Appendix: Analysis

A0 Introduction

This appendix is intended for students using this book for a class in analysis, and for the occasional student in a beginning course who has mastered the statement of the theorem and wishes to delve further.

In addition to proofs of statements not proved in the main text, it includes a justification of arithmetic (Appendix A1), a discussion of cubic and quartic equations (Appendix A2), the Heine-Borel theorem (Appendix A3), a definition of “big O” (Appendix A11), Stirling’s formula (Appendix A16), and a definition of Lebesgue measure and a discussion of what sets are measurable (Appendix A21).

A1 Arithmetic of real numbers

It is harder than one might think to define arithmetic for the reals – addition, multiplication, subtraction, and division – and to show that the usual rules of arithmetic hold. Addition and multiplication as taught in elementary school always start at the right, and for reals there is no right.

Recall that in Section 0.5 we defined the reals as “the set of infinite decimals”. For rigor’s sake we will now spell out exactly what this means; to avoid making special conventions, we will write our infinite decimals with leading 0’s.

Because you learned to add, subtract, divide, and multiply in elementary school, the algorithms used may seem obvious. But understanding how computers simulate real numbers is not nearly as routine as you might imagine. A real number involves an infinite amount of information, and computers cannot handle such things: they compute with finite decimals. This inevitably involves rounding off, and writing arithmetic subroutines that minimize round-off errors is a whole art in itself. In particular, computer addition and multiplication are not commutative or associative. Anyone who really wants to understand numerical problems has to take a serious interest in “computer arithmetic”.

Most equivalence classes consist of a single expression, but the equivalence class 0 has two:

$$+\dots 00.00\dots \text{ and } -\dots 00.00\dots$$

as do the finite decimals. For instance, the equivalence class 1 consists of

$$+\dots 00.99\dots \text{ and } +\dots 01.00\dots .$$

Definition A1.1 (Real numbers). The set of real numbers is the set of equivalence classes of expressions

$$\pm\dots 000a_na_{n-1}\dots a_0\underset{\uparrow}{.}a_{-1}a_{-2}\dots, \qquad A1.1$$

where all a_i are in $\{0,1,2,3,4,5,6,7,8,9\}$, and (as indicated by the arrow) a decimal point separates a_0 and a_{-1}. Two such expressions

$$a = \pm\dots 0a_na_{n-1}\dots a_0.a_{-1}\dots \text{ and } b = \pm\dots 0b_mb_{m-1}\dots b_0.b_{-1}\dots$$

are equivalent if and only if any of the following conditions is met:

1. They are equal.
2. All a_i and all b_i are 0, and the signs are opposite (this equivalence class is called 0).
3. a and b have the same sign; there exists k such that $a_k \neq 9$ and $a_{k-1} = a_{k-2} = \cdots = 9$, and

$$b_j = a_j \text{ for } j > k, \quad b_k = a_k + 1,\ b_{k-1} = b_{k-2} = \cdots = 0.$$

Definition A1.2 (k-truncation). The *k-truncation* of a real number $a = \dots 000 a_n a_{n-1} \cdots a_0 \,.\, a_{-1} a_{-2} \dots$ is the finite decimal

$$[a]_k \stackrel{\text{def}}{=} \dots a_n \dots a_k 000 \dots \qquad \text{A1.2}$$

Equation AA1.2 contains a decimal point but we don't know where to put it. It is among the a_i or among the 0's, depending on whether k is negative or positive.

For instance, if $a = 21.3578$, then $[a]_{-2} = 21.35$.

It is tempting to say that if you take two reals, truncate (cut) them further and further to the right and add them (or multiply them, or subtract them, etc.) and look only at the digits to the left of any fixed position, the digits we see will not be affected by where the truncation takes place, once it is well beyond where we are looking. The problem with this is that it isn't quite true.

Example A1.3 (Addition). Consider adding the following two numbers:

$$\begin{aligned} &.222222\dots 222\dots \\ &.777777\dots 778\dots. \end{aligned} \qquad \text{A1.3}$$

If we truncate before the position of the 8, the sum of the truncated numbers will be $.9999\dots 9$; if we truncate after the 8, it will be $1.0000\dots 0$. So there cannot be any rule which says, "the 100th digit will stay the same if you truncate after the Nth digit, however large N is." The "carry" can come from arbitrarily far to the right. $\triangle$

Definition A1.5: Since we don't yet have a notion of subtraction in $\mathbb{R}$, we can't write $|x - y| < \epsilon$ for $x, y \in \mathbb{R}$, much less

$$\sum (x_i - y_i)^2 < \epsilon^2,$$

which involves addition and multiplication besides. Our definition of k-close uses only subtraction of finite decimals.

It is possible to define the arithmetic of real numbers in terms of digits, but it is quite involved. Even showing that addition is associative involves at least six different cases. None is hard, but keeping straight what you are doing is quite delicate. Exercise A1.6 should give you enough of a taste of this approach. We will use a different approach, based on Definition A1.5, which says what it means for two points to be "k-close".

Let us denote by $\mathbb{D}$ the set of finite decimals.

For instance, if $a = 1.23000013$ and $b = 1.22999903$, then a and b are not 7-close, since

$$[a]_{-7} - [b]_{-7} = 11 \times 10^{-7} > 10^{-7}$$

but they are 6-close, since

$$[a]_{-6} - [b]_{-6} = 10^{-6}.$$

Definition A1.4 (Finite decimal continuity). A map $f : \mathbb{D}^n \to \mathbb{D}$ is called *finite decimal continuous* (or $\mathbb{D}$-continuous) if for all integers N and k, there exists an integer l such that if $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ are two elements of $\mathbb{D}^n$ with all $|x_i|$, $|y_i| < N$, and if $|x_i - y_i| < 10^{-l}$ for all $i = 1, \dots, n$, then

$$|f(x_1, \dots, x_n) - f(y_1, \dots, y_n)| < 10^{-k}. \qquad \text{A1.4}$$

The notion of k-close is the correct way of saying that two numbers agree to k digits after the decimal point. It takes into account the convention by which a number ending in all 9's is equal to the rounded up number ending in all 0's.

The numbers .9998 and 1.0001 are 3-close (but not 4-close).

Exercise A1.2 asks you to show that the functions $A(x, y) = x + y$, $M(x, y) = xy$, $S(x, y) = x - y$, and $\text{Assoc}(x, y, z) = (x + y) + z$ are $\mathbb{D}$-continuous and that $1/x$ is not.

To see why Definition A1.4 is the right definition, we need to say what it means for two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ to be close.

Definition A1.5 (k-close). Two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are *k-close* if $\big|[x_i]_{-k} - [y_i]_{-k}\big| \le 10^{-k}$ for each $i = 1, \dots, n$.

Notice that if two numbers are k-close for all k, then they are equal (see Exercise A1.1).

If $f : \mathbb{D}^n \to \mathbb{D}$ is $\mathbb{D}$-continuous, then define $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}$ by the formula

$$\widetilde{f}(\mathbf{x}) = \sup_k \inf_{l \le -k} f([x_1]_l, \ldots, [x_n]_l). \qquad A1.5$$

Proposition A1.6. *The function $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}$ is the unique function that coincides with f on $\mathbb{D}^n$ and satisfies the continuity condition that for all $k \in \mathbb{N}$ and all $N \in \mathbb{N}$, there exists $l \in \mathbb{N}$ such that when $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are l-close and all coordinates x_i of $\mathbf{x}$ satisfy $|x_i| < N$, then $\widetilde{f}(\mathbf{x})$ and $\widetilde{f}(\mathbf{y})$ are k-close.*

The functions $\widetilde{A}$ and $\widetilde{M}$ satisfy the conditions of Proposition A1.6; thus they apply to the real numbers, while A and M without tildes apply to finite decimals.

The proof is the object of Exercise A1.4. Now setting up arithmetic for the reals is plain sailing: we can define addition and multiplication of reals by setting

$$x + y = \widetilde{A}(x, y) \quad \text{and} \quad xy = \widetilde{M}(x, y), \qquad A1.6$$

where $A(x, y) = x + y$ and $M(x, y) = xy$. It isn't harder to show that the basic laws of arithmetic hold:

$x + y = y + x$	Addition is commutative.
$(x + y) + z = x + (y + z)$	Addition is associative.
$x + (-x) = 0$	Existence of additive inverse.
$xy = yx$	Multiplication is commutative.
$(xy)z = x(yz)$	Multiplication is associative.
$x(y + z) = xy + xz$	Multiplication is distributive over addition.

These are all proved the same way. Let us prove the last. Consider the function $\mathbb{D}^3 \to \mathbb{D}$ given by

$$F(x, y, z) = \overbrace{M\big(x, A(y, z)\big)}^{x(y+z)} - \overbrace{A\big(M(x, y), M(x, z)\big)}^{xy+xz}. \qquad A1.7$$

We leave it to you to check that F is $\mathbb{D}$-continuous, and that

$$\widetilde{F}(x, y, z) = \widetilde{M}\big(x, \widetilde{A}(y, z)\big) - \widetilde{A}\big(\widetilde{M}(x, y), \widetilde{M}(x, z)\big). \qquad A1.8$$

But F is identically 0 on $\mathbb{D}^3$, and the identically 0 function on $\mathbb{R}^3$ coincides with 0 on $\mathbb{D}^3$ and satisfies the continuity condition of Proposition A1.6, so $\widetilde{F}$ vanishes identically by the uniqueness part of Proposition A1.6.

It is one of the basic irritants of elementary school math that division is not defined in the world of finite decimals.

This sets up almost all of arithmetic; the missing piece is division. Exercise A1.3 asks you to define division in the reals.

Exercises for Section A1

A1.1 Show that if two numbers are k-close for all k, then they are equal.