TRIM 38

Tao

Analysis II

trim

TEXTS AND READINGS IN MATHEMATICS 38

Analysis II

Terence Tao

HINDUSTAN BOOK AGENCY

TEXTS AND READINGS IN MATHEMATICS 38

Analysis II

Analysis II

Terence Tao
University of California
Los Angeles

HINDUSTAN
BOOK AGENCY

Published by

Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: hba@vsnl.com
http://www.hindbook.com

Produced from camera ready copy supplied by the Author.

ISBN 81-85931-63-1

To my parents, for everything

Contents

Volume 1

Volume 2

Preface

This text originated from the lecture notes I gave teaching the honours undergraduate-level real analysis sequence at the University of California, Los Angeles, in 2003. Among the undergraduates here, real analysis was viewed as being one of the most difficult courses to learn, not only because of the abstract concepts being introduced for the first time (e.g., topology, limits, measurability, etc.), but also because of the level of rigour and proof demanded of the course. Because of this perception of difficulty, one was often faced with the difficult choice of either reducing the level of rigour in the course in order to make it easier, or to maintain strict standards and face the prospect of many undergraduates, even many of the bright and enthusiastic ones, struggling with the course material.

Faced with this dilemma, I tried a somewhat unusual approach to the subject. Typically, an introductory sequence in real analysis assumes that the students are already familiar with the real numbers, with mathematical induction, with elementary calculus, and with the basics of set theory, and then quickly launches into the heart of the subject, for instance the concept of a limit. Normally, students entering this sequence do indeed have a fair bit of exposure to these prerequisite topics, though in most cases the material is not covered in a thorough manner. For instance, very few students were able to actually *define* a real number, or even an integer, properly, even though they could visualize these numbers intuitively and manipulate them algebraically. This seemed to me to be a missed opportunity. Real analysis is one of the first subjects (together with linear algebra and abstract algebra) that a student encounters, in which one truly has to grapple with the subtleties of a truly rigorous mathematical proof. As such, the course offered an excellent chance to go back to the foundations of mathematics, and in particular the opportunity to do a proper and thorough construction of the real numbers.

Thus the course was structured as follows. In the first week, I described some well-known "paradoxes" in analysis, in which standard laws of the subject (e.g., interchange of limits and sums, or sums and integrals) were applied in a non-rigorous way to give nonsensical results such as $0 = 1$. This motivated the need to go back to the very beginning of the subject, even to the very definition of the natural numbers, and check all the foundations from scratch. For instance, one of the first homework assignments was to check (using only the Peano axioms) that addition was associative for natural numbers (i.e., that $(a+b)+c = a+(b+c)$ for all natural numbers a, b, c: see Exercise 2.2.1). Thus even in the first week, the students had to write rigorous proofs using mathematical induction. After we had derived all the basic properties of the natural numbers, we then moved on to the integers (initially defined as formal differences of natural numbers); once the students had verified all the basic properties of the integers, we moved on to the rationals (initially defined as formal quotients of integers); and then from there we moved on (via formal limits of Cauchy sequences) to the reals. Around the same time, we covered the basics of set theory, for instance demonstrating the uncountability of the reals. Only then (after about ten lectures) did we begin what one normally considers the heart of undergraduate real analysis - limits, continuity, differentiability, and so forth.

The response to this format was quite interesting. In the first few weeks, the students found the material very easy on a conceptual level, as we were dealing only with the basic properties of the standard number systems. But on an intellectual level it was very challenging, as one was analyzing these number systems from a foundational viewpoint, in order to rigorously derive the more advanced facts about these number systems from the more primitive ones. One student told me how difficult it was to explain to his friends in the non-honours real analysis sequence (a) why he was still learning how to show why all rational numbers are either positive, negative, or zero (Exercise 4.2.4), while the non-honours sequence was already distinguishing absolutely convergent and conditionally convergent series, and (b) why, despite this, he thought his homework was significantly harder than that of his friends. Another student commented to me, quite wryly, that while she could obviously *see* why one could always divide a natural number n into a positive integer q to give a quotient a and a remainder r less than q (Exercise 2.3.5), she still had, to her frustration, much difficulty in writing down a proof of this fact. (I told her that later in the course she would have to prove statements for which it would not be as obvious to see that the statements were true; she did not seem to be particularly consoled by this.) Nevertheless, these students greatly enjoyed the homework, as when they did persevere and obtain a rigorous proof of an intuitive fact, it solidifed the link in their minds between the abstract manipulations of formal mathematics and their informal intuition of mathematics (and of the real world), often in a very satisfying way. By the time they were assigned the task of giving the infamous "epsilon and delta" proofs in real analysis, they had already had so much experience with formalizing intuition, and in discerning the subtleties of mathematical logic (such as the distinction between the "for all" quantifier and the "there exists" quantifier), that the transition to these proofs was fairly smooth, and we were able to cover material both thoroughly and rapidly. By the tenth week, we had caught up with the non-honours class, and the students were verifying the change of variables formula for Riemann-Stieltjes integrals, and showing that piecewise continuous functions were Riemann integrable. By the conclusion of the sequence in the twentieth week, we had covered (both in lecture and in homework) the convergence theory of Taylor and Fourier series, the inverse and implicit function theorem for continuously differentiable functions of several variables, and established the dominated convergence theorem for the Lebesgue integral.

In order to cover this much material, many of the key foundational results were left to the student to prove as homework; indeed, this was an essential aspect of the course, as it ensured the students truly appreciated the concepts as they were being introduced. This format has been retained in this text; the majority of the exercises consist of proving lemmas, propositions and theorems in the main text. Indeed, I would strongly recommend that one do as many of these exercises as possible - and this includes those exercises proving "obvious" statements - if one wishes to use this text to learn real analysis; this is not a subject whose subtleties are easily appreciated just from passive reading. Most of the chapter sections have a number of exercises, which are listed at the end of the section.

To the expert mathematician, the pace of this book may seem somewhat slow, especially in early chapters, as there is a heavy emphasis on rigour (except for those discussions explicitly marked "Informal"), and justifying many steps that would ordinarily be quickly passed over as being self-evident. The first few chapters develop (in painful detail) many of the "obvious" properties of the standard number systems, for instance that the sum of two positive real numbers is again positive (Exercise 5.4.1), or that given any two distinct real numbers, one can find rational number between them (Exercise 5.4.5). In these foundational chapters, there is also an emphasis on *non-circularity* - not using later, more advanced results to prove earlier, more primitive ones. In particular, the usual laws of algebra are not used until they are derived (and they have to be derived separately for the natural numbers, integers, rationals, and reals). The reason for this is that it allows the students to learn the art of abstract reasoning, deducing true facts from a limited set of assumptions, in the friendly and intuitive setting of number systems; the payoff for this practice comes later, when one has to utilize the same type of reasoning techniques to grapple with more advanced concepts (e.g., the Lebesgue integral).

The text here evolved from my lecture notes on the subject, and thus is very much oriented towards a pedagogical perspective; much of the key material is contained inside exercises, and in many cases I have chosen to give a lengthy and tedious, but instructive, proof instead of a slick abstract proof. In more advanced textbooks, the student will see shorter and more conceptually coherent treatments of this material, and with more emphasis on intuition than on rigour; however, I feel it is important to know how to do analysis rigorously and "by hand" first, in order to truly appreciate the more modern, intuitive and abstract approach to analysis that one uses at the graduate level and beyond.

The exposition in this book heavily emphasizes rigour and formalism; however this does not necessarily mean that lectures based on this book have to proceed the same way. Indeed, in my own teaching I have used the lecture time to present the intuition behind the concepts (drawing many informal pictures and giving examples), thus providing a complementary viewpoint to the formal presentation in the text. The exercises assigned as homework provide an essential bridge between the two, requiring the student to combine both intuition and formal understanding together in order to locate correct proofs for a problem. This I found to be the most difficult task for the students, as it requires the subject to be genuinely *learnt*, rather than merely memorized or vaguely absorbed. Nevertheless, the feedback I received from the students was that the homework, while very demanding for this reason, was also very rewarding, as it allowed them to connect the rather abstract manipulations of formal mathematics with their innate intuition on such basic concepts as numbers, sets, and functions. Of course, the aid of a good teaching assistant is invaluable in achieving this connection.

With regard to examinations for a course based on this text, I would recommend either an open-book, open-notes examination with problems similar to the exercises given in the text (but perhaps shorter, with no unusual trickery involved), or else a take-home examination that involves problems comparable to the more intricate exercises in the text. The subject matter is too vast to force the students to memorize the definitions and theorems, so I would not recommend a closed-book examination, or an examination based on regurgitating extracts from the book. (Indeed, in my own examinations I gave a supplemental sheet listing the key definitions and theorems which were relevant to the examination problems.) Making the examinations similar to the homework assigned in the course will also help motivate the students to work through and understand their homework problems as thoroughly as possible (as opposed to, say, using flash cards or other such devices to memorize material), which is good preparation not only for examinations but for doing mathematics in general.

Some of the material in this textbook is somewhat peripheral to the main theme and may be omitted for reasons of time constraints. For instance, as set theory is not as fundamental to analysis as are the number systems, the chapters on set theory (Chapters 3, 8) can be covered more quickly and with substantially less rigour, or be given as reading assignments. The appendices on logic and the decimal system are intended as optional or supplemental reading and would probably not be covered in the main course lectures; the appendix on logic is particularly suitable for reading concurrently with the first few chapters. Also, Chapter 16 (on Fourier series) is not needed elsewhere in the text and can be omitted.

For reasons of length, this textbook has been split into two volumes. The first volume is slightly longer, but can be covered in about thirty lectures if the peripheral material is omitted or abridged. The second volume refers at times to the first, but can also be taught to students who have had a first course in analysis from other sources. It also takes about thirty lectures to cover.

I am deeply indebted to my students, who over the progression of the real analysis course corrected several errors in the lectures notes from which this text is derived, and gave other valuable feedback. I am also very grateful to the many anonymous referees who made several corrections and suggested many important improvements to the text.

Terence Tao

Chapter 12

Metric spaces

12.1 Definitions and examples

In Definition 6.1.5 we defined what it meant for a sequence $(x_n)_{n=m}^{\infty}$ of real numbers to converge to another real number x; indeed, this meant that for every $\varepsilon > 0$, there exists an $N \geq m$ such that $|x - x_n| \leq \varepsilon$ for all $n \geq N$. When this is the case, we write $\lim_{n\to\infty} x_n = x$.

Intuitively, when a sequence $(x_n)_{n=m}^{\infty}$ converges to a limit x, this means that somehow the elements x_n of that sequence will eventually be as close to x as one pleases. One way to phrase this more precisely is to introduce the *distance function* $d(x, y)$ between two real numbers by $d(x, y) := |x - y|$. (Thus for instance $d(3, 5) = 2$, $d(5, 3) = 2$, and $d(3, 3) = 0$.) Then we have

Lemma 12.1.1. *Let $(x_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let x be another real number. Then $(x_n)_{n=m}^{\infty}$ converges to x if and only if $\lim_{n\to\infty} d(x_n, x) = 0$.*

Proof. See Exercise 12.1.1. □

One would now like to generalize this notion of convergence, so that one can take limits not just of sequences of real numbers, but also sequences of complex numbers, or sequences of vectors, or sequences of matrices, or sequences of functions, even sequences of sequences. One way to do this is to redefine the notion of convergence each time we deal with a new type of object. As you can guess, this will quickly get tedious. A more efficient way is to work *abstractly*, defining a very general class of spaces - which includes such standard spaces as the real numbers, complex numbers, vectors, etc. - and define the notion of convergence on this entire class of spaces at once. (A *space* is just the set of all objects of a certain type - the space of all real numbers, the space of all 3×3 matrices, etc. Mathematically, there is not much distinction between a space and a set, except that spaces tend to have much more structure than what a random set would have. For instance, the space of real numbers comes with operations such as addition and multiplication, while a general set would not.)

It turns out that there are two very useful classes of spaces which do the job. The first class is that of *metric spaces*, which we will study here. There is a more general class of spaces, called *topological spaces*, which is also very important, but we will only deal with this generalization briefly, in Section 13.5.

Roughly speaking, a metric space is any space X which has a concept of *distance* $d(x, y)$ - and this distance should behave in a reasonable manner. More precisely, we have

Definition 12.1.2 (Metric spaces). A *metric space* (X, d) is a space X of objects (called *points*), together with a *distance function* or *metric* $d : X \times X \to [0, +\infty)$, which associates to each pair x, y of points in X a non-negative real number $d(x, y) \geq 0$. Furthermore, the metric must satisfy the following four axioms:

(a) For any $x \in X$, we have $d(x, x) = 0$.

(b) (Positivity) For any *distinct* $x, y \in X$, we have $d(x, y) > 0$.

(c) (Symmetry) For any $x, y \in X$, we have $d(x, y) = d(y, x)$.

(d) (Triangle inequality) For any $x, y, z \in X$, we have $d(x, z) \leq d(x, y) + d(y, z)$.

In many cases it will be clear what the metric d is, and we shall abbreviate (X, d) as just X.

Remark 12.1.3. The conditions (a) and (b) can be rephrased as follows: for any $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$. (Why is this equivalent to (a) and (b)?)

Example 12.1.4 (The real line). Let $\mathbf{R}$ be the real numbers, and let $d : \mathbf{R} \times \mathbf{R} \to [0, \infty)$ be the metric $d(x, y) := |x - y|$ mentioned earlier. Then $(\mathbf{R}, d)$ is a metric space (Exercise 12.1.2). We refer to d as the *standard metric* on $\mathbf{R}$, and if we refer to $\mathbf{R}$ as a metric space, we assume that the metric is given by the standard metric d unless otherwise specified.

Example 12.1.5 (Induced metric spaces). Let (X, d) be any metric space, and let Y be a subset of X. Then we can restrict the metric function $d : X \times X \to [0, +\infty)$ to the subset $Y \times Y$ of $X \times X$ to create a restricted metric function $d|_{Y \times Y} : Y \times Y \to [0, +\infty)$ of Y; this is known as the metric on Y *induced* by the metric d on X. The pair $(Y, d|_{Y \times Y})$ is a metric space (Exercise 12.1.4) and is known the *subspace* of (X, d) induced by Y. Thus for instance the metric on the real line in the previous example induces a metric space structure on any subset of the reals, such as the integers $\mathbf{Z}$, or an interval $[a, b]$, etc.

Example 12.1.6 (Euclidean spaces). Let $n \geq 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1, \ldots, x_n \in \mathbf{R}\}.$$

We define the *Euclidean metric* (also called the *l^2 metric*) $d_{l^2} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$\begin{aligned} d_{l^2}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &:= \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \\ &= (\sum_{i=1}^{n} (x_i - y_i)^2)^{1/2}. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^2}((1, 6), (4, 2)) = \sqrt{3^2 + 4^2} = 5$. This metric corresponds to the geometric distance between the two points $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$ as given by Pythagoras' theorem. (We remark however that while geometry does give some very important examples of metric spaces, it is possible to have metric spaces which have no obvious geometry whatsoever. Some examples are given below.) The verification that $(\mathbf{R}^n, d)$ is indeed a metric space can be seen geometrically (for instance, the triangle inequality now asserts that the length of one side of a triangle is always less than or equal to the sum of the lengths of the other two sides), but can also be proven algebraically (see Exercise 12.1.6). We refer to $(\mathbf{R}^n, d_{l^2})$ as the *Euclidean space* of *dimension* n.

Example 12.1.7 (Taxi-cab metric). Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^1}, the so-called *taxicab metric* (or l^1 *metric*), defined by

$$\begin{aligned} d_{l^1}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) &:= |x_1 - y_1| + \ldots + |x_n - y_n| \\ &= \sum_{i=1}^{n} |x_i - y_i|. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^1}((1,6),(4,2)) = 5 + 2 = 7$. This metric is called the taxi-cab metric, because it models the distance a taxi-cab would have to traverse to get from one point to another if the cab was only allowed to move in cardinal directions (north, south, east, west) and not diagonally. As such it is always at least as large as the Euclidean metric, which measures distance "as the crow flies", as it were. We claim that the space $(\mathbf{R}^n, d_{l^1})$ is also a metric space (Exercise 12.1.7). The metrics are not quite the same, but we do have the inequalities

$$d_{l^2}(x, y) \leq d_{l^1}(x, y) \leq \sqrt{n} d_{l^2}(x, y) \qquad (12.1)$$

for all x, y (see Exercise 12.1.8).

Remark 12.1.8. The taxi-cab metric is useful in several places, for instance in the theory of error correcting codes. A string of n binary digits can be thought of as an element of $\mathbf{R}^n$, for instance the binary string 10010 can be thought of as the point $(1, 0, 0, 1, 0)$ in $\mathbf{R}^5$. The taxi-cab distance between two binary strings is then the number of bits in the two strings which do not match, for instance $d_{l^1}(10010, 10101) = 3$. The goal of error-correcting codes is to encode each piece of information (e.g., a letter of the alphabet) as a binary string in such a way that all the binary strings are as far away in the taxicab metric from each other as possible; this minimizes the chance that any distortion of the bits due to random noise can accidentally change one of the coded binary strings to another, and also maximizes the chance that any such distortion can be detected and correctly repaired.

Example 12.1.9 (Sup norm metric). Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^∞}, the so-called *sup norm metric* (or *l^∞ metric*), defined by

$$d_{l^\infty}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) := \sup\{|x_i - y_i| : 1 \leq i \leq n\}.$$

Thus for instance, if $n = 2$, then $d_{l^\infty}((1,6), (4,2)) = \sup(5,2) = 7$. The space $(\mathbf{R}^n, d_{l^\infty})$ is also a metric space (Exercise 12.1.9), and is related to the l^2 metric by the inequalities

$$\frac{1}{\sqrt{n}} d_{l^2}(x,y) \leq d_{l^\infty}(x,y) \leq d_{l^2}(x,y) \qquad (12.2)$$

for all x, y (see Exercise 12.1.10).

Remark 12.1.10. The l^1, l^2, and l^∞ metrics are special cases of the more general *l^p metrics*, where $p \in [1, +\infty]$, but we will not discuss these more general metrics in this text.

Example 12.1.11 (Discrete metric). Let X be an arbitrary set (finite or infinite), and define the *discrete metric* d_{disc} by setting $d_{disc}(x,y) := 0$ when $x = y$, and $d_{disc}(x,y) := 1$ when $x \neq y$. Thus, in this metric, all points are equally far apart. The space (X, d_{disc}) is a metric space (Exercise 12.1.11). Thus every set X has at least one metric on it.

Example 12.1.12 (Geodesics). (Informal) Let X be the sphere $\{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\}$, and let $d((x,y,z),(x',y',z'))$ be the length of the shortest curve in X which starts at (x,y,z) and ends at (x',y',z'). (This curve turns out to be an arc of a