TOBIAS HOLCK COLDING

Lecture 1

Two key topics for this class:
- How to write a mathematical proof.
- How to prove theorems.

Here is an example:

Intermediate value theorem:
- Suppose that $f : [a, b] \to \mathbf{R}$ is a continuous functions.
- Assume that $f(a) < 0$ and $f(b) > 0$.

The intermediate value theorem says that there exists a c between a and b where $f(c) = 0$.

- How do we prove this?
- If we draw a picture, then it seems obvious, but how to we actually prove this?

.

That a function is continuous basically means that when you draw the graph of the function the pencil is not allowed to leave the paper.

- How do we make this into a proper proof?

- What properties of the real values are needed for a proof?

This leads to several questions:
- Q1: What is a real number?
- Q2: Why is $\sqrt{2}$ a real number?
- Q3: What is $\sqrt{2}$?

The answer to these questions: $\mathbf{R}$ is a complete ordered field that contains the rational numbers $\mathbf{Q}$.

Here is some notation:

(1) $\mathbf{N}$ is the natural numbers. This means that $\mathbf{N} = \{1, 2, 3, \cdots\}$.
(2) $\mathbf{Z}$ is the integers. This means that $\mathbf{Z} = \{0, \pm 1, \pm 2, \pm 3, \cdots\}$.
(3) $\mathbf{Q}$ is the rational numbers. So all numbers of the form $\frac{m}{n}$, where $m \in \mathbf{Z}$ and $n \in \mathbf{N}$.

Properties:

Rational numbers:
Rational numbers $\mathbf{Q}$ are numbers of the form $\frac{m}{n}$, where $m \in \mathbf{Z}$ and $n \in \mathbf{N}$.

(1) When are two numbers the same?
$$\frac{m_1}{n_1} = \frac{m_2}{n_2} \iff m_1 \, n_2 = m_2 \, n_1 \, .$$

(2) How do we add two numbers?
$$\frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_1 \, n_2 + m_2 \, n_1}{n_1 \, n_2} \, .$$

(3) How do we multiply two numbers?
$$\frac{m_1}{n_1} \frac{m_2}{n_2} = \frac{m_1 \, m_2}{n_1 \, n_2} \, .$$

(4) When is one number less than another?
$$\frac{m_1}{n_1} < \frac{m_2}{n_2} \iff m_1 \, n_2 < m_2 \, n_1 \, .$$

For this to make sense we need (for instance) to show that multiplication is well-defined:

This means that if we have two representations of the same rational number
$$\frac{m_1}{n_1} = \frac{m_2}{n_2}$$
and likewise
$$\frac{p_1}{q_1} = \frac{p_2}{q_2} \, ,$$
then
$$\frac{m_1}{n_1} \frac{p_1}{q_1} = \frac{m_2}{n_2} \frac{p_2}{q_2} \, .$$

Proof. We have that $m_1 \, n_2 = m_2 \, n_1$ and $p_1 \, q_2 = p_2 \, q_1$. Therefore,
$$m_1 \, p_1 \, n_2 \, q_2 = m_1 \, n_2 \, p_1 \, q_2 = m_2 \, n_1 \, p_2 \, q_1 \, .$$

$\square$

This illustrate how detailed a proof should be.

A Field:

Definition:
A Field $\mathbf{F}$ is a set with two operations that we are denoting suggestively by "+" and "$\cdot$". Those two operations satisfies the following axioms:

Additive properties:

(1) $x, y \in \mathbf{F}$, then $x + y \in \mathbf{F}$.
(2) $x + y = y + x$.
(3) $(x + y) + z = x + (y + z)$.
(4) There exists an element $0 \in \mathbf{F}$ such that $0 + x = x$ for all $x \in \mathbf{F}$.
(5) For all $x \in \mathbf{F}$ there exists an element, suggestively, denoted by $(-x)$ such that $x + (-x) = 0$.

Multiplicative properties:

(1) $x, y \in \mathbf{F}$, then $x\,y \in \mathbf{F}$.
(2) $x\,y = y\,x$.
(3) $(x\,y)\,z = x\,(y\,z)$.
(4) There exists an element, suggestively, denoted by 1 such that $1\,x = x$ for all $x \in \mathbf{F}$.
(5) For all $x \in \mathbf{F} \setminus \{0\}$ there exists an element, suggestively, denoted by $\frac{1}{x}$ such that $x\,\frac{1}{x} = 1$.

The final axion that we need is an axiom that chains addition and multiplication together:

-
$$(x + y)\,z = x\,z + y\,z\,.$$

Theorem: For any field 'zero' is unique.

Proof. Suppose there are two. Let us denote them by 0_1 and 0_2. Then
$$0_1 + 0_2 = 0_2$$
since 0_1 is a 'zero' and
$$0_1 + 0_2 = 0_1$$
since 0_2 is a 'zero' so $0_1 = 0_2$. $\qquad\square$

Examples: $\mathbf{Q}$ is a Field, whereas $\mathbf{N}$ and $\mathbf{Z}$ are not Fields.

Ordered set: An ordered set $\mathbf{S}$ is a set with a relation $<$ with the following properties:
(1) For an x, $y \in \mathbf{S}$, one of the following holds: $x < y$ or $y < x$ or $x = y$.
(2) If x, y, $z \in \mathbf{S}$ with $x < y$ and $y < z$, then $x < z$.

Ordered Field:
An ordered Field is an ordered set that is also Field and has the following two additional properties that chains the operations in the Field together with the ordering:
(1) If $x < y$, then $x + z < y + z$.
(2) If $x > 0$ and $y > 0$, then $x\,y > 0$.

Example: $\mathbf{Q}$ is an ordered Field.

Theorem: If $x < y$ and $z > 0$, then $x\,z < y\,z$.

Proof. We need to show that $x\,z < y\,z$ or equivalently $y\,z - x\,z > 0$. The latter can be rewritten as $y\,z - x\,z = (y - x)\,z$. Since $y > x$ we have that $y - x > 0$ and the claim therefore follows since $z > 0$. $\qquad\square$

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, DEPT. OF MATH., 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 2

The real numbers $\mathbf{R}$ is a complete ordered Field that contains $\mathbf{Q}$.

Question: What is the difference between $\mathbf{R}$ and $\mathbf{Q}$?
One difference is that $\mathbf{R}$ contains $\sqrt{2}$ and $\mathbf{Q}$ does not.

$\sqrt{2}$ is a number x so that $x > 0$ and $x^2 = 2$.

Theorem: There does not exists a rational number x so that $x^2 = 2$.

Proof. We will argue by contradiction. So suppose that there exists a rational number $x = \frac{m}{n}$, where $m \in \mathbf{Z}$ and $n \in \mathbf{N}$, so that $x^2 = \frac{m^2}{n^2} = 2$. We can assume m and n does not have a common factor (other than one). We have that $m^2 = 2\,n^2$ and so 2 is a factor in m^2 and therefore in m itself. This means that $m = 2\,m_1$, where m_1 is also an integer. It follows that $m^2 = 4\,m_1^2 = 2\,n^2$ and therefore $2\,m_1 = n$ and so n is also even. We have now that both m and n are even and so have 2 as a common factor. This is the desired contradiction. This show that there is no rational number x with the property that $x^2 = 2$. $\square$

How do we add $\sqrt{2}$ to the number system?

$\sqrt{2} = 1.4142136\cdots$

So 1, 1.4, 1.41, 1.414, 1.4142, 1.41421, $\rightarrow \sqrt{2}$.

$\sqrt{2}$ is the limit of a sequence of numbers.

Completeness of **R**. (Least upper bound property.)

Completeness is: If a subset A of **R** has an upper bound, then A has a least upper bound.

Suppose that **S** is an ordered set and A is a subset of **S**, then M is an upper bound for A if for all $a \in A$ we have that $a \leq M$.

Example: If $A = \{1, 2, 3\} \subset \mathbf{Z}$, then 4 is an upper bound, whereas 2 is not an upper bound.

Example: If $\mathbf{S} = \mathbf{Q}$, then **N** as a subset does not have an upper bound (we will return to this shortly).

Least upper bound: Suppose that **S** is an ordered set and A is a subset that has an upper bound. We say that M is a least upper bound for A if M is an upper bound for A and for any other upper bound M_1 we have that $M \leq M_1$.

Complete ordered set: We say that an ordered set is complete if any subset that has an upper bound has a least upper bound.

Theorem: There exists a complete ordered Field that contains **Q**.

This Field is denoted by **R**.

   We will not prove this, as a proof would take us too far a field, rather we will take it for granted.

Theorem: $\sqrt{2} \in \mathbf{R}$.

Proof. Let $A = (0, \sqrt{2}) \cap \mathbf{Q}$. That is A consists of all the positive rational numbers a so that $a^2 < 2$. Let x be the least upper bound for A. Note that A is nonempty (since $1 \in A$) and that 2 is an upper bound for A. Note also that $x \geq 1 > 0$ since it is an upper bound. We need to show that $x^2 = 2$.

   We will first show that $x^2 \leq 2$. Suppose not; so assume that $x^2 > 2$. We will show that this leads to a contradiction. Consider
$$(x - h)^2 = x^2 - 2\,x\,h + h^2 > x^2 - 2\,h\,x\,.$$
As long as $h > 0$ is chosen so that
$$2\,h\,x < x^2 - 2$$

or, equivalently, that

$$h < \frac{x^2 - 2}{2\,x}$$

then

$$(x - h)^2 > 2$$

and therefore $x - h$ is also an upper bound for A. This contradict that x is the least upper bound. We therefore have that if x is the least upper bound for A, then $x^2 \leq 2$.

To show the reverse inequality (that $x^2 \geq 2$) we argue similarly. Assume that for the least upper bound x we have that $x^2 < 2$. Consider $x + h$, where $0 < h < 1$. We have that

$$(x + h)^2 = x^2 + 2\,x\,h + h^2 < x^2 + 2\,x\,h + h = x^2 + h\,(2\,x + 1)\,.$$

Since we are assuming that $x^2 < 2$ we can choose h positive so that

$$h < \frac{2 - x^2}{2\,x + 1}\,.$$

We therefore have that

$$(x + h)^2 < x^2 + 2 - x^2 < 2\,.$$

This is the desired contradiction and show that $x^2 \geq 2$. Together with the first step we have that $x^2 = 2$. $\qquad\square$

Corollary: Q is not complete.

Proof. If **Q** was complete, then $\sqrt{2} \in \mathbf{Q}$ but we have already proven that there is no rational number with the property that $x^2 = 2$. $\qquad\square$

Archimedean property: For all $x \in \mathbf{R}$, there exists a natural $n \in \mathbf{N}$ so that $x < n$.

Proof. If this was not the case, then **N** would be bounded. To see that **N** is not bounded we argue as follows. Assume it is bounded and let α be the least upper bound for **N**. We would now have that for all $n \in \mathbf{N}$ that $n \leq \alpha$. Since $n + 1$ is also a natural number we would have that $n + 1 \leq \alpha$ as well. So, in fact, $n \leq \alpha - 1$ or in other words, since n was any natural number, $\alpha - 1$ would be an upper bound contradicting that α was the least upper bound. $\qquad\square$

As a corollary of the Archimedean property we get the following:

Corollary: If $x < y$, then there exists a rational number $\frac{m}{n}$ such that

$$x < \frac{m}{n} < y\,.$$

Proof. Set $\beta = \frac{1}{y-x}$. From the Archimedean property we have that there exists a natural number n with $n > \beta$. It follows that

$$0 < \frac{1}{n} < \frac{1}{\beta}\,.$$

Now let $m - 1$ be the largest integer so that

$$m - 1 \leq x\,n\,.$$

It follows that $\frac{m}{n}$ has the desired property.   □

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
     TBB can be downloaded at:
     https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
     (screen-optimized)
     https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
     (print-optimized)

MIT, DEPT. OF MATH., 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 3

Theorem: $\mathbf{R}$ is a complete ordered Field that contains $\mathbf{Q}$.

$\mathbf{S}$ is an ordered set. A non-empty subset A of $\mathbf{S}$ is said to have an upper bound if there exists an $M \in \mathbf{S}$ such that for all $a \in A$ we have that $a \leq M$.

Completeness is the property that every bounded non-empty subset has a least upper bound.

We denote by $\sup A$ the smallest upper bound of A.

Lower bound: A non-empty subset A is said to have a lower bound if there exists $m \in \mathbf{S}$ such that for all $a \in A$ we have that $m \leq a$.

The greatest lower bound is a lower bound that is greater or equal to all other lower bounds.

The greatest lower bound is denoted by $\inf A$.

From now on we will concentrate of the case of $\mathbf{R}$.

How to write a mathematical proof?

This lecture we will look at how to write a mathematical proof.
We will explain this in two results that we talked about last time.

Let us return to the example of showing that $\sqrt{2} \in \mathbf{R}$.

We already showed this but we did not write it as a "proper proof". That is what we will
do next.

Theorem: There exists $\alpha > 0$ such that $\alpha^2 = 2$.

Proof. Define a set A by

$$A = \left\{ x \in \mathbf{R} \,|\, x > 0 \text{ and } x^2 \leq 2 \right\}.$$

We will show that A is a non-empty bounded subset and that $\alpha = \sup A$ has the property
that $\alpha > 0$ and $\alpha^2 = 2$.

   Observe first that $1 \in A$, so A is non-empty. Moreover, 2 is an upper bound for A so A is
bounded from above. Let $\alpha = \sup A$, we need to show that $\alpha > 0$ and that $\alpha^2 = 2$. Since
$1 \in A$ it follows that $0 < 1 < \alpha$. To show that $\alpha^2 = 2$ we divide the proof into two parts.

   Part 1: We will show that $\alpha^2 \leq 2$. Suppose not; we will see that this lead to a contra-
diction. Indeed, we will show that that if this was the case, then there exists an $0 < \alpha_0 < \alpha$
such that $\alpha_0^2 > 2$ so α_0 is an upper bound that is smaller than α. To show this we set

$$h = \frac{\alpha^2 - 2}{4\,\alpha}$$

and set

$$\alpha_0 = \alpha - h\,.$$

Note that since we are assuming that $\alpha > 2$, then we have that $h > 0$ and therefore $\alpha_0 < \alpha$.
Note also that since $1 \leq \alpha \leq 2$ we have that

$$h \leq \frac{1}{2\,\alpha} \leq \frac{1}{2}\,.$$

In particular, $0 < \alpha_0$. Next

$$\alpha_0^2 = \alpha^2 + h^2 - 2\,h\,\alpha > \alpha^2 - \frac{\alpha^2 - 2}{2} = \frac{\alpha^2}{2} + 1 \geq 2\,.$$

This is the desired contradiction and show that $\alpha^2 \leq 2$.

Part 2: We will next show that $\alpha^2 \geq 2$. Suppose not; we will see that this lead to a contradiction. Indeed, we will show that if this was the case, then there exists an $\alpha_1 > \alpha$ such that $\alpha_1^2 < 2$ contradicting that α was an upper bound for A. So assume that $\alpha^2 < 2$. This time we will set

$$h = \frac{2 - \alpha^2}{4\,\alpha}\,.$$

Note that $1 > h > 0$ (the first inequality follows from that $1 \leq \alpha$). Set $\alpha_1 = \alpha + h$. It follows that

$$\alpha_1^2 = \alpha^2 + h^2 + 2\,h\,\alpha < \alpha^2 + h + \frac{2 - \alpha^2}{2} \leq \alpha^2 + 2\frac{2 - \alpha^2}{2} = 2\,.$$

Together parts 1 and 2 show that $\alpha^2 = 2$; completing the proof. $\square$

Archimedean property:

Formal proof:

Theorem: The set of natural number is not bounded from above.

Proof. If **N** is bounded from above, then we can let M be the least upper bound. We now have that for all $n \in \mathbf{N}$
$$n \leq \alpha$$
We claim that also $\alpha - 1$ is an upper bound contradicting that α was the least upper bound. Namely, for a given n since α is an upper bound for all natural numbers we have that
$$n + 1 \leq \alpha$$
but this implies that
$$n \leq \alpha - 1$$
showing that $\alpha - 1$ is an upper bound. That is the desired contradiction.   □

Corollary: For any $\epsilon > 0$, there exists an $n \in \mathbf{N}$ such that $\frac{1}{n} < \epsilon$.

Proof. Set $\alpha = \frac{1}{\epsilon}$, By the Archimedean property we know that there exists an $n \in \mathbf{N}$ with $n > \alpha$. It follows that $\frac{1}{n} < \epsilon$.   □

Sequences:

$\sqrt{2}$ can be thought of a limit of a sequence of decimal numbers as follows.
$$1 < 1.4 < 1.41 < 1.414 \cdots .$$

When does a limit exist?

A sequence of real numbers is a function $f : \mathbf{N} \to \mathbf{R}$.

We usually use the notation $a_n = f(n)$.

Example 1: $\sqrt{2}$ is the limit of $a_1 = 1$, $a_2 = 1.4$, $a_3 = 1.41$, $a_4 = 1.414$ etc.

Example 2: $a_n = (-1)^n$. This sequence has NO limit. The a_n's alternates between -1 and 1.

Example 3: The sequence $a_n = \frac{1}{n}$ has zero as its limit.

Limit: Let a_n be a sequence and a a real number. We say that a_n converges to a if for all $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that if $n \geq N$, then
$$|a_n - a| < \epsilon .$$

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 4

Theorem: $\mathbf{R}$ is a complete ordered Field that contains $\mathbf{Q}$.

$\mathbf{S}$ is an ordered set. A non-empty subset A of $\mathbf{S}$ is said to have an upper bound if there exists an $M \in \mathbf{S}$ such that for all $a \in A$ we have that $a \leq M$.

Completeness is the property that every bounded non-empty subset has a least upper bound.

We denote by $\sup A$ the greatest lower bound of A.

Lower bound: A non-empty subset A is said to have a lower bound if there exists $m \in \mathbf{S}$ such that for all $a \in A$ we have that $m \leq a$.

The greatest lower bound is a lower bound that is greater or equal to all other lower bounds.

Sequences:

$\sqrt{2}$ can be thought of a limit of a sequence of decimal numbers as follows.

$$1 < 1.4 < 1.41 < 1.414 \cdots .$$

A sequence of real numbers is a function $f : \mathbf{N} \to \mathbf{R}$.

We usually use the notation $a_n = f(n)$.

Example 1: $\sqrt{2}$ is the limit of $a_1 = 1$, $a_2 = 1.4$, $a_3 = 1.41$, $a_4 = 1.414$ etc.

Example 2: $a_n = (-1)^n$. This sequence has NO limit. The a_n's alternates between -1 and 1.

Example 3: The sequence $a_n = \frac{1}{n}$ has zero as its limit.

Limit: Let a_n be a sequence and a a real number. We say that a_n converges to a if for all $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that if $n \geq N$, then

$$|a_n - a| < \epsilon\,.$$

If this is the case, then we also say that a is the limit of the sequence and we say that the sequence is convergenet.

A sequence that is not covergent is said to be divergent.

Example:
$$0.999999999 \cdots = 1\,.$$
What does the left hand side mean?

Define a sequence a_n as follows: Set

$$a_1 = 0.9\,,$$

$$a_2 = 0.99\,,$$

$$a_3 = 0.999\,,$$

$$a_4 = 0.9999\,,$$

etc.

The left hand side above is then defined as the limit of the sequence a_n.

Claim:

$$\lim_{n\to\infty} a_n = 1.$$

Proof. We need to show that for all $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that for $n \geq N$ we have that

$$|a_n - 1| < \epsilon.$$

By the Archimedean property we can choose N such that $\frac{1}{N} < \epsilon$. We also have that

$$|a_n - 1| = 10^{-n}.$$

Therefore, for $n \geq N$ we have that

$$|a_n - 1| = 10^{-n} < \frac{1}{n} \leq \frac{1}{N} < \epsilon.$$

This proves the claim. $\qquad\qquad\square$

Theorem If a_n is a convergent sequence, then the set $\{a_n\}$ is a bounded subset of $\mathbf{R}$.

Proof. Since a_n is convergent to a we can find N such that for $n \geq N$ we have that

$$|a - a_n| < 1.$$

Note also that the set $\{a_1, \cdots, a_{N-1}\}$ is bounded so there exists $C \in \mathbf{R}$ such that for $n = 1, \cdots, N-1$ we have that

$$|a_n| \leq C.$$

To see that the larger set $\{a_n\}$ is bounded we will use that for $n \geq N$

$$|a_n| \leq |a| + |a_n - a| \leq |a| + 1.$$

From this we have that for all n

$$|a_n| \leq \max\{C, |a| + 1\}.$$

$\qquad\qquad\square$

Basic algebraic properties of limits:

Theorem Suppose that a_n and b_n are convergenet sequences with $\lim a_n = a$, $\lim b_n = b$ and $C \in \mathbf{R}$, then

(1) $c_n = C\, a_n$ is convergenet and $\lim_{n\to\infty} c_n = C\, a$.
(2) $c_n = a_n + b_n$ is convergent and $\lim_{n\to\infty} c_n = a + b$.
(3) $c_n = a_n b_n$ is convergent with $\lim_{n\to\infty} c_n = a\, b$.
(4) If $b_n \neq 0$, $b \neq 0$ and $c_n = \frac{a_n}{b_n}$, then c_n is convergent and $\lim_{n\to\infty} c_n = \frac{a}{b}$.

Proof. (of the first property.) If $c = 0$, then the claim is obviously true so we need only show the claim for $C \neq 0$. Given $\epsilon > 0$, there exists an N such that if $n \geq N$, then

$$|a - a_n| < \frac{\epsilon}{|C|}.$$

Multiplying both sides by $|C|$ gives that

$$|C\,a - C\,a_n| < \epsilon.$$

for $n \geq N$. This show the first property.

(Of the second of these properties.) Observe that

$$|c_n - (a + b)| = |a_n + b_n - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|.$$

Since $a_n \to a$, given $\epsilon > 0$ we can find a N_a such that if $n \geq N_a$, then

$$|a_n - a| < \frac{\epsilon}{2}.$$

Likewise since $b_n \to b$ we can find N_b such that if $n \geq N_b$, then

$$|b_n - b| < \frac{\epsilon}{2}.$$

We now set $N = \max\{Na, N_b\}$ and observe that if $n \geq N$, then

$$|c_n - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This proves the second property.

(Outline of how to show the third property.) To prove the third property we will use that

$$|a\,b - a_n\,b_n| \leq |a\,b - a_n\,b| + |a_n\,b - a_n\,b_n| = |b|\,|a - a_n| + |a_n|\,|b - b_n|.$$

We then combine it with the theorem above that show that the set

$$\{|a_n|\,|\,n \in \mathbf{N}\}$$

is bounded. This is the main idea of the proof of the third property. There are the details to fill in to make it a proof.

(Outline of how to show the fourth property.) To prove the fourth property we will assume that $a_n = 1$. The general case indeed will follow from this together with the third property. We will use that

$$\frac{1}{b_n} - \frac{1}{b} = \frac{|b - b_n|}{|b|\,|b_n|},$$

together with that

$$|b_n| \leq |b| + |b_n - b|.$$

and therefore

$$|b_n| \geq |b| - |b_n - b|.$$

We then want to use this to bound the denominator (when n is sufficiently large) from below in

$$\frac{|b - b_n|}{|b|\,|b_n|}.$$

Like for the third property there are details to fill in but these are the main ideas.   □

Subsequence:

Example 1: Suppose that $a_n = (-1)^n$.
This is a sequence of 1's and -1's that is alternating between -1 and 1.
The sequence $b_n = 1$ for all n is a subsequence.
Another subsequence is where $c_n = -1$.
Also the sequence $c_n = (-1)^{n+1}$ is a subsequence of a_n.
Another example of a subsequence is

$$1, 1, -1, -1, 1, 1, -1, -1, \cdots .$$

Example 2: Suppose $a_n = n$. So a_n is:

$$1, 2, 3, 4, 5, 6, \cdots .$$

The sequence of increasing odd numbers

$$1, 3, 5, 7, 9, \cdots .$$

is a subsequence.
The sequence of increasing even numbers

$$2, 4, 6, 8, 10, \cdots .$$

is another subsequence.
The sequence

$$1, 1, 2, 2, 3, 3, 4, 4, 5, 5 \cdots ,$$

is NOT s subsequence.

Formel definition: Recall that a sequence a_n is a function $f : \mathbf{N} \to \mathbf{R}$ where we set $a_n = f(n)$. A subsequence b_n of a_n is a composition of functions $f \circ g$ where $g : \mathbf{N} \to \mathbf{N}$ is a strictly increasing function. So $b_n = f(g(n))$. Sometimes a subsequence of the sequence a_n also denoted by a_{n_k}.

Theorem: A sequence a_n is convergent with limit a if and only if all subsequences of a_n are also convergent with limit a.

Proof. We need to show two implications.

First we need to show that is all subsequences of a_n are convergent with limit a, then the sequence a_n is convergent with limit a. However, this is trivially so since a_n itself is a (trivial) subsequence of a_n.

Next we need to show that any subsequence of a convergent sequence is convergent with the same limit. Suppose therefore that $\epsilon > 0$ is given and choose N so large so that for

$n \geq N$

$$|a_n - a| < \epsilon \,.$$

For $k \geq N$ we have that $n_k \geq k \geq N$ and therefore

$$|a_{n_k} - a| < \epsilon \,.$$

This proves the second implication. $\square$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
       TBB can be downloaded at:
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
       (screen-optimized)
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
       (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 5

A sequence of real numbers is a function $f : \mathbf{N} \to \mathbf{R}$.

We usually use the notation $a_n = f(n)$.

Limit: Let a_n be a sequence and a a real number. We say that a_n converges to a if for all $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that if $n \geq N$, then

$$|a_n - a| < \epsilon \,.$$

If this is the case, then we also say that a is the limit of the sequence and we say that the sequence is convergenet. A sequence that is not covergent is said to be divergent.

Theorem If a_n is a convergent sequence, then the set $\{a_n\}$ is a bounded subset of $\mathbf{R}$.

Basic algebraic properties of limits:

Theorem Suppose that a_n and b_n are convergenet sequences with $\lim a_n = a$, $\lim b_n = b$ and $C \in \mathbf{R}$, then

(1) $c_n = C \, a_n$ is convergenet and $\lim_{n \to \infty} c_n = C \, a$.
(2) $c_n = a_n + b_n$ is convergent and $\lim_{n \to \infty} c_n = a + b$.
(3) $c_n = a_n \, b_n$ is convergent with $\lim_{n \to \infty} c_n = a \, b$.
(4) If $b_n \neq 0$, $b \neq 0$ and $c_n = \frac{1}{b_n}$, then c_n is convergent and $\lim_{n \to \infty} c_n = \frac{1}{b}$.

Subsequence: A subsequence b_k of a_n is a composition of functions $f \circ g$ where $g : \mathbf{N} \to \mathbf{N}$ is a strictly increasing function. So $b_k = f(g(k))$. We often write a_{n_k} for b_k.

Example: A sequence is defined by a_n

$$a_n = \frac{n^2 + 1}{n^2 + n + 1} \, .$$

We will show that a_n is convergent with limit 1. We have

$$a_n = \frac{1 + \frac{1}{n^2}}{1 + \frac{1}{n} + \frac{1}{n^2}} \, .$$

Since $\frac{1}{n} \to 0$ and $\frac{1}{n^2} \to 0$ we have that

$$1 + \frac{1}{n^2} \to 1$$

and

$$1 + \frac{1}{n} + \frac{1}{n^2} \to 1 \, .$$

It now follows from the algebraic properties of limits that

$$a_n \to 1 \, .$$

Example: To show that $\sqrt{2}$ is the limit of $a_1 = 1$, $a_2 = 1.4$, $a_3 = 1.41$, $a_4 = 1.414$, ... we need something else. We need the monotone convergence theorem.

Monotone convergence theorem: Increasing version. Let a_n be a monotone increasing sequence. This means that $a_1 \leq a_2 \leq a_3 \leq \cdots$. (Which we can also write this as $a_n \leq a_{n+1}$). If the sequence is bounded from above so that there exists A with

$$a_n \leq A \, ,$$

then a_n is convergent with limit $\sup \{a_n\}$.

Monotone convergence theorem: Decreasing version. Similarly, for a bounded monotone decreasing sequence a_n where $a_{n+1} \leq a_n$, we have that a_n converges and

$$\lim_{n \to \infty} a_n = \inf \{a_n\} \, .$$

Using the monotone convergence theorem we can now show that $\sqrt{2}$ is the limit of $a_1 = 1$, $a_2 = 1.4$, $a_3 = 1.41$, $a_4 = 1.414$, $a_5 = \cdots$.

To show this we argue as follows:

Let
$$a_n = \frac{b_n}{10^{n-1}}$$
where b_n is the largest integer so that
$$b_n^2 \leq 2 \cdot 10^{2n-2}.$$

We will show that a_n is an increasing and bounded sequence and that the limit a has the property that $a^2 = 2$.

So suppose that b_n is an integer and that
$$b_n^2 \leq 2 \cdot 10^{2n-2},$$
then obviously
$$(10\,b_n)^2 \leq 2 \cdot 10^{2(n+1)-2}.$$

This implies that $b_{n+1} \geq 10\,b_n$; so $a_{n+1} \geq a_n$ and the sequence is increasing.

It is clear that $b_n \leq 2 \cdot 10^{n-1}$, since $(2 \cdot 10^{2n-2})^2 > 2 \cdot 10^{2n-2}$.

Therefore the sequence $a_n = \frac{b_n}{10^{n-1}}$ is bounded and by the monotone convergence theorem $a_n \to a$.

To show that $a^2 = 2$ we argue as follows:

By the algebraic properties of limits we have that $a^2 = \lim_{n\to\infty} a_n^2$, but for each n we have that $a_n^2\, 10^{2\,n-2} = b_n^2 \leq 2 \cdot 10^{2\,n-2}$ so $a_n^2 \leq 2$.

This show that $a^2 \leq 2$.

Similarly, since $(b_n + 1)^2 > 2 \cdot 10^{2\,n-2}$ we have that $(a_n + 10^{1-n})^2 = \frac{(b_n+1)^2}{10^{2\,n-2}} > 2$.

Therefore, $2 \leq \lim_{n\to\infty}(a_n + 10^{1-n})^2 = \lim_{n\to\infty} a_n^2 = a^2$.

Proof of the monotone convergence theorem (Increasing version):

The decreasing version is proven similarly.

So suppose that we have a sequence with
$$a_n \leq a_{n+1} \leq A$$
and set
$$a = \sup \{a_n\}.$$
We want to show that
$$a_n \to a.$$

Given $\epsilon > 0$, since $a - \epsilon < a$ we have that $a - \epsilon$ is not an upper bound for the sequence, therefore there exists N such that
$$a_N > a - \epsilon.$$

Since the sequence is increasing we have for $n \geq N$ that
$$a - \epsilon < a_N \leq a_n \leq a$$
Here the last inequality used that a is an upper bound for the sequence.

We now have that for $n \geq N$
$$0 \leq a - a_n < \epsilon.$$

This shows that the sequence is convergent with limit a.

Cauchy sequence: A sequence a_n is said to be a Cauchy sequence if for all $\epsilon > 0$, there exists an N such that if m, $n \geq N$, then

$$|a_n - a_m| < \epsilon.$$

(Tail of the sequence bunch together.)

From Wikipedia: Baron Augustin-Louis Cauchy (1789 – 1857) was a French mathematician, engineer, and physicist. He was one of the first to rigorously state and prove the key theorems of calculus (thereby creating real analysis), pioneered the field complex analysis, and the study of permutation groups in abstract algebra. Cauchy also contributed to a number of topics in mathematical physics, notably continuum mechanics.

Theorem: A sequence is convergent if and only if it is a Cauchy sequence.

Application: Existence of fixed points for a maps. If $T : \mathbf{R} \to \mathbf{R}$ is a map, then $x_0 \in \mathbf{R}$ is a fixed point if

$$T(x_0) = x_0.$$

Definition A contracting map is a map $T : \mathbf{R} \to \mathbf{R}$ such that there exists $c < 1$ so for all x, $y \in \mathbf{R}$ we have that

$$|T(x) - T(y)| \leq c\,|x - y|.$$

(Points are squeezed together under the map.)

Contracting mapping theorem: Any contracting map has a fixed point.

Application of contracting mapping theorem: Existence of solutions to ODEs.

More on all of this next time......

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)

https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 6

Last time:

Basic algebraic properties of limits.

Monotone convergence theorem.

Cauchy sequence.

Cauchy sequence: A sequence a_n is said to be a Cauchy sequence if for all $\epsilon > 0$, there exists an N such that if m, $n \geq N$, then

$$|a_n - a_m| < \epsilon \, .$$

(Tail of the sequence bunch together.)

Theorem (Cauchy convergence theorem): A sequence is convergent if and only if it is a Cauchy sequence.

Application: Existence of fixed points for a maps.
If $T : \mathbf{R} \to \mathbf{R}$ is a map, then $x_0 \in \mathbf{R}$ is a fixed point if

$$T(x_0) = x_0 \, .$$

Definition A contracting map is a map $T : \mathbf{R} \to \mathbf{R}$ such that there exists $c < 1$ so for all x, $y \in \mathbf{R}$ we have that

$$|T(x) - T(y)| \leq c \, |x - y| \, .$$

(Points are squeezed together under the map.)

Contracting mapping theorem: Any contracting map has a fixed point.

For a contracting map the fix point is unique.

Suppose that x and y are two fixed point we want to show that $x = y$. We have

$$|x - y| = |T(x) - T(y)| \leq c\,|x - y|.$$

Since $c < 1$ this implies that $|x - y| = 0$ and so $x = y$.

On Pset 3 you will be asked to show that for a contracting map T and any $a_1 \in \mathbf{R}$ the sequence $a_{n+1} = T(a_n)$ is a Cauchy sequence. By the Cauchy theorem we then have that a_n is convergent.

Let a denote the limit. We claim that $T(a) = a$. Observe that $T(a_n) = a_{n+1} \to a$. If we can show that if $x_n \to x$, then $T(x_n) \to T(x)$, then

$$T(a_n) \to T(a)$$

but we already have that $T(a_n) = a_{n+1} \to a$ so we would have that $T(a) = a$ and thus a is a fixed point.

We need therefore show that if $x_n \to x$, then $T(x_n) \to T(x)$. To do that observe that

$$|T(x_n) - T(x)| \leq c\,|x_n - x|.$$

Since $x_n \to x$ we have that $|x_n - x| \to 0$ and so $|T(x_n) - T(x)| \to 0$. It follows that $T(x_n) \to T(x)$. Applying this to the sequence a_n shows that a is a fixed point for T.

Applications of contracting mapping theorem:

Existence of solutions to ODEs. We will return to this later as this needs a version of the contracting mapping theorem where T is defined on a more general space than the real numbers.

Newton's method: Finding a zeroth of a function $f : \mathbf{R} \to \mathbf{R}$. (So find a solution x to $f(x) = 0$.)

Suppose that x_1 is a "good" initial guess, so $f(x_1)$ is sufficiently small. Assume also that $f' \neq 0$. Define a map

$$T(x) = x - \frac{f(x)}{f'(x)}.$$

We have

$$T'(x) = 1 - \frac{f'}{f'} + \frac{f\,f''}{(f')^2} = f\,\frac{f''}{(f')^2}.$$

So as long as x stay close to the initial guess and for the initial guess $f(x)$ is small compared with $\frac{f''}{(f')^2}$, then T is a contracting map. By the contracting mapping theorem the sequence $x_{n+1} = T(x_n)$ is a Cauchy sequence that converges to a fixed point of T.

For a fixed point for T we have $T(x) = x$ so $x - \frac{f(x)}{f'(x)} = x$ and therefore $f(x) = 0$.

Back to Cauchy sequences.

Bolzano -Weirstrass theorem: Any bounded sequence has a convergent subsequence.

Once we have the Bolzano-Weirstrass theorem we can prove the Cauchy theorem.

Proof. (of the Cauchy theorem.) So suppose that a_n is a Cauchy sequence. We will first show that a_n is bounded. From the definition of a Cauchy sequence we have that there exists N such that for m, $n \geq N$, then

$$|a_n - a_m| < 1.$$

It follows, in particular, that for all $n \geq N$, we have that

$$|a_n - a_N| < 1,$$

and so

$$|a_n| = |(a_n - a_N) + a_N| \leq 1 + |a_N|.$$

Therefore,

$$|a_n| \leq \max\{|a_N| + 1, |a_1|, \cdots, |a_{N-1}|\}.$$

So the sequence is bounded.

From the Bolzano-Weirstrass theorem it follows that a_n has a convergent subsequence a_{n_k} with limit a. We want to show that a_n is convergent with limit a. Given $\epsilon > 0$, there exists an N_1 such that if m, $n \geq N_1$, then

$$|a_n - a_m| < \frac{\epsilon}{2}.$$

Moreover, there exist an N_2 such that if $k \geq N_2$, then $|a_{n_k} - a| < \frac{\epsilon}{2}$. Set $N = \max\{N_1, N_2\}$. It follows that if $n \geq N$ and $k \geq N$, then

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

This show that $a_n \to a$ as claimed. $\qquad\square$

Another application of the Bolzano-Weirstrass theorem is the Extreme value theorem.

Before stating this we need another key notion:

A function $f : \mathbf{R} \to \mathbf{R}$ is said to be continuous at a point $x_0 \in \mathbf{R}$, if for all $\epsilon > 0$, there exists a $\delta > 0$ such that if

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon.$$

A function is said to be continuous if it is continuous at all points in the domain.

Theorem: If $f : \mathbf{R} \to \mathbf{R}$ is continuous and x_n is a sequence with $x_n \to x_0$, then $f(x_n) \to f(x_0)$.

Proof. Given $\epsilon > 0$, since f is continuous, there exists a $\delta > 0$, such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. Since $x_n \to x_0$, there exists N such that if $n \geq N$, then $|x_n - x_0| < \delta$ and therefore $|f(x_n) - f(x_0)| < \epsilon$. This show that $f(x_n) \to f(x_0)$ as claimed. $\square$

Extreme value theorem: Let f be a continues function on an interval $[a, b]$. The extreme value theorem says that the sup and inf are achieved. That is, there exist $x \in [a, b]$ such that $f(x) = \sup f$. Likewise for $\inf f$.

Proof. We will show that the supremum is achieved. The proof that the infimum is the same with obvious modification. Let $x_n \in [a, b]$ be a sequence where $f(x_n) \to \sup f$. Since the sequence is contained in $[a, b]$ it is bounded and therefore by the Bolzano - Weirstrass theorem has a convergent subsequence $x_{n_k} \to x$. Note that $x \in [a, b]$. By the theorem above $f(x_{n_k}) \to f(x)$ and since we also have that $f(x_{n_k}) \to \sup f$ it follows that $\sup f = f(x)$. This proves that the supremum is achieved. $\square$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 7

Last time:

Cauchy convergence theorem: Every Cauchy sequence is convergent.

Bolzano-Weirstrass theorem.

Last time we showed that the Bolzano-Weirstrass implies the Cauchy theorem.

Bolzano-Weirstrass theorem: Any bounded sequence has a convergent subsequence.

Proof. (of the Bolzano-Weirstrass theorem.) Suppose that a_n is a bounded sequence. For simplicity assume that $a_n \in [0, 1]$.

Defining the subsequence a_{n_k}. Either there are infinite many n such that $a_n \in [0, \frac{1}{2}]$ or there are infinite many such a_n in $[\frac{1}{2}, 1]$ (or both). Assume that there are infinitely many in $[0, \frac{1}{2}]$. Set $a_{n_1} = a_1$. Let a_{n_2} be the next a_n such that $a_n \in [0, \frac{1}{2}]$. We have

$$n_2 > n_1 = 1\,,$$

$$a_{n_1} \in [0, 1]\,,$$

$$a_{n_2} \in \left[0, \frac{1}{2}\right]\,.$$

Next either infinitely many a_n lies in $[0, \frac{1}{4}]$ or infinitely many a_n lies in $[\frac{1}{4}, \frac{1}{2}]$. Assume that infinitely many lies in $[\frac{1}{2}, \frac{1}{4}]$. Pick an $n > n_2$ such that $a_n \in [\frac{1}{4}, \frac{1}{2}]$ and set $a_{n_3} = a_n$. We continue this way.

Convergence of a_{n_k}. Note that for $k_1, k_2 \geq k$ we have that

$$|a_{n_{k_1}} - a_{n_{k_2}}| \leq 2^{1-k}\,.$$

Since $2^{-k} \to 0$ as $k \to \infty$ this shows that the subsequence a_{n_k} is a Cauchy sequence. However, more is true. Squeezed between two other sequences $b_k \leq a_{n_k} \leq c_k$. We will define sequences b_k and c_k as follows. The sequence b_k will be the left endpoint of the interval of length 2^{1-k} that all the element a_{n_i} will lie in when $i \geq k$ and c_k will be the right end point of the same interval. We have now that the sequence b_k is increasing and the sequence c_k is decreasing and the a_{n_k} are squeezed between the two. It follows that b_k is convergent (as it is also bounded) and likewise for c_k. Since $c_k - b_k = 2^{1-k}$ it follows that b_k and c_k converges to the

same number and since the later a_{n_k} are all squeezed between the two they also converges to that same number.   □

From Wikipedia: Karl Theodor Wilhelm Weierstrass (1815 – 1897) was a German mathematician often cited as the "father of modern analysis". Despite leaving university without a degree, he studied mathematics and trained as a school teacher, eventually teaching mathematics, physics, botany and gymnastics. Among many other contributions, Weierstrass formalized the definition of the continuity of a function and complex analysis, proved the intermediate value theorem and the Bolzano–Weierstrass theorem, and used the latter to study the properties of continuous functions on closed bounded intervals.

Series: Suppose that a_n is a sequence, we can form a new sequence s_n as follows. We let

$$s_1 = a_1 \,,$$

$$s_2 = a_1 + a_2$$

$$s_3 = a_1 + a_2 + a_3 \,,$$

and in general set

$$s_n = a_1 + \cdots + a_n = \sum_{i=1}^{n} a_i \,.$$

A series $\sum_{i=1}^{\infty} a_i$ converges if the sequence s_n converges and if it do we also write $\sum_{i=1}^{\infty} a_i$ for the limit.

Geometric series: Suppose now that $a_n = c^n$ so the series is

$$s_n = \sum_{i=0}^{n} c^i \,.$$

This is the geometric series. It is convergent precisely when $|c| < 1$. Moreover, when $|c| < 1$, then the limit (infinite sum) is

$$\sum_{i=0}^{\infty} c^i = \frac{1}{1-c} \,.$$

To see this observe that

$$(1-c) \sum_{i=0}^{n} c^i = 1 - c^{n+1} \,.$$

It follows from this that if $c \neq 1$, then

$$s_n = \sum_{i=0}^{n} c^i = \frac{1 - c^{n+1}}{1 - c} \,.$$

Therefore, s_n converges with limit $\frac{1}{1-c}$ if $|c| < 1$ and diverges if $|c| > 1$ or $c = -1$. One easily checks, separately, that it also diverges for $c = 1$.

Harmonic series: For the harmonic series $a_n = \frac{1}{n}$ so the series is $\sum_{i=1}^{\infty} \frac{1}{2}$. This series is divergent. To see this we will show that

$$s_{2^n-1} \geq \frac{n}{2}.$$

This is true for $n = 1$ as $s_1 = 1 \geq \frac{1}{2}$.

Assume that it is true for n we will show that it is also true for $n + 1$. Namely,

$$s_{2^{n+1}-1} \geq s_{2^n-1} + \sum_{i=2^n}^{2^{n+1}-1} \frac{1}{i} \geq \frac{n}{2} + 2^n \frac{1}{2^{n+1}-1} \geq \frac{n}{2} + \frac{1}{2} = \frac{n+1}{2}.$$

So the formula also holds for $n + 1$ and therefore for all n.

Since $\frac{n}{2} \to \infty$ it follows that the subsequence s_{2^n-1} is divergent and therefore so is the original series.

Absolutely convergent; We say that a series

$$\sum_{n=0}^{\infty} a_n$$

is absolutely convergent if the series

$$\sum_{n=0}^{\infty} |a_n|$$

is convergent. Absolutely convergent implies convergent but not the other way around.

Example: We will see later that the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

is convergent but if we take the absolute values of the a_n's, then we get the harmonic series which is divergent.

Proof. (of why absolutely convergent implies convergent.) By the Cauchy convergence theorem we only need to show that the sequence

$$\sum_{i=0}^{\infty} a_i$$

is Cauchy sequence if the series

$$\sum_{n=0}^{\infty} |a_n|$$

is convergent. Set

$$s_n = \sum_{i=0}^{n} a_i \,,$$

$$\mathbf{s}_n = \sum_{i=0}^{n} |a_i| \,.$$

For $m < n$ we have

$$|s_n - s_m| = |a_n + \cdots + a_{m+1}| \leq |a_n| + \cdots + |a_{m+1}| = |\mathbf{s}_n - \mathbf{s}_m| \,.$$

Since the sequence $\mathbf{s}_n$ is a Cauchy sequence it now follows that s_n is.   □

Theorem: A series of non-negative numbers

$$\sum_{i=0}^{\infty} a_i \,,$$

where $a_n \geq 0$, is convergent if and only if the sequence s_n is bounded from above.

Proof. The sequence s_n is monotone nondecreasing since

$$s_{n+1} = s_n + a_n \geq s_n \,.$$

The claim now follows from the monotone convergence theorem.   □

Example: The series

$$\sum_{i=1}^{\infty} \frac{1}{i^2}$$

is convergent. This is a sequence of non-negative numbers so we only need to show that there exist M such that for all n

$$s_n = \sum_{i=1}^{n} \frac{1}{i^2} \leq M \,.$$

Claim:

$$s_{2^n - 1} \leq \sum_{i=0}^{n-1} \left(\frac{1}{2} \right)^i \,.$$

This would be enough because the last is a convergent geometric series so in particular bounded.

We will show this by induction. For $n = 1$ we have that

$$s_1 = 1 = \left(\frac{1}{2}\right)^0 .$$

So it is correct for $n = 1$. Assume next that it is true for n; we will show that it also holds for $n + 1$.

$$s_{2^{n+1}-1} = s_{2^n-1} + \sum_{i=2^n}^{2^{n+1}-1} \frac{1}{i^2} \leq \sum_{i=0}^{n-1} \left(\frac{1}{2}\right)^i + 2^n \frac{1}{(2^n)^2}$$

$$\leq \sum_{i=0}^{n-1} \left(\frac{1}{2}\right)^i + \frac{1}{2^n} = \sum_{i=0}^{n} \left(\frac{1}{2}\right)^i .$$

This show the induction step and completes the proof.

To help determine whether or not a series converges there are a number of tests:
- Comparison test.
- Ratio test.
- Root test.

Comparison test; version 1: Suppose that a_n and b_n are two sequences with

$$0 \leq a_n \leq b_n .$$

If

$$\sum_{n=1}^{\infty} b_n$$

is convergent, then so is

$$\sum_{n=1}^{\infty} a_n .$$

Example: The series

$$\sum_{n=1}^{\infty} \frac{2^{-n}}{n}$$

is convergent. Namely, if we set

$$a_n = \frac{2^{-n}}{n}$$

and

$$b_n = 2^{-n} ,$$

then $0 \leq a_n \leq b_n$ and since the series $\sum_{n=1}^{\infty} b_n$ is convergent, then by the comparison test so is the series $\sum_{n=1}^{\infty} a_n$.

Comparison test; version 2: Suppose that a_n and b_n are two sequences with $b_n \neq 0$ and

$$\lim_{n \to \infty} \frac{a_n}{b_n} = L \neq 0 \,,$$

The series

$$\sum_{n=1}^{\infty} a_n$$

is convergent if and only if

$$\sum_{n=1}^{\infty} b_n$$

is.

Example: The series

$$\sum_{n=2}^{\infty} \frac{1}{n^2 - 1}$$

is convergent since

$$\frac{n^2}{n^2 - 1} \to 1 \,,$$

and the series

$$\sum_{n=2}^{\infty} \frac{1}{n^2}$$

is convergent.

Ratio test: Let $a_n \geq 0$ and assume that

$$\frac{a_{n+1}}{a_n} \to a$$

.

If

- $a < 1$, then the series $\sum a_n$ is convergent.
- $a > 1$, then the series $\sum a_n$ is divergent.
- $a = 1$, it is inconclusive.

Example 1:

$$a_n = \frac{1}{n} \,.$$

In this case

$$\frac{a_{n+1}}{a_n} \to 1$$

so the test is inconclusive, but the series is divergent.

Example 2:

$$a_n = \frac{1}{n^2} \,.$$

In this case

$$\frac{a_{n+1}}{a_n} \to 1$$

so the test is inconclusive, but the series is convergent.

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, DEPT. OF MATH., 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 8

Series: Suppose that a_n is a sequence, we can form a new sequence s_n as follows. We let

$$s_1 = a_1\,,$$

$$s_2 = a_1 + a_2$$

$$s_3 = a_1 + a_2 + a_3\,,$$

and in general set

$$s_n = a_1 + \cdots + a_n = \sum_{i=1}^{n} a_i\,.$$

A series $\sum_{i=1}^{\infty} a_i$ converges if the sequence s_n converges and if it do we also write $\sum_{i=1}^{\infty} a_i$ for the limit.

Geometric series:

$$\sum_{i=0}^{\infty} c^i n\,.$$

Convergent precisely when $|c| < 1$.

Harmonic series:

$$\sum_{i=1}^{\infty} \frac{1}{n}\,.$$

This series is divergent.

Absolutely convergent; We say that a series

$$\sum_{n=0}^{\infty} a_n$$

is absolutely convergent if the series

$$\sum_{n=0}^{\infty} |a_n|$$

is convergent. Absolutely convergent implies convergent but not the other way around.

Theorem: A series of non-negative numbers $a_n \geq 0$

$$\sum_{i=0}^{\infty} a_i \,,$$

is convergent if and only if the sequence s_n is bounded from above.

To help determine whether or not a series converges there are a number of tests:
- Comparison test.
- Ratio test.
- Root test.
- Other tests that we will discuss later.

Comparison test; version 1: Suppose that a_n and b_n are two sequences with

$$0 \leq a_n \leq b_n \,.$$

If

$$\sum_{n=1}^{\infty} b_n$$

is convergent, then so is

$$\sum_{n=1}^{\infty} a_n \,.$$

Example: The series

$$\sum_{n=1}^{\infty} \frac{2^{-n}}{n}$$

is convergent. Namely, if we set

$$a_n = \frac{2^{-n}}{n}$$

and

$$b_n = 2^{-n} \,,$$

then $0 \leq a_n \leq b_n$ and since the series $\sum_{n=1}^{\infty} b_n$ is convergent, then by the comparison test so is the series $\sum_{n=1}^{\infty} a_n$.

Comparison test; version 2: Suppose that a_n and b_n are two sequences with $b_n \neq 0$ and

$$\lim_{n \to \infty} \frac{a_n}{b_n} = L \neq 0 \,,$$

The series

$$\sum_{n=1}^{\infty} a_n$$

is convergent if and only if

$$\sum_{n=1}^{\infty} b_n$$

is.

Example: The series

$$\sum_{n=2}^{\infty} \frac{1}{n^2 - 1}$$

is convergent since

$$\frac{n^2}{n^2 - 1} \to 1 \,,$$

and the series

$$\sum_{n=2}^{\infty} \frac{1}{n^2}$$

is convergent.

Ratio test: Let $a_n \geq 0$ and assume that

$$\frac{a_{n+1}}{a_n} \to a$$

.

If

- $a < 1$, then the series $\sum a_n$ is convergent.
- $a > 1$, then the series $\sum a_n$ is divergent.
- $a = 1$, it is inconclusive.

Example 1:

$$a_n = \frac{1}{n} \,.$$

In this case

$$\frac{a_{n+1}}{a_n} \to 1$$

so the test is inconclusive, but the series is divergent.

Example 2:

$$a_n = \frac{1}{n^2}\,.$$

In this case

$$\frac{a_{n+1}}{a_n} \to 1$$

so the test is inconclusive, but the series is convergent.

Root test: Let $a_n \geq 0$ be a sequence of non-negative numbers. Suppose $\lim_{n\to\infty}(a_n)^{\frac{1}{n}} = r$.
If

- $r < 1$, then the series $\sum_{n=0}^{\infty} a_n$ is convergent.
- $r > 1$, then the series $\sum_{n=0}^{\infty} a_n$ is divergent.
- $r = 1$, then it is inconclusive.

Proof. (of root test.) Suppose that $r < 1$. It follows that for $r < r_0 < 1$, there exists N such that if $n \geq N$, then

$$(a_n)^{\frac{1}{n}} \leq r_0\,.$$

Therefore,

$$0 \leq a_n \leq r_0^n\,.$$

However, the series $\sum_{n=0}^{\infty} r_0^n$ is a geometric series that is convergent since $r_0 < 1$. We now have by the first version of the comparison test that also the series $\sum_{n=0}^{\infty} a_n$ is convergent.

Suppose that $r > 1$. In that case we have that for $1 < r_0 < r$, there exists N such that if $n \geq N$, then

$$(a_n)^{\frac{1}{n}} \geq r_0\,.$$

Hence, for $n \geq N$

$$a_n \geq r_0^n\,,$$

where, the series $\sum_{n=0}^{\infty} r_0^n$ is a divergent geometric series. Therefore, by the comparison test the original series is divergent. $\qquad\square$

Power series:

-
$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}\,.$$

-
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp x\,.$$

-
$$\sum_{n=0}^{\infty}(-1)^n\,\frac{x^{2\,n}}{(2\,n)!} = \cos\,x\,.$$

-
$$\sum_{n=0}^{\infty}(-1)^n\,\frac{x^{2\,n+1}}{(2\,n+1)!} = \sin\,x\,.$$

Formel definition: Let c_n be a sequence, then $\sum_{n=0}^{\infty} c_n\,x^n$ is a power series.

When does a power series converge?

Why does it give familiar functions?

We will answer the second question next time for the exponential function.

The answer to the first question comes from the root test or the ratio test.

Example: Consider the power series:

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}\,.$$

By the ratio test with $a_n = \frac{x^n}{n!}$ we have

$$\frac{a_{n+1}}{a_n} = \frac{n!\,x^{n+1}}{(n+1)!\,x^n} = \frac{x}{n+1} \to 0\,.$$

It follows that the power series is convergent for all x.

Example:

$$\sum_{n=0}^{\infty} x^n\,.$$

This series is convergent for $|x| < 1$ and divergent otherwise.

To talk about convergence of a general power series we need the notion of $\limsup$ of a sequence. This is defined as follows.

Let a_n be a sequence. If it is not bounded from above, then we set $\limsup a_n$ to be ∞. Otherwise we will define a new sequence b_n from a_n as follows.

$$b_n = \sup \{a_n, a_{n+1}, a_{n+2}, \cdots\}.$$

Note that since we are assuming that the a_n's are bounded from above the b_n's are real numbers and the sequence b_n is decreasing. – It is decreasing since

$$b_n = \sup \{a_n, a_{n+1}, a_{n+2}, \cdots\} \geq \sup \{a_{n+1}, a_{n+2}, \cdots\} = b_{n+1}.$$

(For b_{n+1} supremum is taken over a smaller set.)

Since the sequence b_n is decreasing it is converging with limit b that possibly could be $-\infty$ if the sequence b_n is not bounded from below.

Definition (of $\limsup$):

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} b_n = b.$$

Back to power series. Suppose that

$$\sum_{n=0}^{\infty} a_n x^n$$

is a power series. Set

$$R = \frac{1}{\limsup_{n \to \infty} |a_n|^{\frac{1}{n}}}.$$

R is said to be the radius of convergence.

Convention: If $\limsup_{n \to \infty} |a_n|^{\frac{1}{n}} = 0$, then the radius of convergence is said to be ∞. If $\limsup_{n \to \infty} |a_n|^{\frac{1}{n}} = \infty$, then we set $R = 0$.

From the root test one can now show the following:

The power series in convergent if $|x| < R$ and divergent if $|x| > R$.

The case of where $|x| = R$ has to be examined on a case by case basis.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 9

Power series: Suppose that a_n is a sequence, for each x we can form a series

$$\sum_{n=0}^{\infty} a_n \, x^n \, .$$

Exponential map as a power series: Define $E(x)$ as the power series

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \, .$$

Step 0: The power series converges for all x. Namely, since

$$\frac{|a_{n+1} \, x^{n+1}|}{|a_n \, x^n|} = \frac{|n! \, x^{n+1}|}{|(n+1)! \, x^n|} = \frac{x}{n+1} \to 0 \, ,$$

the claim follows from the ratio test.

Step 1: Define $\mathrm{e} = E(1)$ and $\mathrm{e}^2 = \mathrm{e}\,\mathrm{e}$, $\mathrm{e}^3 = \mathrm{e}\,\mathrm{e}\,\mathrm{e}$ etc. This way e^k is defined for all $k \in \mathbf{N}$.

We also define $\mathrm{e}^{-k} = \frac{1}{\mathrm{e}^k}$. (The idea is that we would like to have $\mathrm{e}^{x+y} = \mathrm{e}^x \, \mathrm{e}^y$.)

We set $\mathrm{e}^0 = 1$.

If $q = \frac{m}{n} \in \mathbf{Q}$, where $m \in \mathbf{Z}$ and $n \in \mathbf{N}$, then we let e^q be the positive number α so that $\alpha^n = \mathrm{e}^m$. (Again the idea is that we would like to have that $\mathrm{e}^{x+y} = \mathrm{e}^x \, \mathrm{e}^y$.)

This way e^q is defined for all rational numbers.

What about the irrational numbers like $\sqrt{2}$?

Step 2: Next time we will show that

$$E(x+y) = E(x)\,E(y)\,.$$

Here we claim that $E(x) > 0$ for all x. If $x \geq 0$, then this is clear since

$$E(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \geq 1\,.$$

If $x < 0$, then using the formula we will show next time we have that

$$1 = E(0) = E(x)\,E(-x)\,.$$

Therefore,

$$E(x) = \frac{1}{E(-x)} > 0\,.$$

Using that $E(x+y) = E(x)\,E(y)$ we now claim that $E(q) = e^q$ for all rational numbers q.

For integers m this is how we defined e^m. For $m = -k$, where $k \in \mathbf{N}$ we defined

$$e^{-k} = \frac{1}{e^k} = \frac{1}{E(k)} = E(-k)\,.$$

For a general rational number $q = \frac{m}{n}$, where $m \in \mathbf{N}$ and $n \in \mathbf{Z}$, we have

$$\left(E\left(\frac{m}{n}\right)\right)^n = E\left(\frac{m}{n}\right) \cdots E\left(\frac{m}{n}\right) = E(m)\,,$$

and

$$E\left(\frac{m}{n}\right) > 0\,.$$

This gives us that $E(q) = e^q$ for all rational numbers q.

Step 3: We now have $E(x)$ is defined for all x whereas e^x is defined for all rational numbers.

What other properties would we want of the exponential function?

We would want it to be continuous!

Reminder: A function $f : A \to \mathbf{R}$ on some set $A \subset \mathbf{R}$ is said to be continuous if for all $x_0 \in A$ we have:

For all $\epsilon > 0$, there exists a $\delta = \delta(x_0) > 0$ such that if $|x - x_0| < \delta$ $(x \in A)$, then $|f(x) - f(x_0)| < \epsilon$.

On Pset 5 you will be asked to show that $E(x)$ is continuous at all points.

Step 4: We will next show that $E(x)$ is the unique continuous function where $E(q) = \mathrm{e}^q$ for all rational numbers q.

Theorem: Let f and g be two continuous function on $\mathbf{R}$ that agrees on all rational numbers, then $f = g$.

We will show this theorem next time. For now here are some more about what it means for a function to be continuous.

Example 1: Suppose $f(x) = c$, where c is a constant. We will show that f is continuous. Given $x_0 \in \mathbf{R}$ and $\epsilon > 0$, set $\delta = 1$. We then have that if $|x - x_0| < \delta = 1$, then $|f(x) - f(x_0)| = 0 < \epsilon$. This show that f is continuous.

Example 2: Suppose $f(x) = x$, we will show that f is continuous. Given $x_0 \in \mathbf{R}$ and $\epsilon > 0$, set $\delta = \epsilon$. We then have that if $|x - x_0| < \delta = \epsilon$, then $|f(x) - f(x_0)| = |x - x_0| < \epsilon$. This show that f is continuous.

Algebraic properties of continuous functions:
- If f and g are continuous functions, then so is $f + g$.
- If f is continuous and c is a constant, then $c\,f$ is continuous.
- If f and g are continuous, then $f\,g$ is also continuous.
- If f is continuous and $f \neq 0$, then $\frac{1}{f}$ is continuous.
- If $f(x)$ and $g(x)$ are continuous, then $f(g(x))$ is continuous.

Proof. (the proof is very similar to the one we gave for the algebraic properties of limits.) $\square$

Theorem: All polynomials are continuous.

Example 3: If $f(x) = x^2 + 1$, then f is continuos. We have already proven that $g(x) = x$ is continuous so by the algebraic properties we have that $x^2 = g\,g$ is continuous. We have also already shown that the constant functions are continuous so $h(x) = 1$ is continuous and therefore by the algebraic properties we have that $f = g^2 + h$ is continuous.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 10

Power series: Suppose that a_n is a sequence. For each x we can form a series

$$\sum_{n=0}^{\infty} a_n \, x^n \, .$$

Exponential function as a power series: Define $E(x)$ by the power series

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \, .$$

Step 0: The power series converges for all x.

Step 1: Define e^q for all rational numbers q.

Step 2: Need to show that

$$E(x + y) = E(x) \, E(y) \, .$$

Step 3: $E(x)$ is defined for all x, whereas e^x is defined for all rational numbers, and $E(q) = e^q$ for all rational numbers.

Step 4: E is continuous on all of $\mathbf{R}$. (Pset 5.)

Step 5: If f and g are continuous functions on $\mathbf{R}$ that agrees on $\mathbf{Q}$, then $f = g$ everywhere.

Suppose that we have two convergent series

$$\sum_{n=0}^{\infty} a_n \text{ and } \sum_{n=0}^{\infty} b_n$$

of non-negative numbers $a_n, b_n \geq 0$.

Form the "product series"

$$\sum_{n=0}^{\infty} c_n \,,$$

where

$$c_n = \sum_{i=0}^{n} a_i \, b_{n-i} \,.$$

Note that each $c_n \geq 0$ so by the monotone convergence theorem the series

$$\sum_{n=0}^{\infty} c_n$$

is convergent if it is bounded.

Theorem 1: If $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are as above, then the series

$$\sum_{n=0}^{\infty} c_n$$

is convergent with limit

$$\sum_{n=0}^{\infty} a_n \sum_{n=0}^{\infty} b_n \,.$$

Proof. Denote

$$s_n^a = \sum_{i=0}^{n} a_i \text{ and } s_n^b = \sum_{i=0}^{n} b_i \text{ and } s_n^c = \sum_{i=0}^{n} c_i \,.$$

The idea here is that

$$(*) \qquad \left(\sum_{\ell=0}^{n} a_\ell \right) \left(\sum_{\ell=0}^{n} b_\ell \right) = \sum_{k=0}^{n} \sum_{i+j=k} a_i \, b_j + \sum_{i+j>n \text{ and } i,j \leq n} a_i \, b_j \leq \sum_{k=0}^{2n} \sum_{i+j=k} a_i \, b_j \,.$$

In other words

$$(**) \qquad s_n^c \leq s_n^a \, s_n^b \leq s_{2n}^c \,.$$

This is because $(*)$ is

$$s_n^a \, s_n^b = s_n^c + \sum_{i+j>n \text{ and } i,j \leq n} a_i \, b_j \leq s_{2n}^c \,,$$

and

$$0 \leq \sum_{i+j>n \text{ and } i,j\leq n} a_i\, b_j\,.$$

Note that the first inequality in $(**)$ implies that the sequence s_n^c is bounded and therefore since a_n, b_n, $c_n \geq 0$ we have that

$$s_n^a \uparrow s^a\,, \qquad s_n^b \uparrow s^b\,, \qquad s_n^c \uparrow s^c$$

by the monotone convergence theorem for sequences. Since the product $s_n^a\, s_n^b$ is squeezed between s_n^c and s_{2n}^c by $(**)$ we have that

$$s^c \leq s^a\, s^b \leq s^c\,.$$

From this the claim follows. □

Applying Theorem 1 to the power series $E(x)$ we can now prove the following:

Theorem 2:
$$E(x+y) = E(x)\, E(y)\,.$$

Proof. We will show this assuming that x, $y \geq 0$. Once we have shown the theorem for x, $y \geq 0$ the general case is not too difficult but we will not prove that here. The idea is that $E(x+y)$ will play the role of

$$\sum_{n=0}^{\infty} c_n$$

above. So set

$$c_n = \frac{(x+y)^n}{n!}\,,$$

By the "binomial" formula

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i\, y^{n-i}\,.$$

So

$$c_n = \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} x^i\, y^{n-i}\,.$$

Since

$$\binom{n}{i} = \frac{n!}{i!\,(n-i)!}\,,$$

we have that

$$c_n = \sum_{i=0}^{n} \frac{x^i}{i!}\, \frac{y^{n-i}}{(n-i)!}\,.$$

This shows that

$$c_n = \sum_{i=0}^{n} a_i\, b_{n-i}\,,$$

where

$$a_i = \frac{x^i}{i!}$$

and

$$b_i = \frac{y^i}{i!} \, .$$

The claim now follows from Theorem 1. □

Coming back to the functions E and e. We have that they agree on all rational numbers and that E is defined for all real numbers.

We would want the exponential function to be continuous!

Reminder: A function $f : A \to \mathbf{R}$ on some set $A \subset \mathbf{R}$ is said to be continuous if for all $x_0 \in A$ we have:

For all $\epsilon > 0$, there exists a $\delta = \delta(x_0) > 0$ such that if $|x - x_0| < \delta$ $(x \in A)$, then $|f(x) - f(x_0)| < \epsilon$.

On Pset 5 you will be asked to show that $E(x)$ is continuous at all points.

Step 5: We will show that $E(x)$ is the unique continuous function where $E(q) = e^q$ for all rational numbers q.

Theorem 3: (On Pset 5.) Let f and g be two continuous function on $\mathbf{R}$ that agrees on all rational numbers, then $f = g$.

We will next see that there are functions on $\mathbf{R}$ that are not continuous at any point!

Before defining such a function recall that we already proved that $\sqrt{2}$ is a irrational number and thus for all $\delta > 0$, there exists an N such that if $n \geq N$, then

$$0 < \frac{\sqrt{2}}{n} < \delta \, .$$

So arbitrarily close to zero there are irrational numbers. Likewise by the Archimedean property we have that arbitrarily close to any irrational number there is a rational number.

SPRING 2025 - 18.100B/18.1002

On $\mathbf{R}$ define a function f as follows

$$f(x) = \begin{cases} 1 & x \in \mathbf{Q} \\ 0 & \text{otherwise} \end{cases}$$

We claim that f is nowhere continuous. Suppose first that x_0 is rational and let $0 < \epsilon < 1$. We have that $f(x_0) = 1$ and for any $\delta > 0$, there exists a irrational number x with $|x - x_0| < \delta$ but we also have that

$$\epsilon < 1 = |f(x) - f(x_0)|\,.$$

This show that f is discontinuous at x_0.

Likewise suppose x_0 is an irrational number. We have that $f(x_0) = 0$. Given $0 < \epsilon < 1$ for any $\delta > 0$, there exists a rational number x with $|x - x_0| < \delta$. On the other hand

$$\epsilon < 1 = |f(x) - f(x_0)|\,.$$

This show that f is discontinuous at x_0.

This gives an example of a function that is discontinuous at all points. On the other hand recall from last time how to generate continuous functions from known continuous functions:

Algebraic properties of continuous functions:
- If f and g are continuous functions, then so is $f + g$.
- If f is continuous and c is a constant, then $c\,f$ is continuous.
- If f and g are continuous, then $f\,g$ is also continuous.
- If f is continuous and $f \neq 0$, then $\frac{1}{f}$ is continuous.
- If $f(x)$ and $g(x)$ are continuous, then $f(g(x))$ is continuous.

Proof. (The proof is very similar to the one we gave for the algebraic properties of limits of sequences.) $\square$

Theorem: All polynomials are continuous.

REFERENCES

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, DEPT. OF MATH., 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 11

Reminder: A function $f : A \to \mathbf{R}$ on some set $A \subset \mathbf{R}$ is said to be continuous if for all $x_0 \in A$ we have:

For all $\epsilon > 0$, there exists a $\delta = \delta(x_0) > 0$ such that if $|x - x_0| < \delta$ ($x \in A$), then $|f(x) - f(x_0)| < \epsilon$.

Two theorems about continuous functions:

Extreme Value Theorem: Suppose that $f : [a, b] \to \mathbf{R}$ is a continuous function, then there exist $x_M \in [a, b]$ such that $f(x_M) \geq f(x)$ for all $x \in [a, b]$. Similarly, there exists $x_m \in [a, b]$ such that $f(x_m) \leq f(x)$ for all $x \in [a, b]$.

Intermediate Value Theorem: Suppose that $f : [a, b] \to \mathbf{R}$ is a continuous function, then for all y between $f(a)$ and $f(b)$, there exists $x \in [a, b]$ such that $f(x) = y$.

We will show these theorems using a lemma that connects sequences and continuous functions. This is the following:

Lemma: Suppose that $f : [a, b] \to \mathbf{R}$ is a continuous function and $x_n \to x_\infty$ a sequence, then $f(x_n) \to f(x_\infty)$. We can also write this as

$$\lim_{n \to \infty} f(x_n) = f\left(\lim_{n \to \infty} x_n\right).$$

Proof. To show that $f(x_n) \to f(x_\infty)$ let $\epsilon > 0$ be given. Since f is continuous at x_∞, there exists $\delta > 0$ such that if $|x - x_\infty| < \delta$, then $|f(x) - f(x_\infty)| < \epsilon$. Since $x_n \to x_\infty$ there exists N such that if $n \geq N$, then $|x_n - x_\infty| < \delta$ and therefore $|f(x_n) - f(x_\infty)| < \epsilon$. This shows the lemma. $\square$

Using this lemma we can now prove the extreme value theorem:

Proof. (of EVT.) Let $E = f([a, b])$ and set $M = \sup E$. We will show that $M < \infty$ and that $M = f(x)$ for some $x \in [a, b]$. We show first that M is finite. Otherwise for each n there exists an $x_n \in [a, b]$ such that $f(x_n) > n$. Since the sequence $\{x_n\}$ is bounded by the Bolzano-Weirstrass theorem it has a convergent subsequence. Let us denote that by x_{n_k}. We have $x_{n_k} \to x_\infty \in [a, b]$. By the lemma above $f(x_{n_k}) \to f(x_\infty)$ but we assumed that the sequence $f(x_{n_k})$ is unbounded which is the desired contradiction.

For each integer n we can now choose $x_n \in [a, b]$ such that $f(x_n) > M - \frac{1}{n}$. Again since this sequence is bounded by the Bolzano-Weirstrass theorem it has a convergent subsequence $x_{n_k} \to x_\infty \in [a, b]$. By the lemma above $f(x_{n_k}) \to f(x_\infty) \geq M$. Since $M = \sup f([a, b])$ we have that $f(x_\infty) = M$. This show the EVT. □

Proof. (of IVT.) We will assume that $f(a) < 0 < f(b)$ and show that there exists $x \in [a, b]$ such that $f(x) = 0$. The general case is similar. Let $A = \{y \mid$ for all $x \leq y$ we have that $f(x) \leq 0\}$. Note that $a \in A$ so the set is non-empty. Set $M = \sup A$ and let x_n be a sequence with $x_n < M$ and $x_n \to M$. It follows that $f(x_n) < 0$ and so by the lemma above we must have that $f(M) \leq 0$. We are done if $f(M) = 0$ so assume that $f(M) < 0$. We have that $M < b$ and by continuity there exist a whole interval around M where $f < 0$. This contradict that M was the supremum of the set A. Showing the IVT. □

Abstract metric space.

Definition: Metric space A metric space is a set X with a function $d : X \times X \to \mathbf{R}$ with the following three properties:

(1) $d(x,y) \geq 0$ for all x, $y \in X$ and $d(x,y) = 0$ if and only if $x = y$. (Distances ≥ 0.)
(2) $d(x,y) = d(y,x)$. (Symmetric.)
(3) $d(x,z) \leq d(x,y) + d(y,z)$. (Triangle inequality.)

Examples:

(1) $X = \mathbf{R}$ and
$$d(x,y) = |x - y| \, .$$
(2) $X = \mathbf{R}^2$ and for $x = (x_1, x_2)$ and $y = (y_1, y_2)$
$$d(x,y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} \, .$$
(3) $X = \mathbf{R}^3$ and for $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$
$$d(x,y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2} \, .$$

Example: Continuous function on an interval $[a,b]$. Let $X = C([a,b])$ where $C([a,b])$ is the set of continuous functions on $[a,b]$. The distance between two continuous functions f and g is then
$$d(f,g) = \max_{x \in [a,b]} |f(x - g(x)| \, .$$
Since $f - g$ is also a continuous function the EVT theorem guarantees that the max is achieved for some $x \in [a,b]$.

Metric spaces plays the role of generalised real numbers. A lot of the discussion that we have had in the class holds also for metric spaces and this is useful in many circumstances. For instance, we will see in a later class that we can use it to solve ODEs.

Sequences in a metric space: A sequence in a metric space (X, d) is a map $f : \mathbf{N} \to X$. We typically denote the image $f(n)$ by x_n. Similarly we define a **subsequence** as the composition of a strictly increasing map $g : \mathbf{N} \to \mathbf{N}$ with f and $x_{n_k} = f(g(k))$.

It is not all results that we know from $\mathbf{R}$ that generalises to general metric spaces. For instance, in general there are no algebraic properties, no squeeze theorem, no monotone

convergence theorem. On the other hand the statement of both the Cauchy convergence theorem and the Bolzano-Weirstrass theorems makes sense in a general metric space.

Example (Box distance): The space is $X = \mathbf{R}^2$ and if $\underline{x} = (x_1, x_2)$ and $\underline{y} = (y_1, y_2)$, then

$$d(\underline{x}, \underline{y}) = |x_1 - y_1| + |x_2 - y_2|.$$

Example (Strange metric on integers): The space is $X = \mathbf{N}$ and if m, n are integers, then

$$d(m, n) = \left| \frac{1}{n} - \frac{1}{m} \right|.$$

Here is a wild example of a metric space:

Example (French railway metric): The space is $X = \mathbf{R}^2$ and if $\underline{x} = (x_1, x_2)$ and $\underline{y} = (y_1, y_2)$, then

$$d(\underline{x}, \underline{y}) = \begin{cases} |\underline{x} - \underline{y}| & \text{if } \underline{x} = c\,\underline{y} \text{ or } \underline{y} = c\,\underline{x} \text{ for some } c \in \mathbf{R} \\ |\underline{x}| + |\underline{y}| & \text{otherwise} \end{cases}.$$

Here

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

and likewise for $|\underline{x}|$ and $|\underline{y}|$.

Definition: Convergent sequence in a metric space If (X, d) is a metric space and x_n is a sequence in X, then we say that x_n converges to x and write $x_n \to x$ or $x = \lim_{n \to \infty} x_n$ if for all $\epsilon > 0$, there exists an N such that if $n \geq N$, then

$$d(x, x_n) < \epsilon.$$

This is equivalent to that the sequence $d(x_n, x_\infty) \to 0$.

Definition: Cauchy sequence in a metric space If (X, d) is a metric space and x_n is a sequence in X, then we say that x_n is a Cauchy sequence if for all $\epsilon > 0$, there exists an N, such that if m, $n \geq N$, then

$$d(x_m, x_n) < \epsilon.$$

Theorem: In any metric space (X, d) a convergent sequence is also a Cauchy sequence.

Proof. So suppose that $x_n \in X$ is a sequence and $x_n \to x$. Given $\epsilon > 0$, convergence means that there exists N such that if $n \geq N$, then $d(x, x_n) < \frac{\epsilon}{2}$. If both m, $n \geq N$, then we have by the triangle inequality that

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \,.$$

This show the theorem. $\hfill\square$

The converse is not always the case: If $X = (0, 1) \subset \mathbf{R}$ with $d(x, y) = |x - y|$, then the sequence $x_n = \frac{1}{n}$ is a Cauchy sequence but since 0 is not in X, it is not convergent. We sometimes express this by saying that in this case X is not Cauchy complete.

Definition: Continuous function on a metric space (X, d) Suppose that $F : X \to \mathbf{R}$ is a function. We say that f is continuous at $x_0 \in X$, if for all $\epsilon > 0$, there exists a $\delta > 0$, such that if $x \in X$ with $d(x, x_0) < \delta$, then

$$|F(x) - F(x_0)| < \epsilon \,.$$

Example: Let again $X = C([0, 1])$ be the set of continuous functions on $[0, 1]$. Equip X with the distance described above. So the distance between to continuous functions f and g is then

$$d(f, g) = \max_{x \in [a,b]} |f(x - g(x)| \,.$$

Define F on X to be the function $F(f) = f(0)$ where $f \in C([0, 1])$. F is easily seen to be a continuous function on the metric space X.

We can now extend one of the earlier lemmas to general metric spaces.

Lemma: Let (X, d) be a general metric space. Suppose that $f : X \to \mathbf{R}$ is a continuous function and x_n is a sequence in X with $x_n \to x_\infty$, then $f(x_n) \to f(x_\infty)$.

Proof. To show that $f(x_n) \to f(x_\infty)$ let $\epsilon > 0$ be given. Since f is continuous at x_∞, there exists $\delta > 0$ such that if $d(x, x_\infty) < \delta$, then $|f(x) - f(x_\infty)| < \epsilon$. Since $x_n \to x_\infty$ there exists N such that if $n \geq N$, then $d(x_n, x_\infty) < \delta$ and therefore $|f(x_n) - f(x_\infty)| < \epsilon$. This shows the lemma. $\hfill\square$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 12

Review (discussed in lectures so far):
  (1) $\mathbf{R}$ is the complete ordered field that contains $\mathbf{Q}$.
  (2) Sequences and limits.
  (3) Series.
  (4) Continuous functions.
  (5) Metric spaces.

Definition: Metric space. A metric space is a set X with a function $d : X \times X \to \mathbf{R}$ with the following three properties:
  (1) $d(x, y) \geq 0$ for all x, $y \in X$ and $d(x, y) = 0$ if and only if $x = y$. (Distances ≥ 0.)
  (2) $d(x, y) = d(y, x)$. (Symmetric.)
  (3) $d(x, z) \leq d(x, y) + d(y, z)$. (Triangle inequality.)

Examples (Euclidean distance):
  (1) $X = \mathbf{R}$ and
$$d(x, y) = |x - y|.$$
  (2) $X = \mathbf{R}^2$ and for $x = (x_1, x_2)$ and $y = (y_1, y_2)$
$$d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2}.$$
  (3) $X = \mathbf{R}^3$ and for $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$
$$d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2}.$$

Example: Continuous function on an interval $[a, b]$. Let $X = C([a, b])$ where $C([a, b])$ is the set of continuous functions on $[a, b]$. The distance between two continuous functions f and g is then
$$d(f, g) = \max_{x \in [a,b]} |f(x - g(x)|.$$

Since $f - g$ is also a continuous function the EVT theorem guarantees that the max is achieved for some $x \in [a, b]$.

Example (Box distance): The space is $X = \mathbf{R}^2$ and if $\underline{x} = (x_1, x_2)$ and $\underline{y} = (y_1, y_2)$, then

$$d(\underline{x}, \underline{y}) = |x_1 - y_1| + |x_2 - y_2|\,.$$

Sequences in a metric space: A sequence in a metric space (X, d) is a map $f : \mathbf{N} \to X$. We typically denote the image $f(n)$ by x_n. Similarly we define a **subsequence** as the composition of a strictly increasing map $g : \mathbf{N} \to \mathbf{N}$ with f and $x_{n_k} = f(g(k))$.

It is not all results that we know from $\mathbf{R}$ that generalises to general metric spaces. For instance, in general there are no algebraic properties, no squeeze theorem, no monotone convergence theorem. On the other hand the statement of both the Cauchy convergence theorem and the Bolzano-Weirstrass theorems makes sense in a general metric space.

Definition: Convergent sequence in a metric space If (X, d) is a metric space and x_n is a sequence in X, then we say that x_n converges to x and write $x_n \to x$ or $x = \lim_{n \to \infty} x_n$ if for all $\epsilon > 0$, there exists an N such that if $n \geq N$, then

$$d(x, x_n) < \epsilon\,.$$

This is equivalent to that the sequence $d(x_n, x) \to 0$.

Definition: Cauchy sequence in a metric space If (X, d) is a metric space and x_n is a sequence in X, then we say that x_n is a Cauchy sequence if for all $\epsilon > 0$, there exists an N, such that if $m, n \geq N$, then

$$d(x_m, x_n) < \epsilon\,.$$

Theorem: In any metric space (X, d) a convergent sequence is also a Cauchy sequence.

Proof. So suppose that $x_n \in X$ is a sequence and $x_n \to x$. Given $\epsilon > 0$, convergence means that there exists N such that if $n \geq N$, then $d(x, x_n) < \frac{\epsilon}{2}$. If both $m, n \geq N$, then we have by the triangle inequality that

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.$$

This show the theorem.   □

The converse is not always the case: If $X = (0, 1) \subset \mathbf{R}$ with $d(x, y) = |x - y|$, then the sequence $x_n = \frac{1}{n}$ is a Cauchy sequence but since 0 is not in X, it is not convergent. We sometimes express this by saying that in this case X is not Cauchy complete.

A metric space is said to be Cauchy complete if every Cauchy sequence is convergent.

Definition: (metric) ball. If (X, d) is a metric space, $x \in X$ and $r > 0$, then

$$B_r(x) = \{y \in X \,|\, d(x, y) < r\}$$

is said to be the ball with center x and radius r.

Definition: Bounded subset. If (X, d) is a metric space and $A \subset X$, then we say that A is bounded if A is contained in some metric ball $B_r(x)$.

Theorem: In a metric space (X, d) any Cauchy sequence is bounded.

Proof. Suppose that x_n is a Cauchy sequence. By definition of a Cauchy sequence, there exists some N such that if $m,\, n \geq N$, then

$$d(x_n, x_m) < 1 \,.$$

Set

$$r = 1 + \max\{d(x_N, x_i) \,|\, i < N\} \,.$$

We claim that

$$\{x_n\} \subset B_r(x_N) \,.$$

Since $r \geq 1$ and $d(x_N, x_n) < 1$ for $n \geq N$ we only need to see that $x_n \in B_r(x_N)$ for $n < N$. This follows from that $d(x_N, x_n) < r$ when $n < N$ by definition of r. $\qquad\square$

Bolzano - Weirstrass theorem: Any bounded sequence of real numbers have a convergent subsequence. This theorem does not hold for a general metric space but it holds if the metric space is compact. To discuss this we need the notion of what an open subset of a metric space is.

Definition (Open subset): Let (X, d) be a metric space. We say that O is an open subset of X if for all $x \in O$, there exists an $r > 0$ such that $B_r(x) \subset O$.

Note that $\emptyset$ (the empty set) and X are both open.

On subsets of a set X we have the following operations.

- Union of two or more subsets.
  If U_1 and U_2 are subsets, then $U_1 \cup U_2$ is the union. So
  $$U_1 \cup U_2 = \{x \in X \mid x \in U_1 \text{ or } x \in U_2 \text{ or both}\}.$$
  Similarly, for union of more than two subsets.
- Intersection of two or more subsets.
  If U_1 and U_2 are subsets, then $U_1 \cap U_2$ is the intersection. So
  $$U_1 \cap U_2 = \{x \in X \mid x \in U_1 \text{ and } x \in U_2\}.$$
  Similarly, for intersection of more than two subsets.
- Complement of a subset U.
  $X \setminus U$ is all the elements of X that are not in U.

Example: $X = \mathbf{R}$, $A = (0, 3)$, $B = (-1, 2)$ and $C = (0, 2)$.
$A \cup B = (0, 3)$.
$A \cap B = (0, 2)$.
$X \setminus A = (-\infty, 0] \cup [3, \infty)$.
$C \subset B$.

Union and intersections of families of subsets

- Union of families.
  If U_α is a family of subsets, then $\cup_\alpha U_\alpha$ is the union of all the subsets. So
  $$\cup_\alpha U_\alpha = \{x \in X \mid x \in U_\alpha \text{ for some } \alpha\}.$$
- Intersection of families.
  If U_α is a family of subsets, then $\cap_\alpha U_\alpha$ is the intersection of all the subsets. So
  $$\cap_\alpha U_\alpha = \{x \in X \mid x \in U_\alpha \text{ for all } \alpha\}.$$

Example: $X = \mathbf{R}$, $U_n = (-\frac{1}{n}, \frac{1}{n})$, where $n \in \mathbf{N}$, then
$$\cup_n U_n = (-1, 1) \text{ and } \cap_n U_n = \{0\}.$$

Lemma: For a set X and subsets A, B we have $A = B$ if and only if $A \subset B$ and $B \subset A$.

Lemma: For a set and subset A, B and A_α we have
(1) $X \setminus (X \setminus A) = A$.
(2) $X \setminus \cup_\alpha A_\alpha = \cap_\alpha (X \setminus A_\alpha)$.
(3) $X \setminus \cap_\alpha A_\alpha = \cup_\alpha (X \setminus A_\alpha)$.

Proof. To prove the first of these claim that $X \setminus (X \setminus A) = A$ we need to show two directions. Suppose $x \in A$, then $x \notin X \setminus A$ and therefore $x \in X \setminus (X \setminus A)$. Conversely, if $x \in X \setminus (X \setminus A)$, then $x \notin X \setminus A$ and therefore $x \in A$.

To prove the second claim observe that if $x \in X \setminus \cup_\alpha A_\alpha$, then $x \notin \cup_\alpha A_\alpha$ so x is not in any of the A_α's. Therefore x must be in all the $X \setminus A_\alpha$ and hence in the intersection of those so $x \in \cap_\alpha (X \setminus A_\alpha)$. This show that $X \setminus \cap_\alpha A_\alpha \subset \cap_\alpha (X \setminus A_\alpha)$. To show the other direction suppose that $x \in \cap_\alpha (X \setminus A_\alpha)$. This means that for all α we have that $x \notin A_\alpha$. Therefore, $x \notin \cup_\alpha A_\alpha$ and hence $x \in X \setminus (\cup_\alpha A_\alpha)$. This show the other direction.

Finally, to prove the third claim observe that if $x \in X \setminus \cap_\alpha A_\alpha$, then $x \notin \cap_\alpha A_\alpha$ and so there exists some α so that $x \in X \setminus A_\alpha$. In other words, $x \in \cup_\alpha (X \setminus A_\alpha)$. This show one direction. To see the other direction observe that if $x \in \cup_\alpha (X \setminus A_\alpha)$, then there exists some α so that $x \in X \setminus A_\alpha$. It follows that $x \notin A_\alpha$ and hence $x \notin \cap_\alpha A_\alpha$ but instead $x \in X \setminus \cap_\alpha A_\alpha$. This show the other direction and completes the proof of the lemma. $\qquad\square$

Lemma: In a metric space any ball $B_r(x)$ is an open subset.

Proof. Suppose that $y \in B_r(x)$, and let $s = r - d(x, y)$. Note that since $y \in B_r(x)$ we have that $d(x, y) < r$ and so $s > 0$. We will show that $B_s(y) \subset B_r(x)$. To see that assume that $z \in B_s(y)$ we then have that $d(y, z) < s$ and so by the triangle inequality

$$d(z, x) \leq d(z, y) + d(y, x) < s + d(y, x) = (r - d(x, y)) + d(y, x) = r.$$

This shows the claim. $\qquad\square$

Lemma: In a metric space if O_α are open subsets, then

$$\cup_\alpha O_\alpha$$

is open.

Proof. See Pset. $\qquad\square$

Lemma: In a metric space if $O_1, \cdots, O_n$ are finitely many open subsets, then

$$O_1 \cap \cdots \cap O_n$$

is open.

Proof. Suppose that $x \in O_1 \cap \cdots \cap O_n$, then x lies in each O_i. For each i, there exists an $r_i > 0$, such that $B_{r_i}(x) \subset O_i$. Let $r = \min r_i$, then for each i we have that $B_r(x) \subset B_{r_i}(x) \subset O_i$ so $B_r(x)$ is a subset of each O_i and hence $B_r(x) \subset O_1 \cap \cdots \cap O_n$. This shows the claim. $\qquad\square$

Warning: Intersection of infinitely many open subsets may not be open!!!!

Example: $X = \mathbf{R}$ and for each natural number let O_n be the open set $O_n = (-\frac{1}{n}, \frac{1}{n})$, then

$$\cap_n O_n = \{0\}\,.$$

So the intersection of these infinitely many open subsets is not open.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 13

Definition (Closed subsets): Let (X, d) be a metric space. We say that C is a closed subset of X if the complement $X \setminus C$ is open.

Note that $\emptyset$ (the empty set) and X are both closed.

Examples:
- $(0, 1)$ is not a closed subset of $\mathbf{R}$.
- $\{0\}$ is a closed subset of $\mathbf{R}$.
- $[0, 1]$ is a closed subset of $\mathbf{R}$.
- $[0, 1] \times [0, 1]$ is a closed subset of $\mathbf{R}^2$.

Lemma: Let (X, d) be a metric space and $r > 0$, then
$$A_r = \{y \mid d(x, y) > r\}$$
is open. Equivalently, $\bar{B}_r(x) = \{y \mid d(x, y) \leq r\}$ is closed.

Proof. Suppose that $y \in A_r$, then $d(y, x) > r$ and if we set $s = d(y, x) - r$, then $s > 0$. Moreover, if $z \in B_s(y)$, then by the triangle inequality
$$d(x, y) \leq d(y, z) + d(z, x).$$
So
$$r < r + s - d(y, z) \leq d(x, y) - d(y, z) \leq d(z, x).$$
This show that $B_s(z) \subset A_r$ and so A_r is open. $\qquad\qquad \square$

There is an equivalent way of defining closed subsets and that comes from the next theorem.

Theorem: A subset C of a metric space (X, d) is closed if and only if for all convergent sequences x_n with all x_n in C also the limit is in C.

Proof. Suppose first that A is closed and let x_n be a convergent sequence is A with limit x we need to show that $x \in A$. Since A is closed the complement is open and if $x \in X \setminus A$, then there exists some $r > 0$ so $B_r(x) \subset X \setminus A$ and therefore for all $y \in A$ we would have that $d(y, x) \geq r$. This contradict that $x_n \to x$ and $x_n \in A$.

We also need to show the converse. So suppose that A is a subset with the property that for all sequences in A that are convergent in X the limit is in A. We will show that A is closed or equivalent that the complement is open. If the complement is not open, then there exists an $x \in X \setminus A$ such that no ball around x is entirely contained in the complement. Therefore for each n there exists an $x_n \in A$. This sequence converges to x which was assumed not to be in A contradicting that A contained all limits of sequences in A and therefore the complement must be open and A itself closed. □

For union and intersection of closed subsets we have the following:

Theorem:
- Union: If C_α is a family of closed subsets, then $\cap_\alpha C_\alpha$ is also closed.
- Intersection: If $C_1, \cdots, C_n$ are closed subsets, then $C_1 \cup \cdots \cup C_n$ is also closed.

Proof. There are several ways of proving this. The easiest is probably straight from the definition using the operations on sets. For the first claim we need to show that the complement of $\cap_\alpha C_\alpha$ is open. Using the operations of sets we have that

$$X \setminus \cap_\alpha C_\alpha = \cup_\alpha (X \setminus C_\alpha).$$

Since each $X \setminus C_\alpha$ are open this is the union of open sets and therefore open. This shows the first claim.

To see the second claim we argue similarly. We want to show that $C_1 \cup \cdots \cup C_n$ is closed or, equivalently, $X \setminus (C_1 \cup \cdots \cup C_n)$ is open. However,

$$X \setminus (C_1 \cup \cdots \cup C_n) = (X \setminus C_1) \cap \cdots \cap (X \setminus C_n),$$

where the last is the intersection of finitely many open sets and therefore open. This show the second claim. □

Warning: Union of infinitely many closed sets may not be closed!!!

Definition (Cover, open cover and finite sub-cover): If A is a subset of X, then a cover of A is a collection collection of subsets U_α of X so that

$$A \subset \cup_\alpha U_\alpha.$$

We say that a $U_{\alpha_1}, \cdots, U_{\alpha_n}$ is a finite sub-cover if also $\{U_{\alpha_i}\}_i$ is a cover.

If (X, d) is a metric space and all the U_α are open, then we say that $\{U_\alpha\}_\alpha$ is an open cover.

Example: If $X = \mathbf{R} \times \mathbf{R}$, then $A_n = (-n, n) \times (-n, n)$ is an open cover of X.

Example: Note that in the example where $X = \mathbf{R} \times \mathbf{R}$ and $A_n = (-n, n) \times (-n, n)$ is an open cover, then there is no finite sub-cover. On the other hand if $A \subset \mathbf{R} \times \mathbf{R}$ is bounded, then for n sufficiently large $A \subset A_n$ so for A, there is a finite sub-cover of this cover.

Definition (Compact subset): If (X, d) is a metric space and A is a subset, then we say that A is compact if each open cover has a finite sub-cover.

Example: If (X, d) is $\mathbf{R}$ with the usual metric and $A = (0, 1)$, then $A_n = (\frac{1}{n}, 1)$ is an open cover of A but there is no finite sub-cover of $\{A_n\}_n$ that covers A.

Theorem: $[a, b] \subset \mathbf{R}$ is compact.

Proof. We will show this next time. □

Theorem: If (X, d) is a metric space and A a compact subset, then A is closed and bounded.

Proof. Suppose first that A is not closed. We will show that this leads to a contradiction. If it is not closed, then there exists a convergent sequence $x_n \in A$ with limit x not in A. Set

$$O_n = \left\{ y \mid d(x, y) > \frac{1}{n} \right\}$$

. By the earlier lemma these are open sets. Since $\cup_n A_n = X \setminus \{x\}$ and x is assumed not to be in A we indeed have that A_n is an open cover of A. Since $A_n \subset A_{n+1}$ any finite cover of A_n's would be contained in A_N for some large N but this would imply that for all $y \in A$ we would have that $d(x, y) > \frac{1}{N}$ contradicting that $x_n \in A$ and $x_n \to x$. This show that the limit x is in A.

Since A is compact,

$$X = \cup_y B_r(y)$$

and each $B_1(x)$ is open, then finitely many of these covers A. Say $A \subset B_1(y_1) \cup \cdots \cup B_1(y_n)$. Set $r = 1 + \max_i \{d(y_1, y_i)\}$. It follows by the triangle inequality that $A \subset B_r(y_1)$. Hence, A is bounded. □

Warning: The converse is not the case!!! There are closed a bounded subsets of metric spaces that are not compact.

If $(X, d) = (0, 1)$ with the usual metric, then X is closed and bounded but it is not compact.

Here is a more illuminating example:

Example: Let $X = C([0, 1])$ be the set of continuous functions on the unit interval $[0, 1]$. We equip X with the metric where

$$d(f, g) = \max_x |f(x) - g(x)|.$$

Let $f_n(x)$ be the sequence of continuous functions on $[0, 1]$ given by that

$$f_n(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \frac{1}{n+1} \\ 1 - n(n+1)\left(x - \frac{1}{n+1}\right) & \text{if } \frac{1}{n+1} \leq x \leq \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

We have the f_n is a bounded sequence. After all they all lies in the metric ball $B_2(0)$ where 0 is the zero function. That is, the function on $[0, 1]$ that is identically equal to zero. However, the sequence f_n does not have a convergent subsequence (and does not even have a subsequence that is a Cauchy sequence). Indeed, for any $m \neq m$ we have that

$$d(f_m, f_n) = 1.$$

Note also that the (closed) ball $A = \bar{B}_1(0)$ is closed and bounded but not compact. It is not compact because for the balls $\cup_f B_{\frac{1}{2}}(f)$ finitely many does not cover A. If finitely many did cover A, then for one such ball say $B_{\frac{1}{2}}(f)$ infinitely many f_n's would lie in it but any two elements in such a ball would have distance < 1 showing that there could at most be one f_n in such a ball.

Theorem: If (X, d) is a metric space and A a compact subset, then any closed subset C contained in A is also compact.

Proof. Let O_α be a open cover of C. Since C is closed $X \setminus C$ is open and so $\{O_\alpha\}$ together with $X \setminus C$ is an open cover of A and hence finitely many of those say $O_1, \cdots, O_n, X \setminus C$ covers A. Since $X \setminus C$ contains no elements in C it follows that $C \subset O_1 \cup \cdots \cup O_n$ and thus C is compact. $\qquad\square$

Bolzano-Weirstrass theorem for metric spaces.

Theorem: If (X, d) is a metric space and A a compact subset, then any sequence in A has a convergent subsequence.

Proof. We will show this next time.   □

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
       TBB can be downloaded at:
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
       (screen-optimized)
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
       (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

SPRING 2025 - 18.100B/18.1002

TOBIAS HOLCK COLDING

Lecture 14

Definition (Compact subset): If (X, d) is a metric space and A is a subset, then we say that A is compact if each open cover has a finite sub-cover.

Theorem 0: If (X, d) is a metric space and A a compact subset, then A is closed and bounded.

Warning: The converse is not the case!!! There are closed a bounded subsets of metric spaces that are not compact.

Example: If $(X, d) = (0, 1)$ with the usual metric, then X is closed and bounded but it is not compact.

Here is a more illuminating example:

Example: Let $X = C([0, 1])$ be the set of continuous functions on the unit interval $[0, 1]$. We equip X with the metric where

$$d(f, g) = \max_x |f(x) - g(x)|.$$

Let $f_n(x)$ be the sequence of continuous functions on $[0, 1]$ given by that

$$f_n(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \frac{1}{n+1} \\ 1 - n(n+1)\left(x - \frac{1}{n+1}\right) & \text{if } \frac{1}{n+1} \leq x \leq \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

We have the f_n is a bounded sequence. After all they all lies in the metric ball $B_2(0)$ where 0 is the zero function. That is, the function on $[0, 1]$ that is identically equal to zero. However, the sequence f_n does not have a convergent subsequence (and does not even have

a subsequence that is a Cauchy sequence). Indeed, for any $m \neq m$ we have that

$$d(f_m, f_n) = 1 \,.$$

Note also that the (closed) ball $A = \bar{B}_1(0)$ is closed and bounded but not compact. It is not compact because for the balls $\cup_f B_{\frac{1}{2}}(f)$ finitely many does not cover A. If finitely many did cover A, then for one such ball say $B_{\frac{1}{2}}(f)$ infinitely many f_n's would lie in it but any two elements in such a ball would have distance < 1 showing that there could at most be one f_n in such a ball.

Using what we have shown in earlier lectures one can show the following:

Theorem 1: In $\mathbf{R}^n$, a subset is compact if and only if it is closed and bounded.

In a general metric space this is not the case as the above examples shows.

We won't show this theorem here but instead we will show a version of the Bolzano-Weirstrass theorem for metric spaces. This is the next theorem.

Theorem 2: If (X, d) is a metric space and A a compact subset, then any sequence in A has a convergent subsequence.

Before proving Theorem 2 we will need some results:

Lemma: Let (X, d) be a compact metric space if C_α is a family of closed (decreasing) nested subsets. That is, closed subsets so that $C_{n+1} \subset C_n$. If all C_n are non-empty, then

$$\cap_n C_n \neq \emptyset \,.$$

Proof. Set $O_\alpha = X \setminus C_\alpha$, then each O_α is open. If $\cap_n C_n \neq \emptyset$, then

$$\cup_n O_\alpha = X \,.$$

Therefore, finitely many of the O_n's cover X by compactness. Denote these by O_i for $i = 1, \cdots, k$. Since

$$O_1 \cup \cdots \cup O_k = X$$

it follows that

$$C_1 \cap \cdots \cap C_k = \emptyset \,.$$

However, by the nested property one of these k closed subsets is the smallest, say C_k and therefore $C_1 \cap \cdots \cap C_k = C_k$. Contradicting that the intersection is empty. $\qquad\square$

Before stating the next results recall that in a metric space (X, d) the set $\bar{B}_r(x) = \{y \in X \mid d(x, y) \leq r\}$ is closed and is referred to as the closed ball. The above lemma gives the following useful corollary:

Corollary: Let (X, d) be a compact metric space and suppose that $B_{r_n}(x_n)$ is a family of balls with centres x_n and radii $r_n > 0$, where $r_n \to 0$ and $\bar{B}_{r_{n+1}}(x_{n+1}) \subset \bar{B}_{r_n}(x_n)$. Then

$$\cap_n \bar{B}_{r_n}(x_n) = \{x\}\,.$$

That is, the intersection is non-empty and consists of a single point.

Proof. Set

$$A = \cap_n \bar{B}_{r_n}(x_n)\,.$$

Observe first that for each n we have that $x_n \in \bar{B}_{r_n}(x_n)$ so from the lemma above we have that A is non-empty. We claim that A consists of just one element. Suppose that $x, y \in A$, for any integer n we have that

$$x,\, y \in \bar{B}_{r_n}(x_n)\,,$$

and so by the triangle inequality

$$d(x, y) \leq d(x, x_n) + d(x_n, y) \leq r_n + r_n = 2\, r_n\,.$$

Since this holds for all n we see that $d(x, y) = 0$ and so there is at most one such point. $\qquad\square$

Proof. (of Theorem 2.) Suppose that x_n is a sequence in a compact subset A of a metric space. Fix $r > 0$ and write

$$A \subset_{x \in A} B_r(x)\,.$$

Since A is compact finitely many of these cover A. This means that in one of these balls, say $B_r(y_1)$, there are infinitely many x_n's. From here on and out we will focus on this ball. Since $A \cap \bar{B}_r(y)$ is a closed subset of a compact set we can now cover $\bar{B}_r(y)$ by balls of radius $\frac{r}{4}$. By compactness finitely many of these sub-balls cover the ball $\bar{B}_r(y)$. In one of those sub-balls there are also infinitely many x_n's. Fix such a sub-ball and call it $\bar{B}_{\frac{r}{4}}(y_2)$. We have that

$$\bar{B}_{\frac{r}{2}}(y_2) \subset \bar{B}_{2r}(y_1)$$

and that infinitely many x_n's belongs to $B_{\frac{r}{4}}(y_2)$. If the original $r = 1$ gives after repeating this process i times balls $B_{4^{1-i}}(y_i)$ so that

$$\cdots \subset \bar{B}_{2\,4^{1-i}}(y_i) \subset \cdots \bar{B}_{2\,4^{-1}}(y_2) \subset \bar{B}_2(y_1)\,.$$

where each of these balls contains infinitely many elements from the original sequence. Since the radii of this sequence converges to zero this sequence satisfies the assumptions of the corollary we have from the corollary that

$$\cap B_{2\,4^{1-i}}(y_i) = \{x\}\,.$$

Moreover, we can pick a subsequence x_{n_k} of the original sequence such that

$$x_{n_k} \subset \bar{B}_{2\,4^{1-k}}(y_k)\,.$$

It follows that this subsequence converges to x.   □

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
    TBB can be downloaded at:
    https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
    (screen-optimized)
    https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
    (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 15

Definition: If $f : \mathbf{R} \to \mathbf{R}$ is a function, then we say that f is differentiable at x_0 if the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. (Note that in this fraction x is assumed to be $\neq x_0$.) When the limit exists, then we say that the function f is differentiable at x_0 and that its derivative at x_0 is the limit. In this case we denote the derivative at x_0 by $f'(x_0)$.

Examples:

(1) Constant functions. Suppose that $f(x) = c$ for some constant $c \in \mathbf{R}$, then

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0 \,.$$

It follows that the limit exists and is zero and so f is differentiable at all points and the derivative is zero.

(2) Linear functions. Suppose that $f(x) = x$, then

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1 \,.$$

It follows that the limit exists and is one and so f is differentiable at all points and the derivative is one.

These are just two examples where we computed the derivative directly from the definition. How do we compute the derivative of a general function?

For that there are some tools:

- Sum rule.
- Product rule.
- Quotient rule.
- Chain rule.

Once we know how to compute the derivative of a function, then we would like to understand the function using information about its derivative. For that we have the following tools:

- Mean value theorem.
- L'Hopital's rule.
- Taylor expansion.

Before getting to how to use the derivative we need to be able to compute it. For that it will be useful to note the following:

Lemma: If f is differentiable at x_0, then f is continuous at x_0.

Proof. Since f is differentiable at x_0 we have that

$$\frac{f(x) - f(x_0)}{x - x_0} \to f'(x_0).$$

Therefore, there exist $\delta_1 > 0$ such that if $|x - x_0| < \delta_1$, then

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < 1$$

or, equivalently,

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| < |x - x_0|.$$

Therefore, for $|x - x_0| < \delta_1$ we have

$$|f(x) - f(x_0)| < (|f'(x_0)| + 1)|x - x_0|.$$

Given $\epsilon > 0$, set

$$\delta = \min\left\{ \delta_1, \frac{\epsilon}{|f'(x_0)| + 1} \right\}.$$

It follows that if $|x - x_0| < \delta$, then

$$|f(x) - f(x_0)| < \epsilon.$$

This show that f is continuous at x_0. $\qquad\square$

Example: On the real line suppose that f is the function given by that $f(0) = 0$ and for all other x

$$f(x) = x \sin \frac{1}{x}.$$

This is an example of a function that is continuous at zero but not differentiable at zero. It is not differentiable at zero because it fluctuate too much near zero. To see that it is continuous at zero we will use that $|\sin t| \leq 1$ for all t. Indeed using that it is easy to see

that f is continuous at zero. Next we will see that it is not differentiable at zero. To see that we look at the difference quotient

$$\frac{f(x) - f(0)}{x - 0} = \frac{x \sin \frac{1}{x} - 0}{x - 0} = \sin \frac{1}{x}.$$

As $x \to 0$ this function fluctuate between -1 and 1 so it does not have a limit and therefore the original function f is not differentiable at zero.

Example: If we dampen the fluctuation of the function given in the previous example further, then we get a differentiable function at zero even if it still fluctuate but just not as much. This is done in the following example. Suppose that f is the function that is given by that $f(0) = 0$ and for all other x

$$f(x) = x^2 \sin \frac{1}{x}.$$

Again we form the difference quotient

$$\frac{f(x) - f(0)}{x - 0} = \frac{x^2 \sin \frac{1}{x} - 0}{x - 0} = x \sin \frac{1}{x}.$$

In this case we see that as $x \to 0$, then $x \sin \frac{1}{x} \to 0$ and so the function is differentiable at zero and the derivative there is zero.

The following is very useful to compute the derivative of many functions:

Theorem: If f, g are functions on $\mathbf{R}$ that both are differentiable at x_0, then
- (Sum rule.)
$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$
- (Leibniz's rule.)
$$(f\,g)(x_0) = f'(x_0)\,g(x_0) + f(x_0)\,g'(x_0).$$
- (Quotient rule.) If also $g(x_0) \neq 0$, then
$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)\,g(x_0) - f(x_0)\,g'(x_0)}{g^2(x_0)}.$$

Proof. To prove the sum rule consider the difference quotient

$$\frac{(f + g)(x) - (f + g)(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \to f'(x_0) + g'(x_0)$$

This show the sum rule.

To prove the Leibniz rule we form the difference quotient

$$\frac{(f\,g)(x) - (f\,g)(x_0)}{x - x_0}.$$

We rewrite this using a trick we have used before in other settings. Namely, we can write this as

$$\frac{(f\,g)(x) - (f\,g)(x_0)}{x - x_0} = \frac{f(x)\,g(x) - f(x)\,g(x_0) + f(x)\,g(x_0) - f(x_0)\,g(x_0)}{x - x_0}$$

$$= f(x)\,\frac{g(x) - g(x_0)}{x - x_0} + \frac{f(x) - f(x_0)}{x - x_0}\,g(x_0) \to f(x_0)\,g'(x_0) + f'(x_0)\,g(x_0)\,.$$

(Here we used that by the continuity lemma above $f(x) \to f(x_0)$.) This proves Leibniz's rule.

Finally, to prove the quotient rule we observe first that since g is differentiable at x_0 it is continuous at x_0 and therefore (since $g(x_0) \neq 0$) when x is close to x_0 we have that $g(x) \neq 0$. Moreover, we have that

$$\frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} = \frac{f(x)\,g(x_0) - f(x_0)\,g(x)}{(x - x_0)\,g(x)\,g(x_0)}$$

$$= \frac{f(x)\,g(x_0) - f(x_0)\,g(x_0)}{(x - x_0)\,g(x)\,g(x_0)} + \frac{f(x_0)\,g(x_0) - f(x_0)\,g(x)}{(x - x_0)\,g(x)\,g(x_0)}$$

$$= \frac{1}{g(x)}\,\frac{f(x) - f(x_0)}{x - x_0} + \frac{f(x_0)}{g(x)\,g(x_0)}\,\frac{g(x) - g(x)}{x - x_0}$$

$$\to \frac{f'(x_0)}{g(x_0)} + \frac{f(x_0)\,g'(x_0)}{g^2(x_0)}\,.$$

From this the claim easily follows.   □

Leibniz's rule is named after Gottfried Wilhelm Leibniz (1646 - 1716). Leibniz [from Wikipedia] was a German polymath active as a mathematician, philosopher, scientist and diplomat who is credited, alongside Sir Isaac Newton, with the creation of calculus in addition to many other branches of mathematics, such as binary arithmetic and statistics. Leibniz has been called the "last universal genius" due to his vast expertise across fields, which became a rarity after his lifetime with the coming of the Industrial Revolution and the spread of specialised labor.

Theorem: (Chain rule.) If $f : [a, b] \to [c, d]$ and $g : [c, d] \to \mathbf{R}$ are functions, where f is differentiable at x_0 and g differentiable at $y_0 = f(x_0)$, then the composition $g \circ f$ is differentiable at x_0 and the derivative at x_0 is

$$(g \circ f)'(x_0) = g'(y_0)\,f'(x_0)\,.$$

Proof. Set $y = f(x)$ and $y_0 = f(x_0)$. Assume first that $f'(x_0) \neq 0$. In this case for $x \neq x_0$ but close to x_0 we have that $y \neq y_0$ and we can write the difference quotient as follows. We have that

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(y) - g(y_0)}{y - y_0}\,\frac{f(x) - f(x_0)}{x - x_0}\,.$$

Since f is differentiable at x_0 as $x \to x_0$ we have that

$$\frac{f(x) - f(x_0)}{x - x_0} \to f'(x_0) \, .$$

Moreover, when $x \to x_0$ we have that $f(x) = y \to f(x_0) = y_0$ by the continuity lemma above. It follows that when $x \to x_0$ we have that

$$\frac{g(y) - g(y_0)}{y - y_0} \to g'(y_0) \, .$$

Combining this gives that

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(y) - g(y_0)}{y - y_0} \frac{f(x) - f(x_0)}{x - x_0} \to g'(y_0) \, f'(x_0) \, .$$

This proves the chain rule when $f'(x_0) \neq 0$. When $f'(x_0) = 0$ we argue as above but have to be more careful as in this case we can have that $y = y_0$ even when $x \neq x_0$. For x where $y = y_0$ the difference quotient is zero and where $y \neq y_0$ we can argue as above and rewrite the difference quotient as the product of two factors. In either case we get that the limit is zero proving the remaining case of the chain rule. $\qquad\square$

Lemma: Let $f : [a, b] \to \mathbf{R}$ be a differentiable function and suppose that $a < x_0 < b$ and that f has a local maximum or minimum at x_0, then

$$f'(x_0) = 0 \, .$$

Proof. Suppose that x_0 is a local maximum. The proof when x_0 is a local minimum. It follows from the assumption that for all x near x_0

$$f(x) - f(x_0) \leq 0 \, .$$

Therefore, when $x > x_0$ we have that

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \, ,$$

whereas when $x < x_0$ we have that for the difference quotient

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0 \, .$$

Since the limit is the same whether x converges to x_0 from the the left (negative side) or from the right (positive side) it follows that $f'(x_0) = 0$ as claimed. $\qquad\square$

Theorem: (Rolle's theorem.) Let $f : [a, b] \to \mathbf{R}$ be a differentiable function with $f(a) = f(b)$, then there exists a x_0 between a and b such that

$$f'(x_0) = 0 \, .$$

Proof. There are three cases to consider:

(1) f is constant equal to $f(a)$.

(2) For some x between a and b we have that $f(x) > f(a)$.

(3) For some x between a and b we have that $f(x) < f(a)$.

In the first case the function is constant and the derivate is zero everywhere. The second and third cases are similar so we will just argue in the second case. In the second case by the extreme value theorem there exists some x_0 such that $f(x_0) = \max f > f(a)$. It now follows from the previous lemma that $f'(x_0) = 0$. □

Theorem: (Mean value theorem.) Let $f : [a, b] \to \mathbf{R}$ be a differentiable function, then there exists a x_0 between a and b such that

$$f'(x_0) = \frac{f(b) - f(a)}{b - a} \,.$$

.

Proof. Consider the function g given by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} \, (x - a) \,.$$

Observe that for g we have $g(a) = g(b)$ and so Rolle's theorem applies and we have that there exists some x_0 where $g'(x_0) = 0$. Since

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \,,$$

the claim follows. □

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 16

Last time we defined what it means for a function to be differentiable. This is the following:

Definition: If $f : \mathbf{R} \to \mathbf{R}$ is a function, then we say that f is differentiable at x_0 if the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. (Note that in this fraction x is assumed to be $\neq x_0$.) When the limit exists, then we say that the function f is differentiable at x_0 and that its derivative at x_0 is the limit. In this case we denote the derivative at x_0 by $f'(x_0)$.

One of the first things we showed about differentiable function was that they are continuous:

Lemma: If f is differentiable at x_0, then f is continuous at x_0.

We also established some very useful rules for computing the derivative of functions that are constructed from other functions whose derivative we know:

Theorem: If f, g are functions on $\mathbf{R}$ that both are differentiable at x_0, then
- (Sum rule.)
$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$
- (Leibniz's rule.)
$$(f\,g)(x_0) = f'(x_0)\,g(x_0) + f(x_0)\,g'(x_0).$$
- (Quotient rule.) If also $g(x_0) \neq 0$, then
$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)\,g(x_0) - f(x_0)\,g'(x_0)}{g^2(x_0)}.$$

Finally, for the composition of functions we have the chain rule:

Theorem: (Chain rule.) If $f : [a,b] \to [c,d]$ and $g : [c,d] \to \mathbf{R}$ are functions, where f is differentiable at x_0 and g differentiable at $y_0 = f(x_0)$, then the composition $g \circ f$ is differentiable at x_0 and the derivative at x_0 is

$$(g \circ f)'(x_0) = g'(y_0)\, f'(x_0)\,.$$

Now that we know how to compute the derivative of many functions we will be interested in using the derivative to describe the growth or decay of a function. The first step towards this is the next lemma.

Before stating it recall that a function $f : \mathbf{R} \to \mathbf{R}$ has a local maximum at x_0 if there exists a $\delta > 0$ such that

$$f(x_0) = \max_{[x_0-\delta,x_0+\delta]} f\,,$$

and similarly for a local minimum.

Lemma: Let $f : [a,b] \to \mathbf{R}$ be a differentiable function and suppose that $a < x_0 < b$ and that f has a local maximum or minimum at x_0, then

$$f'(x_0) = 0\,.$$

Proof. Suppose that x_0 is a local maximum. The proof when x_0 is a local minimum is similar. It follows from the assumption that for all x near x_0

$$f(x) - f(x_0) \leq 0\,.$$

Therefore, when $x > x_0$ we have that

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0\,,$$

whereas when $x < x_0$ we have that for the difference quotient

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0\,.$$

Since the limit is the same whether x converges to x_0 from the the the left (negative side) or from the right (positive side) it follows that $f'(x_0) = 0$ as claimed.   □

We can now use this lemma to establish the following very useful result:

Theorem: (Rolle's theorem.) Let $f : [a, b] \to \mathbf{R}$ be a differentiable function with $f(a) = f(b)$, then there exists a x_0 between a and b such that

$$f'(x_0) = 0 \,.$$

Proof. There are three cases to consider:

(1) f is constant equal to $f(a)$.
(2) For some x between a and b we have that $f(x) > f(a)$.
(3) For some x between a and b we have that $f(x) < f(a)$.

In the first case the function is constant and the derivate is zero everywhere. The second and third cases are similar so we will just argue in the second case. In the second case by the extreme value theorem there exists some x_0 such that $f(x_0) = \max f > f(a)$. It now follows from the previous lemma that $f'(x_0) = 0$. $\square$

Rolle's theorem can then be used to show both the mean value theorem and the Cauchy mean value theorem:

Theorem: (Mean value theorem.) Let $f : [a, b] \to \mathbf{R}$ be a differentiable function, then there exists a x_0 between a and b such that

$$f'(x_0) = \frac{f(b) - f(a)}{b - a} \,.$$

.

Proof. Consider the function g given by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} (x - a) \,.$$

Observe that for g we have $g(a) = g(b)$ and so Rolle's theorem applies and we have that there exists some x_0 where $g'(x_0) = 0$. Since

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \,,$$

the claim follows. $\square$

Theorem: (Cauchy mean value theorem.) Let $f, g : [a, b] \to \mathbf{R}$ be differentiable functions, then there exists a x_0 between a and b such that

$$f'(x_0) \left[g(b) - g(a) \right] = g'(x_0) \left[f(b) - f(a) \right] \,.$$

In particular, if $g(b) - g(a) \neq 0$, then

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)} \,.$$

Proof. Consider the function

$$h(x) = f(x) \left[g(b) - g(a) \right] - g(x) \left[f(b) - f(a) \right].$$

Note that

$$h(a) = f(a) \left[g(b) - g(a) \right] - g(a) \left[f(b) - f(a) \right] = f(a) \, g(b) - g(a) \, f(b) \,.$$
$$h(b) = f(b) \left[g(b) - g(a) \right] - g(b) \left[f(b) - f(a) \right] = f(a) \, g(b) - g(a) \, f(b) \,.$$

Therefore, by Rolle's theorem, there exists x_0 between a and b such that $h'(x_0) = 0$. Since

$$h'(x) = f'(x) \left[g(b) - g(a) \right] - g'(x) \left[f(b) - f(a) \right]$$

this shows the claim. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

We observe that the Cauchy mean value theorem implies the earlier mean value theorem. Namely, if we let the second function g be $g(x) = x$, then $g'(x) = 1$ and $g(b) - g(a) = b - a$. Therefore, the Cauchy mean value theorem becomes

$$f'(x_0) \, (b - a) = f'(x_0) \, (g(b) - g(a)) = g'(x_0) \, (f(b) - f(a)) = f(b) - f(a) \,,$$

which is the earlier mean value theorem.

The next two rules are useful to establishing the limit of a faction of function when the denominator either tend to zero or infinity.

Theorem: (L'Hopital's rule, version 1.) Let f, $g : (a, b) \to \mathbf{R}$ be differentiable functions with $g(x) \neq 0$ and $g'(x) \neq 0$ for all x, assume that

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0 \,.$$

If

$$\lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists, then

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

Proof. We will see that this is an easy consequence of the Cauchy mean value theorem. By assumption given $\epsilon > 0$, there exists $\delta > 0$ such that if $a < x < \delta$, then

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \epsilon \,.$$

By the Cauchy mean value theorem we have for any y with $a < y < x$ that there exist z with $y < z < x$ so that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)} \,.$$

We therefore have that

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - L \right| < \epsilon \,.$$

By letting $y \to 0$ we see that

$$\left| \frac{f(x)}{g(x)} - L \right| \le \epsilon \,.$$

Since this holds for all ϵ we get that

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

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Theorem: (L'Hopital's rule, version 2.) Let $f, g : (a, b) \to \mathbf{R}$ be differentiable functions with $g(x) \ne 0$ and $g'(x) \ne 0$ for all x, assume that

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

If

$$\lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists, then

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

Proof. Given $\epsilon > 0$, since $\frac{f'(x)}{g'(x)} \to L$ as as $x \to a$ we have that there exists a $\delta > 0$ such that if $a < x < a + 2\,\delta$, then

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \epsilon \,.$$

Set $x_1 = a + \delta$. For a given $x \in (a, x_1)$, there exists $x_0 \in (x, x_1)$ such that

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(x_1) - f(x)}{g(x_1) - g(x)} \,.$$

It follows that

$$\left| \frac{f(x_1) - f(x)}{g(x_1) - g(x)} - L \right| < \epsilon \,.$$

By dividing the nominator and denominator of the fraction in this expression by $g(x)$ we get

$$\left| \frac{\frac{f(x)}{g(x)} - \frac{f(x_1)}{g(x)}}{1 - \frac{g(x_1)}{g(x)}} - L \right| < \epsilon \,.$$

This implies that

$$\left| \frac{f(x)}{g(x)} - \frac{f(x_1)}{g(x_1)} - L \left(1 - \frac{g(x_1)}{g(x)} \right) \right| < \epsilon \left(1 - \frac{g(x_1)}{g(x)} \right) \,.$$

Since this holds for all $x \in (a, a + \delta)$ and $g(x) \to \infty$ as $x \to a$ we have that for $x > a$ but sufficiently close to a that

$$\frac{f(x)}{g(x)} - L \le \epsilon.$$

Since this holds for all ϵ we see that

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

This proves the claim.   □

Finally, we have the following key fact that show that any differentiable function can be approximated by a polynomial and give a way of estimate the difference between the function and the approximating polynomial.

Theorem: (Taylor expansion.) Let $f : [a, b] \to \mathbf{R}$ be a function and k a positive integer. Assume that $f, f', f^{(2)}, \cdots, f^{(k-1)}$ exists on $[a, b]$ and are continuous and that $f^{(k)}$ is defined on (a, b), then there exists c between a and b such that

$$f(b) = f(a) + f'(a)(b - a) + \frac{f^{(2)}(a)}{2}(b - a)^2 + \cdots + \frac{f^{(k-1)}(a)}{(k - 1)!}(b - a)^{k-1}$$

$$+ \frac{f^{(k)}(c)}{(k)!}(b - a)^k.$$

Proof. Define the Taylor polynomial by

$$P(x) = f(a) + f'(a)(x - a) + \frac{f^{(2)}(a)}{2}(x - a)^2 + \cdots + \frac{f^{(k-1)}(a)}{(k - 1)!}(x - a)^{k-1}$$

and define a number M by that

$$f(b) = P(b) + \frac{M}{k!}(b - a)^k.$$

We want to show that there exists some c between a and b such that

$$M = f^{(k)}(c).$$

To do that we set

$$R(x) = f(x) - P(x) - \frac{M}{k!}(x - a)^k.$$

We have that $R(a) = R(b) = 0$ and so by Rolle's theorem, there exists some c_1 between a and b with $R'(c_1) = 0$. Next observe that $R'(a) = R'(c_1) = 0$ and so again by Rolle's theorem, there exists c_2 between a and c_1 with $R^{(2)}(c_2) = 0$. Since $R^{(i)}(a) = 0$ for $i = 0, \cdots, k - 1$ we can continue this process k times and find some $c = c_k$ such that $R^{(k)}(c) = 0$. However, $0 = R^{(k)}(c) = f^k(c) - M$ Therefore, $M = f^{(k)}(c)$ as claimed.   □

<div align="center">

REFERENCES

</div>

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 17

Recall that last time we showed the Taylor expansion theorem:

Theorem: (Taylor expansion.) Let $f : [a, b] \to \mathbf{R}$ be a function and k a positive integer. Assume that f, f', $f^{(2)}, \cdots, f^{(k-1)}$ exists on $[a, b]$ and are continuous and that $f^{(k)}$ is defined on (a, b), then there exists c between a and b such that

$$f(b) = f(a) + f'(a)(b - a) + \frac{f^{(2)}(a)}{2}(b - a)^2 + \cdots + \frac{f^{(k-1)}(a)}{(k-1)!}(b - a)^{k-1}$$

$$+\frac{f^{(k)}(c)}{(k)!}(b - a)^k.$$

For and infinitely differentiable function f on $\mathbf{R}$ we define the $(k-1)$ Taylor polynomial at a by

$$P_{k-1}(x) = f(a) + f'(a)(x - a) + \frac{f^{(2)}(a)}{2}(x - a)^2 + \cdots + \frac{f^{(k-1)}(a)}{(k-1)!}(x - a)^{k-1}.$$

Question: One naturally wonders how well does this polynomial approximate f when x is near a?

Answer: This depend on the value of the remainder

$$R_k(x) = \frac{f^{(k)}(c)}{k!}(x - a)^k.$$

Example 1: Suppose that $f(x) = \mathrm{e}^x$ so $f^{(k)}(x) = f(x)$ for all k. This means that the Taylor expansion near $a = 0$ becomes

$$P_{k-1}(x) = \sum_{i=0}^{k-1} \frac{x^i}{i!}\,.$$

By the Taylor expansion theorem we have that

$$f(x) = P_{k-1}(x) + \frac{f^{(k)}(c)}{k!}\, x^k\,.$$

Since $f^{(k)}(x) = f(x)$ for all k, it follows from the Taylor expansion theorem that we have

$$|f(x) - P_{k-1}(x)| \le \frac{\mathrm{e}^{|x|}}{k!}\,.$$

We conclude that for k large the polynomial P_{k-1} gives a pretty good approximation to f. For instance, if $|x| \le 1$, then we have that

$$|f(x) - P_{k-1}(x)| \le \frac{\mathrm{e}}{k!}\,.$$

Example 2: On $\mathbf{R}$ define a function f by

$$f(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \mathrm{e}^{-\frac{1}{x^2}} & \text{otherwise} \end{cases}$$

It is easy to see that f is infinitely differentiable and that $f^{(k)}(0) = 0$ for all k. It follows that for all k the Taylor polynomial at 0 is $P_{k-1} \equiv 0$. Thus in this case $f(x) = R_k(x)$.

Riemann integrals

Partition: Let $[a, b]$ be an interval. A partition $\mathcal{P}$ of the interval $[a, b]$ is a number of sub-divisions x_i such that

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b\,.$$

The partition is then the sub-intervals $[x_{i-1}, x_i]$. We will set $\Delta x_i = x_i - x_{i-1}$.

Upper and lower sums: Suppose now that $f : [a, b] \to \mathbf{R}$ is a bounded function and that $\mathcal{P} = \{x_i\}$ is a partition of the interval $[a, b]$. We define upper and lower sums as follows. Set

$$M_i = \sup_{[x_{i-1}, x_i]} f\,,$$

$$m_i = \inf_{[x_{i-1}, x_i]} f\,,$$

and upper $U(f, \mathcal{P})$ and lower sums $L(f, \mathcal{P})$ by

$$U(f, \mathcal{P}) = \sum_{i=1}^{n} M_i \, \Delta x_i \,,$$

$$L(f, \mathcal{P}) = \sum_{i=1}^{n} m_i \, \Delta x_i \,.$$

Example 3: Suppose that the function is $f(x) = x^2 + 1$ on the interval $[-2, 2]$ and that the partition is $\mathcal{P}$ is $\{-2, -1, 0, 1, 2\}$. We have

$$m_1 = 2 \text{ and } M_1 = 5 \,,$$
$$m_2 = 1 \text{ and } M_2 = 2 \,,$$
$$m_3 = 1 \text{ and } M_3 = 2 \,,$$
$$m_4 = 2 \text{ and } M_4 = 5 \,.$$

For the lower and upper sums we have

$$L(f, \mathcal{P}) = 2 + 1 + 1 + 2 = 6 \,,$$
$$U(f, \mathcal{P}) = 5 + 2 + 2 + 5 = 14 \,.$$

The following lemma is immediate (from that $M_i \geq m_i$):

Lemma 1: We always have that

$$U(f, \mathcal{P}) \geq L(f, \mathcal{P}) \,.$$

Sub-partition: Let $[a, b]$ be an interval and $\mathcal{P}_1$ and $\mathcal{P}_2$ two partitions of the interval $[a, b]$. We say that $\mathcal{P}_2$ is a sub-partition (or refinement) of $\mathcal{P}_1$ if all the dividing points in $\mathcal{P}_1$ are also in $\mathcal{P}_1$ (and then presumable some additional dividing points).

Example 4: Suppose that the interval is $[-2, 2]$ and the given partition $\mathcal{P}_1$ is

$$\{-2, -1, 0, 1, 2\} \,.$$

Then the partition

$$\mathcal{P}_2 = \left\{ -2, -1\frac{1}{2}, -1, 0, \frac{1}{2}, 1, 2 \right\}$$

is a refinement (or sub-division) of $\mathcal{P}_1$. Indeed, $\mathcal{P}_2$ has the same dividing points as $\mathcal{P}_1$ in addition to some more.

We now have the following:

Lemma 2: Suppose now that $f : [a, b] \to \mathbf{R}$ is a bounded function and that $\mathcal{P}_1$ is a partition of the interval $[a, b]$ and $\mathcal{P}_2$ is a refinement of $\mathcal{P}_1$, then

$$L(f, \mathcal{P}_1) \leq L(f, \mathcal{P}_2) \leq U(f, \mathcal{P}_2) \leq U(f, \mathcal{P}_1) \,.$$

Proof. The middle inequality is the previous lemma. The inequality to the right follows from that if $\mathcal{P}_2$ is a subdivision of $\mathcal{P}_1$. Namely, suppose that $a = x_0 < x_1 < \cdots < x_n = b$ are the dividing points for $\mathcal{P}_1$ and that between say x_{i-1} and x_i there is an extra dividing point in $\mathcal{P}_2$ say y so $x_{i-1} < y < x_i$, then we have

$$\sup_{[x_{i-1}, y]} f \leq M_i$$

and

$$\sup_{[y, x_i]} f \leq M_i$$

so

$$\left[\sup_{[x_{i-1}, y]} f \right] (y - x_{i-1}) + \left[\sup_{[y, x_i]} f \right] (x_i - y) \leq M_i \, \Delta \, x_i \,.$$

From this it follows easily that

$$U(f, \mathcal{P}_2) \leq U(f, \mathcal{P}_1) \,.$$

Similarly, for the inequality to the left. $\square$

Upper and lower integrals: Suppose now that $f : [a, b] \to \mathbf{R}$ is a bounded function. Define the upper Riemann integral of f by

$$\overline{\int_a^b} f \, dx = \inf_{\mathcal{P}} U(f, \mathcal{P}) \,.$$

Here the infimum is taken over all partitions of $[a, b]$. Likewise, we define the lower Riemann integral by

$$\underline{\int_a^b} f \, dx = \sup_{\mathcal{P}} L(f, \mathcal{P}) \,.$$

Riemann integral: Suppose that $f : [a, b] \to \mathbf{R}$ is a bounded function, then we say that f is Riemann integrable if

$$\overline{\int_a^b} f \, dx = \underline{\int_a^b} f \, dx \,.$$

If the function is Riemann integrable, then the Riemann integral is

$$\int_a^b f\,dx = \overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx\,.$$

The Riemann integrable functions is denoted by $\mathcal{R}\left([a,b]\right)$.

From Wikipedia: Georg Friedrich Bernhard Riemann $(1826 - 1866)$ was a German mathematician who made profound contributions to analysis, number theory, and differential geometry. Riemann held his first lectures in 1854, which founded the field of Riemannian geometry and thereby set the stage for Albert Einstein's general theory of relativity. In the field of real analysis, he is mostly known for the first rigorous formulation of the integral, the Riemann integral, and his work on Fourier series. His contributions to complex analysis include most notably the introduction of Riemann surfaces, breaking new ground in a natural, geometric treatment of complex analysis. His 1859 paper on the prime-counting function, containing the original statement of the Riemann hypothesis, is regarded as a foundational paper of analytic number theory. He is considered by many to be one of the greatest mathematicians of all time.

Example 5: Let $f : [0,1] \to \mathbf{R}$ be given by

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1] \cap \mathbf{Q} \\ 1 & \text{otherwise} \end{cases}$$

For this function and all partitions $\mathcal{P}$ we have that

$$L(f,\mathcal{P}) = 0 \text{ and } U(f,\mathcal{P}) = 1\,.$$

Thus, f is not Riemann integrable.

We will be interested in the questions: "What kind of functions are Riemann integrable?"

.... and "How do we compute the integral?"

The answer to the second question will be the fundamental theorem of calculus. This will be the topic of a later lecture.

Lemma 3: Suppose now that $f : [a, b] \to \mathbf{R}$ is a bounded function, then $f \in \mathcal{R}\left([a, b]\right)$ if and only if for all $\epsilon > 0$, there exists a partition $\mathcal{P}$ such that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.$$

Proof. Suppose that $f \in \mathcal{R}\left([a, b]\right)$, then

$$\sup_{\mathcal{P}} L(f, \mathcal{P}) = \underline{\int_a^b} f \, dx = \overline{\int_a^b} f \, dx = \inf_{\mathcal{P}} U(f, \mathcal{P}).$$

This means that given $\epsilon > 0$, there exists partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ such that

$$\int_a^b f \, dx - \frac{\epsilon}{2} < L(f, \mathcal{P}_1)$$

and

$$U(f, \mathcal{P}_2) \le \int_a^b f \, dx + \frac{\epsilon}{2}.$$

Let $\mathcal{P}$ be the partition that has all the dividing points of both $\mathcal{P}_1$ and $\mathcal{P}_2$. So $\mathcal{P}$ is a refinement of both $\mathcal{P}_1$ and $\mathcal{P}_2$. It follows that

$$\int_a^b f \, dx - \frac{\epsilon}{2} < L(f, \mathcal{P}_1) \le L(f, \mathcal{P}) \le U(f, \mathcal{P}) \le U(f, \mathcal{P}_2) \le \int_a^b f \, dx + \frac{\epsilon}{2}.$$

This proves the claim.

To see the converse, suppose that for some $\epsilon > 0$, there exists a partition $\mathcal{P}$ such that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.$$

Since

$$L(f, \mathcal{P}) \le \underline{\int_a^b} f \, dx$$

and

$$\overline{\int_a^b} f \, dx \le U(f, \mathcal{P})$$

we have that

$$\overline{\int_a^b} f \, dx - \underline{\int_a^b} f \, dx \le U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.$$

Since this holds for all $\epsilon > 0$ we get the claim. $\square$

We now get to a key theorem that gives a simple criterium for a function to be Riemann integrable:

Theorem: Any continuous function on $[a, b]$ is in $\mathcal{R}\left([a, b]\right)$.

Proof. We will show this next time once we have shown that a continuous function on a closed and bounded interval is, in fact, uniformly continuous. $\square$

The proof of this theorem needs the following key concept.

Definition: Uniformly continuous. Suppose that $f : I \to \mathbf{R}$ is a function, where I is an interval. We say that f is uniformly continuous if for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon \text{ if } |x - y| < \delta \,.$$

Note that being uniformly continuous is stronger than being continuous. It means that for a given $\epsilon > 0$, the same δ can be used for all x.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
    TBB can be downloaded at:
    https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
    (screen-optimized)
    https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
    (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 18

Definition: Uniformly continuous. Suppose that $f : I \to \mathbf{R}$ is a function, where I is an interval. We say that f is uniformly continuous if for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon \text{ if } |x - y| < \delta.$$

Note that being uniformly continuous is stronger than being continuous. It means that for a given $\epsilon > 0$, the same δ can be used for all x.

Example 1: Suppose that

$$f(x) = x^2$$

on $\mathbf{R}$, then f is NOT uniformly continuous. To see this, let $\epsilon > 0$ be given if f was uniformly continuous, then there would exists $\delta > 0$ such that

$$f(x + \delta) - f(x) < \epsilon,$$

for all x. This would mean that

$$2\,\delta\,x < (x + \delta)^2 - x^2 < \epsilon$$

for all x, which is clearly not the case.

Example 2: Suppose that

$$f(x) = \frac{1}{x}$$

on $(0, 1]$, then f is NOT uniformly continuous. To see this, consider $x_n = \frac{1}{n}$ and $y_n = \frac{1}{2n}$, then

$$|f(x_n) - f(y_n)| = n$$

and

$$|x_n - y_n| < \frac{1}{n}.$$

From this it easily follows that f is not uniformly continuous.

Theorem 1: Any continuous function on $[a, b]$ is uniformly continuous.

Proof. Suppose not; then there exists $\epsilon > 0$ such that for all $n > 0$, there are x_n and y_n with

$$|x_n - y_n| < \frac{1}{n}$$

and so that

$$|f(x_n) - f(y_n)| \geq \epsilon.$$

Since the interval $[a, b]$ is compact we can choose a subsequence of x_n say x_{n_k} so that

$$x_{n_k} \to x.$$

Since

$$|x - y_{n_k}| \leq |x - x_{n_k}| + |x_{n_k} - y_{n_k}|$$

we have that $y_{n_k} \to x$ as well. Since f is continuous we have that $f(x_{n_k}) \to f(x)$ and $f(y_{n_k}) \to f(x)$. However, this contradict that

$$|f(x_{n_k} - f(y_{n_k})| \geq \epsilon,$$

. □

We now get to a key theorem that gives a simple criterium for a function to be Riemann integrable:

Theorem 2: Any continuous function on $[a, b]$ is in $\mathcal{R}([a, b])$.

Proof. Given $\epsilon > 0$, since f is uniformly continuous by Theorem 1 it follows that there exists $\delta > 0$ such that if $|x - y| < \delta$, then

$$|f(x) - f(y)| < \frac{\epsilon}{b - a}.$$

Let $\mathcal{P}$ be a partition so that for all i we have $\Delta x_i < \delta$, then on each interval of the partition of the form $[x_{i-1}, x_i]$ we have that

$$M_i - m_i < \frac{\epsilon}{b - a}.$$

It follows that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_i M_i \, \Delta x_i - \sum_i m_i \, \Delta x_i$$

$$= \sum_i [M_i - m_i] \, \Delta x_i < \frac{\epsilon}{b - a} \sum_i \Delta x_i = \epsilon.$$

Since this holds for all $\epsilon > 0$ we have that f is integrable. □

Basic properties of integrals.

Theorem 3: We have the following basic formulas for integrals:

   (1) If $f \in \mathcal{R}\left([a,b]\right)$ and $c \in \mathbf{R}$, then $c\,f \in \mathcal{R}\left([a,b]\right)$ and

$$\int_a^b (c\,f)\,dx = c \int_a^b f\,dx\,.$$

   (2) If $f,\,g \in \mathcal{R}\left([a,b]\right)$, then $f+g \in \mathcal{R}\left([a,b]\right)$ and

$$\int_a^b (f+g)\,dx = \int_a^b f\,dx + \int_a^b g\,dx\,.$$

   (3) If $f,\,g \in \mathcal{R}\left([a,b]\right)$ and $f \le g$, then

$$\int_a^b f\,dx \le \int_a^b g\,dx\,.$$

   (4) If $f \in \mathcal{R}\left([a,b]\right)$ and $c \in (a,b)$, then $f \in \mathcal{R}\left([a,c]\right)$ and $f \in \mathcal{R}\left([c,b]\right)$ and

$$\int_a^c f\,dx + \int_c^b f\,dx = \int_a^b f\,dx\,.$$

Proof. The first claim follow from that if $\mathcal{P}$ is a partition, then

$$L(c\,f,\mathcal{P}) = c\,L(f,\mathcal{P})$$

and

$$U(c\,f,\mathcal{P}) = c\,U(f,\mathcal{P})\,.$$

   To prove the second claim. Given $\epsilon > 0$, let $\mathcal{P}_1$ and $\mathcal{P}_2$ be partitions so that

$$U(f,\mathcal{P}_1) - L(f,\mathcal{P}_1) < \frac{\epsilon}{2}$$

and

$$U(g,\mathcal{P}_2) - L(g,\mathcal{P}_2) < \frac{\epsilon}{2}\,.$$

Let $\mathcal{P}$ be the partition that has the combined dividing points of $\mathcal{P}_1$ and $\mathcal{P}_2$. It follows that

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\epsilon}{2}$$

and

$$U(g,\mathcal{P}) - L(g,\mathcal{P}) < \frac{\epsilon}{2}\,.$$

Therefore,

$$U(f+g,\mathcal{P}) - L(f+g,\mathcal{P}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.$$

From this the second claim follows.

   To see the third claim let $\mathcal{P}$ be any partition of $[a,b]$. It follows that

$$U(g,\mathcal{P}) \le U(f,\mathcal{P})\,.$$

Since

$$\int_a^b g\,dx = \inf_{\mathcal{P}} U(g,\mathcal{P})$$

and likewise for f the claim now follows.

Finally, to see the fourth claim. Let $\mathcal{P}$ be any partition of $[a, b]$ and let $\mathcal{P}_0$ be the refinement of $\mathcal{P}$ that in addition to the dividing points of $\mathcal{P}$ also have c as a dividing point. It follows that
$$U(f, \mathcal{P}_0 \cap [a, c]) + U(f, \mathcal{P}_0 \cap [c, b]) = U(f, \mathcal{P}_0) \,.$$
Likewise,
$$L(f, \mathcal{P}_0 \cap [a, c]) + L(f, \mathcal{P}_0 \cap [c, b]) = L(f, \mathcal{P}_0) \,.$$
Therefore,
$$U(f, \mathcal{P}_0 \cap [a, c]) - L(f, \mathcal{P}_0 \cap [a, c]) + U(f, \mathcal{P}_0 \cap [c, b]) - L(f, \mathcal{P}_0 \cap [c, b])$$
$$= U(f, \mathcal{P}_0) - L(f, \mathcal{P}_0) \,.$$
From this the fourth claim easily follows.   □

Corollary: Suppose that $f, |f| \in \mathcal{R}([a, b])$, then
$$\int_x^b f \, dx \ \leq \ \int_a^b |f| \, dx \,.$$

Proof. This follows from the lemma since $f \leq |f|$ and $-f \leq |f|$. Namely, from the first of these inequalities together with the lemma we get that
$$\int_a^b f \, dx \leq \int_a^b |f| \, dx \,,$$
whereas from the second we get that
$$-\int_a^b f \, dx = \int_a^b (-f) \, dx \leq \int_a^b |f| \, dx \,.$$
Together these gives the claim.   □

Fundamental theorem of calculus, version 1: Let f be a continuous function on $[a, b]$ and define F on $[a, b]$ by
$$F(x) = \int_a^x f(s) \, ds \,.$$
The function F is differentiable with derivative f.

Proof. Fix $x_0 \in [a, b]$ and assume first that $x > x_0$. We then have that
$$F(x) = \int_a^x f(s) \, ds = \int_a^{x_0} f(s) \, ds + \int_{x_0}^x f(s) \, ds = F(x_0) + \int_{x_0}^x f(s) \, ds \,.$$
It follows that
$$F(x) - F(x_0) = \int_{x_0}^x f(s) \, ds \,.$$
Therefore,
$$(x - x_0) \min_{[x_0, x]} f \leq F(x) - F(x_0) \leq (x - x_0) \max_{[x_0, x]} f$$

and hence

$$\min_{[x_0,x]} f \le \frac{F(x) - F(x_0)}{x - x_0} \le \max_{[x_0,x]} f \,.$$

Since f is continuous at x_0 as $x \to x_0$ both the left and right hand side of this string of inequalities converges to $f(x_0)$. This proves the claim when $x > x_0$. When $x < x_0$ we can write $F(x)$ as

$$F(x) + \int_x^{x_0} f(s)\,ds = F(x_0)\,.$$

Therefore,

$$F(x) - F(x_0) = -\int_x^{x_0} f(s)\,ds\,.$$

Arguing as above gives the claim also in this case.   □

Fundamental theorem of calculus, version 2: Suppose that $F : [a,b] \to \mathbf{R}$ is differentiable and that $F' = f \in \mathcal{R}\left([a,b]\right)$, then

$$F(b) - F(a) = \int_a^b f(s)\,ds\,.$$

Proof. Since f is integrable, then for all $\epsilon > 0$, there exists a partition $\mathcal{P}$ of $[a,b]$ such that

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon\,.$$

For a given partition $\mathcal{P}$ with dividing points x_i we have

$$L(f,\mathcal{P}) = \sum_i m_i \left(x_i - x_{i-1}\right),$$

$$U(f,\mathcal{P}) = \sum_i M_i \left(x_i - x_{i-1}\right),$$

Moreover, by the mean value inequality

$$F(x_i) - F(x_{i-1}) = f(y_i)\left(x_i - x_{i-1}\right).$$

We now have that

$$m_i \left(x_i - x_{i-1}\right) \le F(x_i) - F(x_{i-1}) \le M_i \left(x_i - x_{i-1}\right).$$

It follows that

$$L(f,\mathcal{P}) \le \sum_{i=1}^n [F(x_i) - F(x_{i-1})] \le U(f,\mathcal{P})\,.$$

Finally, the claim follows from the above by observing that

$$F(b) - F(a) = \sum_{i=1}^n [F(x_i) - F(x_{i-1})]\,.$$

□

Example 3: To compute

$$\int_0^1 x^2 \, dx \, ,$$

we use the second version of the fundamental theorem of calculus. Namely, observe that the derivative of the function

$$F(x) = \frac{x^3}{3}$$

is x^2 and therefore, by the second version of the fundamental theorem of calculus we have that

$$\int_0^1 x^2 \, dx = F(1) - F(0) = \frac{1}{3} - 0 = \frac{1}{3} \, .$$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 19

Question: What kind of functions are integrable?

Theorem: Any continuous function on $[a, b]$ is in $\mathcal{R}([a, b])$.

Basic properties of integrals.

Theorem: We have the following basic formulas for integrals:

(1) If $f \in \mathcal{R}([a, b])$ and $c \in \mathbf{R}$, then $c\,f \in \mathcal{R}([a, b])$ and

$$\int_a^b (c\,f)\,dx = c \int_a^b f\,dx \,.$$

(2) If $f,\,g \in \mathcal{R}([a, b])$, then $f + g \in \mathcal{R}([a, b])$ and

$$\int_a^b (f + g)\,dx = \int_a^b f\,dx + \int_a^b g\,dx \,.$$

(3) If $f,\,g \in \mathcal{R}([a, b])$ and $f \leq g$, then

$$\int_a^b f\,dx \leq \int_a^b g\,dx \,.$$

(4) If $f \in \mathcal{R}([a, b])$ and $c \in (a, b)$, then $f \in \mathcal{R}([a, c])$ and $f \in \mathcal{R}([c, b])$ and

$$\int_a^c f\,dx + \int_c^b f\,dx = \int_a^b f\,dx \,.$$

Corollary: Suppose that $f,\,|f| \in \mathcal{R}([a, b])$, then

$$\int_x^b f\,dx \;\leq\; \int_a^b |f|\,dx \,.$$

101

Fundamental theorem of calculus, version 1: Let f be a continuous function on $[a, b]$ and define F on $[a, b]$ by

$$F(x) = \int_a^x f(s)\,ds\,.$$

The function F is differentiable with derivative f.

Fundamental theorem of calculus, version 2: Suppose that $F : [a, b] \to \mathbf{R}$ is differentiable and that $F' = f \in \mathcal{R}\,([a, b])$, then

$$F(b) - F(a) = \int_a^b f(s)\,ds\,.$$

Application of integrals: arclength.

Suppose that f and $g : [a, b] \to \mathbf{R}$ are differentiable functions and their derivatives are continuous, then we define the arclength of the curve

$$s \to (f(s), g(s))$$

by

$$L = \int_a^b \sqrt{(f'(s))^2 + (g'(s))^2}\,ds\,.$$

Example 1: Suppose that $f(s) = s$ and $g(s) = s^2$, then $f' = 1$ and $s' = 2\,s$. Therefore, the arclength of the curve (s, s^2), where $s \in [0, 1]$ is

$$L = \int_0^1 \sqrt{1 + (2\,s)^2}\,ds = \int_0^1 \sqrt{1 + 4\,s^2}\,ds\,.$$

Improper integrals.

Unbounded interval.

Suppose that $f \in \mathcal{R}([a, b])$ for all $b > a$. If

$$\lim_{b \to \infty} \int_a^b f(x)\,dx$$

exists, then we say that the improper integral

$$\int_a^\infty f(x)\,dx$$

exists and that

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty}\int_a^b f(x)\,dx$$

Example 2: On $[1,\infty)$, set

$$f(x) = \frac{1}{x^2}\,,$$

then

$$\int_1^c \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_c^1 = -\frac{1}{c}+1\,.$$

Since $-\frac{1}{c}+1 \to 1$ as $c\to\infty$, the improper integral

$$\int_1^\infty \frac{1}{x^2}\,dx$$

exist and is equation to 1.

Example 3: On $[1,\infty)$, set

$$f(x) = \frac{1}{x}\,,$$

then

$$\int_1^c \frac{1}{x}\,dx = [\log x]_1^c = \log c\,.$$

The improper integral

$$\int_1^\infty \frac{1}{x}\,dx$$

does not exist.

Unbounded function.

Suppose that $f \in \mathcal{R}([c,b])$ for all $c > a$. If

$$\lim_{c\to a}\int_c^b f(x)\,dx$$

exists, then we say that the improper integral

$$\int_a^b f(x)\,dx$$

exists and that

$$\int_a^b f(x)\, dx = \lim_{c \to a} \int_c^b f(x)\, dx$$

Example 4: On $(0, 1]$, set

$$f(x) = \frac{1}{\sqrt{x}},$$

then

$$\int_c^1 \frac{1}{\sqrt{x}}\, dx = \left[2\sqrt{x}\right]_c^1 = 2 - 2\sqrt{c}.$$

Since $2 - 2\sqrt{c} \to 2$ as $c \to 0$, the improper integral exists and is equal to

$$\int_0^1 \frac{1}{\sqrt{x}}\, dx = 2.$$

Example 5: On $(0, 1]$, set

$$f(x) = \frac{1}{x},$$

then

$$\int_c^1 \frac{1}{x}\, dx = \left[\log x\right]_c^1 = -\log c.$$

Note that $-\log c \to \infty$ as $c \to 0$ so the improper integral does not exist.

Question: How do we define angle?

Answer: We define it through arclength.

On the unit circle

$$\{(x, y) \mid x^2 + y^2 = 1\}$$

we define angle and the arclength. That is, suppose that (x, y) lies on the unit circle. The angle θ between $(1, 0)$ and (x, y) is the arclength of the part of the unit circle from $(1, 0)$ to (x, y). This part of the circle is parametrized by $(f(s), g(s)) = (s, \sqrt{1 - s^2})$ and where $x \leq s \leq 1$. Since $f'(s) = 1$ and $g'(s) = -\frac{s}{\sqrt{1 - s^2}}$ we get that

$$\theta = \int_x^1 \sqrt{1 + \frac{s^2}{1 - s^2}}\, ds = \int_x^1 \frac{1}{\sqrt{1 - s^2}}\, ds.$$

The function $\arcsin x$ is defined by

$$\arcsin x = \int_0^x \frac{1}{\sqrt{1-s^2}}\, ds\,.$$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 20

Application of integrals: arclength.

Suppose that f and $g : [a, b] \to \mathbf{R}$ are differentiable functions and their derivatives are continuous, then we define the arclength of the curve

$$s \to (f(s), g(s))$$

by

$$L = \int_a^b \sqrt{(f'(s))^2 + (g'(s))^2} \, ds \, .$$

Example 1: Suppose that $f(s) = s$ and $g(s) = s^2$, then $f' = 1$ and $s' = 2\,s$. Therefore, the arclength of the curve (s, s^2), where $s \in [0, 1]$ is

$$L = \int_0^1 \sqrt{1 + (2\,s)^2} \, ds = \int_0^1 \sqrt{1 + 4\,s^2} \, ds \, .$$

Question: How do we define angle?

Answer: We define it through arclength.

On the unit circle

$$\{(x, y) \mid x^2 + y^2 = 1\}$$

we define angle and the arclength. That is, suppose that (x, y) lies on the unit circle. The angle θ between $(1, 0)$ and (x, y) is the arclength of the part of the unit circle from $(1, 0)$ to (x, y). This part of the circle is parametrized by $(f(s), g(s)) = (s, \sqrt{1 - s^2})$ and where $x \le s \le 1$. Since $f'(s) = 1$ and $g'(s) = -\frac{s}{\sqrt{1-s^2}}$ we get that

$$\theta = \int_x^1 \sqrt{1 + \frac{s^2}{1 - s^2}} \, ds = \int_x^1 \frac{1}{\sqrt{1 - s^2}} \, ds \, .$$

The function $\arccos x$ is defined by

$$\arccos x = \int_x^1 \frac{1}{\sqrt{1-s^2}} \, ds \, .$$

By the fundamental theorem of calculus we see that

$$\arccos x = -\frac{1}{\sqrt{1-x^2}} \, .$$

Pointwise convergence: Suppose that f_n is a sequence of functions on an interval I, then we say that f_n convergences pointwise to a function f if for all x we have

$$f_n(x) \to f(x) \, .$$

Example 1: Suppose that $f_n(x) = x^n$ on $[0,1]$, then f_n converges pointwise to f where

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

Suppose first that $0 \le x < 1$, then $f_n(x) = x^n \to 0$. If $x = 1$, then $f_n(x) = 1$ for all n and so $f_n(x) \to 1$. This show the claim.

Example 2: If $E_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$, then $E_n(x) \to \exp x$ pointwise. We have already proven that the radius of convergence for the power series

$$\sum_{k=0}^\infty \frac{x^k}{k!}$$

is infinity. From this the claim follows.

Uniform convergence: Suppose that f_n is a sequence of functions on an interval I, then we say that f_n convergences uniformly to a function f if for all $\epsilon > 0$, there exists an N such that if $n \ge N$, then for all x

$$|f(x) - f_n(x)| < \epsilon \, .$$

Lemma 1: Suppose that I is an interval and f_n is a sequence of functions on I that converges uniformly to a function f, then f_n also converges pointwise to f.

Proof. This is immediate from the definition of uniform convergence. □

Example 1A: Suppose again that $f_n(x) = x^n$ on $[0,1]$, then f_n converges pointwise but NOT uniformly to f where

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

To see this observe that for each n,since f_n is continuous by the intermediate value theorem there exists x_n with $0 < x_n < 1$ such that $f_n(x) = \frac{1}{2}$. It now follows that

$$\frac{1}{2} = |f(x_n) - f_n(x_n)| \leq \sup_{x \in [0,1]} |f(x) - f_n| \,.$$

Thus we see that the convergence is not uniform. We already saw in Example 1 that the convergence is pointwise.

Example 2A: If $E_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$, then $E_n(x) \to \exp x$ uniformly on any interval of the form $[-L, L]$. This will be a consequence of of Weirstrass M-test that we will discuss next.

Lemma 2 [Weirstrass M-test]: Suppose that I is an interval and f_n is a sequence of functions on I. Suppose also that M_n is a sequence of non-negative numbers with

$$|f_n(x)| \leq M_n \text{ for all } x \in I \,.$$

If the series

$$\sum_{n=1}^{\infty} M_n$$

converges, then the sequence of functions

$$S_n(x) = \sum_{k=0}^n f_k(x)$$

converges uniformly.

Proof. For each fixed x we have that that the sequence

$$\sum_{k=0}^{\infty} f_k(x)$$

converges. Moreover, we have that for all x and $m < n$ we have

$$|S_n(x) - S_m(x)| \leq |f_n(x)| + |f_{n-1}(x)| + \cdots + |f_{m+1}(x)| \leq M_n + \cdots + M_{m+1} \,.$$

For m fixed and since $S_n(x) \to S(x)$ it follows that

$$|S(x) - S_m(x)| \le \sum_{k=m+1}^{\infty} M_k \,.$$

Since $\sum_{k=0} M_k$ is convergent it implies that given $\epsilon > 0$, there exists N such that if $m \ge N$, then $\sum_{k=m+1}^{\infty} M_k < \epsilon$. Therefore, for $m \ge N$ and all x

$$|S(x) - S_m(x)| < \epsilon \,.$$

This proves the claim. $\qquad\square$

Example 2A: On the interval $I = [-L, L]$ suppose

$$f_n = \frac{x^n}{n!} \,.$$

Then

$$|f_n| \le \frac{L^n}{n!} \,.$$

Since

$$\sum_n \frac{L^n}{n!}$$

is convergent Weirstrass M-test gives that the series

$$\sum_{n=0}^{\infty} f_n$$

converges uniformly on I.

Theorem: If

$$\sum_{k=0}^{\infty} a_k \, x^k$$

is a power series and R is its radius of convergence. Then it converges uniformly on any (finite) interval of the form $[-L, L]$ where $L < R$.

Proof. Recall that if $M = \limsup_{n \to \infty} |a_n|^{\frac{1}{n}}$, then the radius of convergence is $R = \frac{1}{M}$. It follows that if $|x| \le L < R$, then

$$\limsup |a_n \, x^n|^{\frac{1}{n}} = |x| \limsup |a_n|^{\frac{1}{n}} \le L\,M < 1 \,.$$

Choose $1 > \alpha > L\,M$. For n sufficiently large $|a_n \, x^n| \le M_n = \alpha^n$. Since the geometric series $\sum_n \alpha^n$ is convergent, Weirstrass M-test gives the claim. $\qquad\square$

Example 3: The geometric power series

$$\sum_{k=0}^{\infty} x^k$$

converges uniformly to $\frac{1}{1-x}$ on all intervals of the form $[-L, L]$ where $L < 1$. Since the radius of convergence of the power series is one the claim therefore follows from the theorem above.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 21

Theorem 1: Suppose that I is an interval and f_n is a sequence of continuous functions on I. If f_n converges uniformly to f, then f is also continuous.

Proof. Let x_0 in I be arbitrary but fixed. We will show that f is continuous at x_0. Given $\epsilon > 0$, since $f_n \to f$ uniformly, there exists a N such that if $n \geq N$, then for all x in I

$$|f(x) - f_n(x)| < \frac{\epsilon}{3} \, .$$

Since f_N is continuous at x_0, there exists $\delta > 0$ such that if $|x - x_0| < \delta$, then

$$|f_N(x) - f_N(x_0)| < \frac{\epsilon}{3} \, .$$

Combining this gives that for $|x - x_0| < \delta$

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \, .$$

This gives the claim. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Example 1: Set

$$E_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \, ,$$

$$E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \, .$$

In the previous lecture we showed that Weirstrass M-test implies that $E_n \to E$ uniformly on $[-L, L]$. Since each E_n is continuous we have from Theorem 1 that E is continuous.

Here is another useful way of thinking of uniform convergence. Recall that on the space of continuous functions $C(I)$ on an interval $I = [a, b]$ there is a natural metric given by that

$$d(f, g) = \max_{x \in I} \left\{ |f(x) - g(x)| \, | \, x \in I \right\} \, .$$

We have the following:

Proposition: Let I be an interval $[a, b]$ and f_n, $f \in C(I)$, then $f_n \to f$ in the metric space if and only if f_n converges to f uniformly.

Proof. To see this note that

$$|f(x) - f_n(x)| \leq \epsilon \text{ for all } x \in I$$

if and only if

$$d(f, f_n) \leq \epsilon \,.$$

To say that $f_n \to f$ uniformly is therefore equivalent to that $d(f, f_n) \to 0$ giving the claim.
$\square$

From this we get:

Corollary: $C([a, b])$ is Cauchy complete.

Proof. Suppose that f_n is a Cauchy sequence in $C([a, b])$ we need to find a $f \in C([a, b])$ such that $f_n \to f$ uniformly. For each x fixed, the sequence $f_n(x)$ is a Cauchy sequence in $\mathbf{R}$. This follows since

$$|f_n(x) - f_m(x)| \leq d(f_n, f_m) \,.$$

Therefore, since $\mathbf{R}$ is Cauchy complete, for each x there exists a $f(x)$ such that $f_n(x) \to f(x)$. This defines the function f and show that $f_n \to f$ converges pointwise. We need to show that the convergence is uniform. To see that observe that given $\epsilon > 0$ since f_n is a Cauchy sequence, there exists N such that if n and $m \geq N$, then

$$|f_n(x) - f_m(x)| < \frac{\epsilon}{2} \text{ for all } x \in I \,.$$

Therefore, for $f(x) = \lim_{m \to \infty} f_m(x)$ we have

$$|f_n(x) - f(x)| \leq \frac{\epsilon}{2} < \epsilon \text{ for all } x \in I \,.$$

This show that the convergence is uniform.
$\square$

Theorem 2: If $f_n \in \mathcal{R}([a, b])$ and $f_n \to f$ uniformly, then $f \in \mathcal{R}([a, b])$ and

$$\int_a^b f_n \, dx \to \int_a^b f \, dx \,.$$

Proof. We need to first show that $f \in \mathcal{R}([a, b])$ and so we need to show that given $\epsilon > 0$, there exists a partition $\mathcal{P}$ of the interval $[a, b]$ such that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon \,.$$

Since $f_n \to f$ uniformly we have that there exists a N such that if $n \geq N$, then

$$|f(x) - f_n(x)| < \frac{\epsilon}{3\,(b-a)}\,.$$

We have therefore that for any partition $\mathcal{P}$ that

$$|m_i^{f_n} \leq m_i^f| \leq \frac{\epsilon}{3\,(b-a)}\,,$$

$$|M_i^{f_n} \leq M_i^f| \leq \frac{\epsilon}{3\,(b-a)}\,,$$

It follows that for any partition when $n \geq N$, then

$$|U(f,\mathcal{P}) - U(f_n,\mathcal{P})| < \frac{\epsilon}{3}\,,$$

$$|L(f,\mathcal{P}) - L(f_n,\mathcal{P})| < \frac{\epsilon}{3}\,.$$

We can now use that since $f_N \in \mathcal{R}\,([a,b])$ we have that there exists a partition $\mathcal{P}$ such that

$$U(f_N,\mathcal{P}) - L(f_N,\mathcal{P}) < \frac{\epsilon}{3}\,.$$

Combining it all gives that

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < U(f,\mathcal{P}) - U(f_N,\mathcal{P}) + U(f_N,\mathcal{P}) - L(f_N,\mathcal{P}) + L(f_N,\mathcal{P}) - L(f,\mathcal{P})$$

$$< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}\,.$$

This show that $f \in \mathcal{R}\,([a,b])$. We also need to see that

$$\int_a^b f \, dx = \lim_{n\to\infty} \int_a^b f_n \, dx\,.$$

This, however, follows from that

$$L(f,\mathcal{P}) \leq \int_a^b f \, dx \leq U(f,\mathcal{P})\,,$$

$$L(f_n,\mathcal{P}) \leq \int_a^b f_n \, dx \leq U(f_n,\mathcal{P})\,.$$

and that for $n \geq N$

$$|U(f,\mathcal{P}) - U(f_n,\mathcal{P})| < \frac{\epsilon}{3}\,,$$

$$|L(f,\mathcal{P}) - L(f_n,\mathcal{P})| < \frac{\epsilon}{3}\,.$$

Namely, we now have that also

$$L(f,\mathcal{P}) - \frac{\epsilon}{3} \leq \int_a^b f_n \, dx \leq U(f,\mathcal{P}) + \frac{\epsilon}{3}$$

and therefore

$$\int_a^b f \, dx - \int_a^b f \, dx \; < \epsilon\,.$$

$\square$

Example 1A: Set

$$E_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!},$$

$$E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Then we have from Example 1 that $E_n \to E$ uniformly on $[-L, L]$. We now have from Theorem 2 that

$$\sum_{k=0}^{n} \int_0^1 \frac{x^k}{k!}\, dx \to \int_0^1 E(x)\, dx.$$

Theorem 3: Suppose that f_n are differentiable functions on $[a, b]$ and $x_0 \in [a, b]$. If

- $f_n(x_0) \to c$,
- $f_n' \to g$ uniformly,
- f_n' are continuous on $[a, b]$,

then there exists a differentiable function f with

- $f_n \to f$ uniformly,
- $f_n' \to f'$ uniformly.

Proof. Define a function F on $[a, b]$ by

$$f(x) = c + \int_{x_0}^{x} g\, dx,$$

and note since f_n' are continuous and that $f_n' \to g$ uniformly, it follows from Theorem 1 that g is also continuous. Therefore, by the fundamental theorem of calculus f is differentiable and $f' = g$. Moreover, by the fundamental theorem of calculus we have that

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'\, dx.$$

We are done provided we can show that $f_n \to f$. To see that note that

$$|f(x) - f_n(x)| = \left| c + \int_{x_0}^{x} g\, dx - f_n(x_0) - \int_{x_0}^{x} f_n'\, dx \right|$$

$$\leq |c - f_n(x_0)| + \left| \int_{x_0}^{x} (g - f_n')\, dx \right| \leq |c - f_n(x_0)| + \int_{x_0}^{x} |g - f_n'|\, dx$$

$$\leq |c - f_n(x_0)| + (b - a)\, d(g, f_n').$$

The claim now follows since $f_n(x_0) \to c$ and $d(g, f_n') \to 0$. $\qquad\square$

Example 1B: Set

$$E_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!},$$

$$E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}\,.$$

Then

$$E_n' = \sum_{k=1}^{n} k\,\frac{x^{k-1}}{k!} = \sum_{k=0}^{n-1} \frac{x^{k-1}}{(k-1)!} = E_{n-1}\,.$$

From Example 1 that $E_{n-1} \to E$ uniformly on $[-L, L]$ and each E_n are continuous. Moreover, for all n we have that

$$E_n(0) = 1 = E(0)\,.$$

It follows therefore from Theorem 3 that

$$E_n' = E_{n-1} \to E'$$

uniformly and since $E_n' = E_{n-1}$, then we have that $E' = E$.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
TBB can be downloaded at:
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
(screen-optimized)
https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
(print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 22

Suppose that a_n is a sequence and

$$\sum_{n=0}^{\infty} a_n \, x^n \,,$$

is a power series, the radius of convergence R is

$$R = \frac{1}{M} \text{ where } M = \limsup |a_n|^{\frac{1}{n}} \,.$$

Lemma: The radius of convergence is the same for the power series

$$\sum_{n=0}^{\infty} a_n \, x^n$$

as the power series

$$\sum_{n=1}^{\infty} n \, a_n \, x^{n-1} \,.$$

.

Proof. Since

$$n^{\frac{1}{n-1}} = e^{\frac{\log n}{n-1}} \to 1 \,,$$

and

$$\limsup_{n \to \infty} |a_n|^{\frac{1}{n}} = \limsup_{n \to \infty} |a_n|^{\frac{1}{n-1}}$$

we have that

$$\limsup_{n \to \infty} |n \, a_n|^{\frac{1}{n-1}} = \limsup_{n \to \infty} |a_n|^{\frac{1}{n}} \,.$$

From this the claim follows. $\qquad\qquad\square$

Iterating this gives:

Corollary: The power series

$$\sum_{n=0}^{\infty} a_n\, x^n$$

has the same radius of convergence as the power series

$$\sum_{n=k}^{\infty} \frac{n!}{(n-k)!}\, a_n\, x^{n-k}\ .$$

We now get the following:

Theorem: Suppose that

$$f(x) = \sum_{n=0}^{\infty} a_n\, x^n\ ,$$

is a power series with radius of convergence R, then

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!}\, a_n\, x^{n-k}$$

and

$$\int f(x)\, dx = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n}\, x^n\ .$$

Proof. Let us first argue for $= 1$. We will see that this is a consequence of Theorem 3 from Lecture 21. Set

$$f_n(x) = \sum_{k=0}^{n} a_k\, x^k$$

and

$$f(x) = \sum_{k=0}^{\infty} a_k\, x^k\ .$$

Moreover, let R be the radius of convergence for the power series f. We have the following three properties

(1)
$$f_n(0) = a_0 = f(0)\ .$$

(2) On each interval $[-L, L]$, where $L < R$, we have uniform convergence

$$f_n' \to \sum_{k=1}^{\infty} k\, a_k\, x^{k-1}\ .$$

(3) Each f_n' is continuous.

We see that Theorem 3 applies and show that

$$f' = \sum_{k=1}^{\infty} k\, a_k\, x^{k-1}\,.$$

Iterating this gives the claim for all k. Finally, the claim about the integral

$$\int f(x)\, dx$$

follows easily from Theorem 2 from Lecture 21. □

Ordinary differential equations: A differential equation is an equation that involves an unknown function and its derivative.

Example: Here are some examples of differential equations

$$f'(x) = x\,.$$
$$f'(x) - f^2(x) = 0\,.$$
$$f(x)\, f'(x)\, f''(x) = 1\,.$$

For the first of these and each constant c, the function

$$f_c(x) = \frac{1}{2}\, x^2 + c$$

is a solution. For the second

$$f(x) = \frac{1}{1-x}$$

is a solution. For the third $y = 0$ is a solution and so is $y = x$.

We will be interested in an ordinary differential equation (ODE) of the form

$$y' = f(y) + g(x)\,.$$

Here $y = y(x)$ is the unknown function and f, g are given functions. Note that while g only depend on x the function f also depend on the unknown function y.

We are interested in whether there exist solutions and when they exist if they are unique.

More precisely, suppose that we have the following:
- f be a continuously differentiable function on $\mathbf{R}$.
- g be a continuous function on $\mathbf{R}$.
- a is a real number.

We are intersted in existence and uniqueness of the ODE:

$$\begin{cases} y'(x) & = f(y(x)) + g(x) \\ y(0) & = a \,. \end{cases}$$

We will show next time the following:

Picard-Lindelöf theorem: There exists $\delta > 0$ such that there is a unique solution to this ODE on $(-\delta, \delta)$.

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
       TBB can be downloaded at:
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
       (screen-optimized)
       https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
       (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

TOBIAS HOLCK COLDING

Lecture 23

Ordinary differential equations: A differential equation is an equation that involves an unknown function and its derivative.

Suppose that we have the following:

- f be a continuously differentiable function on $\mathbf{R}$.
- g be a continuous function on $\mathbf{R}$.
- a is a real number.

We will be intersted in existence and uniqueness of the ODE:

$$(\dagger) \quad \begin{cases} y'(x) & = f(y(x)) + g(x)\,, \\ y(0) & = a\,. \end{cases}$$

We say that this is a first order equation since it only involves the function and its derivative and not higher derivatives.

The following theorem gives a satisfying answer to the question of existence and uniqueness for this ODE:

Picard-Lindelöf theorem: There exists $\delta > 0$ such that there is a unique solution to $(\dagger)$ on $(-\delta, \delta)$.

Before we prove this theorem let us recall a result that we have proven earlier. Suppose that $[a, b]$ is an interval and let $C([a, b])$ be the space of continuous functions on $[a, b]$. We equip this space with the metric d given by that if h_1, $h_2 \in C([a, b])$, then

$$d(h_1, h_2) = \max_{x \in [a,b]} |h_1(x) - h_2(x)|\,.$$

We proved earlier the following theorem:

Theorem 1: The metric space $(C([a,b]), d)$ is Cauchy complete.

We will also need to recall what it means for a map from a metric space to itself is contracting. A map T is a contracting map on a metric space (X, d) if for some $c < 1$ and all $x, y \in X$

$$d(T(x), T(y)) \leq c \, d(x, y).$$

We shall also use that we have proven the following fact:

Theorem 2: If (X, d) is a Cauchy complete metric space and $T : X \to X$ is a contracting map, then T has a unique fix point.

Indeed this theorem was proven by showing that for any $x \in X$, the sequence $x, T(x), T^2(x), T^3(x), \cdots$ is a Cauchy sequence and the limit is the unique fix point of T. The proof of this used that

$$d(T^{n+1}(x), T^n(x)) \leq c^n \, d(T(x), x),$$

and therefore by the triangle inequality

$$d(T^{n+k}(x), T^n(x)) \leq \sum_{i=1}^{k} d(T^{i+n}(x), T^{i-1+n}(x)) \leq c^{i-1+n} \, d(T(x), x).$$

Which is easily seen to imply that the sequence $T^n(x)$ is a Cauchy sequence.

We will also use the following lemma:

Lemma 1: Suppose that u_1 and u_2 are continuous functions on an interval I. Assume also that

- $u_1(x_0) = u_2(x_0)$.
- If $u_1(x) = u_2(x)$, then $u_1 = u_2$ in a neighborhood of x.

then $u_1 = u_2$.

Proof. Let

$$J_+ = \{z \in I \mid z \geq x_0 \text{ and } u_1(x) = u_2(x) \text{ for all } x \in [x_0, z]\}.$$

Then $x_0 \in J_+$ so $J_+ \neq \emptyset$. Let $z_0 = \sup J_+$, if $z_0 \in I$, then $u_1(z_0) = u_2(z_0)$ by continuity. Since also u_1 and u_2 agrees in a neighborhood of z_0 it follows that z_0 must be the right end point of I. Similarly one can show that $u_1 = u_2$ everywhere to the left of z_0. This proves the lemma. $\square$

Finally, in the proof of the Picard-Lindelöf theorem we will also need the next lemma. In this lemma f is a function on $\mathbf{R}$ as above, so f differentiable and the derivative of f is continuous and g will be a continuous function on $\mathbf{R}$. For $\delta > 0$, on the space of continuous functions on $[-\delta, \delta]$ we define a map T on functions y as follows

$$T(y)(x) = a + \int_0^x [f(z(s)) + g(s)]\, ds\,.$$

Note that when y is continuous, then so is $T(y)$.

Lemma 2: Let a be a constant and set $R = |a| + 2$. There exists a $\delta > 0$ such that:
- The map T maps the ball (in the metric space $(C([-\delta, \delta]), d)$) of radius R and with center the constant function zero into itself. We write $B_R(0)$ for this ball and so have that $T : B_R(0) \to B_R(0)$.
- The map T is contracting on $B_R(0)$.

Proof. Let

$$L_1 = \max_{|z| \leq R} |f(z)|\,,$$

$$L_2 = \max_{|x| \leq 1} |g(x)|\,.$$

We will first show that if that if we choose $\delta_0 > 0$ small enough, then T maps $B_R(0)$ into itself. That is, we will show that if $|y| \leq R$ on $[-\delta_0, \delta_0]$, then

$$|T(y)| \leq R\,.$$

To see this set

$$\delta_0 = \min\left\{1, \frac{1}{L_1 + 1}, \frac{1}{L_2 + 1}\right\}\,.$$

Now suppose that $|y| \leq R$ and $|x| \leq \delta_0$, then

$$|T(y)(x)| \leq |a| + \int_0^x |f(y(s))|\, ds + \int_0^x |g(s)|\, ds$$

$$\leq |a| + \delta_0\, L_1 + \delta_0\, L_2 \leq |a| + 2 = R\,.$$

This show that T maps $B_R(0)$ into itself.

Next set

$$M = \max_{|z| \leq R} |f'(z)|\,,$$

and

$$\delta = \min\left\{\delta_0, \frac{1}{2\,M + 1}\right\}\,.$$

Suppose that y_1 and y_2 are two continuous functions on $[-\delta, \delta]$ in $B_R(0)$, then

$$|T(y_1)(x) - T(y_2)(x)| = \int_0^x [f(y_1(s)) - f(y_2(s))]\, ds\ \ .$$

By the mean value theorem applied to f for each s we have a z_s between $y_1(s)$ and $y_2(s)$ such that

$$f(y_1(s)) - f(y_2(s)) = f'(z_s)\, (y_1(s) - y_2(s))\,.$$

Since $|y_i| \leq R$ we have that we have that for each s

$$|f(y_1(s)) - f(y_2(s))| \leq M\,|y_1(s) - y_2(s)| \leq M \max |y_1 - y_2| = M\,d(y_1, y_2)\,,$$

and therefore

$$|T(y_1)(x) - T(y_2)(x)| = \int_0^x [f(y_1(s)) - f(y_2(s))]\,ds \, \leq M\,\delta\,d(y_1, y_2) < \frac{1}{2}\,d(y_1, y_2)\,.$$

$\square$

We are now ready to show the Picard-Lindelöf theorem:

Proof. (of the Picard-Lindelöf theorem.) Let T be defined as above and R and $\delta > 0$ be given by Lemma 2. A fixed point for T is a function y such that $T(y) = y$. By the fundamental theorem of calculus if y is a fix point of T, then we have that

$$y'(x) = (T(y))'(x) = f(y(x)) + g(x)\,.$$

Moreover, $y(0) = a$. In other words any fix point of T is a solution to the ODE.

We need to show that the solution is unique. Suppose that y^* is any other solution, then by the fundamental theorem of calculus

$$y^*(x) = a + \int_0^x (y^*)'(s)\,ds = T(y^*(s))\,.$$

Note that this holds even if the interval I that y^* is defined on (containing 0) is different from $[-\delta, \delta]$. We have from this that any solution is a fix point of T. Since T is contracting on $B_R(0)$ it follows that for any fix point with $|y| \leq R$, then y is unique. In general, suppose that y_1 and y_2 are two solutions defined on intervals I_1 and I_2 both containing 0. We have from the above that they agree in a neighborhood of 0. The argument in Lemma 2 that proved uniqueness in a small neigborhood of 0 works equally well in a neighborhood of any other point. It now follow from Lemma 1 that y_1 and y_2 agrees everywhere. $\square$

References

[TBB] B.S. Thomson, J.B. Bruckner, and A.M. Bruckner, *Elementary Real Analysis, 2nd edition*
   TBB can be downloaded at:
   https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Landscape.pdf
   (screen-optimized)
   https://classicalrealanalysis.info/com/documents/TBB-AllChapters-Portrait.pdf
   (print-optimized)

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

18.100B Real Analysis

Spring 2025