$$

Then $\mathscr{P}(^*K^n)$ embeds in $\mathscr{P}(^*\tilde{K}^n)$ under $X \mapsto i[X]$, and the following diagram commutes.

$$\begin{tikzcd}
\mathscr{P}(K^n) \arrow[r, "*"] \arrow[d, "\subseteq"'] & \mathscr{P}(^*K^n) \arrow[d, "i"] \\
\mathscr{P}(\tilde{K}^n) \arrow[r, "*"'] & \mathscr{P}(^*\tilde{K}^n)
\end{tikzcd}$$

In particular, if $S \subseteq K^n$, then

$$i[S^{(^*K)}] = S^{(^*\tilde{K})}. \qquad (*)$$

Proof. Everything follows from the observation that the congruence of two sequences depends only on the sequences themselves, by Theorem 83. In particular, if a and b are in K^ω, and $\{x \in \tilde{K}^\omega \colon x \equiv a\} = \{x \in \tilde{K}^\omega \colon x \equiv b\}$, then $a \equiv b$, so $\{x \in K^\omega \colon x \equiv a\} = \{x \in K^\omega \colon x \equiv b\}$; thus i is injective. Then we have $(*)$ since

$$\begin{aligned}
S^{(^*K)} &= \bigl\{\{y \in K^\omega \colon y \equiv x\} \colon x \in S^\omega\bigr\},\\
S^{(^*\tilde{K})} &= \bigl\{\{y \in \tilde{K}^\omega \colon y \equiv x\} \colon x \in S^\omega\bigr\}.
\end{aligned}$$

□

Now suppose τ is a type $n\tau_0\cdots\tau_{n-1}$, where $n>0$, and $S\in K_\tau$, so that

$$S\subseteq K_{\tau(0)}\times\cdots\times K_{\tau(n-1)}.$$

Then S is both an element of $\tilde{K}$ and an n-ary relation on $\tilde{K}$. As, by Theorem 83, we may and do assume $K\subseteq {}^*K$, so we may assume $\tilde{K}\subseteq {}^*\tilde{K}$. In particular, S is now both an element of ${}^*\tilde{K}$ and an n-ary relation on ${}^*\tilde{K}$. But we have also the n-ary relation ${}^*S$ on ${}^*\tilde{K}$, and this is different from S. Nonetheless, when S is considered as an element of ${}^*\tilde{K}$, we shall want to identify it with the relation ${}^*S$ on ${}^*\tilde{K}$, and *not* with the relation S. We shall be able to do this by Theorem 92 below.

Given the nonzero type τ as above, we define

$$E_\tau=\{(\boldsymbol{a},S)\in\tilde{K}^n\times K_\tau\colon\boldsymbol{a}\in S\},$$

an element of K_υ, where $\upsilon=r\tau_0\cdots\tau_{n-1}\tau$, where $r=n+1$. So E_τ is a a relation of membership. Using S as a constant, we can construct the formula $E_\tau\boldsymbol{x}S$, and we have then

$$(E_\tau\boldsymbol{x}S)^{\tilde{K}}=S^{\tilde{K}}=S. \tag{†}$$

Using S also as a predicate, we can construct also the sentence

$$\forall\boldsymbol{x}\;(E_\tau\boldsymbol{x}S\iff S\boldsymbol{x});$$

by (†), the sentence is true in $\tilde{K}$. Therefore, by Theorem 89, the sentence is true in ${}^*\tilde{K}$, which means

$$(E_\tau\boldsymbol{x}S)^{{}^*\tilde{K}}=S^{{}^*\tilde{K}}={}^*S.$$

Thus S, when considered as an element of ${}^*\tilde{K}$, interacts with the other elements as if it were the relation ${}^*S$ on ${}^*\tilde{K}$. This is not a contradiction: $(E_\tau\boldsymbol{x}S)^{{}^*\tilde{K}}$ is not literally the set of elements of S, since ${}^*E_\tau$ is not literally a relation of membership.

In the following, if $\boldsymbol{a}=(a^0,\dots,a^{n-1})$, we use the notation

$$\iota(\boldsymbol{a})=(\iota(a^0),\dots,\iota(a^{n-1})).$$

Theorem 92. *For each type τ, there is an embedding ι of ${}^*(K_\tau)$ in $({}^*K)_\tau$ such that, if $\tau=0$, then ι is the identity, while if $\tau>0$, then*

$$(\boldsymbol{a},R)\in{}^*E_\tau\iff\iota(\boldsymbol{a})\in\iota(R).$$

Proof. The sentence

$$\forall \boldsymbol{x}\ \forall y\ (E_\tau \boldsymbol{x} y \Rightarrow K_{\tau(0)} x^0\ \&\ \cdots\ \&\ K_{\tau(n-1)} x^{n-1}\ \&\ K_\tau y)$$

is true in $\tilde{K}$, so by Theorem 89, it is true in ${}^*\tilde{K}$. By Theorem 88, for all types σ,

$$(K_\sigma x)^{({}^*\tilde{K})} = {}^*(K_\sigma).$$

Hence, if $R \in {}^*(K_\tau)$, then

$$(E_\tau \boldsymbol{x} R)^{{}^*\tilde{K}} \subseteq {}^*(K_{\tau(0)}) \times \cdots \times {}^*(K_{\tau(n-1)}).$$

We can now define ι recursively:

1. $\iota(x) = x$ if $x \in {}^*(K_0)$,
2. if $\tau = n\tau_0 \cdots \tau_{n-1}$, and $R \in {}^*(K_\tau)$, then

$$\iota(R) = \iota[(E_\tau \boldsymbol{x} R)^{({}^*\tilde{K})}].$$

The sentence

$$\forall y\ \forall z\ \exists \boldsymbol{x}\ (y \neq z \Rightarrow (E_\tau \boldsymbol{x} y \Leftrightarrow \neg E_\tau \boldsymbol{x} z))$$

is true in $\tilde{K}$, hence in ${}^*\tilde{K}$; therefore ι is injective. □

As we shall see in Chapter 7, ι is not generally surjective. Also, even though every element of $\tilde{K}$ is an element of some K_τ, not every element of ${}^*\tilde{K}$ is an element of some ${}^*(K_\tau)$.

7. Analysis

Limits of sequences were defined in § 4.4. The aim of the definition is to formalize the understanding that, if a is a sequence $(a_n\colon n \in \mathbb{N})$ of real numbers, and L is a real number, then L is a **limit** of a if a_n is *close* to L when n is *large.* In a first attempt to make this define precise, we may say that a_n is **close** to L if $|a_n - L| < \varepsilon$, where $\varepsilon > 0$; and n is **large** if $n > M$. Of course these definitions are still imprecise, since they depend on the unquantified variables ε and M. In the usual or 'standard' definition (given on p. 64) of when L is a limit of a, the variables are quantified thus, in the symbolism of § 6.3:

$$\forall\varepsilon\ \big(\varepsilon > 0 \Rightarrow \exists M\ \forall n\ (n \in \mathbb{N}\ \&\ n > M \Rightarrow |a_n - L| < \varepsilon)\big).$$

The 'non-standard' alternative is to quantify the variables in the individual definitions of closeness and largeness: a_n is close to L if $|a_n - L| < \varepsilon$ for *all* positive real numbers ε, and n is large if $n > M$ for *all* real numbers M. Such definitions cannot be satisfied by ordinary or 'standard' real numbers a_n and n; however, they can be satisfied by 'non-standard' real numbers.

The original inspiration for this chapter is Robinson's book [25]; but I have made use of [12] and [2].

7.1. Non-standard numbers and relations

Everything we do now will be based on Ch. 6 in case $K = \mathbb{R}$. We fix a non-principal maximal ideal $\mathfrak{m}$ of $\mathscr{P}(\omega)$: this gives us ${}^*\mathbb{R}$, namely the ultrapower $\mathbb{R}^\omega/\mathfrak{m}$. Then $\mathbb{R}$ embeds properly in ${}^*\mathbb{R}$, by Theorem 84 (p. 90), and we have the function $S \mapsto {}^*S$, defined in (‡) on p. 91, from each $\mathscr{P}(\mathbb{R}^n)$ to $\mathscr{P}({}^*\mathbb{R}^n)$; it is an embedding by Theorem 86 (p. 91). Every ${}^*S$ that arises thus is called a **standard relation.** Also, elements of $\mathbb{R}^n$ are called **standard elements.** Note then that a standard relation might have nonstandard elements. The standard relation ${}^*S$ is the **extension** of S.

We can think of the standard relation ${}^*S$ as what we see when we look at S more closely. Then ${}^*S$ will be the main object of interest, although we can recover S by restricting attention to the standard real numbers, again by Theorem 86. The first-order properties of S in $\mathbb{R}$ are the same as those of ${}^*S$ in ${}^*\mathbb{R}$, by Theorem 88 (p. 97). The usefulness of ${}^*S$ lies in its potentially having non-standard elements.

Now, we cannot talk about these non-standard elements in the full signature of $\mathbb{R}$. Still, we *can* talk about them in ${}^*\mathbb{R}$, given a predicate for the subset $\mathbb{R}$ of ${}^*\mathbb{R}$. In this way, there are statements about $\mathbb{R}$ in ${}^*\mathbb{R}$ that cannot be expressed in $\mathbb{R}$.

In general, the first-order properties of $\mathbb{R}$ are also properties of ${}^*\mathbb{R}$, by Theorem 89 (p. 98). One such feature is being an ordered field, by Theorem 94 below. Towards proving this, suppose $S \subseteq \mathbb{R}^n$, and f is a function from S to $\mathbb{R}$. The **graph** of f is the relation $\{(\boldsymbol{x}, f(\boldsymbol{x}))\colon \boldsymbol{x} \in S\}$; by identifying f with this relation, we obtain the $n+1$-ary relation ${}^*f$ on ${}^*\mathbb{R}$.

Theorem 93. *If $S \subseteq \mathbb{R}^n$, and $f\colon S \to \mathbb{R}$, then ${}^*f\colon {}^*S \to {}^*\mathbb{R}$, and*

$$ {}^*f \restriction S = f. \tag{$*$}$$

Proof. Let T be the graph of f. Then the sentences

$$\begin{gathered}\forall \boldsymbol{x}\ \forall y\ (T\boldsymbol{x}y \Rightarrow S\boldsymbol{x}),\\ \forall \boldsymbol{x}\ \exists y\ (S\boldsymbol{x} \Rightarrow T\boldsymbol{x}y),\\ \forall \boldsymbol{x}\ \forall y\ \forall z\ (T\boldsymbol{x}y\ \&\ T\boldsymbol{x}z \Rightarrow y = z))\end{gathered}$$

are true in $\mathbb{R}$; by Theorem 89, they are true in ${}^*\mathbb{R}$. Therefore ${}^*T$ is the graph of a function—namely ${}^*f$—from ${}^*S$ to ${}^*\mathbb{R}$, and $(*)$ follows since $T \subseteq {}^*T$. □

The function f in the theorem is a **standard function.** In place of an atomic formula $T\boldsymbol{x}y$ as in the proof of the theorem, we may now write

$$f(\boldsymbol{x}) = y.$$

In case $n = 2$, instead of $Txyz$ or $f(x, y) = z$ we usually write

$$x\ f\ y = z,$$

as for example in $x + y = z$. In the general case, if in addition U_k is the graph of a function g_k, then the expression

$$f(g_0(\boldsymbol{x}_0), \dots, g_{n-1}(\boldsymbol{x}_{n-1})) = y$$

can be used to stand for the formula

$$\exists \boldsymbol{z}\ (f(\boldsymbol{z}) = y\ \&\ U_0\boldsymbol{x}_0z^0\ \&\ \cdots\ \&\ U_{n-1}\boldsymbol{x}_{n-1}z^{n-1}),$$

where $\exists \boldsymbol{z}$ means $\exists z^0 \cdots \exists z^{n-1}$. This means we can use a polynomial equation, such as $x(-y + z) = w$ in place of a more complicated formula, like

$$\exists u\ \exists v\ (-y = u\ \&\ u + z = v\ \&\ xv = w).$$

Theorem 94. *$^*\mathbb{R}$ is a non-archimedean ordered field with respect to $^*{<}$, $^*{+}$, $^*{-}$, and $^*{\cdot}$, and $\mathbb{R}$ is an ordered subfield of $^*\mathbb{R}$.*

Proof. There is a first-order sentence σ saying that $\mathbb{R}$ is an ordered field; but then $^*\mathbb{R} \vDash \sigma$ by Theorem 89. By Theorems 86 and 93, $\mathbb{R}$ is an ordered subfield of $^*\mathbb{R}$. Since $\mathbb{R}$ is a *proper* subset of $^*\mathbb{R}$, the latter must be non-archimedean as an ordered field by Theorem 55 (p. 63). □

Corollary 95. *Being archimedean is not a first-order property of fields.*

To understand how the corollary can be true, note that every subset S of $\mathbb{R}$ with an element y and an upper bound z has a least upper bound w: in the notation of § 6.4, since $\mathscr{P}(\mathbb{R}) = \mathbb{R}_{10}$, we can write this statement formally as

$$\forall S \Bigl(\exists y\ E_{10}yS\ \&\ \exists z\ \forall y\ (E_{10}yS \Rightarrow y \leqslant z) \Rightarrow \\ \exists w\ \bigl(\forall y\ (E_{10}yS \Rightarrow y \leqslant w)\ \&\ \forall z\ \bigl(\forall y\ (E_{10}yS \Rightarrow y \leqslant z) \Rightarrow w \leqslant z\bigr)\bigr)\Bigr).$$

This is true in $\tilde{\mathbb{R}}$, so the same sentence is true in $^*\tilde{\mathbb{R}}$. But more precisely, in $\tilde{\mathbb{R}}$, the sentence is not about *subsets* S of $\mathbb{R}$, but about *elements* S of $\mathbb{R}_{10}$, which is $\mathscr{P}(\mathbb{R})$. In $^*\tilde{\mathbb{R}}$, the sentence says that every nonempty element of $^*(\mathbb{R}_{10})$ with an upper bound has a least upper bound. By Theorem 92 (p. 100), we may assume $^*(\mathbb{R}_{10}) \subseteq (^*\mathbb{R})_{10}$, which is $\mathscr{P}(^*\mathbb{R})$. Some nonempty elements of this do have upper bounds, but no least upper bound, since $^*\mathbb{R}$ is non-archimedean. Therefore the inclusion of $^*(\mathbb{R}_{10})$ in $(^*\mathbb{R})_{10}$ is proper, and moreover the latter is non-standard as a relation on $^*\tilde{\mathbb{R}}$.

In the terminology of § 5.1, the *finite* elements of $^*\mathbb{R}$ are those x such that, for some *standard* natural number n, we have $|x| < n$; the non-finite elements are *infinite.* We show now that the elements of $^*\mathbb{N} \smallsetminus \mathbb{N}$ are infinite. In doing so, we may write $x \in \mathbb{N}$ in place of the formula $\mathbb{N}x$; this means $x \in {}^*\mathbb{N}$ in $^*\mathbb{R}$.

Theorem 96. *$\mathbb{N}$ is a proper initial segment of $^*\mathbb{N}$. In particular, $\mathbb{N}$ consists of the finite elements of $^*\mathbb{N}$.*

Proof. For each n in $\mathbb{N}$, the sentence

$$\forall x\ (x \in \mathbb{N} \Rightarrow x = 1 \lor x = 2 \lor \cdots \lor x = n \lor x > n)$$

is true in $\mathbb{R}$, hence in $^*\mathbb{R}$, so that $\{1, 2, \dots, n\}$ is an initial segment of $^*\mathbb{N}$. Therefore $\mathbb{N}$ itself is an initial segment of $^*\mathbb{N}$. The elements of $\mathbb{N}$ are finite by definition. The sentence

$$\forall x\ \forall y\ (x \in \mathbb{N}\ \&\ y \in \mathbb{N}\ \&\ x \leqslant y < x + 1 \Rightarrow x = y)$$

is true in $\mathbb{R}$, hence in ${}^*\mathbb{R}$, so all finite elements of ${}^*\mathbb{N}$ are in $\mathbb{N}$. Finally, the sentence

$$\forall x\ \exists y\ (y \in \mathbb{N}\ \&\ x < y)$$

is true in $\mathbb{R}$ and hence in ${}^*\mathbb{R}$. In particular, let a be a positive infinite element of ${}^*\mathbb{R}$. Then there is n in ${}^*\mathbb{N}$ such that $a < n$. Such n must be infinite, so they are not in $\mathbb{N}$. Therefore $\mathbb{N}$ is a proper subset of ${}^*\mathbb{N}$. □

Corollary 97. *Satisfying the Peano Axioms is not a first-order property of iterative structures.*

Theorem 98. *A standard relation has nonstandard elements if and only if it is infinite.*

Proof. Let $S \subseteq \mathbb{R}^m$. If S is finite, then $S = \{\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1}\}$ for some $\boldsymbol{a}_k$; but then the sentence

$$\forall \boldsymbol{x}\ (\boldsymbol{x} \in S \Leftrightarrow \boldsymbol{x} = \boldsymbol{a}_0 \lor \dots \lor \boldsymbol{a}_{n-1})$$

is true in $\mathbb{R}$ and ${}^*\mathbb{R}$, so ${}^*S = S$. Suppose now S is infinite, so that there is an injective function f from $\mathbb{N}$ into S. By a generalization of Theorem 93, we have an injective function ${}^*f$ from ${}^*\mathbb{N}$ into ${}^*S$. If, for some n in ${}^*\mathbb{N}$, the element ${}^*f(n)$ of ${}^*S$ is an element $\boldsymbol{a}$ of S, then the sentence

$$\exists x\ (x \in \mathbb{N}\ \&\ f(x) = \boldsymbol{a}),$$

being true in ${}^*\mathbb{R}$, is true in $\mathbb{R}$, so $n \in \mathbb{N}$ by injectivity of f. Thus, if $n \in {}^*\mathbb{N} \smallsetminus \mathbb{N}$, then ${}^*f(n) \in {}^*S \smallsetminus S$. □

7.2. Sequences

A sequence $(a_n\colon n \in \mathbb{N})$ of standard real numbers is **bounded** if none of its entries are large. In standard terms, this means

$$\exists M\ \forall n\ (n \in \mathbb{N} \Rightarrow |a_n| < M). \tag{$*$}$$

This is a first-order condition, so it is true of $(a_n\colon n \in \mathbb{N})$ in $\mathbb{R}$ if and only if it is true of the extension ${}^*(a_n\colon n \in \mathbb{N})$ in ${}^*\mathbb{R}$. Now, this extension is a sequence $(a_n\colon n \in {}^*\mathbb{N})$, by Theorem 93; the extension can be called a **standard sequence.** If we denote this by a, then the original sequence $(a_n\colon n \in \mathbb{N})$ is $a \restriction \mathbb{N}$. Now we can say that a is **bounded** if $(*)$ holds in ${}^*\mathbb{R}$; then a is bounded if and only if $a \restriction \mathbb{N}$ is bounded. However, a simpler definition of boundedness is possible:

Theorem 99. *A standard sequence is bounded if and only if each of its terms is finite.*

Proof. If $(*)$ holds in $\mathbb{R}$, then for some standard M, the sentence

$$\forall n\ (n \in \mathbb{N} \Rightarrow |a_n| < M)$$

is true in $\mathbb{R}$; then the sentence is true in ${}^*\mathbb{R}$, so every entry in a is bounded by M, hence finite.

Suppose $(*)$ fails in $\mathbb{R}$. This means the sentence

$$\forall M\ \exists n\ (n \in \mathbb{N}\ \&\ |a_n| \geqslant M)$$

is true in $\mathbb{R}$, hence in ${}^*\mathbb{R}$. In particular, if M is positive and infinite, then there is n in ${}^*\mathbb{N}$ such that $|a_n| \geqslant M$, so a_n is infinite. □

By Theorem 64 (p. 70), the finite elements of ${}^*\mathbb{R}$ compose a valuation ring. This ring has the maximal ideal (which is unique by Theorem 65) consisting of the *infinitesimal* elements of ${}^*\mathbb{R}$, namely those x in ${}^*\mathbb{R}$ such that $|x| < 1/n$ for all *standard* natural numbers n. In the notation introduced in (§) on p. 70, this maximal ideal is $\{x \colon x \simeq 0\}$, and

$$a \simeq b \iff a - b \simeq 0.$$

A sequence $(a_n \colon n \in \mathbb{N})$ of standard real numbers **converges** to the standard real number L, and L is a **limit** of the sequence, if a_n is close to L when n is large. In standard terms (as noted on p. 102), this means

$$\forall \varepsilon\ \big(\varepsilon > 0 \Rightarrow \exists M\ \forall n\ (n \in \mathbb{N}\ \&\ n > M \Rightarrow |a_n - L| < \varepsilon)\big). \qquad (\dagger)$$

Again, the same definition applies to standard sequences, so that such a sequence a converges to L if and only if $a \restriction \mathbb{N}$ converges; but a simpler definition is possible:

Theorem 100. *A standard sequence a has the standard limit L if and only if, for all infinite n in ${}^*\mathbb{N}$,*

$$a_n \simeq L.$$

Proof. Suppose $(\dagger)$ holds in $\mathbb{R}$. Then for all standard positive ε, there is a standard M such that the sentence

$$\forall n\ (n \in \mathbb{N}\ \&\ n > M \Rightarrow |a_n - L| < \varepsilon)$$

holds in $\mathbb{R}$, hence in $^*\mathbb{R}$. Suppose now n is an infinite element of $^*\mathbb{N}$. since M is standard, we have $n > M$, and therefore $|a_n - L| < \varepsilon$. This is so for *all* standard positive ε. therefore $a_n \simeq L$.

Suppose (†) fails in $\mathbb{R}$. Then there is some standard positive ε such that the sentence

$$\forall M\ \exists n\ (n \in \mathbb{N}\ \&\ n > M\ \&\ |a_n - L| \geqslant \varepsilon)$$

holds in $\mathbb{R}$, hence in $^*\mathbb{R}$. In particular, when M is positive and infinite, we have in $^*\mathbb{R}$

$$\exists n\ (n \in {}^*\mathbb{N}\ \&\ n > M\ \&\ |a_n - L| \geqslant \varepsilon);$$

in particular, for some infinite n, $|a_n - L| \geqslant \varepsilon$, so $a_n \not\simeq L$. □

The theorem fails if a is not standard: such is the case when, for example,

$$a_n = \begin{cases} 0, & \text{if } n \text{ is finite;} \\ n, & \text{if } n \text{ is infinite.} \end{cases}$$

Here $(a_n \colon n \in \mathbb{N})$ converges to 0, but $a_n \not\simeq 0$.

An arbitrary function from $^*\mathbb{N}$ to $^*\mathbb{R}$ might have a nonstandard limit. However, such is not the case for standard sequences:

Theorem 101. *A standard sequence has at most one limit.*

Proof. L and M are standard, and If $a_n \simeq L$ and $a_n \simeq M$, then $L \simeq M$, so $L = M$. Hence it is a theorem of $\mathbb{R}$ that a sequence has at most one limit. Therefore a standard sequence has at most one limit. □

If a standard sequence a converges to L, we now know that L is **the limit** of a, and we can write one of

$$\lim_{n\to\infty} a_n = L, \qquad\qquad \lim(a) = L.$$

Now for example $\lim_{n\to\infty} 1/n = 0$, simply because $1/n$ is infinitesimal when n is infinite.

Theorem 102. *Convergent standard sequences are bounded.*

Proof. If a converges to the limit L, then L is standard and, in particular, finite; also, when n is infinite, $a_n \simeq L$, and therefore a_n is also finite. By Theorem 99, a is bounded. □

An **vector space** over a field is an abelian group together with an embedding of the field in the ring of endomorphisms of the group; in particular, if a and b are in the field, then

$$a(x+y)=ax+ay, \qquad 1(x)=x, \qquad (a+b)x=ax+bx, \qquad (ab)x=a(bx).$$

Vector-spaces over $\mathbb{R}$ include the powers $\mathbb{R}^n$. An **algebra** over a field is both a vector-space over the field, and a ring, with the same underlying abelian group in each case, such that

$$a(xy)=(ax)y.$$

Examples of algebras over $\mathbb{R}$ are $\mathbb{C}$ and also $\mathbb{H}$ (the *quaternions*).

Theorem 103. *The convergent standard sequences compose an algebra over $\mathbb{R}$, and the function $a\mapsto\lim(a)$ is a homomorphism from this algebra onto $\mathbb{R}$. That is, if a and b are convergent standard sequences, and $r\in\mathbb{R}$, then $a+b$, ra, and ab converge, and*

$$\lim(a+b)=\lim(a)+\lim(b), \tag{‡}$$
$$\lim(ra)=r\lim(a), \tag{§}$$
$$\lim(ab)=\lim(a)\lim(b). \tag{¶}$$

Moreover, if $\lim(a)\neq 0$, and a_n is never 0 when n is standard, then it is never 0 for any n in ${}^*\mathbb{N}$, and $({a_n}^{-1}\colon n\in{}^*\mathbb{N})$ converges, and

$$\lim_{n\to\infty}\frac{1}{a_n}=\frac{1}{L}. \tag{‖}$$

Proof. We use that the infinitesimals compose an ideal of the ring A of finite members of ${}^*\mathbb{R}$. Suppose $\lim(a)=L$ and $\lim(b)=M$; so these are in A. If n is infinite, then a_n-L and b_n-M are infinitesimal, hence so are $(a_n+b_n)-(L+M)$ and ra_n-rL. This shows (‡) and (§). For (¶), note

$$|a_nb_n-LM|=|a_nb_n-a_nM+a_nM-LM|\leqslant|a_n|\,|b_n-M|+|a_n-L|\,|M|\,.$$

But the last is infinitesimal since $|a_n|\in A$ by Theorems 99 and 102, and $|M|\in A$. For (‖), if $a_n\neq 0$, then $1/L$ and $1/a_n$ are both finite, so, since $a_n\simeq L$, we have $a_n/L\simeq 1$ and therefore

$$\left|\frac{1}{a_n}-\frac{1}{L}\right|=\frac{|L-a_n|}{|a_nL|}\simeq\frac{|L-a_n|}{L^2}\simeq 0. \qquad \square$$

The following theorem should be compared with Theorem 58 (p. 65). Recall that, by Theorem 66 (p. 70) and the ensuing discussion, every finite element a of ${}^*\mathbb{R}$ has a unique *standard part,* namely the standard real number b such that $a \simeq b$.

Theorem 104. *A standard sequence a converges if and only if, for all infinite m and n,*

$$a_m \simeq a_n.$$

Proof. If a converges to L, then $a_n \simeq L$ for all infinite n, and therefore $a_m \simeq a_n$ for all infinite m and n, since $\simeq$ is an equivalence relation.

Suppose conversely $a_m \simeq a_n$ for all infinite m and n. If each a_n is *finite,* and m is infinite, then a converges to the standard part of a_m. Suppose however that some a_n is infinite. Then by Theorem 99, $a \upharpoonright \mathbb{N}$ is unbounded. Hence the sentence

$$\forall m\ \exists n\ (m \in \mathbb{N} \Rightarrow n \in \mathbb{N}\ \&\ m < n\ \&\ |a_m| + 1 \leqslant |a_n|)$$

is true in $\mathbb{R}$ and ${}^*\mathbb{R}$, so a_m and a_n fail to be infinitely close for some infinite m and n. □

7.3. Topology

We may refer to real numbers as **points.** A standard real number b is an **accumulation point**[*] of a subset A of $\mathbb{R}$ if A has elements distinct from, but close to, b. In standard terms, this means

$$\forall \varepsilon\ \big(\varepsilon > 0 \Rightarrow \exists x\ (x \in A\ \&\ 0 < |x - b| < \varepsilon)\big). \tag{$*$}$$

So 0 is an accumulation point of $\{1/n \colon n \in \mathbb{N}\}$, but not of $\mathbb{Z}$. As usual, since the definition is first-order, it holds for a set of real numbers if and only if holds for the extension of the set. Such an extension can be called a **standard set.** If A is a standard set, we may let A denote also $A \cap \mathbb{R}$, as in the proof of the following.

Theorem 105. *A standard point b is an accumulation point of a standard set A if and only if A has an element c such that*

$$c \neq b\ \&\ c \simeq b. \tag{$\dagger$}$$

[*] The term *limit point* is also used, but this can be confused with *limit.* I follow Apostol [2] in using *accumulation point.*

Proof. If b is a standard accumulation point of A, then the sentence $(*)$ is true in $\mathbb{R}$ and ${}^*\mathbb{R}$, so for an infinitesimal ε there is c in A such that $0 < |c - b| < \varepsilon$ and hence $(\dagger)$.

Suppose b is not an accumulation point of A. Then there is some standard positive ε such that the sentence

$$\forall c\ (c \in A \Rightarrow c = b \vee |c - b| \geqslant \varepsilon)$$

is true in $\mathbb{R}$ and ${}^*\mathbb{R}$. Since ε is standard, $|c - b| \geqslant \varepsilon$ implies $c \not\simeq b$. □

The theorem may fail if b is not standard. For example, if b is a positive infinitesimal, then b is not an accumulation point of $\{x\colon x \leqslant 0\}$, although the set contains 0 and $b \simeq 0$.

An accumulation point of a set may, but need not, be an element of the set. Also, an element of the set need not be an accumulation point: if it is not, then it is an **isolated point** of the set. So, a standard point b is an isolated point of a standard set A if and only if

$$\{x\colon x \simeq b\} \cap A = \{b\}.$$

Theorem 106 (Bolzano–Weierstraß). *Every bounded infinite standard set has a standard accumulation point.*

Proof. A bounded infinite standard set includes the range of a bounded standard non-repeating sequence a. Let n be infinite. By Theorem 99, a_n is finite, so it has a standard part, b. If $a_n = b$, then the sentence $\exists m\ (m \in \mathbb{N}\ \&\ a_m = b)$ is true in ${}^*\mathbb{R}$, so it is true also in $\mathbb{R}$, contradicting that a is non-repeating. So $a_n \neq b$. Therefore b is an accumulation point of the original set by Theorem 105. □

In the standard proof of this theorem, if X is an infinite subset of the interval $[a, b]$, then there is a sequence $(e_1, e_2, \dots)$, where each e_k is 0 or 1, and if

$$t_n = \frac{e_1}{2} + \frac{e_2}{4} + \dots + \frac{e_n}{2^n},$$

then the interval

$$[t_n a + (1 - t_n)b, t_n a + (1 - t_n)b + 2^{-n}]$$

contains infinitely many points of X; then $\lim_{n\to\infty}\bigl(t_n a + (1 - t_n)b\bigr)$ is an accumulation point of X.

A point b is an **interior point** of a set A if all points close to b, including b itself, are in A. This means precisely that b is not an accumulation point of the complement of B. By Theorem 105 then we have

Theorem 107. *A standard point b is an interior point of a standard set A if and only if*

$$\{x\colon x \simeq a\} \subseteq B.$$

Recall that in the *order topology* defined in 4.3.1, an **open subset** of $\mathbb{R}$, or simply an **open set,** is the union of a family of open intervals.

Theorem 108. *A subset of $\mathbb{R}$ is open if and only if each of its points is an interior point.*

Proof. Each point of an open interval is an interior point; therefore the same is true for open sets in general. Suppose conversely every point b of O is interior. Then for such b there is some positive ε_b such that $(b-\varepsilon_b, b+\varepsilon_b) \subseteq O$. Hence

$$O = \bigcup_{b \in O} (b-\varepsilon_b, b+\varepsilon_b),$$

so O is open. □

A subset A of an order is **convex** if $z \in A$ whenever $x < z < y$ and x and y are in A. An **interval** of an order is a convex subset that, if it has an upper bound, has a supremum, and, if it has a lower bound, has an infimum. Then open intervals in the earlier sense are intervals. We have the usual notation, so that for example $[a,b) = \{x\colon a \leqslant x < b)$, and $(-\infty, b] = \{x\colon x \leqslant b\}$.

Theorem 109. *The intervals of $\mathbb{R}$ are precisely the convex subsets.*

Proof. Completeness of $\mathbb{R}$. □

Now an open set can be understood as the union of a family of convex open sets. There are nonstandard subsets of ${}^*\mathbb{R}$, such as $\{x\colon x \simeq a\}$, that are convex and open, but are not intervals.

The complement of an open set is a **closed set;** so a closed set is just the intersection of a set of closed intervals. We have immediately from Theorem 108:

Theorem 110. *A subset of $\mathbb{R}$ is closed if and only if it contains all of its accumulation points.*

Theorem 111 (Cantor Intersection Theorem)**.** *Suppose $(F_n\colon n \in \mathbb{N})$ is a sequence of bounded nonempty closed subsets of $\mathbb{R}$ such that*

$$F_1 \supseteq F_2 \supseteq \cdots$$

Then $\bigcap_{n\in\mathbb{N}} F_n$ is nonempty.

Proof. There is a sequence $(a_n \colon n \in \mathbb{N})$ such that $a_n \in F_n$. By the Bolzano–Weierstraß Theorem, being bounded, the sequence has an accumulation point b. By Theorem 110, b is in each of the sets F_n, so it is in the intersection. □

For a non-standard proof, by Theorem 92 we can consider ${}^*(F_n \colon n \in \mathbb{N})$ as $(G_n \colon n \in {}^*\mathbb{N})$, where $G_n = {}^*F_n$ when n is finite. Then G_n is nonempty for every n. For some infinite n, let $a \in G_n$. Then a is finite and is an element of each ${}^*F_k$ (where k is finite). Then the standard part of a is an accumulation point of ${}^*F_k$, so it belongs to this set and therefore to F_k.

The remainder of this section makes no new use of non-standard methods; but the Heine–Borel Theorem will be used in the (non-standard) proof of the Heine–Cantor Theorem in the next section.

A subset $\mathcal{A}$ of $\mathscr{P}(\mathbb{R})$ is a **covering** of a subset B of $\mathbb{R}$, and $\mathcal{A}$ **covers** B, if

$$B \subseteq \bigcup \mathcal{A}.$$

In this case, $\mathcal{A}$ is an **open covering** of B if each element of $\mathcal{A}$ is open.

Theorem 112 (Lindelöf Covering Theorem)**.** *If $\mathcal{A}$ is an open covering of a subset B of $\mathbb{R}$, then some countable subset of $\mathcal{A}$ covers B.*

Proof. Each point c of B belongs to some O in $\mathcal{A}$, and then there is a positive number ε_c such that $(c - \varepsilon_c, c + \varepsilon_c) \subseteq O$. There are *rational* numbers a and b such that

$$c - \varepsilon_c \leqslant a < c < b \leqslant c + \varepsilon,$$

so that $c \in (a, b)$ and $(a, b) \subseteq O$. Denote (a, b) by I_c. Let $\mathcal{D}$ be the set $\{I_c \colon c \in B\}$. Then $\mathcal{D}$ is countable, simply because there are only countably many rational numbers and therefore only countably many intervals with rational endpoints. But $\mathcal{D}$ also covers B. For each I in $\mathcal{D}$, let O_I be an element of $\mathcal{A}$ that includes it. Then $\{O_I \colon I \in \mathcal{D}\}$ is a countable subset of $\mathcal{A}$ that covers B. □

A subset A of $\mathbb{R}$ is **compact** if every open covering of A has a finite subset that also covers A.

Theorem 113 (Heine–Borel Theorem)**.** *A subset of $\mathbb{R}$ is compact if and only if it is closed and bounded.*

Proof. Suppose A is compact. Since A is covered by $\{(-n, n) \colon n \in \mathbb{N}\}$, it is included in some $(-n, n)$, so it is bounded. Let b be an accumulation point of A. Then $\{A \smallsetminus (b - 1/n, b + 1/n) \colon n \in \mathbb{N}\}$ is an open covering of $A \smallsetminus \{b\}$, but no finite subset covers it. Therefore $A \smallsetminus \{b\} = A$; in particular, $b \in A$.

Suppose now A is not compact. By the Lindelöf Covering Theorem, there is a countable open covering $\{B_n \colon n \in \mathbb{N}\}$ of A of which no finite subset covers A. For each n in $\mathbb{N}$, there is an element a_n of $A \smallsetminus (B_1 \cup \cdots \cup B_n)$. If $\{a_n \colon n \in \mathbb{N}\}$ is unbounded, then so is A. Suppose it is bounded. Since it is infinite, it has an accumulation point b, by the Bolzano–Weierstraß Theorem. Then b is an accumulation point of each set $\{a_n \colon n \geqslant k\}$ and hence of $A \smallsetminus (B_1 \cup \cdots \cup B_k)$. Since $(B_1 \cup \cdots \cup B_k)^{\mathrm{c}}$ is closed, it contains b. Then $\left(\bigcup_{n \in \mathbb{N}} B_n\right)^{\mathrm{c}}$ contains b, so $b \notin A$. Thus A does not contain all of its accumulation points, so it is not closed. □

7.4. Continuity

Henceforth the domain of every standard function is a subset of ${}^*\mathbb{R}$. Suppose f is a standard function, and c and L are standard real numbers. Then L is a **limit** of f at c if c is an accumulation point of the domain, and $f(x)$ is close to L whenever $x \in \operatorname{dom}(f)$ and is close, but not equal, to c. In traditional terms, the latter condition is

$$\forall \varepsilon \ \bigl(\varepsilon > 0 \Rightarrow \exists \delta \ \forall x \ (x \in \operatorname{dom}(f) \ \& \ 0 < |x - c| < \delta \Rightarrow |f(x) - L| < \varepsilon)\bigr).$$

We proceed just as in § 7.2.

Theorem 114. *A standard function f has the standard limit L at a standard accumulation point a of the domain of f if and only if, when $x \in \operatorname{dom}(f) \smallsetminus \{c\}$ and $x \simeq c$, then $f(x) \simeq L$.*

Theorem 115. *A standard function has at most one limit at a point.*

If f has a limit L at c, then L is now **the limit** of f at a, and we may write one of

$$\lim_{x \to c} f(x) = L, \qquad\qquad \lim_{c} f = L.$$

There is now an analogue of Theorem 103:

Theorem 116. *The standard functions on a given domain with limits at a standard accumulation point c of this domain compose an algebra over $\mathbb{R}$, and the function $f \mapsto \lim_c(f)$ is a homomorphism from the algebra onto $\mathbb{R}$. If f is in the algebra and is never 0, and $\lim_c(f) \neq 0$, then $1/f$ is in the algebra, and $\lim_c(1/f) = 1/\lim_c(f)$.*

The function f is **continuous at** a non-isolated point c of its domain if $\lim_c(f) = f(c)$, equivalently, $f(x) \simeq f(c)$ whenever $x \simeq c$ (and x is in the domain). Then f is **continuous on** a subset of its domain, if continuous at every point of that subset.

Theorem 117 (Intermediate Value Theorem). *If a standard function f is continuous on $[a, b]$, and d lies between $f(a)$ and $f(b)$, then for some c in (a, b),*

$$f(c) = d.$$

Proof. Suppose $f(a) < d < f(b)$. For all n in $\mathbb{N} \smallsetminus \{0\}$, there is some j in $\mathbb{N}$ such that

$$f\bigl(a + \frac{j}{n}(b-a)\bigr) < d \leqslant f\bigl(a + \frac{j+1}{n}(b-a)\bigr). \tag{$*$}$$

Let n be an *infinite* element of ${}^*\mathbb{N}$. Then $(*)$ holds for some j in ${}^*\mathbb{N}$. Let c be the standard part of $a + (j/n)(b-a)$. Then

$$a + \frac{j}{n}(b-a) \simeq c \simeq a + \frac{j+1}{n}(b-a),$$

so by continuity

$$f\bigl(a + \frac{j}{n}(b-a)\bigr) \simeq f(c) \simeq f\bigl(a + \frac{j+1}{n}(b-a)\bigr).$$

By $(*)$ then we must also have

$$f\bigl(a + \frac{j}{n}(b-a)\bigr) \simeq d,$$

so $f(c) \simeq d$. Therefore $f(c) = d$ since both are standard. □

Corollary 118. *The image of a convex set under a continuous function is a convex set.*

Theorem 119. *If f is monotone on a convex set I, and $f[I]$ is convex, then f is continuous on I.*

Proof. We may assume f is increasing. Suppose a and x are in I, and a is standard, but $f(a) \not\simeq f(x)$. We may assume $a < x$, so $f(a) < f(x)$. If $f(x)$ is infinite, then there is some standard c such that

$$f(a) < c < f(x). \tag{$\dagger$}$$

If $f(x)$ is finite, let d be its standard part; then $f(a) < d$, so there is some standard c such that $f(a) < c < d$, and then again $(\dagger)$ holds. But then $c = f(b)$ for some standard b such that $a < b < x$. In particular, $a \not\simeq x$. □

Corollary 120. *If f is continuous and monotone on a convex set I, then f^{-1} is continuous on $f[I]$.*

Theorem 121 (Extreme Value)**.** *If f is continuous on $[a, b]$, then it attains a maximum and minimum value on the interval.*

Proof. For all positive natural numbers n, for some natural number j such that $j \leqslant n$, the value of

$$f(a + \frac{j}{n}(b - a))$$

is maximized. In particular, this is so when n is infinite. If $i \leqslant n$, we now have

$$f(a + \frac{i}{n}(b - a)) \leqslant f(a + \frac{j}{n}(b - a)).$$

Let d be the standard part of $a + (j/n)(b - a)$. For every c in $[a, b]$, there is a natural number i such that

$$a + \frac{i}{n}(b - a) \leqslant c < a + \frac{i + 1}{n}(b - a).$$

Then these three numbers are infinitely close, so

$$f(c) \simeq f(a + \frac{i}{n}(b - a)).$$

Therefore $f(c) \leqslant f(d)$. □

Again, a standard function f is continuous on a standard convex set I if

$$\forall\varepsilon\ \forall x\ \Big(\varepsilon > 0\ \&\ x \in I \Rightarrow \\ \exists\delta\ \big(\delta > 0\ \&\ \forall y\ (y \in I\ \&\ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)\big)\Big).$$

If we make a slight change, we get a stronger condition:

$$\forall\varepsilon\ \Big(\varepsilon > 0 \Rightarrow \\ \exists\delta\ \big(\delta > 0\ \&\ \forall x\ \forall y\ (x \in I\ \&\ y \in I\ \&\ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)\big)\Big).$$

If f satisfies this, it is **uniformly continuous** on I.

Theorem 122. *A standard function f is uniformly continuous on a standard convex set I if and only if, for all x and y in I,*

$$x \simeq y \implies f(x) \simeq f(y). \tag{‡}$$

Proof. Suppose f is uniformly continuous on I. Then for every standard positive ε, there is a standard positive δ such that, for all x and y in I,

$$|x-y|<\delta \implies |f(x)-f(y)|<\varepsilon.$$

In particular, for every standard positive ε, for all x and y in I,

$$x \simeq y \implies |f(x)-f(y)|<\varepsilon.$$

Then (‡) follows. Suppose now f is *not* uniformly continuous on I. Then for some standard positive ε, if δ is a positive infinitesimal, there are x and y in I such that $|x-y|<\delta$, but $|f(x)-f(y)|\geqslant\varepsilon$. Then $x\simeq y$, but $f(x)\not\simeq f(y)$. □

For example, the function $x\mapsto 1/x$ is continuous on $(0,1]$, but not uniformly continuous, since if x is a positive infinitesimal, then x and $2x$ are in $(0,1]$, and $x\simeq 2x$, but

$$\frac{1}{x}-\frac{1}{2x}=\frac{1}{2x},$$

which is infinite. The function $x\mapsto x^2$ is continuous on $[0,\infty)$, but not uniformly continuous, since if x is infinitesimal, then $1/x$ and $1/x+x$ are in $[0,\infty)$, but

$$\left(\frac{1}{x}+x\right)^2-\frac{1}{x^2}=2+x^2,$$

which is not infinitesimal.

Theorem 123 (Heine–Cantor)**.** *If f is continuous on a standard convex set I, and I is compact, then f is uniformly continuous on I.*

Proof. Suppose x and y are in I, and $x\simeq y$. Since I is compact, it is bounded, by the Heine–Borel Theorem (Theorem 113). Therefore x and y are finite, so they have standard parts; indeed, they have the same standard part, a. Then a is an accumulation point of I, so $a\in I$ since I is closed, again by the Heine–Borel Theorem. Since f is continuous at a, we have

$$f(x)\simeq f(a)\simeq f(y).$$

Thus f is uniformly continuous by Theorem 122. □

7.5. Derivatives

If f is a standard function whose domain contains a non-isolated point c, and

$$\lim_{x\to c}\frac{f(x)-f(c)}{x-c}=d,$$

then we write

$$f'(c)=d,$$

saying f is **differentiable** at c, with **derivative** d at c. So $f'(c)$ is that standard real number d such that, whenever $x\in\operatorname{dom}(f)\smallsetminus\{c\}$ and $x\simeq c$,

$$\frac{f(x)-f(c)}{x-c}\simeq d.$$

Theorem 124. *A standard function differentiable at c is continuous at c.*

Proof. If f is differentiable at c and $x\simeq c$, then

$$f(x)-f(c)\simeq(x-c)f'(c)\simeq 0;$$

so f is continuous at c. □

Theorem 125. *If f and g are differentiable at c, then so are $f+g$ and fg, and*

$$(f+g)'(c)=f'(c)+g'(c),\qquad (fg)'(c)=f'(c)g(c)+f(c)g'(c).$$

Proof. The former equation is easy; for the latter, if $x\simeq c$, then

$$\begin{aligned}\frac{(fg)(x)-(fg)(c)}{x-c}&=\frac{f(x)-f(c)}{x-c}g(c)+f(x)\frac{g(x)-g(c)}{x-c}\\&\simeq f'(c)g(c)+f(x)g'(c)\\&\simeq f'(c)g(c)+f(c)g'(c)\end{aligned}$$

by Theorem 124. □

Theorem 126 (Chain Rule). *If g is differentiable at c, and f differentiable at $g(c)$, while $g(c)$ is an interior point of the domain of f, then $f\circ g$ is differentiable at c, and*

$$(f\circ g)'(c)=f'(g(c))\cdot g'(c).$$

Proof. The conditions ensure that c is a non-isolated point of the domain of $f \circ g$. Suppose $x \in \operatorname{dom}(f \circ g)$ and $x \simeq c$, so $g(x) \simeq g(c)$. We want to show

$$\frac{(f \circ g)(x) - (f \circ g)(c)}{x - c} \simeq f'(g(c)) \cdot g'(c).$$

This holds if $g(x) = g(c)$, since then $g'(c) = 0$. It holds also if $g(x) \neq g(c)$, since in this case

$$\frac{(f \circ g)(x) - (f \circ g)(c)}{x - c} = \frac{f(g(x)) - f(g(c))}{g(x) - g(c)} \cdot \frac{g(x) - g(c)}{x - c}.$$ □

A **neighborhood** of a standard real number a is a set of which a is an interior point.

Theorem 127. *A standard set N is a open neighborhood of a standard real number a if and only if*

$$\{x \colon x \simeq a\} \subseteq N.$$

Theorem 128 (Inverse Function Theorem)**.** *If f is monotone and continuous on a neighborhood of a standard real number a, and f is differentiable at a, but $f'(a) \neq 0$, then f^{-1} is differentiable at $f(a)$ and*

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Proof. By Corollary 120, f^{-1} is continuous at $f(a)$. Suppose $y \simeq f(a)$, but $y \neq f(a)$. Then $f^{-1}(y) \simeq a$, but $f^{-1}(y) \neq a$ by monotonicity, so

$$f'(a) \simeq \frac{y - f(a)}{f^{-1}(y) - a},$$

since $f(f^{-1}(y)) = y$. Since $f'(a) \neq 0$, we have

$$\frac{f^{-1}(y) - f^{-1}(f(a))}{y - f(a)} = \frac{f^{-1}(y) - a}{y - f(a)} \simeq \frac{1}{f'(a)}$$

□

A standard function f has a **local maximum** at a if a is an interior point of the domain of f and, for some neighborhood N of a, $f(a)$ is the greatest element of $\{f(x) \colon x \in N\}$.

Theorem 129. *If a is an interior point of the domain of f, then f has a local maximum at a if and only if*

$$f(a) = \max\{f(x) \colon x \simeq a\}.$$

Proof. If a is an interior point of $\operatorname{dom}(f)$, but f does not have a local maximum at a, then the sentence

$$\forall\varepsilon \left(\varepsilon > 0 \Rightarrow \exists x \left(|a - x| < \varepsilon \,\&\, f(a) < f(x)\right)\right)$$

is true in $\mathbb{R}$ and ${}^*\mathbb{R}$, so $f(a) < f(x)$ for some x such that $x \simeq a$. □

Theorem 130. *If f has a local maximum at a and is differentiable at c, then*

$$f'(c) = 0.$$

Proof. By Theorem 129, if $x \simeq c$, but $x \neq c$, then $f(x) \leqslant f(c)$, so

$$\frac{f(x) - f(c)}{x - c} \begin{cases} \geqslant 0, & \text{if } x < c, \\ \leqslant 0, & \text{if } x > c. \end{cases}$$

Since $(f(x) - f(c))/(x - c) \simeq f'(c)$, we can conclude that $f'(c) = 0$. □

Theorem 131 (Rolle's Theorem)**.** *If f is continuous on $[a, b]$ and differentiable on (a, b), and $f(a) = f(b)$, then, for some c in (a, b),*

$$f'(c) = 0.$$

Proof. Theorems 121 (the Extreme Value Theorem) and 130. □

Theorem 132 (Mean Value Theorem)**.** *If f is continuous on $[a, b]$ and differentiable on (a, b), then, for some c in (a, b),*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Apply Rolle's Theorem to the function

$$x \mapsto f(x) - f(a) - \frac{x - a}{b - a} \cdot (f(b) - f(a)).$$ □

Corollary 133. *If f is continuous on $[a, b]$, and $f'(x) > 0$ for all x in (a, b), then f is increasing on $[a, b]$.*

In particular, for the hypothesis of the Inverse Function Theorem (Theorem 128), it is sufficient that f have a continuous derivative on a neighborhood of a, and $f'(a) \neq 0$.

Corollary 134. *If f is continuous on $[a, b]$, and $f'(x) = 0$ for all x in (a, b), then f is constant on $[a, b]$.*

7.6. Integrals

Suppose f is a standard function defined on $[a,b]$. A **partition** of $[a,b]$ is a list $(a_0,\xi_1,a_1,\dots,a_{n-1},\xi_n,a_n)$ of real numbers such that

$$a=a_0\leqslant\xi_1\leqslant a_1\leqslant\dots\leqslant a_{n-1}\leqslant\xi_n\leqslant a_n=b. \tag{$*$}$$

Then an **integral** of f on $[a,b]$ is a standard real number I such that, for all n in $\mathbb{N}$, for all partitions $(a_0,\xi_1,a_1,\dots,a_{n-1},\xi_n,a_n)$ of $[a,b]$ such that the differences a_k-a_{k-1} are small, the sum

$$\sum_{i=1}^{n}f(\xi_i)(a_i-a_{i-1})$$

is close to I. If it does exist, then such I is indeed unique and is denoted by

$$\int_a^b f,$$

and we say f is **integrable** on $[a,b]$. In standard terms then, $\int_a^b f$—if it exists—is the real number I such that, for all positive ε, there is a positive δ such that, for all n in $\mathbb{N}$, for all partitions $(a_0,\xi_1,a_1,\dots,a_{n-1},\xi_n,a_n)$ of $[a,b]$ such that $\min(a_1-a_0,\dots,a_n-a_{n-1})\leqslant\delta$, we have

$$\Bigl|I-\sum_{i=1}^{n}f(\xi_i)(a_i-a_{i-1})\Bigr|<\varepsilon.$$

This definition is not a first-order statement in $\mathbb{R}$, so we move to $\tilde{\mathbb{R}}$. Let $A_{[a,b]}$ be the set of partitions $(a_0,\xi_1,a_1,\dots,a_{n-1}\xi_n,a_n)$ of $[a,b]$. Such sequences can be understood as binary relations on $\mathbb{R}$, so that $A_{[a,b]}\in\mathbb{R}_{200}$. If f is a bounded function on $[a,b]$, let $S_{f,a,b}$ be the function

$$(a_0,\xi_1,a_1,\dots,a_{n-1},\xi_n,a_n)\mapsto\sum_{i=1}^{n}f(\xi_i)\cdot(a_i-a_{i-1})$$

on $A_{[a,b]}$. So $S_{f,a,b}\in\mathbb{R}_{22000}$. An element of ${}^*A_{[a,b]}$ also takes the form

$$(a_0,\xi_1,a_1,\dots,a_{n-1},\xi_n,a_n),$$

where again ($*$) holds; but now $n\in{}^*\mathbb{N}$. Such an element can be called **fine** if $a_{i-1}\simeq a_i$ for each i in $\{1,\dots,n\}$. It must be noted that fine elements of ${}^*A$ do exist: for example,

$$\Bigl(a,a+\frac{1}{2n}(b-a),a+\frac{1}{n}(b-a),\dots,a+\frac{2n-1}{2n}(b-a),b\Bigr),$$

where n is infinite.

Theorem 135. *A bounded function f on $[a,b]$ is integrable on $[a,b]$ if and only if, for any two fine elements P and P' of ${}^*A_{[a,b]}$,*

$$ {}^*S_{f,a,b}(P) \simeq {}^*S_{f,a,b}(P'). $$

In this case, $\int_a^b f$ is the standard part of either of these sums.

Theorem 136. *A function continuous on an interval is differentiable there.*

Proof. Say f is continuous on $[a,b]$, and let P and P' be fine elements of ${}^*A_{[a,b]}$. We may write

$$ P = (a_0, \xi_1, a_1, \dots, a_{\ell-1}, \xi_\ell, a_\ell), \qquad P' = (a'_0, \xi'_1, a'_1, \dots, a'_{m-1}, \xi'_m, a'_m). $$

Then there is a partition $(c_0, \dots, c_n)$ of $[a,b]$ such that

$$ \{c_0, \dots, c_n\} = \{a_0, \dots, a_\ell, a'_0, \dots, a'_m\}. $$

If $1 \leqslant i \leqslant n$, let $\eta_i = \xi_j$, where $[c_{i-1}, c_i] \subseteq [a_{j-1}, a_j]$; likewise, $\eta'_i = \xi'_k$, where $[c_{i-1}, c_i] \subseteq [a'_{k-1}, a'_k]$. Since the intervals $[a_{j-1}, a_j]$ and $[a'_{k-1}, a'_k]$ are overlapping, their union is an interval of length no greater than $a_j - a_{j-1} + a'_k - a'_{k-1}$, which is infinitesimal. Hence η_i and η'_i are infinitesimally close. If ε is a standard positive real number, then

$$ \begin{aligned} \left|{}^*S_{f,a,b}(P) - {}^*S_{f,a,b}(P')\right| &= \left|\sum_{i=1}^{\ell} f(\xi_i)\cdot(a_i - a_{i-1}) - \sum_{i=1}^{m} f(\xi'_i)\cdot(a'_i - a'_{i-1})\right| \\ &= \left|\sum_{i=1}^{n} f(\eta_i)\cdot(c_i - c_{i-1}) - \sum_{i=1}^{n} f(\eta'_i)\cdot(c_i - c_{i-1})\right| \\ &= \left|\sum_{i=1}^{n} (f(\eta_i) - f(\eta'_i))\cdot(c_i - c_{i-1})\right| \\ &\leqslant \sum_{i=1}^{n} \left|f(\eta_i) - f(\eta'_i)\right|\cdot(c_i - c_{i-1}) \\ &\leqslant \varepsilon\cdot\sum_{i=1}^{n}(c_i - c_{i-1}) \\ &= \varepsilon\cdot(b-a). \end{aligned} $$

Thus ${}^*S_{f,a,b}(P) \simeq {}^*S_{f,a,b}(P')$. □

Theorem 137. *If f and g are integrable on $[a,b]$, and $f(x) \leqslant g(x)$ for all x in $[a,b]$, then $\int_a^b f \leqslant \int_a^b g$.*

If $\int_a^b f$ exists, we can write

$$\int_b^a f = -\int_a^b f.$$

Theorem 138. *If a, b, and c belong to an interval on which f is integrable, then*

$$\int_a^b f + \int_b^c f = \int_a^c f.$$

If f is differentiable on an interval I, and f has the derivative g (that is, $f' = g$), then f is a **primitive** of g.

Theorem 139 (Fundamental Theorem of Calculus)**.** *If f is continuous on $[a,b]$, then the function*

$$x \mapsto \int_a^x f$$

is a primitive of f. If also G is a primitive of f, then

$$\int_a^b f = G(b) - G(a).$$

Proof. Suppose c and x are distinct but infinitesimally close elements of $[a,b]$. Then

$$\frac{\int_a^x f - \int_a^c f}{x - c} = \frac{\int_c^x f}{x - c}.$$

Let m be the minimum, and M the maximum, value that f takes on the interval bounded by x and c; then

$$m \leqslant \frac{\int_c^x f}{x - c} \leqslant M.$$

Since $m \simeq f(c) \simeq M$, the first claim follows. For the second claim, we know $x \mapsto \int_a^x f - G(x)$ is constant by Corollary 134, so

$$\int_a^b f - G(b) = \int_a^a f - G(a) = -G(a). \qquad \square$$

7.7. Sequences of functions

If $(f_n \colon n \in \mathbb{N})$ is a sequence of functions on a convex set I, and f is a function on I, and for each x in I, the sequence $(f_n(x) \colon n \in \mathbb{N})$ converges to $f(x)$, then the sequence of functions can be said to **converge** to the **limit** f. Formally the condition is

$$\forall\varepsilon\ \forall x\ \bigl(\varepsilon > 0 \mathbin{\&} x \in I \Rightarrow \exists M\ \forall n\ (n \in \mathbb{N} \mathbin{\&} n > M \Rightarrow |f_n(x) - f(x)| < \varepsilon)\bigr);$$

equivalently, for infinite n in ${}^*\mathbb{N}$, for all *standard* x in I,

$$f_n(x) \simeq f(x). \tag{$*$}$$

This is not a strong property. For example, suppose $f_n(x) = x^n$, and $I = [0,1]$, and f is given by

$$f(x) = \begin{cases} 0, & \text{if } 0 \leqslant x < 1; \\ 1, & \text{if } x = 1. \end{cases}$$

Then the sequence of f_n converges to f on I, and each f_n is continuous, even uniformly continuous by the Heine–Cantor Theorem (Theorem 123); but f is not continuous on I. Note that, in this example, if M is infinite, we have

$$f_M\Bigl(1 - \frac{1}{2M}\Bigr) \simeq \lim_{n\to\infty}\Bigl(1 - \frac{1}{2n}\Bigr)^n = \mathrm{e}^{-1/2} \neq 0 = f\Bigl(1 - \frac{1}{2M}\Bigr).$$

In general, the sequence $(f_n \colon n \in \mathbb{N})$ **converges uniformly** to f on I if

$$\forall\varepsilon\ \Bigl(\varepsilon > 0 \Rightarrow \exists M\ \forall n\ \bigl(n \in \mathbb{N} \mathbin{\&} n > M \Rightarrow \forall x\ (x \in I \Rightarrow |f_n(x) - f(x)| < \varepsilon)\bigr)\Bigr).$$

Theorem 140. *The sequence $(f_n \colon n \in \mathbb{N})$ converges uniformly to f on I if and only if $(*)$ holds for all x in I and all infinite n in ${}^*\mathbb{N}$.*

To prove Theorem 141 by nonstandard methods, we pass to ${}^*({}^*\mathbb{R})$. If x and y are elements of this, and $|x - y|$ is less than every positive element of ${}^*\mathbb{R}$, we may write

$$x \approxeq y.$$

This is a stronger condition than $x \simeq y$.

Lemma. *A standard function f is continuous at a standard point a if and only if*

$$x \approxeq a \implies f(x) \simeq f(a).$$

Proof. The forward direction follows immediately from the definition and Theorem 114. For the reverse, suppose f is not continuous at a. Then for some positive ε in ${}^*\mathbb{R}$,

$$\forall\delta\ \big(\delta > 0 \Rightarrow \exists x\ (|x-a| < \delta\ \&\ |f(x)-f(a)| \geqslant \varepsilon)\big).$$

This holds also in ${}^*({}^*\mathbb{R})$; hence for some positive δ such that $\delta \approxeq 0$, there is x such that $|x-a| < \delta$, but $|f(x)-f(a)| \geqslant \varepsilon$. In particular, $x \approxeq a$, but $f(x) \not\simeq f(a)$. □

Theorem 141. *If* $(f_n \colon n \in \mathbb{N})$ *converges uniformly to* f *on* I*, and each* f_n *is continuous, then so is* f*.*

Proof. Given a and x in I such that $x \approxeq a$, we show $f(x) \simeq f(a)$. By Theorem 140, if n is infinite,

$$f(x) \simeq f_n(x), \qquad\qquad f_n(a) \simeq f(a).$$

Now, f_n is not a standard function, unless we put ${}^*\mathbb{R}$ in place of $\mathbb{R}$. When we do this, then, since f_n is continuous at a, we have $f_n(x) \approxeq f_n(a)$, so $f(x) \simeq f(a)$. By the lemma, f is continuous at a. □

Under the same hypothesis, the limit of the integrals is the integral of the limits, and the limit of the derivatives is the derivative of the limit…

A. The Greek alphabet

capital	minuscule	transliteration	name
Α	α	a	alpha
Β	β	b	beta
Γ	γ	g	gamma
Δ	δ	d	delta
Ε	ε	e	epsilon
Ζ	ζ	z	zeta
Η	η	ê	eta
Θ	ϑ	th	theta
Ι	ι	i	iota
Κ	ϰ	k	kappa
Λ	λ	l	lambda
Μ	μ	m	mu
Ν	ν	n	nu
Ξ	ξ	x	xi
Ο	ο	o	omicron
Π	π	p	pi
Ρ	ρ	r	rho
Σ	σ, ς	s	sigma
Τ	τ	t	tau
ϒ	υ	y, u	upsilon
Φ	φ	ph	phi
Χ	χ	ch	chi
Ψ	ψ	ps	psi
Ω	ω	ô	omega

The following remarks pertain to *ancient* Greek. The vowels are

α, ε, η, ι, ο, υ, ω,

where η is a long ε, and ω is a long ο; the other vowels (α, ι, υ) can be long or short. Some vowels may be given tonal accents (ά, ᾶ, ὰ). An initial vowel takes either a rough-breathing mark (as in ἁ) or a smooth-breathing mark (ἀ): the former mark is transliterated by a preceding h, and the latter can be ignored, as in

ὑπερβολή hyperbolê *hyperbola,*
ὀρθογώνιον orthogônion *rectangle.*

Likewise, ῥ is transliterated as rh, as in

ῥόμβος rhombos *rhombus.*

A long vowel may have an iota subscript (ᾳ, ῃ, ῳ), especially in case-endings of nouns. Of the two forms of minuscule sigma, the ς appears at the ends of words; elsewhere, σ appears, as in

βάσις basis *base.*

Bibliography

[1] Apollonius of Perga, *Conics. Books I–III*, revised ed., Green Lion Press, Santa Fe, NM, 1998, Translated and with a note and an appendix by R. Catesby Taliaferro, With a preface by Dana Densmore and William H. Donahue, an introduction by Harvey Flaumenhaft, and diagrams by Donahue, Edited by Densmore. MR MR1660991 (2000d:01005)

[2] Tom M. Apostol, *Mathematical analysis*, second ed., Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1974. MR 49 #9123

[3] Archimedes, *The works of Archimedes*, Dover Publications Inc., Mineola, NY, 2002, Reprint of the 1897 edition and the 1912 supplement, Edited by T. L. Heath. MR MR2000800 (2005a:01003)

[4] ______, *The works of Archimedes. Vol. I*, Cambridge University Press, Cambridge, 2004, The two books on the sphere and the cylinder, Translated into English, together with Eutocius' commentaries, with commentary, and critical edition of the diagrams by Reviel Netz. MR MR2093668 (2005g:01006)

[5] A. H. Clifford, *Totally ordered commutative semigroups*, Bull. Amer. Math. Soc. **64** (1958), 305–316. MR MR0100641 (20 #7070)

[6] Paul J. Cohen, *Set theory and the continuum hypothesis*, W. A. Benjamin, Inc., New York-Amsterdam, 1966. MR MR0232676 (38 #999)

[7] R. G. Collingwood, *An autobiography*, Clarendon Press, c. 1938, Reprinted 2002.

[8] Richard Dedekind, *Essays on the theory of numbers. I: Continuity and irrational numbers. II: The nature and meaning of numbers*, authorized translation by Wooster Woodruff Beman, Dover Publications Inc., New York, 1963. MR MR0159773 (28 #2989)

[9] John Dyer-Bennet, *A theorem on partitions of the set of positive integers*, Amer. Math. Monthly **47** (1940), 152–154. MR MR0001234 (1,201b)

[10] Euclid, *The thirteen books of Euclid's Elements translated from the text of Heiberg. Vol. I: Introduction and Books I, II. Vol. II: Books III–IX. Vol. III: Books X–XIII and Appendix*, Dover Publications Inc., New York, 1956, Translated with introduction and commentary by Thomas L. Heath, 2nd ed. MR 17,814b

[11] ———, *Euclid's Elements*, Green Lion Press, Santa Fe, NM, 2002, All thirteen books complete in one volume, the Thomas L. Heath translation, edited by Dana Densmore. MR MR1932864 (2003j:01044)

[12] James M. Henle and Eugene M. Kleinberg, *Infinitesimal calculus*, Dover Publications Inc., Mineola, NY, 2003, Reprint of the 1979 original [MIT Press, Cambridge, MA; MR0564651 (82b:26026)]. MR MR1999278

[13] O. Hölder, *Die Axiome der Quantität and die Lehre vom Mass*, Ber. uber d. Verh. d. K. Sächsischen Ges. d. Wiss. zu Leipzig, Math.-Phys. Cl. **53** (1901), 1–64.

[14] Otto Hölder, *The axioms of quantity and the theory of measurement*, J. Math. Psych. **40** (1996), no. 3, 235–252, Translated from the 1901 German original and with notes by Joel Michell and Catherine Ernst, With an introduction by Michell. MR MR1423724 (98c:00001)

[15] ———, *The axioms of quantity and the theory of measurement*, J. Math. Psych. **41** (1997), no. 4, 345–356, Translated from the 1901 German original and with notes by Joel Michell and Catherine Ernst, With an introduction by Michell. MR MR1609942 (99i:00003)

[16] Edward V. Huntington, *A complete set of postulates for the theory of absolute continuous magnitude*, Trans. Amer. Math. Soc. **3** (1902), no. 2, 264–279. MR MR1500598

[17] ———, *Complete sets of postulates for the theories of positive integral and positive rational numbers*, Trans. Amer. Math. Soc. **3** (1902), no. 2, 280–284. MR MR1500599

[18] Casimir Kuratowski, *Sur la notion d'ordre dans la théorie des ensembles*, Fundamenta Mathematicae (1921), 161–71.

[19] Edmund Landau, *Foundations of analysis. The arithmetic of whole, rational, irrational and complex numbers*, third ed., Chelsea Publishing Company, New York, N.Y., 1966, translated by F. Steinhardt; first edition 1951; first German publication, 1929. MR 12,397m

[20] Serge Lang, *Introduction to algebraic geometry*, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1972, Third printing, with corrections. MR MR0344244 (49 #8983)

[21] Jerzy Łoś, *Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres*, Mathematical interpretation of formal systems, North-Holland Publishing Co., Amsterdam, 1955, pp. 98–113. MR MR0075156 (17,700d)

[22] Angus Macintyre, *Finite fields and model theory*, lecture notes, `http://www.msri.org/communications/ln/hosted/ucb/1998/macintyre/1/index.html`, 1998, MSRI/Evans Hall Lectures.

[23] Murray et al. (eds.), *The compact edition of the Oxford English Dictionary*, Oxford University Press, 1973.

[24] Giuseppe Peano, *The principles of arithmetic, presented by a new method (1889)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 83–97.

[25] Abraham Robinson, *Non-standard analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1996, Reprint of the second (1974) edition, With a foreword by Wilhelmus A. J. Luxemburg. MR MR1373196 (96j:03090)

[26] Lucio Russo, *The forgotten revolution*, Springer-Verlag, Berlin, 2004, How science was born in 300 BC and why it had to be reborn, Translated from the 1996 Italian original by Silvio Levy. MR MR2038833 (2004k:01006)

[27] Michael Spivak, *Calculus*, 2nd ed., Publish or Perish, Berkeley, California, 1980.

[28] Robert R. Stoll, *Set theory and logic*, Dover Publications Inc., New York, 1979, corrected reprint of the 1963 edition. MR 83e:04002

[29] Ivor Thomas (ed.), *Selections illustrating the history of Greek mathematics. Vol. II. From Aristarchus to Pappus*, Harvard University Press, Cambridge, Mass, 1951, With an English translation by the editor. MR 13,419b

Index