MT 2016

ANALYSIS I

Webnotes by H.A. Priestley

Contents

A summary of the axioms for the real numbers is given in a separate reference sheet. References [**BS** ...] *are to the textbook by Bartle and Sherbert.*

---

0. Setting the scene

Introductory remarks on the content, aims and style of the Analysis I course, and how it relates to school mathematics and to other Prelims courses.

---

1. The Real Number System: Field Axioms

[**BS, Section 2.1**]

1.1. Real numbers and axioms for them.

What is a real number? The need for clear assumptions about the real numbers and the role of these axioms in the course.

Notation: the set of real numbers is denoted $\mathbb{R}$.

We introduce the axioms for $\mathbb{R}$ and explore how the familiar rules of arithmetic can be obtained as consequences of the axioms (and of properties derived from the axioms).

See reference sheet Axioms for the Real Numbers for the list of axioms.

Addition

The operation $+$ of addition and the axioms, **A1**–**A4**, for addition.

Multiplication and avoiding collapse

The operation $\cdot$ of multiplication and the axioms, M1–**M4**, for multiplication. The need to assume $0 \neq 1$ (Axiom **Z**).

Distributive Law

Linking addition and multiplication together via the distributive law **D**.

1.2. Properties of arithmetic.

Properties (1)–(5) require **A1–A4** only, (6)–(8) require **M1–M4** only. (9)–(13) together draw on all the axioms. Here a, b, c, x and y are real numbers.

(1) If $a + x = a$ for all a then $x = 0$ (uniqueness of zero element).
(2) If $a + x = a + y$ then $x = y$ (cancellation law for addition, which implies uniqueness of additive inverse of a).
(3) $-0 = 0$.
(4) $-(-a) = a$.
(5) $-(a + b) = (-a) + (-b)$.

(6) If $a \cdot x = a$ for all $a \neq 0$ then $x = 1$ (uniqueness of multiplicative identity).
(7) If $a \neq 0$ and $a \cdot x = a \cdot y$ then $x = y$ (cancellation law for multiplication, which implies uniqueness of multiplicative inverse of a).
(8) If $a \neq 0$ then $1/(1/a) = a$.

(9) $(a + b) \cdot c = a \cdot c + b \cdot c$.

(10) $a \cdot 0 = 0$.
(11) $a \cdot (-b) = -(a \cdot b)$. In particular $(-1) \cdot a = -a$.
(12) $(-1) \cdot (-1) = 1$.
(13) If $a \cdot b = 0$ then either $a = 0$ or $b = 0$ (or both). Moreover, if $a \neq 0$ and $b \neq 0$ then $1/(a \cdot b) = (1/a) \cdot (1/b)$.

Full proofs are given below but only a small selection of these will be given in lectures, to illustrate how derivations from the axioms should look. You are recommended to work through some of the others by yourself, referring to the notes only to check that you haven't cut corners or made correct assertions you have omitted to validate.

Proof. One step at a time!

(1) We have

$$\begin{aligned} x &= x + 0 && \text{by } \mathbf{A3} \\ &= 0 + x && \text{by } \mathbf{A1} \\ &= 0 && \text{by the hypothesis, with } a = 0. \end{aligned}$$

(2) We have

$$\begin{aligned} y &= y + 0 && \text{by } \mathbf{A3} \\ &= y + (a + (-a)) && \text{by } \mathbf{A4} \\ &= (y + a) + (-a) && \text{by } \mathbf{A2} \\ &= (a + y) + (-a) && \text{by } \mathbf{A1} \\ &= (a + x) + (-a) && \text{by hypothesis} \\ &= (x + a) + (-a) && \text{by } \mathbf{A1} \\ &= x + (a + (-a)) && \text{by } \mathbf{A2} \\ &= x + 0 && \text{by } \mathbf{A4} \\ &= x && \text{by } \mathbf{A3}. \end{aligned}$$

(3) $0 + 0 = 0$ by **A3** with $a = 0$. Hence $-0 = 0$ by **A4** and (2).

(4) We have

$$\begin{aligned} (-a) + a &= a + (-a) && \text{by } \mathbf{A1} \\ &= 0 && \text{by } \mathbf{A4}\text{, and} \\ (-a) + -(-a) &= 0 && \text{by } \mathbf{A4}. \end{aligned}$$

Now appeal to (2) (cancellation law for addition).

(5) Exercise on Problem sheet 1.

Claims (6)–(8) do for multiplication what (1)–(3) do for addition, but with inverses necessarily considered only for non-zero a, as **M4** requires. The proofs go the same way.

(9) Note that the statement is like Axiom **D** but with multiplication on the other side. We have

$$\begin{aligned} (a + b) \cdot c &= c \cdot (a + b) && \text{by } \mathbf{M1} \\ &= c \cdot a + c \cdot b && \text{by } \mathbf{D} \\ &= a \cdot c + b \cdot c && \text{by } \mathbf{M1} \text{ twice.} \end{aligned}$$

(10) We have

$$\begin{aligned} a \cdot 0 + 0 &= a \cdot 0 && \text{by } \mathbf{A3} \\ &= a \cdot (0 + 0) && \text{by } \mathbf{A3} \\ &= a \cdot 0 + a \cdot 0 && \text{by } \mathbf{D}. \end{aligned}$$

Now appeal to (2) (cancellation law for addition).

(11) We have

$$\begin{aligned} (a \cdot b) + (a \cdot (-b)) &= a \cdot (b + (-b)) && \text{by } \mathbf{D} \\ &= a \cdot 0 && \text{by } \mathbf{A4} \\ &= 0 && \text{by (10). Also} \\ (a \cdot b) + (-(a \cdot b)) &= 0 && \text{by } \mathbf{A4}. \end{aligned}$$

Now appeal to (2).

(12) We have

$$\begin{aligned} (-1) \cdot (-1) &= -(-1) && \text{by (11)} \\ &= 1 && \text{by (4).} \end{aligned}$$

(13) Assume for a contradiction that $a, b \neq 0$ but $a \cdot b = 0$. Then

$$\begin{aligned} 0 &= (1/a \cdot 1/b) \cdot 0 && \text{by (10)} \\ &= 0 \cdot (1/a \cdot 1/b) && \text{by } \mathbf{M1} \\ &= (a \cdot b) \cdot (1/a \cdot 1/b) && \text{by hypothesis} \\ &= ((b \cdot a) \cdot 1/a) \cdot 1/b && \text{by } \mathbf{M2} \\ &= (b \cdot (a \cdot 1/a)) \cdot 1/b && \text{by } \mathbf{M2} \\ &= (b \cdot 1) \cdot 1/b && \text{by } \mathbf{M4} \\ &= b \cdot (1/b) && \text{by } \mathbf{M3} \\ &= 1 && \text{by } \mathbf{M4}, \end{aligned}$$

This contradicts Axiom **Z**. The second assertion has been proved along the way. □

We now have established that, according to the axioms we have set up, arithmetic behaves as we expect it to. We shall henceforth not spell out uses of the axioms in such detail, and are ready to revert to more familiar notation.

Notation Henceforth we shall adopt the customary notational shortcuts. We write

$a - b$	in place of	$a + (-b)$,
ab	in place of	$a \cdot b$,
a/b	in place of	$a \cdot (1/b)$,
a^{-1}	as alternative notation for	$1/a$.

We may, thanks to **A2** and **M2**, omit brackets in iterated sums and products, and write for example $a + b + c$ without ambiguity.

1.3. **Powers.**

Let $a \in \mathbb{R} \setminus \{0\}$. As usual, we define $a^0 = 1$. We then define

$$\begin{cases} a^{k+1} = a^k \cdot a & \text{for } k = 0, 1, 2, \dots \text{ (an `inductive' or `recursive' definition)}, \\ a^{\ell} = 1/(a^{-\ell}) & \text{for } \ell = -1, -2, \dots . \end{cases}$$

Then in particular $a^1 = a$ and a^2 serves (as expected) as shorthand for $a \cdot a$. Problem sheet 1, Q. 3 asks for a proof of the familiar **law of indices**.

1.4. **Stocktaking so far, and looking further afield: $\mathbb{R}$ compared to other systems.**

Substitute a set F in place of $\mathbb{R}$ in the above, and assume that the operations $+$ and $\cdot$, now defined on F, satisfy the axioms

A1–A4,
M1–M4,
Z,
D.

Then F, or more precisely $(F; +, \cdot)$, is a **field**. So our assumptions about $\mathbb{R}$ so far can be summed up as

$(\mathbb{R}; +, \cdot)$ **is a field.**

The number systems $\mathbb{Q}$ (rational numbers) and $\mathbb{C}$ (complex numbers) are also fields. But $\mathbb{N}$ (the set of natural numbers, $0, 1, 2, \dots$) and $\mathbb{Z}$ (the integers) have weaker arithmetic properties and are not fields. [*In the Prelims Linear Algebra courses, in the definition of a vector space, the scalars are assumed to be drawn from any field.*]

---

2. The Real Number System: Order Axioms

[BS, Section 2.2]

2.1. **Positive numbers; order relations; the real numbers as a 'number line'.**

There is a subset $\mathbb{P}$ (the **(strictly) positive numbers**) of $\mathbb{R}$ such that, for $a, b \in \mathbb{R}$,

P1 $a, b \in \mathbb{P} \implies a + b \in \mathbb{P}$;
P2 $a, b \in \mathbb{P} \implies a \cdot b \in \mathbb{P}$;
P3 exactly one of $a \in \mathbb{P}$, $a = 0$ and $-a \in \mathbb{P}$ holds.

We write $a < b$ (or $b > a$) iff $b - a \in \mathbb{P}$ and $a \leqslant b$ (or $b \geqslant a$) iff $b - a \in \mathbb{P} \cup \{0\}$ (the **non-negative numbers**).

We subsequently make use of $\leqslant$ or of $<$, as convenient.

2.2. **Properties of the order on $\mathbb{R}$.**

The following properties justify our visualising $\mathbb{R}$ as a 'number line'.

Reflexivity: $a \leqslant a$.

Proof. $a - a = 0 \in \mathbb{P} \cup \{0\}$, by **A4**.

Antisymmetry: $a \leqslant b$ and $b \leqslant a$ together imply $a = b$.

Proof. If $a - b = 0$ or $b - a = 0$, then $a = b$ by properties of addition. Otherwise $\mathbb{P} \ni a - b$ and $\mathbb{P} \ni b - a = -(a - b)$ (by properties of addition). Now apply **P3**.

Transitivity: Assume $a \leqslant b$ and $b \leqslant c$. Then $a \leqslant c$, and likewise with $<$ in place of $\leqslant$.

Proof. We have $c-a = c+(-a) = c+0+(-a) = c+(-b)+b+(-a) = (c-b)+(b-a)$ by properties of addition. So the result for $<$ follows from the definition of $<$ and **P3**.

The proof for $\leqslant$ is the same except for the need to allow also for the trivial cases $a = b$ and/or $b = c$.

Trichotomy: Exactly one of $a < b$, $a = b$ and $b < a$ holds.

Proof. This follows from **P3** and the definition of $<$.

2.3. **Interaction of order and inequalities with arithmetic.**

[*The arithmetic operations interact as expected with inequalities—indeed the axioms for order are set up to ensure this.*]

The following statements hold.

(1) $0 < 1$ (equivalently, $1 \in \mathbb{P}$),
(2) $a < b$ if and only if $-b < -a$. In particular $a > 0$ iff $-a < 0$.
(3) $a < b$ and $c \in \mathbb{R}$ implies $a + c < b + c$.
(4) $a < b$ and $0 < c$ implies $ac < bc$.
(5) $a^2 \geqslant 0$, with equality iff $a = 0$.
(6) $a > 0$ iff $1/a > 0$.
(7) If $a, b > 0$ and $a < b$ then $1/b < 1/a$.

Claims (2)–(4) also hold with $\leqslant$ replacing $<$.

Proof. We use freely the properties of arithmetic established in Section 1.

(1) By trichotomy, exactly one of (i) $1 < 0$, (ii) $1 = 0$, (iii) $0 < 1$ holds. Axiom **Z** rules out (ii). Suppose for a contradiction that $1 < 0$. Then $-1 = 0 + (-1) \in \mathbb{P}$. We deduce that $(-1)\cdot(-1) \in \mathbb{P}$ by **P2**. But $1 = (-1)\cdot(-1)$ (by 1.2(12)), so $0 < 1$ and we have a contradiction to trichotomy.
(2) By properties of addition,

$$\begin{aligned} a < b &\iff b - a \in \mathbb{P} \\ &\iff (-a) - (-b) \in \mathbb{P} \\ &\iff (-a) > (-b). \end{aligned}$$

(3) $a + c < b + c$ iff $0 < (b + c) - (a + c) = b - a$ iff $a < b$.
(4) $a < b$ and $c > 0$ implies $bc - ac = (b - a)c > 0$, by **P2**.
(5) Note $a^2 = 0$ iff $a = 0$, by 1.2(13). Assume $a \neq 0$. Then we have $a^2 = a \cdot a = (-a)\cdot(-a)$. Since either $a > 0$ or $-a > 0$ it follows that $a^2 > 0$ by P2.
(6) Assume for contradiction $a > 0$ but $1/a < 0$. Then $-1 = -(a \cdot (1/a)) = a \cdot (-1/a) > 0$ by **P2** which is impossible by (1). Likewise we obtain a contradiction if $a < 0$ and $1/a > 0$.
(7) Use (4) and (6). □

2.4. **A useful result: Bernoulli's Inequality.**

Let x be a real number with $x > -1$ and let n be a positive integer. Then

$$(1 + x)^n \geqslant 1 + nx.$$

[*The Mean Value Theorem* (*from Analysis* II) *allows one to extend the result by replacing* n *by any real number* $\geqslant 1$.]

Proof. We shall prove the inequality by induction—note that the inequality is trivially true when $n = 1$.

Suppose that, for $k \in \mathbb{N}$,
$$(1+x)^k \geqslant 1 + kx$$
holds for all real $x > -1$. Then $1 + x > 0$ by 2.3(3) and $kx^2 \geqslant 0$ as $k > 0$ and $x^2 \geqslant 0$ by 2.3(5).

$$\begin{aligned}
(1+x)^{k+1} &= (1+x)\,(1+x)^k && \text{by definition}\\
&\geqslant (1+x)\,(1+kx) && \text{by hypothesis and 2.3(4) ($\leqslant$ version)}\\
&= 1 + (k+1)\,x + kx^2 && \text{by } \mathbf{A1–A4}\\
&\geqslant 1 + (k+1)\,x && \text{by 2.3.}
\end{aligned}$$

Hence the result follows by induction. □

2.5. **Modulus (real case).** [BS 2.2]

The **modulus** $|a|$ of $a \in \mathbb{R}$ is defined by
$$|a| = \begin{cases} a & \text{if } a > 0,\\ 0 & \text{if } a = 0,\\ -a & \text{if } a < 0. \end{cases}$$

The definition makes sense by **P3**.

Basic facts about modulus: for any a, b, c,

(1) $|-a| = |a|$;
(2) $|a| \geqslant 0$;
(3) $|a|^2 = a^2$;
(4) $|ab| = |a||b|$;
(5) $-|a| \leqslant a \leqslant |a|$;
(6) if $c \geqslant 0$, then $|a| \leqslant c$ iff $-c \leqslant a \leqslant c$. if $c > 0$, then $|a| < c$ iff $-c < a < c$.

Proof. (1) and (2) are immediate from the definition and the fact that $a > 0$ iff $-a < 0$.

We can prove (3) and also (4), by using **P3** to enumerate cases; recall too that $(-a)(-a) = a^2$ for any a.

For (5), note that
$$\begin{cases} -|a| \leqslant 0 \leqslant a = |a| & \text{if } a \geqslant 0,\\ -|a| = a < 0 \leqslant |a| & \text{if } a < 0. \end{cases}$$

Now consider (6). Assume first that $|a| \leqslant c$, Then, by (5) and transitivity of $\leqslant$, we get $-c \leqslant a \leqslant c$. Conversely, assume $-c \leqslant a \leqslant c$. Then $-a \leqslant c$ and $a \leqslant c$. Since $|a|$ equals either a or $-a$, we obtain $|a| \leqslant c$.

The case with $<$ in place of $\leqslant$ is handled similarly. □

2.6. **The Triangle Law and the Reverse Triangle Law.** [BS 2.2.3 and 2.2.4] Thw Triangle Law is also known as the **Triangle Inequality**.

(1) Let $a, b \in \mathbb{R}$. Then
$$|a+b| \leqslant |a| + |b|.$$
(2) Let $a, b \in \mathbb{R}$. Then
$$|a+b| \geqslant ||a| - |b||\,.$$

Proof. (1): We have

$$-|a| \leqslant a \leqslant |a| \quad \text{and} \quad -|b| \leqslant b \leqslant |b|.$$

Adding (see Problem sheet 1, Q2(a)) and using properties of addition we get

$$-(|a| + |b|) \leqslant a + b \leqslant (|a| + |b|).$$

Now use 2.5(6).

(2): By the Triangle Law,

$$|a| = |a + b + (-b)| \leqslant |a + b| + |(-b)| = |a + b| + |b|,$$

so $|a|-|b| \leqslant |a+b|$, and likewise, reversing the roles of a and b, we get $|b|-|a| \leqslant |b+a| = |a+b|$. Now use the fact that $|c|$ is either c or $-c$ always, and apply this with $c = |a| - |b|$. □

---

3. The Complex Numbers, Briefly

Properties of the complex numbers are covered in the course Introduction to Complex Numbers and not in Analysis I. But we shall deal with complex numbers occasionally, and it is useful to record what does, and does not, hold in relation to arithmetic and inequalities.

We have noted earlier that

- $(\mathbb{C}; +, \cdot)$ is a field.

Here addition and multiplication are defined in the usual way, in terms of real and imaginary parts, or in the case of multiplication, alternatively via polar representation. The axioms **A1–A4** follow from the corresponding axioms for $\mathbb{R}$. The multiplication axioms **M1–M4** are most easily verified using polar coordinates. Axiom **Z** holds since $0, 1$ are real. Axiom **D** holds by a straightforward, but tedious, calculation.

We highlight what does **not** transfer from $\mathbb{R}$ to $\mathbb{C}$. The key point to note is that, unlike $\mathbb{R}$, the complex numbers do not carry a total order relation: we cannot define $<$ on $\mathbb{C}$ in a way which is compatible with arithmetic and such that, for any $w, z \in \mathbb{C}$, exactly one of $w < z$, $w = z$ or $z < w$ holds. **Exercise:** Prove this by considering $\mathrm{i} \cdot \mathrm{i}$. Therefore

inequalities are off limits unless the quantities being compared are real.

Thus, for $z = 3 + 4\mathrm{i}$ and $w = 4 + \mathrm{i}$ we may correctly say that $\operatorname{Re} z < \operatorname{Re} w$, that $\operatorname{Im} z > \operatorname{Im} w$, and that $|z| = 5 > \sqrt{17} = |w|$.

Both the Triangle Law and the Reverse Triangle Law extend to complex numbers (note that the modulus of a complex number is real): for $z, w \in \mathbb{C}$,

$$|z + w| \leqslant |z| + |w| \quad \text{and} \quad |z + w| \geqslant ||z| - |w||.$$

[Proofs: Introduction to Complex Numbers notes, Proposition 10.]

---

4. The Completeness Axiom for the Real Numbers

This section focuses on the order structure of $\mathbb{R}$ and exploits its interaction with the arithmetic structure. In 4.3 we introduce the last of our axioms for $\mathbb{R}$, the **Completeness Axiom**. Throughout this section, and beyond, we shall make use of the axioms and results in Sections 1 and 2, without spelling out the details.

On the basis of our axioms for arithmetic and order, we can't distinguish between $\mathbb{Q}$ and $\mathbb{R}$. But you would claim they are different:

- in $\mathbb{Q}$ there is not a square root for 2 (standard proof, not included in lecture); but
- (you almost certainly believe that) there is a number $\sqrt{2}$ in $\mathbb{R}$.

It will follow from the Completeness Axiom that 2 does have a square root in $\mathbb{R}$ (Theorem 4.10). Hence this axiom does distinguish $\mathbb{R}$ from $\mathbb{Q}$ (see 4.11).

Before presenting the Completeness Axiom we need some order-theoretic preliminaries.

4.1. **Upper and lower bounds.**

Definitions: Let $S \subseteq \mathbb{R}$ and $b \in \mathbb{R}$. Then

b is an **upper bound** of S if $s \leqslant b$ for all $s \in S$;

b is a **lower bound** of S if $b \leqslant s$ for all $s \in S$;

S is **bounded above** if it has an upper bound;

S is **bounded below** if it has a lower bound;

S is **bounded** if it is bounded above and below.

Examples:

(a) $\mathbb{R}$ is not bounded above.
(b) Every element of $\mathbb{R}$ is an upper bound for the empty set.
(c) $\{\, s \in \mathbb{R} \mid -4 \leqslant s < 3 \,\}$: the numbers $3, 11, 10^{37}$ are upper bounds; 2.999999 is not an upper bound. The set of lower bounds is $(-\infty, -4]$.
(d) $\{\, 1/n \mid n = 1, 2, \dots \,\}$: b is an upper bound for S iff $b \geqslant 1$ and b is a lower bound iff $b \leqslant 0$ (but can you justify this?).
(e) $\{1\}$: any $b \geqslant 1$ is an upper bound and any $b \leqslant 1$ is a lower bound.
(f) $\mathbb{Q}$: no upper bounds and no lower bounds.
(g) Let $S := \{\, x \in \mathbb{Q} \mid x^2 < 2 \,\}$. Then certainly $S \neq \emptyset$ and, for example, 3 is an upper bound of S (why?).

Note that we can't yet in every case here convincingly identify all the bounds.

4.2. **Supremum (= least upper bound).**

Let $S \subseteq \mathbb{R}$. Then α is the **supremum** of S, denoted $\sup S$, if

(sup1) $s \leqslant \alpha$ for all $s \in S$ [α is *an* upper bound of S]

(sup2) $s \leqslant b$ for all $s \in S$ implies $\alpha \leqslant b$ [α is the *least* upper bound of S]

Note: $\sup S$ is unique if it exists. Why this definition? *We'll go straight to the Completeness Axiom, then provide examples and commentary.*

4.3. **The Completeness Axiom for the Real Numbers.**

Let S be a **non-empty** subset of $\mathbb{R}$ which is **bounded above**. Then $\sup S$ exists.

Note the necessity for the exclusions here. The empty set has no supremum because it has no *least* upper bound ((sup2) fails). A set which is not bounded above cannot have a supremum because it has *no* upper bound ((sup1) must fail).

4.4. **Supremum: examples.**

(a) Consider $S = [1, 2)$. Then 2 is an upper bound. And if b is any upper bound for S then $b \geqslant 2$: otherwise, $b < 2$ and $b < s := (2+b)/2$ and $s \in S$. So 2 is the least upper bound of S. Note that $\sup S \notin S$.
(b) Consider $S = (1, 2]$. Here $\sup S = 2$. Note that $\sup S \in S$.

4.5. **Infimum (= greatest lower bound).**

Analogous definitions apply here, by replacing $\leqslant$ by $\geqslant$ in the definitions above. Let S be a subset of $\mathbb{R}$ and $\alpha \in \mathbb{R}$. Then α is the infimum of S, written $\inf S$, if

(inf1) $\alpha \leqslant s$ for all $s \in S$ [α is *a* lower bound for S]
(inf2) if $b \leqslant s$ for all $s \in S$ then $b \leqslant \alpha$ [α is the *greatest* lower bound of S]

4.6. **Two theoretical examples.**

(a) Let $\emptyset \neq S \subseteq T$ and assume that T is bounded above. We claim that S is bounded above and that $\sup S \leqslant \sup T$.

Proof. Let b be such that $t \leqslant b$ for all $t \in T$. Then $s \leqslant b$ for all $s \in S$, so S is indeed bounded above. By the Completeness Axiom, $\sup S$ and $\sup T$ exist. Then $\sup T$ is an upper bound for S and hence $\sup S \leqslant \sup T$. □

(b) Let T be non-empty and bounded below. Define $S := \{-t \mid t \in T\}$. Then S is non-empty and bounded above and $\inf T$ exists and equals $-\sup S$.

Proof. If b is any lower bound for T then $-b$ is an upper bound for S, so S is bounded above. Also S is nonempty because T is nonempty. So by the completeness axiom $\sup S$ exists, and $-b \geqslant \sup S$ since $\sup S$ is the *least* upper bound for S. So $b \leqslant -\sup S$ for every lower bound b for T.

Furthermore, if $t \in T$ then $-t \in S$ and so $-t \leqslant \sup S$ and therefore $t \geqslant -\sup S$. Thus $-\sup S$ is a lower bound for T, and we have already seen that $-\sup S$ is greater than or equal to every lower bound for T. Hence $\inf T$ exists and equals $-\sup S$. □

[*This, together with a corresponding argument with sups and infs interchanged, implies that we could equivalently have formulated the Completeness Axiom as 'Every non-empty subset of* $\mathbb{R}$ *which is bounded below has an infimum'.*]

4.7. **Maximum (greatest element) and minimum (least element).**

Assume $\emptyset \neq S \subseteq \mathbb{R}$ and let $s_0 \in \mathbb{R}$. We say s_0 is the **maximum** of S, and write $s_0 = \max S$, if

(max1) $s_0 \in S$ [s_0 belongs to S]
(max2) $s \leqslant s_0$ for all $s \in S$ [s_0 is an upper bound of S]

If a set S is empty or is not bounded above then $\max S$ cannot exist.

Example (sup and max compared): Let $\emptyset \neq S \subseteq \mathbb{R}$ be such that $\sup S$ exists. Then S has a maximum iff $\sup S \in S$ and then $\sup S = \max S$.

Similarly we say a non-empty set S which is bounded below has a **minimum**, $\min S$, if there exists $s_0 \in S$ such that $s_0 \leqslant s$ for all $s \in S$.

We now want to explore the notion of supremum more closely.

4.8. **Modulus measures distance.**

Given $a, x \in \mathbb{R}$ we may interpret $|x - a|$ as the distance from x to a.

Important fact: For $b > 0$, and $a, x \in \mathbb{R}$,

$$|x - a| < b \iff a - b < x < a + b.$$

Proof. 2.5(6) gives $|x - a| < b$ iff $-b < (x - a) < b$ and this holds iff $a - b < x < a + b$. □

So by considering whether this holds for different values of b we can assess how good an approximation x is to a.

4.9. **Approximation Property: capturing supremum via approximation.**

Let S be non-empty and bounded above (so $\sup S$ exists). Then, given $\varepsilon > 0$, there exists s_ε (in general depending on ε) in S such that

$$\sup S - \varepsilon < s_\varepsilon \leqslant \sup S.$$

4.10. **Theorem (existence of $\sqrt{2}$).**

There exists a unique positive real number α such that $\alpha^2 = 2$.

Proof. **Strategy**: let

$$S := \{\, s \in \mathbb{R} \mid s > 0 \text{ and } s^2 < 2 \,\}.$$

Then we seek to show that S is non-empty and bounded above, so that, by the Completeness Axiom, $\alpha := \sup S$ exists. To show $\alpha^2 = 2$, it will suffice (by trichotomy) to show that each of (a) $\alpha^2 > 2$ and (b) $\alpha^2 < 2$ leads to a contradiction. We shall use the Approximation Property to handle (b).

Step 1: $S \neq \emptyset$ and S is bounded above, by 2.

To prove this, first note that $1 \in S$. Now let $x > 2$, so, by order properties,

$$x^2 = x \cdot x > 4 > 2,$$

which means $x \notin S$. Hence $2 \geqslant s$ for all $s \in S$.

Step 2: Assume for contradiction that $\alpha^2 < 2$. Note also that $\alpha \geqslant 1 > 0$ since $1 \in S$. Let $h > 0$. Then

$$\begin{aligned}(\alpha + h)^2 - 2 &= \alpha^2 + 2\alpha h + h^2 - 2 \\ &< \alpha^2 - 2 + 3\alpha h && \text{if } h < \alpha \\ &< 0 && \text{if } h < (2 - \alpha^2)/(3\alpha).\end{aligned}$$

Taking h such that $0 < h < \min\left(\alpha, \dfrac{2 - \alpha^2}{3\alpha}\right)$ we find that $\alpha + h \in S$, contradicting the assumption that α is the least upper bound.

Step 3: Assume for contradiction that $\alpha^2 > 2$. Let $h > 0$. Then $\alpha - h < \alpha$. By the Approximation Property we can find $s \in S$ such that $\alpha - h < s$. But then

$$\alpha^2 - 2\alpha h < \alpha^2 - 2\alpha h + h^2 < s^2 < 2.$$

Take $h \leqslant (\alpha^2 - 2)/(2\alpha)$ to get the required contradiction.

Finally we must show that α is unique. If also $\beta > 0$ and $\beta^2 = 2$ then

$$0 = \alpha^2 - \beta^2 = (\alpha - \beta)(\alpha + \beta).$$

Since $\alpha, \beta > 0$ this contradicts properties of order. □

4.11. **Incompleteness of $\mathbb{Q}$.** The set $\mathbb{Q}$ of rationals does not satisfy the Completeness Axiom with respect to the order it inherits from $\mathbb{R}$. If it did,

$$T := \{ q \in \mathbb{Q} \mid q > 0 \text{ and } q^2 < 2 \}$$

would have a supremum in $\mathbb{Q}$. The proof in 4.10 works just as well for T as it does for S. But we know there is no rational square root of 2.

4.12. **Theorem (existence of nth roots).** Any positive real number has a real nth root, for any $n = 2, 3, 4, \ldots$.

[Case of cube root of 2 is an exercise on Problem sheet 2.]

4.13. **Theorem (the Archimedean Property of the Natural Numbers).**

(i) $\mathbb{N}$ is not bounded above.
(ii) Let $\varepsilon > 0$. Then there exists $n \in \mathbb{N}$ such that $0 < \dfrac{1}{n} < \varepsilon$.

Proof. (i) Assume for a contradiction that $\mathbb{N}$ is bounded above. Then $\sup N$ exists by the Completeness Axiom. By the Approximation Property with $\varepsilon = 1/2$, there exists $k \in \mathbb{N}$ with

$$\sup \mathbb{N} - \frac{1}{2} < k \leqslant \sup \mathbb{N}.$$

But then $k + 1 \in \mathbb{N}$ and $k + 1 > \sup \mathbb{N} + \dfrac{1}{2}$, a contradiction.

For (ii), exploit the fact that $1/\varepsilon$ cannot be an upper bound for $\mathbb{N}$. □

4.14. **Theorem (compare with the well-ordered property of $\mathbb{N}$).**

(i) Every nonempty subset S of $\mathbb{Z}$ which is bounded below has a minimum (least element).
(ii) Every nonempty subset S of $\mathbb{Z}$ which is bounded above has a maximum (greatest element).

Proof. i) We know that $\inf S$ exists (by applying the completeness axiom to $\{-s : s \in S\}$ as in 4.6(b)). So by the approximation property with $\varepsilon = 1$ there is some $n \in S$ such that

$$\inf S \leqslant n < \inf S + 1.$$

It is enough to show that $\inf S = n$, since then $\inf S \in S$ and so $\inf S = \min S$. Assume for a contradiction that $n \neq \inf S$, so that $n > \inf S$ and hence $n = \inf S + \varepsilon$ where $0 < \varepsilon < 1$. By the approximation property again, there exists some $m \in S$ such that

$$\inf S \leqslant m < \inf S + \varepsilon = n.$$

Since $n > m$ we have $n - m > 0$ and so $n - m \geqslant 1$ because $n - m$ is an integer[1], so $n \geqslant \inf S + 1$, which contradicts our first inequality for n.

The proof of (ii) is similar. □

[1]Strictly speaking we have not proved this property of the integers. It can be proved by observing that a strictly positive integer is a natural number and using induction to deduce that it must be at least 1.

4.15. **Density properties.**

(i) Given $a, b \in \mathbb{R}$ with $a < b$ there exists $x \in \mathbb{Q}$ such that $a < x < b$.
(ii) Given $a, b \in \mathbb{R}$ with $a < b$ there exists $y \in \mathbb{R} \setminus \mathbb{Q}$ such that $a < y < b$.

Proof: Exercise (Problem sheet 2, Q. 6).

4.16. $\mathbb{R}$**: SUMMING UP.**

We assume that $\mathbb{R}$ satisfies the axioms set out on the reference sheet Axioms for the Real Numbers:

- the arithmetic axioms **A1–A4, M1–M4, Z, D** (as discussed in Section 1);
- the order axioms **P1–P3** (as discussed in Section 2);
- the Completeness Axiom (discussed above).

In summary, we assume $\mathbb{R}$ is a **complete ordered field**.

All of real analysis stems from $\mathbb{R}$ being a complete ordered field, and no additional assumptions.

---

5. An aside: countability

The density properties recorded in 4.15 are important, the more so because, as we shall see shortly, $\mathbb{R}$ is 'a bigger set' than $\mathbb{Q}$: the set $\mathbb{Q}$ is countable whereas the set $\mathbb{R}$ is uncountable.

The objective in Analysis I is not to present a crash course in set theory but to give the minimum amount of information on countability necessary to distinguish between countable and uncountable sets in the context of the real numbers. See the supplementary notes on Countability for an informal account of this topic which goes beyond the Analysis I syllabus. These notes may be of interest to those who want to go deeper than Section 5 does and in particular to Maths/Phil students.

5.1. **Comparing sizes of sets.**

Let A and B be sets. We say A and B are **equinumerous** (notation $A \approx B$) if there is a bijection $f\colon A \to B$. Note $\approx$ has the properties of an equivalence relation.

Given sets A and B we shall write $A \preccurlyeq B$ if there is an injection $f\colon A \to B$. Intuitively this says that B is at least as big as A.

5.2. **Finite, countable and uncountable sets.**

Let A be a set. We call A **finite** if either $A = \emptyset$ or there exists $n \in \mathbb{N}$ such that $A \approx \{0, \ldots, n-1\}$ (or equivalently if $A \approx \{1, \ldots, n\}$). A set which is not finite is said to be **infinite**. [*Another way to capture the notion of finiteness is to say that A is finite iff every injective map from A to A is surjective: the Pigeonhole Principle holds for A.*] Any subset of a finite set is finite.

Note that any non-empty finite subset of $\mathbb{R}$ must be bounded above (in fact, it contains a largest element (why?)). Hence any subset of $\mathbb{R}$ which is not bounded above must be infinite. By the Archimedean property, $\mathbb{N}$ is not bounded above, and hence is infinite.

We call a set A

- **countably infinite** if $A \approx \mathbb{N}$;
- **countable** if $A \preccurlyeq \mathbb{N}$;
- **uncountable** if A is not countable.

[**Warning**: Some authors use 'countable' to mean what we call 'countably infinite'.]

FACTS: (see supplementary material on countability)

(1) A is countable (that is, $A \preccurlyeq \mathbb{N}$) iff A is finite or countably infinite.
(2) If $A \preccurlyeq B$ and $B \preccurlyeq A$ then $A \approx B$.

5.3. **Some familiar sets which are countably infinite.**

(a) $\mathbb{N}$;
(b) $\mathbb{N}^{>0} := \mathbb{N} \setminus \{0\}$;
(c) $\{\, 2k+1 \mid k \in \mathbb{N}\,\}$ (the odd natural numbers);
(d) $\mathbb{Z}$;
(e) $\mathbb{N} \times \mathbb{N}$.

[Note that (c) shows that a countably infinite set may be equinumerous with a proper subset of itself.]

Proof. (a) is immediate. For (b): the successor function, $n \mapsto n+1$, is a bijection from $\mathbb{N}$ to $\mathbb{N} \setminus \{0\}$. To prove (c), note that the map $2k+1 \mapsto k$ is injective and maps the given set onto $\mathbb{N}$. Now consider (d). We can define a bijection f from $\mathbb{Z}$ to $\mathbb{N}$ by

$$f(k) = \begin{cases} -2k & \text{if } k \leqslant 0, \\ 2k-1 & \text{if } k > 0. \end{cases}$$

We prove (e) by setting up a bijection $f \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Define f by

$$f((m,n)_= 2^m(2n+1) - 1.$$

Injectivity of f: $2^{m_1}(2n_1+1) = 2^{m_2}(2n_2+1)$ implies (by uniqueness of factorisation in $\mathbb{N}$ (assumed)) that $2^{m_1} = 2^{m_2}$ so that $m_1 = m_2$ (make use of laws of indices for the last step) and $2n_1+1 = 2n_2+1$, whence $n_1 = n_2$. So $(m_1, n_1) = (m_2, n_2)$.

Surjectivity of f: take $k \in \mathbb{N}$. Assume first that $2 \nmid (k+1)$. Then k is even and so $k = 2^0(2n+1) - 1$, for some $n \in \mathbb{N}$. Now assume $2|(k+1)$. Then there exists m such that $2^m \mid (k+1)$ and $2^{m+1} \nmid (k+1)$ (use the fact that $\{\, 2^m \mid m \in \mathbb{N}\,\}$ is not bounded above (Problem sheet 2, Q. 2). Then $k+1 = 2^m(2n+1)$, for some $n \in \mathbb{N}$. □

5.4. **New countable sets from old.**

Let A and B be countable sets. Let $f \colon A \to \mathbb{N}$ and $g \colon \mathbb{B} \to \mathbb{N}$ be injective.

Claim 1: Assume A and B are disjoint. Then $A \cup B$ is countable. [Disjointness not essential, but it simplifies the proof.]

Proof. Define

$$h(x) = \begin{cases} 2f(x) & \text{if } x \in A, \\ 2g(x)+1 & \text{if } x \in B. \end{cases}$$

Then h is an injection from $A \cup B$ to $\mathbb{N}$. □

Claim 2: $A \times B$ is countable.

Proof. Define $h\colon A \times B \to \mathbb{N}$ by

$$h((a,b)) = 2^{f(a)+1}3^{g(b)+1} \quad \text{for } a \in A, b \in B.$$

Then uniqueness of factorisation in $\mathbb{N}^{>0}$ implies h is injective. □

5.5. **Theorem: $\mathbb{Q}$ is countable.**

Proof. We write $\mathbb{Q}$ as the disjoint union

$$\mathbb{Q}^{>0} \cup \{0\} \cup \mathbb{Q}^{<0}.$$

By 5.4, Claim 1, it will be enough to prove that $\mathbb{Q}^{>0}$ (and so, likewise, $\mathbb{Q}^{<0}$) is countable. We can write each element of $\mathbb{Q}^{>0}$ as p/q where $p, q \in \mathbb{N}$, $p > 0$, $q > 0$ and p/q is expressed in its lowest terms. Then $p/q \mapsto 2^p3^q$ is an injection into $\mathbb{N}$. So $\mathbb{Q}^{>0} \preccurlyeq \mathbb{N}$ as claimed. □

5.6. **Theorem: $\mathbb{R}$ is uncountable.**

We shall be able to prove this further on in the course using facts concerning decimal representation of real numbers (see Supplementary note on decimals)—the classic proof due to Cantor. Meantime, an alternative proof, using nested intervals and the Completeness Axiom, can be found in **[BS 2.5.4]**.

z

A note on familiar functions

We want our treatment of sequences, and a bit later, of series not to be divorced from the functions you were introduced to at school and regularly encounter in other courses: trigonometric functions, exponential functions, logarithms, and general powers. Accordingly we'd like to involve such functions in examples and exercises. So, for now, we shall take the existence and the properties of these functions for granted, and use them freely. Later you will see formal definitions of these various functions and rigorous derivations of their (familiar) properties.

When we use logarithms these will always be to base e. We adopt the notation $\log x$ for $\log_{\mathrm{e}} x$, rather than $\ln x$.

Recall that, for $a > 0$ and $x \in \mathbb{R}$, one defines $a^x = \mathrm{e}^{x \log a}$.

---

6. Sequences

[BS, Sections 3.1 and 3.2]

This section covers the rudiments of the theory of convergence of sequences. An ample supply of worked examples is included in these webnotes. Examples omitted from lectures are recommended for self-study.

6.1. **Real numbers via approximations.**

Examples:

(a) $\dfrac{3}{10}, \quad \dfrac{33}{100}, \quad \dfrac{333}{1000}, \quad \ldots, \quad \dfrac{\overbrace{33\ldots3}^{n \text{ digits}}}{10^n}$ are progressively better approximations to $1/3$.

(b) $\dfrac{14}{10}, \quad \dfrac{141}{100}, \quad \dfrac{1414}{1000}, \quad \dfrac{14142}{10000}$ are well-known approximations to $\sqrt{2}$.

(c) [The Archimedean Property revisited] For any $\varepsilon > 0$ there exists $N \geqslant 1$ such that $0 < 1/N < \varepsilon$. And then $0 < 1/n \leqslant 1/N < \varepsilon$ for all $n \geqslant N$. This says that, apart from a finite number of terms at the start, the terms of the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$$

are all within a distance ε of 0. Here ε can be any positive real number whatsoever.

(d) Consider the string of real numbers

$$1, -1, 2, -2, 3, -3, 4, -4, \ldots$$

Your intuition tells you there is no limiting value in this case.

6.2. **The formal definition of a real or complex sequence.**

Officially, a **sequence** of real numbers (a **real sequence**) is an assignment

$$n \mapsto \alpha(n)$$

of a real number $\alpha(n)$ to each $n = 1, 2, \ldots$. Thus a sequence is a function $\alpha\colon \mathbb{N}^{\geqslant 1} \to \mathbb{R}$ (where $\mathbb{N}^{\geqslant 1} = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \ldots\}$) and we call $\alpha(n)$ the nth term of the sequence. (Sometimes later we shall work with sequences with terms labelled by $n = 0, 1, 2, \ldots$ instead.)

We shall usually write a_n in place of $\alpha(n)$ and then say that α defines the sequence

$$(a_n) = (a_1, a_2, a_3, \ldots)$$

also written $(a_n)_{n \geqslant 1}$ — that is, we specify the sequence by its terms. Note that the terms *in order* determine the sequence.

A **complex sequence** (a_n) is defined in the same way, but now with the terms a_n drawn from the complex numbers, $\mathbb{C}$. This section concentrates on convergence of real sequences. We defer discussion of complex sequences until 6.21, but note that most of the results we obtain for real sequences do have analogues for complex sequences. The exceptions will be those whose statement or proof relies on inequalities involving the terms; recall the comments in Section 3 about complex numbers and order. We adopt the convention that only in statements of results applicable *only* to real sequences do we include the word 'real'. Thus 'sequence' on its own can be read as 'real or complex sequence' once 6.21 has been studied.

Examples

(a) Let $a_n = \alpha(n)$, where $\alpha(n) = \sin n/(2n + 1)$. Then the sequence looks like

$$(\frac{1}{3}\sin 1, \ \frac{1}{5}\sin 2, \ \frac{1}{7}\sin 3, \ldots).$$

(b) Let $a_n = \alpha(n)$, where $\alpha(n) = (-1)^n$. Then the sequence looks like

$$(-1, 1, -1, 1, -1, 1, -1, \ldots).$$

6.3. **Manufacturing new sequences.** We can form new sequences from given ones 'a term at a time': given sequences (a_n) and (b_n), we have sequences $(a_n + b_n)$, $(-a_n)$, $(a_n b_n)$ and, provided every b_n is non-zero, (a_n/b_n). Also we can form (ca_n), for any constant c, and $(|a_n|)$.

Examples: let $a_n = (-1)^n$ and $b_n = 1$ for all n. Then

$$\begin{aligned}(a_n + b_n) &= (0, 2, 0, 2, 0, 2, \ldots);\\ (-a_n) &= ((-1)^{n+1});\\ (|a_n|) &= (1, 1, 1, 1, 1, \ldots).\end{aligned}$$

The key notion in this section is that of ***convergence*** *of a sequence (a_n). We want to analyse how the terms a_n of the sequence behave as n gets 'arbitrarily large' and specifically whether or not the terms approach 'arbitrarily closely' some 'limiting value' L. For this we need to convert these informal ideas into precise, formal, ones.*

6.4. **Tails.** As regards the long-run behaviour of the terms of a sequence (a_n) as n gets arbitrarily large we don't care what the values of the first few terms are, or what the first 10 million terms are, ... The notion of a tail of (a_n) will allow us to capture this idea.

Given a sequence (a_n) and any $k \in \mathbb{N}$ we can form a new sequence (b_n) by chopping off the first k terms $a_1, \ldots, a_k$ of (a_n) and relabelling. That is, $b_n = a_{n+k}$ for all n. We call (b_n) a **tail** of (a_n).

6.5. **Capturing 'arbitrarily close to' via ε.** From 2.5(6) (see also 4.8), we have, for $x, a \in \mathbb{R}$ and $b > 0$,

$$|x - a| < b \iff a - b < x < a + b$$

and this condition captures the statement that x lies within a distance b of a. In formulating notions of limits and convergence we are interested in allowing b to be very small—arbitrarily small—and shall use the customary symbol ε to denote a strictly positive real number playing this role.

6.6. **Convergence of a real sequence.** Let (a_n) be a sequence of real numbers and let $L \in \mathbb{R}$. Then we say that (a_n) **converges to** L (notation: $a_n \to L$ (as $n \to \infty$)) if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n > N \quad |a_n - L| < \varepsilon.$$

Here N can, and almost always will, depend on ε. Note that we can replace '$n > N$' by $n \geqslant N$ and/or '$|a_n - L| < \varepsilon$' by $|a_n - L| \leqslant \varepsilon$ in this definition without changing the meaning (WHY?). However it is crucial that ε should be *s*trictly greater than 0.

When $a_n \to L$ we say that L is the **limit** of (a_n) [in 6.13 we'll show it must be unique] and we write

$$L = \lim_{n\to\infty} a_n \quad \text{or just} \quad L = \lim a_n.$$

We say (a_n) **converges** if there exists $L \in \mathbb{R}$ such that $a_n \to L$ as $n \to \infty$. We say (a_n) **diverges** if it does not converge.

The following technical, but intuitively plausible, result reinforces the notion that convergence, or not, depends only on the long-run behaviour of a sequence.

6.7. **Tails Lemma.** Let (a_n) be a sequence.

(i) If (a_n) converges to a limit L then each tail of (a_n) converges, to the same limit L.
(ii) Assume some tail $(b_n) = (a_{n+k})$ converges. Then (a_n) converges.

Proof. (i) Let $k \in \mathbb{N}$ and let $b_n = a_{n+k}$ for $n \geqslant 1$. Assume (a_n) converges to L. Then

$$\forall \varepsilon > 0 \quad \exists N \quad \forall n \geqslant N \quad |a_n - L| < \varepsilon.$$

This implies $|b_n - L| = |a_{n+k} - L| < \varepsilon$ for all $n \geqslant N$ because then $n + k \geqslant n \geqslant N$.

(ii) Since (b_n) converges there exists L such that

$$\forall \varepsilon > 0 \quad \exists N' \quad \forall p \geqslant N' \quad |b_p - L| < \varepsilon.$$

That is, $p \geqslant N'$ implies $|a_{p+k} - L| < \varepsilon$. Define $N = N' + k$. Then $n \geqslant N$ implies that $n = p + k$ where $p \geqslant N'$ and hence $|a_n - L| < \varepsilon$. So (a_n) converges to L too. □

6.8. **Examples: convergence established directly (but messily) from the definition.**

(a) The basic fact that

$$1/n \to 0 \text{ as } n \to \infty$$

is simply the Archimedean Property in new clothes. To see this note that the Archimedean Property is the statement that, given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then for $n \geqslant N$ we have $|1/n - 0| = 1/n \leqslant 1/N < \varepsilon$.

See 6.20 for a discussion of the limiting behaviour of arbitrary powers of n.

(b) Let $a_n = 1 + (-1)^n \dfrac{1}{\sqrt{n}}$. We claim $a_n \to L$ where $L = 1$. Let $\varepsilon > 0$. Then

$$|a_n - 1| = \left|(-1)^n \frac{1}{\sqrt{n}}\right| < \varepsilon \Longleftarrow \frac{1}{\sqrt{n}} < \varepsilon$$
$$\Longleftarrow n \geqslant N, \text{ where we choose } N \in \mathbb{N} \text{ with } N > \frac{1}{\varepsilon^2}.$$

Here , and likewise below, the first $\Longleftarrow$ may be read as ‘if’ and any subsequent ones as ‘and hence if’.

(c) Let $a_n = \dfrac{1}{n^2 - n + 1}$. Let $\varepsilon > 0$. Then

$$|a_n - 0| = \frac{1}{n^2 - n + 1} < \varepsilon \Longleftarrow n^2 - n + 1 > \varepsilon^{-1}$$
$$\Longleftarrow (n - \tfrac{1}{2})^2 > \varepsilon^{-1} - \tfrac{3}{4}$$
$$\Longleftarrow n > N, \text{ where } N \in \mathbb{N} \text{ and } N \geqslant \tfrac{1}{2} + \sqrt{\varepsilon^{-1} - \tfrac{3}{4}},$$

where we may assume without loss of generality (see 6.9(4)) that $\varepsilon < 4/3$. So $a_n \to 0$.

(d) Let $a_n = \dfrac{n \sin n^2}{3n^3 - n - 1}$. Let $\varepsilon > 0$. Then

$$|a_n - 0| = \frac{n|\sin n^2|}{3n^3 - n - 1} < \varepsilon \Longleftarrow \frac{n}{3n^3 - n - 1} < \varepsilon \qquad \text{since } |\sin x| \leqslant 1 \text{ for all } x$$
$$\Longleftarrow n/n^3 < \varepsilon \qquad \text{since } n^3 \geqslant n \text{ and } n^3 \geqslant 1$$
$$\Longleftarrow n \geqslant N \text{ where } N > 1/\sqrt{\varepsilon}.$$

So $a_n \to 0$. Note the need for care in handling inequalities in this example.

6.9. **Testing a sequence for convergence: remarks and technical tips.** Consider the situation in which we have a sequence (a_n) and a candidate limit L, and we wish to show that $a_n \to L$, that is, that it satisfies the ε-N condition in 6.6.

(1) We require N such that

$$n \geqslant N \Longrightarrow |a_n - L| < \varepsilon.$$

We do NOT need

$$n \geqslant N \Longleftrightarrow |a_n - L| < \varepsilon.$$

This means we do not need to find the **smallest** N **possible** when establishing convergence. *Any N that works will do.* This allows us in many cases to simplify calculations by replacing complicated expressions by simpler ones before trying to write down a suitable N.

(2) Beginners often give back-to-front arguments when seeking to prove that a sequence (a_n) tends to some limit L. Note the direction of the implication signs in Examples 6.8. Reading downwards, we are working towards finding a suitable N. Once such an N has been identified, the argument can be re-presented, going from bottom to top and using forward implication signs. Let's carry this out for the sequence (a_n) in 6.8(b), where $a_n = ((-1)^n/\sqrt{n})$, treating the original presentation as rough work. Take $\varepsilon > 0$. Choose $N \in \mathbb{N}$ with $N > 1/\varepsilon^2$. Then

$$\begin{aligned} n \geqslant N &\Longrightarrow n > \frac{1}{\varepsilon^2} \\ &\Longrightarrow \frac{1}{n} < \varepsilon^2 \\ &\Longrightarrow \left| (-1)^n \frac{1}{\sqrt{n}} \right| < \varepsilon \\ &\Longrightarrow |a_n - 1| < \varepsilon. \end{aligned}$$

In the examples in 6.8 most instances of $\Longleftarrow$ could be replaced by $\Longleftrightarrow$, but not all can be (see (d)); (1) says that none needs to be.

(3) We have asked that $N \in \mathbb{N}$. But it's good enough to find $X \in \mathbb{R}$ such that $n \geqslant X$ implies $|a_n - L| < \varepsilon$. If X exists then we can choose $N \geqslant X$ and $N \in \mathbb{N}$ (since $\mathbb{N}$ is not bounded above, as we proved in 4.13(i)).

(4) The smaller ε is, the greater challenge we have to find a corresponding N in general. Turning this around, we see that in establishing convergence we may without loss of generality restrict to values of ε such that $\varepsilon < 1$ (or $\varepsilon < \eta$, where η is some fixed positive number). We did in this in 6.8(c)). It's **small** values of ε that matter.

(5) Facility with inequalities is a valuable skill!

It is already clear from Examples 6.8(c),(d) that finding an explicit N (or X) for a given $\varepsilon > 0$ can be tiresome and messy. The following result is elementary. It is useful in two ways: it allows us

- to simplify ε-N proofs;
- to take advantage of known limits to find the limiting values of other sequences.

6.10. **Sandwiching Lemma (simple form).** Let (b_n) and (c_n) be real sequences. Assume $c_n \to 0$ and that $0 \leqslant b_n \leqslant c_n$ for all n. Then $b_n \to 0$.

Proof. Let $\varepsilon > 0$ and pick N so that $|c_n - 0| < \varepsilon$ for all $n \geqslant N$. Then, for $n \geqslant N$,

$$-\varepsilon < 0 \leqslant b_n \leqslant c_n = |c_n| < \varepsilon.$$

and so $|b_n - 0| < \varepsilon$ for all $n \geqslant N$. □

[*What this says is that, given* ε, *an* N *that works for* (c_n) *also works for* (b_n).]

6.11. **Examples, by sandwiching.**

(a) [Example 6.8(c) revisited] Note that $(n-1)^2 \geqslant 0$, so that $n^2 - n + 1 \geqslant n$. Hence

$$0 < \frac{1}{n^2 - n + 1} \leqslant \frac{1}{n} \to 0 \text{ as } n \to \infty.$$

Now apply the Sandwiching Lemma.

(b) $0 \leqslant 2^{-n} \leqslant 1/n$ for all n (by induction) and $1/n \to 0$. Hence (2^{-n}) converges to 0 by the Sandwiching Lemma.

(c) Let $a_n = \dfrac{n}{\sqrt{n^2+1}}$. Then $a_n \to 1$. To prove this, note that

$$0 \leqslant |a_n - 1| = \frac{\sqrt{n^2+1} - n}{\sqrt{n^2+1}} \leqslant \frac{\sqrt{n^2+2n+1} - n}{\sqrt{n^2+1}} = \frac{(n+1) - n}{\sqrt{n^2+1}} = \frac{1}{\sqrt{n^2+1}} < \frac{1}{n}.$$

Now apply the Sandwiching Lemma, once again using the fact that $1/n \to 0$.

(d) [Example 6.8(d) revisited] The proof given earlier is a sandwiching argument from scratch. The Sandwiching Lemma can be applied with $b_n = |a_n|$ and $c_n = 1/n$.

6.12. **More examples** (two important limits, employing some useful techniques)**.**

(a) Let $|c| < 1$, where c is constant. We claim $c^n \to 0$. We can write $|c| = 1/(1+y)$ where $y > 0$. Take $\varepsilon > 0$. By Bernoulli's inequality,

$$|c|^n = \frac{1}{(1+y)^n} \leqslant \frac{1}{1+ny} < \frac{1}{ny} < \varepsilon \quad \text{if } n \geqslant N,$$

where $N \in \mathbb{N}$ is chosen such that $N \geqslant 1/(y\varepsilon)$.

(b) Let $a_n = \dfrac{n}{2^n}$. We surmise that $a_n \to 0$. Let $\varepsilon > 0$. Now

$$\begin{aligned} |a_n - 0| = \frac{n}{2^n} = \frac{n}{(1+1)^n} &= \frac{n}{1 + n + \binom{n}{2} + \cdots + 1} && \text{(by the binomial theorem)} \\ &\leqslant \frac{2n}{n(n-1)} && \text{(if } n \geqslant 2\text{, by retaining the } \tbinom{n}{2} \text{ term)} \\ &< \varepsilon && \text{provided } n - 1 > 2/\varepsilon. \end{aligned}$$

So we choose $N \in \mathbb{N}$ so that $N > 1 + 2/\varepsilon$.

Now we give a significant theoretical result. The proof gives an illustration of working with the ε-N *definition of convergence.*

6.13. **Theorem (uniqueness of limits).** Let (a_n) be a sequence and suppose that $a_n \to L_1$ and $a_n \to L_2$ as $n \to \infty$. Then $L_1 = L_2$.

Proof. Suppose $L_1 \neq L_2$. Take $\varepsilon := |L_1 - L_2|$. Then $\varepsilon > 0$. So $\varepsilon/2 > 0$ and hence there exist N_1 and N_2 such that

$$n \geqslant N_1 \Longrightarrow |a_n - L_1| < \varepsilon/2,$$
$$n \geqslant N_2 \Longrightarrow |a_n - L_2| < \varepsilon/2.$$

Then for $n \geqslant \max(N_1, N_2)$ we have

$$\begin{aligned} |L_1 - L_2| &= |(L_1 - a_n) + (a_n - L_2)| \\ &\leqslant |L_1 - a_n| + |a_n - L_2| \qquad \text{(by Triangle Law)} \\ &< \varepsilon/2 + \varepsilon/2 = |L_1 - L_2| \end{aligned}$$

which is the required contradiction. □

The next group of results concerns the interaction of limits with modulus and inequalities.

6.14. **Proposition (limits and modulus).** Assume that (a_n) is a sequence which converges to L. Then $(|a_n|)$ converges too, to $|L|$.

Proof. By the Reverse Triangle Law, 2.6(2),

$$||a_n| - |L|| \leqslant |a_n - L|.$$

Now apply the Sandwiching Lemma, or argue directly from the convergence definition. □

6.15. **Limits and inequalities.** Let (a_n) and (b_n) be **real** sequences.

Preservation of weak inequalities: Assume that $a_n \to L$ and $b_n \to M$ and that $a_n \leqslant b_n$ for all n. Then $L \leqslant M$.

Proof. We argue by contradiction. Assume $L \not\leqslant M$. Then (by trichotomy) $L > M$. Let $\varepsilon := (L - M)/2$. We now can find N_1 and N_2 such that

$$\begin{aligned} \forall n \geqslant N_1 \quad & |a_n - L| < \varepsilon, \\ \forall n \geqslant N_2 \quad & |b_n - M| < \varepsilon. \end{aligned}$$

Then for $n \geqslant \max(N_1, N_2)$,

$$L - \varepsilon < a_n \leqslant b_n < M + \varepsilon.$$

Hence $L - M < 2\varepsilon$, a contradiction. □

An example of non-preservation of strict inequalities. Let $a_n = 1/n$. Then $a_n > 0$ for all n, and (by the Archimedean Property), $a_n \to 0$. So $\lim a_n > 0$ is false.

6.16. **A general Sandwiching Lemma.** Assume that (x_n), (y_n) and (a_n) are real sequences such that $x_n \leqslant a_n \leqslant y_n$ for all n. Assume that $\lim x_n = \lim y_n = L$. Then (a_n) converges to L.

Proof. (Outline) Given $\varepsilon > 0$ we can find N such that for all $n \geqslant N$ we have $|x_n - L| < \varepsilon$ and $|y_n - L| < \varepsilon$. Then, for $n \geqslant N$,

$$L - \varepsilon < x_n \leqslant a_n \leqslant y_n < L + \varepsilon,$$

so $|a_n - L| < \varepsilon$. □

The next result shows us that any convergent sequence has a special property, that of being bounded. The result will be useful in some technical proofs later, and also, in its contrapositive form, provides a way to show that certain sequences fail to converge.

6.17. **Proposition (a convergent sequence is bounded).** Assume (a_n) converges. Then

$$\exists M \in \mathbb{R} \;\; \forall n \;\; |a_n| \leqslant M.$$

(This says that (a_n) is **bounded**, meaning that its set $\{\, a_n \mid n = 1, 2, \dots \}$ of terms is bounded.)

If (a_n) is not bounded then (a_n) diverges.

Proof. Assume $a_n \to L$. Take $\varepsilon = 1$. Then there exists N such that

$$\begin{aligned} n \geqslant N &\Longrightarrow |a_n - L| < 1 \\ &\Longrightarrow |a_n| \leqslant |L| + 1 \qquad \text{(by Triangle Law).} \end{aligned}$$

(This says (a_n) has a bounded tail.) Hence

$$\forall n \;\; |a_n| \leqslant M := \max\{|a_1|, \dots, |a_{N-1}|, |L| + 1\}.$$

The second statement is just the contrapositive. □

6.18. **Examples: divergence.**

(a) **Unboundedness implies divergence** by 6.17. So (2^n) diverges.

(b) **Unboundedness not necessary for divergence**. Let $a_n = (-1)^n$. Let $\varepsilon = 1$ and let L be any real number. Assume there exists N such that $n \geqslant N$ implies $|(-1)^n - L| < 1$. Then

$$n \geqslant N \implies -1 < (-1)^n - L < 1.$$

Taking $n = 2N$ gives $L > 0$ and taking $n = 2N + 1$ gives $L < 0$, so we have a contradiction to the assumption that the sequence converges to L, and this holds for every L. So the sequence must diverge.

This example also shows that $(|a_n|)$ convergent need not imply (a_n) convergent, So Proposition 6.14 does not have a converse.

(b) Let $a_n = (-1)^n \dfrac{n^2}{n^2 + 1}$. Then (a_n) diverges.

The result is plausible, but this is an awkward example to handle slickly from first principles. We'll give a proof later, once we have developed an efficient method.

6.19. **Infinity.** Let a_n be a sequence of real numbers. We say 'a_n tends to infinity' and write $a_n \to \infty$ as $n \to \infty$ if

$$\forall M \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad a_n > M.$$

Similarly we write $b_n \to -\infty$ if

$$\forall M \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad b_n < M.$$

(Here we tend to think of M as being a very large positive/negative number.)

Here the symbol ∞ provides a convenient notational shorthand. Do NOT treat ∞ as though it were a real number. Remember that in our convergence definition in 6.6 we demanded that the limit L belong to $\mathbb{R}$. See further discussion of infinite limits in Section 7.

Examples

(a) Let $a_n = n^2 - 1000$. Then $a_n \to \infty$. To prove this let $M \in \mathbb{R}^{>0}$. Then $a_n > M$ for all $n \geqslant N$ if we choose $N \in \mathbb{N}$ such that $N \geqslant \sqrt{M + 1000}$—possible by 4.13(i).

(b) Let $a_n = \begin{cases} n & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$

If it were true that $a_n \to \infty$, then (taking $M = 1$) we could find N such that $n \geqslant N$ implies $a_n > 1$. But $a_{2N} = 0$ and we have a contradiction.

6.20. **Important examples: limits involving powers.**

(a) **Powers of n.** In this example we assume the familiar properties of the exponential function exp and the natural logarithm log; recall that these functions will be defined and their properties proved later in the year. For α a real number we define n^α to be $\exp(\alpha \log n)$. We have

(i) if $\alpha < 0$, then $n^\alpha \to 0$ as $n \to \infty$;
(ii) if $\alpha > 0$, then $n^\alpha \to \infty$ as $n \to \infty$.

Proof. (i) Write $\beta = -\alpha$ so that $n^\alpha = 1/n^\beta$. Let $\varepsilon > 0$ and assume $\varepsilon < 1$. We have

$$1/n^\beta < \varepsilon \Longleftrightarrow n^\beta > 1/\varepsilon \Longleftrightarrow \beta \log n > -\log \varepsilon \Longleftrightarrow n > \mathrm{e}^{-\log \varepsilon/\beta}.$$

[Here, by properties of exponentials and logs, it's just as easy to get $\Longleftrightarrow$ as to get $\Longleftarrow$.]

(ii) Note that to check that $n^\alpha \to \infty$ it suffices to consider the case when $M > 0$. Then

$$n^\alpha > M \Longleftrightarrow \alpha \log n > \log M \Longleftrightarrow n > \mathrm{e}^{\log M/\alpha}.$$

This gives the result. □

(b) **Powers with exponent n.** Let c be a positive constant.

(i) If $c < 1$ then (c^n) converges to 0.
(ii) If $c = 1$ then (c^n) converges to 1.
(iii) If $c > 1$ then $c_n \to \infty$.

Proof. For (i), see Example 6.12(a); (iii) can be handled in a similar way, while (ii) is immediate. Alternatively consider $\log c^n$. □

6.21. **Complex sequences.** As noted already, the definition of convergence and much of the theory of convergence of real sequences carry over in the obvious way to sequences of complex numbers. In particular we say a sequence (z_n) **converges to** L (where now $L \in \mathbb{C}$) if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad |z_n - L| < \varepsilon.$$

Moreover, the limit is unique if it exists. Uniqueness is proved as in 6.13.

Recall the remarks in Section 3 concerning complex numbers.

- Inequalities: don't try to write $w < z$ when $w, z \in \mathbb{C}$, not both real !
- The Triangle and Reverse Triangle Laws do hold for complex numbers.
- Simple sandwiching is valid in the following form: suppose (w_n) and (z_n) are complex sequences such that $|w_n| \leqslant |z_n|$ and $z_n \to 0$, then $w_n \to 0$. [*Note the moduli !*]

6.22. **Theorem (convergence of complex sequences).** Let (z_n) be a sequence of complex numbers and write $z_n = x_n + \mathrm{i}y_n$, so (x_n) and (y_n) are real sequences. Then (z_n) converges if and only if (x_n) and (y_n) both converge.

Proof. The proof of uniqueness is exactly the same as in the real case.

Now consider the first assertion. For $\Longrightarrow$ use the fact that $|x_n| \leqslant |z_n|$ and $|y_n| \leqslant |z_n|$ and sandwiching. The proof of $\Leftarrow$ is a definition-chase and left as an exercise (ideas from 8.3 are useful). □

6.23. **Examples: complex sequences.**

(a) Let $z_n = \sqrt{n^2+1}/(n+\mathrm{i})$. Then $z_n = \dfrac{n-\mathrm{i}}{\sqrt{n^2+1}}$. The real sequences $(n/\sqrt{n^2+1})$ and $(-1/\sqrt{n^2+1})$ both converge, the former to 1 and the latter to 0. So $z_n \to 1+0\cdot\mathrm{i} = 1$.

(b) Let $z_n = \mathrm{i}^n/n$. Then $|z_n| = 1/n$ and hence $z_n \to 0$.

(c) Let $z_n = \mathrm{i}^n$. Then the sequence is

$$\mathrm{i}, -1, -\mathrm{i}, 1, \mathrm{i}, -1, \dots \quad \text{and} \quad \operatorname{Re} z_n = 0, -1, 0, 1, 0, -1, \dots .$$

Thus $(\operatorname{Re} z_n)$ does not converge and so (z_n) cannot converge.

(d) $(\cos n)$ and $(\sin n)$ diverge (see Problem Sheet 4). More generally it is poossible to show that $(\cos n\theta)$ and $(\sin n\theta)$ diverge for $0 < \theta < 2\pi$. Hence if $z \in \mathbb{C}$ and $|z| = 1$ then (z^n) diverges except when $z = 1$.

---

7. Subsequences

This short section introduces the notion of a subsequence of a (real or complex) sequence. The results we obtain here usefully enlarge our armoury of techniques for establishing convergence/divergence. Deeper results involving subsequences are given in Section 10.

When we say that $(a_n)_{n\geqslant 1}$ is a sequence, or that $a_n \to L$ as $n \to \infty$, the symbol n is a 'dummy', standing for a strictly positive integer. We are allowed to replace n here with any other symbol so long as we are consistent, so we might write $(a_m)_{m\geqslant 1}$ or $(a_k)_{k\geqslant 1}$ or $(a_r)_{r\geqslant 1}$ to mean exactly the same as $(a_n)_{n\geqslant 1}$, while $a_m \to L$ as $m \to \infty$ or $a_k \to L$ as $k \to \infty$ or $a_r \to L$ as $r \to \infty$ mean exactly the same as $a_n \to L$ as $n \to \infty$. On the other hand it is worth remembering that it makes maths much easier to read when we use the traditional symbol n or letters close to (or similar to) n in the alphabet in these circumstances, especially if we want to use the shorthand (a_n) or $a_n \to L$, dropping the subscript $n \geqslant 1$ or the phrase 'as $n \to \infty$'.

7.1. **Subsequences.** Let $(a_n)_{n\geqslant 1}$ be a (real or complex) sequence. Informally, a subsequence of $(a_n)_{n\geqslant 1}$ is a sequence $(b_r)_{r\geqslant 1}$ whose terms are obtained by taking infinitely many terms from $(a_n)_{n\geqslant 1}$, in order. So, for example, if $(a_n) = (1, 2, 3, 4, \dots)$ then:

$(2, 4, 6, \dots)$	is a subsequence of (a_n)	(it is the sequence (a_{2n})),
$(2, 2^2, 2^4, \dots)$	is a subsequence of (a_n)	(it is the sequence (a_{2^n})),
$(6, 4 \dots)$	is not a subsequence of (a_n)	(terms not in correct order),
$(2, 4, 0, 0, \dots)$	is not a subsequence of (a_n)	(not all terms are terms of (a_n)),
$(1, 2, 3, 4, \dots, 2015)$	is not a subsequence of (a_n)	(not a sequence).

Formally, a **subsequence** $(b_r)_{r\geqslant 1}$ of the sequence $(a_n)_{n\geqslant 1}$ is defined by a map $f\colon \mathbb{N} \to \mathbb{N}$ such that f is **strictly increasing** (meaning that $r < s$ implies $f(r) < f(s)$), so that

$$b_r := a_{n_r}, \text{ where } n_r = f(r).$$

Expressing this another way, we have a infinite sequence of natural numbers

$$n_1 < n_2 < n_3 < \dots$$

and the sequence $(b_r) = (a_{n_r})$ has terms

$$a_{n_1}, a_{n_2}, a_{n_3}, \dots .$$

Points to note:

- Being formal: if (a_n) is defined by a function $\alpha\colon \mathbb{N} \to \mathbb{R}$ (or $\mathbb{C}$) then the subsequences of (a_n) are defined by the functions of the form $\alpha \circ f$, where $f\colon \mathbb{N} \to \mathbb{N}$ is strictly increasing.
- The variable r used to label the terms of (b_r) is a dummy variable; any other variable name (except n) would have done equally well.
- $r \leqslant n_r$ for all $r \in \mathbb{N}$ (proved by induction on r).

7.2. **Proposition (subsequences of a convergent sequence).**

Let (a_n) be a sequence.

(i) Assume (a_n) converges to L. Then every subsequence (a_{n_r}) of (a_n) converges, to the same limit L.
(ii) Assume (a_n) has subsequences which converge to limits L and M where $L \neq M$. Then (a_n) does not converge,

Proof. (i) Take $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geqslant N$ implies $|a_n - L| < \varepsilon$ for all $n \geqslant N$. In particular, $n_r \geqslant N$ implies $|a_{n_r} - L| < \varepsilon$. Since $r \leqslant n_r$ this holds whenever $r \geqslant N$. Thus we have proved that

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall r \geqslant N \quad |a_{n_r} - L| < \varepsilon.$$

(ii) is just the contrapositive of (i) (using uniqueness of limits 6.13). □

7.3. **Example (using subsequences to establish divergence).** Take

$$a_n = (-1)^n \left(\frac{n^2}{n^2+1} \right).$$

Then (a_{2n}) converges to 1 and (a_{2n+1}) converges to -1. By 7.2(ii), (a_n) does not converge.

---

8. The Algebra of Limits (AOL)

In Section 6 we concentrated on the limit definition and familiarisation examples, and explored the relationship between limits and order. We did not consider how limits interact with arithmetic operations. We now remedy this omission. This section is unashamedly technical, with lots of ε-N proofs. It supplies machinery needed for working correctly with limits of sequences, in this course and beyond. [*Only a selection of the proofs will be presented in lectures.*]

8.1. **epsilon-handling.** Recall the definition of $a_n \to L$ as $n \to \infty$:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad |a_n - L| < \varepsilon.$$

If we want to *prove* $a_n \to L$ as $n \to \infty$:–

- we must take an *arbitrary* $\varepsilon > 0$ and show there exists a corresponding N.

If we *know* $a_n \to L$ as $n \to \infty$:–

- we can make use of the convergence definition with any ε that suits us.

8.2. **Facilitating the construction of ε-N proofs: making proofs less fiddly.**

(1) As noted earlier, in proving convergence it is sufficient to consider, for example, $\varepsilon < 1$; it is only *small* values of ε that we need to consider.

(2) (a_n) converges to L if, for some *constant* $K > 0$,

$$\forall \eta > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad |a_n - L| < K\eta.$$

Here it is crucially important that K is constant—it must not involve n.

Proof. Take $\varepsilon > 0$. Apply the given condition with $\eta := \varepsilon/K$ to get the standard condition for convergence. □

(3) Remember that (a_n) converges provided it has a convergent tail; recall the Tails Lemma, 6.7. [*We needn't even fuss if we have a sequence whose first few terms are not defined.*]

We'll split the AOL results into two groups. Here are the ones which are easier to prove.

8.3. **Theorem: (AOL), Part I.**

Assume that (a_n) and (b_n) are real or complex sequences and assume that (a_n) converges to L and (b_n) converges to M. Then the following hold as $n \to \infty$.

(i) (constant) If $a_n = a$ (constant) for all n then $a_n \to a$.
(ii) (addition) $a_n + b_n \to L + M$.
(iii) (scalar multiplication) $ca_n \to cL$ for any constant c.
(iv) (subtraction) $a_n - b_n \to L - M$.

Not strictly algebra, but recalled here for convenience:

(v) $a_n \to L$ as $n \to \infty$ implies $|a_n| \to |L|$ as $n \to \infty$.

Proof. (i) Immediate from convergence definition.

(ii) If $c = 0$ use (i). Now assume $c \neq 0$ and let $\varepsilon > 0$. Apply the convergence definition with ε replaced by $\varepsilon/|c|$ to find N such that $n \geqslant N$ implies $|a_n - L| < \varepsilon/|c|$. Then

$$n \geqslant N \implies |ca_n - cL| = |c||a_n - L| < |c|\frac{\varepsilon}{|c|} = \varepsilon.$$

(iii) Choose N_1 and N_2 so that

$$n \geqslant N_1 \implies |a_n - L| < \frac{\varepsilon}{2} \quad \text{and} \quad n \geqslant N_2 \implies |b_n - M| < \frac{\varepsilon}{2}.$$

Then, by the Triangle Law, $n \geqslant N := \max(N_1, N_2)$ implies

$$|(a_n + b_n) - (L + M)| = |(a_n - L) + (b_n - M)| \leqslant |a_n - L| + |b_n - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(iv) Use (iii) and (ii) (with $c = -1$).

(v) Use the Reverse Triangle Law; see 6.14. □

8.4. **Examples:(AOL) Part I.**

(a) $2^{-n} + 1000n^{-5} + \dfrac{n}{\sqrt{n^2+1}} \to 1$ as $n \to \infty$, by (AOL) (addition) and earlier examples.

(b) Let $a_n = n^{-2} + (-1)^n$. Then (a_n) diverges.

Proof. Argue by contradiction. Note (n^{-2}) converges. If (a_n) converged, then $((-1)^n)$ would converge, by (AOL) (subtraction). □

8.5. **Facilitating the construction of ε-N proofs: useful technical facts.**

(1) Assume $a_n \to 0$ and (b_n) is bounded. Then $a_n b_n \to 0$ as $n \to \infty$. (Easy exercise on Problem sheet 3.)
(2) Assume $a_n \to L$ and $L \neq 0$. Then there exists N such that $n \geqslant N$ implies $|a_n| > |L|/2$ (and in particular $a_n \neq 0$) and hence

$$\exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \frac{1}{|a_n|} < \frac{2}{|L|}$$

(so $(1/|a_n|)$ has a bounded tail).

Proof. Apply the convergence definition with $\varepsilon = |L|/2$ to find N such that

$$n \geqslant N \implies |a_n - L| < \frac{1}{2}|L|.$$

By the Triangle Law,

$$n \geqslant N \implies |L| = |L - a_n + a_n| \leqslant |a_n - L| + |a_n| < \frac{1}{2}|L| + |a_n|.$$ □

8.6. **Theorem: (AOL), Part II.** Assume that (a_n) and (b_n) are (real or complex) sequences and assume that (a_n) converges to L and (b_n) converges to M. Then the following hold as $n \to \infty$.

(vi) (product) $a_n b_n \to LM$.

(vii) (reciprocal) If $M \neq 0$, then $1/b_n \to 1/M$.

(viii) (quotient) $a_n/b_n \to L/M$ if $M \neq 0$.

[In (vii) and (viii), we may need to restrict to a tail of, respectively, $(1/b_n)$ and (a_n/b_n) to get well-defined terms. This will always be possible by 8.5(2).]

Proof. Consider (vi). By arithmetic and the Triangle Law,

$$\begin{aligned} |a_n b_n - LM| &= |(a_n - L)(b_n - M) + L(b_n - M) + (a_n - L)M| \\ &\leqslant |(a_n - L)(b_n - M)| + |L||b_n - M| + |a_n - L||M|. \end{aligned}$$

Fix $\varepsilon < 1$ (no loss of generality) and choose N_1, N_2 such that

$$n \geqslant N_1 \implies |a_n - L| < \varepsilon \quad \text{and} \quad n \geqslant N_2 \implies |b_n - M| < \varepsilon.$$

Then

$$n \geqslant N := \max(N_1, N_2) \implies |(a_n - L)(b_n - M)| < \varepsilon^2 < \varepsilon.$$

Hence

$$n \geqslant N \implies |a_n b_n - LM| < \varepsilon(1 + |M| + |L|).$$

Since $1 + |M| + |L|$ is a positive constant, we are done, by 8.2(3).

Now consider (vii). Given $\varepsilon > 0$, we can find N_1 such that $n \geqslant N_1$ implies $|b_n - M| < \varepsilon$ and we can find N_2 such that $n \geqslant N_2$ implies $|b_n| > |M|/2$ (and so also $b_n \neq 0$), by 8.5(2) applied to (b_n). Then

$$n \geqslant \max(N_1, N_2) \implies \left| \frac{1}{b_n} - \frac{1}{M} \right| = \frac{|M - b_n|}{|b_n||M|} \leqslant \frac{2}{|M|^2}\varepsilon.$$

This gives $1/b_n \to 1/M$ as $n \to \infty$ (using 8.2(2) again).

To obtain (viii), combine (vi) and (vii). □

8.7. **Example: a typical (AOL) application.**

[*A simple example of this type will be discussed in lectures before the technical proofs of the* (AOL) *results.*]

Let $a_n = \dfrac{n^2+n+1}{3n^2+4}$. Then

$$a_n = \frac{1}{3}\cdot\frac{n^2\left(1+\dfrac{1}{n}+\dfrac{1}{n^2}\right)}{n^2\left(1+\dfrac{4}{3n^2}\right)} = \frac{1}{3}\left(\frac{1+\dfrac{1}{n}+\dfrac{1}{n^2}}{1+\dfrac{4}{3n^2}}\right) \to \frac{1}{3}\left(\frac{1+0+0}{1+0}\right) = \frac{1}{3},$$

by the fact that $1/n \to 0$ as $n \to \infty$ and the (AOL) results (constant, scalar multiplication, product, quotient).

8.8. **Proposition (reciprocals and infinite/zero limits).** Let (a_n) be a sequence of positive real numbers. The following are equivalent:

(a) $a_n \to \infty$ as $n \to \infty$;

(b) $1/a_n \to 0$ as $n \to \infty$.

The proof is a simple definition-chase. Results in 6.20 are special cases.

8.9. **More worked examples, (AOL).**

(a) [**A general fact worth noting**] Suppose that

$$p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0 \quad\text{and}\quad q(x) = x^\ell + b_{\ell-1}x^{\ell-1} + \cdots + b_1x + b_0$$

are polynomials. Then

$$\frac{p(n)}{q(n)} \to \begin{cases} 1 & \text{if } k = \ell, \\ 0 & \text{if } \ell > k, \\ \infty & \text{if } k > \ell. \end{cases}$$

Proof. Write $p(n)/q(n)$ in the form $n^{k-\ell}[\cdots]$. The quantity in square brackets tends to 1, by the (AOL) results for constants, scalar multiplication, product and quotient. If $k \leqslant \ell$ the required result then follows by (AOL) for product. For the case $k > \ell$, note $q(n)/p(n) \to 0$, by interchanging the roles of p and q above. Then use Proposition 8.8, noting that $q(n)/p(n)$ is positive for large n. □

(b) [**A useful limit**] Let $a > 1$ and $m \in \mathbb{N}^{>0}$. Then $n^m/a^n \to 0$ as $n \to \infty$. [The sequence $(n2^{-n})_{n\geqslant 1}$ considered in 6.7(b) is a special case.]

Proof. Write $a = 1 + b$, where $b > 0$. Then, by the binomial theorem,

$$\frac{n^m}{(1+b)^n} = \frac{n^m}{1+\cdots+\binom{n}{k}b^k+\cdots+b^n}.$$

All the terms in the denominator are positive so if we drop all but one of these we make the fraction bigger. We elect to retain the term with $k = m+1$, where we assume $n > m+1$. We have

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Then

$$0 \leqslant \frac{n^m}{(1+b)^n} \leqslant \frac{(m+1)!}{b^{m+1}}\left(\frac{n^m}{n(n-1)\cdots(n-m)}\right).$$

The expression on the right-hand side tends to 0: the first term is a constant, and the second tends to 0 by (a). Now use simple sandwiching. □

8.10. **A 'true-or-false?' worked example.** Let (a_n) be a real sequence such that $a_n \to \infty$ as $n \to \infty$. To decide, with a proof or counterexample as appropriate, whether $a_n/b_n \to \infty$ under each of the following assumptions.

(a) Assume that (b_n) is a bounded sequence with $b_n \neq 0$. Consider $a_n = n$ and $b_n = (-1)^n$. Then $a_n/b_n \not\to \infty$.
(b) Assume that (b_n) is a bounded sequence with $b_n > 0$ for all n (or for $n \geqslant k$ for some k). Then $a_n/b_n \to \infty$ is true. There exists K (constant) such that if $0 < b_n \leqslant K$, and so also $1/b_n > 1/K$, for a tail of (b_n). Since $a_n \to \infty$, given M, we have $a_n > MK$ for large n and hence $a_n/b_n > (MK)/K = M$.
(c) Assume that (b_n) is a sequence which converges to $L > 0$. Then $a_n/b_n \to \infty$. Note that (b_n) is bounded because it is convergent and that $b_n > 0$ for large n by 8.5(2). Now appeal to (b).

Don't expect AOL results to work when infinite limits come into play. Note also Problem sheet 3, Question 5.

8.11. **Orders of magnitude.** When looking for a candidate limit for a given sequence (a_n), an obvious strategy is to look for the dominate components in a_n, which can be expected to dictate the sequence's behaviour in the long term. Usually for this we need to appreciate the relative magnitudes of the terms as n becomes large.

Examples

(a) Let $a_n = \dfrac{n^3 - 10^7 n}{n^5 + 6n + 1}$. For large n, the dominant term in the numerator of a_n is n^3 and that in the denominator is n^5. So we expect the sequence to behave like (n^{-2}), and to have limit 0. Indeed, this is what exactly what an argument using the Algebra of Limits formalises.

(b) We proved in 6.12(b) by a proper ε-N-argument that $2^{-n}n \to 0$ as $n \to \infty$.
You might try to say simply that 2^n grows 'much faster' than does n, and hence that $a_n \to 0$ as $n \to \infty$. *But this is MUCH too imprecise to constitute a proof.*

(c) Let $a_n = \sin(n^n)/\sqrt{n}$. The oscillatory sine function does not assist in getting convergence of a_n to 0 but because $|\sin x| \leqslant 1$ for all x, the rapid growth of n^n is an irrelevance. To get the limit, just use sandwiching: $|a_n| \leqslant 1/\sqrt{n} \to 0$ as $n \to \infty$.

As a rule of thumb, when it comes to the behaviour of functions $f(x)$ for large x:

trig functions & constants $<$ logarithms
$<$ polynomials $<$ positive exponentials & hyperbolic functions.

More precisely:

- $|\cos n| \leqslant 1$ and $|\sin n| \leqslant 1$ for all n.
- For any rational number $q > 0$, $\log n/n^q \to 0$ as $n \to \infty$ [discussion of case $q = 1$ given as an example in next section].
- For any $a > 1$ and polynomial p then $p(n)/a^n \to 0$ as $n \to \infty$.

8.12. **The O and o notation.** Let (a_n) and (b_n) be real or complex sequences. We write $a_n = \mathrm{O}(b_n)$ as $n \to \infty$ if there exists c such that for some N

$$n \geqslant N \Longrightarrow |a_n| \leqslant c|b_n|.$$

Note that, if $a_n = \mathrm{O}(b_n)$ and $b_n \to 0$ as $n \to \infty$, then $a_n \to 0$ as $n \to \infty$ too. We write $a_n = \mathrm{o}(b_n)$ as $n \to \infty$ if a_n/b_n is defined and

$$\frac{a_n}{b_n} \to 0$$

as $n \to \infty$.

Examples

(a) $n + 10^6 = \mathrm{O}\left(n^2\right)$ is true but $n^2 = O(n)$ is false;
(b) $n = \mathrm{o}\left(n^2\right)$ is true;
(c) $\sin n = \mathrm{O}\left(1\right)$ is true;
(d) by 8.9(b) when $a > 1$ and $m \in \mathbb{N}$ we have $n^m = \mathrm{o}\left(a^n\right)$ as $n \to \infty$.

The symbol $\approx$, used to indicate 'approximately the same size as', is not precise (how close an approximation is intended?) and so $\approx$ is outlawed in this course. O, by contrast, has a precise meaning.

The O notation is useful in particular for simplifying calculations while maintaining rigour. Given a complicated a_n, consideration of dominant terms in the expression defining a_n) can help us a simpler expression b_n with $a_n = \mathrm{O}(b_n)$.

---

9. Monotonic Sequences

So far, in order to prove a sequence converges, we have had to identify a candidate limit at the outset. There is a very important class of *real* sequences which can be guaranteed to converge to a limit $L \in \mathbb{R}$, or to tend to ∞ or to $-\infty$.

9.1. **Monotonic sequences: definitions.** Let (a_n) be a **real** sequence.

- (a_n) is **monotonic increasing** if $a_n \leqslant a_{n+1}$ for all n;
- (a_n) is **monotonic decreasing** if $a_n \geqslant a_{n+1}$ for all n;
- (a_n) is **monotonic** if it is either monotonic decreasing or monotonic increasing.

[Some authors use the term 'monotone' instead of 'monotonic'.]

9.2. **Monotonic Sequence Theorem.** Let (a_n) be a real sequence.

(i) Assume (a_n) is monotonic increasing. Then (a_n) converges if and only if it is bounded above (that is, there exists a finite constant M such that $a_n \leqslant M$ for all n).
(ii) Assume (a_n) is monotonic decreasing. Then (a_n) converges if and only if it is bounded below.

Proof. We prove (i). Assume (a_n) is bounded above. Then $S := \{\, a_n \mid n \in \mathbb{N}\}$ is non-empty and bounded above. By the Completeness Axiom, $\sup S$ exists. We shall prove $a_n \to \sup S$. Take $\varepsilon > 0$. By the Approximation Property for sups (4.8), there exists N such that

$$\sup S - \varepsilon < a_N \leqslant \sup S.$$

But then

$$n \geqslant N \implies \sup S - \varepsilon < a_N \leqslant a_n \leqslant \sup S \implies |a_n - \sup S| < \varepsilon.$$

For the converse we use the fact that any convergent sequence is bounded (6.17).

For (ii), note that (a_n) is monotonic decreasing and bounded below iff $(-a_n)$ is monotonic increasing and bounded above. □

Points to note

- A real sequence which has a tail which is monotonic increasing converges iff it is bounded above.
- A real sequence which is monotonic increasing and not bounded above tends to ∞.

  Proof. Given M there exists N such that $a_N > M$. But then $n \geqslant N$ implies $a_n > M$, since (a_n) is monotonic increasing. $\square$

9.3. **Worked examples: Monotonic Sequence Theorem.**

(a) **An exercise revisited**: on Problem sheet 1, Q.5 you were asked to prove that the sequence (a_n) with

$$a_n = \left(1 + \frac{1}{n}\right)^n$$

is monotonic increasing and bounded above by 3. Hence the Monotonic Sequence Theorem implies that there exists L such that

$$\left(1 + \frac{1}{n}\right)^n \to L \leqslant 3.$$

Also $L \geqslant a_n$ for every n so that, in particular, $L \geqslant 2$. In fact the limit is e, but we need more information about the exponential function to prove this than we currently have available.

(b) Let c be a non-negative real constant. Let (a_n) be defined by

$$a_1 = 1, \qquad a_{n+1} = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \quad (n \geqslant 1).$$

(Note that a_{n+1} is defined so long as $a_n \neq 0$. A trivial induction shows that this holds for all n.) We claim that (a_n) is convergent to a limit $L = \sqrt{c}$. The proof proceeds in stages:

(1) **Find possible limit(s)**: IF $\exists\, L \in \mathbb{R}$ such that $a_n \to L$, then

$$a_{n+1} \to L \qquad \text{(tail)},$$

$$\frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \to \frac{1}{2}\left(L + \frac{c}{L}\right) \qquad \text{(by (AOL) (scalar multiple, sum,quotient))}.$$

Hence, by uniqueness of limits,

$$L = \frac{1}{2}\left(L + \frac{c}{L}\right).$$

This gives $L^2 = c$. Therefore the only possibilities for L are $\pm\sqrt{c}$. Since $a_n > 0$ for all n, we must have $L \geqslant 0$ so $L = -\sqrt{2}$ is ruled out.

(2) **Compare terms to the candidate limit**: Consider

$$a_{n+1}^2 - c = \frac{1}{4}\left(a_n^2 + 2c + \frac{1}{a_n^2}\right) - c = \frac{1}{4a_n^2}\left(a_n - \frac{c}{a_n}\right)^2 \geqslant 0.$$

Hence $a_n \geqslant \sqrt{c}$ for $n \geqslant 2$, so we conjecture it is monotonic decreasing, heading downwards to its potential limit.

In this example Stage (2) is particularly simple. Normally an inductive proof would be needed at this point.

(3) **Test for monotonicity**:

$$a_n - a_{n+1} = a_n - \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) = \frac{1}{2a_n}\left(a_n^2 - c\right).$$

By (2), we have $a_n - a_{n+1} \geqslant 0$ for $n \geqslant 2$.

(4) **Putting the pieces together**: The sequence (a_n) is bounded below by 0 and $(a_n)_{n\geqslant 2}$ is monotonic decreasing by (3). By the Monotonic Sequence Theorem, $a_n \to L$ for some (finite) limit L. By (1), L must be $\sqrt{c}$.

This example proves the existence of $\sqrt{c}$. Where this argument differs from the fiddly one used earlier to show $\sqrt{2}$ exists is in the availability now of the Algebra of Limits. Observe that both the new proof and the old one use the Completeness Axiom.

(c) **Engineering a recurrence relation**: In (b), we started from a sequence given by a recurrence relation, and used the recurrence relation to identify a possible limit for our sequence. Here we illustrate how we can sometimes use this idea when we do not have a recurrence relation given. We consider $a_n = \log n/n$ where log denotes $\log_{\mathrm{e}}$. We shall assume properties of the log function.
Certainly $a_n \geqslant 0$, so (a_n) is bounded below. To show (a_n) is monotonic decreasing we look at the gradient of the function $f(x) = \log x/x$. We have

$$f'(x) = \frac{1}{x^2} - \frac{\log x}{x^2} \leqslant 0 \quad (\text{for } x \geqslant \mathrm{e}).$$

So the function f is decreasing, and in particular $f(n+1) \leqslant f(n)$, for $n \geqslant 3$. Applying the Monotonic Sequence Theorem, there exists $L \geqslant 0$ such that

$$a_n = \frac{\log n}{n} \to L.$$

To find L we elect to relate a_{2n} to a_n (note that we have no useful relationship between $\log(n+1)$ and $\log n$ so it's not helpful to compare a_{n+1} with a_n). We have

$$a_{2n} = \frac{\log 2n}{2n} = \frac{\log 2 + \log n}{2n}.$$

Now we let $n \to \infty$ on both sides: (a_{2n}) is a subsequence of (a_n) and so converges to L. By (AOL) results, the right-hand side tends to $0 + L/2$. By uniqueness of limits, $L = L/2$, so $L = 0$. We have proved that

$$\frac{\log n}{n} \to 0.$$

---

10. The Bolzano–Weierstrass Theorem, and the Cauchy Convergence Criterion

In this section we obtain important general results about sequences which are not monotonic. Since monotonic real sequences behave so well, the following theorem provides welcome information.

10.1. **The Scenic Viewpoint Theorem.** Let (a_n) be a real sequence. Then (a_n) has a monotonic subsequence.

Proof. We consider the set $V = \{\, k \in \mathbb{N} \mid m > k \Longrightarrow a_m < a_k \,\}$. This is the set of "scenic viewpoints" (also known as "peaks")—given $k \in V$, looking towards infinity from a point at height a_k, no higher point would impede our view.

Case 1: V is infinite. Then the elements of V can be enumerated as $k_1 < k_2 < \dots$. Then (a_{k_r}) is a subsequence of (a_n) and

$$r > s \Longrightarrow k_r > k_s \Longrightarrow a_{k_r} < a_{k_s}$$

that is, (a_{k_r}) is monotone decreasing.

Case 2: V is finite. Let N be such that every element of V is $< N$. Then $m_1 = N$ is such that $m_1 \notin V$, so there exists $m_2 > m_1$ with $a_{m_2} \geqslant a_{m_1}$. Since $m_2 \notin V$, there exists $m_3 > m_2$ such that $a_{m_3} \geqslant a_{m_2}$. Proceeding in this way we can (inductively) generate a monotonic increasing sequence (a_{m_k}). □

10.2. **Bolzano–Weierstrass Theorem (for real sequences).** Let (a_n) be a bounded real sequence. Then (a_n) has a convergent subsequence.

Proof. By the Scenic Viewpoint Theorem 10.1, (a_n) has a monotonic subsequence which is also bounded. By the Monotonic Sequence Theorem, this subsequence converges. □

Neither the Monotonic Sequence Theorem nor the Scenic Viewpoint Theorem extends to complex sequences, but the Bolzano–Weierstrass Theorem does.

10.3. **Bolzano–Weierstrass Theorem (for complex sequences).** Let (z_n) be a bounded sequence in $\mathbb{C}$. Then (z_n) has a convergent subsequence.

Proof. Write $z_n = x_n + \mathrm{i}y_n$. Let M be such that $|z_n| \leqslant M$ for all $n \geqslant 1$. Then by the definition of modulus in $\mathbb{C}$, we have $|x_n| \leqslant M$ and $|y_n| \leqslant M$, for all $n \geqslant 1$. Therefore by the real BW Theorem, $(x_n)_{n\geqslant 1}$ has a convergent subsequence, say $(x_{n_r})_{r\geqslant 1}$. Let $w_r := y_{n_r}$, for $r \geqslant 1$. Then the subsequence $(w_r)_{r\geqslant 1}$, being a subsequence of the bounded sequence $(y_n)_{n\geqslant 1}$, is itself bounded. Choose a subsequence $(w_{r_s})_{s\geqslant 1}$ of $(w_r)_{r\geqslant 1}$ which is convergent. Then $(y_{n_{r_s}})_{s\geqslant 1}$ converges and so does $(x_{n_{r_s}})_{s\geqslant 1}$, since it is a subsequence of a convergent sequence. Finally, $(z_{n_{r_s}})_{s\geqslant 1}$ is convergent since the sequences of real and imaginary parts converge. □

10.4. **Cauchy sequences.** The idea behind Cauchy sequences is that if the terms of a sequence (a_n) are ultimately arbitrarily close to one another, then there should be a value L to which they must converge. This would be valuable to know, since up till now we have needed to identify a candidate limit in advance (except for bounded monotonic sequences, where we may be able to find the limit from a given, or constructed, recurrence relation).

Definition. Let (a_n) be a real or complex sequence. Then (a_n) is a **Cauchy sequence** if it satisfies the **Cauchy condition**:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall m, n \geqslant N \quad |a_n - a_m| < \varepsilon.$$

It is not sufficient in the Cauchy condition just to consider adjacent terms, that is, only to consider $m = n + 1$.

10.5. **Facts about Cauchy sequences.**

(1) A Cauchy sequence (a_n) is bounded.
Proof. Choose N such that $m, n \geqslant N$ implies $|a_n - a_m| < 1$. Thus (considering $m = N$) we get $|a_n| \leqslant |a_N| + 1$ for $n \geqslant N$. Now

$$|a_n| \leqslant \max(|a_1|, \dots, |a_{N-1}|, |a_N| + 1). \qquad \square$$

[*This argument is very like the proof in* 6.17 that a convergent sequence is bounded.]

(2) A convergent sequence is a Cauchy sequence.
Proof. Assume $a_n \to L$. Take $\varepsilon > 0$ and choose N such that $k \geqslant N$ implies $|a_k - L| < \varepsilon/2$. Then, using the Triangle Law, $m, n \geqslant N$ implies

$$|a_n - a_m| = |(a_n - L) - (a_m - L)| \leqslant |a_n - L| + |a_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

(3) Let (a_n) be a Cauchy sequence and assume that (a_n) has a subsequence (a_{n_r}) which converges, to L say. Then (a_n) converges to L.
Proof. Take $\varepsilon > 0$ and pick N so that

$$m, n \geqslant N \implies |a_n - a_m| < \frac{\varepsilon}{2}.$$

Since $a_{n_r} \to L$, there exists $N' \in \mathbb{N}$ such that

$$r \geqslant N' \implies |a_{n_r} - L| < \frac{\varepsilon}{2}.$$

Fix r such that $r \geqslant \max(N, N')$. Then $n_r \geqslant N$ (since $n_r \geqslant r$). Then applying Cauchy condition with $m = n_r$,

$$n \geqslant N \implies |a_n - L| \leqslant |a_n - a_{n_r}| + |a_{n_r} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}. \qquad \square$$

10.6. **Theorem (Cauchy Convergence Criterion).** Let (a_n) be a (real or complex) sequence. Then

$$(a_n) \text{ is convergent} \iff (a_n) \text{ is a Cauchy sequence.}$$

Proof. $\implies$: By 10.5(2).

$\impliedby$: By 10.5(1) and the BW Theorem, 10.2 or 10.3, (a_n) has a convergent subsequence. Now appeal to 10.5(3). $\square$

Applications of Theorem 10.6 will come later.

---

11. Convergence of Series

11.1. **Series: introductory examples.**

(a) Infinite geometric progressions: for $r \neq 1$,

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

and, if $|r| < 1$,

$$\lim_{n \to \infty} \left(1 + r + r^2 + \cdots + r^n\right) = \frac{1}{1 - r}.$$

(b) Decimal expansions: for

$$\frac{1}{3} = 0{\cdot}333333333333333\cdots, \text{ meaning } \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \cdots.$$

Decimal representations exist for all real numbers. See Supplementary Notes for an account, and a proof via decimals that $\mathbb{R}$ is uncountable.

(c) $\mathrm{e} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$ *by definition.*

(d) Functions given by (or defined by?) infinite sums: for example,

$$\mathrm{e}^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots.$$

11.2. **Series: definitions.** Let (a_k) be a real or complex sequence. Let

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^{n} a_k \quad (n \geqslant 1).$$

Then we say that **the series** $\sum_{k\geqslant 1} a_k$ **converges** (or as shorthand that the series $\sum a_k$ converges) if the sequence (s_n) of **partial sums** converges. If $s_n \to s$, then we write $\sum_{k=1}^{\infty} a_k = s$. If (s_n) fails to converge, then we say $\sum_{k\geqslant 1} a_k$ **diverges**.

A trivial but important observation:

$$a_n = s_n - s_{n-1} \qquad \text{for } n \geqslant 2;$$

Note on notation: We have two sequences in play at the same time here: the sequence (a_k) of the terms of the series and the sequence (s_n) of partial sums. To avoid writing (erroneously!) $\sum_{n=1}^{n} a_n$, we have introduced a new (dummy) variable k to label the terms.

11.3. **Convergence of series: first examples.**

(a) **Geometric series**: Let $z \in \mathbb{C}$. Let $a_k = z^k$, so

$$s_n = a_1 + \cdots + a_n = \begin{cases} \dfrac{z(1-z^n)}{1-z} & \text{if } z \neq 1, \\ n & \text{if } z = 1. \end{cases}$$

Hence $\sum_{k\geqslant 1} z^k$ converges if $|z| < 1$ and fails to converge if $|z| \geqslant 1$ (recall 6.23). (Note: it might have been more natural here to start from $k = 0$ rather than $k = 1$ and to consider $\sum_{k\geqslant 0} z^k$.)

(b) **A telescoping series**: Let $a_k = 1/(k(k+1))$. Then

$$\begin{aligned} s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} &= \sum_{k=1}^{n} \Big(\frac{1}{k} - \frac{1}{k+1}\Big) \\ &= \Big(1 - \frac{1}{2}\Big) + \Big(\frac{1}{2} - \frac{1}{3}\Big) + \cdots + \Big(\frac{1}{n} - \frac{1}{n+1}\Big) = 1 - \frac{1}{n+1} \to 1. \end{aligned}$$

Hence $\sum_{k\geqslant 1} \dfrac{1}{k(k+1)}$ converges.

Note that we are working here with *finite* sums of real numbers and all the normal rules of arithmetic apply, in particular associativity of addition. This is always the case when we work with partial sums.

(c) Let $a_k = (-1)^k$. Then
$$s_n = \begin{cases} 0 & \text{if } n \text{ is even,} \\ -1 & \text{if } n \text{ is odd.} \end{cases}$$
Hence $\sum(-1)^k$ diverges. But, naively,
$$\begin{aligned} \sum_{k=1}^{\infty}(-1)^k &= (-1) + 1 + (-1) + 1 + (-1) + 1 + \dots \\ &= ((-1)+1) + ((-1)+1) + \cdots = 0 + 0 + \cdots = 0. \end{aligned}$$
This is clearly wrong. The error lies in the implicit assumption that infinite sums satisfy the same arithmetic properties as finite sums do. They don't, in general.

Deciding whether a *series* $\sum_{k\geqslant 1} a_k$ converges requires us to decide whether the *sequence* $(s_n) = \left(\sum_{k=1}^n a_k\right)$ converges. We can bring everything we know about convergence of **sequences** to bear on this problem. This leads immediately to a clutch of results sufficiently important to be recorded as theorems.

11.4. **Theorem (terms of a convergent series).**

(i) Assume $\sum_{k\geqslant 1} a_k$ converges. Then $a_k \to 0$ as $k \to \infty$.
(ii) A **sufficient** condition for $\sum_{k\geqslant 1} a_k$ to diverge is that $a_k \nrightarrow 0$.

Proof. (i) Assume $s_n \to s$. For $n \geqslant 2$ we have $a_n = s_n - s_{n-1} \to s - s = 0$.
(ii) is just the contrapositive of (i). □

11.5. **Example: the harmonic series, $\sum_{k\geqslant 1} \dfrac{1}{k}$.** Let
$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \qquad (n \geqslant 1).$$
We claim (s_n) is not a Cauchy sequence, so (s_n) does not converge and therefore
$$\sum_{k\geqslant 1} \frac{1}{k} \text{ diverges.}$$

Proof. Consider
$$\begin{aligned} \left|s_{2^{n+1}} - s_{2^n}\right| &= \frac{1}{2^n+1} + \cdots + \frac{1}{2^{n+1}} \\ &\geqslant \frac{1}{2^{n+1}} \cdot (2^{n+1} - 2^n) \qquad \text{((smallest term)} \times \text{(number of terms))} \\ &= \frac{1}{2}. \end{aligned}$$
Therefore (s_n) is not Cauchy, so fails to converge (by 10.6)). □

This example shows also that $a_k \to 0$ does not imply $\sum_{k\geqslant 1} a_k$ converges: $\sum_{k\geqslant 1} \dfrac{1}{k}$ provides a counterexample.

Later (12.13(c)) we'll see that s_n grows about as fast as $\log n$, in that $s_n - \log n$ tends to a finite constant.

A particularly amenable class of series will be those whose partial sum sequences are monotonic, thanks to the Monotonic Sequence Theorem.

11.6. **Theorem (series of non-negative terms).** Assume a_k is real. Then

$$(s_n) \text{ is monotonic increasing if and only if } a_k \geqslant 0 \text{ (for } k \geqslant 2).$$

Moreover, if a_k is non-negative, then $\sum_{k\geqslant 1} a_k$ converges if and only if its partial sum sequence (s_n) is bounded above.

Proof. By the fact that $s_n - s_{n-1} = a_n$ $(n \geqslant 2)$ and Monotonic Sequence Theorem, 9.2. □

11.7. **Theorem (Comparison Test, simple form).** Assume $0 \leqslant a_k \leqslant Cb_k$, where C is a positive constant. Then $\sum_{k\geqslant 1} b_k$ convergent implies $\sum_{k\geqslant 1} a_k$ convergent. Also $\sum_{k\geqslant 1} a_k$ divergent implies $\sum_{k\geqslant 1} b_k$ divergent

Proof. Let

$$s_n = a_1 + \cdots + a_n \qquad \text{and} \qquad t_n = b_1 + \cdots + b_n.$$

Then $s_n \leqslant Ct_n$ for all n. Since (t_n) converges, it is bounded above, by T say. Hence $s_n \leqslant T$ for all n, and Theorem 11.6 applies. □

11.8. **Examples: simple Comparison Test.**

(a) Take $a_k = 1/k^2$ and $b_k = 1/(k(k+1))$. Then $0 \leqslant a_k \leqslant 2b_k$ and we proved earlier that $\sum_{k\geqslant 1} b_k$ converges. Hence $\sum_{k\geqslant 1} \dfrac{1}{k^2}$ converges.

(b) Let $a_k = 1/k!$ and $b_k = 1/(k(k+1))$. Then $0 \leqslant a_k \leqslant b_k$ and hence the series $\sum_{k\geqslant 1} 1/k!$ converges by comparison with $\sum_{k\geqslant 1} 1/(k(k+1))$. Hence the series defining e converges (note this series is not exactly $\sum_{k\geqslant 1} a_k$ because the latter has one fewer term at the front but this does not affect convergence).

So far in considering series we have restricted attention to

- series whose partial sums we can compute explicitly (for example, geometric or telescoping series);
- series of *non-negative terms*, whose partial sum sequences are *monotonic increasing*.

In general, given a series $\sum_{k\geqslant 1} a_k$ we won't be able to calculate $s_n = \sum_{k=1}^{n} a_k$ explicitly. Indeed we may want, as with e, to use the sum of a series to *define* a number and in such cases we won't have a convenient formula for the partial sums. The Cauchy Criterion, 10.6, provides a way of testing a sequence for convergence without knowing a candidate limit. We now apply this to the partial sum sequence of a series.

11.9. **Theorem (Cauchy criterion for convergence of a series).** Let (a_k) be a real or complex sequence with partial sum sequence (s_n). Then $\sum_{k\geqslant 1} a_k$ converges if and only if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n > m \geqslant N \quad |a_{m+1} + \cdots + a_n| = |s_n - s_m| < \varepsilon.$$

11.10. **Absolute convergence.** Let (a_k) be a sequence of real or complex numbers. We say $\sum_{k\geqslant 1} a_k$ **converges absolutely** if $\sum_{k\geqslant 1} |a_k|$ converges.

Note that $\sum_{k\geqslant 1} |a_k|$ is a series of *non-negative* terms, to which Theorem 11.6 applies. Thus the following theorem is very useful.

11.11. **Theorem (absolute convergence implies convergence).** Let (a_k) be a real or complex sequence. Then

$$\sum_{k\geqslant 1} |a_k| \text{ converges} \implies \sum_{k\geqslant 1} a_k \text{ converges.}$$

Proof. Let $s_n = \sum_{k=1}^n a_k$ and $S_n = \sum_{k=1}^n |a_k|$. For $n > m$,

$$|s_n - s_m| = |a_{m+1} + \cdots + a_n| \leqslant |a_{m+1}| + \cdots + |a_n| = |S_n - S_m|,$$

by the Triangle Law (as it extends by induction to finite sums). Hence

$$\begin{aligned} \sum_{k\geqslant 1} |a_k| \text{ convergent } &\implies \{S_n\} \text{ is a Cauchy sequence} && \text{(Cauchy Criterion, easy direction)} \\ &\implies \{s_n\} \text{ is a Cauchy sequence} && \text{(from above)} \\ &\implies (s_n) \text{ converges} && \text{(Cauchy Criterion, harder direction).} \end{aligned}$$

□

Example Let $a_k = (-1)^{3k}\dfrac{\sin^3(k^2)}{k^2+1}$. Here we do not have $a_k \geqslant 0$. But $0 \leqslant |a_k| \leqslant \dfrac{1}{k^2}$. Hence $\sum_{k\geqslant 1} |a_k|$ converges by comparison with $\sum 1/k^2$ and so $\sum_{k\geqslant 1} a_k$ converges too.

w

11.12. **A standard example: the series $\sum_{k\geqslant 1} k^{-p}$ for $p \in \mathbb{R}$.**

(a) $\sum_{k\geqslant 1} k^{-p}$ diverges if $p \leqslant 1$;
Proof. For $p \leqslant 0$, the terms do not tend to 0. For $p = 1$ we have the harmonic series. For $0 < p < 1$, use the contrapositive of the simple Comparison Test, with $a_k = k^{-1}$ and $b_k = k^{-p}$. □

(b) $\sum_{k\geqslant 1} k^{-p}$ converges if $p \geqslant 2$.
Proof. Use simple Comparison Test, comparing with $\sum_{k\geqslant 1} 1/k^2$, which we proved earlier is convergent. □

(c) [Looking ahead] $\sum_{k\geqslant 1} k^{-p}$ converges if $1 < p < 2$.

Proof deferred until we have available the Integral Test, which allows us to handle all values of $p \geqslant 0$ in a uniform way; see 12.13(a).

In summary:

$\sum k^{-p}$ diverges if $p \leqslant 1$ and converges if $p > 1$.

---

12. Convergence Tests for Series

In this section we shall obtain some very useful tests for convergence/divergence of series.

For testing series of non-negative terms and so for testing for absolute convergence (from which convergence follows):

- **Comparison Test (limit form)**;
- **D'Alembert's Ratio Test**;
- **Integral Test**.

For testing series $\sum(-1)^{k-1}u_k$, where $u_k \geqslant 0$:

- **Leibniz' Alternating Series Test**.

We shall present the Alternating Series Test first. It is quite special and there are close connections with material in the previous section.

12.1. **Leibniz' Alternating Series Test.** The series $\sum(-1)^{k-1}u_k$ converges if

(i) $u_k \geqslant 0$,
(ii) $u_{k+1} \leqslant u_k$,
(iii) $u_k \to 0$ as $k \to \infty$.

Proof. Consider the subsequence (s_{2n}). We have

$$\begin{aligned} s_{2n} &= u_1 - u_2 + u_3 - u_4 + \cdots + u_{2n-1} - u_{2n} \\ &= u_1 - (u_2 - u_3) - \cdots - (u_{2n-2} - u_{2n-1}) - u_{2n} \leqslant u_1, \end{aligned}$$

by (i) and (ii). Also, by (ii) again,

$$s_{2(n+1)} - s_{2n} = u_{2n+1} - u_{2n+2} \geqslant 0.$$

Hence (s_{2n}) is monotonic increasing and bounded above, so converges, to s say, by the Monotonic Sequence Theorem. Now consider (s_{2n-1}). We then have, using (iii) and (AOL),

$$s_{2n-1} = s_{2n} - u_{2n} \to s - 0 = s.$$

Hence (s_{2n}) and (s_{2n-1}) converge to the same limit. Hence (see Problem sheet 4, Q.2(a)), (s_n) converges. $\square$

12.2. **Examples: AST.**

(a) By AST, $\sum_{k\geqslant 1}(-1)^{k-1}\dfrac{1}{k}$ converges.

Here we have an example of a series which converges but fails to converge absolutely, since $\sum_{k\geqslant 1} 1/k$ diverges (recall 11.5).

(b) By AST, $\sum_{k\geqslant 1}(-1)^{k-1}\dfrac{1}{\sqrt{k}}$ converges. This is a useful series for counterexamples.

12.3. **Comparison Test, simple form (recap of 11.7 and remarks).**

Assume a_k, b_k are real numbers and that

$$0 \leqslant a_k \leqslant Cb_k \quad \text{for all } k, \qquad \text{where } C > 0 \text{ is a constant.}$$

Then

$$\begin{cases} \sum a_k \text{ converges} & \text{if } \sum b_k \text{ converges;} \\ \sum b_k \text{ diverges} & \text{if } \sum a_k \text{ diverges.} \end{cases}$$

Suppose we wish to use the Comparison Test to prove that $\sum a_k$ is convergent and have a candidate series $\sum b_k$ with which to compare. In practice it can be awkward to get a valid inequality $a_k \leqslant Cb_k$ so we can apply the test: consider for example

$$a_k = \frac{k^2 + k + 1}{4k^4 - k^2 - 1}.$$

We'd want to try taking $b_k = 1/k^2$, but valid applications of the Triangle Law (on the numerator of a_k) and Reverse Triangle Law (on the denominator of a_k) don't make it very easy to find a suitable C. Fortunately there's a version of the Comparison Test which is very easy to apply, without the need to manipulate inequalities.

12.4. **Comparison Test, limit form.** Let a_k and b_k be strictly positive and assume that

$$\frac{a_k}{b_k} \to L, \quad \text{where } 0 < L < \infty.$$

Then

$$\sum a_k \text{ converges} \quad \Longleftrightarrow \quad \sum b_k \text{ converges.}$$

Proof. In the limit definition take $\varepsilon = L/2$ and choose N such that, for $k \geqslant N$,

$$\left|\frac{a_k}{b_k} - L\right| < \frac{1}{2}L \quad \text{and hence} \quad \frac{1}{2}L < \frac{a_k}{b_k} < \frac{3}{2}L.$$

Then, restricting to tails, $\sum b_k$ convergent implies that $\sum a_k$ is convergent, by the simple Comparison Test with $C = 3L/2$. In the other direction, $\sum a_k$ convergent implies $\sum b_k$ converges by comparison, because $0 < b_k < 2L^{-1}a_k$ for large k. □

Note: It is crucial that the terms a_k and b_k are (ultimately) strictly positive and that L is non-zero and finite.

12.5. **Examples: Comparison Test (limit form).**

(a) Take $a_k = \dfrac{k^2+k+1}{4k^4-k^2-1}$. Take $b_k = k^{-2}$ and use (AOL) to get $a_k/b_k \to 1/4$. Hence $\sum a_k$ converges because $\sum b_k$ does.

(b) Let $a_k = \dfrac{1}{k}\left(1+\frac{1}{k}\right)^{-k}$. Take $b_k = \dfrac{1}{k}$. Then

$$\frac{a_k}{b_k} = \left(1+\frac{1}{k}\right)^{-k} \to \mathrm{e}^{-1}$$

(standard limit). So $\sum a_k$ diverges by comparison with the divergent series $\sum b_k$.

The Comparison Test in either form is not 'internal', in that we have to produce a suitable series with which to compare. We next obtain some important tests which don't have this disadvantage.

12.6. **D'Alembert's Ratio Test, for series of strictly positive terms.** Let $a_k > 0$. Assume that

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} \text{ exists and equals } L.$$

Then

$$\begin{aligned} 0 \leqslant L < 1 \quad &\Longrightarrow \quad \textstyle\sum a_k \text{ converges;} \\ L > 1 \quad &\Longrightarrow \quad \textstyle\sum a_k \text{ diverges;} \\ L = 1 \quad &\Longrightarrow \quad \text{the test gives no result.} \end{aligned}$$

(Here it is permissible, and useful, to allow $L = \infty$ as a possible limit and to treat it as > 1 for the purposes of stating the test.)

Proof. The idea will be to compare with a suitable geometric series.

Case 1: $L < 1$. Fix K such that $L < K < 1$ and apply the limit definition with $\varepsilon = K - L$. Then there exists $N \in \mathbb{N}$ such that

$$k \geqslant N \implies \left|\frac{a_{k+1}}{a_k} - L\right| < K - L \text{ and hence } \frac{a_{k+1}}{a_k} < L + (K-L) = K.$$

So, for $k \geqslant N$, we get

$$0 < a_{k+1} < K a_k.$$

By induction, we obtain, for $k \geqslant 1$,

$$0 < a_{k+N} < a_N K^k.$$

Hence, by the simple Comparison Test, comparing with the convergent geometric series $\sum K^k$, the series $\sum a_k$ has a convergent tail and so converges.

Case 2: $L > 1$**.** Assume first that L is finite. Then pick K such that $L > K > 1$ and apply the limit definition with $\varepsilon = L - K$. Then there exists N such that $k \geqslant N$ implies

$$\left| \frac{a_{k+1}}{a_k} - L \right| < L - K$$

and hence, by the Reverse Triangle Inequality,

$$\frac{a_{k+1}}{a_k} > L - (L - K) = K,$$

If $L = \infty$, then we can, by definition of an infinite limit, certainly find a constant K such that $a_{m+1}/a_m > K$ for all suitably large m.

Now, using an argument similar to that in Case 1, we get, for some N,

$$a_{k+N} > a_N K^k \text{ for all } k \geqslant 1.$$

This shows that $a_k \not\to 0$ so $\sum a_k$ diverges. □

12.7. **Ratio Test, preliminary examples.**

(a) Consider $\sum a_k$, where $a_k = \dfrac{2^k}{k!}$. Here we have $a_k > 0$ and

$$\lim \frac{a_{k+1}}{a_k} = \lim \frac{k!}{(k+1)!} \frac{2^{k+1}}{2^k} = \lim \frac{2}{k+1} = L, \text{where } L = 0.$$

Hence the series converges,

(b) **Don't forget to take the limit!** Consider $\sum a_k$ where $a_k = \dfrac{1}{k^p}$, where $p > 0$. Here we have

$$\lim \frac{a_{k+1}}{a_k} = \lim \left(\frac{k}{k+1} \right)^p = 1.$$

Hence the Ratio Test, correctly applied, gives no result.
Consider $a_k = 1/k$ (the case $p = 1$). Here we have

$$\frac{a_{k+1}}{a_k} = \frac{k}{k+1} < 1.$$

Note that we already know that $\sum 1/k$ diverges. Hence having $a_{k+1}/a_k < 1$, rather than $\lim a_{k+1}/a_k = L < 1$, does *not* give a sufficient condition for a series $\sum a_k$ of positive terms to converge.

12.8. **Ratio Test: tips and warnings.**

- It is quite possible that $\lim a_{k+1}/a_k$ fails to exist: consider for example $\sum a_k$ given by

$$1 + 1 + \frac{1}{3} + \frac{1}{2} + \frac{1}{3^2} + \frac{1}{2^2} + \frac{1}{3^3} + \frac{1}{2^3} + \ldots.$$

(This series does converge: can you prove this by the simple Comparison Test?)

- For a series with 'gaps', leave out the zero terms and relabel before applying the test. For example, consider

$$a_k = \begin{cases} \dfrac{1}{2^{k!}} & \text{if } k \text{ is of the form } 2^m \text{ for some } m \geqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

so $\sum a_k$ (starting from $k = 1$) looks like

$$0 + \frac{1}{2^{1!}} + 0 + \frac{1}{2^{4!}} + 0 + 0 + 0 + \frac{1}{2^{8!}} + 0 + \dots .$$

We cannot apply the Ratio Test directly to $\sum_{k \geqslant 1} a_k$ but we can apply it to $\sum m \geqslant 11/2^{(2^m)!}$.

- We cannot prove that a geometric series $\sum r^k$ is convergent/divergent by applying the Ratio Test. Why not?

12.9. **Testing for absolute convergence: a corollary to the Ratio Test 12.6.** Let a_k be non-zero real or complex numbers and assume that

$$\lim_{k \to \infty} \left| \frac{a_{k+1}}{a_k} \right| \text{ exists and equals } L.$$

(Here we allow $L = \infty$.) Then

$$\begin{aligned} L < 1 &\implies \textstyle\sum a_k \text{ converges absolutely, and hence converges;} \\ L > 1 &\implies \textstyle\sum a_k \text{ diverges;} \\ L = 1 &\implies \text{the test gives no result.} \end{aligned}$$

Proof. For the case $L < 1$ we apply the Ratio Test 12.6, to $\sum |a_k|$. If $L > 1$, again considering $\sum |a_k|$, we note that the proof given above implies that $|a_k| \not\to 0$, so $a_k \not\to 0$ and hence $\sum a_k$ diverges. □

The (corollary to the) Ratio Test is particularly useful for testing series

$$\sum c_k x^k \quad (x \text{ real}) \quad \text{and} \quad \sum c_k z^k \quad (z \text{ complex})$$

for convergence. Here we regard x or z as a variable, so that the series, provided it is convergent, will define a function. The whole of Section 13 is devoted to such **power series** and their properties. There we give further examples of the use of the Ratio Test in important particular cases.

12.10. **The Integral Test: preamble.** Here we shall analyse the the behaviour of the partial sum sequence (s_n) of a series $\sum f(k)$ by comparing it with the sequence (I_n) where

$$I_n = \int_1^n f(x)\,\mathrm{d}x,$$

where $f \colon [1, \infty) \to [0, \infty)$ is a suitable function. We shall need to make use of standard properties of integrals (integration is treated in Analysis III in Trinity Term). We assume the following. On a suitable class of integrable functions:

(a) integration preserves $\leqslant$;
(b) $\int_k^{k+1} c\,\mathrm{d}x = c$, for any constant c;
(c) intervals slot together: $\int_a^c = \int_a^b + \int_b^c$ for $a < b < c$.

12.11. **Integral Test Theorem.** Assume that f is a real-valued function defined on $[1, \infty)$ with the following properties:

(i) f is non-negative and decreasing;

(ii) $\int_k^{k+1} f(x)\,\mathrm{d}x$ exists for each $k \geqslant 1$ [for future reference: this holds if f is continuous].

Let

$$s_n = \sum_{k=1}^{n} f(k) \quad \text{and} \quad I_n = \int_1^n f(x)\,\mathrm{d}x.$$

Let $\sigma_n = s_n - I_n$. Then (σ_n) converges to a limit σ, where $0 \leqslant \sigma \leqslant f(1)$.

Proof. Note that, because f is decreasing, $f(k) \geqslant f(x) \geqslant f(k+1)$ for all $x \in [k, k+1]$. Now, by properties (a) and (b) above,

$$f(k+1) \leqslant \int_k^{k+1} f(x)\,\mathrm{d}x \leqslant f(k) \quad (k = 1, 2, \ldots).$$

So we have

$$\begin{aligned} f(2) &\leqslant \int_1^2 f(x)\,\mathrm{d}x &&\leqslant f(1) \\ f(3) &\leqslant \int_2^3 f(x)\,\mathrm{d}x &&\leqslant f(2) \\ \cdots & \qquad \cdots && \quad \cdots \\ f(n) &\leqslant \int_{n-1}^n f(x)\,\mathrm{d}x &&\leqslant f(n-1). \end{aligned}$$

Add these inequalities and use property (c) to get

$$\sum_{r=1}^{n} f(r) - f(1) = \sum_{k=1}^{n-1} f(k+1) \leqslant \int_1^n f(x)\,\mathrm{d}x \leqslant \sum_{k=1}^{n} f(k) - f(n).$$

Then

$$0 \leqslant f(n) \leqslant \sum_{k=1}^{n} f(k) - \int_1^n f(x)\,\mathrm{d}x \leqslant f(1),$$

so that $0 \leqslant \sigma_n \leqslant f(1)$. Also

$$\sigma_{n+1} - \sigma_n = f(n+1) - \int_n^{n+1} f(x)\,\mathrm{d}x \leqslant 0.$$

Therefore (σ_n) is monotonic decreasing and bounded below, and hence converges to a limit σ, where $0 \leqslant \sigma \leqslant f(1)$. □

Some sources work with a different choice of monotonic sequence.

12.12. **Corollary: the Integral Test.** Assume that (as in 12.11), $f\colon [1, \infty) \to [0, \infty)$ is monotonic decreasing and such that $\int_k^{k+1} f(x)\,\mathrm{d}x$ exists for each k. Then $\sum f(k)$ converges if and only if (I_n) converges.

Proof. This is immediate from elementary properties of limits and the fact that $(s_n - I_n)$ converges. □

12.13. **Applications of the Integral Test and Integral Test Theorem.** Here we assume not just properties of integrals but also how to evaluate standard integrals.

(a) $\sum k^{-p}$ ($p > 0$), definitively. Take $f(x) = x^{-p}$. Then f is non-negative and decreasing on $[1, \infty)$. We have

$$I_n = \int_1^n x^{-p}\,\mathrm{d}x = \begin{cases} \dfrac{1}{-p+1}\left[x^{-p+1}\right]_1^n = \dfrac{1}{-p+1}(n^{-p+1} - 1) & \text{if } p \neq 1, \\ \log n & \text{if } p = 1. \end{cases}$$

Hence (I_n) converges (to a finite limit) iff $p > 1$. Therefore, as claimed in 11.12, $\sum k^{-p}$ converges if $p > 1$ and diverges if $0 < p \leqslant 1$. [If $p < 0$ the Integral Test does not apply but $\sum k^{-p}$ diverges because $k^{-p} \nrightarrow 0$.]

(b) The Integral Test applied with $f(x) = 1/(x \log x)$ (note that the conditions for the test are met, except that we need to start from $x = 2$ rather than $x = 1$) implies that $\sum_{k \geqslant 2} 1/(k \log k)$ diverges. This is, perhaps, a little surprising.

(c) **Euler's constant,** γ. Apply the Integral Test Theorem in the special case that $f(x) = 1/x$; certainly f is non-negative and monotonic decreasing, and its integral exists over any interval $[k, k+1]$. We have that

$$\gamma_n := 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n \to \gamma,$$

where γ is a constant between 0 and 1. This shows that the partial sums of the divergent harmonic series tend to infinity slowly—about as fast as $\log n$. The constant γ, known as Euler's constant, is rather mysterious: it remains unknown whether γ is rational or irrational.

(d) **A series for** $\log 2$. Let $s_n = 1 - \dfrac{1}{2} + \dfrac{1}{3} + \cdots + (-1)^{n-1}\dfrac{1}{n}$. Then

$$\begin{aligned} s_{2n} &= 1 - \frac{1}{2} + \frac{1}{3} + \cdots - \frac{1}{2n} \\ &= \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2n}\right) - 2\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right) \\ &= (\gamma_{2n} + \log(2n)) - (\gamma_n + \log n) \\ &= \log 2 + \gamma_{2n} - \gamma_n \\ &\to \log 2. \end{aligned}$$

12.14. **Example: exploiting the existence of Euler's constant.** [details omitted from lectures] Consider

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots \qquad \text{and}$$
$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \frac{1}{4} + \ldots,$$

so the second series contains the same terms as the first, but in a different order. We analyse the limiting behaviour of the second series. The terms come in groups of four, three positive terms followed by one negative one, so we first look at the sum of $4n$ terms, in n groups:

$$\begin{aligned} s_{4n} &= \left(1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2}\right) + \left(\frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{6n-5} + \frac{1}{6n-3} + \frac{1}{6n-1} - \frac{1}{2n}\right) \\ &= \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{6n}\right) - \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{6n}\right) - \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right) \\ &= (\gamma_{6n} + \log(6n)) - \frac{1}{2}(\gamma_{3n} + \log(3n)) - \frac{1}{2}(\gamma_n + \log n) \end{aligned}$$

$$= \frac{1}{2}\log\left(\frac{36n^2}{3n^2}\right) + \gamma_{6n} - \frac{1}{2}\gamma_{3n} - \frac{1}{2}\gamma_{2n}$$
$$\to s := \frac{1}{2}\log 12 = \log 2 + \frac{1}{2}\log 3.$$

So far we have only looked at a particular subsequence of (s_n). Not good enough! But it is easy to remedy the omission:

$$s_{4n+1} = s_{4n} + \frac{1}{6n+7} \to s,$$
$$s_{4n+2} = s_{4n+1} + \frac{1}{6n+9} \to s,$$
$$s_{4n+3} = s_{4n+2} + \frac{1}{6n+11} \to s.$$

We deduce that $s_n \to s = \log 2 + \frac{1}{2}\log 3$ (using an extension of the result of Problem sheet 4, Q.2(a)).

So, by changing the order of the terms we have changed the value of the sum! We showed that the alternating series $\sum(-1)^{k-1}1/k$ has sum $\log 2$, whereas the '3 pluses, 1 minus' rearrangement has sum $\log 2 + \frac{1}{2}\log 3$.

12.15. **Definition: rearrangement.** Take any series $\sum a_k$ and let $g\colon \mathbb{N}^{>0} \to \mathbb{N}^{>0}$ be a bijection. Let $b_k = a_{g(k)}$. Then $\sum b_k$ is said to be a **rearrangement** of $\sum a_k$.

Facts about rearrangement

(1) **Theorem**: If $\sum a_k$ is *absolutely convergent* then all rearrangements of $\sum a_k$ converge, and the value of the sum is not changed by rearrangement.

(2) If $\sum a_k$ is convergent but not absolutely convergent, then rearrangement may result in a series converging to a different sum, or to one which diverges. (See, for example, Scott & Tims, Mathematical Analysis, for an example; the idea is quite simple—we aim to rearrange so that the positive terms predominate strongly in the partial sums of the rearranged series.

12.16. **Which test? A miscellany of examples in outline.** In all cases, restriction to tails is permissible or necessary.

(a) $\sum(-1)^{k-1}u_k$, where $u_k = \dfrac{\log(k^2+1)}{\sqrt{k+5}}$.

Apply AST. Consider the derivative of $f(x) = \log(x^2+1)/\sqrt{x+5}$ to show (u_k) is monotonic decreasing and sandwiching techniques to show $u_k \to 0$.

(b) $\sum \dfrac{1}{\sqrt{k^2+k}}$.

Comparison Test, limit form, works easily. (Ratio Test gives no result here.)

(c) $\sum \dfrac{1}{k\log k(\log\log k)^2}$.

A classic case for the Integral Test.

(d) $\sum \dfrac{(4k)!}{(k!)^2}$.

Slog it out with the Ratio Test.

12.17. **Tips for finding counterexamples.**

(1) Sometimes you are given a statement that looks rather like one you've seen before. Ask yourself if all the conditions needed to make the statement true have been included. If not, you're looking for a counterexample which fails to satisfy some missing condition (such as non-negative terms).

(2) Series which serve as counterexamples to plausible assertions often 'only just' diverge or converge. The following can be useful:
- series which diverge even more slowly than $\sum 1/k$: an example is provided by (a tail of) $\sum 1/(k \log k)$;
- alternating series such as $\sum(-1)^{k-1}/\sqrt{k}$—the terms are large in modulus, and (non-absolute) convergence occurs because of the alternating signs.

12.18. **A 'true-or-false?' worked example.** There are many conjectures one might make about convergence of series in general. A few, some true, some false, can be decided even on our limited treatment of series, and limited examples, thus far.

(a) For a_k, b_k real, $a_k \leqslant b_k$ and $\sum b_k$ convergent implies $\sum a_k$ convergent.

FALSE: Consider for example $a_k = -1$, $b_k = 0$. Comparison Test needs **non-negative terms**.

(b) $a_k > 0$ and $ka_k \to 0$ implies $\sum a_k$ converges.

FALSE: Take for example $a_k = 1/((k+1)\log(k+1))$.

(c) If $\sum a_k$ is a convergent series of positive terms then $0 < a_k < 1/k$ for k sufficiently large.

FALSE: Take

$$a_k = \begin{cases} 1/m^2 & \text{if } k = 2^m \text{ for some } m, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\sum a_k$ converges (its partial sum sequence behaves like that for $\sum 1/k^2$). But, for $k = 2^m$,

$$|ka_k| = \frac{2^m}{m^2},$$

and this is not bounded by 1 for large k—in fact it tends to infinity, as 8.9(b) shows.

Morals from (c):

(1) Series with gaps can be useful for counterexamples.

(2) A series $\sum 1/k^p$ does not provide a counterexample here. It is naive to think that behaviour of series of this special form is typical.

12.19. **Postscript: other tests for convergence.** Not every series you may meet can be handled by applying directly one of the tests presented above. In particular you should be aware that the Ratio Test is rather crude and quite often fails to give a result.

The list of tests for convergence, or for absolute convergence, that we have given is far from exhaustive, and many other convergence tests exist. We note in particular

- **Cauchy's nth Root Test** and **Raabe's Test**—both useful alternatives to the Ratio Test, the first because it works neatly on many power series (see Section 13) and the second because it provides a backstop to the Ratio Test, giving a result in a number of cases where the Ratio Test does not;
- **Cauchy's Condensation Test** Under the same conditions on f as in the Integral Test, the test asserts that $\sum f(k)$ converges if and only if $\sum 2^m f(2^m)$ converges.

- **Abel's Test** and **Dirichlet's Test**, for conditional convergence.

Statements and proofs can be found in various analysis textbooks and no doubt on the internet too.

You will not encounter series in Prelims for which any one of these additional tests is needed, but they could nevertheless provide you with techniques you might sometimes find useful, now or later.

---

13. Power Series

13.1. **Power series.** A series of the form

$$\sum c_k x^k, \quad \text{where } c_k \ (k \geqslant 0) \text{ are real constants and } x \in \mathbb{R},$$

is a **real power series**. Likewise a series

$$\sum c_k z^k, \quad \text{where } c_k \ (k \geqslant 0) \text{ are complex constants and } z \in \mathbb{C},$$

is a **complex power series**. Where subsequently we are able to handle the complex case we shall do so; this subsumes the real case.

Mostly we shall be concerned with the real case, but the theory of convergence works in the same way for complex power series, and even for elementary applications it is sometimes convenient to venture into the complex plane.

We think of x (or z) as a variable, and want to investigate the functions we can define by power series.

Note that a real or complex power series always converges at 0. So we can exclude consideration of $x = 0$ (or $z = 0$) when testing for convergence.

13.2. **Elementary functions: exponential, trigonometric and hyperbolic functions.**

- **Exponential:** Consider $\sum \frac{z^k}{k!}$. By convention, here and elsewhere, $0! = 1$. When $z \neq 0$ we can test for absolute convergence using (the corollary to) the Ratio Test:

  $$\left| \frac{z^{k+1}/(k+1)!}{z^k/k!} \right| = \frac{|z|}{k+1} \to 0 < 1 \quad \text{as } k \to \infty.$$

  Hence the exponential series converges absolutely, and so converges, *for all* $z \in \mathbb{C}$ and we **define**

  $$\mathrm{e}^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

  This is also written $\exp(z)$, when convenient.

- **Sine and cosine:** Consider $\sum (-1)^k \frac{z^{2k+1}}{(2k+1)!}$ and $\sum (-1)^k \frac{z^{2k}}{(2k)!}$. Applying the Ratio Test (for $z \neq 0$) we can prove (do it for yourself!) that each of these series converges absolutely for all $z \in \mathbb{C}$ and we **define**

  $$\sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \quad \text{and} \quad \cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.$$

- **Sinh and cosh:** We define
$$\sinh z = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} \quad \text{and} \quad \cosh z = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}$$
—noting that we make the definitions legitimate by using the Ratio Test to prove that the associated power series converge absolutely for all $z \in \mathbb{C}$.

Note: We would not expect to be able to define tan, cot, sec and cosec, or their hyperbolic analogues, as power series converging absolutely for all $z \in \mathbb{C}$.

We'd like to work our way towards showing that, for a real variable anyway, these series definitions do capture the properties we expect of the functions familiar from elementary mathematics. We begin with some elementary facts.

13.3. **AOL for series in general and power series in particular.** Recall the key fact a series $\sum a_k$ converges if and only if its partial sum sequence (s_n) converges. Given two series $\sum a_k$ and $\sum b_k$ with partial sum sequences (s_n) and (t_n), we have
$$(a_1 + b_1) + \cdots + (a_n + b_n) = (a_1 + \cdots + a_n) + (b_1 + \cdots + b_n) = s_n + t_n$$
(by properties of arithmetic—only a FINITE number of terms involved!). Hence, if $\sum a_k$ converges to s and $\sum b_k$ converges to t, then $\sum(a_k+b_k)$ converges, to $s+t$. Similar arguments apply to $\sum(a_k - b_k)$ and to $\sum ca_k$, where c is a constant.

It follows in particular that we can add/subtract and scalar-multiply convergent power series in the expected 'term-by-term' way.

Examples of elementary connections

$$\cos z = \frac{1}{2}(\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}), \qquad \cosh z = \frac{1}{2}(\mathrm{e}^{z} + \mathrm{e}^{-z}),$$
$$\sin z = \frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}), \qquad \sinh z = \frac{1}{2}(\mathrm{e}^{z} - \mathrm{e}^{-z}),$$

$$\mathrm{e}^{\mathrm{i}z} = \cos z + \mathrm{i}\sin z.$$

One can also derive, straight from the definitions, the familiar Osborn's Rules linking cos and cosh and linking sin and sinh. For example, $\cos \mathrm{i}z = \cosh z$, and this is valid for complex z and not just when z is a real number.

13.4. **The radius of convergence of a power series.**

A power series which fails to converge is useless for defining a function. So it's critically important to know when power series do converge.

We define the **radius of convergence** of $\sum c_k z^k$ to be R, where
$$R = \begin{cases} \sup\{\,|z| \in \mathbb{R} \mid \sum |c_k z^k| \text{ converges}\,\} & \text{if the sup exists,} \\ \infty & \text{otherwise.} \end{cases}$$

(Remember that $\sup E$ (for $E \subseteq \mathbb{R}$) exists if and only if E is non-empty and bounded above. Since $\sum |c_k 0^k|$ converges, non-emptyness is not at issue here.)

13.5. **Proposition on radius of convergence** (*technical but important*)**.** Let $\sum c_k z^k$ be a power series with radius of convergence $R\,(>0)$. Then

(i) $\sum |c_k z^k|$ converges for $|z| < R$, and hence $\sum c_k z^k$ converges for $|z| < R$.

(ii) $\sum c_k z^k$ diverges if $|z| > R$.

Proof. Assume $R \neq \infty$ (minor adaptation needed for the case $R = \infty$).

(i) Fix z with $|z| < R$ and pick S such that $|z| < S < R$. Then $\varepsilon = R - S > 0$ and by Approximation Property for supremum we can find ρ with $R - \varepsilon = S < \rho < R$ such that $\sum |c_k \rho^k|$ converges. But since $|z| < \rho$, this implies $\sum |c_k z^k|$ converges, by the simple Comparison Test.

(ii) Assume for contradiction that there exists z with $|z| > R$ such that $\sum c_k z^k$ converges. Then $c_k z^k \to 0$. Therefore (because a convergent sequence is bounded) there exists a constant M such that $|c_k z^k| \leqslant M$. Pick ρ with $|z| > \rho > R$. Then

$$0 \leqslant |c_k \rho^k| = |c_k z^k| \left|\frac{\rho}{z}\right|^k \leqslant M \left|\frac{\rho}{z}\right|^k.$$

But $\sum |\rho/z|^k$ is convergent, since $|\rho/z| < 1$. So $\sum |c_k \rho^k|$ converges by comparison, and this contradicts the definition of R. □

Notes

- Part (i) of the proposition is not just the definition of R.
- Some sources replace $\sum |c_k z^k|$ by $\sum c_k z^k$ in the definition of R. The proposition implies that the two versions are equivalent. We prefer to phrase the definition with moduli in, as a reminder that we can find R by using tests for convergence of series of non-negative terms.
- We refer to $\{\, z \in \mathbb{C} \mid |z| < R \,\}$ as the **disc of convergence**. For real power series we have an **interval of convergence**.
- If $\sum c_k x^k$ is a real power series with radius of convergence R then the series may converge or diverge at R and $-R$, and similarly for $|z| = R$ in the complex case.

13.6. **Examples: radius of convergence.**

(a) The exponential, trigonometric and hyperbolic series all have $R = \infty$. Proof: use Ratio Test to test for absolute convergence.

(b) The geometric series $\sum x^k$ has $R = 1$ (from Example 11.3).

(c) Consider $\sum c_k x^k$ where $c_k = \dfrac{k!}{k^k}$. We apply the Ratio Test to test for absolute convergence, for $x \neq 0$.

$$\lim \left| \frac{(k+1)!x^{k+1}/(k+1)^{k+1}}{k!x^k/k^k} \right| = \lim \left(1 + \frac{1}{k}\right)^{-k} |x| = \mathrm{e}^{-1}|x|.$$

Hence $\sum |c_k x^k|$ converges for $|x| < \mathrm{e}$ and diverges for $|x| > \mathrm{e}$, so $R = \mathrm{e}$.

(d) Consider $\sum c_k$ where $c_k = 1$ if k is prime and $c_k = 0$ otherwise. We cannot use the Ratio Test here. Note that $c_k x^k \not\to 0$ if $|x| \geqslant 1$ (because there are infinitely many primes). Hence $R \leqslant 1$. For $|x| < 1$, consider $\sum_{r \geqslant 1} x^{p_r}$, where $p_1 < p_2 < p_3 < \ldots$ is an enumeration of the primes. Then $r \leqslant p_r$, so $|x^{p_r}| \leqslant |x^r|$ and so $\sum |x^{p_r}|$ converges by comparison with $\sum |x^r|$. Therefore $R = 1$.

(e) Problem sheet 6, Q. 4 is set as an exercise on convergence tests in the context of real power series. With the definition of R in place it can be seen as asking you to calculate R for various power series. Working through this exercise is intended to provide practice in finding radius of convergence and to serve also to illustrate the properties of power series captured by Proposition 13.5.

Important notes

- In determining the radius of convergence of $\sum c_k x^k$ it is not correct to say, for example, '$\sum |c_k x^k|$ converges for $|x| < 3$, and hence $R = 3$.' Convergence for $|x| < 3$ only implies $R \geqslant 3$. You need to say where the series diverges as well as where it converges. (Note the way Problem sheet 6, Q. 4 is worded—to prevent you making this common beginners' mistake.)

- The last example shows that the Ratio Test cannot necessarily be used to find R because the limit one needs to consider may not exist. *More advanced texts give a formula for R involving* $\liminf |c_{k+1}/c_k|$. *Here* $\liminf$ *is a notion which can sometimes be exploited in place of* $\lim$ *when the latter fails to exist. This formula is not needed for Prelims, and is frequently abused. DO NOT CLAIM that* $R = 1/\lim |c_{k+1}/c_k|$ *gives R for a general power series.*

If we are to make good use of power series **as functions** we would like to know they have good behaviour, in particular that it is legitimate to differentiate them. The following Differentiation Theorem is very important, but a first-principles proof is technical (you'll see this in Analysis II) and a better proof for real power series uses integration theory (Analysis III). So we present the theorem here without proof.

13.7. **Differentiation Theorem for (real) power series.** Let $\sum c_k x^k$ be a real power series and assume that the series has radius of convergence R where $0 < R \leqslant \infty$. Let

$$f(x) := \sum_{k=0}^{\infty} c_k x^k \qquad (|x| < R).$$

Then $f(x)$ is well defined for each x with $|x| < R$ and moreover the derivative $f'(x)$ exists for each such x and is given by

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^{\infty} c_k x^k \overset{*}{=} \sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\left(c_k x^k\right) = \sum_{k=1}^{\infty} k c_k x^{k-1}.$$

In words, we say that $f'(x)$ is obtained by **term-by-term differentiation**.

Note: The theorem is very powerful and it is far from obvious that it is true. Note that differentiation is carried out by a limiting process, that summing an infinite series also involves a limiting process, and that iterated limits may not commute (recall Problem sheet 4, Points to Ponder B.). Saying we can differentiate term-by-term is exactly saying that we can interchange the order of two limits: in $*$ above

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^{\infty} c_k x^k \quad \text{and} \quad \sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\left(c_k x^k\right)$$

we have iterated limits taken in different orders.

13.8. **Some applications of the (real) Differentiation Theorem.** The general idea here is to use the Differentiation Theorem to derive suitable differential equations whose solution leads to formulae connecting functions defined by power series.

(a) Because the power series defining the functions are absolutely convergent for all real x, from above, the Differentiation Theorem implies that the derivatives below exist and are given by the expected formulae:

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x,$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\sinh x = \cosh x,$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos x = -\sin x, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\cosh x = \sinh x.$$

To illustrate:

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{k=0}^{\infty}\frac{x^k}{k!} \stackrel{*}{=} \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^k}{k!} = \sum_{k=1}^{\infty}\frac{kx^{k-1}}{k!} = \sum_{m=0}^{\infty}\frac{x^m}{m!} = \mathrm{e}^x.$$

Here the equality marked $*$ holds by the Differentiation Theorem.

(b) To prove $\sin^2 x + \cos^2 x = 1$ for all $x \in \mathbb{R}$.

Proof. Let $h(x) = \sin^2 x + \cos^2 x$. Then by the usual rules for differentiation (proved in Analysis II) $h'(x)$ exists for all x and

$$h'(x) = 2\sin x\cos x - 2\sin x\cos x = 0.$$

Hence $h(x)$ is a constant, A say (see Analysis II for justification). Putting $x = 0$, we get $A = 1$ and the result follows. □

We now consider examples in which we have formulae involving two variables.

(c) $\mathrm{e}^{a+b} = \mathrm{e}^a\mathrm{e}^b$ for $a, b \in \mathbb{R}$.

Proof. We let c be a real constant and define a function g on $\mathbb{R}$ by

$$g(x) = \mathrm{e}^x\mathrm{e}^{c-x}.$$

Then, differentiating (using the product rule and chain rule, proved in Analysis II),

$$g'(x) = \mathrm{e}^x\mathrm{e}^{c-x} - \mathrm{e}^x\mathrm{e}^{c-x} = 0.$$

It follows (see Analysis II for justification) that $g(x) = A$, where A is a constant, which will depend on c. But $A = g(0) = \mathrm{e}^c$, since $\mathrm{e}^0 = 1$. We have shown that $\mathrm{e}^c = \mathrm{e}^x\mathrm{e}^{c-x}$ for all $x, c \in \mathbb{R}$, Now put $x = a$, $c = a + b$ to get

$$\mathrm{e}^{a+b} = \mathrm{e}^a\mathrm{e}^b.$$ □

As a special case of the addition formula we get $\mathrm{e}^x\mathrm{e}^{-x} = e^0 = 1$. Certainly $\mathrm{e}^x > 0$ when $x > 0$. We now see that $\mathrm{e}^x > 0$ for all $x \in \mathbb{R}$.

(d) Addition formulae for the trigonometric and hyperbolic functions are proved by adapting the strategy used for (c).

Note: You might be tempted to try to prove (b) by taking the power series for $\sin x$ and squaring it, and likewise for $\cos x$. Not recommended! Justifying multiplying power series in the same way as one would multiply polynomials requires serious work. After all, you've seen that infinite sums do not necessarily behave the same way as regards arithmetic as finite sums do—beware $\cdots$! (See Problem sheet 7, Point to Ponder A. for more on this.)

13.9. **Trigonometric functions: what became of π.** You will probably have met the cosine and sine series before, as Maclaurin expansions, but you will have been introduced to the functions $\cos x$ and $\sin x$ geometrically. We have turned things around and used the series to *define* these functions. One can capture the number known as π in either of the following (equivalent) ways:

- π is the smallest number $x > 0$ for which $\sin x = 0$;
- $\pi/2$ is the smallest number $x > 0$ for which $\cos x = 0$.

[In Analysis II you'll discover why these smallest zeros exist.] One can then use the addition formulae to prove the periodicity results

$$\cos(x + 2\pi) = \cos x, \qquad \sin(x + 2\pi) = \sin x.$$

13.10. **Looking ahead: the importance of complex power series.** We have worked with complex power series where the arguments are the same as for the real case. We have restricted consideration of differentiation to the real case, because you have only learned about differentiation of functions of a real variable so far. In the second year you will study **complex analysis**, 'complex' meaning 'in $\mathbb{C}$' and not 'more complicated'.

One fact worth noting, and quite easy to prove from the definitions, is that

$$\mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^{x}(\cos y + \mathrm{i}\sin y) \qquad (\text{for } x, y \in \mathbb{R}).$$

As a corollary: the facts about π then give

$$\boxed{\mathrm{e}^{2\pi\mathrm{i}} = 1.}$$

[The remaining examples are on the boundary between Analysis I and Analysis II and III, and won't be covered in lectures. They provide additional illustrations of the use of differentiation of power series to obtain valuable information about important functions.]

13.11. **The binomial expansion.** Here we shall freely use arbitrary powers of real numbers and formulae for their derivatives, noting that such powers are defined using exponentials (and logs).

Let $\alpha \in \mathbb{R} \setminus \mathbb{N}$. (The case that $\alpha \in \mathbb{N}$ was covered in Introduction to University Mathematics, and doesn't involve convergence issues.) Consider the infinite sum

$$B_\alpha(x) = \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} x^k.$$

We apply the Ratio Test to show the series has radius of convergence 1:

$$\lim \left| \frac{\alpha(\alpha-1)\dots(\alpha-k)x^{k+1}/(k+1)!}{\alpha(\alpha-1)\dots(\alpha-k+1)x^k/k!} \right| = \lim \left| \frac{\alpha-k}{k+1} \right| |x| = |x|.$$

The Differentiation Theorem tells us we can differentiate term-by-term with respect to x for $|x| < 1$, This gives

$$\begin{aligned} B'_\alpha(x) &= \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{(k-1)!} x^{k-1} \\ &= \sum_{r=0}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-r)}{r!} x^r \qquad (\text{writing } k = r+1). \end{aligned}$$

Now multiply by $(1+x)$, treating x as constant, and collect together coefficients associated with the same power of x. We get

$$(1+x)B'_\alpha(x) = \sum_{s=0}^{\infty}\left(\frac{\alpha(\alpha-1)\dots(\alpha-s)}{s!} + \frac{\alpha(\alpha-1)\dots(\alpha-s+1)}{(s-1)!}\right)x^s$$

$$= \sum_{s=0}^{\infty}\frac{\alpha(\alpha-1)\dots(\alpha-s+1)}{s!}\left((\alpha-s)+s\right)x^s.$$

Therefore we have a differential equation for B_α:

$$(1+x)B'_\alpha(x) = \alpha B_\alpha(x).$$

Solving, we get $B_\alpha(x) = C(1+x)^\alpha$, where C is a constant (see Analysis II for justification). Putting $x=0$ we get $C=1$.

13.12. **The logarithmic series.**

Consider

$$L(x) = \sum_{n=1}^{\infty}\frac{x^k}{k} = x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots .$$

The series $\sum \frac{x^k}{k}$ converges for $|x| < 1$ and diverges for $|x| < 1$ (easy application of Ratio Test). By the Differentiation Theorem, we can differentiate term-by-term for $|x| < 1$ to get

$$L'(x) = 1 + x + x^2 + \dots = \frac{1}{1-x}.$$

For $-1 < x < 1$, define

$$M(x) = (1-x)\exp(L(x)).$$

Computing $M'(x)$ by the chain and product rules (proved in Analysis II), we get

$$M'(x) = (-1)\exp(L(x)) + (1-x)L'(x)\exp(L(x)) = 0.$$

Hence $M(x)$ is constant, equal to $M(0) = 1$ (see Analysis II for justification). Hence

$$\exp(L(x)) = \frac{1}{1-x}. \qquad (-1 < x < 1).$$

So, from the addition formula for the exponential function, we have $\exp(-L(x)) = 1 - x$. Let us define a function log to satisfy

$$-\log(1-x) = \log\left(\frac{1}{1-x}\right) = L(x);\ \text{i.e.}\ \log(1+x) = -L(-x).$$

This then gives us a log function acting as the inverse to the exponential function, and a series expansion for $\log(1+x)$, valid for $-1 < x < 1$. All this is in line with what we want and expect the logarithm to do. The full picture will emerge in Analysis II (which considers inverse functions) and Analysis III (which considers the definition of the log function by an integral).