A ProblemText in Advanced Calculus

John M. Erdman
Portland State University

Version July 1, 2014

E-mail address: `erdman@pdx.edu`

.

To Argentina

Contents

PREFACE

In American universities two distinct types of courses are often called "Advanced Calculus": one, largely for engineers, emphasizes advanced computational techniques in calculus; the other, a more "theoretical" course, usually taken by majors in mathematics and physical sciences (and often called *elementary analysis* or *intermediate analysis*), concentrates on conceptual development and proofs. This *ProblemText* is a book of the latter type. It is not a place to look for post-calculus material on Fourier series, Laplace transforms, and the like. It is intended for students of mathematics and others who have completed (or nearly completed) a standard introductory calculus sequence and who wish to understand where all those rules and formulas come from.

Many advanced calculus texts contain more topics than this *ProblemText*. When students are encouraged to develop much of the subject matter for themselves, it is not possible to "cover" material at the same breathtaking pace that can be achieved by a truly determined lecturer. But, while no attempt has been made to make the book encyclopedic, I do think it nevertheless provides an integrated overview of Calculus and, for those who continue, a solid foundation for a first year graduate course in Real Analysis.

As the title of the present document, *ProblemText in Advanced Calculus*, is intended to suggest, it is as much an extended problem set as a textbook. The proofs of most of the major results are either *exercises* or *problems*. The distinction here is that solutions to *exercises* are written out in a separate chapter in the ProblemText while solutions to *problems* are not given. I hope that this arrangement will provide flexibility for instructors who wish to use it as a text. For those who prefer a (modified) Moore-style development, where students work out and present most of the material, there is a quite large collection of problems for them to hone their skills on. For instructors who prefer a lecture format, it should be easy to base a coherent series of lectures on the presentation of solutions to thoughtfully chosen problems.

I have tried to make the *ProblemText* (in a rather highly qualified sense discussed below) "self-contained". In it we investigate how the edifice of calculus can be grounded in a carefully developed substrata of sets, logic, and numbers. Will it be a "complete" or "totally rigorous" development of the subject? Absolutely not. I am not aware of any serious enthusiasm among mathematicians I know for requiring rigorous courses in *Mathematical Logic* and *Axiomatic Set Theory* as prerequisites for a first introduction to analysis. In the use of the tools from set theory and formal logic there are many topics that because of their complexity and depth are cheated, or not even mentioned. (For example, though used often, the *axiom of choice* is mentioned only once.) Even everyday topics such as "arithmetic," see appendix G, are not developed in any great detail.

Before embarking on the main ideas of Calculus proper one *ideally* should have a good background in all sorts of things: quantifiers, logical connectives, set operations, writing proofs, the arithmetic and order properties of the real numbers, mathematical induction, least upper bounds, functions, composition of functions, images and inverse images of sets under functions, finite and infinite sets, countable and uncountable sets. On the one hand all these are technically prerequisite to a careful discussion of the foundations of calculus. On the other hand any attempt to do all this business systematically at the beginning of a course will defer the discussion of anything concerning calculus proper to the middle of the academic year and may very well both bore and discourage students. Furthermore, in many schools there may be students who have already studied much of this material (in a "proofs" course, for example). In a spirit of compromise and flexibility I have relegated this material to appendices. Treat it any way you like. I teach in a large university where students show up for Advanced Calculus with a wide variety of backgrounds, so it is my practice to go over the appendices first, covering many of them in a quite rapid and casual way, my goal being to provide just enough detail so that everyone will know where to find the relevant material when they need it later in the course. After a rapid traversal of the appendices I start Chapter 1.

For this text to be useful a student should have previously studied introductory calculus, more for mathematical maturity than anything else. Familiarity with properties of elementary functions and techniques of differentiation and integration may be assumed and made use of in *examples*—but is never relied upon in the logical development of the material.

One motive for my writing this text is to make available in fairly simple form material that I think of as "calculus done right." For example, differential calculus as it appears in many texts is a morass of tedious epsilon-delta arguments and partial derivatives, the net effect of which is to almost totally obscure the beauty and elegance which results from a careful and patient elaboration of the concept of tangency. On the other hand texts in which things are done right (for example Loomis and Sternberg [8]) tend to be rather forbidding. I have tried to write a text which will be helpful to a determined student with an average background. (I seriously doubt that it will be of much use to the chronically lazy or totally unengaged.)

In my mind one aspect of doing calculus "correctly" is arrange things so that there is nothing to unlearn later. For example, in this text topological properties of the real line are discussed early on. Later, topological things (continuity, compactness, connectedness, and so on) are discussed in the context of metric spaces (because they unify the material conceptually and greatly simplify subsequent arguments). But the important thing is the *definitions* in the single variable case and the metric space case are the same. Students do not have to "unlearn" material as they go to more general settings. Similarly, the differential calculus is eventually developed in its natural habitat of normed linear spaces. But here again, the student who has mastered the one-dimensional case, which occurs earlier in the text, will encounter definitions and theorems and proofs that are virtually identical to the ones with which (s)he is already familiar. There is nothing to unlearn.

In the process of writing this *ProblemText* I have rethought the proofs of many standard theorems. Although some, perhaps most, results in advanced calculus have reached a final, optimal form, there are many others that, despite dozens of different approaches over the years, have proofs that are genuinely confusing to most students. To mention just one example there is the theorem concerning change of order of differentiation whose conclusion is

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}\,.$$

This well-known result to this day still receives clumsy, impenetrable, and even incorrect proofs. I make an attempt, here and elsewhere in the text, to lead students to proofs that are conceptually clear, even when they may not be the shortest or most elegant. And throughout I try very hard to show that mathematics is about ideas and not about manipulation of symbols.

There are of course a number of advantages and disadvantages in consigning a document to electronic life. One slight advantage is the rapidity with which links implement cross-references. Hunting about in a book for *lemma 3.14.23* can be time-consuming (especially when an author engages in the entirely logical but utterly infuriating practice of numbering lemmas, propositions, theorems, corollaries, *etc.* separately). A perhaps more substantial advantage is the ability to correct errors, add missing bits, clarify opaque arguments, and remedy infelicities of style in a timely fashion. The correlative disadvantage is that a reader returning to the web page after a short time may find everything (pages, definitions, theorems, sections) numbered differently. (LaTeXis an amazing tool.) I will change the date on the title page to inform the reader of the date of the last nontrivial update (that is, one that affects numbers or cross-references).

The most serious disadvantage of electronic life is impermanence. In most cases when a web page vanishes so, for all practical purposes, does the information it contains. For this reason (and the fact that I want this material to be freely available to anyone who wants it) I am making use of a "Share Alike" license from Creative Commons. It is my hope that anyone who finds this text useful will correct what is wrong, add what is missing, and improve what is clumsy. To make this possible I am also including the LaTeXsource code on my web page. For more information on creative commons licenses see

http://creativecommons.org/

I extend my gratitude to the many students over the years who have endured various versions of this ProblemText and to my colleagues who have generously nudged me when they found me napping. I want especially to thank Dan Streeter who provided me with so much help in the technical aspects of getting this document produced. The text was prepared using AMS-LaTeX. For the diagrams I used the macro package XY-pic by Kristoffer H. Rose and Ross Moore supplemented by additional macros in the diagxy package by Michael Barr.

Finally it remains only to say that I will be delighted to receive, and will consider, any comments, suggestions, or error reports. My e-mail address is

erdman@pdx.edu

FOR STUDENTS: HOW TO USE THIS PROBLEMTEXT

Some years ago at a national meeting of mathematicians many of the conventioneers went about wearing pins which proclaimed, “Mathematics is Not a Spectator Sport”. It is hard to overemphasize the importance of this observation. The idea behind it has been said in many ways by many people; perhaps it was said best by Paul Halmos [5]: *The only way to learn mathematics is to do mathematics.*

In most respects learning mathematics is more like learning to play tennis than learning history. It is principally an activity, and only secondarily a body of knowledge. Although no one would try to learn tennis just by watching others play the game and perhaps reading a book or two about it, thousands of mathematics students in American universities every year attempt to master mathematical subjects by reading textbooks and passively watching their instructors do mathematics on a blackboard. There are, of course, reasons why this is so, but it is unfortunate nevertheless. This book is designed to encourage *you* to do mathematics.

When you sit down to work it is important to have a sizeable block of time at your disposal during which you will not be interrupted. As you read pay especially close attention to definitions. (After all, before you can think about a mathematical concept you must know what it means.) Read until you arrive at a result (results are labeled “theorem”, “proposition”, “example”, “problem”, “lemma”, *etc.*). Every result requires justification. The proof of a result may appear in the body of the text, or it may be left to you as an *exercise* or a *problem.*

When you reach a result stop and try to prove it. Make a serious attempt. If a hint appears after the statement of the result, at first do not read it. Do not try to find the result elsewhere; and do not ask for help. Halmos [5] points out: “To the passive reader a routine computation and a miracle of ingenuity come with equal ease, and later, when he must depend on himself, he will find that they went as easily as they came.” Of course, it is true that sometimes, even after considerable effort, you will not have discovered a proof. What then?

If a hint is given, and if you have tried seriously but unsuccessfully to derive the result, then (and only then) should you read the hint. Now try again. Seriously.

What if the hint fails to help, or if there is no hint? If you are stuck on a result whose proof is labeled “exercise”, then follow the link to the solution. Turning to the solution should be regarded as a last resort. Even then do not read the whole proof; read just the first line or two, enough to get you started. Now try to complete the proof on your own. If you can do a few more steps, fine. If you get stuck again in midstream, read some more of the proof. Use as little of the printed proof as possible.

If you are stuck on a result whose proof is a “problem”, you will not find a solution in the text. After a really serious attempt to solve the problem, go on. You can’t bring your mathematical education to a halt because of one refractory problem. Work on the next result. After a day or two go back and try again. Problems often “solve themselves”; frequently an intractably murky result, after having been allowed to “rest” for a few days, will suddenly, and inexplicably become entirely clear. In the worst case, if repeated attempts fail to produce a solution, you may have to discuss the problem with someone else—instructor, friend, mother, ...

A question that students frequently ask is, “When I’m stuck and I have no idea at all what to do next, how can I continue to work on a problem?” I know of only one really good answer. It is advice due to Polya. *If you can't solve a problem, then there is an easier problem you can't solve: find it.*

Consider examples. After all, mathematical theorems are usually generalizations of things that happen in interesting special cases. Try to prove the result in some concrete cases. If you succeed, try to generalize your argument. Are you stuck on a theorem about general metric spaces? Try to prove it for Euclidean n-space. No luck? How about the plane? Can you get the result for the real line? The unit interval?

Add hypotheses. If you can't prove the result as stated, can you prove it under more restrictive assumptions? If you are having no success with a theorem concerning general matrices, can you prove it for symmetric ones? How about diagonal ones? What about the 2×2 case?

Finally, one way or another, you succeed in producing a proof of the stated result. Is that the end of the story? By no means. Now you should look carefully at the hypotheses. Can any of them be weakened? Eliminated altogether? If not, construct counterexamples to show that each of the hypotheses is necessary. Is the conclusion true in a more general setting? If so, prove it. If not, give a counterexample. Is the converse true? Again, prove or disprove. Can you find applications of the result? Does it add anything to what you already know about other mathematical objects and facts?

Once you have worked your way through several results (a section or a chapter, say) it is a good idea to consider the organization of the material. Is the order in which the definitions, theorems, *etc.* occur a good one? Or can you think of a more perspicuous ordering? Rephrase definitions (being careful, of course, not to change meanings!), recast theorems, reorganize material, add examples. Do anything you can to make the results into a clear and coherent body of material. In effect you should end up writing your own advanced calculus text.

The fruit of this labor is understanding. After serious work on the foregoing items you will begin to feel that you "understand" the body of material in question. This quest for understanding, by the way, is pretty much what mathematicians do with their lives.

If you don't enjoy the activities outlined above, you probably don't very much like mathematics.

Like most things that are worth doing, learning advanced calculus involves a substantial commitment of time and energy; but as one gradually becomes more and more proficient, the whole process of learning begins to give one a great sense of accomplishment, and, best of all, turns out to be lots of fun.

CHAPTER 1

INTERVALS

The three great realms of calculus are differential calculus, integral calculus, and the theory of infinite series. Central to each of these is the notion of *limit*: derivatives, integrals, and infinite series can be defined as limits of appropriate objects. Before entering these realms, however, one must have some background. There are the basic facts about the algebra, the order properties, and the topology of the real line $\mathbb{R}$. One must know about functions and the various ways in which they can be combined (addition, composition, and so on). And, finally, one needs some basic facts about continuous functions, in particular the *intermediate value theorem* and the *extreme value theorem*.

Much of this material appears in the appendices. The remainder (the topology of $\mathbb{R}$, continuity, and the *intermediate* and *extreme value theorems*) occupies the first six chapters. After these have been completed we proceed with limits and the differential calculus of functions of a single variable.

As you will recall from beginning calculus we are accustomed to calling certain intervals "open" (for example, $(0, 1)$ and $(-\infty, 3)$ are open intervals) and other intervals "closed" (for example, $[0, 1]$ and $[1, \infty)$ are closed intervals). In this chapter and the next we investigate the meaning of the terms "open" and "closed". These concepts turn out to be rather more important than one might at first expect. It will become clear after our discussion in subsequent chapters of such matters as continuity, connectedness, and compactness just how important they really are.

1.1. DISTANCE AND NEIGHBORHOODS

1.1.1. Definition. If x and a are real numbers, the DISTANCE BETWEEN x and a, which we denote by $d(x, a)$, is defined to be $|x - a|$.

1.1.2. Example. There are exactly two real numbers whose distance from the number 3 is 7.

PROOF. We are looking for all numbers $x \in \mathbb{R}$ such that $d(x, 3) = 7$. In other words we want solutions to the equation $|x-3| = 7$. There are two such solutions. If $x-3 \geq 0$, then $|x-3| = x-3$; so $x = 10$ satisfies the equation. On the other hand, if $x - 3 < 0$, then $|x - 3| = -(x - 3) = 3 - x$, in which case $x = -4$ is a solution. $\square$

1.1.3. Exercise. Find the set of all points on the real line that are within 5 units of the number -2. (Solution Q.1.1.)

1.1.4. Problem. Find the set of all real numbers whose distance from 4 is greater than 15.

1.1.5. Definition. Let a be a point in $\mathbb{R}$ and $\epsilon > 0$. The open interval $(a-\epsilon, a+\epsilon)$ centered at a is called the ϵ-NEIGHBORHOOD of a and is denoted by $J_\epsilon(a)$. Notice that this neighborhood consists of all numbers x whose distance from a is less than ϵ; that is, such that $|x - a| < \epsilon$.

1.1.6. Example. The $\frac{1}{2}$-neighborhood of 3 is the open interval $\left(\frac{5}{2}, \frac{7}{2}\right)$.

PROOF. We have $d(x, 3) < \frac{1}{2}$ only if $|x - 3| < \frac{1}{2}$. Solve this inequality to obtain $J_{\frac{1}{2}}(3) = \left(\frac{5}{2}, \frac{7}{2}\right)$. $\square$

1.1.7. Example. The open interval $(1, 4)$ is an ϵ-neighborhood of an appropriate point.

PROOF. The midpoint of $(1, 4)$ is the point $\frac{5}{2}$. The distance from this point to either end of the interval is $\frac{3}{2}$. Thus $(1, 4) = J_{\frac{3}{2}}\left(\frac{5}{2}\right)$. $\square$

1.1.8. Problem. Find, if possible, a number ϵ such that the ϵ-neighborhood of $\frac{1}{3}$ contains both $\frac{1}{4}$ and $\frac{1}{2}$ but does not contain $\frac{17}{30}$. If such a neighborhood does not exist, explain why.

1.1.9. Problem. Find, if possible, a number ϵ such that the ϵ-neighborhood of $\frac{1}{3}$ contains $\frac{11}{12}$ but does not contain either $\frac{1}{2}$ or $\frac{5}{8}$. If such a neighborhood does not exist, explain why.

1.1.10. Problem. Let $U = \left(\frac{1}{4}, \frac{2}{3}\right)$ and $V = \left(\frac{1}{2}, \frac{6}{5}\right)$. Write U and V as ϵ-neighborhoods of appropriate points. (That is, find numbers a and ϵ such that $U = J_\epsilon(a)$ and find numbers b and δ such that $V = J_\delta(b)$.) Also write the sets $U \cup V$ and $U \cap V$ as ϵ-neighborhoods of appropriate points.

1.1.11. Problem. Generalize your solution to the preceding problem to show that the union and the intersection of any two ϵ-neighborhoods that overlap is itself an ϵ-neighborhood of some point. *Hint.* Since ϵ-neighborhoods are open intervals of finite length, we can write the given neighborhoods as (a, b) and (c, d). There are really just two distinct cases. One neighborhood may contain the other; say, $a \le c < d \le b$. Or each may have points that are not in the other; say $a < c < b < d$. Deal with the two cases separately.

1.1.12. Proposition. *If $a \in \mathbb{R}$ and $0 < \delta \le \epsilon$, then $J_\delta(a) \subseteq J_\epsilon(a)$.*

Proof. Exercise. (Solution Q.1.2.)

1.2. INTERIOR OF A SET

1.2.1. Definition. Let $A \subseteq \mathbb{R}$. A point a is an INTERIOR POINT of A if some ϵ-neighborhood of a lies entirely in A. That is, a is an interior point of A if and only if there exists $\epsilon > 0$ such that $J_\epsilon(a) \subseteq A$. The set of all interior points of A is denoted by A° and is called the INTERIOR of A.

1.2.2. Example. Every point of the interval $(0, 1)$ is an interior point of that interval. Thus $(0, 1)^\circ = (0, 1)$.

Proof. Let a be an arbitrary point in $(0, 1)$. Choose ϵ to be the smaller of the two (positive) numbers a and $1 - a$. Then $J_\epsilon(a) = (a - \epsilon, a + \epsilon) \subseteq (0, 1)$ (because $\epsilon \le a$ implies $a - \epsilon \ge 0$, and $\epsilon \le 1 - a$ implies $a + \epsilon \le 1$). □

1.2.3. Example. If $a < b$, then every point of the interval (a, b) is an interior point of the interval. Thus $(a, b)^\circ = (a, b)$.

Proof. Problem.

1.2.4. Example. The point 0 is not an interior point of the interval $[0, 1)$.

Proof. Argue by contradiction. Suppose 0 belongs to the interior of $[0, 1)$. Then for some $\epsilon > 0$ the interval $(-\epsilon, \epsilon) = J_\epsilon(0)$ is contained in $[0, 1)$. But this is impossible since the number $-\frac{1}{2}\epsilon$ belongs to $(-\epsilon, \epsilon)$ but not to $[0, 1)$. □

1.2.5. Example. Let $A = [a, b]$ where $a < b$. Then $A^\circ = (a, b)$.

Proof. Problem.

1.2.6. Example. Let $A = \{x \in \mathbb{R} \colon x^2 - x - 6 \ge 0\}$. Then $A^\circ \ne A$.

Proof. Exercise. (Solution Q.1.3.)

1.2.7. Problem. Let $A = \{x \in \mathbb{R} \colon x^3 - 2x^2 - 11x + 12 \le 0\}$. Find A°.

1.2.8. Example. The interior of the set $\mathbb{Q}$ of rational numbers is empty.

Proof. No open interval contains only rational numbers. □

1.2.9. Proposition. *If A and B are sets of real numbers with $A \subseteq B$, then $A^\circ \subseteq B^\circ$.*

Proof. Let $a \in A^\circ$. Then there is an $\epsilon > 0$ such that $J_\epsilon(a) \subseteq A \subseteq B$. This shows that $a \in B^\circ$. □

1.2.10. Proposition. *If A is a set of real numbers, then $A^{\circ\circ} = A^{\circ}$.*

PROOF. Problem.

1.2.11. Proposition. *If A and B are sets of real numbers, then*

$$(A \cap B)^{\circ} = A^{\circ} \cap B^{\circ}.$$

PROOF. Exercise. *Hint.* Show separately that $(A \cap B)^{\circ} \subseteq A^{\circ} \cap B^{\circ}$ and that $A^{\circ} \cap B^{\circ} \subseteq (A \cap B)^{\circ}$. (Solution Q.1.4.)

1.2.12. Proposition. *If A and B are sets of real numbers, then*

$$(A \cup B)^{\circ} \supseteq A^{\circ} \cup B^{\circ}.$$

PROOF. Exercise. (Solution Q.1.5.)

1.2.13. Example. Equality may fail in the preceding proposition.

PROOF. Problem. *Hint.* See if you can find sets A and B in $\mathbb{R}$ both of which have empty interior but whose union is all of $\mathbb{R}$.

CHAPTER 2

TOPOLOGY OF THE REAL LINE

It is clear from the definition of "interior" that the interior of a set is always contained in the set. Those sets for which the reverse inclusion also holds are called *open sets*.

2.1. OPEN SUBSETS OF $\mathbb{R}$

2.1.1. Definition. A subset U of $\mathbb{R}$ is OPEN if $U^\circ = U$. That is, a set is open if and only if every point of the set is an interior point of the set. If U is an open subset of $\mathbb{R}$ we write $U \overset{\circ}{\subseteq} \mathbb{R}$.

Notice, in particular, that the empty set is open. This is a consequence of the way implication is defined in section D.2: the condition that each point of $\emptyset$ be an interior point is *vacuously satisfied* because there *are no points* in $\emptyset$. (One argues that *if* an element x belongs to the empty set, *then* it is an interior point of the set. The hypothesis is false; so the implication is true.) Also notice that $\mathbb{R}$ itself is an open subset of $\mathbb{R}$.

2.1.2. Example. Bounded open intervals are open sets. That is, if $a < b$, then the open interval (a, b) is an open set.

PROOF. Example 1.2.3. □

2.1.3. Example. The interval $(0, \infty)$ is an open set.

PROOF. Problem.

One way of seeing that a set is open is to verify that each of its points is an interior point of the set. That is what the definition says. Often it is easier to observe that the set can be written as a union of bounded open intervals. That this happens exactly when a set is open is the point of the next proposition.

2.1.4. Proposition. *A nonempty subset of $\mathbb{R}$ is open if and only if it is a union of bounded open intervals.*

PROOF. Let $U \subseteq \mathbb{R}$. First, let us suppose U is a nonempty open subset of $\mathbb{R}$. Each point of U is then an interior point of U. So for each $x \in U$ we may choose a bounded open interval $J(x)$ centered at x which is entirely contained in U. Since $x \in J(x)$ for each $x \in U$, we see that

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} J(x). \tag{2.1}$$

On the other hand, since $J(x) \subseteq U$ for each $x \in U$, we have (see proposition F.1.8)

$$\bigcup_{x \in U} J(x) \subseteq U. \tag{2.2}$$

Together (2.1) and (2.2) show that U is a union of bounded open intervals.

For the converse suppose $U = \bigcup \mathfrak{J}$ where $\mathfrak{J}$ is a family of open bounded intervals. Let x be an arbitrary point of U. We need only show that x is an interior point of U. To this end choose an interval $J \in \mathfrak{J}$ which contains x. Since J is a bounded open interval we may write $J = (a, b)$ where $a < b$. Choose ϵ to be the smaller of the numbers $x - a$ and $b - x$. Then it is easy to see that $\epsilon > 0$ and that $x \in J_\epsilon(x) = (x - \epsilon, x + \epsilon) \subseteq (a, b)$. Thus x is an interior point of U. □

2.1.5. Example. Every interval of the form $(-\infty, a)$ is an open set. So is every interval of the form (a, ∞). (Notice that this and example 2.1.2 give us the very comforting result that the things we are accustomed to calling open intervals are indeed open sets.)

PROOF. Problem.

The study of calculus has two main ingredients: algebra and topology. Algebra deals with operations and their properties, and with the resulting structure of groups, fields, vector spaces, algebras, and the like. Topology, on the other hand, is concerned with closeness, ϵ-neighborhoods, open sets, and with the associated structure of metric spaces and various kinds of topological spaces. Almost everything in calculus results from the interplay between algebra and topology.

2.1.6. Definition. The word "topology" has a technical meaning. A family $\mathfrak{T}$ of subsets of a set X is a TOPOLOGY on X if

(1) $\emptyset$ and X belong to $\mathfrak{T}$;
(2) if $\mathfrak{S} \subseteq \mathfrak{T}$ (that is, if $\mathfrak{S}$ is a subfamily of $\mathfrak{T}$), then $\bigcup \mathfrak{S} \in \mathfrak{T}$; and
(3) if $\mathfrak{S}$ is a *finite* subfamily of $\mathfrak{T}$, then $\bigcap \mathfrak{S} \in \mathfrak{T}$.

We can paraphrase this definition by saying that a family of subsets of X, which contains both $\emptyset$ and X, is a topology on X if it is closed under arbitrary unions and finite intersections. If this definition doesn't make any sense to you at first reading, don't fret. This kind of abstract definition, although easy enough to remember, is irritatingly difficult to understand. Staring at it doesn't help. It appears that a bewildering array of entirely different things might turn out to be topologies. And this is indeed the case. An understanding and appreciation of the definition come only gradually. You will notice as you advance through this material that many important concepts such as continuity, compactness, and connectedness are defined by (or characterized by) open sets. Thus theorems which involve these ideas will rely on properties of open sets for their proofs. This is true not only in the present realm of the real line but in the much wider world of metric spaces, which we will shortly encounter in all their fascinating variety. You will notice that two properties of open sets are used over and over: that unions of open sets are open and that finite intersections of open sets are open. Nothing else about open sets turns out to be of much importance. Gradually one comes to see that these two facts completely dominate the discussion of continuity, compactness, and so on. Ultimately it becomes clear that nearly everything in the proofs goes through in situations where *only* these properties are available—that is, in topological spaces.

Our goal at the moment is quite modest: we show that the family of all open subsets of $\mathbb{R}$ is indeed a topology on $\mathbb{R}$.

2.1.7. Proposition. *Let $\mathfrak{S}$ be a family of open sets in $\mathbb{R}$. Then*

(a) *the union of $\mathfrak{S}$ is an open subset of $\mathbb{R}$; and*
(b) *if $\mathfrak{S}$ is finite, the intersection of $\mathfrak{S}$ is an open subset of $\mathbb{R}$.*

PROOF. Exercise. (Solution Q.2.1.)

2.1.8. Example. The set $U = \{x \in \mathbb{R} \colon x < -2\} \cup \{x > 0 \colon x^2 - x - 6 < 0\}$ is an open subset of $\mathbb{R}$.

PROOF. Problem.

2.1.9. Example. The set $\mathbb{R} \setminus \mathbb{N}$ is an open subset of $\mathbb{R}$.

PROOF. Problem.

2.1.10. Example. The family $\mathfrak{T}$ of open subsets of $\mathbb{R}$ is not closed under arbitrary intersections.s (That is, there exists a family $\mathfrak{S}$ of open subsets of $\mathbb{R}$ such that $\bigcap \mathfrak{S}$ is *not* open.)

PROOF. Problem.

2.2. CLOSED SUBSETS OF $\mathbb{R}$

Next we will investigate the closed subsets of $\mathbb{R}$. These will turn out to be the complements of open sets. But initially we will approach them from a different perspective.

2.2.1. Definition. A point b in $\mathbb{R}$ is an ACCUMULATION POINT of a set $A \subseteq \mathbb{R}$ if every ϵ-neighborhood of b contains at least one point of A distinct from b. (We do *not* require that b belong to A, although, of course, it may.) The set of all accumulation points of A is called the DERIVED SET of A and is denoted by A'. The CLOSURE of A, denoted by $\overline{A}$, is $A \cup A'$.

2.2.2. Example. Let $A = \{1/n \colon n \in \mathbb{N}\}$. Then 0 is an accumulation point of A. Furthermore, $\overline{A} = \{0\} \cup A$.

PROOF. Problem.

2.2.3. Example. Let A be $(0,1) \cup \{2\} \subseteq \mathbb{R}$. Then $A' = [0,1]$ and $\overline{A} = [0,1] \cup \{2\}$.

PROOF. Problem.

2.2.4. Example. Every real number is an accumulation point of the set $\mathbb{Q}$ of rational numbers (since every open interval in $\mathbb{R}$ contains infinitely many rationals); so $\overline{\mathbb{Q}}$ is all of $\mathbb{R}$.

2.2.5. Exercise. Let $A = \mathbb{Q} \cap (0,\infty)$. Find A°, A', and $\overline{A}$. (Solution Q.2.2.)

2.2.6. Problem. Let $A = (0,1] \cup \big([2,3] \cap \mathbb{Q}\big)$. Find:

(a) A°;
(b) $\overline{A}$;
(c) $\overline{A^\circ}$;
(d) $\big(\overline{A}\big)^\circ$;
(e) $\overline{A^c}$;
(f) $\big(\overline{A^c}\big)^\circ$;
(g) $\big(A^c\big)^\circ$; and
(h) $\overline{\big(A^c\big)^\circ}$.

2.2.7. Example. Let A be a nonempty subset of $\mathbb{R}$. If A is bounded above, then $\sup A$ belongs to the closure of A. Similarly, if A is bounded below, then $\inf A$ belongs to $\overline{A}$.

PROOF. Problem.

2.2.8. Problem. Starting with a set A, what is the greatest number of *different* sets you can get by applying successively the operations of closure, interior, and complement? Apply them as many times as you wish and in any order. For example, starting with the empty set doesn't produce much. We get only $\emptyset$ and $\mathbb{R}$. If we start with the closed interval $[0,1]$, we get four sets: $[0,1]$, $(0,1)$, $(-\infty,0] \cup [1,\infty)$, and $(-\infty,0) \cup (1,\infty)$. By making a more cunning choice of A, how many different sets can you get?

2.2.9. Proposition. *Let $A \subseteq \mathbb{R}$. Then*

(a) $\big(A^\circ\big)^c = \overline{A^c}$; *and*
(b) $\big(A^c\big)^\circ = \big(\overline{A}\big)^c$.

PROOF. Exercise. *Hint.* Part (b) is a very easy consequence of (a). (Solution Q.2.3.)

2.2.10. Definition. A subset A of $\mathbb{R}$ is CLOSED if $\overline{A} = A$.

2.2.11. Example. Every closed interval $\big($that is, intervals of the form $[a,b]$ or $(-\infty,a]$ or $[a,\infty)$ or $(-\infty,\infty)$ $\big)$ are closed.

PROOF. Problem.

CAUTION. It is a common mistake to treat subsets of $\mathbb{R}$ as if they were doors or windows, and to conclude, for example, that a set is closed because it is not open, or that it cannot be closed because it is open. These "conclusions" are wrong! A subset of $\mathbb{R}$ may be open and not closed, or closed and not open, or both open and closed, or neither. For example, in $\mathbb{R}$:

(1) $(0, 1)$ is open but not closed;
(2) $[0, 1]$ is closed but not open;
(3) $\mathbb{R}$ is both open and closed; and
(4) $[0, 1)$ is neither open nor closed.

This is not to say, however, that there is no relationship between these properties. In the next proposition we discover that the correct relation has to do with complements.

2.2.12. Proposition. *A subset of* $\mathbb{R}$ *is open if and only if its complement is closed.*

PROOF. Problem. *Hint.* Use proposition 2.2.9.

2.2.13. Proposition. *The intersection of an arbitrary family of closed subsets of* $\mathbb{R}$ *is closed.*

PROOF. Let $\mathfrak{A}$ be a family of closed subsets of $\mathbb{R}$. By *De Morgan's law* (see proposition F.3.5) $\bigcap \mathfrak{A}$ is the complement of $\bigcup\{A^c \colon A \in \mathfrak{A}\}$. Since each set A^c is open (by 2.2.12), the union of $\{A^c \colon A \in \mathfrak{A}\}$ is open (by 2.1.7(a)); and its complement $\bigcap \mathfrak{A}$ is closed (2.2.12 again). $\square$

2.2.14. Proposition. *The union of a finite family of closed subsets of* $\mathbb{R}$ *is closed.*

PROOF. Problem.

2.2.15. Problem. Give an example to show that the union of an arbitrary family of closed subsets of $\mathbb{R}$ need not be closed.

2.2.16. Definition. Let a be a real number. Any open subset of $\mathbb{R}$ which contains a is a NEIGHBORHOOD of a. Notice that an ϵ-neighborhood of a is a very special type of neighborhood: it is an interval and it is symmetric about a. For most purposes the extra internal structure possessed by ϵ-neighborhoods is irrelevant to the matter at hand. To see that we can operate as easily with general neighborhoods as with ϵ-neighborhoods do the next problem.

2.2.17. Problem. Let A be a subset of $\mathbb{R}$. Prove the following.

(a) A point a is an interior point of A if and only if some neighborhood of a lies entirely in A.
(b) A point b is an accumulation point of A if and only if every neighborhood of b contains at least one point of A distinct from b.

CHAPTER 3

CONTINUOUS FUNCTIONS FROM $\mathbb{R}$ TO $\mathbb{R}$

The continuous functions are perhaps the most important single class of functions studied in calculus. Very roughly, a function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at a point a in $\mathbb{R}$ if $f(x)$ can be made arbitrarily close to $f(a)$ by insisting that x be sufficiently close to a. In this chapter we define continuity for real valued functions of a real variable and derive some useful necessary and sufficient conditions for such a function to be continuous. Also we will show that composites of continuous functions are themselves continuous. We will postpone until the next chapter proofs that other combinations (sums, products, and so on) of continuous functions are continuous. The first applications of continuity will come in chapters 5 and 6 on connectedness and compactness.

3.1. CONTINUITY—AS A LOCAL PROPERTY

The definition of continuity uses the notion of the inverse image of a set under a function. It is a good idea to look at appendices L and M, at least to fix notation.

3.1.1. Definition. A function $f\colon \mathbb{R} \to \mathbb{R}$ is CONTINUOUS AT a point a in $\mathbb{R}$ if $f^{\leftarrow}(V)$ contains a neighborhood of a whenever V is a neighborhood of $f(a)$. Here is another way of saying exactly the same thing: f is continuous at a if every neighborhood of $f(a)$ contains the image under f of a neighborhood of a. (If it is not entirely clear that these two assertions are equivalent, use propositions M.1.22 (a) and M.1.23 (a) to prove that $U \subseteq f^{\leftarrow}(V)$ if and only if $f^{\rightarrow}(U) \subseteq V$.)

As we saw in chapter 2, it seldom matters whether we work with general neighborhoods (as in the preceding definition) or with the more restricted ϵ-neighborhoods (as in the next proposition).

3.1.2. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at $a \in \mathbb{R}$ if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that*

$$J_\delta(a) \subseteq f^{\leftarrow}\big(J_\epsilon(f(a))\big) \tag{3.1}$$

Before starting on a proof it is always a good idea to be certain that the meaning of the proposition is entirely clear. In the present case the "if and only if" tells us that we are being given a condition *characterizing* continuity at the point a; that is, a condition which is both necessary and sufficient in order for the function to be continuous at a. The condition states that no matter what positive number ϵ we are given, we can find a corresponding positive number δ such that property (3.1) holds. This property is the heart of the matter and should be thought of conceptually rather than in terms of symbols. Learning mathematics is easiest when we regard the content of mathematics to be ideas; it is hardest when mathematics is thought of as a game played with symbols. Thus property (3.1) says that if x is any number in the open interval $(a-\delta, a+\delta)$, then the corresponding value $f(x)$ lies in the open interval between $f(a)-\epsilon$ and $f(a)+\epsilon$. Once we are clear about what it is that we wish to establish, it is time to turn to the proof.

PROOF. Suppose f is continuous at a. If $\epsilon > 0$, then $J_\epsilon(f(a))$ is a neighborhood of $f(a)$ and therefore $f^{\leftarrow}\big(J_\epsilon(f(a))\big)$ contains a neighborhood U of a. Since a is an interior point of U, there exists $\delta > 0$ such that $J_\delta(a) \subseteq U$. Then

$$J_\delta(a) \subseteq U \subseteq f^{\leftarrow}\big(J_\epsilon(f(a))\big)\,.$$

Conversely, suppose that for every $\epsilon > 0$ there exists $\delta > 0$ such that $J_\delta(a) \subseteq f^{\leftarrow}\big(J_\epsilon(f(a))\big)$. Let V be a neighborhood of $f(a)$. Then $J_\epsilon(f(a)) \subseteq V$ for some $\epsilon > 0$. By hypothesis, there exists

$\delta > 0$ such that $J_\delta(a) \subseteq f^{\leftarrow}\big(J_\epsilon(f(a))\big)$. Then $J_\delta(a)$ is a neighborhood of a and $J_\delta(a) \subseteq f^{\leftarrow}(V)$; so f is continuous at a. □

We may use the remarks preceding the proof of 3.1.2 to give a second characterization of continuity at a point. Even though the condition given is algebraic in form, it is best to think of it geometrically. Think of it in terms of distance. In the next corollary read "$|x - a| < \delta$" as "x is within δ units of a" (not as "the absolute value of x minus a is less than δ"). If you follow this advice, statements about continuity, when we get to metric spaces, will sound familiar. Otherwise, everything will appear new and strange.

3.1.3. Corollary. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at a if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that*

$$|x - a| < \delta \implies |f(x) - f(a)| < \epsilon\,. \tag{3.2}$$

Proof. This is just a restatement of proposition 3.1.2 because condition (3.1) holds if and only if

$$x \in J_\delta(a) \implies f(x) \in J_\epsilon(f(a))\,.$$

But x belongs to the interval $J_\delta(a)$ if and only if $|x - a| < \delta$, and $f(x)$ belongs to $J_\epsilon(f(a))$ if and only if $|f(x) - f(a)| < \epsilon$. Thus (3.1) and (3.2) say exactly the same thing. □

Technically, 3.1.1 is the definition of continuity at a point, while 3.1.2 and 3.1.3 are characterizations of this property. Nevertheless, it is not altogether wrong-headed to think of them as alternative (but equivalent) definitions of continuity. It really doesn't matter which one we choose to be *the* definition. Each of them has its uses. For example, consider the result: if f is continuous at a and g is continuous at $f(a)$, then $g \circ f$ is continuous at a (see 3.2.17). The simplest proof of this uses 3.1.1. On the other hand, when we wish to verify that some particular function is continuous at a point (see, for example, 3.2.2), then, usually, 3.1.3 is best. There are other characterizations of continuity (see 3.2.12–3.2.16). Before embarking on a particular problem involving this concept, it is wise to take a few moments to reflect on which choice among the assorted characterizations is likely to produce the simplest and most direct proof. This is a favor both to your reader and to yourself.

3.2. CONTINUITY—AS A GLOBAL PROPERTY

3.2.1. Definition. A function $f\colon \mathbb{R} \to \mathbb{R}$ is CONTINUOUS if it is continuous at every point in its domain.

The purpose of the next few items is to give you practice in showing that particular functions are (or are not) continuous. If you did a lot of this in beginning calculus, you may wish to skip to the paragraph preceding proposition 3.2.12.

3.2.2. Example. The function $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto -2x + 3$ is continuous.

Proof. We use corollary 3.1.3. Let $a \in \mathbb{R}$. Given $\epsilon > 0$ choose $\delta = \frac{1}{2}\epsilon$. If $|x - a| < \delta$,then

$$\begin{aligned} |f(x) - f(a)| &= |(-2x + 3) - (-2a + 3)| \\ &= 2|x - a| < 2\delta = \epsilon. \end{aligned}$$

□

3.2.3. Example. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{for } x \le 0, \\ 1 & \text{for } x > 0 \end{cases}$$

is not continuous at $a = 0$.

PROOF. Use proposition 3.1.2; the denial of the condition given there is that there exists a number ϵ such that for all $\delta > 0$ property 3.1 fails. (See example D.4.5.) Let $\epsilon = \frac{1}{2}$. Then $f^{\leftarrow}\big(J_{1/2}(f(0))\big) = f^{\leftarrow}(J_{1/2}(0)) = f^{\leftarrow}(-\frac{1}{2}, \frac{1}{2}) = (-\infty, 0]$. Clearly this contains no δ-neighborhood of 0. Thus 3.1 is violated, and f is not continuous at 0. □

3.2.4. Example. The function $f : \mathbb{R} \to \mathbb{R} : x \mapsto 5x - 8$ is continuous.

PROOF. Exercise. (Solution Q.3.1.)

3.2.5. Example. The function $f : x \mapsto x^3$ is continuous at the point $a = -1$.

PROOF. Exercise. (Solution Q.3.2.)

3.2.6. Example. The function $f : x \mapsto 2x^2 - 5$ is continuous.

PROOF. Exercise. (Solution Q.3.3.)

3.2.7. Example. Let $f(x) = 7 - 5x$ for all $x \in \mathbb{R}$. Then f is continuous at the point $a = -2$.

PROOF. Problem.

3.2.8. Example. Let $f(x) = \sqrt{|x+2|}$ for all $x \in \mathbb{R}$. Then f is continuous at the point $a = 0$.

PROOF. Problem.

3.2.9. Example. If $f(x) = 3x - 5$ for all $x \in \mathbb{R}$, then f is continuous.

PROOF. Problem.

3.2.10. Example. The function f defined by

$$f(x) = \begin{cases} -x^2 & \text{for } x < 0, \\ x + \frac{1}{10} & \text{for } x \geq 0 \end{cases}$$

is not continuous.

PROOF. Problem.

3.2.11. Example. For $x > 0$ sketch the functions $x \mapsto \sin \frac{1}{x}$ and $x \mapsto x \sin \frac{1}{x}$. Then verify the following.

(a) The function f defined by

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ \sin \frac{1}{x}, & \text{for } x > 0 \end{cases}$$

is not continuous.

(b) The function f defined by

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ x \sin \frac{1}{x}, & \text{for } x > 0 \end{cases}$$

is continuous at 0.

PROOF. Problem.

The continuity of a function f at a point is a LOCAL property; that is, it is entirely determined by the behavior of f in arbitrarily small neighborhoods of the point. The continuity of f, on the other hand, is a GLOBAL property; it can be determined only if we know how f behaves everywhere on its domain. In 3.1.1–3.1.3 we gave three equivalent conditions for local continuity. In 3.2.12–3.2.16 we give equivalent conditions for the corresponding global concept. The next proposition gives the most useful of these conditions; it is the one that becomes the *definition* of continuity in arbitrary topological spaces. It says: a necessary and sufficient condition for f to be continuous is that the inverse image under f of open sets be open. This shows that continuity is a purely TOPOLOGICAL PROPERTY; that is, it is entirely determined by the topologies (families of all open sets) of the domain and the codomain of the function.

3.2.12. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if $f^{\leftarrow}(U)$ is open whenever U is open in $\mathbb{R}$.*

Proof. Exercise. (Solution Q.3.4.)

3.2.13. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if $f^{\leftarrow}(C)$ is a closed set whenever C is a closed subset of $\mathbb{R}$.*

Proof. Problem.

3.2.14. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if*

$$f^{\leftarrow}(B^{\circ}) \subseteq (f^{\leftarrow}(B))^{\circ}$$

for all $B \subseteq \mathbb{R}$.

Proof. Problem.

3.2.15. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if*

$$f^{\rightarrow}(\,\overline{A}\,) \subseteq \overline{f^{\rightarrow}(A)}$$

for all $A \subseteq \mathbb{R}$.

Proof. Problem. *Hint.* This isn't so easy. Use problem 3.2.13 and the fact (see propositions M.1.22 and M.1.23) that for any sets A and B

$$f^{\rightarrow}(f^{\leftarrow}(B)) \subseteq B \quad \text{and} \quad A \subseteq f^{\leftarrow}(f^{\rightarrow}(A)).$$

Show that if f is continuous, then $\overline{A} \subseteq f^{\leftarrow}\big(\overline{f^{\rightarrow}(A)}\big)$. Then apply $f^{\rightarrow}$. For the converse, apply the hypothesis to the set $f^{\leftarrow}(C)$ where C is a closed subset of $\mathbb{R}$. Then apply $f^{\leftarrow}$. □

3.2.16. Proposition. *A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if*

$$\overline{f^{\leftarrow}(B)} \subseteq f^{\leftarrow}(\,\overline{B}\,)$$

for all $B \subseteq \mathbb{R}$.

Proof. Problem.

3.2.17. Proposition. *Let $f, g\colon \mathbb{R} \to \mathbb{R}$. If f is continuous at a and g is continuous at $f(a)$, then the composite function $g \circ f$ is continuous at a.*

Proof. Let W be a neighborhood of $g(f(a))$. We wish to show that the inverse image of W under $g \circ f$ contains a neighborhood of a. Since g is continuous at $f(a)$, the set $g^{\leftarrow}(W)$ contains a neighborhood V of $f(a)$. And since f is continuous at a, the set $f^{\leftarrow}(V)$ contains a neighborhood U of a. Then

$$\begin{aligned}(g \circ f)^{\leftarrow}(W) &= f^{\leftarrow}(g^{\leftarrow}(W)) \\ &\supseteq f^{\leftarrow}(V) \\ &\supseteq U\,,\end{aligned}$$

which is what we wanted. □

This important result has an equally important but entirely obvious consequence.

3.2.18. Corollary. *The composite of two continuous functions is continuous.*

3.2.19. Problem. Give a direct proof of corollary 3.2.18. (That is, give a proof which does not rely on proposition 3.2.17.)

3.3. FUNCTIONS DEFINED ON SUBSETS OF $\mathbb{R}$

The remainder of this chapter is devoted to a small but important technical problem. Thus far the definitions and propositions concerning continuity have dealt with functions whose domain is $\mathbb{R}$. What do we do about functions whose domain is a proper subset of $\mathbb{R}$? After all, many old friends—the functions $x \mapsto \sqrt{x}$, $x \mapsto \frac{1}{x}$, and $x \mapsto \tan x$, for example—have domains which are not all of $\mathbb{R}$. The difficulty is that if we were to attempt to apply proposition 3.1.2 to the square root function $f : x \mapsto \sqrt{x}$ (which would of course be improper since the hypothesis $f : \mathbb{R} \to \mathbb{R}$ is not satisfied), we would come to the unwelcome conclusion that f is not continuous at 0: if $\epsilon > 0$ then the set $f^{\leftarrow}\big(J_\epsilon(f(0))\big) = f^{\leftarrow}\big(J_\epsilon(0)\big) = f^{\leftarrow}(-\epsilon, \epsilon) = [0, \epsilon^2)$ contains no neighborhood of 0 in $\mathbb{R}$.

Now this can't be right. What we must do is to provide an appropriate definition for the continuity of functions whose domains are proper subsets of $\mathbb{R}$. And we wish to do it in such a way that we make as few changes as possible in the resulting propositions.

The source of our difficulty is the demand (in definition 3.1.1) that $f^{\leftarrow}(V)$ contain a neighborhood of the point a—and neighborhoods have been defined only in $\mathbb{R}$. But why $\mathbb{R}$? That is not the domain of our function; the set $A = [0, \infty)$ is. We *should* be talking about neighborhoods in A. So the question we now face is: how should we define neighborhoods in (and open subsets of) proper subsets of $\mathbb{R}$? The best answer is astonishingly simple. An open subset of A is the intersection of an open subset of $\mathbb{R}$ with A.

3.3.1. Definition. Let $A \subseteq \mathbb{R}$. A set U contained in A is OPEN IN A if there exists an open subset V of $\mathbb{R}$ such that $U = V \cap A$. Briefly, the open subsets of A are restrictions to A of open subsets of $\mathbb{R}$. If U is an open subset of A we write $U \overset{\circ}{\subseteq} A$. A NEIGHBORHOOD OF a IN A is an open subset of A which contains a.

3.3.2. Example. The set $[0, 1)$ is an open subset of $[0, \infty)$.

PROOF. Let $V = (-1, 1)$. Then $V \overset{\circ}{\subseteq} \mathbb{R}$ and $[0, 1) = V \cap [0, \infty)$; so $[0, 1) \overset{\circ}{\subseteq} [0, \infty)$. □

Since, as we have just seen, $[0, 1)$ is open in $[0, \infty)$ but is not open in $\mathbb{R}$, there is a possibility for confusion. Openness is not an intrinsic property of a set. When we say that a set is open, the answer to the question "open in what?" must be either clear from context or else specified. Since the topology (that is, collection of open subsets) which a set A inherits from $\mathbb{R}$ is often called the RELATIVE TOPOLOGY on A, emphasis may be achieved by saying that a subset B of A is *relatively open* in A. Thus, for example, we may say that $[0, 1)$ is relatively open in $[0, \infty)$; or we may say that $[0, 1)$ is a relative neighborhood of 0 (or any other point in the interval). The question here is emphasis and clarity, not logic.

3.3.3. Example. The set $\{1\}$ is an open subset of $\mathbb{N}$.

PROOF. Problem.

3.3.4. Example. The set of all rational numbers x such that $x^2 \leq 2$ is an open subset of $\mathbb{Q}$.

PROOF. Problem.

3.3.5. Example. The set of all rational numbers x such that $x^2 \leq 4$ is *not* an open subset of $\mathbb{Q}$.

PROOF. Problem.

3.3.6. Definition. Let $A \subseteq \mathbb{R}$, $a \in A$, and $\epsilon > 0$. The ϵ-NEIGHBORHOOD OF a IN A is $(a - \epsilon, a + \epsilon) \cap A$.

3.3.7. Example. In $\mathbb{N}$ the $\frac{1}{2}$-neighborhood of 1 is $\{1\}$.

PROOF. Since $(1 - \frac{1}{2}, 1 + \frac{1}{2}) \cap \mathbb{N} = (\frac{1}{2}, \frac{3}{2}) \cap \mathbb{N} = \{1\}$, we conclude that the ϵ-neighborhood of 1 is $\{1\}$. □

Now is the time to show that the family of relatively open subsets of A (that is, the relative topology on A) is in fact a topology on A. In particular, we must show that this family is closed under unions and finite intersections.

3.3.8. Proposition. *Let $A \subseteq \mathbb{R}$. Then*

(i) *$\emptyset$ and A are relatively open in A;*
(ii) *if $\mathfrak{U}$ is a family of relatively open sets in A, then $\bigcup \mathfrak{U}$ is relatively open in A; and*
(iii) *if $\mathfrak{U}$ is a finite family of relatively open sets in A, then $\bigcap \mathfrak{U}$ is relatively open in A.*

Proof. Problem.

3.3.9. Example. Every subset of $\mathbb{N}$ is relatively open in $\mathbb{N}$.

Proof. Problem.

Now we are in a position to consider the continuity of functions defined on proper subsets of $\mathbb{R}$.

3.3.10. Definition. Let $a \in A \subseteq \mathbb{R}$ and $f\colon A \to \mathbb{R}$. The function f is CONTINUOUS AT a if $f^{\leftarrow}(V)$ contains a neighborhood of a in A whenever V is a neighborhood of $f(a)$. The function f is CONTINUOUS if it is continuous at each point in its domain.

It is important to notice that we discuss the continuity of a function only at points where it is defined. We will not, for example, make the claim found in so many beginning calculus texts that the function $x \mapsto 1/x$ is discontinuous at zero. Nor will we try to decide whether the sine function is continuous at the Bronx zoo.

The next proposition tells us that the crucial characterization of continuity in terms of inverse images of open sets (see 3.2.12) still holds under the definition we have just given. Furthermore codomains don't matter; that is, it doesn't matter whether we start with open subsets of $\mathbb{R}$ or with sets open in the range of the function.

3.3.11. Proposition. *Let A be a subset of $\mathbb{R}$. A function $f\colon A \to \mathbb{R}$ is continuous if and only if $f^{\leftarrow}(V)$ is open in A whenever V is open in $\operatorname{ran} f$.*

Proof. Problem. *Hint.* Notice that if $W \overset{\circ}{\subseteq} \mathbb{R}$ and $V = W \cap \operatorname{ran} f$, then $f^{\leftarrow}(V) = f^{\leftarrow}(W)$.

3.3.12. Problem. Discuss the changes (if any) that must be made in 3.1.2, 3.1.3, 3.2.17, and 3.2.18, in order to accommodate functions whose domain is not all of $\mathbb{R}$.

3.3.13. Problem. Let $A \subseteq \mathbb{R}$. Explain what "(relatively) closed subset of A" should mean. Suppose further that $B \subseteq A$. Decide how to define "the closure of B in A" and "the interior of B with respect to A". Explain what changes these definitions will require in propositions 3.2.13–3.2.16 so that the results hold for functions whose domains are proper subsets of $\mathbb{R}$.

3.3.14. Problem. Let $A = (0,1) \cup (1,2)$. Define $f\colon A \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{for } 0 < x < 1, \\ 1 & \text{for } 1 < x < 2\,. \end{cases}$$

Is f a continuous function?

3.3.15. Example. The function f defined by

$$f(x) = \frac{1}{x} \qquad \text{for } x \neq 0$$

is continuous.

Proof. Problem.

3.3.16. Example. The function f defined by

$$f(x) = \sqrt{x+2} \qquad \text{for } x \geq -2$$

is continuous at $x = -2$.

Proof. Problem.

3.3.17. Example. The square root function is continuous.

Proof. Problem.

3.3.18. Problem. Define $f\colon (0,1) \to \mathbb{R}$ by setting $f(x) = 0$ if x is irrational and $f(x) = 1/n$ if $x = m/n$ where m and n are natural numbers with no common factors. Where is f continuous?

3.3.19. Example. Inclusion mappings between subsets of $\mathbb{R}$ are continuous. That is, if $A \subseteq B \subseteq \mathbb{R}$, then the inclusion mapping $\iota\colon A \to B\colon a \mapsto a$ is continuous. (See definition L.3.1.)

Proof. Let $U \overset{\circ}{\subseteq} B$. By the definition of the relative topology on B (see 3.3.1), there exists an open subset V of $\mathbb{R}$ such that $U = V \cap B$. Then $\iota^{\leftarrow}(U) = \iota^{\leftarrow}(V \cap B) = V \cap B \cap A = V \cap A \overset{\circ}{\subseteq} A$. Since the inverse image under ι of each open set is open, ι is continuous. □

It is amusing to note how easy it is with the preceding example in hand to show that restrictions of continuous functions are continuous.

3.3.20. Proposition. *Let $A \subseteq B \subseteq \mathbb{R}$. If $f\colon B \to \mathbb{R}$ is continuous, then $f|_A$, the restriction of f to A, is continuous.*

Proof. Recall (see appendix L section L.5) that $f|_A = f \circ \iota$ where ι is the inclusion mapping of A into B. Since f is continuous (by hypothesis) and ι is continuous (by example 3.3.19), their composite $f|_A$ is also continuous (by the generalization of 3.2.18 in 3.3.12). □

We conclude this chapter with the observation that if a continuous function is positive at a point, it is positive nearby.

3.3.21. Proposition. *Let $A \subseteq \mathbb{R}$ and $f\colon A \to \mathbb{R}$ be continuous at the point a. If $f(a) > 0$, then there exists a neighborhood J of a in A such that $f(x) > \frac{1}{2}f(a)$ for all $x \in J$.*

Proof. Problem.

CHAPTER 4

SEQUENCES OF REAL NUMBERS

Sequences are an extremely useful tool in dealing with topological properties of sets in $\mathbb{R}$ and, as we will see later, in general metric spaces. A major goal of this chapter is to illustrate this usefulness by showing how sequences may be used to characterize open sets, closed sets, and continuous functions.

4.1. CONVERGENCE OF SEQUENCES

A SEQUENCE is a function whose domain is the set $\mathbb{N}$ of natural numbers. In this chapter the sequences we consider will all be sequences of real numbers, that is, functions from $\mathbb{N}$ into $\mathbb{R}$. If a is a sequence, it is conventional to write its value at a natural number n as a_n rather than as $a(n)$. The sequence itself may be written in a number of ways:

$$a = (a_n)_{n=1}^{\infty} = (a_n) = (a_1, a_2, a_3, \dots).$$

Care should be exercised in using the last of these notations. It would not be clear, for example, whether $(\frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \dots)$ is intended to be the sequence of reciprocals of odd primes (in which case the next term would be $\frac{1}{11}$) or the sequence of reciprocals of odd integers greater than 1 (in which case the next term would be $\frac{1}{9}$). The element a_n in the range of a sequence is the n^{th} TERM of the sequence.

It is important to distinguish in one's thinking between a sequence and its range. Think of a sequence $(x_1, x_2, x_3, \dots)$ as an *ordered* set: there is a first element x_1, and a second element x_2, and so on. The range $\{x_1, x_2, x_3, \dots\}$ is just a set. There is no "first" element. For example, the sequences $(1, 2, 3, 4, 5, 6, \dots)$ and $(2, 1, 4, 3, 6, 5, \dots)$ are different whereas the sets $\{1, 2, 3, 4, 5, 6, \dots\}$ and $\{2, 1, 4, 3, 6, 5, \dots\}$ are exactly the same (both are $\mathbb{N}$).

Remark. Occasionally in the sequel it will be convenient to alter the preceding definition a bit to allow the domain of a sequence to be the set $\mathbb{N} \cup \{0\}$ of all positive integers. It is worth noticing as we go along that this in no way affects the correctness of the results we prove in this chapter.

4.1.1. Definition. A sequence (x_n) of real numbers is EVENTUALLY IN a set A if there exists a natural number n_0 such that $x_n \in A$ whenever $n \geq n_0$.

4.1.2. Example. For each $n \in \mathbb{N}$ let $x_n = 3 + \dfrac{7-n}{2}$. Then the sequence $(x_n)_{n=1}^{\infty}$ is eventually strictly negative; that is, the sequence is eventually in the interval $(-\infty, 0)$.

PROOF. If $n \geq 14$, then $\dfrac{7-n}{2} \leq -\dfrac{7}{2}$ and $x_n = 3+\dfrac{7-n}{2} \leq -\dfrac{1}{2} < 0$. So $x_n \in (-\infty, 0)$ whenever $n \geq 14$. □

4.1.3. Example. For each $n \in \mathbb{N}$ let $x_n = \dfrac{2n-3}{n}$. Then the sequence $(x_n)_{n=1}^{\infty}$ is eventually in the interval $(\frac{3}{2}, 2)$.

PROOF. Problem.

4.1.4. Definition. A sequence $(x_n)_{n=1}^{\infty}$ of real numbers CONVERGES to $a \in \mathbb{R}$ if it is eventually in every ϵ-neighborhood of a. When the sequence converges to a we write

$$x_n \to a \text{ as } n \to \infty. \tag{4.1}$$

These symbols may be read, "x_n approaches a as n gets large." If (x_n) converges to a, the number a is the LIMIT of the sequence (x_n). It would not be proper to refer to *the* limit of a sequence if it were possible for a sequence to converge to two different points. We show now that this cannot happen; limits of sequences are unique.

4.1.5. Proposition. *Let (x_n) be a sequence in $\mathbb{R}$. If $x_n \to a$ and $x_n \to b$ as $n \to \infty$, then $a = b$.*

PROOF. Argue by contradiction. Suppose $a \neq b$, and let $\epsilon = |a - b|$. Then $\epsilon > 0$. Since (x_n) is eventually in $J_{\epsilon/2}(a)$, there exists $n_0 \in \mathbb{N}$ such that $x_n \in J_{\epsilon/2}(a)$ for $n \geq n_0$. That is, $|x_n - a| < \frac{\epsilon}{2}$ for $n \geq n_0$. Similarly, since (x_n) is eventually in $J_{\epsilon/2}(b)$, there is an $m_0 \in \mathbb{N}$ such that $|x_n - b| < \frac{\epsilon}{2}$ for $n \geq m_0$. Now if n is any integer larger than both n_0 and m_0, then

$$\epsilon = |a - b| = |a - x_n + x_n - b| \leq |a - x_n| + |x_n - b| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

But $\epsilon < \epsilon$ is impossible. Therefore, our initial supposition was wrong, and $a = b$. □

Since limits are unique, we may use an alternative notation to (4.1): if (x_n) converges to a we may write

$$\lim_{n\to\infty} x_n = a.$$

(Notice how inappropriate this notation would be if limits were not unique.)

4.1.6. Definition. If a sequence (x_n) does not converge (that is, if there exists no $a \in \mathbb{R}$ such that (x_n) converges to a), then the sequence DIVERGES. Sometimes a divergent sequence (x_n) has the property that it is eventually in every interval of the form (p, ∞) where $p \in \mathbb{N}$. In this case we write

$$x_n \to \infty \text{ as } n \to \infty \qquad \text{or} \qquad \lim_{n\to\infty} x_n = \infty.$$

If a divergent sequence (x_n) is eventually in every interval of the form $(-\infty, -p)$ for $p \in \mathbb{N}$, we write

$$x_n \to -\infty \text{ as } n \to \infty \qquad \text{or} \qquad \lim_{n\to\infty} x_n = -\infty.$$

CAUTION. It is *not* true that every divergent sequence satisfies either $x_n \to \infty$ or $x_n \to -\infty$. See 4.1.9 below.

It is easy to rephrase the definition of convergence of a sequence in slightly different language. The next problem gives two such variants. Sometimes one of these alternative "definitions" is more convenient than 4.1.4.

4.1.7. Problem. Let (x_n) be a sequence of real numbers and $a \in \mathbb{R}$.

(a) Show that $x_n \to a$ if and only if for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $|x_n - a| < \epsilon$ whenever $n \geq n_0$.
(b) Show that $x_n \to a$ if and only if (x_n) is eventually in every neighborhood of a.

4.1.8. Example. The sequence $(\frac{1}{n})_{n=1}^{\infty}$ converges to 0.

PROOF. Let $\epsilon > 0$. Use the *Archimedean property* J.4.1 of the real numbers to choose $N \in \mathbb{N}$ large enough that $N > \frac{1}{\epsilon}$. Then

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{N} < \epsilon.$$

whenever $n \geq N$. □

4.1.9. Example. The sequence $\big((-1)^n\big)_{n=0}^{\infty}$ diverges.

Proof. Argue by contradiction. If we assume that $(-1)^n \to a$ for some $a \in \mathbb{R}$, then there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $|(-1)^n - a| < 1$. Thus for every $n \geq N$

$$\begin{aligned} 2 &= |(-1)^n - (-1)^{n+1}| \\ &= |(-1)^n - a + a - (-1)^{n+1}| \\ &\leq |(-1)^n - a| + |a - (-1)^{n+1}| \\ &< 1 + 1 = 2 \end{aligned}$$

which is not possible. □

4.1.10. Example. The sequence $\left(\dfrac{n}{n+1}\right)_{n=1}^{\infty}$ converges to 1.

Proof. Problem.

4.1.11. Proposition. *Let (x_n) be a sequence in $\mathbb{R}$. Then $x_n \to 0$ if and only if $|x_n| \to 0$.*

Proof. Exercise. (Solution Q.4.1.)

4.2. ALGEBRAIC COMBINATIONS OF SEQUENCES

As you will recall from beginning calculus, one standard way of computing the limit of a complicated expression is to find the limits of the constituent parts of the expression and then combine them algebraically. Suppose, for example, we are given sequences (x_n) and (y_n) and are asked to find the limit of the sequence $(x_n + y_n)$. What do we do? First we try to find the limits of the individual sequences (x_n) and (y_n). Then we add. This process is justified by a proposition that says, roughly, that *the limit of a sum is the sum of the limits.* Limits with respect to other algebraic operations behave similarly.

The aim of the following problem is to develop the theory detailing how limits of sequences interact with algebraic operations.

4.2.1. Problem. Suppose that (x_n) and (y_n) are sequences of real numbers, that $x_n \to a$ and $y_n \to b$, and that $c \in \mathbb{R}$. For the case where a and b are real numbers derive the following:

(a) $x_n + y_n \to a + b$,
(b) $x_n - y_n \to a - b$,
(c) $x_n y_n \to ab$,
(d) $c x_n \to ca$, and
(e) $\dfrac{x_n}{y_n} \to \dfrac{a}{b}$ if $b \neq 0$.

Then consider what happens in case $a = \pm\infty$ or $b = \pm\infty$ (or both). What can you say (if anything) about the limits of the left hand expressions of (a)–(e)? In those cases in which nothing can be said, give examples to demonstrate as many outcomes as possible. For example, if $a = \infty$ and $b = -\infty$, then nothing can be concluded about the limit as n gets large of $x_n + y_n$. All of the following are possible:

(i) $x_n + y_n \to -\infty$. [Let $x_n = n$ and $y_n = -2n$.]
(ii) $x_n + y_n \to \alpha$, where α is any real number. [Let $x_n = \alpha + n$ and $y_n = -n$.]
(iii) $x_n + y_n \to \infty$. [Let $x_n = 2n$ and $y_n = -n$.]
(iv) None of the above. [Let $(x_n) = (1, 2, 5, 6, 9, 10, 13, 14, \dots)$ and $y_n = (0, -3, -4, -7, -8, -11, -12, -15, \dots)$

4.2.2. Example. If $x_n \to a$ in $\mathbb{R}$ and $p \in \mathbb{N}$, then $x_n{}^p \to a^p$.

Proof. Use induction on p. The conclusion obviously holds when $p = 1$. Suppose $x_n{}^{p-1} \to a^{p-1}$. Apply part (c) of the preceding problem:

$$x_n{}^p = x_n{}^{p-1} \cdot x_n \to a^{p-1} \cdot a = a^p.$$

□

4.2.3. Example. If $x_n = \dfrac{2 - 5n + 7n^2 - 6n^3}{4 - 3n + 5n^2 + 4n^3}$ for each $n \in \mathbb{N}$, then $x_n \to -\frac{3}{2}$ as $n \to \infty$.

PROOF. Problem.

4.2.4. Problem. Find $\lim_{n\to\infty}(\sqrt{n^2 + 5n} - n)$.

Another very useful tool in computing limits is the "sandwich theorem", which says that a sequence sandwiched between two other sequences with a common limit has that same limit.

4.2.5. Proposition (Sandwich theorem)**.** *Let a be a real number or one of the symbols $+\infty$ or $-\infty$. If $x_n \to a$ and $z_n \to a$ and if $x_n \le y_n \le z_n$ for all n, then $y_n \to a$.*

PROOF. Problem.

4.2.6. Example. If $x_n = \dfrac{\sin\left(3 + \pi^{n^2}\right)}{n^{3/2}}$ for each $n \in \mathbb{N}$, then $x_n \to 0$ as $n \to \infty$.

PROOF. Problem. *Hint.* Use 4.1.11, 4.1.8, and 4.2.5.

4.2.7. Example. If (x_n) is a sequence in $(0, \infty)$ and $x_n \to a$, then $\sqrt{x_n} \to \sqrt{a}$.

PROOF. Problem. *Hint.* There are two possibilities: treat the cases $a = 0$ and $a > 0$ separately. For the first use problem 4.1.7(a). For the second use 4.2.1(b) and 4.1.11; write $\sqrt{x_n} - \sqrt{a}$ as $|x_n - a|/(\sqrt{x_n} + \sqrt{a})$. Then find an inequality that allows you to use the *sandwich theorem*(proposition 4.2.5).

4.2.8. Example. The sequence $\left(n^{1/n}\right)$ converges to 1.

PROOF. Problem. *Hint.* For each n let $a_n = n^{1/n} - 1$. Apply the *binomial theorem* I.1.17 to $(1 + a_n)^n$ to obtain the inequality $n > \frac{1}{2}n(n-1)\,{a_n}^2$ and hence to conclude that $0 < a_n < \sqrt{\dfrac{2}{n-1}}$ for every n. Use the *sandwich theorem* 4.2.5.

4.3. SUFFICIENT CONDITION FOR CONVERGENCE

An amusing situation sometimes arises in which we know what the limit of a sequence must be *if* it exists, but we have no idea whether the sequence actually converges. Here is an example of this odd behavior. The sequence will be recursively defined. Sequences are said to be RECURSIVELY defined if only the first term or first few terms are given explicitly and the remaining terms are defined by means of the preceding term(s).

Consider the sequence (x_n) defined so that $x_1 = 1$ and for $n \in \mathbb{N}$

$$x_{n+1} = \frac{3(1 + x_n)}{3 + x_n}\,. \tag{4.2}$$

It is obvious that *IF the sequence converges* to a limit, say ℓ, then ℓ must satisfy

$$\ell = \frac{3(1 + \ell)}{3 + \ell}\,.$$

(This is obtained by taking limits as $n \to \infty$ on both sides of (4.2).) Cross-multiplying and solving for ℓ leads us to the conclusion that $\ell = \pm\sqrt{3}$. Since the first term x_1 is positive, equation (4.2) makes it clear that all the terms will be positive; so ℓ cannot be $-\sqrt{3}$. Thus it is entirely clear that if the limit of the sequence exists, it must be $\sqrt{3}$. What is *not* clear is whether the limit exists at all. This indicates how useful it is to know some conditions sufficient to guarantee convergence of a sequence. The next proposition gives one important example of such a condition: it says that bounded monotone sequences converge.

First, the relevant definitions.

4.3.1. Definition. A sequence (x_n) of real numbers is BOUNDED if its range is a bounded subset of $\mathbb{R}$. Another way of saying the same thing: a sequence (x_n) in $\mathbb{R}$ is bounded if there exists a number M such that $|x_n| \le M$ for all $n \in \mathbb{N}$.

4.3.2. Definition. A sequence (a_n) of real numbers is INCREASING if $a_{n+1} \ge a_n$ for every $n \in \mathbb{N}$; it is STRICTLY INCREASING if $a_{n+1} > a_n$ for every n. A sequence is DECREASING if $a_{n+1} \le a_n$ for every n, and is STRICTLY DECREASING if $a_{n+1} < a_n$ for every n. A sequence is MONOTONE if it is either increasing or decreasing.

4.3.3. Proposition. *Every bounded monotone sequence of real numbers converges. In fact, if a sequence is bounded and increasing, then it converges to the least upper bound of its range. Similarly, if it is bounded and decreasing, then it converges to the greatest lower bound of its range.*

PROOF. Let (x_n) be a bounded increasing sequence. Let R be the range of the sequence and ℓ be the least upper bound of R; that is, $\ell = \sup\{x_n : n \in \mathbb{N}\}$. We have seen in example 2.2.7 that the least upper bound of a nonempty subset of $\mathbb{R}$ belongs to the closure of that set. In particular, $\ell \in \overline{R}$. Thus given any $\epsilon > 0$ we can find a number x_{n_0} of R that lies in the interval $J_\epsilon(\ell)$. In fact, $x_{n_0} \in (\ell - \epsilon, \ell]$. Since the sequence is increasing and bounded above by ℓ, we have $x_n \in (\ell - \epsilon, \ell] \subseteq J_\epsilon(\ell)$ for every $n \ge n_0$. What we have just proved is that the sequence (x_n) is eventually in every ϵ-neighborhood of ℓ. That is, $x_n \to \ell$.

If (x_n) is bounded and decreasing, then the sequence $(-x_n)$ is bounded and increasing. If R is the range of (x_n) and g is the greatest lower bound of R, then $-R$ is the range of $(-x_n)$ and $-g$ is the least upper bound of $-R$. By what we have just proved, $-x_n \to -g$. So $x_n \to g$ as desired. □

4.3.4. Example. Let (x_n) be the sequence recursively defined above: $x_1 = 1$ and (4.2) holds for all $n \ge 1$. Then $x_n \to \sqrt{3}$ as $n \to \infty$.

PROOF. We have argued previously that the limit of the sequence is $\sqrt{3}$ *if* it exists. So what we must show is that the sequence does converge. We have thus far one tool—proposition 4.3.3. Thus we *hope* that we will be able to prove that the sequence is bounded and monotone. If the sequence starts at 1, is monotone, and approaches $\sqrt{3}$, then it must be increasing. How do we prove that the sequence is bounded? Bounded by what? We observed earlier that $x_1 \ge 0$ and equation (4.2) then guarantees that all the remaining x_n's will be positive. So 0 is a lower bound for the sequence. If the sequence is increasing and approaches $\sqrt{3}$, then $\sqrt{3}$ will be an upper bound for the terms of the sequence. Thus it appears that the way to make use of proposition 4.3.3 is to prove two things: (1) $0 \le x_n \le \sqrt{3}$ for all n; and (2) (x_n) is increasing. If we succeed in establishing both of these claims, we are done.

Claim 1. $0 \le x_n \le \sqrt{3}$ for all $n \in \mathbb{N}$.

PROOF. We have already observed that $x_n \ge 0$ for all n. We use mathematical induction to show that $x_n \le \sqrt{3}$ for all n. Since $x_1 = 1$, it is clear that $x_1 \le \sqrt{3}$. For the inductive hypothesis suppose that for some particular k we have

$$x_k \le \sqrt{3}\,. \tag{4.3}$$

We wish to show that $x_{k+1} \le \sqrt{3}$. To start, multiply both sides of (4.3) by $3 - \sqrt{3}$. (If you are wondering, "How did you know to do that?" consult the next problem.) This gives us

$$3x_k - \sqrt{3}x_k \le 3\sqrt{3} - 3.$$

Rearrange so that all terms are positive

$$3 + 3x_k \le 3\sqrt{3} + \sqrt{3}x_k = \sqrt{3}(3 + x_k).$$

But then clearly

$$x_{k+1} = \frac{3 + 3x_k}{3 + x_k} \le \sqrt{3},$$

which was to be proved.

Claim 2. The sequence (x_n) is increasing.

Proof. We will show that $x_{n+1} - x_n \geq 0$ for all n. For each n

$$x_{n+1} - x_n = \frac{3+3x_n}{3+x_n} - x_n = \frac{3-{x_n}^2}{3+x_n} \geq 0.$$

(We know that $3 - {x_n}^2 \geq 0$ from claim 1.) □

4.3.5. Problem. A student, Fred R. Dimm, tried on an exam to prove the claims made in example 4.3.4. For the inductive proof of claim 1 that $x_n \leq \sqrt{3}$ he offered the following "proof":

$$\begin{aligned} x_{k+1} &\leq \sqrt{3} \\ \frac{3+3x_k}{3+x_k} &\leq \sqrt{3} \\ 3+3x_k &\leq \sqrt{3}(3+x_k) \\ 3x_k - \sqrt{3}x_k &\leq 3\sqrt{3} - 3 \\ (3-\sqrt{3})x_k &\leq (3-\sqrt{3})\sqrt{3} \\ x_k &\leq \sqrt{3}, \qquad \text{which is true by hypothesis.} \end{aligned}$$

(a) Aside from his regrettable lack of explanation, Fred seems to be making a serious logical error. Explain to poor Fred why his offering is not a proof. *Hint.* What would you say about a proof that $1 = 2$, that goes as follows?

$$\begin{aligned} 1 &= 2 \\ 0 \cdot 1 &= 0 \cdot 2 \\ 0 &= 0, \qquad \text{which is true.} \end{aligned}$$

(b) Now explain why Fred's computation in (a) is really quite useful scratch work, even if it is not a proof. *Hint.* In the correct proof of claim 1, how *might* its author have been inspired to "multiply both sides of 4.3 by $3 - \sqrt{3}$"?

4.3.6. Example. The condition (bounded and monotone) given in proposition 4.3.3, while sufficient to guarantee the convergence of a sequence, is not necessary.

Proof. Problem. (Give an explicit example.)

Expressed as a conditional, proposition 4.3.3 says that if a sequence is bounded and monotone, then it converges. Example 4.3.6 shows that the converse of this conditional is not correct. A partial converse does hold however: if a sequence converges, it must be bounded.

4.3.7. Proposition. *Every convergent sequence in $\mathbb{R}$ is bounded.*

Proof. Exercise. (Solution Q.4.2.)

We will encounter many situations when it is important to know the limit as n gets large of r^n where r is a number in the interval $(-1, 1)$ and the limit of $r^{1/n}$ where r is a number greater than 0. The next two propositions show that the respective limits are always 0 and 1.

4.3.8. Proposition. *If $|r| < 1$, then $r^n \to 0$ as $n \to \infty$. If $|r| > 1$, then (r^n) diverges.*

Proof. Suppose that $|r| < 1$. If $r = 0$, the proof is trivial, so we consider only $r \neq 0$. Since $0 < |r| < 1$ whenever $-1 < r < 0$, proposition 4.1.11 allows us to restrict our attention to numbers lying in the interval $(0, 1)$. Thus we suppose that $0 < r < 1$ and prove that $r^n \to 0$. Let R be the range of the sequence (r^n). That is, let $R = \{r^n : n \in \mathbb{N}\}$. Let $g = \inf R$. Notice that $g \geq 0$ since 0 is a lower bound for R. We use an argument by contradiction to prove that $g = 0$. Assume to the contrary that $g > 0$. Since $0 < r < 1$, it must follow that $gr^{-1} > g$. Since g is the greatest lower bound of R, the number gr^{-1} is not a lower bound for R. Thus there exists a member, say r^k, of

R such that $r^k < gr^{-1}$. But then $r^{k+1} < g$, which contradicts the choice of g as a lower bound of R. We conclude that $g = 0$.

The sequence (r^n) is bounded and decreasing. Thus by proposition 4.3.3 it converges to the greatest lower bound of its range; that is, $r^n \to 0$ as $n \to \infty$.

Now suppose $r > 1$. Again we argue by contradiction. Suppose that (r^n) converges. Then its range R is a bounded subset of $\mathbb{R}$.

Let $\ell = \sup R$. Since $r > 1$, it is clear that $\ell r^{-1} < \ell$. Since ℓ is the least upper bound of R, there exists a number r^k of R such that $r^k > \ell r^{-1}$. Then $r^{k+1} > \ell$, contrary to the choice of ℓ as an upper bound for R.

Finally, suppose that $r < -1$. If (r^n) converges then its range is bounded. In particular, $\{r^{2n} : n \in \mathbb{N}\}$ is bounded. As in the preceding paragraph, this is impossible. □

4.3.9. Proposition. *If $r > 0$, then $r^{1/n} \to 1$ as $n \to \infty$.*

Proof. Problem. *Hint.* Show that $\frac{1}{n} < r < n$ for some natural number n. Then use example 4.2.8 and proposition 4.2.5. (You will also make use of a standard arithmetic fact—one that arises in problem J.4.5—that if $0 < a < b$, then $a^{1/n} < b^{1/n}$ for every natural number n.)

4.3.10. Problem. Find $\displaystyle\lim_{n\to\infty} \frac{5^n + 3n + 1}{7^n - n - 2}$.

4.4. SUBSEQUENCES

As example 4.1.9 shows, boundedness of a sequence of real numbers does not suffice to guarantee convergence. It is interesting to note, however, that although the sequence $\big((-1)^n\big)_{n=1}^{\infty}$ does not converge, it does have subsequences that converge. The odd numbered terms form a constant sequence $(-1, -1, -1, \dots)$, which of course converges. The even terms converge to $+1$. It is an interesting, if not entirely obvious, fact that every bounded sequence has a convergent subsequence. This is a consequence of our next proposition.

Before proving proposition 4.4.3 we discuss the notion of *subsequence*. The basic idea here is simple enough. Let $a = (a_1, a_2, a_3, \dots)$ be a sequence of real numbers, and suppose that we construct a new sequence b by taking some but not necessarily all of the terms of a and listing them in the same order in which they appear in a. Then we say that this new sequence b is a *subsequence* of a. We might, for example, choose every fifth member of a thereby obtaining the subsequence $b = (b_1, b_2, b_3, \dots) = (a_5, a_{10}, a_{15}, \dots)$. The following definition formalizes this idea.

4.4.1. Definition. Let (a_n) be a sequence of real numbers. If $(n_k)_{k=1}^{\infty}$ is a strictly increasing sequence in $\mathbb{N}$, then the sequence

$$(a_{n_k}) = (a_{n_k})_{k=1}^{\infty} = (a_{n_1}, a_{n_2}, a_{n_3}, \dots)$$

is a SUBSEQUENCE of the sequence (a_n).

4.4.2. Example. If $a_k = 2^{-k}$ and $b_k = 4^{-k}$ for all $k \in \mathbb{N}$, then (b_k) is a subsequence of (a_k). Intuitively, this is clear, since the second sequence $(\frac{1}{4}, \frac{1}{16}, \frac{1}{64}, \dots)$ just picks out the even-numbered terms of the first sequence $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots)$. This "picking out" is implemented by the strictly increasing function $n(k) = 2k$ (for $k \in \mathbb{N}$). Thus $b = a \circ n$ since

$$a_{n_k} = a(n(k)) = a(2k) = 2^{-2k} = 4^{-k} = b_k$$

for all k in $\mathbb{N}$.

4.4.3. Proposition. *Every sequence of real numbers has a monotone subsequence.*

Proof. Exercise. *Hint.* A definition may be helpful. Say that a term a_m of a sequence in $\mathbb{R}$ is a PEAK TERM if it is greater than or equal to every succeeding term (that is, if $a_m \geq a_{m+k}$ for all $k \in \mathbb{N}$). There are only two possibilities: either there is a subsequence of the sequence (a_n) consisting of peak terms, or else there is a last peak term in the sequence. (Include in this second case the situation in which there are no peak terms.) Show in each case how to select a monotone subsequence of (a_n). (Solution Q.4.3.)

4.4.4. Corollary. *Every bounded sequence of real numbers has a convergent subsequence.*

PROOF. This is an immediate consequence of propositions 4.3.3 and 4.4.3. □

Our immediate purpose in studying sequences is to facilitate our investigation of the topology of the set of real numbers. We will first prove a key result 4.4.9, usually known as *Bolzano's theorem*, which tells us that bounded infinite subsets of $\mathbb{R}$ have accumulation points. We then proceed to make available for future work sequential characterizations of open sets 4.4.10, closed sets 4.4.12, and continuity of functions 4.4.13.

4.4.5. Definition. A sequence $(A_1, A_2, A_3, \dots)$ of sets is NESTED if $A_{k+1} \subseteq A_k$ for every k.

4.4.6. Proposition. *The intersection of a nested sequence of nonempty closed bounded intervals whose lengths approach* 0 *contains exactly one point.*

PROOF. Problem. *Hint.* Suppose that $J_n = [a_n, b_n] \neq \emptyset$ for each $n \in \mathbb{N}$, that $J_{n+1} \subseteq J_n$ for each n, and that $b_n - a_n \to 0$ as $n \to \infty$. Show that $\cap_{n=1}^{\infty} J_n = \{c\}$ for some $c \in \mathbb{R}$.

4.4.7. Problem. Show by example that the intersection of a nested sequence of nonempty closed intervals may be empty.

4.4.8. Problem. Show by example that proposition 4.4.6 no longer holds if "closed" is replaced by "open".

4.4.9. Proposition (Bolzano's theorem)**.** *Every bounded infinite subset of* $\mathbb{R}$ *has at least one accumulation point.*

PROOF. Let A be a bounded infinite subset of $\mathbb{R}$. Since it is bounded it is contained in some interval $J_0 = [a_0, b_0]$. Let c_0 be the midpoint of J_0. Then at least one of the intervals $[a_0, c_0]$ or $[c_0, b_0]$ contains infinitely many members of A (see O.1.17). Choose one that does and call it J_1. Now divide J_1 in half and choose J_2 to be one of the resulting closed subintervals whose intersection with A is infinite. Proceed inductively to obtain a nested sequence of closed intervals

$$J_0 \supseteq J_1 \supseteq J_2 \supseteq J_3 \supseteq \dots$$

each one of which contains infinitely many points of A. By proposition 4.4.6 the intersection of all the intervals J_k consists of exactly one point c. Every open interval about c contains some interval J_k and hence infinitely many points of A. Thus c is an accumulation point of A. □

4.4.10. Proposition. *A subset* U *of* $\mathbb{R}$ *is open if and only if every sequence that converges to an element of* U *is eventually in* U.

PROOF. Suppose U is open in $\mathbb{R}$. Let (x_n) be a sequence that converges to a point a in U. Since U is a neighborhood of a, (x_n) is eventually in U by 4.1.7(b).

Conversely, suppose that U is not open. Then there is at least one point a of U that is not an interior point of U. Then for each $n \in \mathbb{N}$ there is a point x_n that belongs to $J_{1/n}(a)$ but *not* to U. Then the sequence (x_n) converges to a but no term of the sequence belongs to U. □

4.4.11. Proposition. *A point* b *belongs to the closure of a set* A *in* $\mathbb{R}$ *if and only if there exists a sequence* (a_n) *in* A *that converges to* b.

PROOF. Exercise. (Solution Q.4.4.)

4.4.12. Corollary. *A subset* A *of* $\mathbb{R}$ *is closed if and only if* b *belongs to* A *whenever there is a sequence* (a_n) *in* A *that converges to* b.

4.4.13. Proposition. *Let* $A \subseteq \mathbb{R}$. *A function* $f \colon A \to \mathbb{R}$ *is continuous at* a *if and only if* $f(x_n) \to f(a)$ *whenever* (x_n) *is a sequence in* A *such that* $x_n \to a$.

Proof. Problem.

4.4.14. Problem. Discuss in detail the continuity of algebraic combinations of continuous real valued functions defined on subsets of $\mathbb{R}$. Show, for example, that if functions f and g are continuous at a point a in $\mathbb{R}$, then such combinations as $f+g$ and fg are also continuous at a. What can you say about the continuity of polynomials? *Hint.* Use problem 4.2.1 and proposition 4.4.13.

We conclude this chapter with some problems. The last six of these concern the convergence of recursively defined sequences. Most of these are pretty much like example 4.3.4 and require more in the way of patience than ingenuity.

4.4.15. Problem. If A is a nonempty subset of $\mathbb{R}$ that is bounded above, then there exists an increasing sequence of elements of A that converges to $\sup A$. Similarly, if A is nonempty and bounded below, then there is a decreasing sequence in A converging to $\inf A$.

4.4.16. Problem (Geometric Series)**.** Let $|r| < 1$ and $a \in \mathbb{R}$. For each $n \geq 0$ let $s_n = \sum_{k=0}^{n} ar^k$.

(a) Show that the sequence $(s_n)_{n=0}^{\infty}$ converges. *Hint.* Consider $s_n - rs_n$.
(b) The limit found in part (a) is usually denoted by $\sum_{k=0}^{\infty} ar^k$. (This is the sum of a geometric series.) Use (a) to find $\sum_{k=0}^{\infty} 2^{-k}$.
(c) Show how (a) may be used to write the decimal $0.15242\overline{424}\dots$ as the quotient of two natural numbers.

4.4.17. Exercise. Suppose a sequence (x_n) of real numbers satisfies

$$4x_{n+1} = {x_n}^3$$

for all $n \in \mathbb{N}$. For what values of x_1 does the sequence (x_n) converge? For each such x_1 what is $\lim_{n\to\infty} x_n$? *Hint.* First establish that *if* the sequence (x_n) converges, its limit must be $-2, 0,$ or 2. This suggests looking at several special cases: $x_1 < -2$, $x_1 = -2$, $-2 < x_1 < 0$, $x_1 = 0$, $0 < x_1 < 2$, $x_1 = 2$, and $x_1 > 2$. In case $x_1 < -2$, for example, show that $x_n < -2$ for all n. Use this to show that the sequence (x_n) is decreasing and that it has no limit. The other cases can be treated in a similar fashion. (Solution Q.4.5.)

4.4.18. Problem. Suppose a sequence (x_n) in $\mathbb{R}$ satisfies

$$5x_{n+1} = 3x_n + 4$$

for all $n \in \mathbb{N}$. Show that (x_n) converges. *Hint.* First answer an easy question: If $x_n \to \ell$, what is ℓ? Then look at three cases: $x_1 < \ell$, $x_1 = \ell$, and $x_1 > \ell$. Show, for example, that if $x_1 < \ell$, then (x_n) is bounded and increasing.

4.4.19. Problem. Suppose a sequence (x_n) in $\mathbb{R}$ satisfies

$$x_{n+1} = \sqrt{2 + x_n}$$

for all $n \in \mathbb{N}$. Show that (x_n) converges. To what does it converge?

4.4.20. Problem. Suppose that a sequence (x_n) in $\mathbb{R}$ satisfies

$$x_{n+1} = 1 - \sqrt{1 - x_n}$$

for all $n \in \mathbb{N}$. Show that (x_n) converges. To what does it converge? Does $\left(\dfrac{x_{n+1}}{x_n}\right)$ converge?

4.4.21. Problem. Suppose that a sequence (x_n) of real numbers satisfies

$$3x_{n+1} = {x_n}^3 - 2$$

for all $n \in \mathbb{N}$. For what choices of x_1 does the sequence converge? To what? *Hint.* Compute $x_{n+1} - x_n$.

4.4.22. Problem. Let $a > 1$. Suppose that a sequence (x_n) in $\mathbb{R}$ satisfies

$$(a + x_n)x_{n+1} = a(1 + x_n)$$

for all n. Show that if $x_1 > 0$, then (x_n) converges. In this case find $\lim x_n$.

CHAPTER 5

CONNECTEDNESS AND THE INTERMEDIATE VALUE THEOREM

5.1. CONNECTED SUBSETS OF $\mathbb{R}$

There appears to be no way of finding an exact solution to such an equation as

$$\sin x = x - 1. \tag{5.1}$$

No algebraic trick or trigonometric identity seems to help. To be entirely honest we must ask what reason we have for thinking that (5.1) *has* a solution. True, even a rough sketch of the graphs of the functions $x \mapsto \sin x$ and $x \mapsto x - 1$ seems to indicate that the two curves cross not far from $x = 2$. But is it not possible in the vast perversity of the-way-things-are that the two curves might somehow skip over one another without actually having a point in common? To say that it seems unlikely is one thing, to say that we *know* it can not happen is another. The solution to our dilemma lies in a result familiar from nearly every beginning calculus course: the *intermediate value theorem.* Despite the complicated symbol-ridden formulations in most calculus texts what this theorem really says is that continuous functions from $\mathbb{R}$ to $\mathbb{R}$ take intervals to intervals. In order to prove this fact we will first need to introduce an important topological property of all intervals, *connectedness*. Once we have proved the *intermediate value theorem* not only will we have the intellectual satisfaction of saying that we *know* (5.1) has a solution, but we can utilize this same theorem to find approximations to the solution of (5.1) whose accuracy is limited only by the computational means at our disposal.

So, first we define "connected". Or rather, we define "disconnected", and then agree that sets are connected if they are not disconnected.

5.1.1. Definition. A subset A of $\mathbb{R}$ is DISCONNECTED if there exist disjoint nonempty sets U and V both open in A whose union is A. In this case we say that the sets U and V *disconnect* A. A subset A of $\mathbb{R}$ is CONNECTED if it is not disconnected.

5.1.2. Proposition. *A set $A \subseteq \mathbb{R}$ is connected if and only if the only subsets of A which are both open in A and closed in A are the null set and A itself.*

PROOF. Exercise. (Solution Q.5.1.)

5.1.3. Example. The set $\{1, 4, 8\}$ is disconnected.

PROOF. Let $A = \{1, 4, 8\} \subseteq \mathbb{R}$. Notice that the sets $\{1\}$ and $\{4, 8\}$ are open subsets of A. Reason: $(-\infty, 2) \cap A = \{1\}$ and $(2, \infty) \cap A = \{4, 8\}$; so $\{1\}$ is the intersection of an open subset of $\mathbb{R}$ with A, and so is $\{4, 8\}$ (see definition 3.3.1). Thus $\{1\}$ and $\{4, 8\}$ are disjoint nonempty open subsets of A whose union is A. That is, the sets $\{1\}$ and $\{4, 8\}$ disconnect A. □

5.1.4. Example. The set $\mathbb{Q}$ of rational numbers is disconnected.

PROOF. The sets $\{x \in \mathbb{Q} : x < \pi\}$ and $\{x \in \mathbb{Q} : x > \pi\}$ disconnect $\mathbb{Q}$. □

5.1.5. Example. The set $\{\frac{1}{n} : n \in \mathbb{N}\}$ is disconnected.

PROOF. Problem.

If A is a subset of $\mathbb{R}$, it is somewhat awkward that "connected" is defined in terms of (relatively) open subsets of A. In general, it is a nuisance to deal with relatively open subsets. It would be much more convenient to deal only with the familiar topology of $\mathbb{R}$. Happily, this can be arranged.

It is an elementary observation that A is disconnected if we can find disjoint sets U and V, both open in $\mathbb{R}$, whose intersections with A are nonempty and whose union contains A. (For then the sets $U \cap A$ and $V \cap A$ disconnect A.) Somewhat less obvious is the fact that A is disconnected if and only if it can be written as the union of two nonempty subsets C and D of $\mathbb{R}$ such that

$$C \cap \overline{D} = \overline{C} \cap D = \emptyset. \tag{5.2}$$

(The indicated closures are in $\mathbb{R}$.)

5.1.6. Definition. Two nonempty subsets C and D of $\mathbb{R}$ which satisfy equation (5.2) are said to be MUTUALLY SEPARATED in $\mathbb{R}$.

5.1.7. Proposition. *A subset A of $\mathbb{R}$ is disconnected if and only if it is the union of two nonempty sets mutually separated in $\mathbb{R}$.*

PROOF. If A is disconnected, it can be written as the union of two disjoint nonempty sets U and V which are open in A. (These sets need not, of course, be open in $\mathbb{R}$.) We show that U and V are mutually separated. It suffices to prove that $U \cap \overline{V}$ is empty, that is, $U \subseteq \mathbb{R} \setminus \overline{V}$. To this end suppose that $u \in U$. Since U is open in A, there exists $\delta > 0$ such that

$$A \cap J_\delta(u) = \{x \in A : |x - u| < \delta\} \subseteq U \subseteq \mathbb{R} \setminus V.$$

The interval $J_\delta(u) = (u - \delta, u + \delta)$ is the union of two sets: $A \cap J_\delta(u)$ and $A^c \cap J_\delta(u)$. We have just shown that the first of these belongs to $\mathbb{R} \setminus V$. Certainly the second piece contains no points of A and therefore no points of V. Thus $J_\delta(u) \subseteq \mathbb{R} \setminus V$. This shows that u does not belong to the closure (in $\mathbb{R}$) of the set V; so $u \in \mathbb{R} \setminus \overline{V}$. Since u was an arbitrary point of U, we conclude that $U \subseteq \mathbb{R} \setminus \overline{V}$.

Conversely, suppose that $A = U \cup V$ where U and V are nonempty sets mutually separated in $\mathbb{R}$. To show that the sets U and V disconnect A, we need only show that they are open in A, since they are obviously disjoint.

Let us prove that U is open in A. Let $u \in U$ and notice that since $U \cap \overline{V}$ is empty, u cannot belong to $\overline{V}$.

Thus there exists $\delta > 0$ such that $J_\delta(u)$ is disjoint from V. Then certainly $A \cap J_\delta(u)$ is disjoint from V. Thus $A \cap J_\delta(u)$ is contained in U. Conclusion: U is open in A. □

The importance of proposition 5.1.7 is that it allows us to avoid the definition of "connected", which involves *relatively* open subsets, and replace it with an equivalent condition which refers only to closures in $\mathbb{R}$. There is no important idea here, only a matter of convenience. We will use this characterization in the proof of our next proposition, which identifies the connected subsets of the real line. (To appreciate the convenience that proposition 5.1.7 brings our way, try to prove the next result using only the definition of "connected" and not 5.1.7.) As mentioned earlier we need to know that in the real line the intervals are the connected sets. This probably agrees with your intuition in the matter. It is plausible, and it is true; but it is not obvious, and its proof requires a little thought.

5.1.8. Definition. A subset J of $\mathbb{R}$ is an INTERVAL provided that it satisfies the following condition: if c, $d \in J$ and $c < z < d$, then $z \in J$. (Thus, in particular, the empty set and sets containing a single point are intervals.)

5.1.9. Proposition. *A subset J of $\mathbb{R}$ is connected if and only if it is an interval.*

PROOF. To show that intervals are connected argue by contradiction. Assume that there exists an interval J in $\mathbb{R}$ which is not connected. By proposition 5.1.7 there exist nonempty sets C and D in $\mathbb{R}$ such that $J = C \cup D$ and

$$C \cap \overline{D} = \overline{C} \cap D = \emptyset$$

(closures are taken in $\mathbb{R}$). Choose $c \in C$ and $d \in D$. Without loss of generality suppose $c < d$. Let $A = (-\infty, d) \cap C$ and $z = \sup A$. Certainly $c \leq z \leq d$. Now $z \in \overline{C}$. (We know from example 2.2.7 that z belongs to $\overline{A}$ and therefore to $\overline{C}$.) Furthermore $z \in \overline{D}$. (If $z = d$, then $z \in D \subseteq \overline{D}$. If $z < d$, then the interval (z, d) is contained in J and, since z is an upper bound for A, this interval contains no point of C. Thus $(z, d) \subseteq D$ and $z \in \overline{D}$.) Finally, since z belongs to J it is a member of either C or D. But $z \in C$ implies $z \in C \cap \overline{D} = \emptyset$, which is impossible; and $z \in D$ implies $z \in \overline{C} \cap D = \emptyset$, which is also impossible.

The converse is easy. If J is not an interval, then there exist numbers $c < z < d$ such that c, $d \in J$ and $z \notin J$. It is easy to see that the sets $(-\infty, z) \cap J$ and $(z, \infty) \cap J$ disconnect J. $\square$

5.2. CONTINUOUS IMAGES OF CONNECTED SETS

Some of the facts we prove in this chapter are quite specific to the real line $\mathbb{R}$. When we move to more complicated metric spaces no reasonable analog of these facts is true. For example, even in the plane $\mathbb{R}^2$ nothing remotely like proposition 5.1.9 holds. While it is not unreasonable to guess that the connected subsets of the plane are those in which we can move continuously between any two points of the set without leaving the set, this conjecture turns out to be wrong. The latter property, *arcwise connectedness*, is sufficient for connectedness to hold—but is not necessary. In chapter 17 we will give an example of a connected set which is not arcwise connected.

Despite the fact that some of our results are specific to $\mathbb{R}$, others will turn out to be true in very general settings. The next theorem, for example, which says that continuity preserves connectedness, is true in $\mathbb{R}$, in $\mathbb{R}^n$, in metric spaces, and even in general topological spaces. More important, the same proof works in all these cases! Thus when you get to chapter 17, where connectedness in metric spaces is discussed, you will already know the proof that the continuous image of a connected set is itself connected.

5.2.1. Theorem. *The continuous image of a connected set is connected.*

PROOF. Exercise. *Hint.* Prove the contrapositive. Let $f\colon A \to \mathbb{R}$ be a continuous function where $A \subseteq \mathbb{R}$. Show that if $\operatorname{ran} f$ is disconnected, then so is A. (Solution Q.5.2.)

The important *intermediate value theorem*, is an obvious corollary of the preceding theorem.

5.2.2. Theorem (Intermediate Value Theorem: Conceptual Version)**.** *The continuous image in $\mathbb{R}$ of an interval is an interval.*

PROOF. Obvious from 5.1.9 and 5.2.1. $\square$

A slight variant of this theorem, familiar from beginning calculus, helps explain the *name* of the result. It says that if a continuous real valued function defined on an interval takes on two values, then it takes on every intermediate value, that is, every value between them. It is useful in establishing the existence of solutions to certain equations and also in approximating these solutions.

5.2.3. Theorem (Intermediate Value Theorem: Calculus Text Version)**.** *Let $f\colon J \to \mathbb{R}$ be a continuous function defined on an interval $J \subseteq \mathbb{R}$. If a, $b \in \operatorname{ran} f$ and $a < z < b$, then $z \in \operatorname{ran} f$.*

PROOF. Exercise. (Solution Q.5.3.)

Here is a typical application of the *intermediate value theorem.*

5.2.4. Example. The equation

$$x^{27} + 5x^{13} + x = x^3 + x^5 + \frac{2}{\sqrt{1 + 3x^2}} \tag{5.3}$$

has at least one real solution.

PROOF. Exercise. *Hint.* Consider the function f whose value at x is the left side minus the right side of (5.3). What can you say without much thought about $f(2)$ and $f(-2)$? (Solution Q.5.4.)

As another application of the intermediate value theorem we prove a fixed point theorem. (**Definition:** Let $f\colon S \to S$ be a mapping from a set S into itself. A point c in S is a FIXED POINT of the function f if $f(c) = c$.) The next result is a (very) special case of the celebrated Brouwer fixed point theorem, which says that every continuous map from the closed unit ball of $\mathbb{R}^n$ into itself has a fixed point. The proof of this more general result is rather complicated and will not be given here.

5.2.5. Proposition. *Let $a < b$ in $\mathbb{R}$. Every continuous map of the interval $[a, b]$ into itself has a fixed point.*

PROOF. Exercise. (Solution Q.5.5.)

5.2.6. Example. The equation

$$x^{180} + \frac{84}{1 + x^2 + \cos^2 x} = 119$$

has at least two solutions in $\mathbb{R}$.

PROOF. Problem.

5.2.7. Problem. Show that the equation

$$\frac{1}{\sqrt{4x^2 + x + 4}} - 1 = x - x^5$$

has at least one real solution. Locate such a solution between consecutive integers.

5.2.8. Problem. We return to the problem we discussed at the beginning of this chapter. Use the *intermediate value theorem* to find a solution to the equation

$$\sin x = 1 - x$$

accurate to within 10^{-5}. *Hint.* You may assume, for the purposes of this problem, that the function $x \mapsto \sin x$ is continuous. You will not want to do the computations by hand; write a program for a computer or programmable calculator. Notice to begin with that there is a solution in the interval $[0, 1]$. Divide the interval in half and decide which half, $[0, \frac{1}{2}]$ or $[\frac{1}{2}, 1]$, contains the solution. Then take the appropriate half and divide *it* in half. Proceed in this way until you have achieved the desired accuracy. Alternatively, you may find it convenient to divide each interval into tenths rather than halves.

5.3. HOMEOMORPHISMS

5.3.1. Definition. Two subsets A and B of $\mathbb{R}$ are HOMEOMORPHIC if there exists a continuous bijection f from A onto B such that f^{-1} is also continuous. In this case the function f is a HOMEOMORPHISM.

Notice that if two subsets A and B are homeomorphic, then there is a one-to-one correspondence between the open subsets of A and those of B. In terms of topology, the two sets are identical. Thus if we know that the open intervals $(0, 1)$ and $(3, 7)$ are homeomorphic (see the next problem), then we treat these two intervals for all topological purposes as indistinguishable. (Of course, a concept such as distance is another matter; it is *not* a topological property. When we consider the distance between points of our two intervals, we can certainly distinguish between them: one is four times the length of the other.) A homeomorphism is sometimes called a TOPOLOGICAL ISOMORPHISM.

5.3.2. Problem. Discuss the homeomorphism classes of intervals in $\mathbb{R}$. That is, tell, as generally as you can, which intervals in $\mathbb{R}$ are homeomorphic to which others—and, of course, explain why. It might be easiest to start with some examples. Show that

(a) the open intervals $(0, 1)$ and $(3, 7)$ are homeomorphic; and
(b) the three intervals $(0, 1)$, $(0, \infty)$, and $\mathbb{R}$ are homeomorphic.

Then do some counterexamples. Show that

(c) no two of the intervals $(0,1)$, $(0,1]$, and $[0,1]$ are homeomorphic.

Hint. For (a) consider the function $x \mapsto 4x+3$. For part of (c) suppose that $f\colon (0,1] \to (0,1)$ is a homeomorphism. What can you say about the restriction of f to $(0,1)$?

When you feel comfortable with the examples, then try to prove more general statements. For example, show that any two bounded open intervals are homeomorphic.

Finally, try to find the *most general* possible homeomorphism classes. (A HOMEOMORPHISM CLASS is a family of intervals any two of which are homeomorphic.)

5.3.3. Problem. Describe the class of all continuous mappings from $\mathbb{R}$ into $\mathbb{Q}$.

CHAPTER 6

COMPACTNESS AND THE EXTREME VALUE THEOREM

One of the most important results in beginning calculus is the *extreme value theorem*: a continuous function on a closed and bounded subset of the real line achieves both a maximum and a minimum value. In the present chapter we prove this result. Central to understanding the *extreme value theorem* is a curious observation: while neither boundedness nor the property of being closed is preserved by continuity (see problems 6.1.1 and 6.1.2), the property of being closed *and* bounded is preserved. Once we have proved this result it is easy to see that a continuous function defined on a closed and bounded set attains a maximum and minimum on the set.

Nevertheless, there are some complications along the way. To begin with, the proof that continuity preserves the property of being closed and bounded turns out to be awkward and unnatural. Furthermore, although this result can be generalized to $\mathbb{R}^n$, it does not hold in more general metric spaces. This suggests—even if it is not conclusive evidence—that we are looking at the wrong concept. One of the mathematical triumphs of the early twentieth century was the recognition that indeed the very concept of closed-and-bounded is a manifestation of the veil of *māyā*, a seductively simple vision which obscures the "real" topological workings behind the scenes. Enlightenment, at this level, consists in piercing this veil of illusion and seeing behind it the "correct" concept—compactness. There is now overwhelming evidence that compactness *is* the appropriate concept. First and most rewarding is the observation that the proofs of the preservation of compactness by continuity and of the *extreme value theorem* now become extremely natural. Also, the same proofs work not only for $\mathbb{R}^n$ but for general metric spaces and even arbitrary topological spaces. Furthermore, the property of compactness is an intrinsic one—if $A \subseteq B \subseteq C$, then A is compact in B if and only if it is compact in C. (The property of being closed is not intrinsic: the interval $(0, 1]$ is closed in $(0, \infty)$ but not in $\mathbb{R}$.) Finally, there is the triumph of products. In the early 1900's there were other contenders for the honored place ultimately held by compactness—sequential compactness and countable compactness. Around 1930 the great Russian mathematician Tychonov was able to show that arbitrary products of compact spaces are compact, a powerful and useful property not shared by the competitors.

There is, however, a price to be paid for the wonders of compactness—a frustratingly unintuitive definition. It is doubtless this lack of intuitive appeal which explains why it took workers in the field so long to come up with the optimal notion. Do not be discouraged if you feel you don't understand the definition. I'm not sure anyone really "understands" it. What is important is to be able to *use* it. And that is quite possible—with a little practice. In this chapter we first define compactness (6.1.3) and give two important examples of compact sets: finite sets (see 6.2.1) and the interval $[0, 1]$ (see 6.2.4). Then we give a number of ways of creating new compact sets from old ones (see 6.2.2, 6.2.8, 6.2.9(b), and 6.3.2). In 6.3.2 we show that the continuous image of a compact set is compact and in 6.3.3 we prove the *extreme value theorem*. Finally (in 6.3.6) we prove the *Heine-Borel theorem* for $\mathbb{R}$: the compact subsets of $\mathbb{R}$ are those which are both closed and bounded.

6.1. COMPACTNESS

6.1.1. Problem. Give an example to show that if f is a continuous real valued function of a real variable and A is a closed subset of $\mathbb{R}$ which is contained in the domain of f, then it is not necessarily the case that $f^{\rightarrow}(A)$ is a closed subset of $\mathbb{R}$.

6.1.2. Problem. Give an example to show that if f is a continuous real valued function of a real variable and A is a bounded subset of the domain of f, then it is not necessarily the case that $f^{\rightarrow}(A)$ is bounded.

6.1.3. Definition. A family $\mathfrak{U}$ of sets is said to COVER a set A if $\bigcup \mathfrak{U} \supseteq A$. The phrases "$\mathfrak{U}$ covers A", "$\mathfrak{U}$ is a cover for A", "$\mathfrak{U}$ is a covering of A", and "A is covered by $\mathfrak{U}$" are used interchangeably. If A is a subset of $\mathbb{R}$ and $\mathfrak{U}$ is a cover for A which consists entirely of open subsets of $\mathbb{R}$, then $\mathfrak{U}$ is an OPEN COVER for A. If $\mathfrak{U}$ is a family of sets which covers A and $\mathfrak{V}$ is a subfamily of $\mathfrak{U}$ which also covers A, then $\mathfrak{V}$ is a SUBCOVER of $\mathfrak{U}$ for A. A subset A of $\mathbb{R}$ is COMPACT if *every* open cover of A has a finite subcover.

6.1.4. Example. Let $A = [0,1]$, $U_1 = (-3, \frac{2}{3})$, $U_2 = (-1, \frac{1}{2})$, $U_3 = (0, \frac{1}{2})$, $U_4 = (\frac{1}{3}, \frac{2}{3})$, $U_5 = (\frac{1}{2}, 1)$, $U_6 = (\frac{9}{10}, 2)$, and $U_7 = (\frac{2}{3}, \frac{3}{2})$. Then the family $\mathfrak{U} = \{U_k \colon 1 \le k \le 7\}$ is an open cover for A (because each U_k is open and $\bigcup_{k=1}^{7} U_k = (-3, 2) \supseteq A$). The subfamily $\mathfrak{V} = \{U_1, U_5, U_6\}$ of $\mathfrak{U}$ is a subcover for A (because $U_1 \cup U_5 \cup U_6 = (-3, 2) \supseteq A$). The subfamily $\mathfrak{W} = \{U_1, U_2, U_3, U_7\}$ of $\mathfrak{U}$ is *not* a subcover for A (because $U_1 \cup U_2 \cup U_3 \cup U_7 = (-3, \frac{2}{3}) \cup (\frac{2}{3}, \frac{3}{2})$ does not contain A).

6.1.5. Problem. Let J be the open unit interval $(0,1)$. For each a let $U_a = \left(a, a + \frac{1}{4}\right)$, and let $\mathfrak{U} = \{U_a \colon 0 \le a \le \frac{3}{4}\}$. Then certainly $\mathfrak{U}$ covers J.

(a) Find a finite subfamily of $\mathfrak{U}$ which covers J.
(b) Explain why a solution to (a) does not suffice to show that J is compact.
(c) Show that J is not compact.

6.2. EXAMPLES OF COMPACT SUBSETS OF $\mathbb{R}$

It is easy to give examples of sets that are *not* compact. For example, $\mathbb{R}$ itself is not compact. To see this consider the family $\mathfrak{U}$ of all open intervals of the form $(-n, n)$ where n is a natural number. Then $\mathfrak{U}$ covers $\mathbb{R}$ (what property of $\mathbb{R}$ guarantees this?); but certainly no finite subfamily of $\mathfrak{U}$ does.

What is usually a lot trickier, because the definition is hard to apply directly, is proving that some particular compact set really is compact. The simplest examples of compact spaces are the finite ones. (See example 6.2.1.) Finding nontrivial examples is another matter.

In this section we guarantee ourselves a generous supply of compact subsets of $\mathbb{R}$ by specifying some rather powerful methods for creating new compact sets from old ones. In particular, we will show that a set is compact if it is

(1) a closed subset of a compact set (6.2.2),
(2) a finite union of compact sets (6.2.9(b)), or
(3) the continuous image of a compact set (6.3.2).

Nevertheless it is clear that we need something to start with. In example 6.2.4 we prove directly from the definition that the closed unit interval $[0,1]$ is compact. It is fascinating to see how this single example together with conditions (1)–(3) above can be used to generate a great variety of compact sets in general metric spaces. This will be done in chapter 15.

6.2.1. Example. Every finite subset of $\mathbb{R}$ is compact.

PROOF. Let $A = \{x_1, x_2, \dots, x_n\}$ be a finite subset of $\mathbb{R}$. Let $\mathfrak{U}$ be a family of open sets which covers A. For each $k = 1, \dots, n$ there is at least one set U_k in $\mathfrak{U}$ which contains x_k. Then the family $\{U_1, \dots, U_n\}$ is a finite subcover of $\mathfrak{U}$ for A. □

6.2.2. Proposition. *Every closed subset of a compact set is compact.*

PROOF. Problem. *Hint.* Let A be a closed subset of a compact set K and $\mathfrak{U}$ be a cover of A by open sets. Consider $\mathfrak{U} \cup \{A^c\}$.

6.2.3. Proposition. *Every compact subset of* $\mathbb{R}$ *is closed and bounded.*

Proof. Exercise. *Hint.* To show that a compact set A is closed, show that its complement is open. To this end let y be an arbitrary point in A^c. For each x in A take disjoint open intervals about x and y. (Solution Q.6.1.)

As we will see later, the preceding result holds in arbitrary metric spaces. In fact, it is true in very general topological spaces. Its converse, the celebrated *Heine-Borel theorem* (see 6.3.6), although also true in $\mathbb{R}^n$, (see 16.4.1) does *not* hold in all metric spaces (see problems 6.3.9, 16.4.7 and 16.4.8). Thus it is important at a conceptual level not to regard the property closed-and-bounded as being identical with compactness.

Now, finally, we give a nontrivial example of a compact space. The proof requires verifying some details, which at first glance may make it seem complicated. The basic idea behind the proof, however, is quite straightforward. It repays close study since it involves an important technique of proof that we will encounter again. The first time you read the proof, try to see its structure, to understand its basic logic. Postpone for a second reading the details which show that the conditions labelled (1) and (2) hold. Try to understand instead why verification of these two conditions is really all we need in order to prove that $[0, 1]$ is compact.

6.2.4. Example. The closed unit interval $[0, 1]$ is compact.

Proof. Let $\mathfrak{U}$ be a family of open subsets of $\mathbb{R}$ which covers $[0, 1]$ and let A be the set of all x in $[0, 1]$ such that the closed interval $[0, x]$ can be covered by finitely many members of $\mathfrak{U}$. It is clear that A is nonempty (since it contains 0), and that if a number y belongs to A then so does any number in $[0, 1]$ less than y. We prove two more facts about A:

(1) If $x \in A$ and $x < 1$, then there exists a number $y > x$ such that $y \in A$.
(2) If y is a number in $[0, 1]$ such that $x \in A$ for all $x < y$, then $y \in A$.

To prove (1) suppose that $x \in A$ and $x < 1$. Since $x \in A$ there exists sets $U_1, \dots, U_n$ in $\mathfrak{U}$ which cover the interval $[0, x]$. The number x belongs to at least one of these sets, say U_1. Since U_1 is an open subset of $\mathbb{R}$, there is an interval (a, b) such that $x \in (a, b) \subseteq U_1$. Since $x < 1$ and $x < b$, there exists a number $y \in (0, 1)$ such that $x < y < b$. From $[0, x] \subseteq \bigcup_{k=1}^{n} U_k$ and $[x, y] \subseteq (a, b) \subseteq U_1$ it follows that $U_1, \dots, U_n$ cover the interval $[0, y]$. Thus $y > x$ and y belongs to A.

The proof of (2) is similar. Suppose that $y \in [0, 1]$ and that $[0, y) \subseteq A$. The case $y = 0$ is trivial so we suppose that $y > 0$. Then y belongs to at least one member of $\mathfrak{U}$, say V. Choose an open interval (a, b) in $[0, 1]$ such that $y \in (a, b) \subseteq V$. Since $a \in A$ there is a finite collection of sets $U_1, \dots, U_n$ in $\mathfrak{U}$ which covers $[0, a]$. Then clearly $\{U_1, \dots, U_n, V\}$ is a cover for $[0, y]$. This shows that y belongs to A.

Finally, let $u = \sup A$. We are done if we can show that $u = 1$. Suppose to the contrary that $u < 1$. Then $[0, u) \subseteq A$. We conclude from (2) that $u \in A$ and then from (1) that there is a point greater than u which belongs to A. This contradicts the choice of u as the supremum of A. ☐

6.2.5. Problem. Let $A = \{0\} \cup \{1/n : n \in \mathbb{N}\}$ and $\mathfrak{U}$ be the family $\{U_n : n \geq 0\}$ where $U_0 = (-1, 0.1)$ and $U_n = \left(\dfrac{5}{6n}, \dfrac{7}{6n}\right)$ for $n \geq 1$.

(a) Find a finite subfamily of $\mathfrak{U}$ which covers A.
(b) Explain why a solution to (a) does not suffice to show that A is compact.
(c) Use the definition of compactness to show that A is compact.
(d) Use proposition 6.2.2 to show that A is compact.

6.2.6. Example. Show that the set $A = \{1/n : n \in \mathbb{N}\}$ is not a compact subset of $\mathbb{R}$.

Proof. Problem.

6.2.7. Problem. Give two proofs that the interval $[0, 1)$ is not compact—one making use of proposition 6.2.3 and one not.

6.2.8. Proposition. *The intersection of a nonempty collection of compact subsets of $\mathbb{R}$ is itself compact.*

Proof. Problem.

6.2.9. Problem. Let $\mathfrak{K}$ be the family of compact subsets of $\mathbb{R}$.

(a) Show that $\bigcup \mathfrak{K}$ need not be compact.
(b) Show that if $\mathfrak{K}$ contains only finitely many sets, then $\bigcup \mathfrak{K}$ is compact.

6.3. THE EXTREME VALUE THEOREM

6.3.1. Definition. A real-valued function f on a set A is said to have a MAXIMUM at a point a in A if $f(a) \geq f(x)$ for every x in A; the number $f(a)$ is the MAXIMUM VALUE of f. The function has a MINIMUM at a if $f(a) \leq f(x)$ for every x in A; and in this case $f(a)$ is the MINIMUM VALUE of f. A number is an EXTREME VALUE of f if it is either a maximum or a minimum value. It is clear that a function may fail to have maximum or minimum values. For example, on the open interval $(0, 1)$ the function $f : x \mapsto x$ assumes neither a maximum nor a minimum.

The concepts we have just defined are frequently called GLOBAL (or ABSOLUTE) MAXIMUM and GLOBAL (or ABSOLUTE) MINIMUM. This is to distinguish them from two different ideas *local* (or *relative*) *maximum* and *local* (or *relative*) *minimum*, which we will encounter later. In this text, "maximum" and "minimum" without qualifiers will be the global concepts defined above.

Our goal now is to show that every continuous function on a compact set attains both a maximum and a minimum. This turns out to be an easy consequence of the fact, which we prove next, that the continuous image of a compact set is compact.

6.3.2. Theorem. *Let A be a subset of $\mathbb{R}$. If A is compact and $f \colon A \to \mathbb{R}$ is continuous, then $f^{\to}(A)$ is compact.*

Proof. Exercise. (Solution Q.6.2.)

6.3.3. Theorem (Extreme Value Theorem)**.** *If A is a compact subset of $\mathbb{R}$ and $f \colon A \to \mathbb{R}$ is continuous, then f assumes both a maximum and a minimum value on A.*

Proof. Exercise. (Solution Q.6.3.)

6.3.4. Example. The closed interval $[-3, 7]$ is a compact subset of $\mathbb{R}$.

Proof. Let $A = [0, 1]$ and $f(x) = 10x - 3$. Since A is compact and f is continuous, theorem 6.3.2 tells us that the set $[-3, 7] = f^{\to}(A)$ is compact. □

6.3.5. Example. If $a < b$, then the closed interval $[a, b]$ is a compact subset of $\mathbb{R}$.

Proof. Problem.

6.3.6. Example (Heine-Borel Theorem for $\mathbb{R}$)**.** Every closed and bounded subset of $\mathbb{R}$ is compact.

Proof. Problem. *Hint.* Use 6.3.5 and 6.2.2.

6.3.7. Example. Define $f \colon [-1, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

The set of all x such that $f(x) = 0$ is a compact subset of $[-1, 1]$.

Proof. Problem.

6.3.8. Problem. Let $f \colon A \to B$ be a continuous bijection between subsets of $\mathbb{R}$.

(a) Show by example that f need not be a homeomorphism.
(b) Show that if A is compact, then f must be a homeomorphism.

6.3.9. Problem. Find in $\mathbb{Q}$ a set which is both relatively closed and bounded but which is not compact.

6.3.10. Problem. Show that the interval $[0, \infty)$ is not compact using each of the following:

(a) the definition of compactness;
(b) proposition 6.2.3;
(c) the *extreme value theorem*.

6.3.11. Problem. Let f and g be two continuous functions mapping the interval $[0, 1]$ into itself. Show that if $f \circ g = g \circ f$, then f and g agree at some point of $[0, 1]$. *Hint.* Argue by contradiction. Show that we may suppose, without loss of generality, that $f(x) - g(x) > 0$ for all x in $[0, 1]$. Now try to show that there is a number $a > 0$ such that $f^n(x) \geq g^n(x) + na$ for every natural number n and every x in $[0, 1]$. (Here $f^n = f \circ f \circ \cdots \circ f$ (n copies of f); and g^n is defined similarly.)

CHAPTER 7

LIMITS OF REAL VALUED FUNCTIONS

In chapter 4 we studied limits of sequences of real numbers. In this very short chapter we investigate limits of real valued functions of a real variable. Our principal result (7.2.3) is a characterization of the continuity of a function f at a point in terms of the limit of f at that point.

Despite the importance of this characterization, there is one crucial difference between a function being continuous at a point and having a limit there. If f is continuous at a, then a must belong to the domain of f. In order for f to have a limit at a, it is not required that a be in the domain of f.

7.1. DEFINITION

To facilitate the definition of "limit" we introduce the notion of a *deleted neighborhood* of a point.

7.1.1. Definition. If $J = (b, c)$ is a neighborhood of a point a (that is, if $b < a < c$), then J^*, the DELETED NEIGHBORHOOD associated with J, is just J with the point a deleted. That is, $J^* = (b, a) \cup (a, c)$. In particular, if $J_\delta(a)$ is the δ-neighborhood of a, then $J^*_\delta(a)$ denotes $(a - \delta, a) \cup (a, a + \delta)$.

7.1.2. Definition. Let A be a subset of $\mathbb{R}$, let $f\colon A \to \mathbb{R}$, let a be an accumulation point of A, and let l be a real number. We say that l is the LIMIT OF f AS x APPROACHES a (or the LIMIT OF f AT a) if: for every $\epsilon > 0$ there exists $\delta > 0$ such that $f(x) \in J_\epsilon(l)$ whenever $x \in A \cap J^*_\delta(a)$.

Using slightly different notation we may write this condition as

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in A)\, 0 < |x - a| < \delta \implies |f(x) - l| < \epsilon\,.$$

If this condition is satisfied we write

$$f(x) \to l \text{ as } x \to a$$

or

$$\lim_{x \to a} f(x) = l\,.$$

(Notice that this last notation is a bit optimistic. It would not make sense if f could have two distinct limits as x approaches a. We will show in proposition 7.1.3 that this cannot happen.)

The first thing we notice about the preceding definition is that the point a at which we take the limit need not belong to the domain A of the function f. Very often in practice it does not. Recall the definition in beginning calculus of the derivative of a function $f\colon \mathbb{R} \to \mathbb{R}$ at a point a. It is the limit as $h \to 0$ of the Newton quotient $\dfrac{f(a+h) - f(a)}{h}$. This quotient is not defined at the point $h = 0$. Nevertheless we may still take the limit as h approaches 0.

Here is another example of the same phenomenon. The function on $(0, \infty)$ defined by $x \mapsto (1 + x)^{1/x}$ is not defined at $x = 0$. But its limit at 0 exists: recall from beginning calculus that $\lim_{x \to 0}(1 + x)^{1/x} = e$.

One last comment about the definition: even if a function f is defined at a point a, the value of f at a is irrelevant to the question of the existence of a limit there. According to the definition we consider only points x satisfying $0 < |x - a| < \delta$. The condition $0 < |x - a|$ says just one thing: $x \neq a$.

7.1.3. Proposition. *Let $f: A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$, and let a be an accumulation point of A. If $f(x) \to b$ as $x \to a$, and if $f(x) \to c$ as $x \to a$, then $b = c$.*

PROOF. Exercise. (Solution Q.7.1.)

7.2. CONTINUITY AND LIMITS

There is a close connection between the existence of a limit of a function at a point a and the continuity of the function at a. In proposition 7.2.3 we state the precise relationship. But first we give two examples to show that in the absence of additional hypotheses neither of these implies the other.

7.2.1. Example. The inclusion function $f: \mathbb{N} \to \mathbb{R}: n \mapsto n$ is continuous (because *every* subset of $\mathbb{N}$ is open in $\mathbb{N}$, and thus every function defined on $\mathbb{N}$ is continuous). But the limit of f exists at no point (because $\mathbb{N}$ has no accumulation points).

7.2.2. Example. Let

$$f(x) = \begin{cases} 0, & \text{for } x \neq 0 \\ 1, & \text{for } x = 0\,. \end{cases}$$

Then $\lim_{x\to 0} f(x)$ exists (and equals 0), but f is not continuous at $x = 0$.

We have shown in the two preceding examples that a function f may be continuous at a point a without having a limit there and that it may have a limit at a without being continuous there. If we require the point a to belong to the domain of f and to be an accumulation point of the domain of f (these conditions are independent!), then a necessary and sufficient condition for f to be continuous at a is that the limit of f as x approaches a (exist and) be equal to $f(a)$.

7.2.3. Proposition. *Let $f: A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$, and let $a \in A \cap A'$. Then f is continuous at a if and only if*

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

PROOF. Exercise. (Solution Q.7.2.)

7.2.4. Proposition. *If $f: A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$, and $a \in A'$, then*

$$\lim_{h\to 0} f(a+h) = \lim_{x\to a} f(x)$$

in the sense that if either limit exists, then so does the other and the two limits are equal.

PROOF. Exercise. (Solution Q.7.3.)

7.2.5. Problem. Let $f(x) = 4 - x$ if $x < 0$ and $f(x) = (2+x)^2$ if $x > 0$. Using the definition of "limit" show that $\lim_{x\to 0} f(x)$ exists.

7.2.6. Proposition. *Let $f: A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$ and let $a \in A'$. Then*

$$\lim_{x\to a} f(x) = 0 \qquad \textit{if and only if} \qquad \lim_{x\to a} |f(x)| = 0\,.$$

PROOF. Problem.

7.2.7. Proposition. *Let f, g, $h: A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$, and let $a \in A'$. If $f \le g \le h$ and*

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = l,$$

then $\lim_{x\to a} g(x) = l$.

PROOF. Problem. *Hint.* A slight modification of your proof of proposition 4.2.5 should do the trick.

7.2.8. Problem (Limits of algebraic combinations of functions)**.** Carefully formulate and prove the standard results from beginning calculus on the limits of sums, constant multiples, products, and quotients of functions.

7.2.9. Problem. Let A, $B \subseteq \mathbb{R}$, $a \in A$, $f\colon A \to B$, and $g\colon B \to \mathbb{R}$.

(a) If $l = \lim_{x\to a} f(x)$ exists and g is continuous at l, then

$$\lim_{x\to a}(g \circ f)(x) = g(l)\,.$$

(b) Show by example that the following assertion need not be true: If $l = \lim_{x\to a} f(x)$ exists and $\lim_{y\to l} g(y)$ exists, then $\lim_{x\to a}(g \circ f)(x)$ exists.

CHAPTER 8

DIFFERENTIATION OF REAL VALUED FUNCTIONS

Differential calculus is a highly geometric subject—a fact which is not always made entirely clear in elementary texts, where the study of derivatives as numbers often usurps the place of the fundamental notion of linear approximation. The contemporary French mathematician Jean Dieudonné comments on the problem in chapter 8 of his magisterial multivolume treatise on the *Foundations of Modern Analysis*[3]

> ... the fundamental idea of calculus [is] the "local" approximation of functions by *linear* functions. In the classical teaching of Calculus, this idea is immediately obscured by the accidental fact that, on a one-dimensional vector space, there is a one-to-one correspondence between linear forms and numbers, and therefore the derivative at a point is defined as a *number* instead of a *linear form*. This slavish subservience to the shibboleth of numerical interpretation at any cost becomes much worse when dealing with functions of several variables ...

The goal of this chapter is to display as vividly as possible the geometric underpinnings of the differential calculus. The emphasis is on "tangency" and "linear approximation", not on number.

8.1. THE FAMILIES $\mathfrak{O}$ AND $\mathfrak{o}$

8.1.1. Notation. Let $a \in \mathbb{R}$. We denote by $\mathcal{F}_a$ the family of all real valued functions defined on a neighborhood of a. That is, f belongs to $\mathcal{F}_a$ if there exists an open set U such that $a \in U \subseteq \operatorname{dom} f$.

Notice that for each $a \in \mathbb{R}$, the set $\mathcal{F}_a$ is closed under addition and multiplication. (We define the sum of two functions f and g in $\mathcal{F}_a$ to be the function $f+g$ whose value at x is $f(x)+g(x)$ whenever x belongs to $\operatorname{dom} f \cap \operatorname{dom} g$. A similar convention holds for multiplication.)

Among the functions defined on a neighborhood of zero are two subfamilies of crucial importance; they are $\mathfrak{O}$ (the family of "*big-oh*" functions) and $\mathfrak{o}$ (the family of "*little-oh*" functions).

8.1.2. Definition. A function f in $\mathcal{F}_0$ belongs to $\mathfrak{O}$ if there exist numbers $c > 0$ and $\delta > 0$ such that

$$|f(x)| \leq c\,|x|$$

whenever $|x| < \delta$.

A function f in $\mathcal{F}_0$ belongs to $\mathfrak{o}$ if for every $c > 0$ there exists $\delta > 0$ such that

$$|f(x)| \leq c\,|x|$$

whenever $|x| < \delta$. Notice that f belongs to $\mathfrak{o}$ if and only if $f(0) = 0$ and

$$\lim_{h \to 0} \frac{|f(h)|}{|h|} = 0\,.$$

8.1.3. Example. Let $f(x) = \sqrt{|x|}$. Then f belongs to neither $\mathfrak{O}$ nor $\mathfrak{o}$. (A function belongs to $\mathfrak{O}$ only if in some neighborhood of the origin its graph lies between two lines of the form $y = cx$ and $y = -cx$.)

8.1.4. Example. Let $g(x) = |x|$. Then g belongs to $\mathfrak{O}$ but not to $\mathfrak{o}$.

8.1.5. Example. Let $h(x) = x^2$. Then h is a member of both $\mathfrak{O}$ and $\mathfrak{o}$.

Much of the elementary theory of differential calculus rests on a few simple properties of the families $\mathfrak{O}$ and $\mathfrak{o}$. These are given in propositions 8.1.8–8.1.14.

8.1.6. Definition. A function $L\colon \mathbb{R} \to \mathbb{R}$ is LINEAR if

$$L(x+y) = L(x) + L(y)$$

and

$$L(cx) = cL(x)$$

for all x, y, $c \in \mathbb{R}$. The family of all linear functions from $\mathbb{R}$ into $\mathbb{R}$ will be denoted by $\mathfrak{L}$.

The collection of linear functions from $\mathbb{R}$ into $\mathbb{R}$ is not very impressive, as the next problem shows. When we get to spaces of higher dimension the situation will become more interesting.

8.1.7. Example. A function $f\colon \mathbb{R} \to \mathbb{R}$ is linear if and only if its graph is a (nonvertical) line through the origin.

PROOF. Problem.

CAUTION. Since linear functions must pass through the origin, straight lines are not in general graphs of linear functions.

8.1.8. Proposition. *Every member of $\mathfrak{o}$ belongs to $\mathfrak{O}$; so does every member of $\mathfrak{L}$. Every member of $\mathfrak{O}$ is continuous at 0.*

PROOF. Obvious from the definitions. □

8.1.9. Proposition. *Other than the constant function zero, no linear function belongs to $\mathfrak{o}$.*

PROOF. Exercise. (Solution Q.8.1.)

8.1.10. Proposition. *The family $\mathfrak{O}$ is closed under addition and multiplication by constants.*

PROOF. Exercise. (Solution Q.8.2.)

8.1.11. Proposition. *The family $\mathfrak{o}$ is closed under addition and multiplication by constants.*

PROOF. Problem.

The next two propositions say that the composite of a function in $\mathfrak{O}$ with one in $\mathfrak{o}$ (in either order) ends up in $\mathfrak{o}$.

8.1.12. Proposition. *If $g \in \mathfrak{O}$ and $f \in \mathfrak{o}$, then $f \circ g \in \mathfrak{o}$.*

PROOF. Problem.

8.1.13. Proposition. *If $g \in \mathfrak{o}$ and $f \in \mathfrak{O}$, then $f \circ g \in \mathfrak{o}$.*

PROOF. Exercise. (Solution Q.8.3.)

8.1.14. Proposition. *If $\phi, f \in \mathfrak{O}$, then $\phi f \in \mathfrak{o}$.*

PROOF. Exercise. (Solution Q.8.4.)

Remark. The preceding facts can be summarized rather concisely. (Notation: $\mathcal{C}_0$ is the set of all functions in $\mathcal{F}_0$ which are continuous at 0.)

$$\begin{aligned}
&(1) \qquad \mathfrak{L} \cup \mathfrak{o} \subseteq \mathfrak{O} \subseteq \mathcal{C}_0\,.\\
&(2) \qquad \mathfrak{L} \cap \mathfrak{o} = 0\,.\\
&(3) \qquad \mathfrak{O} + \mathfrak{O} \subseteq \mathfrak{O}\,; \qquad \alpha\,\mathfrak{O} \subseteq \mathfrak{O}\,.\\
&(4) \qquad \mathfrak{o} + \mathfrak{o} \subseteq \mathfrak{o}\,; \qquad \alpha\,\mathfrak{o} \subseteq \mathfrak{o}\,.\\
&(5) \qquad \mathfrak{o} \circ \mathfrak{O} \subseteq \mathfrak{o}\,.\\
&(6) \qquad \mathfrak{O} \circ \mathfrak{o} \subseteq \mathfrak{o}\,.\\
&(7) \qquad \mathfrak{O} \cdot \mathfrak{O} \subseteq \mathfrak{o}\,.
\end{aligned}$$

8.1.15. Problem. Show that $\mathfrak{O} \circ \mathfrak{O} \subseteq \mathfrak{O}$. That is, if $g \in \mathfrak{O}$ and $f \in \mathfrak{O}$, then $f \circ g \in \mathfrak{O}$. (As usual, the domain of $f \circ g$ is taken to be $\{x \colon g(x) \in \operatorname{dom} f\}$.)

8.2. TANGENCY

The fundamental idea of differential calculus is the local approximation of a "smooth" function by a translate of a linear one. Certainly the expression "local approximation" could be taken to mean many different things. One sense of this expression which has stood the test of usefulness over time is "tangency". Two functions are said to be tangent at zero if their difference lies in the family $\mathfrak{o}$. We can of course define tangency of functions at an arbitrary point (see project 8.2.12 below); but for our purposes, "tangency at 0" will suffice. All the facts we need to know concerning this relation turn out to be trivial consequences of the results we have just proved.

8.2.1. Definition. Two functions f and g in $\mathcal{F}_0$ are TANGENT AT ZERO, in which case we write $f \simeq g$, if $f - g \in \mathfrak{o}$.

8.2.2. Example. Let $f(x) = x$ and $g(x) = \sin x$. Then $f \simeq g$ since $f(0) = g(0) = 0$ and $\lim_{x\to 0} \dfrac{x - \sin x}{x} = \lim_{x\to 0} \left(1 - \dfrac{\sin x}{x}\right) = 0$.

8.2.3. Example. If $f(x) = x^2 - 4x - 1$ and $g(x) = \left(3x^2 + 4x - 1\right)^{-1}$, then $f \simeq g$.

PROOF. Exercise. (Solution Q.8.5.)

8.2.4. Proposition. *The relation "tangency at zero" is an equivalence relation on $\mathcal{F}_0$.*

PROOF. Exercise. (Solution Q.8.6.)

The next result shows that at most one linear function can be tangent at zero to a given function.

8.2.5. Proposition. *Let $S, T \in \mathfrak{L}$ and $f \in \mathcal{F}_0$. If $S \simeq f$ and $T \simeq f$, then $S = T$.*

PROOF. Exercise. (Solution Q.8.7.)

8.2.6. Proposition. *If $f \simeq g$ and $j \simeq k$, then $f + j \simeq g + k$, and furthermore, $\alpha f \simeq \alpha g$ for all $\alpha \in \mathbb{R}$.*

PROOF. Problem.

Suppose that f and g are tangent at zero. Under what circumstances are $h \circ f$ and $h \circ g$ tangent at zero? And when are $f \circ j$ and $g \circ j$ tangent at zero? We prove next that sufficient conditions are: h is linear and j belongs to $\mathfrak{O}$.

8.2.7. Proposition. *Let $f, g \in \mathcal{F}_0$ and $T \in \mathfrak{L}$. If $f \simeq g$, then $T \circ f \simeq T \circ g$.*

PROOF. Problem.

8.2.8. Proposition. *Let $h \in \mathfrak{O}$ and $f, g \in \mathcal{F}_0$. If $f \simeq g$, then $f \circ h \simeq g \circ h$.*

PROOF. Problem.

8.2.9. Example. Let $f(x) = 3x^2 - 2x + 3$ and $g(x) = \sqrt{-20x + 25} - 2$ for $x \le 1$. Then $f \simeq g$.

PROOF. Problem.

8.2.10. Problem. Let $f(x) = x^3 - 6x^2 + 7x$. Find a linear function $T \colon \mathbb{R} \to \mathbb{R}$ which is tangent to f at 0.

8.2.11. Problem. Let $f(x) = |x|$. Show that there is no linear function $T \colon \mathbb{R} \to \mathbb{R}$ which is tangent to f at 0.

8.2.12. Problem. Let $T_a\colon x \mapsto x + a$. The mapping T_a is called TRANSLATION BY a. Note that it is *not* linear (unless, of course, $a = 0$). We say that functions f and g in $\mathcal{F}_a$ are TANGENT AT a if the functions $f \circ T_a$ and $g \circ T_a$ are tangent at zero.

(a) Let $f(x) = 3x^2 + 10x + 13$ and $g(x) = \sqrt{-20x - 15}$. Show that f and g are tangent at -2.
(b) Develop a theory for the relationship "tangency at a" which generalizes our work on "tangency at 0".

8.2.13. Problem. Each of the following is an abbreviated version of a proposition. Formulate precisely and prove.

(a) $\mathcal{C}_0 + \mathfrak{O} \subseteq \mathcal{C}_0$.
(b) $\mathcal{C}_0 + \mathfrak{o} \subseteq \mathcal{C}_0$.
(c) $\mathfrak{O} + \mathfrak{o} \subseteq \mathfrak{O}$.

8.2.14. Problem. Suppose that $f \simeq g$. Then the following hold.

(a) If g is continuous at 0, so is f.
(b) If g belongs to $\mathfrak{O}$, so does f.
(c) If g belongs to $\mathfrak{o}$, so does f.

8.3. LINEAR APPROXIMATION

One often hears that differentiation of a smooth function f at a point a in its domain is the process of finding the best "linear approximation" to f at a. This informal assertion is not quite correct. For example, as we know from beginning calculus, the tangent line at $x = 1$ to the curve $y = 4 + x^2$ is the line $y = 2x + 3$, which is not a linear function since it does not pass through the origin. To rectify this rather minor shortcoming we first translate the graph of the function f so that the point $(a, f(a))$ goes to the origin, and *then* find the best linear approximation at the origin. The operation of translation is carried out by a somewhat notorious acquaintance from beginning calculus Δy. The source of its notoriety is two-fold: first, in many texts it is inadequately defined; and second, the notation Δy fails to alert the reader to the fact that under consideration is a function of *two* variables. We will be careful on both counts.

8.3.1. Definition. Let $f \in \mathcal{F}_a$. Define the function Δf_a by

$$\Delta f_a(h) := f(a + h) - f(a)$$

for all h such that $a + h$ is in the domain of f. Notice that since f is defined in a neighborhood of a, the function Δf_a is defined in a neighborhood of 0; that is, Δf_a belongs to $\mathcal{F}_0$. Notice also that $\Delta f_a(0) = 0$.

8.3.2. Problem. Let $f(x) = \cos x$ for $0 \le x \le 2\pi$.

(a) Sketch the graph of the function f.
(b) Sketch the graph of the function Δf_π.

8.3.3. Proposition. *If $f \in \mathcal{F}_a$ and $\alpha \in \mathbb{R}$, then*

$$\Delta(\alpha f)_a = \alpha\,\Delta f_a\,.$$

PROOF. To show that two functions are equal show that they agree at each point in their domain. Here

$$\begin{aligned}\Delta(\alpha f)_a(h) &= (\alpha f)(a + h) - (\alpha f)(a)\\ &= \alpha f(a + h) - \alpha f(a)\\ &= \alpha(f(a + h) - f(a))\\ &= \alpha\,\Delta f_a(h)\end{aligned}$$

for every h in the domain of Δf_a. □

8.3.4. Proposition. *If $f, g \in \mathcal{F}_a$, then*
$$\Delta(f+g)_a = \Delta f_a + \Delta g_a\,.$$

Proof. Exercise. (Solution Q.8.8.)

The last two propositions prefigure the fact that differentiation is a linear operator; the next result will lead to *Leibniz's rule* for differentiating products.

8.3.5. Proposition. *If $\phi, f \in \mathcal{F}_a$, then*
$$\Delta(\phi f)_a = \phi(a)\cdot\Delta f_a \,+\, \Delta\phi_a\cdot f(a) \,+\, \Delta\phi_a\cdot\Delta f_a\,.$$

Proof. Problem.

Finally, we present a result which prepares the way for the *chain rule*.

8.3.6. Proposition. *If $f \in \mathcal{F}_a$, $g \in \mathcal{F}_{f(a)}$, and $g\circ f \in \mathcal{F}_a$, then*
$$\Delta(g\circ f)_a = \Delta g_{f(a)} \circ \Delta f_a\,.$$

Proof. Exercise. (Solution Q.8.9.)

8.3.7. Proposition. *Let $A \subseteq \mathbb{R}$. A function $f\colon A\to\mathbb{R}$ is continuous at the point a in A if and only if Δf_a is continuous at 0.*

Proof. Problem.

8.3.8. Proposition. *If $f\colon U\to U_1$ is a bijection between subsets of $\mathbb{R}$, then for each a in U the function $\Delta f_a\colon U-a\to U_1-f(a)$ is invertible and*
$$\big(\Delta f_a\big)^{-1} = \Delta\big(f^{-1}\big)_{f(a)}\,.$$

Proof. Problem.

8.4. DIFFERENTIABILITY

We now have developed enough machinery to talk sensibly about *differentiating* real valued functions.

8.4.1. Definition. Let $f \in \mathcal{F}_a$. We say that f is differentiable at a if there exists a linear function which is tangent at 0 to Δf_a. If such a function exists, it is called the differential of f at a and is denoted by df_a. (Don't be put off by the slightly complicated notation; df_a is just a member of $\mathfrak{L}$ satisfying $df_a \simeq \Delta f_a$.) We denote by $\mathcal{D}_a$ the family of all functions in $\mathcal{F}_a$ which are differentiable at a.

The next proposition justifies the use of the definite article which modifies "differential" in the preceding paragraph.

8.4.2. Proposition. *Let $f \in \mathcal{F}_a$. If f is differentiable at a, then its differential is unique. (That is, there is at most one linear map tangent at 0 to Δf_a.)*

Proof. Proposition 8.2.5. □

8.4.3. Example. It is instructive to examine the relationship between the differential of f at a, which we defined in 8.4.1, and the derivative of f at a as defined in beginning calculus. For $f \in \mathcal{F}_a$ to be differentiable at a it is necessary that there be a linear function $T\colon\mathbb{R}\to\mathbb{R}$ which is tangent at 0 to Δf_a. According to 8.1.7 there must exist a constant c such that $Tx = cx$ for all x in $\mathbb{R}$. For T to be tangent to Δf_a, it must be the case that
$$\Delta f_a - T \in \mathfrak{o}\,;$$
that is,
$$\lim_{h\to 0}\frac{\Delta f_a(h) - ch}{h} = 0\,.$$

Equivalently,

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=\lim_{h\to 0}\frac{\Delta f_a(h)}{h}=c\,.$$

In other words, the function T, which is tangent to Δf_a at 0, must be a line through the origin whose slope is

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\,.$$

This is, of course, the familiar "derivative of f at a" from beginning calculus. Thus for any real valued function f which is differentiable at a in $\mathbb{R}$

$$df_a(h)=f'(a)\cdot h$$

for all $h\in\mathbb{R}$.

8.4.4. Problem. Explain carefully the quotation from Dieudonné given at the beginning of the chapter.

8.4.5. Example. Let $f(x)=3x^2-7x+5$ and $a=2$. Then f is differentiable at a. (Sketch the graph of the differential df_a.)

Proof. Problem.

8.4.6. Example. Let $f(x)=\sin x$ and $a=\pi/3$. Then f is differentiable at a. (Sketch the graph of the differential df_a.)

Proof. Problem.

8.4.7. Proposition. *Let $T\in\mathfrak{L}$ and $a\in\mathbb{R}$. Then $dT_a=T$.*

Proof. Problem.

8.4.8. Proposition. *If $f\in\mathcal{D}_a$, then $\Delta f_a\in\mathfrak{O}$.*

Proof. Exercise. (Solution Q.8.10.)

8.4.9. Corollary. *Every function which is differentiable at a point is continuous there.*

Proof. Exercise. (Solution Q.8.11.)

8.4.10. Proposition. *If f is differentiable at a and $\alpha\in\mathbb{R}$, then αf is differentiable at a and*

$$d(\alpha f)_a=\alpha\,df_a\,.$$

Proof. Exercise. (Solution Q.8.12.)

8.4.11. Proposition. *If f and g are differentiable at a, then $f+g$ is differentiable at a and*

$$d(f+g)_a=df_a+dg_a\,.$$

Proof. Problem.

8.4.12. Proposition (Leibniz's Rule.)**.** *If $\phi, f\in\mathcal{D}_a$, then $\phi f\in\mathcal{D}_a$ and*

$$d(\phi f)_a=d\phi_a\cdot f(a)\;+\;\phi(a)\,df_a\,.$$

Proof. Exercise. (Solution Q.8.13.)

8.4.13. Theorem (The Chain Rule)**.** *If $f\in\mathcal{D}_a$ and $g\in\mathcal{D}_{f(a)}$, then $g\circ f\in\mathcal{D}_a$ and*

$$d(g\circ f)_a=dg_{f(a)}\circ df_a\,.$$

Proof. Exercise. (Solution Q.8.14.)

8.4.14. Problem (A Problem Set on Functions from $\mathbb{R}$ into $\mathbb{R}$)**.** We are now in a position to derive the standard results, usually contained in the first term of a beginning calculus course, concerning the differentiation of real valued functions of a single real variable. Having at our disposal the machinery developed earlier in this chapter, we may derive these results quite easily; and so the proof of each is a problem.

8.4.15. Definition. If $f \in \mathcal{D}_a$, the DERIVATIVE OF f AT a, denoted by $f'(a)$ or $Df(a)$, is defined to be $\lim_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$. By 7.2.4 this is the same as $\lim_{x\to a} \dfrac{f(x)-f(a)}{x-a}$.

8.4.16. Proposition. *If $f \in \mathcal{D}_a$, then $Df(a) = df_a(1)$.*

PROOF. Problem.

8.4.17. Proposition. *If $f, g \in \mathcal{D}_a$, then*

$$D(fg)(a) = Df(a)\cdot g(a) + f(a)\cdot Dg(a)\,.$$

PROOF. Problem. *Hint.* Use *Leibniz's rule* (8.4.12) and proposition 8.4.16.

8.4.18. Example. Let $r(t) = \dfrac{1}{t}$ for all $t \neq 0$. Then r is differentiable and $Dr(t) = -\dfrac{1}{t^2}$ for all $t \neq 0$.

PROOF. Problem.

8.4.19. Proposition. *If $f \in \mathcal{D}_a$ and $g \in \mathcal{D}_{f(a)}$, then $g \circ f \in \mathcal{D}_a$ and*

$$D(g\circ f)(a) = (Dg)(f(a))\cdot Df(a)\,.$$

PROOF. Problem.

8.4.20. Proposition. *If $f, g \in \mathcal{D}_a$ and $g(a) \neq 0$, then*

$$D\left(\frac{f}{g}\right)(a) = \frac{g(a)\,Df(a) - f(a)\,Dg(a)}{(g(a))^2}\,.$$

PROOF. Problem. *Hint.* Write $\dfrac{f}{g}$ as $f\cdot(r\circ g)$ and use 8.4.16, 8.4.19, and 8.4.18.

8.4.21. Proposition. *If $f \in \mathcal{D}_a$ and $Df(a) > 0$, then there exists $r > 0$ such that*

(i) *$f(x) > f(a)$ whenever $a < x < a + r$, and*
(ii) *$f(x) < f(a)$ whenever $a - r < x < a$.*

PROOF. Problem. *Hint.* Define $g(h) = h^{-1}\,\Delta f_a(h)$ if $h \neq 0$ and $g(0) = Df(a)$. Use proposition 7.2.3 to show that g is continuous at 0. Then apply proposition 3.3.21.

8.4.22. Proposition. *If $f \in \mathcal{D}_a$ and $Df(a) < 0$, then there exists $r > 0$ such that*

(i) *$f(x) < f(a)$ whenever $a < x < a + r$, and*
(ii) *$f(x) > f(a)$ whenever $a - r < x < a$.*

PROOF. Problem. *Hint.* Of course it is possible to obtain this result by doing 8.4.21 again with some inequalities reversed. That is the hard way.

8.4.23. Definition. Let $f\colon A \to \mathbb{R}$ where $A \subseteq \mathbb{R}$. The function f has a LOCAL (or RELATIVE) MAXIMUM at a point $a \in A$ if there exists $r > 0$ such that $f(a) \geq f(x)$ whenever $|x - a| < r$ and $x \in \operatorname{dom} f$. It has a LOCAL (or RELATIVE) MINIMUM at a point $a \in A$ if there exists $r > 0$ such that $f(a) \leq f(x)$ whenever $|x - a| < r$ and $x \in \operatorname{dom} f$.

Recall from chapter 6 that $f\colon A \to \mathbb{R}$ is said to attain a MAXIMUM at a if $f(a) \geq f(x)$ for all $x \in \operatorname{dom} f$. This is often called a GLOBAL (or ABSOLUTE) MAXIMUM to help distinguish it from the local version just defined. It is clear that every global maximum is also a local maximum but not *vice versa.* (Of course a similar remark holds for minima.)

8.4.24. Proposition. *If $f \in \mathcal{D}_a$ and f has either a local maximum or a local minimum at a, then $Df(a) = 0$.*

Proof. Problem. *Hint.* Use propositions 8.4.21 and 8.4.22.)

8.4.25. Proposition (Rolle's Theorem)**.** *Let $a < b$. If $f : [a, b] \to \mathbb{R}$ is continuous, if it is differentiable on (a, b), and if $f(a) = f(b)$, then there exists a point c in (a, b) such that $Df(c) = 0$.*

Proof. Problem. *Hint.* Argue by contradiction. Use the *extreme value theorem* 6.3.3 and proposition 8.4.24.

8.4.26. Theorem (Mean Value Theorem)**.** *Let $a < b$. If $f : [a, b] \to \mathbb{R}$ is continuous and if it is differentiable on (a, b), then there exists a point c in (a, b) such that*

$$Df(c) = \frac{f(b) - f(a)}{b - a} .$$

Proof. Problem. *Hint.* Let $y = g(x)$ be the equation of the line which passes through the points $(a, f(a))$ and $(b, f(b))$. Show that the function $f - g$ satisfies the hypotheses of *Rolle's theorem* (8.4.25)

8.4.27. Proposition. *Let J be an open interval in $\mathbb{R}$. If $f : J \to \mathbb{R}$ is differentiable and $Df(x) = 0$ for every $x \in J$, then f is constant on J.*

Proof. Problem. *Hint.* Use the *mean value theorem* (8.4.26).

CHAPTER 9

METRIC SPACES

Underlying the definition of the principal objects of study in calculus—derivatives, integrals, and infinite series—is the notion of "limit". What we mean when we write

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

is that $f(x)$ can be made arbitrarily close to L by choosing x sufficiently close to a. To say what we mean by "closeness" we need the notion of the distance between two points. In this chapter we study "distance functions", also known as "metrics".

In the preceding chapters we have looked at such topics as limits, continuity, connectedness, and compactness from the point of view of a single example, the real line $\mathbb{R}$, where the distance between two points is the absolute value of their difference. There are other familiar distance functions (the Euclidean metric in the plane $\mathbb{R}^2$ or in three-space $\mathbb{R}^3$, for example, where the distance between points is usually taken to be the square root of the sum of the squares of the differences of their coordinates), and there are many less familiar ones which are also useful in analysis. Each of these has its own distinctive properties which merit investigation. But that would be a poor place to start. It is easier to study first those properties they have in common. We list four conditions we may reasonably expect any distance function to satisfy. If x and y are points, then the distance between x and y should be:

(i) greater than or equal to zero;
(ii) greater than zero if x and y are distinct;
(iii) the same as the distance between y and x; and
(iv) no larger than the sum of the distances produced by taking a detour through a point z.

We formalize these conditions to define a "metric" on a set.

9.1. DEFINITIONS

9.1.1. Definition. Let M be a nonempty set. A function $d\colon M \times M \to \mathbb{R}$ is a METRIC (or DISTANCE FUNCTION) on M if for all x, y, $z \in M$

$$\begin{aligned} &(1) \quad d(x,y) = d(y,x) \\ &(2) \quad d(x,y) \le d(x,z) + d(z,y) \\ &(3) \quad d(x,y) = 0 \text{ if and only if } x = y. \end{aligned}$$

If d is a metric on a set M, then we say that the pair (M,d) is a METRIC SPACE.

It is standard practice—although certainly an abuse of language—to refer to "the metric space M" when in fact we mean "the metric space (M,d)". We will adopt this convention when it appears that no confusion will result. We must keep in mind, however, that there are situations where it is clearly inappropriate; if, for example, we are considering two different metrics on the same set M, a reference to "the metric space M" would be ambiguous.

In our formal definition of "metric", what happened to condition (i) above, which requires a metric to be nonnegative? It is an easy exercise to show that it is implied by the remaining conditions.

9.1.2. Proposition. *If d is a metric on a set M, then $d(x,y) \ge 0$ for all x, $y \in M$.*

PROOF. Exercise. (Solution Q.9.1.)

9.1.3. Definition. For each point a in a metric space (M, d) and each number $r > 0$ we define $B_r(a)$, the OPEN BALL about a of radius r, to be the set of all those points whose distance from a is less than r. That is,

$$B_r(a) := \{x \in M : d(x, a) < r\}.$$

9.2. EXAMPLES

9.2.1. Example. The absolute value of the difference of two numbers is a metric on $\mathbb{R}$. We will call this the usual metric on $\mathbb{R}$. Notice that in this metric the open ball about a of radius r is just the open interval $(a-r, a+r)$. (Proof: $x \in B_r(a)$ if and only if $d(x, a) < r$ if and only if $|x - a| < r$ if and only if $a - r < x < a + r$.)

9.2.2. Problem. Define $d(x, y) = |\arctan x - \arctan y|$ for all x, $y \in \mathbb{R}$.

(a) Show that d is a metric on $\mathbb{R}$.
(b) Find $d(-1, \sqrt{3})$.
(c) Solve the equation $d(x, 0) = d(x, \sqrt{3})$.

9.2.3. Problem. Let $f(x) = \dfrac{1}{1+x}$ for all $x \geq 0$. Define a metric d on $[0, \infty)$ by $d(x, y) = |f(x) - f(y)|$. Find a point $z \neq 1$ in this space whose distance from 2 is equal to the distance between 1 and 2.

9.2.4. Example. Define $d(x, y) = |x^2 - y^2|$ for all x, $y \in \mathbb{R}$. Then d is *not* a metric on $\mathbb{R}$.

PROOF. Problem.

9.2.5. Example. Let $f(x) = \dfrac{x}{1+x^2}$ for $x \geq 0$. Define a function d on $[0, \infty) \times [0, \infty)$ by $d(x, y) = |f(x) - f(y)|$. Then d is *not* a metric on $[0, \infty)$.

PROOF. Problem.

For our next example we make use of a (special case of) an important fact known as *Minkowski's inequality*. This we derive from another standard result, *Schwarz's inequality*.

9.2.6. Proposition (Schwarz's Inequality)**.** *Let $u_1, \dots, u_n, v_1, \dots, v_n \in \mathbb{R}$. Then*

$$\left(\sum_{k=1}^{n} u_k v_k\right)^2 \leq \left(\sum_{k=1}^{n} {u_k}^2\right)\left(\sum_{k=1}^{n} {v_k}^2\right).$$

PROOF. To simplify notation make some abbreviations: let $a = \sum_{k=1}^{n} {u_k}^2$, $b = \sum_{k=1}^{n} {v_k}^2$, and $c = \sum_{k=1}^{n} u_k v_k$. Then

$$\begin{aligned} 0 &\leq \sum_{k=1}^{n} \left(\sqrt{b}\, u_k - \frac{c}{\sqrt{b}}\, v_k\right)^2 \\ &= ab - 2c^2 + c^2 \\ &= ab - c^2. \end{aligned}$$

□

9.2.7. Proposition (Minkowski's Inequality)**.** *Let $u_1, \dots, u_n, v_1, \dots, v_n \in \mathbb{R}$. Then*

$$\left(\sum_{k=1}^{n} (u_k + v_k)^2\right)^{\frac{1}{2}} \leq \left(\sum_{k=1}^{n} {u_k}^2\right)^{\frac{1}{2}} + \left(\sum_{k=1}^{n} {v_k}^2\right)^{\frac{1}{2}}.$$

PROOF. Let a, b, and c be as in 9.2.6. Then

$$\begin{aligned}\sum_{k=1}^{n}(u_k+v_k)^2 &= a+2c+b \\ &\le a+2|c|+b \\ &\le a+2\sqrt{ab}+b \qquad \text{(by 9.2.6)} \\ &= \left(\sqrt{a}+\sqrt{b}\right)^2.\end{aligned}$$

□

9.2.8. Example. For points $x=(x_1,\dots,x_n)$ and $y=(y_1,\dots,y_n)$ in $\mathbb{R}^n$ let

$$d(x,y) := \left(\sum_{k=1}^{n}(x_k-y_k)^2\right)^{\frac{1}{2}}.$$

Then d is the USUAL (or EUCLIDEAN) METRIC on $\mathbb{R}^n$. The only nontrivial part of the proof that d is a metric is the verification of the *triangle inequality* (that is, condition (2) of the definition):

$$d(x,y) \le d(x,z)+d(z,y).$$

To accomplish this let $x=(x_1,\dots,x_n)$, $y=(y_1,\dots,y_n)$, and $z=(z_1,\dots,z_n)$ be points in $\mathbb{R}^n$. Apply *Minkowski's inequality* (9.2.7) with $u_k=x_k-z_k$ and $v_k=z_k-y_k$ for $1\le k\le n$ to obtain

$$\begin{aligned}d(x,y) &= \left(\sum_{k=1}^{n}(x_k-y_k)^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{k=1}^{n}\big((x_k-z_k)+(z_k-y_k)\big)^2\right)^{\frac{1}{2}} \\ &\le \left(\sum_{k=1}^{n}(x_k-z_k)^2\right)^{\frac{1}{2}}+\left(\sum_{k=1}^{n}(z_k-y_k)^2\right)^{\frac{1}{2}} \\ &= d(x,z)+d(z,y).\end{aligned}$$

9.2.9. Problem. Let d be the usual metric on $\mathbb{R}^2$.

(a) Find $d(x,y)$ when $x=(3,-2)$ and $y=(-3,1)$.
(b) Let $x=(5,-1)$ and $y=(-3,-5)$. Find a point z in $\mathbb{R}^2$ such that $d(x,y)=d(y,z)=d(x,z)$.
(c) Sketch $B_r(a)$ when $a=(0,0)$ and $r=1$.

The Euclidean metric is by no means the only metric on $\mathbb{R}^n$ which is useful. Two more examples follow (9.2.10 and 9.2.12).

9.2.10. Example. For points $x=(x_1,\dots,x_n)$ and $y=(y_1,\dots,y_n)$ in $\mathbb{R}^n$ let

$$d_1(x,y) := \sum_{k=1}^{n}|x_k-y_k|.$$

It is easy to see that d_1 is a metric on $\mathbb{R}^n$. When $n=2$ this is frequently called the TAXICAB METRIC. (Why?)

9.2.11. Problem. Let d_1 be the taxicab metric on $\mathbb{R}^2$ (see 9.2.10).

(a) Find $d_1(x,y)$ where $x=(3,-2)$ and $y=(-3,1)$.
(b) Let $x=(5,-1)$ and $y=(-3,-5)$. Find a point z in $\mathbb{R}^2$ such that $d_1(x,y)=d_1(y,z)=d_1(x,z)$.
(c) Sketch $B_r(a)$ for the metric d_1 when $a=(0,0)$ and $r=1$.

9.2.12. Example. For $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $\mathbb{R}^n$ let

$$d_u(x, y) := \max\{|x_k - y_k| \colon 1 \le k \le n\}\,.$$

Then d_u is a metric on $\mathbb{R}^n$. The triangle inequality is verified as follows:

$$\begin{aligned} |x_k - y_k| &\le |x_k - z_k| + |z_k - y_k| \\ &\le \max\{|x_i - z_i| : 1 \le i \le n\} + \max\{|z_i - y_i| : 1 \le i \le n\} \\ &= d_u(x, z) + d_u(z, y) \end{aligned}$$

whenever $1 \le k \le n$. Thus

$$\begin{aligned} d_u(x, y) &= \max\{|x_k - y_k| \colon 1 \le k \le n\} \\ &\le d_u(x, z) + d_u(z, y)\,. \end{aligned}$$

The metric d_u is called the UNIFORM METRIC. The reason for this name will become clear later.

Notice that on the real line the three immediately preceding metrics agree; the distance between points is just the absolute value of their difference. That is, when $n = 1$ the metrics given in 9.2.8, 9.2.10, and 9.2.12 reduce to the one given in 9.2.1.

9.2.13. Problem. This problem concerns the metric d_u (defined in example 9.2.12) on $\mathbb{R}^2$.

(a) Find $d_u(x, y)$ when $x = (3, -2)$ and $y = (-3, 1)$.
(b) Let $x = (5, -1)$ and $y = (-3, -5)$. Find a point z in $\mathbb{R}^2$ such that $d_u(x, y) = d_u(y, z) = d_u(x, z)$.
(c) Sketch $B_r(a)$ for the metric d_u when $a = (0, 0)$ and $r = 1$.

9.2.14. Example. Let M be any nonempty set. For x, $y \in M$ define

$$d(x, y) = \begin{cases} 1 & \text{if } x \ne y \\ 0 & \text{if } x = y\,. \end{cases}$$

It is easy to see that d is a metric; this is the DISCRETE METRIC on M. Although the discrete metric is rather trivial it proves quite useful in constructing counterexamples.

9.2.15. Example. Let d be the usual metric on $\mathbb{R}^2$ and 0 be the origin. Define a function ρ on $\mathbb{R}^2$ as follows:

$$\rho(x, y) := \begin{cases} d(x, y), & \text{if } x \text{ and } y \text{ are collinear with } 0\text{, and} \\ d(x, 0) + d(0, y), & \text{otherwise.} \end{cases}$$

The function ρ is a metric on $\mathbb{R}^2$. This metric is sometimes called the GREEK AIRLINE METRIC.

PROOF. Problem.

9.2.16. Problem. Let ρ be the Greek airline metric on $\mathbb{R}^2$.

(a) Let $x = (-1, 2)$, $y = (-3, 6)$, and $z = (-3, 4)$. Find $\rho(x, y)$ and $\rho(x, z)$. Which point, y or z, is closer to x with respect to ρ?
(b) Let $r = 1$. Sketch $B_r(a)$ for the metric ρ when $a = (0, 0)$, $a = (\frac{1}{4}, 0)$, $a = (\frac{1}{2}, 0)$, $a = (\frac{3}{4}, 0)$, $a = (1, 0)$, and $a = (3, 0)$.

9.2.17. Proposition. *Let (M, d) be a metric space and x, y, $z \in M$. Then*

$$|d(x, z) - d(y, z)| \le d(x, y)\,.$$

PROOF. Problem.

9.2.18. Proposition. *If a and b are distinct points in a metric space, then there exists a number $r > 0$ such that $B_r(a)$ and $B_r(b)$ are disjoint.*

PROOF. Problem.

9.2.19. Proposition. *Let a and b be points in a metric space and r, $s > 0$. If c belongs to $B_r(a) \cap B_s(b)$, then there exists a number $t > 0$ such that $B_t(c) \subseteq B_r(a) \cap B_s(b)$.*

PROOF. Problem.

9.2.20. Problem. Let $f(x) = \dfrac{1}{1+x^2}$ for all $x \geq 0$, and define a metric d on the interval $[0, \infty)$ by

$$d(x, y) = |f(x) - f(y)| .$$

(a) With respect to this metric find the point halfway between 1 and 2.
(b) Find the open ball $B_{\frac{3}{10}}(1)$.

9.3. STRONGLY EQUIVALENT METRICS

Ahead of us lie many situations in which it will be possible to replace a computationally complicated metric on some space by a simpler one without affecting the fundamental properties of the space. As it turns out, a sufficient condition for this process to be legitimate is that the two metrics be "strongly equivalent". For the moment we content ourselves with the *definition* of this term and an example; applications will be discussed later when we introduce the weaker notion of "equivalent metrics" (see 11.2.2).

9.3.1. Definition. Two metrics d_1 and d_2 on a set M are STRONGLY EQUIVALENT if there exist numbers α, $\beta > 0$ such that

$$\begin{aligned} d_1(x, y) &\leq \alpha\, d_2(x, y) \quad \text{and} \\ d_2(x, y) &\leq \beta\, d_1(x, y) \end{aligned}$$

for all x and y in M.

9.3.2. Proposition. *On $\mathbb{R}^2$ the three metrics d, d_1, and d_u, defined in examples 9.2.8, 9.2.10, and 9.2.12, are strongly equivalent.*

PROOF. Exercise. *Hint.* First prove that if a, $b \geq 0$, then

$$\max\{a, b\} \leq a + b \leq \sqrt{2}\,\sqrt{a^2 + b^2} \leq 2 \max\{a, b\} .$$

(Solution Q.9.2.)

9.3.3. Problem. Let d and ρ be strongly equivalent metrics on a set M. Then every open ball in the space (M, d) contains an open ball of the space (M, ρ) (and *vice versa*).

9.3.4. Problem. Let a, b, c, $d \in \mathbb{R}$. Establish each of the following:
(a) $(\frac{1}{3}a + \frac{2}{3}b)^2 \leq \frac{1}{3}a^2 + \frac{2}{3}b^2$.
(b) $(\frac{1}{2}a + \frac{1}{3}b + \frac{1}{6}c)^2 \leq \frac{1}{2}a^2 + \frac{1}{3}b^2 + \frac{1}{6}c^2$.
(c) $(\frac{5}{12}a + \frac{1}{3}b + \frac{1}{6}c + \frac{1}{12}d)^2 \leq \frac{5}{12}a^2 + \frac{1}{3}b^2 + \frac{1}{6}c^2 + \frac{1}{12}d^2$.

Hint. If (a), (b), and (c) are all special cases of some general result, it may be easier to give one proof (of the general theorem) rather than three proofs (of the special cases). In each case what can you say about the numbers multiplying a, b, c, and d? Notice that if $x > 0$, then $xy = \sqrt{x}\,(\sqrt{x}\,y)$. Use Schwarz's inequality 9.2.6.

9.3.5. Proposition. *Let (M, d) be a metric space. The function ρ defined on $M \times M$ by*

$$\rho(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

is a metric on M.

PROOF. Problem. *Hint.* Show first that $\dfrac{u}{1+u} \leq \dfrac{v}{1+v}$ whenever $0 \leq u \leq v$.

9.3.6. Problem. In problem 9.3.5 take M to be the real line $\mathbb{R}$ and d to be the usual metric on $\mathbb{R}$ (see 9.2.1).

(a) Find the open ball $B_{\frac{3}{5}}(1)$ in the metric space $(\mathbb{R}, \rho)$.

(b) Show that the metrics d and ρ are *not* strongly equivalent on $\mathbb{R}$.

CHAPTER 10

INTERIORS, CLOSURES, AND BOUNDARIES

10.1. DEFINITIONS AND EXAMPLES

10.1.1. Definition. Let (M, d) be a metric space and M_0 be a nonempty subset of M. If d_0 is the restriction of d to $M_0 \times M_0$, then, clearly, (M_0, d_0) is a metric space. It is a METRIC SUBSPACE of (M, d). In practice the restricted function (often called the INDUCED METRIC) is seldom given a name of its own; one usually writes, "(M_0, d) is a (metric) subspace of (M, d)". When the metric on M is understood, this is further shortened to, "M_0 is a subspace of M".

10.1.2. Example. Let M be $\mathbb{R}^2$ equipped with the usual Euclidean metric d and $M_0 = \mathbb{Q}^2$. The induced metric d_0 agrees with d where they are both defined:

$$d(x, y) = d_0(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. The only difference is that $d_0(x, y)$ is defined only when both x and y have rational coordinates.

10.1.3. Exercise. Regard $M_0 = \{-1\} \cup [0, 4)$ as a subspace of $\mathbb{R}$ under its usual metric. In this subspace find the open balls $B_1(-1)$, $B_1(0)$, and $B_2(0)$. (Solution Q.10.1.)

10.1.4. Definition. Let A be a subset of a metric space M. A point a is an INTERIOR POINT of A if some open ball about a lies entirely in A. The INTERIOR of A, denoted by A°, is the set of all interior points of A. That is,

$$A^\circ := \{x \in M : B_r(x) \subseteq A \text{ for some } r > 0\}.$$

10.1.5. Example. Let M be $\mathbb{R}$ with its usual metric and A be the closed interval $[0, 1]$. Then $A^\circ = (0, 1)$.

10.1.6. Example. Let M be $\mathbb{R}^2$ with its usual metric and A be the unit disk $\{(x, y) \colon x^2 + y^2 \leq 1\}$. Then the interior of A is the open disk $\{(x, y) \colon x^2 + y^2 < 1\}$.

10.1.7. Example. Let $M = \mathbb{R}^2$ with its usual metric and $A = \mathbb{Q}^2$. Then $A^\circ = \emptyset$.

PROOF. No open ball in $\mathbb{R}^2$ contains only points both of whose coordinates are rational. □

10.1.8. Example. Consider the metrics d, d_1, and d_u on $\mathbb{R}^2$. Let $A = \{x \in \mathbb{R}^2 \colon d(x, 0) \leq 1\}$, $A_1 = \{x \in \mathbb{R}^2 \colon d_1(x, 0) \leq 1\}$, and $A_u = \{x \in \mathbb{R}^2 \colon d_u(x, 0) \leq 1\}$. The point $\left(\frac{2}{3}, \frac{3}{8}\right)$ belongs to A° and A_u°, but not to A_1°.

PROOF. Problem.

10.1.9. Definition. A point x in a metric space M is an ACCUMULATION POINT of a set $A \subseteq M$ if every open ball about x contains a point of A distinct from x. (We do *not* require that x belong to A.) We denote the set of all accumulation points of A by A'. This is sometimes called the DERIVED SET of A. The CLOSURE of the set A, denoted by $\overline{A}$, is $A \cup A'$.

10.1.10. Example. Let $\mathbb{R}^2$ have its usual metric and A be $\big[(0, 1) \times (0, 1)\big] \cup \{(2, 3)\} \subseteq \mathbb{R}^2$. Then $A' = [0, 1] \times [0, 1]$ and $\overline{A} = \big([0, 1] \times [0, 1]\big) \cup \{(2, 3)\}$.

10.1.11. Example. The set $\mathbb{Q}^2$ is a subset of the metric space $\mathbb{R}^2$. Every ordered pair of real numbers is an accumulation point of $\mathbb{Q}^2$ since every open ball in $\mathbb{R}^2$ contains (infinitely many) points with both coordinates rational. So the closure of $\mathbb{Q}^2$ is all of $\mathbb{R}^2$.

10.1.12. Definition. The BOUNDARY of a set A in a metric space is the intersection of the closures of A and its complement. We denote it by ∂A. In symbols,

$$\partial A := \overline{A} \cap \overline{A^c}\,.$$

10.1.13. Example. Take M to be $\mathbb{R}$ with its usual metric. If $A = (0,1)$, then $\overline{A} = A' = [0,1]$ and $\overline{A^c} = A^c = (-\infty, 0] \cup [1, \infty)$; so $\partial A = \{0, 1\}$.

10.1.14. Exercise. In each of the following find A°, A', $\overline{A}$, and ∂A.

(a) Let $A = \{\frac{1}{n} : n \in \mathbb{N}\}$. Regard A as a subset of the metric space $\mathbb{R}$.
(b) Let $A = \mathbb{Q} \cap (0, \infty)$. Regard A as a subset of the metric space $\mathbb{R}$.
(c) Let $A = \mathbb{Q} \cap (0, \infty)$. Regard A as a subset of the metric space $\mathbb{Q}$ (where $\mathbb{Q}$ is a subspace of $\mathbb{R}$).

(Solution Q.10.2.)

10.2. INTERIOR POINTS

10.2.1. Lemma. *Let M be a metric space, $a \in M$, and $r > 0$. If $c \in B_r(a)$, then there is a number $t > 0$ such that*

$$B_t(c) \subseteq B_r(a)\,.$$

PROOF. Exercise. (Solution Q.10.3.)

10.2.2. Proposition. *Let A and B be subsets of a metric space.*

(a) *If $A \subseteq B$, then $A^\circ \subseteq B^\circ$.*
(b) $A^{\circ\circ} = A^\circ$. ($A^{\circ\circ}$ means $\big(A^\circ\big)^\circ$.)

PROOF. Exercise. (Solution Q.10.4.)

10.2.3. Proposition. *If A and B are subsets of a metric space, then*

$$(A \cap B)^\circ = A^\circ \cap B^\circ\,.$$

PROOF. Problem.

10.2.4. Proposition. *Let $\mathfrak{A}$ be a family of subsets of a metric space. Then*

(a) $\bigcup\{A^\circ : A \in \mathfrak{A}\} \subseteq \big(\bigcup \mathfrak{A}\big)^\circ$.
(b) *Equality need not hold in* (a).

PROOF. Exercise. (Solution Q.10.5.)

10.2.5. Proposition. *Let $\mathfrak{A}$ be a family of subsets of a metric space. Then*

(a) $\big(\bigcap \mathfrak{A}\big)^\circ \subseteq \bigcap\{A^\circ : A \in \mathfrak{A}\}$.
(b) *Equality need not hold in* (a).

PROOF. Problem.

10.3. ACCUMULATION POINTS AND CLOSURES

In 10.2.1–10.2.5 some of the fundamental properties of the interior operator $A \mapsto A^\circ$ were developed. In the next proposition we study accumulation points. Once their properties are understood it is quite easy to derive the basic facts about the closure operator $A \mapsto \overline{A}$.

10.3.1. Proposition. *Let A and B be subsets of a metric space.*

(a) *If $A \subseteq B$, then $A' \subseteq B'$.*
(b) $A'' \subseteq A'$.
(c) *Equality need not hold in* (b).
(d) $(A \cup B)' = A' \cup B'$.
(e) $(A \cap B)' \subseteq A' \cap B'$.

(f) *Equality need not hold in* (e).

Proof. (a) Let $x \in A'$. Then each open ball about x contains a point of A, hence of B, distinct from x. Thus $x \in B'$.

(b) Let $a \in A''$. If $r > 0$ then $B_r(a)$ contains a point, say b, of A' distinct from a. By lemma 10.2.1 there exists $s > 0$ such that $B_s(b) \subseteq B_r(a)$. Let $t = \min\{s, d(a,b)\}$. Note that $t > 0$. Since $b \in A'$, there is a point $c \in B_t(b) \subseteq B_r(a)$ such that $c \in A$. Since $t \le d(a,b)$, it is clear that $c \ne a$. Thus every open ball $B_r(a)$ contains a point c of A distinct from a. This establishes that $a \in A'$.

(c) Problem.

(d) Problem.

(e) Since $A \cap B \subseteq A$, part (a) implies that $(A \cap B)' \subseteq A'$. Similarly, $(A \cap B)' \subseteq B'$. Conclusion: $(A \cap B)' \subseteq A' \cap B'$.

(f) In the metric space $\mathbb{R}$ let $A = \mathbb{Q}$ and $B = \mathbb{Q}^c$. Then $(A \cap B)' = \emptyset' = \emptyset$ while $A' \cap B' = \mathbb{R} \cap \mathbb{R} = \mathbb{R}$. ☐

10.3.2. Proposition. *Let A and B be subsets of a metric space with $A \subseteq B$. Then*

(a) $\overline{A} \subseteq \overline{B}$.

(b) $\overline{\overline{A}} = \overline{A}$.

Proof. Problem.

10.3.3. Proposition. *If A and B are subsets of a metric space, then*

$$\overline{A \cup B} = \overline{A} \cup \overline{B}\,.$$

Proof. Problem.

10.3.4. Proposition. *Let $\mathfrak{A}$ be a family of subsets of a metric space. Then*

(a) $\bigcup\{\overline{A} : A \in \mathfrak{A}\} \subseteq \overline{\bigcup \mathfrak{A}}$.

(b) *Equality need not hold in* (a).

Proof. Problem.

10.3.5. Proposition. *Let $\mathfrak{A}$ be a family of subsets of a metric space. Then*

(a) $\overline{\bigcap \mathfrak{A}} \subseteq \bigcap\{\overline{A}\colon A \in \mathfrak{A}\}$.

(b) *Equality need not hold in* (a).

Proof. Problem.

10.3.6. Proposition. *Let A be a subset of a metric space. Then*

(a) $\big(A^\circ\big)^c = \overline{A^c}$.

(b) $\big(A^c\big)^\circ = \big(\,\overline{A}\,\big)^c$.

Proof. Problem.

10.3.7. Problem. Use proposition 10.3.6 and proposition 10.2.2 (but not proposition 10.3.1) to give another proof of proposition 10.3.2.

10.3.8. Problem. Use proposition 10.3.6 and proposition 10.2.3 (but not proposition 10.3.1) to give another proof of proposition 10.3.3.

CHAPTER 11

THE TOPOLOGY OF METRIC SPACES

11.1. OPEN AND CLOSED SETS

11.1.1. Definition. A subset A of a metric space M is OPEN in M if $A^\circ = A$. That is, a set is open if it contains an open ball about each of its points. To indicate that A is open in M we write $A \overset{\circ}{\subseteq} M$.

11.1.2. Example. Care must be taken to claim that a particular set is open (or not open) *only* when the metric space in which the set "lives" is clearly understood. For example, the assertion "the set $[0, 1)$ is open" is false if the metric space in question is $\mathbb{R}$. It is true, however, if the metric space being considered is $[0, \infty)$ (regarded as a subspace of $\mathbb{R}$). The reason: In the space $[0, \infty)$ the point 0 is an interior point of $[0, 1)$; in $\mathbb{R}$ it is not.

11.1.3. Example. In a metric space every open ball is an open set. Notice that this is exactly what lemma 10.2.1 says: each point of an open ball is an interior point of that ball.

The fundamental properties of open sets may be deduced easily from information we already possess concerning interiors of sets. Three facts about open sets are given in 11.1.4–11.1.6. The first of these is very simple.

11.1.4. Proposition. *Every nonempty open set is a union of open balls.*

PROOF. Let U be an open set. For each a in U there is an open ball $B(a)$ about a contained in U. Then clearly

$$U = \bigcup\{B_a \colon a \in U\}\,.$$

□

11.1.5. Proposition. *Let M be a metric space.*

(a) *The union of any family of open subsets of M is open.*
(b) *The intersection of any* finite *family of open subsets of M is open.*

PROOF. (a) Let $\mathfrak{U}$ be a family of open subsets of M. Since the interior of a set is always contained in the set, we need only show that $\bigcup\mathfrak{U} \subseteq \left(\bigcup\mathfrak{U}\right)^\circ$. By 10.2.4

$$\begin{aligned}\bigcup\mathfrak{U} &= \bigcup\{U \colon U \in \mathfrak{U}\}\\ &= \bigcup\{U^\circ \colon U \in \mathfrak{U}\}\\ &\subseteq \left(\bigcup\mathfrak{U}\right)^\circ.\end{aligned}$$

(b) It is enough to show that the intersection of *two* open sets is open. Let $U, V \overset{\circ}{\subseteq} M$. Then by 10.2.3

$$(U \cap V)^\circ = U^\circ \cap V^\circ = U \cap V\,.$$

□

11.1.6. Proposition. *The interior of a set A is the largest open set contained in A.* (Precisely: *A° is the union of all the open sets contained in A.*)

PROOF. Exercise. (Solution Q.11.1.)

11.1.7. Definition. A subset A of a metric space is CLOSED if $\overline{A} = A$. That is, a set is closed if it contains all its accumulation points.

11.1.8. Example. As is the case with open sets, care must be taken when affirming or denying that a particular set is closed. It must be clearly understood in which metric space the set "lives". For example the interval $(0, 1]$ is not closed in the metric space $\mathbb{R}$, but it is a closed subset of the metric space $(0, \infty)$ (regarded as a subspace of $\mathbb{R}$).

REMINDER. Recall the remarks made after example 2.2.11: *sets are not like doors or windows; they are not necessarily either open or closed.* One can not show that a set is closed, for example, by showing that it fails to be open.

11.1.9. Proposition. *A subset of a metric space is open if and only if its complement is closed.*

PROOF. Exercise. *Hint.* Use problem 10.3.6. (Solution Q.11.2.)

Facts already proved concerning closures of sets give us one way of dealing with closed sets; the preceding proposition gives us another. To illustrate this, we give two proofs of the next proposition.

11.1.10. Proposition. *The intersection of an arbitrary family of closed subsets of a metric space is closed.*

FIRST PROOF. Let $\mathfrak{A}$ be a family of closed subsets of a metric space. Then $\bigcap \mathfrak{A}$ is the complement of $\bigcup \{A^c \colon A \in \mathfrak{A}\}$. Since each set A^c is open (by 11.1.9), the union of $\{A^c \colon A \in \mathfrak{A}\}$ is open (by 11.1.5(a)); and its complement $\bigcap \mathfrak{A}$ is closed (11.1.9 again). □

SECOND PROOF. Let $\mathfrak{A}$ be a family of closed subsets of a metric space. Since a set is always contained in its closure, we need only show that $\overline{\bigcap \mathfrak{A}} \subseteq \bigcap \mathfrak{A}$. Use problem 10.3.5(a):

$$\begin{aligned} \overline{\bigcap \mathfrak{A}} &\subseteq \bigcap \{\, \overline{A} \colon A \in \mathfrak{A} \,\} \\ &= \bigcap \{\, A \colon A \in \mathfrak{A} \,\} \\ &= \bigcap \mathfrak{A}\,. \end{aligned}$$

□

11.1.11. Problem. The union of a finite family of closed subsets of a metric space is closed.

(a) Prove this assertion using propositions 11.1.5(b) and 11.1.9.
(b) Prove this assertion using problem 10.3.3.

11.1.12. Problem. Give an example to show that the intersection of an arbitrary family of open subsets of a metric space need not be open.

11.1.13. Problem. Give an example to show that the union of an arbitrary family of closed subsets of a metric space need not be closed.

11.1.14. Definition. Let M be a metric space, $a \in M$, and $r > 0$. The CLOSED BALL $C_r(a)$ about a of radius r is $\{x \in M \colon d(a, x) \le r\}$. The SPHERE $S_r(a)$ about a of radius r is $\{x \in M \colon d(a, x) = r\}$.

11.1.15. Problem. Let M be a metric space, $a \in M$, and $r > 0$.

(a) Give an example to show that the closed ball about a of radius r need not be the same as the closure of the open ball about a of radius r. That is, the sets $C_r(a)$ and $\overline{B_r(a)}$ may differ.
(b) Show that every closed ball in M is a closed subset of M.
(c) Show that every sphere in M is a closed subset of M.

11.1.16. Proposition. *In a metric space the closure of a set A is the smallest closed set containing A.* (Precisely: *$\overline{A}$ is the intersection of the family of all closed sets which contain A.*)

PROOF. Problem.

11.1.17. Proposition. *If A is a subset of a metric space, then its boundary ∂A is equal to $\overline{A} \setminus A^\circ$. Thus ∂A is closed.*

Proof. Problem.

11.1.18. Proposition. *Let A be a subset of a metric space M. If A is closed in M or if it is open in M, then $(\partial A)^{\circ} = \emptyset$.*

Proof. Problem.

11.1.19. Problem. Give an example of a subset A of the metric space $\mathbb{R}$ the interior of whose boundary is all of $\mathbb{R}$.

11.1.20. Definition. Let $A \subseteq B \subseteq M$ where M is a metric space. We say that A is DENSE in B if $\overline{A} \supseteq B$. Thus, in particular, A is dense in the space M if $\overline{A} = M$.

11.1.21. Example. The rational numbers are dense in the reals; so are the irrationals.

Proof. That $\overline{Q} = \mathbb{R}$ was proved in 2.2.4. The proof that $\overline{Q^c} = \mathbb{R}$ is similar. □

The following proposition gives a useful and easily applied criterion for determining when a set is dense in a metric space.

11.1.22. Proposition. *A subset D of a metric space M is dense in M if and only if every open ball contains a point of D.*

Proof. Exercise. (Solution Q.11.3.)

11.1.23. Problem. Let M be a metric space. Prove the following.

(a) If $A \subseteq M$ and $U \overset{\circ}{\subseteq} M$, then $U \cap \overline{A} \subseteq \overline{U \cap A}$.

(b) If D is dense in M and $U \overset{\circ}{\subseteq} M$, then $U \subseteq \overline{U \cap D}$.

11.1.24. Proposition. *Let $A \subseteq B \subseteq C \subseteq M$ where M is a metric space. If A is dense in B and B is dense in C, then A is dense in C.*

Proof. Problem.

11.2. THE RELATIVE TOPOLOGY

In example 11.1.2 we considered the set $A = [0, 1)$ which is contained in both the metric spaces $M = [0,\infty)$ and $N = \mathbb{R}$. We observed that the question "Is A open?" is ambiguous; it depends on whether we mean "open in M" or "open in N". Similarly, the notation $B_r(a)$ is equivocal. In M the open ball $B_{\frac{1}{2}}(0)$ is the interval $[0, \frac{1}{2})$ while in N it is the interval $(-\frac{1}{2}, \frac{1}{2})$. When working with sets which are contained in two different spaces, considerable confusion can be created by ambiguous choices of notation or terminology. In the next proposition, where we examine the relationship between open subsets of a metric space N and open subsets of a subspace M, it is necessary, in the proof, to consider open balls in both M and N. To avoid confusion we use the usual notation $B_r(a)$ for open balls in M and a different one $D_r(a)$ for those in N.

The point of the following proposition is that even if we are dealing with a complicated or badly scattered subspace of a metric space, its open sets are easily identified. When an open set in the larger space is intersected with the subspace M what results is an open set in M; and, less obviously, *every* open set in M can be produced in this fashion.

11.2.1. Proposition. *Let M be a subspace of a metric space N. A set $U \subseteq M$ is open in M if and only if there exists a set V open in N such that $U = V \cap M$.*

Proof. Let us establish some notation. If $a \in M$ and $r > 0$ we write $B_r(a)$ for the open ball about a of radius r in the space M. If $a \in N$ and $r > 0$ we write $D_r(a)$ for the corresponding open ball in the space N. Notice that $B_r(a) = D_r(a) \cap M$. Define a mapping f from the set of all open balls in M into the set of open balls in N by

$$f(B_r(a)) = D_r(a)\,.$$

Thus f is just the function which associates with each open ball in the space M the corresponding open ball (same center, same radius) in N; so $f(B) \cap M = B$ for each open ball B in M.

Now suppose U is open in M. By proposition 11.1.4 there exists a family $\mathfrak{B}$ of open balls in M such that $U = \bigcup \mathfrak{B}$. Let $\mathfrak{D} = \{f(B) \colon B \in \mathfrak{B}\}$ and $V = \bigcup \mathfrak{D}$. Then V, being a union of open balls in N, is an open subset of N and

$$\begin{aligned} V \cap M &= \left(\bigcup \mathfrak{D}\right) \cap M \\ &= \left(\bigcup \{f(B) : B \in \mathfrak{B}\}\right) \cap M \\ &= \bigcup \{f(B) \cap M : B \in \mathfrak{B}\} \\ &= \bigcup \mathfrak{B} \\ &= U\,. \end{aligned}$$

(For the third equality in the preceding string, see proposition F.2.10.)

The converse is even easier. Let V be an open subset of N and $a \in V \cap M$. In order to show that $V \cap M$ is open in M, it suffices to show that, in the space M, the point a is an interior point of the set $V \cap M$. Since V is open in N, there exists $r > 0$ such that $D_r(a) \subseteq V$. But then

$$B_r(a) \subseteq D_r(a) \cap M \subseteq V \cap M\,. \qquad \square$$

The family of all open subsets of a metric space is called the TOPOLOGY on the space. As was the case for the real numbers, the concepts of continuity, compactness, and connectedness can be characterized entirely in terms of the open subsets of the metric spaces involved and without any reference to the specific metrics which lead to these open sets. Thus we say that continuity, compactness, and connectedness are topological concepts. The next proposition 11.2.3 tells us that strongly equivalent metrics on a set produce identical topologies. Clearly, no topological property of a metric space is affected when we replace the given metric with another metric which generates the same topology.

11.2.2. Definition. Two metrics d_1 and d_2 on a set M are EQUIVALENT if they induce the same topology on M.

We now prove that strongly equivalent metrics are equivalent.

11.2.3. Proposition. *Let d_1 and d_2 be metrics on a set M and $\mathfrak{T}_1$ and $\mathfrak{T}_2$ be the topologies on (M, d_1) and (M, d_2), respectively. If d_1 and d_2 are strongly equivalent, then $\mathfrak{T}_1 = \mathfrak{T}_2$.*

PROOF. Exercise. (Solution Q.11.4.)

11.2.4. Problem. Give an example to show that equivalent metrics need not be strongly equivalent.

11.2.5. Definition. If M is a subspace of the metric space (N, d), the family of open subsets of M induced by the metric d is called the RELATIVE TOPOLOGY on M. According to proposition 11.2.1, the relative topology on M is $\{V \cap M \colon V \overset{\circ}{\subseteq} N\}$.

CHAPTER 12

SEQUENCES IN METRIC SPACES

In chapter 4 we were able to characterize several topological properties of the real line $\mathbb{R}$ by means of sequences. The same sort of thing works in general metric spaces. Early in this chapter we give sequential characterizations of open sets, closed sets, dense sets, closure, and interior. Later we discuss products of metric spaces.

12.1. CONVERGENCE OF SEQUENCES

Recall that a SEQUENCE is any function whose domain is the set $\mathbb{N}$ of natural numbers. If S is a set and $x\colon \mathbb{N} \to S$, then we say that x is a sequence of members of S. A map $x\colon \mathbb{N} \to \mathbb{R}$, for example, is a sequence of real numbers. In dealing with sequences one usually (but not always) writes x_n for $x(n)$. The element x_n in the range of a sequence is the n^{th} TERM of the sequence. Frequently we use the notations $(x_n)_{n=1}^{\infty}$ or just (x_n) to denote the sequence x.

12.1.1. Definition. A NEIGHBORHOOD of a point in a metric space is any open set containing the point. Let x be a sequence in a set S and B be a subset of S. The sequence x is EVENTUALLY in the set B if there is a natural number n_0 such that $x_n \in B$ whenever $n \geq n_0$. A sequence x in a metric space M CONVERGES to a point a in M if x is eventually in every neighborhood of a (equivalently, if it is eventually in every open ball about a). The point a is the LIMIT of the sequence x. (In proposition 12.2.4 we find that limits of sequences are unique, so references to "the" limit of a sequence are justified.)

If a sequence x converges to a point a in a metric space we write

$$x_n \to a \quad \text{as} \quad n \to \infty$$

or

$$\lim_{n\to\infty} x_n = a\,.$$

It should be clear that the preceding definition may be rephrased as follows: The sequence x converges to the point a if for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(x_n, a) < \epsilon$ whenever $n \geq n_0$. It follows immediately that $x_n \to a$ if and only if $d(x_n, a) \to 0$ as $n \to \infty$. Notice that in the metric space $\mathbb{R}$ the current definition agrees with the one given in chapter 4.

12.2. SEQUENTIAL CHARACTERIZATIONS OF TOPOLOGICAL PROPERTIES

Now we proceed to characterize some metric space concepts in terms of sequences. The point of this is that sequences are often easier to work with than arbitrary open sets.

12.2.1. Proposition. *A subset U of a metric space M is open if and only if every sequence in M which converges to an element of U is eventually in U.*

PROOF. Suppose that U is an open subset of M. Let (x_n) be a sequence in M which converges to a point a in U. Then U is a neighborhood of a. Since $x_n \to a$, the sequence (x_n) is eventually in U.

To obtain the converse suppose that U is not open. Some point, say a, of U is not an interior point of U. Then for every n in $\mathbb{N}$ we may choose an element x_n in $B_{1/n}(a)$ such that $x_n \notin U$. Then the sequence (x_n) converges to $a \in U$, but it is certainly not true that (x_n) is eventually in U. $\square$

12.2.2. Proposition. *A subset A of a metric space is closed if and only if b belongs to A whenever (a_n) is a sequence in A which converges to b.*

PROOF. Exercise. (Solution Q.12.1.)

12.2.3. Proposition. *A subset D of a metric space M is dense in M if and only if every point of M is the limit of a sequence of elements of D.*

PROOF. Problem.

12.2.4. Proposition. *In metric spaces limits of sequences are unique. (That is, if $a_n \to b$ and $a_n \to c$ in some metric space, then $b = c$.)*

PROOF. Problem.

12.2.5. Problem. Show that a point p is in the closure of a subset A of a metric space if and only if there is a sequence of points in A which converges to p. Also, give a characterization of the interior of a set by means of sequences.

12.2.6. Problem. Since the rationals are dense in $\mathbb{R}$, it must be possible, according to proposition 12.2.3, to find a sequence of rational numbers which converges to the number π. Identify one such sequence.

12.2.7. Proposition. *Let d_1 and d_2 be strongly equivalent metrics on a set M. If a sequence of points in M converges to a point b in the metric space (M, d_1), then it also converges to b in the space (M, d_2).*

PROOF. Problem.

12.2.8. Problem. Use proposition 12.2.7 to show that the metric ρ defined in example 9.2.15 is not strongly equivalent on $\mathbb{R}^2$ to the usual Euclidean (example 9.2.8). You gave a (presumably different) proof of this in problem 9.3.6.

12.2.9. Problem. Let M be a metric space with the discrete metric. Give a simple characterization of the convergent sequences in M.

12.2.10. Definition. Let A and B be nonempty subsets of a metric space M. The DISTANCE between A and B, which we denote by $d(A, B)$, is defined to be $\inf\{d(a, b) \colon a \in A \text{ and } b \in B\}$. If $a \in M$ we write $d(a, B)$ for $d(\{a\}, B)$.

12.2.11. Problem. Let B be a nonempty subset of a metric space M.

(a) Show that if $x \in B$, then $d(x, B) = 0$.
(b) Give an example to show that the converse of (a) may fail.
(c) Show that if B is closed, the converse of (a) holds.

12.3. PRODUCTS OF METRIC SPACES

12.3.1. Definition. Let (M_1, ρ_1) and (M_2, ρ_2) be metric spaces. We define three metrics, d, d_1, and d_u, on the product $M_1 \times M_2$. For $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in $M_1 \times M_2$ let

$$d(x, y) = \left((\rho_1(x_1, y_1))^2 + (\rho_2(x_2, y_2))^2\right)^{\frac{1}{2}},$$

$$d_1(x, y) = \rho_1(x_1, y_1) + \rho_2(x_2, y_2), \quad \text{and}$$

$$d_u(x, y) = \max\{\rho_1(x_1, y_1), \rho_2(x_2, y_2)\}.$$

It is not difficult to show that these really are metrics. They are just generalizations, to arbitrary products, of the metrics on $\mathbb{R} \times \mathbb{R}$ defined in 9.2.8, 9.2.10, and 9.2.12.

12.3.2. Proposition. *The three metrics on $M_1 \times M_2$ defined in 12.3.1 are strongly equivalent.*

Proof. Exercise. *Hint.* Review proposition 9.3.2. (Solution Q.12.2.)

In light of the preceding result and proposition 11.2.3 the three metrics defined on $M_1 \times M_2$ (in 12.3.1) all give rise to exactly the same topology on $M_1 \times M_2$. Since we will be concerned primarily with *topological properties* of product spaces, it makes little difference which of these metrics we officially adopt as "the" product metric. We choose d_1 because it is arithmetically simple (no square roots of sums of squares).

12.3.3. Definition. If M_1 and M_2 are metric spaces, then we say that the metric space $(M_1 \times M_2, d_1)$, where d_1 is defined in 12.3.1, is the PRODUCT (METRIC) SPACE of M_1 and M_2; and the metric d_1 is the PRODUCT METRIC. When we encounter a reference to "the metric space $M_1 \times M_2$" we assume, unless the contrary is explicitly stated, that this space is equipped with *the* product metric d_1.

A minor technical point, which is perhaps worth mentioning, is that the usual (Euclidean) metric on $\mathbb{R}^2$ is not (according to the definition just given) the product metric. Since these two metrics *are* equivalent and since most of the properties we consider are topological ones, this will cause little difficulty.

It is easy to work with sequences in product spaces. This is a consequence of the fact, which we prove next, that a necessary and sufficient condition for the convergence of a sequence in a product space is the convergence of its coordinates.

12.3.4. Proposition. *Let M_1 and M_2 be metric spaces. A sequence $\big((x_n, y_n)\big)_{n=1}^{\infty}$ in the product space converges to a point (a, b) in $M_1 \times M_2$ if and only if $x_n \to a$ and $y_n \to b$.*

Proof. For $k = 1, 2$ let ρ_k be the metric on the space M_k. The product metric d_1 on $M_1 \times M_2$ is defined in 12.3.1.

Suppose $(x_n, y_n) \to (a, b)$. Then

$$\begin{aligned}\rho_1(x_n, a) &\le \rho_1(x_n, a) + \rho_2(y_n, b)\\ &= d_1\big((x_n, y_n)\,,\,(a, b)\big) \to 0;\end{aligned}$$

so $x_n \to a$. Similarly, $y_n \to b$.

Conversely, suppose $x_n \to a$ in M_1 and $y_n \to b$ in M_2. Given $\epsilon > 0$ we may choose $n_1, n_2 \in \mathbb{N}$ such that $\rho_1(x_n, a) < \frac{1}{2}\epsilon$ when $n \ge n_1$ and $\rho_2(y_n, b) < \frac{1}{2}\epsilon$ when $n \ge n_2$. Thus if $n \ge \max\{n_1, n_2\}$,

$$d_1\big((x_n, y_n)\,,\,(a, b)\big) = \rho_1(x_n, a) + \rho_2(y_n, b) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon\,;$$

so $(x_n, y_n) \to (a, b)$ in $M_1 \times M_2$. □

Remark. By virtue of proposition 12.2.7, the truth of the preceding proposition would not have been affected had either d or d_u (as defined in 12.3.1) been chosen as the product metric for $M_1 \times M_2$.

12.3.5. Problem. Generalize definitions 12.3.1 and 12.3.3 to $\mathbb{R}^n$ where $n \in \mathbb{N}$. That is, write appropriate formulas for $d(x, y)$, $d_1(x, y)$, and $d_u(x, y)$ for $x, y \in \mathbb{R}^n$, and explain what we mean by the product metric on an arbitrary finite product $M_1 \times M_2 \times \cdots \times M_n$ of metric spaces.

Also state and prove generalizations of propositions 12.3.2 and 12.3.4 to arbitrary finite products.

CHAPTER 13

UNIFORM CONVERGENCE

13.1. THE UNIFORM METRIC ON THE SPACE OF BOUNDED FUNCTIONS

13.1.1. Definition. Let S be a nonempty set. A function $f\colon S \to \mathbb{R}$ is BOUNDED if there exists a number $M \geq 0$ such that

$$|f(x)| \leq M \qquad \text{for all } x \in S\,.$$

We denote by $\mathcal{B}(S,\mathbb{R})$ (or just by $\mathcal{B}(S)$) the set of all bounded real valued functions on S.

13.1.2. Proposition. *If f and g are bounded real valued functions on a nonempty set S and α is a real number, then the functions $f+g$, αf, and fg are all bounded.*

PROOF. There exist numbers M, $N \geq 0$ such that $|f(x)| \leq M$ and $|g(x)| \leq N$ for all $x \in S$. Then, for all $x \in S$

$$\begin{aligned}
|(f+g)(x)| &\leq |f(x)| + |g(x)| \leq M + N,\\
|(\alpha f)(x)| &= |\alpha|\,|f(x)| \leq |\alpha|\,M, \text{ and}\\
|(fg)(x)| &= |f(x)|\,|g(x)| \leq MN.
\end{aligned}$$

□

13.1.3. Definition. Let S be a nonempty set. We define a metric d_u on the set $\mathcal{B}(S,\mathbb{R})$ by

$$d_u(f,g) \equiv \sup\{|f(x)-g(x)|\colon x \in S\}$$

whenever f, $g \in \mathcal{B}(S,\mathbb{R})$. The metric d_u is the UNIFORM METRIC on $\mathcal{B}(S,\mathbb{R})$.

13.1.4. Example. Let $S = [-1,1]$ and for all $x \in S$ let $f(x) = |x|$ and $g(x) = \frac{1}{2}(x-1)$. Then $d_u(f,g) = 2$.

PROOF. It is clear from the graphs of f and g that the functions are farthest apart at $x = -1$. Thus

$$\begin{aligned}
d_u(f,g) &= \sup\{|f(x)-g(x)|\colon -1 \leq x \leq 1\}\\
&= |f(-1) - g(-1)| = 2.
\end{aligned}$$

□

13.1.5. Example. Let $f(x) = x^2$ and $g(x) = x^3$ for $0 \leq x \leq 1$. Then $d_u(f,g) = 4/27$.

PROOF. Let $h(x) = |f(x)-g(x)| = f(x) - g(x)$ for $0 \leq x \leq 1$. To maximize h on $[0,1]$ use elementary calculus to find critical points. Since $h'(x) = 2x - 3x^2 = 0$ only if $x = 0$ or $x = \frac{2}{3}$, it is clear that the maximum value of h occurs at $x = \frac{2}{3}$. Thus

$$d_u(f,g) = \sup\{h(x)\colon 0 \leq x \leq 1\} = h\left(\frac{2}{3}\right) = \frac{4}{27}\,.$$

□

13.1.6. Exercise. Suppose f is the constant function defined on $[0,1]$ whose value is 1. Asked to describe those functions in $\mathcal{B}([0,1])$ which lie in the open ball about f of radius 1, a student replies (somewhat incautiously) that $B_1(f)$ is the set of all real-valued functions g on $[0,1]$ satisfying $0 < g(x) < 2$ for all $x \in [0,1]$. Why is this response wrong? (Solution Q.13.1.)

13.1.7. Problem. Let $f(x) = \sin x$ and $g(x) = \cos x$ for $0 \leq x \leq \pi$. Find $d_u(f,g)$ in the set of functions $\mathcal{B}([0,\pi])$.

13.1.8. Problem. Let $f(x) = 3x - 3x^3$ and $g(x) = 3x - 3x^2$ for $0 \le x \le 2$. Find $d_u(f,g)$ in the set of functions $\mathcal{B}([0,2])$.

13.1.9. Problem. Explain why it is reasonable to use the same notation d_u (and the same name) for both the metric in example 9.2.12 and the one defined in 13.1.3.

The terminology in 13.1.3 is somewhat optimistic. We have not yet verified that the "uniform metric" is indeed a metric on $\mathcal{B}(S,\mathbb{R})$. We now remedy this.

13.1.10. Proposition. *Let S be a nonempty set. The function d_u defined in 13.1.3 is a metric on the set of functions $\mathcal{B}(S,\mathbb{R})$.*

Proof. Exercise. (Solution Q.13.2.)

13.1.11. Problem. Let $f(x) = x$ and $g(x) = 0$ for all $x \in [0,1]$. Find a function h in $\mathcal{B}([0,1])$ such that

$$d_u(f,h) = d_u(f,g) = d_u(g,h)\,.$$

13.1.12. Definition. Let (f_n) be a sequence of real valued functions on a nonempty set S. If there is a function g in $\mathcal{F}(S,\mathbb{R})$ such that

$$\sup\{|f_n(x) - g(x)| \colon x \in S\} \to 0 \qquad \text{as } n \to \infty$$

we say that the sequence (f_n) CONVERGES UNIFORMLY to g and we write

$$f_n \to g\,(\text{unif})\,.$$

The function g is the UNIFORM LIMIT of the sequence (f_n). Notice that if g and all the f_n's belong to $\mathcal{B}(S,\mathbb{R})$, then uniform convergence of (f_n) to g is the same thing as convergence of f_n to g in the uniform metric.

13.1.13. Example. For each $n \in \mathbb{N}$ and $x \in \mathbb{R}$ let

$$f_n(x) = \frac{1}{n}\sin(nx)\,.$$

Then $f_n \to \mathbf{0}\,(\text{unif})$. (Here $\mathbf{0}$ is the constant function zero.)

Proof.

$$\begin{aligned} d_u(f_n, 0) &= \sup\left\{\tfrac{1}{n}|\sin nx| \colon x \in \mathbb{R}\right\} \\ &= \frac{1}{n} \to 0 \qquad \text{as } n \to \infty\,. \end{aligned}$$

□

13.1.14. Example. Let $g(x) = x$ and $f_n(x) = x + \frac{1}{n}$ for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Then $f_n \to g\,(\text{unif})$ since

$$\sup\{|f_n(x) - g(x)| \colon x \in \mathbb{R}\} = \frac{1}{n} \to 0 \qquad \text{as } n \to \infty.$$

It is not correct (although perhaps tempting) to write $d_u(f_n, g) \to 0$. This expression is meaningless since the functions f_n and g do not belong to the metric space $\mathcal{B}(\mathbb{R})$ on which d_u is defined.

13.2. POINTWISE CONVERGENCE

Sequences of functions may converge in many different and interesting ways. Another mode of convergence that is frequently encountered is "pointwise convergence".

13.2.1. Definition. Let (f_n) be a sequence of real valued functions on a nonempty set S. If there is a function g such that

$$f_n(x) \to g(x) \qquad \text{for all } x \in S$$

then (f_n) CONVERGES POINTWISE to g. In this case we write

$$f_n \to g\,(\text{ptws})\,.$$

The function g is the POINTWISE LIMIT of the sequence f_n.

In the following proposition we make the important, if elementary, observation that uniform convergence is stronger than pointwise convergence.

13.2.2. Proposition. *Uniform convergence implies pointwise convergence, but not conversely.*

Proof. Exercise. (Solution Q.13.3.)

13.2.3. Problem. Find an example of a sequence (f_n) of functions in $\mathcal{B}([0,1])$ which converges pointwise to the zero function $\mathbf{0}$ but satisfies

$$d_u(f_n, \mathbf{0}) \to \infty \qquad \text{as } n \to \infty\,.$$

Next we show that the uniform limit of a sequence of bounded functions is itself bounded.

13.2.4. Proposition. *Let (f_n) be a sequence in $\mathcal{B}(S)$ and g be a real valued function on S.*

(a) *If $f_n \to g$ (unif), then $g \in \mathcal{B}(S)$.*
(b) *The assertion in (a) does not hold if uniform convergence is replaced by pointwise convergence.*

Proof. Exercise. (Solution Q.13.4.)

13.2.5. Exercise. Let $f_n(x) = x^n - x^{2n}$ for $0 \le x \le 1$ and $n \in \mathbb{N}$. Does the sequence (f_n) converge pointwise on $[0,1]$? Is the convergence uniform? (Solution Q.13.5.)

13.2.6. Problem. Given in each of the following is the n^{th} term of a sequence of real valued functions defined on $[0,1]$. Which of these converge pointwise on $[0,1]$? For which is the convergence uniform?

(a) $x \mapsto x^n$.
(b) $x \mapsto nx$.
(c) $x \mapsto xe^{-nx}$.

13.2.7. Problem. Given in each of the following is the n^{th} term of a sequence of real valued functions defined on $[0,1]$. Which of these converge pointwise on $[0,1]$? For which is the convergence uniform?

(a) $x \mapsto \dfrac{1}{nx+1}$.
(b) $x \mapsto \dfrac{x}{nx+1}$.
(c) $x \mapsto \dfrac{x^2}{n} - \dfrac{x}{n^2}$.

13.2.8. Problem. Let $f_n(x) = \dfrac{(n-1)x + x^2}{n+x}$ for all $x \ge 1$ and $n \in \mathbb{N}$. Does the sequence (f_n) have a pointwise limit on $[1,\infty)$? A uniform limit?

CHAPTER 14

MORE ON CONTINUITY AND LIMITS

14.1. CONTINUOUS FUNCTIONS

As is the case with real valued functions of a real variable, a function $f\colon M_1 \to M_2$ between two metric spaces is continuous at a point a in M_1 if $f(x)$ can be made arbitrarily close to $f(a)$ by insisting that x be sufficiently close to a.

14.1.1. Definition. Let (M_1, d_1) and (M_2, d_2) be metric spaces. A function $f\colon M_1 \to M_2$ is CONTINUOUS AT a point a in M_1 if every neighborhood of $f(a)$ contains the image under f of a neighborhood of a. Since every neighborhood of a point contains an open ball about the point and since every open ball about a point *is* a neighborhood of that point, we may restate the definition as follows. The function f is CONTINUOUS AT a if every open ball about $f(a)$ contains the image under f of an open ball about a; that is, if the following condition is satisfied: for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$f^{\rightarrow}(B_\delta(a)) \subseteq B_\epsilon(f(a)). \tag{14.1}$$

There are many equivalent ways of expressing (14.1). Here are three:

$$B_\delta(a) \subseteq f^{\leftarrow}(B_\epsilon(f(a)))$$
$$x \in B_\delta(a) \text{ implies } f(x) \in B_\epsilon(f(a))$$
$$d_1(x,a) < \delta \text{ implies } d_2(f(x), f(a)) < \epsilon$$

Notice that if f is a real valued function of a real variable, then the definition above agrees with the one given at the beginning of chapter 3.

14.1.2. Definition. A function $f\colon M_1 \to M_2$ between two metric spaces is CONTINUOUS if it is continuous at each point of M_1.

In proving propositions concerning continuity, one should not slavishly insist on specifying the radii of open balls when these particular numbers are of no interest. As an illustration, the next proposition, concerning the composite of continuous functions, is given two proofs—one with the radii of open balls specified, and a smoother one in which they are suppressed.

14.1.3. Proposition. *Let $f\colon M_1 \to M_2$ and $g\colon M_2 \to M_3$ be functions between metric spaces. If f is continuous at a in M_1 and g is continuous at $f(a)$ in M_2, then the composite function $g \circ f$ is continuous at a.*

PROOF 1. Let $\eta > 0$. We wish to show that there exists $\delta > 0$ such that

$$B_\delta(a) \subseteq (g \circ f)^{\leftarrow}(B_\eta(g(f(a)))).$$

Since g is continuous at $f(a)$ there exists $\epsilon > 0$ such that

$$B_\epsilon(f(a)) \subseteq g^{\leftarrow}(B_\eta(g(f(a)))).$$

Since f is continuous at a there exists $\delta > 0$ such that

$$B_\delta(a) \subseteq f^{\leftarrow}(B_\epsilon(f(a))).$$

Thus we have

$$\begin{aligned}(g \circ f)^{\leftarrow}(B_\eta(g(f(a)))) &= f^{\leftarrow}(g^{\leftarrow}(B_\eta(g(f(a)))))\\ &\supseteq f^{\leftarrow}(B_\epsilon(f(a)))\\ &\supseteq B_\delta(a).\end{aligned}$$

□

PROOF 2. Let B_3 be an arbitrary open ball about $g(f(a))$. We wish to show that the inverse image of B_3 under $g \circ f$ contains an open ball about a. Since g is continuous at $f(a)$, the set $g^{\leftarrow}(B_3)$ contains an open ball B_2 about $f(a)$. And since f is continuous at a, the set $f^{\leftarrow}(B_2)$ contains an open ball B_1 about a. Thus we have

$$\begin{aligned}(g \circ f)^{\leftarrow}(B_3) &= f^{\leftarrow}(g^{\leftarrow}(B_3))\\ &\supseteq f^{\leftarrow}(B_2)\\ &\supseteq B_1.\end{aligned}$$

□

The two preceding proofs are essentially the same. The only difference is that the first proof suffers from a severe case of clutter. It certainly is not more rigorous; it is just harder to read. It is good practice to relieve proofs (and their readers) of extraneous detail. The following corollary is an obvious consequence of the proposition we have just proved.

14.1.4. Corollary. *The composite of two continuous functions is continuous.*

Next we prove a result emphasizing that continuity is a topological notion; that is, it can be expressed in terms of open sets. A necessary and sufficient condition for a function to be continuous is that the inverse image of every open set be open.

14.1.5. Proposition. *A function $f : M_1 \to M_2$ between metric spaces is continuous if and only if $f^{\leftarrow}(U)$ is an open subset of M_1 whenever U is open in M_2.*

PROOF. Exercise. (Solution Q.14.1.)

14.1.6. Example. As an application of the preceding proposition we show that the function

$$f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto 2x - 5y$$

is continuous. One approach to this problem is to find, given a point (a, b) in $\mathbb{R}^2$ and $\epsilon > 0$, a number $\delta > 0$ sufficiently small that $\sqrt{(x-a)^2 + (y-b)^2} < \delta$ implies $|(2x - 5y) - (2a - 5b)| < \epsilon$. This is not excessively difficult, but it is made somewhat awkward by the appearance of squares and a square root in the definition of the usual metric on $\mathbb{R}^2$. A simpler approach is possible. We wish to prove continuity of f with respect to the usual metric d on $\mathbb{R}^2$ (defined in 9.2.8).

We know that the metric d_1 (defined in 9.2.10) on $\mathbb{R}^2$ is (strongly) equivalent to d (9.3.2) and that equivalent metrics produce identical topologies. Thus $(\mathbb{R}^2, d_1)$ and $(\mathbb{R}^2, d)$ have the same topologies. Since the continuity of a function is a topological concept (this was the point of 14.1.5), we know that f will be continuous with respect to d if and only if it is continuous with respect to d_1. Since the metric d_1 is algebraically simpler, we prove continuity with respect to d_1. To this end, let $(a, b) \in \mathbb{R}^2$ and $\epsilon > 0$. Choose $\delta = \epsilon/5$. If $d_1((x, y), (a, b)) = |x - a| + |y - b| < \delta$, then

$$\begin{aligned}|f(x, y) - f(a, b)| &= |2(x - a) - 5(y - b)|\\ &\le 5(|x - a| + |y - b|)\\ &< 5\delta\\ &= \epsilon.\end{aligned}$$

14.1.7. Example. The function

$$f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto 5x - 7y$$

is continuous.

PROOF. Problem.

The principle used in example 14.1.6 works generally: replacing a metric on the domain of a function by an equivalent metric does not affect the continuity of the function. The same assertion is true for the codomain of a function as well. We state this formally.

14.1.8. Proposition. *Let $f\colon M_1 \to M_2$ be a continuous function between two metric spaces (M_1, d_1) and (M_2, d_2). If ρ_1 is a metric on M_1 equivalent to d_1 and ρ_2 is equivalent to d_2, then f considered as a function from the space (M_1, ρ_1) to the space (M_2, ρ_2) is still continuous.*

Proof. This is an immediate consequence of propositions 11.2.3 and 14.1.5. □

14.1.9. Example. Multiplication is a continuous function on $\mathbb{R}$. That is, if we define

$$M\colon \mathbb{R}^2 \to \mathbb{R}\colon (x, y) \mapsto xy\,,$$

then the function M is continuous.

Proof. Exercise. *Hint.* Use 14.1.8. (Solution Q.14.2.)

14.1.10. Example. Addition is a continuous function on $\mathbb{R}$. That is, if we define

$$A\colon \mathbb{R}^2 \to \mathbb{R}\colon (x, y) \mapsto x + y\,,$$

then the function A is continuous.

Proof. Problem.

14.1.11. Problem. Let d be the usual metric on $\mathbb{R}^2$, let ρ be the metric on $\mathbb{R}^2$ defined in example 9.2.15, and let $f\colon (\mathbb{R}^2, d) \to (\mathbb{R}^2, \rho)$ be the identity function on $\mathbb{R}^2$. Show that f is *not* continuous.

14.1.12. Problem. Let $\mathbb{R}_d$ be the set of real numbers with the discrete metric, and let M be an arbitrary metric space. Describe the family of all continuous functions $f\colon \mathbb{R}_d \to M$.

14.1.13. Proposition. *Let $f\colon M_1 \to M_2$ where M_1 and M_2 are metric spaces. Then f is continuous if and only if $f^{\leftarrow}(C)$ is a closed subset of M_1 whenever C is a closed subset of M_2.*

Proof. Problem.

14.1.14. Proposition. *Let $f\colon M_1 \to M_2$ be a function between metric spaces. Then f is continuous if and only if*

$$f^{\leftarrow}(B^\circ) \subseteq \big(f^{\leftarrow}(B)\big)^\circ$$

for all $B \subseteq M_2$.

Proof. Problem.

14.1.15. Proposition. *Let $f\colon M_1 \to M_2$ be a function between metric spaces. Then f is continuous if and only if*

$$f^{\rightarrow}\big(\,\overline{A}\,\big) \subseteq \overline{f^{\rightarrow}(A)}$$

for every $A \subseteq M_1$.

Proof. Problem.

14.1.16. Proposition. *Let $f\colon M_1 \to M_2$ be a function between metric spaces. Then f is continuous if and only if*

$$\overline{f^{\leftarrow}(B)} \subseteq f^{\leftarrow}\big(\,\overline{B}\,\big)$$

for every $B \subseteq M_2$.

Proof. Problem.

14.1.17. Proposition. *Let f be a real valued function on a metric space M and $a \in M$. If f is continuous at a and $f(a) > 0$, then there exists a neighborhood B of a such that $f(x) > \frac{1}{2}f(a)$ for all $x \in B$.*

Proof. Problem.

14.1.18. Proposition. *Let N be a metric space and M be a subspace of N.*

(a) *The inclusion map $\iota\colon M \to N\colon x \mapsto x$ is continuous.*
(b) *Restrictions of continuous functions are continuous. That is, if $f\colon M_1 \to M_2$ is a continuous mapping between metric spaces and $A \subseteq M_1$, then $f|_A$ is continuous.*

Proof. Problem.

14.1.19. Problem. Show that alteration of the codomain of a continuous function does not affect its continuity. Precisely: If $f\colon M_0 \to M$ and $g\colon M_0 \to N$ are functions between metric spaces such that $f(x) = g(x)$ for all $x \in M_0$ and if their common image is a metric subspace of both M and N, then f is continuous if and only if g is.

In the next proposition we show that if two continuous functions agree on a dense subset of a metric space, they agree everywhere on that space.

14.1.20. Proposition. *If f, $g\colon M \to N$ are continuous functions between metric spaces, if D is a dense subset of M, and if $f|_D = g|_D$, then $f = g$.*

Proof. Problem. *Hint.* Suppose that there is a point a where f and g differ. Consider the inverse images under f and g, respectively, of disjoint neighborhoods of $f(a)$ and $g(a)$. Use proposition 11.1.22.

14.1.21. Problem. Suppose that M is a metric space and that $f\colon M \to M$ is continuous but is *not* the identity map. Show that there exists a proper closed set $C \subsetneq M$ such that

$$C \cup f^{\leftarrow}(C) = M\,.$$

Hint. Choose x so that $x \neq f(x)$. Look at the complement of $U \cap f^{\leftarrow}(V)$ where U and V are disjoint neighborhoods of x and $f(x)$, respectively.

There are two ways in which metric spaces may be regarded as "essentially the same": They may be isometric (having essentially the same distance function); or they may be topologically equivalent (having essentially the same open sets).

14.1.22. Definition. Let(M, d) and (N, ρ) be metric spaces. A bijection $f\colon M \to N$ is an ISOMETRY if

$$\rho(f(x), f(y)) = d(x, y)$$

for all x, $y \in M$. If an isometry exists between two metric spaces, the spaces are said to be ISOMETRIC.

14.1.23. Definition. A bijection $g\colon M \to N$ between metric spaces is a HOMEOMORPHISM if both g and g^{-1} are continuous. Notice that if g is a homeomorphism, then $g^{\rightarrow}$ establishes a one-to-one correspondence between the family of open subsets of M and the family of open subsets of N. For this reason two metric spaces are said to be (TOPOLOGICALLY) EQUIVALENT or HOMEOMORPHIC if there exists a homeomorphism between them. Since the open sets of a space are determined by its metric, it is clear that every isometry is automatically a homeomorphism. The converse, however, is not correct (see example 14.1.25 below).

14.1.24. Problem. Give an example of a bijection between metric spaces which is continuous but is not a homeomorphism.

14.1.25. Example. The open interval $(0, 1)$ and the real line $\mathbb{R}$ (with their usual metrics) are homeomorphic but not isometric.

Proof. Problem.

We have seen (in chapter 12) that certain properties of sets in metric spaces can be characterized by means of sequences. Continuity of functions between metric spaces also has a simple and useful sequential characterization.

14.1.26. Proposition. *Let $f\colon M_1 \to M_2$ be a function between metric spaces and a be a point in M_1. The function f is continuous at a if and only if $f(x_n) \to f(a)$ whenever $x_n \to a$.*

Proof. Exercise. (Solution Q.14.3.)

14.1.27. Problem. Give a second solution to proposition 14.1.20, this time making use of propositions 12.2.3 and 14.1.26.

14.1.28. Problem. Use examples 14.1.10 and 14.1.9 to show that if $x_n \to a$ and $y_n \to b$ in $\mathbb{R}$, then $x_n + y_n \to a + b$ and $x_n y_n \to ab$. (Do not give an "ϵ-δ proof".)

14.1.29. Problem. Let c be a point in a metric space M. Show that the function

$$f\colon M \to \mathbb{R}\colon x \mapsto d(x, c)$$

is continuous. *Hint.* Use problem 9.2.17.

14.1.30. Problem. Let C be a nonempty subset of a metric space. Then the function

$$g\colon M \to \mathbb{R}\colon x \mapsto d(x, C)$$

is continuous. (See 12.2.10 for the definition of $d(x, C)$.)

14.1.31. Proposition (Urysohn's lemma)**.** *Let A and B be nonempty disjoint closed subsets of a metric space M. Then there exists a continuous function $f\colon M \to \mathbb{R}$ such that* $\operatorname{ran} f \subseteq [0, 1]$, *$f^{\to}(A) = \{0\}$, and $f^{\to}(B) = \{1\}$.*

Proof. Problem. *Hint.* Consider $\dfrac{d(x, A)}{d(x, A) + d(x, B)}$. Use problems 12.2.11(c) and 14.1.30.

14.1.32. Problem. (Definition. Disjoint sets A and B in a metric space M are said to be separated by open sets if there exist $U, V \overset{\circ}{\subseteq} M$ such that $U \cap V = \emptyset$, $A \subseteq U$, and $B \subseteq V$.) Show that in a metric space every pair of disjoint closed sets can be separated by open sets.

14.1.33. Problem. If f and g are continuous real valued functions on a metric space M, then $\{x \in M : f(x) \neq g(x)\}$ is an open subset of M.

14.1.34. Problem. Show that if f is a continuous real valued function on a metric space, then $|f|$ is continuous. (We denote by $|f|$ the function $x \mapsto |f(x)|$.)

14.1.35. Problem. Show that metrics are continuous functions. That is, show that if M is a set and $d\colon M \times M \to \mathbb{R}$ is a metric, then d is continuous. Conclude from this that if $x_n \to a$ and $y_n \to b$ in a metric space, then $d(x_n, y_n) \to d(a, b)$.

14.1.36. Problem. Show that if f and g are continuous real valued functions on a metric space M and $f(a) = g(a)$ at some point $a \in M$, then for every $\epsilon > 0$ there exists a neighborhood U of a such that $f(x) < g(x) + \epsilon$ for all $x \in U$.

14.2. MAPS INTO AND FROM PRODUCTS

Let (M_1, ρ_1) and (M_2, ρ_2) be metric spaces. Define the coordinate projections π_1 and π_2 on the product $M_1 \times M_2$ by

$$\pi_k\colon M_1 \times M_2 \to M_k\colon (x_1, x_2) \mapsto x_k \qquad \text{for } k = 1, 2.$$

If $M_1 \times M_2$ has the product metric d_1 (see 12.3.1 and 12.3.3), then the coordinate projections turn out to be continuous functions.

14.2.1. Exercise. Prove the assertion made in the preceding sentence. (Solution Q.14.4.)

14.2.2. Notation. Let S_1, S_2, and T be sets. If $f\colon T \to S_1 \times S_2$, then we define functions $f^1 := \pi_1 \circ f$ and $f^2 := \pi_2 \circ f$. These are the COMPONENTS of f.

If, on the other hand, functions $g\colon T \to S_1$ and $h\colon T \to S_2$ are given, we define the function (g,h) by

$$(g,h)\colon T \to S_1 \times S_2 \colon x \mapsto (g(x), h(x)) .$$

Thus it is clear that whenever $f\colon T \to S_1 \times S_2$, we have

$$f = (f^1, f^2) .$$

14.2.3. Proposition. *Let M_1, M_2, and N be metric spaces and $f\colon N \to M_1 \times M_2$. The function f is continuous if and only if its components f^1 and f^2 are.*

PROOF. Exercise. (Solution Q.14.5.)

14.2.4. Proposition. *Let f and g be continuous real valued functions on a metric space.*

(a) *The product fg is a continuous function.*
(b) *For every real number α the function $\alpha g\colon x \mapsto \alpha g(x)$ is continuous.*

PROOF. Exercise. (Solution Q.14.6.)

14.2.5. Problem. Let f and g be continuous real valued functions on a metric space and suppose that g is never zero. Show that the function f/g is continuous.

14.2.6. Proposition. *If f and g are continuous real valued functions on a metric space, then $f+g$ is continuous.*

PROOF. Problem.

14.2.7. Problem. Show that every polynomial function on $\mathbb{R}$ is continuous. *Hint.* An induction on the degree of the polynomial works nicely.

14.2.8. Definition. Let S be a set and f, $g\colon S \to \mathbb{R}$. Then $f \vee g$, the SUPREMUM (or MAXIMUM) of f and g, is defined by

$$(f \vee g)(x) := \max\{f(x), g(x)\}$$

for every $x \in S$. Similarly, $f \wedge g$, the INFIMUM (or MINIMUM) of f and g, is defined by

$$(f \wedge g)(x) := \min\{f(x), g(x)\}$$

for every $x \in S$.

14.2.9. Problem. Let $f(x) = \sin x$ and $g(x) = \cos x$ for $0 \le x \le 2\pi$. Make a careful sketch of $f \vee g$ and $f \wedge g$.

14.2.10. Problem. Show that if f and g are continuous real valued functions on a metric space, then $f \vee g$ and $f \wedge g$ are continuous. *Hint.* Consider things like $f + g + |f - g|$.

In proposition 14.2.3 we have dealt with the continuity of functions which map *into* products of metric spaces. We now turn to functions which map *from* products; that is, to functions of several variables.

14.2.11. Notation. Let S_1, S_2, and T be sets and $f\colon S_1 \times S_2 \to T$. For each $a \in S_1$ we define the function

$$f(a, \cdot\,)\colon S_2 \to T \colon y \mapsto f(a,y)$$

and for each $b \in S_2$ we define the function

$$f(\,\cdot\,, b)\colon S_1 \to T \colon x \mapsto f(x,b) .$$

Loosely speaking, $f(a, \cdot\,)$ is the result of regarding f as a function of only its second variable; $f(\,\cdot\,, b)$ results from thinking of f as depending on only its first variable.

14.2.12. Proposition. *Let M_1, M_2, and N be metric spaces and $f\colon M_1 \times M_2 \to N$. If f is continuous, then so are $f(a, \cdot\,)$ and $f(\,\cdot\,, b)$ for all $a \in M_1$ and $b \in M_2$.*

Proof. Exercise. (Solution Q.14.7.)

This proposition is sometimes paraphrased as follows: Joint continuity implies separate continuity. The converse is *not* true. (See problem 14.2.13.)

Remark. It should be clear how to extend the results of propositions 14.2.3 and 14.2.12 to products of any finite number of metric spaces.

14.2.13. Problem. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} xy(x^2 + y^2)^{-1}, & \text{for } (x, y) \neq (0, 0) \\ 0, & \text{for } x = y = 0. \end{cases}$$

(a) Show that f is continuous at each point of $\mathbb{R}^2$ except at $(0, 0)$, where it is *not* continuous.
(b) Show that the converse of proposition 14.2.12 is not true.

14.2.14. Notation. Let M and N be metric spaces. We denote by $\mathcal{C}(M, N)$ the family of all continuous functions f taking M into N.

In proposition 13.2.4 we showed that the uniform limit of a sequence of bounded real valued functions is bounded. We now prove an analogous result for continuous real valued functions: The uniform limit of a sequence of continuous real valued functions is continuous.

14.2.15. Proposition. *If (f_n) is a sequence of continuous real valued functions on a metric space M which converges uniformly to a real valued function g on M, then g is continuous.*

Proof. Exercise. (Solution Q.14.8.)

14.2.16. Problem. If the word "pointwise" is substituted for "uniformly" in proposition 14.2.15, the conclusion no longer follows. In particular, find an example of a sequence (f_n) of continuous functions on $[0, 1]$ which converges pointwise to a function g on $[0, 1]$ which is *not* continuous.

14.3. LIMITS

We now generalize to metric spaces the results of chapter 7 on limits of real valued functions. Most of this generalization is accomplished quite simply: just replace open intervals on the real line with open balls in metric spaces.

14.3.1. Definition. If $B_r(a)$ is the open ball of radius r about a point a in a metric space M, then $B_r^*(a)$, the DELETED OPEN BALL of radius r about a, is just $B_r(a)$ with the point a deleted. That is, $B_r^*(a) = \{x \in M : 0 < d(x, a) < r\}$.

14.3.2. Definition. Let (M, d) and (N, ρ) be metric spaces, $A \subseteq M$, $f\colon A \to N$, a be an accumulation point of A, and $l \in N$. We say that l is the LIMIT OF f AS x APPROACHES a (of the LIMIT OF f AT a) if: for every $\epsilon > 0$ there exists $\delta > 0$ such that $f(x) \in B_\epsilon(l)$ whenever $x \in A \cap B_\delta^*(a)$. In slightly different notation:

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in A)\ 0 < d(x, a) < \delta \implies \rho(f(x), l) < \epsilon.$$

When this condition is satisfied we write

$$f(x) \to l \quad \text{as } x \to a$$

or

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

As in chapter 7 this notation is optimistic. We will show in the next proposition that limits, if they exist, are unique.

14.3.3. Proposition. *Let $f\colon A \to N$ where M and N are metric spaces and $A \subseteq M$, and let $a \in A'$. If $f(x) \to b$ as $x \to a$ and $f(x) \to c$ as $x \to a$, then $b = c$.*

Proof. Exercise. (Solution Q.14.9.)

For a function between metric spaces the relationship between its continuity at a point and its limit there is exactly the same as in the case of real valued functions. (See proposition 7.2.3 and the two examples which precede it.)

14.3.4. Proposition. *Let M and N be metric spaces, let $f\colon A \to N$ where $A \subseteq M$, and let $a \in A \cap A'$. Then f is continuous at a if and only if*

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

Proof. Problem. *Hint.* Modify the proof of 7.2.3.

14.3.5. Proposition. *If M is a metric space, $f\colon A \to M$ where $A \subseteq \mathbb{R}$, and $a \in A'$, then*

$$\lim_{h\to 0} f(a+h) = \lim_{x\to a} f(x)$$

in the sense that if either limit exists, then so does the other and the two limits are equal.

Proof. Problem. *Hint.* Modify the proof of 7.2.4.

We conclude this chapter by examining the relationship between "double" and "iterated" limits of real valued functions of two real variables. A limit of the form

$$\lim_{(x,y)\to(a,b)} f(x,y)$$

is a DOUBLE LIMIT; limits of the form

$$\lim_{x\to a}\bigl(\lim_{y\to b} f(x,y)\bigr) \quad \text{and} \quad \lim_{y\to b}\bigl(\lim_{x\to a} f(x,y)\bigr)$$

are ITERATED LIMITS. The meaning of the expression $\lim_{x\to a}\bigl(\lim_{y\to b} f(x,y)\bigr)$ should be clear: it is $\lim_{x\to a} h(x)$ where h is the function defined by $h(x) = \lim_{y\to b} f(x,y)$.

14.3.6. Example. Let $f(x,y) = x\sin(1+x^2y^2)$ for all $x, y \in \mathbb{R}$. Then $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Proof. The function f maps $\mathbb{R}^2$ into $\mathbb{R}$. We take the usual Euclidean metric on both of these spaces. Given $\epsilon > 0$, choose $\delta = \epsilon$. If $(x,y) \in B^*_\delta(0,0)$, then

$$|f(x,y) - 0| = |x||\sin(1+x^2y^2)| \le |x| \le \sqrt{x^2+y^2} = d\bigl((x,y),(0,0)\bigr) < \delta = \epsilon\,. \qquad \square$$

14.3.7. Example. Let $f(x,y) = x\sin(1+x^2y^2)$ for all x, $y \in \mathbb{R}$. Then $\lim_{x\to 0}\bigl(\lim_{y\to 0} f(x,y)\bigr) = 0$.

Proof. Compute the inner limit first: $\lim_{x\to 0}\bigl(\lim_{y\to 0}\bigl(x\sin(1+x^2y^2)\bigr) = \lim_{x\to 0}(x\sin 1) = 0$. $\square$

Because of the intimate relationship between continuity and limits (proposition 14.3.4) and because of the fact that joint continuity implies separate continuity (proposition 14.2.12), many persons wrongly conclude that the existence of a double limit implies the existence of the corresponding iterated limits. One of the last problems in this chapter will provide you with an example of a function having a double limit at the origin but failing to have one of the corresponding iterated limits. In the next proposition we prove that if in addition to the existence of the double limit we assume that $\lim_{x\to a} f(x,y)$ and $\lim_{y\to b} f(x,y)$ always exist, then both iterated limits exist and are equal.

14.3.8. Proposition. *Let f be a real valued function of two real variables. If the limit*

$$l = \lim_{(x,y)\to(a,b)} f(x,y)$$

exists and if $\lim_{x\to a} f(x,y)$ *and* $\lim_{y\to b} f(x,y)$ *exist for all* y *and* x, *respectively, then the iterated limits*

$$\lim_{x\to a}\bigl(\lim_{y\to b} f(x,y)\bigr) \qquad \textit{and} \qquad \lim_{y\to b}\bigl(\lim_{x\to a} f(x,y)\bigr)$$

both exist and are equal to l.

Proof. Exercise. *Hint.* Let $g(x) = \lim_{y\to b} f(x,y)$ for all $x \in \mathbb{R}$. We wish to show that $\lim_{x\to a} g(x) = l$. The quantity $|g(x) - l|$ is small whenever both $|g(x) - f(x,y)|$ and $|f(x,y) - l|$ are. Since $\lim_{(x,y)\to(a,b)} f(x,y) = l$ it is easy to make $|f(x,y) - l|$ small: insist that (x,y) lie in some sufficiently small open ball about (a,b) of radius, say, δ. This can be accomplished by requiring, for example, that

$$|x - a| < \delta/2 \tag{14.2}$$

and that

$$|y - b| < \delta/2. \tag{14.3}$$

Since $g(x) = \lim_{y\to b} f(x,y)$ for every x, we can make $|g(x) - f(x,y)|$ small (for fixed x) by supposing that

$$|y - b| < \eta \tag{14.4}$$

for some sufficiently small η. So the proof is straightforward: require x to satisfy (14.2) and for such x require y to satisfy (14.3) and (14.4). (Solution Q.14.10.)

It is sometimes necessary to show that certain limits do *not* exist. There is a rather simple technique which is frequently useful for showing that the limit of a given real valued function does not exist at a point a. Suppose we can find two numbers $\alpha \neq \beta$ such that in *every* neighborhood of a the function f assumes (at points other than a) both the values α and β. (That is, for every $\delta > 0$ there exist points u and v in $B_\delta(a)$ distinct from a such that $f(u) = \alpha$ and $f(v) = \beta$.) Then it is easy to see that f cannot have a limit as x approaches a. Argue by contradiction: suppose $\lim_{x\to a} f(x) = l$. Let $\epsilon = |\alpha - \beta|$. Then $\epsilon > 0$; so there exists $\delta > 0$ such that $|f(x) - l| < \epsilon/2$ whenever $0 < d(x,a) < \delta$. Let u and v be points in $B^*_\delta(a)$ satisfying $f(u) = \alpha$ and $f(v) = \beta$. Since $|f(u) - l| < \epsilon/2$ and $|f(v) - l| < \epsilon/2$, it follows that

$$\begin{aligned}\epsilon &= |\alpha - \beta| \\ &= |f(u) - f(v)| \\ &\leq |f(u) - l| + |l - f(v)| \\ &< \epsilon\end{aligned}$$

which is an obvious contradiction.

14.3.9. Example. Let $f(x,y) = \dfrac{x^2y^2}{x^2y^2 + (x+y)^4}$ if $(x,y) \neq (0,0)$. Then $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Proof. Exercise. (Solution Q.14.11.)

14.3.10. Example. The limit as $(x,y) \to (0,0)$ of $\dfrac{x^3y^3}{x^{12} + y^4}$ does not exist.

Proof. Problem.

14.3.11. Problem. Let $f\colon A \to \mathbb{R}$ where $A \subseteq M$ and M is a metric space, and let $a \in A'$. Show that

$$\lim_{x\to a} f(x) = 0 \quad \text{ if and only if } \quad \lim_{x\to a} |f(x)| = 0\,.$$

14.3.12. Problem. Let f, g, $h\colon A \to \mathbb{R}$ where $A \subseteq M$ and M is a metric space, and let $a \in A'$. Show that if $f \leq g \leq h$ and

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = l\,,$$

then $\lim_{x\to a} g(x) = l$.

14.3.13. Problem. Let M, N, and P be metric spaces, $a \in A \subseteq M$, $f: A \to N$, and $g: N \to P$.

(a) Show that if $l = \lim_{x\to a} f(x)$ exists and g is continuous at l, then $\lim_{x\to a}(g \circ f)(x)$ exists and is equal to $g(l)$.

(b) Show by example that the following assertion need not be true: If $l = \lim_{x\to a} f(x)$ exists and $\lim_{y\to l} g(y)$ exists, then $\lim_{x\to a}(g \circ f)(x)$ exists.

14.3.14. Problem. Let a be a point in a metric space. Show that

$$\lim_{x\to a} d(x, a) = 0 \,.$$

14.3.15. Problem. In this problem $\mathbb{R}^n$ has its usual metric; in particular,

$$d(x, 0) = \left(\sum_{k=1}^{n} x_k^2\right)^{1/2}$$

for all $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

(a) Show that

$$\lim_{x\to 0} \frac{x_j x_k}{d(x, 0)} = 0$$

whenever $1 \le j \le n$ and $1 \le k \le n$.

(b) For $1 \le k \le n$ show that $\lim_{x\to 0} \dfrac{x_k}{d(x, 0)}$ does not exist.

14.3.16. Problem. Let $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$ for $(x, y) \ne (0, 0)$. Find the following limits, if they exist.

(a) $\lim_{(x,y)\to(0,0)} f(x, y)$

(b) $\lim_{x\to 0}\big(\lim_{y\to 0} f(x, y)\big)$

(c) $\lim_{y\to 0}\big(\lim_{x\to 0} f(x, y)\big)$

14.3.17. Problem. Same as problem 14.3.16, but $f(x, y) = \dfrac{xy}{x^2 + y^2}$.

14.3.18. Problem. Same as problem 14.3.16, but $f(x, y) = \dfrac{x^2 y^2}{x^2 + y^2}$.

14.3.19. Problem. Same as problem 14.3.16, but $f(x, y) = y \sin(1/x)$ if $x \ne 0$ and $f(x, y) = 0$ if $x = 0$.

CHAPTER 15

COMPACT METRIC SPACES

15.1. DEFINITION AND ELEMENTARY PROPERTIES

15.1.1. Definition. Recall that a family $\mathfrak{U}$ of sets is said to COVER a set S if $\bigcup \mathfrak{U} \supseteq S$. The phrases "$\mathfrak{U}$ covers S", "$\mathfrak{U}$ is a cover for S", and "$\mathfrak{U}$ is a covering of S" are used interchangeably. If a cover $\mathfrak{U}$ for a metric space M consists entirely of open subsets of M, then $\mathfrak{U}$ is an OPEN COVER for M. If $\mathfrak{U}$ is a family of sets which covers a set S and $\mathfrak{V}$ is a subfamily of $\mathfrak{U}$ which also covers S, then $\mathfrak{V}$ is a SUBCOVER of $\mathfrak{U}$ for S. A metric space M is COMPACT if *every* open cover of M has a finite subcover.

We have just defined what we mean by a compact *space*. It will be convenient also to speak of a compact *subset* of a metric space M. If $A \subseteq M$, we say that A is a COMPACT SUBSET of M if, regarded as a subspace of M, it is a compact metric space.

Notice that the definition of compactness is identical to the one given for subsets of the real line in definition 6.1.3. Recall also that we have proved that every closed bounded subset of $\mathbb{R}$ is compact (see example 6.3.6).

Remark (A Matter of Technical Convenience)**.** Suppose we wish to show that some particular subset A of a metric space M is compact. Is it really necessary that we work with coverings made up of open subsets of A (as the definition demands) or can we just as well use coverings whose members are open subsets of M? Fortunately, either will do. This is an almost obvious consequence of proposition 11.2.1. Nevertheless, providing a careful verification takes a few lines, and it is good practice to attempt it.

15.1.2. Proposition. *A subset A of a metric space M is compact if and only if every cover of A by open subsets of M has a finite subcover.*

PROOF. Exercise. (Solution Q.15.1.)

An obvious corollary of the preceding proposition: If M_1 is a subspace of a metric space M_2 and $K \subseteq M_1$, then K is a compact subset of M_1 if and only if it is a compact subset of M_2.

15.1.3. Problem. Generalize the result of proposition 6.2.2 to metric spaces. That is, show that every closed subset of a compact metric space is compact.

15.1.4. Definition. A subset of a metric space is BOUNDED if it is contained in some open ball.

15.1.5. Problem. Generalize the result of proposition 6.2.3 to metric spaces. That is, show that every compact subset of a metric space is closed and bounded.

As we will see, the converse of the preceding theorem holds for subsets of $\mathbb{R}^n$ under its usual metric; this is the *Heine-Borel theorem.* It is most important to know, however, that this converse is not true in arbitrary metric spaces, where sets which are closed and bounded may fail to be compact.

15.1.6. Example. Consider an infinite set M under the discrete metric. Regarded as a subset of itself M is clearly closed and bounded. But since the family $\mathfrak{U} = \big\{\{x\} : x \in M\big\}$ is a cover for M which has no proper subcover, the space M is certainly not compact.

15.1.7. Example. With its usual metric $\mathbb{R}^2$ is not compact.

Proof. Problem.

15.1.8. Example. The open unit ball $\{(x, y, z) : x^2 + y^2 + z^2 < 1\}$ in $\mathbb{R}^3$ is not compact.

Proof. Problem.

15.1.9. Example. The strip $\{(x, y) : 2 \le y \le 5\}$ in $\mathbb{R}^2$ is not compact.

Proof. Problem.

15.1.10. Example. The closed first quadrant $\{(x, y) : x \ge 0 \text{ and } y \ge 0\}$ in $\mathbb{R}^2$ is not compact.

Proof. Problem.

15.1.11. Problem. Show that the intersection of an arbitrary nonempty collection of compact subsets of a metric space is itself compact.

15.1.12. Problem. Show that the union of a finite collection of compact subsets of a metric space is itself compact. *Hint.* What about *two* compact sets? Give an example to show that the union of an arbitrary collection of compact subsets of a metric space need not be compact.

15.2. THE EXTREME VALUE THEOREM

15.2.1. Definition. A real valued function f on a metric space M is said to have a (GLOBAL) MAXIMUM at a point a in M if $f(a) \ge f(x)$ for every x in M; the number $f(a)$ is the MAXIMUM VALUE of f. The function has a (GLOBAL) MINIMUM at a if $f(a) \le f(x)$ for every x in M; and in this case $f(a)$ is the MINIMUM VALUE of f. A number is an EXTREME VALUE of f if it is either a maximum or a minimum value. It is clear that a function may fail to have maximum or minimum values. For example, on the open interval $(0, 1)$ the function $f : x \mapsto x$ assumes neither a maximum nor a minimum.

Our next goal is to show that every continuous function on a compact metric space attains both a maximum and a minimum. This turns out to be an easy consequence of the fact that the continuous image of a compact set is compact. All this works exactly as it did for $\mathbb{R}$.

15.2.2. Problem. Generalize the result of theorem 6.3.2 to metric spaces. That is, show that if M and N are metric spaces, if M is compact, and if $f : M \to N$ is continuous, then $f^{\to}(M)$ is compact.

15.2.3. Problem (The Extreme Value Theorem.)**.** Generalize the result of theorem 6.3.3 to metric spaces. That is, show that if M is a compact metric space and $f : M \to \mathbb{R}$ is continuous, then f assumes both a maximum and a minimum value on M.

In chapter 13 we defined the uniform metric on the family $\mathcal{B}(S, \mathbb{R})$ of all bounded real valued functions on S and agreed to call convergence of sequences in this space "uniform convergence". Since S was an arbitrary set (not a metric space), the question of continuity of members of $\mathcal{B}(S, \mathbb{R})$ is meaningless. For the moment we restrict our attention to functions defined on a compact metric space, where the issue of continuity is both meaningful and interesting.

A trivial, but crucial, observation is that if M is a compact metric space, then the family $\mathcal{C}(M) = \mathcal{C}(M, \mathbb{R})$ of continuous real valued functions on M is a subspace of the metric space $\mathcal{B}(M) = \mathcal{B}(M, \mathbb{R})$. This is an obvious consequence of the *extreme value theorem* (15.2.3) which says, in particular, that every continuous real valued function on M is bounded. Furthermore, since the uniform limit of continuous functions is continuous (see proposition 14.2.15), it is clear that $\mathcal{C}(M, \mathbb{R})$ is a closed subset of $\mathcal{B}(M, \mathbb{R})$ (see proposition 12.2.2 for the sequential characterization of closed sets). For future reference we record this formally.

15.2.4. Proposition. *If M is a compact metric space, then $\mathcal{C}(M)$ is a closed subset of $\mathcal{B}(M)$.*

15.2.5. Example. The circle $x^2 + y^2 = 1$ is a compact subset of $\mathbb{R}^2$.

PROOF. Problem. *Hint.* Parametrize.

15.2.6. Example. The ellipse $\dfrac{x^2}{16} + \dfrac{y^2}{9} = 1$ is a compact subset of $\mathbb{R}^2$.

PROOF. Problem.

15.2.7. Example. Regard $\mathbb{R}^2$ as a metric space under the uniform metric (see example 9.2.12). Then the boundary of the unit ball in this space is compact.

PROOF. Problem.

15.2.8. Problem. Let $f\colon M \to N$ be a continuous bijection between metric spaces.

(a) Show by example that f need not be a homeomorphism.
(b) Show that if M is compact, then f must be a homeomorphism.

15.3. DINI'S THEOREM

15.3.1. Definition. A family $\mathfrak{F}$ of sets is said to have the FINITE INTERSECTION PROPERTY if every finite subfamily of $\mathfrak{F}$ has nonempty intersection.

15.3.2. Problem. Show that a metric space M is compact if and only if every family of closed subsets of M having the finite intersection property has nonempty intersection.

15.3.3. Proposition (Dini's Theorem)**.** *Let M be a compact metric space and (f_n) be a sequence of continuous real valued functions on M such that $f_n(x) \geq f_{n+1}(x)$ for all x in M and all n in $\mathbb{N}$. If the sequence (f_n) converges pointwise on M to a continuous function g, then it converges uniformly to g.*

PROOF. Problem. *Hint.* First establish the correctness of the assertion for the special case where $g = 0$. For $\epsilon > 0$ consider the sets $A_n = f_n{}^{\leftarrow}([\epsilon, \infty))$. Argue by contradiction to show that $\cap_{n=1}^{\infty} A_n$ is empty. Then use problem 15.3.2.

15.3.4. Example. *Dini's theorem* (problem 15.3.3) is no longer true if we remove the hypothesis that

(a) the sequence (f_n) is decreasing;
(b) the function g is continuous; or
(c) the space M is compact.

PROOF. Problem.

15.3.5. Example. On the interval $[0, 1]$ the square root function $x \mapsto \sqrt{x}$ is the uniform limit of a sequence of polynomials.

PROOF. Problem. *Hint.* Let p_0 be the zero function on $[0, 1]$, and for $n \geq 0$ define p_{n+1} on $[0, 1]$ by

$$p_{n+1}(t) = p_n(t) + \tfrac{1}{2}\big(t - (p_n(t))^2\big)\,,$$

and verify that

$$0 \leq \sqrt{t} - p_n(t) \leq \frac{2\sqrt{t}}{2 + n\sqrt{t}} \leq \frac{2}{n}$$

for $0 \leq t \leq 1$ and $n \in \mathbb{N}$.

CHAPTER 16

SEQUENTIAL CHARACTERIZATION OF COMPACTNESS

We have previously characterized open sets, closed sets, closure, and continuity by means of sequences. Our next goal is to produce a characterization of compactness in terms of sequences. This is achieved in theorem 16.2.1 where it is shown that compactness in metric spaces is equivalent to something called *sequential compactness*. For this concept we need to speak of *subsequences* of sequences of points in a metric space. In 4.4.1 we defined "subsequence" for sequences of real numbers. There is certainly no reason why we cannot speak of subsequences of arbitrary sequences.

16.1. SEQUENTIAL COMPACTNESS

16.1.1. Definition. If a is a sequence of elements of a set S and $n\colon \mathbb{N} \to \mathbb{N}$ is strictly increasing, then the composite function $a \circ n$ is a SUBSEQUENCE of the sequence a. (Notice that $a \circ n$ is itself a sequence of elements of S since it maps $\mathbb{N}$ into S.) The k^{th} term of $a \circ n$ is frequently written as a_{n_k}; other acceptable notations are $a_{n(k)}$ and $a(n(k))$.

Notice that it is possible for a sequence which fails to converge to have subsequences which do converge. For example, if $a_n = (-1)^n + (1/n)$ for each n, then the subsequence (a_{2n}) converges while the sequence (a_n) itself does not.

16.1.2. Definition. A metric space M is SEQUENTIALLY COMPACT if every sequence in M has a convergent subsequence.

16.1.3. Example. It is important to understand that for a space to be sequentially compact the preceding definition requires that every sequence in the space have a subsequence which converges to a point *in that space*. It is not enough to find a subsequence which converges in some larger space. For example, we know that the metric space $(0, 1]$ regarded as a subspace of $\mathbb{R}$ is not sequentially compact because the sequence $\left(\frac{1}{n}\right)$ has no subsequence which converges to something in $(0, 1]$. That $\left(\frac{1}{n}\right)$ happens to converge to 0 in $\mathbb{R}$ is completely irrelevant.

A major goal of this chapter is to demonstrate that in metric spaces compactness and sequential compactness are the same thing. This is done in theorem 16.2.1. Be aware however that in general topological spaces this is *not* true. An essential ingredient of the proof is the following chain of implications

$$\text{sequentially compact} \Longrightarrow \text{totally bounded} \Longrightarrow \text{separable}.$$

So the next order of business is to define the last two concepts and prove that the preceding implications do hold.

16.1.4. Definition. A metric space M is TOTALLY BOUNDED if for every $\epsilon > 0$ there exists a finite subset F of M such that for every $a \in M$ there is a point $x \in F$ such that $d(x, a) < \epsilon$. This definition has a more or less standard paraphrase: A space is totally bounded if it can be kept under surveillance by a finite number of arbitrarily near-sighted policemen.

16.1.5. Proposition. *Every totally bounded metric space is bounded.*

PROOF. Problem.

16.1.6. Example. The converse of the preceding proposition is false. Any infinite set with the discrete metric is an example of a bounded metric space which is not totally bounded. (Why?)

16.1.7. Proposition. *Every sequentially compact metric space is totally bounded.*

Proof. Exercise. *Hint.* Assume that a metric space M is not totally bounded. Inductively construct a sequence in M no two terms of which are closer together than some fixed distance $\epsilon > 0$. (Solution Q.16.1.)

16.1.8. Definition. A metric space is SEPARABLE if it possesses a countable dense subset.

16.1.9. Example. The space $\mathbb{R}^n$ is separable. The set of points $(q_1, \dots, q_n)$ such that each coordinate q_k is rational is a countable dense subset of $\mathbb{R}^n$.

16.1.10. Example. It follows easily from proposition 11.1.22 that the real line (or indeed any uncountable set) with the discrete metric is *not* separable. (Consider the open balls of radius 1 about each point.)

16.1.11. Proposition. *Every totally bounded metric space is separable.*

Proof. Problem. *Hint.* Let M be a totally bounded metric space. For each n in $\mathbb{N}$ choose a finite set F_n such that for each $a \in M$ the set $F_n \cap B_{\frac{1}{n}}(a)$ is nonempty.

16.1.12. Problem. The metric space $\mathbb{R}_d$ comprising the real numbers under the discrete metric is not separable.

16.1.13. Corollary. *Every sequentially compact metric space is separable.*

Proof. Propositions 16.1.7 and 16.1.11. □

16.2. CONDITIONS EQUIVALENT TO COMPACTNESS

We are now ready for a major result of this chapter—a sequential characterization of compactness. We show that a space M is compact if and only if every sequence in M has a convergent subsequence. In other words, a space is compact if and only if it is sequentially compact. We will see later in the section how useful this result is when we use it to prove that a finite product of compact spaces is compact.

The following theorem also provides a second characterization of compactness: a space M is compact if and only if every infinite subset of M has a point of accumulation in M. The proof of the theorem is fairly straightforward except for one complicated bit. It is not easy to prove that every sequentially compact space is compact. This part of the proof comes equipped with an lengthy hint.

16.2.1. Theorem. *If M is a metric space then the following are equivalent:*

(1) *M is compact;*
(2) *every infinite subset of M has an accumulation point in M;*
(3) *M is sequentially compact.*

Proof. Exercise. *Hint.* Showing that (3) implies (1) is not so easy. To show that a sequentially compact metric space M is compact start with an arbitrary open cover $\mathfrak{U}$ for M and show first that

(A) $\mathfrak{U}$ has a *countable* subfamily $\mathfrak{V}$ which covers M.

Then show that

(B) there is a finite subfamily of $\mathfrak{V}$ which covers M.

The hard part is (A). According to corollary 16.1.13 we may choose a countable dense subset A of M. Let $\mathfrak{B}$ be the family of all open balls $B_r(a)$ such that

(i) $a \in A$;
(ii) $r \in \mathbb{Q}$ and;
(iii) $B_r(a) \subseteq U$ for some $U \in \mathfrak{U}$.

For each B in $\mathfrak{B}$ choose a set U_B in $\mathfrak{U}$ such that $B \subseteq U_B$ and let

$$\mathfrak{V} = \{U_B \colon B \in \mathfrak{B}\}.$$

Then verify that $\mathfrak{V}$ is a countable subfamily of $\mathfrak{U}$ which covers M.

To show that $\mathfrak{V}$ covers M, start with an arbitrary point $x \in M$ and a set $U \in \mathfrak{U}$ containing x. All that is needed is to find an open ball $B_s(a)$ in $\mathfrak{B}$ such that $x \in B_s(a) \subseteq U$. In order to do this the point $a \in A$ must be taken sufficiently close to x so that it is possible to choose a rational number s which is both

(i) small enough for $B_s(a)$ to be a subset of U, and

(ii) large enough for $B_s(a)$ to contain x.

To establish (B) let $(V_1, V_2, V_3, \dots)$ be an enumeration of $\mathfrak{V}$ and $W_n = \cup_{k=1}^{n} V_k$ for each $n \in \mathbb{N}$. If no one set W_n covers M, then for every $k \in \mathbb{N}$ there is a point $x_k \in {W_k}^c$. (Solution Q.16.2.)

16.2.2. Problem. Use theorem 16.2.1 to give three different proofs that the metric space $[0, 1)$ (with the usual metric inherited from $\mathbb{R}$) is not compact.

CAUTION. It is a common (and usually helpful) mnemonic device to reduce statements of complicated theorems in analysis to brief paraphrases. In doing this considerable care should be exercised so that crucial information is not lost. Here is an example of the kind of thing that can go wrong.

Consider the two statements:

(1) In the metric space $\mathbb{R}$ every infinite subset of the open unit interval has a point of accumulation; and
(2) A metric space is compact if every infinite subset has a point of accumulation.

Assertion (1) is a special case of proposition 4.4.9; and (2) is just part of theorem 16.2.1. The unwary tourist might be tempted to conclude from (1) and (2) that the open unit interval $(0, 1)$ is compact, which, of course, it is not. The problem here is that (1) is a correct assertion about $(0, 1)$ regarded as a subset of the space $\mathbb{R}$; every infinite subset of $(0, 1)$ does have a point of accumulation *lying in* $\mathbb{R}$.

If, however, the metric space under consideration is $(0, 1)$ itself, then (1) is no longer true. For example, the set of all numbers of the form $1/n$ for $n \geq 2$ has no accumulation point in $(0, 1)$. When we use (2) to establish the compactness of a metric space M, what we must verify is that every infinite subset of M has a point of accumulation *which lies in* M. Showing that these points of accumulation exist in some space which contains M just does not do the job. The complete statements of 4.4.9 and 16.2.1 make this distinction clear; the paraphrases (1) and (2) above do not.

16.3. PRODUCTS OF COMPACT SPACES

It is possible using just the definition of compactness to prove that the product of two compact metric spaces is compact. It is a pleasant reward for the effort put into proving theorem 16.2.1 that it can be used to give a genuinely simple proof of this important result.

16.3.1. Theorem. *If M_1 and M_2 are compact metric spaces, then so is $M_1 \times M_2$.*

Proof. Problem.

16.3.2. Corollary. *If $M_1, \dots, M_n$ are compact metric spaces, then the product space $M_1 \times \dots \times M_n$ is compact.*

Proof. Induction. □

16.3.3. Problem. Let A and B be subsets of a metric space. Recall from definition 12.2.10 that the distance between A and B is defined by

$$d(A, B) := \inf\{d(a, b) \colon a \in A \text{ and } b \in B\}.$$

Prove or disprove:

(a) If $A \cap B = \emptyset$, then $d(A, B) > 0$.
(b) If A and B are closed and $A \cap B = \emptyset$, then $d(A, B) > 0$.
(c) If A is closed, B is compact, and $A \cap B = \emptyset$, then $d(A, B) > 0$. *Hint.* If $d(A, B) = 0$ then there exist sequences a in A and b in B such that $d(a_n, b_n) \to 0$.
(d) If A and B are compact and $A \cap B = \emptyset$, then $d(A, B) > 0$.

16.3.4. Problem. Let a, $b > 0$. The elliptic disk

$$D := \left\{(x, y) \colon \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}$$

is a compact subset of $\mathbb{R}^2$. *Hint.* Write the disk as a continuous image of the unit square $[0, 1] \times [0, 1]$.

16.3.5. Problem. The unit sphere

$$S^2 \equiv \{(x, y, z) \colon x^2 + y^2 + z^2 = 1\}$$

is a compact subset of $\mathbb{R}^3$. *Hint.* Spherical coordinates.

16.3.6. Problem. Show that the interval $[0, \infty)$ is not a compact subset of $\mathbb{R}$ using each of the following:

(a) The definition of compactness.
(b) Proposition 15.1.2.
(c) Proposition 15.1.5.
(d) The *extreme value theorem* (15.2.3).
(e) Theorem 16.2.1, condition (2).
(f) Theorem 16.2.1, condition (3).
(g) The *finite intersection property.* (See problem 15.3.2.)
(h) *Dini's theorem.* (See problem 15.3.3.)

16.4. THE HEINE-BOREL THEOREM

We have seen in proposition 15.1.5 that in an arbitrary metric space compact sets are always closed and bounded. Recall also that the converse of this is not true in general (see problem 6.3.9). In $\mathbb{R}^n$, however, the converse does indeed hold. This assertion is the *Heine-Borel theorem.* Notice that in example 6.3.6 we have already established its correctness for the case $n = 1$. The proof of the general case is now just as easy—we have done all the hard work in proving that the product of finitely many compact spaces is compact (corollary 16.3.2).

16.4.1. Theorem (Heine-Borel Theorem)**.** *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

PROOF. Exercise. (Solution Q.16.3.)

16.4.2. Example. The triangular region T whose vertices are $(0, 0)$, $(1, 0)$, and $(0, 1)$ is a compact subset of $\mathbb{R}^2$.

PROOF. Define functions $f, g, h : \mathbb{R}^2 \to \mathbb{R}$ by $f(x, y) = x$, $g(x, y) = y$, and $h(x, y) = x + y$. Each of these functions is continuous. Thus the sets

$$\begin{aligned} A &:= f^{\leftarrow}[0, \infty) = \{(x, y) : x \geq 0\}, \\ B &:= g^{\leftarrow}[0, \infty) = \{(x, y) : y \geq 0\}, \quad \text{and} \\ C &:= h^{\leftarrow}(-\infty, 1] = \{(x, y) : x + y \leq 1\} \end{aligned}$$

are all closed sets. Thus $T = A \cap B \cap C$ is closed. It is bounded since it is contained in the open ball about the origin with radius 2. Thus by the *Heine-Borel theorem* (16.4.1), T is compact. □

16.4.3. Problem. Do problem 16.3.4 again, this time using the *Heine-Borel theorem.*

16.4.4. Problem (Bolzano-Weierstrass Theorem)**.** Every bounded infinite subset of $\mathbb{R}^n$ has at least one point of accumulation in $\mathbb{R}^n$ (compare proposition 4.4.9).

16.4.5. Problem (Cantor Intersection Theorem)**.** If (A_n) is a nested sequence of nonempty closed bounded subsets of $\mathbb{R}^n$, then $\bigcap_{n=1}^{\infty} A_n$ is nonempty. Furthermore, if $\operatorname{diam} A_n \to 0$, then $\bigcap_{n=1}^{\infty} A_n$ is a single point.

16.4.6. Problem. Use the *Cantor intersection theorem* (problem 16.4.5) to show that the medians of a triangle are concurrent.

16.4.7. Problem. Let m be the set of all bounded sequences of real numbers under the uniform metric:

$$d_u(a, b) = \sup\{|a_n - b_n| \colon n \in \mathbb{N}\}$$

whenever a, $b \in m$.

(a) Show by example that the *Heine-Borel theorem* (16.4.1) does not hold for subsets of m.
(b) Show by example that the *Bolzano-Weierstrass theorem* does not hold for subsets of m. (See problem 16.4.4.)
(c) Show by example that the *Cantor intersection theorem* does not hold for subsets of m. (See problem 16.4.5.)

16.4.8. Problem. Find a metric space M with the property that every infinite subset of M is closed and bounded but not compact.

16.4.9. Problem. Prove or disprove: If both the interior and the boundary of a set $A \subseteq \mathbb{R}$ are compact, then so is A.

CHAPTER 17

CONNECTEDNESS

In chapter 5 we discussed connected subsets of the real line. Although they are easily characterized (they are just the intervals), they possess important properties, most notably the *intermediate value property.* Connected subsets of arbitrary metric spaces can be somewhat more complicated, but they are no less important.

17.1. CONNECTED SPACES

17.1.1. Definition. A metric space M is DISCONNECTED if there exist disjoint nonempty open sets U and V whose union is M. In this case we say that the sets U and V DISCONNECT M. A metric space is CONNECTED if it is not disconnected. A subset of a metric space M is CONNECTED (respectively, DISCONNECTED) if it is connected (respectively, disconnected) as a subspace of M. Thus a subset N of M is disconnected if there exist nonempty disjoint sets U and V open in the *relative* topology on N whose union is N.

17.1.2. Example. Every discrete metric space with more than one point is disconnected.

17.1.3. Example. The set $\mathbb{Q}^2$ of points in $\mathbb{R}^2$ both of whose coordinates are rational is a disconnected subset of $\mathbb{R}^2$.

PROOF. The subspace $\mathbb{Q}^2$ is disconnected by the sets $\{(x,y) \in \mathbb{Q}^2 \colon x < \pi\}$ and $\{(x,y) \in \mathbb{Q}^2 \colon x > \pi\}$. (Why are these sets open in the relative topology on $\mathbb{Q}^2$?) $\square$

17.1.4. Example. The following subset of $\mathbb{R}^2$ is not connected.

$$\{(x, x^{-1}) \colon x > 0\} \cup \{(0, y) \colon y \in \mathbb{R}\}$$

PROOF. Problem.

17.1.5. Proposition. *A metric space M is disconnected if and only if there exists a continuous surjection from M onto a two element discrete space, say $\{0, 1\}$. A metric space M is connected if and only if every continuous function f from M into a two element discrete space is constant.*

PROOF. Problem.

Proposition 5.1.2 remains true for arbitrary metric spaces; and the same proof works.

17.1.6. Proposition. *A metric space M is connected if and only if the only subsets of M which are both open and closed are the null set and M itself.*

PROOF. Exercise. (Solution Q.17.1.)

Just as in chapter 5, dealing with the relative topology on a subset of a metric space can sometimes be a nuisance. The remedy used there is available here: work with mutually separated sets.

17.1.7. Definition. Two nonempty subsets C and D of a metric space M are said to be MUTUALLY SEPARATED if

$$C \cap \overline{D} = \overline{C} \cap D = \emptyset\,.$$

17.1.8. Proposition. *A subset N of a metric space M is disconnected if and only if it is the union of two nonempty sets mutually separated in M.*

Proof. Exercise. *Hint.* Make a few changes in the proof of proposition 5.1.7. (Solution Q.17.2.)

17.1.9. Proposition. *If A is a connected subset of a metric space, then any set B satisfying $A \subseteq B \subseteq \overline{A}$ is also connected.*

Proof. Problem. *Hint.* Use proposition 17.1.8

If a metric space is disconnected, it is often a rather simple job to demonstrate this fact. All one has to do is track down two subsets which disconnect the space. If the space is connected, however, one is confronted with the unenviable task of showing that *every* pair of subsets fails for some reason to disconnect the space. How, for example, does one go about showing that the unit square $[0, 1] \times [0, 1]$ is connected? Or the unit circle? Or the curve $y = \sin x$? In 17.1.10, 17.1.13, and 17.2.2 we give sufficient conditions for a metric space to be connected. The first of these is that the space be the continuous image of a connected space .

17.1.10. Theorem. *A metric space N is connected if there exist a connected metric space M and a continuous surjection from M onto N.*

Proof. Change "$\mathbb{R}$" to "M" in the proof of theorem 5.2.1. □

17.1.11. Example. The graph of the curve $y = \sin x$ is a connected subset of $\mathbb{R}^2$.

Proof. Exercise. (Solution Q.17.3.)

17.1.12. Example. The unit circle $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ is a connected subset of the plane.

Proof. Problem.

17.1.13. Proposition. *A metric space is connected if it is the union of a family of connected subsets with nonempty intersection.*

Proof. Exercise. (Argue by contradiction. Use definition 17.1.1.) (Solution Q.17.4.)

17.1.14. Problem. Use proposition 17.1.5 to give a second proof of proposition 17.1.13.

17.1.15. Example. The closed unit square $[0, 1] \times [0, 1]$ in $\mathbb{R}^2$ is connected.

Proof. Exercise. (Solution Q.17.5.)

17.2. ARCWISE CONNECTED SPACES

A concept closely related to (but stronger than) connectedness is *arcwise connectedness.*

17.2.1. Definition. A metric space is ARCWISE CONNECTED (or PATH CONNECTED) if for every $x, y \in M$ there exists a continuous map $f : [0, 1] \to M$ such that $f(0) = x$ and $f(1) = y$. Such a function f is an ARC (or PATH , or CURVE) connecting x to y. It is very easy to prove that arcwise connected spaces are connected (proposition 17.2.2). The converse is false (example 17.2.7). If, however, we restrict our attention to open subsets of $\mathbb{R}^n$, then the converse does hold (proposition 17.2.8).

17.2.2. Proposition. *If a metric space is arcwise connected, then it is connected.*

Proof. Problem.

17.2.3. Example. The following subset of $\mathbb{R}^2$ is not connected.

$$\{(x, y) : (x - 1)^2 + (y - 1)^2 < 4\} \cup \{(x, y) : x < 0\} \\ \cup \{(x, y) : (x - 10)^2 + (y - 1)^2 < 49\}$$

Proof. Problem.

17.2.4. Example. The following subset of $\mathbb{R}^2$ is connected.

$$\{(x,y)\colon (x-1)^2+(y-1)^2<4\}\cup\{(x,y)\colon y<0\}$$
$$\cup\,\{(x,y)\colon (x-10)^2+(y-1)^2<49\}$$

Proof. Problem.

17.2.5. Example. The following subset of $\mathbb{R}^2$ is connected.

$$\{(x,x^3+2x)\colon x\in\mathbb{R}\}\cup\{(x,x^2+56)\colon x\in\mathbb{R}\}$$

Proof. Problem.

17.2.6. Example. Every open ball in $\mathbb{R}^n$ is connected. So is every closed ball.

Proof. Problem.

17.2.7. Example. Let $B=\{(x,\sin x^{-1})\colon 0<x\le 1\}$. Then $\overline{B}$ is a connected subset of $\mathbb{R}^2$ but is not arcwise connected.

Proof. Exercise. *Hint.* Let $M=\overline{B}$. To show that M is not arcwise connected argue by contradiction. Assume there exists a continuous function $f\colon[0,1]\to M$ such that $f(1)\in B$ and $f(0)\notin B$. Prove that $f^2=\pi_2\circ f$ is discontinuous at the point $t_0=\sup f^{\leftarrow}(M\setminus B)$. To this end show that $t_0\in f^{\leftarrow}(M\setminus B)$. Then, given $\delta>0$, choose t_1 in $[0,1]$ so that $t_0<t_1<t_0+\delta$. Without loss of generality one may suppose that $f^2(t_0)\le 0$. Show that $\big(f^1\big)^{\rightarrow}[t_0,t_1]$ is an interval containing 0 and $f^1(t_1)$ (where $f^1=\pi_1\circ f$). Find a point t in $[t_0,t_1]$ such that $0<f^1(t)<f^1(t_1)$ and $f^2(t)=1$. (Solution Q.17.6.)

17.2.8. Proposition. *Every connected open subset of $\mathbb{R}^n$ is arcwise connected.*

Proof. Exercise. *Hint.* Let A be a connected open subset of $\mathbb{R}^n$ and $a\in A$. Let U be the set of all points in A which can be joined to a by an arc in A. Show that $A\setminus U$ is empty by showing that U and $A\setminus U$ disconnect A. (Solution Q.17.7.)

17.2.9. Problem. Does there exist a continuous bijection from a closed disk in $\mathbb{R}^2$ to its circumference? Does there exist a continuous bijection from the interval $[0,1]$ to the circumference of a disk in $\mathbb{R}^2$?

17.2.10. Problem. Let x be a point in a metric space M. Define the COMPONENT of M containing x to be the largest connected subset of M which contains x. Discover as much about components of metric spaces as you can. First, of course, you must make sure that the definition just given makes sense. (How do we know that there really *is* a "largest" connected set containing x?)

Here are some more things to think about.

(1) The components of a metric space are a disjoint family whose union is the whole space.
(2) It is fairly clear that the components of a discrete metric space are the points of the space. If the components are points must the space be discrete?
(3) Components of a metric space are closed sets; must they be open?
(4) Distinct components of a metric space are mutually separated.
(5) If a metric space M is the union of two mutually separated sets C and D and if points x and y belong to the same component of M, then both points are in C or both are in D. What about the converse? Suppose x and y are points in M such that whenever M is written as the union of two mutually separated sets C and D, both points lie in C or both lie in D. Must x and y lie in the same component?

17.2.11. Problem. A function f in $\mathcal{C}(M,\mathbb{R})$, where M is a metric space, is IDEMPOTENT if $(f(x))^2=f(x)$ for all $x\in M$. The constant functions 0 and 1 are the TRIVIAL idempotents of $\mathcal{C}(M,\mathbb{R})$. Show that $\mathcal{C}(M,\mathbb{R})$ possesses a nontrivial idempotent if and only if the underlying metric space is disconnected. (This is one of a large number of results which link algebraic properties of $\mathcal{C}(M,\mathbb{R})$ to topological properties of the underlying space M.)

CHAPTER 18

COMPLETE SPACES

18.1. CAUCHY SEQUENCES

18.1.1. Definition. A sequence (x_n) in a metric space is a CAUCHY SEQUENCE if for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(x_m, x_n) < \epsilon$ whenever $m, n \geq n_0$. This condition is often abbreviated as follows: $d(x_m, x_n) \to 0$ as $m, n \to \infty$ (or $\lim_{m,n\to\infty} d(x_m, x_n) = 0$).

18.1.2. Example. In the metric space $\mathbb{R}$ the sequence $(1/n)$ is Cauchy.

PROOF. Given $\epsilon > 0$ choose $n_0 > 2/\epsilon$. If $m, n > n_0$, then $d(1/m, 1/n) = |(1/m) - (1/n)| \leq (1/m) + (1/n) \leq 2/n_0 < \epsilon$. Notice that in $\mathbb{R}$ this sequence is also convergent. □

18.1.3. Example. In the metric space $\mathbb{R} \setminus \{0\}$ the sequence $(1/n)$ is Cauchy. (The proof is exactly the same as in the preceding example.) Notice, however, that this sequence does not converge in $\mathbb{R} \setminus \{0\}$.

18.1.4. Proposition. *In a metric space every convergent sequence is Cauchy.*

PROOF. Exercise. (Solution Q.18.1.)

18.1.5. Proposition. *Every Cauchy sequence which has a convergent subsequence is itself convergent (and to the same limit as the subsequence).*

PROOF. Exercise. (Solution Q.18.2.)

18.1.6. Proposition. *Every Cauchy sequence is bounded.*

PROOF. Exercise. (Solution Q.18.3.)

Although every convergent sequence is Cauchy (proposition 18.1.4), the converse need not be true (example 18.1.3). Those spaces for which the converse is true are said to be *complete*.

18.2. COMPLETENESS

18.2.1. Definition. A metric space M is COMPLETE if every Cauchy sequence in M converges to a point of M.

18.2.2. Example. The metric space $\mathbb{R}$ is complete.

PROOF. Let (x_n) be a Cauchy sequence in $\mathbb{R}$. By proposition 18.1.6 the sequence (x_n) is bounded; by corollary 4.4.4 it has a convergent subsequence; and so by proposition 18.1.5 it converges. □

18.2.3. Example. If (x_n) and (y_n) are Cauchy sequences in a metric space, then $\big(d(x_n, y_n)\big)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$.

PROOF. Problem. *Hint.* Proposition 9.2.17.

18.2.4. Example. The set $\mathbb{Q}$ of rational numbers (regarded as a subspace of $\mathbb{R}$) is not complete.

PROOF. Problem.

18.2.5. Proposition. *Every compact metric space is complete.*

Proof. Problem. *Hint.* Theorem 16.2.1 and proposition 18.1.5.

18.2.6. Problem. Let M be a metric space with the discrete metric.

(a) Which sequences in M are Cauchy?
(b) Show that M is complete.

18.2.7. Problem. Show that completeness is not a topological property.

18.2.8. Proposition. *Let M be a complete metric space and M_0 be a subspace of M. Then M_0 is complete if and only if it is a closed subset of M.*

Proof. Problem.

18.2.9. Proposition. *The product of two complete metric spaces is complete.*

Proof. Exercise. (Solution Q.18.4.)

18.2.10. Proposition. *If d and ρ are strongly equivalent metrics on a set M, then the space (M, d) is complete if and only if (M, ρ) is.*

Proof. Exercise. (Solution Q.18.5.)

18.2.11. Example. With its usual metric the space $\mathbb{R}^n$ is complete.

Proof. Since $\mathbb{R}$ is complete (18.2.2), proposition 18.2.9 and induction show that $\mathbb{R}^n$ is complete under the metric d_1 (defined in 9.2.10). Since the usual metric is strongly equivalent to d_1, we may conclude from proposition 18.2.10 that $\mathbb{R}^n$ is complete under its usual metric. □

Here is one more example of a complete metric space.

18.2.12. Example. If S is a set, then the metric space $\mathcal{B}(S, \mathbb{R})$ is complete.

Proof. Exercise. (Solution Q.18.6.) □

18.2.13. Example. If M is a compact metric space, then $\mathcal{C}(M, \mathbb{R})$ is a complete metric space.

Proof. Problem.

18.2.14. Problem. Give examples of metric spaces M and N, a homeomorphism $f\colon M \to N$, and a Cauchy sequence (x_n) in M such that the sequence $\big(f(x_n)\big)$ is *not* Cauchy in N.

18.2.15. Problem. Show that if D is a dense subset of a metric space M and every Cauchy sequence in D converges to a point of M, then M is complete.

18.3. COMPLETENESS VS. COMPACTNESS

In proposition 18.2.5 we saw that every compact metric space is complete. The converse of this is not true without additional assumptions. (Think of the reals.) In the remainder of this chapter we show that adding total boundedness to completeness will suffice. For the next problem we require the notion of the "diameter" of a set in a metric space.

18.3.1. Definition. The DIAMETER of a subset A of a metric space is defined by

$$\operatorname{diam} A := \sup\{d(x, y)\colon x, y \in A\}$$

if this supremum exists. Otherwise $\operatorname{diam} A := \infty$.

18.3.2. Problem. Show that $\operatorname{diam} A = \operatorname{diam} \overline{A}$ for every subset A of a metric space.

18.3.3. Proposition. *In a metric space M the following are equivalent:*

(1) *M is complete.*
(2) *Every nested sequence of nonempty closed sets in M whose diameters approach 0 has nonempty intersection.*

Proof. Problem. *Hint.* For the definition of "nested" see 4.4.5. To show that (1) implies (2), let (F_k) be a nested sequence of nonempty closed subsets of M. For each k choose $x_k \in F_k$. Show that the sequence (x_k) is Cauchy. To show that (2) implies (1), let (x_k) be a Cauchy sequence in M. Define $A_n = \{x_k \colon k \geq n\}$ and $F_n = \overline{A_n}$. Show that (F_n) is a nested sequence of closed sets whose diameters approach 0 (see the preceding problem). Choose a point a in $\cap F_n$. Find a subsequence (x_{n_k}) of (x_n) such that $d\,(a, x_{n_k}) < 2^{-k}$.

18.3.4. Problem. Since $\mathbb{R}$ is complete, the preceding problem tells us that every nested sequence of nonempty closed subsets of $\mathbb{R}$ whose diameters approach 0 has nonempty intersection.

(a) Show that this statement is no longer correct if the words "whose diameters approach 0" are deleted.
(b) Show that the statement is no longer correct if the word "closed" is deleted.

18.3.5. Proposition. *In a totally bounded metric space every sequence has a Cauchy subsequence.*

Proof. Problem. *Hint.* Let (x_n) be a sequence in a totally bounded metric space M. For every $n \in \mathbb{N}$ the space M can be covered by a finite collection of open balls of radius $1/n$. Thus, in particular, there is an open ball of radius 1 which contains infinitely many of the terms of the sequence (x_n). Show that it is possible inductively to choose subsets $N_1, N_2, \dots$ of $\mathbb{N}$ such that for every $m, n \in \mathbb{N}$

(i) $n > m$ implies $N_n \subseteq N_m$,
(ii) N_n is infinite, and
(iii) $\{x_k : k \in N_n\}$ is contained in some open ball of radius $1/n$.

Then show that we may choose (again, inductively) $n_1, n_2, \dots$ in $\mathbb{N}$ such that for every $j, k \in \mathbb{N}$

(iv) $k > j$ implies $n_k > n_j$, and
(v) $n_k \in N_k$.

Finally, show that the sequence (x_n) is Cauchy.

18.3.6. Proposition. *A metric space is compact if and only if it is complete and totally bounded.*

Proof. Problem.

18.3.7. Problem. Let (x_n) be a sequence of real numbers with the property that each term of the sequence (from the third term on) is the average of the two preceding terms. Show that the sequence converges and find its limit. *Hint.* Proceed as follows.

(a) Compute the distance between x_{n+1} and x_n in terms of the distance between x_n and x_{n-1}.
(b) Show (inductively) that

$$|x_{n+1} - x_n| = 2^{1-n}|x_2 - x_1|\,.$$

(c) Prove that (x_n) has a limit by showing that for $m < n$

$$|x_n - x_m| \leq 2^{2-m}|x_2 - x_1|\,.$$

(d) Show (again inductively) that $2x_{n+1} + x_n = 2x_2 + x_1$.

18.3.8. Problem. Show that if (x_n) is a sequence lying in the interval $[-1, 1]$ which satisfies

$$|x_{n+1} - x_n| \leq \tfrac{1}{4}|x_n^2 - x_{n-1}^2| \qquad \text{for } n \geq 2\,,$$

then (x_n) converges.

CHAPTER 19

APPLICATIONS OF A FIXED POINT THEOREM

We now have enough information about metric spaces to consider some interesting applications. We will first prove a result known as the *contractive mapping theorem* and then use it to find solutions to systems of simultaneous linear equations and to certain integral equations. Since this chapter contains mostly examples, we will make liberal use of computations from beginning calculus. Although it would perhaps be more logical to defer these matters until we have developed the necessary facts concerning integrals, derivatives, and power series, there is nevertheless much to be said for presenting some nontrivial applications relatively early.

19.1. THE CONTRACTIVE MAPPING THEOREM

19.1.1. Definition. A mapping $f\colon M \to N$ between metric spaces is CONTRACTIVE if there exists a constant c such that $0 < c < 1$ and

$$d(f(x), f(y)) \leq c\, d(x, y)$$

for all x, $y \in M$. Such a number c is a CONTRACTION CONSTANT for f. A contractive map is also called a CONTRACTION.

19.1.2. Exercise. Show that every contractive map is continuous. (Solution Q.19.1.)

19.1.3. Example. The map $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$f(x, y) = \left(1 - \tfrac{1}{3}x, 1 + \tfrac{1}{3}y, 2 + \tfrac{1}{3}x - \tfrac{1}{3}y\right)$$

is a contraction.

PROOF. Exercise. (Solution Q.19.2.)

19.1.4. Example. The map

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto \left(\tfrac{1}{2}(1 + y), \tfrac{1}{2}(3 - x)\right)$$

is a contraction on $\mathbb{R}^2$, where $\mathbb{R}^2$ has its usual (Euclidean) metric.

PROOF. Problem.

The next theorem is the basis for a number of interesting applications. Also it will turn out to be a crucial ingredient of the extremely important *inverse function theorem* (in chapter 29). Although the statement of theorem 19.1.5 is important in applications, its proof is even more so. The theorem guarantees the existence (and uniqueness) of solutions to certain kinds of equations; its proof allows us to approximate these solutions as closely as our computational machinery permits. Recall from chapter 5 that a FIXED POINT of a mapping f from a set S into itself is a point $p \in S$ such that $f(p) = p$.

19.1.5. Theorem (Contractive Mapping Theorem)**.** *Every contraction from a complete metric space into itself has a unique fixed point.*

PROOF. Exercise. *Hint.* Let M be a complete metric space and $f\colon M \to M$ be contractive. Start with an arbitrary point x_0 in M. Obtain a sequence $(x_n)_{n=0}^{\infty}$ of points in M by letting $x_1 = f(x_0)$, $x_2 = f(x_1)$, and so on. Show that this sequence is Cauchy. (Solution Q.19.3.) □

19.1.6. Example. We use the *contractive mapping theorem* to solve the following system of equations

$$\begin{cases} 9x - 2y = 7 \\ 3x + 8y = 11\,. \end{cases} \tag{19.1}$$

Define

$$S\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (9x - 2y, 3x + 8y)\,.$$

The system (19.1) may be written as a single equation

$$S(x,y) = (7,11)\,,$$

or equivalently as

$$(x,y) - S(x,y) + (7,11) = (x,y)\,. \tag{19.2}$$

[*Definition.* Addition and subtraction on $\mathbb{R}^2$ are defined coordinatewise. That is, if (x,y) and (u,v) are points in $\mathbb{R}^2$, then $(x,y)+(u,v) := (x+u, y+v)$ and $(x,y)-(u,v) := (x-u, y-v)$. Similar definitions hold for $\mathbb{R}^n$ with $n > 2$.] Let $T(x,y)$ be the left hand side of (19.2); that is, define

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (x,y) - S(x,y) + (7,11)\,.$$

With this notation (19.2) becomes

$$T(x,y) = (x,y)\,.$$

Thus, and this is the crucial point, to solve (19.1) we need only find a fixed point of the mapping T. If T is contractive, then the preceding theorem guarantees that T has a unique fixed point and therefore that the system of equations (19.1) has a unique solution.

Unfortunately, as things stand, T is *not* contractive with respect to the product metric on $\mathbb{R}^2$. (It is for convenience that we use the product metric d_1 on $\mathbb{R}^2$ rather than the usual Euclidean metric. Square roots are a nuisance.) To see that T is not contractive notice that $d_1\big((1,0),(0,0)\big) = 1$ whereas $d_1\big(T(1,0), T(0,0)\big) = d_1\big((-1,8),(7,11)\big) = 11$. All is not lost however. One simple-minded remedy is to divide everything in (19.1) by a large constant c. A little experimentation shows that $c = 10$ works. Instead of working with the system (19.1) of equations, consider the system

$$\begin{cases} 0.9x - 0.2y = 0.7 \\ 0.3x + 0.8y = 1.1 \end{cases} \tag{19.3}$$

which obviously has the same solutions as (19.1). Redefine S and T in the obvious fashion. Let

$$S\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (0.9x - 0.2y, 0.3x + 0.8y)$$

and

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (x,y) - S(x,y) + (0.7, 1.1)\,.$$

Thus redefined, T is contractive with respect to the product metric. *Proof:* Since

$$T(x,y) = (0.1x + 0.2y + 0.7, -0.3x + 0.2y + 1.1) \tag{19.4}$$

for all $(x,y) \in \mathbb{R}^2$, we see that

$$\begin{aligned} d_1\big(T(x,y), T(u,v)\big) &= 10^{-1}\big(|(x+2y) - (u+2v)| + |(-3x+2y) - (-3u+2v)|\big) \\ &\le 10^{-1}\big(|x-u| + 2|y-v| + 3|x-u| + 2|y-v|\big) \\ &= 0.4\,(|x-u| + |y-v|) \\ &= 0.4\, d_1\big((x,y),(u,v)\big) \end{aligned} \tag{19.5}$$

for all points (x,y) and (u,v) in $\mathbb{R}^2$.

Now since T is contractive and $\mathbb{R}^2$ is complete (with respect to the product metric—see 18.2.9), the *contractive mapping theorem* (19.1.5) tells us that T has a unique fixed point. But a fixed point of T is a solution for the system (19.3) and consequently for (19.1).

The construction used in the proof of 19.1.5 allows us to approximate the fixed point of T to any desired degree of accuracy. As in that proof choose x_0 to be any point whatever in $\mathbb{R}^2$. Then the points $x_0, x_1, x_2, \dots$ (where $x_n = T\big(x_{n-1}\big)$ for each n) converge to the fixed point of T. This is a technique of SUCCESSIVE APPROXIMATION.

For the present example let $x_0 = (0,0)$. (The origin is chosen just for convenience.) Now use (19.4) and compute.

$$\begin{aligned} x_0 &= (0,0) \\ x_1 &= T(0,0) = (0.7, 1.1) \\ x_2 &= T(x_1) = (1.021, 1.025) \\ x_3 &= T(x_2) = (1.0071, 0.9987) \\ &\ \vdots \end{aligned}$$

It is reasonable to conjecture that the system (19.1) has a solution consisting of rational numbers and then to guess that the points $x_0, x_1, x_2, \dots$ as computed above are converging to the point $(1,1)$ in $\mathbb{R}^2$. Putting $x = 1$ and $y = 1$ in (19.4), we see that the point $(1,1)$ is indeed the fixed point of T and therefore the solution to (19.1).

In the preceding example we discovered an exact solution to a system of equations. In general, of course, we cannot hope that a successive approximation technique will yield exact answers. In those cases in which it does not, it is most important to have some idea how accurate our approximations are. After n iterations, how close to the true solution are we? How many iterations must be computed in order to achieve a desired degree of accuracy? The answer to these questions in an easy consequence of the proof of theorem 19.1.5.

19.1.7. Corollary. *Let the space M, the mapping f, the sequence (x_n), the constant c, and the point p be as in theorem 19.1.5 and its proof. Then for every $m \geq 0$*

$$d(x_m, p) \leq d(x_0, x_1) \frac{c^m}{(1-c)}\,.$$

PROOF. Inequality (Q.11) in the proof of 19.1.5 says that for $m < n$

$$d(x_m, x_n) \leq d(x_0, x_1) c^m (1-c)^{-1}\,.$$

Take limits as $n \to \infty$. □

19.1.8. Definition. Notation as in the preceding corollary. If we think of the point x_n as being the n^{th} approximation to p, then the distance $d(x_n, p)$ between x_n and p is the ERROR associated with the n^{th} approximation.

Notice that because the product metric d_1 was chosen for $\mathbb{R}^2$ in example 19.1.6, the word "error" there means the *sum* of the errors in x and y. Had we wished for "error" to mean the *maximum* of the errors in x and y, we would have used the uniform metric d_u on $\mathbb{R}^2$. Similarly, if *root-mean-square* "error" were desired (that is, the square root of the sum of the squares of the errors in x and y), then we would have used the usual Euclidean metric on $\mathbb{R}^2$.

19.1.9. Exercise. Let (x_n) be the sequence of points in $\mathbb{R}^2$ considered in example 19.1.6. We showed that (x_n) converges to the point $p = (1,1)$.

(a) Use corollary 19.1.7 to find an upper bound for the error associated with the approximation x_4.
(b) What is the actual error associated with x_4?
(c) According to 19.1.7 how many terms of the sequence (x_n) should we compute to be sure of obtaining an approximation which is correct to within 10^{-4}?

(Solution Q.19.4.)

19.1.10. Problem. Show by example that the conclusion of the *contractive mapping theorem* fails if:

(a) the contraction constant is allowed to have the value 1; or
(b) the space is not complete.

19.1.11. Problem. Show that the map

$$g\colon [0,\infty) \to [0,\infty)\colon x \mapsto \frac{1}{x+1}$$

is not a contraction even though

$$d\big(g(x), g(y)\big) < d(x,y)$$

for all x, $y \geq 0$ with $x \neq y$.

19.1.12. Problem. Let $f(x) = (x/2) + (1/x)$ for $x \geq 1$.

(a) Show that f maps the interval $[1,\infty)$ into itself.
(b) Show that f is contractive.
(c) Let $x_0 = 1$ and for $n \geq 0$ let

$$x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}\,.$$

Show that the sequence (x_n) converges.
(d) Find $\lim_{n\to\infty} x_n$.
(e) Show that the distance between x_n and the limit found in (d) is no greater than 2^{-n}.

19.1.13. Problem. Solve the system of equations

$$\begin{aligned} 9x \quad -y \;\; +2z &= 37 \\ x + 10y \;\; -3z &= -69 \\ -2x \;\; +3y + 11z &= 58 \end{aligned}$$

following the procedure of example 19.1.6. *Hint.* As in 19.1.6 divide by 10 to obtain a contractive mapping. Before guessing at a rational solution, compute 10 or 11 successive approximations. Since this involves a lot of arithmetic, it will be helpful to have some computational assistance—a programmable calculator, for example.

19.1.14. Problem. Consider the following system of equations.

$$\begin{aligned} 75x + 16y - 20z &= 40 \\ 33x + 80y + 30z &= -48 \\ -27x + 32y + 80z &= 36 \end{aligned}$$

(a) Solve the system following the method of example 19.1.6. *Hint.* Because the contraction constant is close to 1, the approximations converge slowly. It may take 20 or 30 iterations before it is clear what the exact (rational) solutions should be. So as in the preceding example, it will be desirable to use computational assistance.
(b) Let (x_n) be the sequence of approximations in $\mathbb{R}^3$ converging to the solution of the system in (a). Use corollary 19.1.7 to find an upper bound for the error associated with the approximation x_{25}.
(c) To 4 decimal places, what is the actual error associated with x_{25}?
(d) According to 19.1.7, how many terms must be computed to be sure that the error in our approximation is no greater than 10^{-3}?

19.1.15. Problem. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the rotation of the plane about the point $(0,1)$ through an angle of π radians. Let $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the map which takes the point (x,y) to the midpoint of the line segment connecting (x,y) and $(1,0)$.

(a) Prove that $g \circ f$ is a contraction on $\mathbb{R}^2$ (with its usual metric).
(b) Find the unique fixed point of $g \circ f$.

(c) Let $(x_0, y_0) = (0, 1)$ and define (as in the proof of 19.1.5)
$$(x_{n+1}, y_{n+1}) = (g \circ f)(x_n, y_n)$$
for all $n \geq 0$. For each n compute the exact Euclidean distance between (x_n, y_n) and the fixed point of $g \circ f$.

19.2. APPLICATION TO INTEGRAL EQUATIONS

19.2.1. Exercise. Use theorem 19.1.5 to solve the integral equation
$$f(x) = \tfrac{1}{3}x^3 + \int_0^x t^2 f(t)\, dt\,. \tag{19.6}$$

Hint. We wish to find a continuous function f which satisfies (19.6) for all $x \in \mathbb{R}$. Consider the mapping T which takes each continuous function f into the function Tf whose value at x is given by
$$Tf(x) = \tfrac{1}{3}x^3 + \int_0^x t^2 f(t)\, dt\,.$$
[It is important to keep in mind that T acts on functions, not on numbers. Thus $Tf(x)$ is to be interpreted as $(T(f))(x)$ and not $T(f(x))$.] In order to make use of theorem 19.1.5, the mapping T must be contractive. One way to achieve this is to restrict our attention to continuous functions on the interval $[0, 1]$ and use the uniform metric on $\mathcal{C}([0, 1], \mathbb{R})$. Once a continuous function f is found such that (19.6) is satisfied for all x in $[0, 1]$, it is a simple matter to check whether (19.6) holds for all x in $\mathbb{R}$. Consider then the map
$$T \colon \mathcal{C}([0, 1], \mathbb{R}) \to \mathcal{C}([0, 1], \mathbb{R}) \colon f \mapsto Tf$$
where
$$Tf(x) := \tfrac{1}{3}x^3 + \int_0^x t^2 f(t)\, dt$$
for all $x \in [0, 1]$. The space $\mathcal{C}([0, 1], \mathbb{R})$ is complete. (Why?) Show that T is contractive by estimating $|Tf(x) - Tg(x)|$, where f and g are continuous functions on $[0, 1]$, and taking the supremum over all x in $[0, 1]$. What can be concluded from theorem 19.1.5 about (19.6)?

To actually find a solution to (19.6), use the *proof* of 19.1.5 (that is, successive approximations). For simplicity start with the zero function in $\mathcal{C}([0, 1], \mathbb{R})$: let $g_0(x) = 0$ for $0 \leq x \leq 1$. For $n \geq 0$ let $g_{n+1}(x) = Tg_n(x)$ for $0 \leq x \leq 1$. Compute g_1, g_2, g_3, and g_4. You should be able to guess what g_n will be. (It is easy to verify the correctness of your guess by induction, but it is not necessary to do this.) Next, let f be the function which is the uniform limit of the sequence (g_n). That is, f is the function whose power series expansion has g_n as its n^{th} partial sum. This power series expansion should be one with which you are familiar from beginning calculus; what elementary function does it represent?

Finally, show by direct computation that this elementary function does in fact satisfy (19.6) for all x in $\mathbb{R}$. (Solution Q.19.5.)

19.2.2. Problem. Give a careful proof that there exists a unique continuous real valued function f on $[0, 1]$ which satisfies the integral equation
$$f(x) = x^2 + \int_0^x t^2 f(t)\, dt\,.$$
(You are not asked to *find* the solution.)

19.2.3. Problem. Use theorem 19.1.5 to solve the integral equation
$$f(x) = x + \int_0^x f(t)\, dt\,.$$

Hint. Follow the procedure of exercise 19.2.1. Keep in mind that the only reason for choosing the particular interval $[0, 1]$ in 19.2.1 was to make the map T contractive.

19.2.4. Problem. For every f in $\mathcal{C}([0,\pi/4],\mathbb{R})$ define

$$Tf(x) = x^2 - 2 - \int_0^x f(t)\,dt$$

where $0 \leq x \leq \pi/4$. Show that T is a contraction. Find the fixed point of T. What integral equation have you solved?

CHAPTER 20

VECTOR SPACES

Most introductory calculus texts, for pedagogical reasons, do calculus twice, once for a single variable and then for either two or three variables, leaving the general finite dimensional and infinite dimensional cases for future courses. It is our goal eventually (in chapter 25) to develop differential calculus in a manner that is valid for any number of variables (even infinitely many).

A certain amount of algebra always underlies analysis. Before one studies the calculus of a single variable, a knowledge of arithmetic in $\mathbb{R}$ is required. For the calculus of a finite number of variables it is necessary to know something about $\mathbb{R}^n$. In this chapter and the next we lay the algebraic foundations for the differential calculus of an arbitrary number of variables; we study vector spaces and the operation preserving maps between them, called linear transformations.

20.1. DEFINITIONS AND EXAMPLES

20.1.1. Definition. A (REAL) VECTOR SPACE is a set V together with a binary operation $(x, y) \mapsto x + y$ (called ADDITION) from $V \times V$ into V and a mapping $(\alpha, x) \mapsto \alpha x$ (called SCALAR MULTIPLICATION) from $\mathbb{R} \times V$ into V satisfying the following conditions:

(1) Addition is associative. That is,
$$x + (y + z) = (x + y) + z \qquad \text{for all } x,\ y,\ z \in V\,.$$
(2) In V there is an element $\mathbf{0}$ (called the ZERO VECTOR) such that
$$x + \mathbf{0} = x \qquad \text{for all } x \in V\,.$$
(3) For each x in V there is a corresponding element $-x$ (the ADDITIVE INVERSE of x) such that
$$x + (-x) = \mathbf{0}\,.$$
(4) Addition is commutative. That is,
$$x + y = y + x \qquad \text{for all } x,\ y \in V\,.$$
(5) If $\alpha \in \mathbb{R}$ and $x,\ y \in V$, then
$$\alpha(x + y) = (\alpha x) + (\alpha y)\,.$$
(6) If $\alpha,\ \beta \in \mathbb{R}$ and $x \in V$, then
$$(\alpha + \beta)x = (\alpha x) + (\beta x)\,.$$
(7) If $\alpha,\ \beta \in \mathbb{R}$ and $x \in V$, then
$$\alpha(\beta x) = (\alpha\beta)x\,.$$
(8) If $x \in V$, then
$$1 \cdot x = x\,.$$

An element of V is a VECTOR; an element of $\mathbb{R}$ is, in this context, often called a SCALAR. Concerning the order of performing operations, we agree that scalar multiplication takes precedence over addition. Thus, for example, condition (5) above may be unambiguously written as
$$\alpha(x + y) = \alpha x + \alpha y\,.$$
(Notice that the parentheses on the left may not be omitted.)

If x and y are vectors, we define $x - y$ to be $x + (-y)$. If A and B are subsets of a vector space, we define

$$A + B := \{a + b \colon a \in A, b \in B\};$$

and if $\alpha \in \mathbb{R}$,

$$\alpha A := \{\alpha a \colon a \in A\}.$$

Condition (3) above is somewhat optimistic. No uniqueness is asserted in (2) for the zero vector $\mathbf{0}$, so one may well wonder whether (3) is supposed to hold for *some* zero vector $\mathbf{0}$ or for *all* such vectors. Fortunately, the problem evaporates since we can easily show that the zero vector is in fact unique.

20.1.2. Exercise. A vector space has exactly one zero vector. That is, if $\mathbf{0}$ and $\mathbf{0}'$ are members of a vector space V which satisfy $x + \mathbf{0} = x$ and $x + \mathbf{0}' = x$ for all $x \in V$, then $\mathbf{0} = \mathbf{0}'$. (Solution Q.20.1.)

In a vector space not only is the zero vector unique but so are additive inverses.

20.1.3. Problem. For every vector x in a vector space V there exists only one vector $-x$ such that

$$x + (-x) = \mathbf{0}.$$

In 20.1.4 to 20.1.7 we state four useful, if elementary, facts concerning the arithmetic of vectors.

20.1.4. Exercise. If x is a vector (in some vector space) and $x + x = x$, then $x = \mathbf{0}$. *Hint.* Add $\mathbf{0}$ to x; then write $\mathbf{0}$ as $x + (-x)$. (Solution Q.20.2.)

20.1.5. Exercise. Let x be a vector (in a some vector space) and let α be a real number. Then $\alpha x = \mathbf{0}$ if and only if $x = \mathbf{0}$ or $\alpha = 0$. *Hint.* Show three things:

(a) $\alpha\mathbf{0} = \mathbf{0}$,
(b) $0\,x = \mathbf{0}$, and
(c) If $\alpha \neq 0$ and $\alpha x = \mathbf{0}$, then $x = \mathbf{0}$.

(If it is not clear to you that proving (a), (b), and (c) is the same thing as proving 20.1.5, see the remark following this hint.) To prove (a) write $\mathbf{0} + \mathbf{0} = \mathbf{0}$, multiply by α, and use 20.1.4. For (c) use the fact that if $\alpha \in \mathbb{R}$ is not zero, it has a reciprocal. What happens if we multiply the vector αx by the scalar $1/\alpha$? (Solution Q.20.3.)

Remark. It should be clear that proving (a) and (b) of the preceding hint proves that:

$$\text{if } x = \mathbf{0} \text{ or } \alpha = 0, \text{ then } \alpha x = \mathbf{0}.$$

What may not be clear is that proving (c) is enough to establish:

$$\text{if } \alpha x = \mathbf{0}, \text{ then either } x = \mathbf{0} \text{ or } \alpha = 0. \tag{20.1}$$

Some students feel that in addition to proving (c) it is also necessary to prove that:

$$\text{if } x \neq \mathbf{0} \text{ and } \alpha x = 0, \text{ then } \alpha = 0.$$

To see that this is unnecessary recognize that there are just two possible cases: either α is equal to zero, or it is not. In case α *is* equal to zero, then the conclusion of (20.1) is certainly true. The other case, where α is *not* zero, is dealt with by (c).

20.1.6. Exercise. If x is a vector, then $-(-x) = x$. *Hint.* Show that $(-x) + x = \mathbf{0}$. What does 20.1.3 say about x? (Solution Q.20.4.)

20.1.7. Problem. If x is a vector, then $(-1)x = -x$. *Hint.* Show that $x + (-1)x = \mathbf{0}$. Use 20.1.3.

20.1.8. Problem. Using nothing about vector spaces other than the definitions, prove that if x is a vector, then $3x - x = x + x$. Write your proof using at most one vector space axiom (or definition) at each step.

We now give some examples of vector spaces.

20.1.9. Example. Let $V = \mathbb{R}^n$. We make a standard notational convention. If x belongs to $\mathbb{R}^n$, then x is an n-tuple whose coordinates are $x_1, x_2, \dots, x_n$; that is,

$$x = (x_1, x_2, \dots, x_n) .$$

It must be confessed that we do not *always* use this convention. For example, the temptation to denote a member of $\mathbb{R}^3$ by (x, y, z), rather than by (x_1, x_2, x_3), is often just too strong to resist. For n-tuples $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ in V define

$$x + y := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) .$$

Accordingly we say that addition in $\mathbb{R}^n$ is defined coordinatewise. Scalar multiplication is also defined in a coordinatewise fashion. That is, if $x = (x_1, x_2, \dots, x_n) \in V$ and $\alpha \in \mathbb{R}$, then we define

$$\alpha x := (\alpha x_1, \alpha x_2, \dots, \alpha x_n) .$$

Under these operations $\mathbb{R}^n$ becomes a vector space.

PROOF. Problem. *Hint.* Just verify conditions (1)–(8) of definition 20.1.1.

20.1.10. Problem. Let s be the family of all sequences of real numbers. Explain how to define addition and scalar multiplication on s in such a way that it becomes a vector space.

20.1.11. Example. Here is one of the ways in which we construct new vector spaces from old ones. Let V be an arbitrary vector space and S be a nonempty set. Let $\mathcal{F}(S, V)$ be the family of all V valued functions defined on S. That is, $\mathcal{F}(S, V)$ is the set of all functions f such that $f \colon S \to V$. We make $\mathcal{F}(S, V)$ into a vector space by defining operations in a pointwise fashion. For functions f, $g \in \mathcal{F}(S, V)$ define

$$(f + g)(x) := f(x) + g(x) \qquad \text{for all } x \in S.$$

It should be clear that the two "+" signs in the preceding equation denote operations in different spaces. The one on the left (which is being defined) represents addition in the space $\mathcal{F}(S, V)$; the one on the right is addition in V. Because we specify the value of $f + g$ at each point x by adding the values of f and g at that point, we say that we add f and g POINTWISE.

We also define scalar multiplication to be a pointwise operation. That is, if $f \in \mathcal{F}(S, V)$ and $\alpha \in \mathbb{R}$, then we define the function αf by

$$(\alpha f)(x) := \alpha(f(x)) \qquad \text{for every } x \in S.$$

Notice that according to the definitions above, both $f + g$ and αf belong to $\mathcal{F}(S, V)$. Under these pointwise operations $\mathcal{F}(S, V)$ is a vector space. (Notice that the family of real valued functions on a set S is a special case of the preceding. Just let $V = \mathbb{R}$.)

PROOF. Problem

Most of the vector spaces we encounter in the sequel are subspaces of $\mathcal{F}(S, V)$ for some appropriate set S and vector space V: so we now take up the topic of subspaces of vector spaces.

20.1.12. Definition. A subset W of a vector space V is a SUBSPACE of V if it is itself a vector space under the operations it inherits from V. In subsequent chapters we will regularly encounter objects which are simultaneously vector spaces and metric spaces. (One obvious example is $\mathbb{R}^n$). We will often use the term VECTOR SUBSPACE (or LINEAR SUBSPACE) to distinguish a subspace of a vector space from a metric subspace. (Example: the unit circle $\{(x, y) \colon x^2 + y^2 = 1\}$ is a metric subspace of $\mathbb{R}^2$ but not a vector subspace thereof.)

Given a subset W of a vector space V we do not actually need to check all eight vector space axioms to establish that it is a subspace of V. We need only know that W is nonempty and that it is closed under addition and scalar multiplication.

20.1.13. Proposition. *Let W be a subset of a vector space V. Then W is a subspace of V provided that*

(a) *$W \neq \emptyset$,*
(b) *$x + y \in W$ whenever $x \in W$ and $y \in W$, and*
(c) *$\alpha x \in W$ whenever $x \in W$ and $\alpha \in \mathbb{R}$.*

Proof. Exercise. (Solution Q.20.5.)

20.1.14. Example. Let S be a nonempty set. Then the family $\mathcal{B}(S, \mathbb{R})$ of all bounded real valued functions on S is a vector space because it is a subspace of $\mathcal{F}(S, \mathbb{R})$.

Proof. That it *is* a subspace of $\mathcal{F}(S, \mathbb{R})$ is clear from the preceding proposition: every constant function is bounded, so the set $\mathcal{B}(S, \mathbb{R})$ is nonempty; that it is closed under addition and scalar multiplication was proved in proposition 13.1.2. □

20.1.15. Example. The x-axis (that is, $\{(x, 0, 0) \colon x \in \mathbb{R}\}$) is a subspace of $\mathbb{R}^3$. So is the xy-plane (that is, $\{(x, y, 0) \colon x, y \in \mathbb{R}\}$). In both cases it is clear that the set in question is nonempty and is closed under addition and scalar multiplication.

20.1.16. Example. Let M be a metric space. Then the set $\mathcal{C}(M, \mathbb{R})$ of all continuous real valued functions on M is a vector space.

Proof. Problem.

20.1.17. Problem. Let $a < b$ and $\mathcal{F}$ be the vector space of all real valued functions on the interval $[a, b]$. Consider the following subsets of $\mathcal{F}$:

$$\begin{aligned}
\mathcal{K} &= \{f \in \mathcal{F} \colon f \text{ is constant}\} \\
\mathcal{D} &= \{f \in \mathcal{F} \colon f \text{ is differentiable}\} \\
\mathcal{B} &= \{f \in \mathcal{F} \colon f \text{ is bounded}\} \\
\mathcal{P}_3 &= \{f \in \mathcal{F} \colon f \text{ is a polynomial of degree } 3\} \\
\mathcal{Q}_3 &= \{f \in \mathcal{F} \colon f \text{ is a polynomial of degree less than or equal to } 3\} \\
\mathcal{P} &= \{f \in \mathcal{F} \colon f \text{ is a polynomial}\} \\
\mathcal{C} &= \{f \in \mathcal{F} \colon f \text{ is continuous}\}
\end{aligned}$$

Which of these are subspaces of which? *Hint.* There is a ringer in the list.

20.1.18. Example. The family of all solutions of the differential equation

$$xy'' + y' + xy = 0$$

is a subspace of $\mathcal{C}(\mathbb{R}, \mathbb{R})$.

Proof. Problem.

Let A be a subset of a vector space V. Question: What is meant by the phrase "the smallest subspace of V which contains A"? Answer: The intersection of all the subspaces of V which contain A. It is important to realize that in order for this answer to make sense, it must be known that the intersection of the family of subspaces containing A is itself a subspace of V. This is an obvious consequence of the fact (proved below) that the intersection of any family of subspaces is itself a subspace.

20.1.19. Proposition. *Let $\mathfrak{S}$ be a nonempty family of subspaces of a vector space V. Then $\bigcap \mathfrak{S}$ is a subspace of V.*

Proof. Exercise. *Hint.* Use 20.1.13. (Solution Q.20.6.)

20.1.20. Example. Let V and W be vector spaces. If addition and scalar multiplication are defined on $V \times W$ by

$$(v, w) + (x, y) := (v + x, w + y)$$

and

$$\alpha(v, w) := (\alpha v, \alpha w)$$

for all v, $x \in V$, all w, $y \in W$, and all $\alpha \in \mathbb{R}$, then $V \times W$ becomes a vector space. (This is called the PRODUCT or (EXTERNAL) DIRECT SUM of V and W. It is frequently denoted by $V \oplus W$.)

PROOF. Problem.

20.2. LINEAR COMBINATIONS

20.2.1. Definition. Let V be a vector space. A LINEAR COMBINATION of a finite set $\{x_1, \dots, x_n\}$ of vectors in V is a vector of the form $\sum_{k=1}^{n} \alpha_k x_k$ where $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. If $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, then the linear combination is TRIVIAL; if at least one α_k is different from zero, the linear combination is NONTRIVIAL.

20.2.2. Exercise. Find a nontrivial linear combination of the following vectors in $\mathbb{R}^3$ which equals zero: $(1, 0, 0)$, $(1, 0, 1)$, $(1, 1, 1)$, and $(1, 1, 0)$. (Solution Q.20.7.)

20.2.3. Problem. Find, if possible, a nontrivial linear combination of the following vectors in $\mathbb{R}^3$ which equals zero: $(4, 1, 3)$, $(-1, 1, -7)$, and $(1, 2, -8)$.

20.2.4. Problem. Find, if possible, a nontrivial linear combination of the following vectors in $\mathbb{R}^3$ which equals zero: $(1, 2, -3)$, $(1, -1, 4)$, and $(5, 4, -1)$.

20.2.5. Problem. Find a nontrivial linear combination of the polynomials p^1, p^2, p^3, and p^4 which equal zero, where

$$\begin{aligned}
p^1(x) &= x + 1 \\
p^2(x) &= x^3 - 1 \\
p^3(x) &= 3x^3 + 2x - 1 \\
p^4(x) &= -x^3 + x\,.
\end{aligned}$$

20.2.6. Example. Define vectors $e^1, \dots, e^n$ in $\mathbb{R}^n$ by

$$\begin{aligned}
e^1 &:= (1, 0, 0, \dots, 0) \\
e^2 &:= (0, 1, 0, \dots, 0) \\
&\ \vdots \\
e^n &:= (0, 0, \dots, 0, 1).
\end{aligned}$$

In other words, for $1 \le j \le n$ and $1 \le k \le n$, the k^{th} coordinate of the vector e^j (denote it by $(e^j)_k$ or e^j_k) is 1 if $j = k$ and 0 if $j \ne k$. The vectors $e^1, \dots, e^n$ are the STANDARD BASIS VECTORS in $\mathbb{R}^n$. (Note that the superscripts here have nothing to do with powers.) In $\mathbb{R}^3$ the three standard basis vectors are often denoted by $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ rather than e^1, e^2, and e^3, respectively.

Every vector in $\mathbb{R}^n$ is a linear combination of the standard basis vectors in that space. In fact, if $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, then

$$x = \sum_{k=1}^{n} x_k e^k\,.$$

PROOF. The proof is quite easy:

$$\begin{aligned}
x &= (x_1, x_2, \dots, x_n) \\
&= (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n) \\
&= x_1 e^1 + x_2 e^2 + \dots + x_n e^n \\
&= \sum_{k=1}^{n} x_k e^k\,.
\end{aligned}$$

□

20.2.7. Definition. A subset A (finite or not) of a vector space is LINEARLY DEPENDENT if the zero vector $\mathbf{0}$ can be written as a nontrivial linear combination of elements of A; that is, if there exist vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in A$ and scalars $\alpha_1, \dots, \alpha_n$, **not all zero,** such that $\sum_{k=1}^{n} \alpha_k \mathbf{x}_k = \mathbf{0}$. A subset of a vector space is LINEARLY INDEPENDENT if it is not linearly dependent.

20.2.8. Definition. A set A of vectors in a vector space V SPANS the space if every member of V can be written as a linear combination of members of A.

20.2.9. Problem. Let $e^1, e^2, \dots e^n$ be the standard basis vectors in $\mathbb{R}^n$ (see example 20.2.6).

(a) Show that the set of standard basis vectors in $\mathbb{R}^n$ is a linearly independent set.
(b) Show that the standard basis vectors span $\mathbb{R}^n$.
(c) Show that in part (b) the representation of a vector in $\mathbb{R}^n$ as a linear combination of standard basis vectors is unique. (That is, show that if $x = \sum_{k=1}^{n} \alpha_k e^k = \sum_{k=1}^{n} \beta_k e^k$, then $\alpha_k = \beta_k$ for each k.)

20.3. CONVEX COMBINATIONS

20.3.1. Definition. A linear combination $\sum_{k=1}^{n} \alpha_k x_k$ of the vectors $x_1, \dots, x_n$ is a CONVEX COMBINATION if $\alpha_k \geq 0$ for each k ($1 \leq k \leq n$) and if $\sum_{k=1}^{n} \alpha_k = 1$.

20.3.2. Exercise. Write the vector $(2, 1/4)$ in $\mathbb{R}^2$ as a convex combination of the vectors $(1, 0)$, $(0, 1)$, and $(3, 0)$. (Solution Q.20.8.)

20.3.3. Problem. Write the vector $(1, 1)$ as a convex combination of the vectors $(-2, 2)$, $(2, 2)$, and $(3, -3)$ in $\mathbb{R}^2$.

20.3.4. Definition. If x and y are vectors in the vector space V, then the CLOSED SEGMENT between x and y, denoted by $[x, y]$, is $\{(1-t)x + ty \colon 0 \leq t \leq 1\}$. (Note: In the vector space $\mathbb{R}$ this is the same as the closed interval $[x, y]$ provided that $x \leq y$. If $x > y$, however, the closed segment $[x, y]$ contains all numbers z such that $y \leq z \leq x$, whereas the closed interval $[x, y]$ is empty.)

A set $C \subseteq V$ is CONVEX if the closed segment $[x, y]$ is contained in C whenever $x, y \in C$.

20.3.5. Example. A disk is a convex subset of $\mathbb{R}^2$. The set $\{(x, y) \colon 1 \leq x^2 + y^2 \leq 2\}$ is not a convex subset of $\mathbb{R}^2$.

PROOF. Problem.

20.3.6. Example. Every subspace of a vector space is convex.

PROOF. Problem.

20.3.7. Example. The set

$$\{(x, y) \in \mathbb{R}^2 \colon x \geq 0,\ y \geq 0, \text{ and } x + y \leq 1\}$$

is a convex subset of $\mathbb{R}^2$.

PROOF. Problem.

20.3.8. Example. Every convex subset of $\mathbb{R}^n$ is connected.

PROOF. Problem.

20.3.9. Definition. Let A be a subset of a vector space V. The CONVEX HULL of A is the smallest convex set containing A; that is, it is the intersection of the family of all convex subsets of V which contain A.

20.3.10. Exercise. What fact must we know about convex sets in order for the preceding definition to make sense? Prove this fact. *Hint.* Review proposition 20.1.19 and the discussion which precedes it. (Solution Q.20.9.)

20.3.11. Problem. Consider the following subsets of $\mathbb{R}^2$:

$$\begin{aligned} A &= \{(x, y)\colon x \geq 0\} \\ B &= \{(x, y)\colon 0 \leq y \leq 2\} \\ C &= \{(x, y)\colon x + y \leq 4\} \\ D &= A \cap B \cap C. \end{aligned}$$

The set D can be described as the convex hull of four points. Which four?

20.3.12. Problem. The concepts of convex combination and convex hull have been introduced. The point of this problem is to explain the way in which these two ideas are related. Start with a set A. Let C be the set of all convex combinations of elements of A. Let H be the convex hull of A. What relation can you find between A and H. It might not be a bad idea to experiment a little; see what happens in some very simple concrete cases. For example, take A to be a pair of points in $\mathbb{R}^3$. Eventually you will have to consider the following question: Are convex sets closed under the taking of convex combinations?

CHAPTER 21

LINEARITY

21.1. LINEAR TRANSFORMATIONS

Linear transformations are central to our study of calculus. Functions are differentiable, for example, if they are smooth enough to admit decent approximation by (translates of) linear transformations. Thus before tackling differentiation (in chapter 25) we familiarize ourselves with some elementary facts about linearity.

21.1.1. Definition. A function $T\colon V \to W$ between vector spaces is LINEAR if

$$T(x+y) = Tx + Ty \qquad \text{for all } x, y \in V \tag{21.1}$$

and

$$T(\alpha x) = \alpha Tx \qquad \text{for all } x \in V \text{ and } \alpha \in \mathbb{R}. \tag{21.2}$$

A linear function is most commonly called a LINEAR TRANSFORMATION, sometimes a LINEAR MAPPING. If the domain and codomain of a linear transformation are the same vector space, then it is often called a LINEAR OPERATOR, and occasionally a VECTOR SPACE ENDOMORPHISM. The family of all linear transformations from V into W is denoted by $\mathfrak{L}(V,W)$. Two oddities of notation concerning linear transformations deserve comment. First, the value of T at x is usually written Tx rather than $T(x)$. Naturally the parentheses are used whenever their omission would create ambiguity. For example, in (21.1) above $Tx+y$ is not an acceptable substitute for $T(x+y)$. Second, the symbol for composition of two linear transformations is ordinarily omitted. If $S \in \mathfrak{L}(U,V)$ and $T \in \mathfrak{L}(V,W)$, then the composite of T and S is denoted by TS (rather than by $T \circ S$). This will cause no confusion since we will define no other "multiplication" of linear maps. As a consequence of this convention if T is a linear operator, then $T \circ T$ is written as T^2, $T \circ T \circ T$ as T^3, and so on. One may think of condition (21.1) in the definition of linearity in the following fashion. Let $T \times T$ be the mapping from $V \times V$ into $W \times W$ defined by

$$(T \times T)(x,y) = (Tx, Ty).$$

Then condition (21.1) holds if and only if the diagram

$$\begin{tikzcd}
V \times V \arrow[r, "T\times T"] \arrow[d, "+"'] & W \times W \arrow[d, "+"] \\
V \arrow[r, "T"'] & W
\end{tikzcd}$$

commutes. (The vertical maps are addition in V and in W.)

Condition (21.2) of the definition can similarly be thought of in terms of a diagram. For each scalar a define the function M_α, MULTIPLICATION BY α, from a vector space into itself by

$$M_\alpha(x) = \alpha x\,.$$

(We use the same symbol for multiplication by α in both of the spaces V and W.) Then condition (21.2) holds if and only if for every scalar α the following diagram commutes.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "M_\alpha"'] & W \arrow[d, "M_\alpha"] \\ V \arrow[r, "T"'] & W \end{tikzcd}$$

21.1.2. Example. If $T\colon \mathbb{R}^3 \to \mathbb{R}^2\colon x \mapsto (x_1 + x_3, x_1 - 2x_2)$, then T is linear.

PROOF. Exercise. (Solution Q.21.1.)

21.1.3. Problem. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^4$ be defined by

$$T(x, y) = (x + 2y, 3x - y, -2x, -x + y).$$

Show that T is linear.

21.1.4. Exercise. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation which satisfies $T(e^1) = (1, 0, 1)$, $T(e^2) = (0, 2, -1)$, and $T(e^3) = (-4, -1, 3)$ (where e^1, e^2, and e^3 are the standard basis vectors for $\mathbb{R}^3$ defined in example 20.2.6). Find $T(2, 1, 5)$. *Hint.* Use problem 20.2.9. (Solution Q.21.2.)

21.1.5. Problem. Suppose that $T \in \mathfrak{L}(\mathbb{R}^3, \mathbb{R}^3)$ satisfies

$$\begin{aligned} Te^1 &= (1, 2, -3) \\ Te^2 &= (1, -1, 0) \\ Te^3 &= (-2, 0, 1) \end{aligned}$$

Find $T(3, -2, 1)$.

21.1.6. Proposition. *Let $T\colon V \to W$ be a linear transformation between two vector spaces. Then*

(a) $T(0) = 0$.
(b) $T(x - y) = Tx - Ty$ *for all* $x, y \in V$.

PROOF. Exercise. (Solution Q.21.3.)

21.1.7. Example. The identity map from a vector space into itself is linear.

PROOF. Obvious. □

21.1.8. Example. Each coordinate projection defined on $\mathbb{R}^n$

$$\pi_k\colon \mathbb{R}^n \to \mathbb{R}\colon x \mapsto x_k$$

is linear.

PROOF. For $1 \le k \le n$, we have $\pi_k(x+y) = \pi_k(x_1+y_1, \dots, x_n+y_n) = x_k + y_k = \pi_k(x) + \pi_k(y)$ and $\pi_k(\alpha x) = \pi_k(\alpha x_1, \dots, \alpha x_n) = \alpha x_k = \alpha \pi_k(x)$. □

21.1.9. Example. Let $\mathcal{F} = \mathcal{F}((a, b), \mathbb{R})$ be the family of all real valued functions defined on the open interval (a, b) and let $\mathcal{D} = \mathcal{D}((a, b), \mathbb{R})$ be the set of all members of $\mathcal{F}$ which are differentiable at each point of (a, b). Then $\mathcal{D}$ is a vector subspace of $\mathcal{F}$ and that the differentiation operator

$$D\colon \mathcal{D} \to \mathcal{F}\colon f \mapsto f'$$

(where f' is the derivative of f) is linear.

PROOF. We know from example 20.1.11 that $\mathcal{F}$ is a vector space. To show that $\mathcal{D}$ is a vector subspace of $\mathcal{F}$ use proposition 20.1.13. That $\mathcal{D}$ is nonempty is clear since constant functions are differentiable. That the space $\mathcal{D}$ of differentiable functions is closed under addition and scalar multiplication and that the operation D of differentiation is linear (that is, $D(\alpha f) = \alpha Df$ and $D(f + g) = Df + Dg$) are immediate consequences of propositions 8.4.10 and 8.4.11. □

21.1.10. Example. We have not yet discussed integration of continuous functions, but recalling a few basic facts about integration from beginning calculus shows us that this is another example of a linear transformation. Let $\mathcal{C} = \mathcal{C}([a,b],\mathbb{R})$ be the family of all members of $\mathcal{F} = \mathcal{F}([a,b],\mathbb{R})$ which are continuous. We know from beginning calculus that any continuous function on a closed and bounded interval is (Riemann) integrable, that

$$\int_a^b \alpha f(x)\,dx = \alpha \int_a^b f(x)\,dx$$

where $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}$, and that

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

where $f, g \in \mathcal{C}$. It then follows easily (again from 20.1.13) that $\mathcal{C}$ is a vector subspace of $\mathcal{F}$ and that the function $K\colon \mathcal{C} \to \mathbb{R}\colon f \mapsto \int_a^b f(x)\,dx$ is linear.

An important observation is that the composite of two linear transformations is linear.

21.1.11. Proposition. *Let U, V, and W be vector spaces. If $S \in \mathfrak{L}(U,V)$ and $T \in \mathfrak{L}(V,W)$ then $TS \in \mathfrak{L}(U,W)$.*

Proof. Problem.

21.1.12. Definition. If $T\colon V \to W$ is a linear transformation between vector spaces, then the KERNEL (or NULL SPACE) of T, denoted by $\ker T$, is $T^{\leftarrow}\{\mathbf{0}\}$. That is,

$$\ker T := \{x \in V : Tx = \mathbf{0}\}.$$

21.1.13. Exercise. Let $T \in \mathfrak{L}(\mathbb{R}^3,\mathbb{R}^3)$ satisfy

$$\begin{aligned} Te^1 &= (1,-2,3)\\ Te^2 &= (0,0,0)\\ Te^3 &= (-2,4,-6) \end{aligned}$$

where e^1, e^2, and e^3 are the standard basis vectors in $\mathbb{R}^3$. Find and describe geometrically both the kernel of T and the range of T. (Solution Q.21.4.)

21.1.14. Problem. Let T be the linear transformation of example 21.1.2. Find and describe geometrically the kernel and the range of T.

21.1.15. Problem. Let D be the linear transformation defined in 21.1.9. What is the kernel of D?

It is useful to know that the kernel of a linear transformation is always a vector subspace of its domain, that its range is a vector subspace of its codomain, and that a necessary and sufficient condition for a linear transformation to be injective is that its kernel contain only the zero vector.

21.1.16. Proposition. *If $T\colon V \to W$ is a linear transformation between vector spaces, then* $\ker T$ *is a subspace of V.*

Proof. Problem. *Hint.* Use proposition 20.1.13.

21.1.17. Proposition. *If $T\colon V \to W$ is a linear transformation between vector spaces, then* $\operatorname{ran} T$ *is a subspace of W.*

Proof. Exercise. *Hint.* Use proposition 20.1.13. (Solution Q.21.5.)

21.1.18. Proposition. *A linear transformation T is injective if and only if* $\ker\ T = \{0\}$.

Proof. Problem. *Hint.* First show that if T is injective and $x \in \ker T$, then $x = 0$. For the converse, suppose that $\ker T = \{\mathbf{0}\}$ and that $Tx = Ty$. Show that $x = y$.

21.1.19. Problem. Let W be a vector space.

(a) Show that a linear transformation $T: \mathbb{R}^n \to W$ is completely determined by its values on the standard basis vectors $e^1, \dots e^n$ of $\mathbb{R}^n$. *Hint.* Use problem 20.2.9(b).
(b) Show that if $S, T \in \mathfrak{L}(\mathbb{R}^n, W)$ and $Se^k = Te^k$ for $1 \le k \le n$, then $S = T$.
(c) Let $w^1, \dots, w^n \in W$. Show that there exists a unique $T \in \mathfrak{L}(\mathbb{R}^n)$ such that $Te^k = w^k$ for $1 \le k \le n$.

Injective linear mappings take linearly independent sets to linearly independent sets

21.1.20. Proposition. *If $T \in \mathfrak{L}(V, W)$ is injective and A is a linearly independent subset of V, then $T^{\to}(A)$ is a linearly independent set in W.*

Proof. Problem. *Hint.* Start with vectors $y^1, \dots y^n$ in $T^{\to}(A)$ and suppose that some linear combination of them $\sum_{k=1}^{n} \alpha_k x^k$ is 0. Show that all the scalars α_k are 0. Use proposition 21.1.18.

Linear mappings take convex sets to convex sets.

21.1.21. Proposition. *If V and W are vector spaces, C is a convex subset of V, and $T: V \to W$ is linear, then $T^{\to}(C)$ is a convex subset of W.*

Proof. Problem. *Hint.* Let u, $v \in T^{\to}(C)$ and $0 \le t \le 1$. Show that $(1-t)u + tv \in T^{\to}(C)$.

21.1.22. Problem. Let $T \in \mathfrak{L}(\mathbb{R}^3, \mathbb{R}^3)$ satisfy :

$$\begin{aligned} Te^1 &= (0, 1, 0) \\ Te^2 &= (0, 0, 1) \\ Te^3 &= (3, -2, 0) \end{aligned}$$

Show that T is bijective. *Hint.* To show that T is injective use proposition 21.1.18.

21.1.23. Problem. Let $\mathcal{C}^1 = \mathcal{C}^1([a, b], \mathbb{R})$ be the set of all functions f in $\mathcal{F} = \mathcal{F}([a, b], \mathbb{R})$ such that f' exists on $[a, b]$ in the usual sense of having one-sided derivatives at a and b) and is continuous. (A function belonging to $\mathcal{C}^1$ is said to be continuously differentiable.) It is easy to see that the set of all C^1 functions is a vector subspace of $\mathcal{F}$. Let $\mathcal{C} = \mathcal{C}([a, b], \mathbb{R})$ be the family of all continuous members of $\mathcal{F}$. For every $f \in \mathcal{C}$ and every $x \in [a, b]$ let

$$(Jf)(x) := \int_a^x f(t)\, dt\,.$$

(a) Why does Jf belong to $\mathcal{C}^1$?
(b) Show that the map $J: \mathcal{C} \to \mathcal{C}^1 : f \mapsto Jf$ is linear.

21.1.24. Problem. (Products) Recall (see Appendix N) that for every pair of functions $f^1 : T \to S_1$ and $f^2 : T \to S_2$ having the same domain there exists a unique map, namely $f = (f^1, f^2)$, mapping T into the product space $S^1 \times S^2$ which satisfies $\pi_1 \circ f = f^1$ and $\pi_2 \circ f = f^2$. (See in particular exercise N.1.4.) Now suppose that T, S_1, and S_2 are vector spaces and that f^1 and f^2 are linear. Then $S_1 \times S_2$ is a vector space (see example 20.1.20). Show that the function $f = (f^1, f^2)$ is linear.

21.2. THE ALGEBRA OF LINEAR TRANSFORMATIONS

The set $\mathfrak{L}(V, W)$ of linear transformations between two vector spaces is contained in the vector space $\mathcal{F}(V, W)$ of *all* W-valued functions whose domain is V. (That $\mathcal{F}$ is a vector space was proved in example 20.1.11.) It is easy to show that $\mathfrak{L}(V, W)$ is a vector space; just show that it is a subspace of $\mathcal{F}(V, W)$.

21.2.1. Proposition. *Let V and W be vector spaces. Then $\mathfrak{L}(V, W)$ with pointwise operations of addition and scalar multiplication is a vector space.*

Proof. Exercise. *Hint.* Use example 20.1.11 and proposition 20.1.13. (Solution Q.21.6.)

Let $T\colon V \to W$ be a function between two sets. We say that T is INVERTIBLE if there exists a function T^{-1} mapping W to V such that $T \circ T^{-1}$ is the identity function on V and $T^{-1} \circ T$ is the identity function on W. (For details about this see appendix M.) Now, suppose that V and W are vector spaces and that $T\colon V \to W$ is a linear transformation. In this context what do we mean when we say that T is *invertible*? For a linear transformation to be invertible we will require two things: the transformation must possess an inverse function, and this function must itself be linear. It is a pleasant fact about linear transformations that the second condition is automatically satisfied whenever the first is.

21.2.2. Proposition. *If $T \in \mathfrak{L}(V, W)$ is bijective, then its inverse $T^{-1}\colon W \to V$ is linear.*

PROOF. Exercise (Solution Q.21.7.)

21.2.3. Definition. A linear transformation $T\colon V \to W$ is INVERTIBLE (or is an ISOMORPHISM) if there exists a linear transformation T^{-1} such that $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$.

The point of the preceding proposition is that this definition is somewhat redundant. In particular, the following are just different ways of saying the same thing about a linear transformation T.

(a) T is invertible.
(b) T is an isomorphism.
(c) As a function T has an inverse.
(d) T is bijective.

21.2.4. Definition. Vector spaces V and W are ISOMORPHIC if there exists an isomorphism from V onto W.

21.2.5. Problem. Let s be the set of all sequences of real numbers. Regard s as a vector space under pointwise operations. That is,

$$x + y := (x_1 + y_1, x_2 + y_2, \dots)$$
$$\alpha x := (\alpha x_1, \alpha x_2, \dots)$$

whenever $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ belong to s and α is a scalar. Define the UNILATERAL SHIFT OPERATOR $U\colon s \to s$ by

$$U(x_1, x_2, x_3, \dots) := (0, x_1, x_2, \dots).$$

(a) Show that $U \in \mathfrak{L}(s, s)$.
(b) Does U have a right inverse? If so, what is it?
(c) Does U have a left inverse? If so, what is it?

21.2.6. Definition. Suppose that a vector space V is equipped with an additional operation $(x, y) \mapsto xy$ from $V \times V$ into V (we will call it *multiplication*) which satisfies

(a) $x(y + z) = xy + xz$,
(b) $(x + y)z = xz + yz$,
(c) $(xy)z = x(yz)$, and
(d) $\alpha(xy) = x(\alpha y)$

whenever $x, y, z \in V$ and $\alpha \in \mathbb{R}$. Then V is an ALGEBRA. (Sometimes it is called a LINEAR ASSOCIATIVE ALGEBRA.) If an algebra possesses a multiplicative identity (that is, a vector $\mathbf{1}$ such that $\mathbf{1}\,x = x\,\mathbf{1} = x$ for all $x \in V$), then it is a UNITAL ALGEBRA. A subset of an algebra A which is closed under the operations of addition, multiplication, and scalar multiplication is a SUBALGEBRA of A. If A is a unital algebra and B is a subalgebra of A which contains the multiplicative identity of A, then B is a UNITAL SUBALGEBRA of A.

21.2.7. Example. If M is a compact metric space, then the vector space $\mathcal{B}(M, \mathbb{R})$ of bounded functions on M is a unital algebra under pointwise operations. (The constant function 1 is its multiplicative identity.) We have already seen in chapter 15 that the space $\mathcal{C}(M, \mathbb{R})$ of continuous functions on M is a vector subspace of $\mathcal{B}(M,\mathbb{R})$. Since the product of continuous functions is continuous (and constant functions are continuous) $\mathcal{C}(M,\mathbb{R})$ is a unital subalgebra of $\mathcal{B}(M,\mathbb{R})$.

21.2.8. Problem. It has already been shown (in proposition 21.2.1) that if V is a vector space, then so is $\mathfrak{L}(V,V)$. Show that with the additional operation of composition serving as multiplication $\mathfrak{L}(V,V)$ is a unital algebra.

21.2.9. Problem. If $T \in \mathfrak{L}(V,W)$ is invertible, then so is T^{-1} and $\left(T^{-1})\right)^{-1} = T$.

21.2.10. Problem. If $S \in \mathfrak{L}(U,V)$ and $T \in \mathfrak{L}(V,W)$ are both invertible, then so is TS and $(TS)^{-1} = S^{-1}T^{-1}$.

21.2.11. Problem. If $T \in \mathfrak{L}(V,V)$ satisfies the equation

$$T^2 - T + I = \mathbf{0}$$

then it is invertible. What is T^{-1}?

21.2.12. Problem. Let V be a vector space; let W be a set which is provided with operations $(u,v) \mapsto u+v$ from $W \times W$ into W and $(\alpha,u) \mapsto \alpha u$ from $\mathbb{R} \times W$ into W; and let $T\colon V \to W$. If T is bijective and it preserves operations (that is, $T(x+y) = Tx + Ty$ and $T(\alpha x) = \alpha Tx$ for all x, $y \in V$ and $\alpha \in \mathbb{R}$), then W is a vector space which is isomorphic to V. *Hint.* Verify the eight defining axioms for a vector space. The first axiom is associativity of addition. Let u, v, $w \in W$. Write $(u+v)+w$ as $\left(T(T^{-1}u) + T(T^{-1}v)\right) + T(T^{-1}w)$ and use the hypothesis that T is operation preserving.

21.2.13. Problem. Let V be a vector space, W be a set, and $T\colon V \to W$ be a bijection. Explain carefully how W can be made into a vector space isomorphic to V. *Hint.* Use problem 21.2.12.

21.3. MATRICES

The purpose of this section and the next two is almost entirely computational. Many (but by no means all!) of the linear transformations we will consider in the sequel are maps between various Euclidean spaces; that is, between $\mathbb{R}^n$ and $\mathbb{R}^m$ where m, $n \in \mathbb{N}$. Such transformations may be represented by matrices. This of great convenience in dealing with specific examples because matrix computations are so very simple. We begin by reviewing a few elementary facts about matrices and matrix operations.

For each $n \in \mathbb{N}$ let $\mathbb{N}_n$ be $\{1,\dots,n\}$. An $m \times n$ (read "m by n") MATRIX is a function whose domain is $\mathbb{N}_m \times \mathbb{N}_n$. We deal here only with matrices of real numbers; that is, with real valued functions on $\mathbb{N}_m \times \mathbb{N}_n$. If $a\colon \mathbb{N}_m \times \mathbb{N}_n \to \mathbb{R}$ is an $m \times n$ matrix, its value at $(i,j) \in \mathbb{N}_m \times \mathbb{N}_n$ will be denoted by a^i_j. (Occasionally we use the notation a_{ij} instead.) The matrix a itself may be denoted by $\left[a^i_j\right]_{i=1}^{m}\,{}_{j=1}^{n}$, by $[a^i_j]$, or by a rectangular array whose entry in row i and column j is a^i_j.

$$\begin{bmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^m_1 & a^m_2 & \dots & a^m_n \end{bmatrix}$$

In light of this notation it is reasonable to refer to the index i in the expression a^i_j as the ROW!INDEX and to call j the COLUMN INDEX. (If you are accustomed to thinking of a matrix as *being* a rectangular array, no harm will result. The reason for defining a matrix as a function is to make good on the boast made in appendix B that everything in the sequel can be defined ultimately in terms of sets.) We denote the family of all $m \times n$ matrices of real numbers by $M_{m\times n}$. For families of square matrices we shorten $M_{n\times n}$ to M_n.

Two $m \times n$ matrices a and b may be added. Addition is done pointwise. The sum $a+b$ is the $m \times n$ matrix whose value at (i,j) is $a^i_j + b^i_j$, That is,

$$(a+b)^i_j = a^i_j + b^i_j$$

for $1 \le i \le m$ and $1 \le j \le n$. Scalar multiplication is also defined pointwise. If a is an $m \times n$ matrix and $\alpha \in \mathbb{R}$, then αa is the $m \times n$ matrix whose value at (i, j) is αa^i_j. That is,

$$(\alpha a)^i_j = \alpha a^i_j$$

for $1 \le i \le m$ and $1 \le j \le n$. We may also subtract matrices. By $-b$ we mean $(-1)b$, and by $a - b$ we mean $a + (-b)$.

21.3.1. Exercise. Let

$$\begin{bmatrix} 4 & 2 & 0 & -1 \\ -1 & -3 & 1 & 5 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & -5 & 3 & -1 \\ 3 & 1 & 0 & -1 \end{bmatrix}.$$

Find $a + b$, $3a$, and $a - 2b$. (Solution Q.21.8.)

If a is an $m \times n$ matrix and b is an $n \times p$ matrix, the product of a and b is the $m \times p$ matrix whose value at (i, j) is $\sum_{k=1}^n a^i_k b^k_j$. That is,

$$(ab)^i_j = \sum_{k=1}^n a^i_k b^k_j$$

for $1 \le i \le m$ and $1 \le j \le p$. Notice that in order for the product ab to be defined the number of columns of a must be the same as the number of rows of b. Here is a slightly different way of thinking of the product of a and b. Define the INNER PRODUCT (or DOT PRODUCT) of two n-tuples $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_n)$ to be $\sum_{k=1}^n x_k y_k$. Regard the rows of the matrix a as n-tuples (read from left to right) and the columns of b as n-tuples (read from top to bottom). Then the entry in the i^{th} row and j^{th} column of the product ab is the dot product of the i^{th} row of a and the j^{th} column of b.

21.3.2. Example. Matrix multiplication is not commutative. If a is a 2×3 matrix and b is a 3×4 matrix, then ab is defined but ba is not. Even in situations where both products ab and ba are defined, they need not be equal. For example, if $a = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}$, then $ab = \begin{bmatrix} 3 & 7 \\ -1 & 1 \end{bmatrix}$ whereas $ba = \begin{bmatrix} 0 & -2 \\ 5 & 4 \end{bmatrix}$.

21.3.3. Exercise. Let $a = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & 4 \end{bmatrix}$ and $b = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 1 & - \end{bmatrix}$. Find ab. (Solution Q.21.9.)

21.3.4. Problem. Let $a = \begin{bmatrix} 4 & 3 & 1 & 2 \\ 0 & -1 & -1 & 1 \\ 2 & 0 & 1 & 3 \end{bmatrix}$ and $b = \begin{bmatrix} 2 & -1 \\ 0 & 1 \\ 1 & 0 \\ -3 & 2 \end{bmatrix}$.

(a) Find the product ab (if it exists).
(b) Find the product ba (if it exists).

21.3.5. Definition. Let a be an $m \times n$ matrix. The TRANSPOSE of a, denoted by a^t, is the $n \times m$ matrix obtained by interchanging the rows and columns of a. That is, if $b = a^t$, then $b^i_j = a^j_i$ for $1 \le i \le n$ and $1 \le j \le m$.

21.3.6. Example. Let $a = \begin{bmatrix} 1 & 2 & 0 & -4 \\ 3 & 0 & -1 & 5 \end{bmatrix}$. Then $a^t = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & -1 \\ -4 & 5 \end{bmatrix}$.

For material in the sequel the most important role played by matrices will be as (representations of) linear transformations on finite dimensional vector spaces. Here is how it works.

21.3.7. Definition. We define the *action* of a matrix on a vector. If $a \in M_{m\times n}$ and $x \in \mathbb{R}^n$, then ax, the RESULT OF a ACTING ON x, is defined to be the vector in $\mathbb{R}^m$ whose j^{th} coordinate is $\sum_{k=1}^n a^j_k x_k$ (this is just the dot product of the j^{th} row of a with x). That is,

$$(ax)_j := \sum_{k=1}^n a^j_k x_k$$

for $1 \le j \le m$. Here is another way of saying the same thing: Regard x as an $n \times 1$ matrix

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

(sometimes called a COLUMN VECTOR). Now multiply the $m \times n$ matrix a by the $n \times 1$ matrix x. The result will be an $m \times 1$ matrix (another column vector), say

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Then ax is the m-tuple $(y_1, \dots, y_m)$. Thus a may be thought of as a mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$.

21.3.8. Exercise. Let $a = \begin{bmatrix} 3 & 0 & -1 & -4 \\ 2 & 1 & -1 & -2 \\ 1 & -3 & 0 & 2 \end{bmatrix}$ and $x = (2, 1, -1, 1)$. Find ax. (Solution Q.21.10.)

21.3.9. Problem. Let $a = \begin{bmatrix} 2 & 0 \\ 1 & -3 \\ 5 & 1 \end{bmatrix}$ and $x = (1, -2)$. Find ax.

From the definition of the action of a matrix on a vector, we derive several formulas which will be useful in the sequel. Each is a simple computation.

21.3.10. Proposition. *Let a, $b \in M_{m\times n}$; $c \in M_{n\times p}$; x, $y \in \mathbb{R}^n$; $z \in \mathbb{R}^p$; and $\alpha \in \mathbb{R}$. Then*

(a) $a(x + y) = ax + ay$;
(b) $a(\alpha x) = \alpha(ax)$;
(c) $(a + b)x = ax + bx$;
(d) $(\alpha a)x = \alpha(ax)$;
(e) $(ac)z = a(cz)$.

PROOF. Part (a) is an exercise. (Solution Q.21.11.) Parts (b)–(e) are problems.

Next we show that a sufficient (and obviously necessary) condition for two $m \times n$ matrices to be equal is that they have the same action on the standard basis vectors in $\mathbb{R}^n$.

21.3.11. Proposition. *Let a and b be $m \times n$ matrices and, as usual, let $e^1, \dots, e^n$ be the standard basis vectors in $\mathbb{R}^n$. If $ae^k = be^k$ for $1 \le k \le n$, then $a = b$.*

PROOF. Problem. *Hint.* Compute $(ae^k)_j$ and $(be^k)_j$ for $1 \le j \le m$ and $1 \le k \le n$. Remember that $(e^k)_l = 0$ if $k \ne l$ and that $(e^k)_k = 1$.

Remark. The definition 21.3.7 of the action of a matrix on a vector technically requires us to think of vectors as "column vectors". It is probably more likely that most of us think of vectors in $\mathbb{R}^n$ as "row vectors", that is, as n-tuples or as $1 \times n$ matrices. Then for the matrix multiplication ax to make sense and for the result to again be a "row vector" we really should write

$$\big(a(x^t)\big)^t$$

for the action of the matrix $a \in M_{m\times n}$ on the vector $x \in \mathbb{R}^n$. We won't do this. We will regard vectors as "row vectors" or "column vectors" as convenience dictates.

21.3.12. Definition. In our later work we will have occasion to consider the action of a SQUARE MATRIX (one with the same number of rows as columns) on a pair of vectors. Let $a \in M_n$ and $x, y \in \mathbb{R}^n$. We denote by xay the number $\sum_{j,k=1}^n a^j_k x_j y_k$.

Since

$$\sum_{j,k=1}^n a^j_k x_j y_k = \sum_{j=1}^n x_j \sum_{k=1}^n a^j_k y_k$$

and since $\sum_{k=1}^n a^j_k y_k$ is just $(ay)_j$, we may write

$$xay = \sum_{j=1}^n x_j (ay)_j .$$

In other words xay is just the dot product of the vectors x and ay. If we identify n-tuples (row vectors) with $1 \times n$ matrices, then xay is the product of the three matrices x, a, and y^t. That is,

$$xay = \begin{bmatrix} x_1 \dots x_n \end{bmatrix} \begin{bmatrix} a^1_1 & \dots & a^1_n \\ \vdots & \ddots & \vdots \\ a^n_1 & \dots & a^n_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} .$$

21.3.13. Exercise. Let $a = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 4 \\ 1 & -1 & 1 \end{bmatrix}$, $x = (1, -2, 0)$, and $y = (3, 0, 1)$. Find the action of a on the pair of vectors x and y; that is, find xay. (Solution Q.21.12.)

21.3.14. Problem. Let $a = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & -3 & 1 & 0 \\ 2 & 0 & 1 & -4 \\ -1 & 1 & -1 & 1 \end{bmatrix}$, $x = (1, -1, 0, 2)$, and $y = (1, 0, 3, 1)$. Find xay.

21.3.15. Definition. The MAIN (or PRINCIPAL) DIAGONAL of a square matrix is the diagonal running from the upper left corner to the lower right corner. That is, it consists of all the elements of the form a^k_k. If each entry on the main diagonal of an $n \times n$ matrix is 1 and all its other entries are 0, then the matrix is the $n \times n$ IDENTITY MATRIX. This matrix is denoted by I_n (or just by I if no confusion will result). If c is a real number, it is conventional to denote the matrix cI_n by c. It is clear that $a\, I_n = I_n\, a = a$ for every $n \times n$ matrix a. The $m \times n$ ZERO MATRIX is the $m \times n$ matrix all of whose entries are 0. It is denoted by $\mathbf{0}_{m\times n}$ or just by $\mathbf{0}$. Certainly $\mathbf{0} + a = a + \mathbf{0} = a$ for every $m \times n$ matrix a.

21.3.16. Definition. A square matrix a in $\mathfrak{M}_{n\times n}$ is invertible if there exists an $n \times n$ matrix a^{-1} such that

$$aa^{-1} = a^{-1}a = I_n .$$

The matrix a^{-1} is the INVERSE of a.

21.3.17. Proposition. *An $n \times n$-matrix has at most one inverse.*

PROOF. Exercise. (Solution Q.21.13.)

21.3.18. Exercise. Show that the matrix $b = \begin{bmatrix} -1/2 & 1/2 & 3/2 \\ 1/4 & 1/4 & -1/4 \\ 3/4 & -1/4 & -3/4 \end{bmatrix}$ is the inverse of the matrix $a = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & -1 \\ 1 & -1 & 1 \end{bmatrix}$. (Solution Q.21.14.)

21.3.19. Problem. Show that the matrix $a = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \\ 1 & -2 & 1 \end{bmatrix}$ satisfies the equation

$$a^3 - 4a^2 + 8a - 9 = \mathbf{0}.$$

Use this fact to find the inverse of a.

21.4. DETERMINANTS

A careful development of the properties of the determinant function on $n \times n$ matrices is not a central concern of this course. In this section we record without proof some of its elementary properties. (Proofs of these facts can be found in almost any linear algebra text. Two elegant (if not entirely elementary) presentations can be found in [4] and [6].)

21.4.1. Fact. Let $n \in \mathbb{N}$. There is exactly one function

$$\det\colon M_{n\times n} \to \mathbb{R}\colon a \mapsto \det a$$

which satisfies

(a) $\det I_n = 1$.
(b) If $a \in \mathfrak{M}_{n\times n}$ and a' is the matrix obtained by interchanging two rows of a, then $\det a' = -\det a$.
(c) If $a \in \mathfrak{M}_{n\times n}$, $c \in \mathbb{R}$, and a' is the matrix obtained by multiplying each element in one row of a by c, then $\det a' = c \det a$.
(d) If $a \in \mathfrak{M}_{n\times n}$, $c \in \mathbb{R}$, and a' is the matrix obtained from a by multiplying one row of a by c and adding it to another row of a (that is, choose $i, j \in \mathbb{N}_n$ with $i \neq j$ and replace a^j_k by $a^j_k + ca^i_k$ for each k in N_n), then $\det a' = \det a$.

21.4.2. Definition. The unique function $\det\colon \mathfrak{M}_{n\times n} \to \mathbb{R}$ described above is the $n \times n$ DETERMINANT FUNCTION.

21.4.3. Fact. If $a \in \mathbb{R}$ ($= \mathfrak{M}_{1\times 1}$), then $\det a = a$; if $a \in \mathfrak{M}_{2\times 2}$, then $\det a = a^1_1 a^2_2 - a^1_2 a^2_1$.

21.4.4. Fact. If $a, b \in \mathfrak{M}_{n\times n}$, then $\det(ab) = (\det a)(\det b)$.

21.4.5. Fact. If $a \in \mathfrak{M}_{n\times n}$, then $\det a^t = \det a$. (An obvious corollary of this: in conditions (b), (c), and (d) of fact 21.4.1 the word “columns” may be substituted for the word “rows”.)

21.4.6. Definition. Let a be an $n \times n$ matrix. The MINOR of the element a^j_k, denoted by M^j_k, is the determinant of the $(n-1) \times (n-1)$ matrix which results from the deletion of the j$^{\text{th}}$ row and k$^{\text{th}}$ column of a. The COFACTOR of the element a^j_k, denoted by C^j_k is defined by

$$C^j_k := (-1)^{j+k} M^j_k\,.$$

21.4.7. Fact. If $a \in \mathfrak{M}_{n\times n}$ and $1 \le j \le n$, then

$$\det a = \sum_{k=1}^{n} a^j_k C^j_k\,.$$

This is the (LAPLACE) EXPANSION of the determinant along the j$^{\text{th}}$ row.

In light of fact 21.4.5, it is clear that expansion along columns works as well as expansion along rows. That is,

$$\det a = \sum_{j=1}^{n} a^j_k C^j_k$$

for any k between 1 and n. This is the (LAPLACE) EXPANSION of the determinant along the k$^{\text{th}}$ column.

21.4.8. Fact. An $n \times n$ matrix a is invertible if and only if $\det a \neq 0$. If a is invertible, then

$$a^{-1} = (\det a)^{-1} C^t$$

where $C = \left[C^j_k\right]$ is the matrix of cofactors of elements of a.

21.4.9. Exercise. Let $a = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & -1 \\ 1 & -1 & 1 \end{bmatrix}$. Use the preceding facts to show that a is invertible and to compute the inverse of a. (Solution Q.21.15.)

21.4.10. Problem. Let a be the matrix given in problem 21.3.19. Use the facts stated in section 21.4 to show that a is invertible and to compute a^{-1}.

21.5. MATRIX REPRESENTATIONS OF LINEAR TRANSFORMATIONS

We are now in a position to represent members of $\mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$ by means of matrices. This will simplify computations involving such linear transformations.

21.5.1. Definition. If $T \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$, we define $[T]$ to be the $m \times n$ matrix whose entry in the j^{th} row and k^{th} column is $(Te^k)_j$, the j^{th} component of the vector Te^k in $\mathbb{R}^m$. That is, if $a = [T]$, then $a^j_k = (Te^k)_j$. The matrix $[T]$ is the MATRIX REPRESENTATION of T (with respect to the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$).

21.5.2. Example. Let $T \colon \mathbb{R}^4 \to \mathbb{R}^3 \colon (w, x, y, z) \mapsto (w + 2x + 3y, 5w + 6x + 7y + 8z, -2x - 3y - 4z)$. Then T is linear and

$$\begin{aligned} Te^1 &= T(1,0,0,0) = (1,5,0) \\ Te^2 &= T(0,1,0,0) = (2,6,-2) \\ Te^3 &= T(0,0,1,0) = (3,7,-3) \\ Te^4 &= T(0,0,0,1) = (0,8,-4). \end{aligned}$$

Having computed $Te^1, \dots, Te^4$, we use these as the successive columns of $[T]$. Thus

$$T = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 5 & 6 & 7 & 8 \\ 0 & -2 & -3 & -4 \end{bmatrix}.$$

21.5.3. Example. If $I \colon \mathbb{R}^n \to \mathbb{R}^n$ is the identity map on $\mathbb{R}^n$, then its matrix representation $[I]$ is just the $n \times n$ identity matrix I_n.

21.5.4. Exercise. Let $T \colon \mathbb{R}^2 \to \mathbb{R}^4 \colon (x, y) \mapsto (x - 3y, 7y, 2x + y, -4x + 5y)$. Find $[T]$. (Solution Q.21.16.)

The point of the representation just defined is that if we compute the action of the matrix $[T]$ on a vector x (as defined in 21.3.7), what we get is the value of T at x. Moreover, this representation is unique; that is, two distinct matrices cannot represent the same linear map.

21.5.5. Proposition. *If $T \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$, then for all x in $\mathbb{R}^n$*

$$Tx = [T]x\,.$$

Furthermore, if a is any $m \times n$ matrix which satisfies

$$Tx = ax \qquad \text{for all } x \in \mathbb{R}^n,$$

then $a = [T]$.

Proof. Exercise. *Hint.* For simplicity of notation let $b = [T]$. The map $S \colon \mathbb{R}^n \to \mathbb{R}^m \colon x \mapsto bx$ is linear. Why? To show that $Sx = Tx$ for all x in $\mathbb{R}^n$ it suffices to show that $(Se^k)_j = (Te^k)_j$ for $1 \leq k \leq n$ and $1 \leq j \leq m$. Why? (Solution Q.21.17.)

21.5.6. Proposition. *Let $m, n \in \mathbb{N}$. The map $T \mapsto [T]$ from $\mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$ into $\mathfrak{M}_{m\times n}$ is a bijection.*

Proof. Exercise. (Solution Q.21.18.)

21.5.7. Proposition. *Let $m, n \in \mathbb{N}$; let $S, T \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$; and let $\alpha \in \mathbb{R}$. Then*

(a) $[S+T] = [S] + [T]$, *and*
(b) $[\alpha T] = \alpha[T]$.

Proof. Exercise. *Hint.* For (a) use propositions 21.3.11, 21.5.5, and 21.3.10(c). (Solution Q.21.19.)

21.5.8. Theorem. *Under the operations of addition and scalar multiplication (defined in section 21.3) $\mathfrak{M}_{m\times n}$ is a vector space and the map $T \mapsto [T]$, which takes a linear transformation to its matrix representation, is an isomorphism between $\mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $\mathfrak{M}_{m\times n}$.*

Proof. Problem. *Hint.* Use problem 21.2.12.

21.5.9. Problem. Define $T : \mathbb{R}^3 \to \mathbb{R}^4$ by

$$T(x, y, z) = (x - 2y, x + y - 3z, y + 4z, 3x - 2y + z)\,.$$

Find $[T]$.

21.5.10. Problem. Define $T : \mathbb{R}^4 \to \mathbb{R}^3$ by

$$T(w, x, y, z) = (w - 3x + z, 2w + x + y - 4z, w + y + z)\,.$$

(a) Find $[T]$.
(b) Use proposition 21.5.5 to calculate $T(4, 0, -3, 1)$.

21.5.11. Problem. Let $f : \mathbb{R}^2 \to \mathbb{R}^4$ be defined by

$$f(x) = (x_1x_2, (x_1)^2 - 4(x_2)^2, (x_1)^3, x_1 \sin(\pi x_2))$$

for all $x = (x_1, x_2)$ in $\mathbb{R}^2$, and let $T : \mathbb{R}^2 \to \mathbb{R}^4$ be the linear transformation whose matrix representation is

$$\begin{bmatrix} 4 & 0 \\ 2 & -1 \\ 5 & -8 \\ -1 & 2 \end{bmatrix}.$$

Find $f(a + h) - f(a) - Th$ when $a = (-2, 1/2)$ and $h = (-1, 1)$.

21.5.12. Proposition. *If $S \in \mathfrak{L}(\mathbb{R}^p, \mathbb{R}^n)$ and $T \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$, then*

$$[TS] = [T][S]\,.$$

Proof. Problem. *Hint.* Why does it suffice to show that $[TS]x = ([T][S])x$ for all x in $\mathbb{R}^p$? Use propositions 21.5.5 and 21.3.10(e).

21.5.13. Problem. Let

$$T : \mathbb{R}^3 \to \mathbb{R}^4 : (x, y, z) \mapsto (2x + y, x - z, y + z, 3x)$$

and

$$S : \mathbb{R}^4 \to \mathbb{R}^3 : (w, x, y, z) \mapsto (x - y, y + z, z - w).$$

(a) Use proposition 21.5.12 to find $[TS]$.
(b) Use proposition 21.5.12 to find $[ST]$.

21.5.14. Problem. Show that matrix multiplication is associative; that is, show that if $a \in \mathfrak{M}_{m\times n}$, $b \in \mathfrak{M}_{n\times p}$, and $c \in \mathfrak{M}_{p\times r}$, then $(ab)c = a(bc)$. *Hint.* Don't make a complicated and messy computation of this by trying to prove it directly. Use propositions L.2.3, 21.5.6, and 21.5.12.

21.5.15. Problem. Show that $\mathfrak{M}_{n\times n}$ is a unital algebra. *Hint.* Use the definition 21.2.6. Notice that problem 21.5.14 establishes condition (c) of 21.5.14. Verify the other conditions in a similar fashion.

21.5.16. Proposition. *A linear map $T \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^n)$ is invertible if and only if* $\det[T] \neq 0$. *If T is invertible, then* $[T^{-1}] = [T]^{-1}$.

Proof. Problem. *Hint.* Show that T is invertible if and only if its matrix representation is. Then use 21.4.8.

21.5.17. Problem. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3\colon (x, y, z) \mapsto (x + 2z, y - z, x + y)$.

(a) Compute $[T]$ by calculating T and then writing down its matrix representation.
(b) Use proposition 21.5.16 to find $[T]$.

21.5.18. Problem. Let $\mathcal{P}_4$ be the family of all polynomial functions on $\mathbb{R}$ with degree (strictly) less than 4.

(a) Show that (under the usual pointwise operations) $\mathcal{P}_4$ is a vector space which is isomorphic to $\mathbb{R}^4$. *Hint.* Problem 21.2.12.
(b) Let $D\colon \mathcal{P}_4 \to \mathcal{P}_4\colon f \mapsto f'$ (where f' is the derivative of f). Using part (a) to identify the spaces $\mathcal{P}_4$ and $\mathbb{R}^4$, find a matrix representation for the (obviously linear) differentiation operator D.
(c) Use your answer to part (b) to differentiate the polynomial $7x^3 - 4x^2 + 5x - 81$.

21.5.19. Problem. Let $\mathcal{P}_4$ be as in problem 21.5.18. Consider the map

$$K\colon \mathcal{P}_4 \to \mathbb{R}\colon f \mapsto \int_0^1 f(x)\,dx\,.$$

(a) Show that K is linear.
(b) Find a way to represent K as a matrix. *Hint.* Use problem 21.5.18(a).
(c) Use your answer to part (b) to integrate the polynomial $8x^3 - 5x^2 - 4x + 6$ over the interval $[0, 1]$.
(d) Let D be as in problem 21.5.18. Find $[KD]$ by two different techniques.

21.5.20. Problem. Let $T \in \mathfrak{L}(\mathbb{R}^4, \mathbb{R}^4)$ satisfy:

$$\begin{aligned} Te^1 &= (1, 2, 0, -1) \\ Te^2 &= (1, 0, -3, 2) \\ Te^3 &= (1, -1, -1, 1) \\ Te^4 &= (0, 2, -1, 0). \end{aligned}$$

Also let $x = (1, -2, 3, -1)$ and $y = (0, 1, 2, 1)$. Find $x[T]y$.

CHAPTER 22

NORMS

22.1. NORMS ON LINEAR SPACES

The last two chapters have been pure algebra. In order to deal with topics in analysis (*e.g.* differentiation, integration, infinite series) we need also a notion of convergence; that is, we need topology as well as algebra. As in earlier chapters we consider only those topologies generated by metrics, in fact only those which arise from norms on vector spaces. Norms, which we introduce in this chapter, are very natural objects; in many concrete situations they abound. Furthermore, they possess one extremely pleasant property: a norm on a vector space generates a metric on the space, and this metric is compatible with the algebraic structure in the sense that it makes the vector space operations of addition and scalar multiplication continuous. Just as metric is a generalization of ordinary Euclidean distance, the concept of norm generalizes on vector spaces the idea of length.

22.1.1. Definition. Let V be a vector space. A function $\| \quad \|\colon V \to \mathbb{R}\colon x \mapsto \|x\|$ is a NORM on V if

(1) $\|x + y\| \le \|x\| + \|y\|$ for all x, $y \in V$,
(2) $\|\alpha x\| = |\alpha|\, \|x\|$ for all $x \in V$ and $\alpha \in \mathbb{R}$, and
(3) If $\|x\| = 0$, then $x = \mathbf{0}$.

The expression $\|x\|$ may be read as "the norm of x" or "the LENGTH of x". A vector space on which a norm has been defined is a NORMED LINEAR SPACE (or NORMED VECTOR SPACE). A vector in a normed linear space which has norm 1 is a UNIT VECTOR.

22.1.2. Example. The absolute value function is a norm on $\mathbb{R}$.

22.1.3. Example. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ let $\|x\| = \left(\sum_{k=1}^n {x_k}^2\right)^{1/2}$. The only nonobvious part of the proof that this defines a norm on $\mathbb{R}^n$ is the verification of the TRIANGLE INEQUALITY (that is, condition (1) in the preceding definition). But we have already done this: it is just Minkowski's inequality 9.2.7. This is the USUAL NORM (or EUCLIDEAN NORM) on $\mathbb{R}^n$; unless the contrary is explicitly stated, $\mathbb{R}^n$ when regarded as a normed linear space will always be assumed to possess this norm.

22.1.4. Example. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ let $\|x\|_1 = \sum_{k=1}^n |x_k|$. The function $x \mapsto \|x\|_1$ is easily seen to be a norm on $\mathbb{R}^n$. It is sometimes called the 1-NORM on $\mathbb{R}^n$.

22.1.5. Example. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ let $\|x\|_u = \max\{|x_k|\colon 1 \le k \le n\}$. Again it is easy to see that this defines a norm on $\mathbb{R}^n$; it is the UNIFORM NORM on $\mathbb{R}^n$.

22.1.6. Exercise. Let $f\colon \mathbb{R}^3 \to \mathbb{R}^4\colon (x, y, z) \mapsto (xz, x^2 + 3y, -x + y^2 - 3z, xyz - \sqrt{2}\,x)$. Find $\|f(a + \lambda h)\|$ when $a = (4, 2, -4)$, $h = (2, 4, -4)$, and $\lambda = -1/2$. (Solution Q.22.1.)

22.1.7. Exercise. Let $f\colon \mathbb{R}^3 \to \mathbb{R}^2 : x \mapsto (3{x_1}^2, x_1x_2 - x_3)$ and let $m = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$. Find $\|f(a + h) - f(a) - mh\|$ when $a = (1, 0, -2)$ and h is an arbitrary vector in $\mathbb{R}^3$. (Solution Q.22.2.)

22.1.8. Problem. Let $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$f(x, y, z) = (xy^3 + yz^2, x \sin(3\pi y), 2z)\,.$$

Find $\|f(a)\|$ when $a = (16, 1/2, 2)$.

22.1.9. Example. Let S be a nonempty set. For f in $\mathcal{B}(S, \mathbb{R})$ let

$$\|f\|_u := \sup\{|f(x)| : x \in S\}.$$

This is the UNIFORM NORM on $\mathcal{B}(S, \mathbb{R})$. Notice that example 22.1.5 is a special case of this one. [An n-tuple may be regarded as a function on the set $\{1, \ldots, n\}$. Thus $\mathbb{R}^n = \mathcal{B}(S, \mathbb{R})$ where $S = \{1, \ldots, n\}$.]

22.1.10. Exercise. Define $f, g \colon [0, 2\pi] \to \mathbb{R}$ by $f(x) = \sin x$ and $g(x) = \cos x$. Find $\|f + g\|_u$. (Solution Q.22.3.)

22.1.11. Problem. Let $f(x) = x + x^2 - x^3$ for $0 \le x \le 3$. Find $\|f\|_u$.

The following proposition lists some almost obvious properties of norms.

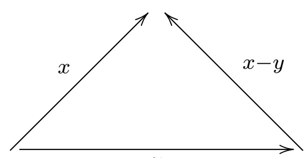
22.1.12. Proposition. *If V is a normed linear space, then*

(a) $\|\mathbf{0}\| = 0$;
(b) $\|-x\| = \|x\|$ *for all* $x \in V$; *and*
(c) $\|x\| \ge 0$ *for all* $x \in V$.

PROOF. Exercise. *Hint.* For part (a) use proposition 20.1.5. For part (b) use proposition 20.1.7; and for (c) use (a) and (b) together with the fact that $x + (-x) = \mathbf{0}$. (Solution Q.22.4.)

22.2. NORMS INDUCE METRICS

We now introduce a crucial fact: *every normed linear space is a metric space.* That is, the norm on a normed linear space induces a metric d defined by $d(x, y) = \|x - y\|$. The distance between two vectors is the length of their difference.



If no other metric is specified we always regard a normed linear space as a metric space under this induced metric. Thus the concepts of compactness, open sets, continuity, completeness, and so on, make sense on any normed linear space.

22.2.1. Proposition. *Let V be a normed linear space. Define $d \colon V \times V \to \mathbb{R}$ by $d(x, y) = \|x - y\|$. Then d is a metric on V.*

PROOF. Problem.

The existence of a metric on a normed linear space V makes it possible to speak of neighborhoods of points in V. These neighborhoods satisfy some simple algebraic properties.

22.2.2. Proposition. *Let V be a normed linear space, $x \in V$, and $r, s > 0$. Then*

(a) $B_r(0) = -B_r(0)$,
(b) $B_{rs}(0) = r\, B_s(0)$,
(c) $x + B_r(0) = B_r(x)$, *and*
(d) $B_r(0) + B_r(0) = 2B_r(0)$.

PROOF. Part (a) is an exercise. (Solution Q.22.5.) Parts (b), (c), and (d) are problems. *Hint.* For (d) divide the proof into two parts: $B_r(0) + B_r(0) \subseteq 2B_r(0)$ and the reverse inclusion. For the first, suppose x belongs to $B_r(0) + B_r(0)$. Then there exist $u, v \in B_r(0)$ such that $x = u + v$. Show that $x = 2w$ for some w in $B_r(0)$. You may wish to use problem 22.2.5.)

22.2.3. Proposition. *If V is a normed linear space then the following hold.*

(a) $\big|\, \|x\| - \|y\| \,\big| \le \|x - y\|$ *for all* $x, y \in V$.

(b) *The norm* $x \mapsto \|x\|$ *is a continuous function on* V.
(c) *If* $x_n \to a$ *in* V, *then* $\|x_n\| \to \|a\|$.
(d) $x_n \to 0$ *in* V *if and only if* $\|x_n\| \to 0$ *in* $\mathbb{R}$.

PROOF. Problem.

22.2.4. Problem. Give an example to show that the converse of part (c) of proposition 22.2.3 does not hold.

22.2.5. Problem. Prove that in a normed linear space every open ball is a convex set. And so is every closed ball.

22.3. PRODUCTS

In this and the succeeding three sections we substantially increase our store of examples of normed linear spaces by creating new spaces from old ones. In particular, we will show that each of the following can be made into a normed linear space:

(i) a vector subspace of a normed linear space,
(ii) the product of two normed linear spaces,
(iii) the set of bounded functions from a nonempty set into a normed linear space, and
(iv) the set of continuous linear maps between two normed linear spaces.

It is obvious that (i) is a normed linear space: if V is a normed linear space with norm $\|\ \ \|$ and W is a vector subspace of V, then the restriction of $\|\ \ \|$ to W is a norm on W. Now consider (ii). Given normed linear spaces V and W, we wish to make the product vector space (see example 20.1.20) into a normed linear space. As a preliminary we discuss *equivalent norms.*

22.3.1. Definition. Two norms on a vector space are EQUIVALENT if they induce equivalent metrics. If two norms on a vector space V are equivalent, then, since they induce equivalent metrics, they induce identical topologies on V. Thus properties such as continuity, compactness, and connectedness are unaltered when norms are replaced by equivalent ones (see proposition 11.2.3 and the discussion preceding it). In the next proposition we give a very simple necessary and sufficient condition for two norms to be equivalent.

22.3.2. Proposition. *Two norms* $\|\ \ \|_1$ *and* $\|\ \ \|_2$ *on a vector space* V *are equivalent if and only if there exist numbers* α, $\beta > 0$ *such that*

$$\|x\|_1 \le \alpha\|x\|_2 \qquad \text{and} \qquad \|x\|_2 \le \beta\|x\|_1$$

for all $x \in V$.

PROOF. Exercise. (Solution Q.22.6.)

If V and W are normed linear spaces with norms $\|\ \ \|_V$ and $\|\ \ \|_W$ respectively, how do we provide the product vector space $V \times W$ (see example 20.1.20) with a norm? There are at least three more or less obvious candidates: for $v \in V$ and $w \in W$ let

$$\begin{aligned}\|(v,w)\| &:= \left(\|v\|_V{}^2 + \|w\|_W{}^2\right)^{1/2},\\ \|(v,w)\|_1 &:= \|v\|_V + \|w\|_W, \qquad \text{and}\\ \|(v,w)\|_u &:= \max\{\|v\|_V, \|w\|_W\}.\end{aligned}$$

First of all, are these really norms on $V \times W$? Routine computations show that the answer is *yes.* By way of illustration we write out the three verifications required for the first of the candidate norms. If v, $x \in V$, if w, $y \in W$, and if $\alpha \in \mathbb{R}$, then

$$\begin{aligned}
\text{(a)} \qquad \|(v,w)+(x,y)\| &= \|(v+x,w+y)\| \\
&= \left(\|v+x\|_V{}^2 + \|w+y\|_W{}^2\right)^{1/2} \\
&\le \left((\|v\|_V + \|x\|_V)^2 + (\|w\|_W + \|y\|_W)^2\right)^{1/2} \\
&\le \left(\|v\|_V{}^2 + \|w\|_W{}^2\right)^{1/2} + \left(\|x\|_V{}^2 + \|y\|_W{}^2\right)^{1/2} \\
&= \|(v,w)\| + \|(x,y)\|.
\end{aligned}$$

The last inequality in this computation is, of course, Minkowski's inequality 9.2.7.

$$\begin{aligned}
\text{(b)} \qquad \|\alpha(v,w)\| &= \|(\alpha v, \alpha w)\| \\
&= \left(\|\alpha v\|_V{}^2 + \|\alpha w\|_W{}^2\right)^{1/2} \\
&= \left(\big(|\alpha|\,\|v\|_V\big)^2 + \big(|\alpha|\,\|w\|_W\big)^2\right)^{1/2} \\
&= |\alpha|\left(\|v\|_V{}^2 + \|w\|_W{}^2\right)^{1/2} \\
&= |\alpha|\,\|(v,w)\|.
\end{aligned}$$

(c) If $\|(v,w)\| = 0$, then $\|v\|_V{}^2 + \|w\|_W{}^2 = 0$. This implies that $\|v\|_V$ and $\|w\|_W$ are both zero. Thus v is the zero vector in V and w is the zero vector in W; so $(v,w) = (0,0)$, the zero vector in $V \times W$.

Now which of these norms should we choose to be the product norm on $V \times W$? The next proposition shows that at least as far as topological considerations (continuity, compactness, connectedness, *etc.*) are concerned, it really doesn't matter.

22.3.3. Proposition. *The three norms on $V \times W$ defined above are equivalent.*

Proof. Notice that the norms $\|\quad\|$, $\|\quad\|_1$, and $\|\quad\|_u$ defined above induce, respectively, the metrics d, d_1, and d_u defined in chapter 9. In proposition 9.3.2 we proved that these three metrics are equivalent. Thus the norms which induce them are equivalent. □

22.3.4. Definition. Since d_1 was chosen (in 12.3.3) as our "official" product metric, we choose $\|\quad\|_1$, which induces d_1, as the PRODUCT NORM on $V \times W$. In proposition 22.3.8 you are asked to show that with this definition of the product norm, the operation of addition on a normed linear space is continuous. In the next proposition we verify that scalar multiplication (regarded as a map from $\mathbb{R} \times V$ into V) is also continuous.

22.3.5. Proposition. *If V is a normed linear space, then the mapping $(\beta, x) \mapsto \beta x$ from $\mathbb{R} \times V$ into V is continuous.*

Proof. Exercise. *Hint.* To show that a map $f \colon U \to W$ between two normed linear space is continuous at a point a in U, it must be shown that for every $\epsilon > 0$ there exists $\delta > 0$ such that $\|u - a\|_U < \delta$ implies $\|f(u) - f(a)\|_W < \epsilon$. (Solution Q.22.7.)

22.3.6. Corollary. *Let (β_n) be a sequence of real numbers and (x_n) be a sequence of vectors in a normed linear space V. If $\beta_n \to \alpha$ in $\mathbb{R}$ and $x_n \to a$ in V, then $\beta_n x_n \to \alpha a$ in V.*

Proof. Exercise. (Solution Q.22.8.)

22.3.7. Corollary. *If V is a normed linear space and α is a nonzero scalar, the map*

$$M_\alpha \colon V \to V \colon x \mapsto \alpha x$$

is a homeomorphism.

Proof. Problem.

22.3.8. Proposition. *Let V be a normed linear space. The operation of addition*

$$A\colon V\times V\to V\colon (x,y)\mapsto x+y$$

is continuous.

Proof. Problem.

22.3.9. Problem. Let V be a normed linear space. Prove the following:

(a) If $x_n\to a$ and $y_n\to b$ in V, then $x_n+y_n\to a+b$.
(b) If S is a vector subspace of V, then so is $\overline{S}$.

22.3.10. Problem. If K and L are compact subsets of a normed linear space, then the set

$$K+L:=\{k+l\colon k\in K \text{ and } l\in L\}$$

is compact. *Hint.* Let A be as in proposition 22.3.8. What is $A^{\rightarrow}(K\times L)$?

22.3.11. Problem. Let B and C be subsets of a normed linear space and $\alpha\in\mathbb{R}$. Prove the following:

(a) $\overline{\alpha B}=\alpha\overline{B}$.
(b) $\overline{B}+\overline{C}\subseteq\overline{B+C}$.
(c) $B+C$ need not be closed even if B and C are; thus equality need not hold in (b). *Hint.* In $\mathbb{R}^2$ try part of the curve $y=1/x$ and the negative x-axis.

22.3.12. Problem. Show that a linear bijection $f\colon V\to W$ between normed linear spaces is an isometry if and only if it is norm preserving (that is, if and only if $\|f(x)\|_W=\|x\|_V$ for all $x\in V$).

22.3.13. Definition. Let a be a vector in a vector space V. The map

$$T_a\colon V\to V\colon x\mapsto x+a$$

is called TRANSLATION by a.

22.3.14. Problem. Show that every translation map on a normed linear space is an isometry and therefore a homeomorphism.

22.3.15. Problem. Let U be a nonempty open set in a normed linear space. Then $U-U$ contains a neighborhood of 0. (By $U-U$ we mean $\{u-v\colon u,v\in U\}$.) *Hint.* Consider the union of all sets of the form $\left(T_{-v}\right)^{\rightarrow}(U)$ where $v\in U$. (As in problem 22.3.14 T_{-v} is a translation map.)

22.3.16. Problem. Show that if B is a closed subset of a normed linear space V and C is a compact subset of V, then $B+C$ is closed. (Recall that part (c) of problem 22.3.11 showed that this conclusion cannot be reached by assuming only that B and C are closed.) *Hint.* Use the sequential characterization of "closed" given in proposition 12.2.2. Let (a_n) be a sequence in $B+C$ which converges to a point in V. Write $a_n=b_n+c_n$ where $b_n\in B$ and $c_n\in C$. Why does (c_n) have a subsequence $\left(c_{n_k}\right)$ which converges to a point in C? Does $\left(b_{n_k}\right)$ converge?

22.3.17. Problem. Let V and W be normed linear spaces, $A\subseteq V$, $a\in A'$, $\alpha\in\mathbb{R}$, and $f,g\colon A\to W$. Prove the following:

(a) If the limits of f and g exist as x approaches a, then so does the limit of $f+g$ and
$$\lim_{x\to a}(f+g)(x)=\lim_{x\to a}f(x)+\lim_{x\to a}g(x)\,.$$
(b) If the limit of f exists as x approaches a, then so does the limit of αf and
$$\lim_{x\to a}(\alpha f)(x)=\alpha\lim_{x\to a}f(x)\,.$$

22.3.18. Problem. Let V and W be normed linear spaces, $A\subseteq V$, $a\in A'$, and $f\colon A\to W$. Show that

(a) $\lim_{x\to a}f(x)=0$ if and only if $\lim_{x\to a}\|f(x)\|=0$; and

(b) $\lim_{h\to 0} f(a+h) = \lim_{x\to a} f(x)$.

Hint. These require only the most trivial modifications of the solutions to problem 14.3.11 and proposition 14.3.5.

22.3.19. Problem. Let V and W be normed linear spaces, $A \subseteq V$, and $f : A \to W$. Suppose that a is an accumulation point of A and that $l = \lim_{x\to a} f(x)$ exists in W.

(a) Show that if the norm on V is replaced by an equivalent one, then a is still an accumulation point of A.

(b) Show that if both the norm on V and the one on W are replaced by equivalent ones, then it is still true that $f(x) \to l$ as $x \to a$.

22.3.20. Problem. Let $f : U \times V \to W$ where U, V, and W are normed linear spaces. If the limit

$$l := \lim_{(x,y)\to(a,b)} f(x,y)$$

exists and if $\lim_{x\to a} f(x,y)$ and $\lim_{y\to b} f(x,y)$ exist for all $y \in V$ and $x \in U$, respectively, then the iterated limits

$$\lim_{x\to a}\bigl(\lim_{y\to b} f(x,y)\bigr) \quad \text{and} \quad \lim_{y\to b}\bigl(\lim_{x\to a} f(x,y)\bigr)$$

exist and are equal to l.

22.3.21. Problem. All norms on $\mathbb{R}^n$ are equivalent. *Hint.* It is enough to show that an arbitrary norm $|||\ \ |||$ on R^n is equivalent to $\|\ \ \|_1$ (where $\|x\|_1 = \sum_{k=1}^n |x_k|$). Use proposition 22.3.2. To find $\alpha > 0$ such that $|||x||| \le \alpha \|x\|_1$ for all x write $x = \sum_{k=1}^n x_k e^k$ (where $e^1, \dots, e^n$ are the standard basis vectors on $\mathbb{R}^n$). To find $\beta > 0$ such that $\|x\|_1 \le \beta |||x|||$ let $\mathbb{R}^n_1$ be the normed linear space $\mathbb{R}^n$ under the norm $\|\ \ \|_1$. Show that the function $x \mapsto |||x|||$ from $\mathbb{R}^n_1$ into $\mathbb{R}$ is continuous. Show that the unit sphere $S = \{x \in \mathbb{R}^n : \|x\|_1 = 1\}$ is compact in $\mathbb{R}^n_1$.

22.4. THE SPACE $\mathcal{B}(S,V)$

Throughout this section S will be a nonempty set and V a normed linear space. In the first section of this chapter we listed $\mathcal{B}(S,\mathbb{R})$ as an example of a normed linear space. Here we do little more than observe that the fundamental facts presented in chapter 13 concerning pointwise and uniform convergence in the space $\mathcal{B}(S,\mathbb{R})$ all remain true when the set $\mathbb{R}$ is replaced by an arbitrary normed linear space. It is very easy to generalize these results: replace absolute values by norms.

22.4.1. Definition. Let S be a set and V be a normed linear space. A function $f : S \to V$ is BOUNDED if there exists a number $M > 0$ such that

$$\|f(x)\| \le M$$

for all x in S. We denote by $\mathcal{B}(S,V)$ the family of all bounded V valued functions on S.

22.4.2. Exercise. Under the usual pointwise operations $\mathcal{B}(S,V)$ is a vector space. (Solution Q.22.9.)

22.4.3. Definition. Let S be a set and V be a normed linear space. For every f in $\mathcal{B}(S,V)$ define

$$\|f\|_u := \sup\{\|f(x)\| : x \in S\}\,.$$

The function $f \mapsto \|f\|_u$ is called the UNIFORM NORM on $\mathcal{B}(S,V)$. This is the usual norm on $\mathcal{B}(S,V)$.

In the following problem you are asked to show that this function really is a norm. The metric d_u induced by the uniform norm $\|\ \ \|_u$ on $\mathcal{B}(S,V)$ is the UNIFORM METRIC on $\mathcal{B}(S,V)$.

22.4.4. Problem. Show that the uniform norm defined in 22.4.3 is in fact a norm on $\mathcal{B}(S,V)$.

22.4.5. Definition. Let (f_n) be a sequence of functions in $\mathcal{F}(S,V)$. If there is a function g in $\mathcal{F}(S,V)$ such that

$$\sup\{\|f_n(x) - g(x)\| : x \in S\} \to 0 \quad \text{as } n \to \infty,$$

then we say that the sequence (f_n) CONVERGES UNIFORMLY to g and write $f_n \to g\,(\text{unif})$. The function g is the UNIFORM LIMIT of the sequence (f_n). Notice that if g and all the f_n's belong to

$\mathcal{B}(S, V)$, then uniform convergence of (f_n) to g is just convergence of (f_n) to g with respect to the uniform metric. Notice also that the preceding repeats verbatim definition 13.1.12, except that $\mathbb{R}$ has been replaced by V and absolute values by norms. We may similarly generalize definition 13.2.1.

22.4.6. Definition. Let (f_n) be a sequence in $\mathcal{F}(S, V)$. If there is a function g such that

$$f_n(x) \to g(x) \qquad \text{for all } x \in S,$$

then (fn) CONVERGES POINTWISE to g. In this case we write

$$f_n \to g\,(\text{ptws}).$$

The function g is the POINTWISE LIMIT of the f_n's. Problem 22.4.7 repeats proposition 13.2.2—uniform convergence implies pointwise convergence—except that it now holds for V valued functions (not just real valued ones). Problem 22.4.8 generalizes proposition 13.2.4(a).

22.4.7. Problem. If a sequence (f_n) in $\mathcal{F}(S, V)$ converges uniformly to a function g in $\mathcal{F}(S, V)$, then $f_n \to g\,(\text{ptws})$.

22.4.8. Problem. Let $f_n)$ be a sequence in $\mathcal{B}(S, V)$ and g be a member of $\mathcal{F}(S, V)$. If $f_n \to g\,(\text{unif})$, then g is bounded.

22.4.9. Problem. Define

$$f(t) = \begin{cases} (t, 0), & \text{if } 0 \le t \le 1; \\ (1, t-1) & \text{if } 1 < t \le 2; \\ (3-t, 1) & \text{if } 2 < t \le 3; \\ (0, 4-t) & \text{if } 3 < t \le 4. \end{cases}$$

Regarding f as a member of the space $\mathcal{B}([0, 4], \mathbb{R}^2)$ find $\|f\|_u$.

22.4.10. Example. If M is a compact metric space, then the family $\mathcal{C}(M, V)$ of all continuous V-valued functions on M is a normed linear space.

PROOF. Problem.

22.4.11. Problem. Let M be a compact metric space. Show that the family $\mathcal{C}(M, \mathbb{R})$ of all continuous real valued functions on M is a unital algebra and that $\|fg\|_u \le \|f\|_u \|g\|_u$. Show also that if A is a subalgebra of $\mathcal{C}(M, \mathbb{R})$, then so is $\overline{A}$.

22.4.12. Proposition. *If (f_n) is a sequence of continuous V valued functions on a metric space M and if this sequence converges uniformly to a V valued function g on M, then g is continuous.*

PROOF. Problem. *Hint.* Modify the proof of proposition 14.2.15.

CHAPTER 23

CONTINUITY AND LINEARITY

23.1. BOUNDED LINEAR TRANSFORMATIONS

Normed linear spaces have both algebraic and topological structure. It is therefore natural to be interested in those functions between normed linear spaces which preserve both types of structure, that is, which are both linear and continuous. In this section we study such functions.

23.1.1. Definition. A linear transformation $T\colon V \to W$ between normed linear spaces is BOUNDED if there exists a number $M > 0$ such that

$$\|Tx\| \le M\|x\|$$

for all $x \in V$. The family of all bounded linear transformations from V into W is denoted by $\mathfrak{B}(V, W)$.

CAUTION. There is a possibility of confusion. Here we have defined bounded linear transformations; in section 22.4 we gave a quite different definition for "bounded" as it applies to arbitrary vector valued functions; and certainly linear transformations are such functions. The likelihood of becoming confused by these two different notions of boundedness is very small once one has made the following observation: Except for the zero function, it is impossible for a linear transformation to be bounded in the sense of section 22.4. (Proof. Let $T\colon V \to W$ be a nonzero linear transformation between normed linear spaces. Choose a in V so that $Ta \ne 0$. Then $\|Ta\| > 0$, so that by the Archimedean principle (proposition J.4.1) the number $\|T(na)\| = n\|Ta\|$ can be made as large as desired by choosing n sufficiently large. Thus there is certainly no $M > 0$ such that $\|Tx\| \le M$ for all x.) Since nonzero linear transformations cannot be bounded in the sense of section 22.4, an assertion that a linear map is bounded should always be interpreted in the sense of boundedness introduced in this section.

A function $f\colon S \to W$, which maps a set S into a normed linear space W, is bounded if and only if it maps *every* subset of S into a bounded subset of W. However, a *linear* map $T\colon V \to W$ from one normed linear space into another is bounded if and only if it maps every bounded subset of V into a bounded subset of W.

23.1.2. Proposition. *A linear transformation $T\colon V \to W$ between normed linear spaces is bounded if and only if $T^{\to}(A)$ is a bounded subset of W whenever A is a bounded subset of V.*

PROOF. Problem.

23.1.3. Example. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3\colon (x, y) \mapsto (3x + y, x - 3y, 4y)$. It is easily seen that T is linear. For all (x, y) in $\mathbb{R}^2$

$$\begin{aligned}
\|T(x,y)\| &= \|(3x + y, x - 3y, 4y)\| \\
&= \big((3x + y)^2 + (x - 3y)^2 + (4y)^2\big)^{\frac{1}{2}} \\
&= (10x^2 + 26y^2)^{\frac{1}{2}} \\
&\le \sqrt{26}(x^2 + y^2)^{\frac{1}{2}} \\
&= \sqrt{26}\|(x, y)\| .
\end{aligned}$$

So the linear transformation T is bounded.

Why is boundedness of linear transformations an important concept? Because it turns out to be equivalent to continuity and is usually easier to establish.

23.1.4. Proposition. *Let $T\colon V \to W$ be a linear transformation between normed linear spaces. The following are equivalent:*

(a) *T is continuous.*
(b) *There is at least one point at which T is continuous.*
(c) *T is continuous at $\mathbf{0}$.*
(d) *T is bounded.*

Proof. Exercise. *Hint.* To prove that (c) implies (d) argue by contradiction. Show that for each n there exists x_n in V such that $\|Tx_n\| > n\|x_n\|$. Let $y_n = \big(n\|x_n\|\big)^{-1}x_n$. Give one argument to show that the sequence (Ty_n) converges to zero as $n \to \infty$. Give another to show that it does not. (Solution Q.23.1.)

23.1.5. Definition. Let $T \in \mathfrak{B}(V, W)$ where V and W are normed linear spaces. Define

$$\|T\| := \inf\{M > 0\colon \|Tx\| \le M\|x\| \text{for all } x \in V\}.$$

This number is called the norm of T. We show in proposition 23.1.14 that the map $T \mapsto \|T\|$ really is a norm on $\mathfrak{B}(V, W)$.

There are at least four ways to compute the norm of a linear transformation T. Use the definition or any of the three formulas given in the next lemma.

23.1.6. Lemma. *Let T be a bounded linear map between nonzero normed linear spaces. Then*

$$\begin{aligned}\|T\| &= \sup\{\|x\|^{-1}\,\|Tx\|\colon x \ne \mathbf{0}\}\\ &= \sup\{\|Tu\|\colon \|u\| = 1\}\\ &= \sup\{\|Tu\|\colon \|u\| \le 1\}.\end{aligned}$$

Proof. Exercise. *Hint.* To obtain the first equality, use the fact that if a subset A of $\mathbb{R}$ is bounded above then $\sup A = \inf\{M\colon M \text{ is an upper bound for } A\}$. (Solution Q.23.2.)

23.1.7. Corollary. *If $T \in \mathfrak{B}(V, W)$ where V and W are normed linear spaces, then*

$$\|Tx\| \le \|T\|\,\|x\|$$

for all x in V.

Proof. By the preceding lemma $\|T\| \ge \|x\|^{-1}\|Tx\|$ for all $x \ne 0$. Thus $\|Tx\| \le \|T\|\,\|x\|$ for all x. □

The following example shows how to use lemma 23.1.6 (in conjunction with the definition) to compute the norm of a linear transformation.

23.1.8. Example. Let T be the linear map defined in example 23.1.3. We have already seen that $\|T(x, y)\| \le \sqrt{26}\|(x, y)\|$ for all (x, y) in $\mathbb{R}^2$. Since $\|T\|$ is defined to be the infimum of the set of all numbers M such that $\|T(x, y)\| \le M\|(x, y)\|$ for all $(x, y) \in \mathbb{R}^2$, and since $\sqrt{26}$ is such a number, we know that

$$\|T\| \le \sqrt{26} \tag{23.1}$$

On the other hand lemma 23.1.6 tells us that $\|T\|$ is the supremum of the set of all numbers $\|Tu\|$ where u is a unit vector. Since $(0, 1)$ is a unit vector and $\|T(0, 1)\| = \|(1, -3, 4)\| = \sqrt{26}$, we conclude that

$$\|T\| \ge \sqrt{26} \tag{23.2}$$

Conditions (23.1) and (23.2) imply that $\|T\| = \sqrt{26}$.

23.1.9. Problem. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by $T(x, y) = (3x, x + 2y, x - 2y)$. Find $\|T\|$.

23.1.10. Problem. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^4\colon x \mapsto (x_1 - 4x_2, 2x_1 + 3x_3, x_1 + 4x_2, x_1 - 6x_3)$. Find $\|T\|$.

23.1.11. Exercise. Find the norm of each of the following.

(a) The identity map on a normed linear space.
(b) The zero map in $\mathfrak{B}(V, W)$.
(c) A coordinate projection $\pi_k : V_1 \times V_2 \to V_k$ $(k = 1, 2)$ where V_1 and V_2 are nontrivial normed linear spaces (that is, normed linear spaces which contain vectors other than the zero vector.)

(Solution Q.23.3.)

23.1.12. Exercise. Let $\mathcal{C} = \mathcal{C}([a, b], \mathbb{R})$. Define $J\colon \mathcal{C} \to \mathbb{R}$ by

$$Jf = \int_a^b f(x)\,dx\,.$$

Show that $J \in \mathfrak{B}(\mathcal{C}, \mathbb{R})$ and find $\|J\|$. (Solution Q.23.4.)

23.1.13. Problem. Let $\mathcal{C}^1$ and $\mathcal{C}$ be as in problem 21.1.23. Let D be the differentiation operator

$$D\colon \mathcal{C}^1 \to \mathcal{C}\colon f \mapsto f'$$

(where f' is the derivative of f). Let both $\mathcal{C}^1$ and $\mathcal{C}$ have the uniform norm. Is the linear transformation D bounded? *Hint.* Let $[a, b] = [0, 1]$ and consider the functions $f_n(x) = x^n$ for $n \in \mathbb{N}$ and $0 \le x \le 1$.

Next we show that the set $\mathfrak{B}(V, W)$ of all bounded linear transformations between two normed linear spaces is itself a normed linear space.

23.1.14. Proposition. *If V and W are normed linear spaces then under pointwise operations $\mathfrak{B}(V, W)$ is a vector space and the map $T \mapsto \|T\|$ from $\mathfrak{B}(V, W)$ into $\mathbb{R}$ defined above is a norm on $\mathfrak{B}(V, W)$.*

Proof. Exercise. (Solution Q.23.5.)

One obvious fact that we state for future reference is that the composite of bounded linear transformations is bounded and linear.

23.1.15. Proposition. *If $S \in \mathfrak{B}(U, V)$ and $T \in \mathfrak{B}(V, W)$, then $TS \in \mathfrak{B}(U, W)$ and $\|TS\| \le \|T\|\,\|S\|$*

Proof. Exercise. (Solution Q.23.6.)

In propositions 21.1.16 and 21.1.17 we saw that the kernel and range of a linear map T are vector subspaces, respectively, of the domain and codomain of T. It is interesting to note that the kernel is always a closed subspace while the range need not be.

23.1.16. Proposition. *If V and W are normed linear spaces and $T \in \mathfrak{B}(V, W)$, then $\ker T$ is a closed linear subspace of V.*

Proof. Problem.

23.1.17. Example. Let c_0 be the vector space of all sequences x of real numbers (with pointwise operations) which converge to zero. Give c_0 the uniform norm (see example 22.1.9 so that for every $x \in c_0$

$$\|x\|_u = \sup\{x_k : k \in \mathbb{N}\}\,.$$

The family l of sequences of real numbers which have only finitely many nonzero coordinates is a vector subspace of c_0, but it is not a closed subspace. Thus the range of the inclusion map of l into c_0 does not have closed range.

Proof. Problem.

In general the calculus of infinite dimensional spaces is no more complicated than calculus on $\mathbb{R}^n$. One respect in which the Euclidean spaces $\mathbb{R}^n$ turn out to be simpler, however, is the fact that every linear map from $\mathbb{R}^n$ into $\mathbb{R}^m$ is automatically continuous. Between finite dimensional spaces there are no discontinuous linear maps. And this is true regardless of the particular norms placed on these spaces.

23.1.18. Proposition. *Let $\mathbb{R}^n$ and $\mathbb{R}^m$ have any norms whatever. If $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear, then it is continuous.*

PROOF. Problem. *Hint.* Let $\left[t^j_k\right] = [T]$ be the matrix representation of T and

$$M = \max\{\left|t^j_k\right| : 1 \leq j \leq m \text{ and } 1 \leq k \leq n\}\,.$$

Let $\mathbb{R}^m_u$ be $\mathbb{R}^m$ provided with the uniform norm

$$\|x\|_u := \max\{|x_k| : 1 \leq k \leq m\}$$

and $\mathbb{R}^n_1$ be $\mathbb{R}^n$ equipped with the norm

$$\|x\|_1 := \sum_{k=1}^{n} |x_k|\,.$$

Show that T regarded as a map from $\mathbb{R}^n_1$ to $\mathbb{R}^m_u$ is bounded (with $\|T\| \leq M$). Then use problem 22.3.21.

23.1.19. Problem. Let V and W be normed linear spaces and $x \in V$. Define a map E_x (called EVALUATION AT x) by

$$E_x \colon \mathfrak{B}(V,W) \to W \colon T \mapsto Tx\,.$$

Show that $E_x \in \mathfrak{B}\big(\mathfrak{B}(V,W), W\big)$ and that $\|E_x\| \leq \|x\|$.

23.1.20. Problem. What changes in the preceding problem if we let M be a compact metric space and W be a nonzero normed linear space and consider the evaluation map $E_x\colon \mathcal{C}(M,W) \to W \colon f \mapsto f(x)$?

23.1.21. Problem. Let S be a nonempty set and $T\colon V \to W$ be a bounded linear transformation between normed linear spaces. Define a function C_T on the normed linear space $\mathcal{B}(S,V)$ by

$$C_T(f) := T \circ f$$

for all f in $\mathcal{B}(S,V)$.

(a) Show that $C_T(f)$ belongs to $\mathcal{B}(S,W)$ whenever f is a member of $\mathcal{B}(S,V)$.
(b) Show that the map $C_T\colon \mathcal{B}(S,V) \to \mathcal{B}(S,W)$ is linear and continuous.
(c) Find $\|C_T\|$.
(d) Show that if $f_n \to g$ (unif) in $\mathcal{B}(S,V)$, then $T \circ f_n \to T \circ g$ (unif) in $\mathcal{B}(S,W)$.
(e) Show that C_T is injective if and only if T is.

23.1.22. Problem. Let M and N be compact metric spaces and $\phi\colon M \to N$ be continuous. Define T_ϕ on $\mathcal{C}(N,\mathbb{R})$ by

$$T_\phi(g) := g \circ \phi$$

for all g in $\mathcal{C}(N,\mathbb{R})$.

(a) T_ϕ maps $\mathcal{C}(N,\mathbb{R})$ into $\mathcal{C}(M,\mathbb{R})$.
(b) T_ϕ is a bounded linear transformation.
(c) $\|T_\phi\| = 1$.
(d) If ϕ is surjective, then T_ϕ is injective.
(e) If T_ϕ is injective, then ϕ is surjective. *Hint.* Suppose ϕ is not surjective. Choose y in $N \setminus \operatorname{ran}\phi$. Show that problem 14.1.31 can be applied to the sets $\{y\}$ and $\operatorname{ran}\phi$.

(f) If T_ϕ is surjective, then ϕ is injective. *Hint.* Here again problem 14.1.31 is useful. [It is also true that if ϕ is injective, then T_ϕ is surjective. But more machinery is needed before we can prove this.]

23.1.23. Problem. Let (S_k) be a sequence in $\mathfrak{B}(V,W)$ and U be a member of $\mathfrak{B}(W,X)$ where V, W, and X are normed linear spaces. If $S_k \to T$ in $\mathfrak{B}(V,W)$, then $US_k \to UT$. Also, state and prove a similar result whose conclusion is, "then $S_kU \to TU$."

23.1.24. Definition. A family $\mathfrak{T}$ of linear maps from a vector space into itself is a COMMUTING FAMILY if $ST = TS$ for all $S, T \in \mathfrak{T}$.

23.1.25. Problem (Markov-Kakutani Fixed Point Theorem)**.** Prove: If $\mathfrak{T}$ is a commuting family of bounded linear maps from a normed linear space V into itself and K is a nonempty convex compact subset of V which is mapped into itself by every member of $\mathfrak{T}$, then there is at least one point in K which is fixed under every member of $\mathfrak{T}$. *Hint.* For every T in $\mathfrak{T}$ and n in $\mathbb{N}$ define

$$T_n = n^{-1}\sum_{j=0}^{n-1} T^j$$

(where $T^0 := I$). Let $\mathfrak{U} = \{T_n \colon T \in \mathfrak{T} \text{ and } n \in \mathbb{N}\}$. Show that $\mathfrak{U}$ is a commuting family of bounded linear maps on V each of which maps K into itself and that if $U_1, \dots, U_n \in \mathfrak{U}$, then

$$(U_1 \dots U_n)^{\to}(K) \subseteq \bigcap_{j=1}^{n} U_j^{\to}(K)\,.$$

Let $\mathfrak{C} = \{U^{\to}(K) \colon U \in \mathfrak{U}\}$ and use problem 15.3.2 to show that $\bigcap \mathfrak{C} \neq \emptyset$.

Finally, show that every element of $\bigcap \mathfrak{C}$ is fixed under each T in $\mathfrak{T}$; that is, if $a \in \bigcap \mathfrak{C}$ and $T \in \mathfrak{T}$, then $Ta = a$. To this end argue that for every $n \in \mathbb{N}$ there exists $c_n \in K$ such that $a = T_n c_n$ and therefore $Ta - a$ belongs to $n^{-1}(K - K)$ for each n. Use problems 22.3.10 and 15.1.5 to show that every neighborhood of 0 contains, for sufficiently large n, sets of the form $n^{-1}(K - K)$. What do these last two observations say about $Ta - a$?

23.2. THE STONE-WEIERSTRASS THEOREM

In example 15.3.5 we found that the square root function can be uniformly approximated on $[0,1]$ by polynomials. In this section we prove the remarkable *Weierstrass approximation theorem* which says that *every* continuous real valued function can be uniformly approximated on compact intervals by polynomials. We will in fact prove an even stronger result due to M. H. Stone which generalizes the Weierstrass theorem to arbitrary compact metric spaces.

23.2.1. Proposition. *Let A be a subalgebra of $\mathcal{C}(M,\mathbb{R})$ where M is a compact metric space. If $f \in A$, then $|f| \in \overline{A}$.*

PROOF. Exercise. *Hint.* Let (p_n) be a sequence of polynomials converging uniformly on $[0,1]$ to the square root function. (See 15.3.5.) What can you say about the sequence $(p_n \circ g^2)$ where $g = f/\|f\|$? (Solution Q.23.7.)

23.2.2. Corollary. *If A is a subalgebra of $\mathcal{C}(M,\mathbb{R})$ where M is a compact metric space, and if f, $g \in A$, then $f \vee g$ and $f \wedge g$ belong to $\overline{A}$.*

PROOF. As in the solution to problem 14.2.10, write $f \vee g = \frac{1}{2}(f + g + |f - g|)$ and $f \wedge g = \frac{1}{2}(f + g - |f - g|)$; then apply the preceding proposition. □

23.2.3. Definition. A family $\mathcal{F}$ of real valued functions defined on a set S is a SEPARATING family if corresponding to every pair of distinct points x and y and S there is a function f in $\mathcal{F}$ such that $f(x) \neq f(y)$. In this circumstance we may also say that the family $\mathcal{F}$ SEPARATES POINTS of S.

23.2.4. Proposition. *Let A be a separating unital subalgebra of $\mathcal{C}(M,\mathbb{R})$ where M is a compact metric space. If a and b are distinct points in M and $\alpha, \beta \in \mathbb{R}$, then there exists a function $f \in A$ such that $f(a) = \alpha$ and $f(b) = \beta$.*

Proof. Problem. *Hint.* Let g be any member of A such that $g(a) \neq g(b)$. Notice that if k is a constant, then the function $f\colon x \mapsto \alpha + k(g(x) - g(a))$ satisfies $f(a) = \alpha$. Choose k so that $f(b) = \beta$.

Suppose that M is a compact metric space. The *Stone-Weierstrass theorem* says that any separating unital subalgebra of the algebra of continuous real valued functions on M is dense. That is, if A is such a subalgebra, then we can approximate each function $f \in \mathcal{C}(M,\mathbb{R})$ arbitrarily closely by members of A. The proof falls rather naturally into two steps. First (in lemma 23.2.5) we find a function in $\overline{A}$ which does not exceed f by much; precisely, given $a \in M$ and $\epsilon > 0$ we find a function g in $\overline{A}$ which agrees with f at a and satisfies $g(x) < f(x) + \epsilon$ elsewhere. Then (in theorem 23.2.6) given $\epsilon > 0$ we find $h \in \overline{A}$ such that $f(x) - \epsilon < h(x) < f(x) + \epsilon$.

23.2.5. Lemma. *Let A be a unital separating subalgebra of $\mathcal{C}(M,\mathbb{R})$ where M is a compact metric space. For every $f \in \mathcal{C}(M,\mathbb{R})$, every $a \in M$, and every $\epsilon > 0$ there exists a function $g \in \overline{A}$ such that $g(a) = f(a)$ and $g(x) < f(x) + \epsilon$ for all $x \in M$.*

Proof. Exercise. *Hint.* For each $y \in M$ find a function ϕ_y which agrees with f at a and at y. Then $\phi_y(x) < f(x) + \epsilon$ for all x in some neighborhood U_y of y. Find finitely many of these neighborhoods $U_{y_1}, \dots, U_{y_n}$ which cover M. Let $g = \phi_{y_1} \wedge \dots \wedge \phi_{y_n}$. (Solution Q.23.8.)

23.2.6. Theorem (Stone-Weierstrass Theorem)**.** *Let A be a unital separating subalgebra of $\mathcal{C}(M,\mathbb{R})$ where M is a compact metric space. Then A is dense in $\mathcal{C}(M,\mathbb{R})$.*

Proof. All we need to show is that $\mathcal{C}(M,\mathbb{R}) \subseteq \overline{A}$. So we choose $f \in \mathcal{C}(M,\mathbb{R})$ and try to show that $f \in \overline{A}$. It will be enough to show that for every $\epsilon > 0$ we can find a function $h \in \overline{A}$ such that $\|f - h\|_u < \epsilon$. [Reason: then f belongs to $\overline{\overline{A}}$ (by proposition11.1.22) and therefore to $\overline{A}$ (by 10.3.2(b)).]

Let $\epsilon > 0$. For each $x \in M$ we may, according to lemma 23.2.5, choose a function $g_x \in \overline{A}$ such that $g_x(x) = f(x)$ and

$$g_x(y) < f(y) + \epsilon \tag{23.3}$$

for every $y \in M$. Since both f and g_x are continuous and they agree at x, there exists an open set U_x containing x such that

$$f(y) < g_x(y) + \epsilon \tag{23.4}$$

for every $y \in U_x$. (Why?)

Since the family $\{U_x \colon x \in M\}$ covers M and M is compact, there exist points $x_1, \dots, x_n$ in M such that $\{U_{x_1}, \dots U_{x_n}\}$ covers M. Let $h = g_{x_1} \vee \dots \vee g_{x_n}$. We know from 22.4.11 that $\overline{A}$ is a subalgebra of $\mathcal{C}(M,\mathbb{R})$. So according to corollary 23.2.2, $h \in \overline{\overline{A}} = \overline{A}$.

By inequality (23.3)

$$g_{x_k}(y) < f(y) + \epsilon$$

holds for all $y \in M$ and $1 \le k \le n$. Thus

$$h(y) < f(y) + \epsilon \tag{23.5}$$

for all $y \in M$. Each $y \in M$ belongs to at least one of the open sets U_{x_k}. Thus by (23.4)

$$f(y) < g_{x_k}(y) + \epsilon < h(y) + \epsilon \tag{23.6}$$

for every $y \in M$. Together (23.5) and (23.6) show that

$$-\epsilon < f(y) - h(y) < \epsilon$$

for all $y \in M$. That is, $\|f - h\|_u < \epsilon$. $\square$

23.2.7. Problem. Give the missing reason for inequality (23.4) in the preceding proof.

23.2.8. Theorem (Weierstrass Approximation Theorem)**.** *Every continuous real valued function on $[a, b]$ can be uniformly approximated by polynomials.*

Proof. Problem.

23.2.9. Problem. Let $\mathcal{G}$ be the set of all functions f in $\mathcal{C}\big([0,1]\big)$ such that f is differentiable on $(0,1)$ and $f\,'(\frac{1}{2}) = 0$. Show that $\mathcal{G}$ is dense in $\mathcal{C}\big([0,1]\big)$.

23.2.10. Proposition. *If M is a closed and bounded subset of $\mathbb{R}$, then the normed linear space $\mathcal{C}(M,\mathbb{R})$ is separable.*

Proof. Problem.

23.3. BANACH SPACES

23.3.1. Definition. A Banach space is a normed linear space which is complete with respect to the metric induced by its norm.

23.3.2. Example. In example 18.2.2 we saw that $\mathbb{R}$ is complete; so it is a Banach space.

23.3.3. Example. In example 18.2.11 it was shown that the space $\mathbb{R}^n$ is complete with respect to the Euclidean norm. Since all norms on $\mathbb{R}^n$ are equivalent (problem 22.3.21) and since completeness of a space is not affected by changing to an equivalent metric (proposition 18.2.10), we conclude that $\mathbb{R}^n$ is a Banach space under all possible norms.

23.3.4. Example. If S is a nonempty set, then (under the uniform norm) $\mathcal{B}(S,\mathbb{R})$ is a Banach space (see example 18.2.12).

23.3.5. Example. If M is a compact metric space, then (under the uniform norm) $\mathcal{C}(M,\mathbb{R})$ is a Banach space (see example 18.2.13).

At the beginning of section 22.3 we listed 4 ways of making new normed linear spaces from old ones. Under what circumstances do these new spaces turn out to be Banach spaces?

(i) A closed vector subspace of a Banach space is a Banach space. (See proposition 18.2.8.)
(ii) The product of two Banach space is a Banach space. (See proposition 18.2.9.)
(iii) If S is a nonempty set and E is a Banach space, then $\mathcal{B}(S, E)$ is a Banach space. (See problem 23.3.7.)
(iv) If V is a normed linear space and F is a Banach space, then $\mathfrak{B}(V, F)$ is a Banach space. (See the following proposition.)

23.3.6. Proposition. *Let V and W be normed linear spaces. Then $\mathfrak{B}(V,W)$ is complete if W is.*

Proof. Exercise. *Hint.* Show that if (T_n) is a Cauchy sequence in $\mathfrak{B}(V,W)$, then $\lim\limits_{n\to\infty} T_n x$ exists for every x in V. Let $Sx = \lim\limits_{n\to\infty} T_n x$. Show that the map $S\colon x \mapsto Sx$ is linear. If $\epsilon > 0$ then for m and n sufficiently large $\|T_m - T_n\| < \frac{1}{2}\epsilon$. For such m and n show that

$$\|Sx - T_n x\| \le \|Sx - T_m x\| + \tfrac{1}{2}\epsilon\|x\|$$

and conclude from this that S is bounded and that $T_n \to S$ in $\mathfrak{B}(V,W)$. (Solution Q.23.9.)

23.3.7. Problem. Let S be a nonempty set and V be a normed linear space. If V is complete, so is $\mathcal{B}(S, V)$. *Hint.* Make suitable modifications in the proof of example 18.2.12.

23.3.8. Problem. Let M be a compact metric space and V be a normed linear space. If V is complete so is $\mathcal{C}(M, V)$. *Hint.* Use proposition 22.4.12.

23.3.9. Problem. Let m be the set of all bounded sequences of real numbers. (That is, a sequence $(x_1, x_2, \dots)$ of real numbers belongs to m provided that there exists a constant $M > 0$ such that $|x_k| \le M$ for all $k \in \mathbb{N}$.) If $x = (x_k)$ and $y = (y_k)$ belong to m and α is a scalar define

$$x + y = (x_1 + y_1, x_2 + y_2, \dots)$$

and

$$\alpha x = (\alpha x_1, \alpha x_2, \dots)\,.$$

Also, for each sequence x in m define

$$\|x\|_u = \sup\{|x_k| \colon k \in \mathbb{N}\}\,.$$

(a) Show that m is a Banach space. *Hint.* The proof should be very short. This is a special case of a previous example.
(b) The space m is not separable.
(c) The closed unit ball of the space m is closed and bounded but not compact.

23.4. DUAL SPACES AND ADJOINTS

In this section U, V, and W are normed linear spaces.

Particularly important among the spaces of bounded linear transformations introduced in section 23.1 are those consisting of maps from a space V into its scalar field $\mathbb{R}$. The space $\mathfrak{B}(V, \mathbb{R})$ is the DUAL SPACE of V and is usually denoted by V^*. Members of V^* are called BOUNDED LINEAR FUNCTIONALS. (When the word "function" is given the suffix "al" we understand that the function is scalar valued.)

23.4.1. Example. Let $f \colon \mathbb{R}^3 \to \mathbb{R} \colon (x, y, z) \mapsto x + y + z$. Then f is a member of $(\mathbb{R}^3)^*$. (It is obviously linear and is bounded by proposition 23.1.18.)

23.4.2. Example. The familiar Riemann integral of beginning calculus is a bounded linear functional on the space $\mathcal{C} = \mathcal{C}([a, b], \mathbb{R})$. That is, the functional J defined in exercise 23.1.12 belongs to $\mathcal{C}^*$.

23.4.3. Example. Let M be a compact metric space and $\mathcal{C} = \mathcal{C}(M, \mathbb{R})$. For each x in M the evaluation functional $E_x \colon f \mapsto f(x)$ belongs to $\mathcal{C}^*$. (Take $W = \mathbb{R}$ in problem 23.1.20.)

23.4.4. Definition. Let $T \in \mathfrak{B}(V, W)$. Define T^* by $T^*(g) = g \circ T$ for every g in W^*. The map T^* is the ADJOINT of T. In the next two propositions and problem 23.4.7 we state only the most elementary properties of the adjoint map $T \mapsto T^*$. We will see more of it later.

23.4.5. Proposition. *If $T \in \mathfrak{B}(V, W)$, then T^* maps W^* into V^*. Furthermore, if $g \in W^*$, then $\|T^* g\| \le \|T\|\,\|g\|$.*

PROOF. Exercise. (Solution Q.23.10.)

23.4.6. Proposition. *Let $S \in \mathfrak{B}(U, V)$, $T \in \mathfrak{B}(V, W)$, and I_V be the identity map on V. Then*

(a) *T^* belongs to $\mathfrak{B}(W^*, V^*)$ and $\|T^*\| \le \|T\|$;*
(b) *$(I_V)^*$ is I_{V^*}, the identity map on V^*; and*
(c) *$(TS)^* = S^* T^*$.*

PROOF. Problem.

23.4.7. Problem. The adjoint map $T \mapsto T^*$ from $\mathfrak{B}(V, W)$ into $\mathfrak{B}(W^*, V^*)$ is itself a bounded linear transformation, and it has norm not exceeding 1.

CHAPTER 24

THE CAUCHY INTEGRAL

In this chapter we develop a theory of integration for vector valued functions. In one way our integral, the Cauchy integral, will be more general and in another way slightly less general than the classical Riemann integral, which is presented in beginning calculus. The Riemann integral is problematic: the derivation of its properties is considerably more complicated than the corresponding derivation for the Cauchy integral. But very little added generality is obtained as a reward for the extra work. On the other hand the Riemann integral is not nearly general enough for advanced work in analysis; there the full power of the Lebesgue integral is needed. Before starting our discussion of integration we derive some standard facts concerning uniform continuity.

24.1. UNIFORM CONTINUITY

24.1.1. Definition. A function $f\colon M_1 \to M_2$ between metric spaces is UNIFORMLY CONTINUOUS if for every $\epsilon > 0$ there exists $\delta > 0$ such that $d(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$.

Compare the definitions of "continuity" and "uniform continuity". A function $f\colon M_1 \to M_2$ is continuous if

$$\forall a \in M_1\, \forall \epsilon > 0\, \exists \delta > 0\, \forall x \in M_1\, d(x,a) < \delta \implies d(f(x), f(a)) < \epsilon\,.$$

We may just as well write this reversing the order of the first two (universal) quantifiers.

$$\forall \epsilon > 0\, \forall a \in M_1\, \exists \delta > 0\, \forall x \in M_1\, d(x,a) < \delta \implies d(f(x), f(a)) < \epsilon\,.$$

The function f is uniformly continuous if

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall a \in M_1\, \forall x \in M_1\, d(x,a) < \delta \implies d(f(x), f(a)) < \epsilon\,.$$

Thus the difference between continuity and uniform continuity is the order of two quantifiers. This makes the following result obvious.

24.1.2. Proposition. *Every uniformly continuous function between metric spaces is continuous.*

24.1.3. Example. The function $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto 3x - 4$ is uniformly continuous.

PROOF. Given $\epsilon > 0$, choose $\delta = \epsilon/3$. If $|x - y| < \delta$, then $|f(x) - f(y)| = |(3x-4) - (3y-4)| = 3|x - y| < 3\delta = \epsilon$. □

24.1.4. Example. The function $f\colon [1, \infty) \to \mathbb{R}\colon x \mapsto x^{-1}$ is uniformly continuous.

PROOF. Exercise. (Solution Q.24.1.)

24.1.5. Example. The function $g\colon (0, 1] \to \mathbb{R}\colon x \mapsto x^{-1}$ is not uniformly continuous.

PROOF. Exercise. (Solution Q.24.2.)

24.1.6. Problem. Let M be an arbitrary positive number. The function

$$f\colon [0, M] \to \mathbb{R}\colon x \mapsto x^2$$

is uniformly continuous. Prove this assertion using only the definition of "uniform continuity".

24.1.7. Problem. The function

$$g\colon [0, \infty) \to \mathbb{R}\colon x \mapsto x^2$$

is *not* uniformly continuous.

24.1.8. Proposition. *Norms on vector spaces are always uniformly continuous.*

Proof. Problem.

We have already seen in proposition 24.1.2 that uniform continuity implies continuity. Example 24.1.5 shows that the converse is not true in general. There are however two special (and important!) cases where the concepts coincide. One is linear maps between normed linear spaces, and the other is functions defined on compact metric spaces.

24.1.9. Proposition. *A linear transformation between normed linear spaces is continuous if and only if it is uniformly continuous.*

Proof. Problem.

Of course, the preceding result does not hold in general metric spaces (where "linearity" makes no sense). The next proposition, for which we give a preparatory lemma, is a metric space result.

24.1.10. Lemma. *Let (x_n) and (y_n) be sequences in a compact metric space. If $d(x_n, y_n) \to 0$ as $n \to \infty$, then there exist convergent subsequences of (x_n) and (y_n) which have the same limit.*

Proof. Exercise. (Solution Q.24.3.)

24.1.11. Proposition. *Let M_1 be a compact metric space and M_2 be an arbitrary metric space. Every continuous function $f\colon M_1 \to M_2$ is uniformly continuous.*

Proof. Exercise. (Solution Q.24.4.)

In section 24.3 of this chapter, where we define the Cauchy integral, an important step in the development is the extension of the integral from an exceedingly simple class of functions, the step functions, to a class of functions large enough to contain all the continuous functions. The two basic ingredients of this extension are the density of the step functions in the large class and the uniform continuity of the integral. Theorem 24.1.15 below is the crucial result which allows this extension. First, two preliminary results.

24.1.12. Proposition. *If $f\colon M_1 \to M_2$ is a uniformly continuous map between two metric spaces and (x_n) is a Cauchy sequence in M_1, then $\big(f(x_n)\big)$ is a Cauchy sequence in M_2.*

Proof. Problem.

24.1.13. Problem. Show by example that proposition 24.1.12 is no longer true if the word "uniformly" is deleted .

24.1.14. Lemma. *Let M_1 and M_2 be metric spaces, $S \subseteq M_1$, and $f\colon S \to M_2$ be uniformly continuous. If two sequences (x_n) and (y_n) in S converge to the same limit in M_1 and if the sequence $\big(f(x_n)\big)$ converges, then the sequence $\big(f(y_n)\big)$ converges and $\lim f(x_n) = \lim f(y_n)$.*

Proof. Exercise. *Hint.* Consider the "interlaced" sequence

$$(x_1, y_1, x_2, y_2, x_3, y_3, \dots) .$$

(Solution Q.24.5.)

We are now in a position to show that a uniformly continuous map f from a subset of a metric space into a complete metric space can be extended in a unique fashion to a continuous function on the closure of the domain of f.

24.1.15. Theorem. *Let M_1 and M_2 be metric spaces, S a subset of M_1, and $f\colon S \to M_2$. If f is uniformly continuous and M_2 is complete, then there exists a unique continuous extension of f to $\overline{S}$. Furthermore, this extension is uniformly continuous.*

Proof. Problem. *Hint.* Define $g\colon \overline{S} \to M_2$ by $g(a) = \lim f(x_n)$ where (x_n) is a sequence in S converging to a. First show that g is well defined. To this end you must show that

(i) $\lim f(x_n)$ does exist, and
(ii) the value assigned to g at a does not depend on the particular sequence (x_n) chosen. That is, if $x_n \to a$ and $y_n \to a$, then $\lim f(x_n) = \lim f(y_n)$.

Next show that g is an extension of f.

To establish the uniform continuity of g, let a and b be points in $\overline{S}$. If (x_n) is a sequence in S converging to a, then $f(x_n) \to g(a)$. This implies that both $d(x_j, a)$ and $d\big(f(x_j), g(a)\big)$ can be made as small as we please by choosing j sufficiently large. A similar remark holds for a sequence (y_n) in S which converges to b. From this show that x_j is arbitrarily close to y_k (for large j and k) provided we assume that a is sufficiently close to b. Use this in turn to show that $g(a)$ is arbitrarily close to $g(b)$ when a and b are sufficiently close.

The uniqueness argument is very easy.

24.1.16. Proposition. *Let $f\colon M \to N$ be a continuous bijection between metric spaces. If M is complete and f^{-1} is uniformly continuous, then N is complete.*

Proof. Problem.

24.2. THE INTEGRAL OF STEP FUNCTIONS

Throughout this section E is a Banach space.

24.2.1. Definition. An $(n+1)$-tuple $(t_0, t_1, \dots, t_n)$ of real numbers is a PARTITION of the interval $[a, b]$ in $\mathbb{R}$ provided that

(i) $t_0 = a$,
(ii) $t_n = b$, and
(iii) $t_{k-1} < t_k$ for $1 \le k \le n$.

If $P = (s_0, \dots, s_m)$ and $Q = (t_0, \dots, t_n)$ are partitions of the same interval $[a, b]$ and if

$$\{s_0, \dots, s_m\} \subseteq \{t_0, \dots, t_n\}$$

then we say that Q is a REFINEMENT of P and we write $P \preceq Q$.

Let $P = (s_0, \dots, s_m)$, $Q = (t_0, \dots, t_n)$, and $R = (u_0, \dots, u_p)$ be partitions of $[a, b] \subseteq \mathbb{R}$. If

$$\{u_0, \dots, u_p\} = \{s_0, \dots, s_m\} \cup \{t_0, \dots, t_n\}$$

then R is the SMALLEST COMMON REFINEMENT of P and Q and is denoted by $P \vee Q$. It is clear that $P \vee Q$ is the partition with fewest points which is a refinement of both P and Q.

24.2.2. Exercise. Consider partitions $P = \big(0, \frac14, \frac13, \frac12, \frac23, \frac34, 1\big)$ and $Q = \big(0, \frac15, \frac13, \frac23, \frac56, 1\big)$ of $[0, 1]$. Find $P \vee Q$. (Solution Q.24.6.)

24.2.3. Definition. Let S be a set and A be a subset of S. We define $\chi_A\colon S \to \mathbb{R}$, the CHARACTERISTIC FUNCTION of A, by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \in A^c. \end{cases}$$

If E is a Banach space, then a function $\sigma\colon [a, b] \to E$ is an E VALUED STEP FUNCTION on the interval $[a, b]$ if

(i) $\operatorname{ran}\sigma$ is finite and
(ii) for every $x \in \operatorname{ran}\sigma$ the set $\sigma^{\leftarrow}(\{x\})$ is the union of finitely many subintervals of $[a, b]$.

We denote by $\mathcal{S}([a, b], E)$ the family of all E valued step functions defined on $[a, b]$. Notice that $\mathcal{S}([a, b], E)$ is a vector subspace of $\mathcal{B}([a, b], E)$.

It is not difficult to see that $\sigma\colon [a,b] \to E$ is a step function if and only if there exists a partition $(t_0,\dots,t_n)$ of $[a,b]$ such that σ is constant on each of the open subintervals (t_{k-1},t_k). If, in addition, we insist that σ be discontinuous at each of the points $t_1,\dots,t_{n-1}$, then this partition is unique. Thus we speak of the PARTITION ASSOCIATED WITH (or INDUCED BY) a step function σ.

24.2.4. Notation. Let σ be a step function on $[a,b]$ and $P=(t_0,\dots,t_n)$ be a partition of $[a,b]$ which is a refinement of the partition associated with σ. We define

$$\sigma_P = (x_1,\dots,x_n)$$

where x_k is the value of σ on the open interval (t_{k-1},t_k) for $1 \le k \le n$.

24.2.5. Exercise. Define $\sigma\colon [0,5] \to \mathbb{R}$ by

$$\sigma = \chi_{[1,4]} - \chi_{(2,5]} - \chi_{\{4\}} - 2\chi_{[2,3)} - \chi_{[1,2)} + \chi_{[4,5)}\,.$$

(a) Find the partition P associated with σ.
(b) Find σ_Q where $Q = (0,1,2,3,4,5)$.

(Solution Q.24.7.)

24.2.6. Definition. Let σ be an E valued step function on $[a,b]$ and $(t_0,\dots,t_n)$ be the partition associated with σ. For $1 \le k \le n$ let $\Delta t_k = t_k - t_{k-1}$ and x_k be the value of σ on the subinterval (t_{k-1},t_k). Define

$$\int \sigma := \sum_{k=1}^{n} (\Delta t_k) x_k\,.$$

The vector $\int \sigma$ is the INTEGRAL of σ over $[a,b]$. Other standard notations for $\int \sigma$ are $\int_a^b \sigma$ and $\int_a^b \sigma(t)\,dt$.

24.2.7. Exercise. Find $\int_0^5 \sigma$ where σ is the step function given in exercise 24.2.5. (Solution Q.24.8.)

24.2.8. Problem. Let $\sigma\colon [0,10] \to \mathbb{R}$ be defined by

$$\sigma = 2\chi_{[1,5)} - 3\chi_{[2,8)} - 5\chi_{\{6\}} + \chi_{[4,10]} + 4\chi_{[9,10]}\,.$$

(a) Find the partition associated with σ.
(b) If $Q = (0,1,2,3,4,5,6,7,8,9,10)$, what is σ_Q?
(c) Find $\int \sigma$.

The next lemma is essentially obvious, but it is good practice to write out a proof anyway. It says that in computing the integral of a step function σ it doesn't matter whether we work with the partition induced by σ or with a refinement of that partition.

24.2.9. Lemma. *Let σ be an E valued step function on $[a,b]$. If $Q = (u_0,\dots,u_m)$ is a refinement of the partition associated with σ and if $\sigma_Q = (y_1,\dots,y_m)$, then*

$$\int \sigma = \sum_{k=1}^{n} (\Delta u_k) y_k\,.$$

PROOF. Exercise. (Solution Q.24.9.)

It follows easily from the preceding lemma that changing the value of a step function at a finite number of points does not affect the value of its integral. Next we show that for a given interval the integral is a bounded linear transformation on the family of all E valued step functions on the interval.

24.2.10. Proposition. *The map*

$$\int : \mathcal{S}([a,b],E) \to E$$

is bounded and linear with $\|\int\| = b - a$. *Furthermore,*

$$\left\|\int \sigma(t)\,dt\right\| \le \int \|\sigma(t)\|\,dt$$

for every E *valued step function* σ *on* $[a,b]$.

PROOF. Problem. *Hint.* To show that $\int(\sigma+\tau) = \int\sigma + \int\tau$, let P and Q be the partitions associated with σ and τ, respectively. Define the partition $R = (t_0,\dots,t_n)$ to be $P \vee Q$. Clearly R is a refinement of the partition associated with $\sigma+\tau$. Suppose $\sigma_R = (x_1,\dots,x_n)$ and $\tau_R = (y_1,\dots,y_n)$. Use lemma 24.2.9 to compute $\int(\sigma+\tau)$. To find $\|\int\|$ use the definition of the norm of a linear map and lemma 23.1.6.

We now show that in the case of real valued step functions the integral is a positive linear functional; that is, it takes positive functions to positive numbers.

24.2.11. Proposition. *If* σ *is a real valued step function on* $[a,b]$ *and if* $\sigma(t) \ge 0$ *for all* t *in* $[a,b]$, *then* $\int \sigma \ge 0$.

PROOF. Problem.

24.2.12. Corollary. *If* σ *and* τ *are real valued step functions on* $[a,b]$ *and if* $\sigma(t) \le \tau(t)$ *for all* t *in* $[a,b]$, *then* $\int\sigma \le \int\tau$.

PROOF. Apply the preceding proposition to $\tau-\sigma$. Then (by 24.2.10) $\int\tau - \int\sigma = \int(\tau-\sigma) \ge 0$. □

Finally we prepare the ground for piecing together and integrating two functions on adjoining intervals.

24.2.13. Proposition. *Let* c *be an interior point of the interval* $[a,b]$. *If* τ *and* ρ *are* E *valued step functions on the intervals* $[a,c]$ *and* $[c,b]$, *respectively, define a function* $\sigma\colon [a,b] \to E$ *by*

$$\sigma(t) = \begin{cases} \tau(t), & \text{if } a \le t \le c \\ \rho(t), & \text{if } c < t \le b. \end{cases}$$

Then σ *is an* E *valued step function on* $[a,b]$ *and*

$$\int_a^b \sigma = \int_a^c \tau + \int_c^b \rho\,.$$

PROOF. Exercise. (Solution Q.24.10.)

Notice that if σ is a step function on $[a,b]$, then $\tau := \sigma|_{[a,c]}$ and $\rho := \sigma|_{[c,b]}$ are step functions and by the preceding proposition

$$\int_a^b \sigma = \int_a^c \tau + \int_c^b \rho\,.$$

In this context one seldom distinguishes notationally between a function on an interval and the restriction of that function to a subinterval. Thus (24.2) is usually written

$$\int_a^b \sigma = \int_a^c \sigma + \int_c^b \sigma.$$

24.2.14. Problem. Let $\sigma\colon [a,b] \to E$ be a step function and $T\colon E \to F$ be a bounded linear transformation from E into another Banach space F. Then $T \circ \sigma$ is an F-valued step function on $[a,b]$ and

$$\int (T \circ \sigma) = T\Big(\int \sigma\Big)\,.$$

24.3. THE CAUCHY INTEGRAL

We are now ready to extend the integral from the rather limited family of step functions to a class of functions large enough to contain all continuous functions (in fact, all piecewise continuous functions).

Following Dieudonné[3] we will call members of this larger class *regulated functions*.

In this section E will be a Banach space, and a and b real numbers with $a < b$.

24.3.1. Definition. Recall that the family $\mathcal{S} = \mathcal{S}([a, b], E)$ of E valued step functions on $[a, b]$ is a subspace of the normed linear space $\mathcal{B} = \mathcal{B}([a, b], E)$ of bounded E-valued functions on $[a, b]$. The closure $\overline{\mathcal{S}}$ of $\mathcal{S}$ in $\mathcal{B}$ is the family of REGULATED FUNCTIONS on $[a, b]$.

It is an interesting fact that the regulated functions on an interval turn out to be exactly those functions which have one-sided limits at every point of the interval. We will not need this fact, but a proof may be found in Dieudonné[3].

According to problem22.3.9(b) the set $\overline{\mathcal{S}}$ is a vector subspace of the Banach space $\mathcal{B}$. Since it is closed in $\mathcal{B}$ it is itself a Banach space. It is on this Banach space that we define the Cauchy integral. The Cauchy integral is not as general as the Riemann integral because the set $\overline{\mathcal{S}}$ is not quite as large as the set of functions on which the Riemann integral is defined. We do not prove this; nor do we prove the fact that when both integrals are defined, they agree. What we are interested in proving is that every continuous function is regulated; that is, every continuous function on $[a, b]$ belongs to $\mathcal{S}$.

24.3.2. Proposition. *Every continuous E-valued function on $[a, b]$ is regulated.*

PROOF. Exercise. *Hint.* Use proposition 24.1.11. (Solution Q.24.11.)

It is not difficult to modify the preceding proof to show that every piecewise continuous function on $[a, b]$ is regulated. (Definition: A function $f : [a, b] \to E$ is PIECEWISE CONTINUOUS if there exists a partition $(t_0, \dots, t_n)$ of $[a, b]$ such that f is continuous on each subinterval (t_{k-1}, t_k).)

24.3.3. Corollary. *Every continuous E valued function on $[a, b]$ is the uniform limit of a sequence of step functions.*

PROOF. According to problem 12.2.5 a function f belongs to the closure of S if and only if there is a sequence of step functions which converges (uniformly) to f. □

Now we are ready to define the Cauchy integral of a regulated function.

24.3.4. Definition. Recall that $\mathcal{S} = \mathcal{S}([a, b], E)$ is a subset of the Banach space $\mathcal{B}([a, b], E)$. In the preceding section we defined the integral of a step function. The map

$$\int : \mathcal{S} \to E$$

was shown to be bounded and linear [proposition 24.2.10]; therefore it is uniformly continuous [proposition 24.1.9]. Thus [theorem 24.1.15] it has a unique continuous extension to $\overline{\mathcal{S}}$. This extension, which we denote also by $\int$, and which is, in fact, uniformly continuous, is the CAUCHY (or CAUCHY-BOCHNER) INTEGRAL. For f in $\overline{\mathcal{S}}$ we call $\int f$ the INTEGRAL OF f (over $[a, b]$). As with step functions we may wish to emphasize the domain of f or the role of a particular variable, in which case we may write $\int_a^b f$ or $\int_a^b f(t)\,dt$ for $\int f$.

24.3.5. Problem. Use the *definition* of the Cauchy integral to show that $\int_0^1 x^2\,dx = 1/3$. *Hint.* Start by finding a sequence of step functions which converges uniformly to the function $x \mapsto x^2$. The *proof* of proposition 24.3.2 may help; so also may problem I.1.15.

Most of the properties of the Cauchy integral are derived from the corresponding properties of the integral of step functions by taking limits. In the remainder of this section it is well to keep in mind one aspect of theorem 24.1.15 (and its proof): When a uniformly continuous function f is extended from a set $\mathcal{S}$ to a function g on its closure $\overline{\mathcal{S}}$, the value of g at a point a in $\overline{\mathcal{S}}$ is the limit of the values $f(x_n)$ where (x_n) is any sequence in $\mathcal{S}$ which converges to a. What does this say in the present context? If h is a regulated function, then there exists a sequence (σ_n) of step functions converging uniformly to h and furthermore

$$\int h = \lim_{n\to\infty} \int \sigma_n .$$

We use this fact repeatedly without explicit reference.

One more simple fact is worthy of notice. Following lemma 24.2.9 we remarked that changing the value of a step function at finitely many points does not affect the value of its integral. The same is true of regulated functions. [Proof. Certainly it suffices to show that changing the value of a regulated function f at a single point c does not alter the value of $\int f$. Suppose (σ_n) is a sequence of step functions converging uniformly to f. Also suppose that g differs from f only at c. Replace each step function σ_n by a function τ_n which is equal to σ_n at each point other than c and whose value at c is $g(c)$. Then $\tau_n \to g$ (unif) and (by the comment in the preceding paragraph) $\int g = \lim \int \tau_n = \lim \int \sigma_n = \int f$.]

The following theorem (and proposition 24.3.10) generalize proposition 24.2.10.

24.3.6. Theorem. *The Cauchy integral is a bounded linear transformation which maps the space of E valued regulated functions on $[a, b]$ into the space E.*

Proof. Exercise. (Solution Q.24.12.)

Next we show that for real valued functions the Cauchy integral is a positive linear functional. This generalizes proposition 24.2.11.

24.3.7. Proposition. *Let f be a regulated real valued function on $[a, b]$. If $f(t) \geq 0$ for all t in $[a, b]$, then $\int f \geq 0$.*

Proof. Problem. *Hint.* Suppose that $f(t) \geq 0$ for all t in $[a, b]$ and that (σ_n) is a sequence of real valued step functions converging uniformly to f. For each $n \in N$ define

$$\sigma_n^+ \colon [a, b] \to \mathbb{R} \colon t \mapsto \max\{\sigma_n(t), 0\} .$$

Show that (σ_n^+) is a sequence of step functions converging uniformly to f. Then use proposition 24.2.11.

24.3.8. Corollary. *If f and g are regulated real valued functions on $[a, b]$ and if $f(t) \leq g(t)$ for all t in $[a, b]$, then $\int f \leq \int g$.*

Proof. Problem.

24.3.9. Proposition. *Let $f \colon [a, b] \to E$ be a regulated function and (σ_n) be a sequence of E-valued step functions converging uniformly to f. Then*

$$\left\| \int f \right\| = \lim \left\| \int \sigma_n \right\| .$$

Furthermore, if $g(t) = \|f(t)\|$ for all t in $[a, b]$ and $\tau_n(t) = \|\sigma_n(t)\|$ for all n in $\mathbb{N}$ and t in $[a, b]$, then (τ_n) is a sequence of real valued step functions which converges uniformly to g and $\int g = \lim \int \tau_n$.

Proof. Problem.

24.3.10. Proposition. *Let f be a regulated E valued function on $[a, b]$. Then*

$$\left\| \int f \right\| \leq \int \|f(t)\| \, dt$$

and therefore

$$\left\| \int f \right\| \le (b-a)\|f\|_u \,.$$

Thus

$$\left\| \int \right\| = b-a$$

where $\int : \overline{\mathcal{S}} \to E$ *is the Cauchy integral and* $\overline{\mathcal{S}}$ *is the family of regulated* E *valued functions on the interval* $[a,b]$*.*

Proof. Problem.

24.3.11. Problem. Explain in one or two brief sentences why the following is obvious: If (f_n) is a sequence of E-valued regulated functions on $[a,b]$ which converges uniformly to g, then g is regulated and $\int g = \lim \int f_n$.

24.3.12. Problem. Let $\sigma, \tau \colon [0,3] \to \mathbb{R}$ be the step functions defined by

$$\sigma = \chi_{[0,2]} \qquad \text{and} \qquad \tau = \chi_{[0,2]} + 2\chi_{(2,3]} \,.$$

Recall from appendix N that the function $(\sigma,\tau)\colon [0,3] \to \mathbb{R}^2$ is defined by $(\sigma,\tau)(t) = \big(\sigma(t),\tau(t)\big)$ for $0 \le t \le 3$. It is clear that (σ,τ) is a step function.

(a) Find the partition R associated with (σ,τ). Find $(\sigma,\tau)_R$. Make a careful sketch of (σ,τ).
(b) Find $\int(\sigma,\tau)$.

24.3.13. Problem. Same as preceding problem except let $\sigma = \chi_{[0,1]}$.

We now generalize proposition 24.2.13 to regulated functions on adjoining intervals.

24.3.14. Proposition. *Let* c *be an interior point of the interval* $[a,b]$*. If* g *and* h *are regulated* E*-valued functions on the intervals* $[a,c]$ *and* $[c,b]$*, respectively, define a function* $f\colon [a,b] \to E$ *by*

$$f(t) = \begin{cases} g(t), & \text{if } a \le c \le c \\ h(t), & \text{if } c < t \le b. \end{cases}$$

Then f *is a regulated function on* $[a,b]$ *and*

$$\int_a^b f = \int_a^c g + \int_c^b h \,.$$

Proof. Problem.

As was remarked after proposition 24.2.13 it is the usual practice to use the same name for a function and for its restriction to subintervals. Thus the notation of the next corollary.

24.3.15. Corollary. *If* $f\colon [a,b] \to E$ *is a regulated function and* c *is an interior point of the interval* $[a,b]$*, then*

$$\int_a^b f = \int_a^c f + \int_c^b f \,. \tag{24.1}$$

Proof. Let $g = f|_{[a,c]}$ and $h = f|_{[c,b]}$. Notice that g and h are regulated. [If, for example, (σ_n) is a sequence of step functions converging uniformly on $[a,b]$ to f, then the step functions $\sigma_n|_{[a,c]}$ converge uniformly on $[a,c]$ to g.] Then apply the preceding proposition. □

It is convenient for formula (24.1) to be correct when a, b, and c are not in increasing order or, for that matter, even necessarily distinct. This can be achieved by means of a simple notational convention.

24.3.16. Definition. For a regulated function f on $[a,b]$ (where $a<b$) define

$$\int_b^a f := -\int_a^b f\,.$$

Furthermore, if g is any function whose domain contains the point a, then

$$\int_a^a g := 0\,.$$

24.3.17. Corollary. *If f is an E valued regulated function whose domain contains an interval to which the points a, b, and c belong, then*

$$\int_a^b f = \int_a^c f + \int_c^b f\,.$$

Proof. The result is obvious if any two of the points a, b, and c coincide; so we suppose that they are distinct. There are six possible orderings. We check one of these. Suppose $c<a<b$. By corollary 24.3.15

$$\int_c^b f = \int_c^a f + \int_a^b f\,.$$

Thus

$$\int_a^b f = -\int_c^a f + \int_c^b f = \int_a^c f + \int_c^b f\,.$$

The remaining five cases are similar. □

Suppose f is an E-valued regulated function and T is a bounded linear map from E into another Banach space F. It is interesting and useful to know that integration of the composite function $T\circ f$ can be achieved simply by integrating f and then applying T. This fact can be expressed by means of the following commutative diagram. Notation: $\overline{\mathcal{S}}_E$ and $\overline{\mathcal{S}}_F$ denote, respectively, the E valued and F valued regulated functions on $[a,b]$; and C_T is the bounded linear transformation discussed in problem 23.1.21. ($C_T(f) = T\circ f$ for every f.)

$$\begin{tikzcd}
\overline{\mathcal{S}}_E \arrow[r, "\int"] \arrow[d, "C_T"'] & E \arrow[d, "T"] \\
\overline{\mathcal{S}}_F \arrow[r, "\int"'] & F
\end{tikzcd}$$

Alternatively it may be expressed by a formula, as in the next proposition.

24.3.18. Proposition. *Let $T\colon E\to F$ be a bounded linear map between Banach spaces and f be a regulated E valued function on the interval $[a,b]$. Then $T\circ f$ is a regulated F valued function on $[a,b]$ and*

$$\int (T\circ f) = T\Big(\int f\Big)\,.$$

Proof. Exercise. *Hint.* Use problem 24.2.14. (Solution Q.24.13.)

24.3.19. Corollary. *If E and F are Banach spaces and*

$$T\colon [a,b]\to \mathfrak{B}(E,F)\colon t\mapsto T_t$$

is continuous, then for every x in E

$$\int T_t(x)\,dt = \Big(\int T\Big)(x)\,.$$

PROOF. Problem. *Hint.* For x in E let $E_x : \mathfrak{B}(E,F) \to F$ be the map (evaluation at x) defined in problem 23.1.19. Write $T_t(x)$ as $(E_x \circ T)(t)$ and apply proposition 24.3.18.

24.3.20. Proposition. *Let $f : [a,b] \to \mathbb{R}$ be a regulated function and $x \in E$. For all t in $[a,b]$ let $g(t) = f(t)\,x$. Then g is a regulated E-valued function and $\int g = (\int f)x$.*

PROOF. Problem. *Hint.* Prove the result first for the case f is a step function. Then take limits. □

24.3.21. Proposition. *If $f : [a,b] \to E$ and $g : [a,b] \to F$ are regulated functions whose ranges lie in Banach spaces, then the function*

$$(f,g) : [a,b] \to E \times F : t \mapsto \big(f(t), g(t)\big)$$

is regulated and

$$\int (f,g) = \left(\int f, \int g\right).$$

PROOF. Problem. *Hint.* Write $\int f$ as $\int \big(\pi_1 \circ (f,g)\big)$ where $\pi_1 : E \times F \to E$ is the usual coordinate projection. Write $\int g$ in a similar fashion. Use proposition 24.3.18. Keep in mind that if p is a point in the product $E \times F$, then $p = \big(\pi_1(p), \pi_2(p)\big)$.

24.3.22. Proposition. *Suppose $h : [a,b] \to E \times F$ is a regulated function from $[a,b]$ into the product of two Banach spaces. Then the components h^1 and h^2 are regulated functions and*

$$\int h = \left(\int h_1, \int h_2\right).$$

PROOF. Problem.

24.3.23. Problem. Suppose $f : [a,b] \to \mathbb{R}^n$ is a regulated function. Express the integral of f in terms of the integrals of its components $f^1, \dots, f^n$. Justify your answer carefully. *Hint.* When $\mathbb{R}^n$ appears (without further qualification) its norm is assumed to be the usual Euclidean norm. This is not the product norm. What needs to be done to ensure that the results of the preceding problem will continue to be true if the product norm on $E \times F$ is replaced by an equivalent norm?

24.3.24. Problem. Suppose $T : [0,1] \to \mathfrak{B}(\mathbb{R}^2, \mathbb{R}^2) : t \mapsto T_t)$ is continuous, and suppose that for each t in $[0,1]$ the matrix representation of T_t is given by

$$[T_t] = \begin{bmatrix} 1 & t \\ t^2 & t^3 \end{bmatrix}.$$

Find $\left[\int T\right]$ the matrix representation of $\int T$. *Hint.* Use 24.3.19, 24.3.20, and 24.3.23.

24.3.25. Problem. For every f in $\mathcal{C}([a,b], E)$ define

$$\|f\|_1 = \int_a^b \|f(t)\|\,dt.$$

(a) Show that $\|\quad\|_1$ is a norm on $\mathcal{C}([a,b], E)$. *Hint.* In showing that $\|f\|_1 = 0$ implies $f = 0$, proposition 3.3.21 may help.
(b) Are the norms $\|\quad\|_1$ and $\|\quad\|_u$ on $\mathcal{C}([a,b], E)$ equivalent?
(c) Let $\mathcal{C}_1$ be the vector space $\mathcal{C}([a,b], E)$ under the norm $\|\quad\|_1$ and $\mathcal{C}_u$ be the same vector space under its usual uniform norm. Does convergence of a sequence in $\mathcal{C}_u$ imply convergence in C_1? What about the converse?

24.3.26. Problem. Show that if $f \in \mathcal{C}([0,1], \mathbb{R})$ and $\int_0^1 x^n f(x)\,dx = 0$ for $n = 0, 1, 2, \dots$, then $f = 0$.

24.3.27. Problem. Find $\lim \frac{\int_0^1 x^n f(x)\,dx}{\int_0^1 x^n\,dx}$ when f is a continuous real valued function on $[0, 1]$. *Hint.* For each n in $\mathbb{N}$ let $L_n(f) = \frac{\int_0^1 x^n f(x)\,dx}{\int_0^1 x^n\,dx}$. Show that L_n is a continuous linear functional on the space $\mathcal{C}([0, 1], \mathbb{R})$ of continuous real valued functions on $[0, 1]$. What is $\lim_{n\to\infty} L_n(p)$ when p is a polynomial? Use the *Weierstrass approximation theorem* 23.2.8.

CHAPTER 25

DIFFERENTIAL CALCULUS

In chapter 8 we studied the calculus of real valued functions of a real variable. We now extend that study to vector valued functions of a vector variable. The first three sections of this chapter repeat almost exactly the material in chapter 8. Absolute values are replaced by norms, and linear maps by continuous linear maps. Other than that there are few differences. In fact, if you have done chapter 8 carefully you may wish just to glance over the first three sections of this chapter and move on to section 25.4.

Throughout this chapter V, W, and X (with or without subscripts) will be normed linear spaces.

25.1. $\mathfrak{O}$ AND $\mathfrak{o}$ FUNCTIONS

25.1.1. Notation. Let $a \in V$. We denote by $\mathcal{F}_a(V,W)$ the family of all functions defined on a neighborhood of a taking values in W. That is, f belongs to $\mathcal{F}_a(V,W)$ if there exists a set U such that $a \in U \overset{\circ}{\subseteq} \operatorname{dom} f \subseteq V$ and if the image of f is contained in W. We shorten $\mathcal{F}_a(V,W)$ to $\mathcal{F}_a$ when no confusion will result. Notice that for each $a \in V$, the set $\mathcal{F}_a$ is closed under addition and scalar multiplication. (As usual, we define the sum of two functions f and g in $\mathcal{F}_a$ to be the function $f+g$ whose value at x is $f(x)+g(x)$ whenever x belongs to $\operatorname{dom} f \cap \operatorname{dom} g$.) Despite the closure of $\mathcal{F}_a$ under these operations, $\mathcal{F}_a$ is *not* a vector space.

25.1.2. Problem. Let $a \in V \neq \{\mathbf{0}\}$. Prove that $\mathcal{F}_a(V,W)$ is not a vector space.

Among the functions defined on a neighborhood of the zero vector in V are two subfamilies of crucial importance; they are $\mathfrak{O}(V,W)$ (the family of *"big-oh"* functions) and $\mathfrak{o}(V,W)$ (the family of *"little-oh"* functions).

25.1.3. Definition. A function f in $\mathcal{F}_{\mathbf{0}}(V,W)$ belongs to $\mathfrak{O}(V,W)$ if there exist numbers $c>0$ and $\delta>0$ such that

$$\|f(x)\| \le c\,\|x\|$$

whenever $\|x\| < \delta$.

A function f in $\mathcal{F}_{\mathbf{0}}(V,W)$ belongs to $\mathfrak{o}(V,W)$ if for every $c>0$ there exists $\delta>0$ such that

$$\|f(x)\| \le c\,\|x\|$$

whenever $\|x\| < \delta$. Notice that f belongs to $\mathfrak{o}(V,W)$ if and only if $f(\mathbf{0}) = \mathbf{0}$ and

$$\lim_{h\to\mathbf{0}} \frac{\|f(h)\|}{\|h\|} = \mathbf{0}\,.$$

When no confusion seems likely we will shorten $\mathfrak{O}(V,W)$ to $\mathfrak{O}$ and $\mathfrak{o}(V,W)$ to $\mathfrak{o}$.

The properties of the families $\mathfrak{O}$ and $\mathfrak{o}$ are given in propositions 25.1.4–25.1.9, 25.1.11, and 25.1.12.

25.1.4. Proposition. *Every member of $\mathfrak{o}(V,W)$ belongs to $\mathfrak{O}(V,W)$; so does every member of $\mathfrak{B}(V,W)$. Every member of $\mathfrak{O}(V,W)$ is continuous at $\mathbf{0}$.*

Proof. Obvious from the definitions. $\square$

25.1.5. Proposition. *Other than the zero transformation, no bounded linear transformation belongs to* $\mathfrak{o}$.

Proof. Exercise. (Solution Q.25.1.)

25.1.6. Proposition. *The family* $\mathfrak{O}$ *is closed under addition and scalar multiplication.*

Proof. Exercise. (Solution Q.25.2.)

25.1.7. Proposition. *The family* $\mathfrak{o}$ *is closed under addition and scalar multiplication.*

Proof. Problem.

The next two propositions say that the composite of a function in $\mathfrak{O}$ with one in $\mathfrak{o}$ (in either order) ends up in $\mathfrak{o}$.

25.1.8. Proposition. *If* $g \in \mathfrak{O}(V, W)$ *and* $f \in \mathfrak{o}(W, X)$*, then* $f \circ g \in \mathfrak{o}(V, X)$.

Proof. Problem.

25.1.9. Proposition. *If* $g \in \mathfrak{o}(V, W)$ *and* $f \in \mathfrak{O}(W, X)$*, then* $f \circ g \in \mathfrak{o}(V, X)$.

Proof. Exercise. (Solution Q.25.3.)

25.1.10. Notation. In the sequel it will be convenient to multiply vectors not only by scalars but also by scalar valued functions. If ϕ is a function in $\mathcal{F}_a(V, \mathbb{R})$ and $w \in W$, we define the function ϕw by

$$(\phi w)(x) = \phi(x) \cdot w$$

for all x in the domain of ϕ. Clearly, ϕw belongs to $\mathcal{F}_a(V, W)$.

Similarly, it is useful to multiply vector valued functions by scalar valued functions. If $\phi \in \mathcal{F}_a(V, \mathbb{R})$ and $f \in \mathcal{F}_a(V, W)$, we define the function ϕf by

$$(\phi f)(x) = \phi(x) \cdot f(x)$$

for all x in $\operatorname{dom} \phi \cap \operatorname{dom} f$. Then ϕf belongs to $\mathcal{F}_a(V, W)$.

25.1.11. Proposition. *If* $\phi \in \mathfrak{o}(V, \mathbb{R})$ *and* $w \in W$*, then* $\phi w \in \mathfrak{o}(V, W)$.

Proof. Exercise. (Solution Q.25.4.)

25.1.12. Proposition. *If* $\phi \in \mathfrak{O}(V, \mathbb{R})$ *and* $f \in \mathfrak{O}(V, W)$*, then* $\phi f \in \mathfrak{o}(V, W)$.

Proof. Exercise. (Solution Q.25.5.)

Remark. The list summarizing these facts is almost the same as the one in 8.1. In (1) and (2) the linear maps $\mathfrak{L}$ have been replaced by continuous linear maps $\mathfrak{B}$; and (7) is new. (As before, $\mathcal{C}_\mathbf{0}$ is the set of all functions in $\mathcal{F}_\mathbf{0}$ which are continuous at $\mathbf{0}$.)

$$\begin{array}{ll}
(1) & \mathfrak{B} \cup \mathfrak{o} \subseteq \mathfrak{O} \subseteq \mathcal{C}_\mathbf{0}\,. \\
(2) & \mathfrak{B} \cap \mathfrak{o} = \{\mathbf{0}\}\,. \\
(3) & \mathfrak{O} + \mathfrak{O} \subseteq \mathfrak{O}\,; \qquad \alpha\,\mathfrak{O} \subseteq \mathfrak{O}\,. \\
(4) & \mathfrak{o} + \mathfrak{o} \subseteq \mathfrak{o}\,; \qquad \alpha\,\mathfrak{o} \subseteq \mathfrak{o}\,. \\
(5) & \mathfrak{o} \circ \mathfrak{O} \subseteq \mathfrak{o}\,. \\
(6) & \mathfrak{O} \circ \mathfrak{o} \subseteq \mathfrak{o}\,. \\
(7) & \mathfrak{o}(V, \mathbb{R}) \cdot W \subseteq \mathfrak{o}(V, W)\,. \\
(8) & \mathfrak{O}(V, \mathbb{R}) \cdot \mathfrak{O}(V, W) \subseteq \mathfrak{o}(V, W)\,.
\end{array}$$

25.1.13. Problem. Find a function in $\mathcal{C}_\mathbf{0}(\mathbb{R}^2, \mathbb{R}) \setminus \mathfrak{O}(\mathbb{R}^2, \mathbb{R})$. Also, find a function in $\mathfrak{O}(\mathbb{R}^2, \mathbb{R}) \setminus \mathfrak{o}(\mathbb{R}^2, \mathbb{R})$.

25.1.14. Problem. Show that $\mathfrak{O} \circ \mathfrak{O} \subseteq \mathfrak{O}$. That is, if $g \in \mathfrak{O}(V, W)$ and $f \in \mathfrak{O}(W, X)$, then $f \circ g \in \mathfrak{O}(V, X)$. (As usual, the domain of $f \circ g$ is taken to be $\{x \in V : g(x) \in \operatorname{dom} f\}$.)

25.1.15. Problem. If $f^1 \in \mathfrak{O}(V_1, W)$ and $f^2 \in \mathfrak{O}(V_2, W)$, then the function g defined on $\operatorname{dom} f^1 \times \operatorname{dom} f^2$ by

$$g(v^1, v^2) = f^1(v^1) + f^2(v^2)$$

belongs to $\mathfrak{O}(V_1 \times V_2, W)$. *Hint.* The simplest proof never mentions the domain elements v^1 and v^2. Instead write g in terms of the projection maps π_1 and π_2 on $V_1 \times V_2$, and apply problem 25.1.14.

25.1.16. Problem. If $\phi \in \mathfrak{O}(V_1, \mathbb{R})$ and $f \in \mathfrak{O}(V_2, W)$, then the function h defined on $\operatorname{dom} \phi \times \operatorname{dom} f$ by

$$h(v^1, v^2) = \phi(v^1)\, f(v^2)$$

belongs to $\mathfrak{o}(V_1 \times V_2, W)$. (Use the hint given in the preceding problem.)

25.1.17. Problem. Show that the membership of the families $\mathfrak{O}(V, W)$ and $\mathfrak{o}(V, W)$ is not changed when the norms on the spaces V and W are replaced by equivalent norms.

25.2. TANGENCY

25.2.1. Definition. Two functions f and g in $\mathcal{F}_{\mathbf{0}}(V, W)$ are TANGENT AT ZERO, in which case we write $f \simeq g$, if $f - g \in \mathfrak{o}(V, W)$.

25.2.2. Proposition. *The relation "tangency at zero" is an equivalence relation on $\mathcal{F}_{\mathbf{0}}$.*

PROOF. Exercise. (Solution Q.25.6.)

The next result shows that at most one bounded linear transformation can be tangent at zero to a given function.

25.2.3. Proposition. *Let $S, T \in \mathfrak{B}$ and $f \in \mathcal{F}_0$. If $S \simeq f$ and $T \simeq f$, then $S = T$.*

PROOF. Exercise. (Solution Q.25.7.)

25.2.4. Proposition. *If $f \simeq g$ and $j \simeq k$, then $f + j \simeq g + k$, and furthermore, $\alpha f \simeq \alpha g$ for all $\alpha \in \mathbb{R}$.*

PROOF. Problem.

In the next proposition we see that if two real valued functions are tangent at zero, multiplication by a vector does not disrupt this relationship. (For notation see 25.1.10.)

25.2.5. Proposition. *Let $\phi, \psi \in \mathcal{F}_0(V, \mathbb{R})$ and $w \in W$. If $\phi \simeq \psi$, then $\phi w \simeq \psi w$.*

PROOF. Exercise. (Solution Q.25.8.)

25.2.6. Proposition. *Let $f, g \in \mathcal{F}_{\mathbf{0}}(V, W)$ and $T \in \mathfrak{B}(W, X)$. If $f \simeq g$, then $T \circ f \simeq T \circ g$.*

PROOF. Problem.

25.2.7. Proposition. *Let $h \in \mathfrak{O}(V, W)$ and $f, g \in \mathcal{F}_0(W, X)$. If $f \simeq g$, then $f \circ h \simeq g \circ h$.*

PROOF. Problem.

25.2.8. Problem. Fix a vector x in V. Define a function $M_x : \mathbb{R} \to V$ by

$$M_x(\alpha) = \alpha x \qquad \text{for all } \alpha \in \mathbb{R}.$$

Notice that $M_x \in \mathfrak{B}(\mathbb{R}, V)$. If $f \in \mathcal{F}_0(\mathbb{R}, V)$ and $f \simeq M_x$, then

(a) $\dfrac{f(\alpha)}{\alpha} \to x$ as $\alpha \to 0$; and
(b) $f(\alpha) \to 0$ as $\alpha \to 0$.

25.2.9. Problem. Each of the following is an abbreviated version of a proposition. Formulate precisely and prove.

(a) $\mathcal{C}_{\mathbf{0}} + \mathfrak{O} \subseteq \mathcal{C}_{\mathbf{0}}$.
(b) $\mathcal{C}_{\mathbf{0}} + \mathfrak{o} \subseteq \mathcal{C}_{\mathbf{0}}$.
(c) $\mathfrak{O} + \mathfrak{o} \subseteq \mathfrak{O}$.

25.2.10. Problem. Suppose that $f \simeq g$. Then the following hold.

(a) If g is continuous at $\mathbf{0}$, so is f.
(b) If g belongs to $\mathfrak{O}$, so does f.
(c) If g belongs to $\mathfrak{o}$, so does f.

25.3. DIFFERENTIATION

25.3.1. Definition. Let $f \in \mathcal{F}_a(V, W)$. Define the function Δf_a by

$$\Delta f_a(h) \equiv f(a + h) - f(a)$$

for all h such that $a + h$ is in the domain of f. Notice that since f is defined in a neighborhood of a, the function Δf_a is defined in a neighborhood of $\mathbf{0}$; that is, Δf_a belongs to $\mathcal{F}_{\mathbf{0}}(V, W)$. Notice also that $\Delta f_a(\mathbf{0}) = \mathbf{0}$.

25.3.2. Proposition. *If $f \in \mathcal{F}_a(V, W)$ and $\alpha \in \mathbb{R}$, then*

$$\Delta(\alpha f)_a = \alpha\, \Delta f_a\,.$$

Proof. No changes need to be made in the proof of 8.3.3. □

25.3.3. Proposition. *If $f,\ g \in \mathcal{F}_a(V, W)$, then*

$$\Delta(f + g)_a = \Delta f_a + \Delta g_a\,.$$

Proof. The proof given (in the *solutions to exercises*) for proposition 8.3.4 needs no alteration. □

25.3.4. Proposition. *If $\phi \in \mathcal{F}_a(V, \mathbb{R})$ and $f \in \mathcal{F}_a(V, W)$, then*

$$\Delta(\phi f)_a = \phi(a) \cdot \Delta f_a \,+\, \Delta\phi_a \cdot f(a) \,+\, \Delta\phi_a \cdot \Delta f_a\,.$$

Proof. Problem.

25.3.5. Proposition. *If $f \in \mathcal{F}_a(V, W)$, $g \in \mathcal{F}_{f(a)}(W, X)$, and $g \circ f \in \mathcal{F}_a(V, X)$, then*

$$\Delta(g \circ f)_a = \Delta g_{f(a)} \circ \Delta f_a\,.$$

Proof. Use the proof given (in the *solutions to exercises*) for proposition 8.3.6. □

25.3.6. Proposition. *A function $f\colon V \to W$ is continuous at the point a in V if and only if Δf_a is continuous at $\mathbf{0}$.*

Proof. Problem.

25.3.7. Proposition. *If $f\colon U \to U_1$ is a bijection between subsets of arbitrary vector spaces, then for each a in U the function $\Delta f_a\colon U - a \to U_1 - f(a)$ is invertible and*

$$\big(\Delta f_a\big)^{-1} = \Delta\big(f^{-1}\big)_{f(a)}\,.$$

Proof. Problem.

25.3.8. Definition. Let $f \in \mathcal{F}_a(V, W)$. We say that f is differentiable at a if there exists a bounded linear map which is tangent at $\mathbf{0}$ to Δf_a. If such a map exists, it is called the differential of f at a and is denoted by df_a. Thus df_a is just a member of $\mathfrak{B}(V, W)$ which satisfies $df_a \simeq \Delta f_a$. We denote by $\mathcal{D}_a(V, W)$ the family of all functions in $\mathcal{F}_a(V, W)$ which are differentiable at a. We often shorten this to $\mathcal{D}_a$.

25.3.9. Proposition. *Let $f \in \mathcal{F}_a(V, W)$. If f is differentiable at a, then its differential is unique. (That is, there is at most one bounded linear map tangent at 0 to Δf_a.)*

Proof. Proposition 25.2.3. $\square$

Remark. If f is a function in $\mathcal{F}_a(V, W)$ which is differentiable at a, its differential df_a has three important properties:

(i) it is linear;
(ii) it is continuous (that is, bounded as a linear map);
(iii) $\displaystyle\lim_{h \to \mathbf{0}} \frac{\Delta f_a(h) - df_a(h)}{\|h\|} = \mathbf{0}$.

(An expression of the form $\dfrac{\Delta f_a(h) - df_a(h)}{\|h\|}$ is called a Newton quotient.)

25.3.10. Exercise. Let $f(x, y) = 3x^2 - xy + 4y^2$. Show that $df_{(1,-1)}(x, y) = 7x - 9y$. Interpret $df_{(1,-1)}$ geometrically. (Solution Q.25.9.)

25.3.11. Exercise. Let

$$f \colon \mathbb{R}^3 \to \mathbb{R}^2 \colon (x, y, z) \mapsto (x^2y - 7, 3xz + 4y)$$

and $a = (1, -1, 0)$. Use the *definition* of "differential" to find df_a. *Hint.* Work with the matrix representation of df_a. Since the differential must belong to $\mathfrak{B}(\mathbb{R}^3, \mathbb{R}^2)$, its matrix representation is a 2×3 matrix $M = \begin{bmatrix} r & s & t \\ u & v & w \end{bmatrix}$. Use the requirement—condition (iii) of the preceding remark—that $\|h\|^{-1} \|\Delta f_a(h) - Mh\| \to \mathbf{0}$ as $h \to \mathbf{0}$ to discover the identity of the entries in M. (Solution Q.25.10.)

25.3.12. Proposition. *If $f \in \mathcal{D}_a$, then $\Delta f_a \in \mathfrak{O}$.*

Proof. Use the proof given (in the *solutions to exercises*) for proposition 8.4.8. Since we are working with bounded linear maps between normed spaces instead of linear maps on $\mathbb{R}$, we must change $\mathfrak{L}$ to $\mathfrak{B}$. $\square$

25.3.13. Corollary. *Every function which is differentiable at a point is continuous there.*

Proof. Use the proof given (in the *solutions to exercises*) for corollary 8.4.9. $\square$

25.3.14. Proposition. *If f is differentiable at a and $\alpha \in \mathbb{R}$, then αf is differentiable at a and*

$$d(\alpha f)_a = \alpha \, df_a .$$

Proof. The proof given (in the *solutions to exercises*) for proposition 8.4.10 works, but three references need to be changed. (What are the correct references?) $\square$

25.3.15. Proposition. *If f and g are differentiable at a, then $f + g$ is differentiable at a and*

$$d(f + g)_a = df_a + dg_a .$$

Proof. Problem.

25.3.16. Proposition (Leibniz's Rule)**.** *If $\phi \in \mathcal{D}_a(V, \mathbb{R})$ and $f \in \mathcal{D}_a(V, W)$, then $\phi f \in \mathcal{D}_a(V, W)$ and*

$$d(\phi f)_a = d\phi_a \cdot f(a) \, + \, \phi(a) \, df_a .$$

Proof. Exercise. (Solution Q.25.11.)

In the chapters of beginning calculus texts devoted to the differential calculus of several variables, the expression "chain rule" refers most frequently to a potentially infinite collection of related results concerning the differentiation of composite functions. Several examples such as the following are usually given:

Let $w = f(x, y, z)$, $x = x(u, v)$, $y = y(u, v)$, and $z = z(u, v)$. If these functions are all differentiable, then the function

$$w = f(x(u, v)\,,\, y(u, v)\,,\, z(u, v))$$

is differentiable and

$$\begin{aligned}\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u}\\ \frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial v}\end{aligned} \tag{$*$}$$

Then the reader is encouraged to invent new "chain rules" for functions having different numbers of variables. Formulations such as ($*$) have many shortcomings the most serious of which is a nearly complete lack of discernible geometric content. The version of the *chain rule* which we will prove says that (after a suitable translation) the best linear approximation to the composite of two functions is the composite of the best linear approximations. In other words, the differential of the composite is the composite of the differentials. Notice that theorem 25.3.17, where this is stated formally, is simpler than ($*$); it has obvious geometric content; it is a single "rule" rather than a family of them; and it holds in arbitrary infinite dimensional normed linear spaces as well as in finite dimensional ones.

25.3.17. Proposition (The Chain Rule.)**.** *If $f \in \mathcal{D}_a(V, W)$ and $g \in \mathcal{D}_{f(a)}(W, X)$, then $g \circ f \in \mathcal{D}_a(V, X)$ and*

$$d(g \circ f)_a = dg_{f(a)} \circ df_a\,.$$

Proof. Exercise. (Solution Q.25.12.)

Each of the preceding propositions concerning differentiation is a direct consequence of a similar result concerning the map Δ. In particular, the linearity of the map $f \mapsto df_a$ (propositions 25.3.14 and 25.3.15) follows from the fact that the function $f \mapsto \Delta f_a$ is linear (propositions 25.3.2 and 25.3.3); *Leibniz's rule* 25.3.16 is a consequence of proposition 25.3.4; and the proof of the *chain rule* 25.3.17 makes use of proposition 25.3.5. It is reasonable to hope that the result given in proposition 25.3.7 concerning $\Delta\big(f^{-1}\big)_{f(a)}$ will lead to useful information concerning the differential of the inverse function. This is indeed the case; but obtaining information about $d\big(f^{-1}\big)_{f(a)}$ from 25.3.7 is a rather involved undertaking. It turns out, happily enough, that (under mild hypotheses) the differential of f^{-1} is the inverse of the differential of f. This result is known as the *inverse function theorem*, and the whole of chapter 29 is devoted to detailing its proof and to examining some of its consequences.

25.3.18. Problem. Show that proposition 25.3.14 is actually a special case of *Leibniz's rule* 25.3.16. Also, suppose that $\phi \in \mathcal{D}_a(V, \mathbb{R})$ and $w \in W$. Prove that $\phi w \in \mathcal{D}_a(V, W)$ and that $d(\phi w)_a = d\phi_a \cdot w$.

25.3.19. Problem. Let $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f(x) = (3x_1 - x_2 + 7, x_1 + 4x_2)$. Show that $\big[df_{(1,0)}\big] = \begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix}$. *Hint.* Let $M = \begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix}$ and $a = (1, 0)$. Show that the Newton quotient $\dfrac{\Delta f_a(h) - Mh}{\|h\|}$ approaches $\mathbf{0}$ as $h \to \mathbf{0}$. Use the uniqueness of differentials (proposition 25.3.9) to conclude that $\big[df_a\big] = M$.

25.3.20. Problem. Let $F \colon \mathbb{R}^2 \to \mathbb{R}^4$ be defined by

$$F(x) = (x_2\,,\, {x_1}^2\,,\, 4 - x_1x_2\,,\, 7x_1)$$

and let $a = (1, 1)$. Use the *definition* of "differential" to show that

$$\big[dF_a\big] = \begin{bmatrix} 0 & 1 \\ 2 & 0 \\ -1 & -1 \\ 7 & 0 \end{bmatrix}.$$

25.3.21. Problem. Let $F\colon \mathbb{R}^3 \to \mathbb{R}^4\colon x \mapsto (x_1+2x_3, x_2-x_3, 4x_2, 2x_1-5x_2)$ and $a = (1, 2, -5)$. Use the *definition* of "differential" to find $\big[dF_a\big]$. *Hint.* Use the technique suggested in exercise 25.3.11.

25.3.22. Problem. Let $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $F(x, y, z) = (xy-3, y+2z^2)$. Use the *definition* of "differential" to find $\big[dF_{(1,-1,2)}\big]$.

25.3.23. Problem. Let $F\colon \mathbb{R}^3 \to \mathbb{R}^3 : x \mapsto (x_1x_2, x_2 - x_3{}^2, 2x_1x_3)$. Use the *definition* of "differential" to find $\big[dF_a\big]$ at a in $\mathbb{R}^3$.

25.3.24. Problem. Let $T \in \mathfrak{B}(V, W)$ and $a \in V$. Find dT_a.

25.4. DIFFERENTIATION OF CURVES

In the preceding section we have discussed the differentiability of functions mapping one normed linear space into another. Here we briefly consider the important and interesting special case which occurs when the domains of the functions are one dimensional.

25.4.1. Definition. A CURVE in a normed linear space V is a continuous mapping from an interval in $\mathbb{R}$ into V. If $c\colon J \to V$ is a curve, if 0 belongs to the interval J, and if $c(0) = a \in V$, then c is a CURVE AT a.

In classical terminology a curve $c\colon J \to V : t \mapsto c(t)$ is usually referred to as a PARAMETRIZED CURVE in V. The interval J is the PARAMETER INTERVAL and the variable t belonging to J is the PARAMETER. If you start with a subset A of V and find a continuous function c from an interval J onto A, then c is called a PARAMETRIZATION of A.

25.4.2. Example. Let

$$c_1\colon [0, 2\pi] \to \mathbb{R}^2\colon t \mapsto (\cos t, \sin t)$$

and

$$c_2\colon [0, 3\pi] \to \mathbb{R}^2\colon t \mapsto (\sin(2t+1), \cos(2t+1)).$$

Then c_1 and c_2 are two different parametrizations of the unit circle

$$\mathbb{S}^1 := \{(x, y) \in \mathbb{R}^2\colon x^2 + y^2 = 1\}.$$

Parameters need have no physical significance, but it is quite common to think of a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$ as representing the motion of a particle in that space: the parameter t is taken to be time and the range value $c(t)$ is the position of the particle at time t. With this interpretation we may view c_1 and c_2 as representing the motion of particles traveling around the unit circle. In the first case the particle starts (when $t = 0$) at the point $(1, 0)$ and makes one complete trip around the circle traveling counterclockwise. In the second case, the particle starts at the point $(\sin 1, \cos 1)$ and traverses $\mathbb{S}^1$ three times moving clockwise.

25.4.3. Example. Let a and $u \neq \mathbf{0}$ be vectors in V. The curve

$$\ell\colon \mathbb{R} \to V : t \mapsto a + tu$$

is the PARAMETRIZED LINE through a in the direction of u. Of course, infinitely many other parametrizations of the range of ℓ are possible, but this is the standard one and we adopt it.

25.4.4. Problem. Find a parametrization of the unit square

$$A := \{(x, y) \in \mathbb{R}^2 \colon d_u\big((x, y)\,,\,(\tfrac{1}{2}, \tfrac{1}{2})\big) = \tfrac{1}{2}\}$$

which starts at $(0, 0)$ and traverses A once in a counterclockwise direction.

25.4.5. Definition. Let $c\colon J \to V$ be a curve in V and suppose that a is a point in the interior of the interval J. Then $Dc(a)$, the DERIVATIVE of c at a, is defined by the formula

$$Dc(a) := \lim_{h \to \mathbf{0}} \frac{\Delta c_a(h)}{h}$$

if the indicated limit exists. Notice that this is just the definition of "derivative" given in beginning calculus. The derivative at a is also called the TANGENT VECTOR to c at a or, in case we are thinking of the motion of a particle, the VELOCITY of c at a. If $Dc(a)$ exists and is not zero, then the parametrized line through $c(a)$ in the direction of $Dc(a)$ is the TANGENT LINE to the image of c at the point $c(a)$.

25.4.6. Exercise. Let $c\colon [0, 2\pi] \to \mathbb{R}^2 \colon t \mapsto (\cos t, \sin t)$.

(a) Find the tangent vector to c at $t = \frac{\pi}{3}$.
(b) Find the tangent line to the range of c at the point $(\frac{1}{2}, \frac{\sqrt{3}}{2})$.
(c) Write the equation of the range in $\mathbb{R}^2$ of the tangent line found in (b).

(Solution Q.25.13.)

25.4.7. Proposition. *If a curve $c\colon J \to V$ is differentiable at a point a in the interior of the interval J, then it has a derivative at a and*

$$Dc(a) = dc_a(1)\,.$$

PROOF. Exercise. *Hint.* Start with the Newton quotient $\dfrac{\Delta c_a(h)}{h}$. Subtract and add $\dfrac{dc_a(h)}{h}$.) (Solution Q.25.14.)

The converse of the preceding proposition is also true. Every curve possessing a derivative is differentiable. This is our next proposition.

25.4.8. Proposition. *If a curve $c\colon J \to V$ has a derivative at a point a in J°, then it is differentiable at a and*

$$dc_a(h) = h\,Dc(a) \qquad \textit{for all } h \in \mathbb{R}.$$

PROOF. Problem. *Hint.* Define $T\colon \mathbb{R} \to V \colon h \mapsto h\,Dc(a)$. Show that $\Delta c_a \simeq T$.

25.4.9. Problem. Suppose that curves c_1 and c_2 in a normed linear space are defined and differentiable in some neighborhood of $a \in \mathbb{R}$. Then

(a) $D(c_1 + c_2)(a) = Dc_1(a) + Dc_2(a)$.
(b) $D(\alpha c_1)(a) = \alpha\, Dc_1(a)$ for all $\alpha \in \mathbb{R}$.

25.4.10. Problem. Let V be a normed linear space and $a \in V$.

(a) Suppose c is a differentiable curve at the zero vector in V. Then $c \simeq \mathbf{0}$ if and only if $Dc(0) = \mathbf{0}$.
(b) Suppose c_1 and c_2 are differentiable curves at a. Then $c_1 \simeq c_2$ if and only if $Dc_1(0) = Dc_2(0)$.

25.4.11. Proposition. *If $c \in \mathcal{D}_t(\mathbb{R}, V)$ and $f \in \mathcal{D}_a(V, W)$, where $a = c(t)$, then $f \circ c \in \mathcal{D}_t(\mathbb{R}, W)$ and*

$$D(f \circ c)(t) = df_a(Dc(t))\,.$$

PROOF. Problem.

Thus far integration and differentiation have been treated as if they belong to separate worlds. In the next theorem, known as the *fundamental theorem of calculus*, we derive the most important link between these two topics.

25.4.12. Theorem (Fundamental Theorem Of Calculus)**.** *Let a belong to an open interval J in the real line, E be a Banach space, and $f\colon J \to E$ be a regulated curve. Define $F(x) = \int_a^x f$ for all $x \in J$. If f is continuous at $c \in J$, then F is differentiable at c and $DF(c) = f(c)$.*

Proof. Exercise. *Hint.* For every $\epsilon > 0$ there exists $\delta > 0$ such that $c+h \in J$ and $\|\Delta f_c(h)\| < \epsilon$ whenever $|h| < \delta$. (Why?) Use the (obvious) fact that $h\, f(c) = \int_c^{c+h} f(c)\, dt$ to show that $\|\Delta F_c(h) - h\, f(c)\| < \epsilon\, |h|$ whenever $0 < |h| < \delta$. From this conclude that $\lim\limits_{h\to 0} \dfrac{1}{h}\Delta F_c(h) = f(c)$. (Solution Q.25.15.)

25.5. DIRECTIONAL DERIVATIVES

We now return to the study of maps between arbitrary normed linear spaces. Closely related to differentiability is the concept of *directional derivative*, an examination of which provides some technical advantages and also throws light on the geometric aspect of differentiation.

25.5.1. Definition. Let f be a member of $\mathcal{F}_a(V, W)$ and v be a nonzero vector in V. Then $D_v f(a)$, the DERIVATIVE OF f AT a IN THE DIRECTION OF v, is defined by

$$D_v f(a) := \lim_{t\to 0} \frac{1}{t}\, \Delta f_a(tv)$$

if this limit exists. This directional derivative is also called the GÂTEAUX DIFFERENTIAL (or GÂTEAUX VARIATION) of f, and is sometimes denoted by $\delta f(a; v)$. Many authors require that in the preceding definition v be a unit vector. We will *not* adopt this convention.

Recall that for $\mathbf{0} \neq v \in V$ the curve $\ell\colon \mathbb{R} \to V$ defined by $\ell(t) = a + tv$ is the parametrized line through a in the direction of v. In the following proposition, which helps illuminate our use of the adjective "directional", we understand the domain of $f \circ \ell$ to be the set of all numbers t for which the expression $f(\ell(t))$ makes sense; that is,

$$\operatorname{dom}(f \circ \ell) = \{t \in \mathbb{R} \colon a + tv \in \operatorname{dom} f\}\,.$$

Since a is an interior point of the domain of f, the domain of $f \circ \ell$ contains an open interval about 0.

25.5.2. Proposition. *If $f \in \mathcal{D}_a(V, W)$ and $\mathbf{0} \neq v \in V$, then the directional derivative $D_v f(a)$ exists and is the tangent vector to the curve $f \circ \ell$ at 0 (where ℓ is the parametrized line through a in the direction of v). That is,*

$$D_v f(a) = D(f \circ \ell)(0)\,.$$

Proof. Exercise. (Solution Q.25.16.)

25.5.3. Exercise. Let $f(x, y) = \ln\bigl(x^2 + y^2\bigr)^{\frac{1}{2}}$. Find $D_v f(a)$ when $a = (1, 1)$ and $v = (\frac{3}{5}, \frac{4}{5})$. (Solution Q.25.17.)

25.5.4. Notation. For $a < b$ let $\mathcal{C}^1([a, b], \mathbb{R})$ be the family of all functions f differentiable on some open subset of $\mathbb{R}$ containing $[a, b]$ whose derivative Df is continuous on $[a, b]$.

25.5.5. Exercise. For all f in $\mathcal{C}^1([0, \frac{\pi}{2}], \mathbb{R})$ define

$$\phi(f) = \int_0^{\frac{\pi}{2}} (\cos x + Df(x))^2\, dx\,.$$

Compute $D_v \phi(a)$ when $a(x) = 1 + \sin x$ and $v(x) = 2 - \cos x$ for $0 \le x \le \frac{\pi}{2}$. (Solution Q.25.18.)

25.5.6. Problem. Let $f(x, y) = e^{x^2+y^2}$, $a = (1, -1)$, and $v = (-1, 2)$. Find $D_v f(a)$.

25.5.7. Problem. Let $f(x,y) = (2xy, y^2 - x^2)$, $a = (-1,2)$, and $v = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Find $D_v f(a)$.

25.5.8. Problem. Let $\phi\colon \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}\colon f \mapsto \int_0^1 \left(\sin^6 \pi x + (f(x))^2\right) dx$, and for $0 \le x \le 1$ let $a(x) = e^{-x} - x + 3$ and $v(x) = e^x$. Find $D_v\phi(a)$.

According to proposition 25.5.2, differentiability implies the existence of directional derivatives in all directions. (In problem 25.5.11 you are asked to show that the converse is not true.) The next proposition makes explicit the relationship between differentials and directional derivatives for differentiable functions.

25.5.9. Proposition. *If $f \in \mathcal{D}_a(V,W)$, then for every nonzero v in V*

$$D_v f(a) = df_a(v)\,.$$

Proof. Exercise. *Hint.* Use problem 25.4.11 (Solution Q.25.19.)

It is worth noticing that even though the domain of a curve is 1-dimensional, it is still possible to take directional derivatives. The relationship between derivatives and directional derivatives is very simple: the derivative of a curve c at a point t is just the directional derivative at t in the direction of the unit vector 1 in $\mathbb{R}$. Proof:

$$\begin{aligned} Dc(t) &= dc_t(1) && \text{(by proposition 25.4.7)} \\ &= D_1 c(t) && \text{(by proposition 25.5.9)} \end{aligned}$$

25.5.10. Problem. Let $\phi\colon \mathcal{C}([0,1],\mathbb{R}) \to \mathcal{C}^1([0,1],\mathbb{R})$ be defined by $\phi f(x) = \int_0^x f(s)\,ds$ for all f in $\mathcal{C}([0,1],\mathbb{R})$ and x in $[0,1]$. For arbitrary functions a and $v \ne 0$ in $\mathcal{C}([0,1],\mathbb{R})$ compute $D_v\phi(a)$ using

(a) the definition of "directional derivative".
(b) proposition 25.5.2.
(c) proposition 25.5.9.

25.5.11. Problem. Show that the converse of proposition 25.5.2 need not hold. A function may have directional derivatives in all directions but fail to be differentiable. *Hint.* Consider the function defined by $f(x,y) = \dfrac{x^3}{x^2+y^2}$ for $(x,y) \ne (0,0)$ and $f(0,0) = 0$.

25.6. FUNCTIONS MAPPING INTO PRODUCT SPACES

In this short section we examine the differentiability of functions which map into the product of normed linear spaces. This turns out to be a very simple matter: a necessary and sufficient condition for a function whose codomain is a product to be differentiable is that each of its components be differentiable. The somewhat more complicated topic of differentiation of functions whose *domain* lies in a product of normed linear spaces is deferred until the next chapter when we have the *mean value theorem* at our disposal.

25.6.1. Proposition. *If $f^1 \in \mathcal{D}_a(V,W_1)$ and $f^2 \in \mathcal{D}_a(V,W_2)$, then the function $f = \left(f^1, f^2\right)$ belongs to $\mathcal{D}_a(V, W_1 \times W_2)$ and $df_a = \left(d(f^1)_a, d(f^2)_a\right)$.*

Proof. Exercise. *Hint.* The function f is defined by $f(x) = \left(f^1(x), f^2(x)\right)$ for all $x \in \operatorname{dom} f^1 \cap \operatorname{dom} f^2$. Check that $f = j_1 \circ f^1 + j_2 \circ f^2$ where $j_1\colon W_1 \to W_1 \times W_2\colon u \mapsto (u,0)$ and $j_2\colon W_2 \to W_1 \times W_2\colon v \mapsto (0,v)$. Use proposition 25.3.15, theorem 25.3.17, and problem 25.3.24. (Solution Q.25.20.)

The preceding proposition says that a function is differentiable if its components are. The converse, which is the next proposition, is also true: the components of a differentiable function are differentiable.

25.6.2. Proposition. *If f belongs to $\mathcal{D}_a(V, W_1 \times W_2)$, then its components f^1 and f^2 belong to $\mathcal{D}_a(V,W_1)$ and $\mathcal{D}_a(V,W_2)$, respectively, and*

$$df_a = \left(d(f^1)_a, d(f^2)_a\right).$$

PROOF. Problem. *Hint.* For $k = 1$, 2 write $f^k = \pi_k \circ f$ where, as usual, $\pi_k(x_1, x_2) = x_k$. Then use the *chain rule* 25.3.17 and problem 25.3.24.

An easy consequence of the two preceding propositions is that curves in product spaces have derivatives if and only if their components do.

25.6.3. Corollary. *Let c^1 and c^2 be curves in W_1 and W_2, respectively. If c^1 and c^2 are differentiable at t, then so is the curve $c = (c^1, c^2)$ in $W_1 \times W_2$ and $Dc(t) = \big(Dc^1(t), Dc^2(t)\big)$. Conversely, if c is a differentiable curve in the product $W_1 \times W_2$, then its components c^1 and c^2 are differentiable and $Dc = \big(Dc^1, Dc^2\big)$.*

PROOF. Exercise. (Solution Q.25.21.)

It is easy to see how to generalize the three preceding results to functions whose codomains are the products of any finite collection of normed linear spaces. Differentiation is done componentwise.

25.6.4. Example. Consider the helix

$$c\colon \mathbb{R} \to \mathbb{R}^3 \colon t \mapsto (\cos t, \sin t, t)\,.$$

Its derivative at t is found by differentiating each component separately. That is, $Dc(t) = (-\sin t, \cos t, 1)$ for all $t \in \mathbb{R}$.

25.6.5. Problem. Let $f(x, y) = (\ln(xy), y^2 - x^2)$, $a = (1, 1)$, and $v = (\frac{3}{5}, \frac{4}{5})$. Find the directional derivative $D_v f(a)$.

CHAPTER 26

PARTIAL DERIVATIVES AND ITERATED INTEGRALS

In this chapter we consider questions which arise concerning a function f whose domain is the product $V_1 \times V_2$ of normed linear spaces. What relationship (if any) exists between the differentiability of f and the differentiability of the functions $x \mapsto f(x,b)$ and $y \mapsto f(a,y)$, where a and b are fixed points in V_1 and V_2, respectively? What happens in the special case $V_1 = V_2 = \mathbb{R}$ if we integrate f first with respect to y (that is, integrate, for arbitrary x, the function $y \mapsto f(x,y)$ over the interval $[c,d]$) and then integrate with respect to x (that is, integrate the function $x \mapsto \int_c^d f(x,y)\,dy$)? Does this produce the same result as integrating first with respect to x and then with respect to y? Before we can answer these and similar questions we must develop a fundamental tool of analysis: the *mean value theorem*.

26.1. THE MEAN VALUE THEOREM(S)

Heretofore we have discussed differentiability only of functions defined on *open* subsets of normed linear spaces. It is occasionally useful to consider differentiability of functions defined on other types of subsets. The business from beginning calculus of right and left differentiability at endpoints of intervals does not extend in any very natural fashion to functions with more complicated domains in higher dimensional spaces. Recall that according to definition 12.1.1 a *neighborhood* of a point in a metric space is an open set containing that point. It will be convenient to expand slightly our use of this word.

26.1.1. Definition. A NEIGHBORHOOD of a subset A of a metric space is any open set which contains A.

26.1.2. Definition. Let V and W be normed linear spaces and $A \subseteq V$. A W valued function f is said to be DIFFERENTIABLE ON A if it is (defined and) differentiable on some neighborhood of A. The function f is CONTINUOUSLY DIFFERENTIABLE ON A if it is differentiable on (a neighborhood of) A and if its differential $df\colon x \mapsto df_x$ is continuous at each point of A. The family of all W valued continuously differentiable functions on A is denoted by $\mathcal{C}^1(A,W)$. (Keep in mind that in order for a function f to belong to $\mathcal{C}^1(A,W)$ its domain must contain a *neighborhood* of A.)

Given a function f for which both "differential" and "derivative" make sense it is natural to ask if there is any difference between requiring df to be continuous and requiring Df to be. It is the point of the next problem to show that there is not.

26.1.3. Problem. Let A be a subset of $\mathbb{R}$ with nonempty interior and W be a normed linear space. A function f mapping a neighborhood of A into W belongs to $\mathcal{C}^1(A,W)$ if and only if its derivative Df (exists and) is continuous on some neighborhood of A.

Not every function that is differentiable is continuously differentiable.

26.1.4. Example. Let $f(x) = x^2 \sin 1/x$ for $x \neq 0$ and $f(0) = 0$. Then f is differentiable on $\mathbb{R}$ but does not belong to $\mathcal{C}^1(\mathbb{R},\mathbb{R})$.

PROOF. Problem.

We have already encountered one version of the *mean value theorem* (see theorem 8.4.26): if $a < b$ and f is a real valued function continuous on the interval $[a,b]$ and differentiable on the interior of that interval, then

$$\frac{f(b)-f(a)}{b-a} = Df(c) \tag{26.1}$$

for at least one number c between a and b.

We consider the problem of generalizing this formula to *scalar fields* (that is, real valued functions of a vector variable), to *curves* (vector valued functions of a real variable), and to *vector fields* (vector valued functions of a vector variable). There is no difficulty in finding an entirely satisfactory variant of (26.1) which holds for scalar fields whose domain lies in $\mathbb{R}^n$. This is done in chapter 27 once we have the notion of *gradient* (see proposition 27.2.17). On the other hand we show in exercise 26.1.5 that for curves (*a fortiori*, vector fields) formula (26.1) does *not* hold. Nevertheless, the most useful aspect of the *mean value theorem*, that changes in f over the interval $[a,b]$ cannot exceed the maximum value of $|Df|$ multiplied by the length of the interval does have a direct generalization to vector fields (see proposition 26.1.6). A somewhat different generalization can be produced by considering the version of the *fundamental theorem of calculus* most used in beginning calculus: if f is a function whose derivative exists and is continuous on $[a,b]$, then

$$f(b)-f(a) = \int_a^b Dft\,dt. \tag{26.2}$$

(We have yet to prove this result. In fact we need the *mean value theorem* to do so. See theorem 26.1.13.) It is conventional to define $(b-a)^{-1}\int_a^b g$ to be the MEAN VALUE (or AVERAGE VALUE) of a function g over the interval $[a,b]$. Thus (26.2) may be regarded as a "mean value theorem" saying that the Newton quotient $(b-a)^{-1}(f(b)-f(a))$ is just the mean value of the *derivative* of f. Since functions between Banach spaces do not in general have "derivatives" it is better for purposes of generalization to rewrite (26.2) in terms of differentials. For a curve f with continuous derivative

$$\int_a^b Df(t)\,dt = \int_a^b df_t(1)\,dt \quad \text{(by 25.4.7)}$$

so (26.2) becomes

$$f(b)-f(a) = \int_a^b df_t(1)\,dt. \tag{26.3}$$

If we divide both sides by $b-a$ this result says, briefly, that the Newton quotient of f is the mean value of the *differential* of f over $[a,b]$.

There is one more thing to consider in seeking to generalize (26.1). In chapter 24 we defined the integral only for vector valued functions of a *real* variable. So if a and b are points in a general Banach space, the expression $\int_a^b df_t(1)\,dt$ in (26.3) is meaningless (and so, of course, is the Newton quotient $(f(b)-f(a))/(b-a)$). This is easily dealt with. In order to integrate over a closed segment $[a,b]$ in a Banach space, parametrize it: let $l(t)=(1-t)a+tb$ for $0\le t\le 1$ and let $g=f\circ l$. Then

$$\begin{aligned} dg_t(1) &= df_{l(t)}(dl_t(1)) \\ &= df_{l(t)}(b-a)\,. \end{aligned} \tag{26.4}$$

Apply (26.3) to the function g over the interval $[0,1]$ to obtain

$$g(1)-g(0) = \int_0^1 dg_t(1)\,dt\,. \tag{26.5}$$

Substituting (26.4) in (26.5) leads to

$$\begin{aligned} f(b)-f(a) &= g(1)-g(0) \\ &= \int_0^1 df_{l(t)}(b-a)\,dt. \end{aligned} \tag{26.6}$$

Notice that if we let $h = b - a$, then (26.6) may be written

$$\begin{aligned}\Delta f_a(h) &= \int_0^1 df_{l(t)}(h)\,dt \\ &= \left(\int_0^1 df_{l(t)}\,dt\right)(h)\end{aligned} \tag{26.7}$$

by corollary 24.3.19. It is in this form (either (26.6) or (26.7)) that the *mean value theorem* holds for a function f between Banach spaces which is continuously differentiable on a segment $[a, b]$. It is worth noticing that this generalization is weaker in three respects than the classical *mean value theorem*: we are not able to conclude that there exists a particular point where the differential of f is equal to $f(b) - f(a)$; we assume differentiability at a and b; and we assume continuity of the differential. Nonetheless, this will be adequate for our purposes.

26.1.5. Exercise. Show that the classical *mean value theorem* (theorem 8.4.26) fails for vector valued functions. That is, find an interval $[a, b]$, a Banach space E, and a continuous function $f\colon [a, b] \to E$ differentiable on (a, b) such that the equation $(b - a)Df(c) = f(b) - f(a)$ holds for *no* point c in (a, b). *Hint.* Consider a parametrization of the unit circle. (Solution Q.26.1.)

Here is our first generalization of the *mean value theorem*. Others will occur in 26.1.7, 26.1.8, and 26.1.14.

26.1.6. Theorem (Mean Value Theorem for Curves)**.** *Let $a < b$ and W be a normed linear space. If a continuous function $f\colon [a, b] \to W$ has a derivative at each point of (a, b) and if there is a constant M such that $\|Df(t)\| \le M$ for all $t \in (a, b)$, then*

$$\|f(b) - f(a)\| \le M(b - a)\,.$$

Proof. Exercise. *Hint.* Given $\epsilon > 0$ define $h(t) = \|f(t) - f(a)\| - (t - a)(M + \epsilon)$ for $a \le t \le b$. Let $A = h^{\leftarrow}(-\infty, \epsilon]$. Show that:

(i) A has a least upper bound, say l;
(ii) $l > a$;
(iii) $l \in A$; and
(iv) $l = b$.

To prove (iv) argue by contradiction. Assume $l < b$. Show that $\|(t - l)^{-1}(f(t) - f(l))\| < M + \epsilon$ for t sufficiently close to and greater than l. For such t show that $t \in A$ by considering the expression

$$\|f(t) - f(l)\| + \|f(l) - f(a)\| - (t - l)(M + \epsilon)\,.$$

Finally, show that the desired conclusion follows from (iii) and (iv). (Solution Q.26.2.)

Next we extend the *mean value theorem* from curves to vector fields.

26.1.7. Proposition. *Let V and W be normed linear spaces and a and h be points in V. If a W valued function f is continuously differentiable on the segment $[a, a + h]$, then*

$$\|\Delta f_a(h)\| \le M\|h\|\,.$$

whenever M is a number such that $\|df_z\| \le M$ for all z in $[a, a + h]$.

Proof. Problem. *Hint.* Use $l\colon t \mapsto a + th$ (where $0 \le t \le 1$) to parametrize the segment $[a, a + h]$. Apply 26.1.6 to the function $g = f \circ l$.

26.1.8. Corollary. *Let V and W be normed linear spaces, a and h be points of V, the operator T belong to $\mathfrak{B}(V, W)$, and g be a W valued function continuously differentiable on the segment $[a, a+h]$. If M is any number such that $\|dg_z - T\| \le M$ for all z in $[a, a + h]$, then*

$$\|\Delta g_a(h) - Th\| \le M\|h\|\,.$$

Proof. Problem. *Hint.* Apply 26.1.7 to the function $g - T$.

The next proposition is an important application of the *mean value theorem*.

26.1.9. Proposition. *Let V and W be normed linear spaces, U be a nonempty connected open subset of V, and $f\colon U \to W$ be differentiable. If $df_x = \mathbf{0}$ for every $x \in U$, then f is constant.*

Proof. Problem. *Hint.* Choose $a \in U$ and set $G = f^{\leftarrow}\{f(a)\}$. Show that G is both an open and closed subset of U. Then use 17.1.6 to conclude that $U = G$.

To prove that G is open in U, take an arbitrary point y in G, find an open ball B about y which is contained in U, and use 26.1.7 to show that $\Delta f_y(w - y) = 0$ for every w in B.

26.1.10. Corollary. *Let V, W, and U be as in the preceding proposition. If f, $g\colon U \to W$ are differentiable on U and have the same differentials at each point of U, then f and g differ by a constant.*

Proof. Problem.

26.1.11. Problem. Show that the hypothesis that U be connected cannot be deleted in proposition 26.1.9.

Corollary 26.1.10 makes possible a proof of the version of the *fundamental theorem of calculus* which is the basis for the procedures of "formal integration" taught in beginning calculus.

26.1.12. Definition. A differentiable curve f in a Banach space whose domain is an open interval in $\mathbb{R}$ is an ANTIDERIVATIVE of a function g if $Df = g$.

26.1.13. Theorem (Fundamental Theorem of Calculus - Version II)**.** *Let a and b be points in an open interval $J \subseteq \mathbb{R}$ with $a < b$. If $g\colon J \to E$ is a continuous map into a Banach space and f is an antiderivative of g on J, then*

$$\int_a^b g = f(b) - f(a)\,.$$

Proof. Problem. *Hint.* Let $h(x) = \int_a^x g$ for x in J. Use corollary 26.1.10 to show that h and f differ by a constant. Find the value of this constant by setting $x = a$.

With this second version of the *fundamental theorem of calculus* in hand, we are in a position to prove the version of the *mean value theorem* discussed in the introduction to this section.

26.1.14. Proposition. *Suppose that E and F are Banach spaces, and a, $h \in E$. If an F valued function f is continuously differentiable on the segment $[a, a + h]$, then*

$$\Delta f_a(h) = \left(\int_0^1 df_{l(t)}\,dt\right)(h) \tag{26.8}$$

where $l(t) = a + th$ for $0 \le t \le 1$.

Proof. Problem. *Hint.* Let $g = f \circ l$. Show that $Dg(t) = df_{l(t)}(h)$. Apply 24.3.19 to the right side of equation (26.8). Use 26.1.13.

It is conventional when we invoke any of the results 26.1.6, 26.1.7, 26.1.8, or 26.1.14 to say that we have used "the" *mean value theorem*.

26.1.15. Problem. Verify proposition 26.1.14 directly for the function $f(x, y) = x^2 + 6xy - 2y^2$ by computing both sides of equation (26.8).

26.1.16. Problem. Let W be a normed linear space and $a < b$. For each f in $\mathcal{C}^1([a, b], W)$ define

$$|||f||| = \sup\{\|f(x)\| + \|Df(x)\|\colon a \le x \le b\}\,.$$

(a) Show that $\mathcal{C}^1([a, b], W)$ is a vector space and that the map $f \mapsto |||f|||$ is a norm on this space.

(b) Let $j(x) = \sqrt{1+x^2}$ for all x in $\mathbb{R}$. Use the *mean value theorem* to show that $|\Delta j_x(y)| \le |y|$ for all $x,\ y \in \mathbb{R}$.

(c) Let k be a continuous real valued function on the interval $[a,b]$. For each f in $\mathcal{C}^1([a,b],\mathbb{R})$ define

$$\Phi(f) = \int_a^b k(x)\sqrt{1+(f'(x))^2}\,dx\,.$$

Show that the function $\Phi\colon \mathcal{C}^1([a,b],\mathbb{R}) \to \mathbb{R}$ is uniformly continuous.

26.1.17. Proposition (Change of Variables)**.** *If $f \in \mathcal{C}(J_2,E)$ and $g \in \mathcal{C}^1(J_1,J_2)$ where J_1 and J_2 are open intervals in $\mathbb{R}$ and E is a Banach space, and if $a,\ b \in J_1$ with $a < b$, then*

$$\int_a^b g'(f\circ g) = \int_{g(a)}^{g(b)} f\,.$$

Proof. Problem. *Hint.* Use 26.1.13. Let $F(x) = \int_{g(a)}^x f$ for all x in J_2. Compute $(F\circ g)'(t)$.

26.1.18. Proposition (Integration by Parts)**.** *If $f \in \mathcal{C}^1(J,\mathbb{R})$ and $g \in \mathcal{C}^1(J,E)$ where J is an open interval in $\mathbb{R}$ and E is a Banach space, and if a and b are points in J with $a < b$, then*

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g\,.$$

Proof. Problem. *Hint.* Differentiate the function $t \mapsto f(t)g(t)$.

26.1.19. Proposition. *If $f \in \mathcal{C}(J,\mathbb{R})$ where J is an interval, and if $a,\ b \in J$ with $a < b$, then there exists $c \in (a,b)$ such that*

$$\int_a^b f = f(c)(b-a)\,.$$

Proof. Problem.

26.1.20. Problem. Show by example that the preceding proposition is no longer true if it is assumed that f is a continuous function from J into $\mathbb{R}^2$.

26.2. PARTIAL DERIVATIVES

Suppose that $V_1, V_2, \dots, V_n$, and W are normed linear spaces. Let $V = V_1 \times \dots \times V_n$. In the following discussion we will need the CANONICAL INJECTION MAPS $j_1, \dots, j_n$ which map the coordinate spaces $V_1, \dots, V_n$ into the product space V. For $1 \le k \le n$ the map $j_k\colon V_k \to V$ is defined by $j_k(x) = (0,\dots,0,x,0,\dots,0)$ (where x appears in the k^{th} coordinate). It is an easy exercise to verify that each j_k is a bounded linear transformation and that if V_k does not consist solely of the zero vector then $\|j_k\| = 1$. Also worth noting is the relationship given in the following exercise between the injections j_k and the projections π_k.

26.2.1. Exercise. Let $V_1, \dots, V_n$ be normed linear spaces and $V = V_1 \times \dots \times V_n$. Then

(a) For $k = 1, \dots, n$ the injection j_k is a right inverse of the projection π_k.

(b) $\sum\limits_{k=1}^{n} (j_k \circ \pi_k) = I_V$.

(Solution Q.26.3.)

26.2.2. Problem. Let $V = V_1 \times \dots \times V_n$ where $V_1, \dots, V_n$ are normed linear spaces. Also let $a \in V$, $r > 0$, and $1 \le k \le n$. Then the image under the projection map π_k of the open ball in V about a of radius r is the open ball in V_k about a_k of radius r.

26.2.3. Definition. A mapping $f\colon M_1 \to M_2$ between metric spaces is OPEN [resp., CLOSED] if $f^{\to}(U)$ is an open [resp., closed] subset of M_2 whenever U is an open [resp., closed] subset of M_1.

26.2.4. Problem. Show that each projection mapping $\pi_k\colon V_1 \times \cdots \times V_n \to V_k$ on the product of normed linear spaces is an open mapping. Construct an example to show that projection mappings need not be closed.

Now suppose that f belongs to $\mathcal{F}_a(V,W)$ and that $1 \le k \le n$. Let B be an open ball about a which is contained in the domain of f and $B_k = (\pi_k^{\rightarrow}(B)) - a_k$. From problem 26.2.2 and the fact that translation by a_k is an isometry (see problem 22.3.14) we see that B_k is an open ball in V_k about the origin (whose radius is the same as the radius of B). Define

$$g\colon B_k \to W\colon x \mapsto f(a + j_k(x))\,.$$

Notice that as x changes, only the k^{th} variable of the domain of f is changing; the other $k-1$ variables are fixed. Also notice that we can write $g = f \circ T_a \circ j_k$, where T_a is translation by a (that is, $T_a\colon x \mapsto x + a$).

26.2.5. Exercise. With notation as above show that

$$\Delta g_{\mathbf{0}} = \Delta f_a \circ j_k\,.$$

in some neighborhood of the origin in V_k. (Solution Q.26.4.)

26.2.6. Proposition. *Let f, g, and a be as above. If f is differentiable at a, then g is differentiable at $\mathbf{0}$ and*

$$dg_{\mathbf{0}} = df_a \circ j_k\,.$$

Proof. Problem. *Hint.* Use exercise 26.2.5 and proposition 25.2.7.

26.2.7. Notation. Suppose that the function $g\colon B_k \to W\colon x \mapsto f(a + j_k(x))$ is differentiable at $\mathbf{0}$. Since g depends only on the function f, the point a, and the index k, it is desirable to have a notation for $dg_{\mathbf{0}}$ which does not require the use of the extraneous letter "g". A fairly common convention is to write $d_k f_a$ for $dg_{\mathbf{0}}$; this bounded linear map is the k^{th} PARTIAL DIFFERENTIAL of f at a. Thus $d_k f_a$ is the unique bounded linear map which is tangent to $\Delta f_a \circ j_k$. We restate the preceding proposition using this notation.

26.2.8. Corollary. *Let $V_1, V_2, \dots, V_n$, and W be normed linear spaces. If the function f belongs to $\mathcal{D}_a(V_1 \times \cdots \times V_n, W)$, then for each $k = 1, \dots, n$ the k^{th} partial differential of f at a exists and*

$$d_k f_a = df_a \circ j_k\,.$$

26.2.9. Corollary. *Let $V_1, V_2, \dots, V_n$, and W be normed linear spaces. If the function f belongs to $\mathcal{D}_a(V_1 \times \cdots \times V_n, W)$, then*

$$df_a(x) = \sum_{k=1}^{n} d_k f_a(x_k)$$

for each x in $V_1 \times \cdots \times V_n$.

Proof. Problem. *Hint.* Write df_a as $df_a \circ I$ (where I is the identity map on $V_1 \times \cdots \times V_n$) and use exercise 26.2.1(b).

The preceding corollaries assure us that if f is differentiable at a point, then it has partial differentials at that point. The converse is not true. (Consider the function defined on $\mathbb{R}^2$ whose value is 1 everywhere except on the coordinate axes, where its value is 0. This function has partial differentials at the origin, but is certainly not differentiable—or even continuous—there.) The following proposition shows that if we assume continuity (as well as the existence) of the partial differentials of a function in an open set, then the function is differentiable (in fact, continuously differentiable) on that set. To avoid complicated notation we prove this only for the product of two spaces.

26.2.10. Proposition. *Let V_1, V_2, and W be normed linear spaces, and let $f: U \to W$ where $U \overset{\circ}{\subseteq} V_1 \times V_2$. If the partial differentials $d_1 f$ and $d_2 f$ exist and are continuous on U, then f is continuously differentiable on U and*

$$df_{(a,b)} = d_1 f_{(a,b)} \circ \pi_1 + d_2 f_{(a,b)} \circ \pi_2 \tag{26.9}$$

at each point (a, b) in U.

Proof. Exercise. *Hint.* To show that f is differentiable at a point (a, b) in U and that its differential there is the expression on the right side of (26.9), we must establish that $\Delta f_{(a,b)}$ is tangent to $R = S \circ \pi_1 + T \circ \pi_2$ where $S = d_1 f_{(a,b)}$ and $T = d_2 f_{(a,b)}$. Let $g: t \mapsto f(a, b+t)$ for all t such that $(a, b+t) \in U$ and $h^t: s \mapsto f(a+s, b+t)$ for all s and t such that $(a+s, b+t) \in U$. Show that $\Delta f_{(a,b)}(s,t)$ is the sum of $\Delta \left(h^t\right)_{\mathbf{0}}(s)$ and $\Delta g_{\mathbf{0}}(t)$. Conclude from this that it suffices to show that: given $\epsilon > 0$ there exists $\delta > 0$ such that if $\|(u,v)\|_1 < \delta$, then

$$\left\|\Delta \left(h^v\right)_{\mathbf{0}}(u) - Su\right\| \le \epsilon \|u\| \tag{26.10}$$

and

$$\|\Delta g_{\mathbf{0}}(v) - Tv\| \le \epsilon \|v\| . \tag{26.11}$$

Find $\delta_1 > 0$ so that (26.11) holds whenever $\|v\| < \delta_1$. (This is easy.) Then find $\delta_2 > 0$ so that (26.10) holds whenever $\|(u,v)\|_1 < \delta_2$. This requires a little thought. Notice that (26.10) follows from the *mean value theorem* (corollary 26.1.8) provided that we can verify that

$$\left\|d\left(h^v\right)_z - S\right\| \le \epsilon$$

for all z in the segment $[\mathbf{0}, u]$. To this end show that $d\left(h^v\right)_z = d_1 f_{(a+z,b+v)}$ and use the fact that $d_1 f$ is assumed to be continuous. (Solution Q.26.5.)

26.2.11. Notation. Up to this point we have had no occasion to consider the problem of the various ways in which a Euclidean space $\mathbb{R}^n$ may be regarded as a product. Take $\mathbb{R}^5$ for example. If nothing to the contrary is specified it is natural to think of $\mathbb{R}^5$ as being the product of 5 copies of $\mathbb{R}$; that is, points of $\mathbb{R}^5$ are 5-tuples $x = (x_1, \dots, x_5)$ of real numbers. If, however we wish to regard $\mathbb{R}^5$ as the product of $\mathbb{R}^2$ and $\mathbb{R}^3$, then a point in $\mathbb{R}^5$ is an ordered pair (x, y) with $x \in \mathbb{R}^2$ and $y \in \mathbb{R}^3$. One good way of informing a reader that you wish $\mathbb{R}^5$ considered as a product in this particular fashion is to write $\mathbb{R}^2 \times \mathbb{R}^3$; another is to write $\mathbb{R}^{2+3}$. (Note the distinction between $\mathbb{R}^{2+3}$ and $\mathbb{R}^{3+2}$.) In many concrete problems the names of variables are given in the statement of the problem: for example, suppose we encounter several equations involving the variables u, v, w, x, and y and wish to solve for the last three variables in terms of the first two. We are then thinking of $\mathbb{R}^5$ as the product of $\mathbb{R}^2$ (the space of independent variables) and $\mathbb{R}^3$ (the space of dependent variables). This particular factorization may be emphasized by writing a point (u, v, w, x, y) of the product as $((u, v), (w, x, y))$. And if you wish to avoid such an abundance of parentheses you may choose to write $(u, v; w, x, y)$ instead. Ordinarily when $\mathbb{R}^n$ appears, it is clear from context what factorization (if any) is intended. The preceding notational devices are merely reminders designed to ease the burden on the reader. In the next exercise a function f of 4 variables is given and you are asked to compute $d_1 f_a$ in several different circumstances; it is important to realize that the value of $d_1 f_a$ *depends on the factorization of $\mathbb{R}^4$ that is assumed.*

26.2.12. Exercise. Let

$$f(t, u, v, w) = tu^2 + 3\,vw \tag{26.12}$$

for all t, u, v, $w \in \mathbb{R}$, and let $a = (1, 1, 2, -1)$. Find $d_1 f_a$ assuming that the domain of f is:

(a) $\mathbb{R}^4$.
(b) $\mathbb{R} \times \mathbb{R}^3$.
(c) $\mathbb{R}^2 \times \mathbb{R}^2$.
(d) $\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}$.
(e) $\mathbb{R}^3 \times \mathbb{R}$.

Hint. First compute df_a. Then use 26.2.8. **Note.** The default case is (a); that is, if we were given only equation (26.12) we would assume that the domain of f is $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. In this case it is possible to compute $d_k f_a$ for $k = 1, 2, 3, 4$. In cases (b), (c), and (e) only $d_1 f_a$ and $d_2 f_a$ make sense; and in (d) we can compute $d_k f_a$ for $k = 1$, 2, 3. (Solution Q.26.6.)

26.2.13. Problem. Let

$$f(t, u, v, w) = tuv - 4\,u^2 w$$

for all t, u, v, $w \in \mathbb{R}$ and let $a = (1, 2, -1, 3)$. Compute $d_k f_a$ for all k for which this expression makes sense, assuming that the domain of f is:

(a) $\mathbb{R}^4$.
(b) $\mathbb{R} \times \mathbb{R}^3$.
(c) $\mathbb{R}^2 \times \mathbb{R}^2$.
(d) $\mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}$.
(e) $\mathbb{R}^3 \times \mathbb{R}$.

26.2.14. Definition. We now consider the special case of a function f which maps an open subset of $\mathbb{R}^n$ into a normed linear space. For $1 \le k \le n$ the injection $j_k \colon \mathbb{R} \to \mathbb{R}^n$ takes the real number t to the vector $t\,e^k$. Thus the function $g \equiv f \circ T_a \circ j_k$ is just $f \circ l$ where as usual l is the parametrized line through a in the direction of e^k. [Proof: $g(t) = f\left(a + j_k(t)\right) = f(a + t\,e^k) = (f \circ l)(t)$.] We define $f_k(a)$ (or $\frac{\partial f}{\partial x_k}(a)$, or $D_k f(a)$), the k^{th} PARTIAL DERIVATIVE of f at a, to be $d_k f_a(1)$. Using propositions 25.4.7 and 25.5.2 we see that

$$\begin{aligned} f_k(a) &= d_k f_a(1) \\ &= dg_0(1) \\ &= Dg(0) \\ &= D(f \circ l)(0) \\ &= D_{e^k} f(a). \end{aligned}$$

That is, the k^{th} partial derivative of f is the directional derivative of f in the direction of the k^{th} coordinate axis of $\mathbb{R}^n$. Thus the notation $D_k f$ for the k^{th} partial derivative can be regarded as a slight abbreviation of the usual notation $D_{e^k} f$ used for directional derivatives of functions on $\mathbb{R}^n$.

It is also useful to note that

$$\begin{aligned} f_k(a) &= Dg(0) \\ &= \lim_{t \to 0} \frac{g(t) - g(0)}{t} \\ &= \lim_{t \to 0} \frac{f(a + t\,e^k) - f(a)}{t}. \end{aligned} \tag{26.13}$$

This is the usual definition given for partial derivatives in beginning calculus. The mechanics of computing partial derivatives is familiar and is justified by (26.13): pretend that a function of n variables is a function only of its k^{th} variable, then take the ordinary derivative.

One more observation: if f is differentiable at a in $\mathbb{R}^n$, then by proposition 25.5.9

$$f_k(a) = D_{e^k} f(a) = df_a(e^k)\,.$$

Let $f \in \mathcal{D}_a(\mathbb{R}^n, W)$ where $W = W_1 \times \cdots \times W_m$ is the product of m normed linear spaces. From proposition 25.6.2 (and induction) we know that

$$df_a = \left(d(f^1)_a, \ldots, d(f^m)_a\right)\,.$$

From this it is an easy step to the next proposition.

26.2.15. Proposition. *If $f \in \mathcal{D}_a(\mathbb{R}^n, W)$ where $W = W_1 \times \cdots \times W_m$ is the product of m normed linear spaces, then*

$$\left(f^j\right)_k = (f_k)^j$$

for $1 \le j \le m$ and $1 \le k \le n$.

Proof. Problem.

The point of the preceding proposition is that no ambiguity arises if we write the expression f_k^j. It may correctly be interpreted either as the k^{th} partial derivative of the j^{th} component function f^j or as the j^{th} component of the k^{th} partial derivative f_k.

26.2.16. Example. Let $f(x, y) = (x^5y^2, x^3 - y^3)$. To find $f_1(x, y)$ hold y fixed and differentiate with respect to x using corollary 25.6.3. Then

$$f_1(x, y) = (5\,x^4y^2, 3\,x^2)\,.$$

Similarly,

$$f_2(x, y) = (2\,x^5y, -3\,y^2)\,.$$

26.2.17. Exercise. Let $f(x, y, z) = (x^3y^2 \sin z, x^2 + y \cos z)$ and $a = (1, -2, \frac{\pi}{2})$. Find $f_1(a)$, $f_2(a)$, and $f_3(a)$. (Solution Q.26.7.)

26.2.18. Problem. Let $f(w, x, y, z) = (wxy^2z^3, w^2 + x^2 + y^2, wx + xy + yz)$ and $a = (-3, 1, -2, 1)$. Find $f_k(a)$ for all appropriate k.

26.2.19. Problem. Let $V_1, \dots, V_n, W$ be normed linear spaces, $U \overset{\circ}{\subseteq} V = V_1 \times \cdots \times V_n$, $\alpha \in \mathbb{R}$, and $f,\, g \colon U \to W$. If the k^{th} partial derivatives of f and g exist at a point a in U, then so do the k^{th} partial differentials of $f + g$ and αf, and

(a) $d_k(f + g)_a = d_kf_a + d_kg_a$;
(b) $d_k(\alpha f)_a = \alpha d_kf_a$;
(c) $(f + g)_k(a) = f_k(a) + g_k(a)$; and
(d) $(\alpha f)_k(a) = \alpha f_k(a)$.

Hint. In (a), consider the function $(f + g) \circ T_a \circ j_k$.

26.2.20. Problem. Let $f,\, g \in \mathcal{F}_a(V, W)$, $\alpha \in \mathbb{R}$, and $\mathbf{0} \ne v \in V$. Suppose that $D_vf(a)$ and $D_vg(a)$ exist.

(a) Show that $D_v(f + g)(a)$ exists and is the sum of $D_vf(a)$ and $D_vg(a)$. *Hint.* Use the definition of directional derivative and proposition 25.3.3. We can *not* use either 25.5.2 or 25.5.9 here because we are *not* assuming that f and g are differentiable at a.
(b) Show that $D_v(\alpha f)(a)$ exists and is equal to $\alpha D_vf(a)$.
(c) Use (a) and (b) to prove parts (c) and (d) of problem 26.2.19.

26.3. ITERATED INTEGRALS

In the preceding section we considered partial differentiation of functions defined on the product of two or more spaces. In this section we consider the integration of such functions. To take the partial derivative of a function $f \colon (x, y) \mapsto f(x, y)$ with respect to x we hold y fixed and differentiate the function $x \mapsto f(x, y)$. Partial integration works in very much the same way. If f is a continuous function mapping a rectangular subset $[a, b] \times [c, d]$ of $\mathbb{R}^2$ into a Banach space E, we may, for each fixed y in $[c, d]$, integrate the function $x \mapsto f(x, y)$ over the interval $[a, b]$. (This function is continuous since it is the composite of the continuous functions $x \mapsto (x, y)$ and f.) The integration will result in a vector which depends on y, call it $h(y)$. We will show shortly that the function $y \mapsto h(y)$ is also continuous and so may be integrated over $[c, d]$. The resulting vector in E is denoted by $\int_c^d \left(\int_a^b f(x, y)\, dx\right) dy$ or just $\int_c^d \int_a^b f(x, y)\, dx\, dy$. The two integrals operating successively are called ITERATED INTEGRALS.

26.3.1. Notation. Let $f\colon (x,y) \mapsto f(x,y)$ be a continuous function defined on a subset of $\mathbb{R}^2$ containing the rectangle $[a,b] \times [c,d]$. Throughout this section we denote by $f^{\,y}$ the function of x which results from holding y fixed and by ${}^{x}\!f$ the function of y resulting from fixing x. That is, for each y

$$f^{\,y}\colon x \mapsto f(x,y)$$

and for each x

$${}^{x}\!f\colon y \mapsto f(x,y)\,.$$

For each y we interpret $\int_a^b f(x,y)\,dx$ to mean $\int_a^b f^{\,y}$, and for each x we take $\int_c^d f(x,y)\,dy$ to be $\int_c^d {}^{x}\!f$. Thus

$$\int_c^d \int_a^b f(x,y)\,dx\,dy = \int_c^d g$$

where $g(y) = \int_a^b f^{\,y}$ for all $y \in [c,d]$. In order for $\int_c^d g$ to make sense we must know that g is a regulated function. It will suffice for our needs to show that if f is continuous, then so is g.

26.3.2. Lemma. *Let $f\colon [a,b]\times[c,d] \to E$ be a continuous function into a Banach space. For each $y \in [c,d]$ let*

$$g(y) = \int_a^b f^{\,y}\,.$$

Then g is uniformly continuous on $[c,d]$.

Proof. Exercise. *Hint.* Use proposition 24.1.11. (Solution Q.26.8.)

Perhaps the most frequently used result concerning iterated integrals is that if f is continuous, then the order of integration does not matter.

26.3.3. Proposition. *If E is a Banach space, if $a < b$ and $c < d$, and if $f\colon [a,b]\times[c,d] \to E$ is continuous, then*

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = \int_c^d \int_a^b f(x,y)\,dx\,dy\,.$$

Proof. Problem. *Hint.* Define functions j and k for all z in $[c,d]$ by the formulas

$$j(z) = \int_a^b \int_c^z f(x,y)\,dy\,dx$$

and

$$k(z) = \int_c^z \int_a^b f(x,y)\,dx\,dy\,.$$

It suffices to show that $j = k$. (Why?) One may accomplish this by showing that $j\,'(z) = k'(z)$ for all z and that $j(c) = k(c)$ (see corollary 26.1.10). Finding $k'(z)$ is easy. To find $j\,'(z)$ derive the formulas

$$\frac{1}{h}\Delta j_z(h) = \int_a^b \frac{1}{h}\int_z^{z+h} f(x,y)\,dy\,dx \tag{26.14}$$

and

$$\int_a^b f(x,z)\,dx = \int_a^b \frac{1}{h}\int_z^{z+h} f(x,z)\,dy\,dx\,. \tag{26.15}$$

Subtract (26.15)) from (26.14) to obtain a new equation. Show that the right side of this new equation can be made arbitrarily small by choosing h sufficiently small. (For this use an argument similar to the one used in the proof of lemma 26.3.2. Conclude that $j\,'(z) = \int_a^b f(x,z)\,dx$.

Now that we have a result (proposition 26.3.3) allowing us to change the order of two integrals, we can prove a result which justifies changing the order of integration and differentiation. We show that if f is continuous and f_2 (exists and) is continuous, then

$$\frac{d}{dy}\int_a^b f(x,y)\,dx = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\,.$$

26.3.4. Proposition. *Let E be a Banach space, $a < b$, $c < d$, and $f\colon [a,b]\times[c,d]\to E$. If f and f_2 are continuous then the function g defined for all $y \in [c,d]$ by*

$$g(y) = \int_a^b f^{\,y}$$

is continuously differentiable in (c,d) and for $c < y < d$

$$g'(y) = \int_a^b f_2(x,y)\,dx\,.$$

Proof. Exercise. *Hint.* Let $h(y) = \int_a^b f_2(x,y)\,dx$ for $c \le y \le d$. Use proposition 26.3.3 to reverse the order of integration in $\int_c^z h$ (where $c < z < d$). Use the version of the *fundamental theorem of calculus* given in Theorem 26.1.13 to obtain $\int_c^z h = g(z) - g(c)$. Differentiate. (Solution Q.26.9.)

26.3.5. Problem. Compute

$$\int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy \qquad \text{and} \qquad \int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\,dx\,.$$

Why does the result not contradict the assertion made in proposition 26.3.3?

26.3.6. Problem. (a) Suppose that the functions $g\colon \mathbb{R}^n\to\mathbb{R}$ and $h\colon\mathbb{R}\to\mathbb{R}$ are differentiable and $f = h\circ g$. Show that

$$f_k(x) = g_k(x)\,(Dh)(g(x))$$

whenever $x\in\mathbb{R}^n$ and $1\le k\le n$.

(b) Let $g\colon\mathbb{R}^n\to\mathbb{R}$ be differentiable and $j\colon\mathbb{R}\to\mathbb{R}$ be continuous. Prove that

$$\frac{\partial}{\partial x_k}\int_c^{g(x)} j(t)\,dt = g_k(x)\,j(g(x))$$

whenever $c\in\mathbb{R}$, $x\in\mathbb{R}^n$, and $1\le k\le n$. *Hint.* The expression on the left denotes $f_k(x)$ where f is the function $x\mapsto\int_c^{g(x)} j(t)\,dt$.

(c) Use part (b) to compute

$$\frac{\partial}{\partial x}\int_{x^3y}^{x^2+y^2}\frac{1}{1+t^2+cos^2t}\,dt\,.$$

26.3.7. Proposition (Leibniz's formula)**.** *Let $f\colon[a,b]\times[c,d]\to\mathbb{R}$ and $h\colon[c,d]\to\mathbb{R}$. If f and f_2 are continuous, if h is continuously differentiable on (c,d), and if $h(y)\in[a,b]$ for every $y\in(c,d)$, then*

$$\frac{d}{dy}\int_a^{h(y)} f(x,y)\,dx = \int_a^{h(y)} f_2(x,y)\,dx + Dh(y)\,f(h(y),y)\,.$$

Proof. Problem.

CHAPTER 27

COMPUTATIONS IN $\mathbb{R}^n$

In the preceding two chapters we have developed some fundamental facts concerning the differential calculus in arbitrary Banach spaces. In the present chapter we restrict our attention to the Euclidean spaces $\mathbb{R}^n$. Not only are these spaces very important historically, but fortunately there are available a variety of powerful yet relatively simple techniques which make possible explicit computations of many of the concepts introduced in chapter 25. The usefulness of these spaces seems to be associated with the emphasis in classical physics on systems having a finite number of degrees of freedom. The computational simplicity stems from two facts: first, differentials of functions between Euclidean spaces are always continuous (see proposition 23.1.18); and second, the usual norm on $\mathbb{R}^n$ is derivable from an inner product. In the first section of this chapter we derive some standard elementary facts about inner products. It is important to appreciate that despite their algebraic appearance, inner products are the source of much of the geometry in Euclidean spaces.

27.1. INNER PRODUCTS

27.1.1. Definition. Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ be vectors in $\mathbb{R}^n$. The INNER PRODUCT (or DOT PRODUCT) of x and y, denoted by $\langle x, y\rangle$, is defined by

$$\langle x, y\rangle := \sum_{k=1}^{n} x_k y_k \,.$$

As a first result we list the most important properties of the inner product.

27.1.2. Proposition. *Let x, y, and z be vectors in $\mathbb{R}^n$ and α be a scalar. Then*

(a) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$;
(b) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$;
(c) $\langle x, y\rangle = \langle y, x\rangle$;
(d) $\langle x, x\rangle \geq 0$;
(e) $\langle x, x\rangle = 0$ *only if* $x = \mathbf{0}$; *and*
(f) $\|x\| = \sqrt{\langle x, x\rangle}$.

Items (a) and (b) say that the inner product is linear *in its first variable; (c) says it is* symmetric*; and (d) and (e) say that it is* positive definite. *It is virtually obvious that an inner product is also linear in its second variable (see exercise 27.1.4). Thus an inner product may be characterized as a positive definite, symmetric, bilinear functional on $\mathbb{R}^n$.*

PROOF. Problem.

27.1.3. Proposition. *If x is in $\mathbb{R}^n$, then*

$$x = \sum_{k=1}^{n} \langle x, e^k\rangle e^k \,.$$

This result is used so frequently that it has been stated formally as a proposition. Its proof, however, is trivial. [It is clear from the definition of the inner product that $\langle x, e^k\rangle = x_k$, where, as usual, $\{e^1, \dots, e^n\}$ is the standard basis for $\mathbb{R}^n$.]

27.1.4. Exercise. Use properties (a)–(c) above, but not the definition of *inner product* to prove that

$$\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$$

and

$$\langle x, \alpha y\rangle = \alpha\langle x, y\rangle$$

for all x, y, $z \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. (Solution Q.27.1.)

27.1.5. Proposition (The Parallelogram Law)**.** *If x, $y \in \mathbb{R}^n$, then*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2\,.$$

PROOF. Problem.

27.1.6. Proposition (Schwarz's Inequality)**.** *If u, $v \in \mathbb{R}^n$, then*

$$|\langle u, v\rangle| \le \|u\|\|v\|\,.$$

PROOF. This has been proved in chapter 9: notice that the left side of the inequality given in proposition 9.2.6 is $\big(\langle u, v\rangle\big)^2$ and the right side is $\|u\|^2\|v\|^2$. □

27.1.7. Definition. If x and y are nonzero vectors in $\mathbb{R}^n$, define $\measuredangle(x, y)$, the ANGLE between x and y, by

$$\measuredangle(x, y) := \arccos\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right)\,.$$

A version of this formula which is perhaps somewhat more familiar is

$$\langle x, y\rangle = \|x\|\|y\|\cos\measuredangle(x, y)\,.$$

27.1.8. Exercise. How do we know that the preceding definition makes sense? (What is the domain of the arccosine function?) (Solution Q.27.2.)

27.1.9. Problem. Prove the *law of cosines*: if x and y are nonzero vectors in $\mathbb{R}^n$ and $\theta = \measuredangle(x, y)$, then

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos\theta\,.$$

27.1.10. Exercise. What is the angle between the vectors $(1, 0, 1)$ and $(0, -1, 1)$ in $\mathbb{R}^3$? (Solution Q.27.3.)

27.1.11. Problem. Find the angle between the vectors $(1, 0, -1, -2)$ and $(-1, 1, 0, 1)$ in $\mathbb{R}^4$.

27.1.12. Problem. The angle of intersection of two curves is by definition the angle between the tangent vectors to the curves at the point of intersection. Find the angle of intersection at the point $(1, -2, 3)$ of the curves C_1 and C_2 where

$$C_1(t) = (t, t^2 + t - 4, 3 + \ln t)$$

and

$$C_2(u) = (u^2 - 8, u^2 - 2u - 5, u^3 - 3u^2 - 3u + 12)\,.$$

27.1.13. Definition. Two vectors x and y in $\mathbb{R}^n$ are PERPENDICULAR (or ORTHOGONAL) if $\langle x, y\rangle = 0$. In this case we write $x \perp y$. Notice that the relationship between perpendicularity and angle is what we expect: if x and y are nonzero vectors then $x \perp y$ if and only if $\measuredangle(x, y) = \pi/2$. The zero vector is perpendicular to all vectors but the angle it makes with another vector is not defined.

27.1.14. Problem. Find a linear combination of the vectors $(1, 0, 2)$ and $(2, -1, 1)$ which is perpendicular to the vector $(2, 2, 1)$ in $\mathbb{R}^3$.

27.1.15. Problem. Prove the *Pythagorean theorem*: if $x \perp y$ in $\mathbb{R}^n$, then

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2\,.$$

Does the converse hold?

27.1.16. Notation. Let $f\colon U \to \mathbb{R}^n$ and $g\colon V \to \mathbb{R}^n$ where U and V are subsets of a normed linear space which are not disjoint. Then we denote by $\langle f, g\rangle$ the real valued function on $U \cap V$ whose value at a point x in $U \cap V$ is $\langle f(x), g(x)\rangle$. That is,

$$\langle f, g\rangle \colon U \cap V \to \mathbb{R}\colon x \mapsto \langle f(x), g(x)\rangle\,.$$

The scalar field $\langle f, g\rangle$ is the INNER PRODUCT (or DOT PRODUCT of f and g.

27.1.17. Proposition. *Suppose that the functions $f\colon U \to \mathbb{R}^n$ and $g\colon V \to \mathbb{R}^n$, defined on subsets of a normed linear space W, are differentiable at a point a in the interior of $U \cap V$. Then $\langle f, g\rangle$ is differentiable at a and*

$$d\langle f, g\rangle_a = \langle f(a), dg_a\rangle + \langle df_a, g(a)\rangle\,.$$

PROOF. Problem. *Hint.* Use propositions 25.3.16 and 25.6.2.

27.1.18. Corollary. *If f and g are curves at a point a in $\mathbb{R}^n$ and are differentiable, then*

$$D\langle f, g\rangle(a) = \langle f(a), Dg(a)\rangle + \langle Df(a), g(a)\rangle\,.$$

PROOF. Use 27.1.17 and 25.4.7.

$$\begin{aligned} D\langle f, g\rangle(a) &= d\langle f, g\rangle_a(1) \\ &= \langle f(a), dg_a(1)\rangle + \langle df_a(1), g(a)\rangle \\ &= \langle f(a), Dg(a)\rangle + \langle Df(a), g(a)\rangle \end{aligned}$$

□

27.1.19. Problem. Let $f = (f^1, f^2, f^3)$ where

$$\begin{aligned} f^1(t) &= t^3 + 2t^2 - 4t + 1 \\ f^2(t) &= t^4 - 2t^3 + t^2 + 3 \\ f^3(t) &= t^3 - t^2 + t - 2 \end{aligned}$$

and let $g(t) = \|f(t)\|^2$ for all t in $\mathbb{R}$. Find $Dg(1)$.

27.1.20. Problem. Let c be a differentiable curve in $\mathbb{R}^n$. Show that the point $c(t)$ moves on the surface of a sphere centered at the origin if and only if the tangent vector $Dc(t)$ at t is perpendicular to the position vector $c(t)$ at each t. *Hint.* Use corollary 27.1.18.

27.2. THE GRADIENT

In beginning calculus texts the gradient of a real valued function of n variables is usually defined to be an n-tuple of partial derivatives. This definition, although convenient for computation, disguises the highly geometric nature of the gradient. Here we adopt a different definition: the gradient of a scalar field is the vector which represents, in a sense to be made precise below, the differential of the function. First we look at an important example of a bounded linear functional on $\mathbb{R}^n$.

27.2.1. Example. Let $b \in \mathbb{R}^n$. Define

$$\psi_b\colon \mathbb{R}^n \to \mathbb{R}\colon x \mapsto \langle x, b\rangle\,.$$

Then ψ_b is a bounded linear functional on $\mathbb{R}^n$ and $\|\psi_b\| = \|b\|$.

PROOF. Exercise. (Solution Q.27.4.)

The reason for the importance of the preceding example is that functions of the form ψ_b turn out to be the *only* bounded linear functionals on $\mathbb{R}^n$. Since on $\mathbb{R}^n$ every linear functional is bounded (see propositions 23.1.18 and 23.1.4), the functions ψ_b are in fact the only real valued linear maps on $\mathbb{R}^n$. Thus we say that every linear functional on $\mathbb{R}^n$ can be represented in the form ψ_b for some vector b in $\mathbb{R}^n$. Furthermore, this representation is unique. These assertions are stated formally in the next theorem.

27.2.2. Theorem (Riesz-Fréchet Theorem)**.** *If $f \in (\mathbb{R}^n)^*$, then there exists a unique vector b in $\mathbb{R}^n$ such that*

$$f(x) = \langle x, b \rangle$$

for all x in $\mathbb{R}^n$.

Proof. Problem. *Hint.* For the existence part, (a) write $\sum_{k=1}^{n} x_k e^k$ for x in the expression $f(x)$, and use the linearity of f. Then, (b) write $\langle x, b \rangle$ as a sum. Comparing the results of (a) and (b), guess the identity of the desired vector b. The uniqueness part is easy: suppose $f(x) = \langle x, a \rangle = \langle x, b \rangle$ for all x in $\mathbb{R}^n$. Show $a = b$.

27.2.3. Definition. If a map $T \colon V \to W$ between two normed linear spaces is both an isometry and a vector space isomorphism, we say that it is an isometric isomorphism and that the spaces V and W are isometrically isomorphic.

27.2.4. Proposition. *Each Euclidean space $\mathbb{R}^n$ is isometrically isomorphic to its dual $(\mathbb{R}^n)^*$.*

Proof. Problem. *Hint.* Consider the map

$$\psi \colon \mathbb{R}^n \to (\mathbb{R}^n)^* \colon b \mapsto \psi_b \,.$$

One thing that must be established is that ψ is linear; don't confuse this with showing that each ψ_b is linear—a task already accomplished in example 27.2.1. Use the Riesz-Fréchet theorem 27.2.2 and problem 22.3.12.

The Riesz-Fréchet theorem 27.2.2 is the crucial ingredient in our definition of the *gradient* of a scalar field.

27.2.5. Definition. Let $U \subseteq \mathbb{R}^n$ and $\phi \colon U \to \mathbb{R}$ be a scalar field. If ϕ is differentiable at a point a in U°, then its differential $d\phi_a$ is a bounded linear map from $\mathbb{R}^n$ into $\mathbb{R}$. That is, $d\phi_a \in (\mathbb{R}^n)^*$. Thus according to the Riesz-Fréchet theorem 27.2.2 there exists a unique vector, which we denote by $\nabla\phi(a)$, representing the linear functional $d\phi_a$. That is, $\nabla\phi(a)$ is the unique vector in $\mathbb{R}^n$ such that

$$d\phi_a(x) = \langle x, \nabla\phi(a) \rangle$$

for all x in $\mathbb{R}^n$. The vector $\nabla\phi(a)$ is the gradient of ϕ at a. If U is an open subset of $\mathbb{R}^n$ and ϕ is differentiable at each point of U, then the function

$$\nabla\phi \colon U \to \mathbb{R}^n \colon u \mapsto \nabla\phi(u)$$

is the gradient of ϕ. Notice two things: first, the gradient of a scalar field is a vector field; and second, the differential $d\phi_a$ is the zero linear functional if and only if the gradient at a, $\nabla\phi(a)$, is the zero vector in $\mathbb{R}^n$.

Perhaps the most useful fact about the gradient of a scalar field ϕ at a point a in $\mathbb{R}^n$ is that it is the vector at a which points in the direction of the most rapid increase of ϕ.

27.2.6. Proposition. *Let $\phi \colon U \to \mathbb{R}$ be a scalar field on a subset U of $\mathbb{R}^n$. If ϕ is differentiable at a point a in U and $d\phi_a$ is not the zero functional, then the maximum value of the directional derivative $D_u\phi(a)$, taken over all unit vectors u in $\mathbb{R}^n$, is achieved when u points in the direction.of the gradient $\nabla\phi(a)$. The minimum value is achieved when u points in the opposite direction $-\nabla\phi(a)$.*

Proof. Exercise. *Hint.* Use proposition 25.5.9 and recall that $\langle x, y \rangle = \|x\|\|y\| \cos\measuredangle(x, y)$. (Solution Q.27.5.)

When a curve c is composed with a scalar field ϕ we obtain a real valued function of a single variable. An easy but useful special case of the *chain rule* says that the derivative of the composite $\phi \circ c$ is the dot product of the derivative of c with the gradient of ϕ.

27.2.7. Proposition. *Suppose that c is a curve in $\mathbb{R}^n$ which is differentiable at a point t in $\mathbb{R}$ and that ϕ belongs to $\mathcal{D}_{c(t)}(\mathbb{R}^n,\mathbb{R})$. Then $\phi\circ c$ is differentiable at t and*

$$D(\phi\circ c)(t) = \langle Dc(t), (\nabla\phi)(c(t))\rangle\,.$$

Proof. Problem. *Hint.* Use proposition 25.4.7 and the *chain rule* 25.3.17.

27.2.8. Problem. If $\phi\colon \mathbb{R}^n\to\mathbb{R}\colon x\mapsto \|x\|^2$, then ϕ is differentiable at each point b of $\mathbb{R}^n$ and

$$d\phi_b = 2\psi_b$$

Furthermore, $\nabla\phi = 2I$ (where I is the identity function on $\mathbb{R}^n$). *Hint.* The definition of ψ_b is given in 27.2.1. Write $\phi = \langle I, I\rangle$ and use 27.1.17.

27.2.9. Problem. For the function ϕ given in the preceding problem verify by direct computation the formula for the *mean value theorem* (proposition 26.1.14).

27.2.10. Definition. A linear transformation $T\colon\mathbb{R}^n\to\mathbb{R}^n$ is SELF-ADJOINT if $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all x, $y\in\mathbb{R}^n$.

27.2.11. Problem. Let $T\in\mathfrak{B}(\mathbb{R}^n,\mathbb{R}^n)$ be self-adjoint and

$$\mu\colon\mathbb{R}^n\to\mathbb{R}\colon x\mapsto\langle Tx, x\rangle.$$

(a) Show that μ is differentiable at each point b in $\mathbb{R}^n$ and find $d\mu_b$.
(b) Find $\nabla\mu$.

27.2.12. Problem. Repeat problem 27.2.9, this time using the function μ given in problem 27.2.11.

27.2.13. Exercise (Conservation of Energy)**.** Consider a particle P moving in $\mathbb{R}^3$ under the influence of a force F. Suppose that the position of P at time t is $x(t)$ where $x\colon\mathbb{R}\to\mathbb{R}^3$ is at least twice differentiable. Let $v := Dx$ be the velocity of P and $a := Dv$ be its acceleration. Assume *Newton's second law*: $F\circ x = ma$, where $F\colon\mathbb{R}^3\to\mathbb{R}^3$ is the force acting on P and m is the mass of P. Suppose further that the force field F is CONSERVATIVE; that is, there exists a scalar field $\phi\colon\mathbb{R}^3\to\mathbb{R}$ such that $F = -\nabla\phi$. (Such a scalar field is a POTENTIAL FUNCTION for F.) The KINETIC ENERGY of P is defined by

$$KE := \tfrac{1}{2}m\|v\|^2\,,$$

its POTENTIAL ENERGY by

$$PE := \phi\circ x\,,$$

and its TOTAL ENERGY by

$$TE := KE + PE\,.$$

Prove, for this situation, the law of *conservation of energy*:

$$TE \text{ is constant.}$$

Hint. Use propositions 26.1.9, 27.1.17, and 27.2.7. (Solution Q.27.6.)

In most circumstances the simplest way of computing the gradient of a scalar field ϕ on $\mathbb{R}^n$ is to calculate the n partial derivatives of ϕ. The n-tuple of these derivatives is the gradient. In most beginning calculus texts this is the *definition* of the gradient.

27.2.14. Proposition. *If ϕ is a scalar field on a subset of $\mathbb{R}^n$ and is differentiable at a point a, then*

$$\nabla\phi(a) = \sum_{k=1}^{n}\phi_k(a)e^k\,.$$

Proof. Exercise. *Hint.* Substitute $\nabla\phi(a)$ for x in 27.1.3. Use 25.5.9. (Solution Q.27.7.)

27.2.15. Exercise. Let $\phi(w,x,y,z) = wz - xy$, $u = (\frac12, -\frac12, \frac12, -\frac12)$, and $a = (1,2,3,4)$. Find the directional derivative $D_u\phi(a)$. *Hint.* Use 25.5.9, the definition of *gradient*, and 27.2.14. (Solution Q.27.8.)

27.2.16. Exercise (Method of Steepest Descent)**.** Let $\phi(x, y) = 2x^2 + 6y^2$ and $a = (2, -1)$. Find the steepest downhill path on the surface $z = \phi(x, y)$ starting at the point a and ending at the minimum point on the surface. *Hints.* (1) It is enough to find the equation of the projection of the curve onto the xy-plane; every curve $t \mapsto \big(x(t), y(t)\big)$ in the xy-plane is the projection along the z-axis of a unique curve $t \mapsto \big(x(t), y(t), \phi(x(t), y(t))\big)$ on the surface $z = \phi(x, y)$. (2) If $c\colon t \mapsto \big(x(t), y(t)\big)$ is the desired curve and we set $c(0) = a$, then according to proposition 27.2.6 the unit vector u which minimizes the directional derivative $D_u\phi(b)$ at a point b in $\mathbb{R}^2$ is the one obtained by choosing u to point in the direction of $-\nabla\phi(b)$. Thus in order for the curve to point in the direction of the most rapid decrease of ϕ at each point $c(t)$, the tangent vector to the curve at $c(t)$ must be some positive multiple $p(t)$ of $-(\nabla\phi)\big(c(t)\big)$. The function p will govern the speed of descent; since this is irrelevant in the present problem, set $p(t) = 1$ for all t. (3) Recall from beginning calculus that on an interval the only nonzero solution to an equation of the form $Dx(t) = kx(t)$ is of the form $x(t) = x(0)e^{kt}$. (4) The parameter t which we have introduced is artificial. Eliminate it to obtain an equation of the form $y = f(x)$. (Solution Q.27.9.)

27.2.17. Proposition (A Mean Value Theorem for Scalar Fields)**.** *Let ϕ be a differentiable scalar field on an open convex subset U of $\mathbb{R}^n$ and suppose that a and b are distinct points belonging to U. Then there exists a point c in the closed segment $[a, b]$ such that*

$$\phi(b) - \phi(a) = \langle b - a, \nabla\phi(c)\rangle .$$

Proof. Problem. *Hint.* Let $l(t) = (1 - t)a + tb$ for $0 \le t \le 1$. Apply the *mean value theorem* for a real valued function of a single variable (8.4.26) to the function $\phi \circ l$. Use proposition 27.2.7.

27.2.18. Problem. Let $c(t) = (\cos t, \sin t, t)$ and $\phi(x, y, z) = x^2y - 3yz$. Find $D(\phi \circ c)(\pi/6)$. *Hint.* Use 27.2.7 and 27.2.14.

27.2.19. Problem. Let $\phi(x, y, z) = xz - 4y$, $u = (\frac{1}{2}, 0, \frac{1}{2}\sqrt{3})$, and $a = (1, 0, -\frac{\pi}{2})$. Find the directional derivative $D_u\phi(a)$.

27.2.20. Problem. Show that if V is a normed linear space, $f \in \mathcal{D}_a(\mathbb{R}^n, V)$, and v is a nonzero vector in $\mathbb{R}^n$, then

$$D_v f(a) = \sum_{k=1}^{n} v_k f_k(a) .$$

Hint. Use proposition 25.5.9.

27.2.21. Problem. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2\colon x \mapsto ({x_1}^2 - {x_2}^2, 3x_1x_2)$, $a = (2, 1)$, and $v = (-1, 2)$. Use the preceding problem to find $D_v f(a)$.

27.2.22. Problem. Find the path of steepest descent on the surface $z = x^6 + 12y^4$ starting at the point whose x-coordinate is 1 and whose y-coordinate is $\frac{1}{2}$.

27.2.23. Problem. Suppose that the temperature $\phi(x, y)$ at points (x, y) on a flat surface is given by the formula

$$\phi(x, y) = x^2 - y^2 .$$

Starting at a point (a, b) on the surface, what path should be followed so that the temperature will increase as rapidly as possible?

27.2.24. Problem. This (like exercise 27.2.16 and problem 27.2.22) is a steepest descent problem; but here, we suppose that for some reason we are unable to solve explicitly the resulting differential equations. Instead we invoke an approximation technique. Let

$$\phi(x) = 13{x_1}^2 - 42x_1 + 13{x_2}^2 + 6x_2 + 10x_1x_2 + 9$$

for all x in $\mathbb{R}^2$. The goal is to approximate the path of steepest descent. Start at an arbitrary point x^0 in $\mathbb{R}^2$ and choose a number $h > 0$. At x^0 compute the gradient of ϕ, take u^0 to be the unit vector pointing in the direction of $-\nabla\phi(x^0)$, and then move h units in the direction of u^0 arriving at a point x^1. Repeat the procedure: find the unit vector u^1 in the direction of $-\nabla\phi(x^1)$, then from x^1 move h units along u^1 to a point x^2. Continue in this fashion. In other words, $x^0 \in \mathbb{R}^2$ and $h > 0$ are arbitrary, and for $n \geq 0$

$$x^{n+1} = x^n + hu^n$$

where $u^n = -\|\nabla\phi(x^n)\|^{-1}\nabla\phi(x^n)$.

(a) Start at the origin $x^0 = (0,0)$ and choose $h = 1$. Compute 25 or 30 values of x^n. Explain geometrically what is happening here. Why is h "too large"? *Hint.* Don't attempt to do this by hand. Write a program for a computer or a programmable calculator. In writing your program don't ignore the possibility that $\nabla\phi(x^n)$ may be zero for some n. (Keep in mind when you write this up that your reader probably has no idea how to read the language in which you write your program. Document it well enough that the reader can easily understand what you are doing at each step.)

(b) Describe what happens when h is "too small". Again start at the origin, take $h = 0.001$ and compute 25 or 30 values of x^n.

(c) By altering the values of h at appropriate times, find a succession of points $x^0, \dots, x^n$ (starting with $x^0 = (0,0)$) such that the distance between x^n and the point where ϕ assumes its minimum value is less than 0.001. (By examining the points $x^0, \dots, x^n$ you should be able to guess, for this particular function, the exact location of the minimum.)

(d) Alter the program in part (a) to eliminate division by $\|\nabla\phi(x^n)\|$. (That is, let $x^{n+1} = x^n - h\nabla\phi(x^n)$.) Explain what happens in this case when h is "too large" (say $h = 1$). Explain why the altered program works better (provided that h is chosen appropriately) than the program in (a) for the present function ϕ.

27.2.25. Problem. Is it possible to find a differentiable scalar field ϕ on $\mathbb{R}^n$ and a point a in $\mathbb{R}^n$ such that $D_u\phi(a) > 0$ for every nonzero u in $\mathbb{R}^n$?

27.2.26. Problem. Is it possible to find a differentiable scalar field ϕ on $\mathbb{R}^n$ and a nonzero vector u in $\mathbb{R}^n$ such that $D_u\phi(a) > 0$ for every a in $\mathbb{R}^n$?

27.3. THE JACOBIAN MATRIX

27.3.1. Definition. Let $U \overset{\circ}{\subseteq} \mathbb{R}^n$ and $f\colon U \to \mathbb{R}^m$ be a differentiable function. Recall that the components $f^1, \dots, f^m$ of f satisfy

$$f(x) = \big(f^1(x), \dots, f^m(x)\big) = \sum_{j=1}^{m} f^j(x)e^j$$

for all x in U. More briefly we may write

$$f = (f^1, \dots, f^m) = \sum_{j=1}^{m} f^j e^j\,.$$

Recall also (from proposition 26.2.15) that

$$(f^j)_k(a) = (f_k)^j(a)$$

whenever $1 \leq j \leq m$, $1 \leq k \leq n$, and $a \in U$. Consequently the notation $f^j_k(a)$ is unambiguous; from now on we use it. As the differential of f at a point a in U is a (bounded) linear map from $\mathbb{R}^n$ into $\mathbb{R}^m$, it may be represented by an $m \times n$ matrix. This is called the JACOBIAN MATRIX of f at a. The entry in the j^{th} row and k^{th} column of this matrix is $f^j_k(a)$.

27.3.2. Proposition. *If $f \in \mathcal{D}_a(\mathbb{R}^n, \mathbb{R}^m)$, then*

$$[df_a] = [f^j_k(a)]\,.$$

PROOF. Exercise. *Hint.* It helps to distinguish notationally between the standard basis vectors in $\mathbb{R}^n$ and those in $\mathbb{R}^m$. Denote the ones in $\mathbb{R}^n$ by $e^1, \dots, e^n$ and those in $\mathbb{R}^m$ by $\hat{e}^1, \dots, \hat{e}^m$. Use 21.3.11. (Solution Q.27.10.)

Note that the j^{th} row of the Jacobian matrix is $f^j_1, \dots, f^j_n$. Thought of as a vector in $\mathbb{R}^n$ this is just the gradient of the scalar field f^j. Thus we may think of the Jacobian matrix $[df_a]$ as being in the form

$$\begin{bmatrix} \nabla f^1(a) \\ \vdots \\ \nabla f^m(a) \end{bmatrix}.$$

27.3.3. Exercise. Let $f\colon \mathbb{R}^4 \to \mathbb{R}^3\colon (w,x,y,z) \mapsto (wxz, x^2+2y^2+3z^2, wy\arctan z)$, let $a = (1,1,1,1)$, and let $v = (0,2,-3,1)$.

(a) Find $[df_a]$.
(b) Find $df_a(v)$.

(Solution Q.27.11.)

27.3.4. Problem. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2\colon x \mapsto ({x_1}^2 - {x_2}^2, 3x_1x_2)$ and $a = (2,1)$.

(a) Find $[df_a]$.
(b) Use part (a) to find $df_a(-1,3)$.

27.3.5. Problem. Let $f\colon \mathbb{R}^3 \to \mathbb{R}^4\colon (x,y,z) \mapsto (xy, y-z^2, 2xz, y+3z)$, let $a = (1,-2,3)$, and let $v = (2,1,-1)$.

(a) Find $[df_a]$.
(b) Use part (a) to calculate $D_vf(a)$.

27.3.6. Problem. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (x^2y, 2xy^2)$, let $a = (2,-1)$, and let $u = (\frac{3}{5}, \frac{4}{5})$. Compute $D_uf(a)$ in three ways:

(a) Use the definition of *directional derivative*.
(b) Use proposition 25.5.2.
(c) Use proposition 25.5.9.

27.3.7. Problem. Suppose that $f \in \mathcal{D}_a(\mathbb{R}^3, \mathbb{R}^4)$ and that the Jacobian matrix of f at a is

$$\begin{bmatrix} b & c & e \\ g & h & i \\ j & k & l \\ m & n & p \end{bmatrix}$$

Find $f_1(a)$, $f_2(a)$, $f_3(a)$, $\nabla f^1(a)$, $\nabla f^2(a)$, $\nabla f^3(a)$, and $\nabla f^4(a)$.

27.3.8. Problem. Let $f \in \mathcal{D}_a(\mathbb{R}^n, \mathbb{R}^m)$ and $v \in \mathbb{R}^n$. Show that

(a) $df_a(v) = \sum_{j=1}^m \langle \nabla f^j(a), v\rangle e^j$, and
(b) $\|df_a\| \le \sum_{j=1}^m \|\nabla f^j(a)\|$.

27.4. THE CHAIN RULE

In some respects it is convenient for scientists to work with variables rather than functions. Variables denote the physical quantities in which a scientist is ultimately interested. (In thermodynamics, for example, T is temperature, P pressure, S entropy, and so on.) Functions usually have no such standard associations. Furthermore, a problem which deals with only a small number of variables may turn out to involve a dauntingly large number of functions if they are specified. The simplification provided by the use of variables may, however, be more apparent than real, and the price paid in increased ambiguity for their suggestiveness is often substantial. Below are a few examples of ambiguities produced by the combined effects of excessive reliance on variables, inadequate (if conventional) notation, and the unfortunate mannerism of using the same name for a function and a dependent variable ("Suppose $x = x(s,t) \dots$").

(A) If $z = f(x,y)$, what does $\frac{\partial}{\partial x} z(y,x)$ mean? (Perhaps $f_1(y,x)$? Possibly $f_2(y,x)$?)
(B) If $z = f(x,t)$ where $x = x(t)$, then what is $\frac{\partial z}{\partial t}$? (Is it $f_2(t)$? Perhaps the derivative of $t \mapsto f(x(t),t)$ is intended?)
(C) Let $f(x,y)$ be a function of two variables. Does the expression $z = f(tx,ty)$ have three partial derivatives $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, and $\frac{\partial z}{\partial t}$? Do $\frac{\partial z}{\partial x}$ and $\frac{\partial f}{\partial x}$ mean the same thing?
(D) Let $w = w(x,y,t)$ where $x = x(s,t)$ and $y = y(s,t)$. A direct application of the chain rule (as stated in most beginning calculus texts) produces
$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial t}\,.$$
Is this correct ? Do the terms of the form $\frac{\partial w}{\partial t}$ cancel?
(E) Let $z = f(x,y) = g(r,\theta)$ where $x = r\cos\theta$ and $y = r\sin\theta$. Do $\frac{\partial z}{\partial r}$ and $\frac{\partial z}{\partial \theta}$ make sense? Do $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$? How about z_1 and z_2? Are any of these equal?
(F) The formulas for changing polar to rectangular coordinates are $x = r\cos\theta$ and $y = r\sin\theta$. So if we compute the partial derivative of the variable r with respect to the variable x we get
$$\frac{\partial r}{\partial x} = \frac{\partial}{\partial x}\sqrt{x^2+y^2} = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} = \cos\theta\,.$$
On the other hand, since $r = \dfrac{x}{\cos\theta}$, we use the *chain rule* to get
$$\frac{\partial r}{\partial x} = \frac{1}{\cos\theta} = \sec\theta\,.$$
Do you suppose something is wrong here? What?

The principal goal of the present section is to provide a reliable formalism for dealing with partial derivatives of functions of several variables in such a way that questions like (A)–(F) can be avoided. The basic strategy is quite simple: when in doubt give names to the relevant functions (especially composite ones!) and then use the chain rule. Perhaps it should be remarked that one need not make a fetish of avoiding variables. Many problems stated in terms of variables can be solved quite simply without the intrusion of the names of functions. (*E.g.* What is $\frac{\partial z}{\partial x}$ if $z = x^3y^2$?) This section is intended as a guide for the perplexed. Although its techniques are often useful in dissipating confusion generated by inadequate notation, it is neither necessary nor even particularly convenient to apply them routinely to every problem which arises. Let us start by writing the chain rule for functions between Euclidean spaces in terms of partial derivatives. Suppose that $f \in \mathcal{D}_a(\mathbb{R}^p,\mathbb{R}^n)$ and $g \in \mathcal{D}_{f(a)}(\mathbb{R}^n,\mathbb{R}^m)$. Then according to theorem 25.3.17
$$d(g\circ f)_a = dg_{f(a)} \circ df_a\,.$$
Replacing these linear transformations by their matrix representations and using proposition 21.5.12 we obtain
$$\big[d(g\circ f)_a\big] = \big[dg_{f(a)}\big]\big[df_a\big]\,. \tag{27.1}$$

27.4.1. Proposition. *If $f \in \mathcal{D}_a(\mathbb{R}^p,\mathbb{R}^n)$ and $g \in \mathcal{D}_{f(a)}(\mathbb{R}^n,\mathbb{R}^m)$, then $\big[d(g\circ f)_a\big]$ is the $m\times p$ matrix whose entry in the j^{th} row and k^{th} column is $\sum_{i=1}^n g^j_i(f(a))f^i_k(a)$. That is,*
$$(g\circ f)^j_k(a) = \sum_{i=1}^n \big(g^j_i\circ f\big)(a)f^i_k(a)$$
for $1 \le j \le m$ and $1 \le k \le p$.

Proof. Multiply the two matrices on the right hand side of (27.1) and use proposition 27.3.2. $\square$

It is occasionally useful to restate proposition 27.4.1 in the following (clearly equivalent) way.

27.4.2. Corollary. *If $f \in \mathfrak{D}_a(\mathbb{R}^p, \mathbb{R}^n)$ and $g \in \mathfrak{D}_{f(a)}(\mathbb{R}^n, \mathbb{R}^m)$, then*

$$\big[d(g \circ f)_a\big] = \big[\langle \nabla g^j(f(a)), f_k(a)\rangle\big]_{j=1}^{m}\,{}_{k=1}^{p}\,.$$

27.4.3. Exercise. This is an exercise in translation of notation. Suppose $y = y(u, v, w, x)$ and $z = z(u, v, w, x)$ where $u = u(s,t)$, $v = v(s,t)$, $w = w(s,t)$, and $x = x(s,t)$. Show that (under suitable hypotheses)

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial t} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial z}{\partial x}\frac{\partial x}{\partial t}\,.$$

(Solution Q.27.12.)

27.4.4. Problem. Suppose that the variables x, y, and z are differentiable functions of the variables α, β, γ, δ, and ϵ, which in turn depend in a differentiable fashion on the variables r, s, and t. As in exercise 27.4.3 use proposition 27.4.1 to write $\frac{\partial z}{\partial r}$ in terms of quantities such as $\frac{\partial z}{\partial \alpha}$, $\frac{\partial \delta}{\partial r}$, *etc.*

27.4.5. Exercise. Let $f(x, y, z) = (xy^2, 3x - z^2, xyz, x^2 + y^2, 4xz + 5)$, $g(s, t, u, v, w) = (s^2 + u^2 + v^2, s^2v - 2tw^2)$, and $a = (1, 0, -1)$. Use the *chain rule* to find $\big[d(g \circ f)_a\big]$. (Solution Q.27.13.)

27.4.6. Problem. Let $f(x, y, z) = (x^3y^2 \sin z, x^2 + y \cos z)$, $g(u, v) = (\sqrt{u}v, v^3)$, $k = g \circ f$, $a = (1, -2, \pi/2)$, and $h = (1, -1, 2)$. Use the *chain rule* to find $dk_a(h)$.

27.4.7. Problem. Let $f(x, y, z) = (x^2y + y^2z, xyz)$, $g(x, y) = (x^2y, 3xy, x - 2y, x^2 + 3)$, and $a = (1, -1, 2)$. Use the *chain rule* to find $\big[d(g \circ f)_a\big]$.

We now consider a slightly more complicated problem. Suppose that $w = w(x, y, t)$ where $x = x(s, t)$ and $y = y(s, t)$ and that all the functions mentioned are differentiable. (This is problem (D) at the beginning of this section.) It is perhaps tempting to write

$$\begin{aligned}\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial t}\frac{\partial t}{\partial t}\\ &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial t}\end{aligned} \tag{27.2}$$

(since $\frac{\partial t}{\partial t} = 1$). The trouble with this is that the $\frac{\partial w}{\partial t}$ on the left is not the same as the one on the right. The $\frac{\partial t}{\partial t}$ on the right refers only to the rate of change of w with respect to t insofar as t appears *explicitly* in the formula for w; the one on the left takes into account the fact that in addition w depends *implicitly* on t *via* the variables x and y. What to do? Use functions. Relate the variables by functions as follows.

$$\begin{matrix} s \\ t \end{matrix} \xrightarrow{\;f\;} \begin{matrix} x \\ y \\ t \end{matrix} \xrightarrow{\;g\;} w \tag{27.3}$$

Also let $h = g \circ f$. Notice that $f^3 = \pi_2$ (that is, $f^3(s,t) = t$). Then according to the *chain rule*

$$h_2 = \sum_{k=1}^{3} (g_k \circ f)\, f_2^k\,.$$

But $f_2^3 = 1$ (that is, $\frac{\partial t}{\partial t} = 1$). So

$$h_2 = (g_1 \circ f) f_2^1 + (g_2 \circ f) f_2^2 + g_3 \circ f. \tag{27.4}$$

The ambiguity of (27.2) has been eliminated in (27.4). The $\frac{\partial w}{\partial t}$ on the left is seen to be the derivative with respect to t of the composite $h = g \circ f$, whereas the $\frac{\partial w}{\partial t}$ on the right is just the derivative with respect to t of the function g.

One last point. Many scientific workers adamantly refuse to give names to functions. What do they do? Look back at diagram (27.3) and remove the names of the functions.

$$\begin{matrix} s \\ t \end{matrix} \longrightarrow \begin{matrix} x \\ y \\ t \end{matrix} \longrightarrow w \tag{27.5}$$

The problem is that the symbol "t" occurs twice. To specify differentiation of the composite function (our h) with respect to t, indicate that the "t" you are interested in is the one in the left column of (27.5). This may be done by listing everything else that appears in that column. That is, specify which variables are held constant. This specification conventionally appears as a subscript outside parentheses. Thus the $\frac{\partial w}{\partial t}$ on the left of (27.2) (our h_2) is written as $\left(\frac{\partial w}{\partial t}\right)_s$ (and is read, "$\frac{\partial w}{\partial t}$ with s held constant"). Similarly, the $\frac{\partial w}{\partial t}$ on the right of (27.2)) (our g_3) involves differentiation with respect to t while x and y are fixed. So it is written $\left(\frac{\partial w}{\partial t}\right)_{x,y}$ (and is read, "$\frac{\partial w}{\partial t}$ with x and y held constant). Thus (27.2) becomes

$$\left(\frac{\partial w}{\partial t}\right)_s = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \left(\frac{\partial w}{\partial t}\right)_{x,y} \tag{27.6}$$

It is not necessary to write, for example, an expression such as $\left(\frac{\partial w}{\partial x}\right)_{t,y}$ because there is no ambiguity; the symbol "x" occurs only once in (27.5). If you choose to use the convention just presented, it is best to use it only to avoid confusion; use it because you *must*, not because you *can*.

27.4.8. Exercise. Let $w = t^3 + 2yx^{-1}$ where $x = s^2 + t^2$ and $y = s \arctan t$. Use the *chain rule* to find $\left(\frac{\partial w}{\partial t}\right)_s$ at the point where $s = t = 1$. (Solution Q.27.14.)

We conclude this section with two more exercises on the use of the *chain rule*. Part of the difficulty here and in the problems at the end of the section is to interpret correctly what the problem says. The suggested solutions may seem longwinded, and they are. Nevertheless these techniques prove valuable in situations complicated enough to be confusing. With practice it is easy to do many of the indicated steps mentally.

27.4.9. Exercise. Show that if $z = xy + x\phi(yx^{-1})$, then $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = xy + z$. *Hint.* Start by restating the exercise in terms of functions. Add suitable hypotheses. In particular, suppose that $\phi\colon \mathbb{R} \to \mathbb{R}$ is differentiable. Let

$$j(x, y) = xy + x\phi(yx^{-1})$$

for x, $y \in \mathbb{R}$, $x \neq 0$. Then for each such x and y

$$xj_1(x, y) + yj_2(x, y) = xy + j(x, y)\,. \tag{27.7}$$

To prove this assertion proceed as follows.

(a) Let $g(x, y) = yx^{-1}$. Find $\left[dg_{(x,y)}\right]$.
(b) Find $\left[d(\phi \circ g)_{(x,y)}\right]$.
(c) Let $G(x, y) = \big(x, \phi(yx^{-1})\big)$. Use (b) to find $\left[dG_{(x,y)}\right]$.
(d) Let $m(x, y) = xy$. Find $\left[dm_{(x,y)}\right]$.
(e) Let $h(x, y) = x\phi(yx^{-1})$. Use (c) and (d) to find $\left[dh_{(x,y)}\right]$.
(f) Use (d) and (e) to find $\left[dj_{(x,y)}\right]$.
(g) Use (f) to prove (27.7).

(Solution Q.27.15.)

27.4.10. Exercise. Show that if $f(u, v) = g(x, y)$ where f is a differentiable real valued function, $u = x^2 - y^2$, and $v = 2xy$, then

$$y\frac{\partial g}{\partial x} - x\frac{\partial g}{\partial y} = 2v\frac{\partial f}{\partial u} - 2u\frac{\partial f}{\partial v} \tag{27.8}$$

Hint. The equations $u = x^2 - y^2$ and $v = 2xy$ give u and v in terms of x and y. Think of the function $h\colon (x, y) \mapsto (u, v)$ as a change of variables in $\mathbb{R}^2$. That is, define

$$h\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x, y) \mapsto (x^2 - y^2, 2xy)\,.$$

Then on the uv-plane (that is, the codomain of h) the function f is real valued and differentiable. The equation $f(u, v) = g(x, y)$ serves only to fix notation. It indicates that g is the composite function $f \circ h$. We may visualize the situation thus.

$$\begin{matrix} x \\ y \end{matrix} \xrightarrow{\ h\ } \begin{matrix} u \\ v \end{matrix} \xrightarrow{\ f\ } w \tag{27.9}$$

where $g = f \circ h$.

Now what are we trying to prove? The conclusion (27.8) is clear enough if we evaluate the partial derivatives at the right place. Recalling that we have defined h so that $u = h^1(x, y)$ and $v = h^2(x, y)$, we may write (27.8) in the following form.

$$yg_1(x, y) - xg_2(x, y) = 2h^2(x, y)f_1(h(x, y)) - 2h^1(x, y)f_2(h(x, y))\,. \tag{27.10}$$

Alternatively we may write

$$\pi_2 g_1 - \pi_1 g_2 = 2h^2 f_1 - 2h^1 f_2$$

(where π_1 and π_2 are the usual coordinate projections). To verify (27.10) use the *chain rule* to find $\big[dg_{(x,y)}\big]$. (Solution Q.27.16.)

27.4.11. Problem. Let $w = \frac{1}{2}x^2y + \arctan(tx)$ where $x = t^2 - 3u^2$ and $y = 2tu$. Find $\left(\frac{\partial w}{\partial t}\right)_u$ when $t = 2$ and $u = -1$.

27.4.12. Problem. Let $z = \frac{1}{16}uw^2xy$ where $w = t^2 - u^2 + v^2$, $x = 2tu + tv$, and $y = 3uv$. Find $\left(\frac{\partial z}{\partial u}\right)_{t,v}$ when $t = 1$, $u = -1$, and $v = -2$.

27.4.13. Problem. If $z = f\left(\frac{x-y}{y}\right)$, then $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = 0$. State this precisely and prove it.

27.4.14. Problem. If ϕ is a differentiable function on an open subset of $\mathbb{R}^2$ and $w = \phi(u^2 - t^2, t^2 - u^2)$, then $t\frac{\partial w}{\partial u} + u\frac{\partial w}{\partial t} = 0$. *Hint.* Let $h(t, u) = (u^2 - t^2, t^2 - u^2)$ and $w = \psi(t, u)$ where $\psi = \phi \circ h$. Compute $\big[dh_{(t,u)}\big]$. Use the *chain rule* to find $\big[d\psi_{(t,u)}\big]$. Then simplify $t\psi_2(t, u) + u\psi_1(t, u)$.

27.4.15. Problem. If $f(u, v) = g(x, y)$ where f is a differentiable real valued function on $\mathbb{R}^2$ and if $u = x^3 + y^3$ and $v = xy$, then

$$x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y} = 3u\frac{\partial f}{\partial u} + 2v\frac{\partial f}{\partial v}\,.$$

27.4.16. Problem. Let $f(x, y) = g(r, \theta)$ where (x, y) are Cartesian coordinates and (r, θ) are polar coordinates in the plane. Suppose that f is differentiable at all (x, y) in $\mathbb{R}^2$.

(a) Show that except at the origin

$$\frac{\partial f}{\partial x} = (\cos\theta)\frac{\partial g}{\partial r} - \frac{1}{r}(\sin\theta)\frac{\partial g}{\partial \theta}\,.$$

(b) Find a similar expression for $\frac{\partial f}{\partial y}$.

Hint. Recall that Cartesian and polar coordinates are related by $x = r\cos\theta$ and $y = r\sin\theta$.

27.4.17. Problem. Let n be a fixed positive integer. A function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is HOMOGENEOUS OF DEGREE n if $f(tx, ty) = t^n f(x, y)$ for all t, x, $y \in \mathbb{R}$. If such a function f is differentiable, then

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = nf\,. \tag{27.11}$$

Hint. Try the following:

(a) Let $G(x, y, t) = (tx, ty)$. Find $\big[dG_{(t,x,y)}\big]$.

(b) Let $h = f \circ G$. Find $\left[dh_{(x,y,t)}\right]$.
(c) Let $H(x, y, t) = (t^n, f(x, y))$. Find $\left[dH_{(x,y,t)}\right]$.
(d) Let $k = m \circ H$ (where $m(u, v) = uv$). Find $\left[dk_{(x,y,t)}\right]$.
(e) By hypothesis $h = k$; so the answers to (b) and (d) must be the same. Use this fact to derive (27.11).

CHAPTER 28

INFINITE SERIES

It is perhaps tempting to think of an infinite series

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \dots$$

as being nothing but "addition performed infinitely often". This view of series is misleading, can lead to curious errors, and should not be taken seriously. Consider the infinite series

$$1 - 1 + 1 - 1 + 1 - 1 + \dots. \tag{28.1}$$

If we think of the "+" and "-" signs as functioning in essentially the same manner as the symbols we encounter in ordinary arithmetic, we might be led to the following "discovery".

$$\begin{align} 0 &= (1-1) + (1-1) + (1-1) + \dots \tag{28.2}\\ &= 1 - 1 + 1 - 1 + 1 - 1 + \dots \tag{28.3}\\ &= 1 - (1-1) - (1-1) - \dots \tag{28.4}\\ &= 1 - 0 - 0 - \dots \tag{28.5}\\ &= 1\,. \tag{28.6}\end{align}$$

One can be more inventive: If S is the sum of the series (28.1), then

$$1 - S = 1 - (1 - 1 + 1 - 1 + \dots) = 1 - 1 + 1 - 1 + 1 - 1 + \dots = S$$

from which it follows that $S = \frac{1}{2}$. This last result, incidentally, was believed (for quite different reasons) by both Leibniz and Euler. See [9]. The point here is that if an intuitive notion of infinite sums and plausible arguments lead us to conclude that $1 = 0 = \frac{1}{2}$, then it is surely crucial for us to exercise great care in defining and working with convergence of infinite series.

In the first section of this chapter we discuss convergence of series in arbitrary normed linear spaces. One reason for giving the definitions in this generality is that doing so is no more complicated than discussing convergence of series in $\mathbb{R}$. A second reason is that it displays with much greater clarity the role of completeness of the underlying space in questions of convergence. (See, in particular, propositions 28.1.17 and 28.3.2.) A final reason is that this generality is actually needed. We use it, for example, in the proofs of the *inverse and implicit function theorems* in the next chapter.

28.1. CONVERGENCE OF SERIES

28.1.1. Definition. Let (a_k) be a sequence in a normed linear space. For each n in $\mathbb{N}$ let $s_n = \sum_{k=1}^{n} a_k$. The vector s_n is the n^{TH} PARTIAL SUM of the sequence (a_k). As is true of sequences, we permit variations of this definition. For example, in section 28.4 on power series we consider sequences $\big(a_k\big)_{k=0}^{n}$ whose first term has index zero. In this case, of course, the proper definition of s_n is $\sum_{k=0}^{n} a_k$.

28.1.2. Exercise. Let $a_k = (-1)^{k+1}$ for each k in $\mathbb{N}$. For $n \in \mathbb{N}$ compute the n^{th} partial sum of the sequence (a_k). (Solution Q.28.1.)

28.1.3. Exercise. Let $a_k = 2^{-k}$ for each k in $\mathbb{N}$. For $n \in \mathbb{N}$ show that the n^{th} partial sum of the sequence (a_k) is $1 - 2^{-n}$. (Solution Q.28.2.)

28.1.4. Definition. Let (a_k) be a sequence in a normed linear space. The INFINITE SERIES $\sum_{k=1}^{\infty} a_k$ is defined to be the sequence (s_n) of partial sums of the sequence (a_k). We may also write $a_1 + a_2 + a_3 + \dots$ for $\sum_{k=1}^{\infty} a_k$. Again we permit variants of this definition. For example, the infinite series associated with the sequence $\big(a_k\big)_{k=0}^{\infty}$ is denoted by $\sum_{k=0}^{\infty} a_k$. Whenever the range of the summation index k is understood from context or is unimportant we may denote a series simply by $\sum a_k$.

28.1.5. Exercise. What are the infinite series associated with the sequences given in exercises 28.1.2 and 28.1.3? (Solution Q.28.3.)

28.1.6. Definition. Let (a_k) be a sequence in a normed linear space V. If the infinite series $\sum_{k=1}^{\infty} a_k$ (that is, the sequence of partial sums of (a_k)) converges to a vector b in V, then we say that the sequence (a_k) is SUMMABLE or, equivalently, that the series $\sum_{k=1}^{\infty} a_k$ is a CONVERGENT SERIES. The vector b is called the SUM of the series $\sum_{k=1}^{\infty} a_k$ and we write

$$\sum_{k=1}^{\infty} a_k = b\,.$$

It is clear that a necessary and sufficient condition for a series $\sum_{k=1}^{\infty} a_k$ to be convergent or, equivalently, for the sequence (a_k) to be summable, is that there exist a vector b in V such that

$$\left\| b - \sum_{k=1}^{n} a_k \right\| \to 0 \qquad \text{as} \qquad n \to \infty. \tag{28.7}$$

If a series does not converge we say that it is a DIVERGENT SERIES (or that it DIVERGES).

CAUTION. It is an old, if illogical, practice to use the same notation $\sum_{k=1}^{\infty} a_k$ for both the sum of a convergent series and the series itself. As a result of this convention, the statements "$\sum_{k=1}^{\infty} a_k$ converges to b" and "$\sum_{k=1}^{\infty} a_k = b$" are interchangeable. It is possible for this to cause confusion, although in practice it is usually clear from the context which use of the symbol $\sum_{k=1}^{\infty} a_k$ is intended. Notice however that since a divergent series has no sum, the symbol $\sum_{k=1}^{\infty} a_k$ for such a series is unambiguous; it can refer only to the series itself.

28.1.7. Exercise. Are the sequences (a_k) given in exercises 28.1.2 and 28.1.3 summable? If (a_k) is summable, what is the sum of the corresponding series $\sum_{k=1}^{\infty} a_k$? (Solution Q.28.4.)

28.1.8. Problem (Geometric Series)**.** Let a and r be real numbers.

(a) Show that if $|r| < 1$, then $\sum_{k=0}^{\infty} ar^k$ converges and

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}\,.$$

Hint. To compute the n^{th} partial sum s_n, use a technique similar to the one used in exercise 28.1.3. See also problem I.1.12.

(b) Show that if $|r| \geq 1$ and $a \neq 0$, then $\sum_{k=0}^{\infty} ar^k$ diverges. *Hint.* Look at the cases $r \geq 1$ and $r \leq -1$ separately.

28.1.9. Problem. Let $\sum a_k$ and $\sum b_k$ be convergent series in a normed linear space.

(a) Show that the series $\sum (a_k + b_k)$ also converges and that

$$\sum (a_k + b_k) = \sum a_k + \sum b_k\,.$$

Hint. Problem 22.3.9(a).

(b) Show that for every $\alpha \in \mathbb{R}$ the series $\sum (\alpha a_k)$ converges and that

$$\sum (\alpha a_k) = \alpha \sum a_k\,.$$

One very easy way of seeing that certain series do not converge is to observe that its terms do not approach 0. (The proof is given in the next proposition.) It is important not to confuse this assertion with its converse. *The condition $a_k \to 0$ does not guarantee that $\sum_{k=1}^{\infty} a_k$ converges.* (An example is given in example 28.1.11.

28.1.10. Proposition. *If $\sum_{k=1}^{\infty} a_k$ is a convergent series in a normed linear space, then $a_k \to 0$ as $k \to \infty$.*

Proof. Exercise. *Hint.* Write a_n in terms of the partial sums s_n and s_{n-1}. (Solution Q.28.5.)

28.1.11. Example. It is possible for a series to diverge even though the terms a_k approach 0. A standard example of this situation in $\mathbb{R}$ is the harmonic series $\sum_{k=1}^{\infty} 1/k$. The harmonic series diverges.

Proof. Exercise. *Hint.* Show that the difference of the partial sums s_{2p} and s_p is at least $1/2$. Assume that (s_n) converges. Use proposition 18.1.4. (Solution Q.28.6.)

28.1.12. Problem. Show that if $0 < p \le 1$, then the series $\sum_{k=1}^{\infty} k^{-p}$ diverges. *Hint.* Modify the argument used in 28.1.11.

28.1.13. Problem. Show that the series $\sum_{k=1}^{\infty} \frac{1}{k^2+k}$ converges and find its sum. *Hint.* $\frac{1}{k^2+k} = \frac{1}{k} - \frac{1}{k+1}$.

28.1.14. Problem. Use the preceding problem to show that the series $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges and that its sum is no greater than 2.

28.1.15. Problem. Show that the series $\sum_{k=4}^{\infty} \frac{1}{k^2-1}$ converges and find its sum.

28.1.16. Problem. Let $p \in \mathbb{N}$. Find the sum of the series $\sum_{k=1}^{\infty} \frac{(k-1)!}{(k+p)!}$. *Hint.* If $a_k = \frac{k!}{(k+p)!}$, what can you say about $\sum_{k=1}^{n}(a_{k-1} - a_k)$?

In *complete* normed linear spaces the elementary fact that a sequence is Cauchy if and only if it converges may be rephrased to give a simple necessary and sufficient condition for the convergence of series in the space.

28.1.17. Proposition (The Cauchy Criterion)**.** *Let V be a normed linear space. If the series $\sum a_k$ converges in V, then for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\left\|\sum_{k=m+1}^{n} a_k\right\| < \epsilon$ whenever $n > m \ge n_0$. If V is a Banach space, then the converse of this implication also holds.*

Proof. Exercise (Solution Q.28.7.)

The principal use of the preceding proposition is to shorten proofs. By invoking the *Cauchy criterion* one frequently can avoid explicit reference to the partial sums of the series involved. See, for example, proposition 28.3.5.

One obvious consequence of the *Cauchy criterion* is that the convergence of an infinite series is unaffected by changing any finite number of its terms. If $a_n = b_n$ for all n greater than some fixed integer n_0, then the series $\sum a_n$ converges if and only if the series $\sum b_n$ does.

The examples of infinite series we have looked at thus far are all series of real numbers. We now turn to series in the Banach space $\mathcal{B}(S, E)$ of bounded E valued functions on a set S (where E is a Banach space). Most of the examples we consider will be real valued functions on subsets of the real line.

First a word of caution: the notations $\sum_{k=1}^{\infty} f_k$ and $\sum_{k=1}^{\infty} f_k(x)$ can, depending on context, mean many different things. There are many ways in which sequences (and therefore series) of functions can converge. There are, among a host of others, uniform convergence, pointwise convergence, convergence in mean, and convergence in measure. Only the first two of these appear in this text. Since we regard $\mathcal{B}(S, E)$ as a Banach space under the uniform norm $\| \ \|_u$, it is not, strictly speaking, necessary for us to write "$f_n \to g\,(\text{unif})$" when we wish to indicate that the sequence (f_n) converges to g in the space $\mathcal{B}(S,E)$; writing "$f_n \to g$" is enough, because unless the contrary is explicitly stated uniform convergence is understood. Nevertheless, in the sequel we will frequently add the redundant "(unif)" just as a reminder that in the space $\mathcal{B}(S,E)$ we are dealing with uniform, and not some other type of, convergence of sequences and series.

It is important to keep in mind that in the space $\mathcal{B}(S,E)$ the following assertions are equivalent:

(a) $\sum_{k=1}^{\infty} f_k$ converges uniformly to g;
(b) $g = \sum_{k=1}^{\infty} f_k$; and
(c) $\|g - \sum_{k=1}^{n} f_k\|_u \to 0$ as $n \to \infty$.

Since uniform convergence implies pointwise convergence, but not conversely, each of the preceding three conditions implies—*but is not implied by*—the following three (which are also equivalent):

(a′) $\sum_{k=1}^{\infty} f_k(x)$ converges to $g(x)$ for every x in S;
(b′) $g(x) = \sum_{k=1}^{\infty} f_k(x)$ for every x in S; and
(c′) For every x in S

$$\left\| g(x) - \sum_{k=1}^{n} f_k(x) \right\| \to 0 \quad \text{as} \quad n \to \infty .$$

One easy consequence of the *Cauchy criterion* (proposition 28.1.17) is called the *Weierstrass M-test*. The rather silly name which is attached to this result derives from the fact that in the statement of the proposition, the constants which appear are usually named M_n.

28.1.18. Proposition (Weierstrass M-test)**.** *Let (f_n) be a sequence of functions in $\mathcal{B}(S,E)$ where S is a nonempty set and E is a Banach space. If there is a summable sequence of positive constants M_n such that $\|f_n\|_u \le M_n$ for every n in $\mathbb{N}$, then the series $\sum f_k$ converges uniformly on S. Furthermore, if the underlying set S is a metric space and each f_n is continuous, then $\sum f_k$ is continuous.*

Proof. Problem.

28.1.19. Exercise. Let $0 < \delta < 1$. Show that the series $\sum_{k=1}^{\infty} \frac{x^k}{1+x^k}$ converges uniformly on the interval $[-\delta, \delta]$. *Hint.* Use problem 28.1.8. (Solution Q.28.8.)

28.1.20. Problem. Show that $\sum_{n=1}^{\infty} \frac{1}{1+n^2x}$ converges uniformly on $[\delta, \infty)$ for any δ such that $0 < \delta < 1$.

28.1.21. Problem. Let $M > 0$. Show that $\sum_{n=1}^{\infty} \frac{n^2x^3}{n^4+x^4}$ converges uniformly on $[-M, M]$.

28.1.22. Problem. Show that $\sum_{n=1}^{\infty} \frac{nx}{n^4+x^4}$ converges uniformly on $\mathbb{R}$.

We conclude this section with a generalization of the alternating series test, familiar from beginning calculus. Recall that an ALTERNATING SERIES in $\mathbb{R}$ is a series of the form $\sum_{k=1}^{\infty}(-1)^{k+1}\alpha_k$ where each $\alpha_k > 0$. The generalization here will not require that the multipliers of the α_k's be $+1$ and -1 in strict alternation. Indeed they need not even be real numbers; they may be the terms of any sequence of vectors in a Banach space for which the corresponding sequence of partial sums is bounded. You are asked in problem 28.1.25 to show that the *alternating series test* actually follows from the next proposition.

28.1.23. Proposition. *Let (α_k) be a decreasing sequence of real numbers, each greater than or equal to zero, which converges to zero. Let (x_k) be a sequence of vectors in a Banach space for which there exists $M > O$ such that $\|\sum_{k=1}^{n} x_k\| \le M$ for all n in $\mathbb{N}$. Then $\sum_{k=1}^{\infty} \alpha_k x_k$ converges and $\|\sum_{k=1}^{\infty} \alpha_k x_k\| \le M\alpha_1$.*

Proof. Problem. *Hint.* This is a bit complicated. Start by proving the following very simple geometrical fact about a normed linear space V: if $[x, y]$ is a closed segment in V, then one of its endpoints is at least as far from the origin as every other point in the segment. Use this to derive the fact that if x and y are vectors in V and $0 \le t \le 1$, then

$$\|x + ty\| \le \max\{\|x\|, \|x + y\| .$$

Next prove the following result.

28.1.24. Lemma. *Let (α_k) be a decreasing sequence of real numbers with $\alpha_k \ge 0$ for every k, let $M > 0$, and let V be a normed linear space. If $x_1, \dots, x_n \in V$ satisfy*

$$\left\| \sum_{k=1}^{m} x_k \right\| \le M \tag{28.8}$$

for all $m \le n$, then

$$\left\| \sum_{k=1}^{n} \alpha_k x_k \right\| \le M\alpha_1 .$$

To prove this result use mathematical induction. Supposing the lemma to be true for $n = p$, let $y_1, \dots, y_{p+1}$ be vectors in V such that $\|\sum_{k=1}^{m} y_k\| \le M$ for all $m \le p + 1$. Let $x_k = y_k$ for $k = 1, \dots, p - 1$ and let $x_p = y_p + (\alpha_{p+1}/\alpha_p)y_{p+1}$. Show that the vectors $x_1, \dots, x_p$ satisfy (28.8) for all $m \le p$ and invoke the inductive hypothesis.

Once the lemma is in hand, apply it to the sequence $\big(x_k\big)_{k=1}^{\infty}$ to obtain $\|\sum_{k=1}^{n} \alpha_k x_k\| \le M\alpha_1$ for all $n \in \mathbb{N}$, and apply it to the sequence $\big(x_k\big)_{k=m+1}^{\infty}$ to obtain $\big\|\sum_{k=m+1}^{n} \alpha_k x_k\big\| \le 2M\alpha_{m+1}$ for $0 < m < n$. Use this last result to prove that the sequence of partial sums of the series $\sum \alpha_k x_k$ is Cauchy.

28.1.25. Problem. Use proposition 28.1.23 to derive the *alternating series test*: If (α_k) is a decreasing sequence of real numbers with $\alpha_k \ge 0$ for all k and if $\alpha_k \to 0$ as $k \to \infty$, then the alternating series $\sum_{k=1}^{\infty} (-1)^{k+1}\alpha_k$ converges. Furthermore, the absolute value of the difference between the sum of the series and its n^{th} partial sum is no greater than α_{n+1}.

28.1.26. Problem. Show that the series $\sum_{k=1}^{\infty} k^{-1}\sin(k\pi/4)$ converges.

An important and interesting result in analysis is the *Tietze extension theorem*. For compact metric spaces it says that any continuous real valued function defined on a closed subset of the space can be extended to a continuous function on the whole space and that this process can be carried out in such a way that the (uniform) norm of the extension does not exceed the norm of the original function. One proof of this uses both the M-*test* and the *approximation theorem* of Weierstrass.

28.1.27. Theorem (Tietze Extension Theorem)**.** *Let M be a compact metric space, A be a closed subset of M, and $g\colon A \to \mathbb{R}$ be continuous. Then there exists a continuous function $w\colon M \to \mathbb{R}$ such that $w\big|_A = g$ and $\|w\|_u = \|g\|_u$.*

Proof. Problem. *Hint.* First of all demonstrate that a continuous function can be truncated without disturbing its continuity. (Precisely: if $f\colon M \to \mathbb{R}$ is a continuous function on a metric space, if $A \subseteq M$, and if $f^{\to}(A) \subseteq [a, b]$, then there exists a continuous function $g\colon M \to \mathbb{R}$ which agrees with f on A and whose range is contained in $[a, b]$.) Let $\mathcal{F} = \{u\big|_A \colon u \in \mathcal{C}(M, \mathbb{R})\}$. Notice that the preceding comment reduces the proof of 28.1.27 to showing that $\mathcal{F} = \mathcal{C}(A, \mathbb{R})$. Use the *Stone-Weierstrass theorem* 23.2.6 to prove that $\mathcal{F}$ is dense in $\mathcal{C}(A, \mathbb{R})$. Next find a sequence (f_n) of functions in $\mathcal{F}$ such that

$$\left\| g - \sum_{k=1}^{n} f_k \right\|_u < \frac{1}{2^n}$$

for every n. Then for each k find a function u_k in $\mathcal{C}(M, \mathbb{R})$ whose restriction to A is f_k. Truncate each u_k (as above) to form a new function v_k which agrees with u_k (and therefore f_k) on A and which satisfies $\|v_k\|_u = \|f_k\|_u$. Use the *Weierstrass M-test* 28.1.18 to show that $\sum_1^\infty v_k$ converges uniformly on M. Show that $w = \sum_1^\infty v_k$ is the desired extension.

Recall that in problem 23.1.22 we showed that if $\phi\colon M \to N$ is a continuous map between compact metric spaces, then the induced map $T_\phi\colon \mathcal{C}(N,\mathbb{R}) \to \mathcal{C}(M,\mathbb{R})$ is injective if and only if ϕ is surjective, and ϕ is injective if T_ϕ is surjective. The "missing part" of this result (T_ϕ is surjective if ϕ is injective) happens also to be true but requires the *Tietze extension theorem* for its proof.

28.1.28. Problem. Let ϕ and T_ϕ be as in problem 23.1.22. Show that if ϕ is injective, then T_ϕ is surjective.

28.2. SERIES OF POSITIVE SCALARS

In this brief section we derive some of the standard tests of beginning calculus for convergence of series of positive numbers.

28.2.1. Definition. Let $S(n)$ be a statement in which the natural number n is a variable. We say that $S(n)$ holds for n SUFFICIENTLY LARGE if there exists N in $\mathbb{N}$ such that $S(n)$ is true whenever $n \geq N$.

28.2.2. Proposition (Comparison Test)**.** *Let (a_k) and (b_k) be sequences in $[0,\infty)$ and suppose that there exists $M > 0$ such that $a_k \leq Mb_k$ for sufficiently large $k \in \mathbb{N}$. If $\sum b_k$ converges, then $\sum a_k$ converges. If $\sum a_k$ diverges, so does $\sum b_k$.*

PROOF. Problem.

28.2.3. Proposition (Ratio Test)**.** *Let (a_k) be a sequence in $(0,\infty)$. If there exists $\delta \in (0,1)$ such that $a_{k+1} \leq \delta a_k$ for k sufficiently large, then $\sum a_k$ converges. If there exists $M > 1$ such that $a_{k+1} \geq M a_k$ for k sufficiently large, then $\sum a_k$ diverges.*

PROOF. Exercise. (Solution Q.28.9.)

28.2.4. Proposition (Integral Test)**.** *Let $f\colon [1,\infty) \to [0,\infty)$ be decreasing; that is, $f(x) \geq f(y)$ whenever $x < y$. If $\lim_{M\to\infty} \int_1^M f$ exists, then $\sum_1^\infty f(k)$ converges. If $\lim_{M\to\infty} \int_1^M f$ does not exist, then $\sum_1^\infty f(k)$ diverges.*

PROOF. Problem. *Hint.* Show that $\int_k^{k+1} f \leq f(k) \leq \int_{k-1}^k f$ for $k \geq 2$.

28.2.5. Proposition (The Root Test)**.** *Let $\sum a_k$ be a series of numbers in $[0,\infty)$. Suppose that the limit $L = \lim_{k\to\infty} (a_k)^{1/k}$ exists.*

(a) *If $L < 1$, then $\sum a_k$ converges.*
(b) *If $L > 1$, then $\sum a_k$ diverges.*

PROOF. Problem.

28.2.6. Problem. Let (a_k) and (b_k) be decreasing sequences in $(0,\infty)$ and let $c_k = \min\{a_k, b_k\}$ for each k. If $\sum a_k$ and $\sum b_k$ both diverge, must $\sum c_k$ also diverge?

28.3. ABSOLUTE CONVERGENCE

It is a familiar fact from beginning calculus that absolute convergence of a series of real numbers implies convergence of the series. The proof of this depends in a crucial way on the completeness of $\mathbb{R}$. We show in the first proposition of this section that for series in a normed linear space V absolute convergence implies convergence if *and only if* V is complete.

28.3.1. Definition. Let (a_k) be a sequence in a normed linear space V. We say that (a_k) is ABSOLUTELY SUMMABLE or, equivalently, that the series $\sum a_k$ CONVERGES ABSOLUTELY if the series $\sum \|a_k\|$ converges in $\mathbb{R}$.

28.3.2. Proposition. *A normed linear space V is complete if and only if every absolutely summable sequence in V is summable.*

Proof. Exercise. *Hint.* If V is complete, the *Cauchy criterion* 28.1.17 may be used. For the converse, suppose that every absolutely summable sequence is summable. Let (a_k) be a Cauchy sequence in V. Find a subsequence (a_{n_k}) such that $\|a_{n_{k+1}} - a_{n_k}\| < 2^{-k}$ for each k. Consider the sequence (y_k) where $y_k := a_{n_{k+1}} - a_{n_k}$ for all k. (Solution Q.28.10.)

One of the most useful consequences of absolute convergence of a series is that the terms of the series may be rearranged without affecting the sum of the series. This is not true of CONDITIONALLY CONVERGENT series (that is, series which converge but do not converge absolutely). One can show, in fact, that a conditionally convergent series of real numbers can, by rearrangement, be made to converge to any real number whatever, or, for that matter, to diverge. We will not demonstrate this here, but a nice proof can be found in [1].

28.3.3. Definition. A series $\sum_{k=1}^{\infty} b_k$ is said to be a REARRANGEMENT of the series $\sum_{k=1}^{\infty} a_k$ if there exists a bijection $\phi\colon \mathbb{N} \to \mathbb{N}$ such that $b_k = a_{\phi(k)}$ for all k in $\mathbb{N}$.

28.3.4. Proposition. *If $\sum b_k$ is a rearrangement of an absolutely convergent series $\sum a_k$ in a Banach space, then $\sum b_k$ is itself absolutely convergent and it converges to the same sum as $\sum a_k$.*

Proof. Problem *Hint.* Let $\beta_n := \sum_{k=1}^{n} \|b_k\|$. Show that the sequence (β_n) is increasing and bounded. Conclude that $\sum b_k$ is absolutely convergent. The hard part of the proof is showing that if $\sum_1^\infty a_k$ converges to a vector A, then so does $\sum_1^\infty b_k$. Define partial sums as usual: $s_n := \sum_1^n a_k$ and $t_n := \sum_1^n b_k$. Given $\epsilon > 0$, you want to show that $\|t_n - A\| < \epsilon$ for sufficiently large n. Prove that there exists a positive N such that $\|s_n - A\| < \frac{1}{2}\epsilon$ and $\sum_n^\infty \|a_k\| \le \frac{1}{2}\epsilon$ whenever $n \ge N$. Write

$$\|t_n - A\| \le \|t_n - s_N\| + \|s_N - A\| .$$

Showing that $\|t_n - s_N\| \le \frac{1}{2}\epsilon$ for n sufficiently large takes a little thought. For an appropriate function ϕ write $b_k = a_{\phi(k)}$. Notice that

$$\|t_n - s_N\| = \left\| \sum_1^n a_{\phi(k)} - \sum_1^N a_k \right\| .$$

The idea of the proof is to choose n so large that there are enough terms $a_{\phi(j)}$ to cancel *all* the terms a_k $(1 \le k \le N)$.

If you have difficulty in dealing with sums like $\sum_{k=1}^n a_{\phi(k)}$ whose terms are not consecutive ($a_1, a_2, \dots$ are consecutive terms of the sequence (a_k); $a_{\phi(1)}, a_{\phi(2)}, \dots$ in general are not), a notational trick may prove useful. For P a *finite* subset of $\mathbb{N}$, write $\sum\limits_P a_k$ for the sum of all the terms a_k such that k belongs to P. This notation is easy to work with. It should be easy to convince yourself that, for example, if P and Q are finite subsets of $\mathbb{N}$ and if they are disjoint, then $\sum\limits_{P \cup Q} a_k = \sum\limits_P a_k + \sum\limits_Q a_k$. (What happens if $P \cap Q \ne \emptyset$?) In the present problem, let $C := \{1, \dots, N\}$ (where N is the integer chosen above). Give a careful proof that there exists an integer p such that the set $\{\phi(1), \dots, \phi(p)\}$ contains C. Now suppose n is any integer greater than p. Let $F := \{\phi(1), \dots, \phi(n)\}$ and show that

$$\|t_n - s_N\| \le \sum_G \|a_k\|$$

where $G := F \setminus C$.

28.3.5. Proposition. *If (α_n) is an absolutely summable sequence in $\mathbb{R}$ and (x_n) is a bounded sequence in a Banach space E, then the sequence $(\alpha_n x_n)$ is summable in E.*

Proof. Problem. *Hint.* Use the Cauchy criterion 28.1.17.

28.3.6. Problem. What happens in the previous proposition if the sequence (α_n) is assumed only to be bounded and the sequence (x_n) is absolutely summable?

28.3.7. Problem. Show that if the sequence (a_n) of real numbers is SQUARE SUMMABLE (that is, if the sequence $({a_n}^2)$ is summable), then the series $\sum n^{-1}a_n$ converges absolutely. *Hint.* Use the *Schwarz inequality* 27.1.6.

28.4. POWER SERIES

According to problem 28.1.8 we may express the reciprocal of the real number $1-r$ as the sum of a power series $\sum_0^\infty r^k$ provided that $|r|<1$. One may reasonably ask if anything like this is true in Banach spaces other than $\mathbb{R}$, in the space of bounded linear maps from some Banach space into itself, for example. If T is such a map and if $\|T\|<1$, is it necessarily true that $I-T$ is invertible? And if it is, can the inverse of $I-T$ be expressed as the sum of the power series $\sum_0^\infty T^k$? It turns out that the answer to both questions is *yes*. Our interest in pursuing this matter is not limited to the fact that it provides an interesting generalization of facts concerning $\mathbb{R}$ to spaces with richer structure. In the next chapter we will need exactly this result for the proof we give of the *inverse function theorem*.

Of course it is not possible to study power series in arbitrary Banach spaces; there are, in general, no powers of vectors because there is no multiplication. Thus we restrict our attention to those Banach spaces which (like the space of bounded linear maps) are equipped with an additional operation $(x,y)\mapsto xy$ (we call it *multiplication*) under which they become linear associative algebras (see 21.2.6 for the definition) and on which the norm is SUBMULTIPLICATIVE (that is, $\|xy\|\le\|x\|\,\|y\|$ for all x and y). We will, for simplicity, insist further that these algebras be unital and that the multiplicative identity $\mathbf{1}$ have norm one. Any Banach space thus endowed is a (UNITAL) BANACH ALGEBRA. It is clear that Banach algebras have many properties in common with $\mathbb{R}$. It is important, however, to keep firmly in mind those properties *not* shared with $\mathbb{R}$. Certainly there is, in general, no linear ordering $<$ of the elements of a Banach algebra (or for that matter of a Banach space). Another crucial difference is that in $\mathbb{R}$ every nonzero element has a multiplicative inverse (its reciprocal); this is not true in general Banach algebras (see, for example, proposition21.5.16). Furthermore, Banach algebras may have nonzero NILPOTENT elements (that is, elements $x\ne0$ such that $x^n=0$ for some natural number n). (Example: the 2×2 matrix $a=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ is not zero, but $a^2=0$.) This,of course, prevents us from requiring that the norm be multiplicative: while $|xy|=|x|\,|y|$ holds in $\mathbb{R}$, all that is true in general Banach algebras is $\|xy\|\le\|x\|\,\|y\|$. Finally, multiplication in Banach algebras need not be commutative.

28.4.1. Definition. Let A be a normed linear space. Suppose there is an operation $(x,y)\mapsto xy$ from $A\times A$ into A satisfying the following: for all x, y, $z\in A$ and $\alpha\in\mathbb{R}$

(a) $(xy)z=x(yz)$,
(b) $(x+y)z=xz+yz$,
(c) $x(y+z)=xy+xz$,
(d) $\alpha(xy)=(\alpha x)y=x(\alpha y)$, and
(e) $\|xy\|\le\|x\|\,\|y\|$.

Suppose additionally that there exists a vector $\mathbf{1}$ in A such that

(f) $x\,\mathbf{1}=\mathbf{1}\,x=x$ for all $x\in A$, and
(g) $\|\mathbf{1}\|=1$.

Then A is a (UNITAL) NORMED ALGEBRA. If A is complete it is a (UNITAL) BANACH ALGEBRA. If all elements x and y in a normed (or Banach) algebra satisfy $xy=yx$, then the algebra A is COMMUTATIVE.

28.4.2. Example. The set $\mathbb{R}$ of real numbers is a commutative Banach algebra. The number 1 is the multiplicative identity.

28.4.3. Example. If S is a nonempty set, then with pointwise multiplication

$$(fg)(x) := f(x)g(x) \qquad \text{for all } x \in S$$

the Banach space $\mathcal{B}(S, \mathbb{R})$ becomes a commutative Banach algebra. The constant function $\mathbf{1}\colon x \mapsto 1$ is the multiplicative identity.

28.4.4. Exercise. Show that if f and g belong to $\mathcal{B}(S, \mathbb{R})$, then $\|fg\|_u \le \|f\|_u \|g\|_u$. Show by example that equality need not hold. (Solution Q.28.11.)

28.4.5. Example. If M is a compact metric space then (again with pointwise multiplication) $\mathcal{C}(M, \mathbb{R})$ is a commutative Banach algebra. It is a SUBALGEBRA of $\mathcal{B}(M, \mathbb{R})$ (that is, a subset of $\mathcal{B}(M, \mathbb{R})$ containing the multiplicative identity which is a Banach algebra under the induced operations).

28.4.6. Example. If E is a Banach space, then $\mathfrak{B}(E, E)$ is a Banach algebra (with composition as "multiplication"). We have proved in problem 21.2.8 that the space of linear maps from E into E is a unital algebra. The same is easily seen to be true of $\mathfrak{B}(E, E)$ the corresponding space of bounded linear maps. The identity transformation I_E is the multiplicative identity. Its norm is 1 by 23.1.11(a). In proposition 23.1.14 it was shown that $\mathfrak{B}(E, E)$ is a normed linear space; the submultiplicative property of the norm on this space was proved in proposition 23.1.15. Completeness was proved in proposition 23.3.6.

28.4.7. Proposition. *If A is a normed algebra, then the operation of multiplication*

$$M\colon A \times A \to A\colon (x, y) \mapsto xy$$

is continuous.

PROOF. Problem. *Hint.* Try to adapt the proof of example 14.1.9.

28.4.8. Corollary. *If $x_n \to a$ and $y_n \to b$ in a normed algebra, then $x_n y_n \to ab$.*

PROOF. Problem.

28.4.9. Definition. An element x of a unital algebra A (with or without norm) is INVERTIBLE if there exists an element x^{-1} (called the MULTIPLICATIVE INVERSE of x) such that $xx^{-1} = x^{-1}x = \mathbf{1}$. The set of all invertible elements of A is denoted by Inv A. We list several almost obvious properties of inverses.

28.4.10. Proposition. *If A is a unital algebra, then*

(a) *Each element of A has at most one multiplicative inverse.*
(b) *The multiplicative identity $\mathbf{1}$ of A is invertible and $\mathbf{1}^{-1} = \mathbf{1}$.*
(c) *If x is invertible, then so is x^{-1} and $\left(x^{-1}\right)^{-1} = x$.*
(d) *If x and y are invertible, then so is xy and $(xy)^{-1} = y^{-1}x^{-1}$.*
(e) *If x and y are invertible, then $x^{-1} - y^{-1} = x^{-1}(y - x)y^{-1}$.*

PROOF. Let $\mathbf{1}$ be the multiplicative identity of A.

(a) If y and z are multiplicative inverses of x, then $y = y\,\mathbf{1} = y(xz) = (yx)z = \mathbf{1}\,z = z$.
(b) $\mathbf{1} \cdot \mathbf{1} = \mathbf{1}$ implies $\mathbf{1}^{-1} = \mathbf{1}$ by (a).
(c) $x^{-1}x = xx^{-1} = 1$ implies x is the inverse of x^{-1} by (a).
(d) $(xy)(y^{-1}x^{-1}) = xx^{-1} = \mathbf{1}$ and $(y^{-1}x^{-1})(xy) = y^{-1}y = \mathbf{1}$ imply $y^{-1}x^{-1}$ is the inverse of xy (again by (a)).
(e) $x^{-1}(y - x)y^{-1} = (x^{-1}y - 1)y^{-1} = x^{-1} - y^{-1}$.

□

28.4.11. Proposition. *If x is an element of a unital Banach algebra and $\|x\| < 1$, then $\mathbf{1} - x$ is invertible and $(\mathbf{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k$.*

PROOF. Exercise. *Hint.* First show that the geometric series $\sum_{k=0}^{\infty} x^k$ converges absolutely. Next evaluate $(\mathbf{1}-x)s_n$ and $s_n(\mathbf{1}-x)$ where $s_n = \sum_{k=0}^{n} x^k$. Then take limits as $n \to \infty$. (Solution Q.28.12.) □

28.4.12. Corollary. *If x is an element of a unital Banach algebra and $\|x\| < 1$, then*

$$\|(\mathbf{1}-x)^{-1} - \mathbf{1}\| \le \frac{\|x\|}{1-\|x\|}\,.$$

PROOF. Problem.

Proposition 28.4.11 says that anything close to $\mathbf{1}$ in a Banach algebra A is invertible. In other words $\mathbf{1}$ is an interior point of $\operatorname{Inv} A$. Corollary 28.4.12 says that if $\|x\|$ is small, that is, if $\mathbf{1}-x$ is close to $\mathbf{1}$, then $(\mathbf{1}-x)^{-1}$ is close to $\mathbf{1}^{-1}$ $(= \mathbf{1})$. In other words, the operation of inversion $x \mapsto x^{-1}$ is continuous at $\mathbf{1}$. These results are actually special cases of much more satisfying results: every point of $\operatorname{Inv} A$ is an interior point of that set (that is, $\operatorname{Inv} A$ is open); and the operation of inversion is continuous at every point of $\operatorname{Inv} A$. We use the special cases above to prove the more general results.

28.4.13. Proposition. *If A is a Banach algebra, then* $\operatorname{Inv} A \overset{\circ}{\subseteq} A$.

PROOF. Problem. *Hint.* Let $a \in \operatorname{Inv} A$, $r = \|a^{-1}\|^{-1}$, $y \in B_r(a)$, and $x = a^{-1}(a-y)$. What can you say about $a(\mathbf{1}-x)$?

28.4.14. Proposition. *If A is a Banach algebra, then the operation of inversion*

$$r\colon \operatorname{Inv} A \to \operatorname{Inv} A\colon x \mapsto x^{-1}$$

is continuous.

PROOF. Exercise. *Hint.* Show that r is continuous at an arbitrary point a in $\operatorname{Inv} A$. Given $\epsilon > 0$ find $\delta > 0$ sufficiently small that if y belongs to $B_\delta(a)$ and $x = \mathbf{1} - a^{-1}y$, then

$$\|x\| < \|a^{-1}\|\delta < \tfrac{1}{2}$$

and

$$\|r(y) - r(a)\| \le \frac{\|x\|\|a^{-1}\|}{1-\|x\|} < \epsilon \tag{28.9}$$

For (28.9) use 28.4.10(e), the fact that $y^{-1}a = \left(a^{-1}y\right)^{-1}$, and 28.4.12. (Solution Q.28.13.)

There are a number of ways of multiplying infinite series. The most common is the *Cauchy product*. And it is the only one we will consider.

28.4.15. Definition. If $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are infinite series in a Banach algebra, then their CAUCHY PRODUCT is the series $\sum_{n=0}^{\infty} c_n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. To see why this definition is rather natural, imagine trying to multiply two power series $\sum_0^\infty a_k x^k$ and $\sum_0^\infty b_k x^k$ just as if they were infinitely long polynomials. The result would be another power series. The coefficient of x^n in the resulting series would be the sum of all the products $a_i b_j$ where $i+j=n$. There are several ways of writing this sum

$$\sum_{i+j=n} a_i b_j = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} a_{n-k} b_k\,.$$

If we just forget about the variable x, we have the preceding definition of the Cauchy product.

The first thing we observe about these products is that convergence of both the series $\sum a_k$ and $\sum b_k$ does *not* imply convergence of their Cauchy product.

28.4.16. Example. Let $a_k = b_k = (-1)^k(k+1)^{-1/2}$ for all $k \geq 0$. Then $\sum_0^\infty a_k$ and $\sum_0^\infty b_k$ converge by the alternating series test (problem 28.1.25). The n^{th} term of their Cauchy product is

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k+1}}\,.$$

Since for $0 \leq k \leq n$

$$\begin{aligned}(k+1)(n-k+1) &= k(n-k) + n + 1 \\ &\leq n^2 + 2n + 1 \\ &= (n+1)^2\end{aligned}$$

we see that

$$|c_n| = \sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k+1}} \geq \sum_{k=0}^{n} \frac{1}{n+1} = 1\,.$$

Since $c_n \nrightarrow 0$, the series $\sum_0^\infty c_k$ does not converge (see proposition 28.1.10).

Fortunately quite modest additional hypotheses do guarantee convergence of the Cauchy product. One very useful sufficient condition is that at least one of the series $\sum a_k$ or $\sum b_k$ converge absolutely.

28.4.17. Theorem (Mertens' Theorem)**.** *If in a Banach algebra $\sum_0^\infty a_k$ is absolutely convergent and has sum a and the series $\sum_0^\infty b_k$ is convergent with sum b, then the Cauchy product of these series converges and has sum ab.*

Proof. Exercise *Hint.* Although this exercise is slightly tricky, the difficulty has nothing whatever to do with Banach algebras. Anyone who can prove *Mertens' theorem* for series of real numbers can prove it for arbitrary Banach algebras.

For each $k \in \mathbb{N}$ let $c_k = \sum_{j=0}^{k} a_j b_{k-j}$. Let s_n, t_n, and u_n be the n^{th} partial sums of the sequences (a_k), (b_k), and (c_k), respectively. First verify that for every n in $\mathbb{N}$

$$u_n = \sum_{k=0}^{n} a_{n-k} t_k \tag{28.10}$$

To this end define, for $0 \leq j, k \leq n$, the vector d_{jk} by

$$d_{jk} = \begin{cases} a_j b_{k-j}, & \text{if } j \leq k \\ 0, & \text{if } j > k. \end{cases}$$

Notice that both the expression which defines u_n and the expression on the right side of equation (28.10) involve only finding the sum of the elements of the matrix $[d_{jk}]$ but in different orders.

From (28.10) it is easy to see that for every n

$$u_n = s_n b + \sum_{k=0}^{n} a_{n-k}(t_k - b).$$

Since $s_n b \to ab$ as $n \to \infty$, the proof of the theorem is reduced to showing that

$$\|\sum_{k=0}^{n} a_{n-k}(t_k - b)\| \to 0 \qquad \text{as } n \to \infty.$$

Let $\alpha_k = \|a_k\|$ and $\beta_k = \|t_k - b\|$ for all k. Why does it suffice to prove that $\sum_{k=0}^{n} \alpha_{n-k}\beta_k \to 0$ as $n \to \infty$? In the finite sum

$$\alpha_n\beta_0 + \alpha_{n-1}\beta_1 + \cdots + \alpha_0\beta_n \tag{28.11}$$

the α_k's towards the left are small (for large n) and the β_k's towards the right are small (for large n). This suggests breaking the sum (28.11) into two pieces

$$p = \sum_{k=0}^{n_1} \alpha_{n-k}\beta_k \qquad \text{and} \qquad q = \sum_{k=n_1+1}^{n} \alpha_{n-k}\beta_k$$

and trying to make each piece small (smaller, say, than $\frac{1}{2}\epsilon$ for a preassigned ϵ).

For any positive number ϵ_1 it is possible (since $\beta_k \to 0$) to choose n_1 in $\mathbb{N}$ so that $\beta_k < \epsilon_1$ whenever $k \geq n_1$. What choice of ϵ_1 will ensure that $q < \frac{1}{2}\epsilon$?

For any $\epsilon_2 > 0$ it is possible (since $\alpha_k \to 0$) to choose n_2 in $\mathbb{N}$ so that $\alpha_k < \epsilon_2$ whenever $k \geq n_2$. Notice that $n - k \geq n_2$ for all $k \leq n_1$ provided that $n \geq n_1 + n_2$. What choice of ϵ_2 will guarantee that $p < \frac{1}{2}\epsilon$? (Solution Q.28.14.)

28.4.18. Proposition. *On a Banach algebra A the operation of multiplication*

$$M\colon A \times A \to A\colon (x, y) \mapsto xy$$

is differentiable.

Proof. Problem. *Hint.* Fix (a, b) in $A \times A$. Compute the value of the function $\Delta M_{(a,b)}$ at the point (h, j) in $A \times A$. Show that $\|(h, j)\|^{-1} hj \to 0$ in A as $(h, j) \to 0$ in $A \times A$. How should $dM_{(a,b)}$ be chosen so that the Newton quotient

$$\frac{\Delta M_{(a,b)}(h, j) - dM_{(a,b)}(h, j)}{\|(h, j)\|}$$

approaches zero (in A) as $(h, j) \to 0$ (in $A \times A$)? Don't forget to show that your choice for $dM_{(a,b)}$ is a bounded linear map.

28.4.19. Proposition. *Let $c \in V$ where V is a normed linear space and let A be a Banach algebra. If f, $g \in \mathcal{D}_c(V, A)$, then their product fg defined by*

$$(fg)(x) := f(x)\, g(x)$$

(for all x in some neighborhood of c) is differentiable at c and

$$d(fg)_c = f(c) dg_c + df_c \cdot g(c)\,.$$

Proof. Problem. *Hint.* Use proposition 28.4.18.

28.4.20. Proposition. *If A is a commutative Banach algebra and $n \in \mathbb{N}$, then the function $f\colon x \mapsto x^n$ is differentiable and $df_a(h) = na^{n-1}h$ for all a, $h \in A$.*

Proof. Problem. *Hint.* A simple induction proof works. Alternatively, you may choose to convince yourself that the usual form of the *binomial theorem* (see I.1.17) holds in every commutative Banach algebra.

28.4.21. Problem. What happens in 28.4.20 if we do not assume that the Banach algebra is commutative? Is f differentiable? *Hint.* Try the cases $n = 2$, 3, and 4. Then generalize.

28.4.22. Problem. Generalize proposition 28.3.5 to produce a theorem concerning the convergence of a series $\sum a_k b_k$ in a Banach algebra.

28.4.23. Definition. If (a_k) is a sequence in a Banach algebra A and $x \in A$, then a series of the form $\sum_{k=0}^{\infty} a_k x^k$ is a power series in x.

28.4.24. Notation. It is a bad (but very common) habit to use the same notation for a polynomial function and for its value at a point x. One often encounters an expression such as "the function $x^2 - x + 5$" when clearly what is meant is "the function $x \mapsto x^2 - x + 5$." This abuse of language is carried over to power series. If (a_k) is a sequence in a Banach algebra A and $D = \{x \in A\colon \sum_{k=0}^{\infty} a_k x^k \text{ converges}\}$, then the function $x \mapsto \sum_{k=0}^{\infty} a_k x^k$ from D into A is usually denoted by

$\sum_{k=0}^{\infty} a_k x^k$. Thus for example one may find the expression, "$\sum_{k=0}^{\infty} a_k x^k$ converges uniformly on the set U." What does this mean? Answer: if $s_n(x) = \sum_{k=0}^{n} a_k x^k$ for all $n \in \mathbb{N}$ and $f(x) = \sum_{k=0}^{\infty} a_k x^k$, then $s_n \to f$ (unif).

28.4.25. Proposition. *Let (a_k) be a sequence in a Banach algebra and $r > 0$. If the sequence $(\|a_k\| r^k)$ is bounded, then the power series $\sum_{k=0}^{\infty} a_k x^k$ converges uniformly on every open ball $B_s(0)$ such that $0 < s < r$. (And therefore it converges on the ball $B_r(0)$.)*

Proof. Exercise. *Hint.* Let $p = s/r$ where $0 < s < r$. Let $f_k(x) = a_k x^k$. Use the *Weierstrass M-test* 28.1.18. (Solution Q.28.15.)

Since the uniform limit of a sequence of continuous functions is continuous (see 14.2.15), it follows easily from preceding proposition that (under the hypotheses given) the function $f\colon x \mapsto \sum_0^{\infty} a_k x^k$ is continuous on $B_r(0)$. We will prove considerably more than this: the function is actually differentiable on $B_r(0)$, and furthermore, the correct formula for its differential is found by differentiating the power series term-by-term. That is, f behaves on $B_r(0)$ just like a "polynomial" with infinitely many terms. This is proved in 28.4.27. We need however a preliminary result.

28.4.26. Proposition. *Let V and W be normed linear spaces, U be an open convex subset of V, and (f_n) be a sequence of functions in $\mathcal{C}^1(U, W)$. If the sequence (f_n) converges pointwise to a function $F\colon U \to W$ and if the sequence $\big(d(f_n)\big)$ converges uniformly on U, then F is differentiable at each point a of U and*

$$dF_a = \lim_{n\to\infty} d\big(f_n\big)_a .$$

Proof. Exercise. *Hint.* Fix $a \in U$. Let $\phi\colon U \to \mathfrak{B}(V, W)$ be the function to which the sequence $\big(d(f_n)\big)$ converges uniformly. Given $\epsilon > 0$ show that there exists $N \in \mathbb{N}$ such that $\big\|d\big(f_n - f_N\big)_x\big\| < \epsilon/4$ for all $x \in U$ and $n \in \mathbb{N}$. Let $g_n = f_n - f_N$ for each n. Use one version of the *mean value theorem* to show that

$$\|\Delta\big(g_n\big)_a(h) - d\big(g_n\big)_a(h)\| \le \tfrac{1}{2}\epsilon\|h\| \tag{28.12}$$

whenever $n \ge N$ and h is a vector such that $a + h \in U$. In (28.12) take the limit as $n \to \infty$. Use the result of this together with the fact that $\Delta\big(f_N\big)_a \simeq d\big(f_N\big)_a$ to show that $\Delta F_a \simeq T$ when $T = \phi(a)$. (Solution Q.28.16.)

28.4.27. Theorem (Term-by-Term Differentiation of Power Series)**.** *Suppose that (a_n) is a sequence in a commutative Banach algebra A and that $r > 0$. If the sequence $(\|a_k\| r^k)$ is bounded, then the function $F\colon B_r(0) \to A$ defined by*

$$F(x) = \sum_{k=0}^{\infty} a_k x^k$$

is differentiable and

$$dF_x(h) = \sum_{k=1}^{\infty} k a_k x^{k-1} h$$

for every $x \in B_r(0)$ and $h \in A$.

Proof. Problem. *Hint.* Let $f_n(x) = \sum_0^n a_k x^k$. Fix x in $B_r(0)$ and choose a number s satisfying $\|x\| < s < r$. Use propositions 28.4.20 and 28.4.26 to show that F is differentiable at x and to compute its differential there. In the process you will need to show that the sequence $\big(d(f_n)\big)$ converges uniformly on $B_s(0)$. Use propositions 28.4.25 and 4.2.8. If $s < t < r$, then there exists $N \in \mathbb{N}$ such that $k^{1/k} < r/t$ for all $k \ge N$. (Why ?)

28.4.28. Problem. Give a definition of and develop the properties of the *exponential function* on a commutative Banach algebra A. Include at least the following:

(a) The series $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$ converges absolutely for all x in A and uniformly on $B_r(0)$ for every $r > 0$.

(b) If $\exp(x) := \sum_{k=0}^{\infty} \frac{1}{k!} x^k$, then $\exp\colon A \to A$ is differentiable and

$$d\exp_x(h) = \exp(x) \cdot h\,.$$

This is the EXPONENTIAL FUNCTION on A.

(c) If x, $y \in A$, then

$$\exp(x) \cdot \exp(y) = \exp(x + y)\,.$$

(d) If $x \in A$, then x is invertible and

$$\big(\exp(x)\big)^{-1} = \exp(-x)\,.$$

28.4.29. Problem. Develop some trigonometry on a commutative Banach algebra A. (It will be convenient to be able to take the DERIVATIVE of a Banach algebra valued function. If $G\colon A \to A$ is a differentiable function, define $DG(a) := dG_a(1)$ for every $a \in A$.) Include at least the following:

(a) The series $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^k$ converges absolutely for all x in A and uniformly on every open ball centered at the origin.

(b) The function $F\colon x \mapsto \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^k$ is differentiable at every $x \in A$. Find $dF_x(h)$.

(c) For every $x \in A$, let $\cos x := F(x^2)$. Let $\sin x := D \cos x$ for every x. Show that

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

for every $x \in A$.

(d) Show that $D \sin x = \cos x$ for every $x \in A$.

(e) Show that $\sin^2 x + \cos^2 x = 1$ for every $x \in A$.

CHAPTER 29

THE IMPLICIT FUNCTION THEOREM

This chapter deals with the problem of solving equations and systems of equations for some of the variables which appear in terms of the others. In a few very simple cases this can be done explicitly. As an example, consider the equation

$$x^2 + y^2 = 25 \tag{29.1}$$

for the circle of radius 5 centered at the origin in $\mathbb{R}^2$. Although (29.1) can *not* be solved *globally* for y in terms of x (that is, there is no function f such that $y = f(x)$ for all points (x, y) satisfying (29.1)), it nevertheless *is* possible at most points on the circle to solve *locally* for y in terms of x. For example, if (a, b) lies on the circle and $b > 0$, then there exist open intervals J_1 and J_2 containing a and b, respectively, and a function $f\colon J_1 \to J_2$ such that every point (x, y) in the rectangular region $J_1 \times J_2$ will satisfy $y = f(x)$ if and only if it satisfies equation (29.1). In particular, we could choose $J_1 = (-5, 5)$, $J_2 = (0, 6)$, and $f\colon x \mapsto \sqrt{25 - x^2}$. In case $b < 0$ the function f would be replaced by $f : x \mapsto -\sqrt{25 - x^2}$. If $b = 0$ then there is *no* local solution for y in terms of x: each rectangular region about either of the points $(5, 0)$ or $(-5, 0)$ will contain pairs of points symmetrically located with respect to the x-axis which satisfy (29.1); and it is not possible for the graph of a function to contain such a pair.

Our attention in this chapter is focused not on the relatively rare cases where it is possible to compute explicit (local) solutions for some of the variables in terms of the remaining ones but on the more typical situation where no such computation is possible. In this latter circumstance it is valuable to have information concerning the *existence* of (local) solutions and the differentiability of such solutions.

The simplest special case is a single equation of the form $y = f(x)$ where f is a continuously differentiable function. The *inverse function theorem*, derived in the first section of this chapter, provides conditions under which this equation can be solved locally for x in terms of y, say $x = g(y)$, and gives us a formula allowing us to compute the differential of g. More complicated equations and systems of equations require the *implicit function theorem*, which is the subject of the second section of the chapter.

29.1. THE INVERSE FUNCTION THEOREM

Recall that in chapter 25 formulas concerning the function $f \mapsto \Delta f$ lead to corresponding formulas involving differentials. For example, $d(f + g)_a = df_a + dg_a$ followed from $\Delta(f + g)_a = \Delta f_a + \Delta g_a$ (see 25.3.15). It is natural to ask whether the formula

$$\Delta\big(f^{-1}\big)_{f(x)} = (\Delta f_x)^{-1}$$

derived for bijective functions f in proposition 25.3.7 leads to a corresponding formula

$$d\big(f^{-1}\big)_{f(x)} = (df_x)^{-1} \tag{29.2}$$

for differentials. Obviously, a necessary condition for (29.2) to hold for all x in some neighborhood of a point a is that the linear map df_a be invertible. The *inverse function theorem* states that for continuously differentiable (but not necessarily bijective) functions this is all that is required. The proof of the *inverse function theorem* is a fascinating application of the *contractive mapping theorem* (theorem 19.1.5). First some terminology.

29.1.1. Definition. Let E and F be Banach spaces and $\emptyset \neq U \overset{\circ}{\subseteq} E$. A function f belonging to $\mathcal{C}^1(U,F)$ is $\mathcal{C}^1$-INVERTIBLE if f is a bijection between U and an open subset V of F and if f^{-1} belongs to $\mathcal{C}^1(V,E)$. Such a function is also called a $\mathcal{C}^1$-ISOMORPHISM between U and V.

29.1.2. Exercise. Find nonempty open subsets U and V of $\mathbb{R}$ and a continuously differentiable bijection $f\colon U \to V$ which is *not* a $\mathcal{C}^1$-isomorphism between U and V. (Solution Q.29.1.)

29.1.3. Definition. Let E and F be Banach spaces. A function f in $\mathcal{F}_a(E,F)$ is LOCALLY $\mathcal{C}^1$-INVERTIBLE (or a LOCAL $\mathcal{C}^1$-ISOMORPHISM) at a point a in E if there exists a neighborhood of a on which the restriction of f is $\mathcal{C}^1$-invertible. The inverse of this restriction is a LOCAL $\mathcal{C}^1$-INVERSE of f at a and is denoted by f_{loc}^{-1}.

29.1.4. Exercise. Let $f(x) = x^2 - 6x + 5$ for all x in $\mathbb{R}$. Find a local $\mathcal{C}^1$-inverse of f at $x = 1$. (Solution Q.29.2.)

29.1.5. Problem. Let $f(x) = x^6 - 2x^3 - 7$ for all x in $\mathbb{R}$. Find local $\mathcal{C}^1$-inverses for f at 0 and at 10.

29.1.6. Problem. Find a nonempty open subset U of $\mathbb{R}$ and a function f in $\mathcal{C}^1(U,\mathbb{R})$ which is *not* $\mathcal{C}^1$-invertible but is locally $\mathcal{C}^1$-invertible at every point in U.

Before embarking on a proof of the *inverse function theorem* it is worthwhile seeing why a naive "proof" of this result using the chain rule fails—even in the simple case of a real valued function of a real variable.

29.1.7. Exercise. If $f \in \mathcal{F}_a(\mathbb{R},\mathbb{R})$ and if $Df(a) \neq 0$, then f is locally $\mathcal{C}^1$-invertible at a and

$$Df_{\text{loc}}^{-1}(b) = \frac{1}{Df(a)} \tag{29.3}$$

where f_{loc}^{-1} is a local $\mathcal{C}^1$-inverse of f at a and $b = f(a)$.

This assertion is correct. Criticize the following "proof" of the result: Since f_{loc}^{-1} is a local $\mathcal{C}^1$-inverse of f at a

$$f_{\text{loc}}^{-1}(f(x)) = x$$

for all x in some neighborhood U of a. Applying the chain rule (proposition 8.4.19) we obtain

$$(Df_{\text{loc}}^{-1})(f(x)) \cdot Df(x) = 1$$

for all x in U. Letting $x = a$ we have

$$(Df_{\text{loc}}^{-1})(b)\, Df(a) = 1$$

and since $Df(a) \neq 0$ equation (29.3) follows. (Solution Q.29.3.)

The *inverse function theorem* (29.1.16) deals with a continuously differentiable function f which is defined on a neighborhood of a point a in a Banach space E and which maps into a second Banach space F. We assume that the differential of f at a is invertible. Under these hypotheses we prove that f is locally $\mathcal{C}^1$-invertible at a and in some neighborhood of a equation (29.2) holds. To simplify the proof we temporarily make some additional assumptions: we suppose that H is a continuously differentiable function which is defined on a neighborhood of 0 in a Banach space E and which maps into this same space E, that $H(\mathbf{0}) = \mathbf{0}$, and that the differential of H at 0 is the identity map on E. Once the conclusion of the *inverse function theorem* has been established in this restricted case the more general version follows easily. The strategy we employ to attack the special case is straightforward, but there are numerous details which must be checked along the way. Recall that in chapter 19 we were able to solve certain systems of simultaneous linear equations by putting them in the form $Ax = b$ where A is a square matrix and x and b are vectors in the Euclidean space of appropriate dimension. This equation was rewritten in the form $Tx = x$

where $Tx := x - b + Ax$, thereby reducing the problem to one of finding a fixed point of the mapping T. When T is contractive a simple application of the *contractive mapping theorem* (19.1.5) is all that is required. We make use of exactly the same idea here. We want a local inverse of H. That is, we wish to solve the equation $H(x) = y$ for x in terms of y in some neighborhood of $\mathbf{0}$. Rewrite the equation $H(x) = y$ in the form $\phi_y(x) = x$ where for each y near $\mathbf{0}$ the function ϕ_y is defined by $\phi_y(x) := x - H(x) + y$. Thus, as before, the problem is to find for each y a unique fixed point of ϕ_y. In order to apply the *contractive mapping theorem* to ϕ_y, the domain of this function must be a *complete* metric space. For this reason we choose temporarily to take the domain of ϕ_y to be a *closed* ball about the origin in E.

In lemma 29.1.8 we find such a closed ball C. It must satisfy two conditions: first, C must lie in the domain of H; and second, if u belongs to C, then dH_u must be close to the identity map on E, say, $\|dH_u - I\| < \frac{1}{2}$. (This latter condition turns out to be a crucial ingredient in proving that ϕ_y is contractive.) In lemma 29.1.9 we show that (for y sufficiently small) ϕ_y maps the closed ball C into itself; and in 29.1.10 the basic task is to show that ϕ_y is contractive and therefore has a unique fixed point. The result of all this is that there exists a number $r > 0$ such that for every y in $B = B_r(0)$ there exists a unique x in the closed ball $C = C_{2r}(\mathbf{0})$ such that $y = H(x)$. Now this is not quite the end of the story. First of all we do *not* know that H restricted to C is injective: some points in C may be mapped to the region outside B, about which the preceding says nothing. Furthermore, the definition of local $\mathcal{C}^1$-invertibility requires a homeomorphism between *open* sets, and C is not open. This suggests we restrict our attention to points lying in the interior of C which map into B. So let $V = C^\circ \cap H^{\leftarrow}(B)$ and consider the restriction of H to V, which we denote by H_{loc}. In lemma 29.1.11 we show that V is a neighborhood of 0 and that H_{loc} is injective. Thus the inverse function $H_{\text{loc}}^{-1} : H^{\rightarrow}(V) \to V$ exists. The succeeding lemma is devoted to showing that this inverse is continuous.

In order to conclude that H is locally C^1-invertible we still need two things: we must know that H_{loc} is a homeomorphism between *open* sets and that H_{loc}^{-1} is continuously differentiable. Lemma 29.1.13 shows that $H^{\rightarrow}(V)$ is open in E. And in lemma 29.1.14 we complete the proof of this special case of the *inverse function theorem* by showing that H_{loc}^{-1} is continuously differentiable and that in the open set V its differential is given by (29.2).

Corollary 29.1.15 shows that the conclusions of the preceding result remain true even when one of the hypotheses is eliminated and another weakened. Here we prove the *inverse function theorem* for a function G whose domain E and codomain F are *not* required to be identical. Of course, if $E \neq F$ we cannot assume that the differential of G at $\mathbf{0}$ is the identity map; we assume only that it is invertible.

Finally, in theorem 29.1.16 we prove our final version of the *inverse function theorem*. Here we drop the requirement that the domain of the function in question be a neighborhood of the origin.

In 29.1.8–29.1.14 the following hypotheses are in force:

(1') $\mathbf{0} \in U_1 \overset{\circ}{\subseteq} E$ (where E is a Banach space);
(2') $H \in \mathcal{C}^1(U_1, E)$;
(3') $H(\mathbf{0}) = \mathbf{0}$; and
(4') $dH_{\mathbf{0}} = I$.

29.1.8. Lemma. *There exists $r > 0$ such that $B_{3r}(\mathbf{0}) \subseteq U_1$ and $\|dH_u - I\| < \frac{1}{2}$ whenever $\|u\| \leq 2r$.*

Proof. Problem.

29.1.9. Lemma. *For $\|y\| < r$ define a function ϕ_y by*

$$\phi_y(x) := x - H(x) + y$$

for all x such that $\|x\| \leq 2r$. Show that ϕ_y maps $C_{2r}(\mathbf{0})$ into itself.

Proof. Problem.

29.1.10. Lemma. *For every y in $B_r(\mathbf{0})$ there exists a unique x in $C_{2r}(\mathbf{0})$ such that $y = H(x)$.*

Proof. Problem. *Hint.* Show that the function ϕ_y defined in lemma 29.1.9 is contractive on the metric space $C_{2r}(\mathbf{0})$ and has $\frac{1}{2}$ as a contraction constant. To find an appropriate inequality involving $\|\phi_y(u) - \phi_y(v)\|$ apply corollary 26.1.8 to $\|H(v) - H(u) - dH_{\mathbf{0}}(v - u)\|$.

29.1.11. Lemma. *Show that if $V := \{x \in B_{2r}(\mathbf{0}) \colon \|H(x)\| < r\}$, then $\mathbf{0} \in V \overset{\circ}{\subseteq} E$. Let H_{loc} be the restriction of H to V. Show that H_{loc} is a bijection between V and $H^{\rightarrow}(V)$.*

Proof. Problem.

29.1.12. Lemma. *The function $H_{loc}^{-1} \colon H^{\rightarrow}(V) \to V$ is continuous.*

Proof. Problem. *Hint.* Prove first that if u, $v \in C_{2r}(\mathbf{0})$, then $\|u - v\| \leq 2\|H(u) - H(v)\|$. In order to do this, look at

$$2\|\phi_{\mathbf{0}}(u) + H(u) - \phi_{\mathbf{0}}(v) - H(v)\| - \|u - v\|$$

and recall (from the proof of 29.1.10) that $\phi_{\mathbf{0}}$ has contraction constant $\frac{1}{2}$. Use this to conclude that if w, $z \in H^{\rightarrow}(V)$, then

$$\|H_{\mathrm{loc}}^{-1}(w) - H_{\mathrm{loc}}^{-1}(z)\| \leq 2\,\|w - z\|$$

where H_{loc} is the restriction of H to V (see lemma 29.1.10.

29.1.13. Lemma. *Show that $H^{\rightarrow}(V) \overset{\circ}{\subseteq} E$.*

Proof. Problem. *Hint.* Show that if a point b belongs to $H^{\rightarrow}(V)$, then so does the open ball $B_{r-\|b\|}(b)$. Proceed as follows: Show that if a point y lies in this open ball, then $\|y\| < r$ and therefore $y = H(x)$ for some (unique) x in $C_{2r}(\mathbf{0})$. Prove that $y \in H^{\rightarrow}(V)$ by verifying $\|x\| < 2r$. To do this look at

$$\|x - H_{\mathrm{loc}}^{-1}(b)\| + \|H_{\mathrm{loc}}^{-1}(b) - H_{\mathrm{loc}}^{-1}(\mathbf{0})\|$$

and use the first inequality given in the hint to the preceding problem.

29.1.14. Lemma. *The function H is locally $\mathcal{C}^1 -$invertible at 0. Furthermore,*

$$d\big(H_{loc}^{-1}\big)_{H(x)} = (dH_x)^{-1}$$

for every x in V.

Proof. Problem. *Hint.* First prove the differentiability of H_{loc}^{-1} on $H^{\rightarrow}(V)$. If $y \in H^{\rightarrow}(V)$, then there exists a unique x in V such that $y = H(x)$. By hypothesis $\Delta H_x \sim dH_x$. Show that multiplication on the right by $\Delta\big(H_{\mathrm{loc}}^{-1}\big)_y$ preserves tangency. (For this it must be established that $\Delta\big(H_{\mathrm{loc}}^{-1}\big)_y$ belongs to $\mathfrak{O}(E, E)$.) Then show that multiplication on the left by $(dH_x)^{-1}$ preserves tangency. (How do we know that this inverse exists for all x in V?) Finally show that the map $y \mapsto d\big(H_{\mathrm{loc}}^{-1}\big)_y$ is continuous on $H^{\rightarrow}(V)$ by using (29.2) to write it as the composite of H_{loc}^{-1}, dH, and the map $T \mapsto T^{-1}$ on $\operatorname{Inv} \mathfrak{B}(E, E)$ (see proposition 28.4.14).

29.1.15. Corollary (A second, more general, version of the inverse function theorem)**.** *Let E and F be Banach spaces. If*

(1") $\mathbf{0} \in U_1 \overset{\circ}{\subseteq} E$,
(2") $G \in \mathcal{C}^1(U_1, F)$,
(3") $G(\mathbf{0}) = \mathbf{0}$, *and*
(4") $dG_{\mathbf{0}} \in \operatorname{Inv} \mathfrak{B}(E, F)$,

then G is locally $\mathcal{C}^1$-invertible at $\mathbf{0}$. Furthermore,

$$d\big(G_{loc}^{-1}\big)_{G(x)} = (dG_x)^{-1}$$

for all x in some neighborhood of $\mathbf{0}$, where G_{loc}^{-1} is a local $\mathcal{C}^1$-inverse of G at $\mathbf{0}$.

PROOF. Problem. *Hint.* Let $H = (dG_{\mathbf{0}})^{-1} \circ G$. Apply lemma 29.1.14.

29.1.16. Theorem (Inverse Function Theorem (third, and final, version))**.** *Let E and F be Banach spaces. If*

(1) $a \in U \overset{\circ}{\subseteq} E$,
(2) $f \in \mathcal{C}^1(U, F)$, *and*
(3) $df_a \in \operatorname{Inv} \mathfrak{B}(E, F)$,

then f is locally $\mathcal{C}^1$-invertible at a. Furthermore,

$$d\big(f_{loc}^{-1}\big)_{f(x)} = (df_x)^{-1}$$

for all x in some neighborhood of a.

PROOF. Problem. *Hint.* Let $U_1 = U - a$ and $G = \Delta f_a$. Write G as a composite of f with translation maps. Apply corollary 29.1.15.

29.1.17. Problem. Let $P\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (r, \theta) \mapsto (r\cos\theta, r\sin\theta)$ and a be a point in $\mathbb{R}^2$ such that $P(a) = (1, \sqrt{3})$.

(a) Show that P is locally $\mathcal{C}^1$-invertible at a by finding a local $\mathcal{C}^1$-inverse P_{loc}^{-1} of P at a. For the inverse you have found, compute $d\big(P_{\text{loc}}^{-1}\big)_{(1,\sqrt{3})}$.
(b) Use the *inverse function theorem* 29.1.16 to show that P is locally $\mathcal{C}^1$-invertible at a. Then use the formula given in that theorem to compute $d\big(P_{\text{loc}}^{-1}\big)_{(1,\sqrt{3})}$ (where P_{loc}^{-1} is a local $\mathcal{C}^1$-inverse of P at a). *Hint.* Use proposition 21.5.16.

29.1.18. Problem. Let $U = \{(x, y, z) \in \mathbb{R}^3 \colon x, y, z > 0\}$ and let $g\colon U \to \mathbb{R}^3$ be defined by

$$g(x, y, z) = \left(\frac{x}{y^2z^2}\,,\ yz\,,\ \ln y\right).$$

Calculate separately $\big[dg_{(x,y,z)}\big]^{-1}$ and $\big[d(g^{-1})_{g(x,y,z)}\big]$.

29.2. THE IMPLICIT FUNCTION THEOREM

In the preceding section we derived the *inverse function theorem*, which gives conditions under which an equation of the form $y = f(x)$ can be solved locally for x in terms of y. The *implicit function theorem* deals with the local solvability of equations that are not necessarily in the form $y = f(x)$ and of systems of equations. The *inverse function theorem* is actually a special case of the *implicit function theorem*. Interestingly, the special case can be used to prove the more general one.

This section consists principally of exercises and problems which illustrate how the *inverse function theorem* can be adapted to guarantee the existence of local solutions for various examples of equations and systems of equations. Once the computational procedure is well understood for these examples it is a simple matter to explain how it works in general; that is, to prove the *implicit function theorem*.

Suppose, given an equation of the form $y = f(x)$ and a point a in the domain of f, we are asked to show that *the equation can be solved for x in terms of y near a*. It is clear what we are being asked: to show that f is locally invertible at a. (Since the function f will usually satisfy some differentiability condition—continuous differentiability, for example—it is natural to ask for the local inverse to satisfy the same condition.)

As we have seen, local invertibility can be established by explicitly computing a local inverse, which can be done only rarely, or by invoking the *inverse function theorem*. Let us suppose we are given an equation of the form

$$f(x, y) = 0 \tag{29.4}$$

and a point (a, b) in $\mathbb{R}^2$ which satisfies (29.4).
Question. What does it mean to say that (29.4) can be solved for y near b in terms of x near a? (Alternative wording: What does it mean to say that (29.4) can be solved for y in terms of x near the point (a, b)?)
Answer. There exist a neighborhood V of a and a function $h\colon V \to R$ which satisfy

$$\begin{aligned} &\text{(i)} \quad h(a) = b \\ &\text{(ii)} \quad f(x, h(x)) = 0 \end{aligned}$$

for all x in V.

29.2.1. Example. In the introduction to this chapter we discussed the problem of solving the equation

$$x^2 + y^2 = 25$$

for y in terms of x. This equation can be put in the form (29.4) by setting $f(x, y) = x^2 + y^2 - 25$. Suppose we are asked to show that (29.4) can be solved for y near 4 in terms of x near 3. As in the introduction, take $V = (-5, 5)$ and $h(x) = \sqrt{25 - x^2}$ for all x in V. Then $h(3) = 4$ and $f(x, h(x)) = x^2 + \left(\sqrt{25 - x^2}\right)^2 - 25 = 0$; so h is the desired local solution to (29.4). If we are asked to show that (29.4) can be solved for y near -4 in terms of x near 3, what changes? We choose $h(x) = -\sqrt{25 - x^2}$. Notice that condition (i) above dictates the choice of h. (Either choice will satisfy (ii).)

As was pointed out in the introduction, the preceding example is atypical in that it is possible to *specify* the solution. We can actually *solve* for y in terms of x. *Much* more common is the situation in which an explicit solution is not possible. What do we do then?

To see how this more complicated situation can be dealt with, let us pretend just for a moment that our computational skills have so totally deserted us that in the preceding example we are unable to specify the neighborhood V and the function h required to solve (29.4). The problem is still the same: show that the equation

$$x^2 + y^2 - 25 = 0 \tag{29.5}$$

can be solved for y near 4 in terms of x near 3. A good start is to define a function $G\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$G(x, y) = (x, f(x, y)) \tag{29.6}$$

where as above $f(x, y) = x^2 + y^2 - 25$, and apply the *inverse function theorem* to G. It is helpful to make a sketch here. Take the xy-plane to be the domain of G and in this plane sketch the circle $x^2 + y^2 = 25$. For the codomain of G take another plane, letting the horizontal axis be called "x" and the vertical axis be called "z". Notice that in this second plane the image under G of the circle drawn in the xy-plane is the line segment $[-5, 5]$ along the x-axis and that the image of the x-axis is the parabola $z = x^2 - 25$. Where do points in the interior of the circle go? What about points outside the circle? If you think of the action of G as starting with a folding of the xy-plane along the x-axis, you should be able to guess the identity of those points where G is not locally invertible. In any case we will find these points using the *inverse function theorem.*

The function G is continuously differentiable on $\mathbb{R}^2$ and (the matrix representation of) its differential at a point (x, y) in $\mathbb{R}^2$ is given by

$$\left[dG_{(x,y)}\right] = \begin{bmatrix} 1 & 0 \\ 2x & 2y \end{bmatrix}.$$

Thus according to the *inverse function theorem* G is locally $\mathcal{C}^1$-invertible at every point (x, y) where this matrix is invertible, that is, everywhere except on the x-axis. In particular, at the point $(3, 4)$ the function G is locally $\mathcal{C}^1$-invertible. Thus there exist a neighborhood W of $G(3, 4) = (3, 0)$ in R^2 and a local $\mathcal{C}^1$-inverse $H\colon W \to \mathbb{R}^2$ of G. Write H in terms of its component functions $H = \left(H^1, H^2\right)$ and set $h(x) = H^2(x, 0)$ for all x in $V := \{x : (x, 0) \in W\}$. Then V is a neighborhood of 3 in $\mathbb{R}$

and the function h is continuously differentiable (because H is). To show that h is a solution of the equation (29.5) for y in terms of x we must show that

$$\begin{aligned} &\text{(i)} \quad h(3) = 4; \qquad \text{and} \\ &\text{(ii)} \quad f(x, h(x)) = 0 \end{aligned}$$

for all x in V.

To obtain (i) equate the second components of the first and last terms of the following computation.

$$\begin{aligned} (3,4) &= H(G(3,4)) \\ &= H(3, f(3,4)) \\ &= H(3,0) \\ &= (H^1(3,0)\,,\ H^2(3,0)) \\ &= (H^1(3,0)\,,\ h(3)). \end{aligned}$$

To obtain (ii) notice that for all x in V

$$\begin{aligned} (x,0) &= G(H(x,0)) \\ &= G(H^1(x,0)\,,\ H^2(x,0)) \\ &= G(H^1(x,0)\,,\ h(x)) \\ &= (H^1(x,0)\,,\ f(H^1(x,0), h(x)))\,. \end{aligned}$$

Equating first components we see that $H^1(x,0) = x$. So the preceding can be written

$$(x,0) = (x, f(x, h(x)))$$

from which (ii) follows by equating second components.

Although the preceding computations demonstrate the existence of a local solution $y = h(x)$ without specifying it, it is nevertheless possible to calculate the value of its derivative $\dfrac{dy}{dx}$ at the point $(3,4)$, that is, to find $h'(3)$. Since $h'(3) = \big(H^2 \circ j_1\big)'(3) = d\big(H^2 \circ j_1\big)_3(1) = \big(dH^2{}_{(3,0)} \circ j_1\big)(1) = dH^2{}_{(3,0)}(1,0) = \frac{\partial H^2}{\partial x}(3,0)$ (where j_1 is the inclusion map $x \mapsto (x,0)$) and since H is a local inverse of G, the *inverse function theorem* tells us that

$$\begin{aligned} \big[dH_{(3,0)}\big] &= \big[dH_{G(3,4)}\big] \\ &= \big[dG_{(3,4)}\big]^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ f_1(3,4) & f_2(3,4) \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 6 & 8 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ -\frac{3}{4} & \frac{1}{8} \end{bmatrix}. \end{aligned}$$

But the entry in the lower left corner is $\frac{\partial H^2}{\partial x}(3,0)$. Therefore, $h'(3) = -\frac{3}{4}$.

In the next exercise we consider an equation whose solution cannot be easily calculated.

29.2.2. Exercise. Consider the equation

$$x^2 y + \sin\left(\frac{\pi}{2} x y^2\right) = 2 \tag{29.7}$$

where $x, y \in R$

(a) What does it mean to say that equation (29.7) can be solved for y in terms of x near the point $(1, 2)$?
(b) Show that it is possible to solve (29.7) as described in (a). *Hint.* Proceed as in the second solution to the preceding example.
(c) Use the *inverse function theorem* to find the value of $\frac{dy}{dx}$ at the point $(1, 2)$.

(Solution Q.29.4.)

29.2.3. Problem. Consider the equation

$$e^{xy^2} - x^2y + 3x = 4 \tag{29.8}$$

where x, $y \in \mathbb{R}$.

(a) What does it mean to say that equation (29.8) can be solved for y near 0 in terms of x near 1?
(b) Show that such a solution does exist.
(c) Use the *inverse function theorem* to compute the value of $\frac{dy}{dx}$ at the point $(1, 0)$.

The preceding examples have all been equations involving only two variables. The technique used in dealing with these examples works just as well in cases where we are given an equation in an arbitrary number of variables and wish to demonstrate the existence of local solutions for one of the variables in terms of the remaining ones.

29.2.4. Exercise. Consider the equation

$$x^2z + yz^2 - 3z^3 = 8 \tag{29.9}$$

for x, y, $z \in \mathbb{R}$.

(a) What does it mean to say that equation (29.9) can be solved for z near 1 in terms of x and y near $(3, 2)$? (Alternatively: that (29.9) can be solved for z in terms of x and y near the point $(x, y, z) = (3, 2, 1)$?)
(b) Show that such a solution does exist. *Hint.* Follow the preceding technique, but instead of using (29.6), define $G(x, y, z) := (x, y, f(x, y, z))$ for an appropriate function f.
(c) Use the *inverse function theorem* to find the values of $\left(\frac{\partial z}{\partial x}\right)_y$ and $\left(\frac{\partial z}{\partial y}\right)_x$ at the point $(3, 2, 1)$.

(Solution Q.29.5.)

29.2.5. Problem. Let $f(x, y, z) = xz + xy + yz - 3$. By explicit computation find a neighborhood V of $(1, 1)$ in $\mathbb{R}^2$ and a function $h\colon V \to \mathbb{R}$ such that $h(1, 1) = 1$ and $f(x, y, h(x, y)) = 0$ for all x and y in V. Find $h_1(1, 1)$ and $h_2(1, 1)$.

29.2.6. Problem. Use the *inverse function theorem*, not direct calculation, to show that the equation

$$xz + xy + yz = 3$$

has a solution for z near 1 in terms of x and y near $(1, 1)$ and to find $\left(\frac{\partial z}{\partial x}\right)_y$ and $\left(\frac{\partial z}{\partial y}\right)_x$ at the point $(1, 1, 1)$.

29.2.7. Problem. (a) What does it mean to say that the equation

$$wx^2y + \sqrt{w}y^2z^4 = 3xz + 6x^3z^2 + 7$$

can be solved for z in terms of w, x, and y near the point $(w, x, y, z) = (4, 1, 2, 1)$?

(b) Show that such a solution exists.
(c) Use the *inverse function theorem* to find $\left(\frac{\partial z}{\partial w}\right)_{x,y}$, $\left(\frac{\partial z}{\partial x}\right)_{w,y}$, and $\left(\frac{\partial z}{\partial y}\right)_{w,x}$ at the point $(4, 1, 2, 1)$.

29.2.8. Problem. Let $U \overset{\circ}{\subseteq} \mathbb{R}^{n-1} \times \mathbb{R}$ and $f \in \mathcal{C}^1(U, \mathbb{R})$. Suppose that the point $(a, b) = (a_1, \dots, a_{n-1}; b)$ belongs to U, that $f(a, b) = 0$, and that $f_n(a, b) \neq 0$. Show that there exists a neighborhood V of a in $\mathbb{R}^{n-1}$ and a function h in $\mathcal{C}^1(V, \mathbb{R})$ such that $h(a) = b$ and $f(x, h(x)) = 0$ for all x in V. *Hint.* Show that the function

$$G\colon U \to\ R^{n-1} \times \mathbb{R} \colon (x, y) = (x_1, \dots, x_{n-1}; y) \mapsto (x, f(x, y))$$

has a local $\mathcal{C}^1$-inverse, say H, at (a, b).

29.2.9. Problem. Show that the equation

$$uvy^2z + w\sqrt{x}z^{10} + v^2e^yz^4 = 5 + uw^2\cos\left(x^3y^5\right)$$

can be solved for x near 4 in terms of u, v, w, y, and z near -3, 2, -1, 0, and 1, respectively. *Hint.* Use the preceding problem.

29.2.10. Problem. (a) Use problem 29.2.8 to make sense of the following "theorem": If $f(x, y, z) = 0$, then

$$\left(\frac{\partial z}{\partial x}\right)_y = -\frac{\left(\frac{\partial f}{\partial x}\right)_{y,z}}{\left(\frac{\partial f}{\partial z}\right)_{x,y}}\,.$$

Hint. After determining that there exists a function h which satisfies the equation $f(x, y, h(x, y)) = 0$ on an appropriate neighborhood of a point, use the chain rule to differentiate both sides of the equation with respect to x.)

(b) Verify part (a) directly for the function f given in problem 29.2.5 by computing each side independently.

(c) Restate and prove the following "theorem": If $f(x, y, z) = 0$, then

$$\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x = -1\,.$$

29.2.11. Problem. A commonly used formula in scientific work is

$$\left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{\left(\frac{\partial y}{\partial x}\right)_z}\,.$$

Recast this as a carefully stated theorem. Then prove the theorem. *Hint.* Use problem 29.2.8 twice to obtain appropriate functions h and j satisfying $f(h(y, z), y, z) = 0$ and $f(x, j(x, z), z) = 0$ on appropriate neighborhoods. Differentiate these equations using the chain rule. Evaluate at a particular point. Solve.

29.2.12. Problem. (a) Make sense of the formula

$$\left(\frac{\partial x}{\partial z}\right)_y = -\frac{\left(\frac{\partial y}{\partial z}\right)_x}{\left(\frac{\partial y}{\partial x}\right)_z}$$

and prove it.

(b) Illustrate the result in part (a) by computing separately $\left(\frac{\partial P}{\partial V}\right)_T$ and $-\frac{\left(\frac{\partial T}{\partial V}\right)_P}{\left(\frac{\partial T}{\partial P}\right)_V}$ from the equation of state $PV = RT$ for an ideal gas. (Here R is a constant.)

We have dealt at some length with the problem of solving a single equation for one variable in terms of the remaining ones. It is pleasant to discover that the techniques used there can be adapted with only the most minor modifications to give local solutions for systems of n equations in p variables (where $p > n$) for n of the variables in terms of the remaining $p - n$ variables. We begin with some examples.

29.2.13. Exercise. Consider the following system of equations:

$$\begin{cases} 2u^3vx^2 + v^2x^3y^2 - 3u^2y^4 = 0 \\ 2uv^2y^2 - uvx^2 + u^3xy = 2\,. \end{cases} \tag{29.10}$$

(a) What does it mean to say that the system (29.10) can be solved for x and y near (c, d) in terms of u and v near (a, b)?

(b) Show that the system (29.10) can be solved for x and y near $(1, 1)$ in terms of u and v near $(1, 1)$. *Hint.* Try to imitate the technique of problem 29.2.8, except in this case define G on an appropriate subset of $\mathbb{R}^2 \times \mathbb{R}^2$.

(Solution Q.29.6.)

29.2.14. Problem. Consider the following system of equations

$$\begin{cases} 4x^2 + 4y^2 &= z \\ x^2 + y^2 &= 5 - z\,. \end{cases} \tag{29.11}$$

(a) What does it mean to say that (29.11) can be solved for y and z near $(1, 4)$ in terms of x near 0?

(b) Show that such a solution exists by direct computation.

(c) Compute $\left(\frac{\partial y}{\partial x}\right)_z$ and $\left(\frac{\partial z}{\partial x}\right)_y$ at the point $x = 0$, $y = 1$, $z = 4$.

29.2.15. Problem. (a) Repeat problem 29.2.14(b), this time using the *inverse function theorem* instead of direct computation.

(b) Use the *inverse function theorem* to find $\left(\frac{\partial y}{\partial x}\right)_z$ and $\left(\frac{\partial z}{\partial x}\right)_y$ at the point $x = 0$, $y = 1$, $z = 4$.

29.2.16. Problem. Discuss the problem of solving the system of equations

$$\begin{cases} ux^2 + vwy + u^2w &= 4 \\ uvy^3 + 2wx - x^2y^2 &= 3 \end{cases} \tag{29.12}$$

for x and y near $(1, 1)$ in terms of u, v, and w near $(1, 2, 1)$.

As the preceding examples indicate, the first step in solving a system of n equations in $n + k$ unknowns (where $k > 0$) for n of the variables in terms of the remaining k variables is to replace the system of equations by a function $f : U \to \mathbb{R}^n$ where $U \overset{\circ}{\subseteq} \mathbb{R}^k \times \mathbb{R}^n$. If the finite dimensional spaces $\mathbb{R}^k$ and $\mathbb{R}^n$ are replaced by arbitrary Banach spaces, the subsequent calculations can be carried out in exactly the same fashion as in the examples we have just considered. The result of this computation is the *implicit function theorem.*

29.2.17. Theorem (The Implicit Function Theorem)**.** *Let E_1, E_2, and F be Banach spaces, $(a, b) \in U \overset{\circ}{\subseteq} E_1 \times E_2$, and $f \in \mathcal{C}^1(U, F)$. If $f(a, b) = \mathbf{0}$ and $d_2f_{(a,b)}$ is invertible in $\mathfrak{B}(E_2, F)$, then there exist a neighborhood V of a in E_1 and a continuously differentiable function $h\colon V \to E_2$ such that $h(a) = b$ and $f(x, h(x)) = \mathbf{0}$ for all $x \in V$.*

Proof. Problem. *Hint.* Let $G(x, y) = (x, f(x, y))$ for all $(x, y) \in U$. Show that $dG_{(a,b)} = (x, Sx + Ty)$ where $S = d_1f_{(a,b)}$ and $T = d_2f_{(a,b)}$. Show that the map $(x, z) \mapsto (x, T^{-1}(z - Sx))$ from $E_1 \times F$ to $E_1 \times E_2$ is the inverse of $dG_{(a,b)}$. (One can guess what the inverse of $dG_{(a,b)}$ should be by regarding $dG_{(a,b)}$ as the matrix $\begin{bmatrix} I_{E_1} & 0 \\ S & T \end{bmatrix}$ acting on $E_1 \times E_2$.) Apply the *inverse function theorem* 29.1.16 and proceed as in exercise 29.2.13 and problems 29.2.15 and 29.2.16.

29.2.18. Problem. Suppose we are given a system of n equations in p variables where $p > n$. What does the *implicit function theorem* 29.2.17 say about the possibility of solving this system locally for n of the variables in terms of the remaining $p - n$ variables?

APPENDIX A

QUANTIFIERS

Certainly "$2 + 2 = 4$" and "$2 + 2 = 5$" are statements—one true, the other false. On the other hand the appearance of the variable x prevents the expression "$x + 2 = 5$" from being a statement. Such an expression we will call an OPEN SENTENCE; its truth is open to question since x is unidentified. There are three standard ways of converting open sentences into statements.

The first, and simplest, of these is to give the variable a particular value. If we "evaluate" the expression "$x + 2 = 5$" at $x = 4$, we obtain the (false) statement "$4 + 2 = 5$".

A second way of obtaining a statement from an expression involving a variable is UNIVERSAL QUANTIFICATION: we assert that the expression is true for all values of the variable. In the preceding example we get, "For all x, $x + 2 = 5$." This is now a statement (and again false). The expression "for all x" (or equivalently, "for every x") is often denoted symbolically by $(\forall x)$. Thus the preceding sentence may be written, $(\forall x)x + 2 = 5$. (The parentheses are optional; they may be used in the interest of clarity.) We call $\forall$ a UNIVERSAL QUANTIFIER.

Frequently there are several variables in an expression. They may all be universally quantified. For example

$$(\forall x)(\forall y)\, x^2 - y^2 = (x - y)(x + y) \tag{A.1}$$

is a (true) statement, which says that for every x and for every y the expression $x^2 - y^2$ factors in the familiar way. The order of consecutive universal quantifiers is unimportant: the statement

$$(\forall y)(\forall x)\, x^2 - y^2 = (x - y)(x + y)$$

says exactly the same thing as (A.1). For this reason the notation may be contracted slightly to read

$$(\forall x, y)\, x^2 - y^2 = (x - y)(x + y)\,.$$

A third way of obtaining a statement from an open sentence $P(x)$ is EXISTENTIAL QUANTIFICATION. Here we assert that $P(x)$ is true for *at least one* value of x. This is often written "$(\exists x)$ such that $P(x)$" or more briefly "$(\exists x)P(x)$", and is read "there exists an x such that $P(x)$" or "$P(x)$ is true for some x." For example, if we existentially quantify the expression "$x + 2 = 5$" we obtain "$(\exists x)$ such that $x + 2 = 5$" (which happens to be true). We call $\exists$ an EXISTENTIAL QUANTIFIER.

As is true for universal quantifiers, the order of consecutive existential quantifiers is immaterial.

CAUTION. It is absolutely essential to realize that the order of an existential and a universal quantifier may *not* in general be reversed. For example,

$$(\exists x)(\forall y)\, x < y$$

says that there is a number x with the property that no matter how y is chosen, x is less than y; that is, there is a smallest real number. (This is, of course, false.) On the other hand

$$(\forall y)(\exists x)\, x < y$$

says that for every y we can find a number x smaller than y. (This is true: take x to be $y - 1$ for example.) *The importance of getting quantifiers in the right order cannot be overestimated.*

There is one frequently used convention concerning quantifiers that should be mentioned. In the statement of definitions, propositions, theorems, *etc.*, missing quantifiers are assumed to be universal; furthermore, they are assumed to be the innermost quantifiers.

A.1.1. Example. Let f be a real valued function defined on the real line $\mathbb{R}$. Many texts give the following definition. The function f is *continuous* at a point a in $\mathbb{R}$ if: for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(a)| < \epsilon \qquad \text{whenever } |x - a| < \delta\,.$$

Here ϵ and δ are quantified; the function f and the point a are fixed for the discussion, so they do not require quantifiers. What about x? According to the convention just mentioned, x is universally quantified and that quantifier is the innermost one. Thus the definition reads: for every $\epsilon > 0$ there exists $\delta > 0$ such that for every x

$$|f(x) - f(a)| < \epsilon \qquad \text{whenever } |x - a| < \delta\,.$$

A.1.2. Example. Sometimes all quantifiers are missing. In this case the preceding convention dictates that all variables are universally quantified. Thus

$$\text{Theorem.} \quad x^2 - y^2 = (x - y)(x + y)$$

is interpreted to mean

$$\text{Theorem.} \quad (\forall x)(\forall y)\, x^2 - y^2 = (x - y)(x + y)\,.$$

APPENDIX B

SETS

In this text everything is defined ultimately in terms of two primitive (that is, undefined) concepts: set and set membership. We assume that these are already familiar to the reader. In particular, it is assumed to be understood that distinct elements (or members, or points) can be regarded collectively as a single set (or family, or class, or collection). To indicate that x belongs to a set A (or that x is a member of A) we write $x \in A$; to indicate that it does not belong to A we write $x \notin A$.

We specify a set by listing its members between braces (for instance, $\{1, 2, 3, 4, 5\}$ is the set of the first five natural numbers), by listing some of its members between braces with an ellipsis (three dots) indicating the missing members (*e.g.* $\{1, 2, 3, \dots\}$ is the set of all natural numbers), or by writing $\{x \colon P(x)\}$ where $P(x)$ is an open sentence which specifies what property the variable x must satisfy in order to be included in the set (*e.g.* $\{x \colon 0 \le x \le 1\}$ is the closed unit interval $[0, 1]$).

B.1.1. Problem. Let $\mathbb{N}$ be the set of natural numbers 1, 2, 3, ... and let

$$S = \{x \colon x < 30 \text{ and } x = n^2 \text{ for some } n \in \mathbb{N}\}\,.$$

List all the elements of S.

B.1.2. Problem. Let $\mathbb{N}$ be the set of natural numbers 1, 2, 3, ... and let

$$S = \{x \colon x = n + 2 \text{ for some } n \in \mathbb{N} \text{ such that } n < 6\}\,.$$

List all the elements of S.

B.1.3. Problem. Suppose that

$$S = \{x \colon x = n^2 + 2 \text{ for some } n \in \mathbb{N}\}$$

and that

$$T = \{3, 6, 11, 18, 27, 33, 38, 51\}\,.$$

(a) Find an element of S that is not in T.
(b) Find an element of T that is not in S.

B.1.4. Problem. Suppose that

$$S = \{x \colon x = n^2 + 2n \text{ for some } n \in \mathbb{N}\}$$

and that

$$T = \{x \colon x = 5n - 1 \text{ for some } n \in \mathbb{N}\}\,.$$

Find an element which belongs to both S and T.

Since all of our subsequent work depends on the notions of set and of set membership, it is not altogether satisfactory to rely on intuition and shared understanding to provide a foundation for these crucial concepts. It is possible in fact to arrive at paradoxes using a naive approach to sets. (For example, ask yourself the question, "If S is the set of all sets which do not contain themselves as members, then does S belong to S?" If the answer is *yes*, then it must be *no*, and *vice versa*.) One satisfactory alternative to our intuitive approach to sets is axiomatic set theory. There are many ways of axiomatizing set theory to provide a secure foundation for subsequent mathematical development. Unfortunately, each of these ways turns out to be extremely intricate, and it is generally felt that the abstract complexities of axiomatic set theory do not serve well as a beginning to an introductory course in advanced calculus.

Most of the paradoxes inherent in an intuitive approach to sets have to do with sets that are too "large". For example, the set S mentioned above is enormous. Thus in the sequel we will assume that in each situation all the mathematical objects we are then considering (sets, functions, *etc.*) belong to some appropriate "universal" set, which is "small" enough to avoid set theoretic paradoxes. (Think of "universal" in terms of "universe of discourse", not "all-encompassing".) In many cases an appropriate universal set is clear from the context. Previously we considered a statement

$$(\forall y)(\exists x)\, x < y\,.$$

The appearance of the symbol "$<$" suggests to most readers that x and y are real numbers. Thus the universal set from which the variables are chosen is the set $\mathbb{R}$ of all real numbers. When there is doubt that the universal set will be properly interpreted, it may be specified. In the example just mentioned we might write

$$(\forall y \in \mathbb{R})(\exists x \in \mathbb{R})\, x < y.$$

This makes explicit the intended restriction that x and y be real numbers.

As another example recall that in appendix A we defined a real valued function f to be continuous at a point $a \in \mathbb{R}$ if

$$(\forall \epsilon > 0)(\exists \delta > 0) \text{ such that } (\forall x)\, |f(x) - f(a)| < \epsilon \text{ whenever } |x - a| < \delta.$$

Here the first two variables, ϵ and δ, are restricted to lie in the open interval $(0, \infty)$. Thus we might rewrite the definition as follows:

$$\forall \epsilon \in (0,\infty)\ \exists \delta \in (0,\infty) \\ \text{such that } (\forall x \in \mathbb{R})\, |f(x) - f(a)| < \epsilon \text{ whenever } |x - a| < \delta.$$

The expressions "$\forall x \in \mathbb{R}$", "$\exists \delta \in (0, \infty)$", *etc.* are called RESTRICTED QUANTIFIERS.

B.1.5. Definition. Let S and T be sets. We say that S is a SUBSET of T and write $S \subseteq T$ (or $T \supseteq S$) if every member of S belongs to T. If $S \subseteq T$ we also say that S is CONTAINED IN T or that T CONTAINS S. Notice that the relation $\subseteq$ is REFLEXIVE (that is, $S \subseteq S$ for all S) and TRANSITIVE (that is, if $S \subseteq T$ and $T \subseteq U$, then $S \subseteq U$). It is also ANTISYMMETRIC (that is, if $S \subseteq T$ and $T \subseteq S$, then $S = T$). If we wish to claim that S is *not* a subset of T we may write $S \not\subseteq T$. In this case there is at least one member of S which does not belong to T.

B.1.6. Example. Since every number in the closed interval $[0, 1]$ also belongs to the interval $[0, 5]$, it is correct to write $[0, 1] \subseteq [0, 5]$. Since the number π belongs to $[0, 5]$ but not to $[0, 1]$, we may also write $[0, 5] \not\subseteq [0, 1]$.

B.1.7. Definition. If $S \subseteq T$ but $S \neq T$, then we say that S is a PROPER SUBSET of T (or that S is PROPERLY CONTAINED IN T, or that T PROPERLY CONTAINS S) and write $S \subsetneqq T$.

B.1.8. Problem. Suppose that $S = \{x \colon x = 2n + 3 \text{ for some } n \in \mathbb{N}\}$ and that T is the set of all odd natural numbers 1, 3, 5, ...

(a) Is $S \subseteq T$? If not, find an element of S which does not belong to T.
(b) Is $T \subseteq S$? If not, find an element of T which does not belong to S.

B.1.9. Definition. The EMPTY SET (or NULL SET), which is denoted by $\emptyset$, is defined to be the set which has no elements. (Or, if you like, define it to be $\{x \colon x \neq x\}$.) It is regarded as a subset of every set, so that $\emptyset \subseteq S$ is always true. (Note: "$\emptyset$" is a letter of the Danish alphabet, not the Greek letter "phi".)

B.1.10. Definition. If S is a set, then the POWER SET of S, which we denote by $\mathfrak{P}(S)$, is the set of all subsets of S.

B.1.11. Example. Let $S = \{a, b, c\}$. Then the members of the power set of S are the empty set, the three one-element subsets, the three two-element subsets, and the set S itself. That is,

$$\mathfrak{P}(S) = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, S\big\} .$$

B.1.12. Problem. In each of the words (a)–(d) below let S be the set of letters in the word. In each case find the number of members of S and the number of members of $\mathfrak{P}(S)$ the power set of S.

(a) lull
(b) appall
(c) attract
(d) calculus

CAUTION. In attempting to prove a theorem which has as a hypothesis "Let S be a set" do not include in your proof something like "Suppose $S = \{s_1, s_2, \dots, s_n\}$" or "Suppose $S = \{s_1, s_2, \dots\}$". In the first case you are tacitly assuming that S is finite and in the second that it is countable. Neither is justified by the hypothesis.

CAUTION. A single letter $\mathfrak{S}$ (an S in fraktur font) is an acceptable symbol in printed documents. Don't try to imitate it in hand-written work or on the blackboard. Use script letters instead.

Finally a word on the use of the symbols $=$ and $:=$. In this text equality is used in the sense of identity. We write $x = y$ to indicate that x and y are two names for the same object. For example, $0.5 = 1/2 = 3/6 = 1/\sqrt{4}$ because 0.5, 1/2, 3/6, and $1/\sqrt{4}$ are different names for the same real number. You have probably encountered other uses of the term *equality*. In many high school geometry texts, for example, one finds statements to the effect that a triangle is isosceles if it has two equal sides (or two equal angles). What is meant of course is that a triangle is isosceles if it has two sides of *equal length* (or two angles of *equal angular measure*). We also make occasional use of the symbol $:=$ to indicate *equality by definition*. Thus when we write $a := b$ we are giving a new name a to an object b with which we are presumably already familiar.

APPENDIX C

SPECIAL SUBSETS OF $\mathbb{R}$

We denote by $\mathbb{R}$ the set of real numbers. Certain subsets of $\mathbb{R}$ have standard names. We list some of them here for reference. The set $\mathbb{P} = \{x \in \mathbb{R} \colon x > 0\}$ of strictly positive numbers is discussed in appendix H. The set $\{1, 2, 3, \dots\}$ of all natural numbers is denoted by $\mathbb{N}$, initial segments $\{1, 2, 3, \dots, m\}$ of this set by $\mathbb{N}_m$, and the set $\{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ of all integers by $\mathbb{Z}$. The set of all rational numbers (numbers of the form p/q where $p, q \in \mathbb{Z}$ and $q \neq 0$) is denoted by $\mathbb{Q}$. There are the *open intervals*

$$\begin{aligned} (a, b) &:= \{x \in \mathbb{R} \colon a < x < b\}, \\ (-\infty, b) &:= \{x \in \mathbb{R} \colon x < b\}, \quad \text{and} \\ (a, \infty) &:= \{x \in \mathbb{R} \colon x > a\}. \end{aligned}$$

There are the *closed intervals*

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} \colon a \le x \le b\}, \\ (-\infty, b] &:= \{x \in \mathbb{R} \colon x \le b\}, \quad \text{and} \\ [a, \infty) &:= \{x \in \mathbb{R} \colon x \ge a\}. \end{aligned}$$

And there are the intervals which (if $a < b$) are neither open nor closed:

$$\begin{aligned} [a, b) &:= \{x \in \mathbb{R} \colon a \le x < b\} \quad \text{and} \\ (a, b] &:= \{x \in \mathbb{R} \colon a < x \le b\}. \end{aligned}$$

The set $\mathbb{R}$ of all real numbers may be written in interval notation as $(-\infty, \infty)$. (As an interval it is considered both open and closed. The reason for applying the words "open" and "closed" to intervals is discussed in chapter 2.)

A subset A of $\mathbb{R}$ is BOUNDED if there is a positive number M such that $|a| \le M$ for all $a \in A$. Thus intervals of the form $[a, b]$, $(a, b]$, $[a, b)$, and (a, b) are bounded. The other intervals are *unbounded*.

If A is a subset of $\mathbb{R}$, then $A^+ := A \cap [0, \infty)$. These are the POSITIVE elements of A. Notice, in particular, that $\mathbb{Z}^+$, the set of positive integers, contains 0, but $\mathbb{N}$, the set of natural numbers, does not.

APPENDIX D

LOGICAL CONNECTIVES

D.1. DISJUNCTION AND CONJUNCTION

The word "or" in English has two distinct uses. Suppose you ask a friend how he intends to spend the evening, and he replies, "I'll" walk home or I'll take in a film." If you find that he then walked home and on the way stopped to see a film, it would not be reasonable to accuse him of having lied. He was using the inclusive "or", which is true when one or both alternatives are. On the other hand suppose that while walking along a street you are accosted by an armed robber who says, "Put up your hands or I'll shoot." You obligingly raise your hands. If he then shoots you, you have every reason to feel ill-used. Convention and context dictate that he used the exclusive "or": either alternative, but not both.

Since it is undesirable for ambiguities to arise in mathematical discourse, the inclusive "or" has been adopted as standard for mathematical (and most scientific) purposes. A convenient way of defining logical connectives such as "or" is by means of a truth table. The formal definition of "or" looks like this.

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Here P and Q are any sentences. In the columns labeled P and Q we list all possible combinations of truth values for P and Q (T for true, F for false). In the third column appears the corresponding truth value for "P or Q". According to the table "P or Q" is true in all cases except when both P and Q are false. The notation "$P \vee Q$" is frequently used for "P or Q". The operation $\vee$ is called DISJUNCTION.

D.1.1. Exercise. Construct a truth table giving the formal definition of "and", frequently denoted by $\wedge$. The operation $\wedge$ is called CONJUNCTION. (Solution Q.30.1.)

We say that two sentences depending on variables $P, Q, \dots$ are LOGICALLY EQUIVALENT if they have the same truth value no matter how truth values T and F are assigned to the variables $P, Q, \dots$. It turns out that truth tables are quite helpful in deciding whether certain statements encountered in mathematical reasoning are logically equivalent to one another. (But do *not* clutter up mathematical proofs with truth tables. Everyone is supposed to argue logically. Truth tables are only scratch work for the perplexed.)

D.1.2. Example. Suppose that P, Q, and R are any sentences. To a person who habitually uses language carefully, it will certainly be clear that the following two assertions are equivalent:

(a) P is true and so is either Q or R.
(b) Either both P and Q are true or both P and R are true.

Suppose for a moment, however, that we are in doubt concerning the relation between (a) and (b). We may represent (a) symbolically by $P \wedge (Q \vee R)$ and (b) by $(P \wedge Q) \vee (P \wedge R)$. We conclude that they are indeed logically equivalent by examining the following truth table. (Keep in mind that since there are 3 variables, P, Q, and R, there are $2^3 = 8$ ways of assigning truth values to them; so we need 8 lines in our truth table.)

(1) P	(2) Q	(3) R	(4) $Q \vee R$	(5) $P \wedge (Q \vee R)$	(6) $P \wedge Q$	(7) $P \wedge R$	(8) $(P \wedge Q) \vee (P \wedge R)$
T	T	T	T	T	T	T	T
T	T	F	T	T	T	F	T
T	F	T	T	T	F	T	T
T	F	F	F	F	F	F	F
F	T	T	T	F	F	F	F
F	T	F	T	F	F	F	F
F	F	T	T	F	F	F	F
F	F	F	F	F	F	F	F

Column (4) is obtained from (2) and (3), column (5) from (1) and (4), column (6) from (l) and (2), column (7) from (1) and (3), and column (8) from (6) and (7). Comparing the truth values in columns (5) and (8), we see that they are exactly the same. Thus $P \wedge (Q \vee R)$ is logically equivalent to $(P \wedge Q) \vee (P \wedge R)$. This result is a DISTRIBUTIVE LAW; it says that conjunction distributes over disjunction.

D.1.3. Problem. Use truth tables to show that the operation of disjunction is associative; that is, show that $(P \vee Q) \vee R$ and $P \vee (Q \vee R)$ are logically equivalent.

D.1.4. Problem. Use truth tables to show that disjunction distributes over conjunction; that is, show that $P \vee (Q \wedge R)$ is logically equivalent to $(P \vee Q) \wedge (P \vee R)$.

One final remark: quantifiers may be "moved past" portions of disjunctions and conjunctions which do not contain the variable being quantified. For example,

$$(\exists y)(\exists x)\,[(y^2 \leq 9) \wedge (2 < x < y)]$$

says the same thing as

$$(\exists y)\,[(y^2 \leq 9) \wedge (\exists x)\,2 < x < y].$$

D.2. IMPLICATION

Consider the assertion, "If 1 = 2, then 2 = 3." If you ask a large number of people (not mathematically trained) about the truth of this, you will probably find some who think it is true, some who think it is false, and some who think it is meaningless (therefore neither true nor false). This is another example of the ambiguity of ordinary language. In order to avoid ambiguity and to insure that "P implies Q" has a truth value whenever P and Q do, we define the operation of IMPLICATION, denoted by $\Rightarrow$, by means of the following truth table.

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

There are many ways of saying that P implies Q. The following assertions are all identical.

$P \Rightarrow Q$.
P implies Q.
If P, then Q.
P is sufficient (or a sufficient condition) for Q.
Whenever P, then Q.
$Q \Leftarrow P$
Q is implied by P.
Q is a consequence of P.
Q follows from P.
Q is necessary (or a necessary condition) for P.
Q whenever P.

The statement $Q \Rightarrow P$ is the CONVERSE of $P \Rightarrow Q$. It is a common mistake to confuse a statement with its converse. This is a grievous error. For example: it is correct to say that if a geometric figure is a square, then it is a quadrilateral; but it is *not* correct to say that if a figure is a quadrilateral it must be a square.

D.2.1. Definition. If P and Q are sentences, we define the logical connective "iff" (read "if and only if") by the following truth table.

P	Q	P iff Q
T	T	T
T	F	F
F	T	F
F	F	T

Notice that the sentence "P iff Q" is true exactly in those cases where P and Q have the same truth values. That is, saying that "P iff Q" is a TAUTOLOGY (true for all truth values of P and Q) is the same as saying that P and Q are equivalent sentences. Thus the connective "iff" is called EQUIVALENCE. An alternative notation for "iff" is "$\Leftrightarrow$".

D.2.2. Example. By comparing columns (3) and (6) of the following truth table, we see that "P iff Q" is logically equivalent to "$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$".

(1) P	(2) Q	(3) P iff Q	(4) $P \Rightarrow Q$	(5) $Q \vee (\sim P)$	(6) $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
T	T	T	T	T	T
T	F	F	F	T	F
F	T	F	T	F	F
F	F	T	T	T	T

This is a very important fact. Many theorems of the form P iff Q are most conveniently proved by verifying separately that $P \Rightarrow Q$ and that $Q \Rightarrow P$.

D.3. RESTRICTED QUANTIFIERS

Now that we have the logical connectives $\Rightarrow$ and $\wedge$ at our disposal, it is possible to introduce restricted quantifiers formally in terms of unrestricted ones. This enables one to obtain properties of the former from corresponding facts about the latter.

(See exercise D.3.2 and problems D.4.8 and D.4.9.)

D.3.1. Definition (of restricted quantifiers)**.** Let S be a set and $P(x)$ be an open sentence. We define $(\forall x \in S)\, P(x)$ to be true if and only if $(\forall x)\,\big((x \in S) \Rightarrow P(x)\big)$ is true; and we define $(\exists x \in S)\, P(x)$ to be true if and only if $(\exists x)\,\big((x \in S) \wedge P(x)\big)$ is true.

D.3.2. Exercise. Use the preceding definition and the fact (mentioned in chapter A) that the order of unrestricted existential quantifiers does not matter to show that the order of restricted existential quantifiers does not matter. That is, show that if S and T are sets and $P(x, y)$ is an open sentence, then $(\exists x \in S)\,(\exists y \in T)\, P(x, y)$ holds if and only if $(\exists y \in T)\,(\exists x \in S)\, P(x, y)$ does. (Solution Q.30.2.)

D.4. NEGATION

If P is a sentence, then $\sim P$, read "the NEGATION of P" or "the DENIAL of P" or just "not P", is the sentence whose truth values are the opposite of P.

P	$\sim P$
T	F
F	T

D.4.1. Example. It should be clear that the denial of the disjunction of two sentences P and Q is logically equivalent to the conjunction of their denials. If we were in doubt about the correctness of this, however, we could appeal to a truth table to prove that $\sim (P \vee Q)$ is logically equivalent to $\sim P \wedge \sim Q$.

(1) P	(2) Q	(3) $P \vee Q$	(4) $\sim (P \vee Q)$	(5) $\sim P$	(6) $\sim Q$	(7) $(\sim P) \wedge (\sim Q)$
T	T	T	F	F	F	F
T	F	T	F	F	T	F
F	T	T	F	T	F	F
F	F	F	T	T	T	T

Columns (4) and (7) have the same truth values: That is, the denial of the disjunction of P and Q is logically equivalent to the conjunction of their denials. This result is one of *De Morgan's laws*. The other is given as problem D.4.2.

D.4.2. Problem (De Morgan's law)**.** Use a truth table to show that $\sim \ (P \wedge Q)$ is logically equivalent to $(\sim P) \vee (\sim Q)$.

D.4.3. Problem. Obtain the result in problem D.4.2 without using truth tables. *Hint.* Use D.4.1 together with the fact that a proposition P is logically equivalent to $\sim\sim P$. Start by writing $(\sim P) \vee (\sim Q)$ iff $\sim\sim ((\sim P) \vee (\sim Q))$.

D.4.4. Exercise. Let P and Q be sentences. Then $P \Rightarrow Q$ is logically equivalent to $Q \vee (\sim P)$. (Solution Q.30.3.)

One very important matter is the process of taking the negation of a quantified statement. Let $P(x)$ be an open sentence. If it is not the case that $P(x)$ holds for every x, then it must fail for some x, and conversely. That is, $\sim (\forall x)P(x)$ is logically equivalent to $(\exists x) \sim P(x)$.

Similarly, $\sim (\exists x)P(x)$ is logically equivalent to $(\forall x) \sim P(x)$. (If it is not the case that $P(x)$ is true for some x, then it must fail for all x, and conversely.)

D.4.5. Example. In chapter 3 we define a real valued function f on the real line to be continuous provided that

$$(\forall a)(\forall \epsilon)(\exists \delta)(\forall x)\,|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

How does one prove that a particular function f is *not* continuous? Answer: find numbers a and ϵ such that for every δ it is possible to find an x such that $|f(x) - f(a)| \geq \epsilon$ and $|x - a| < \delta$. To see that this is in fact what we must do, notice that each pair of consecutive lines in the following argument are logically equivalent.

$$\begin{aligned}
&\sim [(\forall a)(\forall \epsilon)(\exists \delta)(\forall x)\ \ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon]\\
&(\exists a) \sim [(\forall \epsilon)(\exists \delta)(\forall x)\ \ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon]\\
&(\exists a)(\exists \epsilon) \sim [(\exists \delta)(\forall x)\ \ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon]\\
&(\exists a)(\exists \epsilon)(\forall \delta) \sim [(\forall x)\ \ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon]\\
&(\exists a)(\exists \epsilon)(\forall \delta)(\exists x)\ \sim [\ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon]\\
&(\exists a)(\exists \epsilon)(\forall \delta)(\exists x)\ \sim [(|f(x) - f(a)| < \epsilon) \vee \sim (|x-a| < \delta)]\\
&(\exists a)(\exists \epsilon)(\forall \delta)(\exists x)\, [\sim (\ |f(x) - f(a)| < \epsilon) \wedge \sim\sim (|x-a| < \delta)]\\
&(\exists a)(\exists \epsilon)(\forall \delta)(\exists x)\quad [(\ |f(x) - f(a)| \geq \epsilon) \wedge (|x-a| < \delta)]
\end{aligned}$$

To obtain the third line from the end use exercise D.4.4; the penultimate line is a consequence of example D.4.1; and the last line makes use of the obvious fact that a sentence P is always logically equivalent to $\sim\sim P$.

D.4.6. Problem. Two students, Smith and Jones, are asked to prove a mathematical theorem of the form, "If P then if Q then R." Smith assumes that Q is a consequence of P and tries to prove R. Jones assumes both P and Q are true and tries to prove R. Is either of these students doing the right thing? Explain carefully.

D.4.7. Problem. The CONTRAPOSITIVE of the implication $P \Rightarrow Q$ is the implication $(\sim Q) \Rightarrow (\sim P)$. Without using truth tables or assigning truth values show that an implication is logically equivalent to its contrapositive. (This is a very important fact. In many cases when you are asked to prove a theorem of the form $P \Rightarrow Q$, rather than assuming P and proving Q you will find it easier to assume that Q is false and conclude that P must also be false.) *Hint.* Use D.4.4. You may also use the obvious facts that disjunction is a commutative operation ($P \vee Q$ is logically equivalent to $Q \vee P$) and that P is logically equivalent to $\sim\sim P$.)

D.4.8. Problem. Use the formal definition of restricted quantifiers given in section D.3 together with the fact mentioned in appendix A that the order of unrestricted universal quantifiers does not matter to show that the order of restricted universal quantifiers does not matter. That is, show that if S and T are sets and $P(x, y)$ is an open sentence, then $(\forall x \in S)(\forall y \in T)P(x, y)$ holds if and only if $(\forall y \in T)(\forall x \in S)P(x, y)$ does.

D.4.9. Problem. Let S be a set and $P(x)$ be an open sentence. Show that

(a) $\sim (\forall x \in S)P(x)$ if and only if $(\exists x \in S) \sim P(x)$.
(b) $\sim (\exists x \in S)P(x)$ if and only if $(\forall x \in S) \sim P(x)$.

Hint. Use the corresponding facts (given in the two paragraphs following exercise D.4.4) for unrestricted quantifiers.

APPENDIX E

WRITING MATHEMATICS

E.1. PROVING THEOREMS

Mathematical results are called *theorems*—or *propositions*, or *lemmas*, or *corollaries*, or *examples*. All these are intended to be mathematical facts. The different words reflect only a difference of emphasis. Theorems are more important than propositions. A lemma is a result made use of in a (usually more important) subsequent result. The German word for lemma is particularly suggestive: "Hilfsatz," meaning "helping statement." A corollary (to a theorem or proposition) is an additional result we get (almost) for free. All these results are typically packaged in the form, "If P, then Q." The assertion P is the *hypothesis* (or *premise*, or *assumption*, or *supposition*). The assertion Q is the *conclusion*. Notice that the result "Every object of type A is of type B" is in this form. It can be rephrased as, "If x is an object of type A, then x is of type B."

The statements, P and Q themselves may be complicated conjunctions, or disjunctions, or conditionals of simpler statements. One common type of theorem, for example, is, "If P_1, P_2, …, and P_m, then Q_1, Q_2, …, and Q_n." (Personally, I think such a theorem is clumsily stated if m and n turn out to be large.)

A *proof* of a result is a sequence of statements, each with justification, which leads to the conclusion(s) of the desired result. The statements that constitute a proof may be definitions, or hypotheses, or statements which result from applying to previous steps of the proof a valid rule of inference. *Modus ponens* is the basic rule of inference. It says that if you know a proposition P and you also know that P implies Q, then you can conclude that Q is true. Another important rule of inference (sometimes called *universal instantiation*) is that if you know a proposition $P(x)$ to be true for every x in a set S and you know that a is a member of S, then you can conclude that $P(a)$ is true.

Other rules of inference can be derived from *modus ponens*. Let's look at an example. Certainly, if we know that the proposition $P \wedge Q$ is true we should be able to conclude that P is true. The reason is simple we know (or can easily check) that

$$(P \wedge Q) \Rightarrow P \tag{E.1}$$

is a tautology (true for all truth values of P and Q). Since $P \wedge Q$ is known to be true, P follows from (E.1) by *modus ponens*. No attempt is made here to list every rule of inference that it is appropriate to use. Most of them should be entirely obvious. For those that are not, truth tables may help. (As an example consider problem G.1.13: If the product of two numbers x and y is zero, then either x or y must be zero. Some students feel obligated to prove two things: that if $xy = 0$ and $x \neq 0$ then $y = 0$ **AND** that if $xy = 0$ and $y \neq 0$ then $x = 0$. Examination of truth tables shows that this is not necessary.)

A proof in which you start with the hypotheses and reason until you reach the conclusion is a *direct proof*. There are two other proof formats, which are known as *indirect proofs*. The first comes about by observing that the proposition $\sim Q \Rightarrow \sim P$ is logically equivalent to $P \Rightarrow Q$. (We say that $\sim Q \Rightarrow \sim P$ is the *contrapositive* of $P \Rightarrow Q$.) To prove that P implies Q, it suffices to assume that Q is false and prove, using this assumption that P is false. Some find it odd that to prove something is true one starts by assuming it to be false. A slight variant of this is the *proof by contradiction*. Here, to prove that P implies Q, assume two things: that P is true and that Q is false. Then attempt to show that these lead to a contradiction. We like to believe that the mathematical system we work in is consistent (although we know we can't prove it), so when an assumption leads us to a contradiction we reject it. Thus in a proof by contradiction when we find that P and $\sim Q$ can't both be true, we conclude that if P is true Q must also be true.

E.1.1. Problem. Prove that in an inconsistent system, everything is true. That is, prove that if P and Q are propositions and that if both P and $\sim P$ are true, then Q is true. *Hint.* Consider the proposition $(P \wedge \sim P) \Rightarrow Q$.

E.1.2. Problem. What is wrong with the following proof that 1 is the largest natural number.

> Let N be the largest natural number. Since N is a natural number so is N^2. We see that $N^2 = N \cdot N \geq N \cdot 1 = N$. Clearly the reverse inequality $N^2 \leq N$ holds because N is the largest natural number. Thus $N^2 = N$. This equation has only two solutions: $N = 0$ and $N = 1$. Since 0 is not a natural number, we have $N = 1$. That is, 1 is the largest natural number.

E.2. CHECKLIST FOR WRITING MATHEMATICS

After you have solved a problem or discovered a counterexample or proved a theorem, there arises the problem of writing up your result. You want to do this in a way that will be as clear and as easy to digest as possible. Check your work against the following list of suggestions.

(1) *Have you clearly stated the problem you are asked to solve or the result you are trying to prove?*

> Have an audience in mind. Write to someone. And don't assume the person you are writing to remembers the problem. (S)he may have gone on vacation, or been fired; or maybe (s)he just has a bad memory. You need not include every detail of the problem, but there should be enough explanation so that a person not familiar with the situation can understand what you are talking (writing) about.

(2) *Have you included a paragraph at the beginning explaining the method you are going to use to address the problem?*

> No one is happy being thrown into a sea of mathematics with no clue as to what is going on or why. Be nice. Tell the reader what you are doing, what steps you intend to take, and what advantages you see to your particular approach to the problem.

(3) *Have you defined all the variables you use in your writeup?*

> *Never* be so rude as to permit a symbol to appear that has not been properly introduced. You may have a mental picture of a triangle with vertices labelled A, B, and C. When you use those letters no one will know what they stand for unless you tell them. (Even if you have included a graph appropriately labelled, still tell the reader *in the text* what the letters denote.) Similarly, you may be consistent in always using the letter j to denote a natural number. But how would you expect the reader to know?
>
> It is good practice to italicize variables so that they can be easily distinguished from regular text.

(4) *Is the logic of your report entirely clear and entirely correct?*

It is an unfortunate fact of life that the slightest error in logic can make a "solution" to a problem totally worthless. It is also unfortunate that even a technically correct argument can be so badly expressed that no one will believe it.

(5) *In your writeup are your mathematical symbols and mathematical terms all correctly used in a standard fashion? And are all abbreviations standard?*

Few things can make mathematics more confusing than misused or eccentrically used symbols. Symbols should *clarify* arguments not create yet another level of difficulty. By the way, symbols such as "=" and "<" are used *only* in formulas. They are not substitutes for the words "equals" and "less than" in ordinary text. Logical symbols such as $\Rightarrow$ are rarely appropriate in mathematical exposition: write "If A then B," not "$A \Rightarrow B$." Occasionally they may be used in displays.

(6) *Are the spelling, punctuation, diction, and grammar of your report all correct?*

(7) *Is every word, every symbol, and every equation part of a sentence? And is every sentence part of a paragraph?*

For some reason this seems hard for many students. Scratchwork, of course, tends to be full of free floating symbols and formulas. When you write up a result get rid of all this clutter. Keep only what is necessary for a logically complete report of your work. And make sure any formula you keep becomes (an intelligible) part of a sentence. Study how the author of any good mathematics text deals with the problem of incorporating symbols and formulas into text.

(8) *Does every sentence start correctly and end correctly?*

Sentences start with capital letters. *Never* start a sentence with a number or with a mathematical or logical symbol. Every declarative sentence ends with a period. Other sentences may end with a question mark or (rarely) an exclamation mark.

(9) *Is the* function *of every sentence of your report clear?*

Every sentence has a function. It may be a definition. Or it may be an assertion you are about to prove. Or it may be a consequence of the preceding statement. Or it may be a standard result your argument depends on. Or it may be a summary of what you have just proved. **Whatever function a sentence serves, that function should be entirely clear to your reader.**

(10) *Have you avoided all unnecessary clutter?*

Mindless clutter is one of the worst enemies of clear exposition. No one wants to see all the details of your arithmetic or algebra or trigonometry or calculus. Either your reader knows this stuff and could do it more easily than read it, or doesn't know it and will find it meaningless. In either case, get rid of it. If you solve an equation, for example, state what the solutions are; don't show how you used the quadratic formula to find them. Write only things that inform. Logical argument informs, reams of routine calculations do not. Be ruthless in rooting out useless clutter.

(11) *Is the word "any" used unambiguously? And is the order in which quantifiers act entirely clear?*

> Be careful with the word "any", especially when taking negations. It can be surprisingly treacherous. "Any" may be used to indicate universal quantification. (In the assertion, "It is true for any x that $x^2 - 1 = (x+1)(x-1)$", the word "any" means "every".) It may also be used for existential quantification. (In the question, "Does $x^2 + 2 = 2x$ hold for any x?" the word "any" means "for some". Also notice that answering *yes* to this question does not mean that you believe it is true that "$x^2 + 2 = 2x$ holds for any x.") Negations are worse. (What does "It is not true that $x^2 + 2 = 2x$ for any x" mean?) One good way to avoid trouble: don't use "any".
>
> And be careful of word order. The assertions "Not every element of A is an element of B" and "Every element of A is not an element of B" say quite different things. I recommend avoiding the second construction entirely. Putting (some or all) quantifiers at the end of a sentence can be a dangerous business. (The statement, "$P(x)$ or $Q(x)$ fails to hold for all x," has at least two possible interpretations. So does "$x \neq 2$ for all $x \in A$," depending on whether we read $\neq$ as "is not equal to" or as "is different from".)

E.3. FRAKTUR AND GREEK ALPHABETS

Fraktur is a typeface widely used until the middle of the twentieth century in Germany and a few other countries. In this text it is used to denote families of sets and families of linear maps. In general it is not a good idea to try to reproduce these letters in handwritten material or when writing on the blackboard. In these cases English script letters are both easier to read and easier to write. Here is a list of upper case Fraktur letters:

$\mathfrak{A}$ (A)	$\mathfrak{B}$ (B)	$\mathfrak{C}$ (C)	$\mathfrak{D}$ (D)	$\mathfrak{E}$ (E)	$\mathfrak{F}$ (F)	$\mathfrak{G}$ (G)
$\mathfrak{H}$ (H)	$\mathfrak{I}$ (I)	$\mathfrak{J}$ (J)	$\mathfrak{K}$ (K)	$\mathfrak{L}$ (L)	$\mathfrak{M}$ (M)	$\mathfrak{N}$ (N)
$\mathfrak{O}$ (O)	$\mathfrak{P}$ (P)	$\mathfrak{Q}$ (Q)	$\mathfrak{R}$ (R)	$\mathfrak{S}$ (S)	$\mathfrak{T}$ (T)	$\mathfrak{U}$ (U)
$\mathfrak{V}$ (V)	$\mathfrak{W}$ (W)	$\mathfrak{X}$ (X)	$\mathfrak{Y}$ (Y)	$\mathfrak{Z}$ (Z)		

The following is a list of standard Greek letters used in mathematics. Notice that in some cases both upper case and lower case are commonly used:

α (alpha)	β (beta)	Γ, γ (gamma)	Δ, δ (delta)	ϵ (epsilon)	ζ (zeta)
η (eta)	Θ, θ (theta)	ι (iota)	κ (kappa)	Λ, λ (lambda)	μ (mu)
ν (nu)	Ξ, ξ (xi)	Π, π (pi)	ρ (rho)	Σ, σ (sigma)	τ (tau)
Φ, ϕ (phi)	Ψ, ψ (psi)	χ (chi)	Ω, ω (omega)		

APPENDIX F

SET OPERATIONS

F.1. UNIONS

Recall that if S and T are sets, then the UNION of S and T, denoted by $S \cup T$, is defined to be the set of all those elements x such that $x \in S$ or $x \in T$.

That is,

$$S \cup T := \{x \colon x \in S \text{ or } x \in T\}.$$

F.1.1. Example. If $S = [0, 3]$ and $T = [2, 5]$, then $S \cup T = [0, 5]$.

The operation of taking unions of sets has several essentially obvious properties. In the next proposition we list some of these.

F.1.2. Proposition. *Let S, T, U, and V be sets. Then*

(a) $S \cup (T \cup U) = (S \cup T) \cup U$ *(associativity);*
(b) $S \cup T = T \cup S$ *(commutativity);*
(c) $S \cup \emptyset = S$*;*
(d) $S \subseteq S \cup T$*;*
(e) $S = S \cup T$ *if and only if* $T \subseteq S$*; and*
(f) *If* $S \subseteq T$ *and* $U \subseteq V$*, then* $S \cup U \subseteq T \cup V$*.*

PROOF. We prove parts (a), (c), (d), and (e). Ordinarily one would probably regard these results as too obvious to require proof. The arguments here are presented only to display some techniques used in writing formal proofs. Elsewhere in the text references will not be given to this proposition when the facts (a)–(f) are used. When results are considered obvious, they may be mentioned but are seldom cited. The proofs of the remaining parts (b) and (f) are left as problems.

PROOF. (a) A standard way to show that two sets are equal is to show that an element x belongs to one if and only if it belongs to the other. In the present case

$$\begin{aligned} x \in S \cup (T \cup U) &\text{ iff } x \in S \text{ or } x \in T \cup U \\ &\text{ iff } x \in S \text{ or } (x \in T \text{ or } x \in U) \\ &\text{ iff } (x \in S \text{ or } x \in T) \text{ or } x \in U \\ &\text{ iff } x \in S \cup T \text{ or } x \in U \\ &\text{ iff } x \in (S \cup T) \cup U. \end{aligned}$$

Notice that the proof of the associativity of union $\cup$ depends on the associativity of "or" as a logical connective.

Since we are asked to show that two sets are equal, some persons feel it necessary to write a chain of equalities between sets:

$$\begin{aligned} S \cup (T \cup U) &= \{x \colon x \in S \cup (T \cup U)\} \\ &= \{x \colon x \in S \text{ or } x \in T \cup U\} \\ &= \{x \colon x \in S \text{ or } (x \in T \text{ or } x \in U)\} \\ &= \{x \colon (x \in S \text{ or } x \in T) \text{ or } x \in U\} \\ &= \{x \colon x \in S \cup T \text{ or } x \in U\} \\ &= \{x \colon x \in (S \cup T) \cup U\} \\ &= (S \cup T) \cup U. \end{aligned}$$

This second proof is virtually identical to the first; it is just a bit more cluttered. Try to avoid clutter; mathematics is hard enough without it.

PROOF. (c) An element x belongs to $S \cup \emptyset$ if and only if $x \in S$ or $x \in \emptyset$. Since $x \in \emptyset$ is never true, $x \in S \cup \emptyset$ if and only if $x \in S$. That is, $S \cup \emptyset = S$.

PROOF. (d) To prove that $S \subseteq S \cup T$, show that $x \in S$ implies $x \in S \cup T$. Suppose $x \in S$. Then it is certainly true that $x \in S$ or $x \in T$; that is, $x \in S \cup T$.

PROOF. (e) First show that $S = S \cup T$ implies $T \subseteq S$; then prove the converse, if $T \subseteq S$, then $S = S \cup T$. To prove that $S = S \cup T$ implies $T \subseteq S$, it suffices to prove the contrapositive. We suppose that $T \nsubseteq S$ and show that $S \neq S \cup T$. If $T \nsubseteq S$, then there is at least one element t in T which does not belong to S. Thus (by parts (d) and (b))

$$t \in T \subseteq T \cup S = S \cup T;$$

but $t \notin S$. Since t belongs to $S \cup T$ but not to S these sets are not equal.

Now for the converse. Suppose $T \subseteq S$. Since we already know that $S \subseteq S \cup T$ (by part (d)), we need only show that $S \cup T \subseteq S$ in order to prove that the sets S and $S \cup T$ are identical. To this end suppose that $x \in S \cup T$. Then $x \in S$ or $x \in T \subseteq S$. In either case $x \in S$. Thus $S \cup T \subseteq S$.

□

F.1.3. Problem. Prove parts (b) and (f) of proposition F.1.2.

On numerous occasions it is necessary for us to take the union of a large (perhaps infinite) family of sets. When we consider a family of sets (that is, a set whose members are themselves sets), it is important to keep one thing in mind. If x is a member of a set S and S is in turn a member of a family $\mathfrak{S}$ of sets, it does not follow that $x \in \mathfrak{S}$. For example, let $S = \{0, 1, 2\}$, $T = \{2, 3, 4\}$, $U = \{5, 6\}$, and $\mathfrak{S} = \{S, T, U\}$. Then 1 is a member of S and S belongs to $\mathfrak{S}$; but 1 is not a member of $\mathfrak{S}$ (because $\mathfrak{S}$ has only 3 members: S, T, and U).

F.1.4. Definition. Let $\mathfrak{S}$ be a family of sets. We define the UNION of the family $\mathfrak{S}$ to be the set of all x such that $x \in S$ for at least one set S in $\mathfrak{S}$. We denote the union of the family $\mathfrak{S}$ by $\bigcup \mathfrak{S}$ (or by $\bigcup_{S \in \mathfrak{S}} S$, or by $\bigcup \{S : S \in \mathfrak{S}\}$). Thus $x \in \bigcup \mathfrak{S}$ if and only if there exists $S \in \mathfrak{S}$ such that $x \in S$.

F.1.5. Notation. If $\mathfrak{S}$ is a finite family of sets $S_1, \dots, S_n$, then we may write $\bigcup_{k=1}^{n} S_k$ or $S_1 \cup S_2 \cup \dots \cup S_n$ for $\bigcup \mathfrak{S}$.

F.1.6. Example. Let $S = \{0, 1, 3\}$, $T = \{1, 2, 3\}$, $U = \{1, 3, 4, 5\}$, and $\mathfrak{S} = \{S, T, U\}$. Then

$$\bigcup \mathfrak{S} = S \cup T \cup U = \{0, 1, 2, 3, 4, 5\}.$$

The following very simple observations are worthy of note.

F.1.7. Proposition. *If $\mathfrak{S}$ is a family of sets and $T \in \mathfrak{S}$, then $T \subseteq \bigcup \mathfrak{S}$.*

PROOF. If $x \in T$, then x belongs to at least one of the sets in $\mathfrak{S}$, namely T. $\square$

F.1.8. Proposition. *If $\mathfrak{S}$ is a family of sets and each member of $\mathfrak{S}$ is contained in a set U, then $\bigcup \mathfrak{S} \subseteq U$.*

PROOF. Problem.

F.2. INTERSECTIONS

F.2.1. Definition. Let S and T be sets. The INTERSECTION of S and T is the set of all x such that $x \in S$ and $x \in T$.

F.2.2. Example. If $S = [0, 3]$ and $T = [2, 5]$, then $S \cap T = [2, 3]$.

F.2.3. Proposition. *Let S, T, U, and V be sets. Then*

(a) $S \cap (T \cap U) = (S \cap T) \cap U$; *(associativity)*
(b) $S \cap T = T \cap S$; *(commutativity)*
(c) $S \cap \emptyset = \emptyset$;
(d) $S \cap T \subseteq S$;
(e) $S = S \cap T$ *if and only if* $S \subseteq T$;
(f) *If* $S \subseteq T$ *and* $U \subseteq V$, *then* $S \cap U \subseteq T \cap V$.

PROOF. Problem.

There are two distributive laws for sets: union distributes over intersection (proposition F.2.4 below) and intersection distributes over union (proposition F.2.5).

F.2.4. Proposition. *Let S, T, and U be sets. Then*

$$S \cup (T \cap U) = (S \cup T) \cap (S \cup U).$$

PROOF. Exercise. *Hint.* Use problem D.1.4. (Solution Q.31.1.)

F.2.5. Proposition. *Let S, T, and U be sets. Then*

$$S \cap (T \cup U) = (S \cap T) \cup (S \cap U).$$

PROOF. Problem.

Just as we may take the union of an arbitrary family of sets, we may also take its intersection.

F.2.6. Definition. Let $\mathfrak{S}$ be a family of sets. We define the INTERSECTION of the family $\mathfrak{S}$ to be the set of all x such that $x \in \mathfrak{S}$ for every S in $\mathfrak{S}$. We denote the intersection of $\mathfrak{S}$ by $\bigcap \mathfrak{S}$ (or by $\bigcap_{S \in \mathfrak{S}} S$, or by $\bigcap\{S : S \in \mathfrak{S}\}$).

F.2.7. Notation. If S is a finite family of sets $S_1, \dots, S_n$, then we may write $\bigcap_{k=1}^{n} S_k$ or $S_1 \cap S_2 \cap \dots \cap S_n$ for $\bigcap \mathfrak{S}$. Similarly, if $\mathfrak{S} = \{S_1, S_2, \dots\}$, then we may write $\bigcap_{k=1}^{\infty} S_k$ or $S_1 \cap S_2 \cap \dots$ for $\bigcap \mathfrak{S}$.

F.2.8. Example. Let $S = \{0, 1, 3\}$, $T = \{1, 2, 3\}$, $U = \{1, 3, 4, 5\}$, and $S = \{S, T, U\}$. Then

$$\bigcap S = S \cap T \cap U = \{1, 3\}.$$

Proposition F.2.4 may be generalized to say that union distributes over the intersection of an arbitrary family of sets. Similarly there is a more general form of proposition F.2.5 which says that intersection distributes over the union of an arbitrary family of sets. These two facts, which are stated precisely in the next two propositions, are known as GENERALIZED DISTRIBUTIVE LAWS.

F.2.9. Proposition. *Let T be a set and $\mathfrak{S}$ be a family of sets. Then*

$$T \cup \left(\bigcap \mathfrak{S}\right) = \bigcap\{T \cup S : S \in \mathfrak{S}\}.$$

PROOF. Exercise. (Solution Q.31.2.)

F.2.10. Proposition. *Let T be a set and $\mathfrak{S}$ be a family of sets. Then*

$$T \cap \left(\bigcup \mathfrak{S}\right) = \bigcup \{T \cap S : S \in \mathfrak{S}\}.$$

Proof. Problem.

F.2.11. Definition. Sets S and T are said to be DISJOINT if $S \cap T = \emptyset$. More generally, a family $\mathfrak{S}$ of sets is a DISJOINT FAMILY (or a PAIRWISE DISJOINT FAMILY) if $S \cap T = \emptyset$ whenever S and T are distinct (that is, not equal) sets which belong to $\mathfrak{S}$.

CAUTION. Let $\mathfrak{S}$ be a family of sets. Do not confuse the following two statements.

(a) $\mathfrak{S}$ is a (pairwise) disjoint family.
(b) $\bigcap \mathfrak{S} = \emptyset$.

Certainly, if $\mathfrak{S}$ contains more than a single set, then (a) implies (b). But if $\mathfrak{S}$ contains three or more sets the converse need not hold. For example, let $S = \{0, 1\}$, $T = \{3, 4\}$, $U = \{0, 2\}$, and $\mathfrak{S} = \{S, T, U\}$. Then $\mathfrak{S}$ is not a disjoint family (because $S \cap U$ is nonempty), but $\bigcap \mathfrak{S} = \emptyset$.

F.2.12. Example. Let S, T, U, and V be sets.

(a) Then $(S \cap T) \cup (U \cap V) \subseteq (S \cup U) \cap (T \cup V)$.
(b) Give an example to show that equality need not hold in (a).

Proof. Problem. *Hint.* Use propositions F.1.2(d) and F.2.3(f) to show that $S \cap T$ and $U \cap V$ are contained in $(S \cup U) \cap (T \cup V)$. Then use F.1.2(f).

F.3. COMPLEMENTS

Recall that we regard all the sets with which we work in a particular situation as being subsets of some appropriate "universal" set. For each set S we define the COMPLEMENT of S , denoted by S^c, to be the set of all members of our universal set which do not belong to S. That is, we write $x \in S^c$ if and only if $x \notin S$.

F.3.1. Example. Let S be the closed interval $(-\infty, 3]$. If nothing else is specified, we think of this interval as being a subset of the real line $\mathbb{R}$ (our universal set). Thus S^c is the set of all x in $\mathbb{R}$ such that x is not less than or equal to 3. Thus S^c is the interval $(3, \infty)$.

F.3.2. Example. Let S be the set of all points (x, y) in the plane such that $x \geq 0$ and $y \geq 0$. Then S^c is the set of all points (x, y) in the plane such that either $x < 0$ or $y < 0$. That is ,

$$S^c = \{(x, y) \colon x < 0\} \cup \{(x, y) \colon y < 0\}.$$

The two following propositions are De Morgan's laws for sets. As you may expect, they are obtained by translating into the language of sets the facts of logic which go under the same name. (See D.4.1 and D.4.2.)

F.3.3. Proposition. *Let S and T be sets. Then*

$$(S \cup T)^c = S^c \cap T^c.$$

Proof. Exercise. *Hint.* Use example D.4.1. (Solution Q.31.3.)

F.3.4. Proposition. *Let S and T be sets. Then*

$$(S \cap T)^c = S^c \cup T^c.$$

Proof. Problem.

Just as the distributive laws can be generalized to arbitrary families of sets, so too can De Morgan's laws. The complement of the union of a family is the intersection of the complements (proposition F.3.5), and the complement of the intersection of a family is the union of the complements (proposition F.3.6).

F.3.5. Proposition. *Let $\mathfrak{S}$ be a family of sets. Then*

$$(\bigcup \mathfrak{S})^c = \bigcap \{S^c \colon S \in \mathfrak{S}\}\,.$$

Proof. Exercise. (Solution Q.31.4.)

F.3.6. Proposition. *Let $\mathfrak{S}$ be a family of sets. Then*

$$(\bigcap \mathfrak{S})^c = \bigcup \{S^c \colon S \in \mathfrak{S}\}\,.$$

Proof. Problem.

F.3.7. Definition. If S and T are sets we define the COMPLEMENT OF T RELATIVE TO S, denoted by $S \setminus T$, to be the set of all x which belong to S but not to T. That is,

$$S \setminus T := S \cap T^c\,.$$

The operation $\setminus$ is usually called SET SUBTRACTION and $S \setminus T$ is read as "S minus T".

F.3.8. Example. Let $S = [0, 5]$ and $T = [3, 10]$. Then $S \setminus T = [0, 3)$.

It is a frequently useful fact that the union of two sets can be rewritten as a disjoint union (that is, the union of two disjoint sets).

F.3.9. Proposition. *Let S and T be sets. Then $S \setminus T$ and T are disjoint sets whose union is $S \cup T$.*

Proof. Exercise. (Solution Q.31.5.)

F.3.10. Exercise. Show that $(S \setminus T) \cup T = S$ if and only if $T \subseteq S$. (Solution Q.31.6.)

F.3.11. Problem. Let $S = (3, \infty)$, $T = (0, 10]$, $U = (-4, 5)$, $V = [-2, 8]$, and $\mathfrak{S} = \{S^c, T, U, V\}$.

(a) Find $\bigcup \mathfrak{S}$.
(b) Find $\bigcap \mathfrak{S}$.

F.3.12. Problem. If S, T, and U are sets, then

$$(S \cap T) \setminus U = (S \setminus U) \cap (T \setminus U)\,.$$

F.3.13. Problem. If S, T, and U are sets, then

$$S \cap (T \setminus U) = (S \cap T) \setminus (S \cap U)\,.$$

F.3.14. Problem. If S and T are sets, then $T \setminus S$ and $T \cap S$ are disjoint and

$$T = (T \setminus S) \cup (T \cap S)\,.$$

F.3.15. Problem. If S and T are sets, then $S \cap T = S \setminus (S \setminus T)$.

F.3.16. Definition. A family $\mathfrak{S}$ of sets COVERS (or is a COVER FOR, or is a COVERING FOR) a set T if $T \subseteq \bigcup S$.

F.3.17. Problem. Find a family of open intervals which covers the set $\mathbb{N}$ of natural numbers and has the property that the sum of the lengths of the intervals is 1. *Hint.* $\sum_{k=1}^{\infty} 2^{-k} = 1$.

APPENDIX G

ARITHMETIC

G.1. THE FIELD AXIOMS

The set $\mathbb{R}$ of real numbers is the cornerstone of calculus. It is remarkable that all of its properties can be derived from a very short list of axioms. We will not travel the rather lengthy road of deriving from these axioms all the properties (arithmetic of fractions, rules of exponents, *etc.*) of $\mathbb{R}$ which we use in this text. This journey, although interesting enough in itself, requires a substantial investment of time and effort. Instead we discuss briefly one standard set of axioms for $\mathbb{R}$ and, with the aid of these axioms, give sample derivations of some familiar properties of $\mathbb{R}$. In the present chapter we consider the first four axioms, which govern the operations on $\mathbb{R}$ of addition and multiplication. The name we give to the collective consequences of these axioms is *arithmetic*.

G.1.1. Definition. A BINARY OPERATION $*$ on a set S is a rule that associates with each pair x and y of elements in S one and only one element $x * y$ in S. (More precisely, $*$ is a function from $S \times S$ into S. See appendices K and N.)

The first four axioms say that the set $\mathbb{R}$ of real numbers under the binary operations of addition and multiplication (denoted, as usual, by $+$ and $\cdot$) form a FIELD. We will follow conventional practice by allowing xy as a substitute notation for $x \cdot y$.

G.1.2. Axiom (I). *The operations $+$ and $\cdot$ on $\mathbb{R}$ are associative (that is, $x + (y + z) = (x + y) + z$ and $x(yz) = (xy)z$ for all x, y, $z \in \mathbb{R}$) and commutative ($x + y = y + x$ and $xy = yx$ for all x, $y \in \mathbb{R}$).*

G.1.3. Axiom (II). *There exist distinct additive and multiplicative identities (that is, there are elements 0 and 1 in $\mathbb{R}$ with $1 \neq 0$ such that $x + 0 = x$ and $x \cdot 1 = x$ for all $x \in \mathbb{R}$).*

G.1.4. Axiom (III). *Every element x in $\mathbb{R}$ has an additive inverse (that is, a number $-x$ such that $x + (-x) = 0$); and every element $x \in \mathbb{R}$ different from 0 has a multiplicative inverse (that is, a number x^{-1} such that $x\,x^{-1} = 1$).*

G.1.5. Axiom (IV). *Multiplication distributes over addition (that is, $x(y + z) = xy + xz$ for all x, y, $z \in \mathbb{R}$).*

G.1.6. Example. Multiplication is not a binary operation on the set $\mathbb{R}' = \{x \in \mathbb{R} \colon x \neq -1\}$.

PROOF. The numbers 2 and $-\frac{1}{2}$ belong to $\mathbb{R}'$, but their product does not. $\square$

G.1.7. Example. Subtraction is a binary operation on the set $\mathbb{R}$ of real numbers, but is neither associative nor commutative.

PROOF. Problem.

G.1.8. Problem. Let $\mathbb{R}^+$ be the set of all real numbers x such that $x > 0$. On $\mathbb{R}^+$ define

$$x * y = \frac{xy}{x + y}\,.$$

Determine whether $*$ is a binary operation on $\mathbb{R}^+$. Determine whether $*$ is associative and whether it is commutative. Does $\mathbb{R}^+$ have an identity element with respect to $*$? (That is, is there a member e of $\mathbb{R}^+$ such that $x * e = x$ and $e * x = x$ for all x in $\mathbb{R}^+$?)

Subtraction and division are defined in terms of addition and multiplication by

$$x - y := x + (-y)$$

and, for $y \neq 0$,

$$\frac{x}{y} := xy^{-1}\,.$$

We use the familiar rule for avoiding an excess of parentheses: multiplication takes precedence over addition. Thus, for example, $wx + yz$ means $(wx) + (yz)$.

G.1.9. Problem. The rule given in Axiom IV is the *left distributive law.* The *right distributive law,* $(x + y)z = xz + yz$, is also true. Use the axioms to prove it.

G.1.10. Exercise. Show that if x is a real number such that $x+x = x$, then $x = 0$. *Hint.* Simplify both sides of $(x + x) + (-x) = x + (x + (-x))$. (Solution Q.32.1.)

G.1.11. Problem. Show that the additive identity 0 annihilates everything in $\mathbb{R}$ under multiplication. That is, show that $0 \cdot x = 0$ for every real number x. *Hint.* Consider $(0 + 0)x$. Use G.1.9 and G.1.10.

G.1.12. Exercise. Give a careful proof using only the axioms above that if w, x, y, and z are real numbers, then

$$(w + x) + (y + z) = z + (x + (y + w))\,.$$

Hint. Since we are to make explicit use of the associative law, be careful not to write expressions such as $w + x + (y + z)$. Another set of parentheses is needed to indicate the order of operations. Both $(w + x) + (y + z)$ and $w + (x + (y + z))$, for example, do make sense. (Solution Q.32.2.)

G.1.13. Problem. Show that if the product xy of two numbers is zero, then either $x = 0$ or $y = 0$. (Here the word "or" is used in its inclusive sense; both x and y may be 0. It is always used that way in mathematics.) *Hint.* Convince yourself that, as a matter of logic, it is enough to show that if y is not equal to 0 then x must be. Consider $(xy)y^{-1}$ and use G.1.11.

G.2. UNIQUENESS OF IDENTITIES

Axiom II guarantees only the *existence* of additive and multiplicative identities 0 and 1. It is natural to enquire about their *uniqueness.* Could there be two real numbers which act as additive identities? That is, could we have numbers $0' \neq 0$ which satisfy

$$x + 0 = x \tag{G.1}$$

and

$$x + 0' = x \tag{G.2}$$

for all x in $\mathbb{R}$? The answer as you would guess is *no:* there is only one additive identity in $\mathbb{R}$. The proof is very short.

G.2.1. Proposition. *The additive identity in* $\mathbb{R}$ *is unique.*

Proof. Suppose that the real numbers 0 and $0'$ satisfy (G.1) and (G.2) all real numbers x. Then

$$\begin{aligned} 0 &= 0 + 0' \\ &= 0' + 0 \\ &= 0'. \end{aligned}$$

The three equalities are justified, respectively, by (G.2), axiom I, and (G.1). □

G.2.2. Proposition. *The multiplicative identity* 1 *on* $\mathbb{R}$ *is unique.*

Proof. Problem.

G.3. UNIQUENESS OF INVERSES

The question of uniqueness also arises for inverses; only their existence is guaranteed by axiom III. Is it possible for a number to have more than one additive inverse? That is, if x is a real number is it possible that there are two different numbers, say $-x$ and $\overline{x}$, such that the equations

$$x + (-x) = 0 \qquad \text{and} \qquad x + \overline{x} = 0 \tag{G.3}$$

both hold? The answer is *no*.

G.3.1. Proposition. *Additive inverses in $\mathbb{R}$ are unique.*

Proof. Assume that the equations (G.3) are true. We show that $-x$ and $\overline{x}$ are the same number.

$$\begin{aligned}\overline{x} &= \overline{x} + 0\\ &= \overline{x} + (x + (-x))\\ &= (\overline{x} + x) + (-x)\\ &= (x + \overline{x}) + (-x)\\ &= 0 + (-x)\\ &= (-x) + 0\\ &= -x.\end{aligned}$$

□

G.3.2. Problem. The proof of proposition (G.3.1) contains seven equal signs. Justify each one.

G.3.3. Problem. Prove that in $\mathbb{R}$ multiplicative inverses are unique.

G.3.4. Example. Knowing that identities and inverses are unique is helpful in deriving additional properties of the real numbers. For example, the familiar fact that

$$-(-x) = x$$

follows immediately from the equation

$$(-x) + x = 0. \tag{G.4}$$

What proposition G.3.1 tells us is that if $a + b = 0$ then b must be the additive inverse of a. So from (G.4) we conclude that x must be the additive inverse of $-x$; in symbols, $x = -(-x)$.

G.3.5. Problem. Show that if x is a nonzero real number, then

$$\left(x^{-1}\right)^{-1} = x\,.$$

G.4. ANOTHER CONSEQUENCE OF UNIQUENESS

We can use proposition G.3.1 to show that in $\mathbb{R}$

$$-\,(x + y) = -x - y. \tag{G.5}$$

Before looking at the proof of this assertion it is well to note the two uses of the "-" sign on the right side of (G.5). The first, attached to "x", indicates the additive inverse of x; the second indicates subtraction. Thus $-x - y$ means $(-x) + (-y)$. The idea behind the proof is to add the right side of (G.5) to $x + y$. If the result is 0, then the uniqueness of additive inverses, proposition G.3.1, tells us that $-x-y$ is the additive inverse of $x+y$. And that is exactly what we get:

$$\begin{aligned}(x+y)+(-x-y) &= (x+y)+((-x)+(-y))\\ &= (y+x)+((-x)+(-y))\\ &= y+(x+((-x)+(-y)))\\ &= y+((x+(-x))+(-y))\\ &= y+(0+(-y))\\ &= (y+0)+(-y)\\ &= y+(-y)\\ &= 0.\end{aligned}$$

G.4.1. Problem. Justify each step in the proof of equation (G.5).

G.4.2. Problem. Prove that if x and y are nonzero real numbers, then

$$(xy)^{-1} = y^{-1}x^{-1}\,.$$

G.4.3. Problem. Show that

$$(-1)x = -x$$

for every real number x. *Hint.* Use the uniqueness of additive inverses.

G.4.4. Problem. Show that

$$-(xy) = (-x)y = x(-y)$$

and that

$$(-x)(-y) = xy$$

for all x and y in $\mathbb{R}$. *Hint.* For the first equality add $(-x)y$ to xy.

G.4.5. Problem. Use the first four axioms for $\mathbb{R}$ to develop the rules for adding, multiplying, subtracting, and dividing fractions. Show for example that

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$$

if b and d are not zero. (Remember that, by definition, $\dfrac{a}{b}+\dfrac{c}{d}$ is $ab^{-1}+cd^{-1}$ and $\dfrac{ad+bc}{bd}$ is $(ad+bc)(bd)^{-1}$.)

APPENDIX H

ORDER PROPERTIES OF $\mathbb{R}$

The second group of axioms are the order axioms. They concern a subset $\mathbb{P}$ of $\mathbb{R}$ (call this the set of STRICTLY POSITIVE numbers).

H.1.1. Axiom (V)**.** *The set $\mathbb{P}$ is closed under addition and multiplication. (That is, if x and y belong to $\mathbb{P}$, so do $x + y$ and xy.)*

H.1.2. Axiom (VI)**.** *For each real number x exactly one of the following is true: $x = 0$, $x \in \mathbb{P}$, or $-x \in \mathbb{P}$. This is the* axiom of trichotomy.

Define the relation $<$ on $\mathbb{R}$ by

$$x < y \text{ if and only if } y - x \in \mathbb{P}\,.$$

Also define $>$ on $\mathbb{R}$ by

$$x > y \text{ if and only if } y < x\,.$$

We write $x \leq y$ if $x < y$ or $x = y$, and $x \geq y$ if $y \leq x$.

H.1.3. Proposition. *On $\mathbb{R}$ the relation $<$ is transitive (that is, if $x < y$ and $y < z$, then $x < z$).*

PROOF. If $x < y$ and $y < z$, then $y - x$ and $z - y$ belong to $\mathbb{P}$. Thus

$$\begin{aligned}
z - x &= z + (-x) \\
&= (z + 0) + (-x) \\
&= (z + (y + (-y))) + (-x) \\
&= (z + ((-y) + y)) + (-x) \\
&= ((z + (-y)) + y) + (-x) \\
&= (z + (-y)) + (y + (-x)) \\
&= (z - y) + (y - x) \in \mathbb{P}.
\end{aligned}$$

This shows that $x < z$. $\square$

H.1.4. Problem. Justify each of the seven equal signs in the proof of proposition H.1.3.

H.1.5. Exercise. Show that a real number x belongs to the set $\mathbb{P}$ if and only if $x > 0$. (Solution Q.33.1.)

H.1.6. Proposition. *If $x > 0$ and $y < z$ in $\mathbb{R}$, then $xy < xz$.*

PROOF. Exercise. *Hint.* Use problem G.4.4. (Solution Q.33.2.)

H.1.7. Proposition. *If x, y, $z \in \mathbb{R}$ and $y < z$, then $x + y < x + z$.*

PROOF. Problem. *Hint.* Use equation (G.5).

H.1.8. Proposition. *If $w < x$ and $y < z$, then $w + y < x + z$.*

PROOF. Problem.

H.1.9. Problem. Show that $1 > 0$. *Hint.* Keep in mind that 1 and 0 are assumed to be distinct. (Look at the axiom concerning additive and multiplicative identities.) If 1 does not belong to $\mathbb{P}$, what can you say about the number -1? What about $(-1)(-1)$? Use problem G.4.4.

H.1.10. Proposition. *If $x > 0$, then $x^{-1} > 0$.*

PROOF. Problem.

H.1.11. Proposition. *If $0 < x < y$, then $1/y < 1/x$.*

PROOF. Problem.

H.1.12. Proposition. *If $0 < w < x$ and $0 < y < z$, then $wy < xz$.*

PROOF. Exercise. (Solution Q.33.3.)

H.1.13. Problem. Show that $x < 0$ if and only if $-x > 0$.

H.1.14. Problem. Show that if $y < z$ and $x < 0$, then $xz < xy$.

H.1.15. Problem. Show that $x < y$ if and only if $-y < -x$.

H.1.16. Problem. Suppose that x, $y \geq 0$ and $x^2 = y^2$. Show that $x = y$.

H.1.17. Problem. Show in considerable detail how the preceding results can be used to solve the inequality

$$\frac{5}{x+3} < 2 - \frac{1}{x-1}\,.$$

H.1.18. Problem. Let $\mathbb{C} = \{(a, b) \colon a, b \in \mathbb{R}\}$. On $\mathbb{C}$ define two binary operations $+$ and $\cdot$ by:

$$(a, b) + (c, d) = (a + c, b + d)$$

and

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

Show that $\mathbb{C}$ under these operations is a field. (That is, $\mathbb{C}$ satisfies axioms I–IV.) This is the field of COMPLEX NUMBERS.

Determine whether it is possible to make $\mathbb{C}$ into an ordered field. (That is, determine whether it is possible to choose a subset $\mathbb{P}$ of $\mathbb{C}$ which satisfies axioms V and VI.)

The axioms presented thus far define an ordered field. To obtain the particular ordered field $\mathbb{R}$ of real numbers we require one more axiom. We assume that $\mathbb{R}$ is order complete; that is, $\mathbb{R}$ satisfies the *least upper bound axiom.* This axiom will be stated (and discussed in some detail) in chapter J (in particular, see J.3.1).

There is a bit more to the axiomatization of $\mathbb{R}$ than we have indicated in the preceding discussion. For one thing, how do we know that the axioms are consistent? That is, how do we know that they will not yield a contradiction? For this purpose one constructs a *model* for $\mathbb{R}$, that is, a concrete mathematical object which satisfies all the axioms for $\mathbb{R}$. One standard procedure is to define the positive integers in terms of sets: 0 is the empty set $\emptyset$, the number 1 is the set whose only element is 0, the number 2 is the set whose only element is 1, and so on. Using the positive integers we construct the set $\mathbb{Z}$ of all integers $\dots, -2, -1, 0, 1, 2, \dots$. From these we construct the set $\mathbb{Q}$ of rational numbers (that is, numbers of the form p/q where p and q are integers and $q \neq 0$). Finally the reals are constructed from the rationals.

Another matter that requires attention is the use of the definite article in the expression "*the* real numbers". This makes sense only if the axioms are shown to be categorical; that is, if there is "essentially" only one model for the axioms. This turns out to be correct about the axioms for $\mathbb{R}$ given an appropriate technical meaning of "essentially"—but we will not pursue this matter here. More about both the construction of the reals and their uniqueness can be found in [10].

APPENDIX I

NATURAL NUMBERS AND MATHEMATICAL INDUCTION

The *principle of mathematical induction* is predicated on a knowledge of the natural numbers, which we introduce in this section. In what follows it is helpful to keep in mind that the axiomatic development of the real numbers sketched in the preceding chapter says nothing about the natural numbers; in fact, it provides explicit names for only three real numbers: 0, 1, and -1.

I.1.1. Definition. A collection J of real numbers is INDUCTIVE if

(1) $1 \in J$, and
(2) $x + 1 \in J$ whenever $x \in J$.

I.1.2. Example. The set $\mathbb{R}$ is itself inductive; so are the intervals $(0, \infty)$, $[-1, \infty)$, and $[1, \infty)$.

I.1.3. Proposition. *Let $\mathfrak{A}$ be a family of inductive subsets of $\mathbb{R}$. Then $\cap\mathfrak{A}$ is inductive.*

PROOF. Exercise. (Solution Q.34.1.)

I.1.4. Definition. Let J be the family of all inductive subsets of $\mathbb{R}$. Define

$$\mathbb{N} := \cap J\,.$$

We call $\mathbb{N}$ the set of NATURAL NUMBERS.

Notice that according to proposition I.1.3 the set $\mathbb{N}$ is inductive. It is the *smallest* inductive set, in the sense that it is contained in *every* inductive set. The elements of $\mathbb{N}$ have standard names. Define $2 := 1 + 1$. Since 1 belongs to $\mathbb{N}$ and $\mathbb{N}$ is inductive, 2 belongs to $\mathbb{N}$. Define $3 := 2 + 1$. Since 2 belongs to $\mathbb{N}$, so does 3. Define $4 := 3 + 1$; *etc.*

I.1.5. Definition. The set of INTEGERS, denoted by $\mathbb{Z}$, is defined to be

$$-\mathbb{N} \cup \{0\} \cup \mathbb{N}$$

where $-\mathbb{N} := \{-n : n \in \mathbb{N}\}$.

The next proposition is practically obvious; but as it is an essential ingredient of several subsequent arguments (*e.g.* problem I.1.16), we state it formally.

I.1.6. Proposition. *If $n \in \mathbb{N}$, then $n \geq 1$.*

PROOF. Since the set $[1, \infty)$ is inductive, it must contain $\mathbb{N}$. □

The observation made previously that $\mathbb{N}$ is the smallest inductive set clearly implies that no proper subset of $\mathbb{N}$ can be inductive. This elementary fact has a rather fancy name: it is the *principle of mathematical induction.*

I.1.7. Theorem (Principle of Mathematical Induction)**.** *Every inductive subset of $\mathbb{N}$ equals $\mathbb{N}$.*

By spelling out the definition of "inductive set" in the preceding theorem we obtain a longer, perhaps more familiar, statement of the *principle of mathematical induction.*

I.1.8. Corollary. *If S is a subset of $\mathbb{N}$ which satisfies*

(1) $1 \in S$, *and*
(2) $n + 1 \in S$ *whenever* $n \in S$,

then $S = \mathbb{N}$.

Perhaps even more familiar is the version of the preceding which refers to "a proposition (or assertion, or statement) concerning the natural number n".

I.1.9. Corollary. *Let $P(n)$ be a proposition concerning the natural number n. If $P(1)$ is true, and if $P(n+1)$ is true whenever $P(n)$ is true, then $P(n)$ is true for all $n \in \mathbb{N}$.*

PROOF. In corollary I.1.8 let $S = \{n \in \mathbb{N} : P(n) \text{ is true}\}$. Then $1 \in S$ and $n+1$ belongs to S whenever n does. Thus $S = \mathbb{N}$. That is, $P(n)$ is true for all $n \in \mathbb{N}$. □

I.1.10. Exercise. Use *mathematical induction* to prove the following assertion: The sum of the first n natural numbers is $\frac{1}{2}n(n+1)$. *Hint.* Recall that if p and q are integers with $p \leq q$ and if $c_p, c_{p+1}, \dots, c_q$ are real numbers, then the sum $c_p + c_{p+1} + \dots + c_q$ may be denoted by $\sum_{k=p}^{q} c_k$. Using this summation notation, we may write the desired conclusion as $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$.) (Solution Q.34.2.)

It is essentially obvious that there is nothing crucial about starting inductions with $n = 1$. Let m be any integer and $P(n)$ be a proposition concerning integers $n \geq m$. If we prove that $P(m)$ is true and that $P(n+1)$ is true whenever $P(n)$ is true and $n \geq m$, then we may conclude that $P(n)$ is true for all $n \geq m$. [Proof: Apply corollary I.1.9 to the proposition Q where $Q(n) = P(n+m-1)$.]

I.1.11. Problem. Let $a, b \in \mathbb{R}$ and $m \in \mathbb{N}$. Then

$$a^m - b^m = (a-b) \sum_{k=0}^{m-1} a^k b^{m-k-1} .$$

Hint. Multiply out the right hand side. This is not an induction problem.

I.1.12. Problem. If $r \in \mathbb{R}$, $r \neq 1$, and $n \in \mathbb{N}$, then

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r} .$$

Hint. Use I.1.11

I.1.13. Definition. Let m, $n \in \mathbb{N}$. We say that m is a FACTOR of n if $n/m \in \mathbb{N}$. Notice that 1 and n are always factors of n; these are the TRIVIAL FACTORS of n. The number n is COMPOSITE if $n > 1$ and if it has at least one nontrivial factor. (For example, 20 has several nontrivial factors: 2, 4, 5, and 10. Therefore it is composite.) If $n > 1$ and it is not composite, it is PRIME. (For example, 7 is prime; its only factors are 1 and 7.)

I.1.14. Problem. Prove that if $n \in \mathbb{N}$ and $2^n - 1$ is prime, then so is n. (*Hint.* Prove the contrapositive. Use problem I.1.11.) Illustrate your technique by finding a nontrivial factor of $2^{403} - 1$.

I.1.15. Problem. Show that $\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$ for every $n \in \mathbb{N}$.

I.1.16. Problem. Use only the definition of $\mathbb{N}$ and the results given in this section to prove (a) and (b).

(a) If $m, n \in \mathbb{N}$ and $m < n$, then $n - m \in \mathbb{N}$. *Hint.* One proof of this involves an induction within an induction. Restate the assertion to be proved as follows. For every $m \in \mathbb{N}$ it is true that:

$$\text{if } n \in \mathbb{N} \text{ and } n > m, \text{ then } n - m \in \mathbb{N}. \tag{I.1}$$

Prove this assertion by induction on m. That is, show that (I.1) holds for $m = 1$, and then show that it holds for $m = k + 1$ provided that it holds for $m = k$. To show that (I.1) holds for $m = 1$, prove that the set

$$J := \{1\} \cup \{n \in \mathbb{N} : n - 1 \in \mathbb{N}\}$$

is an inductive set.

(b) Let $n \in \mathbb{N}$. There does not exist a natural number k such that $n < k < n+1$. *Hint.* Argue by contradiction. Use part (a).

I.1.17. Problem (The binomial theorem.)**.** If x, $y \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \,.$$

Hint. Use induction. Recall that $0! = 1$, that $n! = n(n-1)(n-2)\dots 1$ for $n \in \mathbb{N}$, and that

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

for $0 \le k \le n$.

The final result of this section is the *principle of well-ordering*. It asserts that every nonempty subset of the natural numbers has a smallest element.

I.1.18. Proposition. *If $\emptyset \neq K \subseteq \mathbb{N}$, then there exists $a \in K$ such that $a \le k$ for every $k \in K$.*

Proof. Exercise. *Hint.* Assume that K has no smallest member. Show that this implies $K = \emptyset$ by proving that the set

$$J \equiv \{n \in \mathbb{N} \colon n < k \text{ for every } k \in K\}$$

is inductive. In the inductive step problem I.1.16(b) may prove useful. (Solution Q.34.3.)

I.1.19. Problem. (This slight modification of the *principle of mathematical induction* is occasionally useful.) Let $P(n)$ be a proposition concerning the natural number n. If $P(n)$ is true whenever $P(k)$ is true for all $k \in \mathbb{N}$ such that $k < n$, then $P(n)$ is true for all n. *Hint.* Use the *well-ordering principle.*

APPENDIX J

LEAST UPPER BOUNDS AND GREATEST LOWER BOUNDS

The last axiom for the set $\mathbb{R}$ of real numbers is the *least upper bound axiom.* Before stating it we make some definitions.

J.1. UPPER AND LOWER BOUNDS

J.1.1. Definition. A number u is an UPPER BOUND for a set A of real numbers if $u \geq a$ for every $a \in A$. If the set A has at least one upper bound, it is said to be BOUNDED ABOVE. Similarly, v is a LOWER BOUND for A if $v \leq a$ for every $a \in A$, and a set with at least one lower bound is BOUNDED BELOW. The set A is BOUNDED if it is bounded both above and below. (Perhaps it should be emphasized that when we say, for example, that A *has* an upper bound we mean only that there is a real number u which is greater than or equal to each member of A; we do *not* mean that u necessarily belongs to A—although of course it may.)

J.1.2. Example. The set $A = \{x \in \mathbb{R} : |x - 2| < 5\}$ is bounded.

PROOF. Problem.

J.1.3. Example. The open interval $(-1, 1)$ has infinitely many upper bounds. In fact, any set which is bounded above has infinitely many upper bounds.

PROOF. Problem.

J.1.4. Example. The set $A = \{x \in \mathbb{R} : x^3 - x \leq 0\}$ is not bounded.

PROOF. Problem.

J.2. LEAST UPPER AND GREATEST LOWER BOUNDS

J.2.1. Definition. A number ℓ is the SUPREMUM(or LEAST UPPER BOUND) of a set A if:

(1) ℓ is an upper bound for A, and
(2) $\ell \leq u$ whenever u is an upper bound for A.

If ℓ is the least upper bound of A, we write $\ell = \sup A$. Similarly, a lower bound g of a set is the INFIMUM(or GREATEST LOWER BOUND) of a set A if it is greater than or equal to every lower bound of the set. If g is the greatest lower bound of A, we write $g = \inf A$.

If A is not bounded above (and consequently, $\sup A$ does not exist), then it is common practice to write $\sup A = \infty$. Similarly, if A is not bounded below, we write $\inf A = -\infty$.

CAUTION. The expression "$\sup A = \infty$" does *not* mean that $\sup A$ exists and equals some object called ∞; it *does* mean that A is not bounded above.

It is clear that least upper bounds and greatest lower bounds, when they exist, are unique. If, for example, ℓ and m are both least upper bounds for a set A, then $\ell \leq m$ and $m \leq \ell$; so $\ell = m$.

J.2.2. Definition. Let $A \subseteq \mathbb{R}$. If there exists a number M *belonging to* A such that $M \geq a$ for every $a \in A$), then this element is the LARGEST ELEMENT (or GREATEST ELEMENT, or MAXIMUM) of A. We denote this element (when it exists) by $\max A$.

Similarly, if there exists a number m *belonging to* A such that $m \leq a$ for every $a \in A$), then this element is the SMALLEST ELEMENT (or LEAST ELEMENT, or MINIMUM) of A. We denote this element (when it exists) by $\min A$.

J.2.3. Example. Although the largest element of a set (when it exists) is always a least upper bound, the converse is not true. It is possible for a set to have a least upper bound but no maximum. The interval $(-2, 3)$ has a least upper bound (namely, 3), but it has no largest element.

J.2.4. Example. If $A = \{x \in \mathbb{R} \colon |x| < 4\}$, then $\inf A = -4$ and $\sup A = 4$. But A has no maximum or minimum. If $B = \{|x| \colon x < 4\}$, then $\inf B = 0$ but $\sup B$ does not exist. (It is correct to write $\sup B = \infty$.) Furthermore, B has a smallest element, $\min B = 0$, but no largest element.

Incidentally, the words "maximum", "supremum", "minimum", and "infimum" are all singular. The preferred plurals are, respectively, "maxima", "suprema", "minima", and "infima".

J.2.5. Problem. For each of the following sets find the least upper bound and the greatest lower bound (if they exist).

(a) $A = \{x \in \mathbb{R} \colon |x - 3| < 5\}$.
(b) $B = \{|x - 3| \colon x < 5\}$.
(c) $C = \{|x - 3| \colon x > 5\}$.

J.2.6. Problem. Show that the set $\mathbb{P}$ of positive real numbers has an infimum but no smallest element.

J.2.7. Exercise. Let $f(x) = x^2 - 4x + 3$ for every $x \in \mathbb{R}$, let $A = \{x \colon f(x) < 3\}$, and let $B = \{f(x) \colon x < 3\}$.

(a) Find $\sup A$ and $\inf A$ (if they exist).
(b) Find $\sup B$ and $\inf B$ (if they exist).

(Solution Q.35.1.)

J.2.8. Example. Let $A = \left\{x \in \mathbb{R} \colon \dfrac{5}{x-3} - 3 \geq 0\right\}$. Then $\sup A = \max A = 14/3$, $\inf A = 3$, and $\min A$ does not exist.

Proof. Problem.

J.2.9. Example. Let $f(x) = -\frac{1}{2} + \sin x$ for $x \in \mathbb{R}$.

(a) If $A = \{f(x) \colon x \in \mathbb{R}\}$, then $\inf A = -\frac{3}{2}$ and $\sup A = \frac{1}{2}$.
(b) If $B = \{|f(x)| \colon x \in \mathbb{R}\}$, then $\inf B = 0$ and $\sup B = \frac{3}{2}$.

Proof. Problem.

J.2.10. Example. Let $f(x) = x^{20} - 2$ for $0 < x < 1$.

(a) If $A = \{f(x) \colon 0 < x < 1\}$, then $\inf A = -2$ and $\sup A = -1$.
(b) If $B = \{|f(x)| \colon 0 < x < 1\}$, then $\inf B = 1$ and $\sup B = 2$.

Proof. Problem.

J.2.11. Example. Let $f(x) = x^{20} - \frac{1}{4}$ for $0 \leq x \leq 1$.

(a) If $A = \{f(x) \colon 0 \leq x \leq 1\}$, then $\inf A = -\frac{1}{4}$ and $\sup A = \frac{3}{4}$.
(b) If $B = \{|f(x)| \colon 0 \leq x \leq 1\}$, then $\inf B = 0$ and $\sup B = \frac{3}{4}$.

Proof. Problem.

J.2.12. Problem. Let $f(x) = -4x^2 - 4x + 3$ for every $x \in \mathbb{R}$, let $A = \{x \in \mathbb{R} \colon f(x) > 0\}$, and let $B = \{f(x) \colon -2 < x < 2\}$.

(a) Find $\sup A$ and $\inf A$ (if they exist).
(b) Find $\sup B$ and $\inf B$ (if they exist).

J.2.13. Problem. For $c > 0$ define a function f on $[0, \infty)$ by $f(x) = x\,e^{-cx}$. Find $\sup\{|f(x)| \colon x \geq 0\}$.

J.2.14. Problem. For each $n = 1, 2, 3, \dots$, define a function f_n on $\mathbb{R}$ by $f_n(x) = \dfrac{x}{1 + nx^2}$. For each $n \in \mathbb{N}$ let $A_n = \{f_n(x) \colon x \in \mathbb{R}\}$. For each n find $\inf A_n$ and $\sup A_n$.

J.3. THE LEAST UPPER BOUND AXIOM FOR $\mathbb{R}$

We now state our last assumption concerning the set $\mathbb{R}$ of real numbers. This is the *least upper bound* (or *order completeness*) *axiom.*

J.3.1. Axiom (VII)**.** *Every nonempty set of real numbers which is bounded above has a least upper bound.*

J.3.2. Notation. If A and B are subsets of $\mathbb{R}$ and $\alpha \in \mathbb{R}$, then

$$\begin{aligned} A + B &:= \{a + b \colon a \in A \text{ and } b \in B\}, \\ AB &:= \{ab \colon a \in A \text{ and } b \in B\}, \\ \alpha B &:= \{\alpha\}B = \{\alpha b \colon b \in B\}, \text{and} \\ -A &:= (-1)A = \{-a \colon a \in A\}. \end{aligned}$$

J.3.3. Proposition. *If A is a nonempty subset of $\mathbb{R}$ which is bounded below, then A has a greatest lower bound. In fact,*

$$\inf A = -\sup(-A)\,.$$

Proof. Let b be a lower bound for A. Then since $b \leq a$ for every $a \in A$, we see that $-b \geq -a$ for every $a \in A$. This says that $-b$ is an upper bound for the set $-A$. By the *least upper bound axiom* (J.3.1) the set $-A$ has a least upper bound, say ℓ. We show that $-\ell$ is the greatest lower bound for A. Certainly it is a lower bound [$\ell \geq -a$ for all $a \in A$ implies $-\ell \leq a$ for all $a \in A$].

Again letting b be an arbitrary lower bound for A, we see, as above, that $-b$ is an upper bound for $-A$. Now $\ell \leq -b$, since ℓ is the least upper bound for $-A$. Thus $-\ell \geq b$. We have shown

$$\inf A = -\ell = -\sup(-A)\,.$$

□

J.3.4. Corollary. *If A is a nonempty set of real numbers which is bounded above, then*

$$\sup A = -\inf(-A)\,.$$

Proof. If A is bounded above, then $-A$ is bounded below. By the preceding proposition $\inf(-A) = -\sup A$. □

J.3.5. Proposition. *Suppose $\emptyset \neq A \subseteq B \subseteq \mathbb{R}$.*

(a) *If B is bounded above, so is A and* $\sup A \leq \sup B$.
(b) *If B is bounded below, so is A and* $\inf A \geq \inf B$.

Proof. Problem.

J.3.6. Proposition. *If A and B are nonempty subsets of $\mathbb{R}$ which are bounded above, then $A + B$ is bounded above and*

$$\sup(A + B) = \sup A + \sup B\,.$$

Proof. Problem. *Hint.* It is easy to show that if ℓ is the least upper bound for A and m is the least upper bound for B, then $\ell + m$ is *an* upper bound for $A + B$.

One way to show that $\ell + m$ is the *least* upper bound for $A + B$, is to argue by contradiction. Suppose there exists an upper bound u for $A + B$ which is strictly less than $\ell + m$. Find numbers a in A and b in B which are close enough to ℓ and m, respectively, so that their sum exceeds u.

An even nicer proof results from taking u to be an arbitrary upper bound for $A+B$ and proving directly that $\ell + m \leq u$. Start by observing that $u - b$ is an upper bound for A for every $b \in B$, and consequently $l \leq u - b$ for every $b \in B$.

J.3.7. Proposition. *If A and B are nonempty subsets of $[0, \infty)$ which are bounded above, then the set AB is bounded above and*

$$\sup(AB) = (\sup A)(\sup B)\,.$$

Proof. Exercise. *Hint.* The result is trivial if $A = \{0\}$ or if $B = \{0\}$. So suppose that both A and B contain elements strictly greater than 0, in which case $\ell := \sup A > 0$ and $m := \sup B > 0$. Show that the set AB is bounded above. (If $x \in AB$, there exist $a \in A$ and $b \in B$ such that $x = ab$.) Then AB has a least upper bound, say c. To show that $\ell m \le c$, assume to the contrary that $c < \ell m$. Let $\epsilon = \ell m - c$. Since ℓ is the least upper bound for A, we may choose $a \in A$ so that $a > \ell - \epsilon(2m)^{-1}$. Having chosen this a, explain how to choose $b \in B$ so that $ab > \ell m - \epsilon$. (Solution Q.35.2.)

J.3.8. Proposition. *If B is a nonempty subset of $[0, \infty)$ which is bounded above and if $\alpha \ge 0$, then αB is bounded above and*

$$\sup(\alpha B) = \alpha \sup B\,.$$

Proof. Problem. *Hint.* This is a *very* easy consequence of one of the previous propositions. □

J.4. THE ARCHIMEDEAN PROPERTY

One interesting property that distinguishes the set of real numbers from many other ordered fields is that for any real number a (no matter how large) and any positive number ϵ (no matter how small) it is possible by adding together enough copies of ϵ to obtain a sum greater than a. This is the *Archimedean property* of the real number system. It is an easy consequence of the order completeness of the reals; that is, it follows from the *least upper bound axiom*(J.3.1).

J.4.1. Proposition (The Archimedean Property of $\mathbb{R}$)**.** *If $a \in \mathbb{R}$ and $\epsilon > 0$, then there exists $n \in \mathbb{N}$ such that $n\epsilon > a$.*

Proof. Problem. *Hint.* Argue by contradiction. Assume that the set $A := \{n\epsilon \colon n \text{ belongs to } \mathbb{N}\}$ is bounded above.

It is worth noting that the preceding proposition shows that the set $\mathbb{N}$ of natural numbers is not bounded above. [Take $\epsilon = 1$.]

Another useful consequence of the *least upper bound axiom* is the existence of n^{th} roots of numbers $a \ge 0$. Below we establish the existence of square roots; but the proof we give can be modified without great difficulty to show that every number $a \ge 0$ has an n^{th} root (see project J.4.5).

J.4.2. Proposition. *Let $a \ge 0$. There exists a unique number $x \ge 0$ such that $x^2 = a$.*

Proof. Exercise. *Hint.* Let $A = \{t > 0 \colon t^2 < a\}$. Show that A is not empty and that it is bounded above. Let $x = \sup A$. Show that assuming $x^2 < a$ leads to a contradiction. [Choose ϵ in $(0,1)$ so that $\epsilon < 3^{-1}x^{-2}(a - x^2)$ and prove that $x(1+\epsilon)$ belongs to A.] Also show that assuming $x^2 > a$ produces a contradiction. [Choose ϵ in $(0,1)$ so that $\epsilon < (3a)^{-1}(x^2 - a)$, and prove that the set $A \cap \big(x(1+\epsilon)^{-1}, x\big)$ is not empty. What can be said about $x(1+\epsilon)^{-1}$?] (Solution Q.35.3.)

J.4.3. Notation. The unique number x guaranteed by the preceding proposition is denoted by $\sqrt{a}$ or by $a^{\frac{1}{2}}$. Similarly, n^{th} roots are denoted by either $\sqrt[n]{a}$ or $a^{\frac{1}{n}}$.

J.4.4. Problem. Prove the following properties of the square root function.

(a) If $x, y \ge 0$, then $\sqrt{xy} = \sqrt{x}\sqrt{y}$.
(b) If $0 < x < y$, then $\sqrt{x} < \sqrt{y}$. *Hint.* Consider $(\sqrt{y})^2 - (\sqrt{x})^2$.
(c) If $0 < x < 1$, then $x^2 < x$ and $x < \sqrt{x}$.
(d) If $x > 1$, then $x < x^2$ and $\sqrt{x} < x$.

J.4.5. Problem. Restate the assertions of the preceding problem for n^{th} roots (and n^{th} powers). Explain what alterations in the proofs must be made to accommodate this change.

J.4.6. Definition. Let $x \in \mathbb{R}$. The ABSOLUTE VALUE of x, denoted by $|x|$, is defined to be $\sqrt{x^2}$. In light of the preceding proposition it is clear that if $x \ge 0$, then $|x| = x$; and if $x < 0$, then $|x| = -x$. From this observation it is easy to deduce two standard procedures for establishing an inequality of the form $|x| < c$ (where $c > 0$). One is to show that $x^2 < c^2$. The other is to show that $-c < x < c$ (or what is the same thing: that both $x < c$ and $-x < c$ hold). Both methods are used extensively throughout the text, especially in chapter 3 when we discuss continuity of real valued functions.

J.4.7. Problem. Prove that if a, $b \in \mathbb{R}$, then

(a) $|ab| = |a|\,|b|$;
(b) $|a + b| \leq |a| + |b|$; and
(c) $\big||a| - |b|\big| \leq |a - b|$.

J.4.8. Problem. Prove that if a, $b \in \mathbb{R}$, then $|ab| \leq \frac{1}{2}(a^2 + b^2)$. *Hint.* Consider the square of $a - b$ and of $a + b$.

APPENDIX K

PRODUCTS, RELATIONS, AND FUNCTIONS

K.1. CARTESIAN PRODUCTS

Ordered pairs are familiar objects. They are used among other things for coordinates of points in the plane. In the first sentence of chapter B it was promised that all subsequent mathematical objects would be defined in terms of sets. So here just for the record is a formal definition of "ordered pair".

K.1.1. Definition. Let x and y be elements of arbitrary sets. Then the ORDERED PAIR (x, y) is defined to be $\{\{x, y\}, \{x\}\}$. This definition reflects our intuitive attitude: an ordered pair is a set $\{x, y\}$ with one of the elements, here x, designated as being "first". Thus we specify two things: $\{x, y\}$ and $\{x\}$.

Ordered pairs have only one interesting property: two of them are equal if and only if both their first coordinates and their second coordinates are equal. As you will discover by proving the next proposition, this fact follows easily from the definition.

K.1.2. Proposition. *Let x, y, u, and v be elements of arbitrary sets. Then $(x, y) = (u, v)$ if and only if $x = u$ and $y = v$.*

PROOF. Exercise. *Hint.* Do not assume that the set $\{x, y\}$ has two elements. If $x = y$, then $\{x, y\}$ has only one element. (Solution Q.36.1.)

K.1.3. Problem. Asked to define an "ordered triple", one might be tempted to try a definition analogous to the definition of ordered pairs: let (a, b, c) be $\{\{a, b, c\}, \{a, b\}, \{a\}\}$. This appears to specify the entries, to pick out a "first" element, and to identify the "first two" elements. Explain why this won't work. (See K.1.8 for a definition that does work.)

K.1.4. Definition. Let S and T be sets. The CARTESIAN PRODUCT of S and T, denoted by $S \times T$, is defined to be $\{(x, y) \colon x \in S \text{ and } y \in T\}$. The set $S \times S$ is often denoted by S^2.

K.1.5. Example. Let $S = \{1, x\}$ and $T = \{x, y, z\}$. Then

$$S \times T = \{(1, x), (1, y), (1, z), (x, x), (x, y), (x, z)\}\,.$$

K.1.6. Problem. Let $S = \{0, 1, 2\}$ and $T = \{1, 2, 3\}$. List all members of $(T \times S) \setminus (S \times T)$.

K.1.7. Problem. Let S, T, U, and V be sets. Then

(a) $(S \times T) \cap (U \times V) = (S \cap U) \times (T \cap V)$;
(b) $(S \times T) \cup (U \times V) \subseteq (S \cup U) \times (T \cup V)$; and
(c) equality need not hold in (b).

The proofs of (a) and (b) in the preceding problem are not particularly difficult. Nonetheless, before one can write down a proof one must have a conjecture as to what is true. How could we have *guessed* initially that equality holds in (a) but not in (b)? The answer is, as it frequently is in mathematics, *by looking at pictures.* Try the following: Make a sketch where S and U are overlapping intervals on the x-axis and T and V are overlapping intervals on the y-axis. Then $S \times T$ and $U \times V$ are overlapping rectangles in the plane. Are not (a) and (b) almost obvious from your sketch?

We will also have occasion to use ordered n-tuples and n-fold Cartesian products for n greater than 2.

K.1.8. Definition. Let $n \geq 3$. We define ORDERED n-TUPLES inductively. Suppose ordered $(n-1)$-tuples $\big(x_1, \dots, x_{n-1}\big)$ have been defined. Let $\big(x_1, \dots, x_n\big) := \big(\big(x_1, \dots, x_{n-1}\big), x_n\big)$. An easy inductive proof shows that $\big(x_1, \dots, x_n\big) = \big(y_1, \dots, y_n\big)$ if and only if $x_k = y_k$ for $k = 1, \dots, n$.

K.1.9. Definition. If $S_1, \dots, S_n$ are sets, we define the CARTESIAN PRODUCT $S_1 \times \dots \times S_n$ to be the set of all ordered n-tuples $\big(x_1, \dots, x_n\big)$ where $x_k \in S_k$ for $1 \leq k \leq n$. We write S^n for $S \times \dots \times S$ (n factors).

K.1.10. Example. The n-fold Cartesian product of the set $\mathbb{R}$ of real numbers is the set $\mathbb{R}^n$ of all n-tuples of real numbers and is often called (Euclidean) n-space.

K.2. RELATIONS

Calculus is primarily about functions. We differentiate functions, we integrate them, we represent them as infinite series. A function is a special kind of relation. So it is convenient before introducing functions to make a few observations concerning the more general concept—relations.

K.2.1. Definition. A RELATION from a set S to a set T is a subset of the Cartesian product $S \times T$. A relation from the set S to itself is often called a relation *on* S or a relation *among members of* S.

There is a notational oddity concerning relations. To indicate that an ordered pair (a, b) belongs to a relation $R \subseteq S \times T$, we almost always write something like aRb rather than $(a, b) \in R$, which we would expect from the definition. For example, the relation "less than" is a relation on the real numbers. (We discussed this relation in appendix H.) Technically then, since $<$ is a subset of $\mathbb{R} \times \mathbb{R}$, we could (correctly) write expressions such as $(3, 7) \in <$. Of course we don't. We write $3 < 7$ instead. And we say, "3 is less than 7", not "the pair $(3, 7)$ belongs to the relation *less than*". This is simply a matter of convention; it has no mathematical or logical content.

K.3. FUNCTIONS

Functions are familiar from beginning calculus. Informally, a function consists of a pair of sets and a "rule" which associates with each member of the first set (the *domain*) one and only one member of the second (the *codomain*). While this informal "definition" is certainly adequate for most purposes and seldom leads to any misunderstanding, it is nevertheless sometimes useful to have a more precise formulation. This is accomplished by defining a function to be a special type of relation between two sets.

K.3.1. Definition. A FUNCTION f is an ordered triple (S, T, G) where S and T are sets and G is a subset of $S \times T$ satisfying:

(1) for each $s \in S$ there is a $t \in T$ such that $(s, t) \in G$, and
(2) if (s, t_1) and (s, t_2) belong to G, then $t_1 = t_2$.

In this situation we say that f is a *function from* S *into* T (or that f *maps* S *into* T) and write $f : S \to T$. The set S is the DOMAIN (or the INPUT SPACE) of f. The set T is the CODOMAIN (or TARGET SPACE, or the OUTPUT SPACE) of f. And the relation G is the GRAPH of f. In order to avoid explicit reference to the graph G it is usual to replace the expression "$(x, y) \in G$" by "$y = f(x)$"; the element $f(x)$ is the IMAGE of x under f. In this text (but not everywhere!) the words "transformation", "map", and "mapping" are synonymous with "function". The domain of f is denoted by $\operatorname{dom} f$.

K.3.2. Example. There are many ways of specifying a function. Statements (1)–(4) below define exactly the same function. We will use these (and other similar) notations interchangeably.

(1) For each real number x we let $f(x) = x^2$.
(2) Let $f = (S, T, G)$ where $S = T = \mathbb{R}$ and $G = \{(x, x^2) : x \in \mathbb{R}\}$.
(3) Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$.
(4) Consider the function $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$.

K.3.3. Notation. If S and T are sets we denote by $\mathcal{F}(S,T)$ the family of all functions from S into T.

K.3.4. Convention. A REAL VALUED function is a function whose codomain lies in $\mathbb{R}$. A function OF A REAL VARIABLE is a function whose domain is contained in $\mathbb{R}$. Some real valued functions of a real variable may be specified simply by writing down a formula. When the domain and codomain are not specified, the understanding is that the domain of the function is the largest set of real numbers for which the formula makes sense and the codomain is taken to be $\mathbb{R}$.

In the case of real valued functions on a set S, we frequently write $\mathcal{F}(S)$ instead of $\mathcal{F}(S,\mathbb{R})$.

K.3.5. Example. Let $f(x) = (x^2 + x)^{-1}$. Since this formula is meaningful for all real numbers except -1 and 0, we conclude that the domain of f is $\mathbb{R} \setminus \{-1, 0\}$.

K.3.6. Example. Let $f(x) = (x^2 + x)^{-1}$ for $x > 0$. Here the domain of f is specified: it is the interval $(0, \infty)$.

K.3.7. Exercise. Let $f(x) = (1 - 2(1 + (1 - x)^{-1})^{-1})^{-1}$.

(a) Find $f(\frac{1}{2})$.
(b) Find the domain of f.

(Solution Q.36.2.)

K.3.8. Exercise. Let $f(x) = (-x^2 - 4x - 1)^{-1/2}$. Find the domain of f. (Solution Q.36.3.)

K.3.9. Problem. Let $f(x) = (1 - (2 + (3 - (1 + x)^{-1})^{-1})^{-1})^{-1}$

(a) Find $f(\frac{1}{2})$.
(b) Find the domain of f.

K.3.10. Problem. Let $f(x) = (-x^2 - 7x - 10)^{-1/2}$.

(a) Find $f(-3)$.
(b) Find the domain of f.

K.3.11. Problem. Let $f(x) = \dfrac{\sqrt{x^2 - 4}}{5 - \sqrt{36 - x^2}}$. Find the domain of f. Express your answer as a union of intervals.

K.3.12. Problem. Explain carefully why two functions f and g are equal if and only if their domains and codomains are equal and $f(x) = g(x)$ for every x in their common domain.

APPENDIX L

PROPERTIES OF FUNCTIONS

L.1. IMAGES AND INVERSE IMAGES

L.1.1. Definition. If $f\colon S \to T$ and $A \subseteq S$, then $f^{\to}(A)$, the IMAGE of A under f, is $\{f(x)\colon x \in A\}$. It is common practice to write $f(A)$ for $f^{\to}(A)$. The set $f^{\to}(S)$ is the RANGE (or IMAGE) of f; usually we write $\operatorname{ran} f$ for $f^{\to}(S)$.

L.1.2. Exercise. Let

$$f(x) = \begin{cases} -1, & \text{for } x < -2 \\ 7 - x^2, & \text{for } -2 \le x < 1 \\ \dfrac{1}{x}, & \text{for } x \ge 1 \end{cases}$$

and $A = (-4, 4)$. Find $f^{\to}(A)$. (Solution Q.37.1.)

L.1.3. Exercise. Let $f(x) = 3x^4 + 4x^3 - 36x^2 + 1$. Find $\operatorname{ran} f$. (Solution Q.37.2.)

L.1.4. Definition. Let $f\colon S \to T$ and $B \subseteq T$. Then $f^{\leftarrow}(B)$, the INVERSE IMAGE of B under f, is $\{x \in S\colon f(x) \in B\}$. In many texts $f^{\leftarrow}(B)$ is denoted by $f^{-1}(B)$. This may cause confusion by suggesting that functions always have inverses (see section M.2 of chapter M).

L.1.5. Exercise. Let $f(x) = \arctan x$ and $B = (\frac{\pi}{4}, 2)$. Find $f^{\leftarrow}(B)$. (Solution Q.37.3.)

L.1.6. Exercise. Let $f(x) = -\sqrt{9 - x^2}$ and $B = (1, 3)$. Find $f^{\leftarrow}(B)$. (Solution Q.37.4.)

L.1.7. Problem. Let

$$f(x) = \begin{cases} -x - 4, & \text{for } x \le 0 \\ x^2 + 3, & \text{for } 0 < x \le 2 \\ (x-1)^{-1}, & \text{for } x > 2 \end{cases}$$

and $A = (-3, 4)$. Find $f^{\to}(A)$.

L.1.8. Problem. Let $f(x) = 4 - x^2$ and $B = (1, 3]$. Find $f^{\leftarrow}(B)$.

L.1.9. Problem. Let $f(x) = \dfrac{x}{1-x}$.

(a) Find $f^{\leftarrow}([0, a])$ for $a > 0$.
(b) Find $f^{\leftarrow}([-\frac{3}{2}, -\frac{1}{2}])$.

L.1.10. Problem. Let $f(x) = -x^2 + 4 \arctan x$. Find $\operatorname{ran} f$.

L.1.11. Problem. Let

$$f(x) = \begin{cases} x + 1, & \text{for } x < 1 \\ 8 + 2x - x^2, & \text{for } x \ge 1. \end{cases}$$

Let $A = (-2, 3)$ and $B = [0, 1]$. Find $f^{\to}(A)$ and $f^{\leftarrow}(B)$.

L.2. COMPOSITION OF FUNCTIONS

Let $f\colon S \to T$ and $g\colon T \to U$. The COMPOSITE of g and f, denoted by $g \circ f$, is the function taking S to U defined by

$$(g \circ f)(x) = g(f(x))$$

for all x in S. The operation $\circ$ is COMPOSITION. We again make a special convention for real valued functions of a real variable: The domain of $g \circ f$ is the set of all x in $\mathbb{R}$ for which the expression $g(f(x))$ makes sense.

L.2.1. Example. Let $f(x) = (x-1)^{-1}$ and $g(x) = \sqrt{x}$. Then the domain of $g \circ f$ is the interval $(1, \infty)$, and for all x in that interval

$$(g \circ f)(x) = g(f(x)) = \frac{1}{\sqrt{x-1}}.$$

PROOF. The square root of $x - 1$ exists only when $x \geq 1$; and since we take its reciprocal, we exclude $x = 1$. Thus $\operatorname{dom}(g \circ f) = (1, \infty)$. □

L.2.2. Exercise. Let

$$f(x) = \begin{cases} 0, & \text{for } x < 0 \\ 3x, & \text{for } 0 \leq x \leq 2 \\ 2, & \text{for } x > 2 \end{cases}$$

and

$$g(x) = \begin{cases} x^2, & \text{for } 1 < x < 3 \\ -1, & \text{otherwise.} \end{cases}$$

Sketch the graph of $g \circ f$. (Solution Q.37.5.)

L.2.3. Proposition. *Composition of functions is associative but not necessarily commutative.*

PROOF. Exercise. *Hint.* Let $f\colon S \to T$, $g\colon T \to U$, and $h\colon U \to V$. Show that $h \circ (g \circ f) = (h \circ g) \circ f$. Give an example to show that $f \circ g$ and $g \circ f$ may fail to be equal. (Solution Q.37.6.)

L.2.4. Problem. Let $f(x) = x^2 + 2x^{-1}$, $g(x) = 2(2x+3)^{-1}$, and $h(x) = \sqrt{2x}$. Find $(h \circ g \circ f)(4)$.

L.2.5. Problem. If $f\colon S \to T$ and $g : T \to U$, then

(a) $(g \circ f)^{\to}(A) = g^{\to}(f^{\to}(A))$ for every $A \subseteq S$.
(b) $(g \circ f)^{\leftarrow}(B) = f^{\leftarrow}(g^{\leftarrow}(B))$ for every $B \subseteq U$.

L.3. The IDENTITY FUNCTION

The family of all functions mapping a set S into a set T is denoted by $\mathcal{F}(S,T)$. One member of $\mathcal{F}(S,S)$ is particularly noteworthy, the IDENTITY FUNCTION on S. It is defined by

$$I_S\colon S \to S\colon x \mapsto x.$$

When the set S is understood from context, we write I for I_S.

The identity function is characterized algebraically by the conditions:

$$\text{if } f\colon R \to S, \text{ then } I_S \circ f = f$$

and

$$\text{if } g\colon S \to T, \text{ then } g \circ I_S = g.$$

L.3.1. Definition. More general than the identity function are the inclusion maps. If $A \subseteq S$, then the INCLUSION MAP taking A into S is defined by

$$\iota_{A,S} : A \to S\colon x \mapsto x.$$

When no confusion is likely to result, we abbreviate $\iota_{A,S}$ to ι. Notice that $\iota_{S,S}$ is just the identity map I_S.

L.4. DIAGRAMS

It is frequently useful to think of functions as arrows in diagrams. For example, the situation $f\colon R \to S$, $h\colon R \to T$, $j\colon T \to U$, $g\colon S \to U$ may be represented by the following diagram.

$$
\begin{tikzcd}
R \arrow[r, "f"] \arrow[d, "h"'] & S \arrow[d, "g"] \\
T \arrow[r, "j"'] \arrow[ur] & U
\end{tikzcd}
$$

The diagram is said to COMMUTE if $j \circ h = g \circ f$.

Diagrams need not be rectangular. For instance,

$$
\begin{tikzcd}
R \arrow[d, "f"'] \arrow[dr, "k"] & \\
S \arrow[r, "g"'] & T
\end{tikzcd}
$$

is a commutative diagram if $k = g \circ f$.

L.4.1. Example. Here is a diagrammatic way of stating the associative law for composition of functions. If the triangles in the diagram

$$
\begin{tikzcd}
R \arrow[r, "j"] \arrow[d, "f"'] & U \arrow[d, "h"] \\
S \arrow[r, "k"'] \arrow[ur, "g"] & T
\end{tikzcd}
$$

commute, then so does the rectangle.

L.5. RESTRICTIONS AND EXTENSIONS

If $f\colon S \to T$ and $A \subseteq S$, then the RESTRICTION of f to A, denoted by $f\big|_A$, is the function $f \circ \iota_{A,S}$. That is, it is the mapping from A into T whose value at each x in A is $f(x)$.

$$
\begin{tikzcd}
S \arrow[dr, "f"] & \\
A \arrow[u, "\iota_{AS}"] \arrow[r, "f|_A"'] & T
\end{tikzcd}
$$

Suppose that $g\colon A \to T$ and $A \subseteq S$. A function $f\colon S \to T$ is an EXTENSION of g to S if $f\big|_A = g$, that is, if the diagram

$$
\begin{tikzcd}
S \arrow[dr, "f"] & \\
A \arrow[u, "\iota_{AS}"] \arrow[r, "g"'] & T
\end{tikzcd}
$$

commutes.

APPENDIX M

FUNCTIONS WHICH HAVE INVERSES

M.1. INJECTIONS, SURJECTIONS, AND BIJECTIONS

A function f is INJECTIVE (or ONE-TO-ONE) if $x = y$ whenever $f(x) = f(y)$. That is, f is injective if no two distinct elements in its domain have the same image. For a real valued function of a real variable this condition may be interpreted graphically: A function is one-to-one if and only if each horizontal line intersects the graph of the function at most once. An injective map is called an INJECTION.

M.1.1. Example. The sine function is not injective (since, for example, $\sin 0 = \sin \pi$).

M.1.2. Example. Let $f(x) = \dfrac{x+2}{3x-5}$. The function f is injective.

PROOF. Exercise. (See Q.38.1.)

M.1.3. Exercise. Find an injective mapping from $\{x \in \mathbb{Q} : x > 0\}$ into $\mathbb{N}$. (Solution Q.38.2.)

M.1.4. Problem. Let $f(x) = \dfrac{2x-5}{3x+4}$. Show that f is injective.

M.1.5. Problem. Let $f(x) = 2x^2 - x - 15$. Show that f is *not* injective.

M.1.6. Problem. Let $f(x) = x^3 - 2x^2 + x - 3$. Is f injective?

M.1.7. Problem. Let $f(x) = x^3 - x^2 + x - 3$. Is f injective?

M.1.8. Definition. A function is SURJECTIVE (or ONTO) if its range is equal to its codomain.

M.1.9. Example. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$ and $g : \mathbb{R} \to [0, \infty)$ by $g(x) = x^2$. Then the function g is surjective while f is not (even though the two functions have the same graph!).

M.1.10. Exercise. Find a surjection (that is, a surjective map) from $[0, 1]$ onto $[0, \infty)$. (Solution Q.38.3.)

M.1.11. Definition. A function is BIJECTIVE if it is both injective and surjective. A bijective map is called a BIJECTION or a ONE-TO-ONE CORRESPONDENCE.

M.1.12. Exercise. Give an explicit formula for a bijection between $\mathbb{Z}$ and $\mathbb{N}$. (Solution Q.38.4.)

M.1.13. Exercise. Give an explicit formula for a bijection between $\mathbb{R}$ and the open interval $(0, 1)$. (Solution Q.38.5.)

M.1.14. Exercise. Give a formula for a bijection between the interval $[0, 1)$ and the unit circle $x^2 + y^2 = 1$ in the plane. (Solution Q.38.6.)

M.1.15. Exercise. Let $f : [1, 2] \to [0, 3]$ be defined by $f(x) = 1/x$. Find an extension g of f to the interval $[0, 3]$ such that $g : [0, 3] \to [0, 3]$ is a bijection. (Solution Q.38.7.)

M.1.16. Exercise. Let $f : R \to [-1, 1]$ be defined by $f(x) = \sin x$. Find a set $A \subseteq R$ such that the restriction of f to A is a bijection from A onto $[-1, 1]$. (Solution Q.38.8.)

M.1.17. Problem. Let $f : S \to T$ and $g : T \to U$. Prove the following.

(a) If f and g are injective, so is $g \circ f$.

(b) If f and g are surjective, so is $g \circ f$.
(c) If f and g are bijective, so is $g \circ f$.

M.1.18. Problem. Find a bijection between the open intervals $(0,1)$ and $(-8,5)$. Prove that the function you use really *is* a bijection between the two intervals.

M.1.19. Problem. Find a bijection between the open intervals $(0,1)$ and $(3,\infty)$. (Proof not required.)

M.1.20. Problem. Find a bijection between the interval $(0,1)$ and the parabola $y = x^2$ in the plane. (Proof not required.)

M.1.21. Problem. Let $f\colon [1,2] \to [0,11]$ be defined by $f(x) = 3x^2 - 1$. Find an extension g of f to $[0,3]$ which is a bijection from $[0,3]$ onto $[0,11]$. (Proof not required.)

It is important for us to know how the image $f^{\to}$ and the inverse image $f^{\leftarrow}$ of a function f behave with respect to unions, intersections, and complements of sets. The basic facts are given in the next 10 propositions. Although these results are quite elementary, we make extensive use of them in studying continuity.

M.1.22. Proposition. *Let $f\colon S \to T$ and $B \subseteq T$.*
(a) $f^{\to}(f^{\leftarrow}(B)) \subseteq B$.
(b) *Equality need not hold in (a).*
(c) *Equality does hold in (a) if f is surjective.*

Proof. Exercise. (Solution Q.38.9.)

M.1.23. Proposition. *Let $f\colon S \to T$ and $A \subseteq S$.*
(a) $A \subseteq f^{\leftarrow}(f^{\to}(A))$.
(b) *Equality need not hold in (a).*
(c) *Equality does hold in (a) if f is injective.*

Proof. Problem.

M.1.24. Problem. Prove the converse of proposition M.1.23. That is, show that if $f\colon S \to T$ and $f^{\leftarrow}(f^{\to}(A)) = A$ for all $A \subseteq S$, then f is injective. *Hint.* Suppose $f(x) = f(y)$. Let $A = \{x\}$. Show that $y \in f^{\leftarrow}(f^{\to}(A))$.

M.1.25. Proposition. *Let $f\colon S \to T$ and $A, B \subseteq S$. Then*
$$f^{\to}(A \cup B) = f^{\to}(A) \cup f^{\to}(B).$$

Proof. Exercise. (Solution Q.38.10.)

M.1.26. Proposition. *Let $f\colon S \to T$ and $C,\, D \subseteq T$. Then*
$$f^{\leftarrow}(C \cup D) = f^{\leftarrow}(C) \cup f^{\leftarrow}(D)\,.$$

Proof. Problem.

M.1.27. Proposition. *Let $f\colon S \to T$ and $C,\, D \subseteq T$. Then*
$$f^{\leftarrow}(C \cap D) = f^{\leftarrow}(C) \cap f^{\leftarrow}(D)\,.$$

Proof. Exercise. (Solution Q.38.11.)

M.1.28. Proposition. *Let $f\colon S \to T$ and $A,\, B \subseteq S$.*
(a) $f^{\to}(A \cap B) \subseteq f^{\to}(A) \cap f^{\to}(B)$.
(b) *Equality need not hold in (a).*
(c) *Equality does hold in (a) if f is injective.*

Proof. Problem.

M.1.29. Proposition. *Let $f\colon S \to T$ and $D \subseteq T$. Then*

$$f^{\leftarrow}(D^c) = \big(f^{\leftarrow}(D)\big)^c .$$

Proof. Problem.

M.1.30. Proposition. *Let $f\colon S \to T$ and $A \subseteq S$.*

(a) *If f is injective, then $f^{\rightarrow}(A^c) \subseteq \big(f^{\rightarrow}(A)\big)^c$.*
(b) *If f is surjective, then $f^{\rightarrow}(A^c) \supseteq \big(f^{\rightarrow}(A)\big)^c$.*
(c) *If f is bijective, then $f^{\rightarrow}(A^c) = \big(f^{\rightarrow}(A)\big)^c$.*

Proof. Problem. *Hints.* For part (a), let $y \in f^{\rightarrow}(A^c)$. To prove that $y \in \big(f^{\rightarrow}(A)\big)^c$, assume to the contrary that $y \in f^{\rightarrow}(A)$ and derive a contradiction. For part (b), let $y \in \big(f^{\rightarrow}(A)\big)^c$. Since f is surjective, there exists $x \in S$ such that $y = f(x)$. Can x belong to A?.

M.1.31. Proposition. *Let $f\colon S \to T$ and $\mathfrak{A} \subseteq \mathfrak{P}(S)$.*

(a) $f^{\rightarrow}(\bigcap \mathfrak{A}) \subseteq \bigcap\{f^{\rightarrow}(A)\colon A \in \mathfrak{A}\}$.
(b) *If f is injective, equality holds in (a).*
(c) $f^{\rightarrow}(\bigcup \mathfrak{A}) = \bigcup\{f^{\rightarrow}(A)\colon A \in \mathfrak{A}\}$.

Proof. Exercise. (Solution Q.38.12.)

M.1.32. Proposition. *Let $f\colon S \to T$ and $\mathfrak{B} \subseteq \mathfrak{P}(T)$.*

(a) $f^{\leftarrow}(\bigcap \mathfrak{B}) = \bigcap\{f^{\leftarrow}(B)\colon B \in \mathfrak{B}\}$.
(b) $f^{\leftarrow}(\bigcup \mathfrak{B}) = \bigcup\{f^{\leftarrow}(B)\colon B \in \mathfrak{B}\}$.

Proof. Problem.

M.2. INVERSE FUNCTIONS

Let $f\colon S \to T$ and $g\colon T \to S$. If $g \circ f = I_S$, then g is a left inverse of f and, equivalently, f is a right inverse of g. We say that f is invertible if there exists a function from T into S which is both a left and a right inverse for f. Such a function is denoted by f^{-1} and is called the inverse of f. (Notice that the last "the" in the preceding sentence requires justification. See proposition M.2.1 below.) A function is invertible if it has an inverse. According to the definition just given, the inverse f^{-1} of a function f must satisfy

$$f \circ f^{-1} = I_T \qquad \text{and} \qquad f^{-1} \circ f = I_S .$$

A simple, but important, consequence of this is that for an invertible function, $y = f(x)$ if and only if $x = f^{-1}(y)$. [Proof: if $y = f(x)$, then $f^{-1}(y) = f^{-1}(f(x)) = I_S(x) = x$. Conversely, if $x = f^{-1}(y)$, then $f(x) = f(f^{-1}(y)) = I_T(y) = y$.]

M.2.1. Proposition. *A function can have at most one inverse.*

Proof. Exercise. (Solution Q.38.13.)

M.2.2. Proposition. *If a function has both a left inverse and a right inverse, then the left and right inverses are equal (and therefore the function is invertible).*

Proof. Problem.

M.2.3. Exercise. The *arcsine* function is defined to be the inverse of what function? (*Hint.* The answer is not *sine.*) What about *arccosine*? *arctangent*? (Solution Q.38.14.)

The next two propositions tell us that a necessary and sufficient condition for a function to have right inverse is that it be surjective and that a necessary and sufficient condition for a function to have a left inverse is that it be injective. Thus, in particular, a function is invertible if and only if it is bijective. In other words, the invertible members of $\mathcal{F}(S,T)$ are the bijections.

M.2.4. Proposition. *Let $S \neq \emptyset$. A function $f\colon S \to T$ has a right inverse if and only if it is surjective.*

PROOF. Exercise. (Solution Q.38.15.)

M.2.5. Proposition. *Let $S \neq \emptyset$. A function $f\colon S \to T$ has a left inverse if and only if it is injective.*

PROOF. Problem.

M.2.6. Problem. Prove: if a function f is bijective, then $\left(f^{-1}\right)^{\leftarrow} = f^{\rightarrow}$.

M.2.7. Problem. Let $f(x) = \dfrac{ax+b}{cx+d}$ where a, b, c, $d \in \mathbb{R}$ and not both c and d are zero.

(a) Under what conditions on the constants $a, b, c,$ and d is f injective?
(b) Under what conditions on the constants $a, b, c,$ and d is f its own inverse?

APPENDIX N

PRODUCTS

The Cartesian product of two sets, which was defined in appendix K, is best thought of not just as a collection of ordered pairs but as this collection together with two distinguished "projection" mappings.

N.1.1. Definition. Let S_1 and S_2 be nonempty sets. For $k = 1, 2$ define the COORDINATE PROJECTIONS $\pi_k \colon S_1 \times S_2 \to S_k$ by $\pi_k(s_1, s_2) = s_k$. We notice two simple facts:

(1) π_1 and π_2 are surjections; and
(2) $z = (\pi_1(z), \pi_2(z))$ for all $z \in S_1 \times S_2$.

If T is a nonempty set and if $g \colon T \to S_1$ and $h \colon T \to S_2$, then we define the function $(g, h) : T \to S_1 \times S_2$ by

$$(g, h)(t) = (g(t), h(t)) .$$

N.1.2. Example. If $g(t) = \cos t$ and $h(t) = \sin t$, then (g, h) is a map from $\mathbb{R}$ to the unit circle in the plane. (This is a *parametrization* of the unit circle.)

N.1.3. Definition. Let S_1, S_2, and T be nonempty sets and let $f \colon T \to S_1 \times S_2$. For $k = 1, 2$ we define functions $f^k \colon T \to S_k$ by $f^k = \pi_k \circ f$; these are the COMPONENTS of f. (The superscripts have nothing to do with powers. We use them because we wish later to attach subscripts to functions to indicate partial differentiation.) Notice that $f(t) = (\pi_1(f(t)), \pi_2(f(t)))$ for all $t \in T$, so that

$$f = (\pi_1 \circ f, \pi_2 \circ f) = (f^1, f^2) .$$

If we are given the function f, the components f^1 and f^2 have been defined so as to make the following diagram commute.

$$\begin{tikzcd}
 & T \arrow[dl, "f^1"'] \arrow[d, "f"] \arrow[dr, "f^2"] & \\
S_1 & S_1 \times S_2 \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & S_2
\end{tikzcd}$$

On the other hand, if the functions $f^1 \colon T \to S_1$ and $f^2 \colon T \to S_2$ are given, then there exists a function f, namely (f^1, f^2), which makes the diagram commute. Actually, (f^1, f^2) is the *only* function with this property, a fact which we prove in the next exercise.

N.1.4. Exercise. Suppose that $f^1 \in \mathcal{F}(T, S_1)$ and $f^2 \in \mathcal{F}(T, S_2)$. Then there exists a unique function $g \in \mathcal{F}(T, S_1 \times S_2)$ such that $\pi_1 \circ g = f^1$ and $\pi_2 \circ g = f^2$. (Solution Q.39.1.)

The following problem, although interesting, is not needed elsewhere in this text, so it is listed as optional. It says, roughly, that any set which "behaves like a Cartesian product" must be in one-to-one correspondence with the Cartesian product.

N.1.5. Problem (optional)**.** Let S_1, S_2, and P be nonempty sets and $\rho_k \colon P \to S_k$ be surjections. Suppose that for every set T and every pair of functions $f^k \in \mathcal{F}(T, S_k)$ $(k = 1, 2)$, there exists a unique function $g \in \mathcal{F}(T, P)$ such that $f^k = \rho_k \circ g$ $(k = 1, 2)$. Then there exists a bijection from P

onto $S_1 \times S_2$. *Hint.* Consider the following diagrams.

$$\begin{tikzcd} & P \arrow[dl, "\rho_1"'] \arrow[d, "\rho"] \arrow[dr, "\rho_2"] & \\ S_1 & S_1 \times S_2 \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & S_2 \end{tikzcd} \tag{N.1}$$

$$\begin{tikzcd} & S_1 \times S_2 \arrow[dl, "\pi_1"'] \arrow[d, "\pi"] \arrow[dr, "\pi_2"] & \\ S_1 & P \arrow[l, "\rho_1"] \arrow[r, "\rho_2"'] & S_2 \end{tikzcd} \tag{N.2}$$

$$\begin{tikzcd} & P \arrow[dl, "\rho_1"'] \arrow[d, "\rho"] \arrow[dr, "\rho_2"] & \\ S_1 & P \arrow[l, "\rho_1"] \arrow[r, "\rho_2"'] & S_2 \end{tikzcd} \tag{N.3}$$

Exercise N.1.4 tells us that there exists a unique map ρ which makes diagram (N.1) commute, and by hypothesis there exists a unique map π which makes diagram (N.2) commute. Conclude from (N.1) and (N.2) that (N.3) commutes when $g = \pi \circ \rho$. It is obvious that (N.3) commutes when $g = I_P$. Then use the uniqueness part of the hypothesis to conclude that π is a left inverse for ρ. Now construct a new diagram replacing P by $S_1 \times S_2$ and ρ_k by π_k in (N.3).

APPENDIX O

FINITE AND INFINITE SETS

There are a number of ways of comparing the "sizes" of sets. In this chapter and the next we examine perhaps the simplest of these, cardinality. Roughly speaking, we say that two sets have the "same number of elements" if there is a one-to-one correspondence between the elements of the sets. In this sense the open intervals $(0,1)$ and $(0,2)$ have the same number of elements. (The map $x \mapsto 2x$ is a bijection.) Clearly this is only one sense of the idea of "size". It is certainly also reasonable to regard $(0,2)$ as being bigger than $(0,1)$ because it is twice as long.

We derive only the most basic facts concerning cardinality. In this appendix we discuss some elementary properties of finite and infinite sets, and in the next we distinguish between countable and uncountable sets. This is all we will need.

O.1.1. Definition. Two sets S and T are CARDINALLY EQUIVALENT if there exists a bijection from S onto T, in which case we write $S \sim T$. It is easy to see that cardinal equivalence is indeed an equivalence relation; that is, it is reflexive, symmetric, and transitive.

O.1.2. Proposition. *Cardinal equivalence is an equivalence relation. Let S, T, and U be sets. Then*

(a) $S \sim S$;
(b) *if* $S \sim T$, *then* $T \sim S$; *and*
(c) *if* $S \sim T$ *and* $T \sim U$, *then* $S \sim U$.

PROOF. Problem.

O.1.3. Definition. A set S is FINITE if it is empty or if there exists $n \in \mathbb{N}$ such that $S \sim \{1, \dots, n\}$. A set is INFINITE if it is not finite. The next few facts concerning finite and infinite sets probably appear obvious, but writing down the proofs may in several instances require a bit of thought. Here is a question which merits some reflection: if one is unable to explain exactly why a result is obvious, is it really obvious?

One "obvious" result is that if two initial segments $\{1, \dots, m\}$ and $\{1, \dots, n\}$ of the set of natural numbers are cardinally equivalent, then $m = n$. We prove this next.

O.1.4. Proposition. *Let $n \in \mathbb{N}$. If $m \in \mathbb{N}$ and $\{1, \dots, m\} \sim \{1, \dots, n\}$, then $m = n$.*

PROOF. Exercise. *Hint.* Use induction on n. (Solution Q.40.1.)

O.1.5. Definition. We define $\operatorname{card} \emptyset$, the cardinal number of the empty set, to be 0. If S is a nonempty finite set, then by the preceding proposition there exists only one positive integer n such that $S \sim \{1, \dots, n\}$. This integer, $\operatorname{card} S$, is the CARDINAL NUMBER of the set S, or the NUMBER OF ELEMENTS in S. Notice that if S is finite with cardinal number n and $T \sim S$, then T is also finite and has cardinal number n. Thus for finite sets the expressions "cardinally equivalent" and "have the same cardinal number" are interchangeable.

O.1.6. Example. Let $S = \{a, b, c, d\}$. Then $\operatorname{card} S = 4$, since $\{a, b, c, d\} \sim \{1, 2, 3, 4\}$.

One simple fact about cardinal numbers is that the number of elements in the union of two disjoint finite sets is the sum of the numbers of elements in each.

O.1.7. Proposition. *If S and T are disjoint finite sets, then $S \cup T$ is finite and*

$$\operatorname{card}(S \cup T) = \operatorname{card} S + \operatorname{card} T\,.$$

Proof. Exercise. (Solution Q.40.2.)

A variant of the preceding result is given in problem O.1.10. We now take a preliminary step toward proving that subsets of finite sets are themselves finite (proposition O.1.9).

O.1.8. Lemma. *If $C \subseteq \{1, \dots, n\}$, then C is finite and* $\operatorname{card} C \leq n$.

Proof. Exercise. *Hint.* Use mathematical induction. If $C \subseteq \{1, \dots, k+1\}$, then $C \setminus \{k+1\} \subseteq \{1, \dots, k\}$. Examine the cases $k+1 \notin C$ and $k+1 \in C$ separately. (Solution Q.40.3.)

O.1.9. Proposition. *Let $S \subseteq T$. If T is finite, then S is finite and* $\operatorname{card} S \leq \operatorname{card} T$.

Proof. Problem. *Hint.* The case $T = \emptyset$ is trivial. Suppose $T \neq \emptyset$. Let $\iota : S \to T$ be the inclusion map of S into T (see chapter L). There exist $n \in \mathbb{N}$ and a bijection $f : T \to \{1, \dots, n\}$. Let $C = \operatorname{ran}(f \circ \iota)$. The map from S to C defined by $x \mapsto f(x)$ is a bijection. Use lemma O.1.8.)

The preceding proposition "subsets of finite sets are finite" has a useful contrapositive: "sets which contain infinite sets are themselves infinite."

O.1.10. Problem. Let S be a set and T be a finite set. Prove that

$$\operatorname{card}(T \setminus S) = \operatorname{card} T - \operatorname{card}(T \cap S)\,.$$

Hint. Use problem F.3.14 and proposition O.1.7.

Notice that it is a consequence of the preceding result O.1.10 that if $S \subseteq T$ (where T is finite), then

$$\operatorname{card}(T \setminus S) = \operatorname{card} T - \operatorname{card} S.$$

How do we show that a set S is infinite? If our only tool were the definition, we would face the prospect of proving that there does *not* exist a bijection from S onto an initial segment of the natural numbers. It would be pleasant to have a more direct approach than establishing the nonexistence of maps. This is the point of our next proposition.

O.1.11. Proposition. *A set is infinite if and only if it is cardinally equivalent to a proper subset of itself.*

Proof. Exercise. *Hint.* Suppose a set S is infinite. Show that it is possible to choose inductively a sequence of distinct elements $a_1, a_2, a_3, \dots$ in S. (Suppose $a_1, \dots, a_n$ have already been chosen. Can $S \setminus \{a_1, \dots, a_n\}$ be empty?) Map each a_k to a_{k+1} and map each member of S which is not an a_k to itself.

For the converse argue by contradiction. Suppose that $T \sim S$ where T is a proper subset of S, and assume that $S \sim \{1, \dots, n\}$. Prove that $S \setminus T \sim \{1, \dots, p\}$ for some $p \in \mathbb{N}$. Write T as $S \setminus (S \setminus T)$ and obtain $n = n - p$ by computing the cardinality of T in two ways. (Make use of problem O.1.10.) What does $n = n - p$ contradict? (Solution Q.40.4.)

O.1.12. Example. The set $\mathbb{N}$ of natural numbers is infinite.

Proof. The map $n \mapsto n + 1$ is a bijection from $\mathbb{N}$ onto $\mathbb{N} \setminus \{1\}$, which is a proper subset of $\mathbb{N}$. □

O.1.13. Exercise. The interval $(0, 1)$ is infinite. (Solution Q.40.5.)

O.1.14. Example. The set $\mathbb{R}$ of real numbers is infinite.

The next two results tell us that functions take finite sets to finite sets and that for injective functions finite sets come from finite sets.

O.1.15. Proposition. *If T is a set, S is a finite set, and $f : S \to T$ is surjective, then T is finite.*

Proof. Exercise. *Hint.* Use propositions M.2.4 and M.2.5. (Solution Q.40.6.)

O.1.16. Proposition. *If S is a set, T is a finite set, and $f : S \to T$ is injective, then S is finite.*

Proof. Exercise. (Solution Q.40.7.)

O.1.17. Problem. Let S and T be finite sets. Prove that $S \cup T$ is finite and that

$$\operatorname{card}(S \cup T) = \operatorname{card} S + \operatorname{card} T - \operatorname{card}(S \cap T)\,.$$

Hint. Use problem O.1.10.

O.1.18. Problem. If S is a finite set with cardinal number n, what is the cardinal number of $\mathfrak{P}(S)$?

APPENDIX P

COUNTABLE AND UNCOUNTABLE SETS

There are many sizes of infinite sets—infinitely many in fact. In our subsequent work we need only distinguish between countably infinite and uncountable sets. A set is countably infinite if it is in one-to-one correspondence with the set of positive integers; if it is neither finite nor countably infinite, it is uncountable. In this section we present some basic facts about and examples of both countable and uncountable sets. This is all we will need. Except for problem P.1.21, which is presented for general interest, we ignore the many intriguing questions which arise concerning various sizes of uncountable sets. For a very readable introduction to such matters see [7], chapter 2.

P.1.1. Definition. A set is COUNTABLY INFINITE (or DENUMERABLE) if it is cardinally equivalent to the set $\mathbb{N}$ of natural numbers. A bijection from $\mathbb{N}$ onto a countably infinite set S is an ENUMERATION of the elements of S. A set is COUNTABLE if it is either finite or countably infinite. If a set is not countable it is UNCOUNTABLE.

P.1.2. Example. The set $\mathbb{E}$ of even integers in $\mathbb{N}$ is countable.

PROOF. The map $n \mapsto 2n$ is a bijection from $\mathbb{N}$ onto $\mathbb{E}$. $\square$

The first proposition of this section establishes the fact that the "smallest" infinite sets are the countable ones.

P.1.3. Proposition. *Every infinite set contains a countably infinite subset.*

PROOF. Problem. *Hint.* Review the proof of proposition O.1.11.

If we are given a set S which we believe to be countable, it may be extremely difficult to prove this by exhibiting an explicit bijection between $\mathbb{N}$ and S. Thus it is of great value to know that certain constructions performed with countable sets result in countable sets. The next five propositions provide us with ways of generating new countable sets from old ones. In particular, we show that each of the following is countable.

(1) Any subset of a countable set.
(2) The range of a surjection with countable domain.
(3) The domain of an injection with countable codomain.
(4) The product of any finite collection of countable sets.
(5) The union of a countable family of countable sets.

P.1.4. Proposition. *If $S \subseteq T$ where T is countable, then S is countable.*

PROOF. Exercise. *Hint.* Show first that every subset of $\mathbb{N}$ is countable. (Solution Q.41.1.)

The preceding has an obvious corollary: if $S \subseteq T$ and S is uncountable, then so is T.

P.1.5. Proposition. *If $f\colon S \to T$ is injective and T is countable, then S is countable.*

PROOF. Problem. *Hint.* Adapt the proof of proposition O.1.16.

P.1.6. Proposition. *If $f\colon S \to T$ is surjective and S is countable, then T is countable.*

PROOF. Problem. *Hint.* Adapt the proof of proposition O.1.15.

P.1.7. Lemma. *The set $\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. Exercise. *Hint.* Consider the map $(m, n) \mapsto 2^{m-1}(2n - 1)$. (Solution Q.41.2.)

P.1.8. Example. The set $\mathbb{Q}^+ \setminus \{0\} = \{x \in \mathbb{Q} : x > 0\}$ is countable.

Proof. Suppose that the rational number m/n is written in lowest terms. (That is, m and n have no common factors greater than 1.) Define $f(m/n) = (m, n)$. It is easy to see that the map $f : \mathbb{Q}^+ \setminus \{0\} \to \mathbb{N} \times \mathbb{N}$ is injective. By proposition P.1.5 and the preceding lemma, Q^+ is countable. □

P.1.9. Proposition. *If S and T are countable sets, then so is $S \times T$.*

Proof. Problem. *Hint.* Either $S \sim \{1, \dots, n\}$ or else $S \sim \mathbb{N}$. In either case there exists an injective map $f : S \to \mathbb{N}$. Similarly there exists an injection $g : T \to \mathbb{N}$. Define the function $f \times g : S \times T \to \mathbb{N} \times \mathbb{N}$ by $(f \times g)(x, y) = (f(x), g(y))$.

P.1.10. Corollary. *If $S_1, \dots, S_n$ are countable sets, then $S_1 \times \cdots \times S_n$ is countable.*

Proof. Proposition P.1.9 and induction. □

Finally we show that a countable union of countable sets is countable.

P.1.11. Proposition. *Suppose that $\mathfrak{A}$ is a countable family of sets and that each member of $\mathfrak{A}$ is itself countable. Then $\bigcup \mathfrak{A}$ is countable.*

Proof. Exercise. *Hint.* Use lemma P.1.7 and proposition P.1.6. (Solution Q.41.3.)

P.1.12. Example. The set $\mathbb{Q}$ of rational numbers is countable.

Proof. Let $A = \mathbb{Q}^+ \setminus \{0\}$ and $B = -A = \{x \in \mathbb{Q} : x < 0\}$. Then $\mathbb{Q} = A \cup B \cup \{0\}$. The set A is countable by example P.1.8. Clearly $A \sim B$ (the map $x \mapsto -x$ is a bijection); so B is countable. Since Q is the union of three countable sets, it is itself countable by the preceding proposition. □

By virtue of P.1.4–P.1.11 we have a plentiful supply of countable sets. We now look at an important example of a set which is not countable.

P.1.13. Example. The set $\mathbb{R}$ of real numbers is uncountable.

Proof. We take it to be known that if we exclude decimal expansions which end in an infinite string of 9's, then every real number has a unique decimal expansion. (For an excellent and thorough discussion of this matter see Stromberg's beautiful text on *classical real analysis* [11], especially Theorem 2.57.) By (the corollary to) proposition P.1.4 it will suffice to show that the open unit interval $(0, 1)$ is uncountable. Argue by contradiction: assume that $(0, 1)$ is countably infinite. (We know, of course, from exercise O.1.13 that it is not finite.) Let $r_1, r_2, r_3, \dots$ be an enumeration of $(0, 1)$. For each $j \in \mathbb{N}$ the number r_j has a unique decimal expansion

$$0.r_{j1}\, r_{j2}\, r_{j3}\, \dots .$$

Construct another number $x = 0.x_1\, x_2\, x_3\, \dots$ as follows. For each k choose $x_k = 1$ if $r_{kk} \neq 1$ and $x_k = 2$ if $r_{kk} = 1$. Then x is a real number between 0 and 1, and it cannot be any of the numbers r_k in our enumeration (since it differs from r_k at the k^{th} decimal place). But this contradicts the assertion that $r_1, r_2, r_3, \dots$ is an enumeration of $(0, 1)$. □

P.1.14. Problem. Prove that the set of irrational numbers is uncountable.

P.1.15. Problem. Show that if S is countable and T is uncountable, then $T \setminus S \sim T$.

P.1.16. Problem. Let ϵ be an arbitrary number greater than zero. Show that the rationals in $[0, 1]$ can be covered by a countable family of open intervals the sum of whose lengths is no greater than ϵ. (Recall that a family $\mathfrak{U}$ of sets is said to cover a set A if $A \subseteq \bigcup \mathfrak{U}$.) Is it possible to cover the set $\mathbb{Q}$ of all rationals in $\mathbb{R}$ by such a family? *Hint.* $\sum_{k=1}^{\infty} 2^{-k} = 1$.

P.1.17. Problem. (Definition: The OPEN DISK in $\mathbb{R}^2$ with radius $r > 0$ and center (p, q) is defined to be the set of all points (x, y) in $\mathbb{R}^2$ such that $(x - p)^2 + (y - q)^2 < r^2$.) Prove that the family of all open disks in the plane whose centers have rational coordinates and whose radii are rational is countable.

P.1.18. Problem. (Definition: A real number is ALGEBRAIC if it is a root of some polynomial of degree greater than 0 with integer coefficients. A real number which is not algebraic is TRANSCENDENTAL. It can be shown that the numbers π and e, for example, are transcendental.) Show that the set of all transcendental numbers in $\mathbb{R}$ is uncountable. *Hint.* Start by showing that the set of polynomials with integer coefficients is countable.

P.1.19. Problem. Prove that the set of all sequences whose terms consist of only 0's and 1's is uncountable. *Hint.* Something like the argument in example P.1.13 works.

P.1.20. Problem. Let $\mathfrak{J}$ be a disjoint family of intervals in $\mathbb{R}$ each with length greater than 0. Show that $\mathfrak{J}$ is countable.

P.1.21. Problem. Find an uncountable set which is not cardinally equivalent to $\mathbb{R}$. *Hint.* Let $\mathcal{F} = \mathcal{F}(\mathbb{R}, \mathbb{R})$. Assume there exists a bijection $\phi \colon \mathbb{R} \to \mathcal{F}$. What about the function f defined by

$$f(x) = 1 + \big(\phi(x)\big)(x)$$

for all $x \in \mathbb{R}$?

APPENDIX Q

SOLUTIONS TO EXERCISES

Q.1. Exercises in chapter 01

Q.1.1. (Solution to 1.1.3) Find those numbers x such that $d(x,-2) \leq 5$. In other words, solve the inequality

$$|x+2| = |x-(-2)| \leq 5.$$

This may be rewritten as

$$-5 \leq x+2 \leq 5,$$

which is the same as

$$-7 \leq x \leq 3.$$

Thus the points in the closed interval $[-7,3]$ are those that lie within 5 units of -2.

Q.1.2. (Solution to 1.1.12) If $x \in J_\delta(a)$, then $|x-a| < \delta \leq \epsilon$. Thus x lies within ϵ units of a; that is, $x \in J_\epsilon(a)$.

Q.1.3. (Solution to 1.2.6) Factor the left side of the inequality $x^2 - x - 6 \geq 0$. This yields $(x+2)(x-3) \geq 0$. This inequality holds for those x satisfying $x \leq -2$ and for those satisfying $x \geq 3$. Thus $A = (-\infty,-2] \cup [3,\infty)$. No neighborhood of -2 or of 3 lies in A. Thus $A^\circ = (-\infty,-2) \cup (3,\infty)$ and $A^\circ \neq A$.

Q.1.4. (Solution to 1.2.11) Since $A \cap B \subseteq A$ we have $(A\cap B)^\circ \subseteq A^\circ$ by 1.2.9. Similarly, $(A\cap B)^\circ \subseteq B^\circ$. Thus $(A\cap B)^\circ \subseteq A^\circ \cap B^\circ$. To obtain the reverse inclusion take $x \in A^\circ \cap B^\circ$. Then there exist $\epsilon_1, \epsilon_2 > 0$ such that $J_{\epsilon_1}(x) \subseteq A$ and $J_{\epsilon_2}(x) \subseteq B$. Then $J_\epsilon(x) \subseteq A \cap B$ where $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. This shows that $x \in (A\cap B)^\circ$.

Q.1.5. (Solution to 1.2.12) Since $A \subseteq \bigcup \mathfrak{A}$ for every $A \in \mathfrak{A}$, we can conclude from proposition 1.2.9 that $A^\circ \subseteq \left(\bigcup \mathfrak{A}\right)^\circ$ for every $A \in \mathfrak{A}$. Thus $\bigcup\{A^\circ : A \in \mathfrak{A}\} \subseteq \left(\bigcup \mathfrak{A}\right)^\circ$. (See F.1.8.)

Q.2. Exercises in chapter 02

Q.2.1. (Solution to 2.1.7) (a) Let $\mathfrak{S}$ be a family of open subsets of $\mathbb{R}$. By proposition 2.1.4 each nonempty member of $\mathfrak{S}$ is a union of bounded open intervals. But then $\bigcup \mathfrak{S}$ is itself a union of bounded open intervals. So $\bigcup \mathfrak{S}$ is open.

(b) We show that if S_1 and S_2 are open subsets of $\mathbb{R}$, then $S_1 \cap S_2$ is open. From this it follows easily by mathematical induction that if $S_1, \dots, S_n$ are all open in $\mathbb{R}$, then $S_1 \cap \dots \cap S_n$ is open. Let us suppose then that S_1 and S_2 are open subsets of $\mathbb{R}$. If $S_1 \cap S_2 = \emptyset$, there is nothing to prove; so we assume that $S_1 \cap S_2 \neq \emptyset$. Let x be an arbitrary point in $S_1 \cap S_2$. Since $x \in S_1 = {S_1}^\circ$, there exists $\epsilon_1 > 0$ such that $J_{\epsilon_1}(x) \subseteq S_1$. Similarly, there exists $\epsilon_2 > 0$ such that $J_{\epsilon_2}(x) \subseteq S_2$. Let ϵ be the smaller of ϵ_1 and ϵ_2. Then clearly $J_\epsilon(x) \subseteq S_1 \cap S_2$, which shows that x is an interior point of $S_1 \cap S_2$. Since every point of the set $S_1 \cap S_2$ is an interior point of the set, $S_1 \cap S_2$ is open.

Q.2.2. (Solution to 2.2.5) Since no point in $\mathbb{R}$ has an ϵ-neighborhood consisting entirely of rational numbers, $A^\circ = \emptyset$. If $x \geq 0$, then every ϵ-neighborhood of x contains infinitely many positive rational numbers. Thus each such x belongs to A'. If $x < 0$, it is possible to find an ϵ-neighborhood of x which contains no positive rational number. Thus $A' = [0,\infty)$ and $\overline{A} = A \cup A' = [0,\infty)$.

Q.2.3. (Solution to 2.2.9) (a) First we show that $\overline{A^c} \subseteq A^{\circ c}$. If $x \in \overline{A^c}$, then either $x \in A^c$ or x is an accumulation point of A^c. If $x \in A^c$, then x is certainly not in the interior of A; that is, $x \in A^{\circ c}$. On the other hand, if x is an accumulation point of A^c, then every ϵ-neighborhood of x contains points of A^c. This means that no ϵ-neighborhood of x lies entirely in A. So, in this case too, $x \in A^{\circ c}$.

For the reverse inclusion suppose $x \in A^{\circ c}$. Since x is not in the interior of A, no ϵ-neighborhood of x lies entirely in A. Thus either x itself fails to be in A, in which case x belongs to A^c and therefore to $\overline{A^c}$, or else every ϵ-neighborhood of x contains a point of A^c different from x. In this latter case also, x belongs to the closure of A^c.

Remark. It is interesting to observe that the proof of (a) can be accomplished by a single string of "iff" statements. That is, each step of the argument uses a reversible implication. One needs to be careful, however, with the negation of quantifiers (see section D.4 of appendix D). It will be convenient to let $J^*_\epsilon(x)$ denote the ϵ-neighborhood of x with x deleted; that is, $J^*_\epsilon(x) = (x-\epsilon, x) \cup (x, x+\epsilon)$. The proof goes like this:

$$\begin{aligned}
x \in A^{\circ c} &\text{ iff } \sim (x \in A^\circ) \\
&\text{ iff } \sim ((\exists \epsilon > 0)\, J_\epsilon(x) \subseteq A) \\
&\text{ iff } (\forall \epsilon > 0)\, J_\epsilon(x) \nsubseteq A \\
&\text{ iff } (\forall \epsilon > 0)\, J_\epsilon(x) \cap A^c \neq \emptyset \\
&\text{ iff } (\forall \epsilon > 0)\, (x \in A^c \text{ or } J^*_\epsilon(x) \cap A^c \neq \emptyset) \\
&\text{ iff } (x \in A^c \text{ or } (\forall \epsilon > 0)\, J^*_\epsilon(x) \cap A^c \neq \emptyset) \\
&\text{ iff } (x \in A^c \text{ or } x \in (A^c)') \\
&\text{ iff } x \in \overline{A^c}.
\end{aligned}$$

Proofs of this sort are not universally loved. Some people admire their precision and efficiency. Others feel that reading such a proof has all the charm of reading computer code.

(b) The easiest proof of (b) is produced by substituting A^c for A in part (a). Then $\overline{A} = \overline{A^{cc}} = A^{c\circ c}$. Take complements to get $\overline{A}^c = A^{c\circ}$.

Q.3. Exercises in chapter 03

Q.3.1. (Solution to 3.2.4) Let $a \in \mathbb{R}$. Given $\epsilon > 0$, choose $\delta = \epsilon/5$. If $|x-a| < \delta$, then $|f(x) - f(a)| = 5|x-a| < 5\delta = \epsilon$.

Q.3.2. (Solution to 3.2.5)Given $\epsilon > 0$, choose $\delta = \min\{1, \epsilon/7\}$, the smaller of the numbers 1 and $\epsilon/7$. If $|x-a| = |x+1| < \delta$, then $|x| = |x+1-1| \le |x+1| + 1 < \delta + 1 \le 2$. Therefore,

$$\begin{aligned}
|f(x) - f(a)| &= |x^3 - (-1)^3| && \text{(i)} \\
&= |x^3 + 1| && \text{(ii)} \\
&= |x+1|\,|x^2 - x + 1| && \text{(iii)} \\
&\le |x+1|(x^2 + |x| + 1) && \text{(iv)} \\
&\le |x+1|\,(4+2+1) && \text{(v)} \\
&= 7|x+1| && \text{(vi)} \\
&< 7\delta && \text{(vii)} \\
&\le \epsilon. && \text{(viii)}
\end{aligned}$$

Remark. How did we know to choose $\delta = \min\{1, \epsilon/7\}$? As scratch work we do steps (i)–(iv), a purely algebraic process, and obtain

$$|f(x) - f(a)| \le |x+1|(x^2 + |x| + 1)\,.$$

Now how do we guarantee that the quantity in parentheses doesn't get "too large". The answer is to require that x be "close to" $a = -1$. What do we mean by close? Almost *anything* will work. Here it was decided, arbitrarily, that x should be no more than 1 unit from -1. In other words, we wish δ to be no larger than 1. Then $|x| \leq 2$ and consequently $x^2 + |x| + 1 \leq 7$; so we arrive at step (vi)

$$|f(x) - f(a)| \leq 7|x+1|\,.$$

Since we assume that $|x - (-1)| = |x+1| < \delta$, we have (vii)

$$|f(x) - f(a)| < 7\delta\,.$$

What we *want* is $|f(x) - f(a)| < \epsilon$. This can be achieved by choosing δ to be no greater than $\epsilon/7$. Notice that we have required two things of δ:

$$\delta \leq 1 \quad \text{ and } \quad \delta \leq \epsilon/7\,.$$

The easiest way to arrange that δ be no larger than each of two numbers is to make it the smaller of the two. Thus our choice is $\delta = \min\{1, \epsilon/7\}$.

A good exercise is to repeat the preceding argument, but at (iv) require x to be within 2 units of -1 (rather than 1 unit as above). This will change some things in the proof, but should not create any difficulty.

Q.3.3. (Solution to 3.2.6) Let $a \in \mathbb{R}$. Given $\epsilon > 0$, choose $\delta = \min\{1, (4|a|+2)^{-1}\epsilon\}$. If $|x - a| < \delta$, then

$$|x| \leq |x-a| + |a| < 1 + |a|.$$

Therefore

$$\begin{aligned}
|f(x) - f(a)| &= |(2x^2 - 5) - (2a^2 - 5)| \\
&= 2|x^2 - a^2| \\
&= 2|x-a|\,|x+a| \\
&\leq 2|x-a|\,(|x| + |a|) \\
&\leq 2|x-a|\,(1 + |a| + |a|) \\
&\leq (4|a|+2)|x-a| \\
&< (4|a|+2)\delta \\
&\leq \epsilon.
\end{aligned}$$

Q.3.4. (Solution to 3.2.12) Suppose f is continuous. Let V be an open subset of $\mathbb{R}$. To show that $f^{\leftarrow}(V)$ is open it suffices to prove that each point of $f^{\leftarrow}(V)$ is an interior point of that set. (Notice that if $f^{\leftarrow}(V)$ is empty, then there is nothing to prove. The null set is open.) If $a \in f^{\leftarrow}(V)$, then V is a neighborhood of $f(a)$. Since f is continuous at a, the set $f^{\leftarrow}(V)$ contains a neighborhood of a, from which we infer that a is an interior point of $f^{\leftarrow}(V)$.

Conversely, suppose that $f^{\leftarrow}(V) \overset{\circ}{\subseteq} \mathbb{R}$ whenever $V \overset{\circ}{\subseteq} \mathbb{R}$. To see that f is continuous at an arbitrary point a in $\mathbb{R}$, notice that if V is a neighborhood of $f(a)$, then $a \in f^{\leftarrow}(V) \overset{\circ}{\subseteq} \mathbb{R}$. That is, $f^{\leftarrow}(V)$ is a neighborhood of a. So f is continuous at a.

Q.4. Exercises in chapter 04

Q.4.1. (Solution to 4.1.11) Let $\epsilon > 0$. Notice that $x_n \in J_\epsilon(0)$ if and only if $|x_n| \in J_\epsilon(0)$. Thus (x_n) is eventually in $J_\epsilon(0)$ if and only if $(|x_n|)$ is eventually in $J_\epsilon(0)$.

Q.4.2. (Solution to 4.3.7) Let $x_n \to a$ where (x_n) is a sequence of real numbers. Then by problem 4.1.7(a) there exists $n_0 \in \mathbb{N}$ such that $n \geq n_0$ implies $|x_n - a| < 1$. Then by J.4.7(c)

$$\big||x_n| - |a|\big| \leq |x_n - a| < 1$$

for all $n \geq n_0$. Thus, in particular,

$$|x_n| < |a| + 1$$

for all $n \geq n_0$. If $M = \max\{|x_1|, |x_2|, \dots, |x_{n_0-1}|, |a| + 1\}$, then $|x_n| \leq M$ for all $n \in \mathbb{N}$.

Q.4.3. (Solution to 4.4.3) Let (a_n) be a sequence in $\mathbb{R}$. As suggested in the hint, consider two cases. First, suppose that there is a subsequence (a_{n_k}) consisting of peak terms. Then for each k,

$$a_{n_k} \geq a_{n_{k+1}} .$$

That is, the subsequence (a_{n_k}) is decreasing.

Now consider the second possibility: there exists a term a_p beyond which there are no peak terms. Let $n_1 = p + 1$. Since a_{n_1} is not a peak term, there exists $n_2 > n_1$ such that $a_{n_2} > a_{n_1}$. Since a_{n_2} is not a peak term, there exists $n_3 > n_2$ such that $a_{n_3} > a_{n_2}$. Proceeding in this way we choose an increasing (in fact, strictly increasing) subsequence (a_{n_k}) of the sequence (a_n).

In both of the preceding cases we have found a monotone subsequence of the original sequence.

Q.4.4. (Solution to 4.4.11) Suppose that $b \in \overline{A}$. There are two possibilities; $b \in A$ or $b \in A'$. If $b \in A$, then the constant sequence $(b, b, b, \dots)$ is a sequence in A which converges to b. On the other hand, if $b \in A'$, then for every $n \in \mathbb{N}$ there is a point $a_n \in J_{1/n}(b)$ such that $a_n \in A$ and $a_n \neq b$. Then (a_n) is a sequence in A which converges to b.

Conversely, suppose there exists a sequence (a_n) in A such that $a_n \to b$. Either $a_n = b$ for some n (in which case $b \in A$) or else a_n is different from b for all n. In the latter case every neighborhood of b contains points of A—namely, the a_n's for n sufficiently large—other than b. Thus in either case $b \in \overline{A}$. □

Q.4.5. (Solution to 4.4.17) If $\ell = \lim_{n\to\infty} x_n$ exists, then taking limits as $n \to \infty$ of both sides of the expression $4x_{n+1} = {x_n}^3$ yields $4\ell = \ell^3$. That is,

$$\ell^3 - 4\ell = \ell(\ell - 2)(\ell + 2) = 0.$$

Thus if ℓ exists, it must be -2, 0, or 2. Next notice that

$$\begin{aligned} x_{n+1} - x_n &= \frac{1}{4}{x_n}^3 - x_n \\ &= \frac{1}{4}x_n(x_n - 2)(x_n + 2). \end{aligned}$$

Therefore

$$x_{n+1} > x_n \quad \text{if} \quad x_n \in (-2, 0) \cup (2, \infty) \tag{Q.1}$$

and

$$x_{n+1} < x_n \quad \text{if} \quad x_n \in (-\infty, -2) \cup (0, 2). \tag{Q.2}$$

Now consider the seven cases mentioned in the hint. Three of these are trivial: if $x_1 = -2$, 0, or 2, then the resulting sequence is constant (therefore certainly convergent).

Next suppose $x_1 < -2$. Then $x_n < -2$ for every n. [The verification is an easy induction: If $x_n < -2$, then ${x_n}^3 < -8$; so $x_{n+1} = \frac{1}{4}{x_n}^3 < -2$.] From this and (Q.2) we see that $x_{n+1} < x_n$ for every n. That is, the sequence (x_n) decreases. Since the only possible limits are -2, 0, and 2, the sequence cannot converge. (It must, in fact, be unbounded.)

The case $x_1 > 2$ is similar. We see easily that $x_n > 2$ for all n and therefore [by (Q.1)] the sequence (x_n) is increasing. Thus it diverges (and is unbounded).

If $-2 < x_1 < 0$, then $-2 < x_n < 0$ for every n. [Again an easy inductive proof: If $-2 < x_n < 0$, then $-8 < {x_n}^3 < 0$; so $-2 < \frac{1}{4}{x_n}^3 = x_{n+1} < 0$.] From (Q.1) we conclude that (x_n) is increasing. Being bounded above it must converge [see proposition 4.3.3] to some real number ℓ. The only available candidate is $\ell = 0$.

Similarly, if $0 < x_1 < 2$, then $0 < x_n < 2$ for all n and (x_n) is decreasing. Again the limit is $\ell = 0$.

We have shown that the sequence (x_n) converges if and only if $x_1 \in [-2, 2]$. If $x_1 \in (-2, 2)$, then $\lim x_n = 0$; if $x_1 = -2$, then $\lim x_n = -2$; and if $x_1 = 2$, then $\lim x_n = 2$.

Q.5. Exercises in chapter 05

Q.5.1. (Solution to 5.1.2) Suppose there exists a nonempty set U which is properly contained in A and which is both open and closed in A. Then, clearly, the sets U and U^c (both open in A) disconnect A. Conversely, suppose that A is disconnected by sets U and V (both open in A). Then the set U is not the null set, is not equal to A (because V, its complement with respect to A, is nonempty), is open in A, and is closed in A (because V is open in A).

Q.5.2. (Solution to 5.2.1) Let A be a subset of $\mathbb{R}$ and $f\colon A \to \mathbb{R}$ be continuous. Suppose that $\operatorname{ran} f$ is disconnected. Then there exist disjoint nonempty sets U and V both open in $\operatorname{ran} f$ whose union is $\operatorname{ran} f$. By 3.3.11 the sets $f^{\leftarrow}(U)$ and $f^{\leftarrow}(V)$ are open subsets of A. Clearly these two sets are nonempty, they are disjoint, and their union is A. This shows that A is disconnected.

Q.5.3. (Solution to 5.2.3) The continuous image of an interval is an interval (by theorem 5.2.2). Thus the range of f (that is, $f^{\rightarrow}(J)$) is an interval. So a point belongs to the range of f if it lies between two other points which belong to the range of f. That is, if a, $b \in \operatorname{ran} f$ and z lies between a and b, then there exists a point x in the domain of f such that $z = f(x)$.

Q.5.4. (Solution to 5.2.4) Let

$$f(x) = x^{27} + 5x^{13} + x - x^3 - x^5 - \frac{2}{\sqrt{1+3x^2}}$$

for all $x \in \mathbb{R}$. The function f is defined on an interval (the real line), and it is easy, using the results of chapter 3, to show that f is continuous. Notice that $f(0) = -2$ and that $f(1) = 4$. Since -2 and 4 are in the range of f and 0 lies between -2 and 4, we conclude from the *intermediate value theorem* that 0 is in the range of f. In fact, there exists an x such that $0 < x < 1$ and $f(x) = 0$. Such a number x is a solution to (5.3). Notice that we have not only shown the existence of a solution to (5.3) but also located it between consecutive integers.

Q.5.5. (Solution to 5.2.5) Let $f\colon [a,b] \to [a,b]$ be continuous. If $f(a) = a$ or $f(b) = b$, then the result is obvious; so we suppose that $f(a) > a$ and $f(b) < b$. Define $g(x) = x - f(x)$. The function g is continuous on the interval $[a,b]$. (Verify the last assertion.) Notice that $g(a) = a - f(a) < 0$ and that $g(b) = b - f(b) > 0$. Since $g(a) < 0 < g(b)$, we may conclude from the *intermediate value theorem* that $0 \in \operatorname{ran} g$. That is, there exists z in (a,b) such that $g(z) = z - f(z) = 0$. Thus z is a fixed point of f.

Q.6. Exercises in chapter 06

Q.6.1. (Solution to 6.2.3) Let A be a compact subset of $\mathbb{R}$. To show that A is closed, prove that A^c is open. Let y be a point of A^c. For each $x \in A$ choose numbers $r_x > 0$ and $s_x > 0$ such that the open intervals $J_{r_x}(x)$ and $J_{s_x}(y)$ are disjoint. The family $\{J_{r_x}(x)\colon x \in A\}$ is an open cover for A. Since A is compact there exist $x_1, \dots, x_n \in A$ such that $\{J_{r_{x_i}}(x_i)\colon 1 \le i \le n\}$ covers A. It is easy to see that if

$$t = \min\{s_{x_1}, \dots, s_{x_n}\}$$

then $J_t(y)$ is disjoint from $\bigcup_{i=1}^n J_{r_{x_i}}(x_i)$ and hence from A. This shows that $y \in A^{c\circ}$. Since y was arbitrary, A^c is open.

To prove that A is bounded, we need consider only the case where A is nonempty. Let a be any point in A. Then $\{J_n(a)\colon n \in \mathbb{N}\}$ covers A (it covers all of $\mathbb{R}$!). Since A is compact a finite subcollection of these open intervals will cover A. In this finite collection there is a largest open interval; it contains A.

Q.6.2. (Solution to 6.3.2) Let $\mathfrak{V}$ be a family of open subsets of $\mathbb{R}$ which covers $f^{\rightarrow}(A)$. The family

$$\mathfrak{U} := \{f^{\leftarrow}(V)\colon V \in \mathfrak{V}\}$$

is a family of open sets which covers A (see 3.2.12). Since A is compact we may choose sets $V_1, \dots, V_n \in \mathfrak{V}$ such that $\bigcup_{k=1}^n f^{\leftarrow}(V_k) \supseteq A$. We complete the proof by showing that $f^{\rightarrow}(A)$ is covered by the finite subfamily $\{V_1, \dots, V_n\}$ of $\mathfrak{V}$. If $y \in f^{\to}(A)$, then $y = f(x)$ for some $x \in A$. This element x belongs to at least one set $f^{\leftarrow}(V_k)$; so (by proposition M.1.22)

$$y = f(x) \in f^{\to}(f^{\leftarrow}(V_k)) \subseteq V_k\,.$$

Thus $f^{\to}(A) \subseteq \bigcup_{k=1}^{n} V_k$.

Q.6.3. (Solution to 6.3.3) We prove that f assumes a maximum on A. (To see that f has a minimum apply the present result to the function $-f$.) By theorem 6.3.2 the image of f is a compact subset of $\mathbb{R}$ and is therefore (see 6.2.3) closed and bounded. Since it is bounded the image of f has a least upper bound, say l. By example 2.2.7 the number l is in the closure of $\operatorname{ran} f$. Since the range of f is closed, $l \in \operatorname{ran} f$. Thus there exists a point a in A such that $f(a) = l \geq f(x)$ for all $x \in A$.

Q.7. Exercises in chapter 07

Q.7.1. (Solution to 7.1.3) Argue by contradiction. If $b \neq c$, then $\epsilon := |b - c| > 0$. Thus there exists $\delta_1 > 0$ such that $|f(x) - b| < \epsilon/2$ whenever $x \in A$ and $0 < |x - a| < \delta_1$ and there exists $\delta_2 > 0$ such that $|f(x) - c| < \epsilon/2$ whenever $x \in A$ and $0 < |x - a| < \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$. Since $a \in A'$, the set $A \cap J_\delta(a)$ is nonempty. Choose a point x in this set. Then

$$\epsilon = |b - c| \leq |b - f(x)| + |f(x) - c| < \epsilon/2 + \epsilon/2 = \epsilon\,.$$

This is a contradiction.

It is worth noticing that the preceding proof cannot be made to work if a is not required to be an accumulation point of A. To obtain a contradiction we must know that the condition $0 < |x-a| < \delta$ is satisfied for at least one x in the domain of f.

Q.7.2. (Solution to 7.2.3) Both halves of the proof make use of the fact that in A the open interval about a of radius δ is just $J_\delta(a) \cap A$ where $J_\delta(a)$ denotes the corresponding open interval in $\mathbb{R}$.

Suppose f is continuous at a. Given $\epsilon > 0$ choose $\delta > 0$ so that $x \in J_\delta(a) \cap A$ implies $f(x) \in J_\epsilon(f(a))$. If $x \in A$ and $0 < |x - a| < \delta$, then $|f(x) - f(a)| < \epsilon$. That is, $\lim_{x \to a} f(x) = f(a)$.

Conversely, suppose that $\lim_{x\to a} f(x) = f(a)$. Given $\epsilon > 0$ choose $\delta > 0$ so that $|f(x)-f(a)| < \epsilon$ whenever $x \in A$ and $0 < |x - a| < \delta$. If $x = a$, then $|f(x) - f(a)| = 0 < \epsilon$. Thus $x \in J_\delta(a) \cap A$ implies $f(x) \in J_\epsilon(f(a))$. This shows that f is continuous at a.

Q.7.3. (Solution to 7.2.4) Let $g\colon h \mapsto f(a+h)$. Notice that $h \in \operatorname{dom} g$ if and only if $a+h \in \operatorname{dom} f$;

$$\operatorname{dom} f = a + \operatorname{dom} g\,. \tag{Q.3}$$

That is, $\operatorname{dom} f = \{a + h\colon h \in \operatorname{dom} g\}$.

First we suppose that

$$l := \lim_{h\to 0} g(h) = \lim_{h\to 0} f(a+h) \text{ exists.} \tag{Q.4}$$

We show that $\lim_{x\to a} f(x)$ exists and equals l. Given $\epsilon > 0$ there exists (by (Q.4)) a number $\delta > 0$ such that

$$|g(h) - l| < \epsilon \text{ whenever } h \in \operatorname{dom} g \text{ and } 0 < |h| < \delta. \tag{Q.5}$$

Now suppose that $x \in \operatorname{dom} f$ and $0 < |x - a| < \delta$. Then by (Q.3)

$$x - a \in \operatorname{dom} g$$

and by (Q.5)

$$|f(x) - l| = |g(x - a) - l| < \epsilon\,.$$

Thus given $\epsilon > 0$ we have found $\delta > 0$ such that $|f(x) - l| < \epsilon$ whenever $x \in \operatorname{dom} f$ and $0 < |x - a| < \delta$. That is, $\lim_{x\to a} f(x) = l$.

The converse argument is similar. Suppose $l := \lim_{x\to a} f(x)$ exists. Given $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x)-l| < \epsilon$ whenever $x \in \operatorname{dom} f$ and $0 < |x-a| < \delta$. If $h \in \operatorname{dom} g$ and $0 < |h| < \delta$, then $a+h \in \operatorname{dom} f$ and $0 < |(a+h)-a| < \delta$. Therefore $|g(h)-l| = |f(a+h)-l| < \epsilon$, which shows that $\lim_{h\to 0} g(h) = l$.

Q.8. Exercises in chapter 08

Q.8.1. (Solution to 8.1.9) Suppose that $T \in \mathfrak{L} \cap \mathfrak{o}$. Let $\epsilon > 0$. Since $T \in \mathfrak{o}$, there exists $\delta > 0$ so that $|Ty| \leq \epsilon|y|$ whenever $|y| < \delta$. Since $T \in \mathfrak{L}$ there exists $m \in \mathbb{R}$ such that $Tx = mx$ for all x (see example 8.1.7). Now, suppose $0 < |y| < \delta$. Then

$$|m|\,|y| = |Ty| \leq \epsilon y$$

so $|m| \leq \epsilon$. Since ϵ was arbitrary, we conclude that $m = 0$. That is, T is the constant function 0.

Q.8.2. (Solution to 8.1.10) Let f, $g \in \mathfrak{O}$. Then there exist positive numbers M, N, δ, and η such that $|f(x)| \leq M|x|$ whenever $|x| < \delta$ and $|g(x)| < N|x|$ whenever $|x| < \eta$. Then $|f(x) + g(x)| \leq (M + N)|x|$ whenever $|x|$ is less than the smaller of δ and η. So $f + g \in \mathfrak{O}$.

If c is a constant, then $|cf(x)| = |c|\,|f(x)| \leq |c|\,M\,|x|$ whenever $|x| < \delta$; so $cf \in \mathfrak{O}$.

Q.8.3. (Solution to 8.1.13) (The domain of $f \circ g$ is taken to be the set of all numbers x such that $g(x)$ belongs to the domain of f; that is, $\operatorname{dom}(f \circ g) = g^{\leftarrow}(\operatorname{dom} f)$.) Since $f \in \mathfrak{O}$ there exist $M, \delta > 0$ such that $|f(y)| \leq M|y|$ whenever $|y| < \delta$. Let $\epsilon > 0$. Since $g \in \mathfrak{o}$ there exists $\eta > 0$ such that $|g(x)| \leq \epsilon M^{-1}|x|$ whenever $|x| \leq \eta$.

Now if $|x|$ is less than the smaller of η and $M\epsilon^{-1}\delta$, then $|g(x)| \leq \epsilon M^{-1}|x| < \delta$, so that

$$|(f \circ g)(x)| \leq M|g(x)| \leq \epsilon|x|.$$

Thus $f \circ g \in \mathfrak{o}$.

Q.8.4. (Solution to 8.1.14) Since ϕ and f belong to $\mathfrak{O}$, there exist positive numbers M, N, δ, and η such that $|\phi(x)| \leq M|x|$ whenever $|x| < \delta$ and $|f(x)| < N|x|$ whenever $|x| < \eta$. Suppose $\epsilon > 0$. If $|x|$ is less than the smallest of $\epsilon M^{-1}N^{-1}$, δ, and η, then

$$|(\phi f)(x)| = |\phi(x)|\,|f(x)| \leq MNx^2 \leq \epsilon|x|\,.$$

Q.8.5. (Solution to 8.2.3) Clearly $f(0) - g(0) = 0$. Showing that $\lim_{x\to 0} \dfrac{f(x) - g(x)}{x} = 0$ is a routine computation:

$$\begin{aligned}\lim_{x\to 0} \frac{f(x) - g(x)}{x} &= \lim_{x\to 0} \frac{1}{x}\left(x^2 - 4x - 1 - \frac{1}{3x^2 + 4x - 1}\right)\\ &= \lim_{x\to 0} \frac{3x^4 - 8x^3 - 20x^2}{x(3x^2 + 4x - 1)} = 0.\end{aligned}$$

Q.8.6. (Solution to 8.2.4) Reflexivity is obvious. Symmetry: if $f \simeq g$, then $f - g \in \mathfrak{o}$; so $g - f = (-1)(f - g) \in \mathfrak{o}$ (by proposition 8.1.11). Thus $g \simeq f$. Transitivity: if $f \simeq g$ and $g \simeq h$, then $g - f \in \mathfrak{o}$ and $h - g \in \mathfrak{o}$; so $h - f = (h - g) + (g - f) \in \mathfrak{o}$ (again by 8.1.11). Thus $f \simeq h$.

Q.8.7. (Solution to 8.2.5) Since $S \simeq f$ and $T \simeq f$, we conclude from proposition 8.2.4 that $S \simeq T$. Then $S - T \in \mathfrak{L} \cap \mathfrak{o}$ and thus $S - T = 0$ by proposition 8.1.9.

Q.8.8. (Solution to 8.3.4) For each h in the domain of $\Delta(f + g)_a$ we have

$$\begin{aligned}\Delta(f + g)_a(h) &= (f + g)(a + h) - (f + g)(a)\\ &= f(a + h) + g(a + h) - f(a) - g(a)\\ &= \Delta f_a(h) + \Delta g_a(h)\\ &= (\Delta f_a + \Delta g_a)(h).\end{aligned}$$

Q.8.9. (Solution to 8.3.6) For every h in the domain of $\Delta(g \circ f)_a$ we have

$$\begin{aligned}\Delta(g \circ f)_a(h) &= g(f(a + h)) - g(f(a))\\ &= g(f(a) + f(a + h) - f(a)) - g(f(a))\\ &= g(f(a) + \Delta f_a(h)) - g(f(a))\\ &= \big(\Delta g_{f(a)} \circ \Delta f_a\big)(h).\end{aligned}$$

Q.8.10. (Solution to 8.4.8) If f is differentiable at a, then

$$\Delta f_a = \big(\Delta f_a - df_a\big) + df_a \in \mathfrak{o} + \mathfrak{L} \subseteq \mathfrak{O} + \mathfrak{O} \subseteq \mathfrak{O}\,.$$

Q.8.11. (Solution to 8.4.9) If $f \in \mathcal{D}_a$, then $\Delta f_a \in \mathfrak{O} \subseteq \mathcal{C}_0$. Since Δf_a is continuous at 0, f is continuous at a.

Q.8.12. (Solution to 8.4.10) Since f is differentiable at a, its differential exists and $\Delta f_a \simeq df_a$. Then

$$\Delta(\alpha f)_a = \alpha \Delta f_a \simeq \alpha df_a$$

by propositions 8.3.3 and 8.2.6. Since αdf_a is a linear function which is tangent to $\Delta(\alpha f)_a$, we conclude that it must be the differential of αf at a (see proposition 8.4.2). That is, $\alpha df_a = d(\alpha f)_a$.

Q.8.13. (Solution to 8.4.12) It is easy to check that $\phi(a)df_a + f(a)d\phi_a$ is a linear function. From $\Delta f_a \simeq df_a$ we infer that $\phi(a)\Delta f_a \simeq \phi(a)df_a$ (by proposition 8.2.6), and from $\Delta\phi_a \simeq d\phi_a$ we infer that $f(a)\Delta\phi_a \simeq f(a)d\phi_a$ (also by 8.2.6). From propositions 8.4.8 and 8.1.14 we see that

$$\Delta\phi_a \cdot \Delta f_a \in \mathfrak{O} \cdot \mathfrak{O} \subseteq \mathfrak{o}\,;$$

that is, $\Delta\phi_a \cdot \Delta f_a \simeq 0$. Thus by propositions 8.3.5 and 8.2.6

$$\begin{aligned}\Delta(\phi f)_a &= \phi(a)\Delta f_a + f(a)\Delta\phi_a + \Delta\phi_a \cdot \Delta f_a \\ &\simeq \phi(a)df_a + f(a)d\phi_a + 0 \\ &= \phi(a)df_a + f(a)d\phi_a.\end{aligned}$$

Q.8.14. (Solution to 8.4.13) By hypothesis $\Delta f_a \simeq df_a$ and $\Delta g_{f(a)} \simeq dg_{f(a)}$. By proposition 8.4.8 $\Delta f_a \in \mathfrak{O}$. Then by proposition 8.2.8

$$\Delta g_{f(a)} \circ \Delta f_a \simeq dg_{f(a)} \circ \Delta f_a\,; \tag{Q.6}$$

and by proposition 8.2.7

$$dg_{f(a)} \circ \Delta f_a \simeq dg_{f(a)} \circ df_a\,. \tag{Q.7}$$

According to proposition 8.3.6

$$\Delta(g \circ f)_a = \Delta g_{f(a)} \circ \Delta f_a\,. \tag{Q.8}$$

From (Q.6), (Q.7) , and (Q.8), and proposition 8.2.4 we conclude that

$$\Delta(g \circ f)_a \simeq dg_{f(a)} \circ df_a\,.$$

Since $dg_{f(a)} \circ df_a$ is a linear function, the desired conclusion is an immediate consequence of proposition 8.4.2.

Q.9. Exercises in chapter 09

Q.9.1. (Solution to 9.1.2) If x, $y \in M$, then

$$0 = d(x,x) \le d(x,y) + d(y,x) = 2\,d(x,y)\,.$$

Q.9.2. (Solution to 9.3.2) It is clear that $\max\{a,b\} \le a+b$. Expanding the right side of $0 \le (a \pm b)^2$ we see that $2|ab| \le a^2 + b^2$. Thus

$$\begin{aligned}(a+b)^2 &\le a^2 + 2|ab| + b^2 \\ &\le 2(a^2+b^2)\,.\end{aligned}$$

Taking square roots we get $a + b \le \sqrt{2}\,\sqrt{a^2+b^2}$. Finally,

$$\begin{aligned}\sqrt{2}\,\sqrt{a^2+b^2} &\le \sqrt{2}\,\sqrt{2(\max\{a,b\})^2} \\ &= 2\max\{a,b\}\,.\end{aligned}$$

We have established the claim made in the hint. Now if $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are points in $\mathbb{R}^2$, then

$$\begin{aligned}\max\{|x_1 - y_1|, |x_2 - y_2|\} &\le |x_1 - y_1| + |x_2 - y_2| \\ &\le \sqrt{2}\,\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \\ &\le 2\max\{|x_1 - y_1|, |x_2 - y_2|\}\,.\end{aligned}$$

In other words

$$d_u(x, y) \le d_1(x, y) \le \sqrt{2}\, d(x, y) \le 2\, d_u(x, y)\,.$$

Q.10. Exercises in chapter 10

Q.10.1. (Solution to 10.1.3) The three open balls are $B_1(-1) = \{-1\}$, $B_1(0) = [0, 1)$, and $B_2(0) = \{-1\} \cup [0, 2)$.

Q.10.2. (Solution to 10.1.14) (a) $A^\circ = \emptyset$; $A' = \{0\}$; $\overline{A} = \partial A = \{0\} \cup A$.
(b) $A^\circ = \emptyset$; $A' = \overline{A} = \partial A = [0, \infty)$.
(c) $A^\circ = A$; $A' = \overline{A} = A \cup \{0\}$; $\partial A = \{0\}$.

Q.10.3. (Solution to 10.2.1) Let $t = r - d(a, c)$. (Note that $t > 0$.) If $x \in B_t(c)$, then $d(a, x) \le d(a, c) + d(c, x) < (r - t) + t = r$; so $x \in B_r(a)$. (Easier solution: This is just a special case of proposition 9.2.19. Take $b = a$ and $s = r$ there.)

Q.10.4. (Solution to 10.2.2) (a) If $x \in A^\circ$, then there exists $r > 0$ such that $B_r(x) \subseteq A \subseteq B$. So $x \in B^\circ$.

(b) Since $A^\circ \subseteq A$ we may conclude from (a) that $A^{\circ\circ} \subseteq A^\circ$. For the reverse inclusion take $a \in A^\circ$. Then there exists $r > 0$ such that $B_r(a) \subseteq A$. By lemma 10.2.1, about each point b in the open ball $B_r(a)$ we can find an open ball $B_s(b)$ contained in $B_r(a)$ and hence in A. This shows that $B_r(a) \subseteq A^\circ$. Since some open ball about a lies inside A°, we conclude that $a \in A^{\circ\circ}$.

Q.10.5. (Solution to 10.2.4) (a) Since $A \subseteq \bigcup \mathfrak{A}$ for every $A \in \mathfrak{A}$, we can conclude from proposition 10.2.2(a) that $A^\circ \subseteq \big(\bigcup \mathfrak{A}\big)^\circ$ for every $A \in \mathfrak{A}$. Thus $\bigcup\{A^\circ \colon A \in \mathfrak{A}\} \subseteq \big(\bigcup \mathfrak{A}\big)^\circ$.

(b) Let $\mathbb{R}$ be the metric space and $\mathfrak{A} = \{A, B\}$ where $A = \mathbb{Q}$ and $B = \mathbb{Q}^c$. Then $A^\circ \cup B^\circ = \emptyset$ while $(A \cup B)^\circ = \mathbb{R}$.

Q.11. Exercises in chapter 11

Q.11.1. (Solution to 11.1.6) Since A° is an open set contained in A (by 10.2.2(b)), it is contained in the union of all such sets. That is,

$$A^\circ \subseteq \bigcup\{\, U \colon U \subseteq A \text{ and } U \text{ is open}\}.$$

On the other hand, if U is an open subset of A, then (by 10.2.2(a)) $U = U^\circ \subseteq A^\circ$. Thus

$$\bigcup\{U \colon U \subseteq A \text{ and } U \text{ is open}\} \subseteq A^\circ\,.$$

Q.11.2. (Solution to 11.1.9) Let A be a subset of a metric space. Using problem 10.3.6 we see that

$$\begin{aligned} A \text{ is open } &\text{ iff } A = A^\circ \\ &\text{ iff } A^c = \big(A^\circ\big)^c = \overline{A^c} \\ &\text{ iff } A^c \text{ is closed.}\end{aligned}$$

Q.11.3. (Solution to 11.1.22) Suppose that D is dense in M. Argue by contradiction. If there is an open ball B which contains no point of D (that is, $B \subseteq D^c$), then

$$B = B^\circ \subseteq \big(D^c\big)^\circ = \big(\,\overline{D}\,\big)^c = M^c = \emptyset\,,$$

which is not possible.

Conversely, suppose that D is not dense in M. Then $\overline{D}$ is a proper closed subset of M, making $\big(\overline{D}\big)^c$ a nonempty open set. Choose an open ball $B \subseteq \big(\overline{D}\big)^c$. Since $\big(\overline{D}\big)^c \subseteq D^c$, the ball B contains no point of D.

Q.11.4. (Solution to 11.2.3) It is enough to show that $\mathfrak{T}_1 \subseteq \mathfrak{T}_2$. Thus we suppose that U is an open subset of (M, d_1) and prove that it is an open subset of (M, d_2). For $a \in M$, $r > 0$, and $k = 1, 2$ let $B^k_r(a)$ be the open ball about a of radius r in the space (M, d_k); that is,

$$B^k_r(a) = \{x \in M : d_k(x, a) < r\}\,.$$

To show that U is an open subset of (M, d_2) choose an arbitrary point x in U and find an open ball $B^2_r(x)$ about x contained in U. Since U is assumed to be open in (M, d_1) there exists $s > 0$ such that $B^1_s(x) \subseteq U$. The metrics d_1 and d_2 are equivalent, so, in particular, there is a constant $\alpha > 0$ such that $d_1(u, v) \leq \alpha\, d_2(u, v)$ for all $u, v \in M$. Let $r = s\,\alpha^{-1}$. Then if $y \in B^2_r(x)$, we see that

$$d_1(y, x) \leq \alpha\, d_2(y, x) < \alpha\, r = s\,.$$

Thus $B^2_r(x) \subseteq B^1_s(x) \subseteq U$.

Q.12. Exercises in chapter 12

Q.12.1. (Solution to 12.2.2) Suppose that A is closed in M. Let (a_n) be a sequence in A which converges to a point b in M. If b is in A^c then, since A^c is a neighborhood of b, the sequence (a_n) is eventually in A^c, which is not possible. Therefore $b \in A$.

Conversely, if A is not closed, there exists an accumulation point b of A which does not belong to A. Then for every $n \in \mathbb{N}$ we may choose a point a_n in $B_{1/n}(b) \cap A$. The sequence (a_n) lies in A and converges to b; but $b \notin A$.

Q.12.2. (Solution to 12.3.2) As in 12.3.1, let ρ_1 be the metric on M_1 and ρ_2 be the metric on M_2. Use the inequalities given in the hint to 9.3.2 with $a = \rho_1(x_1, y_1)$ and $b = \rho_2(x_2, y_2)$ to obtain

$$d_u(x, y) \leq d_1(x, y) \leq \sqrt{2}\, d(x, y) \leq 2\, d_u(x, y)\,.$$

Q.13. Exercises in chapter 13

Q.13.1. (Solution to 13.1.6) Let C be the set of all functions defined on $[0, 1]$ such that $0 < g(x) < 2$ for all $x \in [0, 1]$. It is clear that $B_1(f) \subseteq C$. The reverse inclusion, however, is not correct. For example, let

$$g(x) = \begin{cases} 1, & \text{if } x = 0 \\ x, & \text{if } 0 < x \leq 1. \end{cases}$$

Then g belongs to C; but it does *not* belong to $B_1(f)$ since

$$d_u(f, g) = \sup\{|f(x) - g(x)| : 0 \leq x \leq 1\} = 1\,.$$

Q.13.2. (Solution to 13.1.10) Let f, g, $h \in \mathcal{B}(S)$. There exist positive constants M, N, and P such that $|f(x)| \leq M$, $|g(x)| \leq N$, and $|h(x)| \leq P$ for all x in S.

First show that d_u is real valued. (That is, show that d_u is never infinite.) This is easy:

$$\begin{aligned} d_u(f, g) &= \sup\{|f(x) - g(x)| \colon x \in S\} \\ &\leq \sup\{|f(x)| + |g(x)| \colon x \in S\} \\ &\leq M + N. \end{aligned}$$

Now verify conditions (1)–(3) of the definition of “metric” in 9.1.1. Condition (1) follows from the observation that

$$|f(x) - g(x)| = |g(x) - f(x)| \qquad \text{for all } x \in S.$$

To establish condition (2) notice that for every $x \in S$

$$\begin{aligned}|f(x) - h(x)| &\le |f(x) - g(x)| + |g(x) - h(x)| \\ &\le d_u(f,g) + d_u(g,h);\end{aligned}$$

and therefore

$$d_u(f,h) = \sup\{|f(x) - h(x)| \colon x \in S\} \le d_u(f,g) + d_u(g,h)\,.$$

Finally, condition (3) holds since

$$\begin{aligned}d_u(f,g) = 0 &\text{ iff } |f(x) - g(x)| = 0 \qquad \text{for all } x \in S \\ &\text{ iff } f(x) = g(x) \qquad \text{for all } x \in S \\ &\text{ iff } f = g.\end{aligned}$$

Q.13.3. (Solution to 13.2.2) Let (f_n) be a sequence of functions in $\mathcal{F}(S,\mathbb{R})$ and suppose that $f_n \to g\,(\text{unif})$ in $\mathcal{F}(S,\mathbb{R})$. Then for every $x \in S$

$$|f_n(x) - g(x)| \le \sup\{|f_n(y) - g(y)| \colon y \in S\} \to 0 \qquad \text{as } n \to \infty.$$

Thus $f_n \to g\,(\text{ptws})$.

On the other hand, if we define for each $n \in \mathbb{N}$ a function f_n on $\mathbb{R}$ by

$$f_n(x) = \begin{cases} 1, & \text{if } x \ge n \\ 0, & \text{if } x < n, \end{cases}$$

then it is easy to see that the sequence (f_n) converges pointwise to the zero function $\mathbf{0}$; but since $d_u(f_n, \mathbf{0}) = 1$ for every n, the sequence does not converge uniformly to $\mathbf{0}$.

Q.13.4. (Solution to 13.2.4) (a) Since $f_n \to g\,(\text{unif})$, there exists $m \in \mathbb{N}$ such that

$$|f_n(x) - g(x)| \le \sup\{|f_n(y) - g(y)| \colon y \in S\} < 1$$

whenever $n \ge m$ and $x \in S$. Thus, in particular,

$$|g(x)| \le |f_m(x) - g(x)| + |f_m(x)| < 1 + K$$

where K is a number satisfying $|f_m(x)| \le K$ for all $x \in S$.

(b) Let

$$f_n(x) = \begin{cases} x, & \text{if } |x| \le n \\ 0, & \text{if } |x| > n \end{cases}$$

and $g(x) = x$ for all x in $\mathbb{R}$. Then $f_n \to g\,(\text{ptws})$, each f_n is bounded, but g is not.

Q.13.5. (Solution to 13.2.5) If $0 \le x < 1$, then

$$f_n(x) = x^n - x^{2n} \to 0 \qquad \text{as } n \to \infty\,.$$

This and the obvious fact that $f_n(1) = 0$ for every n tell us that

$$f_n \to \mathbf{0}\,(\text{ptws})\,.$$

Observe that $x^{2n} \le x^n$ whenever $0 \le x \le 1$ and $n \in \mathbb{N}$. Thus $f_n \ge 0$ for each n. Use techniques from beginning calculus to find the maximum value of each f_n. Differentiating we see that

$$f_n'(x) = n\,x^{n-1} - 2n\,x^{2n-1} = n\,x^{n-1}(1 - 2x^n)$$

for $n > 1$. Thus the function f_n has a critical point, in fact assumes a maximum, when $1 - 2x^n = 0$; that is, when $x = 2^{-1/n}$. But then

$$\sup\{|f_n(x) - 0| \colon 0 \le x \le 1\} = f_n(2^{-1/n}) = \tfrac{1}{2} - \tfrac{1}{4} = \tfrac{1}{4}$$

for all $n > 1$. Since $\sup\{|f_n(x) - 0| \colon 0 \le x \le 1\} \nrightarrow 0$ as $n \to \infty$, the convergence is not uniform.

Is it possible that the sequence f_n converges uniformly to some function g other than the zero function? The answer is *no*. According to proposition 13.2.2 if $f_n \to g \ne 0\,(\text{unif})$, then $f_n \to g\,(\text{ptws})$. But this contradicts what we have already shown, namely, $f_n \to 0\,(\text{ptws})$.

Note: Since each of the functions f_n belongs to $\mathcal{B}([0,1])$, it is permissible, and probably desirable, in the preceding proof to replace each occurrence of the rather cumbersome expression $\sup\{|f_n(x)-0|\colon 0 \le x \le 1\}$ by $d_u(f_n, \mathbf{0})$.

Q.14. Exercises in chapter 14

Q.14.1. (Solution to 14.1.5) Suppose f is continuous. Let $U \overset{\circ}{\subseteq} M_2$. To show that $f^{\leftarrow}(U)$ is an open subset of M_1, it suffices to prove that each point of $f^{\leftarrow}(U)$ is an interior point of $f^{\leftarrow}(U)$. If $a \in f^{\leftarrow}(U)$, then $f(a) \in U$. Since f is continuous at a, the set U, which is a neighborhood of $f(a)$, must contain the image under f of a neighborhood V of a. But then

$$a \in V \subseteq f^{\leftarrow}(f^{\rightarrow}(V)) \subseteq f^{\leftarrow}(U)$$

which shows that a lies in the interior of $f^{\leftarrow}(U)$.

Conversely, suppose that $f^{\leftarrow}(U) \overset{\circ}{\subseteq} M_1$ whenever $U \overset{\circ}{\subseteq} M_2$. To see that f is continuous at an arbitrary point a in M_1, notice that if V is a neighborhood of $f(a)$, then $a \in f^{\leftarrow}(V) \overset{\circ}{\subseteq} M_1$. Thus $f^{\leftarrow}(V)$ is a neighborhood of a whose image $f^{\rightarrow}(f^{\leftarrow}(V))$ is contained in V. Thus f is continuous at a.

Q.14.2. (Solution to 14.1.9) Show that M is continuous at an arbitrary point (a,b) in $\mathbb{R}^2$. Since the metric d_1 (defined in 9.2.10) is equivalent to the usual metric on $\mathbb{R}^2$, proposition 14.1.8 assures us that it is enough to establish continuity of the function M with respect to the metric d_1. Let $K = |a| + |b| + 1$. Given $\epsilon > 0$, choose $\delta = \min\{\epsilon/K, 1\}$. If (x,y) is a point in $\mathbb{R}^2$ such that $|x-a| + |y-b| = d_1((x,y),(a,b)) < \delta$, then

$$|x| \le |a| + |x-a| < |a| + \delta \le |a| + 1 \le K.$$

Thus for all such points (x,y)

$$\begin{aligned}|M(x,y) - M(a,b)| &= |xy - xb + xb - ab|\\ &\le |x|\,|y-b| + |x-a|\,|b|\\ &\le K|y-b| + K|x-a|\\ &< K\delta\\ &\le \epsilon.\end{aligned}$$

Q.14.3. (Solution to 14.1.26) Suppose that f is continuous at a. Let $x_n \to a$ and B_2 be a neighborhood of $f(a)$. There exists a neighborhood B_1 of a such that $f^{\rightarrow}(B_1) \subseteq B2$. Choose $n_0 \in \mathbb{N}$ so that $x_n \in B_1$ whenever $n \ge n_0$. Then $f(x_n) \in f^{\rightarrow}(B_1) \subseteq B_2$ whenever $n \ge n_0$. That is, the sequence $\bigl(f(x_n)\bigr)$ is eventually in the neighborhood B_2. Since B_2 was arbitrary, $f(x_n) \to f(a)$.

Conversely, suppose that f is not continuous at a. Then there exists $\epsilon > 0$ such that the image under f of $B_\delta(a)$ contains points which do not belong to $B_\epsilon(f(a))$, no matter how small $\delta > 0$ is chosen. Thus for every $n \in \mathbb{N}$ there exists $x_n \in B_{1/n}(a)$ such that $f(x_n) \notin B_\epsilon(f(a))$. Then clearly $x_n \to a$ but $f(x_n) \not\to f(a)$.

Q.14.4. (Solution to 14.2.1) Let (a,b) be an arbitrary point in $M_1 \times M_2$. If $(x_n, y_n) \to (a,b)$ in $M_1 \times M_2$, then proposition 12.3.4 tells us that $\pi_1(x_n, y_n) = x_n \to a = \pi_1(a,b)$. This shows that π_1 is continuous at (a,b). Similarly, π_2 is continuous at (a,b).

Q.14.5. (Solution to 14.2.3) Suppose f is continuous. Then the components $f^1 = \pi_1 \circ f$ and $f^2 = \pi_2 \circ f$, being composites of continuous functions, are continuous. Conversely, suppose that f^1 and f^2 are continuous. Let a be an arbitrary point in N. If $x_n \to a$, then (by proposition 14.1.26) $f^1(x_n) \to f^1(a)$ and $f^2(x_n) \to f^2(a)$. From proposition 12.3.4 we conclude that

$$f(x_n) = (f^1(x_n), f^2(x_n)) \to (f^1(a), f^2(a)) = f(a)\,;$$

so f is continuous at a.

Q.14.6. (Solution to 14.2.4) (a) The function fg is the composite of continuous functions (that is, $fg = M \circ (f, g)$); so it is continuous by corollary 14.1.4.

(b) This is just a special case of (a) where f is the constant function whose value is α.

Q.14.7. (Solution to 14.2.12) For each $a \in M_1$ let $j_a \colon M_2 \to M_1 \times M_2$ be defined by $j_a(y) = (a, y)$. Then j_a is continuous. (Proof: If $y_n \to c$, then $j_a(y_n) = (a, y_n) \to (a, c) = j_a(c)$.) Since $f(a, \cdot\,) = f \circ j_a$, it too is continuous. The continuity of each $f(\,\cdot\, b)$ is established in a similar manner.

Q.14.8. (Solution to 14.2.15) Show that g is continuous at an arbitrary point a in M. Let $\epsilon > 0$. Since $f_n \to g$ (unif) there exists $n \in \mathbb{N}$ such that

$$|f_m(x) - g(x)| \le \sup\{|f_n(y) - g(y)| \colon y \in M\} < \epsilon/3$$

whenever $m \ge n$ and $x \in M$. Since f_n is continuous at a, there exists $\delta > 0$ such that

$$|f_n(x) - f_n(a)| < \epsilon/3$$

whenever $d(x, a) < \delta$. Thus for $x \in B_\delta(a)$

$$\begin{aligned} |g(x) - g(a)| &\le |g(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - g(a)| \\ &< \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon\,. \end{aligned}$$

This shows that g is continuous at a.

Q.14.9. (Solution to 14.3.3) Argue by contradiction. If $b \ne c$, then $\epsilon = d(b, c) > 0$. Thus there exists $\delta_1 > 0$ such that $d(f(x), b) < \epsilon/2$ whenever $x \in A$ and $0 < d(x, a) < \delta_1$, and there exists $\delta_2 > 0$ such that $d(f(x), c) < \epsilon/2$ whenever $x \in A$ and $0 < d(x, a) < \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$. Since $a \in A'$, the set $A \cap B_\delta(a)$ is nonempty. Choose a point x in this set. Then

$$\epsilon = d(b, c) \le d(b, f(x)) + d(f(x), c) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon\,.$$

This is a contradiction.

It is worth noticing that the preceding proof cannot be made to work if a is not required to be an accumulation point of A. To obtain a contradiction we must know that the condition $0 < d(x, a) < \delta$ is satisfied for at least one x in the domain of f.

Q.14.10. (Solution to 14.3.8) Let $g(x) = \lim_{y\to b} f(x, y)$ for all $x \in \mathbb{R}$. It is enough to show that $\lim_{x\to a}\bigl(\lim_{y\to b} f(x, y)\bigr) = \lim_{x\to a} g(x) = l$. Given $\epsilon > 0$, choose $\delta > 0$ so that $|f(x, y) - l| < \epsilon/2$ whenever $0 < d\bigl((x, y), (a, b)\bigr) < \delta$. (This is possible because $l = \lim_{(x,y)\to(a,b)} f(x, y)$.) Suppose that $0 < |x - a| < \delta/\sqrt{2}$. Then (by the definition of g) there exists $\eta_x > 0$ such that $|g(x) - f(x, y)| < \epsilon/2$ whenever $0 < |y - b| < \eta_x$. Now choose any y such that $0 < |y - b| < \min\{\delta/\sqrt{2}, \eta_x\}$. Then (still supposing that $0 < |x - a| < \delta/\sqrt{2}$) we see that

$$\begin{aligned} 0 &< d\bigl((x, y), (a, b)\bigr) \\ &= \bigl((x - a)^2 + (y - b)^2\bigr)^{1/2} \\ &< \bigl((\delta^2/2 + \delta^2/2\bigr)^{1/2} \\ &= \delta \end{aligned}$$

so

$$|f(x, y) - l| < \epsilon/2$$

and therefore

$$\begin{aligned} |g(x) - l| &\le |g(x) - f(x, y)| + |f(x, y) - l| \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon. \end{aligned}$$

That is, $\lim_{x\to a} g(x) = l$.

Q.14.11. (Solution to 14.3.9) By the remarks preceding this exercise, we need only find two distinct values which the function f assumes in every neighborhood of the origin. This is easy. Every neighborhood of the origin contains points $(x, 0)$ distinct from $(0, 0)$ on the x-axis. At every such point $f(x, y) = f(x, 0) = 0$. Also, every neighborhood of the origin contains points (x, x) distinct from $(0, 0)$ which lie on the line $y = x$. At each such point $f(x, y) = f(x, x) = 1/17$. Thus f has no limit at the origin since in every neighborhood of $(0, 0)$ it assumes both the values 0 and 1/17. (Notice, incidentally, that both iterated limits, $\lim_{x\to 0}\bigl(\lim_{y\to 0} f(x, y)\bigr)$ and $\lim_{y\to 0}\bigl(\lim_{x\to 0} f(x, y)\bigr)$ exist and equal 0.)

Q.15. Exercises in chapter 15

Q.15.1. (Solution to 15.1.2) Suppose that A is compact. Let $\mathfrak{V}$ be a family of open subsets of M which covers A. Then

$$\mathfrak{U} := \{V \cap A \colon V \in \mathfrak{V}\}$$

is a family of open subsets of A which covers A. Since A is compact we may choose a subfamily $\{U_1, \dots, U_n\} \subseteq \mathfrak{U}$ which covers A. For $1 \le k \le n$ choose $V_k \in \mathfrak{V}$ so that $U_k = V_k \cap A$. Then $\{V_1, \dots, V_n\}$ is a finite subfamily of $\mathfrak{V}$ which covers A.

Conversely, Suppose every cover of A by open subsets of M has a finite subcover. Let $\mathfrak{U}$ be a family of open subsets of A which covers A. According to proposition 11.2.1 there is for each U in $\mathfrak{U}$ a set V_U open in M such that $U = V_U \cap A$. Let $\mathfrak{V}$ be $\{V_U \colon U \in \mathfrak{U}\}$. This is a cover for A by open subsets of M. By hypothesis there is a finite subfamily $\{V_1, \dots, V_n\}$ of $\mathfrak{V}$ which covers A. For $1 \le k \le n$ let $U_k = V_k \cap A$. Then $\{U_1, \dots, U_n\}$ is a finite subfamily of $\mathfrak{U}$ which covers A. Thus A is compact.

Q.16. Exercises in chapter 16

Q.16.1. (Solution to 16.1.7) If a metric space M is not totally bounded then there exists a positive number ϵ such that for every finite subset F of M there is a point a in M such that $F \cap B_\epsilon(a) = \emptyset$. Starting with an arbitrary point x_1 in M, construct a sequence (x_k) inductively as follows: having chosen $x_1, \dots, x_n$ no two of which are closer together than ϵ, choose $x_{n+1} \in M$ so that

$$B_\epsilon(x_{n+1}) \cap \{x_1, \dots, x_n\} = \emptyset\,.$$

Since no two terms of the resulting sequence are closer together than ϵ, it has no convergent subsequence.

Q.16.2. (Solution to 16.2.1) (1) $\Rightarrow$ (2): Suppose that there exists an infinite subset A of M which has no accumulation point in M. Then $\overline{A} = A$, so that A is closed. Each point $a \in A$ has a neighborhood B_a which contains no point of A other than a. Thus $\{B_a \colon a \in A\}$ is a cover for A by open subsets of M no finite subfamily of which covers A. This shows that A is not compact. We conclude from proposition 15.1.3 that M is not compact.

(2) $\Rightarrow$ (3): Suppose that (2) holds. Let a be a sequence in M. If the range of a is finite, then a has a constant (therefore convergent) subsequence. Thus we assume that the image of a is infinite. By hypothesis $\operatorname{ran} a$ has an accumulation point $m \in M$. We define $n \colon \mathbb{N} \to \mathbb{N}$ inductively: let $n(1) = 1$; if $n(1), \dots, n(k)$ have been defined, let $n(k+1)$ be an integer greater than $n(k)$ such that $a_{n_{k+1}} \in B_{1/(k+1)}(m)$. (This is possible since $B_{1/(k+1)}(m)$ contains infinitely many distinct points of the range of a.) It is clear that $a \circ n$ is a subsequence of a and that $a_{n_k} \to m$ as $k \to \infty$.

(3) $\Rightarrow$ (1): Let $\mathfrak{U}$ be an open cover for M. By corollary 16.1.13 the space M is separable. Let A be a countable dense subset of M and $\mathfrak{V}$ be the family of all open balls $B(a; r)$ such that

(i) $a \in A$;

(ii) $r \in \mathbb{Q}$ and;

(iii) $B_r(a)$ is contained in at least one member of $\mathfrak{U}$.

Then for each $B \in \mathfrak{B}$ choose a set U_B in $\mathfrak{U}$ which contains B. Let

$$\mathfrak{V} = \{U_B \in \mathfrak{U} : B \in \mathfrak{B}\}\,.$$

It is clear that $\mathfrak{V}$ is countable; we show that V covers M.

Let $x \in M$. There exist U_0 in $\mathfrak{U}$ and $r > 0$ such that $B_r(x) \subseteq U_0$. Since A is dense in M, proposition 11.1.22 allows us to select a point a in $A \cap B_{\frac{1}{3}r}(x)$. Next let s be any rational number such that $\frac{1}{3}r < s < \frac{2}{3}r$. Then $x \in B_s(a) \subseteq B_r(x) \subseteq U_0$. [Proof: if $y \in B_s(a)$, then

$$d(y, x) \le d(y, a) + d(a, x) < s + \frac{1}{3}r < r\,;$$

so $y \in B_r(x)$.] This shows that $B_s(a)$ belongs to $\mathfrak{B}$ and that $x \in U_{B_s(a)} \in \mathfrak{V}$. Thus $\mathfrak{V}$ covers M. Now enumerate the members of $\mathfrak{V}$ as a sequence $(V_1, V_2, V_3, \dots)$ and let $W_n = \cup_{k=1}^{n} V_k$ for each $n \in \mathbb{N}$. To complete the proof it suffices to find an index n such that $W_n = M$. Assume there is no such n. Then for every k we may choose a point x_k in ${W_k}^c$. The sequence (x_k) has, by hypothesis, a convergent subsequence (x_{n_k}). Let b be the limit of this sequence. Then for some m in $\mathbb{N}$ we have $b \in V_m \subseteq W_m$. Thus W_m is an open set which contains b but only finitely many of the points x_{n_k}. (W_m contains at most the points $x_1, \dots, x_{m-1}$.) Since $x_{n_k} \to b$, this is not possible.

Q.16.3. (Solution to 16.4.1) A compact subset of *any* metric space is closed and bounded (by problem 15.1.5). It is the converse we are concerned with here.

Let A be a closed and bounded subset of $\mathbb{R}^n$. Since it is bounded, there exist closed bounded intervals $J_1, \dots, J_n$ in $\mathbb{R}$ such that

$$A \subseteq J \equiv J_1 \times \cdots \times J_n\,.$$

Each J_k is compact by example 6.3.5. Their product J is compact by corollary 16.3.2. Since J is a compact subset of $\mathbb{R}^n$ under the product metric, it is a compact subset of $\mathbb{R}^n$ under its usual Euclidean metric (see proposition 11.2.3 and the remarks preceding it). Since A is a closed subset of J, it is compact by proposition 15.1.3.

Q.17. Exercises in chapter 17

Q.17.1. (Solution to 17.1.6) Suppose there exists a nonempty set U which is properly contained in M and which is both open and closed. Then, clearly, the open sets U and U^c disconnect M. Conversely, suppose that the space M is disconnected by sets U and V. Then the set U is not the null set, is not equal to M (because its complement V is nonempty), is open, and is closed (because V is open).

Q.17.2. (Solution to 17.1.8) If N is disconnected, it can be written as the union of two disjoint nonempty sets U and V which are open in N. (These sets need not, of course, be open in M.) We show that U and V are mutually separated. It suffices to prove that $U \cap \overline{V}$ is empty, that is, that $U \subseteq \overline{V}^c$. To this end suppose that $u \in U$. Since U is open in N, there exists $\delta > 0$ such that

$$N \cap B_\delta(u) = \{x \in N : d(x, u) < \delta\} \subseteq U \subseteq V^c\,.$$

Clearly $B_\delta(u)$ is the union of two sets: $N \cap B_\delta(u)$ and $N^c \cap B_\delta(u)$. We have just shown that the first of these is contained in V^c. The second contains no points of N and therefore no points of V. Thus $B_\delta(u) \subseteq V^c$. This shows that u does not belong to the closure (in M) of the set V; so $u \in \overline{V}^c$. Since u was an arbitrary point of U, we conclude that $U \subseteq \overline{V}^c$.

Conversely, suppose that $N = U \cup V$ where U and V are nonempty sets mutually separated in M. To show that the sets U and V disconnect N, we need only show that they are open in N, since they are obviously disjoint.

We prove that U is open in N. Let $u \in U$ and notice that since $U \cap \overline{V}$ is empty, u cannot belong to $\overline{V}$. Thus there exists $\delta > 0$ such that $B_\delta(u)$ is disjoint from V. Then certainly $N \cap B_\delta(u)$ is disjoint from V. Thus $N \cap B_\delta(u)$ is contained in U. Conclusion: U is open in N.

Q.17.3. (Solution to 17.1.11) Let $G = \{(x, y): y = \sin x\}$. The function $x \mapsto (x, \sin x)$ is a continuous surjection from $\mathbb{R}$ (which is connected by proposition 5.1.9) onto $G \subseteq \mathbb{R}^2$. Thus G is connected by theorem 17.1.10.

Q.17.4. (Solution to 17.1.13) Let the metric space M be the union of a family $\mathfrak{C}$ of connected subsets of M and suppose that $\bigcap \mathfrak{C} \neq \emptyset$. Argue by contradiction. Suppose that M is disconnected by disjoint nonempty open sets U and V. Choose an element p in $\bigcap \mathfrak{C}$. Without loss of generality suppose that $p \in U$. Choose $v \in V$. There is at least one set C in $\mathfrak{C}$ such that $v \in C$. We reach a contradiction by showing that the sets $U \cap C$ and $V \cap C$ disconnect C. These sets are nonempty [p belongs to $C \cap U$ and v to $C \cap V$] and open in C. They are disjoint because U and V are, and their union is C, since

$$(U \cap C) \cup (V \cap C) = (U \cup V) \cap C = M \cap C = C\,.$$

Q.17.5. (Solution to 17.1.15) Between an arbitrary point x in the unit square and the origin there is a straight line segment, denote it by $[0, x]$. Line segments are connected because they are continuous images of (in fact, homeomorphic to) intervals in $\mathbb{R}$. The union of all the segments $[0, x]$ where x is in the unit square is the square itself. The intersection of all these segments is the origin. Thus by proposition 17.1.13 the square is connected.

Q.17.6. (Solution to 17.2.7) The set $B = \{(x, \sin x^{-1}): 0 < x \leq 1\}$ is a connected subset of $\mathbb{R}^2$ since it is the continuous image of the connected set $(0, 1]$ (see theorem 17.1.10). Then by proposition 17.1.9 the set $M := \overline{B}$ is also connected. Notice that $M = A \cup B$ where $A = \{(0, y): |y| \leq 1\}$.

To show that M is not arcwise connected, argue by contradiction. Assume that there exists a continuous function $f: [0, 1] \to M$ such that $f(0) \in A$ and $f(1) \in B$. We arrive at a contradiction by showing that the component function $f^2 = \pi_2 \circ f$ is not continuous at the point $t_0 = \sup f^{\leftarrow}(A)$.

To this end notice first that, since A is closed in M and f is continuous, the set $f^{\leftarrow}(A)$ is closed in $[0, 1]$. By example 2.2.7 the point t_0 belongs to $f^{\leftarrow}(A)$. Without loss of generality we may suppose that $f^2(t_0) \leq 0$. We need only show that for every $\delta > 0$ there exists a number $t \in [0, 1]$ such that $|t - t_0| < \delta$ and $|f^2(t) - f^2(t_0)| \geq 1$.

Let $\delta > 0$. Choose a point t_1 in $(t_0, t_0 + \delta) \cap [0, 1]$. By proposition 5.1.9 the interval $[t_0, t_1]$ is connected, so its image $\left(f^1\right)^{\rightarrow}[t_0, t_1]$ under the continuous function $f^1 = \pi_1 \circ f$ is also a connected subset of $[0, 1]$ (by theorem 17.1.10) and therefore itself an interval. Let $c = f^1(t_1)$. From $t_1 > t_0$ infer that $t_1 \in f^{\leftarrow}(B)$ and that therefore $c > 0$. Since $t_0 \in f^{\leftarrow}(A)$ it is clear that $f^1(t_0) = 0$. Thus the interval $[0, c]$ is not a single point and it is contained in $\left(f^1\right)^{\rightarrow}[t_0, t_1]$. Choose $n \in \mathbb{N}$ sufficiently large that

$$x = \frac{2}{(4n + 1)\pi} < c\,.$$

Since x belongs to $\left(f^1\right)^{\rightarrow}[t_0, t_1]$, there exists $t \in [t_0, t_1]$ such that $x = f^1(t)$. And since $x > 0$ the point $f(t)$ belongs to B. This implies that

$$f(t) = \left(f^1(t), f^2(t)\right) = (x, \sin x^{-1}) = (x, \sin(4n + 1)\tfrac{\pi}{2}) = (x, 1)\,.$$

But then (since $f^2(t_0) \leq 0$)

$$|f^2(t) - f^2(t_0)| = |1 - f^2(t_0)| \geq 1\,.$$

Q.17.7. (Solution to 17.2.8) Let A be a connected open subset of $\mathbb{R}^n$. If A is empty the result is obvious, so suppose that it is not. Choose $a \in A$. Let U be the set of all points x in A for which there exists a continuous function $f: [0, 1] \to A$ such that $f(0) = a$ and $f(1) = x$. The set U is nonempty [it contains a]. Let $V = A \setminus U$. We wish to show that V is empty. Since A is connected it suffices to show that both U and V are open.

To show that U is open let $u \in U$ and let $f: [0, 1] \to A$ be a continuous function such that $f(0) = a$ and $f(1) = u$. Since A is an open subset of $\mathbb{R}^n$ there exists $\delta > 0$ such that $B_\delta(u) \subseteq A$.

Every point b in $B_\delta(u)$ can be joined to u by the parametrized line segment $\ell\colon [0,1] \to A$ defined by

$$\ell(t) = \big((1-t)u_1 + tb_1, \dots, (1-t)u_n + tb_n\big)\,.$$

It is easy to see that since b can be joined to u and u to a, the point b can be joined to a. [Proof: If $f\colon [0,1] \to A$ and $\ell\colon [0,1] \to A$ are continuous functions satisfying $f(0) = a$, $f(1) = u$, $\ell(0) = u$, and $\ell(1) = b$, then the function $g\colon [0,1] \to A$ defined by

$$g(t) = \begin{cases} f(2t), & \text{for } 0 \le t \le \frac{1}{2} \\ \ell(2t-1), & \text{for } \frac{1}{2} < t \le 1 \end{cases}$$

is continuous, $g(0) = a$, and $g(1) = b$.] This shows that $B_\delta(u) \subseteq U$.

To see that V is open let $v \in V$ and choose $\epsilon > 0$ so that $B_\epsilon(v) \subseteq A$. If some point y in $B_\epsilon(v)$ could be joined to a by an arc in A, then v could be so joined to a (*via* y). Since this is not possible, we have that $B_\epsilon(v) \subseteq V$.

Q.18. Exercises in chapter 18

Q.18.1. (Solution to 18.1.4) Suppose (x_n) is a convergent sequence in a metric space and a is its limit. Given $\epsilon > 0$ choose $n_0 \in \mathbb{N}$ so that $d(x_n, a) < \frac{1}{2}\epsilon$ whenever $n \ge n_0$. Then $d(x_m, x_n) \le d(x_m, a) + d(a, x_n) < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon$ whenever $m, n \ge n_0$. This shows that (x_n) is Cauchy.

Q.18.2. (Solution to 18.1.5) Suppose that (x_{n_k}) is a convergent subsequence of a Cauchy sequence (x_n) and that $x_{n_k} \to a$. Given $\epsilon > 0$ choose n_0 such that $d(x_m, x_n) < \frac{1}{2}\epsilon$ whenever $m,\ n \ge n_0$. Next choose $k \in \mathbb{N}$ such that $n_k \ge n_0$ and $d\left(x_{n_k}, a\right) < \frac{1}{2}\epsilon$. Then for all $m \ge n_0$

$$d(x_m, a) \le d\left(x_m, x_{n_k}\right) + d\left(x_{n_k}, a\right) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon\,.$$

Q.18.3. (Solution to 18.1.6) A sequence in a metric space M, being a function, is said to be bounded if its range is a bounded subset of M. If (x_n) is a Cauchy sequence in M, then there exists $n_0 \in \mathbb{N}$ such that $d(x_m, x_n) < 1$ whenever $m, n \ge n_0$. For $1 \le k \le n_0 - 1$, let $d_k = d\left(x_k, x_{n_0}\right)$; and let $r = \max\{d_1, \dots, d_{n_0-1}, 1\}$. Then for every $k \in \mathbb{N}$ it is clear that x_k belongs to $C_r\left(x_{n_0}\right)$ (the closed ball about x_{n_0} of radius r). Thus the range of the sequence (x_n) is bounded.

Q.18.4. (Solution to 18.2.9) Let (M, d) and (N, ρ) be complete metric spaces. Let d_1 be the usual product metric on $M \times N$ (see 12.3.3). If $((x_n, y_n))_{n=1}^{\infty}$ is a Cauchy sequence in $M \times N$, then (x_n) is a Cauchy sequence in M since

$$\begin{aligned} d(x_m, x_n) &\le d(x_m, x_n) + \rho(y_m, y_n) \\ &= d_1\left((x_m, y_m)\,,\,(x_n, y_n)\right) \to 0 \end{aligned}$$

as $m,\ n \to \infty$.

Similarly, (y_n) is a Cauchy sequence in N. Since M and N are complete, there are points a and b in M and N respectively such that $x_n \to a$ and $y_n \to b$. By proposition 12.3.4, $(x_n, y_n) \to (a, b)$. Thus $M \times N$ is complete.

Q.18.5. (Solution to 18.2.10) It suffices to show that if (M, d) is complete, then (M, ρ) is. There exist $\alpha, \beta > 0$ such that $d(x, y) \le \alpha\,\rho(x, y)$ and $\rho(x, y) \le \beta\, d(x, y)$ for all $x, y \in M$. Let (x_n) be a Cauchy sequence in (M, ρ). Then since

$$d(x_m, x_n) \le \alpha\,\rho(x_m, x_n) \to 0 \qquad \text{as } m,\ n \to \infty\,,$$

the sequence (x_n) is Cauchy in (M, d). By hypothesis (M, d) is complete; so there is a point a in M such that $x_n \to a$ in (M, d). But then $x_n \to a$ in (M, ρ) since

$$\rho(x_n, a) \le \beta\, d(x_n, a) \to 0 \qquad \text{as } n \to \infty\,.$$

This shows that (M, ρ) is complete.

Q.18.6. (Solution to 18.2.12) Let (f_n) be a Cauchy sequence in $\mathcal{B}(S, \mathbb{R})$. Since for every $x \in S$

$$|f_m(x) - f_n(x)| \le d_u(f_m, f_n) \to 0 \qquad \text{as } m, n \to \infty,$$

it is clear that $\big(f_n(x)\big)$ is a Cauchy sequence in $\mathbb{R}$ for each $x \in S$. Since $\mathbb{R}$ is complete, there exists, for each $x \in S$, a real number $g(x)$ such that $f_n(x) \to g(x)$ as $n \to \infty$. Consider the function g defined by

$$g \colon S \to \mathbb{R} \colon x \mapsto g(x).$$

We show that g is bounded and that $f_n \to g$ (unif). Given $\epsilon > 0$ choose $n_0 \in \mathbb{N}$ so that $d_u(f_m, f_n) < \epsilon$ whenever $m, n \ge n_0$. Then for each such m and n

$$|f_m(x) - f_n(x)| < \epsilon \quad \text{whenever } x \in S.$$

Take the limit as $m \to \infty$ and obtain

$$|g(x) - f_n(x)| \le \epsilon$$

for every $n \ge n_0$ and $x \in S$. This shows that $g - f_n$ is bounded and that $d_u(g, f_n) \le \epsilon$. Therefore the function

$$g = (g - f_n) + f_n$$

is bounded and $d_u(g, f_n) \to 0$ as $n \to \infty$.

Q.19. Exercises in chapter 19

Q.19.1. (Solution to 19.1.2) **19.1.2** Let $f : M \to N$ be a contraction and let $a \in M$. We show f is continuous at a. Given $\epsilon > 0$, choose $\delta = \epsilon$. If $d(x, a) < \delta$ then $d(f(x), f(a)) \le c d(x, a) \le d(x, a) < \delta = \epsilon$, where c is a contraction constant for f.

Q.19.2. (Solution to 19.1.3) If (x, y) and (u, v) are points in $\mathbb{R}^2$, then

$$\begin{aligned} d\big(f(x,y), f(u,v)\big) &= \tfrac{1}{3}\left[(u-x)^2 + (y-v)^2 + (x-y-u+v)^2\right]^{1/2} \\ &= \tfrac{1}{3}\left[2(x-u)^2 + 2(y-v)^2 - 2(x-u)(y-v)\right]^{1/2} \\ &\le \tfrac{1}{3}\left[2(x-u)^2 + 2(y-v)^2 + 2|x-u|\,|y-v|\right]^{1/2} \\ &\le \tfrac{1}{3}\left[2(x-u)^2 + 2(y-v)^2 + (x-u)^2 + (y-v)^2\right]^{1/2} \\ &= \tfrac{\sqrt{3}}{3}\left[(x-u)^2 + (y-v)^2\right]^{1/2} \\ &= \tfrac{1}{\sqrt{3}} d\big((x,y),(u,v)\big). \end{aligned}$$

Q.19.3. (Solution to 19.1.5) Let M be a complete metric space and $f \colon M \to M$ be contractive. Since f is contractive, there exists $c \in (0,1)$ such that

$$d(f(x), f(y)) \le c\, d(x, y) \tag{Q.9}$$

for all $x, y \in M$. First we establish the existence of a fixed point. Define inductively a sequence $\big(x_n\big)_{n=0}^{\infty}$ of points in M as follows: Let x_0 be an arbitrary point in M. Having chosen $x_0, \dots, x_n$ let $x_{n+1} = f(x_n)$. We show that (x_n) is Cauchy. Notice that for each $k \in \mathbb{N}$

$$d(x_k, x_{k+1}) \le c^k d(x_0, x_1). \tag{Q.10}$$

[Inductive proof: If $k = 1$, then $d(x_1, x_2) = d\,(f(x_0), f(x_1)) \le c\, d(x_0, x_1)$. Suppose that (Q.10) holds for $k = n$. Then $d\big(x_{n+1}, x_{n+2}\big) = d\big(f(x_n), f(x_{n+1})\big) \le c\, d\big(x_n, x_{n+1}\big) \le c \cdot c^n\, d\big(x_0, x_1\big) =$

$c^{n+1}d\big(x_0, x_1\big)$.) Thus whenever $m < n$

$$\begin{aligned} d\big(x_m, x_n\big) &\le \sum_{k=m}^{n-1} d\big(x_k, x_{k+1}\big) \\ &\le \sum_{k=m}^{n-1} c^k d\big(x_0, x_1\big) \\ &\le d\big(x_0, x_1\big) \sum_{k=m}^{\infty} c^k \\ &= d\big(x_0, x_1\big) \frac{c^m}{1-c}\,. \end{aligned} \tag{Q.11}$$

Since $c^m \to 0$ as $m \to \infty$, we see that

$$d(x_m, x_n) \to 0 \qquad \text{as } m,\, n \to \infty\,.$$

That is, the sequence (x_n) is Cauchy. Since M is complete there exists a point p in M such that $x_n \to p$. The point p is fixed under f since f is continuous and therefore

$$f(p) = f(\lim x_n) = \lim f(x_n) = \lim x_{n+1} = p\,.$$

Finally, to show that there is at most one point fixed by f, argue by contradiction. Assume that $f(p) = p$, $f(q) = q$, and $p \ne q$. Then

$$\begin{aligned} d(p,q) &= d\big(f(p), f(q)\big) \\ &\le c\, d(p,q) \\ &< d(p,q) \end{aligned}$$

which certainly cannot be true.

Q.19.4. (Solution to 19.1.9) (a) In inequality (19.5) of example 19.1.6 we found that $c = 0.4$ is a contraction constant for the mapping T. Thus, according to 19.1.7,

$$\begin{aligned} d_1(x_4, p) &\le d_1(x_0, x_1)\frac{c^4}{1-c} \\ &= (0.7 + 1.1)\frac{(0.4)^4}{1-0.4} \\ &= 0.0768. \end{aligned}$$

(b) The d_1 distance between x_4 and p is

$$\begin{aligned} d_1(x_4, p) &= |1.0071 - 1.0000| + |0.9987 - 1.0000| \\ &= 0.0084\,. \end{aligned}$$

(c) We wish to choose n sufficiently large that $d_1(x_n, p) \le 10^{-4}$. According to corollary 19.1.7 it suffices to find n such that

$$d_1(x_0, x_1)\frac{c^n}{1-c} \le 10^{-4}\,.$$

This is equivalent to requiring

$$(0.4)^n \le \frac{10^{-4}(0.6)}{1.8} = \tfrac{1}{3}10^{-4}\,.$$

For this, $n = 12$ suffices.

Q.19.5. (Solution to 19.2.1) Define T on $\mathcal{C}([0,1],\mathbb{R})$ as in the hint. The space $\mathcal{C}([0,1],\mathbb{R})$ is a complete metric space by example 18.2.13. To see that T is contractive, notice that for $f, g \in \mathcal{C}([0,1],\mathbb{R})$ and $0 \le x \le 1$

$$\begin{aligned}
|Tf(x) - Tg(x)| &= \left| \int_0^x t^2 f(t)\,dt - \int_0^x t^2 g(t)\,dt \right| \\
&= \left| \int_0^x t^2 (f(t) - g(t))\,dt \right| \\
&\le \int_0^x t^2 |f(t) - g(t)|\,dt \\
&\le d_u(f,g) \int_0^x t^2\,dt \\
&= \tfrac{1}{3} x^3\, d_u(f,g)\,.
\end{aligned}$$

Thus

$$\begin{aligned}
d_u(Tf,Tg) &= \sup\{|Tf(x) - Tg(x)| : 0 \le x \le 1\} \\
&\le \tfrac{1}{3}\, d_u(f,g)\,.
\end{aligned}$$

This shows that T is contractive.

Theorem 19.1.5 tells us that the mapping T has a unique fixed point in $\mathcal{C}([0,1],\mathbb{R})$. That is, there is a unique continuous function on $[0,1]$ which satisfies (19.6).

To find this function we start, for convenience, with the function $g_0 = 0$ and let $g_{n+1} = Tg_n$ for all $n \ge 0$. Compute g_1, g_2, g_3, and g_4.

$$\begin{aligned}
g_1(x) &= Tg_0(x) = \tfrac{1}{3}x^3, \\
g_2(x) &= Tg_1(x) \\
&= \tfrac{1}{3}x^3 + \int_0^x t^2 \left(\tfrac{1}{3}t^3\right)\,dt \\
&= \tfrac{1}{3}x^3 + \tfrac{1}{3\cdot 6}x^6\,, \\
g_3(x) &= Tg_2(x) \\
&= \tfrac{1}{3}x^3 + \int_0^x t^2 \left(\tfrac{1}{3}t^3 + \tfrac{1}{3\cdot 6}x^6\right)\,dt \\
&= \tfrac{1}{3}x^3 + \tfrac{1}{3\cdot 6}x^6 + \tfrac{1}{3\cdot 6\cdot 9}x^9\,,
\end{aligned}$$

It should now be clear that for every $n \in \mathbb{N}$

$$g_n(x) = \sum_{k=1}^{n} \frac{1}{3^k\,k!}x^{3k} = \sum_{k=1}^{n} \frac{1}{k!}\left(\frac{x^3}{3}\right)^k$$

and that the uniform limit of the sequence (g_n) is the function f represented by the power series

$$\sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{x^3}{3}\right)^k.$$

Recall from elementary calculus that the power series expansion for e^y (also written $\exp(y)$) is

$$\sum_{k=0}^{\infty} \frac{1}{k!}y^k$$

for all y in $\mathbb{R}$; that is,

$$\sum_{k=1}^{\infty} \frac{1}{k!} y^k = e^y - 1\,.$$

Thus

$$\begin{aligned} f(x) &= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\frac{x^3}{3}\right)^k \\ &= \exp\left(\tfrac{1}{3}x^3\right) - 1\,. \end{aligned}$$

Finally we check that this function satisfies (19.6) for all x in $\mathbb{R}$.

$$\begin{aligned} \tfrac{1}{3}x^3 + \int_0^x t^2 f(t)\,dt &= \tfrac{1}{3}x^3 + \int_0^x t^2 \left(\exp\left(\tfrac{1}{3}t^3\right) - 1\right)\,dt \\ &= \tfrac{1}{3}x^3 + \left(\exp\left(\tfrac{1}{3}t^3\right) - \tfrac{1}{3}t^3\right)\Big|_0^x \\ &= \exp\left(\tfrac{1}{3}x^3\right) - 1 \\ &= f(x). \end{aligned}$$

Q.20. Exercises in chapter 20

Q.20.1. (Solution to 20.1.2) Suppose that $\mathbf{0}$ and $\mathbf{0}'$ are vectors in V such that $x + \mathbf{0} = x$ and $x + \mathbf{0}' = x$ for all $x \in V$. Then

$$\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}\,.$$

Q.20.2. (Solution to 20.1.4) The proof takes one line:

$$x = x + \mathbf{0} = x + (x + (-x)) = (x + x) + (-x) = x + (-x) = \mathbf{0}\,.$$

Q.20.3. (Solution to 20.1.5) Establish (a), (b), and (c) of the hint.

(a) Since $\alpha\mathbf{0} + \alpha\mathbf{0} = \alpha(\mathbf{0} + \mathbf{0}) = \alpha\mathbf{0}$, we conclude from 20.1.4 that $\alpha\mathbf{0} = \mathbf{0}$.

(b) Use the same technique as in (a): since $0x + 0x = (0 + 0)x = 0x$, we deduce from 20.1.4 that $0x = \mathbf{0}$.

(c) We suppose that $\alpha \neq 0$ and that $\alpha x = \mathbf{0}$. We prove that $x = \mathbf{0}$. Since the real number α is not zero, its reciprocal a^{-1} exists. Then

$$x = 1 \cdot x = \left(\alpha^{-1}\alpha\right)x = \alpha^{-1}(\alpha x) = \alpha^{-1}\mathbf{0} = \mathbf{0}\,.$$

(The last equality uses part (a).)

Q.20.4. (Solution to 20.1.6) Notice that $(-x) + x = x + (-x) = \mathbf{0}$. According to 20.1.3, the vector x must be the (unique) additive inverse of $-x$. That is, $x = -(-x)$.

Q.20.5. (Solution to 20.1.13) The set W is closed under addition and scalar multiplication; so vector space axioms (1) and (4) through (8) hold in W because they do in V. Choose an arbitrary vector x in W. Then using (c) we see that the zero vector $\mathbf{0}$ of V belongs to W because it is just the result of multiplying x by the scalar 0 (see exercise 20.1.5). To show that (3) holds we need only verify that if $x \in W$, then its additive inverse $-x$ in V also belongs to W. But this is clear from problem 20.1.7 since the vector $-x$ is obtained by multiplying x by the scalar -1.

Q.20.6. (Solution to 20.1.19) Use proposition 20.1.13.

(a) The zero vector belongs to every member of $\mathfrak{S}$ and thus to $\bigcap\mathfrak{S}$. Therefore $\bigcap\mathfrak{S} \neq \emptyset$.

(b) Let x, $y \in \bigcap\mathfrak{S}$. Then x, $y \in S$ for every $S \in \mathfrak{S}$. Since each member of $\mathfrak{S}$ is a subspace, $x + y$ belongs to S for every $S \in \mathfrak{S}$. Thus $x + y \in \bigcap\mathfrak{S}$.

(c) Let $x \in \bigcap\mathfrak{S}$ and $\alpha \in \mathbb{R}$. Then $x \in S$ for every $S \in \mathfrak{S}$. Since each member of $\mathfrak{S}$ is closed under scalar multiplication, αx belongs to S for every $S \in \mathfrak{S}$. Thus $\alpha x \in \bigcap\mathfrak{S}$.

Q.20.7. (Solution to 20.2.2) We wish to find scalars α, β, γ, and δ such that

$$\alpha(1,0,0)+\beta(1,0,1)+\gamma(1,1,1)+\delta(1,1,0)=(0,0,0)\,.$$

This equation is equivalent to

$$(\alpha+\beta+\gamma+\delta,\gamma+\delta,\beta+\gamma)=(0,0,0)\,.$$

Thus we wish to find a (not necessarily unique) solution to the system of equations:

$$\begin{aligned}\alpha{+}\beta{+}\gamma{+}\delta &= 0\\ {+}\gamma{+}\delta &= 0\\ \beta{+}\gamma{+}\quad &= 0\end{aligned}$$

One solution is $\alpha=\gamma=1$, $\beta=\delta=-1$.

Q.20.8. (Solution to 20.3.2) We must find scalars α, β, $\gamma\geq 0$ such that $\alpha+\beta+\gamma=1$ and

$$\alpha(1,0)+\beta(0,1)+\gamma(3,0)=(2,1/4)\,.$$

This last vector equation is equivalent to the system of scalar equations

$$\begin{cases}\alpha+3\gamma=2\\ \qquad\ \beta=\frac{1}{4}\,.\end{cases}$$

From $\alpha+\frac{1}{4}+\gamma=1$ and $\alpha+3\gamma=2$, we conclude that $\alpha=\frac{1}{8}$ and $\gamma=\frac{5}{8}$.

Q.20.9. (Solution to 20.3.10) In order to say that the intersection of the family of all convex sets which contain A is the "smallest convex set containing A", we must know that this intersection is indeed a convex set. This is an immediate consequence of the fact that the intersection of any family of convex sets is convex. (Proof. Let $\mathfrak{A}$ be a family of convex subsets of a vector space and let $x,y\in\bigcap\mathfrak{A}$. Then $x,y\in A$ for every $A\in\mathfrak{A}$. Since each A in $\mathfrak{A}$ is convex, the segment $[x,y]$ belongs to A for every A. Thus $[x,y]\subseteq\bigcap\mathfrak{A}$.)

Q.21. Exercises in chapter 21

Q.21.1. (Solution to 21.1.2) If x, $y\in\mathbb{R}^3$ and $\alpha\in\mathbb{R}$, then

$$\begin{aligned}T(x+y)&=T(x_1+y_1,x_2+y_2,x_3+y_3)\\&=(x_1+y_1+x_3+y_3,x_1+y_1-2x_2-2y_2)\\&=(x_1+x_3,x_1-2x_2)+(y_1+y_3,y_1-2y_2)\\&=Tx+Ty\end{aligned}$$

and

$$\begin{aligned}T(\alpha x)&=T(\alpha x_1,\alpha x_2,\alpha x_3)\\&=(\alpha x_1+\alpha x_3,\alpha x_1-2\alpha x_2)\\&=\alpha(x_1+x_3,x_1-2x_2)\\&=\alpha Tx.\end{aligned}$$

Q.21.2. (Solution to 21.1.4) Write $(2,1,5)$ as $2e^1+e^2+5e^3$. Use the linearity of T to see that

$$\begin{aligned}T(2,1,5)&=T(2e^1+e^2+5e^3)\\&=2Te^1+Te^2+5Te^3\\&=2(1,0,1)+(0,2,-1)+5(-4,-1,3)\\&=(2,0,2)+(0,2,-1)+(-20,-5,15)\\&=(-18,-3,16).\end{aligned}$$

Q.21.3. (Solution to 21.1.6) (a) Let x be any vector in V; then (by proposition 20.1.5) $0\,x = \mathbf{0}$. Thus $T(\mathbf{0}) = T(0x) = 0\,Tx = \mathbf{0}$.

(b) By proposition 20.1.7

$$\begin{aligned} T(x-y) &= T(x+(-y)) \\ &= T(x+(-1)y) \\ &= Tx + (-1)Ty \\ &= Tx - Ty. \end{aligned}$$

Q.21.4. (Solution to 21.1.13) First we determine where T takes an arbitrary vector (x,y,z) in its domain.

$$\begin{aligned} T(x,y,z) &= T(xe^1 + ye^2 + ze^3) \\ &= xTe^1 + yTe^2 + zTe^3 \\ &= x(1,-2,3) + y(0,0,0) + z(-2,4,-6) \\ &= (x-2z, -2x+4z, 3x-6z). \end{aligned}$$

A vector (x,y,z) belongs to the kernel of T if and only if $T(x,y,z) = (0,0,0)$; that is, if and only if $x - 2z = 0$. (Notice that the two remaining equations, $-2x+4z = 0$ and $3x - 6z = 0$, have exactly the same solutions.) Thus the kernel of T is the set of points (x,y,z) in $\mathbb{R}^3$ such that $x = 2z$. This is a plane in $\mathbb{R}^3$ which contains the y-axis. A vector (u,v,w) belongs to the range of T if an only if there exists a vector $(x,y,z) \in \mathbb{R}^3$ such that $(u,v,w) = T(x,y,z)$. This happens if an only if

$$(u,v,w) = (x-2z, -2x+4z, 3x-6z)\,;$$

that is, if an only if

$$\begin{aligned} u &= x - 2z \\ v &= -2x + 4z = -2u \\ w &= 3x - 6z = 3u. \end{aligned}$$

Consequently, only points of the form $(u, -2u, 3u) = u(1,-2,3)$ belong to $\operatorname{ran} T$. Thus the range of T is the straight line in $\mathbb{R}^3$ through the origin which contains the point $(1,-2,3)$.

Q.21.5. (Solution to 21.1.17) According to proposition 20.1.13 we must show that $\operatorname{ran} T$ is nonempty and that it is closed under addition and scalar multiplication. That it is nonempty is clear from proposition 21.1.6(a): $0 = T0 \in \operatorname{ran} T$. Suppose that $u, v \in \operatorname{ran} T$. Then there exist $x, y \in V$ such that $u = Tx$ and $v = Ty$. Thus

$$u + v = Tx + Ty = T(x+y)\,;$$

so $u+v$ belongs to $\operatorname{ran} T$. This shows that $\operatorname{ran} T$ is closed under addition. Finally, to show that it is closed under scalar multiplication let $u \in \operatorname{ran} T$ and $\alpha \in \mathbb{R}$. There exists $x \in V$ such that $u = Tx$; so

$$\alpha u = \alpha Tx = T(\alpha x)$$

which shows that αu belongs to $\operatorname{ran} T$.

Q.21.6. (Solution to 21.2.1) Recall from example 20.1.11 that under pointwise operations of addition and scalar multiplication $\mathcal{F}(V,W)$ is a vector space. To prove that $\mathfrak{L}(V,W)$ is a vector space it suffices to show that it is a vector subspace of $\mathcal{F}(V,W)$. This may be accomplished by invoking proposition 20.1.13, according to which we need only verify that $\mathfrak{L}(V,W)$ is nonempty and is closed under addition and scalar multiplication. Since the zero transformation (the one which takes every x in V to the zero vector in W) is certainly linear, $\mathfrak{L}(V,W)$ is not empty. To prove that it is closed under addition we verify that the sum of two linear transformations is itself linear. To this end let S and T be members of $\mathfrak{L}(V, W)$. Then for all x and y in V

$$\begin{aligned}(S+T)(x+y) &= S(x+y)+T(x+y) \\ &= Sx+Sy+Tx+Ty \\ &= (S+T)x+(S+T)y.\end{aligned} \tag{Q.12}$$

(It is important to be cognizant of the reason for each of these steps. There is no "distributive law" ate work here. The first and last use the definition of addition as a pointwise operation, while the middle one uses the linearity of S and T.) Similarly, for all x in V and α in $\mathbb{R}$

$$\begin{aligned}(S+T)(\alpha x) &= S(\alpha x)+T(\alpha x) \\ &= \alpha Sx+\alpha Tx \\ &= \alpha(Sx+Tx) \\ &= \alpha(S+T)x.\end{aligned} \tag{Q.13}$$

Equations (Q.12) and (Q.13) show that $S+T$ is linear and therefore belongs to $\mathfrak{L}(V, W)$.

We must also prove that $\mathfrak{L}(V, W)$ is closed under scalar multiplication. Let $T \in \mathfrak{L}(V, W)$ and $\alpha \in \mathbb{R}$, and show that the function αT is linear. For all $x, y \in V$

$$\begin{aligned}(\alpha T)(x+y) &= \alpha(T(x+y)) \\ &= \alpha(Tx+Ty) \\ &= \alpha(Tx)+\alpha(Ty) \\ &= (\alpha T)x+(\alpha T)y.\end{aligned} \tag{Q.14}$$

Finally, for all x in V and β in $\mathbb{R}$

$$\begin{aligned}(\alpha T)(\beta x) &= \alpha(T(\beta x)) \\ &= \alpha(\beta(Tx)) \\ &= (\alpha\beta)Tx \\ &= (\beta\alpha)Tx \\ &= \beta(\alpha(Tx)) \\ &= \beta((\alpha T)x)\,.\end{aligned} \tag{Q.15}$$

Equations (Q.14) and (Q.15) show that αT belongs to $\mathfrak{L}(V, W)$.

Q.21.7. (Solution to 21.2.2) Since T is bijective there exists a function $T^{-1}\colon W \to V$ satisfying $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$. We must show that this function is linear. To this end let $u, v \in W$. Then

$$\begin{aligned}T(T^{-1}(u+v)) &= I_W(u+v) \\ &= u+v \\ &= I_W(u)+I_W(v) \\ &= T(T^{-1}(u)+T(T^{-1}(v) \\ &= T(T^{-1}u+T^{-1}v).\end{aligned}$$

Since T is injective the preceding computation implies that

$$T^{-1}(u+v) = T^{-1}u+T^{-1}v.$$

Similarly, from

$$TT^{-1}(\alpha x) = \alpha x = \alpha TT^{-1}x = T(\alpha T^{-1}x)$$

we infer that

$$T^{-1}(\alpha x) = \alpha T^{-1}x.$$

Q.21.8. (Solution to 21.3.1)

$$a+b=\begin{bmatrix}5 & -3 & 3 & -2\\ 2 & -2 & 1 & 4\end{bmatrix}3a \qquad =\begin{bmatrix}12 & 6 & 0 & -3\\ -3 & -9 & 3 & 15\end{bmatrix}a-2b=\begin{bmatrix}2 & 12 & -6 & 1\\ -7 & -5 & 1 & 7\end{bmatrix}.$$

Q.21.9. (Solution to 21.3.3) $ab=\begin{bmatrix}2(1)+3(2)+(-1)1 & 2(0)+3(-1)+(-1)(-2)\\ 0(1)+1(2)+4(1) & 0(0)+1(-1)+4(-2)\end{bmatrix}=\begin{bmatrix}7 & -1\\ 6 & -9\end{bmatrix}.$

Q.21.10. (Solution to 21.3.8) $ax=(1,4,1)$.

Q.21.11. (Solution to 21.3.10(a)) To show that the vectors $a(x+y)$ and $ax+ay$ are equal show that $(a(x+y))_j$ (that is, the j^{th} component of $a(x+y)$) is equal to $(ax+ay)_j$ (the j^{th} component of $ax+ay$) for each j in $\mathbb{N}_m$. This is straight forward

$$\begin{aligned}(a(x+y))_j &= \sum_{k=1}^{n} a^j_k(x+y)_k\\ &= \sum_{k=1}^{n} a^j_k(x_k+y_k)\\ &= \sum_{k=1}^{n}(a^j_kx_k+a^j_ky_k)\\ &= \sum_{k=1}^{n} a^j_kx_k+\sum_{k=1}^{n} a^j_ky_k\\ &= (ax)_j+(ay)_j\\ &= (ax+ay)_j.\end{aligned}$$

Q.21.12. (Solution to 21.3.13)

$$\begin{aligned}xay &= \begin{bmatrix}1 & -2 & 0\end{bmatrix}\begin{bmatrix}1 & 3 & -1\\ 0 & 2 & 4\\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix}3\\ 0\\ 1\end{bmatrix}\\ &= \begin{bmatrix}1 & -2 & 0\end{bmatrix}\begin{bmatrix}2\\ 4\\ 4\end{bmatrix}=-6\,.\end{aligned}$$

Q.21.13. (Solution to 21.3.17) Suppose a is an $n\times n$-matrix with inverses b and c. Then $ab=ba=I_n$ and $ac=ca=I_n$. Thus

$$b=bI_n=b(ac)=(ba)c=I_nc=c\,.$$

Q.21.14. (Solution to 21.3.18) Multiply a and b to obtain $ab=I_3$ and $ba=I_3$. By the uniqueness of inverses (proposition 21.3.17) b is the inverse of a.

Q.21.15. (Solution to 21.4.9) Expanding the determinant of a along the first row (fact 5) we obtain

$$\begin{aligned}\det a &= \sum_{k=1}^{3} a^1_kC^1_k\\ &= 1\cdot(-1)^{1+1}\det\begin{bmatrix}3 & -1\\ -1 & 1\end{bmatrix}+0\cdot(-1)^{1+2}\det\begin{bmatrix}0 & -1\\ 1 & 1\end{bmatrix}+2\cdot(-1)^{1+3}\det\begin{bmatrix}0 & 3\\ 1 & -1\end{bmatrix}\\ &= 2+0-6=-4.\end{aligned}$$

Since $\det a \neq 0$, the matrix a is invertible. Furthermore,

$$\begin{aligned}
a^{-1} &= (\det a)^{-1} \begin{bmatrix} C^1_1 & C^1_2 & C^1_3 \\ C^2_1 & C^2_2 & C^2_3 \\ C^3_1 & C^3_2 & C^3_3 \end{bmatrix}^t \\
&= -\frac{1}{4} \begin{bmatrix} C^1_1 & C^2_1 & C^3_1 \\ C^1_2 & C^2_2 & C^3_2 \\ C^1_3 & C^2_3 & C^3_3 \end{bmatrix}^t \\
&= -\frac{1}{4} \begin{bmatrix} 2 & -2 & -6 \\ -1 & -1 & 1 \\ -3 & 1 & 3 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{3}{4} & -\frac{1}{4} & -\frac{3}{4} \end{bmatrix}.
\end{aligned}$$

Q.21.16. (Solution to 21.5.4) Since $Te^1 = T(1,0) = (1,0,2,-4)$ and $Te^2 = T(0,1) = (-3,7,1,5)$ the matrix representation of T is given by

$$[T] = \begin{bmatrix} 1 & -3 \\ 0 & 7 \\ 2 & 1 \\ -4 & 5 \end{bmatrix}.$$

Q.21.17. (Solution to 21.5.5) Let $a = [T]$. Then $a^j_k = (Te^k)_j$ for each j and k. Notice that the map

$$S \colon \mathbb{R}^n \to \mathbb{R}^m \colon x \mapsto ax$$

is linear by proposition 21.3.10 (a) and (b). We wish to show that $S = T$. According to problem 21.1.19 (b) it suffices to show that $Se^k = Te^k$ for $1 \leq k \leq n$. But this is essentially obvious: for each $j \in \mathbb{N}_m$

$$(Se^k)_j = (ae^k)_j = \sum_{l=1}^n a^j_l e^k_l = a^j_k = (Te^k)_j.$$

To prove the last assertion of the proposition, suppose that $Tx = ax$ for all x in $\mathbb{R}^n$. By the first part of the proposition $[T]x = ax$ for all x in $\mathbb{R}^n$. But then proposition 21.3.11 implies that $[T] = a$.

Q.21.18. (Solution to 21.5.6) First we show that the map $T \mapsto [T]$ is surjective. Given an $m \times n$-matrix a we wish to find a linear map T such that $a = [T]$. This is easy: let $T \colon \mathbb{R}^n \to \mathbb{R}^m \colon x \mapsto ax$. By proposition 21.3.10 (a) and (b) the map T is linear. By proposition 21.5.5

$$[T]x = Tx = ax \qquad \text{for all } x \in \mathbb{R}^n.$$

Then proposition 21.3.11 tells us that $[T] = a$.

Next we show that the map $T \mapsto [T]$ is injective. If $[T] = [S]$, then by proposition 21.5.5

$$Tx = [T]x = [S]x = Sx \qquad \text{for all } x \in \mathbb{R}^n.$$

This shows that $T = S$.

Q.21.19. (Solution to 21.5.7) By proposition 21.3.11 it suffices to show that $[S+T]x = ([S]+[T])x$ for all x in $\mathbb{R}^n$. By proposition 21.5.5

$$[S+T]x = (S+T)x = Sx + Tx = [S]x + [T]x = ([S]+[T])x\,.$$

The last step uses proposition 21.3.10(c).

(b) Show that $[\alpha T]x = (\alpha[T])x$ for all x in $\mathbb{R}^n$.

$$[\alpha T]x = (\alpha T)x = \alpha(Tx) = \alpha([T]x) = (\alpha[T])x\,.$$

The last step uses proposition 21.3.10(d).

Q.22. Exercises in chapter 22

Q.22.1. (Solution to 22.1.6) Here of course we use the usual norm on $\mathbb{R}^4$.

$$\begin{aligned}\|f(a+\lambda h)\| &= \|f\big((4,2,-4)+(-\tfrac{1}{2})(2,4,-4)\big)\| \\ &= \|f(3,0,-2)\| \\ &= \|(-6,9,3,-3\sqrt{2})\| \\ &= 3\big[(-2)^2+3^2+1^2+(-\sqrt{2})^2\big]^{1/2} \\ &= 12\end{aligned}$$

Q.22.2. (Solution to 22.1.7) Since

$$\begin{aligned}f(a+h)-f(a)-mh &= f(1+h_1,h_2,-2+h_3)-f(1,0,-2)-\begin{bmatrix}6&0&0\\0&1&-1\end{bmatrix}\begin{bmatrix}h_1\\h_2\\h_3\end{bmatrix} \\ &= \big(3(1+2h_1+{h_1}^2),h_2+h_1h_2+2-h_3\big)-(3,2)-(6h_1,h_2-h_3) \\ &= (3{h_1}^2,h_1h_2),\end{aligned}$$

we have

$$\|f(a+h)-f(a)-mh\| = (9{h_1}^4+{h_1}^2{h_2}^2)^{1/2} = |h_1|(9{h_1}^2+{h_2}^2)^{1/2}.$$

Q.22.3. (Solution to 22.1.10) Using the second derivative test from beginning calculus (and checking the value of $f+g$ at the endpoints of the interval) we see that $f(x)+g(x)$ assumes a maximum value of $\sqrt{2}$ at $x=\pi/4$ and a minimum value of $-\sqrt{2}$ at $x=5\pi/4$. So

$$\|f+g\|_u = \sup\{|f(x)+g(x)|\colon 0\le x\le 2\pi\} = \sqrt{2}.$$

Q.22.4. (Solution to 22.1.12) Let $x\in V$. Then

(a) $\|\mathbf{0}\| = \|0\cdot x\| = |0|\,\|x\| = 0$.
(b) $\|-x\| = \|(-1)x\| = |-1|\,\|x\| = \|x\|$.
(c) $0 = \|\mathbf{0}\| = \|x+(-x)\| \le \|x\|+\|-x\| = 2\|x\|$.

Q.22.5. (Solution to 22.2.2(a)) By proposition 22.1.12(b) it is clear that $\|x\|<r$ if and only if $\|-x\|<r$. That is, $d(x,\mathbf{0})<r$ if and only $d(-x,\mathbf{0})<r$. Therefore, $x\in B_r(\mathbf{0})$ if and only if $-x\in B_r(\mathbf{0})$, which in turn holds if and only if $x=-(-x)\in -B_r(\mathbf{0})$.

Q.22.6. (Solution to 22.3.2) Suppose $\|\quad\|_1$ and $\|\quad\|_2$ are equivalent norms on V. Let d_1 and d_2 be the metrics induced by $\|\quad\|_1$ and $\|\quad\|_2$, respectively. (That is, $d_1(x,y)=\|x-y\|_1$ and $d_2(x,y)=\|x-y\|_2$ for all x, $y\in V$.) Then d_1 and d_2 are equivalent metrics. Thus there exist $\alpha,\beta>0$ such that $d_1(x,y)\le\alpha d_2(x,y)$ and $d_2(x,y)\le\beta d_1(x,y)$ for all x, $y\in V$. Then in particular

$$\|x\|_1 = \|x-\mathbf{0}\|_1 = d_1(x,\mathbf{0}) \le \alpha d_2(x,\mathbf{0}) = \alpha\|x-\mathbf{0}\|_2 = \alpha\|x\|_2$$

and similarly

$$\|x\|_2 = d_2(x,\mathbf{0}) \le \beta d_1(x,\mathbf{0}) = \beta\|x\|_1$$

for all x in V. Conversely, suppose there exist α, $\beta>0$ such that $\|x\|_1\le\alpha\|x\|_2$ and $\|x\|_2\le\beta\|x\|_1$ for all x in V. Then for all x, $y\in V$

$$d_1(x,y) = \|x-y\|_1 \le \alpha\|x-y\|_2 = \alpha d_2(x,y)$$

and similarly

$$d_2(x,y) \le \beta d_1(x,y).$$

Thus d_1 and d_2 are equivalent metrics.

Q.22.7. (Solution to 22.3.5) Let $f\colon \mathbb{R}\times V \to V\colon (\beta,x)\mapsto \beta x$. We show that f is continuous at an arbitrary point (α,a) in $\mathbb{R}\times V$. Given $\epsilon>0$ let M be any number larger than both $|\alpha|$ and $\|a\|+1$. Choose $\delta=\min\{1,\epsilon/M\}$. Notice that

$$\begin{aligned}|\beta-\alpha|+\|x-a\| &= \|(\beta-\alpha,x-a)\|_1\\ &= \|(\beta,x)-(\alpha,a)\|_1.\end{aligned}$$

Thus whenever $\|(\beta,x)-(\alpha,a)\|_1<\delta$ we have

$$\|x\|\le\|a\|+\|x-a\|<\|a\|+\delta\le\|a\|+1\le M$$

so that

$$\begin{aligned}\|f(\beta,x)-f(\alpha,a)\| &= \|\beta x-\alpha a\|\\ &\le \|\beta x-\alpha x\|+\|\alpha x-\alpha a\|\\ &= |\beta-\alpha|\,\|x\|+|\alpha|\,\|x-a\|\\ &\le M(|\beta-\alpha|+\|x-a\|)\\ &< M\,\delta\\ &\le \epsilon.\end{aligned}$$

Q.22.8. (Solution to 22.3.6) If $\beta_n\to\alpha$ in $\mathbb{R}$ and $x_n\to a$ in V, then $(\beta_n,x_n)\to(\alpha,a)$ in $\mathbb{R}\times V$ by proposition 12.3.4). According to the previous proposition

$$f\colon \mathbb{R}\times V\to V\colon(\beta,x)\mapsto\beta x$$

is continuous. Thus it follows immediately from proposition 14.1.26 that

$$\beta_n x_n = f(\beta_n,x_n)\to f(\alpha,a)=\alpha a\,.$$

Q.22.9. (Solution to 22.4.2) We know from example 20.1.11 that $\mathcal{F}(S,V)$ is a vector space. We show that $\mathcal{B}(S,V)$ is a vector space by showing that it is a subspace of $\mathcal{F}(S,V)$. Since $\mathcal{B}(S,V)$ is nonempty (it contains every constant function), we need only verify that $f+g$ and αf are bounded when f, $g\in\mathcal{B}(S,V)$ and $\alpha\in\mathbb{R}$. There exist constants M, $N>0$ such that $\|f(x)\|\le M$ and $\|g(x)\|\le N$ for all x in S. But then

$$\|(f+g)(x)\|\le\|f(x)+g(x)\|\le\|f(x)\|+\|g(x)\|\le M+N$$

and

$$\|(\alpha f)(x)\|=\|\alpha f(x)\|=|\alpha|\,\|f(x)\|\le|\alpha|\,M\,.$$

Thus the functions $f+g$ and αf are bounded.

Q.23. Exercises in chapter 23

Q.23.1. (Solution to 23.1.4)

(a) $\Longrightarrow$ (b): Obvious.

(b) $\Longrightarrow$ (c): Suppose T is continuous at a point a in V. Given $\epsilon>0$ choose $\delta>0$ so that $\|Tx-Ta\|<\epsilon$ whenever $\|x-a\|<\delta$. If $\|h-\mathbf{0}\|=\|h\|<\delta$ then $\|(a+h)-a\|<\delta$ and $\|Th-T\mathbf{0}\|=\|Th\|=\|T(a+h)-Ta\|<\epsilon$. Thus T is continuous at $\mathbf{0}$.

(c) $\Longrightarrow$ (d): Argue by contradiction. Assume that T is continuous at $\mathbf{0}$ but is not bounded. Then for each $n\in\mathbb{N}$ there is a vector x_n in V such that $\|Tx_n\|>n\|x_n\|$. Let $y_n=\big(n\|x_n\|\big)^{-1}x_n$. Then $\|y_n\|=n^{-1}$; so $y_n\to\mathbf{0}$. Since T is continuous at $\mathbf{0}$ we conclude that $Ty_n\to\mathbf{0}$ as $n\to\infty$. On the other hand

$$\|Ty_n\|=\big(n\|x_n\|\big)^{-1}\|Tx_n\|>1$$

for every n. This shows that $Ty_n\not\to\mathbf{0}$ as $n\to\infty$, which contradicts the preceding assertion.

(d) $\implies$ (a): If T is bounded, there exists $M > 0$ such that $\|Tx\| \le M\|x\|$ for all x in V. It is easy to see that T is continuous at an arbitrary point a in V. Given $\epsilon > 0$ choose $\delta = \epsilon/M$. If $\|x - a\| < \delta$, then

$$\|Tx - Ta\| = \|T(x-a)\| \le M\|x - a\| < M\delta = \epsilon\,.$$

Q.23.2. (Solution to 23.1.6) The first equality is an easy computation.

$$\begin{aligned}\sup\{\|x\|^{-1}\,\|Tx\|\colon x \neq \mathbf{0}\} &= \inf\{M > 0\colon M \ge \|x\|^{-1}\|Tx\| \text{for all } x \neq \mathbf{0}\}\\ &= \inf\{M > 0\colon \|Tx\| \le M\|x\| \text{for all } x\}\\ &= \|T\|\end{aligned}$$

The second is even easier.

$$\begin{aligned}\sup\{\|x\|^{-1}\,\|Tx\|\colon x \neq \mathbf{0}\} &= \sup\{\|T\big(\|x\|^{-1}x\big)\|\}\\ &= \sup\{\|Tu\|\colon \|u\| = 1\}.\end{aligned}$$

To obtain the last equality notice that since

$$\{\|Tu\|\colon \|u\| = 1\} \subseteq \{\|Tx\|\colon \|x\| \le 1\}$$

it is obvious that

$$\sup\{\|Tu\|\colon \|u\| = 1\} \le \sup\{\|Tx\|\colon \|x\| \le 1\}\,.$$

On the other hand, if $\|x\| \le 1$ and $x \neq 0$, then $v := \|x\|^{-1}x$ is a unit vector, and so

$$\begin{aligned}\|Tx\| &\le \|x\|^{-1}\|Tx\|\\ &= \|Tv\|\\ &\le \sup\{\|Tu\|\colon \|u\| = 1\}.\end{aligned}$$

Therefore

$$\sup\{\|Tx\|\colon \|x\| \le 1\} \le \sup\{\|Tu\|\colon \|u\| = 1\}\,.$$

Q.23.3. (Solution to 23.1.11)

(a) Let $I\colon V \to V\colon x \mapsto x$. Then (by lemma 23.1.6)

$$\|I\| = \sup\{\|Ix\|\colon \|x\| = 1\} = \sup\{\|x\|\colon \|x\| = 1\} = 1\,.$$

(b) Let $\widehat{\mathbf{0}}\colon V \to W\colon x \mapsto \mathbf{0}$. Then $\|\widehat{\mathbf{0}}\| = \sup\{\|\widehat{\mathbf{0}}x\|\colon \|x\| = 1\} = \sup\{0\} = 0$.

(c) We suppose $k = 1$. (The case $k = 2$ is similar.) Let x be a nonzero vector in V_1 and $u = \|x\|^{-1}x$. Since $\|(u, \mathbf{0})\|_1 = \|u\| + \|\mathbf{0}\| = \|u\| = 1$, we see (from lemma 23.1.6) that

$$\begin{aligned}\|\pi_1\| &= \sup\{\|\pi_1(x_1, x_2)\|\colon \|(x_1, x_2)\| = 1\}\\ &\ge \|\pi_1(u, \mathbf{0})\|\\ &= \|u\|\\ &= 1.\end{aligned}$$

On the other hand since $\|\pi_1(x_1, x_2)\| = \|x_1\| \le \|(x_1, x_2)\|_1$ for all (x_1, x_2) in $V_1 \times V_2$, it follows from the definition of the norm of a transformation that $\|\pi_1\| \le 1$.

Q.23.4. (Solution to 23.1.12) Let f, $g \in \mathcal{C}$ and $\alpha \in \mathbb{R}$. Then

$$J(f + g) = \int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx = Jf + Jg$$

and

$$J(\alpha f) = \int_a^b \alpha f(x)\,dx = \alpha \int_a^b f(x)\,dx = \alpha Jf.$$

Thus J is linear. If $f \in \mathcal{C}$, then

$$|Jf| = |\int_a^b f(x)\,dx| \le \int_a^b |f(x)|\,dx \le \int_a^b \|f\|_u\,dx = (b - a)\|f\|_u\,.$$

This shows that J is bounded and that $\|J\| \le b - a$. Let $g(x) = 1$ for all x in $[a, b]$. Then g is a unit vector in $\mathcal{C}$ (that is, $\|g\|_u = 1$) and $Jg = \int_a^b g(x)\,dx = b - a$. From lemma 23.1.6 we conclude that $\|J\| \ge b - a$. This and the preceding inequality prove that $\|J\| = b - a$.

Q.23.5. (Solution to 23.1.14) It was shown in proposition 21.2.1 that $\mathfrak{L}(V, W)$ is a vector space. Since $\mathfrak{B}(V, W)$ is a nonempty subset of $\mathfrak{L}(V, W)$ [it contains the zero transformation], we need only show that sums and scalar multiples of bounded linear maps are bounded in order to establish that $\mathfrak{B}(V, W)$ is a vector space. This is done below in the process of showing that the map $T \mapsto \|T\|$ is a norm.

Let S, $T \in \mathfrak{B}(V, W)$ and $\alpha \in \mathbb{R}$. To show that $\|S + T\| \le \|S\| + \|T\|$ and $\|\alpha T\| = |\alpha|\|T\|$ we make use of the characterization $\|T\| = \sup\{\|Tu\| \colon \|u\| = 1\}$ given in lemma 23.1.6. If v is a unit vector in V, then

$$\begin{aligned}\|(S+T)v\| &= \|Sv + Tv\| \\ &\le \|Sv\| + \|Tv\| \\ &\le \sup\{\|Su\| \colon \|u\| = 1\} + \sup\{\|Tv\| \colon \|v\| = 1\} \\ &= \|S\| + \|T\|.\end{aligned}$$

This shows that $S + T$ is bounded and that

$$\begin{aligned}\|S+T\| &= \sup\{\|(S+T)v\| \colon \|v\| = 1\} \\ &\le \|S\| + \|T\|.\end{aligned}$$

Also

$$\begin{aligned}\sup\{\|\alpha Tv\| \colon \|v\| = 1\} &= |\alpha| \sup\{\|Tv\| \colon \|v\| = 1\} \\ &= |\alpha|\|T\|,\end{aligned}$$

which shows that αT is bounded and that $\|\alpha T\| = |\alpha|\|T\|$.

Finally, if $\sup\{\|x\|^{-1}\|Tx\| \colon x \ne \mathbf{0}\} = \|T\| = 0$, then $\|x\|^{-1}\|Tx\| = 0$ for all x in V, so that $Tx = \mathbf{0}$ for all x and therefore $T = \mathbf{0}$.

Q.23.6. (Solution to 23.1.15) The composite of linear maps is linear by proposition 21.1.11. From corollary 23.1.7 we have

$$\|TSx\| \le \|T\|\,\|Sx\| \le \|T\|\,\|S\|\,\|x\|$$

for all x in U. Thus TS is bounded and $\|TS\| \le \|T\|\,\|S\|$.

Q.23.7. (Solution to 23.2.1) First deal with the case $\|f\|_u \le 1$. Let (p_n) be a sequence of polynomials on $[0, 1]$ which converges uniformly to the square root function (see 15.3.5). Given $\epsilon > 0$, choose $n_0 \in \mathbb{N}$ so that $n \ge n_0$ implies $|p_n((t) - \sqrt{t}| \le \epsilon$ for all $t \in [0, 1]$. Since $\|f\|_u \le 1$

$$\big|p_n\big([f(x)]^2\big) - |f(x)|\big| = \left|p_n\big([f(x)]^2\big) - \sqrt{[f(x)]^2}\right| < \epsilon$$

whenever $x \in M$ and $n \ge n_0$. Thus $p_n \circ f^2 \to |f|$ (unif). For every $n \in \mathbb{N}$ the function $p_n \circ f^2$ belongs to A. Consequently, $|f|$ is the uniform limit of functions in A and therefore belongs to $\overline{A}$.

If $\|f\|_u > 1$ replace f in the argument above by $g = f/\|f\|_u$.

Q.23.8. (Solution to 23.2.5) Let $f \in \mathcal{C}(M, \mathbb{R})$, $a \in M$, and $\epsilon > 0$. According to proposition 23.2.4 we can choose, for each $y \ne a$ in M, a function $\phi_y \in A$ such that

$$\phi_y(a) = f(a) \qquad \text{and} \qquad \phi_y(y) = f(y)\,.$$

And for $y = a$ let ϕ_y be the constant function whose value is $f(a)$. Since in either case ϕ_y and f are continuous functions which agree at y, there exists a neighborhood U_y of y such that

$$\phi_y(x) < f(x) + \epsilon$$

for all $x \in U_y$. Clearly $\{U_y \colon y \in M\}$ covers M. Since M is compact there exist points $y_1, \dots, y_n$ in M such that the family $\{U_{y_1}, \dots, U_{y_n}\}$ covers M. Let $g = \phi_{y_1} \wedge \dots \wedge \phi_{y_n}$. By corollary 23.2.2 the function g belongs to $\overline{A}$. Now $g(a) = \phi_{y_1}(a) \wedge \cdots \wedge \phi_{y_n}(a) = f(a)$. Furthermore, given any x in M there is an index k such that $x \in U_{y_k}$. Thus

$$g(x) \leq \phi_{y_k}(x) < f(x) + \epsilon\,.$$

Q.23.9. (Solution to 23.3.6) Let (T_n) be a Cauchy sequence in the normed linear space $\mathfrak{B}(V, W)$. For each x in V

$$\|T_m x - T_n x\| \leq \|T_m - T_n\|\,\|x\| \to 0$$

as m, $n \to \infty$. Thus $(T_n x)$ is a Cauchy sequence in W for each x. Since W is complete, there exists a vector Sx in W such that $T_n x \to Sx$. Define the map

$$S\colon V \to W\colon x \mapsto Sx\,.$$

It is easy to see that S is linear: $S(x+y) = \lim T_n(x+y) = \lim(T_n x + T_n y) = \lim T_n x + \lim T_n y = Sx + Sy$; $S(\alpha x) = \lim T_n(\alpha x) = \alpha \lim T_n x = \alpha Sx$. For every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|T_m - T_n\| < \frac{1}{2}\epsilon$ whenever m, $n \geq N$. Then for all such m and n and for all x in V

$$\begin{aligned}
\|(S - T_n)x\| &= \|Sx - T_n x\| \\
&\leq \|Sx - T_m x\| + \|T_m x - T_n x\| \\
&\leq \|Sx - T_m x\| + \|T_m - T_n\|\|x\| \\
&\leq \|Sx - T_m x\| + \tfrac{1}{2}\epsilon\|x\|.
\end{aligned}$$

Taking limits as $m \to \infty$ we obtain

$$\|(S - T_n)x\| \leq \tfrac{1}{2}\epsilon\|x\|$$

for all $n \geq N$ and $x \in V$. This shows that $S - T_n$ is bounded and that $\|S - T_n\| \leq \frac{1}{2}\epsilon < \epsilon$ for $n \geq N$. Therefore the transformation

$$S = (S - T_n) + T_n$$

is bounded and

$$\|S - Tn\| \to 0$$

as $n \to \infty$. Since the Cauchy sequence (T_n) converges in the space $\mathfrak{B}(V, W)$, that space is complete.

Q.23.10. (Solution to 23.4.5) We wish to show that if $g \in W^*$, then $T^*g \in V^*$. First we check linearity: if x, $y \in V$ and $\alpha \in \mathbb{R}$, then

$$\begin{aligned}
(T^*g)(x+y) &= gT(x+y) \\
&= g(Tx + Ty) \\
&= gTx + gTy \\
&= (T^*g)(x) + (T^*g)(y)
\end{aligned}$$

and

$$(T^*g)(\alpha x) = gT(\alpha x) = g(\alpha Tx) = \alpha gTx = \alpha(T^*g)(x)\,.$$

To see that T^*g is bounded use corollary 23.1.7. For every x in V

$$|(T^*g)(x)| = |gTx| \leq \|g\|\,\|Tx\| \leq \|g\|\,\|T\|\,\|x\|\,.$$

Thus T^*g is bounded and $\|T^*g\| \leq \|T\|\,\|g\|$.

Q.24. Exercises in chapter 24

Q.24.1. (Solution to 24.1.4) Given $\epsilon > 0$ choose $\delta = \epsilon$. Assume $|x - y| < \delta$. Then

$$|f(x) - f(y)| = \left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{y-x}{xy}\right| = \left|\frac{x-y}{xy}\right| \leq |x - y| < \delta = \epsilon\,.$$

Q.24.2. (Solution to 24.1.5) We must show that

$$(\exists \epsilon > 0)(\forall \delta > 0)(\exists x, y \in (0,1])|x - y| < \delta \text{ and } |f(x) - f(y)| \geq \epsilon\,.$$

Let $\epsilon = 1$. Suppose $\delta > 0$. Let $\delta_0 = \min\{1, \delta\}$, $x = \frac{1}{2}\delta_0$, and $y = \delta_0$. Then x, $y \in (0,1]$, $|x - y| = \frac{1}{2}\delta_0 \leq \frac{1}{2}\delta < \delta$, and $|f(x) - f(y)| = |\frac{2}{\delta_0} - \frac{1}{\delta_0}| = \frac{1}{\delta\,0} \geq 1$.

Q.24.3. (Solution to 24.1.10) Since M is compact it is sequentially compact (by theorem 16.2.1). Thus the sequence (x_n) has a convergent subsequence (x_{n_k}). Let a be the limit of this subsequence. Since for each k

$$d\big(y_{n_k}, a\big) \leq d\big(y_{n_k}, x_{n_k}\big) + d\big(x_{n_k}, a\big)$$

and since both sequences on the right converge to zero, it follows that $y_{n_k} \to a$ as $k \to \infty$.

Q.24.4. (Solution to 24.1.11) Assume that f is not uniformly continuous. Then there is a number $\epsilon > 0$ such that for every n in $\mathbb{N}$ there correspond points x_n and y_n in M_1 such that $d(x_n, y_n) < 1/n$ but $d\big(f(x_n), f(y_n)\big) \geq \epsilon$. By lemma 24.1.10 there exist subsequences $\big(x_{n_k}\big)$ of (x_n) and $\big(y_{n_k}\big)$ of (y_n) both of which converge to some point a in M_1. It follows from the continuity of f that for some integer k sufficiently large, $d\big(f(x_{n_k}), f(a)\big) < \epsilon/2$ and $d\big(f(y_{n_k}), f(a)\big) < \epsilon/2$. This contradicts the assertion that $d\big(f(x_n), f(y_n)\big) \geq \epsilon$ for every n in $\mathbb{N}$.

Q.24.5. (Solution to 24.1.14) By hypothesis there exists a point a in M_1 such that $x_n \to a$ and $y_n \to a$. It is easy to see that the "interlaced" sequence $(z_n) = (x_1, y_1, x_2, y_2, x_3, y_3, \dots)$ also converges to a. By proposition 18.1.4 the sequence (z_n) is Cauchy (in M_1 and therefore) in S, and by proposition 24.1.12 (applied to the metric space S) the sequence $\big(f(z_n)\big)$ is Cauchy in M_2. The sequence $\big(f(x_n)\big)$ is a subsequence of $\big(f(z_n)\big)$ and is, by hypothesis, convergent. Therefore, according to proposition 18.1.5, the sequence $\big(f(z_n)\big)$ converges and

$$\lim f(x_n) = \lim f(z_n) = \lim f(y_n)\,.$$

Q.24.6. (Solution to 24.2.2) Just take the union of the sets of points in P and Q and put them in increasing order. Thus

$$P \vee Q = \big(0, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{5}{6}, 1\big)\,.$$

Q.24.7. (Solution to 24.2.5) (a) Either sketch the graph of σ or reduce the function algebraically to obtain

$$\sigma = -\chi_{\{2\}} - 2\chi_{(2,3)} - \chi_{\{5\}}\,.$$

Then the partition associated with σ is $P = (0, 2, 3, 5)$.

(b) $\sigma_Q = (0, 0, -2, 0, 0)$.

Q.24.8. (Solution to 24.2.7) The values of σ on the subintervals of P are given by $\sigma_P = (0, -2, 0)$. Multiply each of these by the length of the corresponding subinterval:

$$\int_0^5 \sigma = (2-0)(0) + (3-2)(-2) + (5-3)(0) = -2\,.$$

Q.24.9. (Solution to 24.2.9) Perhaps the simplest way to go about this is to observe first that we can get from the partition associated with σ to the refinement Q one point at a time. That is, there exist partitions

$$P_1 \preceq P_2 \preceq \cdots \preceq P_r = Q$$

where P_1 is the partition associated with σ and P_{j+1} contains exactly one point more than P_j (for $1 \leq j \leq r-1$).

Thus it suffices to prove the following: If σ is a step function on $[a, b]$, if $P = (s_0, \dots, s_n)$ is a refinement of the partition associated with σ, and if $P \preceq Q = (t_0, \dots, t_{n+1})$, then

$$\sum_{k=1}^{n+1} (\Delta t_k) y_k = \sum_{k=1}^{n} (\Delta s_k) x_k$$

where $\sigma_P = (x_1, \dots, x_n)$ and $\sigma_Q = (y_1, \dots, y_{n+1})$.

To prove this assertion, notice that since the partition Q contains exactly one point more than P, it must be of the form

$$Q = (s_0, \dots, s_{p-1}, u, s_p, \dots, s_n)$$

for some p such that $1 \le p \le n$. Thus

$$y_k = \begin{cases} x_k, & \text{for } 1 \le k \le p \\ x_{k-1}, & \text{for } p+1 \le k \le n+1. \end{cases}$$

Therefore,

$$\begin{aligned} \sum_{k=1}^{n+1} (\Delta t_k) y_k &= \sum_{k=1}^{p-1} (\Delta t_k) y_k + (\Delta t_p) y_p + (\Delta t_{p+1}) y_{p+1} + \sum_{k=p+2}^{n+1} (\Delta t_k) y_k \\ &= \sum_{k=1}^{p-1} (\Delta s_k) x_k + (u - s_{p-1}) x_p + (s_p - u) x_p + \sum_{k=p+2}^{n+1} (\Delta s_{k-1}) x_{k-1} \\ &= \sum_{k=1}^{p-1} (\Delta s_k) x_k + (s_p - s_{p-1}) x_p + \sum_{k=p+1}^{n} (\Delta s_k) x_k \\ &= \sum_{k=1}^{n} (\Delta s_k) x_k \end{aligned}$$

Q.24.10. (Solution to 24.2.13) That σ is an E valued step function on $[a, b]$ is obvious. Let $Q = (u_0, \dots, u_m)$ and $R = (v_0, \dots, v_n)$ be the partitions associated with τ and ρ, respectively; and suppose that $\tau_Q = (y_1, \dots, y_m)$ and $\rho_R = (z_1, \dots, z_n)$. For $1 \le k \le m + n$, let

$$t_k = \begin{cases} u_k, & \text{for } 0 \le k \le m \\ v_{k-m}, & \text{for } m+1 \le k \le m+n \end{cases}$$

and $P = (t_0, \dots, t_{m+n})$. Also define

$$x_k = \begin{cases} y_k, & \text{for } 1 \le k \le m \\ z_{k-m}, & \text{for } m+1 \le k \le m+n. \end{cases}$$

Then P is a partition of $[a,b]$ and $\sigma_P = (x_1, \dots, x_{m+n})$. Furthermore,

$$\begin{aligned}\int_a^b \sigma &= \sum_{k=1}^{m} +n(\Delta t_k)x_k \\ &= \sum_{k=1}^{m}(\Delta t_k)x_k + \sum_{k=m+1}^{m+n}(\Delta t_k)x_k \\ &= \sum_{k=1}^{m}(\Delta u_k)y_k + \sum_{k=m+1}^{m+n}(\Delta v_{k-m})z_{k-m} \\ &= \sum_{k=1}^{m}(\Delta u_k)y_k + \sum_{k=1}^{n}(\Delta v_k)z_k \\ &= \int_a^c \tau + \int_c^b \rho.\end{aligned}$$

(The third equality requires the observation that $\Delta t_{m+1} = v_1 - u_m = v_1 - c = v_1 - v_0$.)

Q.24.11. (Solution to 24.3.2) Let $f\colon [a,b] \to E$ be continuous. Given $\epsilon > 0$ we find a step function σ such that $\|f - \sigma\|_u < \epsilon$. Since the domain of f is compact, proposition 24.1.11 guarantees that f is uniformly continuous. Thus there exists $\delta > 0$ such that $\|f(u) - f(v)\| < \epsilon/2$ whenever u and v are points in $[a,b]$ such that $|u - v| < \delta$. Choose a partition $(t_0, \dots, t_n)$ of $[a,b]$ so that $t_k - t_{k-1} < \delta$ for each $k = 1, \dots, n$.

Define $\sigma\colon [a,b] \to E$ by $\sigma(s) = f(t_{k-1})$ if $t_{k-1} \le s \le t_k$ $(1 \le k \le n)$ and define $\sigma(b) = f(b)$. It is easy to see that $\|f(s) - \sigma(s)\| < \epsilon/2$ for every s in $[a,b]$. Thus $\|f - \sigma\|_u < \epsilon$; so f belongs to the closure of the family of step functions.

Q.24.12. (Solution to 24.3.6) Let $\overline{\mathcal{S}}$ be the closure of $\mathcal{S}([a,b],E)$ in the space $\mathcal{B}([a,b],E)$. If $g, h \in \overline{\mathcal{S}}$, then there exist sequences (σ_n) and (τ_n) of step functions which converge uniformly to g and h, respectively. Then $(\sigma_n + \tau_n)$ is a sequence of step functions and $\sigma_n + \tau_n \to g + h$ (unif); so $g + h \in \overline{\mathcal{S}}$. Thus

$$\begin{aligned}\int (g+h) &= \lim \int (\sigma_n + \tau_n) \\ &= \lim \left(\int \sigma_n + \int \tau_n \right) \\ &= \lim \int \sigma_n + \lim \int \tau_n \\ &= \int g + \int h.\end{aligned}$$

Similarly, if $\alpha \in \mathbb{R}$, then $(\alpha\sigma_n)$ is a sequence of step functions which converges to αg. Thus $\alpha g \in \overline{\mathcal{S}}$ and

$$\int (\alpha g) = \lim \int (\alpha \sigma_n) = \lim \left(\alpha \int \sigma_n \right) = \alpha \lim \int \sigma_n = \alpha \int g\,.$$

The map $\int\colon \overline{\mathcal{S}} \to E$ is bounded since it is both linear and uniformly continuous (see 24.1.15 and 24.1.9).

Q.24.13. (Solution to 24.3.18) The function f, being regulated, is the uniform limit in $\mathcal{B}([a,b],E)$ of a sequence (σ_n) of step functions. Since $\int \sigma_n \to \int f$ in E and T is continuous, we see that

$$T\left(\int f \right) = \lim T\left(\int \sigma_n \right)\,. \tag{Q.16}$$

By problem 24.2.14 each $T \circ \sigma_n$ is an F valued step function and

$$\int (T \circ \sigma_n) = T\left(\int \sigma_n\right) \qquad \text{for each } n. \tag{Q.17}$$

For every t in $[a, b]$

$$\begin{aligned}\|(T \circ f - T \circ \sigma_n)(t)\| & \\ &= \|T\big((f - \sigma_n)(t)\big)\| \\ &\le \|T\|\|(f - \sigma_n)(t)\| \\ &\le \|T\|\|f - \sigma_n\|_u\end{aligned}$$

so

$$\|T \circ f - T \circ \sigma_n\|_u \le \|T\|\|f - \sigma_n\|_u\,.$$

Since $\|f - \sigma_n\|_u \to 0$, we conclude that

$$T \circ \sigma_n \to T \circ f(\text{unif})$$

in $\mathcal{B}([a, b], F)$. Thus $T \circ f$ is regulated and

$$\int (T \circ f) = \lim \int (T \circ \sigma_n). \tag{Q.18}$$

The desired conclusion follows immediately from (Q.16), (Q.17), and (Q.18).

Q.25. Exercises in chapter 25

Q.25.1. (Solution to 25.1.5) Suppose that $T \in \mathfrak{B} \cap \mathfrak{o}$. Then given $\epsilon > 0$, we may choose $\delta > 0$ so that $\|Ty\| \le \epsilon\|y\|$ whenever $\|y\| < \delta$. Let x be an arbitrary unit vector. Choose $0 < t < \delta$. Then $\|tx\| = t < \delta$; so $\|Tx\| = \|T(\frac{1}{t}tx)\| = \frac{1}{t}\|T(tx)\| \le \frac{1}{t}\epsilon\|tx\| = \epsilon$. Since this last inequality holds for every unit vector x, $\|T\| \le \epsilon$. And since ϵ was arbitrary, $\|T\| = 0$. That is, $T = \mathbf{0}$.

Q.25.2. (Solution to 25.1.6) If f, $g \in \mathfrak{O}$, then there exist positive numbers M, N, δ, and η such that $\|f(x)\| \le M\|x\|$ whenever $\|x\| < \delta$ and $\|g(x)\| \le N\|x\|$ whenever $\|x\| < \eta$. Then $\|f(x) + g(x)\| \le (M + N)\|x\|$ whenever $\|x\| < \min\{\delta, \eta\}$. So $f + g \in \mathfrak{O}$.

If $\alpha \in \mathbb{R}$, then $\|\alpha f(x)\| \le |\alpha|M\|x\|$ whenever $\|x\| < \delta$; so $\alpha f \in \mathfrak{O}$.

Q.25.3. (Solution to 25.1.9) The domain of $f \circ g$ is taken to be the set of all x in V such that $g(x)$ belongs to the domain of f; that is, $\operatorname{dom}(f \circ g) = g^{\leftarrow}(\operatorname{dom} f)$. Since $f \in \mathfrak{O}$ there exist M, $\delta > 0$ such that $\|f(y)\| \le M\|y\|$ whenever $\|y\| < \delta$. Given $\epsilon > 0$, choose $\eta > 0$ so that $\|g(x)\| \le \frac{\epsilon}{M}\|x\|$ whenever $\|x\| < \eta$. If $\|x\| < \min\{\eta, \frac{M}{\epsilon}\delta\}$, then $\|g(x)\| \le \frac{\epsilon}{M}\|x\| < \delta$, so that $\|(f \circ g)(x)\|\| \le M\|g(x)\| \le \epsilon\|x\|$.

Q.25.4. (Solution to 25.1.11) Suppose $w \ne \mathbf{0}$. If $\epsilon > 0$, then there exists $\delta > 0$ such that $|\phi(x)| \le \frac{\epsilon}{\|w\|}\|x\|$ whenever $\|x\| < \delta$. Thus $\|(\phi w)(x)\| = |\phi(x)|\|w\| \le \epsilon\|x\|$ when $\|x\| < \delta$.

Q.25.5. (Solution to 25.1.12) There exist positive numbers M, N, δ, and η such that $\|\phi(x)\| \le M\|x\|$ whenever $\|x\| < \delta$ and $\|f(x)\| \le N\|x\|$ whenever $\|x\| < \eta$. Suppose that $\epsilon > 0$. If $x \in V$ and $\|x\| < \min\big\{\epsilon(MN)^{-1}, \delta, \eta\big\}$, then

$$\|(\phi f)(x)\| = |\phi(x)|\|f(x)\| \le MN\|x\|^2 \le \epsilon\|x\|\,.$$

Q.25.6. (Solution to 25.2.2) Reflexivity is obvious. Symmetry: If $f \simeq g$, then $f - g \in \mathfrak{o}$; so $g - f = (-1)(f - g) \in \mathfrak{o}$ by proposition 25.1.7. This proves $g \simeq f$. Transitivity: If $f \simeq g$ and $g \simeq h$, then both $f - g$ and $g - h$ belong to $\mathfrak{o}$; thus $f \simeq h$, since $f - h = (f - g) + (g - h) \in \mathfrak{o} + \mathfrak{o} \subseteq \mathfrak{o}$ (again by 25.1.7).

Q.25.7. (Solution to 25.2.3) By the preceding proposition $S \simeq T$. Thus $S - T \in \mathfrak{B} \cap \mathfrak{o} = \{\mathbf{0}\}$ by proposition 25.1.5.

Q.25.8. (Solution to 25.2.5) This requires only a simple computation: $\phi w - \psi w = (\phi - \psi)w \in \mathfrak{o}(V, \mathbb{R}) \cdot W \subseteq \mathfrak{o}(V, W)$ by proposition 25.1.11.

Q.25.9. (Solution to 25.3.10) The map $(x, y) \mapsto 7x - 9y$ is clearly continuous and linear. So all that needs to be verified is condition (iii) of remark 25.3:

$$\begin{aligned}
&\lim_{(x,y)\to(0,0)} \frac{\Delta f_{(1,-1)}(x, y) - (7x - 9y)}{\sqrt{x^2 + y^2}} \\
&= \lim_{(x,y)\to(0,0)} \frac{3(x + 1)^2 - (x + 1)(y - 1) + 4(y - 1)^2 - 8 - 7x + 9y}{\sqrt{x^2 + y^2}} \\
&= \lim_{(x,y)\to(0,0)} \frac{3x^2 - xy + 4y^2}{\sqrt{x^2 + y^2}} = 0. \qquad \text{(See problem 14.3.15(a).)}
\end{aligned}$$

(The equation $z = 7x - 9y$ represents a plane through the origin.)

Q.25.10. (Solution to 25.3.11) Since

$$\begin{aligned}
\Delta f_{(1,-1,0)}(h, j, k) &= f(h + 1, j - 1, k) - f(1, -1, 0) \\
&= \big((h + 1)^2(j - 1) - 7, 3(h + 1)k + 4(j - 1)\big) - (-8, -4) \\
&= \big(h^2(j - 1) + 2hj - 2h + j, 3hk + 4j + 3k\big)
\end{aligned}$$

and

$$\begin{aligned}
M(h, j, k) &= \begin{bmatrix} r & s & t \\ u & v & w \end{bmatrix} (h, j, k) \\
&= (rh + sj + tk, uh + vj + wk),
\end{aligned}$$

we find that the first coordinate of the Newton quotient

$$\frac{\Delta f_{(1,-1,0)}(h, j, k) - M(h, j, k)}{\|(h, j, k)\|}$$

turns out to be

$$\frac{h^2(j - 1) + 2hj - (2 + r)h + (1 - s)j - tk}{\sqrt{h^2 + j^2 + k^2}}.$$

If we choose $r = -2$, $s = 1$, and $t = 0$, then the preceding expression approaches zero as $(h, j, k) \to (0, 0, 0)$. (See problem 14.3.15(a).) Similarly, the second coordinate of the Newton quotient is

$$\frac{3hk - uh + (4 - v)j + (3 - w)k}{\sqrt{h^2 + j^2 + k^2}},$$

which approaches zero as $(h, j, k) \to (0, 0, 0)$ if we choose $u = 0$, $v = 4$, and $w = 3$. We conclude from the uniqueness of differentials (proposition 25.3.9) that

$$\big[df_{(1,-1,0,)}\big] = \begin{bmatrix} -2 & 1 & 0 \\ 0 & 4 & 3 \end{bmatrix}.$$

Equivalently we may write

$$df_{(1,-1,0)}(h, j, k) = (-2h + j, 4j + 3k).$$

Q.25.11. (Solution to 25.3.16) It is easy to check that $\phi(a)df_a + d\phi_a \cdot f(a)$ is bounded and linear. From our hypotheses $\Delta f_a \simeq df_a$ and $\Delta\phi_a \simeq d\phi_a$ we infer (using propositions 25.2.4 and 25.2.5) that $\phi(a)\Delta f_a \simeq \phi(a)df_a$ and that $\Delta\phi_a \cdot f(a) \simeq d\phi_a \cdot f(a)$. Then from corollary 25.3.13 and proposition 25.1.12 we conclude that $\Delta\phi_a \Delta f_a$ belongs to $\mathfrak{O}(V, \mathbb{R}) \cdot \mathfrak{O}(V, W)$ and therefore to $\mathfrak{o}(V, W)$.

That is, $\Delta\phi_a\,\Delta f_a \simeq 0$. Thus by propositions 25.3.4 and 25.2.4

$$\begin{aligned}\Delta(\phi f)_a &= \phi(a)\cdot\Delta f_a + \Delta\phi_a\cdot f(a) + \Delta\phi_a\cdot\Delta f_a\\ &\simeq \phi(a)df_a + d\phi_a\cdot f(a) + 0\\ &= \phi(a)df_a + d\phi_a\cdot f(a).\end{aligned}$$

Q.25.12. (Solution to 25.3.17) Our hypotheses are $\Delta f_a \simeq df_a$ and $\Delta g_{f(a)} \simeq dg_{f(a)}$. By proposition 25.3.12 $\Delta f_a \in \mathfrak{O}$. Then by proposition 25.2.7

$$\Delta g_{f(a)} \circ \Delta f_a \simeq dg_{f(a)} \circ \Delta f_a \tag{Q.19}$$

and by proposition 25.2.6

$$dg_{f(a)} \circ \Delta f_a \simeq dg_{f(a)} \circ df_a\,. \tag{Q.20}$$

According to proposition 25.3.5

$$\Delta(g\circ f)_a \simeq \Delta g_{f(a)} \circ \Delta f_a\,. \tag{Q.21}$$

From (Q.19), (Q.20),(Q.21), and proposition 25.2.2 it is clear that

$$\Delta(g\circ f)_a \simeq dg_{f(a)} \circ df_a\,.$$

Since $dg_{f(a)} \circ df_a$ is a bounded linear transformation, the desired conclusion is an immediate consequence of proposition 25.3.9.

Q.25.13. (Solution to 25.4.6)
(a) $Dc(\pi/3) = (-\sin(\pi/3), \cos(\pi/3)) = (-\sqrt{3}/2, 1/2)$.
(b) $l(t) = (1/2, \sqrt{3}/2) + t(-\sqrt{3}/2, 1/2) = \frac{1}{2}(1-\sqrt{3}t, \sqrt{3}+t)$.
(c) $x + \sqrt{3}y = 2$.

Q.25.14. (Solution to 25.4.7) If c is differentiable at a, then there exists a bounded linear transformation $dc_a\colon \mathbb{R} \to V$ which is tangent to Δc_a at 0. Then

$$\begin{aligned}Dc(a) &= \lim_{h\to 0}\frac{\Delta c_a(h)}{h}\\ &= \lim_{h\to 0}\frac{\Delta c_a(h) - dc_a(h) + dc_a(h)}{h}\\ &= \lim_{h\to 0}\frac{\Delta c_a(h) - dc_a(h)}{h} + \lim_{h\to 0}\frac{h\,dc_a(1)}{h}\\ &= dc_a(1).\end{aligned}$$

Q.25.15. (Solution to 25.4.12) Let $\epsilon > 0$. Since f is continuous at c and the interval J is open, we may choose $\delta > 0$ so that $c + h \in J$ and $\|\Delta f_c(h)\| < \epsilon$ whenever $|h| < \delta$. Thus if $0 < |h| < \delta$,

$$\begin{aligned}\|\Delta F_c(h) - h\,f(c)\| &= \left\|\int_a^{c+h} f - \int_a^c f - h\,f(c)\right\|\\ &= \left\|\int_c^{c+h} f(t)\,dt - \int_c^{c+h} f(c)\,dt\right\|\\ &= \left\|\int_c^{c+h} \Delta f_c(t-c)\,dt\right\|\\ &\le \left|\int_c^{c+h} \|\Delta f_c(t-c)\|\,dt\right| \qquad \text{(by 24.3.10)}\\ &< \epsilon\,|h|\,.\end{aligned}$$

It follows immediately that

$$\left\|\frac{\Delta F_c(h)}{h} - f(c)\right\| < \frac{1}{|h|}\epsilon|h| = \epsilon$$

whenever $0 < |h| < \delta$; that is,

$$DF(c) = \lim_{h\to 0} \frac{\Delta F_c(h)}{h} = f(c)\,.$$

Q.25.16. (Solution to 25.5.2) If $l(t) = a + tv$, then

$$\begin{aligned} D(f\circ l)(0) &= \lim_{t\to 0} \frac{1}{t}\Delta(f\circ l)_0(t) \\ &= \lim_{t\to 0} \frac{1}{t}((f\circ l)(0+t) - (f\circ l)(0)) \\ &= \lim_{t\to 0} \frac{1}{t}(f(a+tv) - f(a)) \\ &= \lim_{t\to 0} \frac{1}{t}\Delta f_a(tv) \\ &= D_v f(a). \end{aligned}$$

Q.25.17. (Solution to 25.5.3) Let $l(t) = a + tv$. Then

$$\begin{aligned} (f\circ l)(t) &= f(a+tv) \\ &= f((1,1) + t(\tfrac{3}{5},\tfrac{4}{5})) \\ &= f(1+\tfrac{3}{5}t, 1+\tfrac{4}{5}t) \\ &= \tfrac{1}{2}\ln\big((1+\tfrac{3}{5}t)^2 + (1+\tfrac{4}{5}t)^2\big) \\ &= \tfrac{1}{2}\ln\big(2 + \tfrac{14}{5}t + t^2\big). \end{aligned}$$

It follows that $D(f\circ l)(t) = \frac{1}{2}\big(\frac{14}{5} + 2t\big)\big(2 + \frac{14}{5}t + t^2\big)^{-1}$, so that $D_v f(a) = D(f\circ l)(0) = \frac{7}{10}$.

Q.25.18. (Solution to 25.5.5) As usual let $l(t) = a + tv$. Then

$$\begin{aligned} (\phi\circ l)(t) &= \int_0^{\pi/2} (\cos x + Da(x) + tDv(x))^2\,dx \\ &= \int_0^{\pi/2} (2\cos x + t\sin x)^2\,dx \\ &= \int_0^{\pi/2} 4\cos^2 x\,dx + 4t\int_0^{\pi/2} \sin x\cos x\,dx + t^2\int_0^{\pi/2}\sin^2 x\,dx\,. \end{aligned}$$

Differentiating we obtain

$$D(\phi\circ l)(t) = 4\int_0^{\pi/2}\sin x\cos x\,dx + 2t\int_0^{\pi/2}\sin^2 x\,dx\,;$$

so

$$D_v\phi(a) = D(\phi\circ l)(0) = 4\int_0^{\pi/2}\sin x\cos x\,dx = 2\,.$$

Q.25.19. (Solution to 25.5.9) If $l = a + tv$, then, since $l(0) = a$ and $Dl(0) = v$, we have

$$\begin{aligned} D_v f(a) &= D(f\circ l)(0) \\ &= df_{l(0)}(Dl(0)) \quad \text{(by 25.4.11)} \\ &= df_a(v). \end{aligned}$$

Q.25.20. (Solution to 25.6.1) If $x \in \operatorname{dom} f^1 \cap \operatorname{dom} f^2$, then

$$\begin{aligned} \big((j_1\circ f^1) + (j_2\circ f^2)\big)(x) &= j_1\big(f^1(x)\big) + j_2\big(f^2(x)\big) \\ &= \big(f^1(x), 0\big) + \big(0, f^2(x)\big) \\ &= \big(f^1(x), f^2(x)\big) \\ &= f(x)\,. \end{aligned}$$

Being the sum of composites of differentiable functions, f is differentiable, and

$$\begin{aligned} df_a &= d\big((j_1 \circ f^1) + (j_2 \circ f^2)\big)_a \\ &= d\big(j_1 \circ f^1\big)_a + d\big(j_2 \circ f^2\big)_a && \text{(by 25.3.15)} \\ &= d\big(j_1\big)_{f^1(a)} \circ d\big(f^1\big)_a + d\big(j_2\big)_{f^2(a)} \circ d\big(f^2\big)_a && \text{(by 25.3.17)} \\ &= j_1 \circ d\big(f^1\big)_a + j_2 \circ d\big(f^2\big)_a && \text{(by 25.3.24)} \\ &= \big(d\big(f^1\big)_a \,,\, d\big(f^2\big)_a\big)\ . \end{aligned}$$

Q.25.21. (Solution to 25.6.3) By propositions 25.4.7 and 25.4.8 a curve has a derivative at t if and only if it is differentiable at t. Thus the desired result is an immediate consequence of the following easy computation:

$$\begin{aligned} Dc(t) &= dc_t(1) \\ &= \big(d\big(c^1\big)_t(1) \,,\, d\big(c^2\big)_t(1)\big) \\ &= \big(Dc^1(t) \,,\, Dc^2(t)\big)\ . \end{aligned}$$

Q.26. Exercises in chapter 26

Q.26.1. (Solution to 26.1.5) Let $f \colon [0, 2\pi] \to \mathbb{R}^2 \colon t \mapsto (\cos t, \sin t)$. Then f is continuous on $[0, 2\pi]$ and differentiable on $(0, 2\pi)$. Notice that $f(2\pi) - f(0) = (1, 0) - (1, 0) = (0, 0)$. But $Df(t) = (-\sin t, \cos t)$. Certainly there is no number c such that $2\pi(-\sin c, \cos c) = (0, 0)$.

Q.26.2. (Solution to 26.1.6) Given $\epsilon > 0$, define $h(t) = \|f(t) - f(a)\| - (t - a)(M + \epsilon)$ for $a \le t \le b$. Since f is continuous on $[a, b]$, so is h. Let $A = h^{\leftarrow}(-\infty, \epsilon]$. The set A is nonempty (it contains a) and is bounded above (by b). By the *least upper bound axiom* J.3.1 it has a supremum, say l. Clearly $a \le l \le b$. Since h is continuous and $h(a) = 0$, there exists $\eta > 0$ such that $a \le t < a + \eta$ implies $h(t) \le \epsilon$. Thus $[a, a + \eta) \subseteq A$ and $l > a$. Notice that since h is continuous the set A is closed (proposition 14.1.13); and since $l \in \overline{A}$ (see example 2.2.7), l belongs to A.

We show that $l = b$. Assume to the contrary that $l < b$. Since f is differentiable at l, there exists $\delta > 0$ such that if $t \in (l, l + \delta)$ then $\|(t - l)^{-1}(f(t) - f(l))\| < M + \epsilon$. Choose any point t in $(l, l + \delta)$. Then

$$\begin{aligned} h(t) &= \|f(t) - f(a)\| - (t - a)(M + \epsilon) \\ &\le \|f(t) - f(l)\| + \|f(l) - f(a)\| - (t - l)(M + \epsilon) - (l - a)(M + \epsilon) \\ &< (t - l)(M + \epsilon) + h(l) - (t - l)(M + \epsilon) \\ &= h(l) \\ &\le \epsilon\,. \end{aligned}$$

This says that $t \in A$, which contradicts the fact that l is an upper bound for A. Thus $l = b$ and $h(b) \le \epsilon$. That is,

$$\|f(b) - f(a)\| \le (M + \epsilon)(b - a) + \epsilon\,.$$

Since ϵ was arbitrary,

$$\|f(b) - f(a)\| \le M(b - a)\,.$$

Q.26.3. (Solution to 26.2.1)

(a) For every x in V_k

$$(\pi_k \circ j_k)(x) = \pi_k(0, \dots, 0, x, 0, \dots, 0) = x\,.$$

(b) For every x in V

$$\begin{aligned}\sum_{k=1}^{n}(j_k \circ \pi_k)(x) &= \sum_{k=1}^{n} j_k(x_k)\\ &= (x_1, 0, \dots, 0) + \cdots + (0, \dots, 0, x_n)\\ &= (x_1, \dots, x_n)\\ &= x\,.\end{aligned}$$

Q.26.4. (Solution to 26.2.5) For every h in B_k

$$\begin{aligned}(\Delta f_a \circ j_k)(h) &= \Delta f_a(j_k(h))\\ &= f(a + j_k(h)) - f(a)\\ &= g(h) - g(\mathbf{0})\\ &= \Delta g_{\mathbf{0}}(h)\,.\end{aligned}$$

Q.26.5. (Solution to 26.2.10) Fix a point (a, b) in U. We make two observations about the notation introduced in the hint. First,

$$d(h^v)_z = d_1 f_{(a+z,b+v)}\,. \tag{Q.22}$$

[Proof: Since $f(a+z+s, b+v) = h^v(z+s) = (h^v \circ T_z)(s)$, we see that $d_1 f_{(a+z,b+v)} = d(h^v \circ T_z)_{\mathbf{0}} = d(h^v)_{T_z(\mathbf{0})} \circ d(T_z)_{\mathbf{0}} = d(h^v)_z \circ I = d(h^v)_z$.]

Second,

$$\Delta f_{(a,b)}(u, v) = \Delta(h^v)_{\mathbf{0}}(u) + \Delta g_{\mathbf{0}}(v)\,. \tag{Q.23}$$

[Proof:

$$\begin{aligned}\Delta f_{(a,b)}(u, v) &= f(a + u, b + v) - f(a, b)\\ &= f(a + u, b + v) - f(a, b + v) + f(a, b + v) - f(a, b)\\ &= h^v(u) - h^v(\mathbf{0}) + g(v) - g(\mathbf{0})\\ &= \Delta(h^v)_{\mathbf{0}}(u) + \Delta g_{\mathbf{0}}(v)\,.]\end{aligned}$$

Let $\epsilon > 0$. By hypothesis the second partial differential of f exists at (a, b). That is, the function g is differentiable at $\mathbf{0}$ and

$$\Delta g_{\mathbf{0}} \simeq dg_{\mathbf{0}} = d_2 f_{(a,b)} = T\,.$$

Thus there exists $\delta_1 > 0$ such that

$$\|\Delta g_{\mathbf{0}}(v) - Tv\| \le \epsilon\|v\| \tag{Q.24}$$

whenever $\|v\| < \delta_1$.

Since $d_1 f$ is assumed to (exist and) be continuous on U, there exists $\delta_2 > 0$ such that

$$\|d_1 f_{(a+s,b+t)} - d_1 f_{(a,b)}\| < \epsilon \tag{Q.25}$$

whenever $\|(s, t)\|_1 < \delta_2$. Suppose then that (u, v) is a point in U such that $\|(u, v)\|_1 < \delta_2$. For each z in the segment $[\mathbf{0}, u]$

$$\|(z, v)\|_1 = \|z\| + \|v\| \le \|u\| + \|v\| = \|(u, v)\|_1 < \delta_2$$

so by (Q.22) and (Q.25)

$$\|d(h^v)_z - S\| = \|d_1 f_{(a+z,b+v)} - d_1 f_{(a,b)}\| < \epsilon\,.$$

Thus according to the version of the *mean value theorem* given in corollary 26.1.8

$$\|\Delta(h^v)_{\mathbf{0}}(u) - Su\| \le \epsilon\|u\| \tag{Q.26}$$

whenever $\|(u, v)\|_1 < \delta_2$.

Now let $\delta = \min\{\delta_1, \delta_2\}$. Suppose $\|(u,v)\|_1 < \delta$. Then since $\|v\| \le \|u\| + \|v\| = \|(u,v)\|_1 < \delta \le \delta_1$ inequality (Q.24) holds, and since $\|(u,v)\|_1 < \delta \le \delta_2$ inequality (Q.26) holds. Making use of these two inequalities and (Q.23) we obtain

$$\begin{aligned}\|\Delta f_{(a,b)}(u,v) - R(u,v)\| &= \|\Delta (h^v)_{\mathbf{0}}(u) + \Delta g_{\mathbf{0}}(v) - Su - Tv\|\\ &\le \|\Delta (h^v)_{\mathbf{0}}(u) - Su\| + \|\Delta g_{\mathbf{0}}(v) - Tv\|\\ &\le \epsilon\,\|u\| + \epsilon\,\|v\|\\ &= \epsilon\,\|(u,v)\|_1\,.\end{aligned}$$

Thus $\Delta f_{(a,b)} \simeq R$ showing that f is differentiable at (a,b) and that its differential is given by

$$df_{(a,b)} = R = d_1 f_{(a,b)} \circ \pi_1 + d_2 f_{(a,b)} \circ \pi_2\,. \tag{Q.27}$$

That df is continuous is clear from (Q.27) and the hypothesis that $d_1 f$ and $d_2 f$ are continuously differentiable.

Q.26.6. (Solution to 26.2.12) First we compute df_a. A straightforward calculation gives

$$\frac{\Delta f_a(h)}{\|h\|_1} = \frac{h_1 + 2\,h_2 - 3\,h_3 + 6\,h_4 + {h_2}^2 + 2\,h_1h_2 + h_1{h_2}^2 + 3\,h_3h_4}{\|h\|_1}\,.$$

From this it is clear that the desired differential is given by

$$df_a(h) = h_1 + 2\,h_2 - 3\,h_3 + 6\,h_4$$

for then

$$\frac{\Delta f_a(h) - df_a(h)}{\|h\|_1} = \frac{{h_2}^2 + 2\,h_1h_2 + h_1{h_2}^2 + 3\,h_3h_4}{\|h\|_1} \to 0 \qquad \text{as } h \to 0$$

Note. In the preceding computation the use of the product norm $\|\quad\|_1$ for $\mathbb{R}^4$ rather than the usual Euclidean norm is both arbitrary and harmless (see problem 22.3.21.

(a) Compose df_a with the injection

$$j_1 \colon \mathbb{R} \to \mathbb{R} \times R \times \mathbb{R} \times \mathbb{R} \colon x \mapsto (x,0,0,0)\,.$$

Then

$$d_1 f_a(x) = df_a(j_1(x)) = df_a(x,0,0,0) = x$$

for all x in $\mathbb{R}$.

(b) This has exactly the same answer as part (a)—although the rationale is slightly different. The appropriate injection map is

$$j_1 \colon \mathbb{R} \to \mathbb{R}^3 \colon x \mapsto (x,\mathbf{0})$$

(where $\mathbf{0}$ is the zero vector in $\mathbb{R}^3$). We may rewrite (26.12) in this case as

$$f(x,y) = x{y_1}^2 + 3\,y_2y_3$$

for all $x \in \mathbb{R}$ and $y \in \mathbb{R}^3$. Also write $a = (b,c)$ where $b = 1 \in \mathbb{R}$ and $c = (1,2,-1) \in \mathbb{R}^3$, and write $h = (r,s)$ where $r \in \mathbb{R}$ and $s \in \mathbb{R}^3$. Then

$$df_a(h) = df_{(b,c)}(r,s) = r + 2\,s_1 - 3\,s_2 + 6\,s_3$$

so that

$$d_1 f(x) = df_a(j_1(x)) = df_a(x,\mathbf{0}) = x$$

for all x in $\mathbb{R}$.

(c) Here the appropriate injection is

$$j_1 \colon \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^2 \colon x \mapsto (x,\mathbf{0})$$

where $\mathbf{0}$ is the zero vector in $\mathbb{R}^2$. Rewrite (26.12) as

$$f(x,y) = x_1{x_2}^2 + 3\,y_1y_2$$

for all x, $y \in \mathbb{R}^2$. Let $a = (b, c)$ where $b = (1, 1)$ and $c = (2, -1)$; and let $h = (r, s)$ where r, $s \in \mathbb{R}^2$. Then

$$df_a(h) = df_{(b,c)}(r, s) = r_1 + 2\,r_2 - 3\,s_1 + 6\,s_2$$

so that

$$d_1 f_a(x) = df_a(j_1(x)) = df_a(x, \mathbf{0}) = x_1 + 2\,x_2$$

for all x in $\mathbb{R}^2$.

(d) As far as the partial differential d_1 is concerned, this is essentially the same problem as (c). However, in this case the injection j_1 is given by

$$j_1 \colon \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \colon x \mapsto (x, 0, 0)\,.$$

Equation (26.12) may be written

$$f(x, y, z) = x_1 {x_2}^2 + 3\,yz$$

for all $x \in \mathbb{R}^2$ and y, $z \in \mathbb{R}$. Let $a = (b, c, d)$ where $b = (1, 1)$, $c = 2$, and $d = -1$; and let $h = (q, r, s)$ where $q \in \mathbb{R}^2$ and r, $s \in \mathbb{R}$. Then

$$df_a(h) = df_{(b,c,d)}(q, r, s) = q_1 + 2\,q_2 - 3\,r + 6\,s$$

so that

$$d_1 f_a(x) = df_a(j_1(x)) = df_{(b,c,d)}(x, 0, 0) = x_1 + 2\,x_2\,.$$

(e) Here $j_1 \colon \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R} \colon x \mapsto (x, 0)$. Rewrite (26.12) as

$$f(x, y) = x_1 {x_2}^2 + 3\,x_3 y$$

for all $x \in \mathbb{R}^3$ and $y \in \mathbb{R}$. Let $a = (b, c)$ with $b = (1, 1, 2)$ and $c = -1$; and let $h = (r, s)$ where $r \in \mathbb{R}^3$ and $s \in \mathbb{R}$. Then

$$df_a(h) = df_{(b,c)}(r, s) = r_1 + 2\,r_2 - 3\,r_3 + 6\,s$$

so that

$$d_1 f_a(x) = df_a(j_1(x)) = df_{(b,c)}(x, 0) = x_1 + 2\,x_2 - 3\,x_3\,.$$

Q.26.7. (Solution to 26.2.17) Just do what you have always done: hold two of the variables constant and differentiate with respect to the other. (See the paragraph after equation (26.13).)

$$\begin{aligned}
f_1(x, y, z) &= (3\,x^2 y^2 \sin z, 2\,x);\; \text{ so } f_1(a) = (12, 2).\\
f_2(x, y, z) &= (2\,x^3 y \sin z, \cos z);\; \text{ so } f_2(a) = (-4, 0).\\
f_3(x, y, z) &= (x^3 y^2 \cos z, -y \sin z);\; \text{ so } f_3(a) = (0, 2).
\end{aligned}$$

Q.26.8. (Solution to 26.3.2) Let $\epsilon > 0$. Since $[a, b] \times [c, d]$ is compact, the continuous function f must be uniformly continuous (see proposition 24.1.11). Thus there exists $\delta > 0$ such that $\|(x, y) - (u, v)\|_1 < \delta$ implies $\|f(x, y) - f(u, v)\| < \epsilon\,(b - a)^{-1}$. Suppose that y and v lie in $[c, d]$ and that $|y - v| < \delta$. Then $\|(x, y) - (x, v)\|_1 < \delta$ for all x in $[a, b]$; so $\|f(x, y) - f(x, v)\| < \epsilon\,(b - a)^{-1}$

from which it follows that

$$\begin{aligned}
\|g(y) - g(v)\| &= \left\| \int_a^b f^y - \int_a^b f^v \right\| \\
&= \left\| \int_a^b (f^y - f^v) \right\| \\
&\leq \int_a^b \|f^y(x) - f^v(x)\| \, dx \\
&= \int_a^b \|f(x,y) - f(x,v)\| \, dx \\
&< \int_a^b \epsilon \, (b-a)^{-1} \, dx \\
&= \epsilon \, .
\end{aligned}$$

Thus g is uniformly continuous.

Q.26.9. (Solution to 26.3.4) Let $h(y) = \int_a^b f_2(x,y) \, dx$. By lemma 26.3.2 the function h is continuous and therefore integrable on every interval of the form $[c,z]$ where $c \leq z \leq d$. Then by proposition 26.3.3 we have

$$\begin{aligned}
\int_c^z h &= \int_c^z \int_a^b f_2(x,y) \, dx \, dy \\
&= \int_a^b \int_c^z f_2(x,y) \, dy \, dx \\
&= \int_a^b \int_c^z \frac{d}{dy}\big({}^x f(y)\big) \, dy \, dx \\
&= \int_a^b \big({}^x f(z) - {}^x f(c)\big) \, dx \\
&= \int_a^b \big(f(x,z) - f(x,c)\big) \, dx \\
&= g(z) - g(c) \, .
\end{aligned}$$

Differentiating we obtain

$$h(z) = g'(z)$$

for $c < z < d$. This shows that g is continuously differentiable on (c,d) and that

$$\begin{aligned}
\frac{d}{dy} \int_a^b f(x,y) \, dx &= g'(y) \\
&= h(y) \\
&= \int_a^b f_2(x,y) \, dx \\
&= \int_a^b \frac{\partial f}{\partial y}(x,y) \, dx \, .
\end{aligned}$$

Q.27. Exercises in chapter 27

Q.27.1. (Solution to 27.1.4) If x, y, and $z \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, then

$$\begin{aligned}\langle x, y+z\rangle &= \langle y+z, x\rangle && \text{(by (c))}\\ &= \langle y, x\rangle + \langle z, x\rangle && \text{(by (a))}\\ &= \langle x, y\rangle + \langle x, z\rangle && \text{(by (c))}\end{aligned}$$

and

$$\begin{aligned}\langle x, \alpha y\rangle &= \langle \alpha y, x\rangle && \text{(by (c))}\\ &= \alpha\langle y, x\rangle && \text{(by (b))}\\ &= \alpha\langle x, y\rangle && \text{(by (c))}\end{aligned}$$

Q.27.2. (Solution to 27.1.8) The domain of the arccosine function is the closed interval $[-1, 1]$. According to the *Schwarz inequality* $|\langle x, y\rangle| \le \|x\|\|y\|$; equivalently,

$$-1 \le \frac{\langle x, y\rangle}{\|x\|\|y\|} \le 1$$

for nonzero vectors x and y. This shows that $\langle x, y\rangle\|x\|^{-1}\|y\|^{-1}$ is in the domain of arccosine.

Q.27.3. (Solution to 27.1.10) If $x = (1, 0, 1)$ and $y = (0, -1, 1)$, then $\langle x, y\rangle = 1$ and $\|x\| = \|y\| = \sqrt{2}$. So

$$\measuredangle(x, y) = \arccos\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right) = \arccos\frac{1}{2} = \frac{\pi}{3}\,.$$

Q.27.4. (Solution to 27.2.1) The computations

$$\psi_b(x+y) = \langle x+y, b\rangle = \langle x, b\rangle + \langle y, b\rangle = \psi_b(x) + \psi_b(y)$$

and

$$\psi_b(\alpha x) = \langle \alpha x, b\rangle = \alpha\langle x, b\rangle = \alpha\psi_b(x)$$

show that ψ_b is linear. Since

$$|\psi_b(x)| = |\langle x, b\rangle| \le \|b\|\|x\|$$

for every x in $\mathbb{R}^n$, we conclude that ψ_b is bounded and that $\|\psi_b\| \le \|b\|$. On the other hand, if $b \ne \mathbf{0}$, then $\|b\|^{-1}b$ is a unit vector, and since

$$|\psi_b(\|b\|^{-1}b)| = \langle \|b\|^{-1}b, b\rangle = \|b\|^{-1}\langle b, b\rangle = \|b\|$$

we conclude (from lemma 23.1.6) that $\|\psi_b\| \ge \|b\|$.

Q.27.5. (Solution to 27.2.6) By proposition 25.5.9 we have for every unit vector u in $\mathbb{R}^n$

$$\begin{aligned}D_u\phi(a) &= d\phi_a(u)\\ &= \langle u, \nabla\phi(a)\rangle\\ &= \|\nabla\phi(a)\|\cos\theta\end{aligned}$$

where $\theta = \measuredangle(u, \nabla\phi(a))$. Since ϕ and a are fixed we maximize the directional derivative $D_u\phi(a)$ by maximizing $\cos\theta$. But $\cos\theta = 1$ when $\theta = 0$; that is, when u and $\nabla\phi(a)$ point in the same direction. Similarly, to minimize $D_u\phi(a)$ choose $\theta = \pi$ so that $\cos\theta = -1$.

Q.27.6. (Solution to 27.2.13) It suffices, by proposition 26.1.9, to show that the derivative of the total energy TE is zero.

$$\begin{aligned} D(TE) &= D(KE) + D(PE) \\ &= \tfrac{1}{2} m D\langle v, v\rangle + D(\phi \circ x) \\ &= \tfrac{1}{2} m (2\langle v, Dv\rangle) + \langle Dx, (\nabla\phi)\circ x\rangle \\ &= m\langle v, a\rangle + \langle v, -F\circ x\rangle \\ &= m\langle v, a\rangle - m\langle v, a\rangle \\ &= 0. \end{aligned}$$

(The third equality uses 27.1.17 and 27.2.7; the second last uses *Newton's second law.*)

Q.27.7. (Solution to 27.2.14) Using the hint we compute

$$\begin{aligned} \nabla\phi(a) &= \sum_{k=1}^{n} \langle \nabla\phi(a), e^k\rangle e^k \qquad \text{(by 27.1.3)} \\ &= \sum_{k=1}^{n} d\phi_a(e^k) e^k \\ &= \sum_{k=1}^{n} D_{e^k}\phi(a) e^k \qquad\quad \text{(by 25.5.9)} \\ &= \sum_{k=1}^{n} \phi_k(a) e^k. \end{aligned}$$

Q.27.8. (Solution to 27.2.15) By proposition 25.5.9

$$D_u\phi(a) = d\phi_a(u) = \langle u, \nabla\phi(a)\rangle\,.$$

Since

$$\begin{aligned} \nabla\phi(w,x,y,z) &= \sum_{k=1}^{4} \phi_k(w,x,y,z) e^k \\ &= (z, -y, -x, w) \end{aligned}$$

we see that

$$\nabla\phi(a) = (4, -3, -2, 1)\,.$$

Thus

$$D_u\phi(a) = \langle(\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}), (4, -3, -2, 1)\rangle = 2\,.$$

Q.27.9. (Solution to 27.2.16) As suggested in the hint, let $c\colon t \mapsto \big(x(t), y(t)\big)$ be the desired curve and set

$$c(0) = \big(x(0), y(0)\big) = a = (2, -1)\,.$$

At each point $c(t)$ on the curve set the tangent vector $Dc(t)$ equal to $-\big(\nabla\phi\big)\big(c(t)\big)$. Then for every t we have

$$\begin{aligned} \big(Dx(t), Dy(t)\big) &= -\big(\nabla\phi\big)\big(x(t), y(t)\big) \\ &= \big(-4x(t), -12y(t)\big). \end{aligned}$$

The two resulting equations

$$Dx(t) = -4x(t) \qquad \text{and} \qquad Dy(t) = -12y(t)$$

have as their only nonzero solutions

$$x(t) = x(0)e^{-4t} = 2e^{-4t}$$

and

$$y(t) = y(0)e^{-12t} = -e^{-12t}\,.$$

Eliminating the parameter we obtain

$$y(t) = -e^{-12t} = -\big(e^{-4t}\big)^3 = -\big(\frac{1}{2}x(t)\big)^3 = -\frac{1}{8}\big(x(t)\big)^3\,.$$

Thus the path of steepest descent (in the xy-plane) follows the curve $y = -\frac{1}{8}x^3$ from $x = 2$ to $x = 0$ (where ϕ obviously assumes its minimum).

Q.27.10. (Solution to 27.3.2) By proposition 21.3.11 it suffices to show that

$$[df_a]e^l = [f^j_k(a)]e^l$$

for $1 \le l \le n$. Since the i^{th} coordinate $(1 \le i \le m)$ of the vector which results from the action of the matrix $[f^j_k(a)]$ on the vector e^l is

$$\sum_{k=1}^{n} f^i_k(a)(e^l)_k = f^i_l(a)$$

we see that

$$\begin{aligned}
[f^j_k(a)]e^l &= \sum_{i=1}^{m} f^i_l(a)\hat{e}^i \\
&= f_l(a) \qquad\qquad \text{(by proposition 26.2.15)} \\
&= df_a(e^l) \\
&= [df_a]e^l.
\end{aligned}$$

Q.27.11. (Solution to 27.3.3)
(a) By proposition 27.3.2

$$[df_{(w,x,y,z)}] = \begin{bmatrix} xz & wz & 0 & wx \\ 0 & 2x & 4y & 6z \\ y\arctan z & 0 & w\arctan z & wy(1+z^2)^{-1} \end{bmatrix}.$$

Therefore

$$[df_a] = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 4 & 6 \\ \pi/4 & 0 & \pi/4 & 1/2 \end{bmatrix}.$$

(b)

$$\begin{aligned}
df_a(v) &= [df_a]v \\
&= \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 4 & 6 \\ \pi/4 & 0 & \pi/4 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ -3 \\ 1 \end{bmatrix} \\
&= \big(3, -2, \tfrac{1}{4}(2-3\pi)\big)\,.
\end{aligned}$$

Q.27.12. (Solution to 27.4.3) Let

$$g\colon \mathbb{R}^4 \to \mathbb{R}^2 \colon (u,v,w,x) \mapsto \big(y(u,v,w,x), z(u,v,w,x)\big)$$

and

$$f\colon \mathbb{R}^2 \to \mathbb{R}^4 \colon (s,t) \mapsto \big(u(s,t), v(s,t), w(s,t), x(s,t)\big)\,.$$

Here it is appropriate to think of the variables and functions as being arranged in the following fashion.

$$\begin{matrix} & & u & & \\ s & \xrightarrow{\ f\ } & v & \xrightarrow{\ g\ } & y \\ t & & w & & z \\ & & x & & \end{matrix}$$

The expression $\frac{\partial u}{\partial t}$ is then taken to represent the function f_2^1. The expression $\frac{\partial z}{\partial u}$ appearing in the statement of the exercise represents $g_1^2 \circ f$. [One's first impulse might be to let $\frac{\partial z}{\partial u}$ be just g_1^2. But this cannot be correct. The product of $\frac{\partial z}{\partial u}$ and $\frac{\partial u}{\partial t}$ is defined only at points where both are defined. The product of g_1^2 (whose domain lies in $\mathbb{R}^4$) and f_2^1 (whose domain is in $\mathbb{R}^2$) is never defined.] On the left side of the equation the expression $\frac{\partial z}{\partial t}$ is the partial derivative with respect to t of the composite function $f \circ g$. Thus it is expressed functionally as $(g \circ f)_2^2$.

Using proposition 27.4.1 we obtain

$$\begin{aligned} \frac{\partial z}{\partial t} &= (g \circ f)_2^2 \\ &= \sum_{i=1}^{4} (g_i^2 \circ f) f_2^i \\ &= \frac{\partial z}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial t} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} \end{aligned}$$

This equation is understood to hold at all points a in $\mathbb{R}^2$ such that f is differentiable at a and g is differentiable at $f(a)$.

Q.27.13. (Solution to 27.4.5) Since

$$\left[df_{(x,y,z)}\right] = \begin{bmatrix} y^2 & 2xy & 0 \\ 3 & 0 & -2z \\ yz & xz & xy \\ 2x & 2y & 0 \\ 4z & 0 & 4x \end{bmatrix}$$

we see that

$$\left[df_a\right] = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \\ -4 & 0 & 4 \end{bmatrix}.$$

And since

$$\left[dg_{(s,t,u,v,w)}\right] = \begin{bmatrix} 2s & 0 & 2u & 2v & 0 \\ 2sv & -2w^2 & 0 & s^2 & -4tw \end{bmatrix}$$

we see that

$$\left[dg_{f(a)}\right] = \left[dg_{(0,2,0,1,1)}\right] = \begin{bmatrix} 0 & 0 & 0 & 2 & 0 \\ 0 & -2 & 0 & 0 & -8 \end{bmatrix}.$$

Thus by equation (27.1)

$$\begin{aligned}
\big[d(g\circ f)_a\big] &= \big[dg_{f(a)}\big]\big[df_a\big]\\
&= \begin{bmatrix} 0 & 0 & 0 & 2 & 0\\ 0 & -2 & 0 & 0 & -8\end{bmatrix}\begin{bmatrix} 0 & 0 & 0\\ 3 & 0 & 2\\ 0 & -1 & 0\\ 2 & 0 & 0\\ -4 & 0 & 4\end{bmatrix}\\
&= \begin{bmatrix} 4 & 0 & 0\\ 26 & 0 & -36\end{bmatrix}.
\end{aligned}$$

Q.27.14. (Solution to 27.4.8) Use formula (27.6). It is understood that $\frac{\partial x}{\partial t}$ and $\frac{\partial y}{\partial t}$ must be evaluated at the point $(1,1)$; and since $x(1,1)=2$ and $y(1,1)=\pi/4$, the partials $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$, and $\big(\frac{\partial w}{\partial t}\big)_{x,y}$ are to be evaluated at the point $(2,\pi/4,1)$. Calculate the terms appearing on the right hand side of (27.6):

$$\begin{aligned}
\frac{\partial x}{\partial t} &= 2t; &&\text{so } \frac{\partial x}{\partial t}(1,1)=2,\\
\frac{\partial y}{\partial t} &= \frac{s}{1+t^2}; &&\text{so } \frac{\partial y}{\partial t}(1,1)=1/2,\\
\frac{\partial w}{\partial x} &= -\frac{2y}{x^2}; &&\text{so } \frac{\partial w}{\partial x}(2,\pi/4,1)=-\pi/8,\\
\frac{\partial w}{\partial y} &= \frac{2}{x}; &&\text{so } \frac{\partial w}{\partial y}(2,\pi/4,1)=1, \text{ and}\\
\left(\frac{\partial w}{\partial t}\right)_{x,y} &= 3t^2; &&\text{so } \left(\frac{\partial w}{\partial t}\right)_{x,y}(2,\pi/4,1)=3\,.
\end{aligned}$$

Therefore

$$\left(\frac{\partial w}{\partial t}\right)_s(1,1) = -\frac{\pi}{8}\cdot 2 + 1\cdot\frac{1}{2} + 3 = \frac{7}{2} - \frac{\pi}{4}.$$

Q.27.15. (Solution to 27.4.9) We proceed through steps (a)–(g) of the hint.

(a) Define $g(x,y)=y/x$ and compute its differential

$$\begin{aligned}
\big[dg_{(x,y)}\big] &= \big[g_1(x,y) \quad g_2(x,y)\big]\\
&= \big[-yx^{-2} \quad x^{-1}\big].
\end{aligned}$$

(b) Then compute the differential of $\phi\circ g$

$$\begin{aligned}
\big[d(\phi\circ g)_{(x,y)}\big] &= \big[d\phi_{g(x,y)}\big]\big[dg_{(x,y)}\big]\\
&= \phi'\big(g(x,y)\big)\big[dg_{(x,y)}\big]\\
&= \phi'(yx^{-1})\big[-yx^{-2} \quad x^{-1}\big]\\
&= \big[-yx^{-2}\phi'(yx^{-1}) \quad x^{-1}\phi'(yx^{-1})\big].
\end{aligned}$$

(c) Let $G(x,y)=\big(x,\phi(y/x)\big)$ and use (b) to calculate $\big[dG_{(x,y)}\big]$

$$\begin{aligned}
\big[dG_{(x,y)}\big] &= \begin{bmatrix} G^1_1(x,y) & G^1_2(x,y)\\ G^2_1(x,y) & G^2_2(x,y)\end{bmatrix}\\
&= \begin{bmatrix} 1 & 0\\ -yx^{-2}\phi'(yx^{-1}) & x^{-1}\phi'(yx^{-1})\end{bmatrix}
\end{aligned}$$

(d) Let $m(x,y)=xy$ and compute its differential

$$\big[dm_{(x,y)}\big] = \big[m_1(x,y) \quad m_2(x,y)\big] = \big[y \quad x\big]\,.$$

(e) Since $h(x,y) = x\phi(yx^{-1}) = m(G(x,y))$ we see that $h = m \circ G$ and therefore

$$\begin{aligned}
\left[dh_{(x,y)}\right] &= \left[d(m \circ G)_{(x,y)}\right] \\
&= \left[dm_{G(x,y)}\right]\left[dG_{(x,y)}\right] \\
&= \begin{bmatrix} G^2(x,y) & G^1(x,y) \end{bmatrix}\left[dG_{(x,y)}\right] \\
&= \begin{bmatrix} \phi(yx^{-1}) & x \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -yx^{-2}\phi'(yx^{-1}) & x^{-1}\phi'(yx^{-1}) \end{bmatrix} \\
&= \begin{bmatrix} \phi(yx^{-1} - yx^{-1}\phi'(yx^{-1}) & \phi'(yx^{-1}) \end{bmatrix}.
\end{aligned}$$

(f) Since $j(x,y) = xy + x\phi(yx^{-1}) = m(x,y) + h(x,y)$, we see that

$$\begin{aligned}
\left[dj_{(x,y)}\right] &= \left[dm_{(x,y)}\right] + \left[dh_{(x,y)}\right] \\
&= \begin{bmatrix} y + \phi(yx^{-1} - yx^{-1}\phi'(yx^{-1}) & x + \phi'(yx^{-1}) \end{bmatrix}.
\end{aligned}$$

(g) Then finally,

$$\begin{aligned}
xj_1(x,y) + yj_2(x,y) &= x\big(y + \phi(yx^{-1} - yx^{-1}\phi'(yx^{-1})\big) + y\big(x + \phi'(yx^{-1})\big) \\
&= xy + x\phi(yx^{-1}) + yx \\
&= xy + j(x,y)
\end{aligned}$$

Q.27.16. (Solution to 27.4.10) Let h be as in the hint. Then

$$\left[dh_{(x,y)}\right] = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

so

$$\begin{aligned}
\left[dg_{(x,y)}\right] &= \left[d(f \circ h)_{(x,y)}\right] \\
&= \left[df_{h(x,y)}\right]\left[dh_{(x,y)}\right] \\
&= \begin{bmatrix} f_1(h(x,y)) & f_2(h(x,y)) \end{bmatrix} \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix} \\
&= \begin{bmatrix} 2xf_1(h(x,y)) + 2yf_2(h(x,y)) & -2yf_1(h(x,y)) + 2xf_2(h(x,y)) \end{bmatrix}.
\end{aligned}$$

Therefore

$$\begin{aligned}
yg_1(x,y) - xg_2(x,y) &= 2xyf_1(h(x,y)) + 2y^2f_2(h(x,y)) + 2xyf_1(h(x,y)) - 2x^2f_2(h(x,y)) \\
&= 4xyf_1(h(x,y)) - 2(x^2 - y^2)f_2(h(x,y)) \\
&= 2h^2(x,y)f_1(h(x,y)) - 2h^1(x,y)f_2(h(x,y)).
\end{aligned}$$

This computation, incidentally, gives one indication of the attractiveness of notation which omits evaluation of partial derivatives. If one is able to keep in mind the points at which the partials are being evaluated, less writing is required.

Q.28. Exercises in chapter 28

Q.28.1. (Solution to 28.1.2) If n is odd then the n^{th} partial sum s_n is 1; if n is even then $s_n = 0$.

Q.28.2. (Solution to 28.1.3) Use problem 28.1.8. The n^{th} partial sum of the sequence $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right)$ is

$$\begin{aligned} s_n &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots + \frac{1}{2^n} \\ &= \sum_{k=1}^{n} \left(\frac{1}{2}\right)^k \\ &= \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k - 1 \\ &= \frac{1 - (\frac{1}{2})^{n+1}}{1 - \frac{1}{2}} - 1 \\ &= 2 - \left(\frac{1}{2}\right)^n - 1 \\ &= 1 - 2^{-n}. \end{aligned}$$

Q.28.3. (Solution to 28.1.5) For the sequence given in exercise 28.1.2, the corresponding series $\sum_{k=1}^{\infty} a_k$ is the sequence $(1, 0, 1, 0, 1, \dots)$ (of partial sums). For the sequence in exercise 28.1.3, the series $\sum_{k=1}^{\infty} a_k$ is the sequence $\left(\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \dots\right)$ (of partial sums).

Q.28.4. (Solution to 28.1.7) For the sequence (a_k) given in 28.1.2 the corresponding sequence of partial sums $(1, 0, 1, 0, 1, \dots)$ does not converge. Thus the sequence $(1, -1, 1, -1, \dots)$ is not summable. Equivalently, the series $\sum_{k=1}^{\infty} (-1)^{k+1}$ diverges.

For the sequence (a_k) of 28.1.3, the n^{th} partial sum is $1 - 2^{-n}$ (see 28.1.3). Since $\lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - 2^{-n}\right) = 1$ (see proposition 4.3.8), we conclude that the sequence $(1/2, 1/4, 1/8, \dots)$ is summable; in other words, the series $\sum_{k=1}^{\infty} 2^{-k}$ converges. The sum of this series is 1; that is

$$\sum_{k=1}^{\infty} 2^{-k} = 1\,.$$

Q.28.5. (Solution to 28.1.10) Suppose that $\sum_{k=1}^{\infty} a_k = b$. If $s_n = \sum_{k=1}^{n} a_k$, then it is easy to see that for each n we may write a_n as $s_n - s_{n-1}$ (where we let $s_0 = 0$). Take limits as $n \to \infty$ to obtain

$$a_n = s_n - s_{n-1} \to b - b = 0\,.$$

Q.28.6. (Solution to 28.1.11) Assume that the series $\sum_{k=1}^{\infty} k^{-1}$ converges. Let $s_n = \sum_{k=1}^{n} k^{-1}$. Since the sequence (s_n) of partial sums is assumed to converge, it is Cauchy (by proposition 18.1.4). Thus there exists an index p such that $|s_n - s_p| < \frac{1}{2}$ whenever $n \geq p$. We obtain a contradiction by noting that

$$\begin{aligned} |s_{2p} - s_p| &= \sum_{k=p+1}^{2p} \frac{1}{k} \\ &\geq \sum_{k=p+1}^{2p} \frac{1}{2p} \\ &= \frac{p}{2p} \\ &= \frac{1}{2}\,. \end{aligned}$$

Q.28.7. (Solution to 28.1.17) Let $\sum a_k$ be a convergent series in the normed linear space V. For each n in $\mathbb{N}$ let $s_n = \sum_{k=1}^{n} a_k$. Then (s_n) is a convergent sequence. By proposition 18.1.4 it is

Cauchy. Thus given $\epsilon > 0$ we may choose n_0 in $\mathbb{N}$ so that $n > m \geq n_0$ implies

$$\left\| \sum_{k=m+1}^{n} a_k \right\| = \|s_n - s_m\| < \epsilon \,. \tag{Q.28}$$

For the second assertion of the proposition, suppose that V is complete. Suppose further that (a_k) is a sequence in V for which there exists $n_0 \in \mathbb{N}$ such that (Q.28) holds whenever $n > m \geq n_0$. (As above, $s_n = \sum_{k=1}^{n} a_k$.) This says that the sequence (s_n) of partial sums is Cauchy, and since V is complete, the sequence (s_n) converges. That is, the series $\sum a_k$ converges.

Q.28.8. (Solution to 28.1.19) Let $f_n(x) = x^n\big(1 + x^n\big)^{-1}$ for every $n \in \mathbb{N}$ and $x \in [-\delta, \delta]$. Also let $M_n = \delta^n(1-\delta)^{-1}$. Since $0 < \delta < 1$, the series $\sum M_n = \sum \delta^n(1-\delta)^{-1}$ converges (by problem 28.1.8). For $|x| \leq \delta$, we have $-x^n \leq |x|^n \leq \delta^n \leq \delta$; so $x^n \geq -\delta$ and $1 + x^n \geq 1 - \delta$. Thus

$$|f_n(x)| = \frac{|x|^n}{1 + x^n} \leq \frac{\delta^n}{1 - \delta} = M_n \,.$$

Thus

$$\|f_n\|_u = \sup\{|f_n(x)| \colon |x| \leq \delta\} \leq M_n$$

By the *Weierstrass M-test* (proposition 28.1.18), the series $\sum_{k=1}^{\infty} f_k$ converges uniformly.

Q.28.9. (Solution to 28.2.3) As was remarked after proposition 28.1.17, the convergence of a series is not affected by altering any finite number of terms. Thus without loss of generality we suppose that $a_{k+1} \leq \delta a_k$ for all k. Notice that $a_2 \leq \delta a_1$, $a_3 \leq \delta a_2 \leq \delta^2 a_1$, $a_4 \leq \delta^3 a_1$, *etc.* In general, $a_k \leq \delta^{k-1} a_1$ for all k. The geometric series $\sum \delta^{k-1}$ converges by problem 28.1.8. Thus by the *comparison test* (proposition 28.2.2), the series $\sum a_k$ converges. The second conclusion follows similarly from the observations that $a_k \geq M^{k-1} a_1$ and that $\sum M^{k-1}$ diverges.

Q.28.10. (Solution to 28.3.2) Suppose that V is complete and that (a_k) is an absolutely summable sequence in V. We wish to show that (a_k) is summable. Let $\epsilon > 0$. Since $\sum \|a_k\|$ converges in $\mathbb{R}$ and $\mathbb{R}$ is complete, we may invoke the *Cauchy criterion* (proposition 28.1.17) to find an integer n_0 such that $n > m \geq n_0$ implies $\sum_{k=m+1}^{n} \|a_k\| < \epsilon$. But for all such m and n

$$\left\| \sum_{k=m+1}^{n} a_k \right\| \leq \sum_{k=m+1}^{n} \|a_k\| < \epsilon \,.$$

This, together with the fact that V is complete, allows us to apply for a second time the *Cauchy criterion* and to conclude that $\sum a_k$ converges. That is, the sequence (a_k) is summable.

For the converse suppose that every absolutely summable sequence in V is summable. Let (a_k) be a Cauchy sequence in V. In order to prove that V is complete we must show that (a_k) converges. For each k in $\mathbb{N}$ we may choose a natural number p_k such that $\|a_n - a_m\| \leq 2^{-k}$ whenever $n > m \geq p_k$. Choose inductively a sequence (n_k) in $\mathbb{N}$ as follows. Let n_1 be any integer such that $n_1 \geq p_1$. Having chosen integers $n_1 < n_2 < \cdots < n_k$ in $\mathbb{N}$ so that $n_j \geq p_j$ for $1 \leq j \leq k$, choose n_{k+1} to be the larger of p_{k+1} and $n_k + 1$. Clearly, $n_{k+1} > n_k$ and $n_{k+1} \geq p_{k+1}$. Thus $\big(a_{n_k}\big)$ is a subsequence of (a_n) and (since $n_{k+1} > n_k \geq p_k$ for each k) $\big\|a_{n_{k+1}} - a_{n_k}\big\| < 2^{-k}$ for each k in $\mathbb{N}$. Let $y_k = a_{n_{k+1}} - a_{n_k}$ for each k. Then (y_n) is absolutely summable since $\sum \|y_k\| < \sum 2^{-k} = 1$. Consequently (y_k) is summable in V. That is, there exists b in V such that $\sum_{k=1}^{j} y_k \to b$ as $j \to \infty$. However, since $\sum_{k=1}^{j} y_k = a_{n_{j+1}} - a_{n_1}$, we see that

$$a_{n_{j+1}} \to a_{n_1} + b \qquad \text{as } j \to \infty \,.$$

This shows that $\big(a_{n_k}\big)$ converges. Since (a_n) is a Cauchy sequence having a convergent subsequence it too converges (proposition 18.1.5). But this is what we wanted to show.

Q.28.11. (Solution to 28.4.4)

(a) It follows immediately from

$$|(fg)(x)| = |f(x)||g(x)| \le \|f\|_u \|g\|_u \qquad \text{for every } x \in S$$

that

$$\|fg\|_u = \sup\{|(fg)(x)| \colon x \in S\} \le \|f\|_u \, \|g\|_u \,.$$

(b) Define f and g on $[0, 2]$ by

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1, \\ 0 & \text{if } 1 < x \le 2 \end{cases}$$

and $g(x) = 1 - f(x)$. Then $\|f\|_u = \|g\|_u = 1$, but $\|fg\|_u = 0$.

Q.28.12. (Solution to 28.4.11) Since $\|x\| < 1$, the series $\sum_{k=0}^{\infty} \|x\|^k$ converges by problem 28.1.8. Condition (e) in the definition of normed algebras is that $\|xy\| \le \|x\| \, \|y\|$. An easy inductive argument shows that $\|x^n\| \le \|x\|^n$ for all n in $\mathbb{N}$. We know that $\|x\|^n \to 0$ (by proposition 4.3.8); so $\|x^n\| \to 0$ also. Thus (by proposition 22.2.3(d)) $x^n \to 0$ as $n \to \infty$. Furthermore, comparing $\sum_0^{\infty} \|x^k\|$ with the series $\sum_0^{\infty} \|x\|^k$ shows that the former converges (see proposition 28.2.2). But this says just that the series $\sum_0^{\infty} x^k$ converges absolutely. It then follows from proposition 28.3.2 that $\sum_0^{\infty} x^k$ converges. Letting $s_n = \sum_{k=0}^{n} x^k$ we see that

$$\begin{aligned} (\mathbf{1} - x) \sum_{k=0}^{\infty} x^k &= (\mathbf{1} - x) \lim s_n \\ &= \lim\big((\mathbf{1} - x) s_n\big) \\ &= \lim\big(\mathbf{1} - x^{n+1}\big) \\ &= \mathbf{1}. \end{aligned}$$

Similarly, $\big(\sum_0^{\infty} x^k\big)(\mathbf{1} - x) = \mathbf{1}$. This shows that $\mathbf{1} - x$ is invertible and that its inverse $(\mathbf{1} - x)^{-1}$ is the geometric series $\sum_0^{\infty} x^k$.

Q.28.13. (Solution to 28.4.14) Let $a \in \operatorname{Inv} A$. We show that r is continuous at a. Given $\epsilon > 0$ choose δ to be the smaller of the numbers $\frac{1}{2}\|a^{-1}\|^{-1}$ and $\frac{1}{2}\|a^{-1}\|^{-2}\epsilon$. Suppose that $\|y - a\| < \delta$ and prove that $\|r(y) - r(a)\| < \epsilon$. Let $x = \mathbf{1} - a^{-1}y$. Since

$$\|x\| = \|a^{-1}a - a^{-1}y\| \le \|a^{-1}\| \|y - a\| < \|a^{-1}\|\delta \le \|a^{-1}\| \tfrac{1}{2} \|a^{-1}\|^{-1} = \tfrac{1}{2}$$

we conclude from 28.4.11 and 28.4.12 that $\mathbf{1} - x$ is invertible and that

$$\|(\mathbf{1} - x)^{-1} - \mathbf{1}\| \le \frac{\|x\|}{1 - \|x\|} \tag{Q.29}$$

Thus

$$\begin{aligned} \|r(y) - r(a)\| &= \|y^{-1}(a - y)a^{-1}\| && \text{(by 28.4.10(e))} \\ &\le \|y^{-1}a - \mathbf{1}\| \, \|a^{-1}\| \\ &= \|(a^{-1}y)^{-1} - \mathbf{1}\| \, \|a^{-1}\| && \text{(by 28.4.10(d))} \\ &= \|(\mathbf{1} - x)^{-1} - \mathbf{1}\| \, \|a^{-1}\| \\ &\le \frac{\|x\|}{1 - \|x\|} \|a^{-1}\| && \text{(by inequality (Q.29))} \\ &\le 2\|x\| \, \|a^{-1}\| && \text{(because } \|x\| \le \tfrac{1}{2}) \\ &< 2\|a^{-1}\|^2 \\ &\le \epsilon. \end{aligned}$$

Q.28.14. (Solution to 28.4.17) Throughout the proof we use the notation introduced in the hint. To avoid triviality we suppose that (a_k) is not identically zero. Since u_n is defined to be $\sum_{k=0}^{n} c_k = \sum_{k=0}^{n} \sum_{j=0}^{k} a_j b_{k-j}$ it is clear that u_n can be obtained by finding the sum of each column of the matrix $\big[d_{jk}\big]$ and then adding these sums. On the other hand the expression

$$\sum_{k=0}^{n} a_{n-k} t_k = \sum_{k=0}^{n} \sum_{j=0}^{k} a_{n-k} b_j$$

is obtained by finding the sum of each row of the matrix $[d_{jk}]$ and then adding the sums. It is conceivable that someone might find the preceding argument too "pictorial", depending as it does on looking at a "sketch" of the matrix $[d_{jk}]$. It is, of course, possible to carry out the proof in a purely algebraic fashion. And having done so, it is also quite conceivable that one might conclude that the algebraic approach adds more to the amount of paper used than to the clarity of the argument. In any event, here, for those who feel more comfortable with it, is a formal verification of the same result.

$$\begin{aligned}
u_n &= \sum_{k=0}^{n} c_k \\
&= \sum_{k=0}^{n} \sum_{j=0}^{k} a_j b_{k-j} \\
&= \sum_{k=0}^{n} \sum_{j=0}^{k} d_{jk} \\
&= \sum_{k=0}^{n} \sum_{j=0}^{n} d_{jk} \\
&= \sum_{j=0}^{n} \sum_{k=0}^{n} d_{jk} \\
&= \sum_{j=0}^{n} \sum_{k=j}^{n} d_{jk} \\
&= \sum_{j=0}^{n} \sum_{k=j}^{n} a_j b_{k-j} \\
&= \sum_{j=0}^{n} a_j \sum_{r=0}^{n-j} b_r \\
&= \sum_{j=0}^{n} a_j t_{n-j} \\
&= \sum_{k=0}^{n} a_{n-k} t_k.
\end{aligned}$$

Now that equation (28.10) has been established we see that

$$\begin{aligned}
u_n &= \sum_{k=0}^{n} a_{n-k} b + \sum_{k=0}^{n} a_{n-k}(t_k - b) \\
&= s_n b + \sum_{k=0}^{n} a_{n-k}(t_k - b).
\end{aligned}$$

Since $s_n b \to ab$, it remains only to show that the last term on the right approaches 0 as $n \to \infty$. Since

$$\begin{aligned}\left\|\sum_{k=0}^{n} a_{n-k}(t_k - b)\right\| &\le \sum_{k=0}^{n}\|a_{n-k}(t_k - b)\| \\ &\le \sum_{k=0}^{n}\|a_{n-k}\|\,\|(t_k - b)\| \\ &= \sum_{k=0}^{n}\alpha_{n-k}\beta_k\end{aligned}$$

it is sufficient to prove that given any $\epsilon > 0$ the quantity $\sum_{k=0}^{n}\alpha_{n-k}\beta_k$ is less than ϵ whenever n is sufficiently large.

Let $\alpha = \sum_{k=0}^{\infty}\|a_k\|$. Then $\alpha > 0$. Since $\beta_k \to 0$ there exists n_1 in $\mathbb{N}$ such that $k \ge n_1$ implies $\beta_k < \epsilon/(2\alpha)$. Choose $\beta > \sum_{k=0}^{n_1}\beta_k$. Since $\alpha_k \to 0$, there exists n_2 in $\mathbb{N}$ such that $k \ge n_2$ implies $\alpha_k < \epsilon/(2\beta)$.

Now suppose that $n \ge n_1 + n_2$. If $0 \le k \le n_1$, then $n - k \ge n - n_1 \ge n_2$, so that $\alpha_{n-k} < \epsilon/(2\beta)$. This shows that

$$\begin{aligned}p &= \sum_{k=0}^{n_1}\alpha_{n-k}\beta_k \\ &\le \epsilon(2\beta)^{-1}\sum_{k=0}^{n_1}\beta_k \\ &< \epsilon/2.\end{aligned}$$

Furthermore,

$$\begin{aligned}q &= \sum_{k=n_1+1}^{n}\alpha_{n-k}\beta_k \\ &\le \epsilon(2\alpha)^{-1}\sum_{k=n_1+1}^{n}\alpha_{n-k} \\ &\le \epsilon(2\alpha)^{-1}\sum_{j=0}^{\infty}\|a_j\| \\ &= \epsilon/2.\end{aligned}$$

Thus

$$\sum_{k=0}^{n}\alpha_{n-k}\beta_k = p + q < \epsilon\,.$$

Q.28.15. (Solution to 28.4.25) Let $0 < s < r$, let $M > 0$ be such that $\|a_k\|r^k \le M$ for every k in $\mathbb{N}$, and let $\rho = s/r$. Let $f_k(x) = a_k x^k$ for each k in $\mathbb{N}$ and x in $B_s(0)$. For each such k and x

$$\begin{aligned}\|f_k(x)\| = \|a_k x^k\| &\le \|a_k\|\|x\|^k \le \|a_k\|s^k \\ &= \|a_k\|r^k\rho^k \le M\rho^k.\end{aligned}$$

Thus $\|f_k\|_u \le M\rho^k$ for each k. Since $0 < \rho < 1$, the series $\sum M\rho^k$ converges. Then, according to the *Weierstrass M-test* (proposition 28.1.18), the series $\sum a_k x^k = \sum f_k(x)$ converges uniformly on $B_s(0)$. The parenthetical comment in the statement of the proposition is essentially obvious: For $a \in B_r(0)$ choose s such that $\|a\| < s < r$. Since $\sum a_k x^k$ converges uniformly on $B_s(0)$, it converges at a (see problem 22.4.7).

Q.28.16. (Solution to 28.4.26) Let a be an arbitrary point of U. Let $\phi = \lim_{n\to\infty} d(f_n)$. We show that $\Delta F_a \simeq T$ where $T = \phi(a)$. We are supposing that $d(f_n) \to \phi$ (unif) on U. Thus given $\epsilon > 0$ we may choose N in $\mathbb{N}$ so that

$$\sup\{\left\|d(f_n)_x - \phi(x)\right\| : x \in U\} < \tfrac{1}{8}\epsilon$$

whenever $x \in U$ and $n \geq N$. Let $g_n = f_n - f_N$ for all $n \geq N$. Then for all such n and all $x \in U$ we have

$$\left\|d(g_n)_x\right\| \leq \left\|d(f_n)_x - \phi(x)\right\| + \left\|\phi(x) - d(f_N)_x\right\| < \tfrac{1}{4}\epsilon\,.$$

Also it is clear that

$$\left\|d(g_n)_x - d(g_n)_a\right\| \leq \left\|d(g_n)_x\right\| + \left\|d(g_n)_a\right\| < \tfrac{1}{2}\epsilon$$

for $x \in U$ and $n \geq N$. According to corollary 26.1.8

$$\left\|\Delta(g_n)_a(h) - d(g_n)_a(h)\right\| \leq \tfrac{1}{2}\epsilon\|h\|$$

whenever $n \geq N$ and h is a vector such that $a + h \in U$. Thus

$$\left\|\Delta(f_n)_a(h) - d(f_n)_a(h) - \Delta(f_N)_a(h) + d(f_N)_a(h)\right\| \leq \tfrac{1}{2}\epsilon\|h\|$$

when $n \geq N$ and $a + h \in U$. Taking the limit as $n \to \infty$ we obtain

$$\left\|(\Delta F_a(h) - Th) - \left(\Delta(f_N)_a(h) - d(f_N)_a(h)\right)\right\| \leq \frac{1}{2}\epsilon\|h\| \tag{Q.30}$$

for h such that $a + h \in U$. Since f_N is differentiable, $\Delta(f_N)_a \simeq d(f_N)_a$; thus there exists $\delta > 0$ such that $B_\delta(a) \subseteq U$ and

$$\left\|\Delta(f_N)_a(h) - d(f_N)_a(h)\right\| < \tfrac{1}{2}\epsilon\|h\| \tag{Q.31}$$

for all h such that $\|h\| < \delta$. From (Q.30) and (Q.31) it is clear that

$$\|\Delta F_a(h) - Th\| < \epsilon\|h\|$$

whenever $\|h\| < \delta$. Thus $\Delta F_a \simeq T$, which shows that F is differentiable at a and

$$dF_a = T = \lim_{n\to\infty} d(f_n)_a\,.$$

Q.29. Exercises in chapter 29

Q.29.1. (Solution to 29.1.2) Let $U = V = \mathbb{R}$ and $f(x) = x^3$ for all x in $\mathbb{R}$. Although f is continuously differentiable and does have an inverse, it is not $\mathcal{C}^1$-invertible. The inverse function $x \mapsto x^{\frac{1}{3}}$ is not differentiable at 0.

Q.29.2. (Solution to 29.1.4) Set $y = x^2 - 6x + 5$ and solve for x in terms of y. After completing the square and taking square roots we have

$$|x - 3| = \sqrt{y + 4}\,.$$

Thus there are two solutions $x = 3 + \sqrt{y+4}$ and $x = 3 - \sqrt{y+4}$. The first of these produces values of x no smaller than 3 and the second produces values no larger than 3. Thus for $x = 1$ we choose the latter. A local $\mathcal{C}^1$-inverse of f is given on the interval $f^{\rightarrow}(0,2) = (-3,5)$ by

$$f_{\mathrm{loc}}^{-1}(y) = 3 - \sqrt{y+4}\,.$$

Q.29.3. (Solution to 29.1.7) In order to apply the chain rule to the composite function $f_{\mathrm{loc}}^{-1} \circ f$ we need to know that *both* f and f_{loc}^{-1} are differentiable. But differentiability of f_{loc}^{-1} was not a hypothesis. Indeed, the major difficulty in proving the *inverse function theorem* is showing that a local $\mathcal{C}^1$-inverse of a $\mathcal{C}^1$-function is in fact differentiable (at points where its differential does not vanish). Once that is known, the argument presented in 29.1.7 correctly derives the formula for $Df_{\mathrm{loc}}^{-1}(b)$.

Q.29.4. (Solution to 29.2.2) Let $f(x,y) = x^2y + \sin(\frac{\pi}{2}xy^2) - 2$ for all x and y in $\mathbb{R}$.

(a) There exist a neighborhood V of 1 and a function $h\colon V \to \mathbb{R}$ which satisfy

$$\begin{aligned} &\text{(i)} \quad h(1) = 2; \quad \text{and} \\ &\text{(ii)} \quad f(x, h(x)) = 0 \quad \text{for all } x \text{ in } V. \end{aligned}$$

(b) Let $G(x, y) = (x, f(x, y))$ for all $x, y \in \mathbb{R}$. Then G is continuously differentiable and

$$\big[dG_{(1,2)}\big] = \begin{bmatrix} 1 & 0 \\ 4+2\pi & 1+2\pi \end{bmatrix}.$$

Thus $dG_{(1,2)}$ is invertible, so by the *inverse function theorem* G has a local $\mathcal{C}^1$-inverse, say H, defined on some neighborhood W of $(1, 0) = G(1, 2)$. Let $V = \{x : (x, 0) \in W\}$ and $h(x) = H^2(x, 0)$ for all x in V. The function h is in $\mathcal{C}^1$ because H is. Condition (i) is satisfied by h since

$$\begin{aligned} (1,2) &= H(G(1,2)) \\ &= H(1, f(1,2)) \\ &= H(1,0) \\ &= (H^1(1,0)\,,\; H^2(1,0)) \\ &= (H^1(1,0)\,,\; h(1)) \end{aligned}$$

and (ii) holds because

$$\begin{aligned} (x,0) &= G(H(x,0)) \\ &= G(H^1(x,0)\,,\; H^2(x,0)) \\ &= (H^1(x,0)\,,\; f(H^1(x,0), H^2(x,0))) \\ &= (x\,,\; f(x, h(x))) \end{aligned}$$

for all x in V.

(c) Let G, H, and h be as in (b). By the *inverse function theorem*

$$\begin{aligned} \big[dH_{(1,0)}\big] &= \big[dH_{G(1,2)}\big] \\ &= \big[dG_{(1,2)}\big]^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 4+2\pi & 1+2\pi \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ -\frac{4+2\pi}{1+2\pi} & \frac{1}{1+2\pi} \end{bmatrix}. \end{aligned}$$

Then $\frac{dy}{dx}$ at $(1, 2)$ is just $h'(1)$ and

$$h'(1) = H^2_1(1,0) = -\frac{4+2\pi}{1+2\pi}.$$

Q.29.5. (Solution to 29.2.4) Let $f(x, y, z) = x^2 z + yz^2 - 3z^3 - 8$ for all $x,\ y,\ z \in \mathbb{R}$.

(a) There exist a neighborhood V of $(3, 2)$ and a function $h\colon V \to \mathbb{R}$ which satisfy

$$\begin{aligned} &\text{(i)} \quad h(3,2) = 1; \quad \text{and} \\ &\text{(ii)} \quad f(x, y, h(x,y)) = 0 \quad \text{for all } x,\ y \in V. \end{aligned}$$

(b) Let $G(x, y, z) := (x, y, f(x, y, z))$ for all $x,\ y,\ z \in \mathbb{R}$. Then

$$\big[dG_{(3,2,1)}\big] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 6 & 1 & 4 \end{bmatrix}$$

so $dG_{(3,2,1)}$ is invertible. By the *inverse function theorem* G has a local $\mathcal{C}^1$-inverse H defined on some neighborhood W of $(3,2;0) = G(3,2,1)$. Write $H = (H^1, H^2)$ where $\operatorname{ran} H^1 \subseteq \mathbb{R}^2$ and $\operatorname{ran} H^2 \subseteq \mathbb{R}$. Let $V = \{(x,y) : (x,y,0) \in W\}$ and $h(x,y) = H^2(x,y;0)$. The function h belongs to $\mathcal{C}^1$ because H does. Now condition (i) holds because

$$\begin{aligned}(3,2;1) &= H(G(3,2;1))\\ &= H(3,2;f(3,2,1))\\ &= H(3,2;0)\\ &= (H^1(3,2;0)\,;\,H^2(3,2;0))\\ &= (H^1(3,2;0)\,;\,h(3,2))\end{aligned}$$

and condition (ii) follows by equating the third components of the first and last terms of the following computation

$$\begin{aligned}(x,y;0) &= G(H(x,y;0))\\ &= G(H^1(x,y;0);H^2(x,y;0))\\ &= (H^1(x,y;0)\,;\,f(H^1(x,y;0);H^2(x,y;0)))\\ &= (x,y\,;\,f(x,y;h(x,y)))\,.\end{aligned}$$

(c) We wish to find $\left(\frac{\partial z}{\partial x}\right)_y$ and $\left(\frac{\partial z}{\partial y}\right)_x$ at $(3,2,1)$; that is, $h_1(3,2)$ and $h_2(3,2)$, respectively. The *inverse function theorem* tells us that

$$\begin{aligned}\left[dH_{(3,2,0)}\right] &= \left[dH_{G(3,2,1)}\right]\\ &= \left[dG_{(3,2,1)}\right]^{-1}\\ &= \begin{bmatrix}1&0&0\\0&1&0\\6&1&4\end{bmatrix}^{-1}\\ &= \begin{bmatrix}1&0&0\\0&1&0\\-\frac{3}{2}&-\frac{1}{4}&\frac{1}{4}\end{bmatrix}.\end{aligned}$$

Thus at $(3,2,1)$

$$\left(\frac{\partial z}{\partial x}\right)_y = h_1(3,2) = \frac{\partial H^2}{\partial x}(3,2) = -\frac{3}{2}$$

and

$$\left(\frac{\partial z}{\partial y}\right)_x = h_2(3,2) = \frac{\partial H^2}{\partial y}(3,2) = -\frac{1}{4}\,.$$

Q.29.6. (Solution to 29.2.13) Let $f = (f^1, f^2)$ where

$$f^1(u,v;x,y) = 2u^3vx^2 + v^2x^3y^2 - 3u^2y^4$$

and

$$f^2(u,v;x,y) = 2uv^2y^2 - uvx^2 + u^3xy - 2\,.$$

(a) There exist a neighborhood V of (a,b) in $\mathbb{R}^2$ and a function $h\colon V \to \mathbb{R}^2$ which satisfy

(i) $h(a,b) = (c,d)$; and

(ii) $f(u,v;h(u,v)) = (0,0)$ for all $u,\,v \in V$.

(b) Let $G(u,v;x,y) := (u,v;f(u,v;x,y))$ for all $u,\ v,\ x,\ y \in \mathbb{R}$. Then G is continuously differentiable and

$$\left[dG_{(1,1)}\right] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & 7 & -10 \\ 4 & 3 & -1 & 3 \end{bmatrix}.$$

Since $\det\left[dG_{(1,1)}\right] = 11 \neq 0$, we know from the *inverse function theorem* that G is locally $\mathcal{C}^1$-invertible at $(1,1)$. That is, there exist a neighborhood W of $G(1,1;1,1) = (1,1;0,0)$ in $\mathbb{R}^4$ and a local $\mathcal{C}^1$-inverse $H : W \to \mathbb{R}^4$ of G. Write H in terms of its component functions, $H = (H^1, H^2)$ where $\operatorname{ran} H^1$ and $\operatorname{ran} H^2$ are contained in $\mathbb{R}^2$, and set $h(u,v) = H^2(u,v;0,0)$ for all (u,v) in $V := \{(u,v) : (u,v;0,0) \in W\}$. Then V is a neighborhood of $(1,1)$ in $\mathbb{R}^2$ and the function h is continuously differentiable because H is. We conclude that $h(1,1) = (1,1)$ from the following computation.

$$\begin{aligned}(1,1;1,1) &= H(G(1,1;1,1)) \\ &= H(1,1;f(1,1;1,1)) \\ &= (H^1(1,1;f(1,1;1,1))\,;\,H^2(1,1;f(1,1;1,1))) \\ &= (1,1\,;\,H^2(1,1;0,0)) \\ &= (1,1\,;\,h(1,1))\,.\end{aligned}$$

And from

$$\begin{aligned}(u,v;0,0) &= G(H(u,v;0,0)) \\ &= G(H^1(u,v;0,0);H^2(u,v;0,0)) \\ &= (H^1(u,v;0,0)\,;\,f(H^1(u,v;0,0);H^2(u,v;0,0))) \\ &= (u,v\,;\,f(u,v;h(u,v)))\end{aligned}$$

we conclude that (ii) holds; that is,

$$f(u,v;h(u,v)) = (0,0)$$

for all $u,\ v \in V$.

Q.30. Exercises in appendix D

Q.30.1. (Solution to D.1.1)

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Q.30.2. (Solution to D.3.2) First observe that the operation $\wedge$ is commutative and associative. (The former is obvious and the latter may be easily checked by means of a truth table.) Therefore if A, B, and C are propositions

$$\begin{aligned}A \wedge (B \wedge C) &\text{ iff } (A \wedge B) \wedge C \\ &\text{ iff } (B \wedge A) \wedge C \\ &\text{ iff } B \wedge (A \wedge C).\end{aligned} \tag{Q.32}$$

It then follows that

$$\begin{aligned}
(\exists x \in S)(\exists y \in T)\, P(x, y) &\text{ iff } (\exists x \in S)(\exists y)((y \in T) \wedge P(x, y)) \\
&\text{ iff } (\exists x)\big((x \in S) \wedge (\exists y)((y \in T) \wedge P(x, y))\big) \\
&\text{ iff } (\exists x)(\exists y)\big((x \in S) \wedge ((y \in T) \wedge P(x, y))\big) \\
&\text{ iff } (\exists x)(\exists y)\big((y \in T) \wedge ((x \in S) \wedge P(x, y))\big) \quad \text{(by (Q.32))} \\
&\text{ iff } (\exists y)(\exists x)\big((y \in T) \wedge ((x \in S) \wedge P(x, y))\big) \\
&\text{ iff } (\exists y)\big((y \in T) \wedge (\exists x)((x \in S) \wedge P(x, y))\big) \\
&\text{ iff } (\exists y)((y \in T) \wedge (\exists x \in S)\, P(x, y)) \\
&\text{ iff } (\exists y \in T)(\exists x \in S)\, P(x, y).
\end{aligned}$$

Notice that at the third and sixth steps we used the remark made in the last paragraph of section D.1.

Q.30.3. (Solution to D.4.4)

(1) P	(2) Q	(3) $P \Rightarrow Q$	(4) $\sim P$	(5) $Q \vee (\sim P)$
T	T	T	F	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

The third and fifth columns have the same truth values.

Q.31. Exercises in appendix F

Q.31.1. (Solution to F.2.4) If S, T, and U are sets, then

$$\begin{aligned}
x \in S \cup (T \cap U) \quad &\text{iff} \quad x \in S \text{ or } x \in T \cap U \\
&\text{iff} \quad x \in S \text{ or } (x \in T \text{ and } x \in U) \\
&\text{iff} \quad (x \in S \text{ or } x \in T) \text{ and } (x \in S \text{ or } x \in U) \\
&\text{iff} \quad x \in S \cup T \text{ and } x \in S \cup U \\
&\text{iff} \quad x \in (S \cup T) \cap (S \cup U).
\end{aligned}$$

Problem D.1.4 was used to get the third line.

Q.31.2. (Solution to F.2.9) If T is a set and $\mathfrak{S}$ is a family of sets, then

$$\begin{aligned}
x \in T \cup \big(\bigcap \mathfrak{S}\big) \quad &\text{iff} \quad x \in T \text{ or } x \in \bigcap \mathfrak{S} \\
&\text{iff} \quad x \in T \text{ or } (\forall S \in \mathfrak{S})\, x \in S \\
&\text{iff} \quad (\forall S \in \mathfrak{S})(x \in T \text{ or } x \in S) \\
&\text{iff} \quad (\forall S \in \mathfrak{S})\, x \in T \cup S \\
&\text{iff} \quad x \in \bigcap \{T \cup S \colon S \in \mathfrak{S}\}.
\end{aligned}$$

To obtain the third line we used the principle mentioned in the last paragraph of section D.1 of appendix D.

Q.31.3. (Solution to F.3.3) Here is one proof: A necessary and sufficient condition for an element x to belong to the complement of $S \cup T$ is that it not belong to S or to T. This is the equivalent to its belonging to both S^c and T^c, that is, to the intersection of the complements of S and T.

A second more "formalistic" proof looks like this :

$$\begin{aligned}
x \in (S \cup T)^c \quad &\text{iff} \quad x \notin S \cup T \\
&\text{iff} \quad \sim (x \in S \cup T) \\
&\text{iff} \quad \sim (x \in S \text{ or } x \in T) \\
&\text{iff} \quad \sim (x \in S) \text{ and } \sim (x \in T) \\
&\text{iff} \quad x \notin S \text{ and } x \notin T \\
&\text{iff} \quad x \in S^c \text{ and } x \in T^c \\
&\text{iff} \quad x \in S^c \cap T^c.
\end{aligned}$$

This second proof is not entirely without merit: at each step only one definition or fact is used. (For example, the result presented in example D.4.1 justifies the fourth "iff".) But on balance most readers, unless they are very unfamiliar with the material, would probably prefer the first version. After all, it's easier to read English than to translate code.

Q.31.4. (Solution to F.3.5) Here is another formalistic proof. It is a good idea to try and rewrite it in ordinary English.

$$\begin{aligned}
x \in (\bigcup \mathfrak{S})^c \quad &\text{iff} \quad x \notin \bigcup \mathfrak{S} \\
&\text{iff} \quad \sim (x \in \bigcup \mathfrak{S}) \\
&\text{iff} \quad \sim (\exists S \in \mathfrak{S})(x \in S) \\
&\text{iff} \quad (\forall S \in \mathfrak{S}) \sim (x \in S) \\
&\text{iff} \quad (\forall S \in \mathfrak{S})(x \notin S) \\
&\text{iff} \quad (\forall S \in \mathfrak{S})(x \in S^c) \\
&\text{iff} \quad x \in \bigcap \{S^c \colon S \in \mathfrak{S}\}.
\end{aligned}$$

Q.31.5. (Solution to F.3.9) To see that $S \setminus T$ and T are disjoint, notice that

$$\begin{aligned}
(S \setminus T) \cap T &= S \cap T^c \cap T \\
&= S \cap \emptyset \\
&= \emptyset.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
(S \setminus T) \cup T &= (S \cap T^c) \cup T \\
&= (S \cup T) \cap (T^c \cup T) \\
&= S \cup T.
\end{aligned}$$

As usual S and T are regarded as belonging to some universal set, say U. Then $T^c \cup T$ is all of U and its intersection with $S \cup T$ (which is contained in U) is just $S \cup T$.

Q.31.6. (Solution to F.3.10) We know from proposition F.1.2 (e) that $S \cup T = S$ if and only if $T \subseteq S$. But proposition F.3.9 tells us that $S \cup T = (S \setminus T) \cup T$. Thus $(S \setminus T) \cup T = S$ if and only if $T \subseteq S$.

Q.32. Exercises in appendix G

Q.32.1. (Solution to G.1.10) If $x + x = x$, then

$$\begin{aligned} x &= x + 0 \\ &= x + (x + (-x)) \\ &= (x + x) + (-x) \\ &= x + (-x) \\ &= 0. \end{aligned}$$

Q.32.2. (Solution to G.1.12) Use associativity and commutativity of addition.

$$\begin{aligned} (w + x) + (y + z) &= ((w + x) + y) + z \\ &= (w + (x + y)) + z \\ &= ((x + y) + w) + z \\ &= z + ((x + y) + w) \\ &= z + (x + (y + w)). \end{aligned}$$

The first, second, and last equalities use associativity of addition; steps 3 and 4 use its commutativity.

Q.33. Exercises in appendix H

Q.33.1. (Solution to H.1.5) By definition $x > 0$ holds if and only if $0 < x$, and this holds (again by definition) if and only if $x - 0 \in \mathbb{P}$. Since $-0 = 0$ (which is obvious from $0 + 0 = 0$ and the fact that the additive identity is unique), we conclude that $x > 0$ if and only if

$$x = x + 0 = x + (-0) = x - 0 \in \mathbb{P}\,.$$

Q.33.2. (Solution to H.1.6) By the preceding exercise $x > 0$ implies that $x \in \mathbb{P}$; and $y < z$ implies $z - y \in \mathbb{P}$. Since $\mathbb{P}$ is closed under multiplication, $x(z - y)$ belongs to $\mathbb{P}$. Thus

$$\begin{aligned} xz - xy &= xz + (-(xy)) \\ &= xz + x(-y) \qquad \text{by problem G.4.4} \\ &= x(z + (-y)) \\ &= x(z - y) \in \mathbb{P}\,. \end{aligned}$$

This shows that $xy < xz$.

Q.33.3. (Solution to H.1.12) Since $0 < w < x$ and $y > 0$, we may infer from exercise H.1.6 that $yw < yx$. Similarly, we obtain $xy < xz$ from the conditions $0 < y < z$ and $x > 0$ (which holds by the transitivity of $<$, proposition H.1.3). Then

$$wy = yw < yx = xy < xz\,.$$

Thus the desired inequality $wy < xz$ follows (again by transitivity of $<$).

Q.34. Exercises in appendix I

Q.34.1. (Solution to I.1.3) Since 1 belongs to A for every $A \in \mathfrak{A}$, it is clear that $1 \in \cap\mathfrak{A}$. If $x \in \cap\mathfrak{A}$, then $x \in A$ for every $A \in \mathfrak{A}$. Since each set A in $\mathfrak{A}$ is inductive, $x + 1 \in A$ for every $A \in \mathfrak{A}$. That is, $x + 1 \in \cap\mathfrak{A}$.

Q.34.2. (Solution to I.1.10) Let S be the set of all natural numbers for which the assertion is true. Certainly 1 belongs to S. If $n \in S$, then $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$. Therefore

$$\begin{aligned}\sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + (n+1)\\ &= \frac{1}{2}n(n+1) + (n+1)\\ &= \frac{1}{2}(n+1)(n+2),\end{aligned}$$

which shows that $n+1 \in S$. Thus S is an inductive subset of $\mathbb{N}$. We conclude from corollary I.1.8 that $S = \mathbb{N}$. In other words, the assertion holds for all $n \in \mathbb{N}$.

Q.34.3. (Solution to I.1.18) Let K be a subset of $\mathbb{N}$ which has no smallest member. We show $K = \emptyset$. Let

$$J = \{n \in \mathbb{N} \colon n < k \text{ for all } k \in K\}.$$

Certainly 1 belongs to J. [If not, there would exist $c \in K$ such that $1 \geq c$. From proposition I.1.6 we see that $c = 1$. Thus 1 belongs to K and is the smallest member of K, contrary to our assumption.]

Now suppose that $n \in J$ and prove that $n+1 \in J$. If $n+1 \notin J$, then there exists $k \in K$ such that $n+1 \geq k$. By the inductive hypothesis $n < k$. Thus $n < k \leq n+1$. We conclude from problem I.1.16(b) that $k = n+1$. But, since n is smaller than every member of K, this implies that $n+1$ is the smallest member of K. But K has no smallest member. Therefore we conclude that $n+1 \in J$.

We have shown that J is an inductive subset of $\mathbb{N}$. Then $J = \mathbb{N}$ (by theorem I.1.7). If K contains any element at all, say j, then $j \in J$; so in particular $j < j$. Since this is not possible, we conclude that $K = \emptyset$.

Q.35. Exercises in appendix J

Q.35.1. (Solution to J.2.7)

(a) A number x belongs to the set A if $x^2 - 4x + 3 < 3$; that is, if $x(x-4) < 0$. This occurs if and only if $x > 0$ and $x < 4$. Thus $A = (0, 4)$; so $\sup A = 4$ and $\inf A = 0$.

(b) Use beginning calculus to see that $f'(x) = 2x - 4$. Conclude that the function f is decreasing on the interval $(-\infty, 2)$ and is increasing on $(2, 3)$. Thus f assumes a minimum at $x = 2$. Since $f(2) = -1$, we see that $B = [-1, \infty)$. Thus $\sup B$ does not exist and $\inf B = -1$.

Q.35.2. (Solution to J.3.7) As in the hint let $\ell = \sup A$ and $m = \sup B$, and suppose that $\ell, m > 0$. If $x \in AB$, then there exist $a \in A$ and $b \in B$ such that $x = ab$. From $a \leq \ell$ and $b \leq m$ it is clear that $x \leq \ell m$; so ℓm is an upper bound for AB.

Since AB is bounded above it must have a least upper bound, say c. Clearly $c \leq \ell m$; we show that $\ell m \leq c$. Assume, to the contrary, that $c < \ell m$. Let $\epsilon = \ell m - c$. Since $\epsilon > 0$ and ℓ is the least upper bound for A we may choose an element a of A such that $a > \ell - \epsilon(2m)^{-1}$. Similarly, we may choose $b \in B$ so that $b > m - \epsilon(2\ell)^{-1}$. Then

$$\begin{aligned}ab &> \left(\ell - \epsilon(2m)^{-1}\right)\left(m - \epsilon(2\ell)^{-1}\right)\\ &= \ell m - \epsilon + \epsilon^2(4\ell m)^{-1}\\ &> \ell m - \epsilon\\ &= c.\end{aligned}$$

This is a contradiction, since ab belongs to AB and c is an upper bound of AB. We have shown

$$\sup(AB) = c = \ell m = (\sup A)(\sup B)$$

as required.

Remark. It is not particularly difficult to follow the details of the preceding proof. But that is *not* the same thing as *understanding* the proof! It is easy to see, for example, that *if* we choose $a > \ell - \epsilon(2m)^{-1}$ and $b > m - \epsilon(2\ell)^{-1}$, *then* $ab > c$. But that still leaves room to be puzzled. You might reasonably say when shown this proof, "Well, that certainly *is* a proof. And it looks very clever. But what I don't understand is how did you know to choose a and b in just that particular (or should I say 'peculiar'?) way? Do you operate by fits of inspiration, or a crystal ball, or divination of entrails, or what?" The question deserves an answer. Once we have assumed c to be an upper bound smaller than ℓm (and set $\epsilon = \ell m - c$), our hope is to choose $a \in A$ close to ℓ and $b \in B$ close to m in such a way that their product ab exceeds c. It is difficult to say immediately *how* close a should be to ℓ (and b to m). Let's just say that $a > \ell - \delta_1$ and $b > m - \delta_2$, where δ_1 and δ_2 are small positive numbers. We will figure out *how* small they should be in a moment. Then

$$ab > (\ell - \delta_1)(m - \delta_2) = \ell m - m\delta_1 - \ell\delta_2 + \delta_1\delta_2 .$$

Since $\delta_1\delta_2$ is positive, we can simplify the preceding inequality and write

$$ab > \ell m - m\delta_1 - \ell\delta_2 . \tag{Q.33}$$

What we *want* to get at the end of our computation is

$$ab > c = \ell m - \epsilon. \tag{Q.34}$$

Now comparing what we have (Q.33) with what we want (Q.34), we see that all we need to do is choose δ_1 and δ_2 in such a way that

$$m\delta_1 + \ell\delta_2 < \epsilon \tag{Q.35}$$

(for then $\ell m - (m\delta_1 + \ell\delta_2) > \ell m - \epsilon = c$, and we are done). To guarantee that the sum of two numbers is less than ϵ it suffices to choose both of them to be less than $\epsilon/2$. Clearly, we have $m\delta_1 < \epsilon/2$ if we choose $\delta_1 < \epsilon(2m)^{-1}$; and we have $\ell\delta_2 < \epsilon/2$ if we choose $\delta_2 < \epsilon(2\ell)^{-1}$. And that's all we need.

Q.35.3. (Solution to J.4.2) Let $A = \{t > 0 \colon t^2 < a\}$. The set A is not empty since it contains $a(1+a)^{-1}$. $[a^2(1+a)^{-2} < a(1+a)^{-1} < a.]$ It is easy to see that A is bounded above by $M := \max\{1, a\}$. [If $t \in A$ and $t \leq 1$, then $t \leq M$; on the other hand, if $t \in A$ and $t > 1$, then $t < t^2 < a \leq M$.] By the *least upper bound axiom* (J.3.1) A has a supremum, say x. It follows from the *axiom of trichotomy* (H.1.2) that exactly one of three things must be true: $x^2 < a$, $x^2 > a$, or $x^2 = a$. We show that $x^2 = a$ by eliminating the first two alternatives.

First assume that $x^2 < a$. Choose ϵ in $(0,1)$ so that $\epsilon < 3^{-1}x^{-2}(a - x^2)$. Then

$$(1+\epsilon)^2 = 1 + 2\epsilon + \epsilon^2 \tag{Q.36}$$

$$< 1 + 3\epsilon \tag{Q.37}$$

so that

$$x^2(1+\epsilon)^2 < x^2(1+3\epsilon) < a .$$

Thus $x(1+\epsilon)$ belongs to A. But this is impossible since $x(1+\epsilon) > x$ and x is the supremum of A.

Now assume $x^2 > a$. Choose ϵ in $(0,1)$ so that $\epsilon < (3a)^{-1}(x^2 - a)$. Then by (Q.36)

$$a < x^2(1+3\epsilon)^{-1} < x^2(1+\epsilon)^{-2} . \tag{Q.38}$$

Now since $x = \sup A$ and $x(1+\epsilon)^{-1} < x$, there must exist $t \in A$ such that $x(1+\epsilon)^{-1} < t < x$. But then

$$x^2(1+\epsilon)^{-2} < t^2 < a ,$$

which contradicts (Q.38). Thus we have demonstrated the existence of a number $x \geq 0$ such that $x^2 = a$. That there is only one such number has already been proved: see problem H.1.16.

Q.36. Exercises in appendix K

Q.36.1. (Solution to K.1.2) Suppose that $(x, y) = (u, v)$. Then

$$\big\{\{x, y\}, \{x\}\big\} = \big\{\{u, v\}, \{u\}\big\}.$$

We consider two cases.

Case 1: $\{x, y\} = \{u, v\}$ and $\{x\} = \{u\}$. The second equality implies that $x = u$. Then from the first equality we infer that $y = v$.

Case 2: $\{x, y\} = \{u\}$ and $\{x\} = \{u, v\}$. We derive $x = u = y$ from the first equality and $u = x = v$ from the second. Thus $x = y = u = v$. In either case $x = u$ and $y = v$. The converse is obvious.

Q.36.2. (Solution to K.3.7)

(a) $f(\frac{1}{2}) = 3$;

(b) Notice that $(1 - x)^{-1}$ does not exist if $x = 1$, $(1 + (1 - x)^{-1})^{-1}$ does not exist if $x = 2$, and $(1 - 2(1 + (1 - x)^{-1})^{-1})^{-1}$ does not exist if $x = 0$; so $\operatorname{dom} f = \mathbb{R} \setminus \{0, 1, 2\}$.

Q.36.3. (Solution to K.3.8) We can take the square root of $g(x) = -x^2 - 4x - 1$ only when $g(x) \geq 0$, and since we take its reciprocal, it should not be zero. But $g(x) > 0$ if and only if $x^2 + 4x + 1 < 0$ if and only if $(x + 2)^2 < 3$ if and only if $|x + 2| < \sqrt{3}$ if and only if $-2 - \sqrt{3} < x < -2 + \sqrt{3}$. So $\operatorname{dom} f = (-2 - \sqrt{3}, -2 + \sqrt{3})$.

Q.37. Exercises in appendix L

Q.37.1. (Solution to L.1.2) We may write A as the union of three intervals

$$A = (-4, 4) = (-4, -2) \cup [-2, 1) \cup [1, 4)\,.$$

Then

$$f^{\rightarrow}(A) = f^{\rightarrow}\big(\,(-4, 4)\,\big) = f^{\rightarrow}\big(\,(-4, -2)\,\big) \cup f^{\rightarrow}\big(\,[-2, 1)\,\big) \cup f^{\rightarrow}\big(\,[1, 4)\,\big)\,. \tag{Q.39}$$

(This step is justified in the next section by M.1.25.) Since f is constant on the interval $(-4, -2)$ we see that $f^{\rightarrow}\big(\,(-4, -2)\,\big) = \{-1\}$. On the interval $[-2, 1)$ the function increases from $f(-2) = 3$ to $f(0) = 7$ and then decreases to $f(1) = 6$ so that $f^{\rightarrow}\big(\,[-2, 1)\,\big) = [3, 7]$. (This interval is closed because both -2 and 0 belong to $[-2, 1)$.) Finally, since f is decreasing on $[1, 4)$ we see that $f^{\rightarrow}\big(\,[1, 4)\,\big) = \big(f(4), f(1)\big] = (\frac{1}{4}, 1]$. Thus from equation (Q.39) we conclude that

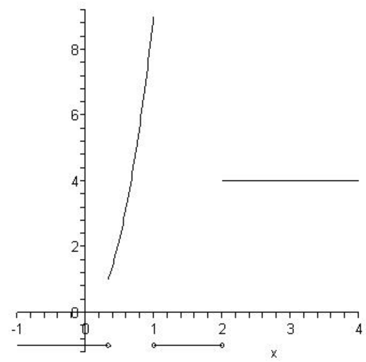
$$f^{\rightarrow}(A) = \{-1\} \cup (\tfrac{1}{4}, 1] \cup [3, 7].$$

Q.37.2. (Solution to L.1.3) Use techniques from beginning calculus. The function is a fourth degree polynomial, so $f(x) \to \infty$ as $x \to -\infty$ and as $x \to \infty$. Thus the range of f is not bounded above. The minimum value of the range will occur at a critical point, that is, at a point where $f'(x) = 0$. But this occurs at $x = -3$, $x = 0$, and $x = 2$. The values of f at these points are, respectively -188, 0, and -63. We conclude that $\operatorname{ran} f = [-188, \infty)$.

Q.37.3. (Solution to L.1.5) Notice that the arctangent function is strictly increasing (its derivative at each x is $(1+x^2)^{-1}$). Its range is $(-\pi/2, \pi/2)$. Thus $f^{\leftarrow}(B) = f^{\leftarrow}\big((\pi/4, 2)\big) = f^{\leftarrow}\big((\pi/4, \pi/2)\big) = (1, \infty)$.

Q.37.4. (Solution to L.1.6) For $-\sqrt{9 - x^2}$ to lie between 1 and 3, we would need $-3 < \sqrt{9 - x^2} < -1$. But since the square root function on $\mathbb{R}$ takes on only positive values, this is not possible. So $f^{\leftarrow}(B) = \emptyset$.

Q.37.5. (Solution to L.2.2) For $x \leq \frac{1}{3}$, $f(x) \leq 1$ which implies $g(f(x)) = -1$. For $x \in (\frac{1}{3}, 1)$, $f(x) \in (1, 3)$, so $g(f(x)) = 9x^2$. For $1 \leq x \leq 2$, $f(x) = 2$, so $g(f(x)) = -1$. Finally, for $x > 2$, $f(x) = 2$ and therefore $g(f(x)) = 4$.



Q.37.6. (Solution to L.2.3) Associativity: for every x

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)\,;$$

so $h \circ (g \circ f) = (h \circ g) \circ f$.

To see that composition is not commutative take, for example, $f(x) = x + 1$ and $g(x) = x^2$. Since $(g \circ f)(1) = 4$ and $(f \circ g)(1) = 2$, the functions $g \circ f$ and $f \circ g$ cannot be equal.

Q.38. Exercises in appendix M

Q.38.1. (Solution to M.1.2) If $f(x) = f(y)$, then $(x + 2)(3y - 5) = (y + 2)(3x - 5)$. Thus $6y - 5x = 6x - 5y$, which implies $x = y$.

Q.38.2. (Solution to M.1.3) Suppose that m and n are positive integers with no common prime factors. Let $f\left(\frac{m}{n}\right) = 2^m 3^n$. Then f is injective by the *unique factorization theorem* (see, for example, [2], page 21).

Q.38.3. (Solution to M.1.10) Let $f(x) = \frac{1}{x} - 1$ for $x \neq 0$ and $f(0) = 3$.

Q.38.4. (Solution to M.1.12) Define $f\colon \mathbb{Z} \to \mathbb{N}$ by

$$f(n) = \begin{cases} 2n + 1, & \text{for } n \geq 0 \\ -2n, & \text{for } n < 0. \end{cases}$$

Q.38.5. (Solution to M.1.13) Define $f\colon \mathbb{R} \to (0, 1)$ by $f(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x$.

Q.38.6. (Solution to M.1.14) Let $\mathbb{S}^1$ be $\{(x, y)\colon x^2 + y^2 = 1\}$. Define $f\colon [0, 1) \to \mathbb{S}^1$ by $f(t) = (\cos(2\pi t), \sin(2\pi t))$.

Q.38.7. (Solution to M.1.15) Let

$$g(x) = \begin{cases} 3 - 2x, & \text{for } 0 \leq x < 1 \\ f(x), & \text{for } 1 \leq x \leq 2 \\ \frac{1}{2}(3 - x), & \text{for } 2 < x \leq 3. \end{cases}$$

Q.38.8. (Solution to M.1.16) $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Q.38.9. (Solution to M.1.22)

(a) We show that if $y \in f^{\rightarrow}(f^{\leftarrow}(B))$, then $y \in B$. Suppose that $y \in f^{\rightarrow}(f^{\leftarrow}(B))$. Then (by the definition of $f^{\rightarrow}$) there exists $x \in f^{\leftarrow}(B)$ such that $y = f(x)$. From $x \in f^{\leftarrow}(B)$ we infer (using the definition of $f^{\leftarrow}$) that $f(x) \in B$. That is, $y \in B$.

(b) Let $f(x) = x^2$ and $B = \{-1\}$. Then $f^{\rightarrow}(f^{\leftarrow}(B)) = f^{\rightarrow}(f^{\leftarrow}\{-1\}) = f^{\rightarrow}(\emptyset) = \emptyset \neq B$.

(c) Suppose that f is surjective. Show that $B \subseteq f^{\rightarrow}(f^{\leftarrow}(B))$ by showing that $y \in B$ implies $y \in f^{\rightarrow}(f^{\leftarrow}(B))$. If $y \in B$, then (since f is surjective) there exists $x \in S$ such that $y = f(x)$. Since $f(x) \in B$, we see that $x \in f^{\leftarrow}(B)$ (by the definition of $f^{\leftarrow}$). From this it follows (using the definition of $f^{\rightarrow}$) that $y = f(x) \in f^{\rightarrow}(f^{\leftarrow}(B))$.

Q.38.10. (Solution to M.1.25) This requires nothing other than the definitions of $\cup$ and $f^{\rightarrow}$:

$$\begin{aligned} y \in f^{\rightarrow}(A \cup B) &\text{ iff there exists } x \in A \cup B \text{ such that } y = f(x) \\ &\text{ iff there exists } x \in A \text{ such that } y = f(x) \text{ or} \\ &\qquad \text{there exists } x \in B \text{ such that } y = f(x) \\ &\text{ iff } y \in f^{\rightarrow}(A) \text{ or } y \in f^{\rightarrow}(B) \\ &\text{ iff } y \in f^{\rightarrow}(A) \cup f^{\rightarrow}(B). \end{aligned}$$

Q.38.11. (Solution to M.1.27) Here the definitions of $\cap$ and $f^{\leftarrow}$ are used:

$$\begin{aligned} x \in f^{\leftarrow}(C \cap D) &\text{ iff } f(x) \in C \cap D \\ &\text{ iff } f(x) \in C \text{ and } f(x) \in D \\ &\text{ iff } x \in f^{\leftarrow}(C) \text{ and } x \in f^{\leftarrow}(D) \\ &\text{ iff } x \in f^{\leftarrow}(C) \cap f^{\leftarrow}(D). \end{aligned}$$

Q.38.12. (Solution to M.1.31)

(a) Show that if $y \in f^{\rightarrow}(\bigcap \mathfrak{A})$, then $y \in \bigcap\{f^{\rightarrow}(A) \colon A \in \mathfrak{A}\}$. Suppose that $y \in f^{\rightarrow}(\bigcap \mathfrak{A})$. Then there exists $x \in \bigcap \mathfrak{A}$ such that $y = f(x)$. Since x belongs to the intersection of the family $\mathfrak{A}$ it must belong to every member of $\mathfrak{A}$. That is, $x \in A$ for every $A \in \mathfrak{A}$. Thus $y = f(x)$ belongs to $f^{\rightarrow}(A)$ for every $A \in \mathfrak{A}$; and so $y \in \bigcap\{f^{\rightarrow}(A) \colon A \in \mathfrak{A}\}$.

(b) Suppose f is injective. If $y \in \bigcap\{f^{\rightarrow}(A) \colon A \in \mathfrak{A}\}$, then $y \in f^{\rightarrow}(A)$ for every $A \in \mathfrak{A}$. Choose a set $A_0 \in \mathfrak{A}$. Since $y \in f^{\rightarrow}(A_0)$, there exists $x_0 \in A_0$ such that $y = f(x_0)$. The point x_0 belongs to every member of $\mathfrak{A}$. To see this, let A be an arbitrary set belonging to $\mathfrak{A}$. Since $y \in f^{\rightarrow}(A)$, there exists $x \in A$ such that $y = f(x)$; and since $f(x) = y = f(x_0)$ and f is injective, we conclude that $x_0 = x \in A$. Thus we have shown that $x_0 \in \bigcap \mathfrak{A}$ and therefore that $y = f(x_0) \in f^{\rightarrow}(\bigcap \mathfrak{A})$.

(c) If $y \in f^{\rightarrow}(\bigcup \mathfrak{A})$, then there exists $x \in \bigcup \mathfrak{A}$ such that $y = f(x)$. Since $x \in \bigcup \mathfrak{A}$ there exists $A \in \mathfrak{A}$ such that $x \in A$. Then $y = f(x) \in f^{\rightarrow}(A)$ and so $y \in \bigcup\{f^{\rightarrow}(A) : A \in \mathfrak{A}\}$. Conversely, if y belongs to $\bigcup\{f^{\rightarrow}(A) : A \in \mathfrak{A}\}$, then it must be a member of $f^{\rightarrow}(A)$ for some $A \in \mathfrak{A}$. Then $y = f(x)$ for some $x \in A \subseteq \bigcup \mathfrak{A}$ and therefore $y = f(x) \in f^{\rightarrow}(\bigcup \mathfrak{A})$.

Q.38.13. (Solution to M.2.1) Let $f \colon S \to T$ and suppose that g and h are inverses of f. Then

$$g = g \circ I_T = g \circ (f \circ h) = (g \circ f) \circ h = I_S \circ h = h\,.$$

Q.38.14. (Solution to M.2.3) *Arcsine* is the inverse of the restriction of the *sine* function to the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The *arccosine* is the inverse of the restriction of *cosine* to $[0, \pi]$. And *arctangent* is the inverse of the restriction of *tangent* to $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Q.38.15. (Solution to M.2.4) Suppose that f has a right inverse f_r. For each $y \in T$ it is clear that $y = I_T(y) = f\big(f_r(y)\big) \in \operatorname{ran} f$; so $\operatorname{ran} f = T$ and f is surjective.

Conversely, suppose that f is surjective. Then for every $y \in T$ the set $f^{\leftarrow}(\{y\})$ is nonempty. For each $y \in T$ let x_y be a member of $f^{\leftarrow}(\{y\})$ and define

$$f_r \colon T \to S \colon y \mapsto x_y\,.$$

Then $f\big(f_r(y)\big) = f(x_y) = y$, showing that f_r is a right inverse of f. (The reader who has studied a bit of set theory will likely have noticed the unadvertised use of the *axiom of choice* in this proof. It is used in this fashion throughout the text.)

Q.39. Exercises in appendix N

Q.39.1. (Solution to N.1.4) The existence of the function has already been demonstrated: if $f = (f^1, f^2)$, then $(\pi_k \circ f)(t) = \pi_k\big(f^1(t), f^2(t)\big) = f^k(t)$ for $k = 1, 2$ and $t \in T$.

To prove uniqueness suppose that there is a function $g \in \mathcal{F}(T, S_1 \times S_2)$ such that $\pi_k \circ g = f^k$ for $k = 1, 2$. Then $g(t) = \big(\pi_1(g(t)), \pi_2(g(t))\big) = \big(f^1(t), f^2(t)\big) = \big(f^1, f^2\big)(t)$ for $k = 1, 2$ and $t \in T$. So $g = \big(f^1, f^2\big)$.

Q.40. Exercises in appendix O

Q.40.1. (Solution to O.1.4) We wish to demonstrate that for all natural numbers m and n if there is a bijection from $\{1, \dots, m\}$ onto $\{1, \dots, n\}$, then $m = n$. To accomplish this use induction on n.

First, suppose that for an arbitrary natural number m we have $\{1, \dots, m\} \sim \{1\}$. That is, we suppose that there exists a bijection f from $\{1, \dots, m\}$ onto $\{1\}$. Then since $f(1) = 1 = f(m)$ and f is injective, we conclude that $m = 1$. This establishes the proposition in the case $n = 1$.

Next, we assume the truth of the result for some particular $n \in \mathbb{N}$: for every $m \in \mathbb{N}$ if $\{1, \dots, m\} \sim \{1, \dots, n\}$, then $m = n$. This is our inductive hypothesis. What we wish to show is that for an arbitrary natural number m if $\{1, \dots, m\} \sim \{1, \dots, n+1\}$, then $m = n+1$. Suppose then that $m \in \mathbb{N}$ and $\{1, \dots, m\} \sim \{1, \dots, n+1\}$. Then there is a bijection f from $\{1, \dots, m\}$ onto $\{1, \dots, n+1\}$. Let $k = f^{-1}(n+1)$. The restriction of f to the set $\{1, \dots, k-1, k + 1, \dots, m\}$ is a bijection from that set onto $\{1, \dots, n\}$. Thus

$$\{1, \dots, k-1, k+1, \dots, m\} \sim \{1, \dots, n\}. \tag{Q.40}$$

Furthermore, it is easy to see that

$$\{1, \dots, m-1\} \sim \{1, \dots, k-1, k+1, \dots, m\}. \tag{Q.41}$$

(The required bijection is defined by $g(j) = j$ if $1 \le j \le k-1$ and $g(j) = j+1$ if $k \le j \le m-1$.) From (Q.40), (Q.41), and proposition O.1.2 we conclude that

$$\{1, \dots, m-1\} \sim \{1, \dots, n\}.$$

By our inductive hypothesis, $m - 1 = n$. This yields the desired conclusion $m = n+1$.

Q.40.2. (Solution to O.1.7) The result is trivial if S or T is empty; so we suppose they are not. Let $m = \operatorname{card} S$ and $n = \operatorname{card} T$. Then $S \sim \{1, \dots, m\}$ and $T \sim \{1, \dots, n\}$. It is clear that

$$\{1, \dots, n\} \sim \{m+1, \dots, m+n\}.$$

(Use the map $j \mapsto j + m$ for $1 \le j \le n$.) Thus $T \sim \{m+1, \dots, m+n\}$. Let $f\colon S \to \{1, \dots, m\}$ and $g\colon T \to \{m+1, \dots, m+n\}$ be bijections. Define $h\colon S \cup T \to \{1, \dots, m+n\}$ by

$$h(x) = \begin{cases} f(x), & \text{for } x \in S \\ g(x), & \text{for } x \in T. \end{cases}$$

Then clearly h is a bijection. So $S \cup T$ is finite and $\operatorname{card}(S \cup T) = m + n = \operatorname{card} S + \operatorname{card} T$.

Q.40.3. (Solution to O.1.8) Proceed by mathematical induction. If $C \subseteq \{1\}$, then either $C = \emptyset$, in which case $\operatorname{card} C = 0$, or else $C = \{1\}$, in which case $\operatorname{card} C = 1$. Thus the lemma is true if $n = 1$.

Suppose then that the lemma holds for some particular $n \in \mathbb{N}$. We prove its correctness for $n+1$. So we assume that $C \subseteq \{1, \dots, n+1\}$ and prove that C is finite and that $\operatorname{card} C \le n+1$. It is clear that $C \setminus \{n+1\} \subseteq \{1, \dots, n\}$. By the inductive hypothesis $C \setminus \{n+1\}$ is finite and $\operatorname{card}(C \setminus \{n+1\}) \le n$. There are two possibilities: $n+1 \notin C$ and $n+1 \in C$. In case $n+1$ does not belong to C, then $C = C \setminus \{n+1\}$; so C is finite and $\operatorname{card} C \le n < n+1$. In the other case, where $n+1$ does belong to C, it is clear that C is finite (because $C \setminus \{n+1\}$ is) and we have (by proposition O.1.7)

$$\begin{aligned}\operatorname{card} C &= \operatorname{card}\big((C \setminus \{n+1\}) \cup \{n+1\}\big)\\ &= \operatorname{card}(C \setminus \{n+1\}) + \operatorname{card}(\{n+1\})\\ &\le n+1.\end{aligned}$$

Q.40.4. (Solution to O.1.11) Suppose that S is infinite. We prove that there exists a proper subset T of S and a bijection f from S onto T. We choose a sequence of distinct elements a_k in S, one for each $k \in \mathbb{N}$. Let a_1 be an arbitrary member of S. Then $S \setminus \{a_1\} \neq \emptyset$. (Otherwise $S \sim \{a_1\}$ and S is finite.) Choose $a_2 \in S \setminus \{a_1\}$. Then $S \setminus \{a_1, a_2\} \neq \emptyset$. (Otherwise $S \sim \{a_1, a_2\}$ and S is finite.) In general, if distinct elements $a_1, \dots, a_n$ have been chosen, then $S \setminus \{a_1, \dots, a_n\}$ cannot be empty; so we may choose $a_{n+1} \in S \setminus \{a_1, \dots, a_n\}$. Let $T = S \setminus \{a_1\}$, and define $f \colon S \to T$ by

$$f(x) = \begin{cases} a_{k+1}, & \text{if } x = a_k \text{ for some } k\\ x, & \text{otherwise.}\end{cases}$$

Then f is a bijection from S onto the proper subset T of S.

For the converse construct a proof by contradiction. Suppose that $S \sim T$ for some proper subset $T \subseteq S$, and assume further that S is finite, so that $S \sim \{1, \dots, n\}$ for some $n \in \mathbb{N}$. Then by proposition O.1.9 the set $S \setminus T$ is finite and, since it is nonempty, is therefore cardinally equivalent to $\{1, \dots, p\}$ for some $p \in \mathbb{N}$. Thus

$$\begin{aligned} n &= \operatorname{card} S\\ &= \operatorname{card} T\\ &= \operatorname{card}(S \setminus (S \setminus T))\\ &= \operatorname{card} S - \operatorname{card}(S \setminus T) \qquad \text{(by problem O.1.10)}\\ &= n - p.\end{aligned}$$

Therefore $p = 0$, which contradicts the earlier assertion that $p \in \mathbb{N}$.

Q.40.5. (Solution to O.1.13) The map $x \mapsto \frac{1}{2}x$ is a bijection from the interval $(0, 1)$ onto the interval $(0, \frac{1}{2})$, which is a proper subset of $(0, 1)$.

Q.40.6. (Solution to O.1.15) Since f is surjective it has a right inverse f_r (see proposition M.2.4). This right inverse is injective, since it has a left inverse (see proposition M.2.5). Let $A = \operatorname{ran} f_r$. The function f_r establishes a bijection between T and A. Thus $T \sim A \subseteq S$. If S is finite, so is A (by proposition O.1.9) and therefore so is T.

Q.40.7. (Solution to O.1.16) Let $B = \operatorname{ran} f$. Then $S \sim B \subseteq T$. If T is finite, so is B (by proposition O.1.9) and therefore so is S.

Q.41. Exercises in appendix P

Q.41.1. (Solution to P.1.4) If S is finite there is nothing to prove; so we suppose that S is an infinite subset of T. Then T is countably infinite. Let $f \colon \mathbb{N} \to T$ be an enumeration of the members of T. The restriction of f to the set $f^{\leftarrow}(S) \subseteq \mathbb{N}$ is a bijection between $f^{\leftarrow}(S)$ and S; so we may conclude that S is countable provided we can prove that $f^{\leftarrow}(S)$ is. Therefore it suffices to show that *every* subset of $\mathbb{N}$ is countable.

Let A be an infinite subset of $\mathbb{N}$. Define inductively elements $a_1 < a_2 < \dots$ in A. (Let a_1 be the smallest member of A. Having chosen $a_1 < a_2 < \dots < a_n$ in A, notice that the set $A \setminus \{a_1, \dots, a_n\}$ is not empty and choose a_{n+1} to be the smallest element of that set.) Let $a \colon \mathbb{N} \to A$ be the function $n \mapsto a_n$. It is clear that $a_k \ge k$ for all k and, since $a_k < a_{k+1}$ for all k, that a is injective. To see that a is surjective, assume that it is not and derive a contradiction. If a is not surjective, then the range of a is a proper subset of A. Let p be the smallest element of $A \setminus \operatorname{ran} a$. Since

$p \in A \setminus \operatorname{ran} a \subseteq A \setminus \{a_1, \dots, a_p\}$, we see from the definition of a_{p+1} that $a_{p+1} \le p$. On the other hand we know that $a_{p+1} \ge p + 1 > p$. This contradiction shows that a is a surjection. Thus $A \sim \mathbb{N}$ proving that A is countable.

Q.41.2. (Solution to P.1.7) To see that the map

$$f \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \colon (m, n) \mapsto 2^{m-1}(2n - 1)$$

is a bijection, we construct its inverse (see propositions M.2.4 and M.2.5). If $p \in \mathbb{N}$ let m be the largest member of $\mathbb{N}$ such that 2^{m-1} divides p. (If p is odd, then $m = 1$.) Then $p/2^{m-1}$ is odd and can be written in the form $2n - 1$ for some $n \in \mathbb{N}$. The map $g \colon p \mapsto (m, n)$ is clearly the inverse of f.

Q.41.3. (Solution to P.1.11) If $\mathfrak{A}$ is infinite let

$$\mathfrak{A} = \{A_1, A_2, A_3, \dots\}\,;$$

while if $\mathfrak{A}$ is finite, say $\operatorname{card} \mathfrak{A} = m$, let

$$\mathfrak{A} = \{A_1, \dots, A_m\}$$

and let $A_n = A_m$ for all $n > m$. For each $j \in \mathbb{N}$ the set A_j is either infinite, in which case we write

$$A_j = \{a_{j1}, a_{j2}, a_{j3}, \dots\}\,,$$

or else it is finite, say $\operatorname{card} A_j = p$, in which case we write

$$A_j = \{a_{j1}, \dots, a_{jp}\}$$

and let $a_{jq} = a_{jp}$ for all $q > p$. Then the map

$$a \colon \mathbb{N} \times \mathbb{N} \to \bigcup \mathfrak{A} \colon (j, k) \mapsto a_{jk}$$

is surjective. Thus $\bigcup \mathfrak{A} = \bigcup_{j,k=1}^{\infty} A_{j,k} = \operatorname{ran} a$ is countable by lemma P.1.7 and proposition P.1.6.

Bibliography

[1] Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley Pub. Co., Reading, Mass., second edition, 1974. 201

[2] Garrett Birkhoff and Saunders Mac Lane. *A Survey of Modern Algebra*. Macmillan, New York, revised edition, 1953. 345

[3] Jean Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York, 1962. 43, 150

[4] Paul R. Halmos. *Finite-dimensional Vector Spaces*. D. Van Nostrand, Princeton, 1958. 124

[5] Paul R. Halmos. *A Hilbert Space Problem Book*. D. Van Nostrand, Princeton, Toronto, London, 1967. xv

[6] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice Hall, Englewood Cliffs, 1971. 124

[7] Irving Kaplansky. *Set Theory and Metric Spaces*. Chelsea Publ. Co., New York, 1977. 277

[8] Lynn H. Loomis and Shlomo Sternberg. *Advanced Calculus*. Jones and Bartlett, Boston, 1990. xii

[9] C. N. Moore. Infinite series and sequences. *Amer. Mathematical Monthly*, 39:62–71, 1932. 195

[10] Michael Spivak. *Calculus*. W. A. Benjamin, New York, 1967. 248

[11] Karl R. Stromberg. *An Introduction to Classical Real Analysis*. Wadsworth, Belmont, 1981. 278

Index