MATH 312

Analysis

2012

Introduction.

In the rather vague division of mathematics into subordinate topics, "analysis" may be crudely described as the branch that deals directly or indirectly with limiting processes, like the taking of limits of infinite sequences of functions or the summation of infinite series, whilst "algebra" is concerned with "finitary" processes like the addition or multiplication of two (or at most only finitely many) things at a time. Such descriptions should not be taken too seriously, since, as one progresses, one finds more and more that the supposedly different branches of mathematics overlap both conceptually and practically. The questions an "analyst" deals with may often be expressed in "algebraic" terms, for instance. But it is indisputable that limiting processes are at the heart of analysis.

[The idea of a limit, when you try to catch it in a definition that may serve for proofs, depends at the outset on *inequalities*. (I have attached Appendix A, which I sometimes provided as a handout—not only in 312—to try to explain the thought that lies behind the standard definition of a limit.) One might say that analysis largely consists of disguised applications of inequalities. In particular, you should be careful in what follows to check that any inequalities I use are in fact true.]

Our first aim, then, is to introduce you to some of the fundamental ideas of rigorous analysis. The first- and second-year courses in "calculus" hint at many of the ideas we shall discuss, but their main emphasis was on "methods"—this is how such-and-such is done. This order of introducing material (methods before proofs) certainly reflects the modern development of the subject, and fits the purposes for which most people study it. The great mathematicians of the not-so-recent past did amazing things, but they never clarified the concepts they were playing with. People like Euler, Gauss, Leibniz, and Newton, and many less famous names, were stupendously talented, and their arguments remain fruitful today; but they rarely *defined* anything with the precision we now expect, partly because there were so many things that seemed to them obviously and automatically true. Indeed, there were from time to time squabbles amongst them that seem bizarre to us because we can see they were really about the meanings of the terms used, which they believed at the time to be uncontentious. A famous example is the debate about the sum of the series $1-1+1-1+\cdots$, which Leibniz maintained—not without some reason—to be $\frac{1}{2}$. (The problem, as we now see it, is what you understand by the sum of a series.) Most theorems in analysis before the nineteenth century were, indeed, stated without exact and explicit hypotheses. It is a historical curiosity that, because of Euclid, geometry was far more logically developed, albeit on somewhat shaky foundations.

Nevertheless, this casual approach to fundamentals, despite its fitness for students who are not primarily interested in the subject for its own sake, has the consequence that the ideas of serious contemporary mathematics (which have important practical consequences) eventually come as a far greater shock than they should. When mathematics had reached a certain stage, it became apparent that it was *necessary* to have rigorous definitions. Some things that had seemed intuitively plausible were turning out to be false in some circumstances, and some assertions that people wished to deny could apparently be "proved". One should not overstate the case; these people weren't idiots by any means. But conceptual confusion was only cleared up slowly, step by step. I am not a historian, and am far from sure whether there was a single crucial moment, but it is arguable that Dirichlet's work on Fourier series (1829) was a turning point. Fourier had stated, and believed, that "every function on the interval $(-\pi, \pi)$ has a Fourier series"; some people found this—again, for reasonable reasons—incredible, and every significant step of one "proof" he offered was false. Like others at the time, he had only a vague idea what he meant by a "function", or by "convergence". Dirichlet gave some fairly strong conditions and proved that a function satisfying them does have a Fourier series in a reasonable, albeit limited, sense. Some more subtle versions of the question remain unanswered to this day.

Another crux was the introduction of non-Euclidean geometry, which suggested that some statements that had been thought 'obviously and automatically true' (the parallels postulate, in particular) need not necessarily correspond to "reality", and so raised the question what the logical foundations of geometry really should be. Putting mathematics on a reasonably firm basis, by establishing exactly what meanings to attach to various vague ideas and what should be assumed about them to deduce various familiar theorems, took rather a long time—seventy years or more—, and progress was not simple because various gaps in the theory were filled in random order as people became aware of them. When this process was more or less finished, soon after 1900, other mathematicians began to give a more abstract and general cast to many of the ideas (and others were already worrying, as they still do, about the logical foundations).

The first aim of the course, then, is to give a necessarily brief but in essence quite rigorous discussion of the foundations of real analysis as they were standardized around the end of the nineteenth century. Our treatment is not, however, at all chronological; we take for granted some of the basic ideas from courses like 151, 161, or 251, such as "set", "function", "field", "equivalence relation", or "partial order". Many of them were first clearly and precisely formulated in the early twentieth century, but they are nowadays familiar. And our proofs of many results are, similarly, of a rather more modern type than those originally invented.

The second part of the course deals with *complex* analysis—the theory of suitable complex-valued functions of a complex variable. It has a rather different flavour from the first part. Firstly, the kind of function it studies is extremely restricted, and secondly, it is in some ways necessarily less rigorous (at least in a first course). The theorems are so startling and have such remarkable consequences that it is undesirable to worry too much whether they are being presented in the best possible form from a logical standpoint. They are certainly true in the rather simple cases we present; but it takes a lot of time to expound good general versions. The whole theory used to be of great importance in many practical contexts: it is the only reasonably general way of deriving many amazing exact formulæ and of carrying out many complicated exact calculations. With the development of computers and numerical methods it has lost some, though by no means most, of its practical value, but it remains indispensable for many more theoretical applications.

From all this, you will have grasped that in some ways this course is more theoretical and abstract than you are perhaps used to. It cannot be too much stressed that *WE SHALL NOT DO ANYTHING VERY DEEP OR DIFFICULT IN ITSELF*. The novelty is the emphasis on

concepts and on *proofs* in arguments about functions. We try to make, and to justify, precise and general statements, instead of just introducing rules of thumb. Oddly enough, *abstract statements, once understood, are often a lot easier to prove*, because the ideas they use are adapted to the ends in view and stripped of irrelevancies. Many great results were originally proved with considerable difficulty in particular circumstances, but turned out to be both more general and more straightforward once an appropriate conceptual framework had been discovered. The difficulty was absorbed into the concepts, and the proofs became trifling. However, this obviously implies that the concepts are fundamental; hence, you must learn the meanings of many unfamiliar mathematical terms.

We try to mark our definitions carefully, but THEY MUST BE LEARNT. If you do not learn them, you cannot possibly understand the course. A fairly significant—though not overwhelming—number of marks in the final examination depends on regurgitating, or at least presenting good formulations of, definitions. Many first-year students these days seem to think it unreasonable to learn definitions in a mathematics course; but you would not expect to study French without learning what large numbers of French words "mean", even if you had a French–English dictionary to hand. The analogy with foreign languages is, of course, very imperfect. A mathematical definition tells you what the concept denoted by a word actually *is* (in terms of more elementary concepts, that is), whereas the dictionary of a foreign language simply lists approximate English equivalents so that you can on the whole use the foreign word correctly, and perhaps, in the long run, react to it in much the same way as a native speaker. This is perhaps why some beginning students have problems with definitions in mathematics, because they believe a definition only attaches a *word* to a pre-existing concept or group of concepts, as a dictionary does, instead of telling you the exact content of an entirely new one.

The best strategy is to **understand** each definition, so that you can reconstruct it in your own words whenever you want. "Understanding" a definition includes, amongst other things, grasping what it excludes and what it includes: trying to invent objects that conform to the definition, or that seem at first glance as though they might but in fact do not. This takes time, and involves very careful and detailed reading of your notes. A mathematical exposition is not an airport novel; our notes are not intended to be skipped through, although I hope they are more accessible, being much shorter, than most books on the subject. In the interim, before the penny drops as it were, learning a definition by heart is the next best thing to understanding it and should not be despised, since small changes in wording may make great alterations in meaning. (Take no notice of the University's ranting about plagiarism, by the way; its only comprehensible purpose is to ensure that students caught cheating can be expelled without legal argument.) Incidentally, my definitions will occasionally differ slightly from Rudin's—generally, I hope, on reasonable grounds. Mathematical definitions are not regulated by any central authority; conventions that were accepted in the 1920s and 1930s in Texas may be odd by the standards of Russia in the 1950s, and so on. In reading a mathematical text, you should always take note of the definitions the author uses.

There is a problem with the wording of definitions. A mathematical definition's purpose, as I said, is to tell us precisely what something is, so that there is a solid and unambiguous basis for later argument. If you are asked for the definition of the inverse of a matrix, it is not correct to answer "If B is an inverse of the $n \times n$ matrix A, then $AB = BA = I_n$". This is the wrong way round; instead of saying what an "inverse" is, it presents a property the inverse has; formally speaking, the sentence would tell us nothing unless we already knew that B is indeed an inverse of A. The intended thought is better expressed by the sentence "If $AB = BA = I_n$, then B is called an inverse of A", which makes it explicit that we are concerned with introducing the concept attached to the word "inverse" and not with a statement of fact.

[The bald sentence just given, though the core of a possible definition, is inadequate on its own. It assumes at least another sentence or two explaining what the symbols represent, such as "Let A and B be $n \times n$ real matrices." Although I_n is a well-known symbol, it would do no harm to add "$\cdots$ and let I_n denote

the $n \times n$ identity matrix". In general, then, a definition, set out in full, may consist of three parts. There may be an introduction, possibly of great length, that, as it were, sets the scene, so far as necessary ("Let A be an $n \times n$ real matrix"). Then there will be a single sentence that is the substance of the definition ("An inverse of A is an $n \times n$ matrix B such that $AB = BA = I_n$"); ideally, it will be of the form "(thing defined) is …". Finally, there may be an appendix, which might be merely a dependent clause or might be quite long, explaining anything that might remain unclear ("where I_n denotes the identity matrix of degree n"). There is no absolute distinction between what goes into the introduction and what is left until the appendix; that is a matter of style and clarity.]

The main formal difficulty with definitions is the use of 'if'. When I say "A matrix is said to be real if all its entries are real numbers", the word 'if' is not being used in a logical or mathematical sense—the statement clearly means that, when I say a matrix is real, I have in mind no more and no less than that its entries are real numbers. It would be perverse to suppose that I am allowing a matrix to be "said to be real" under other circumstances, although the mathematical or logical sense of 'if' permits such an interpretation. The phrase "is said to be" should alert us that this is a definition, and that 'if' is to be understood as introducing the meaning we wish to attach to the previous clause. However, such a clarifying phrase as "is said to be" is often omitted, because the context and form of the statement normally make it obvious that a definition is intended. Some people, aware of this inconsistency in the possible meanings of 'if', try to overcome it, in my opinion rather over-zealously, by writing 'if and only if' instead of 'if' in definitions. I find this objectionable too, because 'if and only if' is also a logical connective; consequently, its use in a definition is still misleading, because it creates a statement that looks like a theorem but is not one.

I have already suggested that, ideally, one would avoid this "defining" use of 'if' altogether, by saying, for instance, "a real matrix is one all of whose entries are real numbers", which tells us at once what a real matrix is. Unfortunately, the slovenly use of 'if' is much more the rule than the exception. It would be finicky to shun it; nor have I tried to. Fortunately, the form (in English at least), and, more unequivocally, the context, of any statement normally indicate whether 'if' has a logical meaning or is part of a definition.

When the main clause comes first and its meaning has not previously been defined, as in "A matrix is real if all its entries are real", where the adjective "real" has not previously been applied to matrices, that is customarily a definition. When the 'if' clause precedes the main clause, as for instance in "If a matrix is real, all its entries are real", that suggests strongly that 'if' is a logical connective and the statement is not a definition; a "real matrix" might, in principle, be required to have other properties not mentioned. For a parallel statement of exactly the same form that cannot possibly be a definition, consider "If B is a real square matrix, then $\det(B^2) \geq 0$". In this case, the stated property, namely $\det(B^2) \geq 0$, is shared by some non-real matrices, such as $\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$.

Much of the material in this course is fundamental in modern pure and applied mathematics—even in contexts like statistics, logic, number theory, algebra, or geometry where you might not at first expect it to have any relevance at all. This is worth emphasizing, because students have been known to complain that the course has no "applications". If you're not interested in what is true and why, then, indeed, much of the course may seem rather superfluous. We shall spend quite a lot of time proving things we all "know" to be true (although such "knowledge", like many things people take for granted, is often just habit, and I hope there are also surprises from time to time). But the tools we shall acquire are essential in many contexts. Along the way, we shall remove gaps in earlier courses by *proving* results that were previously assumed without explanation. Notable examples are Rolle's theorem, the fundamental theorem of algebra, and the method of partial fractions.

And now for the disclaimers. In principle, this was a 30-something-lecture, 15-point course. Even with some handouts, we could not cover all the material we should have liked to, or even what I believe "anybody with a mathematics major ought to know". When I was an undergraduate, the topics we shall be touching on were spread over something like 100 lectures that everyone was supposed to attend. The comparison is unfair; the emphasis of mathematics has shifted and there is a lot more of it; those 100 lectures also included many things that we treat, albeit unrigorously, in 243 or elsewhere; moreover, some things were repeated three times. But it remains true that there is a lot of fairly basic analysis we cannot even cursorily include. We have had to *choose* what to do. Our choice, although deliberately

conventional, must be arbitrary to some extent. I have confessed to some gaps in the later logical development, and, even for things we do discuss, we have left out some proofs that seem relatively uninteresting (you can supply them yourselves, although they may be a little confusing to construct in some cases). Despite this, I have not always tried to take the shortest routes through the theorems.

The notes below are not exactly those presented in any particular year. Some material that only appeared in handouts (or only in some years) has been included. There has been a good deal of tidying-up; I have removed the errors that I have noticed, although I have probably introduced others; and I have also included appendices, including Appendix B that formed a fairly early part of the course in some years and became superfluous when a separate course on topology was introduced. (Its place was taken by §8 and some other matter. I should note that I have numbered the results of Appendix B in sequence with those of the main text, for the convenience of occasional cross-reference.) Thus, these notes are rather fuller than in any given year, and are not, perhaps, in the best logical order.

At first sight, there may seem to be an intimidating quantity of information. So a little advice may not be amiss. Generally speaking, the best way to learn a lemma or theorem at this fairly basic level is to study the *proof* carefully; it usually depends on some idea or insight, which, once you have grasped it, will clarify what the theorem is really about. Simply trying to learn the *statements* of theorems can be confusing, because the precise form of the hypotheses is often extremely important but is not in itself very memorable; those hypotheses are usually forced upon us by the proof. This advice does not apply to more advanced theorems, whose proofs may be very involved and lack a single guiding idea. In such cases the hypotheses are often chosen because of the applications one has in mind rather than out of logical necessity.

1. What *are* the real numbers?

People do not agree whether 0 should be accounted a "natural number" or not. My convention is that it is not: $\mathbb{N} := \{1, 2, 3, \dots\}$. If $a, b \in \mathbb{N}$, their sum $a + b \in \mathbb{N}$ and their product $ab \in \mathbb{N}$ are defined, and they satisfy the usual rules of arithmetic; we can also *compare* a and b, determining whether $a \leq b$ or $b \leq a$, and this "total order" on $\mathbb{N}$ is related to addition and multiplication (if $a \leq b$ and $c \in \mathbb{N}$ too, then $a + c \leq b + c$ and $ac \leq bc$). "Subtraction" and "division" are not *always* possible in $\mathbb{N}$. $m - n$ makes sense only if $m > n$, and $m \div n$ makes sense only if m is a multiple of n.

There are, however, *cancellation laws*: for $a, b, c \in \mathbb{N}$, if $ac \leq bc$, then $a \leq b$, and, if $a + c \leq b + c$, likewise $a \leq b$.

These properties may be derived either directly from basic set theory (my preference), or on the basis of the *Peano axioms*—stated, in essence, by Dedekind in 1888, a year before Peano. (I have never seen the point of the Peano axioms myself.) But we don't have the time for discussing these niceties, when the conclusions are familiar to us all. In particular, we take for granted *well-ordering of* $\mathbb{N}$:

Lemma 1.1. *Any non-empty subset A of $\mathbb{N}$ has a least element.* □

As a "justification", notice that if we count $1, 2, 3, 4, \dots$, the least element of A will be the first number we come to that belongs to A. $\mathbb{N}$ itself has 1 as least element.

1.1 is in effect the principle of mathematical induction, which we shall use freely.

$\mathbb{N}$ is the set of numbers we use for counting (that is why I think it should start at 1). But a three-metre long line (three times the length of a "standard" line) can be bisected, and the result is a line whose length is the same as if we had divided a line of length 9 metres into six equal sections. Thus we obtain the idea of "fractions", which are "quotients" of the form m/n, where $m, n \in \mathbb{N}$. For the ancient Greeks, numbers *were* the lengths of lines and no logical problem—for the very thorny question of the foundations of geometry itself was only noticed much later—was apparent here; but for us, there is a difficulty: what *is* a quotient m/n? $3/2$, for instance, is certainly not a natural number. It must be an object of a new kind. A fraction is *represented by* a symbol m/n that is subject to certain rules; to have a watertight definition, we can employ a standard trick, more or less saying that the fraction *is* the symbol. For brevity (although at the cost of some technical complication), I shall construct not only fractions, but all rational numbers, at one go. The idea is that a rational number "ought" to have the properties we *expect* of the meaningless "quotient" $(m - n)/p$ in which m, n, p are natural numbers, though no such "quotient" exists *a priori*. What we define definitely does exist, and has all the properties we expect of quotients.

In the set $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ consisting of all triples (m, n, p) of natural numbers (none of m, n, p is zero) introduce a relation $\sim$ by

$$(m, n, p) \sim (m', n', p') \quad \text{means} \quad p'm + n'p = m'p + p'n\,.$$

(m', n', p', m, n, p here denote arbitrary natural numbers). This is an equivalence relation (i.e. reflexive, symmetric, and transitive), and so it partitions $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ into equivalence classes.

Definition 1.2. A *rational number* is a $\sim$-equivalence class of $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$. That is, it is a non-empty subset of $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ that contains all the elements $\sim$-equivalent to any one of its elements. We write $[(m, n, p)]$ to denote the equivalence class which contains the element $(m, n, p) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}$; then equality of equivalence classes, $[(m, n, p)] = [(m', n', p')]$,

means the same as $(m,n,p) \sim (m',n',p')$. The class of all rational numbers is denoted by $\mathbb{Q}$ (possibly from "quotient", or rather the German "Quotient").

Definition 1.3. Let $\xi = [(m,n,p)]$, $\eta = [(a,b,c)]$ be rational numbers. We *wish* to define

(a) $\xi \le \eta$ if and only if $cm + bp \le ap + cn$,
(b) $\xi + \eta := [(cm + ap, cn + bp, cp)]$,
(c) $\xi\eta := [(am + bn, an + bm, cp)]$.

These "definitions" are, however, illegitimate. They might not be "well-defined". To construct $\xi + \eta$, for instance, we must *choose* a specific triple (m,n,p) from the equivalence class ξ and a specific triple (a,b,c) from η, and form the triple $(cm + ap, cn + bp, cp)$ whose *equivalence class* is to be $\xi + \eta$. It is conceivable that this resulting equivalence class might depend on our choices of (m,n,p) and of (a,b,c), and in that case we could not justifiably write it as $\xi + \eta$, since the notation would not indicate which choices were used.

This theoretical difficulty does not, however, arise, because:

Lemma 1.4. *If* $(m,n,p) \sim (m',n',p')$ *and* $(a,b,c) \sim (a',b',c')$, *then*

(a) $cm + bp \le ap + cn$ *if and only if* $c'm' + b'p' \le a'p' + c'n'$,
(b) $(cm + ap, cn + bp, cp) \sim (c'm' + a'p', c'n' + b'p', c'p')$,
(c) $(am + bn, an + bm, cp) \sim (a'm' + b'n', a'n' + b'm', c'p')$. □

These statements should be proved by using only addition and multiplication, since one cannot always subtract or divide natural numbers. However, the cancellation laws can be invoked. The trick is to perform suitable manipulations so that permissible cancellations will lead to the required conclusions.

As a consequence of this lemma, we can genuinely add and multiply rational numbers, and "compare" them (i.e. say $\xi \le \eta$ when appropriate).

In fact $\le$ is a *total order* on $\mathbb{Q}$, as one sees easily. I recall (from 161) what this means. Some authors omit 1.5*(c)* in some contexts, but we have no reason to.

Definition 1.5. A relation $\le$ on a set S is a *partial order* if, for any $\xi, \eta, \zeta \in S$,

(a) $\xi \le \xi$ ($\le$ is reflexive),
(b) $(\xi \le \eta \ \&\ \eta \le \zeta) \Longrightarrow \xi \le \zeta$ ($\le$ is transitive), and
(c) $(\xi \le \eta \ \&\ \eta \le \xi) \Longrightarrow \xi = \eta$ ($\le$ is antisymmetric).

(Compare and contrast the idea of an equivalence relation.) Then $\le$ is a *total order* on S if, for any $\xi, \eta \in S$, either $\xi \le \eta$ or $\eta \le \xi$.

We often write $\xi \ge \eta$ to mean the same as $\eta \le \xi$, and $\xi < \eta$ (or $\eta > \xi$) to mean that $\xi \le \eta$ but $\xi \ne \eta$. Then any total order is subject to the *trichotomy*: given $\xi, \eta \in S$, either $\xi < \eta$ or $\xi = \eta$ or $\xi > \eta$. In particular, $\mathbb{Q}$ satisfies the trichotomy.

Lemma 1.6. $\mathbb{Q}$ *is a field.* □

This means that all the field axioms (the 'laws of arithmetic') given in MATH 251 are satisfied. As well as addition and multiplication, subtraction always makes sense, and you can divide by anything except "zero". If you are not familiar with the axioms, look them up.

The field zero '0' of $\mathbb{Q}$ is $[(1,1,1)] = [(m,m,n)]$ for any $m, n \in \mathbb{N}$, and the field multiplicative identity '1' is $[(2,1,1)] = [(m+n,n,m)]$ for any $m, n \in \mathbb{N}$. The negative of $[(m,n,p)]$ is $[(n,m,p)]$. If $\xi := [(m,n,p)] \neq$ '0', then $m \neq n$, and the reciprocal of ξ is either $[(p+1,1,m-n)]$ if $m > n$, or $[(1,1+p,n-m)]$ if $m < n$.

The natural number 1 and the rational number I have just written '1' are conceptually quite different. For any $m \in \mathbb{N}$, we can introduce a rational number 'm' as

$$\text{`}m\text{'} := [(m+1,1,1)] = [(m+n,n,1)],$$

and then it is easily checked that

$$\underset{\text{addition in } \mathbb{Q}}{\text{`}m\text{'} + \text{`}n\text{'}} = \underset{\text{addition in } \mathbb{N}}{\text{`}m+n\text{'}}, \qquad \underset{\text{multiplication in } \mathbb{Q}}{\text{`}m\text{'`}n\text{'}} = \underset{\text{multiplication in } \mathbb{N}}{\text{`}mn\text{'}};$$

further, 'm' $\leq$ 'n' when, and only when, $m \leq n$ in $\mathbb{N}$. In effect, then, there is a copy of $\mathbb{N}$ inside $\mathbb{Q}$, consisting of the rationals 'm'. It has the same addition, multiplication, and order as $\mathbb{N}$ itself. In practice we draw no distinction of notation between $\mathbb{N}$ and this copy, despite the conceptual difference; that is, we regard natural numbers as just a special kind of rational number, and treat $\mathbb{N}$ as a subset of $\mathbb{Q}$. It is then convenient to write the rational number $[(m,n,p)]$ as $(m-n)/p$, which we are now allowed to do since subtraction and division make sense in $\mathbb{Q}$.

We can go further. The copy of $\mathbb{N}$ in $\mathbb{Q}$ consisted of equivalence classes $[(m,n,1)]$ for $m > n$ in $\mathbb{N}$. The equivalence classes $[(m,n,1)]$, for *any* $m, n \in \mathbb{N}$, consist of the 'natural numbers' in $\mathbb{Q}$ as just introduced, of their negatives, and of '0'. They form a *commutative ring with identity* (as defined in 311; all we mean here is that the sum, product, or difference of two such rational numbers is another, and the zero and identity of $\mathbb{Q}$ are also of this form), which is called the *ring of rational integers* (or just *of integers*) and is denoted by $\mathbb{Z}$ (die Zahlen).

The total order in $\mathbb{Q}$ is related to the algebraic structure as follows.

Lemma 1.7. *Suppose* $\alpha, \beta, \gamma \in \mathbb{Q}$.
(i) *If* $\alpha \leq \beta$, *then* $\alpha + \gamma \leq \beta + \gamma$.
(ii) *If* $\alpha \leq \beta$ *and* $\gamma \geq 0$, *then* $\alpha\gamma \leq \beta\gamma$. □

If *(i)* already holds, *(ii)* is equivalent to: if $\alpha \geq 0$ and $\beta \geq 0$, then $\alpha\beta \geq 0$.

Definition 1.8. A field $\mathbb{F}$ is said to be *ordered* by a partial order relation $\leq$ if $\leq$ is in fact a total order, and if, for any $\alpha, \beta, \gamma \in \mathbb{F}$,

$$\textit{(i)} \quad \alpha \leq \beta \Longrightarrow \alpha + \gamma \leq \beta + \gamma, \qquad \textit{(ii)} \quad \alpha, \beta \geq 0 \Longrightarrow \alpha\beta \geq 0.$$

Remark 1.9. These ideas may be expressed differently by introducing the set of *positive rational numbers* $P := \{\xi \in \mathbb{Q} : \xi > 0\}$. We write $-P$ for $\{-\xi : \xi \in \mathbb{Q}\}$. Then:
(1) the trichotomy property becomes the assertion that $\mathbb{Q}$ is the *disjoint* union $P \cup \{0\} \cup (-P)$,
(2) 1.8*(i)* above becomes $P + P := \{\alpha + \beta : \alpha, \beta \in P\} \subseteq P$, and
(3) 1.8*(ii)* above becomes $PP := \{\alpha\beta : \alpha, \beta \in P\} \subseteq P$.

(I use the $:=$ sign occasionally to indicate either that something is a definition—"$P + P$" is *defined* to be the set $\{\alpha + \beta : \alpha \in P \ \&\ \beta \in P\}$—or to point out that two things are equal by definition. Notice, by the way, that the sum and product of two positive numbers must be *positive*, not just non-negative. In any field, $\alpha, \beta \neq 0 \Longrightarrow \alpha\beta \neq 0$.)

In an ordered field, $1 > 0$. (This is a very easy exercise, but not quite obvious.)

An ordered field is necessarily *of characteristic* 0. This means that adding 1 to itself repeatedly gives an infinite sequence of different elements of the field:

$$1 + 1 > 1\,,\ 1 + 1 + 1 > 1 + 1\,,$$

and so on. It follows that any ordered field contains a copy of $\mathbb{N}$ (in which the 1 of $\mathbb{N}$ is also the 1 of the field), and so a copy of $\mathbb{Q}$. In this sense, $\mathbb{Q}$ is the smallest possible ordered field.

In any totally ordered set, one has the greater $\max\{a, b\}$ and the lesser $\min\{a, b\}$ of two elements. By induction, one may also determine the greatest or least element of any finite subset: $\max\{a_1, a_2, \ldots, a_k\}$, for instance. In an ordered field $\mathbb{F}$, one has the "absolute value": $|\alpha| := \max(\alpha, -\alpha)$. (It is common, albeit illogical, to employ parentheses—round brackets—instead of braces, in connection with max and min).

[Jumping ahead: $\mathbb{C}$, not yet defined, is *not* an ordered field; why not?]

Lemma 1.10. *Let $\mathbb{F}$ be an ordered field and $x, y \in \mathbb{F}$. Then $x = 0$ if and only if $|x| = 0$, and one has the* triangle inequality

$$|x + y| \leq |x| + |y|$$

or equivalently $\big||x| - |y|\big| \leq |x - y|$. □

The need to extend $\mathbb{N}$ to a field $\mathbb{Q}$ resulted from mensuration: dividing a line into n equal parts, as I remarked. However, the Greeks discovered early that, on superficially reasonable assumptions about the foundations of geometry, you have Pythagoras's theorem, and, therefore, the hypotenuse of a right-angled triangle whose other sides are integer multiples of a standard length will often not be a rational multiple of that length. For instance, if the adjacent sides are of unit length, the square on the hypotenuse has area 2. But:

Lemma 1.11. *There is no rational number ξ such that $\xi^2 = 2 \in \mathbb{Q}$.*

Proof. Such a ξ cannot be 0, and if $\xi \in -P$ we may substitute it by $-\xi$. So we may suppose $\xi \in P$ and $\xi = m/n$ for some $m, n \in \mathbb{N}$. Consider the set A of all natural numbers n for which there exists some $m \in \mathbb{N}$ such that $\xi = m/n$. By Lemma 1.1, there is a *least* such n, with a corresponding m. Then $n < m < 2n$ (for, otherwise, either $m \leq n$ and $m^2/n^2 \leq 1$, or $m/n \geq 2$ and $m^2/n^2 \geq 4$). Thus $p := 2n - m$ and $q := m - n$ are also in $\mathbb{N}$, and $q < n$. However,

$$p^2 = 4n^2 - 4nm + m^2 = 2m^2 - 4mn + 2n^2 = 2q^2\,,$$

so $q \in A$; this is absurd, as $q < n$ and n was the *least* element of A. □

The above proof (which can be modified slightly for other natural numbers than 2 that are not perfect squares) is not the well-known one given in MATH 161, which used deeper properties of $\mathbb{N}$ to do with divisibility. There are many other proofs.

Pythagoras's theorem, as an empirical fact at least, was known in some form outside Greece before Pythagoras (who seems really to have existed: he was born on Samos around

570 B.C., moved in middle age to Croton in South Italy and founded his school there, but probably was not personally responsible for the theorem). However, it was his school that first understood that it constituted a problem. They believed that natural numbers were somehow the atoms out of which the universe was built, and so the seeming existence of geometrical quantities that cannot possibly be described by rational numbers was a shock. Despite this, Greek mathematics, being fundamentally geometrical, eventually succeeded in assimilating the idea of irrational *ratios* between magnitudes of the same kind. You can see this in the definitions at the beginning of Euclid's Book V, and in his Book X. From a modern point of view, however, the question is what sort of *numbers* we can use to measure irrational lengths; what, in other words, 'is' a real number?

There are several possible answers. The plain man's remedy is to use infinite decimal expansions "with familiar rules for addition, multiplication, subtraction and division". This appears snappy, and is probably historically accurate, but there are serious objections.

The first is that the representation of numbers by decimals is not entirely unique ($1\cdot000\cdots$ is the same as $0\cdot9999\cdots$, and so on). The second is that there seems to be no *a priori* reason to prefer 10 to any other base, or to suppose that the base does not matter; and translating from one base to another is messy. The third objection is that the rules for addition, multiplication, and especially division are also not as easy as you might think. I called them 'familiar'; but one usually operates with *finite* decimal expansions, not infinite ones, and it is not simple to form a rule for dividing or multiplying one *infinite* decimal by another.

An entirely satisfactory definition of real numbers was only achieved in the later 19th century, and then two equivalent ones were proposed. One was invented by Dedekind in 1858, but only published in 1872 as a riposte to Cantor; the other was published by Méray in 1869 and then by Cantor in 1872. (Cantor got most of the credit.) Dedekind's construction is conceptually simpler, so I shall sketch a version of it. It is based on the order structure of $\mathbb{Q}$, and, consequently, is not well adapted to the algebraic structure. The Méray-Cantor construction needs much more preparation, but the algebra, at least, becomes easy. Both constructions have important generalizations, in rather different directions; but both are founded on the idea that a real number is something "approximated" by rational numbers. The two constructions differ in the kind of approximation.

Definition 1.12. A of $\mathbb{Q}$ is a *Dedekind section* or *Dedekind cut* if

$$\left.\begin{array}{ll} \textit{(i)} & \emptyset \neq A \neq \mathbb{Q}, \\ \textit{(ii)} & \text{if } \alpha \in A \text{ and } \alpha \leq \beta \in \mathbb{Q}, \text{ then } \beta \in A, \text{ and} \\ \textit{(iii)} & \text{for any } \alpha \in A, \text{ there exists } \gamma \in A \text{ such that } \gamma < \alpha. \end{array}\right\} \tag{1}$$

We say the cut A is *non-negative* if it does not contain the rational number 0. In view of *(ii)*, this is equivalent to saying $A \subseteq P$.

Condition *(iii)* says that A has no least member. Thus P in particular is a cut; if β is a positive rational, so is $\frac{1}{2}\beta$, which is less than β.

If A is a cut and $\beta \in \mathbb{Q} \setminus A$ (such a β exists, by *(i)*), then $\beta < \alpha$ for any $a \in A$, by *(ii)*, and $A - \beta \coloneqq \{\alpha - \beta : \alpha \in A\}$ is a non-negative cut.

Given a cut A, we can try to define

$$-A \coloneqq \{-\beta : \beta \in \mathbb{Q} \setminus A \ \ \& \ \underbrace{(\exists \gamma \in \mathbb{Q} \setminus A)(\gamma > \beta)}_{(\dagger)}\}. \tag{2}$$

Notice that P is a cut, but its complement $(-P) \cup \{0\}$, the set of all non-positive rationals,

has a largest member 0; reversing the sign of all members of $\mathbb{Q} \setminus P$ gives the set $P \cup \{0\}$, which does not satisfy *(iii)*. Thus the definition of $-A$ includes the rather unexpected condition (†) to ensure that 1.12*(iii)* is indeed satisfied by $-A$.

The *zero cut* is just P. More generally, if $\xi \in \mathbb{Q}$, let the corresponding *principal cut* be

$$S(\xi) \coloneqq \{\alpha \in \mathbb{Q} : \alpha > \xi\}, \tag{3}$$

which is also a cut. *Not all cuts are principal*.

If A and B are both Dedekind cuts, we define

$$A + B \coloneqq \{\alpha + \beta : \alpha \in A \ \ \& \ \ \beta \in B\}.$$

It is easily seen that $A + B$ is also a cut, and that the "addition of cuts" thus defined is commutative and associative. Furthermore, $A + P = A$ and $A + (-A) = P$ for any cut A. (The last statement requires some argument, because of (2).)

If A and B are *non-negative* cuts, we define

$$AB \coloneqq \{\alpha\beta : \alpha \in A \ \ \& \ \ \beta \in B\} = BA, \tag{4}$$

which is also a cut. (This would certainly be false if A or B had a negative member). If A and B are general cuts, *choose* non-positive rationals $\lambda \in \mathbb{Q} \setminus A$, $\mu \in \mathbb{Q} \setminus B$, and then

$$A - \lambda \coloneqq \{\alpha - \lambda : \alpha \in A\}$$

and $B - \mu$ are non-negative cuts. Hence, $(A - \lambda)(B - \mu)$ is a cut, and so is

$$(A - \lambda)(B - \mu) + (-A)S(-\lambda) + (-B)S(-\mu) + S(-\lambda\mu). \tag{5}$$

This cut does not depend on the choice of λ and μ, and we can take it as the definition of AB for general cuts. (I omit the proof, which is rather long, and really amounts to showing that (5) is equivalent to a different and less "natural" definition that was originally given by Dedekind.) For non-negative cuts we could take $\lambda = \mu = 0$, and then (5) reduces to (4). This troublesome definition of multiplication is the main defect of Dedekind's theory.

Finally, for cuts A and B,

$$\text{"}A \leq B\text{"} \quad \text{means} \quad \text{"}B \subseteq A\text{"}. \tag{6}$$

With all these definitions, it may be proved, albeit not without effort, that the set of cuts forms an ordered field, which we agree to be the field of real numbers $\mathbb{R}$. In short,

Definition 1.13. A *real number* is a Dedekind cut of the rationals.

The principal cuts $\{S(\xi) : \xi \in \mathbb{Q}\}$ form a subfield isomorphic to $\mathbb{Q}$; in effect, then, $\mathbb{R}$ includes a copy of $\mathbb{Q}$ (with the same addition, multiplication, *and order*). We denote these "rational reals" by the names of the corresponding rationals—indeed, we often just call them "rationals"—and write $\mathbb{Q}$ for its copy in $\mathbb{R}$ (and $\mathbb{N}$ and $\mathbb{Z}$ for their copies in that copy of $\mathbb{Q}$).

The details of this construction can be varied in many ways, but the governing idea, not unlike Euclid's, is that a real number should be defined as the set of rational numbers that "ought to be bigger" than that real number. For instance, $\sqrt{2}$ should be thought of as the set of all positive rational numbers whose squares are greater than 2. The reason for the condition 1.12*(iii)*, which at first sight may seem superfluous, is that without it the rational real numbers would be represented twice; 3, for instance, would correspond not only to the genuine cut $S(3)$, but also to $\{\alpha \in \mathbb{Q} : \alpha \geq 3\}$, which does not satisfy *(iii)*.

The definition of multiplication of cuts is messy because the construction is founded on the order relation in $\mathbb{R}$ without reference to the arithmetical operations. The Méray-Cantor construction, which I shall briefly mention later, has the opposite difficulty that upper and lower bounds (introduced below) appear slightly unnatural, although multiplication is easy.

It is clear that one could define Dedekind cuts in any partially ordered set.

Apart from the construction itself, Dedekind pointed out a property of $\mathbb{R}$, as thus defined, which is in some sense the foundation of all analysis. This is the *Dedekind completeness axiom*, which fails for $\mathbb{Q}$. Again, it really only uses the order structure of $\mathbb{R}$.

Definition 1.14. Let $(T, \leq)$ be a partially ordered set, and let A be a subset of T. Suppose $a \in A$ and $t \in T$.

(i) a is a (or the) *least element* of A if $(\forall x \in A)\ a \leq x$.
{a is itself in A, and is "below" everything else in A.}

(ii) a is a *minimal element* of A if
$$(\forall x \in A)(x \leq a \Rightarrow x = a).$$
{a is in A, and there is nothing "below" it in A.}

(iii) t is an *upper bound* (in T) for A if $(\forall x \in A)\ x \leq t$.

(iv) A is *bounded above* (in T) if it has an upper bound.

A can have at most one least element. If $a, a' \in A$ and $a \leq a'$, $a' \leq a$, then $a = a'$ by 1.5*(c)*. On the other hand, if $\leq$ is not a total order on A, there may be several minimal elements. For instance, let $\{a, b, c\}$ be given the partial order $a \leq c$, $b \leq c$ (and $a \leq a$, $b \leq b$, $c \leq c$). Then both a and b are minimal in A. But neither of them is a least element in A.

If $\leq$ is changed to $\geq$, one has the definitions of a *greatest* element of A, of a *maximal* element of A, of a *lower* bound for A in T, and of the phrase "A is bounded below in T".

A set is *bounded* in T if it is bounded in T both above and below.

If A has a greatest element a, then a is an upper bound for A, and, conversely, an upper bound for A that also belongs to A must be the greatest element of A. But, for instance, the open interval $(0, 1)$ has no greatest (or least) element. (Whatever $a \in (0, 1)$ you take, $\frac{1}{2}(1 + a)$ is greater and $\frac{1}{2}a$ is less.) It has many upper bounds in $\mathbb{R}$, such as 1 and 2.

It is obvious that a greatest element of A, if there is one, is also a maximal element, and a least element is also minimal. In general, a maximal element need not be greatest (as above for "least" and "minimal").

On the other hand, when A is *totally* ordered by $\leq$, which is the case for any subset of $\mathbb{R}$ (or of $\mathbb{Q}$), the distinction between a greatest element and a maximal element of A disappears. If a is maximal and $x \in A$, then $x \geq a$ is only possible if $x = a$; thus, either $x = a$ or $x < a$ (recall from the remark after 1.5 that this means $x \leq a$ but $x \neq a$); that is, $x \leq a$. As "maximal" elements and "greatest" elements are the same for subsets of $\mathbb{R}$, it is legitimate to speak of a real-valued function's "maximum" value rather of its greatest value (if either exists). But notice that $f : (0, 1) \longrightarrow (0, 1) : x \mapsto x$ has no greatest or least value.

Definition 1.15. A partially ordered set T is *Dedekind-complete* (or *boundedly order-complete*) if every *nonnull* subset of T that is *bounded below* in T has a greatest lower bound in T. In other words: if $\emptyset \neq A \subseteq T$ and the set L of lower bounds for A in T is non-null, then L has a *greatest* element, the "greatest lower bound" for A in T. This greatest lower bound is commonly called the *infimum* of A in T and is usually written $\inf A$. (It is rarely necessary to specify T, because the context makes it clear; indeed, for us T is mostly $\mathbb{R}$.)

In older English-language books $\inf A$ is sometimes denoted g.l.b. A or glb A.

Some special nonnull subsets A of T may, of course, have greatest lower bounds (or infima) in T even if T is not Dedekind-complete. In any case, the infimum of a nonnull set A in T (when it exists) is, as it were, the nearest approach to a least element of A in T. For example, the lower bounds in $\mathbb{R}$ of $(0,1)$ form the interval $(-\infty, 0]$, which has the greatest element 0: thus, $\inf(0,1) = 0 \notin (0,1)$.

It may seem odd that the definition 1.15 mentions only *lower* bounds for A in T.

Lemma 1.16. *If the partially ordered set T is boundedly order-complete in the sense of* 1.15*, then any non-null subset A of T that is bounded above has a least upper bound.*

Proof. Let U be the set of all upper bounds of A. As A is bounded above, $U \neq \emptyset$; and U is bounded below by any element of A, because $A \neq \emptyset$. Thus U has a greatest lower bound u in T, by hypothesis. However, any $a \in A$ is a lower bound for U; thus $u \geq a$. This shows that u is itself a upper bound for A, that is, $u \in U$. As u is in U and is a lower bound for U, it is the least element of U. □

Definition 1.17. Given a nonnull subset A of T, its least upper bound in T (if it exists) is called the *supremum* of A in T.

If T is clear from the context and A has a least upper bound in T, the least upper bound is often written as $\sup A$ (or l.u.b. A, or lub A).

The argument of 1.16 is easily changed to prove that, if any nonnull subset A of T that is bounded above has a least upper bound, then T is Dedekind-complete in the sense of 1.15. The asymmetry of 1.15 (for sets bounded *below* and *lower* bounds)—was, therefore, only apparent; the property holds equally if I use sets bounded above and upper bounds. (In $\mathbb{R}$, this equivalence may be proved by reversing the signs, but it is true in any partially ordered set).

For A to have a *greatest* lower bound in T, it must be bounded below to begin with. However, $\mathbb{R}$ itself has no lower bound in $\mathbb{R}$; less trivially, $\mathbb{Z}$ has no lower bound either. There are partially ordered sets T in which *all* non-null subsets have both a least upper bound and a greatest lower bound. An example is $T := [0,1]$, with its usual partial order as a subset of $\mathbb{R}$. A somewhat less banal example is this.

Let Ω be any set, and let T be the class of all subsets of Ω (that is, the "power class" $\mathcal{P}(\Omega)$ of Ω). There is a natural partial order in T: "$A \leq B$" means "$A \subseteq B$". Then any subset $\mathcal{Q}$ of T (that is, $\mathcal{Q}$ is some class of subsets of Ω) has both a supremum and an infimum. Indeed, $\sup \mathcal{Q} = \bigcup_{Q \in \mathcal{Q}} Q$ and $\inf \mathcal{Q} = \bigcap_{Q \in \mathcal{Q}} Q$.

We can now prove "Dedekind's axiom" for the real numbers. By this, we mean

Theorem 1.18. *$\mathbb{R}$ is Dedekind-complete.*

Proof. Let A be a non-empty subset of $\mathbb{R}$ that is bounded below, with a lower bound b. Now b is, by definition, a Dedekind cut of $\mathbb{Q}$. By the definition of the partial order in $\mathbb{R}$, $a \subseteq b$ for any element $a \in A$.

Let $k := \bigcup_{a \in A} a$; thus $k \subseteq b \neq \mathbb{Q}$, and $k \neq \emptyset$ as each $a \neq \emptyset$, by 1.12*(i)*. Hence, k itself satisfies 1.12*(i)*. If $\alpha \in k$, there is some $a \in A$ with $\alpha \in a$, and, if $\beta \geq \alpha$ in $\mathbb{Q}$, then $\beta \in a$ too; so $\beta \in k$. This means that k satisfies 1.12*(ii)*. Similarly, there exists some $\gamma \in a$ with $\gamma < \alpha$, and, as $\gamma \in k$, k satisfies 1.12*(iii)*. So k is a Dedekind section of $\mathbb{Q}$; and, by its

definition, k is the smallest *set* which includes all $a \in A$. It is, therefore, the greatest lower bound of A. □

For any *finite* subset $\{a_1, a_2, \ldots, a_k\}$ of $\mathbb{R}$, or of $\mathbb{Q}$, or of any *totally ordered* set, one may find the greatest element by comparing elements in pairs, and this greatest element is denoted $\max\{a_1, a_2, \ldots, a_k\}$. It is the supremum of the subset. In a totally ordered set, only infinite subsets (i.e. subsets that have infinitely many elements) can ever lack suprema. Similarly, the infimum of $\{a_1, a_2, \ldots, a_k\}$ is its least element $\min\{a_1, a_2, \ldots, a_k\}$. On the other hand, a set with infinitely many elements need not be bounded, and if it is bounded there need not be a greatest or a least element. (Think of the set $\{1/n : n \in \mathbb{N}\} \cup \{3 - 1/n : n \in \mathbb{N}\}$.)

Remark 1.19. To recapitulate: Dedekind's idea was this. Real numbers have no "simple and natural" definition. So let us agree that a real number a is *by definition* the set of all rational numbers that "ought to be greater than whatever our intuitive notion of a demands"; for instance, $\sqrt{2}$ will be the set of all positive rationals whose squares exceed 2.

This appears strange. We are accustomed to working with real numbers; it seems bizarre to treat them as complicated "compound" entities. There are two replies to this objection.

Firstly, in truth we never *do* work directly with real numbers. Either we define them implicitly (as for instance $\sqrt{2}$ or "the smallest positive zero of the cosine" or "the largest root of $x^5 - 4x^4 + 3x^3 - 2x^2 + x - 1$") or we use approximations (instead of $\sqrt{2}$ we take ten decimal places, $1 \cdot 4142135624$, or whatever). We assume, however, that there is *something*, an "entity", that is being defined by either of these techniques; what can it be? Dedekind proposed one possible answer to this question.

Secondly, we believe, for whatever reasons, that real numbers, whatever they may be, should have certain properties. Until we construct objects with these properties in a logically defensible way, there is some doubt whether the properties are consistent. In talking about the real numbers, we might be appealing to something that cannot exist. You have probably seen those proofs that "all triangles are isosceles" and so on, where absurd conclusions are drawn from a construction that looks plausible but is in fact impossible. Dedekind's construction (or, equally, Cantor's, or others that are also available) assures us that the existence and standard properties of $\mathbb{R}$ are as consistent as the basic ideas of set theory and of the natural numbers, which seem essential for mathematics to exist at all.

(I do not claim that the real numbers, on either construction, "really are" the things we have constructed. That would be a metaphysical, or perhaps psychological, question. As mathematicians, we want merely to construct entities that are precisely defined, so that we can prove results about them, and that behave as we expect real numbers to. Why we think there ought to be such things at all, behaving as we expect, is not our business.)

Granted a good construction of $\mathbb{R}$, all one needs thereafter are certain properties assured by the construction. Logically speaking, Dedekind's real numbers, Cantor's real numbers, and "infinite decimal expansions", are radically different kinds of object; nevertheless, in everything that interests us they have identical properties, and any of them could be employed as the basis of analysis. They are, in fact, "isomorphic" to each other as ordered fields, but I shall not bother to define this idea precisely or to present isomorphisms.

Lemma 1.20. *Let $a, b \in \mathbb{R}$ and $a < b$. There is a rational real ξ such that $a < \xi \leq b$.*

Proof. To say $a < b$ is to assert that $a \supseteq b \neq a$, as subsets of $\mathbb{Q}$ (by (6)). Thus, there exists some $\tau \in a \setminus b$. (I use the slanted minus, as before, to denote set difference, to avoid confusion with arithmetical subtraction.) By 1.12*(ii)* and (3), $S(\tau) \subseteq a$; by 1.12*(iii)*,

$S(\tau) \neq a$, since there is an element of a which is less than τ. Thus $a < S(\tau)$. On the other hand, $\tau \notin b$. Take any $\eta \in b$. By trichotomy, $\eta < \tau$ or $\eta = \tau$ or $\eta > \tau$. The first and second possibilities are excluded by 1.12*(ii)*, which would imply that $\tau \in b$. Thus $\eta \in S(\tau)$ for any $\eta \in b$, or $b \subseteq S(\tau)$, or $S(\tau) \leq b$. Take $\xi := S(\tau)$. □

This result is usually stated in an apparently stronger form:

Corollary 1.21. *If $a, b \in \mathbb{R}$ and $a < b$, there is a rational real ξ such that $a < \xi < b$.*

Proof. As $a < b$, $a < \frac{1}{2}(a+b) < b$. Apply 1.20 to $a, \frac{1}{2}(a+b)$ instead of a, b. There is a rational real number ξ such that $a < \xi \leq \frac{1}{2}(a+b) < b$. □

This Corollary may be expressed by saying that $\mathbb{Q}$ (that is, the copy of $\mathbb{Q}$ inside $\mathbb{R}$) is "order-dense" in $\mathbb{R}$. I have proved it from our construction of $\mathbb{R}$, but it is in fact true for any boundedly order-complete ordered field—I shall sketch the proof in a moment.

We have now derived (from Dedekind's construction) all the properties of $\mathbb{R}$ that we shall ever use, and it is of no subsequent importance what $\mathbb{R}$ actually is.

Lemma 1.22. *Let $\xi \in \mathbb{R}$. There is a real integer n such that $\xi < n$.* □

Definition 1.23. The least real integer n such that $n + 1 > \xi$ (recall 1.1) is called the *integer part of the real number ξ*, frequently denoted $[\xi]$. When $\xi \in \mathbb{Z}$, $[\xi] = \xi$; when $\xi \notin \mathbb{Z}$, $[\xi] < \xi$. The name "integer part" can be misleading when $\xi < 0$, since, for instance, $[-\frac{5}{2}] = -3$. (The unwary might think the "integer part" of $-2\frac{1}{2}$ should be -2.)

A stronger version of 1.22 (and my reason for postponing its proof!) is as follows.

Remark 1.24. An ordered field $\mathbb{F}$ is said to be *Archimedean* if, whenever α, β are positive elements of $\mathbb{F}$, there is a natural number n such that $n\alpha > \beta$.

[Of course $\mathbb{F}$ contains a copy of $\mathbb{N}$; see Remark 1.9. Equivalently, one could think of $n\alpha$ as the result of adding n copies of α. The case $\alpha := 1$ implies 1.21; take $\beta := |\xi|$.]

Any Dedekind-complete ordered field $\mathbb{F}$ is necessarily Archimedean:—

Proof. If not, then $\{n\alpha : n \in \mathbb{N}\}$ is nonempty and bounded above by β, so has a supremum γ. For any $n \in \mathbb{N}$, $(n+1)\alpha \leq \gamma$, and so $n\alpha \leq \gamma - \alpha$. This means that $\gamma - \alpha < \gamma$ is also an upper bound for $\{n\alpha : n \in \mathbb{N}\}$, which is impossible as γ was the *least* upper bound. □

The Archimedean property is popular, because it is true for $\mathbb{R}$ but not for some other well-known ordered fields (such as the hyperreals; but there are others).

Lemma 1.25. *In any Archimedean ordered field, $\mathbb{Q}$ is order-dense.* □

Thus, the copy of $\mathbb{Q}$ inside *any* Dedekind-complete ordered field $\mathbb{F}$ is order-dense in $\mathbb{F}$, and this is the crucial step in proving the "uniqueness theorem", which, roughly, states that

Theorem 1.26. *Any Dedekind-complete ordered field is a copy of $\mathbb{R}$ (as an ordered field).* □

I shall not give the proof. But it means that the particulars of Dedekind's construction can be ignored; all that matters is that $\mathbb{R}$ is a *boundedly complete ordered field*.

Lemma 1.27. *Let A be a bounded subset of $\mathbb{R}$, and let $-A \coloneqq \{-\alpha : \alpha \in A\}$. Then $-A$ is also bounded, and*

$$\sup(-A) = -\inf A\,, \quad \inf(-A) = -\sup A\,. \qquad \square$$

Although I have already used these names several times, let us have the formal definitions:

Definition 1.28. An *open interval* in $\mathbb{R}$ is a set of one of the forms

$$\begin{aligned} (a,b) &\coloneqq \{x \in \mathbb{R} : a < x < b\}\,, \quad \text{where } a,b \in \mathbb{R}\text{, or} \\ (-\infty,b) &\coloneqq \{x \in \mathbb{R} : x < b\}\,, \quad \text{where } b \in \mathbb{R}\text{, or} \\ (a,\infty) &\coloneqq \{x \in \mathbb{R} : a < x\}\,, \quad \text{where } a \in \mathbb{R}\,. \end{aligned}$$

(Notice that $(a,b) \neq \emptyset$ if and only if $a < b$. The symbols "∞" and "$-\infty$" do not, for our present purposes, denote anything; they only appear in certain standard formulæ.)

Similarly, a *closed interval* in $\mathbb{R}$ is a set of one of the forms

$$\begin{aligned} [a,b] &\coloneqq \{x \in \mathbb{R} : a \leq x \leq b\}\,, \quad \text{where } a,b \in \mathbb{R}\text{, or} \\ (-\infty,b] &\coloneqq \{x \in \mathbb{R} : x \leq b\}\,, \quad \text{where } b \in \mathbb{R}\text{, or} \\ [a,\infty) &\coloneqq \{x \in \mathbb{R} : a \leq x\}\,, \quad \text{where } a \in \mathbb{R}\,. \end{aligned}$$

(Here $[a,b] \neq \emptyset$ if and only if $a \leq b$.) These are definitions of the whole phrases "open interval" and "closed interval"—"open" and "closed" on their own have other meanings. (See, for instance, 3.5 and 3.6; more seriously, Appendix B.)

We frequently describe a member of $\mathbb{R}$ as a "point" of $\mathbb{R}$ [and similarly for $\mathbb{R}^2$, and so on].

§2. Sequences.

Definition 2.1. A *sequence* in a set X is a function $a : \mathbb{N} \longrightarrow X$. We usually write a_n instead of $a(n)$ to denote the value of the function a at the natural number n, and often describe the sequence by the symbol (a_n) or $(a_n)_{n=1}^{\infty}$. a_n is called the nth *term* of the sequence, and n is its *term index* or *index*. If a is an injective (one-to-one) function, we may say (a_n) is a sequence *without repeats*—that is, its terms are all different.

Since parentheses (round brackets) are used in so many senses, some people prefer to denote sequences in some other way, for instance by angle brackets, $\langle a_n \rangle$ or $\langle a_n \rangle_{n \in \mathbb{N}}$. But my notation is still the usual one.

The index set is taken to be $\mathbb{N}$ for simplicity. Other choices (for instance, $\{5,6,7,\dots\}$) do not seriously alter the theory. A sequence may sometimes be informally described by writing out its first few terms, $a_1, a_2, a_3, \dots$, especially if they strongly suggest a rule of formation for all the later terms. Similarly, the idea of a *subsequence* can be formulated in various (not entirely equivalent) ways that do not much affect our arguments, so I choose the simplest:

Definition 2.2. A function $\alpha : \mathbb{N} \longrightarrow \mathbb{N}$ is *strictly increasing* (and thus is a *strictly increasing sequence* in $\mathbb{N}$; see below) if $\alpha(m) < \alpha(n)$ whenever $m < n$. Let (a_n) be a sequence in X. A *subsequence* of (a_n) [occasionally called an *infinite* subsequence, to emphasize that infinitely many term indices of the original sequence are also indices of the

subsequence] is the composition $a \circ \alpha$ of the mapping $a : \mathbb{N} \longrightarrow X$ with a strictly increasing function $\alpha : \mathbb{N} \longrightarrow \mathbb{N}$. It may be written $(a_{\alpha(n)})$.

Thus a subsequence of $a_1, a_2, a_3, \dots$ will be of the form $a_{\alpha(1)}, a_{\alpha(2)}, \dots$. For example, let (a_n) be a given sequence; $a_5, a_{11}, a_{18}, a_{26}, \dots$ is a subsequence in which the nth term is $a_{\alpha(n)}$ for $\alpha(n) := \frac{1}{2}n^2 + \frac{9}{2}n$. We might write this subsequence as $\left(a_{\frac{1}{2}n^2+\frac{9}{2}n}\right)$. In the same way, there are subsequences (a_{n+1}) (this is the sequence $a_2, a_3, a_4, \dots$) or (a_{n^2}) (the sequence $a_1, a_4, a_9, \dots$).

If $\alpha : \mathbb{N} \longrightarrow \mathbb{N}$ is strictly increasing, then $\alpha(n) \geq n$ for all $n \in \mathbb{N}$. (This follows from a very simple induction.)

Definition 2.3. A sequence (a_n) in $\mathbb{R}$ is *increasing* if $a_n \leq a_{n+1}$ for all n; it is *decreasing* if $a_n \geq a_{n+1}$ for all n. It is *monotonic* if it is either decreasing or increasing [that is, if either $a_n \leq a_{n+1}$ for all n, or $a_n \geq a_{n+1}$ for all n; most sequences will definitely not be monotonic, for instance $1, 2, 3, 4, 5, 6, 5, 4, 3, 4, 5, 6, 7, \dots$ is not]. It is *strictly increasing* if $a_n < a_{n+1}$ for all n, *strictly decreasing* if $a_n > a_{n+1}$ for all n, and *strictly monotonic* if it is either strictly increasing or strictly decreasing.

These names vary, and, in particular, the phrase "monotonic increasing" is often used instead of "increasing", and "monotonic decreasing" instead of "decreasing", possibly because "increasing" is at some risk of being interpreted as "strictly increasing".

Definition 2.4. Let A be any non-null set; suppose $f : A \longrightarrow \mathbb{R}$. (This is a common way of abbreviating "f is a function from A to $\mathbb{R}$".) We say that f is *bounded above* if its image $f(A) := \{f(x) : x \in A\}$ is bounded above in $\mathbb{R}$; and similarly for the statements that f is *bounded below* or that f is *bounded*, which means that it is bounded both above and below.

Let $(a_n)_{n=1}^{\infty}$ be a bounded sequence in $\mathbb{R}$. By definition 2.4, this means that the set $A_1 := a(\mathbb{N}) = \{a_n : n \in \mathbb{N}\}$ is bounded. Let us suppose it has upper and lower bounds U and L (they might be its supremum and infimum, for instance):

$$(\forall n \in \mathbb{N}) \quad L \leq a_n \leq U\,.$$

It follows that the various obviously nonnull subsets of A_1

$$A_2 := \{a_n : 2 \leq n \in \mathbb{N}\} = \{a_n : n = 2, 3, 4, \dots\}\,, \quad A_3 := \{a_n : 3 \leq n \in \mathbb{N}\}\,,$$
$$A_k := \{a_n : k \leq n \in \mathbb{N}\} = \{a_n : n = k, k+1, k+2, \dots\} \quad \text{for any } k \in \mathbb{N}\,,$$

are also bounded above with U as an upper bound, and bounded below with L as a lower bound. By Dedekind's axiom, each of them has a supremum; let us write

$$b_k := \sup A_k = \sup\{a_n : n \geq k\} = \sup_{n \geq k} a_n = \sup_{n \geq k} a_n\,; \tag{7}$$

all the three last notations are in common use. Notice that, for any $k \in \mathbb{N}$,

$$L \leq a_k \leq b_k \leq U\,,$$

and also that $b_k \geq b_{k+1}$, since $b_k := \sup A_k$ is an upper bound for A_{k+1}. In particular, $\{b_k : k \in \mathbb{N}\}$ is non-null and bounded below (by L), and $(b_k)_{k \in \mathbb{N}}$ is monotonic decreasing.

Definition 2.5. Let (a_n) be a bounded sequence in $\mathbb{R}$. Its *upper limit* (or *limes superior*) $\limsup a_n$, also denoted by $\overline{\lim}\, a_n$, $\overline{\lim}_n a_n$, $\limsup_{n \to \infty} a_n$, and so on, is

$$\inf\{b_k : k \in \mathbb{N}\} = \inf_{k\geq 1} \sup_{n\geq k} a_n \,.$$

In a symmetrical way, one defines the *lower limit* or *limes inferior*:

$$\liminf a_n = \underline{\lim}\, a_n = \sup\{\inf\{a_n : k \leq n \in \mathbb{N}\} : k \in \mathbb{N}\} = \sup_{k\geq 1} \inf_{n\geq k} a_n \,.$$

Here $(\inf_{n\geq k} a_n)_{k\in\mathbb{N}}$ is monotonic increasing.

To define $\overline{\lim}$ and $\underline{\lim}$ for bounded sequences in a partially ordered set T we only need a weak version of Dedekind-completeness; namely, we need suprema and infima only for *countable* bounded subsets of T. When, as in $\mathbb{R}$, we have the full strength of Dedekind's Axiom, upper and lower limits for more general objects than sequences can be defined. But this does not concern us at the moment.

These ideas are really rather natural. Think of the graph of (a_n), consisting of the points (n, a_n) in $\mathbb{R}^2$. The line $y = \sup_{n\geq k} a_n$ is the lowest horizontal line that lies above all the points $(k, a_k), (k+1, a_{k+1}), (k+2, a_{k+2}), \ldots$; as k increases, it descends (in principle) towards the line $y = \overline{\lim}\, a_n$. Examples will make the point clear. But the importance of upper and lower limits is that *they make sense for any bounded sequence of real numbers*, whether or not it converges. One might say vaguely that they are what remains of the notion of limit when there isn't one (but the sequence is still bounded).

Lemma 2.6. *For any bounded sequence* (a_n) *in* $\mathbb{R}$,

$$\inf_{n\in\mathbb{N}} a_n \leq \underline{\lim}\, a_n \leq \overline{\lim}\, a_n \leq \sup_{n\in\mathbb{N}} a_n \,.$$

Proof. Let $l, m \in \mathbb{N}$. Then $\inf_{n\geq l} a_n \leq a_{\max(l,m)} \leq \sup_{n\geq m} a_n$; the term $a_{\max(l,m)}$ belongs to both the sets whose supremum or infimum is being taken. Thus, $\inf_{n\geq l} a_n$, for any specific l, is a lower bound for $\{\sup_{n\geq k} a_n : k \in \mathbb{N}\}$, and consequently

$$\inf_{n\geq l} a_n \leq \inf_{k\geq 1} \sup_{n\geq k} a_n = \overline{\lim}\, a_n \,.$$

This holds for any l; so $\overline{\lim}\, a_n$ is an upper bound for $\{\inf_{n\geq l} a_n : l \in \mathbb{N}\}$. It follows that

$$\underline{\lim}\, a_n = \sup_{l\geq 1} \inf_{n\geq l} a_n \leq \overline{\lim}\, a_n \,. \qquad \square$$

Definition 2.7. The bounded sequence (a_n) in $\mathbb{R}$ is said to be *order-convergent* when its upper limit and its lower limit are equal; their common value may be called the *order-limit* of the sequence.

This notion of convergence applies to sequences in any Dedekind-complete (or even "boundedly *countably* order-complete") partially ordered set. The definitions you have met previously involved real-valued sequences and positive numbers ϵ; so far we have not seen a single ϵ. It is high time to introduce one, and so to return to the real numbers exclusively.

Lemma 2.8. *Let* A *be a non-null subset of* $\mathbb{R}$*, and let* $x \in \mathbb{R}$*. Then* x *is the supremum of* A *if and only if the two following conditions are satisfied:*

(i) *for every* $a \in A$, $a \le x$ [that is, x is an upper bound for A in $\mathbb{R}$], *and*
(ii) *for every* $\epsilon > 0$, *there exists some* $a \in A$ *such that* $x - \epsilon < a$.

Dually, $y \in \mathbb{R}$ *is the infimum of* A *if and only if it is a lower bound for* A *and, for every* $\epsilon > 0$, *there is some* $b \in A$ *such that* $b < y + \epsilon$.

(When I write "for every $\epsilon > 0$", I always mean "for every *positive real number* ϵ".)

Proof. If x is the supremum of A, it is *an* upper bound for A; that is *(i)*. If it is the *least* upper bound, then, for any $\epsilon > 0$, $x - \epsilon$ (being strictly *less* than x) is *not* an upper bound for A—so there must be some $a \in A$ with $a > x - \epsilon$. This is *(ii)*.

Conversely, suppose *(i)* and *(ii)* are satisfied. By *(i)*, x is an upper bound for A. If it is not the least upper bound, there is an upper bound y for A with $y < x$. Take $\epsilon := x - y > 0$; then, by *(ii)*, there is some $a \in A$ with $y = x - \epsilon < a$, and this contradicts the assumption that y is an upper bound. Consequently x is indeed the least upper bound. □

Definition 2.9. Let (a_n) be a sequence in $\mathbb{R}$. It *converges* or *tends* or *is convergent* to a number $b \in \mathbb{R}$, and one writes $a_n \to b$ (or, if necessary, $a_n \to b$ as $n \to \infty$), if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n \in \mathbb{N}) \quad n \ge N \Longrightarrow |a_n - b| < \epsilon. \tag{8}$$

If (8) is satisfied, one also says that b is the *limit* of the sequence (a_n) (as $n \to \infty$).

For some motivation of this definition, see Appendix A.

The sequence (a_n) in $\mathbb{R}$ is said to *converge* or to be *convergent* (without mention of a limit) if there is *some* number b in $\mathbb{R}$ such that $a_n \to b$. To emphasize that this is *not* order-convergence, I might sometimes call it "metric convergence".

If a_n (for all $n \in \mathbb{N}$) and b are complex numbers (that is, if "(a_n) is a sequence in $\mathbb{C}$" and $b \in \mathbb{C}$), and $|a_n - b|$ is understood as the modulus of the complex number $a_n - b$, then (8) is the definition of convergence of the *complex* sequence (a_n) to the *complex* limit b. See §10 (and later sections) for this.

Example 2.10. Let (a_n) be a sequence in $\mathbb{R}$ and $a \in \mathbb{R}$ be such that, for each n, $a - 1/n \le a_n \le a$. Then $a_n \to a$.

Suppose $\epsilon > 0$. Then $1/\epsilon$ is also a positive number; let N be the least integer *exceeding* $1/\epsilon$, which is $[1/\epsilon] + 1$ and so is positive. ([] denotes the "integer part", 1.23.) Then

$$n \ge N \Longrightarrow |a_n - a| \le \frac{1}{n} \le \frac{1}{N} < \frac{1}{1/\epsilon} = \epsilon;$$

thus, $|a_n - a| < \epsilon$. The definition of convergence to a is satisfied.

Lemma 2.11. *Let* A *be a bounded nonempty subset of* $\mathbb{R}$. *There are sequences* $(a_n), (b_n)$ *in* A *such that*
(a) (a_n) *is nondecreasing and* $a_n \to \sup A$, *and*
(b) (b_n) *is nonincreasing and* $b_n \to \inf A$.

Proof. I prove *(a)*; *(b)* follows by reversing signs. Let $a := \sup A$. By 2.8*(ii)*, there is some $a_1 \in A$ such that $a - 1 < a_1 \le a$. Suppose $a_k \le a$ has been defined. If $a_k = a$, let

$a_{k+1} := a_k$; if $a_k < a$, then, by 2.8*(ii)*, there is some $a_{k+1} \in A$ such that

$$\max(a_k, a - 1/(k+1)) < a_{k+1} \leq a\,.$$

Then $a_{k+1} \geq a_k$ and $a_{k+1} > a - 1/(k+1)$, so that the inductive procedure constructs a nondecreasing sequence (a_n), and $a_n \to a$ by 2.10. □

Remark 2.12. $a_n \to b$ if and only if $|a_n - b| \to 0$.

Lemma 2.13. *If (x_n) is a sequence in $\mathbb{R}$, it can have at most one limit. That is, if $x_n \to x$ and $x_n \to y$, where $x, y \in \mathbb{R}$, then necessarily $x = y$.*

Proof. Let $x \neq y$; take $\epsilon := \frac{1}{2}|x - y|$. Then there exist $N_1, N_2 \in \mathbb{N}$ such that

$$n \geq N_1 \Longrightarrow |x_n - x| < \epsilon\,, \quad n \geq N_2 \Longrightarrow |x_n - y| < \epsilon\,.$$

Take $N := \max(N_1, N_2) \in \mathbb{N}$, and then, if $n \geq N$, both $|x_n - x| < \epsilon$ and $|x_n - y| < \epsilon$, so that, from the triangle inequality 1.10,

$$|x - y| = |(x - x_n) + (x_n - y)| \leq |x - x_n| + |x_n - y| < 2\epsilon = |x - y|\,.$$

This is absurd, and we must conclude that $x = y$. □

Remark 2.14. *(a)* Thus, if (a_n) does converge to a limit, the symbol $\lim_{n\to\infty} a_n$ can be employed to denote that limit unambiguously. It is rather undesirable (though some authors do it) to treat $b = \lim_{n\to\infty} a_n$ as entirely equivalent to $a_n \to b$, because, in general, the symbol "$\lim_{n\to\infty} a_n$" is MEANINGLESS. That is, most sequences do not have limits. For instance $1, 0, 1, 0, 1, 0, \ldots$ does not. Similarly, in ordinary language a phrase may make perfect grammatical sense, but not denote anything—for example, the phrase "the reigning King of France" has no denotation at this moment.

(b) To prove that a sequence (a_n) does converge to a given limit b, you have to show that for ANY positive number ϵ there is some value of N such that $n \geq N \Longrightarrow |a_n - b| < \epsilon$. Usually one gives a "formula" for a possible N in terms of ϵ; however, the definition doesn't *require* a formula, or even any explicit value for N, provided you can show that some N exists whatever the choice of positive ϵ. If ϵ is altered, the value of N you take will often also need to be changed; in this weak sense, N "depends on ϵ", and people write $N(\epsilon)$ to indicate "an N which works in (8) for the given ϵ".[1]

(c) The conditions ' $< \epsilon$ ', ' $\geq N$ ' in the definition may be changed (either or both) to ' $\leq \epsilon$ ' or ' $> N$ ' without altering the class of convergent sequences or their limits (though the N required for a given ϵ might have to be changed). Likewise, one could without affecting the sense of the definition restrict ϵ to positive rational values or even, for instance, to numbers of the form $1/k$ for $k \in \mathbb{N}$. You should try to convince yourself of these facts.

(d) $|a_n - b|$ should be thought of as the "distance" of a_n from b. Thus, (8) asserts that, whatever measure ϵ of "closeness to b" you take, the sequence eventually (at stage N and

[1] Some authors state 2.9 as "$a_n \to b$ if there is a *function* $N : (0, \infty) \longrightarrow \mathbb{N}$ such that, for any $\epsilon > 0$, $n \geq N(\epsilon) \Longrightarrow |a_n - b| < \epsilon$". If, for each ϵ, there is *some* $N \in \mathbb{N}$ satisfying (8), there will be by 1.1 a *least possible* such N, which one could take to be the function value $N(\epsilon)$. However, there are analogous definitions where the indices do not form a well-ordered set like $\mathbb{N}$. There may then be no least possible N, and one might need the Axiom of Choice to choose a specific N for each ϵ. It is simpler in all cases to accept that the notation $N(\epsilon)$ here is just an informal reminder: "an N that could be used in (8) for the given ϵ".

later) settles down to be "close to b" in the sense described by ϵ. This is *metric convergence* (to do with distances), as distinct from the order-convergence of 2.7.

(e) If $a_n \to b$, then (a_n) is bounded (that is, the set $\{a_n : n \in \mathbb{N}\}$ of terms of the sequence is bounded). For $(\exists N \in \mathbb{N}) \quad n \geq N \Longrightarrow |a_n - b| \leq 1$, so that

$$(\forall n \in \mathbb{N}) \quad |a_n| \leq C := \max(1 + |b|, |a_1|, |a_2|, \ldots, |a_{n-1}|)\,.$$

In the real case, $-C \leq a_n \leq C$ for all n; in the complex case, $|a_n| \leq C$ for all n.

Lemma 2.15. *If $a_n \to b$, and $(a_{\alpha(n)})$ is a subsequence of (a_n), then (a_n) is bounded, and $a_{\alpha(n)} \to b$. In particular, $a_{n+1} \longrightarrow b$.*

Proof. For any $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $n \geq N \Longrightarrow |a_n - b| < \epsilon$. As $\alpha(n) \geq n$ for all n, it follows that $n \geq N \Longrightarrow |a_{\alpha(n)} - b| < \epsilon$. □

Lemma 2.16. *If $a_n \to b$ and k is a real constant, then $ka_n \to kb$.*

Proof. If $k = 0$, then trivially a sequence of zeros tends to 0. But, if $k \neq 0$, then, for any $\epsilon > 0$, $\epsilon/|k| > 0$ too, and we are told there is some $N \in \mathbb{N}$ such that

$$n \geq N \Longrightarrow |a_n - b| < \frac{\epsilon}{|k|}$$

(in terms of the original sequence (a_n), we might call this "$N(\epsilon/|k|)$"). But then

$$n \geq N \Longrightarrow |ka_n - kb| = |k||a_n - b| < \epsilon\,.$$

Hence, definition 2.9 is satisfied for the sequence (ka_n) and limit kb. □

Theorem 2.17. *Suppose the sequence (a_n) in $\mathbb{R}$ is bounded and order-convergent with order-limit b. Then $a_n \to b$. Conversely, if $a_n \to b$, then (a_n) is bounded and order-convergent with order-limit b.*

Proof. **A.** Take any $\epsilon > 0$. As $b = \inf\{\sup_{n \geq k} a_n : k \in \mathbb{N}\}$, 2.8 tells us that there must be some value of k—call it μ— with $\sup_{n \geq \mu} a_n < b + \epsilon$. This implies that

$$n \geq \mu \Longrightarrow a_n \leq \sup_{n \geq \mu} a_n < b + \epsilon$$

(as the supremum is an upper bound). Dually, from the fact that b is the lower limit we deduce there is a value of k, which we may call ν, such that $\inf_{n \geq \nu} a_n > b - \epsilon$, or $n \geq \nu \Longrightarrow a_n > b - \epsilon$. Let us now define N to be $\max(\mu, \nu)$. Then, if $n \geq N$, it is true that both $n \geq \mu$ and $n \geq \nu$, and so

$$n \geq N \Longrightarrow b - \epsilon < a_n < b + \epsilon\,, \tag{9}$$

which is the same as $-\epsilon < a_n - b < \epsilon$ or $|a_n - b| < \epsilon$. In other words: given $\epsilon > 0$, we have proved the existence of an N such that (8) is true.

B. Conversely, if $a_n \to b$, then, for any $\epsilon > 0$ there exists some $N(\epsilon)$ such that (9) holds, and then

$$b - \epsilon \leq \inf_{n \geq N(\epsilon)} a_n \leq \sup_{n \geq N(\epsilon)} a_n \leq b + \epsilon\,. \tag{10}$$

Hence, in the first place, (a_n) is bounded. For *all* n,

$$\min(b-1, a_1, a_2, \ldots, a_{N(1)-1}) \leq a_n \leq \max(b+1, a_1, a_2, \ldots, a_{N(1)-1})\,. \tag{11}$$

For $n \geq N(1)$, $\min(b-1, a_1, , \ldots, a_{N(1)-1}) \leq b-1 \leq a_n$, from (10), and in the same way $a_n \leq b+1$. For $n \leq N(1)-1$, on the other hand,

$$\min(b-1, a_1, , \ldots, a_{N(1)-1}) \leq a_n \leq \max(b+1, a_1, \ldots, a_{N(1)-1})\,.$$

This proves (11) in all cases, and (11) shows that (a_n) is bounded, and has upper and lower limits. Now, taking any $\epsilon > 0$, (10) also tells us that

$$b - \epsilon \leq \sup_{N \geq 1} \inf_{n \geq N} a_n = \liminf a_n \leq \limsup a_n = \inf_{N \geq 1} \sup_{n \geq N} \leq b + \epsilon \tag{12}$$

(recall 2.6). (For the first inequality, $b - \epsilon \leq \inf_{n \geq N(\epsilon)} a_n$; for the last, $\sup_{n \geq N(\epsilon)} \leq b + \epsilon$). this can only be true if the middle terms are equal. Indeed, (12) implies that

$$\overline{\lim}\, a_n - \underline{\lim}\, a_n \leq (b+\epsilon) - (b-\epsilon) = 2\epsilon\,, \tag{13}$$

an inequality not depending on $N(\epsilon)$, which must, therefore, be true for *any* positive ϵ. If $\overline{\lim}\, a_n > \underline{\lim}\, a_n$, take $\epsilon := \frac{1}{3}(\overline{\lim}\, a_n - \underline{\lim}\, a_n) > 0$, and (13) is false. So $\overline{\lim}\, a_n \leq \underline{\lim}\, a_n$, and 2.6 shows that they must in fact be equal. □

Briefly, a *real* sequence is convergent to a limit if and only if it is bounded and order-convergent with the same limit. To prove this, we have exploited the special properties of $\mathbb{R}$.

One frequently used corollary (which can also be proved directly without difficulty) is

Corollary 2.18. *A bounded increasing sequence in* $\mathbb{R}$ *converges to its supremum. A bounded decreasing sequence converges to its infimum.*

Proof. Let (a_n) be increasing; then $\inf_{n \geq k} a_n = a_k$ for all $k \in \mathbb{N}$, and it follows that $\sup_{k \geq 1} \inf_{n \geq k} a_n = \sup_{k \geq 1} a_k = \sup a_n$. On the other hand, $\sup_{n \geq k} a_n = \sup_{n \geq 1} a_n$ (the earliest terms are no greater than the later ones) and $\inf_{k \geq 1} \sup_{n \geq k} a_n = \sup_{n \geq 1} a_n$. So $\liminf a_n = \limsup a_n = \sup a_n$, and the result follows from Theorem 2.17. The result for a decreasing sequence (b_n) may be proved similarly, or by considering $(-b_n)$ instead. □

Lemma 2.19. *Let* $0 < \xi < 1$ *in* $\mathbb{R}$. *Then* $\xi^n \to 0$ *as* $n \to \infty$.

Proof. Clearly $0 < \xi^{n+1} < \xi^n$ for all $n \in \mathbb{N}$ (simply multiply the inequalities of the hypothesis by $\xi^n > 0$). So (ξ^n) is a decreasing sequence bounded below by 0; by 2.18, it has a limit η, which is its greatest lower bound and so must be non-negative. On the one hand, as $\xi^n \to \eta$, then $\xi^{n+1} \to \xi\eta$ by 2.16; on the other, $\xi^{n+1} \to \eta$ by 2.15. By 20.4, these two limits must be the same: $\xi\eta = \eta$. As $\xi \neq 1$, it follows that $\eta = 0$. □

There are more informative and explicit proofs of this Lemma, but they require rather more discussion, and it is useful to see how one can exploit the "abstract nonsense" we have developed. The result itself is a minor obsession of mine. If you pick up some recent introductory texts on analysis, you will find that they "prove" the result *by assuming the*

existence and properties of the logarithmic function, long before they have defined it. I am not saying their arguments would necessarily be circular if they wrote them down in full; for there are ways of defining the logarithm, and proving its properties, that do not assume the Lemma at any step. But they all require a lot of preparation (quite possibly some theory of integration), which the authors in question have certainly not presented. I suppose their excuse is to give an easily grasped example of the definition (8) for students too incurious to care what a logarithm is.

There is a reason for someone sometime to have worried about this sort of thing. I commented at the outset that some important statements that had seemed obviously true were discovered to depend on unspoken assumptions. It is therefore important to be clear about your assumptions; and the logarithmic function is a whopping thing to take for granted.

Let me also insert here a conventional warning. The words "convergent" and "convergence" are used in many contexts, and their meaning depends on the context. Here, we are discussing convergence of *sequences of real numbers*. This concept differs, for example, from *convergence of series* (which we shall discuss later) or *convergence of integrals.*

Another way of relating upper and lower limits (defined by the order on $\mathbb{R}$) to metric convergence (defined from the distance in $\mathbb{R}$) is as follows.

Proposition 2.20. *Let (x_n) be a bounded sequence in $\mathbb{R}$. Then (x_n) has a convergent subsequence; and $\limsup x_n$ is the greatest number that is the limit of a convergent subsequence of (x_n). Similarly, $\liminf x_n$ is the least number that is the limit of a convergent subsequence.*

[A number that is the limit of a convergent subsequence of (x_n) is sometimes called a "subsequential limit" of (x_n). Thus, the set of subsequential limits of a bounded sequence is nonempty, and has both a greatest and a least member.]

Proof. Let $b := \underline{\lim}\, x_n$. Construct a subsequence $(x_{n(p)})_{p=1}^{\infty}$ by choosing the indices $n(p)$ inductively. Take (for instance) $n(1) := 1$, just to get started; and suppose $n(q)$ has been chosen. Now $b = \sup_{n\geq 1} \inf_{k\geq n} x_k$, by definition 2.5 (the sequence $(\inf_{k\geq n} x_k)_{n=1}^{\infty}$ is increasing, as at 2.4). Thus,

$$(\exists N(q) \in \mathbb{N}) \quad \inf_{k\geq N(q)} x_k > b - \tfrac{1}{q+1}$$

by 2.8*(ii)*. It follows that $x_k > b - \frac{1}{q+1}$ for all $k \geq N(q)$.

Let $M(q) := \max(n(q)+1, N(q))$. Then, $k \geq M(q) \Longrightarrow x_k > b - \frac{1}{q+1}$. However, by 2.5 (or 2.8*(i)*), $b \geq \inf_{k\geq M(q)} x_k$, and, by (the dual of) 2.8*(ii)*, there is some $l \geq M(q)$ such that $x_l < b + \frac{1}{q+1}$. So, $b - \frac{1}{q+1} < x_l < b + \frac{1}{q+1}$ and $l \geq n(q)+1$. Choose $n(q+1)$ to be this l. By the definition of $M(q)$, $n(q+1) \geq n(q)+1$.

This inductive procedure ensures that the sequence $(x_{n(p)})_{p=1}^{\infty}$ is indeed a subsequence of (x_n) and that $b - \frac{1}{p} < x_{n(p)} < b + \frac{1}{p}$ for every $p > 1$, which is sufficient to ensure that $x_{n(p)} \to b$.

A similar argument shows that $\overline{\lim}\, x_n$ is also a subsequential limit; or, alternatively, observe that $\overline{\lim}\, x_n = -\underline{\lim}(-x_n)$, which is the negative of a limit of a subsequence of $(-x_n)$ and so is the limit of a subsequence of (x_n).

Suppose now that $(x_{n(p)})$ is *any* subsequence of (x_n). Then, for any m,

$$\{x_{n(p)} : p \geq m\} \subseteq \{x_p : p \geq m\}$$

(as $n(p) \geq p$ for every p); so $\inf_{p\geq m} x_{n(p)} \geq \inf_{p\geq m} x_p$ for every m, and on taking suprema $\underline{\lim}\, x_{n(p)} \geq \underline{\lim}\, x_n$. If $(x_{n(p)})$ converges, its limit is $\underline{\lim}\, x_{n(p)}$, by 2.17. Thus, a subsequential limit is not less than $\underline{\lim}\, x_n$, and similarly cannot be greater than $\overline{\lim}\, x_n$. □

Remark 2.21. *(a)* Any definition of $\mathbb{R}$ that has the properties we desire must be complicated, as I have remarked. As $\mathbb{R}$ is uncountable, irrational numbers cannot all be "named" in the way that all rational numbers may be; and the definition 2.9 of convergence often cannot be applied directly, because the putative limit b must be specified before (8) can be checked. The upper and lower limits of the sequence, which we know to *exist* if the sequence is bounded, are frequently difficult to describe explicitly (i.e. other than as those upper or lower limits).

This is a serious matter. One commonly wants to *define* a previously unknown number as the limit of a sequence (in effect by "successive approximation"). So the best information we have so far is 2.18: if the sequence is monotonic and bounded, it is order-convergent.

(b) 2.18 is, indeed, precisely how an infinite decimal expansion defines a positive real number; its successive truncations form an increasing sequence of rationals that is bounded above. Similarly for negative real numbers, where the truncations form a decreasing sequence that is bounded below.

(c) However, for non-monotonic sequences, there is no corresponding way of seeing at a glance whether they converge or not. We can, nevertheless, characterize convergent sequences in $\mathbb{R}$ in a way that does not require knowledge of their limits. This was pointed out by Cauchy.

Definition 2.22. Let (a_n) be a sequence in $\mathbb{R}$. It is described as a *Cauchy sequence* if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \quad m, n \geq N \Longrightarrow |a_m - a_n| < \epsilon\,. \tag{14}$$

Nowadays "Cauchy" has become an adjective—'the sequence is Cauchy'. (To avoid this slightly absurd statement, Cauchy sequences were previously called "fundamental".) The terms of a sequence convergent to b get "closer and closer to b"; those of a *Cauchy sequence* get "closer and closer to *one another*".

"$m, n \geq N \Longrightarrow \cdots$" in (14) means "$(m \geq N \;\&\; n \geq N) \Longrightarrow \cdots$". This is fairly standard usage; and (14) is also often expressed by

$$|a_m - a_n| \to 0 \quad \text{as} \quad m, n \to \infty\,.$$

It is important to realise that "as $m, n \to \infty$" means that, for a given ϵ, there is an N for which $|a_m - a_n| < \epsilon$ if *both* m and n are greater than N.

Lemma 2.23. *A convergent sequence in* $\mathbb{R}$ *is Cauchy.*

Proof. Let $a_n \to b$. Given any $\epsilon > 0$, there is an N such that

$$n \geq N \Longrightarrow |a_n - b| < \tfrac{1}{2}\epsilon$$

(taking $\frac{1}{2}\epsilon$ for ϵ in (8)). If $m, n \geq N$, $|a_m - b| < \frac{1}{2}\epsilon$ and $|a_n - b| < \frac{1}{2}\epsilon$, and consequently $|a_m - a_n| \leq |a_m - b| + |b - a_n| < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon$. Thus the N which serves for $\frac{1}{2}\epsilon$ in (8) will serve for ϵ in (14). □

Lemma 2.24. *A Cauchy sequence in* $\mathbb{R}$ *is bounded.*

Proof. Let (a_n) be Cauchy. In (14), take $\epsilon := 1$; thus there is some N such that $|a_m - a_n| < 1$ whenever $m, n \geq N$. Hence, $|a_m - a_N| < 1$ when $m \geq N$, that is,

$$a_N - 1 < a_m < a_N + 1 .$$

Let $A := \max\{a_1, a_2, \ldots, a_{N-1}, a_N + 1\}$. Then $a_m < a_N + 1 \leq A$ when $m \geq N$, and $a_1, a_2, \ldots a_{N-1} \leq A$ by definition; so A is an upper bound for the whole sequence. Similarly, $B := \min\{a_1, a_2, \ldots, a_{N-1}, a_N - 1\}$ is a lower bound for (a_n). □

2.23 and 2.24 tell us that a convergent sequence is bounded—which we already know (it was much the same proof) from 2.17 (see **B** in the proof).

We now come to the big theorem of this section, sometimes called the (or Cauchy's) General Principle of Convergence. Our proof is very like 2.17**B**.

Theorem 2.25. *A sequence (a_n) in $\mathbb{R}$ is convergent if and only if it is Cauchy.*

Proof. One direction, convergent ⇒ Cauchy, we have already seen in 2.23. So suppose (a_n) is Cauchy. Then, by 2.24, (a_n) is bounded; so it has an upper and a lower limit (see 2.5), and, by 2.6, $\liminf a_n \leq \limsup a_n$.

If $\liminf a_n < \limsup a_n$, set $\epsilon := \frac{1}{3}(\limsup a_n - \liminf a_n) > 0$. As (a_n) is Cauchy, there is N such that $m, n \geq N \Longrightarrow |a_m - a_n| < \epsilon$, and so

$$\begin{aligned} m \geq N &\Longrightarrow a_N - \epsilon < a_m < a_N + \epsilon . \qquad \text{Hence,} \\ \underline{\lim}\, a_m &= \sup_{k\geq 1} \inf_{m\geq k} a_m \geq \inf_{m\geq N} a_m \geq a_N - \epsilon \quad \text{and} \\ \overline{\lim}\, a_m &= \inf_{k\geq 1} \sup_{m\geq k} a_m \leq \sup_{m\geq N} a_m \leq a_N + \epsilon , \end{aligned}$$

and it follows that

$$3\epsilon := \limsup a_m - \liminf a_m \leq (a_N + \epsilon) - (a_N - \epsilon) = 2\epsilon ,$$

which is absurd, as $\epsilon > 0$. Thus, $\limsup a_m > \liminf a_m$ is impossible.

Consequently, $\liminf a_m = \limsup a_m$. But 2.17 now proves that

$$a_n \to \liminf a_m = \limsup a_m ,$$

as desired. □

That every Cauchy sequence in $\mathbb{R}$ is convergent is often expressed by saying that $\mathbb{R}$ is *metrically complete*, or even just *complete* (although "complete" is an overused word; that is why I previously used "order-complete" at 1.15).

There are several possible properties an ordered field may possess that are "equivalent", in the sense that an ordered field satisfying any one of them may be proved to have all the others. We have taken Dedekind-completeness, which holds for $\mathbb{R}$ by Dedekind's construction, as our basic property. Cantor's (or Méray's) construction of $\mathbb{R}$ defined a "real number" to be an equivalence class of Cauchy sequences of rational numbers under an equivalence relation that "ought to ensure they have the same limit" if the limits existed; this definition ensured the metric completeness of $\mathbb{R}$ (and made it easy to prove its algebraic properties); Dedekind-completeness could be deduced as a somewhat non-trivial consequence. A third equivalent property is sequential compactness of bounded closed intervals, which we shall discuss soon.

For spaces other than $\mathbb{R}$, metric completeness tends to be more useful and general than order-completeness, partly because the idea of a distance seems to apply more often than the idea of a partial order with suitable properties.

The peculiar advantage of the proof I have given for 2.25 is that the concepts of upper and lower limits are of independent importance, for instance in the theory of stochastic processes.

We can now prove fairly painlessly (exactly as in 161)

Lemma 2.26. $\mathbb{R}$ *is uncountable.* □

§3. Sets of points in $\mathbb{R}$ and $\mathbb{R}^2$.

In earlier years I had here a long digression about metric spaces and topologies; that material has mostly been moved to MATH 313. (For the benefit of anyone who took the earlier course, I have added the "old" §§3–6 as Appendix B; but I have not attempted to sew them into a fully coherent exposition; there are repetitions.) Almost everything I do below will be stated in terms of convergence of sequences, which is enough for our purposes and is also historically accurate. It will, nevertheless, be obvious that some of the arguments ought to generalize—and, if you take 313, you will discover that sequences often give unnecessarily clumsy proofs.

I recall that $\mathbb{R}^2$ is the set of column vectors $\begin{bmatrix} \xi \\ \eta \end{bmatrix}$ where $\xi, \eta \in \mathbb{R}$, but I shall mostly *write* these vectors as ordered pairs (ξ, η) for convenience.

Definition 3.1. Suppose that $((\xi_n, \eta_n))_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}^2$ [the outer parentheses indicate a sequence; the *inner* parentheses signify the ordered pair (ξ_n, η_n), where $\xi_n \in \mathbb{R}$ and $\eta_n \in \mathbb{R}$], and $(\xi, \eta) \in \mathbb{R}^2$. We say that $(\xi_n, \eta_n) \to (\xi, \eta)$ as $n \to \infty$ if $\xi_n \to \xi$ and $\eta_n \to \eta$. (This is equivalent to the definition in terms of the standard Euclidean distance in $\mathbb{R}^2$, which you have seen in MATH 211; and to convergence in $\mathbb{C}$.) It is easy to check that analogues of 2.13 and 2.15 hold for this notion of convergence.

For convenience, let us write Ω for a subset of $\mathbb{R}$ or of $\mathbb{R}^2$ (or $\mathbb{C}$) as the case may be. Indeed, we could also consider subsets of $\mathbb{R}^n$ for any $n \in \mathbb{N}$.

Definition 3.2. Let A be a subset of Ω.

(a) By a sequence in A, we mean a sequence in Ω all of whose terms belong to A. (It is then automatically a sequence in Ω itself.)

(b) A point x of Ω is a *limit point* of A in Ω if there is a sequence (x_n) in A that (as a sequence in Ω) converges to x.

(c) The set of all limit points of A in Ω is called the *closure* of A in Ω.

(b) is perhaps the original meaning of the phrase "limit point", and will cause us no difficulty in this course. In general topological spaces, there are various other phrases that are in use for related ideas, with subtle shades of meaning that need not trouble us here. (To reduce confusion I avoided "limit point" altogether in earlier versions of 312; see Appendix B.) You should always check the definitions an author has in mind, because the same words may not carry quite the same sense in two different books.

Lemma 3.3. *If $A \subseteq \Omega \subseteq \mathbb{R}$, then $x \in \Omega$ is a limit point of A if and only if, for any $\epsilon > 0$, there is a point $a \in A$ such that $|x - a| < \epsilon$.* [Of course a usually "depends on" ϵ, in the same weak sense as at 2.12*(b)*— if you change ϵ, you will commonly need to choose a different a.]

Proof. "Only if": suppose there is a sequence (x_n) in A that tends to x. Then, by the definition 2.9, there is, for any $\epsilon > 0$, some N such that $n \geq N \Longrightarrow |x_n - x| < \epsilon$. So, for this ϵ, one may take $a := x_N$.

"If": suppose that, for each $n \in \mathbb{N}$, there is some $a_n \in A$ such that $|x - a_n| < 1/n$. Then the sequence (a_n) in A converges (in Ω) to x, which is, therefore, a limit point of A. □

As on other occasions, the "countable axiom of choice" is implicitly invoked here.

Example 3.4. *(a)* Any point x of $A \subseteq \Omega$ is also a limit point of A in Ω, since it is the limit of the sequence $x, x, x, \ldots$ with all terms equal to x.

(b) If $a < b$ in $\mathbb{R}$, the open interval (a, b) has a and b as limit points in $\mathbb{R}$. Thus a set A may, indeed, have limit points in $\mathbb{R}$ that are not in A.

(c) If $a \leq b$ in $\mathbb{R}$, any limit point in $\mathbb{R}$ of the closed interval $[a, b]$ must belong to $[a, b]$. Indeed, if (x_n) is a sequence in $[a, b]$ and $x_n \to x \in \mathbb{R}$, then, by 2.6 and 2.17,

$$a \leq \inf x_n \leq \liminf x_n = x = \limsup x_n \leq \sup x_n \leq b\,,$$

so that $x \in [a, b]$ too. (It is also easy, and perhaps simpler, to give a proof directly from the definition 2.9.)

(d) If the set A is finite (i.e. has only finitely many points), any limit point of A in Ω must belong to A. (Why? You should be able to give a simple proof.)

Definition 3.5. A subset A of Ω is *closed* in Ω if every limit point of A in Ω is a point of A.

Thus, a closed *interval* in $\mathbb{R}$ is a closed *set* in $\mathbb{R}$, by 3.4*(c)*, and any finite subset of Ω is closed in Ω, by 3.4*(d)*. Also, if $B \subseteq \Omega$, the closure of B in Ω (cf. 3.2*(c)*) is itself closed in Ω and, therefore, is the smallest closed set in Ω that includes B; for a limit point in Ω of the set of limit points of B in Ω is itself a limit point of B in Ω. This follows from 3.3 above (you should try to see why).

Definition 3.6. *(a)* A subset A of Ω is *open* in Ω if its complement in Ω is closed in Ω.

(b) The *interior* in Ω of the subset A of Ω is the complement in Ω of the closure of A in Ω.

"Open" and "closed" sets are of great importance. For historical reasons, they have been defined in various ways in various contexts. (In MATH 313, it is likely that open sets will be taken as fundamental; that is, certain sets will be *decreed* to be open in Ω.) However, it is important to grasp that, whatever you treat as the most basic idea, an open set is the COMPLEMENT of a closed set, and that, generally speaking, "most" sets in Ω will be neither themselves closed nor have closed complements. For instance, if $a < b$ in $\mathbb{R}$

$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$$

is not *closed* in $\mathbb{R}$, because a is a limit point of the set but not an element thereof, and is not *open* in $\mathbb{R}$, because b is a limit point of the complement but not a point of the complement. (There are also situations—not in $\mathbb{R}^n$—in which sets may be *both* open and closed.)

From 3.4, a "closed interval" is closed in $\mathbb{R}$; similarly, though slightly less obviously, an "open interval" in $\mathbb{R}$ is open in $\mathbb{R}$. So our vocabulary is consistent.

The next definition has in effect already been given (at 1.14) for subsets of $\mathbb{R}$.

Definition 3.7. *(a)* A subset A of $\mathbb{R}$ is *bounded* if there are numbers $L, U \in \mathbb{R}$ such that

$$(\forall x \in A) \quad L \leq x \leq U\,.$$

(L is a *lower bound* for A and U an *upper bound*.) Equivalently, $A \subseteq [L, U]$.

(b) $A \subseteq \mathbb{R}^2$ is *bounded* if there are $L_1, L_2, U_1, U_2 \in \mathbb{R}$ such that

$$(\forall (x_1, x_2) \in \mathbb{R}) \quad L_1 \leq x_1 \leq U_1 \;\;\&\;\; L_2 \leq x_2 \leq U_2\,.$$

Equivalently, $A \subseteq [L_1, U_1] \times [L_2, U_2]$.

A sequence (a_n) is bounded if the set $\{a_n : n \in \mathbb{N}\}$ is bounded.

Lemma 3.8. *If a nonempty subset A of $\mathbb{R}$ is closed and bounded above, then* $\sup A \in A$.

Proof. If $u := \sup A$, there is for each $n \in \mathbb{N}$ some $a_n \in A$ such that

$$u - \frac{1}{n} < a_n \leq u\,.$$

Check that $a_n \to u$. □

[I have again used a form of the (countable) Axiom of Choice, assuming we may select a_n for every n simultaneously. This is difficult to avoid, or to disbelieve, when one is using sequences.]

There is a dual lemma to the above for "bounded below" and inf.

Lemma 3.9. *Any bounded sequence (a_n) in Ω has a convergent subsequence.*

Proof. If $\Omega = \mathbb{R}$, 2.20 shows that there is a subsequence convergent to $\overline{\lim}\, a_n$.

Suppose $\Omega = \mathbb{R}^2$, and that $a_n = (x_n, y_n)$, where $x_n, y_n \in \mathbb{R}$. Then (x_n) is a bounded sequence in $\mathbb{R}$ (see 3.7*(b)*), and so has a convergent subsequence $(x_{n(k)})_{k=1}^{\infty}$. However, the corresponding subsequence $(y_{n(k)})_{k=1}^{\infty}$ of (y_n) is also a bounded sequence in $\mathbb{R}$, and so has a *further* subsequence $(y_{n(k(p))})_{p=1}^{\infty}$ that converges in $\mathbb{R}$. Then $\big((x_{n(k(p))}, y_{n(k(p))})\big)_{p=1}^{\infty}$ is a subsequence of (a_n) that converges in $\mathbb{R}^2$. □

Definition 3.10. A subset A of Ω is said to be *sequentially compact* if every sequence in A has a subsequence that converges *to a point of* A. [Cf. 23.3.]

The idea of "compactness" appears in various guises. In its most general formulations, it is one of the most important and far-reaching concepts in mathematics: a version of "finiteness" that applies to "continuous" mathematics. It will be discussed more seriously in MATH 313. (See also §22 of Appendix B.) For our purposes, sequential compactness will be sufficient; and, *in metric spaces*, sequential compactness is equivalent to the other versions (cf. 23.10). Notice the essential point that the limit of the convergent subsequence has itself to be in A; sequential compactness is—vaguely speaking—intended to be a property of the set A alone, not of the way it lies in Ω.

Theorem 3.11. *A subset A of Ω is sequentially compact if and only if it is closed and bounded in Ω.*

Proof. Suppose A is not closed. Then there is some limit point of A that is not in A; that is, there is a sequence (a_n) in A such that $a_n \to a \notin A$. But then any subsequence of (a_n) also tends to a (and so does *not* tend to any point of A, by 93

the analogues of 2.13 and 2.15).

Suppose A is not bounded; for instance, suppose that there is no number L_1 in 3.7***(b)***. Then, for each $n \in \mathbb{N}$, there is some $a_n = (x_n, y_n) \in A$ with $x_n < -n$. But then every subsequence of (x_n) is unbounded, and cannot converge (by 2.15); thus, (a_n) does not have any convergent subsequence, and A is not sequentially compact. A similar argument applies to each of the other boundedness conditions (U_1, L_2, U_2) of 3.7***(b)***.

So, conversely, suppose A is closed and bounded. Let (a_n) be any sequence in A. It is certainly a bounded sequence in Ω. By 3.9, it has a convergent subsequence (with limit a) in Ω. But, as A is closed, 3.5 assures us that $a \in A$. Thus A is sequentially compact. □

Sequential compactness may be defined for subsets of an arbitrary metric space (cf. Appendix B). But this Theorem does *not* extend to general metric spaces; it is a special property of $\mathbb{R}^n$.

I remarked after 2.25 that Dedekind-completeness, as a property of an ordered field, is equivalent to metric completeness. The sequential compactness of bounded closed intervals (or, indeed, of the interval $[0, 1]$, from which the others may be deduced) is another equivalent property. But sequential compactness is a property that may be enjoyed by some sets in a general metric space, and is useful in many ways, whereas Dedekind-completeness refers to a partial order, and partial orders are not as useful as metrics.

Remark 3.12. If a subset A of $\mathbb{R}$ is sequentially compact and nonempty, then it is bounded and nonempty, so it has a supremum and infimum. Being closed, it must by 3.8 contain both of them. They are, therefore, its greatest and least elements. People sometimes say the supremum or infimum is *attained*, i.e. it is itself an element of A.

§4. Series.

In ordinary language the words "series" and "sequence" are more or less interchangeable. In mathematical analysis, there is a distinction, which, however, has to do rather with the properties we are interested in than with the things themselves. We describe a sequence (a_n) in $\mathbb{R}$ (or in $\mathbb{R}^2$ or $\mathbb{C}$) as a *series* when our focus is on the convergence not of the original sequence itself, but of the associated *sequence of partial sums* (s_n), where

$$(\forall n \in \mathbb{N}) \quad s_n := \sum_{i=1}^{n} a_i .$$

Despite frequent use of the word "series", then, I shall not give a formal definition of a *series* as such. It is just a sequence in which our primary interest is the partial sums. Hence, we often write "the series $\sum a_n$", the a_n being the terms of the sequence and the summation sign suggesting the partial sums. (Recall our tacit assumption that the indices go through $\mathbb{N}$.)

[It's worth noting that this distinction between "series" and "sequence" is not preserved in all other mathematical contexts: a "time series" may often be just a sequence of observations.]

Definition 4.1. The SERIES $\sum a_n$ is said to *converge* or to be *convergent* when its sequence of partial sums (s_n), where $s_n := \sum_{i=1}^{n} a_n$ for each n, converges to a limit s. s is then called the *sum* (or the *sum to infinity*) of the series, and we write

$$s = \sum_{i=1}^{\infty} a_i = \sum_{n=1}^{\infty} a_n \,,$$

or less formally $s = a_1 + a_2 + a_3 + \cdots$.

Remark 4.2. *(a)* This definition is the same whether (a_n) takes values in $\mathbb{R}$, in $\mathbb{C}$, in $\mathbb{R}^2$, or in other cases where "addition" and "convergence" may be defined. "Convergence" of a *series* is in principle different from "convergence" of a *sequence*.

(b) A series that does not converge is said to *diverge*.

(c) I shall try to preserve a distinction between $\sum a_n$ (the $\sum$ is just a reminder that we are dealing with a series) and $\sum_{n=1}^{\infty} a_n$ (which denotes the sum to infinity of the series, if it exists); but many authors confuse the notations or use only one in both senses.

(d) For real series one can distinguish several types of divergence of sequences and consequently of series (via the sequence of partial sums): divergence "to ∞" or "to $-\infty$", "finite oscillation", "infinite oscillation", etc. I shall not go into this—the ideas are pretty obvious; but please note that "∞" and "$-\infty$" are merely symbols that occur in certain abbreviated expressions, and have not been defined as genuine mathematical objects. They are, as it were, metaphorical, at least for our purposes.

(e) The algebraic operation of addition of real or complex numbers, or of vectors in $\mathbb{R}^2$, only adds two numbers or vectors at a time. The operation may be extended by induction to sum any *finite* set of numbers; but no algebraic meaning can be attached to an expression like $\sum_{n=1}^{\infty} a_n$. Our definition is that the sum to infinity, denoted $\sum_{n=1}^{\infty} a_n$ (the symbol may denote a number or vector), will *exist* only when the series $\sum a_n$ converges, and then it will denote the limit of the sequence of partial sums. It is not a *sum* in the algebraic sense at all. As with $\lim_{n\to\infty} a_n$, the expressions $\sum_{n=1}^{\infty} a_n$ or $a_1 + a_2 + a_3 + \cdots$ often have no denotation (do not refer to anything), just like the phrases "the current Emperor of China" or "the navy of Luxemburg" or "the greatest prime number". For instance, the sum to infinity $\sum_{n=1}^{\infty} (-1)^{n+1}$ (or $1 - 1 + 1 - 1 + 1 - \cdots$) has no meaning, since the partial sums of the series $s_1 = 1$, $s_2 = 0$, $s_3 = 1$, $s_4 = 0$, and so on; they form the sequence $1, 0, 1, 0, 1, 0, \ldots$, which does not converge to a limit. (Why not?)

It is interesting to note that the set operations of union and intersection are not "algebraic" in the sense I have described. For *any number* of sets (possibly countably or even uncountably infinitely many) there is no difficulty in defining their union or their intersection.

Lemma 4.3. *Let* $(s_n), (t_n)$ *be* sequences *in* $\mathbb{R}$ [or $\mathbb{C}$ or $\mathbb{R}^2$], *and* $\lambda, \mu \in \mathbb{R}$ [or $\mathbb{C}$]. *If* $s_n \to s$ *and* $t_n \to t$, *then* $\lambda s_n + \mu t_n \to \lambda s + \mu t$. [In effect done in Tut. Ex. 3.]

Proof. Take any $\epsilon > 0$. Then there exist L, M such that

$$\left.\begin{aligned} n \geq L &\Longrightarrow |s_n - s| < \frac{\epsilon}{|\lambda| + |\mu| + 1} \quad \text{and} \\ n \geq M &\Longrightarrow |t_n - t| < \frac{\epsilon}{|\lambda| + |\mu| + 1} \,. \end{aligned}\right\} \qquad (*)$$

Let $N \coloneqq \max(L, M)$. Then, if $n \geq N$,

$$\left.\begin{aligned} |(\lambda s_n + \mu t_n) - (\lambda s + \mu t)| &\le |\lambda||s_n - s| + |\mu||t_n - t| \\ &\le \frac{|\lambda|\epsilon + |\mu|\epsilon}{|\lambda| + |\mu| + 1} < \epsilon\,, \end{aligned}\right\} \qquad (**)$$

and this proves the result. □

I have set out this proof in the customary way. Given ϵ, we want an N, and therefore we arrange the right-hand sides at (*) so that ϵ itself appears on the right of (**). We could, if we wished, use $|\lambda| + |\mu|$ in the denominator instead of $|\lambda| + |\mu| + 1$, but then we should have to deal separately with the (admittedly trivial) case $\lambda = \mu = 0$, for which we cannot divide by $|\lambda| + |\mu|$. Until you are accustomed to it, this sort of fiddling may seem odd, but it is quite usual in textbooks and elsewhere and not entirely pointless.

Proposition 4.4. *Let $\sum a_n$, $\sum b_n$ be convergent real or complex* series, *and let λ, μ be real or complex numbers. Then $\sum (\lambda a_n + \mu b_n)$ is also convergent, and*

$$\sum\nolimits_{n=1}^{\infty} (\lambda a_n + \mu b_n) = \lambda\Big(\sum\nolimits_{n=1}^{\infty} a_n\Big) + \mu\Big(\sum\nolimits_{n=1}^{\infty} b_n\Big)\,.$$

Proof. Let s_n, t_n be the nth partial sums of $\sum a_n$, $\sum b_n$ respectively. Then $\lambda s_n + \mu t_n$ is the nth partial sum of $\sum (\lambda a_n + \mu b_n)$, since that is a *finite* sum which may be rearranged at will, and the result then follows from the preceding Lemma. □

Lemma 4.5. (a) *Suppose that $\sum b_n$ is obtained from $\sum a_n$ by omitting finitely many terms at the beginning (and renumbering). Then each series converges if and only if the other does.*

(b) *Similarly, suppose that $\sum b_n$ is obtained from $\sum a_n$ by omitting certain terms a_k (possibly infinitely many) whose value is 0, and renumbering. Each series converges if and only if the other does.*

Proof. Let $b_n = a_{n+k}$; this means that the terms $a_1, a_2, \ldots, a_k$ are omitted at the beginning of $\sum a_n$. If $\sum_{n=1}^{k} a_n = q$ say, then for any $m \in \mathbb{N}$

$$\sum\nolimits_{n=1}^{m} b_n = \Big(\sum\nolimits_{n=1}^{m+k} a_n\Big) - q\,,$$

so that the partial sums of $\sum b_n$ converge if and only if the partial sums of $\sum a_n$ do. This proves *(a)*, and the proof of *(b)* is similar (though the bookkeeping is a little worse). □

Many criteria have been given to ensure series of certain types converge, and some of them are very ingenious. The question was a very attractive one. There can be no decisive method to determine whether an arbitrary series converges, and all the tests for convergence that mathematicians have found must apply only to restricted (but presumably useful) classes of series. I shall prove here only the most straightforward and most memorable criteria.

Theorem 4.6. *Let $\sum a_n$ be a series of non-negative real terms. It converges if and only if the sequence (s_n) of its partial sums is bounded above.*

Proof. Clearly $s_{n+1} = s_n + a_{n+1} \geq s_n \geq 0$ for all n. The sequence of partial sums is increasing and bounded below by 0; by 2.15 and 2.18, it converges (to its supremum) if and only if it is bounded above. □

Of course 4.5*(a)* implies that it would suffice to suppose the terms are nonnegative "from some stage onwards", that is, $(\exists N \in \mathbb{N})\ n \geq N \Rightarrow a_n \geq 0$.

Corollary 4.7. (The comparison test). *Let $\sum a_n, \sum b_n$ be series of real non-negative terms, $K \geq 0$, and $0 \leq a_n \leq Kb_n$ for each n. If $\sum b_n$ converges, then so does $\sum a_n$, and*

$$0 \leq \sum_{n=1}^{\infty} a_n \leq K \sum_{n=1}^{\infty} b_n . \tag{15}$$

Proof. Letting s_n, t_n be the partial sums as before, we know by 4.6 that (t_n) is bounded and converges to its supremum B. Hence, for any n,

$$0 \leq s_n = \sum_{i=1}^{n} a_i \leq \sum_{i=1}^{n} Kb_i = Kt_n \leq KB ,$$

so that (s_n) is bounded above by KB, and so convergent to its supremum, which cannot exceed the upper bound KB. This proves (15). □

The contrapositive is also useful: if $\sum a_n$ diverges, then so does $\sum b_n$.

Example 4.8. If $0 \leq a < 1$, $\sum a^n$ converges. If $a \geq 1$, it does not.

Proof. Write s_n for the nth partial sum, as before. Then

$$\begin{aligned} (1-a)s_n &= \left(\sum_{i=1}^{n} a^i\right) - a\left(\sum_{i=1}^{n} a^i\right) \\ &= \left(\sum_{i=1}^{n} a^i\right) - \left(\sum_{i=1}^{n} a^{i+1}\right) \\ &= a - a^{n+1} , \end{aligned}$$

as all the intermediate terms $a^2, a^3, \ldots, a^n$ cancel. As $a \neq 1$, we deduce that

$$s_n = \frac{a(1-a^n)}{1-a} .$$

(This is the familiar formula for the sum of a geometrical progression.) If $0 \leq a < 1$, $a^n \to 0$ as $n \to \infty$, by 2.19, and so by 4.3 $s_n \to a/(1-a)$. If $a \geq 1$, then $s_n \geq n$ trivially, so it is unbounded and has no limit as $n \to \infty$. □

Notice that this is a rare case where the "sum to infinity" of a series can be given by a simple formula without a "limit" sign—"in closed form" is the jargon you sometimes see.

Now for an extremely important example.

Example 4.9. The series $\sum \frac{1}{n}$ does *not* converge.

Proof. Some of the partial sums s_n can be estimated as follows:

$$\begin{gathered} s_1 = 1,\ s_2 = 1 + \tfrac{1}{2},\ s_4 = s_2 + \tfrac{1}{3} + \tfrac{1}{4} > s_2 + \tfrac{1}{2}, \\ s_{2^{k+1}} = s_{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} > s_{2^k} + \frac{2^k}{2^{k+1}} = s_{2^k} + \tfrac{1}{2} . \end{gathered}$$

The obvious induction gives $s_{2^k} \geq 1 + \frac{1}{2}k$ for all $k \geq 0$ (indeed, if $k \geq 2$ the inequality is strict). But this shows the partial sums are not bounded. □

This proof is said to have been given by Oresme around 1360 (before Gutenberg was born), but to have been 'lost' (presumably this means that the book containing it was not "published" in the modern sense, and, therefore, not read by competent readers, until much later). The method of proof, when it is formalized and strengthened as follows, is sometimes called *Cauchy's condensation test*.

Lemma 4.10. *Let $\sum a_n$ be a series of non-negative terms, such that the* sequence (a_n) *is decreasing. The series converges if and only if $\sum 2^n a_{2^n}$ converges.*

The idea is that the behaviour of the original series is "condensed" to the series consisting of the terms whose indices are the powers of 2.

Proof. If s_n is the nth partial sum of the original series and t_k the kth partial sum of the condensed series, with $t_0 := a_1$, then, for $n \geq 1$,

$$\begin{aligned} s_{2^n-1} &= a_1 + (a_2 + a_3) + \cdots + (a_{2^{n-1}} + a_{2^{n-1}+1} + \cdots + a_{2^n-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^{n-1}a_{2^{n-1}} = a_1 + t_{n-1}\,, \quad \text{whilst} \\ s_{2^n} &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{n-1}+1} + \cdots + a_{2^n}) \\ &\geq a_1 + a_2 + 2a_4 + \cdots + 2^{n-1}a_{2^n} = a_1 + \tfrac{1}{2}t_n\,. \end{aligned}$$

If $\sum 2^n a_{2^n}$ converges, then, from 4.6, (t_k) is bounded above, say by T, and for any $n \in \mathbb{N}$

$$s_n \leq s_{2^n-1} \leq a_1 + T\,,$$

so that (s_n) is bounded and $\sum a_n$ converges; conversely, if $\sum a_n$ converges, $s_n \leq S$ for all n, and for any k

$$t_k \leq 2(s_{2^k} - a_1) \leq 2(S - a_1)\,,$$

which ensures $\sum 2^n a_{2^n}$ converges. □

Example 4.11. Let us cheat and assume that we know what n^α means, and its basic properties, for any $n \in \mathbb{N}$ and $\alpha \in \mathbb{R}$. (The problem is with the meaning of n^α when α is irrational; it would require a considerable digression to give a watertight definition at this point.) Then, if $\alpha > 0$, $n^{-\alpha}$ decreases as n increases, and by the Lemma $\sum n^{-\alpha}$ converges if and only if $\sum 2^n(2^n)^{-\alpha} = \sum 2^{n(1-\alpha)}$ does. This is a geometric series; by 4.8, it converges when $\alpha > 1$ and diverges when $\alpha \leq 1$.

This fact makes it tempting to suppose that, in some sense, the "harmonic series" $\sum \frac{1}{n}$ constitutes a "boundary" between convergent and divergent series. However, there are series whose terms diminish more rapidly than $1/n$ and which are still divergent.

Consider $\sum a_n$, where $a_1 = 1$ and, for $n \geq 2$, $a_n = \dfrac{1}{n \log n}$. (More cheating, for the logarithm has not yet been defined.) Again, a_n decreases as n increases; so the condensation test applies, and the given series converges if and only if

$$\sum \frac{2^n}{2^n \log(2^n)} = \sum \frac{1}{n \log 2}$$

converges. But it diverges, as we have seen.

WARNING. The condensation test is only occasionally useful, because the terms of the series $\sum 2^n a_{2^n}$ are often very messy. Its main value is in settling the convergence or divergence of a few standard series such as those above.

All so-called "convergence tests" apply to special classes of series, and their importance usually lies in the frequent occurrence of these classes. The test that "every schoolchild knows", sometimes called "d'Alembert's test", may be more precisely stated as follows.

Lemma 4.12. (The ratio test.) *Suppose* $\sum a_n$ *is a series of* positive *terms.*

(a) *If* $\limsup_{n\to\infty} \dfrac{a_{n+1}}{a_n} < 1$, *then* $\sum a_n$ *converges.*

(b) *If* $\liminf_{n\to\infty} \dfrac{a_{n+1}}{a_n} > 1$, *then* $\sum a_n$ *diverges.*

The ratios all make sense, because every term a_n is positive.

The form in which the test is usually applied is more demanding. If the ratio a_{n+1}/a_n *has a limit* as $n \to \infty$ (which need not be the case!) then the series converges when that limit is less than 1 and diverges when the limit is greater than 1. If the limit exists but is equal to 1, the series may diverge or may converge.

Proof. If $\limsup_{n\to\infty} \dfrac{a_{n+1}}{a_n} < 1$, choose a number b such that

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} := \inf_{k\geq 1}\sup_{n\geq k} \frac{a_{n+1}}{a_n} < b < 1 .$$

There is a k such that $m \geq k \Longrightarrow \dfrac{a_{m+1}}{a_m} \leq \sup_{n\geq k} \dfrac{a_{n+1}}{a_n} < b < 1$, by 2.8*(ii)*. Therefore, if $n \geq k$,

$$0 < a_n \leq b^{n-k} a_k$$

(by a simple induction on n). The truncated series $a_k + a_{k+1} + \cdots$ converges by comparison with the geometric series $a_k + a_k b + a_k b^2 + \cdots$ of common ratio b.

If $\liminf_{n\to\infty} \dfrac{a_{n+1}}{a_n} > 1$, there exists N such that $n \geq N \Longrightarrow \dfrac{a_{n+1}}{a_n} > 1$; it follows that, for $n \geq N$, $a_n > a_N$. The truncated series $a_N + a_{N+1} + \cdots$ diverges by comparison with $a_N + a_N + \cdots$.

In both cases, the conclusion now results from 4.5*(a)*. □

The frequent usefulness of the ratio test in many practical situations should not mislead you into believing it is always helpful. A silly but convincing example is $\sum a_n$, where

$$a_n := 2^{-n} \quad \text{when } n \text{ is a prime}, \qquad a_n := n^{-2} \quad \text{otherwise.}$$

It is quite easy (use 4.4 and 4.5*(b)*, and of course the fact that there are infinitely many primes) to see that the series converges, but the ratio test by itself tells us nothing.

In principle, the ratio test, the comparison test, or the condensation test apply only to series of nonnegative real terms. But they can be helpful more generally, because

Proposition 4.13. *Let* (a_n) *be a series of* complex *terms. If* $\sum |a_n|$ *converges, then* $\sum a_n$ *converges; furthermore,* $|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$.

Proof. Let $s_n := \sum_{i=1}^{n} a_i$ and $t_n := \sum_{i=1}^{n} |a_i|$. Thus the hypothesis says $t_n \to t$ as $n \to \infty$. By 2.23, (t_n) is Cauchy. I assert that, as a consequence, (s_n) must also be Cauchy; for, given $\epsilon > 0$, there exists N such that $m, n \geq N \Longrightarrow |t_m - t_n| < \epsilon$, so

$$|s_m - s_n| = \left|\sum_{i=\min(m,n)+1}^{\max(m,n)} a_i\right| \leq \sum_{i=\min(m,n)+1}^{\max(m,n)} |a_i| = |t_m - t_n| < \epsilon \tag{16}$$

by the triangle inequality for $|\cdot|$. In other words, the same N that works for (t_n) and ϵ also works for (s_n) and ϵ. Thus, by 2.25, (s_n) also converges.

Finally, $|s_n| = |\sum_{i=1}^{n} a_i| \leq \sum_{i=1}^{n} |a_i| = t_n$ for all n. Hence,

$$\left|\sum_{n=1}^{\infty} a_n\right| = \lim |s_n| \leq \lim t_n = \sum_{n=1}^{\infty} |a_n| . \qquad \square$$

Thus, for instance, $\sum i^n/n^2$ converges, because $\sum 1/n^2$ does.

Definition 4.14. A series $\sum a_n$ of complex (or real) terms is described as *absolutely convergent* if the series $\sum |a_n|$ is convergent.

Lemma 4.15. *A complex series $\sum a_n$ is absolutely convergent if and only if the series of real parts $\sum \Re a_n$ and of imaginary parts $\sum \Im a_n$ are both absolutely convergent.*

Proof. $|\alpha_n| \leq |\Re \alpha_n| + |\Im \alpha_n|$ and $|\Re \alpha_n| \leq |\alpha_n| \geq |\Im \alpha_n|$. $\square$

Notice that absolute convergence implies convergence only because of the *metric completeness* of $\mathbb{R}$ or $\mathbb{C}$. Non-mathematicians sometimes treat it as obvious (certainly when I was first told about it there was no hint that it was not), but there are analogous situations where a similar statement is false.

Definition 4.16. A series that is convergent but is not absolutely convergent is called *conditionally convergent*.

The method of proof of 4.13 yields another easy and important fact, which could have been pointed out much earlier, but is perhaps less confusing now.

Lemma 4.17. *If the* series $\sum a_n$ *(of real or complex terms) converges, then the* sequence (a_n) *of its terms tends to* 0.

Proof. Let $s_n = \sum_{i=1}^{n} a_i$ as before. The hypothesis means that $s_n \to s$ as $n \to \infty$ for some s. But then $a_n = s_{n+1} - s_n \to s - s = 0$ as $n \to \infty$. $\square$

This lemma shows at a glance that some series do not converge. For instance, there can be no possible choice of terms $\epsilon_n = \pm 1$ that will make $\sum \epsilon_n$ converge, because ϵ_n does not converge to 0. (This settles the divergence of $1 - 1 + 1 - 1 + \cdots$ at a glance.)

IMPORTANT. Politicians, journalists, and managers of all kinds tend (with rare exceptions) to believe that the Lemma is the other way round, i.e. that a series does converge if its terms tend to 0, but this is **OBVIOUSLY** false on the basis of our arguments. The harmonic series $\sum 1/n$ does not converge, by 4.9, although its terms tend to 0 by 2.10.

Lemma 4.18. *If the series* $\sum a_n$ *(of real or complex terms) converges, then the* sequence of partial sums (s_n)*, where* $s_n := \sum_{k=1}^{n} a_k$*, must be Cauchy.* □

In the real case this follows instantly from 2.23, but exactly the same proof as I gave there still works in $\mathbb{C}$. The fact has significant consequences.

The last of the elementary tests is the *alternating series test* (sometimes called *Leibniz's test*), the only easy standard result that enables us to detect conditionally convergent series.

Proposition 4.19. *Let* (a_n) *be a sequence of non-negative real terms, and suppose that it is a decreasing sequence with limit* 0 *(that is,* $a_n \downarrow 0$ *as* $n \to \infty$*). Then the series* $a_1 - a_2 + a_3 - a_4 + \cdots$*, or* $\sum (-1)^{n+1} a_n$*, is convergent. If* s_n *is its* n*th partial sum, and* $s := \lim s_n$ *the sum to infinity, then* $s_1 \geq s_3 \geq s_5 \geq \cdots \geq s \geq \cdots \geq s_6 \geq s_4 \geq s_2$.

Proof. Firstly, notice that an odd partial sum

$$s_{2n+1} = s_{2n-1} - a_{2n} + a_{2n+1} \leq s_{2n-1}$$

(since $a_{2n+1} \leq a_{2n}$ by hypothesis), so that the sequence (s_{2n-1}) of *odd* partial sums is decreasing. Similarly,

$$s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geq s_{2n}\,,$$

so that the sequence (s_{2n}) of *even* partial sums is increasing. From these two facts, we see that, for any choice of m and n, so that $2m \leq 2\max(m,n)$ and $2n+1 \leq 2\max(m,n)+1$,

$$s_{2m} \leq s_{2\max(m,n)} \leq s_{2\max(m,n)} + a_{2\max(m,n)+1} = s_{2\max(m,n)+1} \leq s_{2n+1}\,.$$

The sequence (s_{2n+1}) of odd partial sums is decreasing and bounded below by *any* even partial sum. By 2.18, it converges to its greatest lower bound, c say, and $s_{2n+1} \geq c \geq s_{2m}$ for any m and n. Likewise, the sequence of even partial sums is increasing and bounded above by c; so it converges to its supremum d, and $s_{2n+1} \geq c \geq d \geq s_{2m}$ for any m and n. It follows that $0 \leq c - d \leq s_{2n+1} - s_{2n} = a_{2n+1}$ for any n. However, as $a_{2n+1} \downarrow 0$ by hypothesis, this is only possible if $c = d$. The even partial sums and the odd partial sums converge to the same limit; hence, *all* partial sums converge to the common limit of (s_{2n}) and (s_{2n+1}). [Why?] □

Example 4.20. The harmonic series $\sum 1/n$ diverges by by 4.9, but the superficially similar series $\sum (-1)^{n+1}/n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ *converges* by 4.19. It is, therefore, a "conditionally convergent" series. It converges, but not absolutely.

A more extreme example, since its terms tend only very slowly to 0, is

$$\sum \frac{(-1)^{n+1}}{\text{logloglog}(n+25)}\,.$$

(The '25' ensures that the logloglog makes sense for $n \geq 1$. Perhaps I should add here that I use 'log' to denote *natural* logarithms. This is fairly usual in serious mathematics, because only natural logarithms are mathematically interesting. In lower-level courses we often use 'ln' to avoid confusing people like engineers who may use "common" logarithms to base 10; there's nothing wrong in this, but it isn't what mathematicians customarily do.) This series is convergent by 4.19, but the series of absolute values is "very divergent".

Remark 4.21. It is tempting to suppose that the sum to infinity of an infinite series should not depend on the order in which the terms appear, for instance to assume

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} - \cdots = 1 - \tfrac{1}{2} - \tfrac{1}{4} + \tfrac{1}{3} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{5} - \cdots , \tag{17}$$

since every term of either series appears once and once only in the other. This is certainly true for *finite* sums. But, as I pointed out at 4.2*(e)*, a sum to infinity is not really a sum at all in an algebraic sense; it is only a limit of partial sums. The partial sums depend on the order of the terms, so we have no right to expect their limit always to be the same. Every partial sum after the second for the right-hand series of (17) is significantly less than the corresponding sum for the left-hand series, as it takes into account more negative and fewer positive terms. Let us make this question more precise.

Definition 4.22. Let $\sum a_n$ be a (real or complex) series. A *rearrangement* of $\sum a_n$ is a series $\sum a_{\sigma(n)}$, where $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ is a bijection. [Thus, every number that appears as a term of the original series appears in the rearranged series exactly the same number of times.]

The big theorem on this topic, customarily attributed to Riemann, is

Theorem 4.23. (a) *If a real or complex series $\sum a_n$ is absolutely convergent, then any rearrangement of it is also convergent with the same sum.*

(b) *If a real series $\sum a_n$ is conditionally convergent, then, for any $q \in \mathbb{R}$, there is a rearrangement of $\sum a_n$ that converges to the sum q.*

[In *(b)*, you can also find a rearrangement that will diverge in any desired fashion; see 4.2*(d)*. But I shan't talk about that, although my proof will more or less make the point.]

Proof. A. Suppose $\sum |a_n|$ is convergent. By 4.13, $\sum a_n$ is convergent; call its sum A. Let $\sum a_{\sigma(n)}$ be a rearrangement of the series, and let $\epsilon > 0$.

There exists M such that $n \geq M \Longrightarrow |\sum_{k=1}^{n} a_k - A| < \frac{1}{2}\epsilon$, and there exists L such that $m \geq n \geq L \Longrightarrow \sum_{k=n}^{m} |a_k| < \frac{1}{2}\epsilon$ (as $\sum |a_n|$ converges; cf. (16)). Let $K \coloneqq \max(L, M)$. The indices $1, 2, \ldots, K$ appear in the rearranged series as indices $\sigma(r_1), \sigma(r_2), \ldots, \sigma(r_K)$; that is, $r_i \coloneqq \sigma^{-1}(i)$ for $1 \leq i \leq K$. Let

$$N \coloneqq \max\{r_1, r_2, \ldots, r_K\} .$$

If $n \geq N$, then $\sum_{k=1}^{n} a_{\sigma(k)}$ is the sum, in some order, of all the terms $a_1, a_2, \ldots, a_K$, together with some other terms (probably not consecutive) a_r, where the indices r, all greater than K, form a finite set $T(n)$ (depending on n). If R is the least element of $T(n)$ and S the greatest, then by the triangle inequality

$$\left|\sum\nolimits_{r\in T(n)} a_r\right| \leq \sum\nolimits_{r\in T(n)} |a_r| \leq \sum\nolimits_{k=R}^{S} |a_k| < \tfrac{1}{2}\epsilon ,$$

since $S \geq R > K \geq L$. It follows that, when $n \geq N$,

$$\begin{aligned} \left|\sum\nolimits_{k=1}^{n} a_{\sigma(k)} - A\right| &= \left|\sum\nolimits_{k=1}^{K} a_k - A + \sum\nolimits_{r\in T(n)} a_r\right| \\ &\leq \left|\sum\nolimits_{k=1}^{K} a_k - A\right| + \left|\sum\nolimits_{r\in T(n)} a_r\right| < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon . \end{aligned}$$

This proves *(a)*.

B. For *(b)*, let $a_n^+ := \max(a_n, 0)$, $a_n^- := -\min(a_n, 0) \geq 0$, for each n. Then $a_n = a_n^+ - a_n^-$ (one or other of a_n^+, a_n^- is 0), $|a_n| = a_n^+ + a_n^-$. I show that both the series of nonnegative terms $\sum a_n^+$, $\sum a_n^-$ must diverge.

If $\sum a_n^-$ converges, then $\sum a_n^+ = \sum (a_n + a_n^-)$ also converges (by 4.4, as $\sum a_n$ converges by hypothesis); similarly, if $\sum a_n^+$ converges, so does $\sum a_n^-$. But if *both* converge, so does $\sum |a_n| = \sum (a_n^+ + a_n^-)$, and this is contrary to hypothesis. So neither can converge.

On the other hand, $a_n \to 0$ as $n \to \infty$, since $\sum a_n$ converges. (Recall 4.17.)

For clarity I shall describe the construction of a rearrangement whose sum is q in ordinary language, to avoid complicated notation. Let me call an index k "nonnegative" if $a_k \geq 0$, and "negative" if $a_k < 0$. Proceed as follows to choose $\sigma(1), \sigma(2), \ldots$ inductively.

Let $\sigma(1)$ be the first nonnegative index, and choose $\sigma(2), \sigma(3) \ldots$ as successive nonnegative indices, UNTIL the sum of the corresponding terms exceeds q (this may of course already be true at $\sigma(1)$, but we have seen it must happen sooner or later, as $\sum a_n^+$ diverges). Then take the next term of the rearranged series to be the first negative index, and take a run of successive negative indices until the partial sum falls below q (which must happen, as $\sum a_n^-$ diverges); after which, take the least nonnegative index not already used, and successive nonnegative indices until the sum again exceeds q; and so on. It is left as an exercise that the result is a rearrangement of $\sum a_n$ with partial sums tending to q. □

Note 4.24. *(a)* In 4.23**B**, one could take a run of nonnegative indices until the sum is more than 1, then a negative index, then a run of nonnegative indices until the sum exceeds 2, and so on; then the rearrangement that is obtained "diverges to ∞". And so on.

(b) The property that every rearrangement of a series converges is sometimes called "unconditional convergence" of the series. 4.23*(a)* shows that absolute convergence of a real or complex series implies unconditional convergence; the proof works without much alteration in $\mathbb{R}^n$ or in $\mathbb{C}^n$, and indeed in any Banach space.

Conversely, 4.23**A** above shows that a series in $\mathbb{R}$ that is not absolutely convergent is not unconditionally convergent. The proof depends on the order structure of $\mathbb{R}$; but the statement that an unconditionally convergent series must be absolutely convergent will follow for series in $\mathbb{R}^n$ or $\mathbb{C}^n$ by considering individual coordinates. However, the statement is false in any infinite-dimensional Banach space (Dvoretsky-Rogers, 1950), and is quite trivially false in any infinite-dimensional Hilbert space.

§5. Continuity.

I shall write $\mathbb{K}$ and $\mathbb{L}$ to mean "either $\mathbb{R}$ or $\mathbb{C}$ as the case may be". Many of the same ideas apply to metric spaces in general: see §24.

Definition 5.1. Let X be a subset of $\mathbb{K}$, and $f : X \longrightarrow \mathbb{L}$. f is said to be *continuous at* $x \in X$ (or, equivalently, x is a point of continuity of f) if, for any $\epsilon > 0$, there exists $\delta > 0$ such that, for all $y \in X$ for which $|x - y| < \delta$, $|f(x) - f(y)| < \epsilon$: symbolically,

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall y \in X) \quad |y - x| < \delta \Longrightarrow |f(y) - f(x)| < \epsilon\,. \tag{18}$$

Let me stress here that X is the domain of f.

There is an analogy with the definition 2.9 (formula (8)) of convergence of a sequence. We can make the analogy a little clearer.

Definition 5.2. Suppose $f : X \longrightarrow \mathbb{L}$ as in 5.1, and let a be a limit point of X and $b \in \mathbb{L}$. Then one says "$f(x)$ has limit b at a" or "$f(x)$ tends to b as x tends to a", and one writes $f(x) \to b$ as $x \to a$ or $\lim_{x\to a} f(x) = b$, when, for any $\epsilon > 0$, there is a $\delta > 0$ such that $|f(x) - b| < \epsilon$ for any $x \in X \setminus \{a\}$ such that $|x - a| < \delta$.

Of course, if $a \notin X$, the condition that $x \in X \setminus \{a\}$ reduces to $x \in X$. But, in all cases, the condition could be rephrased

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in X) \quad 0 < |x - a| < \delta \Longrightarrow |f(x) - b| < \epsilon .$$

If $a \in X$, the value of f at a, $f(a)$ itself, does not appear in this definition.

Lemma 5.3. *$f : X \longrightarrow \mathbb{L}$ is continuous at $a \in X$ if and only if $f(x)$ has limit $f(a)$ at a, that is, if and only if $f(x) \to f(a)$ as $x \to a$.* □

It does not make sense to speak of continuity of f at a unless $f(a)$ is already defined; that is why I require $a \in X$. I recall here a general definition from MATH 161:

Definition 5.4. Suppose that $f : X \longrightarrow \mathbb{L}$ as above, and that $Y \subseteq X$. The restriction of f to Y, $f|Y$, is the function $Y \longrightarrow \mathbb{L}$ whose value at each point $y \in Y$ is $f(y)$; that is, $f|Y$ has the same value as f at each point where it is defined, but its domain (the set of points where it is defined) has been cut down to Y.

Formally, a function $f : X \longrightarrow \mathbb{L}$ is a subset of $X \times \mathbb{L}$ satisfying certain properties, and

$$f|Y := f \cap (Y \times \mathbb{L})$$

(where $Y \times \mathbb{L}$ is a subset of $X \times \mathbb{L}$). X is the set of "first coordinates" of f, and Y is the set of "first coordinates" of $f|Y$.

Lemma 5.5. *Let $f : X \longrightarrow \mathbb{L}$ as in* 5.1. *Then f is continuous at $x \in X$ if and only if, for any sequence (x_n) in X that converges to x, the sequence $f(x_n)$ in $\mathbb{L}$ converges to $f(x)$.*

Proof. The condition is *necessary*. Suppose that f is continuous at x and that $x_n \to x$. Let $\epsilon > 0$. There exists $\delta > 0$ such that, for any $y \in X$,

$$|y - x| < \delta \Longrightarrow |f(y) - f(x)| < \epsilon .$$

As $x_n \to x$, there is some N such that $n \geq N \Longrightarrow |x_n - x| < \delta$. Hence

$$n \geq N \Longrightarrow |f(x_n) - f(x)| < \epsilon .$$

We have found an N which "works" for the chosen ϵ in (8).

The condition is *sufficient*. We prove this by contrapositive. Suppose, in fact, that f is *not* continuous at x. Then

$$(\exists \epsilon > 0)(\forall \delta > 0)(\exists y \in X) \quad |y - x| < \delta \ \ \& \ \ |f(y) - f(x)| \geq \epsilon$$

(simply by the rules for negation, applied to (18)). So, if we take $\delta := 1$, there is $y_1 \in X$ such that $|y_1 - x| < 1$ and $|f(y) - f(x)| \geq \epsilon$; if we take $\delta := 1/n$, there is $y_n \in X$ with $|y_n - x| < 1/n$ and $|f(y_n) - f(x)| \geq \epsilon$. But now, the sequence (y_n) tends to x, even though $f(y_n)$ does not tend to $f(x)$. So the condition is not fulfilled. □

Remark 5.6. The condition of 5.5 is *sequential continuity* at x. It is equivalent to continuity in metric spaces, *but not in general topological spaces* (for which see 24.6). We shall use it here to avoid the more general definition.

Those of you who have done or are doing MATH 309 will recognize that in 5.5 and in 3.12 I again casually assumed the "countable axiom of choice"; to define the sequence (y_n) we make infinitely many *choices*—there was no rule telling us how to do it. Having pointed this out, I shall proceed without further comment to make the same assumption many times[2].

There are a lot of elementary facts about functions continuous or sequentially continuous at a point that follow simply from the definitions and may be left to tutorials or assignments.

Definition 5.7. Let $f : X \longrightarrow \mathbb{L}$ as above. We say that f is *continuous* [or sequentially continuous], or sometimes [sequentially] *continuous on* X, if it is continuous [sequentially continuous] at each point of X. Equivalently, we might say x is a point of [sequential] continuity of f.

Lemma 5.8. *Suppose that B is a closed subset of $\mathbb{L}$ and that $f : X \longrightarrow \mathbb{L}$ is* [*sequentially*] *continuous. Then $f^{-1}(B) \coloneqq \{x \in X : f(x) \in B\}$ is closed in X.*

Proof. Let (x_n) be a sequence in $f^{-1}(B)$ that converges to a point $x \in X$. Then $f(x_n)$ converges to $f(x)$ in $\mathbb{L}$. As $f(x_n) \in B$ for each n and B is closed in $\mathbb{L}$, $f(x) \in B$; which means that $x \in f^{-1}(B)$. This proves the result. □

Theorem 5.9. *Suppose $f : X \longrightarrow \mathbb{L}$ is sequentially continuous, and A is a nonempty sequentially compact subset of X. Then $f(A) \coloneqq \{f(x) : x \in A\}$ is a sequentially compact subset of $\mathbb{L}$.*

Proof. Let (y_n) be a sequence in $f(A)$. For each n, there is an $x_n \in A$ such that $f(x_n) = y_n$. By hypothesis, there is a subsequence $(x_{n(k)})_{k=1}^{\infty}$ which converges as $k \to \infty$ to a point $x \in A$. But then, as f is sequentially continuous, $f(x_{n(k)}) \to f(x) \in f(A)$; that is, $y_{n(k)} \to f(x)$ as $k \to \infty$. □

In the case where $\mathbb{L} = \mathbb{R}$, we deduce (cf. 3.12) that $f(A)$ is a closed and bounded subset of $\mathbb{R}$, and contains both its supremum and its infimum: thus, in particular,

Corollary 5.10. *If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, $\{f(t) : a \le t \le b\}$ is a bounded set in $\mathbb{R}$, and there are points $\xi, \eta \in [a, b]$ such that*

$$f(\xi) = \sup\{f(t) : a \le t \le b\} \quad \textit{and} \quad f(\eta) = \inf\{f(t) : a \le t \le b\}. \qquad \square$$

So $f(\xi) \ge f(t) \ge f(\eta)$ for all $t \in [a, b]$; ξ is a point in $[a, b]$ where the value of f is the greatest value it takes on $[a, b]$, i.e. the "maximum" value, and $f(\eta)$ is the "minimum" value. The traditional formulation of this property is that "a continuous function on a closed bounded

[2] I use the Axiom of Choice unashamedly even if it is unnecessary. Littlewood, I am told, always emphasized in lectures how to avoid it when possible, by giving some explicit procedure for the choices; at the time this seemed very important, because Gödel had not yet proved the Axiom was consistent with standard set theory. But nowadays this doubt no longer exists, and proofs without the Axiom, even when they are possible, always involve some careful argument that is immaterial to the main idea.

interval is bounded and attains its bounds". Here "bounds" meant "its least upper bound and its greatest lower bound", and each is a value of f at some point of $[a,b]$. (There may be many points ξ where the value is maximal and many points η where the value is minimal; but, of course, a "ξ" cannot be an "η" unless f is constant, when all points of $[a,b]$ give the same value of f.)

In elementary courses, and historically, this was treated as obvious, without much mention of the conditions of continuity and compactness. It was only seen quite late that something subtle is involved; the paper usually mentioned is by Heine (1871). Some very big names, Dirichlet and Riemann amongst them, fell into error as a result, for the property is indeed not at all automatic.

Suppose $a = 0$ and $b = 1$. The function f defined by

$$f(0) = \tfrac{1}{2} = f(1)\,, \qquad f(t) = t \quad \text{for } 0 < t < 1\,,$$

is bounded, as $0 < f(t) < 1$ for all $t \in [0,1]$, but it has no maximum or minimum value (its supremum is 1 and its infimum is 0, neither of which is a *value*). The explanation, of course, is that it is discontinuous at 0 and at 1. The function g defined by

$$g(\tfrac{1}{2}) = 0\,, \quad g(t) = \frac{1}{\frac{1}{2} - t} \quad \text{for } 0 \le t < \tfrac{1}{2} \text{ and } \tfrac{1}{2} < t \le 1\,,$$

is unbounded; and it is discontinuous at $\frac{1}{2}$. In short, for discontinuous functions, even if the domain considered is a closed bounded interval, the result isn't true.

Similarly, it isn't true if the domain is noncompact, even if the function is continuous. The interval $(0,1]$ is not compact, and $h(t) := 1/t$ is continuous but unbounded on this interval; whilst the function $k(t) := t$ is continuous and bounded, but does not attain its infimum 0.

The other familiar property of continuous real-valued functions of a real variable is the "intermediate value theorem". It, too, has many generalizations, but they go in a rather different direction from ideas of compactness, as you would expect from the proof.

Theorem 5.11. (The intermediate value theorem) *Let $f : [a,b] \longrightarrow \mathbb{R}$ be continuous, where $a < b$. If $\alpha := f(a) < \beta := f(b)$ and $\xi \in (\alpha, \beta)$, then there exists some $x \in (a,b)$ such that $f(x) = \xi$.*

In words, "any number strictly between $f(a)$ and $f(b)$ is the value of f at some point strictly between a and b". This statement of the "intermediate value *property*" allows $f(a) > f(b)$; in that case, apply the theorem as given above to $-f$ instead of f to deduce the conclusion.

Proof. Let $J := \{t \in [a,b] : f(t) \le \xi\}$. Then $a \in J$, $b \notin J$, and J is certainly bounded above by b. Hence J has a supremum, which we call x, and $a \le x \le b$ evidently. I assert that $f(x) = \xi$. I shall prove this directly from the definitions.

If $f(x) \ne \xi$, take $\epsilon := |f(x) - \xi| > 0$. There exists $\delta > 0$ such that, if $t \in [a,b]$ and $|t - x| < \delta$, $|f(t) - f(x)| < \epsilon$.

If $f(x) < \xi$ and $x \ne b$ (in particular if $x = a$), then, for $|t - x| < \delta$, $t \in [a,b]$,

$$f(t) \le f(x) + |f(t) - f(x)| < f(x) + \epsilon = \xi\,,$$

so that $t \in J$. Thus, if for instance $t = x + \frac{1}{2}\min(\delta, b - x)$, $t \in J$ and yet $t > x$; which is absurd (as $x := \sup J$). Hence, $x \ne a$, and either $f(x) \ge \xi$ or $x = b$.

If $f(x) > \xi$ (so $x > a$), then, for any $t \in [a,b]$ with $|t-x| < \delta$,

$$f(t) \geq f(x) - |f(t) - f(x)| > f(x) - \epsilon = \xi\,.$$

This is absurd too, implying there is no point of J in $(x - \min(\delta, x-a), x]$, and contradicting 2.8*(ii)*. We conclude that, in fact, $f(x) = \xi$. □

It is tempting to suppose, on the basis of the vague intuition that continuity of a function should mean the absence of breaks in its graph, that the intermediate value property characterizes continuous real-valued functions of a real variable—but it doesn't. There are (ingenious, but quite elementary) examples of functions with the same property that are wildly discontinuous.

§6. "Uniform".

The words "uniform" and "uniformly" occur very frequently. It is, however, difficult to give a satisfactory general sense to them (Whittaker and Watson tried, on p. 52, without, I suspect, helping anyone very much). In obscure and imprecise language, one might say they are used when some property involves a choice for every element of some set (of functions or of points), and the choice can in the case considered be made *for all elements of the set simultaneously*. The simplest example is (perhaps) uniform continuity.

Recall 5.7 (or 24.4). To say that $f : X \longrightarrow \mathbb{L}$ is continuous at each point of $K \subseteq X$ is to assert that, for each $x \in K$ and for each $\epsilon > 0$, there exists some $\delta > 0$ satisfying (18). Symbolically,

$$\begin{aligned}(\forall \epsilon > 0)(\forall x \in K)(\exists \delta > 0)(\forall y \in X)\\ (|y-x| < \delta \Longrightarrow |f(y) - f(x)| < \epsilon)\,.\end{aligned} \tag{19}$$

In statements, like this, where several quantifiers occur, the order of adjacent universal quantifiers is immaterial (as is the order of adjacent existential quantifiers), but an existential quantifier cannot be swapped with an adjacent universal quantifier without changing the meaning. "For every bottle of mass-produced beer on the shelves today, there is a bottling machine from which it came" is presumably true; "there is a bottling machine from which every bottle on the shelves today came" is false. Similarly (19) asserts that, for any x, there is a δ that "works" for that x and the given ϵ; in principle, δ may have to change if x changes. We say, by contrast, that f is *uniformly* continuous on K if there is a δ that "works" for the given ϵ and for *all* $x \in K$ (and, in that sense, is "independent of x", at least for $x \in K$). This amounts to swapping quantifiers in (19).

Definition 6.1. Let $K \subseteq X \subseteq \mathbb{K}$. A function $f : X \longrightarrow \mathbb{L}$ is *uniformly continuous* on K if

$$\begin{aligned}(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in K)\\ ((y \in X \ \ \&\ \ |y-x| < \delta) \Longrightarrow |f(y) - f(x)| < \epsilon)\,.\end{aligned} \tag{20}$$

There is a "sequential" version of this definition: if (x_n) is a sequence in K, (y_n) is a sequence in X, and $|x_n - y_n| \to 0$, then $|f(x_n) - f(y_n)| \to 0$. I leave it as an exercise to show this is equivalent to (20). (Compare 5.5.)

In many books the definition is (in effect) stated only for $K = X$, by restricting y to lie in K. But such a statement, which puts x and y on exactly the same footing, loses a little of the force of the Theorem following.

Two other points. "Uniform" continuity says that δ may be freed from dependence on $x \in K$. If you think about it for a moment, you will see that it cannot usefully be freed from dependence on ϵ. And it is a condition involving the distance between points x and y where x may be *anywhere* in K; as such it is not really a "topological" condition—you need to be able to compare the distance between x and y to the distance between x' and y', which a topology will not do.

Example 6.2. Consider functions $\mathbb{R} \longrightarrow \mathbb{R}$.

(a) The identity function $f(x) = x$ is uniformly continuous on the whole of $\mathbb{R}$, because, for any $\epsilon > 0$, we can take $\delta = \epsilon$ to ensure

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon\,.$$

(b) On the other hand, if $f(x) = x^2$ for all x, then f is *not* uniformly continuous on $\mathbb{R}$. However small $\delta > 0$ may be, if we take $y = x + \frac{1}{2}\delta$, then $|x - y| < \delta$, but

$$|f(x) - f(y)| = |(x + \tfrac{1}{2}\delta)^2 - x^2| = \delta|x + \tfrac{1}{4}\delta|$$

which we can make as large as we like by taking x to be large enough (for example, it is more than 1 if $x > \delta^{-1}$ or if $x < -\delta^{-1} - \frac{1}{4}\delta$).

(c) A rather less straightforward case is the function $f(x) = x^{1/3}$. We may argue as follows: given $\epsilon > 0$, suppose that $|x - y| < \frac{1}{8}\epsilon^3$. Then

$$\left|x^{1/3} - y^{1/3}\right|\left|x^{2/3} + x^{1/3}y^{1/3} + y^{2/3}\right| < \tfrac{1}{8}\epsilon^3\,, \tag{21}$$

so *either* $\left|x^{1/3} - y^{1/3}\right| < \epsilon$ (as we wish), *or* $\left|x^{2/3} + x^{1/3}y^{1/3} + y^{2/3}\right| < \frac{1}{8}\epsilon^2$. Consider the second case. There is an inequality

$$x^{2/3} + x^{1/3}y^{1/3} + y^{2/3} \geq \tfrac{1}{2}(x^{2/3} + y^{2/3})\,;$$

for $\left|x^{1/3}y^{1/3}\right| \leq \frac{1}{2}(x^{2/3} + y^{2/3})$, the inequality between arithmetic and geometric means, or, more basically, because $\frac{1}{2}(x^{1/3} + y^{1/3})^2 \geq 0$. We deduce that $x^{2/3} + y^{2/3} < \frac{1}{4}\epsilon^2$, and, therefore, that $x^{2/3} < \frac{1}{4}\epsilon^2$ and $y^{2/3} < \frac{1}{4}\epsilon^2$, or $\left|x^{1/3}\right| < \frac{1}{2}\epsilon > \left|y^{1/3}\right|$. But this entails that $\left|x^{1/3} - y^{1/3}\right| \leq \left|x^{1/3}\right| + \left|y^{1/3}\right| < \epsilon$ anyway. Hence, taking $\delta := \frac{1}{8}\epsilon^3$ always suffices.

Consequently, this function f is also uniformly continuous on $\mathbb{R}$, although we had to do some calculations "in the margin" (taking δ to be $\frac{1}{8}\epsilon^3$ was not mere luck; I had to see what number on the right-hand side of (21) would make the argument work) to find a suitable candidate for δ. [This is common in writings about analysis—a seemingly unmotivated choice of some number, here δ, results from considering what is necessary to make the argument that follows valid, but the author does not explain the reasons for his choice beforehand because the argument itself is the explanation.]

In *(a)* and *(b)* the slope of the graph already suggests uniform continuity or its absence, whereas, for $x^{1/3}$, the graph has a vertical tangent at the origin. More complicated examples may not have tangents to the graph in any meaningful sense at all, but still be uniformly continuous. For instance:

Theorem 6.3. *If K is a sequentially compact subset of X, and every point of K is a point of continuity of $f : X \longrightarrow \mathbb{K}$, then f is uniformly continuous on K.*

Proof. Suppose not. Then (taking the negation of (20))

$$(\exists \epsilon > 0)(\forall \delta > 0)(\exists x \in K)(\exists y \in X) \\ |y - x| < \delta \ \ \& \ \ |f(y) - f(x)| \geq \epsilon \,. \tag{22}$$

If we take $\delta := 1/n$, then there are points $x_n \in K$, $y_n \in X$, such that $|x_n - y_n| < 1/n$ and $|f(y_n) - f(x_n)| \geq \epsilon$. K being sequentially compact, there is a subsequence $(x_{n(i)})_{i=1}^{\infty}$ of (x_n) that converges to a point $x_0 \in K$. It follows that

$$|y_{n(i)} - x_0| \leq |x_{n(i)} - y_{n(i)}| + |x_{n(i)} - x_0| < 1/n(i) + |x_{n(i)} - x_0| \to 0$$

as $i \to \infty$; that is, $y_{n(i)} \to x_0$ too. (Recall 2.12.) But x_0 is a point of continuity of f, by hypothesis; so $f(x_{n(i)}) \to f(x_0)$ and $f(y_{n(i)}) \to f(x_0)$ as $i \to \infty$. As a consequence,

$$|f(x_{n(i)}) - f(y_{n(i)})| \to 0 \quad \text{as } i \to \infty \,.$$

But this contradicts the hypothesis that $|f(x_n) - f(y_n)| \geq \epsilon$ for all n. So (22) must be false; and this establishes the result. □

This theorem does not, of course, find a δ for a given ϵ; it is not "constructive". An Intuitionist would claim that such a general proof (by contradiction!) of the existence of a δ is inadmissible, and that one should give a method to *find* a suitable δ.

Another context where the word "uniform" appears is for sequences of functions.

Definition 6.4. Suppose that $K \subseteq X \subseteq \mathbb{K}$, that $f : X \longrightarrow \mathbb{L}$, and that, for each $n \in \mathbb{N}$, $f_n : X \longrightarrow \mathbb{L}$.

(a) We say $f_n \to f$ *pointwise on* K (the sequence (f_n) tends or converges pointwise on K to f) if, for each $x \in X$, the sequence $(f_n(x))$ in $\mathbb{L}$ converges to $f(x)$. [I apologize for the convenient but ugly word "pointwise"; one could say "point by point".]

(b) We say $f_n \to f$ *uniformly on* X if, for any $\epsilon > 0$, there is some $N \in \mathbb{N}$ such that, whenever $n \geq N$, $|f_n(x) - f(x)| < \epsilon$ for any $x \in X$.

Remark 6.5. It is traditional to regard the distinction between the preceding definitions as "difficult". If we put them symbolically, we get

$$\textit{(a)} \quad (\forall x \in K)(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \quad n \geq N \Longrightarrow |f_n(x) - f(x)| < \epsilon \,, \\ \textit{(b)} \quad (\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall x \in K) \quad n \geq N \Longrightarrow |f_n(x) - f(x)| < \epsilon \,.$$

Interchanging two quantifiers of the same kind makes no difference to the sense, so *(a)* could equivalently be written

$$(\forall \epsilon > 0)(\forall x \in K)(\exists N \in \mathbb{N}) \quad n \geq N \Longrightarrow |f_n(x) - f(x)| < \epsilon \,.$$

(We could read it as "for all x and ϵ, ...".) Thus, the essential difference between *(a)* and *(b)* is the swapping of existential and universal quantifiers, as before. What is asserted in *(b)* is the existence, granted ϵ, of an N which "works" simultaneously for all $x \in \Omega$; it "depends" only on ϵ, and might be written $N(\epsilon)$ as a reminder. In *(a)*, there may in principle be, for given ϵ, a different N for each x. One might call it $N(x, \epsilon)$.

Example 6.6. *(a)* Let $K = X = \mathbb{R}$. Define functions $h_n, h : \mathbb{R} \longrightarrow \mathbb{R}$:

$$h_n(x) := 0 \quad \text{for } x \leq n, \quad h_n(x) := 1 \quad \text{for } x > n, \quad h(x) := 0 \quad \text{for all } x.$$

If $\epsilon > 1$, then $|h_n(x) - h(x)| < \epsilon$ for all n and all x. If $0 < \epsilon \le 1$, then clearly $|h_n(x) - h(x)| = |h_n(x)| < \epsilon$ only when $h_n(x) = 0$, that is, when $n \ge x$. So $h_n \to h$ pointwise (for any specific x, we can take $N > x$); but definitely not uniformly. (A graph may clarify the matter.)

(b) A more interesting example is this. Let $K := [0,1]$, which is sequentially compact. Define $g_n(x) := n^2 x \exp\left(-\frac{1}{2}n^2x^2\right)$ and $g(x) := 0$ for all x. Assuming all sorts of things from differential calculus that we have not yet begun to justify, we see that $g_n(0) = 0$, that $g_n(x)$ increases to a maximum value of $ne^{-1/2}$ at $x = \frac{1}{n}$, and that it decreases thereafter (the graph shows a "sliding hump" which squeezes more and more closely to the y-axis as n increases, but also becomes higher and higher). $g_n \to g$ pointwise as $n \to \infty$, because for $x = 0$ the value is always 0, and for $0 < x \le 1$ the hump of g_n will pass to the left of x as n increases (indeed, we know that, when $x > 0$, $n^2 x e^{-n^2x^2/2} \to 0$ as $n \to \infty$). But the convergence is not uniform; indeed,

$$\sup\{|g_n(x) - g(x)| : 0 \le x \le 1\} = \sup\{|g_n(x)| : 0 \le x \le 1\} = ne^{-1/2},$$

which tends to ∞ rather than to 0 as $n \to \infty$.

Theorem 6.7. (Dini's theorem on monotone convergence.) *Let K be a sequentially compact subset of $\mathbb{K}$. Suppose that $f_n : K \longrightarrow \mathbb{R}$ (for each $n \in \mathbb{N}$) and $f : K \longrightarrow \mathbb{R}$ are functions continuous on K, and that $f_n \to f$ pointwise monotonically on Ω (i.e., for each individual $x \in K$, $(f_n(x))$ is a monotone sequence in $\mathbb{R}$). Then $f_n \to f$* uniformly *on K.*

Briefly, a pointwise monotonic sequence of *continuous* functions on a compact set having a *continuous* limit must converge uniformly. The sequence (g_n) in 6.6*(b)* is not pointwise monotonic (for small positive x, $h_n(x)$ increases up to a certain value of n, and then falls for all greater n).

Proof. A. Consider $f_n - f$ instead of f_n; this reduces the question to the special case where f is identically 0. As (f_n) tends pointwise monotonically to 0, $|f_n|$ *decreases* monotonically to 0; if $|f_n|$ tends uniformly to 0, then f_n tends uniformly to 0, from the definition 6.4*(b)*. So all we need to prove is that, if $f_n \downarrow 0$ pointwise, then $f_n \downarrow 0$ uniformly.

B. Suppose not. Then (negating 6.5*(b)*, with $f(x) = 0$ and everything nonnegative)

$$(\exists \epsilon > 0)(\forall N \in \mathbb{N})(\exists x \in K)(\exists n \ge N) \quad f_n(x) \ge \epsilon.$$

Take $N := 1$, and then there are $n(1) \ge 1$ and $x_1 \in K$ with $f_{n(1)}(x_1) \ge \epsilon$. If $n(k), x_k$ have been chosen, take $N := n(k) + 1$; then there are some $n(k+1) \ge n(k) + 1$ and $x_{k+1} \in K$ with $f_{n(k+1)}(x_{k+1}) \ge \epsilon$. We obtain inductively a sequence (x_k) in K such that $f_{n(k)}(x_k) \ge \epsilon$ for each k. For any $l \ge k$, $n(l) \ge n(k)$, so that

$$f_{n(k)}(x_l) \ge f_{n(l)}(x_l) \ge \epsilon \quad \text{for all } l \ge k. \tag{23}$$

However, K is sequentially compact, so there is a subsequence $(x_{k(m)})_{m=1}^{\infty}$ which converges to a point $x \in K$. By (23), $f_{n(k(m))}(x_{k(p)}) \ge \epsilon$ whenever $p \ge m$ (so that $k(p) \ge k(m)$). But $f_{n(k(m))}$ is continuous at x, so $f_{n(k(m))}(x_{k(p)}) \to f_{n(k(m))}(x)$ as $p \to \infty$, and consequently $f_{n(k(m))}(x) \ge \epsilon$. This must hold for all m—so we have found a subsequence of (f_n) for which the values at x do *not* tend to 0.

This contradiction of the hypothesis establishes the theorem. □

My use of sequences here (quite possibly following Dini himself) is not the clearest proof. Nevertheless, the method of "extracting subsequences" is often useful because it seems so "natural" and "elementary", even though it may not be ideal in a theoretical sense. (See 24.7.)

The reason why uniform convergence is rather important is that it is a relatively accessible condition ensuring that you can interchange the order of two limiting processes without changing the result. This was first pointed out, in a rather absurd terminology, by Stokes in 1847; the great mathematicians of that era (and earlier) made no distinction between pure and applied mathematics—Stokes is these days generally regarded as an applied mathematician (Navier-Stokes equation and all that). An important example is the following.

Theorem 6.8. *Let $X \subseteq \mathbb{K}$, and let f_n, for each $n \in \mathbb{N}$, and f be functions $X \longrightarrow \mathbb{L}$. If $f_n \to f$ uniformly on X and each f_n is continuous at $x \in X$, then f is continuous at x.*

Proof. Take $\epsilon > 0$. By uniform convergence, there exists N such that

$$(\forall y \in X) \quad n \geq N \Longrightarrow |f_n(y) - f(y)| < \tfrac{1}{3}\epsilon\,. \tag{24}$$

But, in particular, f_N is continuous at x, and so

$$(\exists \delta > 0) \quad y \in X \;\;\&\;\; |y - x| < \delta \Longrightarrow |f_N(x) - f_N(y)| < \tfrac{1}{3}\epsilon\,. \tag{25}$$

Putting these two facts together, one finds that, if $y \in X$ and $|y - x| < \delta$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f_N(x) - f(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &< \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon = \epsilon\,. \end{aligned}$$

This shows that f is continuous at x. □

This proof should be carefully studied. The reason for imposing the condition of uniform convergence is that, to find a δ for the given ϵ, we have to refer to some *specific* N, but we then also require $|f_N(y) - f(y)|$ to be small without having any correspondingly specific information about y. Uniform convergence is a relatively simple condition that arranges we can do this. (In fact, a rather weaker condition, "subuniform" convergence, is sufficient, since, in the proof, y is not *completely* arbitrary; it has to be within the distance δ of x.)

There is an alternative condition that would imply the limit of a pointwise convergent sequence of functions continuous at x is continuous at x. Instead of being able, given ϵ, to choose N independently of x, we could prove continuity of the limit function f if we could choose δ for the given x independently of n. The principle of the argument is identical, but the details obviously differ somewhat. This alternative condition is called *equicontinuity* of the sequence (f_n) at x. (There is also a concept of *uniform equicontinuity* on a set K, sometimes amusingly known as *equiuniform continuity*.)

As an interesting application of uniform convergence, let me cheat and assume we have already some knowledge of integration.

Theorem 6.9. *Suppose $a, b \in \mathbb{R}$ and $a \leq b$, and (f_n) is a sequence of continuous functions $[a, b] \longrightarrow \mathbb{R}$, where $f_n \to f$ uniformly on $[a, b]$. Then f is continuous and*

$$\int_a^b f_n(t)\,dt \to \int_a^b f(t)\,dt\,. \tag{26}$$

Proof. Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$(\forall t \in [a,b]) \quad n \geq N \Longrightarrow |f_n(t) - f(t)| \leq \frac{\epsilon}{b-a+1},$$

and, therefore, when $n \geq N$,

$$\left| \int_a^b f_n(t)\,dt - \int_a^b f(t)\,dt \right| \leq \int_a^b |f_n(t) - f(t)|\,dt \leq \frac{(b-a)\epsilon}{b-a+1} < \epsilon. \qquad (*)$$

The definition of the convergence (26) is satisfied. □

(The inequalities at (*) obviously depend on properties of the integral. The argument may not work for some kinds of integral and some more general functions or domains.)

The integral results from some sort of limiting process, so this is another theorem about interchange of two such processes: $\lim(\int f_n) = \int(\lim f_n)$. It is not true for general non-uniform limits. For example, in 6.6*(b)*, the pointwise limit is 0: we see that

$$\int_0^1 g_n(t)\,dt = 1 - e^{-n^2/2} \to 1 \text{ as } n \to \infty, \quad \int_0^1 \lim g_n(t)\,dt = 0.$$

In pictorial terms, the pointwise limit is 0 because, for any given $x > 0$, the hump in the graph is eventually to the left of x. But the *height* of the hump grows at such a rate that the area beneath it does not diminish (and indeed increases towards 1).

There are several theorems about taking limits under the integral sign; this is the most elementary, and applies to most reasonable definitions of the integral, provided the functions considered are suitably well-behaved.

§7. Differentiability.

In this section we confine our attention exclusively to a real-valued function f defined on an interval $[a,b]$ in $\mathbb{R}$, where $a < b$.

Definition 7.1. Let $c \in [a,b)$. f is said to be *differentiable on the right*, or *right-differentiable*, at c if the *difference quotient*

$$q^r_{f,c} : (0, b-c] \longrightarrow \mathbb{R} : h \mapsto \frac{f(c+h) - f(c)}{h}$$

has a limit in $\mathbb{R}$ as $h \downarrow 0$. (In 5.2, take $X := [0, b-c]$, $a := 0$. The notation $\downarrow$ seems appropriate.) The limit is called the *right-derivative of f at c*, $f'_r(c)$.

Let $c \in (a,b]$. f is *differentiable on the left*, or *left-differentiable*, at c if the difference quotient

$$q^l_{f,c} : [a-d, 0) \longrightarrow \mathbb{R} : h \mapsto \frac{f(c+h) - f(c)}{h}$$

has a limit in $\mathbb{R}$ as $h \uparrow 0$. The limit is the *left-derivative of f at d*, $f'_l(c)$.

If $x \in (a,b)$, we say that f is *differentiable* at x if it is both left- and right-differentiable at x and $f'_r(x) = f'_l(x)$. This is equivalent to saying that the difference quotient

$$q_{f,c} : [a-x, b-x] \setminus \{0\} \longrightarrow \mathbb{R}$$

has a limit as the variable "h" tends to 0 ("from either side", "through positive and negative values"). The value of this limit is the *derivative of f at x*, $f'(x)$, and we sometimes say "$f'(x)$ exists" as a shorter alternative to "f is differentiable at x with derivative $f'(x)$".

f is described as *differentiable on* (a,b) if it is differentiable at each point of (a,b).

It is conventional to say f is differentiable at a if it is right-differentiable there, and differentiable at b if it is left-differentiable there; with this convention, one may speak of differentiability on $[a,b]$, on $[a,b)$, and on $(a,b]$. The notation $f'(a)$ may then be used for $f'_r(a)$, there being no two-sided derivative available, and likewise $f'(b)$ may be written for $f'_l(b)$. f is described as C^1 on $[a,b]$ if it is differentiable on $[a,b]$ and $f'(t)$, as just defined for all $t \in [a,b]$, is continuous on $[a,b]$.

Functions given by simple algebraic formulæ are always differentiable, and there are familiar methods of finding their derivatives by the "rules for differentiation", which result from the algebraic properties of limits. (You should formulate and prove these rules for yourself.) But, in the abstract, differentiability *even at a single point* is an amazingly strong condition—it is not satisfied by "most" continuous functions, for instance. It involves only the "infinitesimal behaviour of f" near the point. One might not expect the *global* behaviour of f to relate to derivatives in any simple way.

Lemma 7.2. *If $f : [a,b] \longrightarrow \mathbb{R}$ is differentiable at $c \in [a,b]$, it is continuous at c.*

Proof. Take $c \in (a,b)$; minor changes in the argument are needed if c is an end-point. As $f'(c)$ exists, there is $\delta_1 > 0$ such that $(c-\delta_1, c+\delta_1) \subseteq (a,b)$ and

$$0 < |h| < \delta_1 \Longrightarrow \left| \frac{f(c+h)-f(c)}{h} - f'(c) \right| < 1\,. \tag{27}$$

Given $\epsilon > 0$, take $\delta := \min\left(\delta_1, \dfrac{\epsilon}{1+|f'(c)|}\right)$. Then, if $0 < |h| < \delta$,

$$\left| \frac{f(c+h)-f(c)}{h} \right| \leq 1 + |f'(c)| \quad \text{by (27), as } \delta \leq \delta_1\,,$$

But then $|f(c+h)-f(c)| \leq (1+|f'(c)|)|h| < (1+|f'(c)|)\delta \leq \epsilon$. Hence, f satisfies the definition of continuity at c. □

Theorem 7.3. (Rolle's Theorem) *Suppose that f is continuous on $[a,b]$ and differentiable on (a,b), and that $f(a) = f(b)$. Then there is a point $c \in (a,b)$ such that $f'(c) = 0$.*

Michel Rolle did not prove "his" theorem; that was impossible at the time (his dates are 1652–1719). His statement was for polynomial functions only. (I don't know his actual argument, but it is easy to construct one that appeals only to the Intermediate Value Theorem and to the elementary algebra of polynomials.) The theorem is of course "obvious" if you "sketch a graph". Unfortunately, not all graphs, even of decent functions, are "sketchable".

Proof. 5.10 shows f is bounded and attains its bounds on $[a,b]$. Then,

$$\begin{aligned} \sup\{f(t) : t \in [a,b]\} &\geq f(a) = f(b) \\ &\geq \inf\{f(t) : a \leq t \leq b\}\,. \end{aligned} \tag{28}$$

If both inequalities are in fact equalities, f is *constant* on $[a,b]$ and c may be taken to be any

point of (a,b). Suppose, then, that the first inequality is strict. By 5.10, there exists some $c \in [a,b]$ such that

$$f(c) = \sup\{f(t) : t \in [a,b]\} > f(a) = f(b);$$

necessarily, then, $a \neq c \neq b$, and $c \in (a,b)$. Our hypotheses tell us that $f'(c)$ exists. Now, if $h \in (0, b-c]$, $f(c+h) \leq f(c)$ (as $f(c)$ is the supremum), h is positive, and

$$q^r_{f,c}(h) = \frac{f(c+h) - f(c)}{h} \leq 0. \tag{29}$$

Taking the limit as $h \downarrow 0$, we see $f'_r(c) \leq 0$. But, for $h \in [a-c, 0)$, h is negative, and

$$q^l_{f,c}(h) = \frac{f(c+h) - f(c)}{h} \geq 0; \tag{30}$$

as $h \uparrow 0$ we find $f'_l(c) \geq 0$. Since, however, $f'(c) = f'_r(c) = f'_l(c)$, the only possible value is $f'(c) = 0$.

If the second inequality in (28) is strict, we take c to be the point at which the infimum is attained; inequalities (29) and (30) are reversed, but the conclusion remains. □

There may be *many* possible points c. (If *both* inequalities (28) are strict, there are at least two.) The theorem only asserts there is *at least* one that is strictly between a and b. Of course we also have a corollary (of the proof) that everyone knows and that is often applied:

Corollary 7.4. *Assume the hypotheses of the theorem. Any point of (a,b) at which the value of f is either $\sup f([a,b])$ or $\inf f([a,b])$ must also be a point at which the derivative is 0.*

Theorem 7.5. (The mean-value theorem) *Suppose that f is continuous on $[a,b]$ and is differentiable on (a,b), where $a < b$. Then there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = (b-a)f'(c).$$

The name refers to the fact that $f'(c)$ is the "average rate of increase" of f on $[a,b]$. With suitable physical assumptions, it says that your *average* speed over the journey must be your *instantaneous* speed at some moment or other.

Proof. Let $\lambda := \dfrac{f(b) - f(a)}{b-a}$. Then consider $g(t) := f(t) - \lambda(t-a)$:

$$g(a) = f(a), \quad g(b) = f(a),$$

and g is also continuous on $[a,b]$ and differentiable on (a,b). By Rolle's theorem, there is some $c \in (a,b)$ such that $g'(c) = f'(c) - \lambda = 0$, or $f'(c) = \lambda$. □

This Theorem enables us, rather surprisingly, to deduce results about the 'large-scale' behaviour of f from information about its derivative; for instance,

Lemma 7.6. *Suppose g is continuous on $[a,b]$ and differentiable on (a,b), and $g'(t) \neq 0$ for all $t \in (a,b)$. Then, whenever $x, y \in [a,b]$ and $x \neq y$, $g(x) \neq g(y)$.* □

There is, by the way, no reason at all to hope that g' is *continuous* on (a,b). On the other hand, there is also a theorem, which I omit, that it must have the intermediate-value property, so that, here, it cannot be negative at some points of (a,b) and positive at others.

There is a cunning extension of the mean-value theorem due to Cauchy.

Theorem 7.7. (Cauchy's mean-value theorem.) *Suppose that f and g are both continuous on $[a,b]$ and differentiable on (a,b), and also that $g'(t) \neq 0$ for all $t \in (a,b)$. Then there is a point $c \in (a,b)$ such that*

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

Proof. By 7.6, $g(b) \neq g(a)$. So I may define $\lambda := \dfrac{f(b)-f(a)}{g(b)-g(a)}$, and a function $h(t) := f(t) - \lambda g(t)$ on $[a,b]$. Then h is continuous on $[a,b]$, differentiable on (a,b), and

$$h(a) = \frac{f(a)g(b)-f(b)g(a)}{g(b)-g(a)} = h(b). \tag{31}$$

Hence, by Rolle's theorem, there exists some $c \in (a,b)$ such that

$$f'(c) - \lambda g'(c) = h'(c) = 0,$$

which is exactly the result. □

This result *implies* the ordinary MVT—just take $g(t) := t$. However, the implication in the opposite direction is false, since 7.5 tells us only that the numerator is $(b-a)f'(c_1)$ for some $c_1 \in (a,b)$ and the denominator is $(b-a)g'(c_2)$ for some $c_2 \in (a,b)$, where c_1, c_2 need not be the same.

The algebraic manipulations (31) are often presented in terms of determinants. Set

$$k(t) := \begin{vmatrix} f(t) & f(a) & f(b) \\ g(t) & g(a) & g(b) \\ 1 & 1 & 1 \end{vmatrix};$$

then the properties of determinants ensure $k(a) = k(b) = 0$, so that by Rolle's theorem there is some $c \in (a,b)$ for which

$$k'(c) = \begin{vmatrix} f'(c) & f(a) & f(b) \\ g'(c) & g(a) & g(b) \\ 0 & 1 & 1 \end{vmatrix} = 0,$$

and, expanding by the bottom row, we deduce the result again.

Lemma 7.8. (l'Hôpital's rule, first version.) *Let* $f, g : (a, b) \longrightarrow \mathbb{R}$ *be functions differentiable at* $c \in (a, b)$, *and* $g'(c) \neq 0 = f(c) = g(c)$. *There exists* $\epsilon > 0$ *such that* $(c - \epsilon, c + \epsilon) \subseteq (a, b)$, $g(t) \neq 0$ *for* $0 < |t - c| < \epsilon$, *and*

$$\frac{f(t)}{g(t)} \to \frac{f'(c)}{g'(c)}$$

as $t \to c$ *through values such that* $0 < |t - c| < \epsilon$.

Proof. As $\dfrac{g(c + h) - g(c)}{h} \to g'(c) \neq 0$ as $h \to 0$ and $g(c) = 0$, there is $\epsilon > 0$ such that, for $0 < |h| < \epsilon$, $c + h \in (a, b)$ and $\dfrac{g(c + h)}{h} \neq 0$. Then $g(c + h) \neq 0$ too. Thus, if $0 < |t - c| < \epsilon$, $t \in (a, b)$ and $g(t) \neq 0$:

$$\frac{f(t)}{g(t)} = \frac{f(t) - f(c)}{g(t) - g(c)} = \left(\frac{f(t) - f(c)}{t - c}\right) \Bigg/ \left(\frac{g(t) - g(c)}{t - c}\right)$$
$$\to \frac{f'(c)}{g'(c)} \quad \text{as } t \to c.$$ □

This argument depends only on the definition of the derivatives at c, and on basic facts about limits (which, to be fair, I have not proved in these lectures).

Lemma 7.9. (l'Hôpital's rule, second version.) *Suppose* $f, g : (a, b) \longrightarrow \mathbb{R}$, $c \in (a, b)$, *and* $f(c) = g(c) = 0$; *let* f *and* g *be continuous on* (a, b), *and differentiable at all points of* $(a, b) \setminus \{c\}$, *with* $g'(t) \neq 0$ *for all* $t \neq c$. *If* $f'(t)/g'(t) \to l$ *as* $t \to c$, *then* $f(t)/g(t) \to l$ *as* $t \to c$.

Proof. Take $\epsilon > 0$. By hypothesis, there exists $\delta > 0$ such that

$$0 < |\tau - c| < \delta \Longrightarrow \left|\frac{f'(\tau)}{g'(\tau)} - l\right| < \epsilon.$$

Suppose now that $0 < |t - c| < \delta$. Cauchy's MVT applies:

$$\frac{f(t)}{g(t)} = \frac{f(t) - f(c)}{g(t) - g(c)} = \frac{f'(\tau)}{g'(\tau)}$$

for some τ between t and c. However, τ must also satisfy $0 < |\tau - c| < \delta$, and so

$$\left|\frac{f(t)}{g(t)} - l\right| = \left|\frac{f'(\tau)}{g'(\tau)} - l\right| < \epsilon.$$

So, for given ϵ, we have found a δ, as required by the definition. □

Although the hypotheses of this version of the rule are much more demanding (it requires differentiability not at c itself, but everywhere else), it has the advantage that it can sometimes (that is, if $f'(c), g'(c)$ exist and are 0 and so on) be repeated by passing to the second

derivative. For instance, $\dfrac{\sin x - x}{x^3}$ takes the form "$\dfrac{0}{0}$" when $x = 0$; we may seek a limit as $x \to 0$ by differentiating top and bottom: $\dfrac{\cos x - 1}{3x^2}$. But again, we get "$\dfrac{0}{0}$" when $x = 0$, and to find the limit we can try differentiating again: $\dfrac{-\sin x}{6x}$; and again: $\dfrac{-\cos x}{6}$. Here at last there is a limit as $x \to 0$, viz. $-\frac{1}{6}$, and by the Theorem this is also the limit of the previous expression and of the original expression.

The implication is again only one-way. The limit of $f(t)/g(t)$ may exist without the derivatives' existing at all, or without their quotient's having a limit. As an example, let $f(t) := t \sin(1/t)$ for $t \neq 0$, and $g(t) := \sqrt[3]{t}$. Since $|f(t)| \leq |t|$ for all $t \neq 0$,

$$\frac{t \sin(\frac{1}{t})}{\sqrt[3]{t}} \to 0 \quad \text{as } t \to 0.$$

We can extend f by setting $f(0) := 0$, and it then becomes a continuous function. But the quotient of the derivatives is $\dfrac{-\frac{1}{t}\cos(\frac{1}{t}) + \sin(\frac{1}{t})}{\frac{1}{3}t^{-2/3}}$, which has no limit as $t \to 0$.

This is an instance of a general thought. Because the derivative describes the "infinitesimal" character of a function f, it may have behaviour that is very irregular by comparison with f itself. Equivalently, *integrating a function smooths out bumps*.

There are several other versions of l'Hôpital's rule, when either the variable t or the numerator and denominator of $f(t)/g(t)$ tend to ∞ or to $-\infty$, or both. They are proved by similar applications of the Cauchy MVT.

We now come to Taylor's theorem. I commented in my introduction that theorems before about 1800 were rarely stated with the precision we now expect. What Taylor did, I suppose, was to present an expansion, without discussing what it meant or when it was true:

$$f(a+h) = f(a) + \frac{hf'(a)}{1!} + \frac{h^2 f''(a)}{2!} + \cdots. \tag{32}$$

As soon as you look at this statement with a critical eye, it collapses. For a start, f need not be differentiable or even continuous at a (for instance, $f(t) := \sqrt[3]{t}$ is not differentiable at 0; $g(t) := t^{n+\frac{1}{3}}$ is n times differentiable at 0, but not $n+1$ times). If f is differentiable infinitely often at all points, being what is called "C^∞", the infinite series on the right may not converge for any non-zero h. (One may construct a C^∞ function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f^{(n)}(a) = (n!)^2$ for all n; then the Taylor series (32) cannot converge if $h \neq 0$.) And—the final humiliation—even if the derivatives all exist at all points and the series always converges, its sum may not be $f(a+h)$.

There is a standard example. Let $\psi(t) := \exp(-t^{-2})$ for $t \neq 0$, and $\psi(0) := 0$. This is sometimes called the Cauchy function. It is C^∞ on $\mathbb{R}$, i.e. its derivatives of all orders exist at all points of $\mathbb{R}$; they may be calculated at points other than the origin by the chain and product rules, and then one may show by induction that all the derivatives at the origin exist and are zero. So ψ has an infinite Taylor series about 0 all coefficients of which are 0, and which therefore converges (with sum 0) for all values of h. Yet ψ obviously has non-zero value at every point except the origin.

The general moral is that the Taylor expansion (32) cannot be taken seriously without a clear statement of its conditions of validity. There are, in fact, something like five commonly

stated versions of "Taylor's theorem", which differ in the precise assumptions about the function f and in the conclusions they draw. Firstly the version that is the most "natural" and "familiar":

Theorem 7.10. (Lagrange's version of Taylor's theorem.) *Given* $h > 0$, *let the function* $f : [a, a+h] \longrightarrow \mathbb{R}$ *be* $n-1$ *times differentiable at every point of* $[a, a+h]$, *where* $n \geq 1$, *and suppose that* $f^{(n-1)} : [a, a+h] \longrightarrow \mathbb{R}$ *is continuous on* $[a, a+h]$ *and differentiable on* $(a, a+h)$. *Define the "*n*th remainder in Taylor's formula"* $R_n(f,h)$ *by*

$$R_n(f,h) \coloneqq f(a+h) - f(a) - \frac{hf'(a)}{1!} - \frac{h^2 f''(a)}{2!} - \cdots - \frac{h^{n-1} f^{(n-1)}(a)}{(n-1)!}\,.$$

Then there is a point $c \in (a, a+h)$ *such that*

$$R_n(f,h) = \frac{h^n f^{(n)}(c)}{n!}\,.$$

This is often expressed otherwise, by writing $c = a + \theta h$ for the appropriate value $\theta \in (0,1)$. This gives a statement that applies also when $h < 0$, namely: there exists $\theta \in (0,1)$ such that $R_n(f,h) = \dfrac{h^n f^{(n)}(a+\theta h)}{n!}$. (I leave the formulation both of the hypotheses and of the proof in this case to you.) The case $n = 1$ above is just the MVT.

Proof. Define $\phi_n(t) \coloneqq R_n(f,t)$ and $\psi_n(t) \coloneqq \dfrac{h^n}{n!}$ for $0 \leq t \leq h$. Then $\phi_n(0) = \psi_n(0) = 0$, $\phi_n'(0) = \psi_n'(0) = 0$, …, $\phi_n^{(n-1)}(0) = \psi_n^{(n-1)}(0) = 0$, and these derivatives are continuous on $[0,h]$. Cauchy's MVT applies repeatedly:

$$\frac{\phi_n(h)}{\psi_n(h)} = \frac{\phi_n(h) - \phi_n(0)}{\psi_n(h) - \psi_n(0)} = \frac{\phi_n'(h_1)}{\psi_n'(h_1)} \quad \text{for some } h_1 \in (0,h),$$
$$\frac{\phi_n'(h_1)}{\psi_n(h_1)} = \frac{\phi_n'(h_1) - \phi_n'(0)}{\psi_n'(h_1) - \psi_n'(0)} = \frac{\phi_n''(h_2)}{\psi_n''(h_2)} \quad \text{for some } h_2 \in (0,h_1),$$

and so on (notice the successive derivatives of ψ_n are nonzero except at the origin!), up to

$$\frac{\phi_n(h)}{\psi_n(h)} = \frac{\phi_n^{(n-1)}(h_{n-1})}{\psi_n^{(n-1)}(h_{n-1})} = \frac{\phi_n^{(n-1)}(h_{n-1}) - \phi_n^{(n-1)}(0)}{\psi_n^{(n-1)}(h_{n-1}) - \psi_n^{(n-1)}(0)} = \frac{\phi_n^{(n)}(h_n)}{\psi_n^{(n)}(h_n)}$$

for some $h_n \in (0, h_{n-1})$. Take $c \coloneqq h_n$, and then

$$\frac{\phi_n(h)}{\psi_n(h)} = \frac{\phi_n^{(n)}(c)}{\psi_n^{(n)}(c)} = \frac{f^{(n)}(c)}{1}\,,$$

which says exactly that $R_n(f,h) = \dfrac{h^n f^{(n)}(c)}{n!}$. □

Theorem 7.11. (Young's version of Taylor's theorem.) *If* $h > 0$, *let the function* $f : [a, a+h] \longrightarrow \mathbb{R}$ *be* $n-1$ *times differentiable on* $[a, a+h]$, *and let* $f^{(n-1)}$ *be differentiable on the right at* a. *Then*

$$\frac{R_{n+1}(f,h)}{h^n} := \frac{R_n(f,h) - \frac{1}{n!}h^n f^{(n)}(a)}{h^n} \to 0 \tag{33}$$

as $h \downarrow 0$.

Lagrange's version was a generalization of the MVT to higher orders of differentiation. Young's version assumes the nth derivative (on the right, in my statement) exists at a, but not necessarily anywhere else in $[a, a+h]$, and is a generalization of the definition of the (right) derivative at a. (It even suggests a way of describing "nth derivatives" without assuming the previous ones exist.) The condition (33) says that $R_{n+1}(f,h)$, which is the error in taking $R_n(f,h)$ to be $\dfrac{h^n f^{(n)}(a)}{n!}$, is negligible by comparison with h^n when h is small (it is "$o(h^n)$"). There is of course a version for negative h with much the same proof.

Proof. Argue as before, but this time for $\dfrac{\phi_{n+1}(t)}{\psi_n(t)}$. We obtain $h_n \in (0,h)$ such that

$$\begin{aligned}\frac{\phi_{n+1}(h)}{\psi_n(h)} &= \frac{\phi_{n+1}^{(n-1)}(h_n)}{\psi_n^{(n-1)}(h_n)} \quad \text{which by easy calculation} \\ &= \frac{f^{(n-1)}(a+h_n) - f^{(n-1)}(a)}{h_n} - f^{(n)}(a)\,.\end{aligned}$$

But, as $h \downarrow 0$, necessarily $h_n \to 0$ (as $0 < h_n < h$), and, by the definition of the derivative,

$$\frac{f^{(n-1)}(a+h_n) - f^{(n-1)}(a)}{h_n} - f^{(n)}(a) \to 0\,.$$

Hence, $\dfrac{R_{n+1}(f,h)}{h^n} = \dfrac{\phi_{n+1}(h)}{n!\psi_n(h)} \to 0$ as $h \downarrow 0$. □

The proof above could be slightly shortened by quoting 7.9.

Theorem 7.12. (Cauchy's version of Taylor's theorem, or a generalization thereof.) *With the same hypotheses and notation as in* 7.10, *let* $p > 0$. *Then there exists some* $\theta \in (0,1)$ *such that*

$$R_n(f,h) = \frac{h^n(1-\theta)^{n-p}}{(n-1)!\,p} f^{(n)}(a+\theta h)\,.$$

There is, of course, a similar statement for intervals $[a-h, a]$. The result holds for positive exponents p that are not integers, although we have not yet defined such powers.

Proof. Set $b := a+h$, $A := (b-a)^{-p}R_n(f,h)$. Define for $t \in [a,b]$

$$F(t) := f(t) + \frac{(b-t)f'(t)}{1!} + \cdots + \frac{(b-t)^{n-1}f^{(n-1)}(t)}{(n-1)!} + A(b-t)^p\,.$$

F is continuous on $[a,b]$, since $p>0$. It is also differentiable on (a,b), since $f^{(n-1)}$ is. Furthermore, $F(b)=f(b)$ as all the later terms vanish, $F(a)=f(b)$ because A is defined to make this true. By Rolle's theorem, then, there exists $c\in(a,b)$ such that $F'(c)=0$. Calculating $F'(c)$ and writing c as $a+\theta h$, we find

$$\frac{(b-c)^{n-1}f^{(n)}(c)}{(n-1)!}+Ap(b-c)^{p-1}=0\,,\quad\text{and so}$$

$$R_n(f,h)=(b-a)^pA=\frac{h^p(b-c)^{n-p}f^{(n)}(c)}{(n-1)!\,p}=\frac{h^n(1-\theta)^{n-p}}{(n-1)!\,p}f^{(n)}(a+\theta h)\,.$$

This is the result. □

If we take $p=n$, we recover Lagrange's version. Taking $p=1$, we get Cauchy's form of the remainder, $\dfrac{h^n(1-\theta)^{n-1}f^{(n)}(a+\theta h)}{(n-1)!}$. At first sight this seems very strange, but it has some interesting applications, for instance to the binomial series for fractional indices.

If we assume basic properties of the integral, we can get yet another version, from which Cauchy's may be deduced.

Theorem 7.13. *Again with the hypotheses and notation of* 7.10*,*

$$R_n(f,h)=\frac{(b-a)^n}{(n-1)!}\int_0^1(1-\tau)^{n-1}f^{(n)}(a+\tau(b-a))\,d\tau\,.$$

Proof. Define, for $a\le t\le b$, a function

$$G(t):=f(t)+\frac{(b-t)f'(t)}{1!}+\cdots+\frac{(b-t)^{n-1}f^{(n-1)}(t)}{(n-1)!}\,.$$

Here $G(b)-G(a)=R_n(f;h)$, $G'(t)=\dfrac{(b-t)^{n-1}f^{(n)}(t)}{(n-1)!}$ for $a<t<b$ as in previous calculations; hence,

$$R_n(f;h)=G(b)-G(a)=\int_a^bG'(t)\,dt=\int_a^b\frac{(b-t)^{n-1}f^{(n)}(t)}{(n-1)!}\,dt\,.\qquad(*)$$

Now substitute $t=a+\tau(b-a)$ for $0\le\tau\le1$:

$$R_n(f;h)=\frac{(b-a)^n}{(n-1)!}\int_0^1(1-\tau)^{n-1}f^{(n)}(a+\tau(b-a))\,d\tau\,.$$ □

Note 7.14. *(a)* I have stated and proved the above forms of Taylor's Theorem on the assumption that $h>0$ for brevity's sake. There are versions for $h<0$; I leave the requisite changes of statements and proofs to you.

(b) In the proof of 7.13, I have cheated in several ways, principally because we have not yet studied integration. The hypotheses of 7.10 implied only that the integrand $G'(t)$ of (*) is *defined* on $(0,1)$, although $G(t)$ is to be continuous on $[0,1]$. On that basis, the integral at (*) is probably in some sense "improper" because the integrand $G'(t)$ may not even be defined at the end-points. This integrand is also a derivative on $(0,1)$, with no other

obvious properties; in what sense can it be integrated, and satisfy the "fundamental theorem of calculus" that I have appealed to?

If we use the Riemann integral and *add the substantial extra hypothesis* that $f^{(n)}$ is continuous (not just defined) on $(a, a+h)$, the integral at (*) may be interpreted as an improper Riemann integral (that is, as the limit of $G(\beta) - G(\alpha) = \int_\alpha^\beta G'(t)\,dt$ as $\alpha \downarrow a$ and $\beta \uparrow b$). Since G itself is continuous on $[a, b]$, the improper integral is equal to $G(b) - G(a)$. This, however, requires *stronger* hypotheses for 7.13 than for 7.10.

Dieudonné in the first volume of his 'Traité d'Analyse', the one he published in English while in Chicago (*Foundations of Modern Analysis*, Academic Press 1960), spoke sneeringly on p. 142 of the Riemann integral: "a mildly interesting exercise in the general theory of measure and integration". (This is a little unfair, since for numerical analysis you have to use the Riemann integral in some form.) He avoided it for his purposes by introducing an integral (the "Cauchy integral") for functions that are derivatives of a continuous function (the "primitive") except at countably many points; he called these "regulated functions". He had to prove that, for instance, continuous functions are regulated, without using the Riemann integral—which is not difficult, once you see it done, but a little unexpected. The "primitive" for him was the (indefinite) integral, but he offered no means, however theoretical, of finding it. I doubt whether he was aware of the Kurzweil-Henstock integral (invented in 1957 by Kurzweil, and developed independently and extensively by Henstock after 1961), which starts from a construction of a (definite) integral that is a subtle variant of Riemann's and coincides with the Cauchy integral for regulated functions. If the Henstock integral is used (or the Cauchy integral), 7.13 is valid on the hypothesis—*weaker* than the hypothesis of 7.10—that $f^{(n-1)}$ is continuous on $[a, a+h]$ and differentiable except possibly at countably many points thereof. See 8.17 below.

(c) The general point in these four results, 7.10–7.13, is that any rigorous statement on the lines of "Taylor's theorem" must be concerned with the *error* in approximating a function by some combination of terms in its Taylor series. There may be many different estimates of this error in different circumstances. The fifth standard version is the Taylor theorem in complex analysis; it is much more pleasing, but Taylor had nothing to do with it. The reason it looks much better is, of course, that its hypotheses are vastly more restrictive.

§8. Integration.

In this section, as in the last, I shall only consider functions defined on a closed bounded interval $[a, b]$ in $\mathbb{R}$, with values in $\mathbb{R}$ or in $\mathbb{C}$. This is the most fundamental case of integration theory, and all that is seriously needed for the rest of the course.

The theory of integration is extremely extensive. Many important theorems are really quite difficult to prove even for the Riemann integral of a function of a single variable, despite being treated as obvious in elementary courses. The more radical problem with the Riemann integral is that there are many functions (for instance unbounded functions, or functions which are very discontinuous) which it cannot integrate, even when they arise by some simple process from continuous functions and we should expect a reasonable notion of "area under the curve". This difficulty is commonly overcome for first-year students by introducing extensions of the original Riemann definition of the integral, such as improper integrals of various kinds—but they are all *ad hoc* improvisations, methods for *calculating* expressions that we believe ought to possess a numerical value, and they lack a general justification. At least one such technique

(the Daniell integral) can be shown to do most of what one might desire, but all such approaches have a certain whiff of trickery about them. The Henstock-Kurzweil integral, on the other hand, rethinks the Riemann definition *ab initio*, and as such is much more satisfactory. But, despite its relatively simple definition, its deeper properties are not by any means easy to establish.

The "Lebesgue integral" is the standard integral for most uses, partly because it makes mathematical sense of probability theory and partly because it enables us to integrate in abstract spaces rather than just in $\mathbb{R}^n$. But it needs quite a lot of preparation, and its relation to differentiation, and indeed to other integrals, is not transparent. (The Henstock and Lebesgue integrals have the same value in cases for which they both make sense.)

The mathematical theory of integration—Lebesgue or Henstock or Daniell integrals and dozens of others that have been invented from time to time—is mostly concerned with theoretical questions: what it means to integrate a function, which functions can be integrated, and what properties the integrals have. General algorithms for *evaluating* integrals have to assume the integrands have very special properties (for instance, they may rely on the calculation of the values of the integrand at certain specified points). In this sense, the Riemann integral remains fundamental, even though it can only be applied to a very restricted class of functions. The situation is not unlike that with real numbers; in practical calculations, we must always use *rational* numbers, but we need the "abstract" concept of "real number" for other reasons.

Definition 8.1. A *gauge* in $[a,b]$ is a function $\gamma : [a,b] \longrightarrow (0,\infty)$.

We require *nothing* of a gauge except that it be strictly positive at all points of $[a,b]$. (By the way, the word "gauge" has some other quite different mathematical senses.) It may, for instance, be very discontinuous indeed. At the other extreme, it may be constant, i.e. it may take the same positive value δ at all points of $[a,b]$. (In that case I shall give the gauge the same name δ as its constant positive value.) But we can instantly say that

Remark 8.2. If $\gamma_1, \gamma_2, \dots, \gamma_k$ are gauges in $[a,b]$, so is $\min(\gamma_1, \gamma_2, \dots, \gamma_k)$, defined "pointwise", i.e.

$$(\forall t \in [a,b]) \quad (\min(\gamma_1, \gamma_2, \dots, \gamma_k))(t) := \min(\gamma_1(t), \gamma_2(t), \dots, \gamma_k(t)) .$$

(For each t, the smallest of the positive numbers $\gamma_1(t), \gamma_2(t), \dots, \gamma_k(t)$ is positive.) We may also say of gauges γ, γ' that $\gamma \le \gamma'$ if $\gamma(t) \le \gamma'(t)$ for each $t \in [a,b]$.

Definition 8.3. *(a)* A *division* of $[a,b]$ is a finite subset $\{t_0, t_1, \dots, t_n\}$ of $[a,b]$ such that

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b .$$

The points $t_0, t_1, \dots, t_n$ are the *division points*, or just *points*, of the division.

The same idea may be formulated by describing a division as a finite set $\mathcal{J}$ of closed, non-degenerate, non-overlapping subintervals of $[a,b]$ whose union is the whole of $[a,b]$:

$$\mathcal{J} = \{[t_{k-1}, t_k] : 1 \le k \le n\}, \quad [a,b] = \bigcup_{k=1}^{n} [t_{k-1}, t_k] ,$$

(A "nondegenerate" interval is one that is neither empty nor a singleton; "nonoverlapping" means that the intersection of any two of them is either empty or a singleton.) The subintervals $[t_{k-1}, t_k]$ are then the "intervals of the subdivision".

The relation between the two formulations is obvious, and I shall use either. [In higher dimensions, things are not so simple, here or later.]

(b) A division $\mathcal{J}_1$ is a *subdivision* or *refinement* of the division $\mathcal{J}$ of $[a,b]$ if every division point of $\mathcal{J}$ is a division point of $\mathcal{J}_1$, or, equivalently, if every interval of the division $\mathcal{J}_1$ is included in an interval of $\mathcal{J}$. (The equivalence does take a moment's thought.)

(c) A *tagged division* $\mathcal{J}'$ of $[a,b]$ is a division $\mathcal{J}$ of $[a,b]$ together with the assignment to each interval $J \in \mathcal{J}$ of a specific point $\tau_J \in J$. τ_J is called the *tag* of J, and may in principle be anywhere in J. (Two adjacent intervals of $\mathcal{J}$ may, therefore, have the same tag, if it is a common endpoint—and only then.)

Thus, a tagged division is a set $\{([t_{k-1},t_k],\tau_k) : 1 \le k \le n\}$ of ordered pairs, where, for each k, $t_{k-1} \le \tau_k \le t_k$, and $a = t_0 < t_1 < \cdots < t_n = b$; or, equivalently, it is of the form $\{(J,\tau_J\} : J \in \mathcal{J}\}$, where $\mathcal{J}$ is a division of $[a,b]$ and $\tau_J \in J$ for each J.

(d) If γ is a gauge in $[a,b]$, a tagged division $\mathcal{J}'$ of $[a,b]$ is γ*-fine* if, for each $J \in \mathcal{J}$, $J \subseteq (\tau_J - \gamma(\tau_J), \tau_J + \gamma(\tau_J))$. In the formulation with "division points", this means that, for $1 \le k \le n$, $\tau_k - \gamma(\tau_k) < t_{k-1}$ and $t_k < \tau_k + \gamma(\tau_k)$.

Definition 8.4. Suppose that $f : [a,b] \longrightarrow \mathbb{R}$. For any tagged division $\mathcal{J}'$ of $[a,b]$, consisting of a division $\{t_0,t_1,t_2,\ldots,t_n\}$ with tags $\tau_k \in [t_{k-1},t_k]$ as above, there is a corresponding *Riemann sum* of f,

$$\sigma(f,\mathcal{J}') := \sum\nolimits_{k=1}^{n} (t_k - t_{k-1}) f(\tau_k)\,. \tag{34}$$

[This is exactly the kind of *sum* originally proposed by Riemann, but his definition of the integral involved taking a limit of the sums in a different sense, as we shall see.]

Definition 8.5. A number $M \in \mathbb{K}$ is said to be a *generalized Riemann integral* (or Henstock or Henstock-Kurzweil integral, etc.) of f on $[a,b]$, and we write $M = \int_{[a,b]} f$ (or various other familiar notations like $\int_{[a,b]} f(t)\,dt$ or $\int_a^b f(t)\,dt$) if, for every $\epsilon > 0$, there is a gauge γ such that, for any γ-fine tagged division $\mathcal{J}'$ of $[a,b]$,

$$|\sigma(f,\mathcal{J}') - M| < \epsilon\,. \tag{35}$$

If such a number M exists, we say that f is *integrable* (generalized Riemann-integrable, Henstock-integrable, Henstock-Kurzweil integrable, &c.) on $[a,b]$ (and say also that M is the *gauge integral*, etc., of f on $[a,b]$).

Remark 8.6. *(a)* It is clear that, in some sense, M is a to be a "limit" of Riemann sums of tagged divisions as the tagged divisions "get finer and finer". The difference with Riemann's original definition is that he restricted attention to *constant* gauges: f is Riemann-integrable on $[a,b]$ (with a Riemann integral M) if, for every $\epsilon > 0$, there is some $\delta > 0$ such that, for every δ-fine tagged division $\mathcal{J}'$ of $[a,b]$, $|\sigma(f,\mathcal{J}') - M| < \epsilon$. Evidently this requires much more than 8.5, since only constant gauges are allowed. For the Henstock integral, one is free to adjust the gauge for a given ϵ so that (35) is satisfied for all γ-fine tagged divisions. The integral is also a limit of Riemann sums, but the "limit" is taken in a relatively undemanding sense.

(b) Riemann's definition is often stated somewhat differently these days (using suprema and infima of f over the subintervals rather than tags), but that need not concern us.

(c) It is conceivable *a priori* that there may be some gauge γ in $[a,b]$ so strange that no γ-fine tagged division of $[a,b]$ can exist. In that case, (35) would not make sense at all for that γ, and the definition of the integral would be vacuously satisfied for any M.

Given any constant gauge δ, you can take all the subintervals to be of length less than δ (for instance by taking all of them of length exactly $(b-a)/n$ for $n > (b-a)/\delta$), and then any tagging of the division gives a δ-fine tagged division; but for a general gauge γ it is not immediately obvious that γ-fine tagged divisions exist.

The difficulty can be removed by an argument of Goursat (not very different in principle from 23.6) that appears in the complex analysis in the slightly altered form Goursat used for a quite different purpose. (See D2.1 of Appendix D.)

Lemma 8.7. *For any gauge γ on $[a,b]$, there is a γ-fine tagged division $\mathcal{J}'$ of $[a,b]$.*

Proof. Obviously γ defines a gauge (which I shall still call γ) on any closed subinterval of $[a,b]$, by restriction, and, if $a < c < b$ and there are γ-fine tagged divisions of both $[a,c]$ and $[c,b]$, they may be combined to give a γ-fine tagged division of $[a,b]$.

Suppose, then, that there is *no* γ-fine tagged division of $[a,b]$. Then, perforce, one at least of the subintervals $[a, \frac{1}{2}(a+b)]$ and $[\frac{1}{2}(a+b), b]$ admits no γ-fine tagged division. Continue in this way, at each stage splitting the interval in half. One obtains a sequence $([a_n, b_n])$ of subintervals, where $b_n - a_n = 2^{-n+1}(b-a)$, $a_1 = a$, $b_1 = b$, and $[a_n, b_n]$ is either the left- or the right-hand half of the preceding interval and does not have a γ-fine tagged division.

But $[a,b]$ is sequentially compact, by 3.11. Thus there is some subsequence $(a_{n(k)})$ of (a_n) that converges to a point $a_* \in [a,b]$. In fact $a_* \in [a_m, b_m]$ for all m, since $a_{n(k)} \in [a_m, b_m]$ whenever $n(k) \geq m$; and it follows that $|a_m - a_*| \leq 2^{-m+1}(b-a)$, and similarly for b_m. Hence, both sequences (a_n) and (b_n) tend to a_*.

However, $\gamma(a_*) > 0$ by the definition of a gauge. Consequently, if n is so large that $2^{-n+1}(b-a) < \gamma(a_*)$,

$$[a_n, b_n] \subseteq (a_* - \gamma(a_*), a_* + \gamma(a_*)),$$

and so $[a_n, b_n]$ *does* have a tagged γ-fine division consisting of the one interval $[a_n, b_n]$ with a_* as its tag. This is a contradiction of our hypothesis, so the Lemma is established. □

The countable Axiom of Choice was used above, for it may have been necessary to *choose* a half-interval infinitely many times. I could, however, avoid all choices by insisting that the left-hand half of $[a_n, b_n]$ should always be taken if possible (that is, if it has no γ-fine tagged division) and the right-hand half only if it cannot be avoided. Then no "choices" are required. This is the sort of procedure Littlewood used in such cases.

Brouwer had a different objection to arguments like this. We proved the existence of something by contradiction from the assumption that it did not exist—the proof as we gave it was *nonconstructive*. No method was given or suggested to construct a γ-fine tagged division. It is not very important, in my view, whether this is a serious philosophical objection. (Brouwer was worried about Russell's paradox, in which the "set of all sets that are not members of themselves" is not in his opinion "constructive". However, there are less radical ways of avoiding the paradox.) The distinction between "constructive" and merely "existential" proofs is significant, nevertheless, for other reasons. A "constructive" proof suggests possible (though not always very practical) methods of computation, whereas a "nonconstructive" proof is merely airy-fairy. (Any existence proof that depends essentially on the Axiom of Choice must

obviously be nonconstructive[3].) It is, however, possible to formulate a "constructive" proof of 8.7, albeit by a rather different method (cf. 22.3).

At any rate, we now have a watertight definition both of Henstock-integrability and of the Henstock integral. The effort is not worthless:

Lemma 8.8. *If f is Riemann-integrable on $[a,b]$, with a Riemann integral M, it is also Henstock-integrable, with a Henstock integral M.* □

Lemma 8.9. *Let $f,g:[a,b]\longrightarrow\mathbb{K}$ be Henstock- [Riemann-]integrable on $[a,b]$, with integrals M,N. If $\lambda,\mu\in\mathbb{K}$, then $\lambda f+\mu g$ is also Henstock- [Riemann-]integrable on $[a,b]$, with integral $\lambda M+\mu N$.*

Proof. [Hint.] Given $\epsilon>0$, choose gauges γ_1 for f and γ_2 that satisfy (35) for g and for $\epsilon/(|\lambda|+|\mu|+1)$. (Cf. 4.3.) Then take $\gamma:=\min(\gamma_1,\gamma_2)$; a γ-fine division must be both γ_1-fine and γ_2-fine. □

We must also show that the integral has at most one value. Furthermore, the definition, like the definition of the limit of a sequence, has the defect that it appears to be necessary to know the value M of the integral before integrability can be proved. So we must also provide a Cauchy condition for Henstock-integrability (at least if f takes values in $\mathbb{R}$ or $\mathbb{C}$).

Lemma 8.10. *There can be at most one Henstock integral of f on $[a,b]$.*

Proof. Suppose M_1,M_2 are both H-integrals of f on $[a,b]$, and $M_1\neq M_2$. Take $\epsilon:=\frac{1}{2}|M_1-M_2|>0$, and then, by the definition 8.5, there are gauges γ_1,γ_2 such that

$$|\sigma(f,\mathcal{J}')-M_1|<\epsilon \quad \text{whenever the tagged division } \mathcal{J}' \text{ is } \gamma_1\text{-fine, and}$$
$$|\sigma(f,\mathcal{J}')-M_2|<\epsilon \quad \text{whenever the tagged division } \mathcal{J}' \text{ is } \gamma_2\text{-fine.}$$

Take $\gamma:=\min(\gamma_1,\gamma_2)$. If $\mathcal{J}'$ is γ-fine, it is *both* γ_1-fine and γ_2-fine; hence,

$$|M_1-M_2|\leq|\sigma(f,\mathcal{J}')-M_1|+|\sigma(f,\mathcal{J}')-M_2|<\epsilon+\epsilon=|M_1-M_2|\,, \tag{36}$$

which is absurd. We conclude that $M_1=M_2$. □

Lemma 8.11. *$f:[a,b]\longrightarrow\mathbb{K}$ is Henstock-integrable if and only if, for any $\epsilon>0$, there is some gauge γ such that, for any two γ-fine tagged divisions $\mathcal{J}',\mathcal{J}''$ of $[a,b]$, $|\sigma(f,\mathcal{J}')-\sigma(f,\mathcal{J}'')|<\epsilon$. It is Riemann-integrable if, for any $\epsilon>0$, there is a constant gauge with the same property.*

---

[3] There are at least two situations in analysis in which the Axiom of Choice was used to "construct" something that was then proved to be unique (so that the choices were irrelevant). In one case—the existence of Haar measure—an alternative construction without the Axiom was later found. In the other—construction of the Shilov boundary—as far as I know (I am out of touch) it remains a puzzle. In many "practical" instances one can see how to dispense with the Axiom, but the uniqueness proof is quite general.

Proof. [Sketch.] The condition is clearly necessary (cf. 2.23). Suppose it is satisfied. Then, taking $\epsilon := 2^{-n}$, we obtain a gauge γ'_n, and, taking $\gamma_n := \min(\gamma'_1, \dots, \gamma'_n)$, we may suppose, not only that γ_n "works" for 2^{-n}, but also that $\gamma_n \geq \gamma_{n+1}$ for each n. Let $\mathcal{J}'_n$ be a γ_n-fine tagged division of $[a,b]$, for each n; then $(\sigma(f, \mathcal{J}'_n))_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{K}$, so has a limit M. Now check (it is easy!) that 8.5 is satisfied. □

Lemma 8.12. *If $f : [a,b] \longrightarrow \mathbb{K}$ is continuous, it is Riemann-integrable on $[a,b]$.*

Proof. Take any $\epsilon > 0$. By 6.3, there is $\delta > 0$ such that, if $x, y \in [a,b]$ and $|x-y| < \delta$, then $|f(x) - f(y)| < \frac{1}{2}\epsilon/(b-a)$. Suppose that $\mathcal{J}'_1, \mathcal{J}'_2$ are δ-fine tagged divisions of $[a,b]$. Let $\mathcal{J}_3 := \mathcal{J}_1 \cup \mathcal{J}_2$ (the set of all division points for either division) and take any tags whatever for $\mathcal{J}_3$ to get a tagged division $\mathcal{J}'_3$.

I claim $|\sigma(f, \mathcal{J}'_i) - \sigma(f, \mathcal{J}'_3)| < \frac{1}{2}\epsilon$ for $i = 1, 2$. Indeed, if we take a subinterval $[t^{(1)}_{k-1}, t^{(1)}_k]$ of $\mathcal{J}'_1$ with tag $\tau^{(1)}_k$, it is divided into several subintervals $[t^{(3)}_{j-1}, t^{(3)}_j]$ of $\mathcal{J}'_3$, each with its tag $\tau^{(3)}_j$. As $\tau^{(3)}_j \in [t^{(1)}_{k-1}, t^{(1)}_k]$ for each j, certainly $\left|\tau^{(1)}_k - \tau^{(3)}_j\right| < \delta$ and $\left|f(\tau^{(1)}_k) - f(\tau^{(3)}_j)\right| < \frac{1}{2}\epsilon/(b-a)$. So

$$\begin{aligned}
&\left|(t^{(1)}_k - t^{(1)}_{k-1})f(\tau^{(1)}_k) - \sum\nolimits_j (t^{(3)}_j - t^{(3)}_{j-1})f(\tau^{(3)}_j)\right| \\
&\qquad = \left|\sum\nolimits_j (t^{(3)}_j - t^{(3)}_{j-1})(f(\tau^{(1)}_k) - f(\tau^{(3)}_j))\right| \\
&\qquad \leq \sum\nolimits_j (t^{(3)}_j - t^{(3)}_{j-1})\left|f(\tau^{(1)}_k) - f(\tau^{(3)}_j)\right| \\
&\qquad < \frac{\epsilon}{2(b-a)} \sum\nolimits_j (t^{(3)}_j - t^{(3)}_{j-1}) = \frac{\epsilon(t^{(1)}_k - t^{(1)}_{k-1})}{2(b-a)},
\end{aligned}$$

since the $t^{(3)}_j$s constitute a division of $[t^{(1)}_{k-1}, t^{(1)}_k]$. But, summing over k,

$$|\sigma(f, \mathcal{J}'_3) - \sigma(f, \mathcal{J}'_1)| < \sum\nolimits_k \frac{\epsilon(t^{(1)}_k - t^{(1)}_{k-1})}{2(b-a)} = \tfrac{1}{2}\epsilon\,.$$

The same estimate must apply to $\mathcal{J}'_2$: $|\sigma(f, \mathcal{J}'_3) - \sigma(f, \mathcal{J}'_2)| < \frac{1}{2}\epsilon$. Hence,

$$|\sigma(f, \mathcal{J}'_1) - \sigma(f, \mathcal{J}'_2)| \leq |\sigma(f, \mathcal{J}'_3) - \sigma(f, \mathcal{J}'_1)| + |\sigma(f, \mathcal{J}'_3) - \sigma(f, \mathcal{J}'_2)| < \epsilon\,.$$

As this is true for any two δ-fine tagged divisions of $[a,b]$, the condition of 8.11 holds. □

Remark 8.13. *(a)* 6.3 plays an essential rôle here, because the gauges considered must be constant. Relatedly, it is not difficult to see that f can only be Riemann-integrable on $[a,b]$ if it is bounded on $[a,b]$. However, some unbounded functions may be Henstock-integrable, if the gauges γ may be chosen to "counteract the growth of f". It is clumsy to give examples directly from the definition; see qu. 4 of Tut. Ex. 6 (2011), for instance.

(b) The example that is always given of a *bounded* function which is not Riemann-integrable (on any nondegenerate interval) is the so-called Dirichlet function: $f(x) = 1$ for x rational and $f(x) = 0$ for x irrational.

(c) I leave it as an exercise to show that, if $a \leq c < d \leq b$, then the function

$$f(x) \coloneqq \begin{cases} 0 & \text{when } x \notin [c,d]\,, \\ 1 & \text{when } x \in [c,d]\,, \end{cases}$$

is Henstock-integrable with integral $d - c$. The proof is very like that of

Lemma 8.14. *If $f, h : [a,b] \longrightarrow \mathbb{K}$ and $Z \coloneqq \{x \in [a,b] : f(x) \neq h(x)\}$ is countable*[4]*, then, if f is Henstock-integrable with integral M, the same is true of h.*

Proof. Because of 8.9, it suffices to show that a function $g : [a,b] \longrightarrow \mathbb{K}$ such that $Z \coloneqq \{x \in [a,b] : g(x) \neq 0\}$ is countable is Henstock-integrable with integral 0. (For then, take $g \coloneqq h - f$ and apply 8.9 to deduce the result.)

Enumerate the points of Z (without repeats) as $z_1, z_2, z_3, \ldots$. Given $\epsilon > 0$, define

$$\gamma(z_n) \coloneqq \frac{\epsilon}{2^{n+2}(1 + |g(z_n)|)} \tag{37}$$

for each point of Z; and, if $x \in [a,b] \setminus Z$, $\gamma(x)$ may be any positive number, for instance 1. Now, for any γ-fine tagged division $\mathcal{J}'$,

$$\sigma(g, \mathcal{J}') = \sum\nolimits_{k=1}^{n} (t_k - t_{k-1}) g(\tau_k) \tag{38}$$

as at (34). If $\tau_k \notin Z$, the kth term is 0 as $g(\tau_k) = 0$. If $\tau_k \in Z$, $\tau_k = z_{n(k)}$ for some $n(k) \in \mathbb{N}$, and then $t_k - t_{k-1} < 2\gamma(z_{n(k)})$, as $\mathcal{J}'$ is γ-fine, so that

$$|(t_k - t_{k-1}) g(\tau_k)| < \frac{\epsilon}{2^{n(k)+1}(1 + |g(z_{n(k)})|} |g(z_{n(k)})| < 2^{-n(k)-1}\epsilon\,.$$

However, $z_{n(k)}$ can be a tag for at most two subintervals of the division (if it is an endpoint of two abutting subintervals), so that the mapping $k \mapsto n(k)$ is at worst two-to-one and the sum (38) is definitely less than $2\sum_{n=1}^{\infty} 2^{-n-1}\epsilon = \epsilon$. This proves the result. □

The point is, of course, that it is the value of γ at the *tags* that determines the fineness of the tagged division; so we adjust those values appropriately. But the Lemma has the consequence that we can talk of the Henstock-integrability of a function that is defined on all but a countable number of points of $[a,b]$—for instance on (a,b); for we may define the function at the missing points in any way we like, and integrability and the integral will be the same for all such definitions.

Corollary 8.15. *The Dirichlet function is Henstock-integrable with integral* 0 *(on any interval $[a,b]$).* □

This alone should convince you that Henstock's theory is at least interesting, but there is a general warning. One of the properties of integrals that is often used, and that I have often called the "fundamental estimate", is the inequality

[4] It would be sufficient that Z be of measure zero. We do not have any need for this in the course, and it would require a very substantial digression.

$$\left| \int_a^b f(t)\,dt \right| \leq \int_a^b |f(t)|\,dt\,. \tag{39}$$

For Riemann sums, $|\sum_{k=1}^n (t_k - t_{k-1}) f(\tau_k)| \leq \sum_{k=1}^n (t_k - t_{k-1}) |f(\tau_k)|$ is obvious; consequently, if $f(t)$ is continuous, so is $|f(t)|$, both sides of (39) exist and are limits of Riemann sums, and the inequality holds. More generally, if $f(t)$ is Riemann-integrable on $[a,b]$, so is $|f(t)|$ (this is not so obvious), and (39) is again valid. The Riemann theory is an "absolute integration theory", as is the Lebesgue theory; that is, they do not allow "conditionally convergent integrals", which may be introduced later as ad hoc devices.

But the Henstock integral is not "absolute". We shall soon see that $\int_a^b f(t)\,dt$ may be defined as a Henstock integral in cases when $\int_a^b |f(t)|\,dt$ cannot be so defined (and there are also, theoretically, cases where the opposite happens). Some such cases would be regarded in the nineteenth century as "conditionally convergent improper integrals", but they will be quite respectable as Henstock integrals. Thus, the inequality (39) will be true for Henstock integrals only when *both* sides are defined as Henstock integrals.

Definition 8.16. Let $f : [a,b] \longrightarrow \mathbb{K}$. A function $g : [a,b] \longrightarrow \mathbb{K}$ is a *primitive* for f if

(a) g is continuous on $[a,b]$, and

(b) there is a *countable* subset Z of $[a,b]$ such that, if $x \in [a,b] \setminus Z$, g is differentiable at x (recall 7.1) and $g'(x) = f(x)$.

The word "primitive" does not always carry this sense, but see 7.14*(b)*.

I want to finish this brief description of Henstock's theory with a proof of the most startling of its elementary properties, which was already remarked on at 7.14*(b)*: a version of the "fundamental theorem of calculus" which says, amongst other things, that if a function is a derivative, it is Henstock-integrable, with the integral you expect. This is (if you are accustomed to other definitions of the integral, and until you see the proof) amazing, because there is no simple answer to the question *when* a function is a derivative; indeed, as far as I am aware, no really satisfactory answer is known; and it is absolutely certain, from elementary examples, that a derivative may not be Riemann- or even Lebesgue-integrable.

Theorem 8.17. *If $g : [a,b] \longrightarrow \mathbb{K}$ is a primitive for $f : [a,b] \longrightarrow \mathbb{K}$, then f is Henstock-integrable on $[a,b]$, and its integral is $g(b) - g(a)$.*

Proof. Let Z be a countable subset of $[a,b]$ such that $g'(x)$ exists and equals $f(x)$ whenever $x \in [a,b] \setminus Z$. Enumerate the points of Z without repeats as $z_1, z_2, \ldots$. Z may be finite, but I shall for the moment employ notation as if it were infinite.

Take any $\epsilon > 0$. I wish to define a gauge γ so that, for any γ-fine tagged division $\mathcal{J}'$ of $[a,b]$, $|\sigma(f, \mathcal{J}') - (g(b) - g(a))| < \epsilon$.

If $x \in Z$, then $x = z_m$ for some $m \in \mathbb{N}$. g is continuous at x (by 8.16), so there exists some $\kappa_m > 0$ such that

$$|y - z_m| < \kappa_m \Longrightarrow |g(y) - g(z_m)| < \frac{\epsilon}{3.2^m}\,. \tag{40}$$

However, we may also imitate 8.14. Combining (40) and (37), I set

$$\gamma(x) = \gamma(z_m) \coloneqq \min\left(\kappa_m, \frac{\epsilon}{3.2^{m+1}(|f(z_m)|+1)}\right). \tag{41}$$

If $x \notin Z$, $g'(x)$ exists and is equal to $f(x)$. Thus, from 7.1, there is some $\delta(x) > 0$ (specific to this x) such that, if $0 < |h| < \delta(x)$ and $x + h \in [a,b]$, then

$$\left|\frac{g(x+h)-g(x)}{h} - f(x)\right| < \frac{\epsilon}{3(b-a)}. \tag{42}$$

If $t' \le x \le t''$, where $t', t'' \in [a,b]$, and $x - \delta(x) < t' \le x \le t'' < x + \delta(x)$,

$$|g(t') - g(x) - f(x)(t'-x)| \le \frac{\epsilon|t'-x|}{3(b-a)} = \frac{\epsilon(x-t')}{3(b-a)}. \tag{43}$$

(notice that, unlike (42), this holds when $t' = x$, the only case with equality. We excluded $h = 0$ in (42) to ensure that the difference quotient made sense), and similarly

$$|g(t'') - g(x) - f(x)(t''-x)| \le \frac{\epsilon|t''-x|}{3(b-a)} = \frac{\epsilon(t''-x)}{3(b-a)}. \tag{44}$$

(The equality again only arises when $x = t''$.) From (43) and (44),

$$\begin{aligned} &\big|g(t'') - g(t') - (t''-t')f(x)\big|\text{g} \\ &\qquad \le \big|g(t'') - g(x) - f(x)(t''-x)\big| + \big|g(x) - g(t') - f(x)(x-t')\big| \\ &\qquad \le \frac{\epsilon(t''-t')}{3(b-a)}. \end{aligned} \tag{45}$$

(Equality holds only when $t' = x = t''$.)

Let us now define $\gamma(x) \coloneqq \delta(x)$, for each $x \in [a,b] \setminus Z$. This completes the definition of the gauge γ on $[a,b]$.

Suppose $\mathcal{J}' \coloneqq \{t_0, t_1, \ldots, t_n\}$ is a γ-fine tagged division of $[a,b]$. Then

$$\begin{aligned} &\big|\sigma(f,\mathcal{J}') - (g(b)-g(a))\big| \\ &\qquad = \left|\sum_{k=1}^{n} \{(t_k - t_{k-1})f(\tau_k) - (g(t_k) - g(t_{k-1}))\}\right| \\ &\qquad \le \sum_{k=1}^{n} |(t_k - t_{k-1})f(\tau_k) - (g(t_k) - g(t_{k-1}))| \end{aligned} \tag{46}$$

where, for each k, $1 \le k \le n$, $\tau_k - \gamma(\tau_k) < t_{k-1} \le \tau_k \le t_k < \tau_k + \gamma(\tau_k)$.

If $\tau_k \in Z$, we have $\tau_k = z_{m(k)}$ for some $m(k) \in \mathbb{N}$. From (40) and (41),

$$|g(t_k) - g(\tau_k)| < \frac{\epsilon}{3.2^{m(k)}} > |g(t_{k-1}) - g(\tau_k)| \tag{47}$$

(for $\gamma(\tau_k) \le \kappa_m$, and, therefore, $|t_{k-1} - \tau_k| < \kappa_m > |t_k - \tau_k|$); at the same time, since $t_k - t_{k-1} < 2\gamma(\tau_k)$,

$$|(t_k - t_{k-1})f(\tau_k)| \le \frac{2\epsilon}{3.2^{m(k)+1}(|f(z_{m(k)})|+1)}|f(z_{m(k)})| < \frac{\epsilon}{3.2^{m(k)}}. \tag{48}$$

Hence, adding the estimates (47) and (48), when $\tau_k \in Z$

$$|(t_k - t_{k-1})f(\tau_k) - (g(t_k) - g(t_{k-1}))| < 2^{-m(k)}\epsilon\,. \tag{49}$$

On the other hand, if $\tau_k \notin Z$, we have the inequality resulting from (45):

$$|(t_k - t_{k-1})f(\tau_k) - (g(t_k) - g(t_{k-1}))| \le \frac{\epsilon(t_k - t_{k-1})}{3(b-a)}\,.$$

(The inequality is strict unless both sides are 0.) Substituting these estimates in (46), we find

$$\begin{aligned}|\sigma(f,\mathcal{J}') - (g(b) - g(a))| &< \tfrac{1}{3}\epsilon\left(\sum\nolimits_{k=1}^{n}\left\{\frac{t_k - t_{k-1}}{b-a}\right\} + \sum\nolimits_{m=1}^{\infty}\frac{2}{2^m}\right)\\ &< \tfrac{1}{3}\epsilon\left\{\frac{b-a}{b-a}\right\} + \tfrac{2}{3}\epsilon = \epsilon\,.\end{aligned}$$

These inequalities are strict, because the sums include all terms that might arise from a given k, whether $\tau_k \in Z$ or not, and in any case only finitely many k occur; $\sum_{m=1}^{\infty} 2/2^m$ is an estimate for $\sum_{k=1}^{\infty} 1/2^{m(k)}$, because the same tag τ_k may occur for two distinct indices m (for two adjacent intervals) but no more, so that a given value of $m(k)$ may appear at most twice for different k, and only finitely many times at most. If Z is finite, only finitely many subscripts k will be needed, but the estimate remains valid.

Thus, the Henstock integral of f exists and is $g(b) - g(a)$. □

Corollary 8.18. *If $f : [a,b] \longrightarrow \mathbb{K}$ has a primitive $g : [a,b] \longrightarrow \mathbb{K}$, then*

$$(\forall x \in [a,b]) \quad g(x) - g(a) = \int_a^x f(t)\,dt\,,$$

where the integral is the Henstock integral. □

Note 8.19. *(a)* Although the proof of 8.17 seems long and involves some messy estimates, it is fairly straightforward in principle. First we deal, rather as in 8.14, with the points where the derivative of g either does not exist or does not agree with f, and then use the definition of differentiability. It is, indeed, rather surprising that the result was not noticed earlier (if it wasn't). The "Cauchy integral" (as I called it) of 7.14*(b)* is obviously closely related.

(b) In footnote 4 (and in MATH 243) I mentioned "sets of zero measure". It is not overwhelmingly difficult to generalize 8.14, with a rather similar proof, to allow the "exceptional set" Z (i.e. the set of exceptions) to be of zero measure rather than merely countable. One might naturally expect something similar for 8.17; however, the obvious candidate for a generalization (that is, letting $g'(x) = f(x)$ except at the points of a set of zero measure) is *false*—there is a famous example of a function g on $[0,1]$ and a set of measure zero Z in $[0,1]$ such that g is continuous on $[0,1]$ and differentiable with zero derivative on $[0,1] \setminus Z$, but $g(0) = 0$, $g(1) = 1$. The integral of g' (which is defined except on Z) exists and is 0, by the generalized 8.14, so it does *not* give the increment of g.

Something must go wrong in the proof above. In fact, continuity of g on its own is not good enough to control the estimates at uncountably many points of non-differentiability. We made use in 8.14 of $f(z_n)$ in the definition of γ; that would not be possible if we had to handle uncountably many potential tags. There is a well-known "generalization" of 8.17 in the context of the Lebesgue integral, but it imposes a much stronger condition than continuity on g (so-

called absolute continuity), and as a consequence forces the derivative to be absolutely integrable. This excludes some useful cases in which 8.17 is valid.

(c) The Theorem makes clear that all the examples of "improper Riemann integrals" on bounded intervals that we have ever met are in fact Henstock integrals. For instance, in the integral $\int_0^1 \frac{1}{\sqrt{t}}\,dt$ the integrand may be defined to have any desired value at 0, and then has a primitive $2\sqrt{t}$.

(d) The function defined by $\sin\left(\frac{1}{t}\right) - \frac{1}{t}\cos\left(\frac{1}{t}\right)$ for $t > 0$ (and, for instance, taking the value 0 at 0) has a primitive $t\sin\left(\frac{1}{t}\right)$, so is Henstock-integrable on $[0,1]$. So is the function $\sin\left(\frac{1}{t}\right)$ (directly from the definition, or by various other arguments); so $\frac{1}{t}\cos\left(\frac{1}{t}\right)$ is also Henstock-integrable on $[0,1]$. However, its absolute value is not, though this is not altogether obvious without more development of the theory.

§9. Power series.

In §4, I was considering *numerical* series, whose terms are (real or complex) numbers. It is also possible to consider series whose terms are functions, $\sum f_n$. Familiar examples are power series and Fourier series, but there are other kinds (Dirichlet series, wavelet expansions, Mittag-Leffler expansions). Again, to describe such a series as *convergent* is to assert that the sequence of partial sums has a limit; but where functions are involved, the ideas both of a 'partial sum' and of a 'limit' may need further explanation.

We shall only deal with so-called power series, and (prematurely) with the complex case (see §10), of which the real case is an easy corollary.

Definition 9.1. A power series in one variable z with complex coefficients c_n, $n = 0, 1, 2, 3, \ldots$, about the centre $a \in \mathbb{C}$, is a series $\sum c_n(z-a)^n$, whose terms are monomials $c_n(z-a)^n$ in $z-a$.

The indices of all the series will now be numbered from 0 (not from 1) for convenience, and I denote the index set of non-negative integers by $\mathbb{Z}_+$. To begin with, the monomials are just algebraic expressions in an 'unknown' (a mere symbol) z. Associated to such a monomial is a monomial *function* $f_n : \mathbb{C} \longrightarrow \mathbb{C} : z \mapsto c_n(z-a)^n$, which, when z is assigned a complex value ζ, gives the function f_n the complex-numerical value $f_n(\zeta) \coloneqq c_n(\zeta - a)^n$. In analysis, one can usually treat lightly the distinction between the monomial itself (an algebraic expression) and the monomial function it defines; for example, we often write "the function z^2", although z^2 is just an algebraic expression and the function is really $z \mapsto z^2$.

When we substitute for the 'unknown' z the complex value ζ, each term of the power series acquires a complex value, inevitably denoted $c_n(\zeta - a)^n$. (In particular, when $n = 0$, $c_0(\zeta - a)^0$ is understood as c_0 for every choice of ζ. If $\zeta = a$, this is a special convention for power series, because it may be argued that 0^0 should in general be undefined.)

In this way, to each complex number ζ the power series (of monomials) associates a series of complex numbers, denoted $\sum c_n(\zeta - a)^n$. If this numerical series is convergent (in the usual sense that its partial sums in $\mathbb{C}$ form a convergent sequence in $\mathbb{C}$—cf. 2.9), we say that

the power series converges at the complex number ζ. Similarly, if the numerical series is absolutely convergent, that is, if the real series $\sum |c_n(\zeta - a)^n|$ converges in $\mathbb{R}$, we say the power series is *absolutely* convergent at the complex number ζ.

Theorem 9.2. *The power series $\sum c_n(z-a)^n$ converges at a. If it converges at ζ_0, then it converges absolutely when $|\zeta - a| < |\zeta_0 - a|$.*

(The second sentence tells us nothing when $\zeta_0 = a$.)

Proof. The first statement is clear; the numerical series is $c_0 + 0 + 0 + \cdots$. Now suppose that $\sum c_n(\zeta_0 - a)^n$ converges. By 4.18, $c_n(\zeta_0 - a)^n \to 0$ as $n \to \infty$. (Recall that this is true because $c_n z_0^n$ is the difference of consecutive partial sums of the series, and the partial sums converge to a limit by hypothesis). Hence there is a constant $K \geq 0$ such that $|c_n(\zeta_0 - a)^n| \leq K$ for all n (see 2.14*(e)*). Thus, for each n,

$$|c_n(\zeta - a)^n| \leq K \left| \frac{\zeta - a}{\zeta_0 - a} \right|^n .$$

$\sum |(\zeta - a)/(\zeta_0 - a)|^n$ converges (geometric series, common ratio less than 1), so, by 4.7, $\sum |c_n(\zeta - a)^n|$ converges, i.e. $\sum c_n(\zeta - a)^n$ converges absolutely. □

Consider the set E of real numbers $r \geq 0$ such that there is a complex number $\zeta_0 \in \mathbb{C}$ with $|\zeta_0 - a| = r$ and $\sum c_n(\zeta_0 - a)^n$ convergent. 9.2 shows that $0 \in E$ and that, if $r \in E$, then $r' \in E$ for any $r' \in [0, r)$. If E is not bounded above, write $R = \infty$, a symbolic value. If E is bounded above, let $R = \sup E$.

Definition 9.3. R, as above, is the *radius of convergence* of $\sum c_n(z-a)^n$.

Remark 9.4. This definition, with 9.2, has the consequences that

(a) if $R = 0$, $E = \{0\}$ and the series converges only at $\zeta = a$,

(b) if $R = \infty$, $E = [0, \infty)$, and the series converges *absolutely* at all values of ζ. Indeed, if $|\zeta - a| = r'$, there exists $r \in E$ such that $r > r'$ (otherwise r' would be an upper bound for E), and 9.2 assures us that $\sum c_n(\zeta - a)^n$ converges absolutely. Similarly,

(c) if $R \in [0, \infty)$, the series converges *absolutely* when $|\zeta - a| < R$ and diverges when $|\zeta - a| > R$. If $|\zeta - a| = r' < R$, there exists some $r \in E$ such that $r' < r < R$, $\sum c_n(\zeta_0 - a)^n$ converges if $|\zeta_0 - a| = r$ (such a ζ_0 does exist!), and then $\sum c_n(\zeta - a)^n$ converges absolutely by 9.2. If $|\zeta_1 - a| = r > R$, then $r \notin E$, and so $\sum c_n(\zeta_1 - a)^n$ does not converge, by the definition of E.

(d) There is no general statement about convergence if $|\zeta - a| = R$. The series $\sum z^n$ diverges for all z on the "unit circle" $\{z : |z| = 1\}$ (for its terms do not tend to 0), but

$$\frac{z}{1} + \frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \cdots$$

diverges when $z = 1$ (it becomes the harmonic series) and converges at all other points of the unit circle (I omit the proof); whilst

$$\frac{z}{1^2} + \frac{z^2}{2^2} + \frac{z^3}{3^2} + \cdots$$

converges (absolutely) at all points of the circle, since the moduli of the terms form the series $\sum n^{-2}$, for which see 4.11 and 4.13. All three series have 1 as radius of convergence.

(e) 9.2 implies that the radius of convergence of $\sum c_n(z-a)^n$ may be determined by considering only values of z such that $z-a$ is real (convergence being in $\mathbb{C}$).

(f) It is often possible in useful cases (where the coefficients of the series are fairly well-behaved) to find radii of convergence by the ratio test; this is, indeed, why the ratio test has a certain priority amongst convergence tests, as "the first thing you try".

Definition 9.5. Let A be the set of complex numbers ζ at which the series $\sum c_n(z-a)^n$ converges. If $\zeta \in A$, the sum to infinity $\sum_{n=0}^{\infty} c_n(\zeta-a)^n$ is defined in $\mathbb{C}$. It defines a function $f : A \longrightarrow \mathbb{C}$, which is called the *sum-function* or *sum* of the power series: we write

$$f(z) \coloneqq \sum_{n=0}^{\infty} c_n(z-a)^n.$$

Remark 9.6. A complex numerical series $\sum \alpha_n$ is absolutely convergent if and only if $\sum \Re\alpha_n$ and $\sum \Im\alpha_n$ are both absolutely convergent. See 4.15.

Theorem 9.7. *Let $\sum \alpha_n$ be an absolutely convergent series of complex numbers. Then any rearrangement or grouping or rearranged grouping of the terms of the series also yields an absolutely convergent series; and the sum is unchanged.*

Proof. This follows from 4.23, with 4.15. □

Suppose now that $\sum \alpha_n$ and $\sum \beta_n$ are absolutely convergent complex series. We may arrange the terms $\alpha_n\beta_p$ in a single sequence (γ_q), so that to each pair (n,p) there corresponds a unique value of q. (The Cartesian product $\mathbb{Z}_+ \times \mathbb{Z}_+$ is in one-one correspondence with $\mathbb{Z}_+$; in other words, the product of two countable sets is countable.) In this situation

Lemma 9.8. *$\sum \gamma_q$ is absolutely convergent.*

Proof. Since $|\gamma_q|$ is nonnegative for each q, it suffices to show the partial sums $\sum_{n=0}^{Q} |\gamma_q|$ are bounded as Q varies, by 4.6. For any given Q, the terms $\gamma_0, \gamma_1, \ldots, \gamma_Q$ constitute a selection of finitely many of the products $\alpha_n\beta_p$, for which there will be a largest value N of the suffixes "n" and a largest value P of the suffixes "p"; thus $\gamma_0, \gamma_1, \ldots, \gamma_Q$ form a subcollection of

$$\alpha_0\beta_0, \alpha_1\beta_0, \ldots, \alpha_N\beta_0, \alpha_0\beta_1, \alpha_1\beta_1, \ldots, \alpha_N\beta_1, \ldots, \alpha_0\beta_P, \alpha_1\beta_P, \ldots, \alpha_N\beta_P,$$

$$\text{and} \quad \sum_{q=0}^{Q} |\gamma_q| \le \sum_{0\le n\le N,\, 0\le p\le P} |\alpha_n||\beta_p| = \sum_{n=0}^{N} |\alpha_n| \sum_{p=0}^{P} |\beta_p|$$

(for every term on the left of the first equality appears also on the right, and all are non-negative). If $A = \sum_{n=0}^{\infty} |\alpha_n|$, $B = \sum_{p=0}^{\infty} |\beta_p|$, it follows that all partial sums $\sum_{q=0}^{Q} |\gamma_q|$ are bounded by AB; that suffices. □

Now, according to 9.7, the sum $\sum \gamma_q$ is unchanged by arbitrary groupings and rearrangements. One such (not a very pleasant one, since infinitely many terms are grouped together at each step) is

$$\sum_{n=0}^{\infty}\Big(\sum_{p=0}^{\infty} \alpha_n\beta_p\Big) = \sum_{n=0}^{\infty} \alpha_n\Big(\sum_{p=0}^{\infty} \beta_p\Big) = \Big(\sum_{n=0}^{\infty} \alpha_n\Big)\Big(\sum_{p=0}^{\infty} \beta_p\Big).$$

A rather natural way to group terms together is the so-called *Cauchy product*, for which we group together all terms $\alpha_n\beta_p$ with the same total index $n+p$. The Cauchy product of the series $\sum\alpha_n$ and $\sum\beta_p$ will be $\sum\sigma_r$, where, for each $r\in\mathbb{Z}_+$, $\sigma_r=\sum_{n+p=r}\alpha_n\beta_p$.

Theorem 9.9. *If $\sum\alpha_n$ and $\sum\beta_p$ are absolutely convergent complex series, so is their Cauchy product $\sum\sigma_r$, and $\sum\sigma_r=(\sum\alpha_n)(\sum\beta_p)$.*

Proof. The equality has been proved ($\sum\sigma_r$ is a grouping and rearrangement of $\sum\gamma_q$). It only remains to note that $|\sigma_r|\le\sum_{n+p=r}|\alpha_n||\beta_p|$, so that $\sum|\sigma_r|\le\sum_r\sum_{n+p=r}|\alpha_n||\beta_p|$, which is a grouping and rearrangement of $\sum|\gamma_q|$ and therefore converges by 9.7. □

The Cauchy product of two numerical series behaves much more satisfactorily than this; it converges and has the sum $(\sum\alpha_n)(\sum\beta_p)$ when $\sum\alpha_n$ and $\sum\beta_p$ both converge and only one of them is absolutely convergent. There would be little point in giving it a special name if most of its properties were shared by all other methods of grouping the terms $\alpha_n\beta_p$. However, we need no more information than 9.9.

The reason for studying the Cauchy product is that it is well adapted to the multiplication of power series. For given $\zeta\in\mathbb{C}$, the Cauchy product of $\sum c_n(\zeta-a)^n$ and $\sum d_p(\zeta-a)^p$ is $\sum s_r(\zeta-a)^r$, where $s_r=\sum_{n+p=r}c_nd_p$. In symbols rather than numbers, it is natural to describe $\sum s_r(z-a)^n$ as the Cauchy product of the power series $\sum c_n(z-a)^n$ and $\sum d_p(z-a)^p$.

We can also speak of the sum of the power series $\sum c_n(z-a)^n$ and $\sum d_p(z-a)^p$, meaning the power series $\sum(c_n+d_n)(z-a)^n$.

Theorem 9.10. *Suppose that $\sum c_n(z-a)^n$ has radius of convergence R_1 and $\sum d_p(z-a)^p$ has radius of convergence R_2. Then the sum and the product of these power series have radii of convergence not less than* $\min(R_1,R_2)$.

Proof. Take any ζ such that $|\zeta-a|<\min(R_1,R_2)$. Then $|\zeta-a|<R_1$, so that $\sum c_n(\zeta-a)^n$ is absolutely convergent, and $|\zeta-a|<R_2$, so that $\sum d_p(\zeta-a)^p$ is absolutely convergent; it follows from the general properties of series that the sum $\sum(c_n(\zeta-a)^n+d_n(\zeta-a)^n)$ or $\sum(c_n+d_n)(\zeta-a)^n$ is absolutely convergent, and from 9.9 that the product series $\sum s_r(\zeta-a)^r$, as defined above, is absolutely convergent. □

Definition 9.11. *(a)* Let $a\in\mathbb{C}$ and $r>0$ (that is, r is to be a positive real number). The *open disk in $\mathbb{C}$ of radius r about a* is the set

$$B(a;r)\coloneqq\{z\in\mathbb{C}:\,|z-a|<r\},$$

where the B recalls its alternative name as an open *ball*. (The only reason for preferring the name "disk" here is to emphasize that it is a set in two dimensions.) The *closed disk about a of radius r* is

$$C(a;r)\coloneqq\{z\in\mathbb{C}:|z-a|\le r\}.$$

(b) A subset U of $\mathbb{C}$ is *open* if, for any $u\in U$, there is some $r>0$ (specific to u) such that $B(u;r)\subseteq U$. (See §20, in particular 20.10. This is equivalent in the context of $\mathbb{C}$ [or of any metric space] to the previous definition 3.6, as you may easily check.)

(c) Suppose U is an open subset of $\mathbb{C}$, and $f : U \longrightarrow \mathbb{C}$. f is *complex-differentiable* at $a \in U$ if the difference quotient $\dfrac{f(a+h)-f(b)}{h}$ has a limit in $\mathbb{C}$ as $h \to 0$ through complex values. The limit, when it exists, is called the complex-derivative of f at a, and is denoted $f'(a)$ or df/dz or for more precision $\dfrac{df}{dz}(a)$ (etc., etc.) (Notice that there is $r > 0$ such that $a + h \in U$ whenever $|h| < r$, so the difference quotient makes sense for all complex h such that $0 < |h| < r$. The limit means that, for every $\epsilon > 0$, there is $\delta > 0$ such that

$$(0 < |h| < \delta \ \ \& \ \ a + h \in U) \Longrightarrow \left| \frac{f(a+h)-f(a)}{h} - f'(a) \right| < \epsilon .$$

This definition should be compared with 7.1; the essential difference is that here the limit is taken in $\mathbb{C}$ and with respect to a two-dimensional (that is, complex) variable.

Note 9.12. If, in 9.11*(c)*, f is complex-differentiable at $a \in U$, it is continuous at a. The proof is almost the same as at 7.2 (with changes to take account of the changed hypotheses).

Power series are important in this course for two reasons. The first is that they can be used to give precise and logical definitions of the elementary transcendental functions (see below). But the 'holomorphic' functions we shall soon be discussing are, as it will turn out, those which can be *locally* represented as the sums of complex power series. In one direction, this involves proving that the sum-function of a power series is complex-differentiable where possible. There is an 'advanced' proof of this fact using the Cauchy integral formula. The 'elementary' proof I give here is, in fact, better and more general; that is, it can be applied in other contexts, not just for complex power series. To shorten formulæ I shall take the centre of the power series to be 0, but this amounts only to a change of variable (from $z - a$ to z) in the calculation.

Given a power series $\sum c_n z^n$, we call the power series $\sum (n+1)c_{n+1}z^n$ (in which, as usual for power series, n is taken in $\mathbb{Z}_+$) the 'formal derivative' of $\sum c_n z^n$. The word 'formal', which is very common among mathematicians, emphasizes that we are not speaking here of a genuine derivative of a genuine function (neither series is assumed even to converge except at 0, for instance), but only of something that has the 'form' one might hope such a thing to have if it existed.

Lemma 9.13. *The radius of convergence of a power series is the same as the radius of convergence of its formal derivative.*

Proof. Let the radius of convergence of $\sum (n+1)c_{n+1}z^n$ be $R > 0$, and $|\zeta| < R$. Then $\sum (n+1)c_{n+1}\zeta^n$ is absolutely convergent by 9.2, and

$$|(c_{n+1}\zeta^{n+1})/((n+1)c_{n+1}\zeta^n)| = |\zeta|/(n+1) \le |\zeta| \quad \text{for all } n,$$

so that, by 4.7 (the comparison test; take "K" to be $|\zeta|$), $\sum |c_{n+1}\zeta^{n+1}|$ is also convergent. Thus, from 4.5*(a)*, $\sum c_n\zeta^n$ converges absolutely if $|\zeta| < R$, and the radius of convergence R' of $\sum c_n z^n$ cannot be less than R, $R' \ge R$. If $R = 0$ or $R = \infty$, this is true anyway (in the obvious sense).

Suppose $R' > 0$, and $|\zeta| < R'$. Take r such that $|\zeta| < r < R'$. $\sum c_n r^n$ converges absolutely; there is K such that $|c_n r^n| \le K$ for all n. For each n, $|(n+1)c_{n+1}\zeta^n| \le (n+1)Kr^{-n-1}|\zeta|^n$, and $\sum Kr^{-1}(n+1)(|\zeta|/r)^n$ converges by the ratio

test. Hence $\sum (n+1)c_{n+1}\zeta^n$ converges absolutely whenever $|\zeta| < R'$, and the radius of convergence R of $\sum (n+1)c_{n+1}z^n$ cannot be less than R'. If $R' = 0$, $R \geq R'$ automatically. Hence $R = R'$ in all cases (including $R' = \infty$). □

If the radius of convergence of the series is 0, the sum-function is only defined at the origin and cannot, therefore, be differentiated.

Theorem 9.14. *Let $\sum c_n z^n$ be a power series with positive radius of convergence R. For $|z| < R$, let $f(z)$ be the sum-function of the series, and $g(z)$ the sum of the formal derivative series (which is also defined on $B(0;R)$, by* 9.13*). Then, at each point $\zeta \in B(0;R)$, the function f is complex-differentiable, and $f'(\zeta) = g(\zeta)$.*

Proof. Fix $\zeta \in B(0;R)$, so that $|\zeta| < R$ and $f(\zeta), g(\zeta)$ are both defined. Choose some r so that $|\zeta| < r < R$. Then $\sum c_n r^n$ converges, so that its terms are bounded: there is $K \geq 0$ such that $|c_n|\, r^n \leq K$ for all n, or $|c_n| \leq K r^{-n}$.

Take any $h \in B(0; r - |\zeta|)$, with $h \neq 0$. Then $|\zeta + h| \leq |\zeta| + |h| < r$, so that $f(\zeta + h)$ is defined, and furthermore

$$\left| \frac{f(\zeta+h) - f(\zeta)}{h} - g(\zeta) \right| = \left| \frac{\sum c_n(\zeta+h)^n - \sum c_n \zeta^n - \sum (n+1)c_{n+1}\zeta^n h}{h} \right|$$
$$= \left| \sum c_{n+1} \left(\frac{(\zeta+h)^{n+1} - \zeta^{n+1} - (n+1)\zeta^n h}{h} \right) \right|$$

(by general facts about series; the terms for $n = 0$ in the first two series cancel)

$$= \left| \sum\nolimits_{n=1}^{\infty} c_{n+1} \left\{ \sum\nolimits_{k=2}^{n+1} \binom{n+1}{k} \zeta^{n+1-k} h^{k-1} \right\} \right|$$

(by the binomial theorem)

$$\leq \sum K r^{-n-1} \left\{ \sum\nolimits_{k=2}^{n+1} \binom{n+1}{k} |\zeta|^{n+1-k} |h|^{k-1} \right\}$$
$$= \sum\nolimits_{n=0}^{\infty} K \frac{(|\zeta| + |h|)^{n+1} - |\zeta|^{n+1} - (n+1)|\zeta|^n |h|)}{r^{n+1}|h|}$$
$$= K \left\{ \frac{|\zeta| + |h|}{|h|(r - |\zeta| - |h|)} - \frac{|\zeta|}{|h|(r-|\zeta|)} - \frac{r}{(r-|\zeta|)^2} \right\}$$

(by actual summation of the three series involved—which are all standard series)

$$= K \frac{r|h|(r-|\zeta|) - r|h|(r - |\zeta| - |h|)}{|h|(r-|\zeta|)^2(r - |\zeta| - |h|)}$$
$$= K \frac{r|h|}{(r-|\zeta|)^2(r-|\zeta|-|h|)} \to 0 \quad \text{as } h \to 0.$$

This shows that $f'(\zeta)$ exists and is $g(\zeta)$, as required. □

Corollary 9.15. *The function* $f : B(0;R) \longrightarrow \mathbb{C}$ *is complex-differentiable* k *times, for any positive integer* k*, at each point* ζ *of* $B(0;R)$*. Furthermore, for each such* ζ*,*

$$f^{(k)}(\zeta) = \sum\nolimits_{n=0}^{\infty} \frac{(n+k)!}{n!}\, c_{n+k}\zeta^n\,, \tag{50}$$

where the series converges.

Proof. Apply 9.13 and 9.14 inductively. (50) is the kth formal derivative of $\sum c_n\zeta^n$. □

Note 9.16. The results above may be summarized by saying that, provided one remains strictly inside the 'circle of convergence' $\{\zeta : |\zeta - a| = R\}$, where R is the radius of convergence (possibly ∞), one may treat power series about any centre a almost as if they were polynomials; the functions they define may be added, multiplied (using the Cauchy product), and differentiated 'term-by-term' in the natural ways. Let us apply these ideas to give a treatment of the elementary transcendental functions. Later on, knowledge of the (real) exponential and circular functions will be assumed, but the definitions of these functions as you have so far met them are in some ways unsatisfactory.

Those definitions were roughly as follows. The trigonometrical functions were defined from right-angled triangles; relations amongst them, such as the addition and derivation formulæ, were obtained by geometrical arguments; and Maclaurin series (that is, Taylor series about 0) for the sine and cosine were found from the derivation formulæ. For the exponential function, one possible procedure was to define it as the inverse function of the logarithm, which, for positive values x of the variable, was the (Riemann) integral of $1/t$ over the interval $[1,x]$ of values of t (or minus the integral over $[x,1]$).

Definitions of this kind, which (at least for the circular functions) follow historical lines, are open to several "philosophical" objections. To begin with, they are more complicated than one might wish. To define the trigonometrical functions in this way, we must assume rather a lot of Euclidean geometry (up to similar triangles and Pythagoras's theorem); and, since the 1800s, we are aware that the logical structure of Euclidean geometry is not as straightforward as Euclid once thought, and that it has no absolute claim to be an accurate description of the real world. It is undesirable to define the circular functions, which will appear in contexts far removed from geometry, in a way that makes them dependent on a specific, quite complicated, and maybe unrealistic geometrical theory. As for the above definition for the exponential, it requires a fairly full discussion of the definite integral even to state it, and its extension to the complex exponential function (to be discussed later) demands a rather "unnatural" appeal to the trigonometrical functions.

One way of setting up the real trigonometrical functions which avoids the geometry, and allows a parallel definition of the exponential function, is to define them as solutions of suitable differential equations. Thus sin is the solution of the second-order equation

$$y'' + y = 0\,, \quad y(0) = 0\,, \quad y'(0) = 1\,,$$

and cos is its derivative; whilst exp is the solution of the first-order equation

$$y' - y = 0\,, \quad y(0) = 1\,.$$

These definitions are certainly preferable in a logical sense, in that they involve no appeal to anything outside analysis, but they assume that the differential equations do have solutions satisfying the initial-value conditions and that those solutions are unique. (There are differential

equations for which either existence or uniqueness fails.) The equations above do satisfy the hypotheses of Picard's existence and uniqueness theorem for solutions of ordinary differential equations, but the theorem itself would take some time and effort to set up.

Our approach is therefore crude and direct: we *define* the functions by their "Maclaurin series". This approach is given with some historical comments in the Appendix to Whittaker and Watson. It has a small secondary advantage: the various other definitions raise the question of showing that the Maclaurin series do converge to the functions. This involves using a theorem about remainder terms (Lagrange's form of Taylor's theorem will do). If the sum-functions of the series *are* the functions by definition, there is nothing more to prove. Of course there is a loss as well as a gain—at some stage the functions defined by the series must be identified with the ones that arise in geometry—but this difficulty is much less serious; without going into details, one finds that the functions have the same properties, so they must be the same. So we say:

Definition 9.17. The exponential function $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ is defined by

$$(\forall \zeta \in \mathbb{C}) \quad \exp \zeta := 1 + \frac{1}{1!}\zeta + \frac{1}{2!}\zeta^2 + \ldots = \sum\nolimits_{n=0}^{\infty} \frac{\zeta^n}{n!} .$$

It is easily seen (for instance by the ratio test) that this power series has radius of convergence ∞. Thus $\exp$ is complex-differentiable at all points of $\mathbb{C}$, by 9.14. Note that $\exp(0) = 1$.

Proposition 9.18. *For any complex numbers ζ and η,*

$$\exp(\eta + \zeta) = \exp \zeta \,.\, \exp \eta \quad \textit{and} \quad \exp' \zeta = \exp \zeta \,.$$

Proof. By 9.14,

$$\exp' \zeta = \sum\nolimits_{n=0}^{\infty} \frac{(n+1)}{(n+1)!}\zeta^n = \sum\nolimits_{n=0}^{\infty} \frac{1}{n!}\zeta^n = \exp \zeta \,;$$

the formal derivative series (recall 4.5) may be re-indexed in the obvious way. Likewise

$$\exp \zeta \,.\, \exp \eta = \left(\sum \frac{1}{n!}\zeta^n \right) \left(\sum \frac{1}{p!}\eta^p \right) .$$

Here each series is absolutely convergent, so that the product of their sums is the sum of their Cauchy product, by 9.9. Thus

$$\begin{aligned} \exp \zeta \,.\, \exp \eta &= \sum\nolimits_{q=0}^{\infty} \left(\sum\nolimits_{n+p=q} \frac{1}{n!}\frac{1}{p!}\zeta^n \eta^p \right) \\ &= \sum\nolimits_{q=0}^{\infty} \frac{1}{q!} \left(\sum\nolimits_{n+p=q} \frac{q!}{n!p!}\, \zeta^n \eta^p \right) \\ &= \sum\nolimits_{q=0}^{\infty} \frac{1}{q!}(\zeta + \eta)^q \qquad \text{by the binomial theorem} \\ &= \exp(\zeta + \eta) \qquad \text{by definition.} \end{aligned}$$

□

Definition 9.19. The *hyperbolic* functions cosh and sinh are defined by

$$\cosh\zeta := \tfrac{1}{2}(\exp\zeta + \exp(-\zeta)) \qquad \text{and}$$
$$\sinh\zeta := \tfrac{1}{2}(\exp\zeta - \exp(-\zeta))\,.$$

The *trigonometric* (or *circular*) functions cos and sin are defined by

$$\cos\zeta := \cosh(i\zeta)\,, \qquad \sin\zeta := -i\sinh(i\zeta)\,.$$

It follows instantly that these functions have the customary power series expansions:

$$\cosh\zeta = 1 + \frac{1}{2!}\zeta^2 + \frac{1}{4!}\zeta^4 + \ldots, \qquad \cos\zeta = 1 - \frac{1}{2!}\zeta^2 + \frac{1}{4!}\zeta^4 - \ldots,$$
$$\sinh\zeta = \zeta + \frac{1}{3!}\zeta^3 + \frac{1}{5!}\zeta^5 + \ldots, \qquad \sin\zeta = \zeta - \frac{1}{3!}\zeta^3 + \frac{1}{5!}\zeta^5 - \ldots.$$

Furthermore, 9.18 leads to all the addition formulæ, and to the identities

$$\cosh^2\zeta - \sinh^2\zeta = 1\,, \qquad \cos^2\zeta + \sin^2\zeta = 1\,, \quad \text{for all } \zeta \in \mathbb{C}\,.$$

(I leave the derivations of these facts as easy exercises; they were given for the hyperbolic functions in 113). All these functions are defined by series having only real coefficients, so that they take real values for real ζ, and we can treat them as real-valued functions of a real variable when appropriate; but the addition formulæ and the identities are also valid for complex values of the variables. However, if we consider real values, we have

Lemma 9.20. $\cos 0 = 1$ *and* $\cos 2 < 0$.

Proof. The first statement is trivial. For the second, argue as follows. Let

$$\begin{aligned}\Theta &= \frac{1}{4!}\times 2^4 - \frac{1}{6!}\times 2^6 + \frac{1}{8!}\times 2^8 - \frac{1}{10!}\times 2^{10} + \ldots \\ &= \frac{1}{4!}\times 2^4\left(1 + \frac{2^4}{5.6.7.8} + \ldots\right) - \frac{1}{6!}\times 2^6\left(1 + \frac{2^4}{7.8.9.10} + \ldots\right) \\ &< \frac{2}{3}(1 + 2^{-4} + 2^{-8} + \ldots) - \frac{4}{45} = \frac{2}{3}\times\frac{1}{1-\frac{1}{16}} - \frac{4}{45} = \frac{28}{45}\,.\end{aligned}$$

From this, $\cos 2 = 1 - \frac{1}{2!}\times 2^2 + \Theta = -1 + \Theta < -\frac{17}{45}$. □

The argument is to calculate an approximate value for $\cos 2$ from early terms of the series, and then estimate the error in this approximation by comparing the sum of the "tail" with the sum of a geometric series. The computation above is already far more accurate than we need; in fact it is fairly obvious that

$$\Theta < \tfrac{2}{3}\left(1 + \tfrac{2^2}{5.6} + \tfrac{2^4}{5.6.7.8} + \ldots\right) < \tfrac{2}{3}\times\tfrac{1}{1-\frac{1}{4}} = \tfrac{8}{9}\,,$$

which is good enough for the Lemma (for which we only need $\Theta < 1$). On the other hand,

Lemma 9.21. *Suppose* $0 < \xi < 2$. *Then*

$$-1 < 1 - \tfrac{1}{2}\xi^2 < \cos\xi < 1\,. \tag{51}$$

Proof. As previously,

$$\begin{aligned}
\cos\xi &= 1 - \frac{1}{2!}\times\xi^2 + \Theta_1(\xi)\,, \quad \text{where} \\
\Theta_1(\xi) &= \frac{1}{4!}\times\xi^4\left(1+\frac{\xi^4}{5.6.7.8}+\dots\right) - \frac{1}{6!}\times\xi^6\left(1+\frac{\xi^4}{7.8.9.10}+\dots\right) \\
&> \frac{1}{4!}\times\xi^4 - \frac{1}{6!}\times\xi^6\left(1+\frac{\xi^4}{7^4}+\dots\right) \\
&> \frac{1}{4!}\times\xi^4 - \frac{1}{6!}\times\xi^6.\frac{2401}{2385} > \frac{1}{4!}\times\xi^4\left(1-\frac{4}{30}\cdot\frac{2401}{2385}\right) > 0\,, \\
\Theta_1(\xi) &< \frac{1}{4!}\times\xi^4\left(1+\frac{\xi^4}{5.6.7.8}+\dots\right) < \frac{1}{4!}\times\xi^2\left(2^2+\frac{2^6}{5.6.7.8}+\dots\right) \\
&< \frac{1}{3!}\xi^2.\frac{16}{15} < \frac{1}{2}\xi^2\,,
\end{aligned}$$

and these estimates give (51). □

Corollary 9.22. *There is a number θ such that $0<\theta<2$ and $\cos\theta = 0$.*

Proof. By 9.14, cos is complex-differentiable everywhere, and therefore continuous everywhere (as a complex-valued function of a complex variable). It is therefore certainly continuous if one only considers real values of the variable. As $\cos 0 = 1$ and $\cos 2 < 0$, the existence of θ follows from the intermediate value theorem (which is a result about real-valued continuous functions of a real variable). □

Definition 9.23. Let the *least* positive zero of cos be denoted by $\pi/2$.

That is, π is the number such that $\pi > 0$, $\cos(\frac{1}{2}\pi) = 0$, and, for any α with $0<\alpha<\frac{1}{2}\pi$, $\cos\alpha > 0$.

By 9.22, there is *some* positive number θ such that $\cos\theta = 0$. We want the least member of the set $A = \{\tau : \tau\in\mathbb{R} \;\&\; 0\le\tau\le 2 \;\&\; \cos\tau = 0\}$. 9.22 ensures $A\neq\emptyset$, and A is bounded below by 0 by definition; hence it has an infimum $\alpha\ge 0$. For any positive integer n, there is an element $\alpha_n\in A$ such that $\alpha\le\alpha_n<\alpha+\frac{1}{n}$, by 2.8. Hence $\alpha_n\to\alpha$ and $\cos\alpha_n = 0$ for each n; but, as cos is continuous, it follows that $\cos\alpha = 0$. Since $\cos 0 = 1$, this implies that $\alpha > 0$, and therefore $\alpha\in A$. So the infimum of A belongs to A, and must be its least element. (More briefly: A is a closed bounded non-empty subset of $\mathbb{R}$, so must contain its infimum; this follows instantly from 3.8 and 5.8.)

So there is a least positive zero of the cosine. Is the θ of Corollary 9.22 necessarily that zero? By (50), $1>\cos\xi>-1$ for $0<\xi<2$; hence $\sin^2\xi = 1-\cos^2\xi > 0$, and $\sin\xi$ cannot be zero. The mean-value theorem tells us that, whenever $\theta_1\neq\theta_2$, $\cos\theta_1-\cos\theta_2 = (\theta_1-\theta_2)(-\sin\xi)$ for some ξ strictly between θ_1 and θ_2; it follows that cos is one-to-one on the interval $[0,2]$, and it can take the value 0 at only one point of this interval, namely at θ.

Remark 9.24. We have just seen that $\sin\xi\neq 0$ for $0<\xi<2$. However, $\sin 0 = 0$ and $\cos 0 = \sin' 0 = 1$. Thus $\dfrac{\sin\xi}{\xi} = \dfrac{\sin\xi-\sin 0}{\xi}\to 1$ as $\xi\downarrow 0$, so that $\sin\xi>0$ for small

enough positive values of ξ. (This may also be proved directly from the series for $\sin$, by easy estimates like those in 9.20). Hence $\sin\xi > 0$ for $0 < \xi < 2$, as it cannot take a positive value at one point and a negative value at another without taking the value 0 somewhere in between (by the intermediate value theorem). In particular, as

$$\sin^2(\tfrac{1}{2}\pi) = 1 - \cos^2(\tfrac{1}{2}\pi) = 1\,,$$

$\sin(\frac{1}{2}\pi) = \pm 1$; we have shown the sign must be positive, so $\cos(\frac{1}{2}\pi) = 0$ and $\sin(\frac{1}{2}\pi) = 1$. Now apply the addition formulæ:

$$\begin{aligned}
\cos(\zeta + \tfrac{1}{2}\pi) &= \cos\zeta\cos(\tfrac{1}{2}\pi) - \sin\zeta\sin(\tfrac{1}{2}\pi) = -\sin\zeta\,,\\
\sin(\zeta + \tfrac{1}{2}\pi) &= \sin\zeta\cos(\tfrac{1}{2}\pi) + \cos\zeta\sin(\tfrac{1}{2}\pi) = \cos\zeta\,, \qquad \text{and in turn}\\
\sin(\zeta + \pi) &= \cos(\zeta + \tfrac{1}{2}\pi) = -\sin\zeta\,,\\
\cos(\zeta + \pi) &= -\sin(\zeta + \tfrac{1}{2}\pi) = -\cos\zeta\,. \qquad \text{Hence,}\\
\sin(\zeta + 2\pi) &= -\sin(\zeta + \pi) = \sin\zeta \qquad \text{and}\\
\cos(\zeta + 2\pi) &= -\cos(\zeta + \pi) = \cos\zeta\,.
\end{aligned}$$

So the (complex) circular functions $\cos$ and $\sin$ are *periodic*, with period 2π. This is their smallest positive real period. If $\cos\psi = \cos 0 = 1$, where $\psi > 0$, $\sin^2\psi = 1 - \cos^2\psi = 0$ and $\sin\psi = 0$, and $\cos(\zeta + \psi) = \cos\zeta$ for all ζ by the addition formula; ψ is, then, a period of $\cos$. But also

$$2\cos^2(\tfrac{1}{2}\psi) - 1 = \cos\psi = 1\,, \qquad \cos(\tfrac{1}{2}\psi) = \pm 1\,.$$

If the positive sign holds, $\frac{1}{2}\psi$ is a smaller period of $\cos$. If $\cos(\frac{1}{2}\psi) = -1$,

$$2\cos^2(\tfrac{1}{4}\psi) - 1 = -1 \quad \text{and} \quad \cos(\tfrac{1}{4}\psi) = 0\,.$$

Hence $\frac{1}{4}\psi \geq \frac{1}{2}\pi$, by the definition of $\frac{1}{2}\pi$, and $\psi \geq 2\pi$, as asserted. In the same way, π is the least positive number such that $\cos\pi = -1$.)

It is possible to carry the development of the circular functions much further than this, but the method should now be clear. At no point have we applied to geometrical intuition, or, for that matter, to geometrical argument. The word ‘angle’ is quite redundant. To jump ahead a little: the *principal argument* of a non-zero complex number $\zeta = x + iy$ is that number θ such that $-\pi < \theta \leq \pi$ and

$$\cos\theta = \frac{x}{x^2 + y^2}\,, \qquad \sin\theta = \frac{y}{x^2 + y^2}\,.$$

The existence and uniqueness of such a number is now easily established.

§10. Complex numbers.

Our aim from now on is to describe some of the ways in which the introduction of complex numbers leads to new ideas in analysis (and, to a certain extent, in geometry). We shall discover some very striking results, which seem to arise magically in the complex context, have no elementary analogues in real analysis, and often seem far more intuitive. There is a very prosaic reason for this—the functions we study are much more restricted than those which are customarily discussed in the real case; but nevertheless they include all the functions that one meets in standard applications. That is why the subject used to be called the Theory of Functions, on the principle that any respectable function ought to obey it. (The "set-theoretic" definition of a function that we met in 151 was invented very late; for a very long time people thought of functions as things given by "formulæ".) That said, there is also a *poetical* reason for the theory's amazing properties: it looks, quite seriously, although less mysteriously than it at first sounds, as if the universe actually runs on complex numbers and not real ones. Quantum mechanics is based on complex-valued wavefunctions.

I begin by recollecting some basic remarks about complex numbers.

Remark 10.1. Recall that $\mathbb{R}$ is an ordered field (see 1.8). The usual laws of arithmetic concerning addition, multiplication, and division hold; in particular, one can always divide by any real number different from 0. Moreover, we can compare any two *real* numbers x, y to decide (trichotomy) that either $x < y$, or $x = y$, or $x > y$, and there are rules for the behaviour of these inequalities under arithmetical operations.

The set of all ordered pairs (x, y) of real numbers is, of course, called $\mathbb{R}^2$. We can define an addition in $\mathbb{R}^2$ (the usual addition in $\mathbb{R}^2$ as a vector space) by

$$(x, y) + (u, v) := (x + u, y + v). \tag{52}$$

However, we can also define a multiplication of sorts in $\mathbb{R}^2$, by the formula

$$(x, y)(u, v) := (xu - yv, xv + yu), \tag{53}$$

and we call the resulting object ($\mathbb{R}^2$ with this multiplication and addition) the set of complex numbers, $\mathbb{C}$. Notice that, *as a set*, $\mathbb{C}$ is just $\mathbb{R}^2$, and the new symbol therefore indicates the new algebraic structure. The common (not invariable) custom is to use the letters z and w mostly to denote complex numbers, whilst x, y, u, and v are mostly reserved for real numbers. If I write $z = (a, b)$, it is usually to be understood that z is complex and a and b are real.

Remark 10.2. The addition and multiplication just defined in $\mathbb{C}$ are commutative and associative, and multiplication is distributive over addition. I show, for example, that multiplication is associative. Suppose $z_1 := (x, y)$, $z_2 := (u, v)$, $z_3 := (a, b)$. Then

$$
\begin{aligned}
&(x,y)((u,v)(a,b)) = (x,y)(ua-vb, ub+va) \qquad \text{from (53)},\\
&= (x(ua-vb) - y(ub+va), x(ub+va) + y(ua-vb)), \quad \text{again from (53)},\\
&= ((xu-yv)a - (xv+yu)b, (xu-yv)b + (xv+yu)a), \quad \text{rearranging},\\
&= (xu-yv, xv+yu)(a,b) = ((x,y)(u,v))(a,b), \quad \text{from (53)}.
\end{aligned}
$$

The other statements are proved similarly (but are easier).

Remark 10.3. If the complex number z is (x,y), the real number x is called the 'real part' of z, and is written $\Re z$ or $\mathrm{Re}(z)$, and the real number y is its 'imaginary part', written $\Im z$ or $\mathrm{Im}(z)$. Notice that the 'imaginary part', so called, is a real number! The absurd adjectives 'real' and 'imaginary' are historical survivals, without philosophical justification.

There is an operation called (complex) conjugation in $\mathbb{C}$. If, as before, $z = (x,y)$, then its complex conjugate, written $\overline{z}$ or, if it is typographically more convenient, z^-, is $(x,-y)$. Thus z and $\overline{z}$ have the same real part, but 'opposite' imaginary parts. It is obvious that $(\overline{z})^- = z$ for any $z \in \mathbb{C}$. (The jargon is that conjugation is *involutory*, or an *involution*, which means that doing it twice in succession brings you back where you started). Also conjugation is an *automorphism* of the field $\mathbb{C}$, which means that it respects addition and multiplication; for any complex numbers w and z,

$$(w+z)^- = \overline{w} + \overline{z}, \quad (wz)^- = \overline{w}\,\overline{z}. \tag{54}$$

Remark 10.4. If $\Im z = 0$, or equivalently $z = \overline{z}$, so that z is of the form $(x,0)$, we say that z is *purely real* or just 'real'; if $\Re z = 0$, or $z = -\overline{z}$, so that z is of the form $(0,y)$, we say that z is *purely imaginary*. Now, if w and z are purely real, $z \coloneqq (x,0)$ and $w \coloneqq (u,0)$, we see that $z+w = (x+u,0)$ and $zw = (xu,0)$, so that z and w seem to behave exactly like the real numbers x and u. It is therefore not too misleading to abbreviate $(x,0)$ to x, and so on. Furthermore, for any complex number (x,y),

$$(x,y) \coloneqq (x,0) + (0,1)(y,0)$$

by a simple calculation. Thus, if we denote the complex number $(0,1)$ by i, we can write (x,y) as $x+iy$. (When I write $x+iy$, or $u+iv$, it is usually to be understood that x and y, or u and v, are real numbers, and, of course, one usually writes $0+iy$ as iy and $x+i0$ as x). Now $i^2 = (0,1)(0,1) = (-1,0)$, which we write as -1. The general rules for addition and multiplication become

$$
\begin{aligned}
(x+iy) + (u+iv) &= (x+u) + i(y+v),\\
(x+iy)(u+iv) &= xu + i(yu+xv) + i^2 yv.\\
&= (xu-yv) + i(yu+xv)
\end{aligned}
$$

These would be the expected rules for expressions $x+iy$, if ordinary algebraic manipulations are permitted and $i^2 = -1$. The form $x+iy$ of a complex number is therefore very convenient, because it carries with it the easy mnemonic $i^2 = -1$ for the multiplication. In particular the complex number 1, that is $1+i0$, is a multiplicative identity in $\mathbb{C}$: $1z = z$ for all z.

Note 10.5. The word 'imaginary' was originally introduced because, when complex numbers were invented (essentially by Cardano and Bombelli, in the late sixteenth century, in connection with the solution of the cubic; see Appendix C for an explanation why the cubic seemed to require new ideas when the quadratic did not), it was in effect by *pretending* that there were a square root of -1, called i, and seeing what resulted. But, of course, the real number -1 does not have a real square root, so that i was described as 'imaginary'. The procedure is risky from the logical point of view—how can we be sure that no contradiction will ever result from pretending -1 has a square root? (There was a similar problem with the Axiom of Choice, and at 1.19 I point out that even the real numbers are subject to the same sort of criticism until they have been adequately defined.) The method of constructing complex numbers that I have followed avoids the logical difficulty. It was only explicitly proposed by Hamilton in 1835, although ideas of much the same sort had been around since the 1790s, already more than two centuries after complex numbers had been in common use.

Mathematicians write i for an 'imaginary' square root of -1, but sometimes i already has some other customary meaning. Electrical engineers, who often want i to stand for current, may use j for the square root of -1. People who wish to avoid detailed explanations may write simply $\sqrt{-1}$. Similarly, i is also a letter often exploited as a dummy index in summation; there is a tacit convention (mainly due to negligence) that in such contexts it does not mean a square root of -1. For these and similar reasons, MAPLE denotes $\sqrt{-1}$ by I.

Remark 10.6. The zero of $\mathbb{C}$ is $0+i0$, or $(0,0)$, usually written 0; it is the only complex number which is simultaneously purely real and purely imaginary. As above, $1+i0$ or $(1,0)$, usually written 1, is a multiplicative identity. Now, if $(x,y)\neq(0,0)$, then $x^2+y^2\neq 0$, and $x+iy$ has a multiplicative inverse by the calculation:

$$(x,y)\left(\frac{x}{x^2+y^2}\,,\frac{-y}{x^2+y^2}\right)=(1,0)\,.$$

Notice that, if $z=x+iy$, then $z\overline{z}=x^2+y^2$ (that is, the purely real complex number x^2+y^2). Thus $\overline{z}/(x^2+y^2)$ (if suitably defined) *ought* to be the inverse of z, and the calculation above confirms this.

Let me review the facts so far discussed. $\mathbb{C}$ *is* $\mathbb{R}^2$, with the usual vector addition and a suitable multiplication; these operations satisfy the rules of arithmetic (they are both commutative and associative, multiplication is distributive over addition, there are an additive zero and a multiplicative identity, and non-zero elements of $\mathbb{C}$ have multiplicative inverses). Thus $\mathbb{C}$ is a *field*. Complex conjugation respects addition and multiplication (and inverses). The complex numbers that are unchanged by conjugation—the purely real complex numbers—also form a field (a subfield of $\mathbb{C}$) isomorphic to the real numbers, and they are customarily written as if they were real. These are the essential purely algebraic properties of $\mathbb{C}$.

It is natural to conjecture that something similar could be done with $\mathbb{R}^3$ or $\mathbb{R}^4$, or with $\mathbb{R}^n$ in general. This is not true; only $\mathbb{R}^2$ can be furnished with a multiplication turning it into a field with $\mathbb{R}$ as a subfield. (If we abandon commutativity of the multiplication, there is also a suitable multiplication in $\mathbb{R}^4$, making the quaternions, which are a 'skew field' or 'division ring'. If we even abandon associativity, there is a multiplication, with inverses, in $\mathbb{R}^8$, giving the (misnamed) 'Cayley numbers' or octonions. But that is as far as it goes if you want inverses, although a final proof, of amazing subtlety, was only given by Adams in 1958. It since been simplified somewhat, though it remains quite difficult.)

Remark 10.7. Since $\mathbb{C}$ is, except for its multiplication, just $\mathbb{R}^2$, it is natural for us to think of a complex number $z = (x, y)$ or $z = x + iy$ as a point in the Cartesian plane. For the historical reasons summarized at 10.5, this idea was not originally so obvious, and is therefore referred to as 'representing z on the Argand diagram'. As so often happens in such cases, Argand was not in fact the first person to invent it (just as the Cayley numbers were not first discovered by Cayley, but by Graves, who was a friend of Hamilton, a little earlier). He published the idea in 1806, but Wessel had already written about it in 1797, and Gauss, who was always reluctant to publish, and was 20 at the time, knew about it then too. The distance from the origin of the point representing z on the Argand diagram is $\sqrt{x^2 + y^2}$ (recall that the square root sign applied to a nonnegative real number denotes the nonnegative square root), which is called the *modulus* of z and written $|z|$. As $z\overline{z}$ is purely real and may be regarded as identical to the real number $x^2 + y^2$, we can write $|z| = \sqrt{z\overline{z}}$, the non-negative real square root of the non-negative real number $z\overline{z}$.

As $|x| \leq \sqrt{x^2 + y^2} \geq |y|$, where $||$ denotes the absolute value of the real numbers x and y, one has $|\Re z| \leq |z|$ and $|\Im z| \leq |z|$. For a purely real complex number, the modulus agrees with the absolute value, so that the use of the same notation $||$ for both causes no practical difficulties.

The modulus has the following properties, for any $w, z \in \mathbb{C}$.

(a) $|z| = 0$ if and only if $z = 0$.

Since $|z|$ is just the distance from 0 of the point represented by z, this is geometrically obvious. But one can also argue that, if $z \neq 0$, necessarily $\overline{z} \neq 0$ and $\overline{z}$ has an inverse, and so $z\overline{z} \neq 0$ (for otherwise multiplication by the inverse of $\overline{z}$ would give $z = 0$).

(b) $|wz| = |w||z|$.

For $|wz|^2 = wz(wz)^- = wz\overline{wz} = w\overline{w}z\overline{z} = |w|^2\,|z|^2 = (|w||z|)^2$; take non-negative square roots. If $z = x + iy$, $w = u + iv$, the equality thus proved between the squares of the moduli yields the remarkable identity

$$(x^2 + y^2)(u^2 + v^2) = (xu - yv)^2 + (xv + yu)^2\,,$$

which has a striking application: the product of two integers each of which is the sum of two squares is also the sum of two squares.

(c) $|w + z| \leq |w| + |z|$.

This is the triangle inequality for the length in $\mathbb{R}^2$. We can prove it thus:

$$\begin{aligned} |w + z|^2 &= (w + z)(w + z)^- = (w + z)(\overline{w} + \overline{z}) \\ &= w\overline{w} + z\overline{z} + (w\overline{z} + \overline{w}z) \\ &= |w|^2 + |z|^2 + 2\Re(w\overline{z}) \leq |w|^2 + |z|^2 + 2|w\overline{z}| \\ &= |w|^2 + |z|^2 + 2|w||z| = (|w| + |z|)^2\,. \end{aligned}$$

The result follows by taking non-negative square roots. (This argument is related to familiar proofs of the triangle inequality in $\mathbb{R}^n$. The scalar product $(u, v) \cdot (x, y)$, when $z := (x, y)$ and $w := (u, v)$ are regarded as vectors in $\mathbb{R}^2$, is just $\Re(w\overline{z})$.)

For later use, let me note the following version of the triangle inequality:

(d) $|w - z| \geq \big||w| - |z|\big|$.

The outer, longer bars on the right-hand side represent the absolute value in $\mathbb{R}$.

Remark 10.8. Let W be the point on the Argand diagram representing the complex number w, and let Z represent z. Then the vector $\overrightarrow{WZ}$ is represented by $z-w$, and the distance WZ is the length of this vector, $|z-w|$.

In the Euclidean geometry of $\mathbb{R}^2$, there is also a notion of *angle*. This leads to the idea of the *argument*, *amplitude*, or *phase* of a non-zero complex number. (Mathematicians usually call it the argument.) If $z \neq 0$, Z is the corresponding point on the Argand diagram, and O the origin, then an argument or amplitude of z is one of the angles through which the positive x-axis may be rotated in the *anticlockwise* direction to coincide with OZ. This concept makes sense only if $z \neq 0$ (for you can't have an angle between two lines in the plane when one of them reduces to the single point $(0,0)$). But a non-zero complex number has many arguments.

If one possible argument is $\frac{2}{3}\pi$ (more or less), but Ox may be rotated through $2\pi+\frac{2}{3}\pi$, $4\pi+\frac{2}{3}\pi$, $-6\pi+\frac{2}{3}\pi$, etc., with the same result. The fact that the argument of z is not uniquely defined is involved in many profound results we meet later.

Although $\arg z$ is not uniquely defined, one often writes $\arg z$ to denote whichever argument of z is indicated by the context. In particular, the *principal* argument of $z \neq 0$ is that argument θ of z which satisfies $-\pi < \theta \leq \pi$. The principal argument of z has strictly smaller absolute value than any other (unless it is π, when $-\pi$ has the same absolute value). We may denote it by $\operatorname{Arg} z$.

[A function, by the definition nowadays accepted, has only one value at each point of its domain. It is tempting to say that arg is a 'many-valued function', and Victorian mathematicians were happy to do so, but we must now be careful not to use such phrases in serious mathematical reasoning. I have been scrupulous not to describe $\arg z$ as a function of z. By contrast, $\operatorname{Arg} z$ *is* a function of z, for $z \neq 0$. It is not a *continuous* function, since, for any small positive number ϵ, $\operatorname{Arg}(-1+\epsilon i)$ is close to π and $\operatorname{Arg}(-1-\epsilon i)$ is close to $-\pi$.

One way to describe arg legitimately as a function would be to define it as a function whose values are sets of numbers rather than numbers themselves.]

Remark 10.9. Suppose θ is an argument of the non-zero complex number $z = x+iy$. Let $|z| = r$. Then $x = r\cos\theta$ and $y = r\sin\theta$, so that, together, $z = r(\cos\theta + i\sin\theta)$; this is the *polar form* of the complex number. For convenience, we may abbreviate $\cos\theta + i\sin\theta$ to $\operatorname{cis}\theta$. (Some authors prefer to write $\exp(i\theta)$ or $e^{i\theta}$; this accords with 9.19). Notice that

$$|\operatorname{cis}\theta| = |\exp(i\theta)| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$

automatically, so that, from 10.7, $|r\operatorname{cis}\theta| = r$ (provided that $r \geq 0$).

Similarly, let $w = \rho\operatorname{cis}\phi$. Then (de Moivre's theorem)

$$\begin{aligned} wz &= (\rho\cos\phi + i\rho\sin\phi)(r\cos\theta + ir\sin\theta) \\ &= \rho r(\cos\phi\cos\theta - \sin\phi\sin\theta) + i\rho r(\cos\phi\sin\theta + \sin\phi\cos\theta) \\ &= \rho r(\cos(\phi+\theta) + i\sin(\phi+\theta)) = \rho r\operatorname{cis}(\phi+\theta), \end{aligned}$$

using the addition formulæ for cos and sin (again see 9.19), and the definition (53) of multiplication of complex numbers. This again shows that $|wz| = \rho r = |w||z|$, as at 10.7, but also that $\phi+\theta$ is an argument for wz when ϕ is an argument for w and θ for z. Crudely speaking, multiplication of complex numbers multiplies their moduli and adds their arguments,

although care must be taken in interpreting this statement because the argument is not uniquely defined.

If $z = r\,\mathrm{cis}\,\theta$, then $z^2 = r^2\,\mathrm{cis}(2\theta) = w \neq 0$ if and only if $r^2 = |w| > 0$ and 2θ is an argument of w. That is, $2\theta = \mathrm{Arg}\,w + 2k\pi$ for some $k \in \mathbb{N}$. As r is nonnegative by definition, $r = \sqrt{|w|}$, and $\mathrm{cis}\,\theta = \mathrm{cis}(\frac{1}{2}\mathrm{Arg}\,w + k\pi)$ for some $k \in \mathbb{N}$. Both cos and sin reverse sign when the variable increases by π, so that $\mathrm{cis}\,\theta = \pm\,\mathrm{cis}(\frac{1}{2}\mathrm{Arg}\,w)$. In short, there are exactly two square roots of $w \neq 0$, $\pm\sqrt{|w|}\,\mathrm{cis}(\frac{1}{2}\mathrm{Arg}\,w)$. (Of course the complex number 0 has only one square root, 0 itself.)

Remark 10.10. Now consider the non-zero complex numbers w and z as displacement vectors in $\mathbb{R}^2$ (starting at the origin). Then an argument of w/z measures the angle from the vector z to the vector w, taken anticlockwise, and the modulus of w/z is the ratio of the magnitude $|w|$ of the vector w to the magnitude $|z|$ of the vector z.

That finishes our survey of basic facts about complex numbers.

Remark 10.11. The convention, as above, is that angles are measured anticlockwise; that is, they may be described as positive or negative, the anticlockwise sense being positive. This is, of course, merely a convention about the pictorial representation of the complex plane. In 'classical' Euclidean geometry, triangles PQR and ABC may be similar although corresponding angles in the two triangles have opposite senses; the circuit PQR may go clockwise although ABC goes anticlockwise. But it is perfectly possible to set up a plane geometry which distinguishes between 'similarity' and 'antisimilarity'.

Remark 10.12. There are two important cautionary remarks.

(i) We cannot make $\mathbb{C}$ into an ordered field as we did for $\mathbb{R}$. Whether i were positive or negative, i^2 would have to be positive; so i^3 would be positive [*or* negative] if i were positive [*or* negative]. But $i^3 = -i$. This does not mean that $\mathbb{C}$ cannot be "ordered" at all, only that any order that we give it will lack the useful algebraic properties of 1.8.

(ii) For non-negative real numbers x, $\sqrt{x}$ or $x^{1/2}$ is usually understood as denoting the non-negative square root of x. This convention ensures that $\sqrt{x_1x_2} = \sqrt{x_1}\sqrt{x_2}$. There is no corresponding systematic way to select one of the two square roots of an arbitrary non-zero complex number z. The definition in $\mathbb{R}$ obviously uses the order in $\mathbb{R}$. As with *(i)*, the difficulty in $\mathbb{C}$ is not in making a selection in the first place—one could, for instance, always take the square root with an argument in the interval $[0, \pi)$—but in doing it for all z whilst preserving the relation $\sqrt{wz} = \sqrt{w}\sqrt{z}$. To preserve it, we should need

$$1 = \sqrt{1}\sqrt{1} = \sqrt{1 \times 1} = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = -1\,,$$

which is absurd. Thus, if we ever want to use the symbol $\sqrt{z}$ or a square root function, we shall have to state explicitly which square root we have in mind, and to take care about its properties. There is nothing mysterious about this; it is the inevitable result of passing from the ordered field $\mathbb{R}$ to the field $\mathbb{C}$ that lacks an "algebraic" ordering.

In particular, -1 has two square roots, i and $-i$. We have given the name i to $(0, 1)$, but there is no way *intrinsic to* $\mathbb{C}$ of distinguishing between i and $-i$. (This is a slightly less pompous way of saying that complex conjugation is an automorphism of $\mathbb{C}$, so that any purely algebraic property of i must be shared by $-i$).

§11. Functions in $\mathbb{R}^2$.

For the time being, I shall mostly talk about $\mathbb{R}^2$ rather than $\mathbb{C}$, but, where necessary, I write $|(x,y)|$ to mean $\sqrt{x^2+y^2}$ (the length of the vector (x,y) in $\mathbb{R}^2$).

Our discussion will mostly be restricted to functions defined on open sets of $\mathbb{R}^2$. Vaguely speaking, this is because we want to be able to study a function's directional derivatives in all directions at each point of its domain. It is possible to generalize the domains used, but there is remarkably little advantage in doing so at this stage.

Definition 11.1. A *path* in $\mathbb{R}^2$ is a continuous function $\gamma : [a,b] \longrightarrow \mathbb{R}^2$, where $[a,b]$ is a closed interval in $\mathbb{R}$ with $a < b$. The path is said to begin (or start) at the point $\gamma(a)$ and end (or finish) at the point $\gamma(b)$. If $\gamma(a) = \gamma(b)$, the path is called *closed*. (As so often, it would be preferable to use a different word, but it is too late).

For the path $\gamma : [a,b] \longrightarrow \mathbb{R}^2$, there is an equivalent way of describing its continuity. For each $t \in [a,b]$, $\gamma(t)$ is a pair $\begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix}$ in $\mathbb{R}^2$. Here γ_1 and γ_2 are real-valued, and the vector in the codomain $\mathbb{R}^2$ is written as a column vector. Continuity of γ is then equivalent to the continuity of both γ_1 and γ_2 as real-valued functions of t.

Definition 11.2. Let E be a subset of $\mathbb{R}^2$. E is said to be *path-connected* if, for any two points $x, y \in E$, there is a path $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ which begins at x, ends at y, and takes all its values in E. (It is evident that one may assume $a = 0$ and $b = 1$.)

A path-connected open subset of $\mathbb{R}^2$ is called a *region* of $\mathbb{R}^2$, or of $\mathbb{C}$, or of 'the plane'. (Another word in common use to describe a region of $\mathbb{R}^2$ is "domain", but it is undesirably ambiguous).

There are other sorts of 'connectedness', but we are only interested in open subsets of $\mathbb{R}^2$, for which they are all equivalent (though not obviously so) to path-connectedness.

The various ideas of convergence of complex-valued sequences, continuity of complex-valued functions, and so on, are special cases of the concepts for metric spaces (see §§20, 24). For instance, if U is an open subset of $\mathbb{C}$, we can define continuity of a function $f : U \longrightarrow \mathbb{C}$ at a point $z_0 \in U$. All the properties you expect hold; their proofs are word for word as in the real case. Thus, if $f : U \longrightarrow \mathbb{C}$ and $g : U \longrightarrow \mathbb{C}$ are both continuous at $z_0 \in \mathbb{C}$, then so are their product $h = fg$ (defined by $h(z) = f(z)g(z)$ for each $z \in U$) and their sum $f + g$, similarly defined pointwise.

Definition 11.3. Let U be an open subset of $\mathbb{R}^2$; suppose $g : U \longrightarrow \mathbb{R}^2$. We can regard g as a pair of functions $u, v : U \longrightarrow \mathbb{R}$, such that, for each $(x,y) \in U$,

$$g(x,y) = (u(x,y), v(x,y)).$$

Given $(x_0, y_0) \in U$ and $i = 1, 2$, consider the difference quotient of u in the first variable, $\Delta_{11}(x_0, y_0, h) = \dfrac{u(x_0+h, y_0) - u(x_0, y_0)}{h}$. Since U is open, there exists some $\delta > 0$ (varying with the choice of $(x_0, y_0) \in U$) such that $(u,v) \in U$ whenever $|(x_0, y_0) - (u,v)| < \delta$. Hence $\Delta_{11}(x_0, y_0, h)$ makes sense, provided that the real number h satisfies $0 < |h| < \delta$. If $\Delta_{11}(x_0, y_0, h)$ has a limit as $h \to 0$, we say that u is differentiable with respect to the first variable, customarily called 'x', at (x_0, y_0), and the value of the limit

is the partial derivative at (x_0, y_0) with respect to the first variable, $\dfrac{\partial u}{\partial x}(x_0, y_0)$. In a similar way one defines differentiability of u at (x_0, y_0) with respect to the second variable 'y', and the corresponding partial derivative $\dfrac{\partial u}{\partial y}(x_0, y_0)$, starting from the difference quotient

$$\Delta_{12}(x_0, y_0, h) = \frac{u(x_0, y_0 + h) - u(x_0, y_0)}{h}\,.$$

The same applies to v. The function g, which takes values in $\mathbb{R}^2$, may therefore have as many as four numerical derivatives at (x_0, y_0), namely the "first" and "second" partial derivatives of u and of v. If all four of them are defined (each of u and v being differentiable in both x and y), the matrix

$$\begin{pmatrix} \dfrac{\partial u}{\partial x}(x_0, y_0) & \dfrac{\partial u}{\partial y}(x_0, y_0) \\ \dfrac{\partial v}{\partial x}(x_0, y_0) & \dfrac{\partial v}{\partial y}(x_0, y_0) \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix},$$

which is an acceptable abbreviation when there is no doubt which point (x_0, y_0) is meant, is called the *Jacobian matrix* of g at the point (x_0, y_0). Write it as $Dg(x_0, y_0)$.

The existence at (x_0, y_0) of all four numerical derivatives of g need not tell us much about the behaviour of g at nearby points. A function which has the value 0 on the coordinate axes has all four derivatives at the origin, and they are all 0; but its values off the axes may be completely arbitrary. This example emphasizes the obvious fact that partial derivatives involve the behaviour of the function only in the directions of the coordinate axes. This objection was met by Stolz in 1887 by defining differentiability (in the sense of a vector-valued function of two real variables) of g at (x_0, y_0):

g is differentiable at (x_0, y_0) if there is a 2×2 matrix $M := (m_{ij})$ such that

$$\frac{\left| g(x_0 + h, y_0 + k) - g(x_0, y_0) - M\begin{pmatrix} h \\ k \end{pmatrix} \right|}{|(h, k)|} \to 0 \tag{55}$$

as $|(h, k)| \to 0$. (The term $M\begin{pmatrix} h \\ k \end{pmatrix}$ signifies the matrix product). The quotient makes sense as long as $(h, k) \neq (0, 0)$, and, of course, the assertion that it tends to 0 as $|(h, k)| \to 0$ means that, for any $\epsilon > 0$, there exists $\delta > 0$ such that, whenever $0 < |(h, k)| < \delta$, the quotient is less than ϵ.

If this definition is satisfied, all four partial derivatives must be defined at (x_0, y_0), and M, the so-called Fréchet derivative, must be the Jacobian matrix $Dg(x_0, y_0)$. For example, consider what happens as $(h, 0) \to (0, 0)$; then (55) becomes

$$\frac{1}{|h|} \left| \begin{pmatrix} u(x_0 + h, y_0) - u(x_0, y_0) - m_{11}h \\ v(x_0 + h, y_0) - v(x_0, y_0) - m_{21}h \end{pmatrix} \right| \to 0\,,$$

whose upper entry tells us that $(u(x_0 + h, y_0 + k) - u(x_0, y_0))/h - m_{11} \to 0$ as $h \to 0$, and therefore that u is differentiable with respect to x at (x_0, y_0) and $m_{11} = \partial u/\partial x$, whilst the lower entry says likewise that $m_{21} = \partial v/\partial x$. Similarly for derivatives with respect to y.

This notion of 'differentiability' restricts the behaviour of g as (x, y) tends towards (x_0, y_0) in any fashion, not just along straight lines parallel to the coordinate axes, and it

generalizes one version of the definition of differentiability for a real-valued function of a real variable—namely, that $\dfrac{|f(x_0+h)-f(x_0)-f'(x_0)h|}{|h|} \to 0$ as $|h| \to 0$. (The rôle of the derivative $f'(x_0)$ is taken by the Jacobian matrix). The disadvantage is that the definition is often not easy to check in practical cases, but fortunately checking is rarely necessary, because of the standard result that follows. Its proof is an easy application of the mean-value theorem, and I omit it.

Theorem 11.4. *Suppose U is an open subset of $\mathbb{R}^2$, that $g : U \longrightarrow \mathbb{R}^2$ as above, and that the partial derivatives $\partial u/\partial x$, $\partial v/\partial x$, $\partial u/\partial y$, and $\partial v/\partial y$ exist at each point $(x_0, y_0) \in U$. If all four partial derivatives are continuous (as functions of two real variables) at $(x_1, y_1) \in U$, then g is differentiable at (x_1, y_1).* □

Definition 11.5. If U is open in $\mathbb{R}^2$ and $g : U \longrightarrow \mathbb{R}^2$, and all the four partial derivatives exist at each point of U (so that each of them defines a function $U \to \mathbb{R}$) and are all continuous, as real-valued functions of two real variables, on U, then g is said to be C^1 (more precisely, of differentiability class C^1) on U. More generally, g would be C^n if all the nth. partial derivatives, including all the mixed ones, were defined at all points of U and were continuous; and C^∞ if it were C^n for all positive integers n. Thus 11.4 tells us that a C^1 function is differentiable at each point of U.

I shall wish later to recall Green's theorem, which applies to C^1 functions.

§12. Complex-differentiability.

Definition 12.1. Let U be an open set in $\mathbb{C}$, and $g : U \longrightarrow \mathbb{C}$. Let $z \in U$. Then g is *complex-differentiable* at z if the *complex* difference quotient $\dfrac{g(z+w)-g(z)}{w}$ has a limit as $w \to 0$ through complex values. The limit, when it exists, is the (complex) *derivative* $g'(z)$ of g at z. Of course g' is a complex-valued function on the set of points at which g is complex-differentiable. As in the real case, one can write dg/dz instead of $g'(z)$.

Since U is open, there exists $r > 0$ such that $z + w \in U$ whenever $|w| < r$. Thus the difference quotient exists when $0 < |w| < r$, and the definition is an obvious extension of the definition of real-differentiability: the complex derivative at z exists and its value is the complex number q if

$$(\forall \epsilon > 0)(\exists \delta > 0) \quad 0 < |w| < \delta \Longrightarrow \left| \frac{g(z+w)-g(z)}{w} - q \right| < \epsilon .$$

The crucial ingredient of this definition is that $\mathbb{C}$, like $\mathbb{R}$, is a field (algebraically speaking)—so that the "difference quotient" is defined in $\mathbb{C}$,—and also has a topology (indeed a metric, $d(z_1, z_2) := |z_1 - z_2|$) which allows us to speak of limits.

Remark 12.2. Complex-differentiability has the properties you expect, and the proofs are exactly as in the real case. Thus, if $f, g : U \longrightarrow \mathbb{C}$ are both complex-differentiable at z_0, so are $f + g$ and fg, and

$$(f+g)'(z_0) = f'(z_0) + g'(z_0)\,,$$
$$(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0)\,.$$

If V is a second open set in $\mathbb{C}$ and $f(U) \subseteq V$, and $h : V \longrightarrow \mathbb{C}$ is complex-differentiable at $f(z_0)$, then the composition $h \circ f : U \longrightarrow \mathbb{C}$ is complex-differentiable at z_0 and $(h \circ f)'(z_0) = h'(f(z_0))f'(z_0)$.

The question that naturally arises is whether complex-differentiability is a common property or not. So it is appropriate to begin by accumulating some easy examples.

Example 12.3. Let $f : \mathbb{C} \longrightarrow \mathbb{C}$ be constant: $f(z) = a$ for all z. Then $f'(z)$ exists, and is 0, for all z. Indeed, $(f(z+w) - f(z))/w = 0$ for all non-zero w.

Example 12.4. If $f(z) = z$ for all z, then $(f(z+w) - f(z))/w = 1$ for all non-zero w. So f is complex-differentiable everywhere, with derivative 1.

Example 12.5. Let $f : \mathbb{C} \setminus \{0\} \longrightarrow \mathbb{C} : z \mapsto z^{-1}$.

If $z \neq 0$ and $0 < |w| < |z|$, then $w + z \neq 0$ too, and

$$\frac{f(z+w) - f(z)}{w} = \frac{1}{w}\left(\frac{1}{z+w} - \frac{1}{z}\right)$$
$$= \frac{-w}{wz(z+w)} = \frac{-1}{z(z+w)} \to \frac{-1}{z^2}$$

as $w \to 0$; that is, $f'(z) = -z^{-2}$. (Notice that w is to tend to 0 whilst $|w| < |z|$.)

Example 12.6. Consider $f_n : \mathbb{C} \longrightarrow \mathbb{C} : z \mapsto z^n$, where $n \in \mathbb{N}$. When $n = 1$, 12.4 shows that f_n is complex-differentiable everywhere with derivative 1. Suppose that, for any $k \geq 1$ and any z, f_k is complex-differentiable at z and $f_k'(z) = kz^{k-1}$. As $f_{k+1}(z) = f_k(z)f_1(z)$ for all z, the product rule (see 12.2) gives

$$f_{k+1}'(z) = f_k'(z)f_1(z) + f_k(z)f_1'(z)$$
$$= kz^{k-1}.z + z^k.1 = (k+1)z^k\,.$$

Hence, by induction, f_n is complex-differentiable everywhere for $n \in \mathbb{N}$, and $f_n' = nf_n$. In other words, $d(z^n)/dz = nz^{n-1}$.

Hence, any complex polynomial function is complex-differentiable everywhere. Of course we have already proved much more—see 9.14; functions represented by power series are also complex-differentiable within their radius of convergence.

Example 12.7. Consider $f : \mathbb{C} \longrightarrow \mathbb{C} : z \mapsto \overline{z}$. The difference quotient is

$$\frac{f(z+w) - f(z)}{w} = \frac{\overline{w}}{w}$$

(for $w \neq 0$). However, $\overline{w}/w$ has no limit as $w \to 0$. When w is real, $\overline{w}/w = 1$, but when w

is pure imaginary $\overline{w}/w = -1$. There can, therefore, be no limit as $w \to 0$ through all nonzero complex values.

Remark 12.8. Definition 12.1 may be expressed differently. g is complex-differentiable at z_0 if there is a complex number $q = a + ib$ such that

$$\frac{|g(z_0 + w) - g(z_0) - qw|}{|w|} \to 0 \quad \text{as} \quad w \to 0\,.$$

(This is just a complicated way of saying that $(g(z_0 + w) - g(w))/w \to q$). If we translate into the language of $\mathbb{R}^2$, with $z_0 = (x_0, y_0)$, $w = (h, k)$, and so on, this is

$$\frac{|g(x_0 + h, y_0 + k) - g(x_0, y_0) - q(h, k)|}{|(h, k)|} \to 0 \tag{56}$$

as $w \to 0$.

Writing $q(h, k) = (a + ib)(h + ik)$ as a column vector,

$$\begin{pmatrix} ah - bk \\ ak + bh \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

(recall the definition of multiplication of complex numbers!) and comparing the definition (55) of differentiability of vector-valued functions of two variables discussed in 11.3 with (56), one sees that they are identical except that the matrix in (56) has to be of a special form. We can put this as follows: g is complex-differentiable at z_0 if it is differentiable as a vector-valued function of two real variables and if at the same time its Jacobian matrix is of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Since the entries are the partial derivatives, this condition on the Jacobian matrix says that at (x_0, y_0)

$$\begin{aligned} &\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}\,, \quad \text{which is } a; \\ &\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}\,, \quad \text{which is } -b\,. \end{aligned} \tag{57}$$

The conditions (57) are called the Cauchy-Riemann equations. They say that the Jacobian matrix at (x_0, y_0) (a 2×2 real matrix, representing a real linear mapping in $\mathbb{R}^2$) acts as a complex-linear mapping in $\mathbb{C}$ (in effect a 1×1 complex matrix) when $\mathbb{C}$ is identified with $\mathbb{R}^2$ in the usual way. So we can state:

Theorem 12.9. *The function $g : U \longrightarrow \mathbb{C}$ is complex-differentiable at $z_0 = (x_0, y_0) \in U$ if and only if it is differentiable as a vector-valued function of two real variables at (x_0, y_0) and satisfies the Cauchy-Riemann equations there. Furthermore, in that case*

$$f'(z_0) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} + i\frac{\partial v}{\partial x}\,. \tag{58}$$

Proof. See 12.8. The derivative is $q = a + ib$, where $a = \partial u/\partial x$ and $b = \partial v/\partial x$, as in (57). □

Complex-differentiability at a single point is not very interesting.

Definition 12.10. Let U be an open subset of $\mathbb{C}$, and $g : U \longrightarrow \mathbb{C}$. Say that g is *holomorphic* on U if it is complex-differentiable at each point of U. (Occasionally one says f is *holomorphic at* z when it is holomorphic on some open set containing z).

g is *strongly holomorphic* on U if it is holomorphic on U and the complex derivative g' is a continuous function on U.

"Strong holomorphicity" is *not* a standard notion, for a very good reason. We shall see that, if g is holomorphic on U, it is necessarily strongly holomorphic on U, but this is *not obvious* (and will not be fully proved in the lectures; I shall distribute a handout to fill the gap). In the earliest investigations of complex analysis by Cauchy and others, the functions were tacitly assumed to be strongly holomorphic.

The word 'holomorphic' is nowadays the most commonly used in connection with these concepts, and we shall keep to it. However, its popularity is relatively recent; before about 1950, the word was usually 'analytic', which is still frequently used in some circles (in particular by applied mathematicians of rather traditional interests), and before 1940 'regular' was often used. (These words were not always treated as perfect synonyms. But the distinctions between them that some authors introduced need not concern us.) The reason for avoiding 'analytic' and 'regular' is not just fashion. They have many other meanings, whilst the word 'holomorphic' is *fairly* unambiguous. There is, however, the general moral: do not assume without evidence that words used by somebody else have exactly the meanings you expect.

Lemma 12.11. *g is strongly holomorphic on U if and only if it is* C^1 *on U as a vector-valued function of two real variables (see* 11.3*) and satisfies the Cauchy-Riemann equations* (57) *at each point of U.*

Proof. The C^1 condition means that the partial derivatives exist and are continuous at each point of U. By 11.4, this means that g is differentiable at each point as a vector-valued function of two real variables. By 12.9, g is complex-differentiable at each point, and g' is continuous (as $\partial u/\partial x$ and $\partial v/\partial x$ are). □

§13. More examples of holomorphic functions.

At 12.3–12.6 we saw that any complex polynomial function is holomorphic on the whole of $\mathbb{C}$, and that the function $(z-a)^{-1}$, for complex constant a, is holomorphic except at the point a. Using the chain rule, one deduces that any rational function $p_1(z)/p_2(z)$, where p_1, p_2 are polynomials in z (and p_2 is not identically zero), is holomorphic on the whole complex plane minus the finitely many points at which p_2 vanishes.

Example 13.1. From 12.6, $f : z \mapsto z^2$ is holomorphic on $\mathbb{C}$. Now

$$\begin{aligned} f(x+iy) = (x+iy)^2 &= (x^2-y^2) + 2ixy \\ &= u(x,y) + iv(x,y), \quad \text{where} \\ u(x,y) = x^2 - y^2, &\quad v(x,y) = 2xy. \end{aligned}$$

Both u and v have continuous partial derivatives of all orders (in fact their partial derivatives of total order 3 or more are identically zero); that is, they are C^∞ on $\mathbb{R}^2$. And

$$\frac{\partial u(x,y)}{\partial x} = 2x\,, \quad \frac{\partial u(x,y)}{\partial y} = -2y\,,$$
$$\frac{\partial v(x,y)}{\partial x} = 2y\,, \quad \frac{\partial v(x,y)}{\partial y} = 2x\,.$$

Thus the Cauchy-Riemann equations (57) are satisfied at every point of $\mathbb{C}$. This demonstrates again, this time from 12.9, that f is holomorphic, and indeed strongly holomorphic, on $\mathbb{C}$. Furthermore,

$$f'(z) = u_x + iv_x = 2(x+iy) = 2z\,,$$

as at 12.6.

Example 13.2. Define $f : \mathbb{C} \longrightarrow \mathbb{C}$ by

$$f(x+iy) = u(x,y) + iv(x,y)\,, \quad \text{where}$$

$$u(x,y) = \exp x \cos y\,, \quad v(x,y) = \exp x \sin y\,. \tag{59}$$

Both u and v are evidently C^∞ on $\mathbb{R}^2$; that is, all their partial derivatives of any order (mixed or not), exist and are continuous at all points of $\mathbb{R}^2$. And

$$\frac{\partial u(x,y)}{\partial x} = \exp x \cos y\,, \quad \frac{\partial u(x,y)}{\partial y} = -\exp x \sin y\,,$$
$$\frac{\partial v(x,y)}{\partial x} = \exp x \sin y\,, \quad \frac{\partial v(x,y)}{\partial y} = \exp x \cos y\,.$$

Thus the Cauchy-Riemann equations (57) are satisfied at every point of $\mathbb{C}$. The function f is consequently holomorphic, and indeed strongly holomorphic, on $\mathbb{C}$. As

$$f(z) = f(x+iy) = \exp x\,(\cos y + i \sin y) = \exp x \operatorname{cis} y\,, \tag{60}$$

f agrees on the real axis (that is, when $y = 0$) with the usual real exponential $\exp x$. If, as was historically the case, the exponential and trigonometrical functions had been defined *ad hoc*, it would be natural to call f the (complex) exponential function.

We have $f'(z) = u_x + iv_x = \exp x \operatorname{cis} y = f(z)$. Also, it is easy to see that

$$(\forall z, w \in \mathbb{C}) \quad f(w+z) = f(w).f(z)\,.$$

It follows that $f(z) \neq 0$ for any z, since it has a reciprocal $f(-z)$.

A noteworthy property of f, which the real exponential function does not prepare you for, is that the complex function is periodic with (pure imaginary) period $2\pi i$, since

$$\exp(z + 2\pi i) = \exp z \exp(2\pi i) = \exp z \operatorname{cis}(2\pi) = \exp z\,.$$

This is roughly how the complex exponential function was originally developed. I explained at 9.16 why it is not an entirely satisfactory procedure from the mathematical point of view, and at 9.17 et seqq. how one can use power series to define the complex exponential, trigonometric, and hyperbolic functions directly. 9.14 then shows that, on the power series definition, they are everywhere strongly holomorphic.

Example 13.3. Similarly, try now the function

$$g(x+iy) = \sin x \cosh y + i \cos x \sinh y\,. \tag{61}$$

Here $u(x,y) = \sin x \cosh y$, $v(x,y) = \cos x \sinh y$. They are C^∞ on $\mathbb{R}^2$, and

$$\frac{\partial u}{\partial x} = \cos x \cosh y\,, \qquad \frac{\partial u}{\partial y} = \sin x \sinh y\,,$$
$$\frac{\partial v}{\partial x} = -\sin x \sinh y\,, \qquad \frac{\partial v}{\partial y} = \cos x \cosh y\,.$$

The Cauchy-Riemann equations are satisfied, and g is holomorphic on $\mathbb{C}$. Since, once again, g agrees on the real axis with the real sine function, it may be called the (complex) sine function, and written $\sin z$. We have

$$\begin{aligned} g'(z) = \frac{d}{dz}(\sin z) &= u_x + iv_x \\ &= \cos x \cosh y - i \sin x \sinh y\,. \end{aligned} \tag{62}$$

This function is written $\cos z$, and clearly agrees on the real axis with the familiar cosine function for a real variable. It is easy to check, by the same means as for sin, that it is holomorphic on $\mathbb{C}$. (This, and other properties, are left as exercises).

Note 13.4. Although the identity $\sin^2 z + \cos^2 z = 1$ remains true for the complex functions, we cannot deduce that the complex-valued functions $\sin z$ and $\cos z$ are bounded (in modulus) for all $z \in \mathbb{C}$. Indeed, if $z = \frac{1}{2}\pi + iy$, $\sin z = \cosh y \to \infty$ as $y \to \pm\infty$. It is only for real x that $|\sin x| \leq 1$ and so on.

Example 13.5. A like procedure may be used for the "historical" approach to the hyperbolic functions. We define, by analogy with the real case,

$$\begin{aligned} \cosh(x+iy) &= \tfrac{1}{2}(\exp(x+iy) + \exp(-x-iy)) \\ &= \cosh x \cos y + i \sinh x \sin y\,, \end{aligned} \tag{63}$$

from the formulæ (59), and likewise

$$\begin{aligned} \sinh(x+iy) &= \tfrac{1}{2}(\exp(x+iy) - \exp(-x-iy)) \\ &= \sinh x \cos y + i \cosh x \sin y\,. \end{aligned} \tag{64}$$

As $\exp z$ and $\exp(-z)$ are holomorphic on $\mathbb{C}$, so are their sum and difference; thus cosh and sinh are holomorphic. (We do not need to check the Cauchy-Riemann equations again). Now, on these definitions,

$$\begin{aligned}
\cosh(iz) &= \cosh(-y+ix) \\
&= \cosh(-y)\cos x + i\sinh(-y)\sin x \\
&= \cosh y\cos x - i\sinh y\sin x \\
&= \cos z\,, \qquad \text{by (63) and (62), and} \\
\sinh(iz) &= \sinh(-y+ix) \\
&= \sinh(-y)\cos x + i\cosh(-y)\sin x \\
&= i(\cosh y\sin x + i\sinh y\cos x) \\
&= i\sin z\,, \qquad \text{by (64) and (61).}
\end{aligned}$$

(When z is real, the formulæ $\cosh(ix) = \cos x$, $\sinh(ix) = i\sin x$ result immediately from (63) and (64), or from (61) and (62).)

These identities explain the analogy between the real trigonometric and hyperbolic functions; they are the same functions except for factors of i, and that is why identities relating trigonometric functions translate to similar formulæ for hyperbolic functions, with a change of sign if i^2 appears. Some of these identities appear in the exercises. But, of course, the approach of 9.16 *et seqq.* is conceptually preferable.

Remark 13.6. In each of the above cases, we have constructed a function strongly holomorphic on $\mathbb{C}$ which coincides on the real axis with a well-known function ϕ of a real variable. You might wonder whether this can "always" be done (for "reasonable" functions on the real line). We shall see later that it cannot; see 17.8.

To do it in general, we should have to find real-valued functions $u(x,y)$ and $v(x,y)$ such that $u(x,0) = \phi(x)$ and $u(x,y) + iv(x,y)$ is holomorphic. If we assume u is C^2 (we see later that it must be), the Cauchy-Riemann equations $u_x = v_y$, $u_y = -v_x$ show that v, which must be C^1 (as $u+iv$ is holomorphic), is in fact C^2 (and vice versa), and that

$$u_{xx} = v_{yx} = v_{xy} = -u_{yy}\,, \quad v_{xx} = -u_{yx} = -u_{xy} = -v_{yy}\,.$$

Thus the real-valued functions of two variables $u(x,y)$ and $v(x,y)$ both satisfy Laplace's equation in two dimensions, which for a function h is

$$\Delta h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0\,.$$

(Some people prefer to define the Laplace operator Δ with the opposite sign, for very good reasons that we need not discuss here.)

Definition 13.7. Let U be an open set in $\mathbb{R}^2$. A function $h : U \longrightarrow \mathbb{R}$ is *harmonic* on U if it is C^2 on U and satisfies Laplace's equation at each point of U.

The sum of two harmonic functions is harmonic. Thus we have:

Lemma 13.8. *If the real part of a holomorphic function $f : U \longrightarrow \mathbb{C}$ is C^2 on U, the imaginary part must also be C^2 on U, and both the real and imaginary parts are harmonic functions on U. The same applies if the imaginary part of f is C^2.* □

This suggests that holomorphicity is not as similar to real-differentiability as its original definition might have indicated. The obvious question is whether every harmonic C^2 function on U is the real part of a suitable holomorphic function on U.

Suppose $u(x,y)$ is a harmonic C^2 function on $\mathbb{R}^2$. If it is the real part of a holomorphic function $u(x,y)+iv(x,y)$, the Cauchy-Riemann equations must be satisfied:

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}, \qquad \text{at every point of } \mathbb{R}^2.$$

If $x=x_0$ is fixed, the first equation says $\dfrac{\partial v}{\partial y}(x_0,y)=\dfrac{\partial u}{\partial x}(x_0,y)$, which is a known C^1 function of y, as u is given. We can integrate with respect to y:

$$v(x_0,y_0)-v(x_0,0)=\int_0^{y_0}\frac{\partial u}{\partial x}(x_0,\eta)\,d\eta. \tag{65}$$

On the other hand,

$$v(x_0,0)-v(0,0)=\int_0^{x_0}\left(-\frac{\partial u}{\partial y}(\xi,0)\right)d\xi \tag{66}$$

from the second Cauchy-Riemann equation. Thus u determines v completely, except for the choice of $v(0,0)$; and (65) ensures that $\partial v/\partial y=\partial u/\partial x$. (A small change in y_0 only affects (65)). However, it is not clear that $\partial v/\partial x=-\partial u/\partial y$, so all we have done so far is find a small class of possible solutions of the problem (differing from each other by constants); we have not shown that any of them really is a solution. Furthermore, we have made no use of the harmonicity of u.

Recall the material from MATH 206 about integrating vector fields along paths. (I shall have more to say about this later). In effect, we have found $v(x_0,y_0)$ by integrating the vector field

$$X(x,y)=(-u_y(x,y),u_x(x,y))$$

around two sides of the rectangle which has $(0,0)$ and (x_0,y_0) as opposite vertices. The first side was from $(0,0)$ to $(x_0,0)$, where, as the path travels along the x-axis, the integrand reduces to $-u_y$, and the second was from $(x_0,0)$ to (x_0,y_0), giving the integrand u_x. This path of integration may be called p, and the integral of X along p may be written as I. But we could have gone round the other way, from $(0,0)$ to $(0,y_0)$ first and then to (x_0,y_0); call this path q, and the corresponding integral J. If we integrate along the path 'first p, and then the reverse of q' (which we could write as $p*\overline{q}$), the result is $I-J$. But thus $I-J$ is the integral of X round the perimeter of the rectangle. (It might be the integral either clockwise or widdershins, depending on the values of x_0 and y_0, and that means there is an uncertainty of sign below).

Now apply Green's theorem. X is a C^1 vector field on $\mathbb{R}^2$, and consequently its integral around the perimeter of the rectangle is equal to plus or minus the integral of its curl over the rectangle. But

$$\nabla\times X(x,y)=(0,0,-u_{yy}-u_{xx}),$$

which is 0 by the harmonicity of u. Thus X has zero integral round the rectangle, whether the integral is taken clockwise or anticlockwise: $I-J=0$, $I=J$. In other words, instead of calculating $v(x_0,y_0)$ from (65) and (66), we could use the other half-circuit of integration,

$$v(x_0,y_0)=\int_0^{y_0}\frac{\partial u}{\partial x}(0,\eta)\,d\eta-\int_0^{x_0}\frac{\partial u}{\partial y}(\xi,y_0)\,d\xi+v(0,0) \tag{67}$$

and get the same answer as before. But (67) shows that $\partial v/\partial x = -\partial u/\partial y$. (A change in x_0 affects only the second integral). This proves that

Proposition 13.9. *Any* C^2 *harmonic function* u *on* $\mathbb{R}^2$ *is the real part of a holomorphic function* $f := u + iv$ *on* $\mathbb{C}$. *Any two candidates for* f *differ only by a purely imaginary constant; that is, we may choose the real number* $v(0,0)$ *arbitrarily, but once we have done so* f *is completely fixed.* □

Notice, though, that our proof worked in $\mathbb{R}^2$, not in a general open set U, because we needed to integrate round a rectangle and to know that everything inside and on the rectangle is in U.

§14. Harmonic functions.

Example 14.1. Define $u(x,y) = \dfrac{x}{x^2+y^2}$, for $(x,y) \neq (0,0)$. Then

$$\frac{\partial u}{\partial x} = \frac{y^2-x^2}{(x^2+y^2)^2} \quad \text{and} \quad \frac{\partial u}{\partial y} = \frac{-2xy}{(x^2+y^2)^2},$$
$$\frac{\partial^2 u}{\partial x^2} = \frac{2x(x^2-3y^2)}{(x^2+y^2)^3} \quad \text{and} \quad \frac{\partial^2 u}{\partial y^2} = \frac{2x(-x^2+3y^2)}{(x^2+y^2)^3}.$$

So u is a harmonic function on $\mathbb{R}^2 \setminus \{(0,0)\}$. As it is undefined at $(0,0)$, the proof of 13.9 cannot be applied without change. Recall (65) and (66). We find

$$v(x_0,y_0) - v(x_0,0) = \int_0^{y_0} \frac{\eta^2 - x_0^2}{(x_0^2+\eta^2)^2}\, d\eta = \frac{-y_0}{x_0^2+y_0^2},$$

and, if $x_0 > 0$,

$$v(x_0,0) - v(1,0) = \int_0^{x_0} -\frac{\partial u}{\partial y}(\xi,0)\, d\xi = 0$$

(since $\partial u/\partial y = 0$ on the x-axis). Thus $v(x,y) = \dfrac{-y}{x^2+y^2} + v(1,0)$ when $x > 0$, and $v(x,y) = \dfrac{-y}{x^2+y^2} + v(-1,0)$ when $x < 0$ by the same argument; on either half-plane, $-y/(x^2+y^2)$ is a candidate for the imaginary part v of a holomorphic function whose real part is u, and it is in fact defined everywhere except the origin. (Other candidates defined on $\mathbb{C} \setminus \{0\}$ would be obtained by setting $v(1,0) = v(-1,0)$ to be any real value we choose.) It is also possible to argue likewise for the upper or lower half-planes. Since $u+iv$ is $(x-iy)/(x^2+y^2) = z^{-1}$, which is certainly holomorphic on $\mathbb{C} \setminus \{0\}$, we need not verify that the second Cauchy-Riemann equation is satisfied.

Definition 14.2. If u, v are harmonic functions on a region U of $\mathbb{R}^2$, v is said to be *conjugate* to u on U if $u+iv$ is holomorphic on U.

As $-i(u+iv) = v - iu$, it follows that $-u$ is conjugate to v. 13.9 shows that any harmonic function on the whole of $\mathbb{R}^2$ has a conjugate on the whole of $\mathbb{R}^2$, whilst in 14.1 we

had an example of a harmonic function on $\mathbb{R}^2 \setminus \{(0,0)\}$ which also has a conjugate on $\mathbb{R}^2 \setminus \{(0,0)\}$.

Example 14.3. Let $r = \sqrt{x^2+y^2}$. If $(x,y) \neq (0,0)$, $\partial r/\partial x = x/r$ and $\partial r/\partial y = y/r$. Let $\phi(r)$ be a C^2 function of r alone (a "radially symmetric" function). Then $\partial\phi/\partial x = \phi'(r)\,x/r$ by the chain rule and, by the chain and product rules,

$$\frac{\partial^2\phi}{\partial x^2} = \phi''(r)\Big(\frac{x}{r}\Big)^2 + \frac{\phi'(r)}{r} - \phi'(r)\frac{x^2}{r^3} \quad \text{and likewise for } y;$$
$$\Delta\phi = \phi''(r) + 2r^{-1}\phi'(r) - r^{-1}\phi'(r), \quad \text{as } x^2+y^2 = r^2.$$

So ϕ will be harmonic on $\mathbb{R}^2 \setminus \{(0,0)\}$ if and only if $\phi'' + r^{-1}\phi' = 0$. Rewriting this (it is a linear first-order equation in ϕ' as a function of r) as $(r\phi')' = 0$, one finds that $r\phi' = A$ a constant, and $\phi = A\log r + B$. Thus the harmonic functions on $\mathbb{R}^2 \setminus \{(0,0)\}$ that depend only on r are necessarily of this form. (Again I recall that I use log to denote the *natural* logarithm—not that it really matters here.)

For Laplace's equation, dimension 2 is rather special, since no power of the radius r is harmonic; the reason why it is special is that r^{-1} does not have an antiderivative which is a power of r. In dimension $n > 2$, the same argument shows that r^{2-n} is a solution of Laplace's equation on $\mathbb{R}^n \setminus \{0\}$ (the so-called *fundamental solution*).

Example 14.4. The function $\phi = \log r$ is now harmonic on $\mathbb{R}^2 \setminus \{(0,0)\}$. Let us try to find a conjugate harmonic function ψ. It must satisfy

$$\textit{(i)} \quad \frac{\partial\psi}{\partial x} = -\frac{\partial\phi}{\partial y} = -\frac{y}{x^2+y^2}, \qquad \textit{(ii)} \quad \frac{\partial\psi}{\partial y} = \frac{\partial\phi}{\partial x} = \frac{x}{x^2+y^2}.$$

Holding x constant non-zero and integrating *(ii)*, we obtain

$$\psi(x,y) = \tan^{-1}(y/x) + \alpha(x) + C;$$

substituting back in *(i)*, $\alpha'(x) = 0$. Thus in fact ψ must be (for $x > 0$) $\tan^{-1}(y/x) + C_+$, for some constant C_+. For $x < 0$, it is $\tan^{-1}(y/x) + C_-$, for some constant C_-. On each of the left and right half-planes, ψ differs by a constant from the polar coordinate θ.

For $y > 0$, equation *(i)* integrates to $\psi(x,y) = -\tan^{-1}(x/y) + D_+$; and, on the lower half-plane ($y < 0$), $\psi(x,y) = -\tan^{-1}(x/y) + D_-$. However, when x and y are both non-zero, $\tan^{-1}(y/x) + \tan^{-1}(x/y) = \frac{1}{2}\pi$. So, again, ψ differs from θ by a constant on the lower half-plane.

Note 14.5. Thus, if we let θ denote the principal argument of w and r its modulus, the function $f(w) = \log r + i\theta$ is holomorphic on the right half-plane, on the upper half-plane, and on the lower half-plane; but on the non-positive real axis θ is discontinuous; it jumps from $-\pi$ to π as w moves on to the axis from just below.

Example 14.6. Applying (60) of 13.2, with $z = x + iy$, $w = r\,\mathrm{cis}\,\theta$,

$$\exp f(w) = \exp(\log r)\,\mathrm{cis}\,\theta = r\,\mathrm{cis}\,\theta = w, \tag{68}$$

whilst, conversely,

$$\begin{aligned} f(\exp z) &= \log(|\exp z|) + i\,\mathrm{Arg}(\exp z) \\ &= \log(\exp x) + i\,\mathrm{Arg}(\mathrm{cis}\, y) = x + iy = z\,. \end{aligned} \tag{69}$$

So $f(w)$, defined except at $w = 0$, is a function inverse to exp. Since $\exp z$ can never be 0, as we remarked at 13.2, it is not surprising that the inverse function is not defined at 0. It is natural to call f the *principal value* of the logarithm; 'logarithm', because it extends the ordinary logarithmic function (defined only for positive real numbers) and is an inverse to the exponential, and 'principal' because it involves the principal value of the argument. It may be written as Log (or Ln). I emphasize again that it is holomorphic only on the plane minus the nonpositive real axis.

Remark 14.7. There are other inverses to the exponential function. For instance, given any integer n, $\mathrm{Log}\, z + 2ni\pi$ is also inverse to exp, by essentially the same calculation as at (68) and (69). One could also take any ray from the origin (with polar equation $\theta = \theta_0$) and set $\log_1 z = \log|z| + i\theta(z) + 2ni\pi$, where $\theta(z)$ is that argument of z which takes values in $(\theta_0 - 2\pi, \theta_0]$. By this choice, $\log_1 z$ is discontinuous at the points of the ray $\theta = \theta_0$ instead of the points of the negative real axis. The nature of the discontinuities is the same, however; as one leaves the ray anticlockwise, $\log_1 z$ instantly jumps by $-2\pi i$. A minor adjustment would be to choose $\theta(z)$ to take values in $[\theta_0 - 2\pi, \theta_0)$ instead, and then the jump would occur on arriving at the ray rather than on leaving it. Still more generally, one could do something similar with any curve which goes steadily outwards from the origin (for instance a spiral) and 'cuts the plane' so that one cannot go round the origin without crossing the curve.

The important thing is that there is no harmonic function conjugate to $\log r$ and defined on the *whole* of $\mathbb{R}^2 \setminus \{(0,0)\}$; it must "jump" somewhere (and so be discontinuous). The argument just given is already in principle a messy proof of this statement, and there will be a far snappier one soon. The thing that goes wrong, in any case, is that it is impossible to choose an argument continuously for all $z \neq 0$. Any attempt at such a choice fails because it leads inevitably to an increase of 2π after a circuit of the origin. On the other hand, there is a harmonic function conjugate to $\log r$ on any region in which it is impossible to encircle the origin—because such a choice of argument is then possible. But that has not yet been proved.

Definition 14.8. Let U be a region in $\mathbb{C}$, and $f : U \longrightarrow \mathbb{C}$ a holomorphic function such that $\exp f(z) = z$ for all $z \in U$. Then f is called a *branch of the logarithm* in U.

So, if $u : U \longrightarrow \mathbb{R}$ is harmonic, where U is a region in $\mathbb{R}^2$, there may or may not be a conjugate harmonic function defined on U. When U is the whole of $\mathbb{R}^2$, there always is, by 13.9, but 14.4 shows that there exist functions u and regions U for which a conjugate function does not exist on the whole of U. A specific such region is the 'punctured plane' $\mathbb{C} \setminus \{0\}$, and a specific such function is $\log|z|$.

The argument of 13.9 depended only on integrating around the perimeter and over the interior of certain rectangles. It will work without alteration for any region U, provided that there is some fixed point (x_1, y_1) of U such that, for any other point (x_2, y_2) of U, the whole rectangle with corners (x_1, y_1), (x_2, y_1), (x_2, y_2), (x_1, y_2) lies inside U (together with its interior). This will be true if U is a circular disk, and for some other simple figures like coordinate rectangles. So we can say, for instance,

Theorem 14.9. *Let U be a circular disk in $\mathbb{R}^2$. Then any harmonic function u defined on U has a conjugate harmonic function v also defined on U, and this conjugate function is unique except for the possible addition of a constant.* □

Thus any harmonic function u on a region V has a conjugate harmonic function 'locally'—for any point $z \in V$, there is a conjugate harmonic function to u defined on a neighbourhood of z. (Any disk included in V and containing z would do).

Remark 14.10. Let $f = u + iv : V \longrightarrow \mathbb{C}$ have partial derivatives

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}.$$

Then we can define the "formal derivatives"

$$\partial f = \frac{\partial f}{\partial z} := \frac{1}{2}\left\{\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right\}, \quad \overline{\partial} f = \frac{\partial f}{\partial \overline{z}} := \frac{1}{2}\left\{\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right\}.$$

$\partial f/\partial x$ and $\partial f/\partial y$ are in principle *complex* numbers, so that ∂f and $\overline{\partial} f$ are not, in general, conjugate to each other. The Cauchy-Riemann equations are exactly equivalent to $\overline{\partial} f = 0$, and if they are satisfied then $f' = \partial f/\partial x = \partial f + \overline{\partial} f = \partial f$. In this symbolic sense, ∂f can indeed be thought of as the "partial derivative with respect to z" and $\overline{\partial} f$ as the "partial derivative with respect to $\overline{z}$", but, of course, these phrases lack a literal meaning, since z and $\overline{z}$ are not independent variables.

Let $u : U \longrightarrow \mathbb{R}$ be harmonic. A second way of associating a holomorphic function with u, instead of trying to find a conjugate harmonic function, is to take

$$g(z) = \partial u = \frac{1}{2}\left(\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\right).$$

Since u is harmonic, it is easy to check that g satisfies the Cauchy-Riemann equations $\overline{\partial} g = 0$ and so is holomorphic on U. ∂u is always holomorphic if u is harmonic, but u may not have a harmonic conjugate. Indeed, ∂u has a complex antiderivative if and only if u has a harmonic conjugate. Suppose that

$$h(x, y) := p(x, y) + iq(x, y)$$

were a complex antiderivative of ∂u; then h would be holomorphic, with (recall (58)) $\partial p/\partial x = \frac{1}{2}\partial u/\partial x$, $\partial p/\partial y = -\partial q/\partial x = \frac{1}{2}\partial u/\partial y$; thus p would differ from $\frac{1}{2}u$ by a constant, and $2q$ would be a harmonic function conjugate to u. So the search for complex antiderivatives of ∂u is essentially the same as the search for harmonic conjugates of u.

§15. Complex integration.

Since a complex-valued function of a complex variable $f = u + iv$ is a pair of real-valued functions of two real variables, one could integrate each of u and v with respect to two-dimensional area, taking double integrals as in 206. This idea is interesting for some purposes, but it cannot be closely related to the idea of complex differentiation, since it makes no use of the complex numbers as such. The most useful way to integrate f, which leads to the amazing results I mentioned at the outset, is along 'contours'. In effect, our integrals will be taken over one-dimensional objects in $\mathbb{R}^2$.

Definition 15.1. Let $\gamma : [a, b] \longrightarrow \mathbb{C}$ be a path in $\mathbb{C}$ (recall 11.1), where $a < b$. We say that γ is C^1 (continuously differentiable) on a subinterval $[c, d]$ of $[a, b]$ if its both its real part γ_1 and its imaginary part γ_2 are C^1 on $[c, d]$. (Recall from 7.1 what this means.) Then γ is

piecewise C^1 on $[a,b]$ if there exist points $t_0, t_1, t_2, \ldots, t_k \in [a,b]$ such that

$$a = t_0 < t_1 < t_2 < \cdots < t_k = b$$

and γ_1, γ_2 are both C^1 on each of the subintervals $[t_{i-1}, t_i]$ for $1 \leq i \leq k$.

Notice that, at the subdivision points $t_1, \ldots, t_{k-1}$, γ_1 and γ_2 have both left and right derivatives, but these one-sided derivatives are not necessarily equal. We take $\gamma'(t)$ to mean $\gamma_1'(t) + i\gamma_2'(t)$ (and likewise for the one-sided derivatives).

A piecewise C^1 path may be called a *contour*.

Definition 15.2. Suppose that U is a region in $\mathbb{C}$, $\gamma : [a,b] \longrightarrow U$ is a C^1 path, and $f : U \longrightarrow \mathbb{C}$ is a continuous function. We shall define the integral of f along γ, $\int_\gamma f$. As I said above, γ is a 'one-dimensional' object; we are not talking about the integral of f over a two-dimensional region. The definition is concisely stated as

$$\int_\gamma f \coloneqq \int_a^b f(\gamma(t))\,\gamma'(t)\,dt\,. \tag{70}$$

This formula requires explanation. For $t \in [a,b]$, $\gamma(t) \in U$, $\gamma'(t) \in \mathbb{C}$, and $f(\gamma(t)) \in \mathbb{C}$. Multiply these complex numbers: let $g(t) \coloneqq f(\gamma(t))\,\gamma'(t)$. Both γ' and $f \circ \gamma$ are continuous on $[a,b]$, so their product g is a continuous complex-valued function of t. Hence, the right-hand side of (70) is the integral of a complex-valued continuous function g of the real variable t.

Such an integral is conveniently defined by separating its real and imaginary parts:

$$\int_a^b g(t)\,dt \coloneqq \int_a^b \Re(g(t))\,dt + i\int_a^b \Im(g(t))\,dt\,.$$

Thus, if $f = u + iv$ as usual, and $\gamma(t) = \gamma_1(t) + i\gamma_2(t)$, the memorable and convenient formula (70) is interpreted as meaning a more complicated expression in terms of 'ordinary' Riemann integrals of continuous real-valued functions of a real variable:

$$\begin{aligned}\int_\gamma f \coloneqq \int_a^b &\{u(\gamma_1(t),\gamma_2(t))\gamma_1'(t) - v(\gamma_1(t),\gamma_2(t))\gamma_2'(t)\}\,dt \\ &+ i\int_a^b \{u(\gamma_1(t),\gamma_2(t))\gamma_2'(t) + v(\gamma_1(t),\gamma_2(t))\gamma_1'(t)\}\,dt\,.\end{aligned} \tag{71}$$

Notice also that $\int_\gamma f$ does not "depend on the parametrization of γ", in the following sense. If $c < d$ in $\mathbb{R}$ and $\phi : [c,d] \longrightarrow [a,b]$ is C^1 (as a mapping into $\mathbb{R}$; thus it has a right derivative at c and a left derivative at d, and so on) and $\phi(c) = a$, $\phi(d) = b$, then $\gamma_1 \coloneqq \gamma \circ \phi$ is a C^1 path parametrized by $[c,d]$, and

$$\int_c^d f(\gamma_1(t))\,\gamma_1'(t)\,dt = \int_a^b f(\gamma(t))\,\gamma'(t)\,dt\,.$$

This follows directly from the usual formula for change of variable in the integral of a real-valued function of a real variable, by way of (71) [check this]. But it means that, vaguely speaking, it is the "picture of γ" that matters—the direction it is traversed in and the "track" it leaves, rather than the rate at which we travel along it.

Remark 15.3. If $p : [a, b] \longrightarrow \mathbb{C}$ and $q : [b, c] \longrightarrow \mathbb{C}$ are paths, and if $p(b) = q(b)$, they may be concatenated to give a path $p*q : [a, c] \longrightarrow \mathbb{C}$. This idea has already been used, in a casual way, in the proof of 13.9. We define $p*q(t)$ to be $p(t)$ when $a \leq t \leq b$, and $q(t)$ when $b \leq t \leq c$; $p*q$ is then clearly continuous (because $p(b) = q(b)$), and is the path which 'goes first along p and then along q'.

$p*q$ is only defined when the end of p coincides with the beginning of q. On the other hand, concatenation is "associative as far as possible", in the sense that, if $p*q$ and $q*r$ are both defined, then so are $(p*q)*r$ and $p*(q*r)$, and they are equal. A multiple concatenation such as $p_1*p_2*p_3*p_4$ will make unambiguous sense without brackets, provided that the binary concatenations $p_1*p_2, p_2*p_3, p_3*p_4$ are all defined.

Unfortunately, $p*q$ will not usually be C^1 when p and q both are, for $p'(b)$ (a left derivative) and $q'(b)$ (a right derivative) need not agree. This is why we introduced *piecewise* C^1 paths. The concatenation of two C^1 paths is piecewise C^1, and the concatenation of two contours is always a contour.

Definition 15.4. Suppose that U is a region in $\mathbb{C}$, that γ is a contour in U, and that $f : U \longrightarrow \mathbb{C}$ is continuous. Then $\int_\gamma f$ may be defined as follows.

Choose C^1 paths $p_1, p_2, \ldots, p_k$ so that

$$\gamma = p_1*p_2*\ldots*p_k ,$$

which is certainly possible, by the definition of a contour. Now set tentatively

$$\int_\gamma f := \sum_{i=1}^{k} \int_{p_i} f . \tag{72}$$

This is not at first a legitimate definition, because it depends on the choice of $p_1, p_2, \ldots, p_k$. But if piecewise C^1 paths $q_1, q_2, \ldots, q_l$ are also such that $\gamma = q_1*q_2*\ldots*q_l$, then there is a third way of splitting γ (by taking all the subdivision points of its domain $[a, b]$ that appeared as end-points of subintervals that were domains of any p_i or q_j), and the integrals $\int_{p_i} f$ and $\int_{q_j} f$ are sums of the integrals over these smaller subintervals of $[a, b]$; it follows that

$$\sum_{j=1}^{l} \int_{q_j} f := \sum_{i=1}^{k} \int_{p_i} f .$$

The choices involved in (72) do not, therefore, affect the proposed value of $\int_\gamma f$. The formulæ (70) and (71) may, in fact, still be used to define $\int_\gamma f$, but they now involve—if we continue to use the Riemann integral—"formally improper" integrals, because there may be points (only finitely many, however) at which $\gamma'(t)$ (the *two-sided* derivative at t) is undefined. (This problem vanishes if we use the Henstock integral.) I take this for granted henceforth.

A contour $\gamma : [a, b] \longrightarrow \mathbb{C}$ is described as *closed* if $\gamma(a) = \gamma(b)$. (Compare 11.1).

Remark 15.5. The contour integrals just defined have reasonable properties. Specifically, if γ is a contour in the region U, and f and g are both continuous on U, then $\int_\gamma (f + g) = \int_\gamma f + \int_\gamma g$ and $\int_\gamma cf = c\int_\gamma f$ for any complex constant c. Also, when the contours γ_1 and γ_2 may be concatenated, $\int_{\gamma_1*\gamma_2} f = \int_{\gamma_1} f + \int_{\gamma_2} f$.

Recall that, if g is a real-valued function of a real variable t for $a \le t \le b$, there is the 'fundamental estimate' for its integral, $\left|\int_a^b g(t)\,dt\right| \le \int_a^b |g(t)|\,dt$, provided that the integrals exist.

Lemma 15.6. *Suppose that* $f : U \longrightarrow \mathbb{C}$ *is continuous on the region* U *and* $\gamma : [a,b] \longrightarrow U$ *is a contour. Then*

$$\left|\int_\gamma f\right| \le \int_a^b |f(\gamma(t))|\,|\gamma'(t)|\,dt\,. \tag{73}$$

Proof. Choose an argument θ for $\int_\gamma f$. Then

$$\left|\int_\gamma f\right| = \exp(-i\theta)\int_\gamma f = \int_\gamma \exp(-i\theta) f \qquad \text{by 15.5.}$$

Write $f_1 = \exp(-i\theta)f = u_1 + iv_1$, so that $\int_\gamma f_1$ is real and non-negative. In (71), only the real part of the integral of f_1 can survive:

$$\begin{aligned}\left|\int_\gamma f\right| &= \int_a^b \{u_1(\gamma_1(t),\gamma_2(t))\gamma_1'(t) - v_1(\gamma_1(t),\gamma_2(t))\gamma_2'(t)\}\,dt \\ &= \int_a^b \Re\{f_1(\gamma(t))\,\gamma'(t)\}\,dt \le \int_a^b |f_1(\gamma(t))|\,|\gamma'(t)|\,dt\end{aligned}$$

(since, for any complex number w, $\Re w \le |w|$)

$$= \int_a^b |f(\gamma(t))|\,|\gamma'(t)|\,dt\,.$$

This completes the proof. □

Note again that the integral on the right-hand side of (73) is, strictly speaking, an improper Riemann integral, since $\gamma'(t)$ may be undefined at certain points of $[a,b]$. But the impropriety is of the simplest possible kind.

Remark 15.7. The integral

$$\int_a^b |\gamma'(t)|\,dt = \int_a^b \sqrt{(\gamma_1'(t))^2 + (\gamma_2'(t))^2}\,dt$$

(this too exists as an improper Riemann integral) defines the length of the contour in the usual geometrical sense in $\mathbb{R}^2$. Call this length L. If the modulus of f is bounded by M on the contour (that is, $|f(\gamma(t)| \le M$ for $a \le t \le b$), then 15.6 tells us

$$\left|\int_\gamma f\right| \le ML\,. \tag{74}$$

In words: the modulus of the integral of f along a contour cannot exceed the product of the length of the contour and the supremum of the modulus of f on the contour. This is a useful estimate, but 15.6 can potentially yield more information.

Note 15.8. The contour integral $\int_\gamma f$ is more usually written $\int_\gamma f(z)\,dz$, and the right-hand side of (73) is correspondingly denoted $\int_\gamma |f(z)|\,|dz|$. This notation derives from a second sort of contour integral along γ of a continuous complex-valued function g. This is the 'unoriented contour integral' or 'integral with respect to contour length', written $\int_\gamma g(z)\,|dz|$. It is defined as $\int_a^b f(\gamma(t))\,|\gamma'(t)|\,dt$, with the usual conventions about subdivision points and real and imaginary parts. We shall not use it except as a convenient notation. (In (73), the integrand "g" was in fact $|f|$, which is real-valued.)

Example 15.9. *(a)* Suppose that g is a function strongly holomorphic in a region U, and let $\gamma : [a,b] \longrightarrow U$ be a continuously differentiable path. Then the complex derivative g' is continuous on U, so that $\int_\gamma g'$ is defined. But, by the definition (70),

$$\int_\gamma g' = \int_a^b g'(\gamma(t))\gamma'(t)\,dt = \int_a^b (g\circ\gamma)'(t)\,dt$$

(the chain rule holds for complex-valued functions of a complex variable!)

$$= (g\circ\gamma)(b) - (g\circ\gamma)(a) = g(\gamma(a)) - g(\gamma(a))\,.$$

If $g\circ\gamma = u + iv$, the integral is by definition $\int_a^b u'(t)\,dt + i\int_a^b v'(t)\,dt$, where both integrands are continuous, so it is

$$u(b) - u(a) + i(v(b) - v(a)) = g(\gamma(b)) - f(\gamma(a))\,.$$

Thus, the integral of the complex derivative of a strongly holomorphic function along a continuously differentiable path is just the difference in the values of the function at the ends. For the integral along a contour, as defined at 15.4, the same holds (by addition). This is precisely analogous to the situation for real-valued functions of a real variable:

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

if f is C^1 on $[a,b]$. However, contours are defined in two dimensions. There will be many contours in U that start at $\gamma(a)$ and end at $\gamma(b)$, and the integral will be the same, $f(\gamma(b)) - f(\gamma(a))$, over every one; only the ends matter.

(b) In particular, the integral of the complex derivative of a strongly holomorphic function over a *closed* contour must be 0.

Example 15.10. Take the function $f(z) = 1/z$, which is strongly holomorphic on $\mathbb{C}\setminus\{0\}$, and the closed contour $\gamma : [0, 2\pi] \longrightarrow \mathbb{C}\setminus\{0\}$ defined by

$$\gamma(t) := \exp(it) = \cos t + i\sin t \qquad \text{(recall (60))}.$$

As t goes from 0 to 2π, $\gamma(t)$ travels widdershins round the unit circle $\{z : |z| = 1\}$, starting and finishing at 1. We have

$$\gamma'(t) = i\exp(it) = -\sin t + i\cos t\,. \quad \text{Thus,}$$

$$\int_\gamma f = \int_0^{2\pi} \frac{1}{\exp(it)}\, i\exp(it)\,dt = 2\pi i\,. \tag{75}$$

It must be emphasized that this is just a calculation from the definition; there is no mystery about it. But it enables us to give an easy and convincing proof of the fact we found in 14.7, that there can be no branch of the logarithm defined on the whole of $\mathbb{C} \setminus \{0\}$.

If $g(z)$ were a holomorphic function on $\mathbb{C} \setminus \{0\}$ such that $\exp g(z) = z$, then, by the chain rule, $g'(z) \exp g(z) = 1$, $g'(z) = 1/z$; thus $1/z$ would be a complex derivative of the holomorphic function g, which would in fact be strongly holomorphic, as $1/z$ is continuous on $\mathbb{C} \setminus \{0\}$. By 15.9, the integral of $1/z$ round the closed contour γ would be 0. It is not, by (75), so there can be no holomorphic function $g(z)$; that is, there is no holomorphic branch of the logarithm on $\mathbb{C} \setminus \{0\}$. [In its essence this is the same proof as before, the difference being that integrating around a circle is much easier.]

Example 15.11. If we ask for the integral of $1/z$ round a *segment* of the unit circle, such as the contour $\gamma(t) = \exp(it)$ for $\theta_1 \le t \le \theta_2$, the only alteration in 15.10 is that the limits of integration in (75) become θ_1 and θ_2 (instead of 0 and 2π). The integral is therefore $i(\theta_2 - \theta_1)$.

Example 15.12. Let us study the integral of $f(z) = \dfrac{1}{z-a}$, which is holomorphic on $\mathbb{C} \setminus \{a\}$, around the circle $\gamma_r(t) = r \exp(it)$ for $0 \le t \le 2\pi$. I must assume that $r \ne |a|$, since otherwise γ does not map into $\mathbb{C} \setminus \{a\}$. The integral can be evaluated directly by elementary methods, but the details are rather painful (they are rather like 14.4), and it is easier and more informative to proceed as follows.

Suppose firstly that $|a| < r$, and define $f_s(z) \coloneqq (z - sa)^{-1} = F(s, z)$ for $0 \le s \le 1$. Then, for each such s,

$$I(s) \coloneqq \int_{\gamma_r} f_s$$

is defined, because $r > |sa|$ (which ensures that γ_r maps into $\mathbb{C} \setminus \{sa\}$, on which f_s is holomorphic). But, for such values of s,

$$I(s) = \int_0^{2\pi} \frac{ir \exp(it)\, dt}{r \exp(it) - sa}\,.$$

Here s appears as a parameter in the integral. Now I am going to cheat, by taking for granted what is in fact true but has not been proved, that $I(s)$ may be differentiated with respect to the parameter s and that

$$\begin{aligned} I'(s) &= \int_0^{2\pi} \frac{\partial}{\partial s} \left(\frac{ir \exp(it)}{r \exp(it) - sa} \right) dt \\ &= \int_0^{2\pi} \frac{ira \exp(it)\, dt}{(r \exp(it) - sa)^2} = a \int_\gamma g_s\,, \end{aligned}$$

where $g_s(z) = (z - sa)^{-2}$. (This procedure is called 'differentiating under the integral sign', which is a particular case of interchanging a limit with an integration; see 6.9. It is not always correct, and there are various theorems about its applicability. In particular, it suffices to know that the derivative with respect to the parameter s is continuous as a function of (s, t).) The effect of this calculation—which is valid in this case—is that

$$\frac{d}{ds}I(s) = \frac{d}{ds}\int_{\gamma_r} \frac{dz}{z-sa}$$
$$= \int_{\gamma_r} \frac{\partial}{\partial s}\left(\frac{1}{z-sa}\right)dz = \int_{\gamma_r} \frac{a\,dz}{(z-sa)^2}\,.$$

However, the last integral vanishes by 15.9, since γ_r is a closed contour and $(z-sa)^{-2}$ is the complex derivative with respect to z of $-(z-sa)^{-1}$. It follows that $I(s)$ does not change with s, for $0 \le s \le 1$, and therefore that

$$\int_{\gamma_r} \frac{dz}{z-a} = I(1) = I(0) = \int_{\gamma_r} \frac{dz}{z} = 2\pi i\,, \qquad \text{by 15.10.}$$

The condition $|a| < r$ says that the circle γ_r goes round a.

Next, suppose that $|a| > r$. Now the point a lies outside the circle γ_r. As before, $I(s)$ is defined for $s \ge 1$, because then γ_r maps into a region where f_s is holomorphic; and, as in the previous case, $I'(s) = 0$, so that $I(s)$ is constant for all $s \ge 1$. But now

$$\int_{\gamma_r} \frac{dz}{z-a} = I(1) = I(s) = \int_{\gamma_r} \frac{dz}{z-sa}\,. \tag{76}$$

For $s \ge 1$ and $0 \le t \le 2\pi$,

$$|\gamma_r(t) - sa| \ge |sa| - |\gamma(t)| = s|a| - r\,,$$

by the triangle inequality (10.7*(d)*). The length of γ_r is $2\pi r$. Thus, by the fundamental estimate (74) and (76), I have, for any $s \ge 1$,

$$\left|\int_{\gamma_r} \frac{dz}{z-a}\right| = \left|\int_{\gamma_r} \frac{dz}{z-sa}\right| \le \frac{2\pi r}{s|a| - r} \to 0 \qquad \text{as } s \to \infty\,.$$

Hence, $\displaystyle\int_{\gamma_r} \frac{dz}{z-a} = 0$ (otherwise the inequality fails for large enough s).

The conclusion is that $\displaystyle\int_{\gamma_r} \frac{dz}{z-a} = 0$ when a lies outside the circle traced out by γ_r, whilst $\displaystyle\int_{\gamma_r} \frac{dz}{z-a} = 2\pi i$ if a lies inside the circle.

If a lies *on* the circle γ_r, the integrand is undefined at a.

Now suppose that $\gamma(t) = r\exp(it)$ for $0 \le t \le 4\pi$, so that γ goes *twice* round the circle of radius r. The arguments are essentially unchanged, and we find in this case that the integral is $4\pi i$ if a is within the circle and (again) 0 if it is outside.

Example 15.13. Lastly, consider the integral around the same γ_r of the function

$$f(z) = \frac{2}{1-z^2} = \frac{1}{1-z} + \frac{1}{1+z}\,.$$

It is necessary to assume that $r \ne 1$, since f is undefined at 1 and at -1. We have, by 15.5,

$$\int_{\gamma_r} f = \int_{\gamma_r} \frac{dz}{1-z} + \int_{\gamma_r} \frac{dz}{1+z}\,,$$

and, by 15.12, both these integrals are zero when $r < 1$. (Then both -1 and 1 lie outside the circle). On the other hand, if $r > 1$, the circle encloses both 1 and -1, and

$$\int_{\gamma_r} \frac{dz}{1-z} = -\int_{\gamma_r} \frac{dz}{z-1} = -2\pi i\,,$$
$$\int_{\gamma_r} \frac{dz}{1+z} = \int_{\gamma_r} \frac{dz}{z-(-1)} = 2\pi i\,, \quad \int_{\gamma_r} f = 0\,.$$

It is of course possible to integrate round an anticlockwise circle which encloses (say) -1 but not 1; then the integral will take the value $2\pi i$; or round an anticlockwise circle enclosing 1 but not -1, when the integral is $-2\pi i$.

The use of circular contours in each of these examples is peculiarly convenient because $\gamma'(t)$ cancels with the denominator of f, but it is also possible to carry out explicit computations with some other figures. There will be some exercises on this. However, in practice it is rarely necessary to resort to lengthy direct calculation, as we shall see.

Remark 15.14. Let $\gamma : [a, a+m] \longrightarrow U$ be a *closed* contour, and suppose that $f : U \longrightarrow \mathbb{C}$ is continuous. Given $c \in (a, a+m)$, there is another closed contour $\gamma_c : [c, c+m] \longrightarrow U$, defined as follows.

$$\text{If } c \le t \le a+m\,,\ \gamma_c(t) = \gamma(t)\,;$$
$$\text{If } a+m \le t \le c+m\,,\ \gamma_c(t) = t-m\,.$$

The effect is that γ_c parametrizes the same loop as γ, and in the same direction, but it starts and ends at $\gamma(c) = \gamma_c(c) = \gamma_c(c+m)$ instead of $\gamma(a) = \gamma(a+m)$. This might be called 'cyclic reparametrization'. The easiest way to visualize it is by wrapping the interval $[a, a+m]$ once round a circle T, so that the points of the circle corresponding to a and to $a+m$ coincide; γ then determines a mapping of T into U, and γ_c describes the same mapping, but begins the description at a different point.

The significance of the reparametrization γ_c of the closed contour is that $\int_\gamma f = \int_{\gamma_c} f$. This is a trivial consequence of the definitions: the integral round γ is the sum of the integral over parameter values from a to c (which is equal to the γ_c-integral over parameter values from $a+m$ to $c+m$) and the integral over parameter values from c to $a+m$ (equal to the γ_c-integral for the same range of parameter values).

The parametrization of the circle is immaterial, because of the formula for change of variable in an integral in dimension 1. It is important only that $\gamma(t)$ goes round the loop the required number of times in the positive sense, i.e. anticlockwise. In 15.10, for instance, we could take $\exp(it^5/(2\pi)^4)$ instead. If the sense is reversed, the sign of the integral is reversed too. In intuitive terms, the integral is a 'function of the directed circle' only—that is, of the circle counted in the right direction and the right number of times. This is the justification of the notation $\int_\gamma f(z)\,dz$ already introduced, in which the actual 'parameter of integration' t does not appear at all, and of another, $\int_C f(z)\,dz$, in which C denotes the 'oriented curve' (curve with a sense) traced out by γ. Occasionally $-C$ is used to mean 'C with reversed sense', and then $\int_{-C} f(z)\,dz = -\int_C f(z)\,dz$. It is sometimes convenient to use the notation $\int_C f(z)\,dz$ even in cases where C is not a single curve, to mean the sum of the integrals over several curves making up a figure C. In the same spirit, we shall sometimes use such phrases as 'the points of the contour γ', meaning in more rigorous language the values taken by γ. It is not entirely easy to justify these verbal and notational usages, but I hope that they will seem natural in context.

§16. Cauchy's theorem.

The subject of this section is the most fundamental theorem of complex analysis, and leads rather quickly to the startling results I mentioned earlier. As with some other very basic results—for example the fundamental theorem of calculus—it is very easy to state Cauchy's theorem imprecisely, but rather difficult to formulate it in a way that accurately describes the various situations in which it is true. For this reason, you will find several different versions in textbooks. I shall take the course of presenting Cauchy's original version, and then discussing informally how it may be improved.

Definition 16.1. Let E be a subset of $\mathbb{C}$. The frontier of E in $\mathbb{C}$, which we may write as $\operatorname{Fr} E$, is the set of points z such that, for any $\delta > 0$, there is a point of $\mathbb{C} \setminus E$ and there is a point of E which are both within distance δ of z:

$$(\mathbb{C} \setminus E) \cap B(z;\delta) \neq \emptyset \neq E \cap B(z;\delta)\,.$$

[That is, z is an adherent point of both E and its complement, or is a point of the closure both of E and of its complement; cf. 21.1.]

When E is a "simple geometrical figure", the frontier of a set, as thus defined, is precisely what one would expect; for example, the frontier of the circular disk $\{z : |z - z_0| < r\}$ is the circle $\{z : |z - z_0| = r\}$.

Definition 16.2. A contour $\gamma : [a,b] \longrightarrow \mathbb{C}$ (this means, by 15.4, that γ is piecewise C^1) is a *Jordan contour* if it is closed (that is, if $\gamma(a) = \gamma(b)$) and if $\gamma(t_1) = \gamma(t_2)$ only when either $t_1 = t_2$ or $\{t_1, t_2\} = \{a, b\}$. (If $[a,b]$ is wrapped round a circle Γ and γ is regarded as a mapping $\Gamma \longrightarrow \mathbb{C}$, as in 15.14, then γ is Jordan if it is one-to-one as that mapping $\Gamma \longrightarrow \mathbb{C}$. In visual language, a Jordan contour is a deformed circle).

Theorem 16.3. (Cauchy's Theorem, Cauchy's version.) *Let $f : U \longrightarrow \mathbb{C}$ be strongly holomorphic, and let $\gamma : [a,b] \longrightarrow U$ be a Jordan contour such that $\gamma([a,b])$ is the frontier of a region V included in U. Then $\int_\gamma f = 0$.*

Compare 15.9*(b)*. What is true for *any* closed contour and the derivative of a strongly holomorphic function is also true for a Jordan contour and a function that is strongly holomorphic on a region including both the contour and the region "inside" it. The function may be more general, provided that the geometrical condition is stronger.

Remark 16.4. According to Green's theorem, if X is a C^1 vector field defined on an open set U in $\mathbb{R}^2$, and $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ is a Jordan contour whose image is the frontier of a region $V \subseteq U$, so oriented that V always lies to the left of γ (this is saying that V has a piecewise C^1 boundary, and that γ traverses it anticlockwise), then

$$\int_\gamma X(\gamma(t)) \cdot \gamma'(t)\, dt = \int_V (\operatorname{curl} X)\, d(x,y)$$

($d(x,y)$ here indicates the integral with respect to area). For this purpose $\operatorname{curl} X$ is regarded as a scalar (in three dimensions it would be a vector perpendicular to the (x,y)-plane, with its sense determined by a 'handedness rule'). The assertion of the theorem again shows that the parametrization of γ is irrelevant, and for this reason it is often expressed as

$$\int_{\partial V} (X_1\,dx + X_2\,dy) = \iint_V \left(\frac{\partial X_2}{\partial x} - \frac{\partial X_1}{\partial y}\right) dx\,dy\,.$$

Here ∂V is the 'positively oriented boundary of V', where V always lies to the left as one traverses the boundary; that is, one goes around V anticlockwise. Once a piecewise C^1 parametrization $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ of the boundary is fixed, the left-hand side means

$$\int_a^b (X_1(\gamma(t))\gamma_1'(t) + X_2(\gamma(t))\gamma_2'(t))\,dt\,. \tag{77}$$

This integral, of course, is calculated in the 'piecewise' sense, as in 15.4; it is the sum of integrals over subintervals on each of which $\gamma'(t)$ is continuous—one takes the left derivative at the right-hand end, etc. Now compare (77) with (71). Both the real and imaginary parts of $\int_\gamma f$ can be put in the form of integrals of vector fields along γ:

$$\int_\gamma f = \int_a^b \{X_1(\gamma(t))\gamma_1'(t) + X_2(\gamma(t))\gamma_2'(t)\}\,dt + i\int_a^b \{Y_1(\gamma(t))\gamma_1'(t) + Y_2(\gamma(t))\gamma_2'(t)\}\,dt\,,$$

if $X_1(x,y) = p(x,y) = Y_2(x,y)$ is the real part of f, and $X_2 = -q = -Y_1$ its imaginary part. By Green's theorem, therefore,

$$\int_\gamma f = \iint_V \left(-\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx\,dy + i\iint_V \left(\frac{\partial p}{\partial x} - \frac{\partial q}{\partial y}\right) dx\,dy$$

(writing the integral "with respect to area" as a repeated integral). The two Cauchy-Riemann equations show the two integrals to be 0. This establishes Cauchy's theorem when f is strongly holomorphic on U. Notice that f has to be C^1 as a function of two variables to ensure the validity of Green's theorem; once the contour integral has been transformed to an integral over V, the Cauchy-Riemann equations are all that is needed.

Note 16.5. Cauchy's theorem and its proof via Green's theorem, as just given, are due to Cauchy himself. There are reasons, explained at length in §D1 of Appendix D, for some dissatisfaction with it, and I shall now explain in merely intuitive terms the more satisfactory version that I shall assume for the rest of the course. We could easily make do with 16.3, provided that we assumed some "obvious" facts of plane topology, considered only strongly holomorphic functions, and repeatedly exploited some rather tedious geometrical tricks.

Definition 16.6. Suppose $\gamma : [a,b] \longrightarrow \mathbb{C}$ is a closed path, and $z \in \mathbb{C}$ is not a value of γ. The *index* of z with respect to γ, or the *winding number* of γ about z, written $n(\gamma, z)$, is the total number of times $\gamma(t)$, $a \le t \le b$, goes round z in the positive (anticlockwise) direction. As in 15.14, this does not depend on the parametrization of γ. Thus, if γ is a circle with centre z described once anticlockwise, $n(\gamma, z) = 1$; if it is described once clockwise, $n(\gamma, z) = -1$. If it goes round once clockwise, then twice anticlockwise, then once clockwise, $n(\gamma, z) = 0$. I take it for granted that we can determine $n(\gamma, z)$ in all the cases we shall need. (They will all be very simple).

If $n(\gamma, z) = 0$, γ *does not wind* about z; otherwise, γ *winds about* z.

Remark 16.7. Winding numbers have the property

$$n(\gamma_1 * \gamma_2, z) = n(\gamma_1, z) + n(\gamma_2, z)$$

whenever the statement makes sense, i.e. whenever γ_1, γ_2 have the same beginning and finishing points (so that $\gamma_1 * \gamma_2$ is defined and closed) and z is not a value of either. In particular, if $\gamma(t) = a \neq z$ for all $t \in [a, b]$, then $n(\gamma, z) = 0$.

Definition 16.8. Let U be a region in $\mathbb{C}$, and γ a closed path in U. We say that γ is *homologous to zero* in U, or *nullhomologous* in U, if, for any $z \notin U$, $n(\gamma, z) = 0$; that is, if γ does not wind about any point *outside* U.

This depends on the "position" of γ in U and the position of U in $\mathbb{C}$.

Theorem 16.9. (Cauchy's theorem, first "topological" version). *Let U be a region in $\mathbb{C}$, and $f : U \longrightarrow \mathbb{C}$ a holomorphic function. Then $\int_\gamma f = 0$ for any closed contour $\gamma : [a, b] \longrightarrow U$ which is nullhomologous in U.* □

Example 16.10. If $U = \mathbb{C}$, there are no points outside $\mathbb{C}$; consequently, any closed contour is nullhomologous in $\mathbb{C}$, and any function holomorphic on the whole of $\mathbb{C}$ has zero integral round any closed contour.

The formula $g(z) := \dfrac{2z}{1 - z^2} = \dfrac{1}{1 - z} - \dfrac{1}{1 + z}$ defines a function g that is holomorphic on $U := \mathbb{C} \setminus \{\pm 1\}$. It is easy to construct a contour γ (*not* a Jordan contour, for it must cross itself) that has winding number 0 about both $+1$ and -1. Then γ is nullhomologous in U, and the integral of $\dfrac{2z}{1 - z^2}$ around γ must be 0.

A circle about the origin of radius greater than 1 winds once about 1 and once about -1. As in 15.13, $g(z) = -\left(\dfrac{1}{z - 1} + \dfrac{1}{z + 1}\right)$ has integral $-2\pi i - 2\pi i = -4\pi i$ around such a circle.

Remark 16.11. The topological version 16.9 of Cauchy's theorem can be further generalized, following the hint in 15.14. Suppose that, instead of a single closed contour γ (corresponding to an 'oriented closed curve' C), we have a 'closed chain $\mathcal{C}$' (usually called a *cycle*) consisting of several oriented closed curves $C_1, C_2, \ldots, C_n$, not necessarily all different—that is, we are allowed to repeat them (finitely often) if we wish. Here each C_i can be parametrized as a closed contour γ_i. Provided that the point $z \in \mathbb{C}$ is not on any of the C_i, we can define $n(\mathcal{C}, z)$ to be $\sum_{i=1}^{n} n(\gamma_i, z)$. Then we say that $\mathcal{C}$ is nullhomologous in U if $n(\mathcal{C}, z) = 0$ for all $z \notin U$. The generalization of 16.9 is that, if f is holomorphic on U and $\mathcal{C}$ is nullhomologous in U, then

$$\int_{\mathcal{C}} f = \sum\nolimits_{i=1}^{n} \int_{\gamma_i} f = 0\,.$$

I sketch below the proof (from 16.9) of the particular case of this result where $\mathcal{C}$ consists of two closed contours γ_1 and $\overline{\gamma}_2$. Essentially the same argument gives the general case, but the particular case will be used repeatedly.

Theorem 16.12. *Suppose U is a region of $\mathbb{C}$ and $f : U \longrightarrow \mathbb{C}$ is holomorphic. Let γ_1 and γ_2 be closed contours in U such that $n(\gamma_1, z) = n(\gamma_2, z)$ for all $z \notin U$. Then*

$$\int_{\gamma_1} f = \int_{\gamma_2} f .$$

Proof. Since reparametrizations do not affect the integrals (see 15.2; but the reparametrizations here may be very inoffensive), I may assume that $\gamma_1 : [0,1] \longrightarrow U$ and $\gamma_2 : [0,1] \longrightarrow U$. Since U is connected, there is a contour $\gamma_3 : [1,2] \longrightarrow U$ joining $\gamma_1(1)$ to $\gamma_2(1)$: that is, $\gamma_1(1) = \gamma_3(1)$ and $\gamma_3(2) = \gamma_2(1)$. Then $\gamma = \gamma_1 * \gamma_3 * \overline{\gamma}_2 * \overline{\gamma}_3$, where $\overline{\gamma}_3 : [3,4] \longrightarrow U$ is γ_3 reversed and reparametrized ($\overline{\gamma}_3(t) = \gamma_3(5-t)$) and likewise $\overline{\gamma}_2 : [2,3] \longrightarrow U$ is defined by $\overline{\gamma}_2(t) = \gamma_2(3-t)$, is also a closed contour in U. Its winding number $n(\gamma, z)$ about any point z not on itself is $n(\gamma_1, z) - n(\gamma_2, z)$, because the contributions of γ_3 and $\overline{\gamma}_3$ to the winding about z cancel out. Hence γ is nullhomologous in U, so by 16.9

$$\int_{\gamma} f = 0 .$$

On the other hand, it is clear that

$$\int_{\gamma} f = \int_{\gamma_1} f + \int_{\gamma_3} f - \int_{\gamma_2} f - \int_{\gamma_3} f ,$$

and the result follows. □

Remark 16.13. Two cycles in U that have the same winding numbers about any point not in U are described as homologous (to each other) in U; thus, we have in effect shown that the integral of f is the same around either of two closed contours that are homologous in U.

§17. Consequences of Cauchy's theorem.

Theorem 17.1. (The Cauchy integral theorem). *Let $f : U \longrightarrow \mathbb{C}$ be holomorphic on the region U, and let $\gamma : [a,b] \longrightarrow U$ be a closed contour homologous to 0 in U. Then, for any $w \in U$ which is not on γ,*

$$n(\gamma, w) f(w) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)\, dz}{z - w} .$$

Proof. Suppose $n(\gamma, w) \geq 0$. Choose $r > 0$ so small that $B(w; r) \subseteq U$. Given any $\delta \in (0, r)$, let $\gamma[\delta](t) := w + \delta \exp(it)$ for $0 \leq t \leq 2\pi n(\gamma, w)$. If $z \notin U$, $n(\gamma, z) = 0$ because γ is nullhomologous in U, and $n(\gamma[\delta], z) = 0$ as $z \notin B(w; r)$. (A sketch will make this clear). But $n(\gamma[\delta], w) = n(\gamma, w)$ by the very construction of $\gamma[\delta]$. Hence γ and $\gamma[\delta]$ have the same winding numbers about any point not in $U \setminus \{w\}$. (Such a point is either not in U, or is w itself).

Now take $g(z) := \dfrac{f(z)}{z-w}$, which is not defined at w, but is clearly holomorphic on $U \setminus \{w\}$. Applying 16.12 to g and to $U \setminus \{w\}$ in the places of f and U,

$$\begin{aligned}\int_\gamma \frac{f(z)\,dz}{z-w} &= \int_{\gamma[\delta]} \frac{f(z)\,dz}{z-w} \\ &= \int_{\gamma[\delta]} \frac{f(w)\,dz}{z-w} + \int_{\gamma[\delta]} \frac{f(z)-f(w)}{z-w}\,dz\,.\end{aligned} \tag{78}$$

Here, however,

$$\int_{\gamma[\delta]} \frac{f(w)}{z-w}\,dz = \int_0^{2\pi n(\gamma,w)} f(w)\frac{i\delta\exp(it)}{\delta\exp(it)}\,dt = 2\pi i n(\gamma,w) f(w) \tag{79}$$

(the calculation, directly from the definition of the contour integral, is essentially the same as at 15.10). This is true for all values of $\delta \in (0,r)$, so (78) tells us that

$$\int_\gamma \frac{f(z)\,dz}{z-w} = 2\pi i n(\gamma,w) f(w) + \int_{\gamma[\delta]} \frac{f(z)-f(w)}{z-w}\,dz\,. \tag{80}$$

Now, f is complex-differentiable at w: $\dfrac{f(z)-f(w)}{z-w} \to f'(w)$ as $z \to w$. Hence, there is some real number q, $0 < q \le r$, such that

$$\begin{aligned} 0 < |z-w| < q \implies & \left|\frac{f(z)-f(w)}{z-w} - f'(w)\right| < 1 \\ \implies & \left|\frac{f(z)-f(w)}{z-w}\right| < K := 1 + |f'(w)|\,. \end{aligned}$$

The length of $\gamma[\delta]$ is $2\pi\delta\, n(\gamma,w)$. So, by the fundamental estimate (74),

$$\kappa := \left|\int_{\gamma[\delta]} \frac{f(z)-f(w)}{z-w}\,dz\right| \le 2\pi\delta K n(\gamma,w) \tag{81}$$

whenever $0 < \delta < q$. According to (80), the expression inside the modulus sign is

$$\int_\gamma \frac{f(z)\,dz}{z-w} - 2\pi i n(\gamma,w) f(w)\,,$$

which is independent of δ, but the right-hand side of (81) may be made arbitrarily small by suitable choice of δ. Hence, necessarily

$$\int_{\gamma[\delta]} \frac{f(z)-f(w)}{z-w}\,dz = 0 \qquad \text{whenever } 0 < \delta < r\,.$$

(Indeed, if $\kappa > 0$, take $\delta := \min(\frac{1}{2}r, \kappa/(3\pi K n(\gamma,w)))$, and (81) becomes false. So $\kappa = 0$ by contradiction). The Theorem now follows from (80).

If $n(\gamma,w) < 0$, one need only change the definition of $\gamma[\delta]$ to

$$\gamma[\delta](t) := w + \delta\exp(-it) \quad \text{for } 0 \le t \le 2\pi|n(\gamma,w)|\,,$$

and the proof is otherwise unaltered. □

Theorem 17.2. *Suppose $\mathcal{C}$ is a cycle nullhomologous in U, and $w \notin \mathcal{C}$ (that is, w is not a value of any of the contours that appear with nonzero coefficient in $\mathcal{C}$). Then*

$$n(\mathcal{C}, w) f(w) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f(z)\, dz}{z - w} \ .$$

Proof. This follows from 16.11, as 17.1 comes from 16.12. □

Remark 17.3. The most important case of 17.1 is when γ is a Jordan contour in U which is the frontier in $\mathbb{C}$, described anticlockwise, of a region V itself wholly included in U; these were the hypotheses of 16.3. It is intuitively clear that V then consists exactly of the points whose index with respect to γ is 1. (This is, indeed, how Eilenberg's proof of the Jordan curve theorem proceeds; see Ahlfors.) Then, for any $w \in V$, 17.1 gives

$$f(w) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)\, dz}{z - w} \, .$$

This is Cauchy's integral formula as he gave it. It applies, for instance, if γ is a circle small enough for its inside to be wholly within U.

Remark 17.4. The Cauchy integral formula expresses the value of f at any point $w \in V$ as an integral involving the values of f on the boundary contour γ. For *any* continuous function $g : \operatorname{Fr} V \longrightarrow \mathbb{C}$, we can define $Kg : V \longrightarrow \mathbb{C}$ by the formula

$$Kg(w) \coloneqq \int_{\gamma} \mathcal{K}(w, z) g(z)\, dz \, , \quad \text{for any } w \in V \, , \tag{82}$$

where $\mathcal{K}(w, z) \coloneqq \dfrac{1}{2\pi i (z - w)}$. The expression on the right of (82) is analogous to multiplication of a vector g (with "coordinates" $g(z)$ for the various $z \in \gamma$) by a "matrix" $\mathcal{K}$, with integration in the place of summation. $\mathcal{K}(w, z)$ has continuous families of "rows" (indexed by $w \in V$) and "columns" (indexed by $z \in \gamma$). Thus the function Kg may be thought of as obtained by applying a linear operator or matrix K to g, where K, defined by the integral on the right-hand side of (82), is called an 'integral operator K with kernel $\mathcal{K}$'. (This use of the word 'kernel' is quite distinct from, and older than, its usual meaning in algebra; it is unfortunately too late now to clear up the ambiguity).

The kernel $\mathcal{K}$ here, $1 \backslash \{2\pi i (z - w)$, is the "Cauchy kernel"; and Cauchy's integral theorem says that, for the given holomorphic function $f : U \longrightarrow \mathbb{C}$, $K(f | \operatorname{Fr} V) = f | V$. In this way, the values of f on the frontier of V completely determine its values in V. For example, if you know f on a circle nullhomologous in U—the frontier of a closed disk included in U—its behaviour on the whole disk is completely prescribed. This is only one aspect of what we might call the 'rigidity' of holomorphic functions. In general, they are fixed by surprisingly sparse data.

Proposition 17.5. *Again let U be an open set in $\mathbb{C}$, and γ a Jordan curve in U whose image is the frontier of a region $V \subseteq U$. Suppose that $g : \operatorname{Fr} V \longrightarrow \mathbb{C}$ is continuous, let $n \in \mathbb{N}$, and define $K_n g : V \longrightarrow \mathbb{C}$ by*

$$K_n g(w) := \frac{1}{2\pi i}\int_\gamma \frac{g(z)\,dz}{(z-w)^n} \qquad \textit{for all } \ w \in V\,. \tag{83}$$

Then $K_n g$ is holomorphic on V, and, for each $w \in V$,

$$(K_n g)'(w) = n(K_{n+1}g)(w)\,. \tag{84}$$

Proof. Fix $w \in V$, and suppose that h is so small that $w + h \in V$ too. Then

$$\begin{aligned}
&\frac{(K_n g)(w+h) - (K_n g)(w)}{h} \\
&\quad = \int_\gamma \left\{ \frac{1}{(z-w-h)^n} - \frac{1}{(z-w)^n} \right\} \frac{g(z)\,dz}{2\pi i h} \\
&\quad = \int_\gamma \frac{(z-w)^n - (z-w-h)^n}{(z-w-h)^n (z-w)^n} \frac{g(z)\,dz}{2\pi i\, h}\,.
\end{aligned}$$

However, by the binomial theorem,

$$\begin{aligned}
(z-w-h)^n = (z-w)^n - nh(z-w)^{n-1} \\
+ \tfrac{1}{2}n(n-1)h^2(z-w)^{n-2} - \dots\,,
\end{aligned}$$

so there is a polynomial in two variables $P(h, z-w)$ such that

$$\begin{aligned}
\frac{(K_n g)(w+h) - (K_n g)(w)}{h} &= \int_\gamma \frac{nh(z-w)^{n-1} + h^2 P(h, z-w)}{(z-w-h)^n (z-w)^n} \frac{g(z)\,dz}{2\pi i h} \\
&= \int_\gamma \frac{nh(z-w)^{n-1} + h^2 P(h, z-w)}{(z-w-h)^n (z-w)^n} \frac{g(z)\,dz}{2\pi i h} \\
&= \int_\gamma \frac{n}{(z-w-h)^n (z-w)} \frac{g(z)\,dz}{2\pi i} \\
&\qquad + \int_\gamma \frac{hP(h, z-w)}{(z-w-h)^n (z-w)^n} \frac{g(z)\,dz}{2\pi i}\,.
\end{aligned}$$

As $h \to 0$, each integrand here has a limit:

$$\begin{aligned}
\frac{n}{(z-w-h)^n (z-w)} \frac{g(z)}{2\pi i} &\to \frac{n}{2\pi i} \frac{g(z)}{(z-w)^{n+1}}\,, \\
\frac{hP(h, z-w)}{(z-w-h)^n (z-w)^n} \frac{g(z)}{2\pi i} &\to 0\,.
\end{aligned}$$

It is natural to suppose that the integrals tend to the integrals of the corresponding limits; and this follows quite rigorously, if one checks that the convergence of the integrands is uniform for $z \in \operatorname{Fr} V$ so that 6.9 applies. (I omit the details, which are tedious rather than hard; you should be able to supply them.) Consequently,

$$\frac{(K_n g)(w+h) - (K_n g)(w)}{h} \to \frac{n}{2\pi i}\int_\gamma \frac{g(z)\,dz}{(z-w)^{n+1}} \quad \text{as } \ h \to 0\,. \tag{85}$$

So $(K_n g)'(w)$ exists for $w \in V$, and is given by the stated formula. □

The statement (85) is another instance of 'differentiation under the integral sign', as at 15.12. As there, a detailed proof may be obtained by an appeal to 6.9.

Corollary 17.6. *Let r be a positive integer. Then $K_n g$, as defined at* (83)*, is complex-differentiable r times at w, and*

$$(K_n g)^{(r)}(w) = \frac{n(n+1)\cdots(n+r-1)}{2\pi i}\int_\gamma \frac{g(z)\,dz}{(z-w)^{n+r}}\,. \tag{86}$$

Proof. Apply 17.5, specifically (84), inductively. □

We obtain the quite amazing corollary:

Theorem 17.7. *Let U be an open set in $\mathbb{C}$, and suppose that $f : U \longrightarrow \mathbb{C}$ is holomorphic. Then f has complex derivatives of all orders at every point of U.*

Proof. Take any point w_1 in U, and $r > 0$ such that $B(w_1; r) \subseteq U$. Then, if $0 < \delta < r$, let $\gamma(t) = w_1 + \delta \exp(it)$ for $0 \le t \le 2\pi$. This is a Jordan contour in U which goes round the frontier of $V := B(w_1; \delta)$. So we can apply the Cauchy integral formula: for any $w \in V$,

$$f(w) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)\,dz}{z-w} = (Kg)(w)\,, \tag{87}$$

where $g := f|\mathrm{Fr}\, V$. But now, 17.6, with $n = 1$ and any $r = 1, 2, 3, \ldots$, tells us that, for any $w \in V$, the rth complex derivative of f at w exists and is given by the formula

$$f^{(r)}(w) = \frac{r!}{2\pi i}\int_\gamma \frac{f(z)\,dz}{(z-w)^{r+1}}\,, \tag{88}$$

which proves the Theorem. □

Remark 17.8. The existence of the *first* derivative $f'(w)$ at any $w \in U$ is precisely the hypothesis that f is holomorphic on U. But the formula (88) for $f'(z)$ is new; for general r, it is called Cauchy's formula for the derivatives. Notice that Cauchy's version of the Theorem assumed that f is strongly holomorphic (to apply Green's Theorem), but that is unnecessary (see Appendix D, and in particular D2.1, which is the crucial substitute for Green's Theorem).

We have now proved, from the stronger version of Cauchy's Theorem in Appendix D, that, if the complex derivative of f exists at every point of an open set U, all the higher-order complex derivatives of f exist at every point of U. A holomorphic function is strongly holomorphic, since its first complex derivative, being itself complex-differentiable, is continuous; its derivative is itself holomorphic; and so on. I called this an amazing fact, since nothing similar is true for real-valued functions of a real variable.

For a simple example, start from the function

$$f(t) := \begin{cases} 1 - |t| & \text{for } -1 \le t \le 1\,, \\ 0 & \text{elsewhere.} \end{cases}$$

This function is continuous everywhere, but not differentiable at ± 1 and at 0, where the graph has 'corners'. If g is an indefinite integral of f, g is once differentiable everywhere, but not twice differentiable at ± 1 or 0.

There is a famous class of examples (invented by Weierstrass) of functions continuous everywhere on $\mathbb{R}$ but differentiable nowhere. There are even significant senses in which it can be said that 'nearly all' continuous functions are nowhere differentiable. The indefinite integral of any such function will be differentiable everywhere, with continuous derivative, but will be nowhere twice differentiable.

There is an apparent paradox here. A function $\mathbb{C} \longrightarrow \mathbb{C}$ which is everywhere complex-differentiable is everywhere twice complex-differentiable, but a function $\mathbb{R} \longrightarrow \mathbb{R}$ which is everywhere real-differentiable need not be twice real-differentiable anywhere. The resolution of the paradox is that some real-differentiable functions $\mathbb{R} \longrightarrow \mathbb{R}$ cannot be extended to any complex-valued complex-differentiable function on any open set U in $\mathbb{C}$ including $\mathbb{R}$.

For a function of a real variable, there is a second related property that contrasts with the behaviour of holomorphic functions. Even if the real-valued function is differentiable to all orders everywhere on $\mathbb{R}$, the sequence of its derivatives at a specific point x is quite arbitrary, as I mentioned before 7.10. Given *any* sequence (c_n) of real numbers, there is a C^∞ function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f^{(n)}(x) = c_n$ for all n. (This theorem of Denjoy is messy rather than profound, and it has long been superseded by famous but more difficult theorems of Whitney, my mathematical grandfather). The sequence (c_n) can increase as rapidly as we like; one might take $c_n := n^{\exp n}$, or $c_n := \exp(\exp(n!))$. However, for holomorphic functions the successive (complex) derivatives at a fixed point cannot increase too rapidly because of the *Cauchy estimates*.

Lemma 17.9. *Suppose in* 17.5 *that the length of the contour γ is L, and that $|g(z)| \le M$ and $|w - z| \ge D > 0$ for every point z on γ. Then, for $r = 0, 1, 2, \dots$,*

$$|(K_n g)^{(r)}(w)| \le \frac{(n + r - 1)!\, LM}{2\pi (n - 1)!\, D^{n+r}} . \tag{89}$$

Proof. Apply the fundamental estimate (74) to the integral (86). [The formula (89) does make sense, as $n \ge 1$.] □

Theorem 17.10. (The Cauchy estimates). *Let $w \in \mathbb{C}$ and $R > 0$, and suppose that f is defined and holomorphic on the disk $D = \{z : |z - w| < R\}$ and that $|f(z)| \le K$ for $z \in D$. Then, for $r = 0, 1, 2, \dots$,*

$$|f^{(r)}(w)| \le r!\, K R^{-r} . \tag{90}$$

Proof. Let $\rho \in (0, R)$; 17.9 applies with $\gamma(t) = \rho \exp(it)$, $g := f|\mathrm{Fr}\, V$. On γ, $|g(z)| = |f(z)| \le K$, and the length L of γ is $2\pi\rho$. So, by (89) with $n = 1$,

$$|f^{(r)}(w)| \le \frac{r!}{2\pi} \frac{2\pi\rho K}{\rho^{r+1}} = r! K \rho^{-r} .$$

Since this is true for any ρ such that $0 < \rho < R$, it must be the case that

$$|f^{(r)}(w)| \le r!\, K R^{-r} .$$

(If this 'limiting case' were false, $|f^{(r)}(w)| > r!\, K R^{-r}$, it would be possible to choose $\rho < R$ so close to R that $|f^{(r)}(w)| > r!\, K \rho^{-r}$). □

Remark 17.11. *(a)* Let f be a function holomorphic on an open set U containing w. Let $R' > 0$ be such that $C(w;R') := \{z : |z-w| \le R'\} \subseteq U$ (cf. 9.11*(a)*). This set is (sequentially) compact by 3.11. By 5.9, $f(C(w;R'))$ is also sequentially compact, so it is bounded in $\mathbb{C}$: $|f(z)| \le K(R')$ for all $z \in C(w;R')$ and some $K(R') \ge 0$. (Here $K(R')$ depends on R'.)

(b) For that R' and $K(R')$, the sequence of derivatives of f at w satisfies inequalities of the form (90). The derivatives cannot increase too rapidly with r; we cannot, for example, have $|f^{(r)}(w)| \ge r^{\exp r}$. In particular,

$$\left| \frac{f^{(r)}(w)}{r!}(z-w)^r \right| \le K(R') \left| \frac{z-w}{R'} \right|^r \qquad \text{for } r = 0, 1, 2, \ldots .$$

(c) If $|z-w| < R'$, $\sum K(R') \left| \frac{z-w}{R'} \right|^r$ is a geometric series with common ratio less than 1, and, therefore, converges (cf. 4.8). The Taylor series $\sum \frac{f^{(r)}(w)}{r!}(z-w)^r$ converges (absolutely) by the comparison test (and 4.13). Thus: if f is holomorphic on U and $C(w,R') \subseteq U$, the Taylor series of f about w converges for $|z-w| < R'$. If, in fact, the *open* disk $B(w;R) \subseteq U$ (for instance, if U is exactly $B(w;R)$, or if R is the distance from w to the nearest point at which f is undefined or is not complex-differentiable), the above reasoning works for any $R' \in (0,R)$, so that the Taylor series converges for any $z \in B(w,R)$ (given z, take R' so that $|z-w| < R' < R$).

This contrasts with the situation for real-valued C^∞ functions of a real variable, whose Taylor series at a point need not converge (as is clear from Denjoy's theorem, mentioned in 17.8, which says that the Taylor coefficients at a point may be anything at all). *However*, the sum of the Taylor series of f about w has not yet been shown to be $f(z)$; see 18.11.

This rather tough section will end with an interesting and easy "digression". Liouville's theorem, one of three famous results to his name, in three widely divergent branches of mathematics—apart from this, in mechanics and in number theory—, was actually discovered by Cauchy, and Liouville simply included it in some lectures a couple of years later; such are the vagaries of fame. It has many important applications, but for us its immediate interest is to give a proof of the Fundamental Theorem of Algebra.

Theorem 17.12. (Liouville's theorem). *Suppose $f : \mathbb{C} \longrightarrow \mathbb{C}$ is holomorphic and bounded; that is, $|f(z)| \le K$ for all $z \in \mathbb{C}$, for some $K \ge 0$. Then f is constant.*

Proof. Fix $w \in \mathbb{C}$; apply 17.10 to $B(w;R)$. We find that

$$|f'(w)| \le KR^{-1} .$$

But R may be as large as we like, whilst K is given once and for all. So the only possibility is that $|f'(w)| = 0$. As this must be true for all $w \in \mathbb{C}$, f' is identically 0. Not surprisingly, it follows that f is constant. For instance, taking γ to be the contour $\gamma(t) := tz$ for $0 \le t \le 1$,

$$f(z) = f(0) + \int_\gamma f' = f(0)$$

for any choice of $z \in \mathbb{C}$, by 15.9*(a)*. □

There are other proofs—one is in the corresponding assignment.

Theorem 17.13. (The Fundamental Theorem of Algebra). *Any complex polynomial of degree greater than* 0 *has at least one complex root.*

Proof. Suppose the polynomial is

$$p(z) = a_0 + a_1 z + \ldots + a_n z^n,$$

where $n > 0$ and $a_n \neq 0$ (this is what it means to "have degree greater than 0"). Assume, if possible, that p has no complex root. As a function of the complex variable z, p is holomorphic (by 12.3–12.6) and $p(z) \neq 0$ for all $z \in \mathbb{C}$ by hypothesis. Hence $1/p(z)$ is holomorphic on all of $\mathbb{C}$ (by 12.5 and 12.2). But, if $|z| = R$,

$$\begin{aligned} |p(z)| &= |a_n z^n + \ldots + a_1 z + a_0| \\ &\geq |a_n| R^n - |a_{n-1}| R^{n-1} - \ldots - |a_0|, \end{aligned}$$

which tends to ∞ as $R \to \infty$, since the R^n term dominates (you should invent a genuine proof for this). Thus $1/p(z)$ is holomorphic on $\mathbb{C}$ and tends to 0 as $|z| \to \infty$; it is therefore bounded on the whole of $\mathbb{C}$, and Liouville's theorem tells us that it is constant. Furthermore, the constant must obviously be 0. This is absurd, as $1/p(z)$ cannot take the value 0. □

Another contradiction is available: if $1/p(z)$ is constant, so is $p(z)$. But, if p is a constant function (takes only the one value), it must be a polynomial of degree 0 (why? This requires a little thought!), which is contrary to our assumptions.

Remark 17.14. There are many other proofs of the Fundamental Theorem of Algebra, some of which may require less preparation in total—for we have used a lot of analysis—and some of which allow useful generalization. D'Alembert presented a proof of sorts in 1746, and Euler suggested an "algebraic" proof (using, however, the non-"algebraic" facts that $\mathbb{R}$ is an ordered field and that real polynomials of odd degree have a real root) in 1749, and these proofs were legitimized by later mathematicians. It is often said that the first *satisfactory* proof was due to Gauss (around 1799), before complex analysis as such existed. In fact his proof also required later correction (by Ostrowski in—perhaps—1927). The whole question was a sort of watershed in mathematics, being perhaps the first time that a solution of a problem had to be shown to *exist* without its being *exhibited* in a formula of some sort. (We now know, of course, that there are serious difficulties in 'exhibiting' a root of a general polynomial of degree greater than 4 by any algebraic formula.) The proof above, however, is perhaps the easiest to understand if Liouville's Theorem, which is important for other reasons, is already established, and it requires only rather inoffensive versions of Cauchy's theorem; in principle, it is enough to know the theorem for rectangular contours. There are many other proofs that do not require any complex analysis as such.

One naturally feels that there should be a purely 'algebraic' proof of the FTA. The feeling depends on what you understand by "algebra", and so is incapable of precise expression, but, if one accepts the vague pseudo-definition presented in the Introduction, that algebra is concerned exclusively with finitary operations (ones that involve only finitely many arguments or steps, like addition or multiplication), it is unfounded. The FTA is a result *specifically about complex numbers*, not about algebra in general. What it says in algebraic terms is that the algebraic completion of the field $\mathbb{R}$ is the splitting field of the polynomial $x^2 + 1$, a quite specific extension of $\mathbb{R}$. (Other fields have algebraic completions, but usually they are constructed by the Axiom of Choice and cannot be described explicitly.) Now, the complex numbers are constructed from the real numbers, and the construction of the real numbers is

itself not 'algebraic'. Thus any theorem which is directly or indirectly about real numbers, rather than about fields in general, must implicitly involve limiting processes or something equivalent. You will have noticed above that Euler's proposed proof required some form of the Intermediate Value Theorem, which is not "algebraic".

§18. Singularities.

The word 'singularity' is used by mathematicians in many contexts, often without precise explanation, to mean either 'something that goes wrong' or 'the circumstances in which something goes wrong'. In complex analysis, we can make the following

Definition 18.1. Let $A \subseteq \mathbb{C}$. The function $f : A \longrightarrow \mathbb{C}$ has a singularity at the point $w \in \mathbb{C}$, or, equivalently, w is a singularity of f, if either

(a) f is not defined on the *whole* disk $B(w;\delta) = \{z : |z - w| < \delta\}$ for any $\delta > 0$; in other words, w is not an interior point of A; or

(b) although there is some $\delta > 0$ such that f is defined at all points of $B(w;\delta)$, the complex derivative of f does not exist at w.

Definition 18.2. Suppose that, for some $\delta > 0$, f is defined and holomorphic at all points on the so-called 'punctured disk'

$$B(w;\delta) \setminus \{w\} = \{z : 0 < |z - w| < \delta\},$$

but is either undefined or is not complex-differentiable at w itself. Then w is called an *isolated singularity* of f. Notice that in this case w is an interior point of $A \cup \{w\}$.

The importance of isolated singularities is twofold: they appear in interesting circumstances, and they have a simple and illuminating theory. A rough-and-ready classification, taking no account of the questions whether all the types we distinguish are possible or whether the classification is useful, might be as follows.

Definition 18.3. Suppose that w is an isolated singularity of f.

(a) w is a *removable* singularity of f if there exists $\delta > 0$ such that f is defined, holomorphic, and *bounded* on $B(w;\delta) \setminus \{w\}$. [For the moment, the name "removable" is just a convenient word.]

(b) w is a *pole* of f if $|f(z)| \to \infty$ as $z \to w$. (This is the simplest possible behaviour f might have if it is defined and holomorphic, but unbounded, on $B(w;\delta) \setminus \{w\}$, for any small enough δ).

(c) w is an *isolated essential singularity* if it is neither removable nor a pole. (On every small punctured ball about w, f is unbounded; but $|f(z)|$ does *not* tend to ∞ as $z \to w$.)

Remark 18.4. The method we shall employ to study holomorphic functions on a punctured disk works somewhat more generally. Suppose $0 \leq r < R$, where R may take the symbolic value ∞. The set

$$A(z_0; r, R) = \{z : r < |z - z_0| < R\}$$

is the (open) *annulus* centred at z_0 of inner radius r and outer radius R. If $R < \infty$, it is the region between two circles $|z - z_0| = r$ and $|z - z_0| = R$; and, if $r = 0$, it reduces to a punctured disk of radius R about z_0. If $R = \infty$, $A(z_0; r, R)$ is the whole region outside the circle $|z - z_0| = r$.

Lemma 18.5. *Let f be a complex-valued function defined and holomorphic on $A(z_0; r, R)$, where $0 \le r < R \le \infty$.*

(a) *For $\rho \in (r, R)$, let $\gamma[\rho]$ be the contour defined for $0 \le t \le 2\pi$ by*

$$\gamma[\rho](t) = z_0 + \rho \exp(it).$$

If $n \in \mathbb{Z}$, the integral $\displaystyle\int_{\gamma[\rho]} (z - z_0)^n f(z)\, dz$ does not depend on $\rho \in (r, R)$.

(b) *If $r < \rho_1 < |w - z_0| < \rho_2 < R$, then*

$$f(w) = \frac{1}{2\pi i}\int_{\gamma[\rho_2]} \frac{f(z)\, dz}{z - w} - \frac{1}{2\pi i}\int_{\gamma[\rho_1]} \frac{f(z)\, dz}{z - w}.$$

Proof. *(a)* The function $(z - z_0)^n f(z)$ is holomorphic on the annulus, as $(z - z_0)^n$ is, even for negative n. The two concentric anticlockwise circles, $\gamma[\rho_1]$ and $\gamma[\rho_2]$, where ρ_1, ρ_2 are as in *(b)*, are evidently homologous in $A(z_0; r, R)$. (Each has winding number 1 about any $w \in C(z_0; r)$ and does not wind about any $w \notin B(z_0; R)$.) The result follows from 16.12.

(b) $\gamma[\rho_1], \gamma[\rho_2]$ lie in $A(z_0; r, R)$, and the cycle $\mathcal{C} := \gamma[\rho_2] - \gamma[\rho_1]$ is nullhomologous in $A(z_0; r, R)$ and winds once around w. Apply 17.2:

$$f(w) = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{f(z)\, dz}{z - w} = \frac{1}{2\pi i}\int_{\gamma[\rho_2]} \frac{f(z)\, dz}{z - w} - \frac{1}{2\pi i}\int_{\gamma[\rho_1]} \frac{f(z)\, dz}{z - w},$$

as required. □

Remark 18.6. I must now introduce *Laurent series*. First, I recall some facts from §9.

A power series about the centre z_0 takes the form $\sum_{n=0}^{\infty} a_n (z - z_0)^n$, where $(z - z_0)^0$ denotes 1 (even if $z = z_0$) and the coefficients $a_n \in \mathbb{C}$ are arbitrary. It converges when $z = z_0$, with sum a_0, but need not converge for any other values of z; however, it has a radius of convergence R, either a non-negative real number or ∞, which has the property that the series converges absolutely if $|z - z_0| < R$ and diverges if $|z - z_0| > R$. (When $R = \infty$, the first inequality is true and the second is false for all $z \in \mathbb{C}$; if $R = 0$ the first inequality is always false).

By a *Laurent series* about z_0 we mean an extended expression of the form $\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$. The (formal) sum is now taken over all integer powers of $z - z_0$. This raises an obvious question: what can we mean by convergence or summation of such a "series"? For an ordinary power series, convergence means convergence of the sequence of partial sums, but for a Laurent series, which has no 'first term', it is not clear what a 'partial sum' should be. To avoid worrying about this, we define convergence of the Laurent series $\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$ at $z \in \mathbb{C}$ to mean convergence of both the series $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ and $\sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n}$.

The first of these is an ordinary power series (called the *regular part* of the Laurent series). It has a radius of convergence which I may call ρ'.

The second, which reverses the order of the terms in the Laurent series, is called the *singular part* of the Laurent series. If all the coefficients a_n for $n < 0$ are 0, we agree that this singular part converges when $z = z_0$ with sum 0. Otherwise, the singular part is a power series in the variable $(z - z_0)^{-1}$, and, as such, has a radius of convergence ρ. It converges absolutely when $|(z - z_0)^{-1}| < \rho$ and diverges if $|(z - z_0)^{-1}| > \rho$. If $\rho > 0$, this means it converges absolutely for $|z - z_0| > \rho^{-1}$ and diverges for $|z - z_0| < \rho^{-1}$. If $\rho = 0$, it never converges at all, as $(z - z_0)^{-1} = 0$ is impossible. If $\rho = \infty$, it converges for all $z \neq z_0$.

The Laurent series, therefore, converges absolutely when $\rho > 0$ and $\rho^{-1} < |z - z_0| < \rho'$, and diverges if $\rho = 0$, $|z - z_0| < \rho^{-1}$, or $|z - z_0| > \rho'$. That is, the interior of the set of convergence of a Laurent series is an open annulus, called the annulus of convergence, on which the series converges absolutely (and so it is unimportant what rules we use to sum it). We call ρ^{-1} the inner radius of convergence and ρ' the outer radius of convergence. As with a power series, there can be no general statement about convergence at the points of the frontier circles. However, when $\rho^{-1} > \rho'$ the series never converges ($\rho = 0$ is a limiting case of this), and, if $\rho^{-1} = \rho'$, the only points where it may possibly converge lie on the circle $|z - z_0| = \rho'$, so that the annulus of convergence will certainly be empty.

The singular part determines the inner radius ρ^{-1} of convergence, and the regular part the outer radius ρ' of convergence.

Since a Laurent series is in effect the sum of two power series, the standard properties of power series (see §9), after the obvious modifications, apply when appropriate.

(1) the sum of a Laurent series is a holomorphic function on its annulus of convergence, and

(2) if one has two Laurent series in $z - z_0$, say $\sum a_n(z - z_0)^n$ and $\sum b_n(z - z_0)^n$, and they both converge on an annulus $r < |z - z_0| < R$ (which may be a proper subset of both annuli of convergence), then 'they may be multiplied term-by-term' thereon. That is, if we multiply all the individual terms and collect together like powers of $z - z_0$, the resulting Laurent series,

$$\sum\nolimits_{n=-\infty}^{+\infty} \left(\sum\nolimits_{j+k=n} a_j b_k \right) (z - z_0)^n$$

converges on the same annulus, and its sum at each point of the annulus is the product

$$\left(\sum\nolimits_{n=-\infty}^{\infty} a_n (z - z_0)^n \right) \left(\sum\nolimits_{m=-\infty}^{\infty} b_m (z - z_0)^m \right).$$

This is a corollary of the theorem on multiplication of absolutely convergent series (see 9.9), as the Laurent series must converge *absolutely* on their annuli of convergence. (The product is not, however, a Cauchy product, as the terms of the new series are, in principle, infinite sums). This fact is of some importance later.

(3) For the same reason, the sum of a Laurent series may be differentiated term-by-term on its annulus of convergence, the result being, of course, another Laurent series.

Theorem 18.7. (Laurent's theorem). *Let f be a function holomorphic on the annulus $A(z_0; r, R)$, where $0 \leq r < R \leq \infty$. Then there is a unique Laurent series $\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$, whose annulus of convergence includes $A(z_0; r, R)$, such that, for every $w \in A(z_0; r, R)$,*

$$f(w) = \sum\nolimits_{n=-\infty}^{\infty} a_n (w - z_0)^n. \tag{91}$$

Furthermore, the coefficient a_n *is given for each* $n \in \mathbb{Z}$ *by the formula*

$$a_n = \frac{1}{2\pi i}\int_{\gamma[\rho]} (z - z_0)^{-n-1} f(z)\, dz\,, \tag{92}$$

where $\gamma[\rho]$ *is, as in* 18.5(a)*, a circular Jordan contour contained in the annulus:* $\gamma[\rho](t) = z_0 + \rho \exp(it)$ *, for* $0 \le t \le 2\pi$ *, with* $r < \rho < R$ *.*

The choice of $\rho \in (r, R)$ has no effect on the formua (92), by 18.5*(a)*.

Proof. Choose ρ_1 and ρ_2 so that $r < \rho_1 < |w - z_0| < \rho_2 < R$. As in 18.5*(b)*,

$$f(w) = \frac{1}{2\pi i}\int_{\gamma[\rho_2]} \frac{f(z)\,dz}{z - w} - \frac{1}{2\pi i}\int_{\gamma[\rho_1]} \frac{f(z)\,dz}{z - w}\,. \tag{93}$$

Now, as $|w - z_0| < \rho_2 = |z - z_0|$ for each z on the contour $\gamma[\rho_2]$,

$$\begin{aligned}\frac{1}{z - w} &= \frac{1}{(z - z_0) - (w - z_0)} = \frac{1}{z - z_0}\left(1 - \frac{w - z_0}{z - z_0}\right)^{-1} \\ &= \frac{1}{z - z_0}\left(1 + \frac{w - z_0}{z - z_0} + \left(\frac{w - z_0}{z - z_0}\right)^2 + \dots\right) \\ &= \sum\nolimits_{n=0}^{\infty} \frac{(w - z_0)^n}{(z - z_0)^{n+1}}\,,\end{aligned}$$

this being merely a geometrical series with common ratio $(w - z_0)/(z - z_0)$ of modulus less than 1. So

$$\int_{\gamma[\rho_2]} \frac{f(z)\,dz}{z - w} = \int_{\gamma[\rho_2]} \left(\sum\nolimits_{n=0}^{\infty} \frac{(w - z_0)^n f(z)}{(z - z_0)^{n+1}}\right) dz\,. \tag{94}$$

We may interchange the summation and the integration. (This follows from 6.9; the partial sums of the series converge uniformly on the contour to the sum $f(z)/(z - w)$. Indeed, as f is continuous on the contour, there is a constant K such that $|f(z)| \le K$ at every point $z \in \gamma[\rho_2]$ (cf. 17.11*(a)*), and then the nth. term of the series has modulus not exceeding the corresponding term $K|w - z_0|^n \rho_2^{-n-1}$ of a convergent geometric series not depending on z.) So we obtain from (94)

$$\begin{aligned}&\frac{1}{2\pi i}\int_{\gamma[\rho_2]} \frac{f(z)\,dz}{z - w} \\ &\qquad = \sum\nolimits_{n=0}^{\infty} (w - z_0)^n \left\{\frac{1}{2\pi i}\int_{\gamma[\rho_2]} \frac{f(z)}{(z - z_0)^{n+1}}\, dz\right\} \\ &\qquad = \sum\nolimits_{n=0}^{\infty} a_n (w - z_0)^n\,, \qquad \text{by (92) and 18.5(a).}\end{aligned} \tag{95}$$

[Since we now know this series converges for $|w - z_0| < \rho_2$, it converges absolutely there, from 9.2. But one may also use the fundamental estimate to show directly that

$$|a_n| \le \frac{1}{2\pi}.2\pi\rho_2.K\rho_2^{-n-1} = K\rho_2^{-n}\,,$$

which implies absolute convergence of the series (95) for $|w - z_0| < \rho_2$.]

On the other hand, $|w - z_0| > \rho_1 = |z - z_0|$ for each z on the contour $\gamma[\rho_1]$, and

$$\frac{1}{z-w} = \frac{1}{(z_0-w)-(z_0-z)} = \frac{1}{z_0-w}\left(1 - \frac{z_0-z}{z_0-w}\right)^{-1}$$
$$= -(w-z_0)^{-1}\left(1 + \frac{z-z_0}{w-z_0} + \left(\frac{z-z_0}{w-z_0}\right)^2 + \dots\right).$$

Again, the series is a geometric series with common ratio $(z-z_0)/(w-z_0)$ of modulus less than 1. From this,

$$\int_{\gamma[\rho_1]} \frac{f(z)\,dz}{z-w} = -\int_{\gamma[\rho_1]} \left(\sum\nolimits_{k=0}^{\infty} \frac{(z-z_0)^k f(z)}{(w-z_0)^{k+1}}\right) dz\,,$$

where, once again, we can interchange integration and summation (the series converges uniformly on the contour $\gamma[\rho_1]$, much as before), and obtain

$$\frac{1}{2\pi i}\int_{\gamma[\rho_1]} \frac{f(z)\,dz}{z-w}$$
$$= -\sum\nolimits_{k=0}^{\infty} (w-z_0)^{-k-1} \left\{\frac{1}{2\pi i}\int_{\gamma[\rho_1]} (z-z_0)^k f(z)\,dz\right\}, \tag{96}$$

As with (95), this series converges absolutely for $|w - z_0| > \rho_1$, so that it can be rearranged in any fashion without abolishing its convergence or changing its sum. In any case, our conventions 18.6 on Laurent series prescribe summing in the order '-1 downwards', so that, in writing the summation "from $-\infty$ to -1", we really mean the same as (96). In short, we may write n in the place of $-k-1$, and then (96) becomes

$$-\sum\nolimits_{n=-\infty}^{-1} (w-z_0)^n \left\{\frac{1}{2\pi i}\int_{\gamma[\rho_1]} (z-z_0)^{-n-1} f(z)\,dz\right\}$$
$$= -\sum\nolimits_{n=-\infty}^{-1} a_n (w-z_0)^n\,, \quad \text{again by (92) and 18.5(a).}$$

Combining this with (93) and (95) yields (91).

The uniqueness of the Laurent expansion follows by substituting (91) into (92) and justifying the interchange of summation and integration, exactly as above. □

Remark 18.8. *(a)* In the *proof*, the outer contour $\gamma[\rho_2]$ contributes the nonnegative powers of $w - z_0$ in (91) and the inner contour $\gamma[\rho_1]$ contributes the negative powers. But 18.5*(a)* shows that the coefficients a_n may be calculated for all values of n by the integral (92) around any circle $|z - z_0| = \rho$ lying within the annulus.

(b) Although the Laurent series for f exists and its coefficients are given by (92), it is rarely appropriate to employ those formulæ to calculate them. Almost always the relation (92) is used to evaluate integrals on the right. The Laurent series being unique, the coefficients can be found by other methods (term-by-term multiplication, differentiation, or what-not). Recall (2) and (3) of 18.6. Examples will clarify this later.

Remark 18.9. *(a)* Suppose that z_0 is a removable isolated singularity of f (recall 18.3*(a)*). By definition, there are positive numbers δ and K such that f is holomorphic and $|f(z)| \le K$ when $0 < |z - z_0| < \delta$. Applying the fundamental estimate (74) to (92), with

$0 < \rho < \delta$,

$$|a_n| = \left| \frac{1}{2\pi i} \int_{\gamma[\rho]} (z - z_0)^{-n-1} f(z)\, dz \right| \leq \frac{1}{2\pi} \times 2\pi\rho \times \frac{K}{\rho^{n+1}} = K\rho^{-n} \qquad (97)$$

(compare 17.9). This estimate holds for any $n \in \mathbb{Z}$ and $\rho \in (0, \delta)$, and, when $n < 0$, $K\rho^{-n} \to 0$ as $\rho \to 0$. Thus $|a_n| = 0$ and $a_n = 0$ when $n < 0$, and the Laurent series has zero singular part:

$$f(z) = \sum\nolimits_{n=0}^{\infty} a_n (z - z_0)^n \qquad \text{for } 0 < |z - z_0| < R\,.$$

(b) The right-hand side here is an ordinary power series, and defines a function $g(z)$ that is holomorphic for $|z - z_0| < R$. (See 9.14.) Therefore, f only has a singularity at z_0 because it is either undefined or has the "wrong value" there; for g agrees with f on the punctured disk and is holomorphic on the whole disk. A removable isolated singularity may therefore be 'removed' by defining or redefining the value of f at the singularity. The new value must be $\lim\limits_{z \to z_0} f(z)$.

(c) As simple examples of removable singularities, we may mention the singularities at 0 of the functions defined at all other points $z \in \mathbb{C}$ by the formulæ z/z or $(\sin z)/z$. To complete their definitions as holomorphic functions on the whole plane, we need only specify the "correct" value at 0, which is 1 in both cases. The regular parts of the resulting Laurent series are

$$1 + 0z + 0z^2 + \cdots, \quad 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \frac{z^6}{7!} + \cdots.$$

Note 18.10. In 18.9*(a)*, let $M(\rho) := \sup\{|f(z)| : |z - z_0| = \rho\}$ if $\rho \in (0, \delta)$. The calculation (97) then gives $|a_n| \leq \rho M(\rho)/\rho^n$. It follows as before that, if $\rho M(\rho) \to 0$ as $\rho \to 0$ (in effect saying that $(z - z_0)f(z) \to 0$ as $z \to z_0$), then $|a_n| = 0$ and $a_n = 0$ for all $n < 0$; the Laurent series again has zero singular part, and the singularity at z_0 may be removed, as before. In "intuitive" language, $\rho M(\rho) \to 0$ means that "$|f(z)|$ tends to ∞ *genuinely more slowly* than $|z - z_0|^{-1}$ as $z \to z_0$", and so one appears to have a notable strengthening of the result of 18.9*(a)*; but the conclusion is that it only happens if the singularity may be removed, so that $|f(z)|$ is bounded anyway on a punctured ball about z_0. Functions holomorphic on the punctured disk $A(z_0; 0, \delta)$ for which $\rho M(\rho) \to 0$ as $\rho \to 0$ must in fact be bounded on some punctured disk $A(z_0; 0, \delta')$ with $0 < \delta' \leq \delta$.

You might guess there are exceptions like $(z - z_0)^{-4/5}$, for which one may suppose that $\rho M(\rho) \to 0$ as $\rho \to 0$. But recall 14.4 and 14.7. Such a fractional power cannot be defined as a (single-valued) continuous (or holomorphic) function on a whole annulus about z_0, so the Laurent expansion does not apply at all.

Theorem 18.11. *If f is holomorphic on the punctured disk $A(z_0; 0, R)$ and is bounded on $A(z_0; 0, \delta)$ for some $\delta \in (0, R)$, then there is a function g, holomorphic on the disk $B(z_0; R)$, which agrees with f on $A(z_0; 0, R)$. Furthermore, the Taylor series of g about z_0 converges to g at every point of $B(z_0; R)$.*

If f is in fact holomorphic on $B(z_0; R)$, then of course $g = f$.

Proof. In 18.9 we defined g by the Laurent series of f, which turned out to be an ordinary power series. That it is in fact the Taylor series of g about z_0 follows from comparing (88) of

17.7 with (92) of 18.7, or, indeed, by taking successive derivatives of the series for g term-by-term (by 9.14) and evaluating them at z_0. □

Remark 18.12. Removable isolated singularities are often 'removed' without comment. One might carelessly say, for instance, that '$\frac{\sin z}{z}$ is holomorphic on $\mathbb{C}$', because its singularity at 0 is obviously removable. In 18.9*(b)* one would not normally distinguish f from its extension g across the point z_0.

In 18.11, we have proved that a holomorphic function on $B(z_0; R)$ is the sum there of its Taylor series about z_0. In the real case, the question of convergence of the Taylor series (32) of a function to the function itself involved many difficulties, as we saw; the function might be differentiable only a limited number of times, so that the Taylor series lacks its later terms; the Taylor series might not converge at all; it might converge, but not to the function. A proof that the Taylor series does converge to the function usually involves messing with complicated remainder formulæ. The complex case is simpler, because complex-differentiability at all points of $B(z_0; R)$ is a vastly stronger condition than the existence of real derivatives even of all orders.

The radius of convergence of the Taylor series in 18.11 cannot be less than R. If it is $R' > R$, the sum of the series defines a holomorphic function h on the larger disk $B(z_0; R')$, and h agrees with g on $B(z_0; R)$. There can be only one such function h defined on $B(z_0; R')$ (you should consider why this is now obvious). Thus h is uniquely defined by f, and we can think of it as the natural extension of f to the larger disk. Hence, the radius of convergence R'' of the Taylor series of f about z_0 is the distance from z_0 to the nearest 'genuine' singularity of f—meaning that there is a holomorphic extension of f to the disk $B(z_0; R'')$, but not to any larger disk. (If there were a holomorphic extension to a larger disk, its Taylor series would be the same series, but would have a larger radius of convergence, by 18.11).

A function holomorphic on all of $\mathbb{C}$ is called an *entire function* (or an *integral function*; these odd names are translations of the odd original German ganze Funktion). Entire functions as so defined (for the phrase has other senses) have the property that their Taylor series about any point have infinite radius of convergence. Conversely, any such Taylor series defines an entire function. Typical examples are $\exp z$, $\sin z$, $\exp(\sin(z^2 + z))$.

Remark 18.13. *(a)* Next, suppose that $|f(z)| \to \infty$ as $z \to z_0$ (f has a pole at z_0). There exists $\delta > 0$ such that $|f(z)| > 1$ for $|z - z_0| < \delta$. So $1/f(z)$ is holomorphic and bounded for $0 < |z - z_0| < \delta$; its singularity at z_0 is removable, and by 18.11

$$\frac{1}{f(z)} = \sum\nolimits_{n=0}^{\infty} b_n (z - z_0)^n \quad \text{for some coefficients } b_n,$$

where the series converges at least for $0 \leq |z - z_0| < \delta$, although the equality requires $z \neq z_0$. Furthermore, $b_0 = 0$, since it is the limit of $1/f(z)$ as $z \to z_0$. However, $1/f(z)$ cannot be constant (as f isn't), so not all the coefficients b_n can be 0, and there must be a least index $k \geq 1$ such that $b_k \neq 0$. This index k is called the multiplicity of the zero of the sum-function at z_0. For $0 < |z - z_0| < \delta$,

$$\begin{aligned}\frac{1}{f(z)} &= (z-z_0)^k(b_k + b_{k+1}(z-z_0) + b_{k+2}(z-z_0)^2 + \dots) \\ &= (z-z_0)^k g(z)\,,\end{aligned}$$

where $g(z)$ is defined by the series $b_k + b_{k+1}(z-z_0) + \dots$. This series must converge for $|z-z_0| < \delta$ (why? recall 9.2), so that g is holomorphic and $g(z_0) = b_k \neq 0$. Therefore, $g(z) \neq 0$ when z is sufficiently close to z_0, say for $|z-z_0| < \delta' \leq \delta$, and $h(z) = 1/g(z)$ is holomorphic for $|z-z_0| < \delta'$. In turn $h(z)$ has a power series expansion:

$$h(z) = c_{-k} + c_{-k+1}(z-z_0) + c_{-k+2}(z-z_0)^2 + \dots\,.$$

The indices start at $-k$ for convenience, and $c_{-k} = h(0) = b_k^{-1}$. So

$$\begin{aligned}f(z) &= (z-z_0)^{-k}\{c_{-k} + c_{-k+1}(z-z_0) + c_{-k+2}(z-z_0)^2 + \dots\} \\ &= c_{-k}(z-z_0)^{-k} + c_{-k+1}(z-z_0)^{-k+1} + \dots\end{aligned} \tag{98}$$

for $0 < |z-z_0| < \delta'$. We exclude $z = z_0$, as $(z-z_0)^{-k}$ is not defined there. This shows that f has a Laurent series about z_0, valid (in the first instance) for $0 < |z-z_0| < \delta'$, with only finitely many nonzero coefficients for the negative powers of $z - z_0$. Laurent's theorem shows that f has a Laurent series in any punctured disk $0 < |z-z_0| < R$ on which f is holomorphic; by uniqueness, it must be the one we have found on $A(z_0; 0, \delta')$.

Conversely, any Laurent series whose singular part is not zero but contains only finitely many non-zero terms can be written in the form (98), where $c_{-k} \neq 0$. As $z \to z_0$, the series in braces tends to c_{-k}, and therefore $|f(z)| \to \infty$ as $z \to z_0$. Hence poles are completely characterized by this property of Laurent series.

(b) In short, a removable isolated singularity of f is one about which the Laurent series of f has zero singular part, and a pole z_0 is an isolated singularity about which the Laurent series of f has a singular part consisting of at least one, but only finitely many nonzero terms (and therefore the singular part converges everywhere in a punctured disk about z_0 except at z_0 itself). The singular part about a pole is sometimes called the *principal part* of the function. The order of the pole is k when $(z-z_0)^{-k}$ is the lowest power (that is, the largest negative power) of $z - z_0$ that appears with non-zero coefficient in the Laurent series.

A zero of a non-constant function f, say at a, corresponds to a pole of $1/f$ at a, and vice versa; the order of the pole a of $1/f$ is the same as the multiplicity of a as a zero of f. A zero of multiplicity 1 is called a *simple* zero; and a pole of order 1 is a *simple* pole. If the multiplicity or order is 2, one speaks of a *double* zero or pole, and so on.

Remark 18.14. From 18.3*(c)*, it follows, as the remaining possibility, that an isolated essential singularity of f is one about which the Laurent series of f has infinitely many negative powers with non-zero coefficients. This can really happen. For instance, $\exp(1/z)$ is a function holomorphic on $\mathbb{C} \setminus \{0\}$, and its Laurent series about 0 is

$$\dots + \frac{z^{-n}}{n!} + \dots + \frac{z^{-2}}{2!} + \frac{z^{-1}}{1!} + 1\,,$$

which has infinitely many non-zero terms in its singular part.

Note 18.15. *(a)* We have not classified *all* singularities, but only *isolated* singularities, which are rather special. If we define $f(z) := z$ for $|z| < 1$, then f has a singularity at any point w with $|w| \geq 1$, simply because it is not defined at w. In this case, we could extend the

definition of f to the holomorphic function z on the whole of $\mathbb{C}$, and so our singularities are 'removable' even though they are not isolated. But it is easy to invent 'genuine' examples of non-isolated singularities. For instance, $\sin z$ has simple zeros at $n\pi$, for any integer n; $\sin(\pi/z)$ therefore has simple zeros at $\pm 1, \pm\frac{1}{2}, \pm\frac{1}{3}, \pm\frac{1}{4}, \dots$, and the function $f(z) = \operatorname{cosec}(\pi/z)$ has a sequence of simple poles tending to 0. The origin is therefore a limit point of simple poles of f, and, as such, is of course a singularity, though not an isolated one.

(b) Consider the series $1 + z + z^2 + z^3 + \dots$, with radius of convergence 1. Its sum $g(z)$ is defined, and is holomorphic, on the unit disk $D = \{z : |z| < 1\}$. Thus g has a singularity, in the sense of 18.2, at every point of the unit circle $\{z : |z| = 1\}$; for the series does not converge at any point of the circle. However, we know that $g(z) = 1/(1 - z)$ for all $z \in D$, and so there is a function, namely $1/(1 - z)$, which agrees with g where possible and is holomorphic on $\mathbb{C} \setminus \{1\}$. Consequently, all points of the unit circle except 1, although they are not isolated as singularities of g, are 'removable', like the singularities of f above, in the unofficial sense that they only arise from an inadequate definition of g. This function exemplifies the principle of 18.12 that the radius of convergence of the Taylor series at 0 is the distance to the nearest 'genuine' singularity, which is the pole of $\dfrac{1}{1 - z}$ at 1.

(c) Let $h(z)$ be the sum of the power series $\sum z^{n!}$, which has radius of convergence 1. $h(z)$ is defined and holomorphic on D, and has a singularity at every point a of the unit circle. However, none of these singularities can be removed by an improved definition of h. To put it precisely, there is no holomorphic function defined on a whole disk U centred at a, and agreeing with h on $U \cap D$. Such a function would extend the definition of h holomorphically across the unit circle, at least near a. (The proof is neither difficult nor obvious, and I omit it.) One then says that the unit circle is a 'natural boundary' for h.

It is a curious fact that any open set V in $\mathbb{C}$ is the domain of some holomorphic function for which $\operatorname{Fr} V$ is a natural boundary, but that the analogous statement in $\mathbb{C}^2$ is false. That is to say, there are open sets V in $\mathbb{C}^2$ such that any holomorphic function (in the two-dimensional sense, which we have not discussed) defined on V must extend to a holomorphic function on a larger open set.

Remark 18.16. *(a)* We can now explain the method of partial fractions in elementary calculus. (I should say that there are also algebraic explanations, but they are more elaborate, although they are also more general.) In a *proper* rational function f of z (f is the quotient of two polynomials, the numerator being of lower degree than the denominator), each zero of the denominator gives a pole of f. Let g be the sum of the principal parts of f at all these poles. Then $f - g$ has only removable singularities in $\mathbb{C}$; remove them. Both f and g tend to 0 as $z \to \infty$ (this is where we assume f is proper). Then, by Liouville's theorem, $f - g$ (after removal of singularities) is constantly 0, which means that $f = g$ except at the poles of f. g is, indeed, the partial fraction expansion of f. All that remains is to find economical methods of calculating the partial fraction expansion in practice.

In 113, it was necessary to stay in the real domain, so that the denominator of the rational function might have irreducible quadratic factors over $\mathbb{R}$. Over $\mathbb{C}$, the fundamental theorem of algebra makes all the irreducible factors linear.

(b) One can also discuss 'singularities at ∞'. For example, $\exp z$ has an isolated essential singularity at ∞, since $\exp(1/\zeta)$ has an isolated essential singularity at $\zeta = 0$. But I shall not discuss the idea of the complex point at infinity.

We have shown that a function holomorphic on $B(a; r)$ (where $a \in \mathbb{C}$ and $r > 0$) is the sum of a Taylor series $\sum_{n=0}^{\infty} c_n (z - a)^n$ with radius of convergence at least r. This has an interesting consequence.

Lemma 18.17. *Let U be a region, and let $f : U \longrightarrow \mathbb{C}$ be holomorphic. For any $a \in \mathbb{C}$, the set $\{z \in U : f(z) = a\}$ is either the whole of U or has no limit points in U (whether it is empty or not). That is: either f is constant on the region U, or, for any $z_0 \in U$, there is a disc $B(z_0; r) \subseteq U$ of positive radius r about z_0 such that, for each $z \in B(z_0; r) \setminus \{z_0\}$, $f(z) \neq f(z_0)$.*

Proof. Given $z_0 \in U$, $g(z) = f(z) - f(z_0)$ is holomorphic on U. It has a Taylor series $\sum c_n(z - z_0)^n$ about z_0, convergent on $B(z_0; r) \subseteq U$ for some $r > 0$. As $g(z_0) = 0$, certainly $c_0 = 0$. There are two possibilities.

(a) *All* coefficients c_n are zero. Then $g(z) = 0$ for all $z \in B(z_0; r)$. If z_1 is any other point of U, let $\gamma : [0, 1] \longrightarrow U$ be a continuous path with $\gamma(0) = z_0$ and $\gamma(1) = z_1$. Such a path exists—recall 11.2. Let

$$V \coloneqq \{t \in [0, 1] : (\forall s \in [0, t])(\forall n \in \mathbb{N} \cup \{0\})\ g^{(n)}(\gamma(s)) = 0\},$$

which is nonempty (for $0 \in V$) and bounded above by 1. Let $t_1 \coloneqq \sup V$. We have shown that g is 0 for $|z - z_0| < r'$, so that $t_1 > 0$. But, by continuity of $g^{(n)}$ on U, $g^{(n)}(\gamma(t_1)) = 0$ for all n (for t_1 is the limit of a sequence in V; apply 2.11*(a)* and 5.5). Thus, in fact, the Taylor series of g about $\gamma(t_1)$ is also zero; g is zero on $B(\gamma(t_1); r') \subseteq U$ for some $r' > 0$, and it follows that t_1 cannot be the supremum of V unless $t_1 = 1$. Therefore, $g(\gamma(t)) = 0$ for $0 \leq t \leq 1$. In particular, $g(z_1) = g(\gamma(t_1) = 0$, that is, $f(z_1) = f(z_0)$. This holds for any $z_1 \in U$. It follows that f is constant on U.

(b) Alternatively, there is a first nonzero coefficient, say c_k, where $k > 0$. Then $g(z) = (z - z_0)^k \phi(z)$, with $\phi(z) = \sum_{n=0}^{\infty} c_{n+k}(z - z_0)^k$. (This series also converges absolutely for $|z - z_0| < r$, and there defines a holomorphic function.) But, as $\phi(z_0) \neq 0$ and ϕ is continuous, there is $r'' \in (0, r)$ such that $\phi(z) \neq 0$ when $|z - z_0| < r''$, and $g((z) = (z - z_0)^k \phi(z) \neq 0$ if $0 < |z - z_0| < r''$. This proves z_0 is not a limit point of $\{z \in U : g(z) = 0\}$, and, because of *(a)*, this conclusion must hold for any $z_0 \in U$ if f is not constant. □

Corollary 18.18. *If f is holomorphic on the region U, and $a \in \mathbb{C}$ is such that $\{z \in U : f(z) = a\}$ has a limit point in U, then $f(U) = \{a\}$.* □

Of course $\{z \in U : f(z) = a\}$ can have a limit point only if it is nonempty. And it is important that the limit point should be in U (not in $\mathbb{C} \setminus U$, that is, in effect, in $\operatorname{Fr} U$). For example, $\sin(\pi/z)$, holomorphic on $\mathbb{C} \setminus \{0\}$, has zeros at the points $1/n$ for all nonzero integers n, and these points have 0 as a limit point; but $\sin(\pi/z)$ is not constant on $\mathbb{C} \setminus \{0\}$.

Once again we see that holomorphic functions are very "rigid"; they are fixed completely (on regions in $\mathbb{C}$) by surprisingly weak data. Indeed, if g_1, g_2 are holomorphic on the region U and $\{z \in U : g_1(z) = g_2(z)\}$ has a limit point in U (I stress "in U"), then g_1 and g_2 agree at all points of U. (Take $f \coloneqq g_1 - g_2$ and $a \coloneqq 0$ in 18.18.)

§19. Residues.

Definition 19.1. Let b be an isolated singularity of the function f, with a Laurent series

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - b)^n, \tag{99}$$

which converges when $0 < |z-b| < \delta$, for some $\delta > 0$. The *residue* of f at b, which I shall write $\operatorname{res}(f,b)$, is the coefficient of $(z-b)^{-1}$ in the series, that is, a_{-1}.

Lemma 19.2. *Suppose that the function f is holomorphic on the punctured disk $\{z : 0 < |z-b| < R\}$. Then, for any number $r \in (0,R)$, and for the circular contour $\gamma[r](t) = b + r\exp(it)$, $0 \le t \le 2\pi$,*

$$\int_{\gamma[r]} f = 2\pi i \operatorname{res}(f,b).$$

Proof. This is just the case $n = -1$ of the formula (92) of 18.7. □

Indeed, only the term in $(z-b)^{-1}$ in (99) is not a derivative; all the other terms have zero integral around $\gamma[r]$ by 15.9.

Theorem 19.3. (The residue theorem). *Let $\mathcal{C}$ be a cycle homologous to 0 in the region U in $\mathbb{C}$. Suppose the distinct points $a_1, a_2, \ldots, a_n \in U$ do not lie on $\mathcal{C}$, and that f is a function holomorphic on $U \setminus \{a_1, a_2, \ldots, a_n\}$. Then*

$$\int_{\mathcal{C}} f = 2\pi i \sum\nolimits_{k=1}^{n} n(\mathcal{C}, a_k) \operatorname{res}(f, a_k). \tag{100}$$

Note 19.4. The residue theorem is stated in the literature in many different ways—as one would expect, since it is deduced from the equally polymorphous theorem of Cauchy. The above version is more general than one needs for almost all applications.

Proof. Choose $\delta > 0$ to be so small that the disks $B(a_k; 2\delta)$ are disjoint from each other and included in U. Around each a_k, construct a small circular Jordan contour γ_k of radius δ. Then the cycle $\sum n(\mathcal{C}, a_k)\gamma_k$ is homologous in $U \setminus \{a_1, \ldots, a_n\}$ to $\mathcal{C}$ (around any point outside U, both have index 0; and around each a_k, their indices are equal by definition). Thus, by Cauchy's theorem in the form discussed at 16.11,

$$\int_{\mathcal{C}} f = \sum n(\mathcal{C}, a_k) \int_{\gamma_k} f = 2\pi i \sum n(\mathcal{C}, a_k) \operatorname{res}(f, a_k)$$

by 19.2; and this is the result. □

Therefore, the value of an integral around a cycle may be calculated by finding residues. Using formula (92) to find the residue would be a circular argument; one must employ other methods of developing Laurent series. Here are some examples.

Example 19.5. From 13.3, $\sin(x+iy) = \sin x \cosh y + i \cos x \sinh y$, and

$$\begin{aligned} |\sin z|^2 &= \sin^2 x \cosh^2 y + \cos^2 x \sinh^2 y \qquad (\text{here } z = x+iy) \\ &= \sin^2 x\,(1 + \sinh^2 y) + \cos^2 x \sinh^2 y \\ &= \sin^2 x + \sinh^2 y, \quad \text{as } \sin^2 + \cos^2 = 1. \end{aligned} \tag{101}$$

Hence, $\sin z = 0$ if and only if $\sin x = 0$ and $\sinh y = 0$, and the only zeros of the complex sine function are those which are already zeros of the real sine function (that is, those on the

real axis): $z = n\pi$, for integers π. Likewise,

$$\begin{aligned} |\cos z|^2 &= \cos^2 x \cosh^2 y + \sin^2 x \sinh^2 y \\ &= \cos^2 x\,(1 + \sinh^2 y) + \sin^2 x \sinh^2 y \\ &= \cos^2 x + \sinh^2 y\,. \end{aligned} \tag{102}$$

Again, all the zeros are on the real axis, $z = (n + \frac{1}{2})\pi$ for $n \in \mathbb{Z}$.

The Taylor (or Laurent) expansion of sin about $n\pi$ is of the form

$$\sin z = a_0 + a_1(z - n\pi) + a_2(z - n\pi)^2 + \cdots,$$

where, since $\sin(n\pi) = 0$, necessarily $a_0 = 0$. But, if we write $z = \zeta + n\pi$,

$$\begin{aligned} \sin z &= \sin(n\pi + \zeta) \\ &= \sin(n\pi)\cos\zeta + \cos(n\pi)\sin\zeta = \cos(n\pi)\sin\zeta \\ &= (-1)^n\left(\zeta - \frac{\zeta^3}{3!} + \frac{\zeta^5}{5!} - \cdots\right). \end{aligned}$$

So the required Taylor expansion is

$$\begin{aligned} \sin z = (-1)^n(z - n\pi) + (-1)^{n+1}\frac{(z - n\pi)^3}{3!} \\ + (-1)^{n+2}\frac{(z - n\pi)^5}{5!} + \cdots. \end{aligned} \tag{103}$$

Again, this must be the expansion, because it is a power series expansion for sin about $n\pi$ and there can be only one such.

Example 19.6. The complex cosecant function $\operatorname{cosec} z$ is defined exactly as in the real case to be the reciprocal of $\sin z$. It follows that $\operatorname{cosec} z$ is holomorphic and non-zero except at the isolated singularities $n\pi$; as they are isolated zeros of sin, they are poles of cosec. Furthermore, 19.5 shows that the zeros of sin are simple:

$$\sin z = (z - n\pi)(-1)^n\left\{1 - \frac{1}{3!}(z - n\pi)^2 + \cdots\right\}.$$

Thus $n\pi$ is a simple pole of cosec, and

$$\operatorname{cosec} z = \frac{(-1)^n}{z - n\pi}\left\{1 + \frac{1}{3!}(z - n\pi)^2 + \cdots\right\}, \tag{104}$$

so that the residue at this pole is $(-1)^n$. (Only odd powers appear in the series; I am using a general theorem *that I have not stated or proved* about substitution of power series in power series. In this case, I appeal to $(1 - w)^{-1} = 1 + w + w^2 + \cdots$, which has radius of convergence 1, and take $w = \frac{1}{3!}\zeta^2 - \frac{1}{5!}\zeta^4 + \cdots$, which converges for all ζ. The ellipsis $\cdots$ in (104) indicates terms, all even powers, that are not easily calculated.)

Remark 19.7. A function f has a pole of order $k \geq 1$ at the isolated singularity b,

$$f(z) = a_{-k}(z-b)^{-k} + a_{-k+1}(z-b)^{-k+1} + \cdots,$$

with $a_{-k} \neq 0$, if and only if $(z-b)^k f(z)$ has a finite non-zero limit as $z \to b$; and this limit is a_{-k}. In particular, $(z-b)f(z)$ has a non-zero limit as $z \to b$ if and only if b is a simple pole of f, and then the residue of f at b is $\lim_{z\to b}(z-b)f(z)$.

For a pole of order $k > 1$, the residue a_{-1} is the coefficient of $(z-b)^{k-1}$ in the Taylor expansion about b of $(z-b)^k f(z)$. As such it is

$$\lim_{z\to b} \frac{1}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}((z-b)^k f(z)).$$

This formula is rather less useful than the one for a simple pole.

Example 19.8. Let $f(z) = \pi\cot(\pi z)$. It has singularities at the zeros of $\sin(\pi z)$, which are the real integers. Further, f is periodic with period 1; that is, $f(z+1) = f(z)$ for all z:

$$\cos(\zeta+\pi) = -\cos\zeta \quad \text{and} \quad \sin(\zeta+\pi) = -\sin\zeta.$$

Hence the Laurent expansion of f about any of the singularities $n\pi$ will have exactly the same coefficients as the expansion about 0. Now

$$\pi z\cot(\pi z) = \frac{\pi z}{\sin(\pi z)}\cos(\pi z) \to 1.1 \quad \text{as } z \to 0.$$

(The standard limits for $\sin z/z$ and for $\cos z$ as $z \to 0$ hold in the complex case—indeed, they follow directly from the series expansions of sin and cos.) So the singularities of f are simple poles, one at each integer, with residue 1.

Remark 19.9. *(a)* Let f and g be functions defined, non-constant, and holomorphic on $B(a;\delta)\setminus\{a\}$, where a is a pole or removable singularity of both f and g; let $\omega(f)$ denote the least power of $z-a$ that has non-zero coefficient in the Laurent series for f about a, and similarly for $\omega(g)$. It is evident that

$$\omega(fg) = \omega(f) + \omega(g), \tag{105}$$

for the leading term of the product Laurent expansion is the product of the leading terms of the expansions of f and of g. Similarly, if $\omega(f) \neq \omega(g)$,

$$\omega(f+g) = \min(\omega(f), \omega(g)).$$

(If $\omega(f) = \omega(g)$, there may be cancellations, so $\omega(f+g) \geq \omega(f) = \omega(g)$).

From (105), it is clear that a pole of f may be 'cancelled out' by a zero of high enough multiplicity, and vice versa.

(b) Generally, the residue of fg at a is a sum of several products of coefficients of the Laurent expansions, but if $\omega(fg) = -1$ it will be just the product of the leading coefficients, i.e. the coefficients of the lowest powers that have non-zero coefficients. In particular, if f has a simple pole at a, and g has a removable singularity at a with limit $q \neq 0$, then fg has a simple pole at a, with residue $q\,\mathrm{res}(f,a)$.

Briefly, $\mathrm{res}(fg,a) = g(a)\,\mathrm{res}(f,a)$, provided the pole of f at a is simple and g is holomorphic near a (if $g(a) = 0$ this statement remains true).

Lemma 19.10. *Suppose the function f is holomorphic on $B(a;\delta)\setminus\{a\}$, where $\delta>0$, and has a simple pole at a. If $\gamma[r,\kappa](\theta) := a + r\exp(i\theta)$ for $0\le\theta\le\kappa$, where $0<r<\delta$, then $\dfrac{1}{r}\left\{\displaystyle\int_{\gamma[r,\kappa]} f(z)\,dz - \kappa i\,\mathrm{res}(f,a)\right\}$ is bounded as $r\to 0$.*

(For a simple pole a, the integral around a segment of a circle centred at a is approximately i times the angle of the segment times the residue, the error in the approximation tending to 0 as rapidly as the radius r as $r\to 0$. Of course the integral around the whole circle is exactly $2\pi i\,\mathrm{res}(f,a)$, independently of the radius, provided it is small enough.)

Proof. The Laurent series, valid for $0<|z-a|<\delta$, is of the form

$$f(z) = \frac{c_{-1}}{z-a} + c_0 + c_1(z-a) + c_2(z-a)^2 + \cdots,$$

where $c_{-1} = \mathrm{res}(f,a)$. Define

$$g(z) := c_0 + c_1(z-a) + \cdots,$$

the regular part of the series. Then $g(z)\to c_0$ as $z\to a$; there is some number $K>0$ such that $|g(z)|\le K$ whenever $|z-a|\le\frac{1}{2}\delta$, for example. Now

$$\int_{\gamma[r,\kappa]} \frac{c_{-1}}{z-a}\,dz = \int_0^\kappa \frac{c_{-1}\,i\exp(i\theta)}{\exp(i\theta)}\,d\theta = i\kappa c_{-1}, \quad \text{and}$$

$$\left|\int_{\gamma[r,\kappa]} f(z)\,dz - i\kappa\,\mathrm{res}(f,a)\right| = \left|\int_{\gamma[r,\kappa]} g(z)\,dz\right| \le K\kappa r$$

by the fundamental estimate, as the length of $\gamma[r,\kappa]$ is κr. □

This result can be very useful if the contour of integration appears to have to pass through a simple pole. We shall see examples.

Example 19.11. Squaring the Laurent expansion of $\mathrm{cosec}\,z$ about $n\pi$ from (104):

$$\mathrm{cosec}^2 z = (z-n\pi)^{-2}\left\{1 + \frac{2}{3!}(z-n\pi)^2 + \cdots\right\},$$

so that the residue of $\mathrm{cosec}^2 z$ at $n\pi$ is in fact 0. As $\cot^2 z = \mathrm{cosec}^2 z - 1$, the residue of $\cot^2 z$ is also 0. By contrast, the residue of $\mathrm{cosec}^3 z$ is $\frac{1}{2}$, by a similar argument.

Remark 19.12. The residue theorem is the crown of the course, because of its frequent use to evaluate difficult definite integrals and series. But it does not really constitute a *method*, despite phrases such as "the method of residues" or "the calculus of residues". There are, indeed, large classes of integral to which it can be applied in standard ways, but it is also unpredictably useful in many other situations. It often leaves much scope for individual ingenuity, both in the choice of contour and in the manipulation of the integrand. It will be best here to present some few sadly abbreviated examples.

Example 19.13. Let $k \in \mathbb{N}$. The function

$$f(z) = \frac{\pi \cot(\pi z)}{z^{2k}}$$

has simple poles at all non-zero integers, by 19.8 and 19.9, and the residue at $n \neq 0$ is $1/n^{2k}$. At 0 it has a pole of order $2k+1$. For $N \in \mathbb{N}$, let $\Gamma(N)$ be the square, centred at 0, with sides parallel to the axes at $x = \pm(N + \frac{1}{2})$ and $y = \pm(N + \frac{1}{2})$.

$$\begin{aligned} |\cot(\pi(x+iy))|^2 &= \left| \frac{\cos(\pi x)\cosh(\pi y) - i\sin(\pi x)\sinh(\pi y)}{\sin(\pi x)\cosh(\pi y) + i\cos(\pi x)\sinh(\pi y)} \right|^2 \\ &= \frac{\cos^2(\pi x) + \sinh^2(\pi y)}{\sin^2(\pi x) + \sinh^2(\pi y)} \qquad \text{(cf. (101) and (102)).} \end{aligned}$$

It follows that, when $y = \pm(N + \frac{1}{2})$, and for any value of x,

$$|\cot(\pi(x+iy))|^2 \leq \frac{1 + \sinh^2(\pi(N+\frac{1}{2}))}{0 + \sinh^2(\pi(N+\frac{1}{2}))} < 2,$$

since $\sinh^2(\pi(N+\frac{1}{2})) \geq \sinh^2(\frac{1}{2}\pi) > 1$ for all $N \in \mathbb{Z}$. If $x = \pm(N+\frac{1}{2})$, $\cos(\pi x) = 0$ and $\sin^2(\pi x) = 1$, so that

$$|\cot(\pi(x+iy))|^2 = \frac{\sinh^2(\pi y)}{1 + \sinh^2(\pi y)} < 1.$$

Hence $|\pi \cot(\pi z)| < \pi\sqrt{2}$ at all points of $\Gamma(N)$. In addition, $|z| \geq N + \frac{1}{2}$ at all such points. The length of $\Gamma(N)$ is $8(N+\frac{1}{2})$. By the fundamental estimate (74),

$$\begin{aligned} \left| \int_{\Gamma(N)} f(z)\,dz \right| &\leq \frac{\pi\sqrt{2}.\,8(N+\frac{1}{2})}{(N+\frac{1}{2})^{2k}} \\ &= 8\pi\sqrt{2}(N+\tfrac{1}{2})^{1-2k}. \end{aligned}$$

As $k \geq 1$, this shows that $\int_{\Gamma(N)} f \to 0$ as $N \to \infty$.

The poles inside Γ are at the real integers $-N, -N+1, \ldots, N-1, N$, and the residue of f at $\pm n$, where $n \in \mathbb{N}$, is n^{-2k}, by 19.8 and 19.9*(b)*. The residue theorem gives

$$\begin{aligned} \int_{\Gamma(N)} f(z)\,dz &= 2\pi i \sum\nolimits_{n=-N}^{N} \operatorname{res}(f,n) \\ &= 2\pi i \operatorname{res}(f,0) \;+\; 4\pi i \sum\nolimits_{n=1}^{N} n^{-2k}. \end{aligned}$$

As $N \to \infty$, then,

$$\sum\nolimits_{n=1}^{N} \frac{1}{n^{2k}} + \tfrac{1}{2}\operatorname{res}(f,0) = \frac{1}{4\pi i}\int_{\Gamma(N)} f \to 0,$$

from which we draw the simple consequence that

$$\sum\nolimits_{n=1}^{\infty} \frac{1}{n^{2k}} = -\tfrac{1}{2}\operatorname{res}(f,0).$$

In principle, then, to calculate $1 + \dfrac{1}{2^{2k}} + \dfrac{1}{3^{2k}} + \cdots$, all we have to do is to find the residue of $\dfrac{\pi \cot(\pi z)}{z^{2k}}$ at 0. But this is not trivial.

To find the Laurent expansion of $\pi \cot(\pi z)$ about 0, argue as follows.

$$\begin{aligned}
\pi \cot(\pi z) = \frac{\pi \cos(\pi z)}{\sin(\pi z)} &= \frac{\pi(1 - \frac{1}{2!}\pi^2 z^2 + \frac{1}{4!}\pi^4 z^4 - \dots)}{\pi z - \frac{1}{3!}\pi^3 z^3 + \frac{1}{5!}\pi^5 z^5 - \dots} \\
&= z^{-1}\left(1 - \frac{1}{2!}\pi^2 z^2 + \frac{1}{4!}\pi^4 z^4 - \dots\right) \times \\
&\qquad \times \left(1 - \frac{1}{3!}\pi^2 z^2 + \frac{1}{5!}\pi^4 z^4 - \dots\right)^{-1} \\
&= z^{-1}\left(1 - \frac{1}{2!}\pi^2 z^2 + \frac{1}{4!}\pi^4 z^4 - \dots\right) \\
&\qquad \times \left\{1 + \left(\frac{1}{3!}\pi^2 z^2 - \frac{1}{5!}\pi^4 z^4 + \dots\right)\right. \\
&\qquad\qquad \left. + \left(\frac{1}{3!}\pi^2 z^2 - \frac{1}{5!}\pi^4 z^4 + \dots\right)^2 + \dots\right\}
\end{aligned}$$

(as $\dfrac{1}{1-w} = 1 + w + w^2 + \dots$ for $|w| < 1$). $\operatorname{res}(f, 0)$ is the coefficient of z^{2k} (the pole is of order $2k$) in the expansion in powers of z of

$$\begin{aligned}
&\left(1 - \frac{1}{2!}\pi^2 z^2 + \frac{1}{4!}\pi^4 z^4 - \dots\right) \\
&\qquad \times \left\{1 + \left(\frac{1}{3!}\pi^2 z^2 - \frac{1}{5!}\pi^4 z^4 + \dots\right)\right. \\
&\qquad\qquad \left. + \left(\frac{1}{3!}\pi^2 z^2 - \frac{1}{5!}\pi^4 z^4 + \dots\right)^2 + \dots\right\}
\end{aligned}$$

It is impossible to give a general formula for this coefficient. However, the coefficient of z^2 is clearly

$$-\frac{1}{2!}\pi^2 + \frac{1}{3!}\pi^2 = -\frac{1}{3}\pi^2 ,$$

and of z^4 is $\left\{\dfrac{1}{4!} - \dfrac{1}{5!} - \dfrac{1}{2!\,3!} + \left(\dfrac{1}{3!}\right)^2\right\}\pi^4 = -\dfrac{1}{45}\pi^4$. Thus

$$\sum\nolimits_{n=1}^{\infty} n^{-2} = \tfrac{1}{6}\pi^2 \quad \text{and} \quad \sum\nolimits_{n=1}^{\infty} n^{-4} = \tfrac{1}{90}\pi^4 .$$

The computations grow rapidly in difficulty as k increases.

Remark 19.14. The method as given above cannot yield the sums of odd or fractional powers of n. Having even powers, we could extract useful information from a contour of integration that was symmetrical about the origin. Much the same method can be used for such series as $\sum (n^2 + 1)^{-1}$, and a similar trick with $\pi \operatorname{cosec}(\pi z)$ can be done to sum alternating series of the same general type.

The sum $\sum n^{-2} = \frac{1}{6}\pi^2$ can also be deduced from an argument with Fourier series (as, indeed, can some other sums of the kind), and possibly in other ways as well. The Fourier series proof is sometimes given in second-year courses; it is legitimate, but relies at least on Dirichlet's non-trivial result on the convergence of Fourier series, which is not proved in our courses. Our proof above, by contrast, is in principle completely rigorous (granted Cauchy's theorem), except for some minor technicalities which I have not discussed in full. (Dirichlet's proof, by the way, is not really more difficult; it is just different.)

Example 19.15. The Cauchy-Riemann integral $\displaystyle\int_0^\infty \frac{x^{2/3}\,dx}{x^2+x+1}$ is defined as the limit of Riemann integrals $\displaystyle\int_0^R \frac{x^{2/3}\,dx}{x^2+x+1}$ as $R \uparrow \infty$. Here the cube root has its customary real value. That the limit exists can be proved directly by the "comparison test" for integrals (another thing we have not discussed in this course), but it will also result from our arguments below.

The integral may be seen as a 'contour integral' along the 'infinite contour' of the positive real axis. To try to apply the residue theorem, we must first extend the integrand to a holomorphic function, and for that we need to consider fractional powers of complex z.

If α is not an integer, z^α is not well-defined (even $z^{1/2}$ is "two-valued" for $z \neq 0$). We must define a holomorphic branch of z^α, i.e. an appropriate holomorphic function f, defined on a suitable region, which satisfies $zf'(z) = \alpha f(z)$ at all points of the region. This suggests a relation with branches of the logarithm (see 14.8), which satisfy the differential equation $f'(z) = 1/z$. Following this hint, we define z^α to mean $\exp(\alpha \log_1 z)$ where $\log_1$ is a branch of the logarithm chosen to suit our convenience.

[The easiest way to define fractional powers of a positive real number x is by setting

$$x^\alpha := \exp(\alpha \log x)\,,$$

where $\log : (0,\infty) \longrightarrow \mathbb{R}$ is the inverse of the exponential function. This "logarithm" is also the restriction to the positive real axis of the principal value of the complex logarithm.]

In our case, it is essential that $z^{2/3}$ should be real and positive when z is; but it may also help if it is holomorphic on the negative real axis. So let us define $\log_1 z$ by $\log_1 z := \log|z| + i \arg_1 z$, where $\arg_1 z \in (-\frac{1}{2}\pi, \frac{3}{2}\pi)$. Then $z^{2/3}$ is defined and holomorphic except on the negative imaginary axis, where its argument must jump from $\frac{2}{3}\frac{3}{2}\pi$ to $-\frac{2}{3}\frac{1}{2}\pi$ as the axis is crossed in the anticlockwise direction, and at the origin.

Now $f(z) = \dfrac{z^{2/3}}{z^2+z+1}$ is holomorphic except on the nonpositive imaginary axis, and also at the zeros of z^2+z+1, which are the non-trivial cube roots of 1. The only singularity in the upper half-plane is therefore at $\omega = \mathrm{cis}(2\pi/3) = -\frac{1}{2} + \frac{1}{2}i\sqrt{3}$. Since

$$f(z) = \frac{z^{2/3}}{z^2+z+1} = \frac{z^{2/3}}{(z-\omega)(z-\omega^2)}\,,$$
$$\lim_{z\to\omega}(z-\omega)f(z) = \frac{\omega^{2/3}}{\omega-\omega^2}\,,$$

there is a simple pole at ω with residue $\omega^{2/3}/(\omega-\omega^2)$, and $\omega - \omega^2 = i\sqrt{3}$.

Let $\gamma[R,r]$ consist of a semicircle of large radius R centred at the origin, starting at R and going anticlockwise to $-R$, followed by the segment $(-R,-r)$ of the real axis (where r is a small positive number), by a small *clockwise* semicircle of radius r about the origin, and by the

segment (r, R) of the real axis. The small semicircle avoiding the origin is needed because of the lack of holomorphicity there. The integral round $\gamma[R, r]$ will be

$$2\pi i\omega^{2/3}/(\omega - \omega^2) = 2\pi i \operatorname{cis}(\tfrac{4}{9}\pi)/(i\sqrt{3}),$$

by the residue theorem and our definition for the fractional power.

The fundamental estimate (74) for the integral round the large semicircle is $\pi R.R^{2/3}\big/(R^2 - R - 1)$; round the small semicircle there is a similar estimate $\pi r.r^{2/3}/(1 - r - r^2)$. Both estimates may be made as small as we wish by taking large enough R and small enough r. The remaining parts of the integral over γ are $\int_{-R}^{-r} f(x)\,dx = \int_r^R f(-u)\,du$ and $\int_r^R f(x)\,dx$. For $u > 0$, our definition of f implies that

$$f(-u) = u^{2/3}\exp(\tfrac{2}{3}\pi i)/(u^2 - u + 1),$$

and therefore in all

$$\begin{aligned}\int_{\gamma[R,r]} f(z)\,dz &= \int_r^R \Big\{\frac{1}{x^2 + x + 1} + \frac{\exp(\frac{2}{3}\pi i)}{x^2 - x + 1}\Big\} x^{2/3}\,dx + Q(R, r)\\ &= \frac{2\pi \operatorname{cis}(\frac{4}{9}\pi)}{\sqrt{3}};\end{aligned} \tag{106}$$

the remainder term Q consists of the integrals over the semicircles, which, as we observed, may be arbitrarily small if R is sufficiently large and r sufficiently small. Thus, in the limit, taking the imaginary parts on both sides of (106),

$$\begin{aligned}\sin(\tfrac{2}{3}\pi)\int_0^\infty \frac{x^{2/3}\,dx}{x^2 - x + 1} &= \frac{2\pi \sin(\frac{4}{9}\pi)}{\sqrt{3}}, \quad \text{which reduces to}\\ \int_0^\infty \frac{x^{2/3}\,dx}{x^2 - x + 1} &= \frac{4\pi \sin(\frac{4}{9}\pi)}{3}.\end{aligned} \tag{107}$$

Taking the real parts in (106) and using (107), we obtain what we wanted:

$$\begin{aligned}\int_0^\infty \frac{x^{2/3}\,dx}{x^2 + x + 1} &= \frac{2\pi \cos(\frac{4}{9}\pi)}{\sqrt{3}} - \cos(\tfrac{2}{3}\pi).\frac{4\pi \sin(\frac{4}{9}\pi)}{3}\\ &= \frac{4\pi}{3}\big(\sin(\tfrac{2\pi}{3})\cos(\tfrac{4}{9}\pi) - \cos(\tfrac{2}{3}\pi)\sin(\tfrac{4}{9}\pi)\big)\\ &= \frac{4\pi \sin(\frac{2}{9}\pi)}{3}.\end{aligned}$$

This integral can in fact be evaluated by methods taught in MATH 142, but not easily. You substitute $x = t^3$ and then use partial fractions.

Example 19.16. As a last example, consider

$$\int_0^\infty \frac{x \sin(sx)}{1 + x^2}\,dx$$

for $s > 0$.

This is a conditionally convergent Cauchy-Riemann integral, and we need a subtler argument than before. Take $f(z) := z\exp(isz)/(1+z^2)$, and let our closed contour γ consist of the interval $[-R,R]$ on the real axis and the semicircle $C(R)$ of radius $R>0$ described anticlockwise from R to $-R$. Thus, $\int_\gamma f = \int_C f + \int_{-R}^{R} f(x)\,dx$. For real x, the real part of $f(x)$ is $x\cos(sx)/(1+x^2)$, which is an odd function, and the imaginary part, $x\sin(sx)/(1+x^2)$, is even; hence,

$$\int_{-R}^{R} f(x)\,dx = 2i\int_0^R \frac{x\sin(sx)\,dx}{1+x^2}\,.$$

For large R, the only singularity of f enclosed by γ is a simple pole at $z=i$, where the residue is $i\exp(-s)/(2i) = \frac{1}{2}e^{-s}$ by 19.9 (for

$$\frac{1}{1+x^2} = \frac{1}{2i}\left(\frac{1}{i-z}+\frac{1}{i+x}\right)$$

has a simple pole with residue $1/(2i)$). Thus, if we can show that the integral around $C(R)$ tends to 0 as $R\to\infty$, we can deduce that the limit exists:

$$2i\int_0^\infty \frac{x\sin(sx)\,dx}{1+x^2} = 2\pi i.\tfrac{1}{2}e^{-s}\,,$$
$$\int_0^\infty \frac{x\sin(sx)\,dx}{1+x^2} = \tfrac{1}{2}\pi e^{-s}\,.$$

Remark 19.17. The proof that the integral over $C(R)$ tends to 0 must be different from the argument in 19.13 or the similar reasoning in 19.15. In those cases the modulus of the integrand had a bound $R^{-\beta}$, for some $\beta>1$, and as the length of the contour is proportional to R this showed the integral tends to 0 like $R^{1-\beta}$. Now, however, $f(z) = ze^{isz}/(1+z^2)$; when z is near to the real axis, $|e^{isz}|$ is close to 1, and $|f(z)|$ is approximately R^{-1} for large R. As $C(R)$ is of length πR, the "fundamental estimate" alone cannot show the integral tends to 0 as $R\to\infty$.

However, a slightly more delicate argument does work. The fundamental estimate (74) uses the "maximum modulus" M, but the integrand may sufficiently often be significantly smaller than M to make (73) a better estimate. The Lemma below is rather special, but is often useful and has earned a name.

Lemma 19.18. (Jordan's Lemma). *Suppose that $s,T>0$, that $f(z)$ is defined and continuous when $\Im z\ge 0$ and $|z|>T$, and that the maximum value $M(R)$ of $|f(z)|$ on $C(R)$ (defined when $R>T$) tends to 0 as $R\to\infty$. Then $\int_{C(R)} e^{isz}f(z)\,dz\to 0$ as $R\to\infty$.*

Proof. As $|f(z)|$ is defined and continuous on the compact set $C(R)$ when $R>T$, it is indeed bounded. Since $|\exp z| = \exp(\Re z)$, (73) gives

$$\left|\int_{C(R)} e^{isz} f(z)\,dz\right| \le \int_0^\pi |\exp(isR(\cos\theta + i\sin\theta)|\, M(R) R\,d\theta$$
$$= RM(R)\int_0^\pi \exp(-sR\sin\theta)\,d\theta\,. \tag{108}$$

But now $\sin\theta = \sin(\pi-\theta)$, so that

$$\int_0^\pi \exp(-sR\sin\theta)\,d\theta = 2\int_0^{\pi/2} \exp(-sR\sin\theta)\,d\theta\,.$$

For $0 \le \theta \le \pi/2$, $\sin\theta \ge 2\theta/\pi$. (The easiest way to *convince* yourself is graphical: the line $y = 2x/\pi$ lies underneath the curve $y = \sin x$. But how would you *prove* it formally? There are many possible proofs; try to find one.) So the right-hand side is not greater than

$$2\int_0^{\pi/2} \exp(-2sR\theta/\pi)\,d\theta = \frac{2\pi}{2sR}[-\exp(-2sR\theta/\pi)]_0^{\pi/2}$$
$$= \frac{\pi}{sR}\,(1 - e^{-sR}) < \frac{\pi}{sR}\,.$$

Hence $\left|\int_{C(R)} e^{isz} f(z)\,dz\right| \le \dfrac{\pi M(R)}{s}$, and the desired result follows. □

Returning now to 19.16, one may take $f(z) = z/(1+z^2)$, for which

$$M(R) \le R/(R^2-1) \quad \text{if } R > 1\,;$$

so $M(R) \to 0$ as $R \to \infty$. From Jordan's Lemma,

$$\int_{C(R)} \frac{ze^{isz}\,dz}{1+z^2} \to 0 \quad \text{as } R \to \infty\,.$$

This completes the proof in 19.16 that $\displaystyle\int_0^\infty \frac{x\sin(sx)\,dx}{1+x^2} = \tfrac{1}{2}\pi e^{-s}$. □

Remark 19.19. I commented that the "calculus of residues" does not altogether constitute a *method*, and 19.16 perhaps illustrates the point. We changed the integrand to $\dfrac{z\exp(isz)}{1+z^2}$, although it might have seemed more natural to take $\dfrac{z\sin(sz)}{1+z^2}$, because the estimates on $C(R)$ would not have worked for a sine. Similarly, in 19.13, the use of the cotangent is probably rather unexpected. It is, in fact, the only function that could be used in exactly the desired way, because it has simple poles with residue 1 exactly where we want and is bounded on the frontiers of all the squares $\Gamma(N)$. (A small modification of Liouville's theorem is relevant.)

APPENDIX A

Remarks on convergence

The general idea of the definition of convergence of a sequence given at 2.9 dates back at least to Wallis in 1655. It is perhaps the oldest thing in the course, and is possibly the most difficult concept to grasp. The remarks below (some of them are also in the main text) are intended to convince you that, however strange it may seem at first, it is in fact the natural expression of our intuitive ideas of a limit.

When we say that a sequence (ξ_n) converges, or *tends*, to a limit ξ (which we write as $\xi_n \to \xi$ or as $\lim_{n\to\infty} \xi_n = \xi$) we have in mind the vague intuitive idea that

(A) ξ_n gets closer and closer to ξ as n gets larger and larger.

Thus, for instance, 2^{-n} gets closer and closer to 0 as n increases: $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$.

As a precise definition, however, (A) will not do. To begin with, what do we mean by 'closer and closer'? Presumably, we are describing the behaviour of the *distance* of ξ_n from ξ, which is $|\xi_n - \xi|$. That ξ_n gets closer and closer to ξ as n increases would then mean that this distance gets steadily smaller and smaller as n increases:

(B) for each n, $|\xi_{n+1} - \xi| \le |\xi_n - \xi|$.

But that is not good enough to fit all the cases we have in mind; for instance, $1 + 2^{-n}$ gets closer and closer, in this sense, to 0, but we do not want to say it converges to 0, because we believe its limit should be 1. Or, the sequence $\frac{1}{3}, \frac{1}{2}, \frac{1}{9}, \frac{1}{4}, \frac{1}{27}, \frac{1}{8}, \dots$ ought, we feel, to have the limit 0, but it does not *decrease* steadily in absolute value.

Clearly, then, what we really want is not that the distance $|\xi_n - \xi|$ should *decrease*. The important thing is that it should, in the course of time, become "arbitrarily small", or "as small as we like". However we define smallness, it must eventually be as small as that. Furthermore, the distance must be *consistently* "small" in our chosen sense, for all large enough values of n. A sequence such as $\xi_n := 2^{-n}$ for even values of n, $\xi_n := 1$ for odd n, ought not to converge to 0, because ξ_n never settles down within distance $\frac{1}{2}$ of 0 —every second term is at distance 1 from 0.

Suppose we choose a positive number ϵ to be our measure of smallness; $|\xi_n - \xi|$ is "small" if it is smaller than ϵ. Then what we want is that $|\xi_n - \xi| \le \epsilon$ for all *large enough* values of n, which presumably means for all n at and after some definite stage N. That is, there is some N such that $n \ge N \implies |\xi_n - \xi| \le \epsilon$. And finally, we want the same to be true whatever ϵ we choose as our measure of smallness. In short,

(C) for any $\epsilon > 0$, there is some N such that

$$|\xi_n - \xi| \le \epsilon \quad \text{whenever} \quad n \ge N .$$

This is precisely the DEFINITION as I gave it: symbolically,

(D) $(\forall \epsilon > 0)(\exists N \in \mathbb{R}) \quad n \ge N \implies |\xi_n - \xi| \le \epsilon$.

These equivalent statements (C) and (D) are the mathematical *definition* of what we mean by "$\xi_n \to \xi$ as $n \to \infty$" or "(ξ_n) converges to ξ", and are what any mathematician has in mind when he uses those phrases about a sequence.

The discussion preceding (C) and (D) has no mathematical status at all; it is just an attempt at persuading you that (C) and (D) are reasonable expressions in precise language of our

intuition. But once we have a clear concept of convergence for sequences, intuition is superfluous, for logical argument must begin with the definition.

Some clarifications are called for. In the first place, it is unimportant whether N is required to be an integer, or a positive integer, or is allowed to be any real number. If R is a real number such that $n \geq R \implies |\xi_n - \xi| \leq \epsilon$, then the smallest integer which is not less than R, sometimes written $\lceil R \rceil$ (if $R \in \mathbb{Z}$, then $\lceil R \rceil = R$, whilst, if $\mathbb{R} \notin \mathbb{Z}$, $\lceil R \rceil = [R] + 1$) is an integer N such that $n \geq N \implies |\xi_n - \xi| \leq \epsilon$. (If this is not immediately apparent, notice that, since $N \geq R$, $n \geq N$ implies that $n \geq R$.)

The second point is that the number N in (D) usually "depends on" ϵ. By this I mean only that, when the chosen value of ϵ is altered, the N whose existence is asserted by (C) or (D) may also have to be changed. For example, when we say that $\dfrac{1}{n} \to 0$ as $n \to \infty$, we mean that, for any positive ϵ, there is some real number N such that, whenever $n \geq N$, $\left|\dfrac{1}{n}\right| \leq \epsilon$. Since $|1/n| \leq \epsilon$ if and only if $n \geq 1/\epsilon$, a possible value of N in this assertion, in fact the least possible value, is $1/\epsilon$. So a suitable value of N when $\epsilon = \frac{1}{2}$ is 2; but this is not a suitable choice for N if $\epsilon = \frac{1}{4}$, since $\frac{1}{2} > \frac{1}{4}$. To stress the fact that the N in (D) is associated with a particular ϵ, people sometimes write $N(\epsilon)$. This does not mean that N is a "function" of ϵ, for there are (in principle) many possible values of N for each ϵ.

This remark, that the N in (D) is related to the ϵ under consideration, is implicit in the order of the quantifiers $\forall$ and $\exists$. "For every positive ϵ, there exists an N ..." means that N works for that ϵ (and need not work for others). If the order of the quantifiers is reversed, "there exists an N such that for all ϵ ..." is a far stronger, and usually untrue, statement. For another example, "Every person in Wellington has or had a father" may be rephrased as "For every person in Wellington, there exists or has existed someone who is or was that person's father"; which is true. But reversing the order of the quantifiers gives "There is or was someone who is or was the father of every person in Wellington", which is presumably false. (It may be observed that the order of quantifiers doesn't matter if they are of the same kind, both $\forall$ or both $\exists$. But you can't swap round $\forall$ and $\exists$ without changing the meaning).

A third point is that, in the statements (C) and (D), the inequalities $\geq$, $\leq$ may be changed to $>$, $<$, individually or together, without changing the meaning of *these particular* statements. (This is because of the quantifiers in (C) and (D). I am absolutely not saying that $<$ and $\leq$ can usually be substituted for each other!)

If (D) is true, then, given ϵ, there is an N, which we might call $N(\frac{1}{2}\epsilon)$, such that $n \geq N \implies |\xi_n - \xi| \leq \frac{1}{2}\epsilon$. Hence, $n \geq N \implies |\xi_n - \xi| < \epsilon$, and it follows that $n > N \implies |\xi_n - \xi| < \epsilon$. Similarly, if $n > N \implies |\xi_n - \xi| < \epsilon$, then $n \geq N + 1 \implies |\xi_n - \xi| \leq \epsilon$.

Finally: to prove that definition (D) is satisfied, one must show that, for *all possible choices* of the positive number ϵ, there exists an $N(\epsilon)$. Normally, this is done by giving a formula of some sort for N, or at least some procedure to find N, in terms of ϵ (as, for instance, $N = \epsilon^{-1}$ above). But there will not be a single correct formula (for example, $N = 10^{20}\epsilon^{-1} + 10^{10^7}$ is, for definition (D), just as acceptable as $N = \epsilon^{-1}$), and the *definition* does not require a formula, as such, at all. There are even cases where a formula, in any usual sense, cannot easily be presented. This is obviously the same question about "constructive" methods that I mention elsewhere, after 6.3 and after 8.17.

APPENDIX B

I remarked at the beginning of §3 that, in earlier years, these notes had included an introduction to ideas of general topology and that it had become superfluous with the introduction of a separate course. Here are the notes that were removed. (I have changed the numbering to preserve cross-references in this Appendix; but you will notice that some material here repeats things that also appear in the 2012 version.)

§20. Metric spaces and topologies.

The ideas of Cauchy sequence and of convergent sequence, though not of upper and lower limits, depend only on the 'distance' in $\mathbb{R}$, and we can introduce a similar idea more generally.

Definition 20.1. Let Ω be any set. A *metric* (or *distance function*) on Ω is a function $d : \Omega \times \Omega \longrightarrow \mathbb{R}$ such that, for any $x, y, z \in \Omega$,

(a) $d(x, y) = 0$ if and only if $x = y$, and
(b) $d(x, z) \leq d(x, y) + d(z, y)$.

The pair (Ω, d) is called a *metric space*. If d has been unambiguously fixed, one speaks of "the metric space Ω" and suppresses mention of d. *(b)* is the *triangle inequality* for d.

The definition is often stated in slightly different, less concise, and perhaps more natural forms. Taking $x = y$ in *(b)* and applying *(a)*, $d(y, z) \leq d(z, y)$ for any $y, z \in \Omega$; since y and z may be swapped, we deduce $d(y, z) = d(z, y)$ always. Taking $x = z$, we find similarly that $0 \leq 2d(x, y)$, so that d only takes non-negative values. Thus we may add that *(a)* and *(b)* imply the further properties of the metric

(c) $(\forall x, y \in \Omega) \quad d(x, y) \geq 0$,
(d) $(\forall x, y \in \Omega) \quad d(x, y) = d(y, x)$,

which are often taken as part of the definition; and *(b)* may then be written as

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

Theoretical physicists, and some differential geometers, use the word "metric" to denote not the actual distance function d on a manifold but rather its "infinitesimal" version, which is a structure in the tangent bundle. This is not our convention here.

In $\mathbb{R}$ and in $\mathbb{Q}$, there is a *standard metric* given in each case by

$$d(x, y) := |x - y| := \max(x - y, y - x).$$

(In the case of $\mathbb{Q}$, the metric just defined takes only rational values; which is all right.)

Definition 20.2. Let (x_n) be a sequence in the metric space (Ω, d), and let $x \in \Omega$. We say that x_n tends to x as n tends to infinity, or that (x_n) converges to x, or (briefly) that $x_n \to x$, if, for any positive real number ϵ, there exists some natural number N such that

$d(x_n, x) < \epsilon$ whenever $n \geq N$. If there is some $x \in \Omega$ such that $x_n \to x$, we say that (x_n) is convergent or that the limit of x_n is x, $\lim_{n\to\infty} x_n = x$.

The definition is a generalization of (8). The remarks of Appendix A apply.

Lemma 20.3. *The sequence (x_n) in the metric space (Ω, d) converges to $x \in \Omega$ if and only if the numerical sequence $(d(x_n, x))$ in $\mathbb{R}$ tends to 0.*

Lemma 20.4. *If (x_n) is a sequence in a metric space Ω, it can have at most one limit. That is, if $x_n \to x$ and $x_n \to y$, where $x, y \in \Omega$, then necessarily $x = y$.*

Proof. Let $x \neq y$; $d(x, y) > 0$, by 20.1*(c)* and 20.1*(a)*. Let $\epsilon := \frac{1}{2}d(x, y)$. Then there exist $N_1, N_2 \in \mathbb{N}$ such that

$$n \geq N_1 \Longrightarrow d(x_n, x) < \epsilon, \quad n \geq N_2 \Longrightarrow d(x_n, y) < \epsilon.$$

Take $N := \max(N_1, N_2)$. If $n \geq N$, both $d(x_n, x) < \epsilon$ and $d(x_n, y) < \epsilon$, so that, by 20.1*(b)* and *(d)*, $d(x, y) \leq d(x, x_n) + d(y, x_n) < 2\epsilon = d(x, y)$. This is absurd, and we must conclude that $x = y$. (Compare 2.11.) □

Consequently, in a metric space, we may speak of *the* limit of a sequence, if a limit exists at all.

Lemma 20.5. *If a sequence (x_n) in the metric space (Ω, d) converges to $x \in \Omega$, then any subsequence of (x_n) also converges to x.* □

Definition 20.6. A sequence (x_n) in the metric space (Ω, d) is *Cauchy* (that is, it is a *Cauchy sequence*) if, for any positive real number ϵ, there exists some natural number N such that $d(x_m, x_n) < \epsilon$ whenever $m \geq N$ and $n \geq N$.

Again, it would be enough to restrict the values of ϵ to numbers $1/k$ for $k \in \mathbb{N}$. I leave it to you to make the necessary changes in the arguments.

Lemma 20.7. *A convergent sequence in any metric space is Cauchy.* □

This is a straightforward copy of 2.23. It is tempting to suppose the converse should be true as well (that is, that 2.25 holds in any metric space), but it is FALSE in many cases. A Cauchy sequence in $\mathbb{Q}$, with the metric already described, need not have a limit in $\mathbb{Q}$. For a familiar example consider the sequence defined inductively by

$$a_1 = 1, \quad a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right).$$

By induction, $a_n \in \mathbb{Q}$ for all n. The sequence converges in $\mathbb{R}$ to $\sqrt{2}$; it is the usual method (Newton-Raphson) of calculating $\sqrt{2}$ by successive approximations. So the sequence is Cauchy in $\mathbb{R}$, and Cauchy in $\mathbb{Q}$, but has no limit in $\mathbb{Q}$, since we know $\sqrt{2}$ is not in $\mathbb{Q}$ by 1.11.

Definition 20.8. A subset A of a metric space (Ω, d) is *complete* if every Cauchy sequence in A converges to a limit in A. (The commonest case is when $A = \Omega$.)

The word "complete" is grossly over-used in mathematics, but this is its main use in this course. Where it is ambiguous, one might say "metrically complete".

There are very many examples of metric spaces, and some are given in the exercises. For the moment, let me point out two. The *Euclidean space* $\mathbb{R}^k$ (where $k \in \mathbb{N}$) is the real vector space $\mathbb{R}^k$ together with the *Euclidean metric*

$$d((x_1, x_2, \ldots, x_k), (y_1, y_2, \ldots, y_k)) \coloneqq \sqrt{(x_1 - y_1)^2 + \cdots + (x_k - y_k)^2}\,,$$

whilst the *Hermitian space* (or *unitary space*) $\mathbb{C}^k$ is the complex vector space $\mathbb{C}^k$ together with the *Hermitian metric*

$$d((z_1, z_2, \ldots, z_k), (w_1, w_2, \ldots, w_k)) \coloneqq \sqrt{|z_1 - w_1|^2 + \cdots + |z_k - w_k|^2}\,.$$

(The xs and ys were *real* numbers, the zs and ws are *complex*.) When $n = 1$, these formulæ reduce to the previous definitions of the standard metrics in $\mathbb{R}$ and in $\mathbb{C}$.

Notice that we have proved the (metric) completeness of $\mathbb{R}$ with respect to the standard metric at 2.25. The proof depended on Dedekind's axiom.

Theorem 20.9. *$\mathbb{R}^k$ is complete with respect to the Euclidean metric and $\mathbb{C}^k$ is complete with respect to the Hermitian metric, for any $k \in \mathbb{N}$.*

Proof. (Sketch only.) If (x_n) is Cauchy in $\mathbb{R}^k$, where $x_n \coloneqq (\xi_1^{(n)}, \ldots, \xi_k^{(n)})$, then, for each i, the sequence of ith coordinates $(\xi_i^{(n)})_{n=1}^{\infty}$ is also Cauchy, this time in $\mathbb{R}$. So it converges to a limit $\xi_i \in \mathbb{R}$. Set $x \coloneqq (\xi_1, \xi_2, \ldots, \xi_k)$, and then $x_n \to x$ in $\mathbb{R}^k$.

The argument in $\mathbb{C}^k$ is much the same; indeed, $\mathbb{C}^k$ is obviously "isometric with" $\mathbb{R}^{2k}$ in a natural sense. □

Definition 20.10. Let (Ω, d) be a metric space. If $x \in \Omega$ and $r \in \mathbb{R}$, the *open metric ball in Ω of radius r about x* is $B(x; r) \coloneqq \{y \in \Omega : d(x, y) < r\}$. It is the set of points of Ω whose distance from x is strictly *less* than r. The *closed metric ball about x of radius r* is $C(x; r) \coloneqq \{y \in \Omega : d(x, y) \leq r\}$.

Lemma 20.11. *Given $r \in \mathbb{R}$, $B(x; r) \subseteq C(x; r)$. If $r < 0$, $C(x; r) = \emptyset$. If $r' < r$, then $C(x; r') \subseteq B(x; r)$. If $r \leq 0$, then $B(x; r) = \emptyset$. If $r = 0$, then $C(x; r) = \{x\}$.* □

Definition 20.12. *(a)* A subset A of the metric space (Ω, d) is *open* if, for every $x \in A$, there exists some $r_x > 0$ such that $B(x; r_x) \subseteq A$. (From 20.11, an equivalent way of stating the same condition is that there exists $r_x \in \mathbb{R}$ such that $x \in B(x; r_x) \subseteq A$.)

(b) If A is open, it is a union of open metric balls (each $x \in A$ is in $B(x; r_x)$, where r_x is such a number as is required in *(a)*).

In intuitive terms, A is open if every point x of A is not just *in* A but *inside* A; any point of Ω sufficiently close to x also belongs to A. For example, consider the interval

$$(a, b) \coloneqq \{x \in \mathbb{R} : a < x < b\}\,,$$

where we may to begin with assume $a < b$ in $\mathbb{R}$.

Lemma 20.13. *If $a < b$ in $\mathbb{R}$, the open interval (a, b) is open in $\mathbb{R}$ with the usual metric.*

Proof. Suppose $x \in (a,b)$. Then $a < x < b$; $x - a > 0$ and $b - x > 0$. Take $r := \min(b-x, x-a)$, and then, if $y \in B(x;r)$, $|y-x| < r$ and

$$y < x + r \leq x + b - x = b\,,$$
$$y > x - r \geq x - (x - a) = a\,,$$

so that $a < y < b$ too. This proves $B(x;r) \subseteq (a,b)$. As the argument works for any $x \in (a,b)$, (a,b) is an open subset of $\mathbb{R}$ with the usual metric. □

Here r was defined by a formula that obviously varies as x changes, and a moment's thought will reveal that it must do so; r must be very small if x is close to a or to b. Even more obviously, r depends on the choice of interval (a,b) containing x.

The statement that "an open interval is an open subset of $\mathbb{R}$" is equally true for "semi-infinite open intervals". I leave the proof to you.

Lemma 20.14. *For any $a, b \in \mathbb{R}$, both of the semi-infinite open intervals $(-\infty, a) := \{x \in \mathbb{R} : x < a\}$ and $(b, \infty) := \{x \in \mathbb{R} : b < x\}$ are open subsets of $\mathbb{R}$.* □

The symbol ∞ (to repeat the matter) does not in itself denote anything.

Lemma 20.15. *Let (Ω, d) be a metric space, and suppose that $x \in \Omega$ and $r \in \mathbb{R}$. Then $B(x;r)$ is an open subset of Ω.*

This tells us *open metric balls are open sets.*

Proof. Let $y \in B(x;r)$. Then $d(x,y) < r$. Let $\delta := r - d(x,y) > 0$. I claim $B(y;\delta) \subseteq B(x;r)$. Indeed, suppose $z \in B(y;\delta)$. Then $d(y,z) < \delta$, so that

$$d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + \delta = r\,;$$

which means that $z \in B(x;r)$ too, and this shows $B(y;\delta) \subseteq B(x;r)$. As y was any element of $B(x;r)$, it follows that $B(x;r)$ is open. □

Proposition 20.16. *Let (Ω, d) be any metric space, and let $\mathcal{T}$ be the class of all open subsets of Ω. Then $\mathcal{T}$ has the following properties.*

(a) *$\Omega \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$.*

(b) *For any index set Γ, if $A_\gamma \in \mathcal{T}$ for every $\gamma \in \Gamma$, $\bigcup_{\gamma\in\Gamma} A_\gamma \in \mathcal{T}$.*

(c) *If $A_1, A_2, \ldots, A_k \in \mathcal{T}$ for some $k \in \mathbb{N}$, then $\bigcap_{i=1}^{k} A_i \in \mathcal{T}$.*

[In words: the whole space Ω and the null set are both open in Ω; the union of *any* class of open sets is also open; the intersection of *finitely many* open sets is also open. It is very important to grasp that in *(b)* the index set Γ may be entirely arbitrary. It is not restricted to being finite or countable. In *(c)* only finite index sets are allowed.]

Proof. That Ω itself is open is trivial, since $B(x;r) \subseteq \Omega$ for any r and any $x \in \Omega$, directly from 20.10. That $\emptyset$ is open is "vacuously true"—the definition 20.12 is satisfied because there is no point x in A. (We want $x \in A \implies (\exists r > 0)(B(x;r) \subseteq A)$. But recall that $p \implies q$ is logically equivalent to $q \vee \neg p$, and in this case p is "$x \in A$", where $A = \emptyset$; thus $\neg p$ is automatically true.)

Suppose A_γ is open for each $\gamma \in \Gamma$. Let $x \in \bigcup_{\gamma\in\Gamma} A_\gamma$. There is a specific $\delta \in \Gamma$ such that $x \in A_\delta$, and, as A_δ is open, there exists $r > 0$ such that $B(x;r) \subseteq A_\delta$. Then $B(x;r) \subseteq \bigcup_{\gamma\in\Gamma} A_\gamma$. This shows that $\bigcup_{\gamma\in\Gamma} A_\gamma$ is open.

Finally, suppose $A_1, A_2, \dots, A_k$ are all open and $x \in \bigcap_{i=1}^k A_i$. For each i, there is $r_i > 0$ such that $B(x;r_i) \subseteq A_i$. Take $r := \min(r_1, r_2, \dots, r_k) > 0$. (It is essential here to have only finitely many sets A_i.) Then, for each i, $B(x;r) \subseteq B(x;r_i) \subseteq A_i$. It follows that $B(x;r) \subseteq \bigcap_{i=1}^k A_i$. As x was any element of $\bigcap_{i=1}^k A_i$, this proves $\bigcap_{i=1}^k A_i$ is open in Ω. □

With 20.12*(b)*, this tells us that

Corollary 20.17. *Every open set in Ω is a union of open metric balls, and every union of open metric balls is an open set in Ω.*

Definition 20.18. Let Ω be *any* set. A *topology* in (or on) Ω is a class $\mathcal{T}$ of subsets of Ω (that is, $\mathcal{T} \subseteq \mathcal{P}(\Omega)$) which satisfies the three conditions

(a) (as 20.16*(a)*), *(b)* (as 20.16*(b)*), *(c)* (as 20.16*(c)*).

A pair $(\Omega, \mathcal{T})$ of a set Ω together with a topology $\mathcal{T}$ in Ω is called a *topological space*.

Lemma 20.16 may therefore be restated: if (Ω, d) is a metric space, the class of all the subsets of Ω that are open with respect to d constitutes a topology in Ω, and in this way Ω automatically becomes a topological space.

In this course all the topological spaces we meet arise from metric spaces. However, there are several reasons to have a distinct concept of topological space. Firstly, many (though by no means all) of the important ideas we discuss do not depend on the specific metric we use, but only on the open sets it defines, and two different metrics may (and often do) have exactly the same open sets. So it makes sense to concentrate our attention on the open sets rather than on the metric. This can often lead to simpler definitions and proofs (as I said in my Introduction). Secondly, there are also important ways of constructing topologies that do not involve metrics at all. (It was an obvious problem how to characterize those topological spaces that could be obtained from a metric. The problem was finally solved in 1950 by Nagata; but the solution, though very ingenious, is both difficult and of little practical value, because the properties a topology must have to be derived from a metric are quite difficult to check. In practice, when one is given a topology, it is usually easier to construct a suitable metric from the definition of the topology than to show that the topology has the "theoretical" properties required by a metrizability theorem.)

This procedure of abstraction (from $\mathbb{R}$ to metric spaces, then to topological spaces) does have costs. $\mathbb{R}$ has both an order structure and an algebraic structure, neither of which is usually present in a metric space, and similarly a topological space lacks the idea of a specific distance that is the basis of a metric space.

When we start with a topological space, it is customary to describe the sets in $\mathcal{T}$ as the "open sets of Ω". Thus there is an ambiguity: in a *metric* space a metric d is *given* and the "open sets" are *defined* in terms of d by Definition 20.12, whilst in a topological space the topology $\mathcal{T}$ is *given* and an "open set" is just another name for a member of $\mathcal{T}$—it is open by decree, as it were, rather than having any other property. In practice this never causes any difficulty, because the context will make it clear which meaning is intended. If one speaks of a topological space without ever mentioning a metric, obviously the second meaning is the only possible one.

Given a metric space (Ω, d), one often speaks of "the metric space Ω". Likewise, in statements about a topological space $(\Omega, \mathcal{T})$ one frequently suppresses mention of $\mathcal{T}$ and simply calls certain subsets "open".

Definition 20.19. *(a)* Let $(\Omega, \mathcal{T})$ be a topological space, and $x \in \Omega$. A subset M of Ω is a *neighbourhood of x in Ω* if there is an open set U such that $x \in U \subseteq M$.

The word "neighbourhood" is sometimes used in slightly different senses, but this is perhaps the most common one in general topology. However, some authors (such as Rudin) define a neighbourhood of x in a *metric* space to be what I have called a metric ball of positive radius about x.

(b) If Ω is a metric space, M is a neighbourhood of x [in my sense] if and only if there exists some $r > 0$ such that $B(x; r) \subseteq M$.

(c) I write $\mathfrak{N}(x)$ or $\mathfrak{N}_x$ to denote the class of all neighbourhoods of x in Ω.

Definition 20.20. Let $(\Omega, \mathcal{T})$ be a topological space and $A \subseteq \Omega$. The *interior* of A in Ω, denoted by $\operatorname{int}(A)$, $\operatorname{int}_\Omega A$, or A°, is the set of all points of Ω of which A is a neighbourhood.

Lemma 20.21. *For any subset A of the topological space Ω,*

(a) $\operatorname{int}(A)$ *is an open subset of Ω;*

(b) *any open subset of A is included in* $\operatorname{int}(A)$*;*

(c) $\operatorname{int}(A)$ *is the largest open subset of Ω;*

(d) *A is open in Ω if and only if* $\operatorname{int}(A) = A$.

Proof. Let $\mathcal{V} := \{U : U \in \mathcal{T} \ \& \ U \subseteq A\}$. Then $V := \bigcup_{U \in \mathcal{V}} U$ is a union of open sets, so is itself open in Ω by 20.18*(b)*. It is also obviously a subset of A, and includes every open subset of A by definition; thus it is the largest open subset of A, and will be equal to A if and only if A is itself open. To complete the proof I show $V = \operatorname{int}(A)$.

If $x \in \operatorname{int}(A)$, then A is a neighbourhood of x, so there is an open set U with $x \in U \subseteq A$; thus $U \in \mathcal{V}$, and, by definition, $x \in U \subseteq V$. This proves $\operatorname{int}(A) \subseteq V$.

On the other hand, if $x \in V$, then (by the definition of union) there exists some $U \in \mathcal{V}$ such that $x \in U$; as U is open and a subset of A, this shows that $x \in \operatorname{int}(A)$. Hence, $V \subseteq \operatorname{int}(A)$, and, with the previous inclusion, this shows $V = \operatorname{int}(A)$. □

As an example of the way in which other definitions may be reformulated in terms of "topologies", we have

Lemma 20.22. *Suppose that (Ω, d) is a metric space, (x_n) a sequence in Ω, and $x \in \Omega$. Then $x_n \to x$ if and only if, for every neighbourhood U of x, there exists $N \in \mathbb{N}$ such that $x_n \in U$ whenever $n \geq N$.* □

Thus convergence according to 2.9 is equivalent to a statement in which the metric is not mentioned at all, being substituted by the idea of "open sets" or "neighbourhoods". It is therefore natural to define the notion of convergence of a sequence in a *topological* space by the condition given in 20.22:

$$(\forall M \in \mathfrak{N}(x))(\exists N \in \mathbb{N}) \quad n \geq N \Longrightarrow x_n \in M. \tag{109}$$

The definition of a neighbourhood means that it is enough to consider *open* neighbourhoods here; (109) will hold if the following is true.

Definition 20.23. The sequence (x_n) in the topological space Ω is said to converge to $x \in \Omega$ if, for every open set U of Ω that contains x, there exists a natural number N such that $x_n \in U$ whenever $n \geq N$.

However, it is impossible to express the idea of a "Cauchy sequence" by means of open sets alone. For it depends on comparing the distances between varying pairs of points, not the distances from the fixed point x to the varying point x_n.

§21. Closed sets.

Definition 21.1. *(a)* Let A be a subset of a topological space Ω. A point $x \in \Omega$ is called an *adherent point* of A if every neighbourhood of x contains a point of A,

$$(\forall M \in \mathfrak{N}(x)) \quad M \cap A \neq \emptyset .$$

(b) x is called an *accumulation point* of A if every neighbourhood of x contains a point of A *other than* x,

$$(\forall M \in \mathfrak{N}(x)) \quad (M \setminus \{x\}) \cap A \neq \emptyset .$$

Equivalently, x is an accumulation point of A if every open set that contains x also contains a point of A different from x.

(c) An adherent point of A is either a point of A or an accumulation point of A (possibly both). The set of all accumulation points is called the *derived set* A' of A.

This is a rather extreme example of terminological uncertainty; you have to be very careful what meaning an author gives to these terms. Accumulation points are often called, and in many contexts are the same as, "limit points" or "cluster points", but I have tried here to give names you will quite probably never need to unlearn.

Example 21.2. Let Ω be $\mathbb{R}$ with the usual topology. The subset $\mathbb{Z}$ of integers has no accumulation points in Ω at all. (Such a set is sometimes called *isolated*.) Every real number is an accumulation point of the set $\mathbb{Q}$ of rational numbers, by 1.21. An interval (a, b), where $a, b \in \mathbb{R}$ and $a < b$, has derived set $[a, b]$, but $\{1, 2, 3\} \cup (0, \frac{1}{2})$ has derived set $[0, \frac{1}{2}]$.

Example 21.3. Let A be a non-null subset of $\mathbb{R}$ which is bounded above. Then $\sup A$ is an adherent point of A. This follows immediately from 2.8. There is a similar statement for infima.

Lemma 21.4. *Let (Ω, d) be a metric space, and $A \subseteq \Omega$. If x is an accumulation point of A, then every neighbourhood of x in Ω contains* infinitely many *points of A.*

That is: the definition only requires that it contains one point of A (apart from x), but in fact it must contain infinitely many.

Proof. Suppose $U \in \mathfrak{N}(x)$; there exists $r > 0$ such that $B(x;r) \subseteq U$. Assume that $U \cap A$ is finite. Enumerate its members other than x as $a_1, a_2, \ldots, a_k$, and let $\delta := \min\{r, d(x,a_i) : 1 \le i \le k\}$. Then δ is the least of $k+1$ positive numbers, so is itself positive. Now $B(x;\delta)$ is a neighbourhood of x (by 20.15), and $B(x;\delta) \subseteq B(x;r)$, so that $B(x;\delta) \cap A \subseteq U \cap A$ can contain only $a_1, a_2, \ldots, a_k$ apart from x. However, $d(x,a_i) \ge \delta$ for each i, so that $a_i \notin B(x;\delta)$; and therefore $B(x;\delta) \cap A$ must be either $\emptyset$ or the singleton $\{x\}$. This contradicts the assumption that x is an accumulation point of A, and the contradiction proves the Lemma. □

Definition 21.5. Let A be a subset of the topological space Ω. The class of all adherent points of A in Ω is called the *closure* of A and is denoted according to the context by $\mathrm{cl}(A)$, $\mathrm{cl}\,A$, $\mathrm{cl}_\Omega A$, or sometimes $\overline{A}$.

A is described as *closed* in Ω when $A = \mathrm{cl}(A)$.

It follows that $A \subseteq \mathrm{cl}(A) = A \cup A'$ always, and A is closed if and only if $A' \subseteq A$.

Lemma 21.6. *The subset A of the topological space Ω is closed if and only if its complement A^c in Ω is open. In general, the closure in Ω of a set B is the smallest closed set of Ω including B, and is the complement of the interior of the complement of B:*

$$\mathrm{cl}(B) = (\mathrm{int}(B^c))^c \quad \text{or} \quad \overline{B} = ((B^c)^\circ)^c .$$

Proof. Suppose $x \notin \mathrm{cl}(B)$. This means that x is not an adherent point of B, so there is a neighbourhood M of x such that $M \cap B = \emptyset$, or $M \subseteq B^c$. Hence, B^c is a neighbourhood of x, and $x \in \mathrm{int}(B^c)$. The argument reverses. So $(\mathrm{cl}(B))^c = \mathrm{int}(B^c)$. taking complements gives $\mathrm{cl}(B) = (\mathrm{int}(B^c))^c$. This shows that a set A is closed if and only if

$$A = (\mathrm{int}(A^c))^c \quad \text{or} \quad A^c = \mathrm{int}(A^c),$$

that is, if and only if A^c is open. (Recall 20.21*(d)*.)

For any B, $\mathrm{int}(B^c)$ is the largest open subset of B^c (20.21*(c)*), and so its complement $\mathrm{cl}(B)$ must be the smallest closed *superset* of B. □

Remark 21.7. The Lemma shows that closed sets are just the complements of open sets, and this is often taken as the *definition* of a closed set. Then $\mathrm{cl}(A)$ would be *defined* as the intersection of all the closed sets including A. In metric spaces, sets are sometimes defined as closed if they satisfy the property given as Lemma 21.11 below. These various approaches arose as the ideas developed, and you must accustom yourself to the fact that different authors start from different definitions.

Beware of the "standard mistake" of supposing that a subset of a topological space Ω must be either open or closed. Most subsets in most interesting spaces—like $\mathbb{R}$—are neither (for instance, the interval $(1,2]$ in $\mathbb{R}$ is neither). The sets $\emptyset$ and Ω are *both* open and closed; and there are many non-trivial topological spaces in which there are more than these two sets that are both open and closed. When such sets are of special interest, they are sometimes called *clopen* (because it almost sounds like a word).

Lemma 21.8. *Let (Ω, d) be a metric space, $x \in \Omega$ and $r \in \mathbb{R}$. The closed metric ball $C(x;r)$ (see* 20.10*) is a closed set in Ω.*

Proof. Let $y \notin C(x;r)$, so $d(x,y) > r$. Take $\delta(y) := d(x,y) - r > 0$; if $z \in B(y;\delta(y))$, then $d(x,y) \le d(x,z) + d(z,y) < d(x,z) + \delta(y)$, and

$$d(x,z) > d(x,y) - \delta(y) = r\,;$$

thus $z \notin C(y;r)$ too. This shows that $B(y;\delta(y)) \subseteq C(x;r)^c$, for any $y \in C(x;r)^c$. So $C(x;r)^c$ is open. Use 21.6. □

Remark 21.9. A sequence (a_n) (in any set at all) is said to be *without repeats* if $a_n \neq a_m$ whenever the indices m and n differ: $m \neq n \Longrightarrow a_m \neq a_n$. [That is, the sequence is a one-to-one function. I don't call this a formal definition only because it is so obvious.]

Lemma 21.10. *Let A be a subset of the metric space (Ω, d). A point a of Ω is an accumulation point of A if and only if there is a sequence (a_n) in A that has no repeats and converges to a.*

Proof. Let $a \in A'$. By hypothesis, $A \cap (B(a;1) \setminus \{a\}) \neq \emptyset$. Choose an element of this set to be a_1. If $a_1, a_2, \ldots, a_n$ have been constructed, then

$$A \cap \big(B(a;2^{-n+1}) \setminus \{a, a_1, a_2, \ldots, a_n\}\big) \neq \emptyset$$

by 21.4; choose a_{n+1} to be an element of the left-hand set. This inductive construction produces a sequence without repeats, and $d(a,a_n) < 2^{-n+1}$, so that $a_n \to a$ by 20.3 and 2.19.

Conversely, let (a_n) be a sequence in A, without repeats, that converges to $a \in \Omega$. Let M be a neighbourhood of a in Ω. There exists N such that $n \ge N \Longrightarrow a_n \in M$, which shows that M contains infinitely many points of A. So a is an accumulation point of A. □

In this proof we used a version of the axiom of choice (by making countably infinitely many arbitrary choices).

Lemma 21.11. *A subset A of a metric space (Ω, d) is closed if and only if every sequence in A that converges in Ω has its limit in A.*

Proof. Let A be closed. If a sequence in A that converges in Ω has infinitely many different terms, it clearly has a subsequence without repeats. This subsequence is also in A and tends to the same limit; by 21.10, the limit is an accumulation point of A, so belongs to A. If, on the other hand, a sequence in A has only finitely many distinct terms, it has a constant subsequence (all its terms are the same), which converges to the repeated value. Thus the limit is again in A (being actually a term of the sequence).

Conversely, suppose that every sequence in A that converges in Ω has its limit in A. If a is an accumulation point of A, then by 21.10 it is the limit of a sequence in A without repeats; by hypothesis this implies $a \in A$. So A is closed. □

It is interesting to contrast this characterization of closed sets in a metric space with the definition of open sets. At first sight they perhaps seem unrelated; but we now know open sets and closed sets are just complements of each other.

Lemma 21.12. *Let A, B be subsets of the metric space (Ω, d). If A is complete, it is closed in Ω. If A is complete and B is closed in Ω, then $A \cap B$ is complete.*

Proof. Let (x_n) be a sequence in A such that $x_n \to x \in \Omega$. Then it is Cauchy by 20.7, so converges to a limit $y \in A$ by the definition of completeness, 20.8. But then $x = y$ by 20.4. Thus, $x \in A$, and this shows that A is closed, by 21.11.

Similarly, if (y_n) is a Cauchy sequence in $A \cap B$, it is a Cauchy sequence in A, so converges to a limit $y \in A$. But, by 21.11, $y \in B$ too (as B is closed). So $y \in A \cap B$. This shows that $A \cap B$ is complete. □

Definition 21.13. Let (Ω, d) be a metric space. A non-empty subset A of Ω is described as *bounded* **with respect to *d*** if $\{d(x, y) : x, y \in A\}$ is a bounded subset of $\mathbb{R}$. In that case, $\sup\{d(x, y) : x, y \in A\}$ is called the *diameter* of A.

This notion of boundedness is specific to the given metric d.

§22. Compactness.

Definition 22.1. Let B be a subset of a topological space Ω. A class $\mathcal{U}$ of subsets of Ω is described as a *covering* (or just a *cover*) of B if $B \subseteq \bigcup_{U \in \mathcal{U}} U$ (in words: B is included in the union of all the members of $\mathcal{U}$).

A subclass $\mathcal{V}$ of a covering $\mathcal{U}$ of B is called a *subcovering* of B if the union of the members of $\mathcal{V}$ still includes B.

A covering $\mathcal{U}$ of B is *open* (in Ω) if all the members of $\mathcal{U}$ are open sets in Ω. (It is, similarly, a *closed covering* if all its members are closed sets in Ω.) It is a *finite* covering if it has only finitely many members.

Thus, a covering of B is a class of subsets of Ω—that is, a set whose members are subsets of Ω; and it may have *very many* members (even uncountably many).

We now come to one of the most important ideas in the whole of mathematics: compactness. It crops up in quite unexpected places, and, as a vague explanation of its importance, one might say that it is a "topological analogue of finiteness". The usual modern formulation, which was the end of a long evolution, is given below. It needs some effort to grasp what it means, but it is both general and simple to apply in most situations. In metric spaces, however, it is equivalent to several other conditions that may seem rather more natural at first sight than 22.2. I shall discuss them in some detail and with some redundancies in §23.

Definition 22.2. Suppose that B is a subset of the topological space Ω. B is said to be *compact* [or to have the *Heine-Borel property*] if every open covering of B has a finite subcovering.

This definition says that, in any class consisting of infinitely many open sets whose union includes B, nearly all of them are redundant for the purposes of covering—you can throw out all but finitely many and still cover B. This is trivially true if B has only finitely many points:

for each point of B, take one set of the covering containing that point, and the result of these choices will be a finite subcovering[5]; but it is surprising that there are more interesting compact sets. The central fact is as follows.

Lemma 22.3. *Let $a, b \in \mathbb{R}$. Then the closed interval $X := [a, b]$ is compact in $\mathbb{R}$ in the usual topology (defined by the metric $d(x, y) := |x - y|$).*

Proof. If $b \leq a$, X is either $\emptyset$ or a singleton, and the result is trivial; so we assume $a < b$. Let $\mathcal{U}$ be an open covering of X. Say that a point $x \in [a, b]$ is *$\mathcal{U}$-reachable* or just *reachable* if there is some finite subset $\mathcal{S}$ of $\mathcal{U}$ for which $[a, x] \subseteq \bigcup_{U \in \mathcal{S}} U$. ("You only need to pass through finitely many sets of $\mathcal{U}$ to reach x from a".) Evidently a is reachable (a belongs to at least one member of $\mathcal{U}$). So the set of reachable points is non-empty and bounded above by b. By Dedekind's axiom, it has a supremum $q \in [a, b]$. There is some $U_0 \in \mathcal{U}$ for which $q \in U_0$; U_0 is open in $\mathbb{R}$; so, by definition 20.12, there is some $r > 0$ such that $B(q; r) = (q - r, q + r) \subseteq U_0$.

Now, by 2.8*(ii)*, there is a reachable point q' for which $q - r < q' \leq q$ (q is the supremum of the reachable points). Let $q'' := \min(b, q + \frac{1}{2}r) \in U_0$. Then $q'' \in U_0$ too, and is also reachable, since $q' \in U_0$ may be reached via finitely many members of $\mathcal{U}$, and the one additional member U_0 suffices to reach q''. As q was the supremum of the reachable points and $q'' \geq q$, necessarily $q = q''$. But the definition of q'' shows that this can only occur if $\min(b, q + \frac{1}{2}r) = q$, which evidently implies that $q = q'' = b$. We conclude that b is reachable, which is just what is desired. □

You may find it helpful in following the proof to sketch the positions of the points on a line—as often, what is being said is quite easy in terms of geometrical intuition, but clumsy to phrase in rigorous language. Cambridge folklore was that this proof was presented by an undergraduate in an exam, when all that was expected was mere regurgitation of material from a lecture.

The Lemma is often summarized as *bounded closed intervals in $\mathbb{R}$ are compact*. (X has an upper bound b and a lower bound a.) It is another statement that is in some sense equivalent to Dedekind's Axiom. At any rate, we can conclude that compactness is not uselessly uncommon, especially in view of the next Lemma.

Lemma 22.4. *Let Ω be any topological space, and suppose K is a compact subset of Ω and C is a closed set in Ω. Then $K \cap C$ is also compact in Ω.*

Proof. Let $\mathcal{U}$ be an open covering of $K \cap C$. Then $\mathcal{U}' := \mathcal{U} \cup \{\Omega \setminus C\}$ is an open covering of K, so it has a finite subcovering $\mathcal{V}'$. In principle, $\mathcal{V}'$ may or may not have $\Omega \setminus C$ as a member; let $\mathcal{V} := \mathcal{V}' \setminus \{\Omega \setminus C\}$ (i.e. remove $\Omega \setminus C$ if it is in $\mathcal{V}'$). Then $\mathcal{V} \subseteq \mathcal{U}$, and $\mathcal{V}$ is clearly a covering of $K \cap C$ (though perhaps not of K), since $\Omega \setminus C$ is disjoint from C. That is, $\mathcal{V}$ is a finite subcovering of $\mathcal{Q}$. □

In this case, a Venn diagram may clarify the proof, although, of course, Venn diagrams cannot on their own constitute proofs.

Corollary 22.5. *Any bounded closed subset of $\mathbb{R}$ is compact.*

---

[5] This does not require the Axiom of Choice. Only finitely many choices have to be made, and the usual axioms of set theory permit that.

Proof. Let X be any bounded closed subset of $\mathbb{R}$, with an upper bound b and a lower bound a. Then $X \subseteq [a,b]$, which is compact. So, by the preceding Lemma, X is also compact. □

In older books, before the modern concept of "compactness" was invented, many results were stated with the hypothesis that a set was bounded and closed in $\mathbb{R}$ or in $\mathbb{R}^n$.

§23. Notions related to compactness.

There are many properties that are equivalent to compactness, perhaps under suitable additional hypotheses, and I shall give only the commoner ones. Furthermore, those of you who have come across it will notice that I am making free use of the axiom of choice; and I should point out that (at least partly for clarity, but also to introduce some additional concepts) I have not tried to follow the most economical line of exposition.

Throughout this section, Ω is a topological space (at least).

Definition 23.1. A subset B of Ω has the *Bolzano-Weierstrass property* if every infinite subset of B has an accumulation point (in Ω; see 21.1) belonging to B. That is: if $E \subseteq B$ and E has infinitely many elements, then some point of B is an accumulation point of E in Ω.

Lemma 23.2. *Let B be a subset of the topological space Ω. If B is compact, it has the Bolzano-Weierstrass property.*

Proof. Let E be an infinite subset of B, and suppose it has *no* accumulation point in B. Thus, for every $x \in B$, there is an open set $U(x)$ of Ω (see 21.1*(b)*) which contains x but no point of E (except possibly x itself in the case that $x \in E$). Certainly $\{U(x) : x \in B\}$ is an open covering of B, since each point x belongs to the corresponding $U(x)$. But B is compact; hence, there is a finite subcovering $\{U(x_i) : 1 \leq i \leq n\}$. Each of $U(x_1), \ldots, U(x_n)$ contains at most one point (the "x_i" that indexes it) of $E \subseteq \bigcup_{i=1}^{n} U_i$, and so E has no more than n members. This contradiction shows that E must have an accumulation point in B. □

Definition 23.3. A subset B of Ω is said to be *sequentially compact* if any sequence (b_n) in B has a subsequence that converges to a limit in B.

Lemma 23.4. *If the subset B of the topological space Ω is sequentially compact, it has the Bolzano-Weierstrass property.*

Proof. Let E be an infinite subset of B. Then we can construct by repeated choices a sequence $(e_n)_{n=1}^{\infty}$ in E without repeats. By the sequential compactness of B, there is a subsequence $(e_{n(k)})_{k=1}^{\infty}$ that converges to some point e of B. By the definition 20.23, any neighbourhood of e contains infinitely many terms of the sequence, and so b is an accumulation point of E. □

Perhaps surprisingly, the argument does not reverse. The difficulty is to pass from an accumulation point of a *set* $\{x_n\}$ to a subsequence of the *sequence* (x_n). But, for metric spaces:

Lemma 23.5. *Let B be a subset of the* metric *space (Ω, d). If B has the Bolzano-Weierstrass property, it is sequentially compact.*

Proof. Let (b_n) be a sequence in B. If any of its terms is repeated infinitely often (with value b, say), it has a constant subsequence (consisting of those repeated terms) that converges to $b \in B$.

If no term is repeated infinitely many times, there is a subsequence (a_n) entirely without repeats. Indeed, if $a_1 = b_{n(1)}$, $a_2 = b_{n(2)}$, ... , $a_k = b_{n(k)}$ have been chosen, where $n(1) < n(2) < \cdots < n(k)$, take $n(k+1)$ to be the least index for which $b_{n(k+1)}$ differs from all of $a_1, a_2, \ldots, a_k$, and let a_{k+1} be $b_{n(k+1)}$. In this way, we define by induction a subsequence (a_n) without repeats; then $\{a_n : n \in \mathbb{N}\}$ is an infinite subset of B, and so the Bolzano-Weierstrass property would mean it has an accumulation point a in B. As Ω is a metric space, 21.10 applies: there is a sequence without repeats $(a_{n(i)})_{i=1}^{\infty}$ whose terms belong to $\{a_n : n \in \mathbb{N}\}$ and which converges to a. It may not, however, be a sub*sequence* of (a_n), as the $n(i)$ may not increase as i increases.

Nevertheless, we can *select* as follows a subsequence $(a_{m(i)})$ of $(a_{n(i)})$ which is also a subsequence of (a_n). Let $m(1) = n(1)$, and, if $m(i) = n(s_i)$, let $m(i+1) = n(s_{i+1})$, where s_{i+1} is the least index greater than s_i such that $n(s_{i+1}) > n(s_i)$. There must be such an s_{i+1}, as otherwise the whole sequence $(a_{n(i)})$ (without repeats) has no indices greater than $n(s_i)$, which is absurd. And, of course, $(a_{m(i)})$ is a subsequence of $(a_{n(i)})$ and, therefore, converges to a. □

Thus, in METRIC spaces, the Bolzano-Weierstrass property and sequential compactness are equivalent, and are *implied by* compactness. To show they *imply* compactness requires intermediate concepts that are of interest in themselves.

For the rest of this section, Ω will be a metric space with metric d.

Lemma 23.6. *Let B be a sequentially compact subset of Ω. Suppose that $\mathcal{U}$ is an open cover of B. Then there exists $\epsilon > 0$ such that, for any $x \in B$, there is some $U \in \mathcal{U}$ for which $B(x; \epsilon) \subseteq U$.*

Thus, if we shrink every member $U \in \mathcal{U}$ by taking away its "outer shell" of thickness ϵ, leaving $U_\epsilon := \{x \in \Omega : B(x; \epsilon) \subseteq U\}$, the result $\{U_\epsilon : U \in \mathcal{U}\}$ is still a covering of B. A number ϵ with this property is often called a *Lebesgue number* for the covering $\mathcal{U}$ of B.

Proof. Suppose it were untrue. Then, for each $n \in \mathbb{N}$, there must be some $x_n \in B$ such that $B(x_n; 1/n)$ is not included in any member of $\mathcal{U}$. By sequential compactness, the sequence (x_n) in B has a subsequence $(x_{n(i)})_{i=1}^{\infty}$ which converges to some point $x_* \in B$. Thus $x_* \in U$ for some $U \in \mathcal{U}$. As U is open in Ω, there is $r > 0$ such that $B(x_*; r) \subseteq U$. As $x_{n(i)} \to x_*$, there is a J such that $i \geq J \Longrightarrow d(x_{n(i)}, x_*) < \frac{1}{2}r$. Suppose that $i \geq \max(J, 2/r)$, and take any $x \in B(x_{n(i)}, 1/n(i))$; then (recall from 2.2 that $n(i) \geq i$)

$$d(x, x_*) \leq d(x, x_{n(i)}) + d(x_{n(i)}, x_*) < \frac{1}{n(i)} + \tfrac{1}{2}r \leq \tfrac{1}{2}r + \tfrac{1}{2}r = r,$$

so that $B(x_{n(i)}; 1/n(i)) \subseteq U$. This contradicts the construction of the sequence (x_n). We conclude that the Lemma must be true. □

A sketch (possibly in two dimensions) may help to grasp the argument.

Definition 23.7. The subset A of the metric space Ω is *precompact* [or, especially in older texts, *totally bounded*] if, for any $r > 0$, there is a finite subset F of A such that $A \subseteq \bigcup_{f \in F} B(f; r)$.

Equivalently, there is a finite subset F of A such that any point of A is within distance r of some point of F. It clearly implies that A is bounded with respect to d (as at 21.13), since, for any $x \in \Omega$ and any $a \in A$, there is some $f \in F$ for which

$$d(x, a) \leq d(x, f) + d(f, a) < \max\{d(x, f) : f \in F\} + r\,;$$

this is a number that does not depend on the choice of a in A (the maximum is taken over finitely many numbers that are independent of a).

Lemma 23.8. *The subset A of the metric space Ω is precompact if and only if, for any $\delta > 0$, any subset G of A such that $d(x, y) \geq \delta$ for any two distinct points $x, y \in G$ must necessarily be finite.*

Let us say temporarily that a set such as G is "δ-separated".

Proof. Suppose that A is precompact and $\delta > 0$. There exists a finite subset $F(\delta)$ of Ω such that $A \subseteq \bigcup_{f \in F} B(f; \frac{1}{2}\delta)$. If G is any subset of A with more elements than $F(\delta)$, then there must be two elements $x, y \in G$ which belong to the same ball $B(f; \frac{1}{2}\delta)$ (this is the so-called "pigeonholing principle" or Schubfachprinzip). Thus,

$$d(x, y) \leq d(x, f) + d(f, y) < \tfrac{1}{2}\delta + \tfrac{1}{2}\delta = \delta\,.$$

So a δ-separated subset G of A cannot have more elements than $F(\delta)$; in particular, it cannot be infinite.

Suppose that any δ-separated subset G of A must be finite. Let $r > 0$ be given. Choose inductively $x_1, x_2, x_3, \cdots$ in A so that, for each n,

$$x_{n+1} \notin \bigcup_{i=1}^{n} B(x_i; r)\,,$$

if such a choice is possible. In this way, $d(x_m, x_n) = d(x_n, x_m) \geq r$ whenever $n > m$, i.e. $\{x_n\}$ is δ-separated. It is therefore finite, by hypothesis. There must be some n for which no x_{n+1} may be chosen; that is,

$$\bigcup_{i=1}^{n} B(x_i; r) = A\,.$$

As $r > 0$ was arbitrary, this shows A is precompact. □

Lemma 23.9. *A sequentially compact subset A of a metric space Ω is precompact.*

Proof. By 23.8, it will suffice to show that, given $\delta > 0$, any δ-separated set $G \subseteq A$ must be finite. If it were not, we could select a sequence (x_n) without repeats from G. By hypothesis, there is a subsequence $(x_{n(i)})$ convergent in A. It is therefore Cauchy, by 20.7. There exists J such that $0 < d(x_{n(i)}, x_{n(j)}) < \delta$ if $i > j \geq J$. This is absurd, as $\{x_{n(i)}\}$ is δ-separated as a subset of G. □

Theorem 23.10. *The subset A of the metric space Ω is sequentially compact if and only if it is compact.*

Proof. If A is compact, 23.2 shows it has the Bolzano-Weierstrass property, and 23.5 then shows it is sequentially compact. Now suppose it is sequentially compact, and let $\mathcal{U}$ be an open covering of A. By 23.6, there is a Lebesgue number $\epsilon > 0$ for the covering. By 23.9, there is a finite subset F of A such that $A \subseteq \bigcup_{f\in F} B(f;\epsilon)$. For each $f \in F$, there is a $U_f \in \mathcal{U}$ such that $B(f;\epsilon) \subseteq U_f$; and so

$$A \subseteq \bigcup_{f\in F} B(f;\epsilon) \subseteq \bigcup_{f\in F} U_f\,,$$

and $\{U_f : f \in F\}$ is a finite subcovering of $\mathcal{U}$. □

There is yet a third way of characterizing compactness of a subset of a metric space, with the advantage that it generalizes valuably to some other situations. We start with an easy observation that is often quite useful.

Remark 23.11. If (b_n) is a Cauchy sequence in Ω, and has a subsequence $(b_{n(i)})$ that converges to $b \in \Omega$, then $b_n \to b$ too.

Proof. Let $\epsilon > 0$. As (b_n) is Cauchy, there is N such that $d(b_n, b_m) < \frac{1}{2}\epsilon$ when $n \geq N$. Also, there exists J such that $i \geq J \Longrightarrow d(b_{n(i)}, b) < \frac{1}{2}\epsilon$. Hence, if $n \geq M$, where $M := \max(J, N)$,

$$d(b_n, b) \leq d(b_n, b_{n(M)}) + d(b_{n(M)}, b) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon\,.$$

(I have used the fact that $n(M) \geq M \geq N$.) □

Lemma 23.12. *If the subset A of the metric space Ω is sequentially compact, it is complete (see* 20.8*).*

Proof. Let (b_n) be a Cauchy sequence in A. By sequential compactness, there is a subsequence $(b_{n(i)})_{i=1}^{\infty}$ which converges to a point $b \in A$. But $b_n \to b$ by 23.11. □

Lemma 23.13. *The subset A of Ω is precompact if and only if every sequence in A has a Cauchy subsequence.*

Proof. Suppose A is not precompact. By 23.8, there is some $\delta > 0$ for which there is an infinite δ-separated set $G \subseteq A$. G will contain an infinite sequence without repeats, which cannot have a Cauchy subsequence, since, for any subsequence, the distance between terms is always at least δ.

Suppose A is precompact, and (b_n) is a sequence in A. We use a 'diagonal process' to construct a Cauchy subsequence. First, A may be covered by finitely many balls

$$B(f_{11};1),\ B(f_{21};1),\ \ldots,\ B(f_{m(1),1};1)$$

with centres $f_{11}, f_{21}, \ldots, f_{m(1),1} \in A$. At least one of these balls contains infinitely many terms of (b_n), say $B(f_{j(1),1};1)$; we get a subsequence

$$(b_{n_1(1)}, b_{n_1(2)}, \dots, b_{n_1(i)}, \dots)$$

of (b_n). Now, there are finitely many balls $B(f_{12}; \frac{1}{2})$, $\dots$, $B(f_{m(2),2}; \frac{1}{2})$ that cover A, with centres $f_{12}, \dots, f_{m(2),2} \in A$. At least one of these balls, say $B(f_{j(2),2}; \frac{1}{2})$, must contain infinitely many terms of $(b_{n_1(i)})_{i=1}^{\infty}$, constituting a subsequence $(b_{n_2(i)})$ of $(b_{n_1(i)})$. If the sequence $(b_{n_k(i)})$ has been defined, A is covered by finitely many balls

$$B(f_{1,k+1}; 2^{-k}),\ \dots,\ B(f_{m(k+1),k+1}; 2^{-k}),$$

at least one of which, say $B(f_{j(k+1),k+1}; 2^{-k})$, contains infinitely many terms of $(b_{n_k(i)})$. These terms constitute a further subsequence $(b_{n_{k+1}(i)})$. In this way we obtain a sequence $((b_{n_k(i)})_{i=1}^{\infty})_{k=1}^{\infty}$ of subsequences of (b_n), each being a subsequence of its predecessor.

Define $c_m := b_{n_m(m)}$ for $m = 1, 2, 3, \dots$. This "diagonal sequence" (c_n) is certainly a subsequence of (b_n), and all its terms after the mth form a subsequence of $(b_{n_m(i)})$ and lie in $B(f_{j(m),m}; 2^{-m+1})$. Thus

$$k, l \geq m \Longrightarrow d(c_k, c_l) \leq d(c_k, f_{j(m),m}) + d(f_{j(m),m}, c_l) < 2.2^{-m+1} = 2^{-m+2}.$$

Hence, (c_n) is a Cauchy sequence, and a subsequence of (b_n). □

Theorem 23.14. *Let A be a subset of the metric space (Ω, d). A is sequentially compact if and only if it is complete and precompact.*

Proof. Let A be sequentially compact. It is complete by 23.12 and precompact by 23.9.

Now suppose that A is complete and precompact, and let (x_n) be a sequence in A. By 23.13, it has a Cauchy subsequence; by completeness, that Cauchy subsequence converges to a point of A. So A is sequentially compact. □

In short, compactness, the Bolzano-Weierstrass property, sequential compactness, and completeness+precompactness, are all equivalent FOR SUBSETS OF METRIC SPACES. The proofs above are not the shortest possible; you will certainly find different arguments in other sources. There are also various generalizations that I omit.

Theorem 23.15. *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded with respect to the Euclidean metric. A subset of $\mathbb{C}^n$ is compact if and only if it is closed and bounded with respect to the Hermitian metric.*

Proof. Suppose first that A is a compact subset of $\mathbb{R}^n$. Let (a_n) be a sequence in A that converges to $a \in \mathbb{R}^n$. As A is sequentially compact, (a_n) has a subsequence that converges to a point $a' \in A$. By 20.5, $a' = a$, so that $a \in A$. By 21.11, A is closed in Ω. As A is totally bounded, it is bounded (see the remark after 23.7).

Now suppose A is closed and bounded. As $\mathbb{R}^n$ is complete by 23.14, any closed subset is also complete (by 21.12). All we need to show, therefore, is that a bounded set in $\mathbb{R}^n$ must be totally bounded. If A is d-bounded, there exists R such that $d(0, a) < R$ for every $a \in A$, and, if $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$,

$$(\forall i) \quad |\alpha_i| \le \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2} < R\,.$$

Thus A is included in the 'cube' $\{(\alpha_1, \alpha_2, \ldots, \alpha_n) : (\forall i)\ |\alpha_i| \le R\}$.

Given $\epsilon > 0$, choose $N \in \mathbb{N}$ such that $N\epsilon > R$. We may split up the cube of side $2R$ into $(2nN)^n$ smaller cubes of the form

$$Q(j_1, j_2, \ldots, j_n) \coloneqq \{(\alpha_1, \alpha_2, \ldots, \alpha_n) : (\forall i)\ \ j_i\epsilon/n \le \alpha_i \le (j_i + 1)\epsilon/n\}\,,$$

where $(j_1, j_2, \ldots, j_n)$ ranges over all n-tuples of integers such that, for each i, $-nN \le j_i < nN$. But

$$Q(j_1, j_2, \ldots, j_n) \subseteq B((j_1\epsilon/n, \ldots, j_n\epsilon/n); \epsilon)$$

(why?), and so $A \subseteq \bigcup_{(j_1, j_2, \ldots, j_n)} B((j_1\epsilon/n, \ldots, j_n\epsilon/n); \epsilon)$. This proves that A is precompact, as required.

The proof for $\mathbb{C}^n$ is (except for more notation) the same as for $\mathbb{R}^{2n}$. □

There are many other proofs of this Theorem (for instance, compare 3.11). In older texts sequential compactness or the Bolzano-Weierstrass property are often proved directly for bounded closed subsets of $\mathbb{R}^n$. It should be emphasized, though, that the result applies specifically to $\mathbb{R}^n$ or $\mathbb{C}^n$. In more complicated spaces, boundedness certainly does *not* always imply precompactness, and a set that is bounded and complete need not be compact.

§24 Subspaces and continuity

Definition 24.1. Suppose that (Ω, d_Ω) is a metric space. If Ψ is any subset of Ω, the restriction $d_\Psi \coloneqq d_\Omega\big|\Psi \times \Psi$ of d_Ω to $\Psi \times \Psi$ is a metric on Ψ, and we describe (Ψ, d_Ψ) as a *metric subspace* of (Ω, d_Ω). In effect, this means that, if $x, y \in \Psi$, we define the distance between them as points of Ψ, $d_\Psi(x, y)$, to be exactly the same as the distance between them as points of Ω, $d_\Omega(x, y)$.

If we say "the metric space Ω", without explicit mention of the metric d_Ω, we may also speak of "the (metric) subspace Ψ of Ω", taking for granted the metric d_Ψ. We might even denote both metrics by d when notation is needed, since they have the same values for pairs of points in Ψ.

[It's worth noticing that in "real life", we don't always use this "subspace metric". When we speak of the distance from Wellington to Tokyo, we usually mean the distance as measured along a great circle in the subset constituting the surface of the earth, rather than along a straight line through the interior.]

Lemma 24.2. *In* 24.1, *a subset A of Ψ is open* [or *closed*] *in Ψ (with respect to d_Ψ) if and only if there is a subset B of Ω that is open* [or *closed*] *(with respect to d_Ω) and such that $A = B \cap \Psi$.* □

The proof is easy (see 20.12 and 20.15 for the open sets, and take complements for the closed sets; one may also use the "sequence" definition of closed sets). This suggests

Definition 24.3. Suppose $\mathcal{T}$ is a topology on the set Ω, and Ψ is a subset of Ω. Then the *subspace topology* on Ψ (induced from $\mathcal{T}$) is $\{U \cap \Psi : U \in \mathcal{T}\}$.

It is easily checked that the subspace topology is indeed a topology on Ψ, and that, if $\mathcal{T}$ is defined by a metric d_Ω on Ω, the subspace topology on Ψ may be defined by the subspace metric d_Ψ (cf. 24.1).

Definition 24.4. *(a)* Let (Ω, d_Ω) and (Ψ, d_Ψ) be metric spaces. Suppose that $f : \Omega \longrightarrow \Psi$ is a mapping. f is said to be *continuous* at $x \in \Omega$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall y \in \Omega) \quad d_\Omega(x, y) < \delta \Longrightarrow d_\Psi(f(x), f(y)) < \epsilon\,.$$

This is an obvious translation of 5.1 into the language of metric spaces (and, as there, one may say "x is a point of continuity of f"). It may be expressed in terms of metric balls:

$$(\forall \epsilon > 0)(\exists \delta > 0) \quad f(B_\Omega(x; \delta)) \subseteq B_\Psi(f(x); \epsilon)\,.$$

(The notations B_Ω, B_Ψ are, I hope, self-explanatory.)

(b) f is *continuous* (or *continuous on* Ω) if it is continuous at every point of Ω.

Lemma 24.5. *Let $f : \Omega \longrightarrow \Psi$ be a mapping between metric spaces.*

(a) *f is continuous at $x \in \Omega$ if and only if, for every neighbourhood N of $f(x)$ in Ψ, the inverse image $f^{-1}(N)$ is a neighbourhood of x in Ω.*

(b) *f is continuous (on Ω) if and only if, for every open set U in Ψ, $f^{-1}(U)$ is an open set in Ω.*

Proof. *(a)* Suppose f is continuous at $x \in \Omega$, and N is a neighbourhood of $f(x)$ in Ψ, $N \in \mathfrak{N}_\Psi(f(x))$ (in an obvious notation). By 20.19*(b)*, there is, by 20.12*(a)*, some $\epsilon > 0$ such that (again in an obvious notation) $B_\Psi(f(x); \epsilon) \subseteq N$. By 24.4*(a)*, there is $\delta > 0$ so $f(B_\Omega(x; \delta)) \subseteq B_\Psi(f(x); \epsilon)$. Thus, $B_\Omega(x; \delta) \subseteq f^{-1}(N)$; and this means that $f^{-1}(N) \in \mathfrak{N}_\Omega(x)$, again by 20.19*(b)*. This proves "only if".

Conversely, $N_0 := B_\Psi(f(x); \epsilon) \in \mathfrak{N}_\Psi(f(x))$; if $f^{-1}(N_0) \in \mathfrak{N}_\Omega(x)$, there is some $\delta > 0$ such that $B_\Omega(x; \delta) \subseteq f^{-1}(N_0)$, again by 20.19*(b)*, and this implies that $f(B_\Omega(x; \delta)) \subseteq B_\Psi(f(x); \epsilon)$, as required for 24.4*(a)*.

(b) Suppose f continuous and $U \in \mathcal{T}_\Psi$. Take any $x \in f^{-1}(U)$. By definition, $f(x) \in U$ and (U being open) $U \in \mathfrak{N}_\Psi(f(x))$. From *(a)*, $f^{-1}(U) \in \mathfrak{N}_\Omega(x)$; it is a neighbourhood of each of its points, and is open in Ω (cf. 20.21*(d)*).

The converse is now a trivial exercise. □

This Lemma suggests the following definition.

Definition 24.6. *(a)* Let $(\Omega, \mathcal{T}_\Omega)$ and $(\Psi, \mathcal{T}_\Psi)$ be *topological* spaces, and suppose $f : \Omega \longrightarrow \Psi$ a mapping. f is said to be *continuous* if, for every $U \in \mathcal{T}_\Psi$, its inverse image $f^{-1}(U)$ under f belongs to $\mathcal{T}_\Omega$: that is, if the inverse image under f of any open set in Ψ is an open set in Ω.

(b) f is said to be *continuous at* x (where $x \in \Omega$) if, for every neighbourhood N of $f(x)$ in Ψ, $f^{-1}(N)$ is a neighbourhood of x in Ω. In that case, one also says that x is a *point of continuity* of f.

This is the usual definition of continuity in general topology. It is rather remarkable that, starting from ϵs and δs, we arrive at a definition of continuous mapping that is stated in terms of pure set theory (that is, a topology is a class of subsets having certain set-theoretic properties, and the definition of continuity is in terms of those classes.)

If $\Omega_1 \subseteq \Omega$ and $g : \Omega_1 \longrightarrow \Psi$, one may say that g is continuous on Ω_1 if it is continuous with respect to the subspace topology on Ω_1 and $\mathcal{T}_\Psi$ on Ψ. When $g = f\big|\Omega_1$ and $f : \Omega \longrightarrow \Psi$ is continuous, g is certainly continuous on Ω_1. However, $f\big|\Omega_1 : \Omega_1 \longrightarrow \Psi$ may be continuous when $f : \Omega \longrightarrow \Psi$ is not; that is, every point $y \in \Omega_1$ may be a point of continuity of $f\big|\Omega_1$ (with respect to the subspace topology on Ω_1) without being a point of continuity of f. It is easy to construct trivial examples.

As an example of the way in which the idea of compactness can allow relatively snappy proofs of some results, here is a "purely topological" version of 6.7.

Theorem 24.7. (Dini's theorem on monotone convergence.) *Suppose that $(f_n)_{n=1}^{\infty}$ is a sequence of continuous functions $\Omega \longrightarrow \mathbb{R}$, where Ω is a compact topological space; and that, for each $x \in \Omega$, the numerical sequence $(f_n(x))_{n=1}^{\infty}$ tends monotonically to a limit $f(x)$, where $f : \Omega \longrightarrow \mathbb{R}$ is itself continuous. Then $f_n \to f$ uniformly on Ω.*

Proof. As at 6.7**A**, it suffices to assume $f_n \downarrow 0$ pointwise on Ω (and each f_n is continuous). Take any $\epsilon > 0$, and any $x \in \Omega$. There exists $N(x) \in \mathbb{N}$ such that

$$n \geq N(x) \Longrightarrow f_n(x) < \epsilon\,.$$

As $f_{N(x)}$ is continuous at x, $\{y \in \Omega : f_{N(x)}(y) < \epsilon\}$ is an open set $U(x)$ containing x. Thus, $\{U(x) : x \in \Omega\}$ is an open covering of Ω, and has a finite subcovering, say $\{U(x_1), U(x_2), \ldots, U(x_r)\}$.

Take $N := \max(N(x_1), \ldots, N(x_r)\} \in \mathbb{N}$. If $n \geq N$, any point $y \in \Omega$ is in $U(x_k)$ for some k, $1 \leq k \leq r$, and, as $n \geq N(x_k)$, $f_n(y) \leq f_{N(x_k)}(y) < \epsilon$ (as the sequence is decreasing). Thus,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall y \in \Omega) \quad n \geq N \Longrightarrow 0 \leq f_n(y) < \epsilon\,,$$

which is precisely the uniform convergence asserted. □

APPENDIX C

Solution of the cubic

The aim of this note is to amplify the frequently stated remark that Bombelli (together with some contemporaries) invented complex numbers in connection with the solution of the cubic. Solution of the cubic or of the quartic is a topic that has almost been squeezed out of the syllabus, except occasionally for some remarks in first-year courses and sometimes as an application of Galois theory in the fourth year, but it is really quite elementary and extremely ingenious. Of course Galois 'explained' it; he indicated why the method worked, why the similar method for the quartic was also possible, and why the idea could not be extended to the quintic; but the actual solution of the cubic itself was much older, and due in successive refinements to Scipione del Ferro (1465–1526), Niccolà Fontana (usually called Tartaglia, 'the stammerer') (1500–1557), and Gerolamo Cardano (1501–1576), who was so famous throughout Europe that his name itself was translated into other languages and you may find him called Cardan. James VI (as he was at the time) accused him of witchcraft. His fame,

however, was principally in medicine. He claimed—quite falsely—the "rule of Cardan" that if you have gout, you cannot catch tuberculosis or suffer from gallstones. Much medical science until very recently consisted largely of such unsubstantiated assertions, and indeed it seems to be an innate characteristic of human beings to generalize wildly on the basis of minimal evidence. The quartic was solved, with surprisingly little additional difficulty, by Ludovico Ferrari (1522–1565), who was Cardan's pupil.

In the cubic equation $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$, we may as well assume $a_0 \neq 0$, for otherwise it reduces to a lower degree. Dividing the original equation by $a_0 \neq 0$, we can suppose the cubic is $x^3 + b_1x^2 + b_2x + b_3 = 0$. If we now set $y = x + \frac{1}{3}b_1$, $x = y - \frac{1}{3}b_1$, the corresponding equation in y is

$$y^3 + ay + b = 0\,, \tag{1}$$

where $a = b_2 - \frac{1}{3}b_1^2$, $b = \frac{2}{27}b_1^3 - \frac{1}{3}b_1b_2 + b_3$. This is called a *reduced* cubic (without a term in the square of the unknown). Notice that a and b can have arbitrary values.

Cardano worked with the reduced cubic (1). Put $y = z + w$. Then

$$y^3 = z^3 + 3z^2w + 3zw^2 + w^3 = z^3 + 3zwy + w^3,$$
$$y^3 + (-3zw)y + (-z^3 - w^3) = 0\,.$$

The effect of this calculation is that, if z, w can satisfy the simultaneous equations

$$3zw = -a\,, \qquad z^3 + w^3 = -b\,, \tag{1}$$

then $z + w$ will be a solution of (1). But we can solve (2). More precisely, we can solve the related equations, obtained by substituting ξ for z^3, η for w^3,

$$\xi\eta = -a^3/27\,,\ \xi + \eta = -b\,, \tag{2}$$

whose solutions are the roots of the quadratic

$$\lambda^2 + b\lambda - a^3/27 = 0\,. \tag{3}$$

This last equation (4) is the *quadratic resolvent* of the original reduced cubic. Its roots

$$\lambda_1 = -\tfrac{1}{2}b + \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\,, \qquad \lambda_2 = -\tfrac{1}{2}b - \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\,,$$

where the same choice of the complex square root is made in both formulæ, are cubes of the roots of (2). Choose w to be a specific cube root of λ_1; the first equation of (2) will then fix z (unless $\lambda_1 = 0$, which applies only if $a = 0$). So the formula (Cardan's formula, 1573)

$$\left(-\tfrac{1}{2}b + \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\right)^{1/3} + \left(-\tfrac{1}{2}b - \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\right)^{1/3} \tag{4}$$

gives a root of the cubic (1), provided that the two cube roots are chosen so their product is $-\frac{1}{3}a$. It is clear that the prior choice of the square root does not affect the result—reversing it simply interchanges the terms of the sum (5). We might write the formula as

$$\left(-\tfrac{1}{2}b + \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\right)^{1/3} - \tfrac{1}{3}a \Big/ \left(-\tfrac{1}{2}b + \sqrt{\tfrac{1}{4}b^2 + \tfrac{1}{27}a^3}\right)^{1/3}$$

to emphasize the correct choice of cube roots, though at the cost of losing the symmetry of (5).

If one possible choice of cube roots in (5) is z_0, w_0, and if ω is a complex cube root (that is, other than 1) of unity, the other possible choices with product $-\frac{1}{3}a$ are $\omega z_0, \omega^2 w_0$ and $\omega^2 z_0, \omega w_0$; we get solutions of (0)

$$z_0 + w_0\,, \quad \omega z_0 + \omega^2 w_0\,, \quad \omega^2 z_0 + \omega w_0\,. \tag{5}$$

(Recall that $1 + \omega + \omega^2 = 0$, so that the sum of these three solutions is indeed 0, as it must be for a reduced cubic.) Thus "in principle" the formula (5) gives three distinct solutions of (0), which, if genuinely different, must be all the solutions.

[There is an alternative way of setting out the argument, which may appear more natural and which is presented in some books. Given the reduced cubic (0), substitute $y = pz$, where $p^2 = -\frac{1}{3}a$; this puts it in the simpler form $z^3 - 3z + c = 0$. Now, substitute $z = \zeta + \frac{1}{\zeta}$; then the equation becomes

$$\zeta^3 + 3\zeta + 3\zeta^{-1} + \zeta^{-3} - 3\zeta - 3\zeta^{-1} + c = 0\,,$$

i.e. $\zeta^3 + \zeta^{-3} = -c$, which is a quadratic in ζ^3. A moment's thought will convince you that this is essentially the same method as before.]

Let us try to discover what happens if the possibilities (6) are not all distinct. If

$$\begin{aligned} z_0 + w_0 &= \omega z_0 + \omega^2 w_0\,, \qquad \text{then} \\ (1-\omega)z_0 &= (\omega^2 - 1)w_0 = -\,(1-\omega)(1+\omega)w_0\,, \quad \text{and} \\ z_0 &= -\,(1+\omega)w_0 = \omega^2 w_0 \end{aligned}$$

(since $1 - \omega \neq 0$ and $1 + \omega = -\omega^2$). Other pairs from (6) are equal if $z_0 = \omega w_0$ or $z_0 = w_0$. Since z_0^3, w_0^3 have to solve (3), in each of these cases $2z_0^3 = 2w_0^3 = -b$ and $z_0^6 = -a^3/27$. Hence, if two of the formulæ (6) give the same values, necessarily the *discriminant* $\Delta = b^2 + \frac{4}{27}a^3$ must vanish. (Δ is called the discriminant of the *cubic*; but it is also the discriminant of the quadratic resolvent of (4), in the old sense for quadratics).

If all the formulæ (6) give the same values, we have $z_0 = \omega w_0 = \omega^2 w_0$, so that $z_0 = 0 = w_0$; it follows that (again) $\Delta = 0$, but, in fact, $a = b = 0$. Hence, the reduced cubic is just $z^3 = 0$, so that *all* its roots coincide.

Suppose, on the other hand, that $\Delta = 0$, but $a \neq 0$. Then it is easily checked by direct calculation that $\zeta = -\frac{3b}{2a}$ is a root of both the polynomials

$$\zeta^3 + a\zeta + b\,, \quad 3\zeta^2 + a\,,$$

of which the second is the formal derivative of the first. So ζ must be a *repeated* root of the first polynomial. (If $p(y) := y^3 + ay + b$ factorizes as $(y - \alpha)\,q(y)$, its derivative is $q(y) + (y - \alpha)\,q'(y)$, which can only have α as a root if $q(\alpha) = 0$ —that is, if α is a repeated root of p).

The upshot is that, if two of the formulæ (6) have the same value, the cubic (1) anyway has a repeated root; if all the formulæ (6) have the same value, the cubic has three coincident roots (all zero, because it is a reduced cubic). So the formulæ (6) do, in fact, give all the roots of (1),

and, what is more, the vanishing of the cubic discriminant is the necessary and sufficient condition for there to be a repeated root.

In the argument so far I have taken complex numbers for granted, by assuming the existence of *three* cube roots (and of at least one square root). But now, consider what the mathematicians of the sixteenth century, for whom the coefficients a and b had to be real, could understand by Cardan's formula.

When $\Delta > 0$, the quantities whose cube roots appear in the formula are both real, have product $-\frac{1}{27}a^3$, and differ from each other; thus we may choose w_0 and z_0 to be real and different. The formulæ (5) give a single real value. In addition, there are two mutually conjugate complex values (notice that $\overline{\omega} = \omega^2$) which our men of the sixteenth century, recognising only real cube roots of real numbers, could not perceive.

If $\Delta = 0$, we can take z_0 and w_0 to be real and equal. (This needs checking, for recall that $w_0 z_0$ is to be $-\frac{1}{3}a$.) Hence the formula certainly gives us *one* real root, $z_0 + w_0 = 2z_0$. A sixteenth-century mathematician could then use this root to reduce the degree of the equation (dividing by $z - z_0$), and could solve the resulting quadratic to obtain the other two roots (which must coincide, as we saw above, and so be equal to $-z_0$). However, Cardan's formula, in his understanding, would *not* have given him these roots. In *modern* terms it yields the two expressions $\omega z_0 + \omega^2 w_0 = -z_0$ and $\omega^2 z_0 + \omega w_0 = -z_0$, which, although themselves real, arise from non-real cube roots in the formula. This is odd enough.

When $\Delta < 0$, the peculiarity of the formula becomes quite striking. In this case—a and b being real—a must be negative, from the definition of Δ; hence w_0 and z_0 are cube roots of two conjugate complex numbers, and must themselves be conjugate (as their product $-\frac{1}{3}a$ is to be positive). Thus (6) gives $z_0 + \overline{z}_0$, $\omega z_0 + \overline{\omega}\,\overline{z}_0$, $\overline{\omega} z_0 + \omega \overline{z}_0$, all of which are real and different (as we saw above). Notice, by the way, that this means that, when the coefficients a and b of the reduced cubic are real, the sign of Δ determines whether there are three distinct real roots (when $\Delta < 0$), two real roots with one repeating (when $\Delta = 0$), or only one (when $\Delta > 0$). Our argument shows that in the case of three real roots, $\Delta < 0$, *none at all* of them could be derived by a sixteenth-century mathematician from Cardan's formula, since they all involve $\sqrt{\Delta}$.

At the time this was extremely mysterious. The formula seemed to be "correct", in some metaphysical sense; its derivation was convincing when the various roots existed, and it certainly gave solutions of the cubic when it could be interpreted in real terms. But in the case $\Delta < 0$ it did not detect any roots at all. And yet it was clear that any real cubic has at least one real root, since large positive values of x make $x^3 + (\text{quadratic})$ positive and large negative values make it negative. (Of course, the intermediate value theorem had not been rigorously proved at the time—the apparatus did not exist for any modern "proof"; but, like many other theorems of elementary analysis, it was pretty obviously true at least for polynomials.) In fact, it is not difficult to see, by checking the values of the cubic expression at some particular choices of x and applying the intermediate value theorem, that there must even be *three* real roots if $\Delta < 0$. This baffling situation—there is at least one root for quite elementary reasons, and one can even see in a fairly straightforward way that there must be three, but the formula, which seems to be correct in other respects, makes no sense—was therefore called the "irreducible case".

Cardan's suggestion, which Bombelli elaborated to a system of algebra, was to get over this seeming absurdity by 'pretending' that the negative number Δ did have a square root. As

we know, this leads to the existence of three distinct cube roots of non-zero complex numbers. Then Cardan's formula yields all the roots in all cases, as we found above, and the anomalies fall away. But there seemed at the time to be no legitimate explanation for this. I do not know what people really thought then; there have been far more recent examples, however, of concepts that were accepted as useful despite a general awareness that they could not be justified as they stood (one that you may find familiar is Dirac's δ-function—but, now that ideas spread far more rapidly, it was justified, and in a valuable way, after only a few years).

The curious aspect of these arguments, from our point of view, is that the solution of a real quadratic did not present a like paradox to a sixteenth-century mathematician. The formula for the solution of a quadratic would make no sense to him precisely when he would say that there are no solutions anyway, namely when the quadratic discriminant is negative. In graphical terms, a quadratic expression is represented by a parabola that may never cross the x-axis, whilst a cubic curve must cross it at least once and possibly three times. So it was only the solution of the cubic that motivated "complex numbers".

APPENDIX D

25 Cauchy's theorem.

§D0. Introduction.

In the lectures, I adopted the policy of explaining and using a rather strong version of Cauchy's theorem, whilst "proving" only a weak version (my proof relied on Green's theorem, which itself is rather difficult to prove satisfactorily). The purpose of these notes is to discuss a proof of the version I have actually used. In the interests of brevity and readability, I shall not be absolutely rigorous, but all the steps I fudge can be supplied fairly easily by using facts from the earlier part of this course.

Cauchy's original proof of his theorem was more or less the one I gave. That is, he considered only functions strongly holomorphic on a region (recall that these are the functions for which $f'(z)$ not only exists at each point of the region, but is also a continuous function of z), and then applied Green's theorem. The argument should not be despised, because it can be generalized to higher dimensions in various ways, whereas the more precise proof I shall soon give is rather more specific to $\mathbb{C}$. Furthermore, Cauchy's version of his theorem is indeed sufficient, with a few contemptible extra tricks, to prove all the later results of the course, provided that one restricts attention to strongly holomorphic functions. All the functions we shall be interested in are obviously strongly holomorphic. On the other hand:

1. We now know, though Cauchy didn't, that a holomorphic function on a region *must* be strongly holomorphic. A proof of Cauchy's theorem which requires strong holomorphicity is, therefore, making a superfluous assumption. In practical terms this is perhaps of little importance, but the fact is startling, and by far the easiest *proof* that holomorphic functions are strongly holomorphic uses Cauchy's theorem for *holomorphic* functions.

2. There are topological problems involved. If you recall the proof of Green's theorem that you have seen (in whatever version), it amounts ultimately to the fundamental theorem of

calculus, plus the geometrical assumption that any line parallel to one of the axes cuts the contour of integration in a particular way. Crudely speaking, the points that it has in common with the contour can be classified as "points of entry" (into the region enclosed by the contour), "points of exit", or "neither". The fundamental theorem of calculus is applied to the intervals between points of entry and points of exit, which are assumed to occur in pairs. In effect, the assumption is made that a closed contour must have an "inside" and an "outside". For closed contours that do not cross themselves, this looks overwhelmingly obvious, and, in Green's and Cauchy's day, mathematicians were disinclined to worry about it. That it is *not* obvious can be seen if you consider circles on the surface of a torus, some of which do split the surface in two and some of which don't; those which 'go right round' the torus don't. The property that a closed non-self-intersecting contour splits the plane into an "inside" and an "outside" is a property in some measure specific to the plane, and is the *Jordan curve theorem*, which was first proved some decades after Cauchy's work.

The most accessible proof (not Jordan's, but Eilenberg's) of the Jordan curve theorem uses some of the ideas in the proof of Cauchy's theorem below.

That is not all, however. The condition that the contour be a Jordan contour is odd; for a change of parametrization of the contour does not alter the integral, but may easily make the contour non-Jordan. It is thus not the "right" condition. (Of course, it is quite conceivable that it might nevertheless be the only convenient one.) However, there are many situations in which the contour integral must manifestly be 0 though the contour is not Jordan. In 16.10, we mentioned a contour that gives a zero integral for any function holomorphic in $\mathbb{C} \setminus \{-1, 1\}$, but that is not Jordan.

The moral is that Cauchy's statement of his theorem is not ideal, either analytically or topologically. We shall see that his proof (even if supported with a full proof of Green's theorem) is not the ideal proof.

3. In practice, one tends to need Cauchy's theorem in situations to which 16.3 does not apply directly (because the contours are not Jordan contours, etc.), and this leads to various messy techniques where contours are split up into Jordan contours or expressed as limits of Jordan contours. It is better to have a more general version at the outset, and then the applications follow easily and naturally.

§D1. Goursat's Lemma

An amazingly simple and elegant argument of Goursat proves Cauchy's theorem in a case with no topological difficulties. By the 'boundary' ∂R of a rectangle R, I mean the contour obtained by travelling round its sides in the obvious way (anticlockwise); you can easily write down appropriate formulæ. In the proof below, I have employed diagrams to clarify the notation, but it is evident that the argument does not really depend on the geometry at all and could be set out in purely analytical terms.

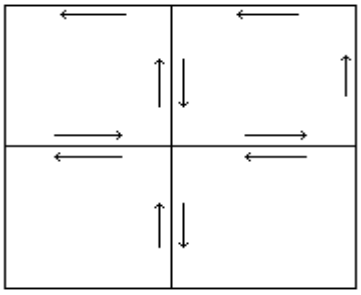
Lemma 25.1. *Let R be a rectangle included in the region U of $\mathbb{C}$, and let $f: U \longrightarrow \mathbb{C}$ be holomorphic. Then the integral of f around the boundary of R is zero.*

Proof. Suppose that $|\int_{\partial R} f| = a > 0$. Let the perimeter of R be L. If R is split into four congruent "quarters" $R(i)$, for $i = 1, 2, 3, 4$ (see the diagram below), then

$$\int_{\partial R(i)} f = \sum\nolimits_{i=1}^{4} \int_{\partial R(i)} f\,,$$

because the integrals over the *internal* line segments cancel two by two. Hence

$$a \le \left| \int f \right| \le \sum\nolimits_{i=1}^{4} \left| \int_{\partial R(i)} f \right|,$$



and there is an index i such that $\left| \int_{\partial R(i)} f \right| \ge \frac{1}{4}a$. The rectangle $R(i)$ may be similarly subdivided, and so on. In this way one may select a sequence of rectangles $J(0) = R$, $J(1) = R(i)$, $J(2), J(3), \dots$, such that each is a quarter of its predecessor, and, for each $n = 0, 1, 2, 3, \dots$,

$$4^{-n} a \le \left| \int_{\partial J(n)} f \right|. \tag{D1}$$

(Incidentally, this is another situation in which Littlewood might have given a scheme for making the choice of $J(n)$ amongst the four quarters of $J(n-1)$; see my remark after 8.17.) The rectangles $J(n)$ must converge to a point $z_0 \in R$. That is to say:

$$(\forall \delta > 0)(\exists N(\delta) \in \mathbb{N}) \quad n \ge N(\delta) \implies J(n) \subseteq B(z_0; \delta)\,.$$

(This follows from any of several related arguments, the details of which you can easily supply. For instance: if one takes a point $z_n \in J(n)$ for each n, then all the (z_n) for $n \ge k$ lie in $J(k)$, and, since the diameter of $J(k)$ is 2^{-k} times the diameter of R, the sequence (z_n) is Cauchy, therefore must converge, and its limit is z_0. Or, take any point z_i in $J(i)$ for each i—or, to avoid the Axiom of Choice, take the bottom left-hand point of $J(i)$—and the sequence (z_i) has a convergent subsequence, by sequential compactness of $J(0)$, with a limit z_0.)

Now f is complex-differentiable at $z_0 \in U$: there exists $\delta > 0$ such that

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| < \frac{a}{L} \qquad \text{when } 0 < |z - z_0| < \delta\,.$$

Hence, if $n \ge N(\delta)$ and $z \in J(n)$,

$$|f(z) - f(z_0) - (z - z_0) f'(z_0)| \le \frac{a}{L} |z - z_0\,|. \tag{D2}$$

(This is true when $z = z_0$ as well.) But

$$\left| \int_{\partial J(n)} f \right| \le \left| \int_{\partial J(n)} (f(z) - f(z_0) - (z - z_0) f'(z_0))\, dz \right| + \\ + \left| \int_{\partial J(n)} (f(z_0) + (z - z_0) f'(z_0))\, dz \right|.$$

But $f(z_0) + (z - z_0) f'(z_0)$ is $\dfrac{d}{dz}\{z f(z_0) + \frac{1}{2}(z - z_0)^2 f'(z_0)\}$, and is continuous. So the

last integral is 0 by 15.9*(b)*, and therefore, using (D2) and the fundamental estimate,

$$\left| \int_{\partial J(n)} f \right| \leq \int_{\partial J(n)} \frac{a}{L} \, |z - z_0| \, dz \, . \tag{D3}$$

The perimeter of $J(n)$, i.e. the length of $\partial J(n)$, is $2^{-n}L$, and, for every point $z \in \partial J(n)$, $|z - z_0| \leq 2^{-n-1}L$; the integral on the right of (D3) cannot exceed $2^{-2n-1}La/L$. Putting this together with (D1),

$$4^{-n}a \leq \left| \int_{\partial J(n)} f \right| \leq 2^{-2n-1}a \, .$$

Hence $a \leq \frac{1}{2}a$, which is absurd; the initial assumption that $a > 0$ must have been false. □

The proof only uses complex-differentiability of f at the single point $z_0 \in U$, but there is no prior information on the location of z_0 in R, so the hypotheses must demand differentiability at all points of R. (It is possible to weaken them slightly, but little is gained in the end by doing so). The really striking fact, however, is that only complex-differentiability is used. There is an analogous argument for Green's theorem on a rectangle (briefly described at 25.16), but the requirement that the curl be integrable over the rectangle forces one to assume more than mere differentiability of the vector field—customarily, although rather too demandingly, that its derivatives be continuous as well.

§D2. The rectangle property

Definition 25.2. A region U has the *rectangle property* with respect to $w \in U$ if, for any $z \in U$, the closed rectangle with w and z as opposite vertices lies entirely in U.

Corollary 25.3. *Let f be holomorphic in a region U having the rectangle property with respect to some $w \in U$. f is the complex derivative of a strongly holomorphic function in U.*

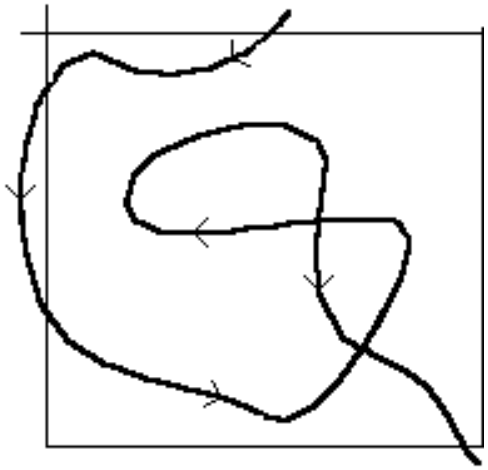
Proof. For any $z \in U$, let $g(z)$ be the integral of f around two sides of the rectangle of which w and z are opposite vertices (starting at w and ending at z). There are usually two choices for the contour of integration, but 25.2 ensures they give the same answer. Calculate $\partial g/\partial x$ by taking that contour of integration which ends with a horizontal segment. Hence $\partial g/\partial x = f(z)$, and is continuous in z as f is. From the other contour of integration, $\partial g/\partial y = if(z)$. Hence, g is C^1 on U and the Cauchy-Riemann equations are satisfied at each point $z \in U$. By 12.11, g is complex-differentiable at z, and $g'(z) = \partial g/\partial x = f(z)$. So g is strongly holomorphic in U. □

Proposition 25.4. *Let f be holomorphic in a region U having the rectangle property with respect to one of its points. The integral of f around any closed contour in U is* 0.

Proof. Put together 25.3 and 15.9. □

Remark 25.5. It is evident that a disk $B(a;r)$, or any open *coordinate* rectangle containing a, has the rectangle property with respect to a. (There are many other examples of such regions).

Suppose f is holomorphic in a region U, and that $\gamma : [\alpha, \beta] \longrightarrow U$ is a closed contour. Choose points $t_0 = \alpha < t_1 < t_2 < \ldots < t_k = \beta$ such that, for $1 \leq i \leq k$, $\gamma(t_{i-1})$ and $\gamma(t_i)$ are points on the frontier of a closed rectangle that is wholly included in U, and $\gamma(t)$ lies in the same rectangle for $t_{i-1} < t < t_i$ —the proof of this involves Lebesgue numbers (see 23.6), and I omit it. Hence, from 25.4, the integral of f over the contour $\gamma|[t_{i-1}, t_i]$ is equal to its integral over parts of the edges of the rectangle (see the diagram, in which $\gamma(t_{i-1})$ is the third point where the contour crosses the edges of the rectangle; there is a choice of two routes over the edges of the rectangle, but this is unimportant, since both choices give the same integral).



Proceeding in this way for each i, we can substitute γ by a new contour consisting of rectilinear segments each parallel to one of the axes, and each traversed at a constant rate. Let us call such a contour a *rectangular contour*.

Lemma 25.6. Let γ be any contour in the region U. Then there is a rectangular contour ρ in U, having the same end-points as γ (in particular, ρ is closed if γ is), such that, for any function f holomorphic on U,

$$\int_\gamma f = \int_\rho f\,. \qquad \square$$

§D3. Cycles and winding numbers

In the lectures I made a brief mention of the notion of a closed chain or cycle. (More precisely, we were interested in closed piecewise C^1 chains with integer coefficients). A fuller definition is as follows: a *cycle*, or *closed chain*, is a formal sum $\sum m_i\gamma_i$ of finitely many terms, where each term consists of an integer coefficient m_i multiplying a closed contour γ_i. More generally, a *chain* is a similar formal linear combination of contours that are not necessarily closed.

At first sight these seem perverse definitions. Why introduce 'linear combinations' that have no geometrical meaning? (The use of the word 'formal' here is in effect the same as after 9.12: these linear combinations are just *expressions*, having no more concrete interpretation;

they are 'formal' because their 'form' is all they have). I shall not try to justify doing so in any detail, but the idea is that a chain is (more or less) the most general object over which we can integrate a function. If $\mathcal{C} = \sum m_i\gamma_i$ as above, one simply defines

$$\int_{\mathcal{C}} f = \sum m_i \int_{\gamma_i} f\,.$$

One *thinks* of the chain as consisting of m_i copies of γ_i, for each i. Just as we write $z \in \gamma$ to mean that z is a value taken by the function γ (when γ is a path or contour), so also we write $z \in \mathcal{C}$ to mean that z is a value of one of the functions γ_i that have a non-zero coefficient m_i, and say that 'z is a point of the chain', although the chain is neither a subset of $\mathbb{C}$ nor even a function. The *complement* of $\mathcal{C}$ is the set $\{z : z \notin \mathcal{C}\}$.

Cycles can be added or subtracted in the obvious way (adding or subtracting the coefficients of a contour that appears in both cycles, etc.; treating terms that have coefficient 0 as algebraically "zero"; and so on.)

Let $\gamma : [\alpha, \beta] \longrightarrow \mathbb{C}$ be a closed contour, with $\gamma(\alpha) = \gamma(\beta)$, and suppose that $a \notin \gamma$. Thus γ traces out a curve that does not pass through a, and, since the function $1/(z-a)$ is holomorphic except at a, its integral around γ exists. In this situation

Lemma 25.7. $\dfrac{1}{2\pi i}\displaystyle\int_\gamma \frac{dz}{z-a}$ *is a real integer.*

Proof. For $\alpha \le t \le \beta$, let $F(t) := \displaystyle\int_\alpha^t \frac{\gamma'(\tau)\,d\tau}{\gamma(\tau)-a}$. For any t at which $\gamma'(t)$ exists, which will also be a point of continuity of γ' (by the definition of a C^1 path),

$$\frac{d}{dt}\{(\gamma(t)-a)\exp(-F(t))\} = \gamma'(t)\exp(-F(t)) - (\gamma(t)-a)F'(t)\exp(-F(t)) = 0\,,$$

since $F'(t) = \gamma'(t)/(\gamma(t)-\alpha)$. Hence $(\gamma(t)-a)\exp(-F(t))$ must be constant on each of the C^1 paths making up γ, and hence on the whole of $[\alpha,\beta]$, with value q. So, as $F(\alpha) = 0$,

$$1 = \exp(-F(\alpha)) = \frac{q}{\gamma(\alpha)-a} = \frac{q}{\gamma(\beta)-a} = \exp(-F(\beta))\,, \quad \text{and}$$
$$\exp(F(\beta)) = 1\,,$$

from which we know that $\dfrac{1}{2\pi i}\displaystyle\int_\gamma \frac{dz}{z-a} = \frac{F(\beta)}{2\pi i}$ is a real integer. □

Definition 25.8. If γ is a closed contour and $a \notin \gamma$, then the *winding number* of γ about a, or the *index* of a with respect to γ, is the real integer

$$n(\gamma, a) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}\,.$$

If $n(\gamma, a) \neq 0$ we say γ *winds* about a. More generally, if $\mathcal{C} = \sum m_i\gamma_i$ is a cycle (the γ_i being closed contours) and $a \notin \mathcal{C}$, the winding number $n(\mathcal{C}, a)$ of $\mathcal{C}$ about a (or the index of a with respect to $\mathcal{C}$) is

$$n(\mathcal{C}, a) = \sum m_i n(\gamma_i, a)\,.$$

Lemma 25.9. *Given a cycle $\mathcal{C}$, $n(\mathcal{C},a)$ is continuous on the complement of $\mathcal{C}$.*

Proof. By 17.5, it is complex-differentiable on the complement of $\mathcal{C}$. □

Lemma 25.10. *For any cycle $\mathcal{C}$, there exists $R \geq 0$ such that $n(\mathcal{C},a) = 0$ when $|a| > R$.*

Proof. Choose $R = \sup\{|z|: z \in \mathcal{C}\} > 0$. Then, if $|a| > R$, $n(\mathcal{C},ta)$ is a continuous function of t for all $t \geq 1$ (since $ta \notin \mathcal{C}$). Since it takes only integer values, it must be constant. Now, however, if $\mathcal{C} = \sum m_i\gamma_i$ and $t > 1$,

$$n(\mathcal{C},ta) = \frac{1}{2\pi i}\sum m_i \int_{\gamma_i} \frac{dz}{z-ta} \quad \text{and}$$

$$|n(\mathcal{C},ta)| \leq \frac{1}{2\pi}\sum \frac{m_i\,\ell(\gamma_i)}{(t-1)R} \qquad \text{by the fundamental estimate}$$

(here $\ell(\gamma_i)$ is the length of γ_i, and $|z-ta| \geq tR - R$ because $z \in \mathcal{C}$). The right-hand side may be made less than 1 by taking a large enough value of t, and therefore the constant value of the integer $n(\mathcal{C},ta)$ must be 0. □

§D4. Homology.

Definition 25.11. Let $\mathcal{C}$ be a cycle in the region U of $\mathbb{C}$. Then $\mathcal{C}$ is *nullhomologous* in U or *homologous to* 0 in U if its winding number about any point *outside* U is 0. This is sometimes written "$\mathcal{C} \sim 0$ in U". Two cycles $\mathcal{C}_1$ and $\mathcal{C}_2$ are homologous in U if $\mathcal{C}_1 - \mathcal{C}_2 \sim 0$ in U, or, equivalently, if $n(\mathcal{C}_1,a) = n(\mathcal{C}_2,a)$ for any $a \notin U$. We say U is *simply-connected* in $\mathbb{C}$ if any cycle in U is nullhomologous in U. It is both necessary and sufficient for this that any closed contour in U should be nullhomologous in U.

In intuitive terms, $\mathcal{C} \sim 0$ in U if it does not wind about any "hole" in U, and U is simply-connected if it has no "holes" for a cycle to wind round. The general notion of "homology" in algebraic topology has to do (speaking very vaguely) with the number and the dimension of "holes" in a space. However, the definitions I have just given, although extremely convenient for the formulation and proof of Cauchy's theorem, are *not* the standard ones of algebraic topology. They are equivalent to those standard ones when one restricts attention to a region U of $\mathbb{C}$, but the equivalences are not by any means obvious and involve lengthy argument. For my purposes, the generally accepted meanings of the phrases "homologous" and "simply-connected" are unnecessary, require far too much preparation, and will not be given.

Remark 25.12. Suppose now that we have a cycle $\mathcal{C} = \sum m_i\gamma_i$ in a region U. According to D2.4, each of the closed contours γ_i appearing with non-zero coefficient in $\mathcal{C}$ may be substituted by a rectangular closed contour ρ_i in U, in such a way that, for any holomorphic function f in U, the integrals of f over ρ_i and over γ_i are the same. Thus the integrals of f over $\mathcal{C}$ and over the 'rectangular cycle' $\sum m_i\rho_i$ are also the same, and, in effect, we need only consider rectangular cycles.

A rectangular cycle $\mathcal{D}$ is a finite sum of integral multiples of closed rectangular contours, each consisting of finitely many rectilinear segments parallel to the coordinate axes. Extend each such segment to a bi-infinite line of the form $x =$ constant or $y =$ constant. (It is

possible that several segments belonging to contours of $\mathcal{D}$ may lie on the same line). This divides the plane into a grid; the closed contours of the cycle go along some of the edges of the grid, and the plane is split up into a finite collection of bounded rectangles and semi-infinite strips (i.e. "rectangles with one edge missing" or "unbounded rectangles"). Let $\mathcal{R}$ denote the class of bounded or unbounded rectangles thus created.

If $R \in \mathcal{R}$, take a point b in the interior of R, and set $n(\mathcal{D}, R) = n(\mathcal{D}, b)$. This index is independent of the choice of the point b, for any two points of the interior of R can be joined by a line segment in the interior of R, which does not meet $\mathcal{D}$; along this line the index with respect to $\mathcal{D}$ therefore varies continuously, by 25.9, and must be constant. When R is an unbounded rectangle, $n(\mathcal{D}, R) = 0$ by 25.10. Now, set

$$\mathcal{E} = \sum\nolimits_{R \in \mathcal{R}} n(\mathcal{D}, R)\, \partial R. \tag{D5}$$

This is also a rectangular cycle (although not necessarily in U).

Lemma 25.13. *If $\mathcal{D}$ is nullhomologous in U, then each $R \in \mathcal{R}$ for which ∂R appears with nonzero coefficient in $\mathcal{E}$ is wholly included (together with ∂R) in U.*

Proof. Suppose that $n(\mathcal{D}, R) \neq 0$. By the definition of $n(\mathcal{D}, R)$, this means that $n(\mathcal{D}, b) \neq 0$ for any $b \in \operatorname{int} R$; since $\mathcal{D}$ is nullhomologous in U, this implies that $b \in U$. So $\operatorname{int} R \subseteq U$. Suppose there is a point $c \in \partial R$ such that $c \notin U$. Then $c \notin \mathcal{D}$ too (for $\mathcal{D}$ is a cycle in U); $n(\mathcal{D}, c)$ is defined, and must be equal to $n(\mathcal{D}, R)$ by continuity, since there are points of $\operatorname{int} R$ as close as one wishes to c. Hence $n(\mathcal{D}, c) \neq 0$, which contradicts the assumptions that $c \notin U$ and $\mathcal{D}$ is nullhomologous in U. It follows that $\partial R \subseteq U$ too. □

Lemma 25.14. *Any function holomorphic on U has the same integral around $\mathcal{E}$ as around $\mathcal{D}$.* [Notice that 25.13 ensures the integral around $\mathcal{E}$ makes sense].

Proof. Let f be holomorphic on U. $\int_{\mathcal{E}} f$ and $\int_{\mathcal{D}} f$ are sums of integral multiples of integrals over parametrized edges of rectangles $R \in \mathcal{R}$. For each such edge I, choose a specific parametrization, to get a contour J going from one end-point of I to the other.

In the cycles $\mathcal{D}$ and $\mathcal{E}$, the edge I appears finitely many times, with various parametrizations. Each appearance contributes to the integral over the cycle in question; the contribution will be $\int_J f$ if the parametrization is in the same sense as J (I may call this an appearance with coefficient $+1$) and $-\int_J f$ otherwise (an appearance with coefficient -1).

I must be an edge of exactly two rectangles $R_1, R_2 \in \mathcal{R}$. For definiteness, suppose that R_1 lies to the left of J; that is, J is part of the closed contour ∂R_1, described anticlockwise. Then R_2 must lie to the right of J. I assert that the sum σ of the coefficients of all appearances of I in $\mathcal{D}$ is $n(\mathcal{D}, R_1) - n(\mathcal{D}, R_2)$.

Firstly, if I does not appear in $\mathcal{D}$ at all, then $n(\mathcal{D}, z)$ is defined and continuous at and near the mid-point d of I, so that $n(\mathcal{D}, R_1) = n(\mathcal{D}, R_2)$, and $\sigma = 0$ as required.

Otherwise, draw a circle Γ about d, of radius so small that it is included in $R_1 \cup R_2$. Let Δ be the open disk with Γ as frontier, and take points $w_1 \in \Delta \cap R_1$, $w_2 \in \Delta \cap R_2$. Construct a new cycle $\mathcal{D}'$ by substituting, for each appearance in $\mathcal{D}$ of the diameter of Γ, a semicircular contour around the half of Γ in R_1. Each substitution changes the winding number about w_1. If, for instance, the appearance of I had coefficient $+1$, the substitution removes the (anticlockwise) diameter and adds the (clockwise) semicircle, in effect subtracting

1 from the winding number. (See the note at the end of the proof). In this way

$$n(\mathcal{D}', w_1) = n(\mathcal{D}, w_1) - \sigma\,.$$

The line segment between w_1 and w_2 does not cross $\mathcal{D}'$, so that, from 25.9,

$$n(\mathcal{D}', w_1) = n(\mathcal{D}', w_2)\,,$$

whilst $n(\mathcal{D}', w_2) = n(\mathcal{D}, w_2)$, since each substitution makes no alteration in these winding numbers. The assertion follows:

$$\begin{aligned}\sigma = n(\mathcal{D}, w_1) - n(\mathcal{D}', w_1) &= n(\mathcal{D}, w_1) - n(\mathcal{D}, w_2)\\ &= n(\mathcal{D}, R_1) - n(\mathcal{D}, R_2)\,.\end{aligned}$$

However, since I can appear in $\mathcal{E}$ only as an edge of R_1 or as an edge of R_2, and in no other way, this proves that σ is precisely the sum of the coefficients of the appearances of I in $\mathcal{E}$. Thus I makes the same contribution to $\int_{\mathcal{E}} f$ as to $\int_{\mathcal{D}} f$.

Adding over all the edges I, one has the result. □

Note. I have treated as obvious the facts that a clockwise circuit around a "semicircle plus diameter" Γ about w_1 has winding number -1 about w_1, and 0 about w_2. The second statement results from 25.9 and 25.10, for w_2 may be joined along a radius of Γ to a point of arbitrarily large modulus. For the first, change the semicircle to a clockwise circle; in effect this means adding the second semicircle, also clockwise, which has winding number 0 by the argument just given. The winding number of the circle about w_1 is the same as about d, which is -1 by explicit calculation (as at 17.12).

Theorem 25.15. (Cauchy's theorem, topological form). *Suppose f is a function holomorphic on the region U. Then, for any cycle $\mathcal{C}$ nullhomologous in U, $\int_{\mathcal{C}} f = 0$.*

Proof. Use 25.12 to substitute $\mathcal{C}$ by a rectangular cycle $\mathcal{D}$ in U over which, by 25.14, all functions holomorphic in U have the same integrals as over $\mathcal{C}$. For $a \notin U$, then,

$$n(\mathcal{D}, a) = \frac{1}{2\pi i}\int_{\mathcal{D}} \frac{dz}{z-a} = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{dz}{z-a} = n(\mathcal{C}, a) = 0\,,$$

since $1/(z-a)$ is holomorphic on U. Thus $\mathcal{D} \sim 0$ in U, and I may construct a cycle $\mathcal{E}$ from $\mathcal{D}$ by (D5); then

$$\int_{\mathcal{C}} f = \int_{\mathcal{D}} f = \int_{\mathcal{E}} f = \sum\nolimits_{R\in\mathcal{R}} n(\mathcal{D}, R)\int_{\partial R} f$$

by definition. By 25.13, the only terms on the right for which $n(\mathcal{D}, R) \neq 0$ have $R \subseteq U$, and therefore $\int f = 0$ by Goursat's lemma 25.1. □

Remarks 25.16. I hinted in the introduction to Appendix D at like arguments for Green's theorem. The idea would be to consider, given the C^1 vector field X, for what cycles $\mathcal{C}$

$$\int_{\mathcal{C}} X\cdot ds = \int n(\mathcal{C}, x)\,\mathrm{curl}\,X(x)\,dS(x)\,,$$

where, of course, "dS" means the integral with respect to area and the integral is in principle

extended over the whole plane (because of 25.10, the integrand is zero outside a sufficiently large disk). curl X is in effect a scalar-valued function. Goursat's argument is readily modified to prove this for rectangles, and the general result follows much as for Cauchy's theorem.

As I remarked previously, elementary proofs of Green's theorem assume something like the Jordan curve theorem (and are in some ways less general than the form above, since they presuppose that the contour does not cross itself). The standard *advanced* proof of Green's theorem, as a special case of the n-dimensional "Stokes's theorem", avoids the Jordan curve theorem by assuming that the contour is the frontier of a 'cell'. The added generality of n dimensions is attained by making suitable geometrical assumptions at the start. Moreover, the formula above does not generalize in a direct way.

Our proof of Cauchy's theorem is essentially due to Emil Artin, who noticed that the rudimentary homology theory (winding numbers) needed in the proof may be developed entirely within complex analysis. I have taken it, with minor modifications, from Ahlfors's well-known textbook. (Eilenberg's proof of the Jordan curve theorem appears in his exercises.) His exposition is quite readable already. If you wish to dismiss the arguments as unnecessarily elaborate, compare some other books—there is a whole paperback by M. H. A. Newman whose avowed aim is to expound the plane topology required for a satisfactory discussion of Cauchy's theorem, and there are textbooks of complex analysis, such as that by Thron, which devote disproportionate space to the topic. It should also be added that there are other versions of Cauchy's theorem with slightly weaker analytical hypotheses on f, but they are of little interest at this stage; indeed, it is surprising how rarely one needs to apply the more refined versions. Even the full strength of the version given above is not really needed anywhere in the course, since we only used quite simple contours.

Another approach is to prove a "homotopy" form of the theorem, rather than the "homology" form studied above. Heins's text does this. However, I cannot see any advantage in doing so.