Math 311 Notes

Erin P. J. Pearse

April 18, 2007
# Contents

0 Course Overview  
  0.1 Logic and inference ..................................................... 10

1 Real Numbers and Monotone Seqs  
  1.1 Introduction. Real Numbers ........................................... 15  
  1.2 Increasing sequences .................................................. 16  
  1.3 Limit of an increasing sequence ...................................... 17  
  1.4 Example: $e$ ............................................................... 18  
  1.5 Harmonic sum and Euler’s gamma ................................... 19  
  1.6 Decreasing seqs, Completeness ...................................... 20

2 Estimates and approximation  
  2.1 Introduction. Inequalities ............................................. 23  
  2.2 Estimations ............................................................... 24  
  2.3 Proving boundedness .................................................... 25  
  2.4 Absolute values. Estimating size .................................... 25  
  2.5 Approximation ........................................................... 27  
  2.6 “for $n$ large” .......................................................... 28

3 The Limit of a Sequence  
  3.1 Definition of limit ..................................................... 29  
  3.2 The uniqueness of limits. The $K - \varepsilon$ principle ............. 30  
  3.3 Infinite limits ............................................................ 31  
  3.4 An important limit ..................................................... 31  
  3.5 Writing limit proofs .................................................... 33  
  3.6 Some limits involving integrals ...................................... 33
3.7 Another limit involving integrals. ........................................ 34

4 Error Term Analysis .................................................. 35
4.1 The error term ..................................................... 35
4.2 Geometric series error term. ..................................... 36
4.3 Newton’s method .................................................. 36
4.4 The Fibonacci numbers .......................................... 37

5 The Limit Theorems .................................................. 41
5.1 Limits of sums, products, quotients ............................ 41
5.2 Comparison Theorems ........................................... 43
5.3 Location theorems ................................................ 44
5.4 Subsequences ..................................................... 45
5.5 Two common mistakes .......................................... 47

6 The Completeness Property ........................................ 49
6.1 Introduction. Nested intervals. ................................. 49
6.2 Cluster points .................................................... 50
6.3 Bolzano-Weierstrass Theorem .................................. 51
6.4 Cauchy sequences ................................................ 52
6.5 Completeness Property for sets ............................... 54

7 Infinite Series .......................................................... 57
7.1 Series and sequences ............................................ 57
7.2 Elementary convergence tests ................................. 59
7.3 Series with negative terms ..................................... 61
7.4 Ratio and Root tests ............................................ 64
7.5 Integral, p-series, and asymptotic comparison ............. 65
7.6 Alternating series test .......................................... 66
7.7 Rearrangements .................................................. 67
7.8 Multiplication of Series ....................................... 68

8 Power Series .......................................................... 71
8.1 Intro, radius of convergence ................................. 71
8.2 Convergence at endpoints, Abel summation .............. 73
8.3 Linearity of power series ................................... 73
CONTENTS

8.4 Multiplication of power series ........................................... 74

9 Functions of One Variable ..................................................... 75
  9.1 Functions ................................................................. 75
  9.2 Algebraic operations on functions ...................................... 77
  9.3 Properties of functions .................................................. 77
  9.4 Elementary functions .................................................... 78

10 Local and Global Behavior .................................................. 79
  10.1 Intervals. Estimating functions ....................................... 79
  10.2 Approximating functions ............................................... 81
  10.3 Local behavior ........................................................ 81
  10.4 Local and global properties .......................................... 82

11 Continuity and limits ......................................................... 83
  11.1 Continuous functions .................................................. 83
     11.1.1 Discontinuities .................................................. 85
  11.2 Limits of functions ..................................................... 85
  11.3 Limit theorems for functions .......................................... 86
  11.4 Limits and continuity .................................................. 86

12 Intermediate Value Theorem ................................................ 91
  12.1 Existence of zeros ..................................................... 91
  12.2 Applications of Bolzano ............................................... 92
  12.3 Monotonicity and the IVP ............................................. 92
  12.4 Inverse functions ..................................................... 93

13 Continuity and Compact Intervals ........................................ 95
  13.1 Compact intervals ..................................................... 95
  13.2 Bounded continuous functions ...................................... 96
  13.3 Extrema of continuous functions .................................... 97
  13.4 The mapping viewpoint .............................................. 98
  13.5 Uniform continuity .................................................. 98

14 Differentiation: Local Properties .......................................... 101
  14.1 The derivative ....................................................... 101
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.2</td>
<td>Differentiation formulas</td>
<td>103</td>
</tr>
<tr>
<td>14.3</td>
<td>Derivatives and local properties</td>
<td>104</td>
</tr>
<tr>
<td>15</td>
<td>Differentiation: Global Properties</td>
<td>107</td>
</tr>
<tr>
<td>15.1</td>
<td>The Mean-Value Theorem</td>
<td>107</td>
</tr>
<tr>
<td>15.4</td>
<td>L’Hôpital’s rule for indeterminate forms</td>
<td>108</td>
</tr>
<tr>
<td>16</td>
<td>Linearization and Convexity</td>
<td>111</td>
</tr>
<tr>
<td>16.1</td>
<td>Linearization</td>
<td>111</td>
</tr>
<tr>
<td>16.2</td>
<td>Applications to convexity</td>
<td>112</td>
</tr>
<tr>
<td>17</td>
<td>Taylor Approximation</td>
<td>115</td>
</tr>
<tr>
<td>17.1</td>
<td>Taylor polynomials</td>
<td>115</td>
</tr>
<tr>
<td>17.2</td>
<td>Taylor’s theorem with Lagrange remainder</td>
<td>116</td>
</tr>
<tr>
<td>17.3</td>
<td>Estimating error in Taylor’s approximation</td>
<td>117</td>
</tr>
<tr>
<td>17.4</td>
<td>Taylor series</td>
<td>117</td>
</tr>
<tr>
<td>18</td>
<td>Integrability</td>
<td>119</td>
</tr>
<tr>
<td>18.1</td>
<td>Introduction. Partitions</td>
<td>119</td>
</tr>
<tr>
<td>18.2</td>
<td>Integrability</td>
<td>119</td>
</tr>
<tr>
<td>18.3</td>
<td>Integrability of monotone or continuous $f$</td>
<td>122</td>
</tr>
<tr>
<td>18.4</td>
<td>Basic properties of integrable functions</td>
<td>123</td>
</tr>
<tr>
<td>19</td>
<td>The Riemann Integral</td>
<td>125</td>
</tr>
<tr>
<td>19.3</td>
<td>Riemann sums</td>
<td>125</td>
</tr>
<tr>
<td>19.4</td>
<td>Basic properties of the integral</td>
<td>126</td>
</tr>
<tr>
<td>19.5</td>
<td>Interval addition property</td>
<td>127</td>
</tr>
<tr>
<td>19.6</td>
<td>Piecewise properties</td>
<td>127</td>
</tr>
<tr>
<td>20</td>
<td>Derivatives and Integrals</td>
<td>131</td>
</tr>
<tr>
<td>20.1</td>
<td>First fundamental theorem of calculus</td>
<td>131</td>
</tr>
<tr>
<td>20.2</td>
<td>Second fundamental theorem of calculus</td>
<td>132</td>
</tr>
<tr>
<td>20.3</td>
<td>Other relations between integrals and derivatives</td>
<td>133</td>
</tr>
<tr>
<td>20.4</td>
<td>Logarithm and exponential</td>
<td>134</td>
</tr>
<tr>
<td>20.5</td>
<td>Stirling’s Formula</td>
<td>136</td>
</tr>
<tr>
<td>20.6</td>
<td>Growth rate of functions</td>
<td>136</td>
</tr>
</tbody>
</table>
21 Improper Integrals 137
   21.1 Basic Definitions ............................................. 137
   21.2 Comparison theorems ........................................... 138
   21.3 The Gamma function ........................................... 140
   21.4 Absolute and conditional convergence ....................... 141

22 Sequences and Series of Functions 143
   22.1 Pointwise and uniform convergence ........................... 143
   22.2 Criteria for uniform convergence .............................. 146
   22.3 Continuity and uniform convergence ......................... 147
   22.4 Term-by-term integration ...................................... 147
   22.5 Term-by-term differentiation .................................. 149
   22.6 Power series and analyticity .................................. 150
Chapter 0

Course Overview

The study of the real numbers, \( \mathbb{R} \), and functions of a real var, \( f(x) = y \) where \( x, y \) real.

Given \( f : \mathbb{R} \to \mathbb{R} \) which describes some system, how to study \( f \)?

- Need rigorous vocab for properties of \( f \) (definitions)
- Need to see when some properties imply others (theorems)

Result: can make inferences about the system.

Limits: the heart & soul of calculus.

Limits provide a rigorous basis for ideas like sequences, series, continuity, derivatives, integrals. More adv: model an arbitrary function as a limit of a sequence of “nice” functions (polys, trigs) or as a sum of “nice” functions (Fourier, wavelets). All of this requires understanding limits of numbers.

Outline:

1. Logic: not, and, or, implication; rules of inference
2. Sets: elements, intersection, union, containment; special sets
3. The real numbers: algebraic properties (+, ×), order properties (<), completeness properties

April 18, 2007
4. Sequences: types of, convergence, basic results (arithmetic, etc), subsequences, convergence, Cauchy sequences

5. Series: convergence tests, absolute convergence, power series

6. Functions: arith, behavior, continuity & limits, IVT, compact domains

7. Differentiation: MVT, L’Hopital, Taylor & linearization

8. Integrals: integrability and the Riemann integral

9. Special functions: exp, log, gamma

10. Seqs and series of functions

0.1 Logic and inference

Most theorems involve proving a statement of the form “if $A$ is true, then $B$ is true.” This is written $A \implies B$ and called if-then or implication. $A$ is the hypothesis and $B$ is the conclusion. To say “the hypothesis is satisfied” means that $A$ is true. In this case, one can make the argument

$$
\begin{array}{c}
A \implies B \\
A \\
B
\end{array}
$$

and infer that $B$ must therefore be true, also.

What does $A \implies B$ mean? We use the more familiar connectives “and” and “or” and “not” ($\neg$) to describe it, via truth tables. Consider:

<table>
<thead>
<tr>
<th>$A$</th>
<th>$B$</th>
<th>$A$ and $B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$A$</th>
<th>$B$</th>
<th>$A$ or $B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$A$</th>
<th>$\neg A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>
A \implies B means that whenever A is true, B must also be true, i.e., it CANNOT be the case that A is true B is false: \((A \implies B) \equiv \neg(A \text{ and } \neg B)\). This means that the truth table for \implies can be found:

\[
\begin{array}{ccc|c|c|c}
A & B & \neg B & A \text{ and } \neg B & \neg(A \text{ and } \neg B) & A \implies B \\
T & T & F & F & T & T \\
T & F & T & T & F & \Rightarrow T & F & F \\
F & T & F & F & T & F & T & T \\
F & F & T & F & T & F & F & T \\
\end{array}
\]

If \(A \implies B\) and \(B \implies A\), then the statements are equivalent and we write “A if and only if B” as \(A \iff B\). This is often used in definitions.

\[
\begin{array}{cccc|c|c|c}
A & B & \neg A & \neg B & A \implies B & \neg(A \text{ and } \neg B) & \neg A \text{ or } \neg B & \neg B \implies \neg A & B \implies A & \neg A \implies \neg B \\
T & T & F & F & T & T & T & T & T & T \\
T & F & F & T & F & F & F & T & F & T \\
F & T & T & F & T & T & T & T & F & F \\
F & F & T & T & T & T & T & T & T & T \\
\end{array}
\]

If you know that \(A \iff B\), then you can replace A with B (or v.v.) wherever it appears. \(A \equiv B\) is like “=” for logical statements.

One last rule (DeMorgan):

\[
\begin{array}{cccc|c|c|c|c|c|c}
A & B & \neg A & \neg B & \neg(A \text{ and } B) & \neg A \text{ or } \neg B & \neg(A \text{ or } B) & \neg A \text{ and } \neg B \\
T & T & F & F & F & F & F & F \\
T & F & F & T & T & F & F & F \\
F & T & T & F & T & F & F & F \\
F & F & T & T & T & T & T & T \\
\end{array}
\]
Thus, \( \neg(A \text{ and } B) \equiv (\neg A \text{ or } \neg B) \) and \( \neg(A \text{ or } B) \equiv (\neg A \text{ and } \neg B) \).

**Example 0.1.1.** Thm: a bounded increasing sequence converges.

This means: If a sequence \( \{a_n\} \) is increasing and bounded, then it converges, i.e.,

\[
(\{a_n\} \text{ increasing}) \text{ and } (\{a_n\} \text{ bounded}) \implies \{a_n\} \text{ converges.}
\]

Suppose we are considering the sequence where \( a_n = 1 - \frac{1}{n} \). We apply the theorem and see that \( a_n \) must converge (to something?).

Suppose we are considering the sequence \( a_n = (-1)^n \), which is known to diverge. The theorem is still helpful; by contrapositive,

\[
\neg(\{a_n\} \text{ converges}) \implies \neg((\{a_n\} \text{ increasing}) \text{ and } (\{a_n\} \text{ bounded}))
\]

\[
\{a_n\} \text{ diverges } \implies \neg(\{a_n\} \text{ increasing}) \text{ or } \neg(\{a_n\} \text{ bounded}),
\]

using DeMorgan. So \( a_n \) is either not increasing or unbounded. However, \( a_n \) is bounded, because every term is contained in the finite interval \([-1, 1]\). Thus, we can infer that \( a_n \) must not be increasing. (Note: not increasing does not imply decreasing!)

**How to prove** \( A \implies B \).

Direct proof.

1. Assume the hypothesis, i.e., assume \( A \) is true, just for now.

2. Apply this “fact” and other basic knowledge.

3. Show that \( B \) is true, based on all this.

**Example 0.1.2** (direct pf). \( n \text{ odd } \implies n^2 \text{ odd.} \)

1. Assume \( n \) is an odd integer.

2. Then \( n = 2k + 1 \), for some integer \( k \), so

\[
n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1 = 2m + 1, \text{ for some } m \in \mathbb{Z}.
\]

3. Thus, \( n^2 \) is odd.
Indirect proof: Proof by contrapositive.

\[(A \implies B) \equiv (\neg B \implies \neg A),\]

so show \(\neg B \implies \neg A\) directly.

**Example 0.1.3** (contrapositive). \(3n + 2\) odd \(\implies n\) odd.
The contrapositive is: \(n\) even \(\implies 3n + 2\) even.

1. Assume \(n\) is an even integer.
2. Then \(n = 2k\), for some integer \(k\), so

\[3n + 2 = 3(2k) + 2 = 6k + 2 = 2(3k + 1) = 2m, \text{ for some } m \in \mathbb{Z}.
\]
3. Thus, \(3n + 2\) is even.

**Example 0.1.4** (contrapositive). \(n^2\) even \(\implies n\) even.
This is just the contrapositive of the prev. example.

Indirect proof: Proof by contradiction.

In order to show that \(A\) is true by contradiction,

1. assume that \(A\) is false (assume \(\neg A\) is true)
2. derive a contradiction (show that \(\neg A\) implies something which is clearly false/impossible)

**Example 0.1.5** (contradiction). \(\sqrt{2}\) is irrational.

1. Assume the negative of the statement: \(\sqrt{2} = \frac{m}{n}\), for some \(m, n \in \mathbb{Z}\).
2. If \(m, n\) have a common factor, we can cancel it out to obtain

\[
\sqrt{2} = \frac{a}{b}, \quad \text{in lowest terms} \quad (*)
\]

\[
2 = \frac{a^2}{b^2}
\]

\[
2b^2 = a^2
\]

This shows \(a^2\) is even. But we just showed in the prev ex that

\[a^2 \text{ even } \implies a \text{ even},\]
so $a$ must be even. This means $a = 2c$ for some integer $c$, so

$$2b^2 = (2c)^2$$

$$b^2 = 2c^2$$

This shows that $b^2$ is even. But then $b$ must also be even. $\rightarrow(*)$

Mathematical (weak) induction: how to prove statements of the form

$$P(n)$$

is true for every $n$.

1. Basis step: show that $P(0)$ or $P(1)$ is true.

2. Induction step: show that $P(n) \implies P(n+1)$.

**Example 0.1.6** (induction). The sum of the first $n$ odd positive integers is $n^2$.

1. The sum of the first 1 positive integers is $1 = 1^2$.

2. Induction step: show that

$$[1 + 3 + 5 + \cdots + (2n - 1) = n^2] \implies [1 + 3 + 5 + \cdots + (2n - 1) + (2n + 1) = (n + 1)^2].$$

This is a statement $A \implies B$ which we show directly, so assume $A$ is true:

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$  

(This is the induction hypothesis.)

$$1 + 3 + 5 + \cdots + (2n - 1) + (2n + 1) = (1 + 3 + 5 + \cdots + (2n - 1)) + (2n + 1)$$

$$= n^2 + (2n + 1)$$

$$= (n + 1)^2.$$  

Thus we have shown that $B$ is true, based on the assumption $A$. Hence, we have proven the statement: $A \implies B$.

**Exercises:** A.4.1, A.4.10  

**Problems:** none  

**Due:** Jan. 29

1. Use the DeMorgan law to argue that $\neg(A \text{ and } \neg B) \equiv (\neg A \text{ or } B)$.

2. Use induction to show $n! \leq n^n$ for every $n \in \mathbb{N}$. 

Chapter 1

Real Numbers and Monotone Seqs

1.1 Introduction. Real Numbers.

\( \mathbb{R} \) is the set of real numbers.

How to define/consider them? Points on a line, decimal expansions, or ... ?

Want elements of \( \mathbb{R} \) to satisfy certain properties. Whenever \( a, b \in \mathbb{R} \), need to know:

1. Arithmetic: \( a + b, a - b, a \times b, a/b \)

2. Order: \( a < b, a \leq b \)

3. Completeness: a bounded sequence of increasing numbers has a limit.

   Other desirable things (which will follow from the above):

4. Archimedean: if \( a > 0 \), then for any \( N \) (no matter how large), we can find \( b \in \mathbb{R} \) such that \( ab > N \).

5. Distance: \( d(a, b) = |a - b| \).

Points on a line and decimal expansions have problems: not so good for computation, nonuniqueness, etc.

\(^{1}\)April 18, 2007
Alternative: we have 1, 2, 4, 5 already for the rational numbers \( \mathbb{Q} \), so start with them.

\[ \mathbb{R} = \mathbb{Q} \cup \{ \text{limits of points of } \mathbb{Q} \}. \]

Recall: for \( \mathbb{Q} = \{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \} \) we have

\[
\frac{p}{q} + \frac{r}{s} = \frac{ps + qr}{qs}, \quad \frac{p}{q} \times \frac{r}{s} = \frac{pr}{qs}.
\]

This will give us (3), and the other.

Example of the main idea: understanding \( a + b \), when \( a, b \in \mathbb{R} \).

1. By the defn of \( \mathbb{R} \), \( a = \lim a_n \), \( b = \lim b_n \), where \( a_n, b_n \in \mathbb{Q} \).

2. We know what \( a_n + b_n \) is, since we known how + works in \( \mathbb{Q} \).

3. If \( \lim a_n + \lim b_n = \lim(a_n + b_n) \), then define

\[ a + b := \lim(a_n + b_n). \]

Same technique works for things more complicated than \( a + b \).

(There are some technicalities, e.g., when is (3) true, indep of limit representation.)

### 1.2 Increasing sequences.

**Definition 1.2.1.** A sequence of numbers is an infinite ordered list

\[ a_1, a_2, \ldots \]

\( a_n \) is the \( n^{th} \) term.

A sequence can be specified by giving

(i) the first few terms: \( \{1, \frac{1}{2}, \frac{1}{3}, \ldots \} \)

(ii) an explicit formula for the \( n^{th} \) term: \( \{ \frac{1}{n} \} \), or

(iii) a recurrence relation for the \( n^{th} \) term: \( a_1 = 1, a_{n+1} = \frac{n-1}{n}a_n \).

**Example 1.2.1.** The Fibonacci numbers can be described by
1.3 Limit of an increasing sequence.

(i) \( \{1, 1, 2, 3, 5, 8, 13, 21, \ldots \} \)

(ii) \( \left\{ \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n \right\} \), or

(iii) \( a_0 = 1, a_1 = 1, a_{n+2} = a_{n+1} + a_n \).

**Definition 1.2.2.** \( \{a_n\} \) is **increasing** iff \( a_n \leq a_{n+1}, \forall n \).

\( \{a_n\} \) is **strictly increasing** iff \( a_n < a_{n+1}, \forall n \).

\( \{a_n\} \) is **decreasing**, (strictly decreasing) iff \( a_n \geq a_{n+1} \) (\( a_n > a_{n+1} \)), \( \forall n \).

1.3 Limit of an increasing sequence.

Suppose we are using decimal representations.

**Definition 1.3.1.** A real number \( L \) is the **limit** of an increasing sequence \( \{a_n\} \) if, given any integer \( k > 0 \), all the \( a_n \) after some point in the sequence agree with \( L \) to \( k \) decimal places:

\[
L = \lim_{n \to \infty} a_n, \quad \text{or} \quad a_n \xrightarrow{n \to \infty} L.
\]

We say \( a_n \) **converges** (to \( L \)).

In symbols,

\[
\forall k \in \mathbb{N}, \exists N \text{ such that, for } n \geq N, a_n \text{ agrees with } L \text{ to } k \text{ decimal places.}
\]

or,

\[
\forall k \in \mathbb{N}, \exists N \text{ such that } n \geq N \implies |a_n - L| < 10^{-k},
\]

or,

\[
\forall \varepsilon > 0, \exists N \text{ such that } n \geq N \implies |a_n - L| < \varepsilon.
\]

This last one doesn’t refer to the decimal expansion.

If a sequence has a limit, that limit is unique. A sequence may not have a limit, like \( \{n\} = \{1, 2, 3, 4, \ldots \} \), or the Fibo.

**Definition 1.3.2.** A sequence \( \{a_n\} \) is **bounded above** if there is a number \( B \in \mathbb{R} \) such that \( a_n \leq B, \forall n \). This \( B \) is an **upper bound** for the sequence \( \{a_n\} \).
Theorem 1.3.3. An increasing sequence which is bounded above has a limit.

\[ \{ \{ a_n \} \text{ is bounded above} \} \text{ and } \{ \{ a_n \} \text{ is increasing} \} \implies \{ a_n \} \text{ has a limit.} \]

We will see why this is true in Ch. 6, using the notion of sup. If \( \beta \in \mathbb{R} \) satisfies

1. \( a_n \leq \beta, \forall n \) (so \( \beta \) is an upper bound of \( \{ a_n \} \), and

2. \( (a_n \leq b, \forall n) \implies \beta \leq b, \)

then we call \( \beta \) the least upper bound (or supremum) of \( \{ a_n \} \):

\[ \beta = \sup_{\{ a_n \}}. \]

We will see that for a pos, incr seq, \( \lim a_n = \sup a_n. \)

1.4 Example: \( e \)

We use two results from discrete math.

Binom formula:

\[ (1 + x)^k = 1 + kx + \cdots + \binom{k}{i} x^i + \cdots + x^n \]

Geometric sum (finite):

\[ 1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}. \]

When \( r = \frac{1}{2} \), this gives \( 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} = 2 \left( 1 - \frac{1}{2^{n+1}} \right) < 2. \)

Theorem 1.4.1. The sequence \( a_n = \left( 1 + \frac{1}{2^n} \right)^{2^n} \) has a limit. (The limit is \( e \)).

Proof. By Thm. 1.3.3, NTS \( a_n \) bdd & incr. Since \( n \to \infty \), suffices to consider \( n \geq 2. \)

\( a_n \) is incr: need \( (1 + \frac{1}{2^n})^{2^n} < (1 + \frac{1}{2^{n+1}})^{2^{n+1}}. \)

\[ b \neq 0 \implies b^2 > 0 \implies (1 + b)^2 > 1 + 2b \implies ((1 + b)^2)^{2^n} > (1 + 2b)^{2^n} \implies (1 + \frac{1}{2^{n+1}})^{2^{n+1}} > (1 + \frac{1}{2^n})^{2^n} \implies \lim a_n = \frac{1}{2^{n+1}}. \]
1.5 Harmonic sum and Euler’s gamma

\(a_n\) is bounded above. First, note that

\[ k(k-1) \ldots (k-i+1) \leq k^i, \quad \text{and} \quad \frac{1}{i!} = \frac{1}{i} \cdot \frac{1}{i-1} \ldots \frac{1}{2} \leq \left( \frac{1}{2} \right)^{i-1}. \]

Then

\[
\left(1 + \frac{1}{k}\right)^k = 1 + k \left(\frac{1}{k}\right) + \cdots + \frac{k(k-1) \ldots (k-i+1)}{i!} \left(\frac{1}{k}\right)^i + \cdots + \frac{k!}{k!} \left(\frac{1}{k}\right)^k \\
= 1 + k \left(\frac{1}{k}\right) + \cdots + k^i \left(\frac{1}{2}\right)^i \left(\frac{1}{k}\right)^i + \cdots + k^k \left(\frac{1}{2}\right)^{k-1} \left(\frac{1}{k}\right)^k \\
< 1 + 1 + \frac{1}{2} + \cdots + \frac{1}{2^{i-1}} + \cdots + \frac{1}{2^{k-1}} \\
< 1 + 2 = 3.
\]

So 3 is an upper bound for \(a_n\). \(\square\)

1.5 Harmonic sum and Euler’s gamma

Definition 1.5.1. The harmonic sum or harmonic series is the infinite sum

\[ 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n}. \]

Proposition 1.5.2. Let \(a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\), so that \(a_n\) is the \(n^{th}\) partial sum of the harmonic series. Then \(a_n\) has no upper bound (hence the infinite sum diverges).

Proof. Write the \((2^k)^{th}\) term

\[
a_{2^k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k}\right) \\
> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^k} + \cdots + \frac{1}{2^k}\right) \\
= 1 + \frac{1}{2} + (k-1) \frac{1}{2}.
\]

So \(a_n\) becomes arbitrarily large. \(\square\)

Theorem 1.5.3. Let \(b_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log(n+1), n \geq 1\). Then \(\{b_n\}\) converges.
Proof. Show \( b_n \) is increasing and bounded above.

\[
b_n = \text{area of rectangles} - \text{area under curve} = T_1 + \cdots + T_2.
\]

Each “trianglet” has positive area, so \( T_i \geq 0 \) implies

\[
b_{n+1} = b_n + T_{n+1} \implies \{b_n\} \text{ increasing.}
\]

Each trianglet can be horizontally translated into the initial rectangle, so \( b_n \) is bounded above by 1.

If the curved “hypotemuses” were replaced by straight lines,

\[
\gamma = T_1 + T_2 + \ldots
\]

would be exactly half the area of the original rectangle, so \( \frac{1}{2} \). Thus, \( \frac{1}{2} \leq \gamma \leq 1 \). However, \( \gamma \approx 0.577\ldots \).

1.6 Decreasing seqs, Completeness

Eventually (Chap. 6), \( \mathbb{R} \) is complete because every nonempty subset \( A \subseteq \mathbb{R} \) which is bounded above has a least upper bound \( \sup A \in \mathbb{R} \). Note: \( \sup A \) is an element of \( \mathbb{R} \); it may not be an element of \( A \).

Until then: we continue to use monotone sequences.

Definition 1.6.1. A real number \( L \) is the limit of a decreasing sequence \( \{a_n\} \) if, given any integer \( k > 0 \), all the \( a_n \) after some point in the sequence agree with \( L \) to \( k \) decimal places:

\[
L = \lim_{n \to \infty} a_n, \quad \text{or} \quad a_n \xrightarrow{n \to \infty} L.
\]

We say \( a_n \) converges (to \( L \)).

When rephrased in symbols, it is identical to previous:

\[
\forall \varepsilon > 0, \exists N \text{ such that } n \geq N \implies |a_n - L| < \varepsilon.
\]
Decreasing seqs, Completeness

Note: for monotonic sequences, sometimes write

\[ a_n \nearrow L, \text{ or } a_n \searrow L. \]

**Definition 1.6.2.** \( \{a_n\} \) is **bounded below** if there is a number \( B \in \mathbb{R} \) such that \( a_n \geq B, \forall n \).
This \( B \) is an **lower bound** for the sequence \( \{a_n\} \).

**Theorem 1.6.3.** A positive decr seq has a limit.

Note: (\( \{a_n\} \) positive) \( \equiv (a_n \geq 0, \forall n) \equiv (\{a_n\} \) is bounded below by 0).

**Definition 1.6.4.** \( \{a_n\} \) is **bounded** iff it is bounded above and bounded below.

**Definition 1.6.5.** \( \{a_n\} \) is **monotone** iff it is increasing or decreasing.

**Theorem 1.6.6** (Completeness). A bounded monotone sequence in \( \mathbb{R} \) has a limit.

**Exercises:** 1.2.1, 1.3.1, 1.4.2, 1.5.1, 1.5.2  **Problems:** 1-1, 1-2  
**Due:** Jan. 29

1. Prove that \( 1.0000 \cdots = 0.9999 \cdots \) using the geometric series formula:

\[ 1 + r + r^2 + r^3 + \cdots = \sum_{n=0}^{\infty} r^n = \frac{1}{1-r}, \quad \text{for } |r| < 1. \]

Hint: use \( r = \frac{1}{10} \).

2. Briefly explain why \( (1 + b)^2 > 1 + 2b \) in the proof of the Thm. in §1.4.1.

3. Briefly explain why the conclusion of the proof of Thm. in §1.5.2 follows from the Archimedean Property of \( \mathbb{R} \).
Chapter 2

Estimates and approximation

2.1 Introduction. Inequalities.

Inequalities: for making comparisons.
Absolute values: for measuring size and distance.
a is (strictly) positive iff \( a > 0 \); a is nonnegative iff \( a \geq 0 \).
Properties of \(<\):
1. \( a \not< a \)
2. \( a < b \implies b \not< a \)
3. \( a < b \) and \( b < c \) implies \( a < c \) (transitivity)
Properties of \(\leq\):
1. \( a \leq a \)
2. Either \( a \leq b \) or \( b \leq a \) is true. If both are true, write \( a = b \).
3. \( a \leq b \) and \( b \leq c \) implies \( a \leq c \) (transitivity)
Note: the negation of \( a < b \) is \( b \leq a \), not \( b < a \).

Arithmetic with inequalities.

- Addition: for any \( a, b \in \mathbb{R} \), \( a < b, c < d \implies a + c < b + d \).
• Multiplication: for any $a, b >, a < b, c < d \implies ac < bd$.

• Negation: for any $a, b \in \mathbb{R}, a < b \implies -a > -b$.

• Reciprocals: for any $a, b >, a < b \implies \frac{1}{a} > \frac{1}{b}$.

Note. This is not a theorem (yet) but it is a handy rule:

1. Functions with everywhere positive derivatives are monotonic increasing, and hence order-preserving. E.g., $f(x) = \log(x)$ has derivative $f'(x) = \frac{1}{x}$. Then
   
   $$a < b \implies \log a < \log b$$

2. Functions with everywhere negative derivatives are monotonic decreasing, and hence order-reversing. E.g., $f(x) = e^{-x}$ has derivative $f'(x) = -e^{-x}$. Then
   
   $$a < b \implies e^{-a} > e^{-b}$$

3. Of course, some function are neither. E.g., $f(x) = \cos x$. Then if $a < b$, you know nothing about $\cos a$ or $\cos b$.

The Negation rule demonstrates $f(x) = -x$, where $f'(x) = -1 < 0, \forall x$.

The Reciprocal rule demonstrates $f(x) = \frac{1}{x}$, where $f'(x) = -\frac{1}{x^2} < 0$, for $x > 0$.

### 2.2 Estimations

**Definition 2.2.1.** If $a < b < c$, then $a$ is a *lower estimate* for $b$, and $c$ is an *upper estimate* for $b$. The same is said if $a \leq b \leq c$.

**Definition 2.2.2.** If $a < a' < b < c' < c$, then $a'$ and $c'$ are *stronger* estimates ($a'$ is stronger than $a$, etc.) and $a, c$ are *weaker*. “stronger” = more info.

**Example 2.2.1.** If $0 < x < 1$, then we have the estimates $0 < x^2 < 1$ and $1 < \frac{1}{x}$.

**Example 2.2.2.** Give upper and lower estimates for $\frac{1}{1+x^2}$.

$$x^2 \geq 0 \implies 1 + x^2 \geq 1 \implies \frac{1}{1+x^2} \leq 1$$

$$x^2 \geq 0 \implies 1 + x^2 \geq 1 > 0 \implies \frac{1}{1+x^2} > 0$$
2.3 Proving boundedness

Example 2.2.3. If $|1 - 3x| < 2$, give upper and lower estimates for $x$.

$$|1 - 3x| < 2 \iff -2 < 1 - 3x < 2 \iff -3 < -3x < 1 \iff 3 > 3x > -1 \iff 1 > x > -\frac{1}{3}.$$ 

Example 2.2.4. Use $2 < \sqrt{5} < 3$ to show $\frac{1}{4} < \frac{\sqrt{5} - 1}{\sqrt{5} + 1} < \frac{2}{3}$.

$$\frac{\sqrt{5} - 1}{\sqrt{5} + 1} < \frac{3 - 1}{\sqrt{5} + 1} = \frac{2}{\sqrt{5} + 1} < \frac{2}{2 + 1} = \frac{2}{3}$$

$$\frac{\sqrt{5} - 1}{\sqrt{5} + 1} > \frac{2 - 1}{\sqrt{5} + 1} = \frac{1}{\sqrt{5} + 1} > \frac{1}{3 + 1} = \frac{1}{4}$$

2.3 Proving boundedness

To show that $\{a_n\}$ is bounded: find one upper estimate $a_n \leq B, \forall n$.

To show that $\{a_n\}$ is not bounded: find lower estimate for each term, $a_n \geq B_n$, with $B_n \to \infty$.

Example 2.3.1. Earlier, we showed $b_k = (1 + \frac{1}{k})^k < 3$.

Example 2.3.2. Earlier, we showed $a_{2^n} = \sum_{k=1}^{2^n} \frac{1}{k} > \frac{3}{2} + (n - 1)/2$.

2.4 Absolute values. Estimating size.

Definition 2.4.1. The absolute value of $a \in \mathbb{R}$ is its size; i.e., its distance from 0:

$$|a| = \begin{cases} a, & a \geq 0 \\ -a, & a < 0. \end{cases}$$

The size of the difference between $a$ and $b$ is the distance from $a$ to $b$:

$$|a - b| = \text{dist}(a, b).$$
Note: $|a| = |a - 0|$.  

In higher dimensions (say in the plane, $\mathbb{R}^2$) the defn remains the same. (sketch):

$S^1 = \{x \in \mathbb{R}^2 : |x| = 1\}$

$B^1 = \{x \in \mathbb{R}^2 : |x| < 1\}$

**Theorem 2.4.2.** Let $a \in \mathbb{R}$.

(i) $|a| \geq 0$ and $|a| = 0 \implies a = 0$.  

(ii) $|a| < M$ is the same thing as $-M < a < M$.

Note: $|a - b| < M$ means that the distance from $a$ to $b$ is less than $M$:  

$$a \in (b - M, b + M) \quad \text{or} \quad b \in (a - M, a + M).$$

If you want to show that $a$ is close to $b$, show that $|a - b| < \varepsilon$, where $\varepsilon > 0$ is very small.

**Theorem 2.4.3** (Absolute Value Laws).

(i) Product law: $|ab| = |a||b|$. This implies division law.  

(ii) Triangle inequality: $|a + b| \leq |a| + |b|$.  

Proof. Use Thm. 2.4.2 (ii) twice:

\[
[Proof 1] -|a| \leq a \leq |a| \\
-|b| \leq b \leq |b| \\
-(|a| + |b|) \leq a + b \leq |a| + |b| \\
|a + b| \leq ||a| + |b|| = |a| + |b|.
\]

\[\square\]

\[Proof 2. \ |a + b|^2 = (a + b)(a + b) = a^2 + 2ab + b^2 \leq a^2 + 2|a||b| + b^2 = (|a| + |b|)^2. \ \square\]

(iii) $|a - b| \geq |a| - |b|$, $|a + b| \geq |a| - |b|$, and $||a| - |b|| \leq |a - b|$.  

**Example 2.4.1.** Fourier analysis:

$$S_n = c_1 \cos t + c_2 \cos 2t + \cdots + c_n \cos nt.$$  

If $c_i = 1/2^i$, give an upper estimate for $S_n$.  


Solution. Since $|\cos x| \leq 1, \forall x \in \mathbb{R}$,

$$|S_n| \leq |c_1||\cos t| + |c_2||\cos 2t| + \cdots + |c_n||\cos nt|$$

$$\leq \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 1 + \cdots + \frac{1}{2^n} \cdot 1$$

$$< 1$$

\[\square\]

**Theorem 2.4.4.** \{a_n\} is bounded $\iff$ there is a $B$ such that $|a_n| \leq B, \forall n$.

Proof. ($\Rightarrow$) The hypothesis means that $K \leq a_n \leq L$ for all $n$. Then $B = \max(|K|, |L|)$.

($\Leftarrow$) $|a_n| \leq B$ implies that $B \leq a_n \leq B$, so \{a_n\} is bounded. \[\square\]

### 2.5 Approximation

**Theorem 2.5.1 (Density of $\mathbb{Q}$ in $\mathbb{R}$).** Let $a < b$ be real numbers. Then

(i) $\exists r \in \mathbb{Q}$ such that $a < r < b$, and

(ii) $\exists s \in \mathbb{R} \setminus \mathbb{Q}$ such that $a < s < b$.

Proof. (i) $b - a > 0$, so we can find $n$ such that $n(b - a) > 1$ by the Archimedean property. Then $1 + na < nb$. Let $m$ be an integer such that $m - 1 \leq na < m$. (This is possible, since $\bigcup_{m \in \mathbb{Z}}[m - 1, m) = \mathbb{R}$.) Then $m \leq 1 + na$.

$$na < m \leq 1 + na < nb$$

$$a < \frac{m}{n} < b.$$

(ii) From (i) we have $a < r < b$, with $r \in \mathbb{Q}$. Since $\sqrt{2}$ is irrational, so is $c\sqrt{2}$ for $c \in \mathbb{R}$.

$$a < r + \frac{\sqrt{2}}{n} < b,$$

for some $n \in \mathbb{N}$. \[\square\]

**Definition 2.5.2.** $a \approx \varepsilon b$ means $|a - b| < \varepsilon$.

Example: $2 \approx_{\varepsilon} 3$ but it is not true that $2 \approx_{\varepsilon} 3$.

Given $x \in \mathbb{R}$ and $\varepsilon > 0$, Thm. 2.5.1 says one can thus always find $r \in \mathbb{Q}$ such that $x \approx_{\varepsilon} r$.

**Theorem 2.5.3.** (i) $a \approx_{\varepsilon} b$ and $b \approx_{\varepsilon'} c \implies a \approx_{\varepsilon + \varepsilon'} c$. (transitivity)
(ii) \( a \approx \epsilon^* a' \) and \( b \approx \epsilon^* b' \) \( \implies \) \( a + b \approx \epsilon^* a' + b' \). (addition)

2.6 “for \( n \) large”

**Definition 2.6.1.** The sequence \( \{a_n\} \) has property \( \mathcal{P} \) for \( n \) large (write “\( n \gg 1 \)”) if there is a number \( N \in \mathbb{N} \) such that \( a_n \) has property \( \mathcal{P} \) for all \( n \geq N \).

**Example 2.6.1.** \( \frac{1}{n} < 0.001 \) for \( n \gg 1 \).

**Example 2.6.2.** \( \frac{n^2}{e^n} \) is decreasing for \( n \gg 1 \).

**Example 2.6.3.** If \( \{a_n\} \) is bounded above for \( n \gg 1 \), then it is bounded above.

*Proof.* The hypothesis mean that there is a \( B \) and an \( N \) such that \( a_n \leq B \) for \( n \geq N \). Let \( M = \max\{a_1, a_2, \ldots, a_n, B\} \). Then \( a_n \leq M < \infty \). \( \square \)

**Example 2.6.4.** \( \{a_n\} \) and \( \{b_n\} \) increasing for \( n \gg 1 \) \( \implies \) \( \{a_n + b_n\} \) is, too.

*Proof.* By hypothesis,

\[
\begin{align*}
a_n &\leq a_{n+1} & \text{for } n \geq N_1, \text{ and} \\
b_n &\leq b_{n+1} & \text{for } n \geq N_2.
\end{align*}
\]

Choose \( N \geq N_1, N_2 \). Then

\[
\begin{align*}
a_n &\leq a_{n+1} & \text{for } n \geq N, \text{ and} \\
b_n &\leq b_{n+1} & \text{for } n \geq N, \text{ so} \\
a_n + b_n &\leq a_{n+1} + b_{n+1} & \text{for } n \geq N. \quad \square
\end{align*}
\]

**Question 1.** What does it mean if \( |a - b| < \frac{1}{n}, \forall n \in \mathbb{N} \)?

**Exercises:** 2.1.2, 2.2.1, 2.4.2, 2.4.7, 2.5.2, 2.6.1

**Problems:** 2-1

**Due:**

Feb. 5

1. Prove \( n! \leq n^n \) for every \( n \in \mathbb{N} \), without using induction.

2. Prove that \( ||x| - |y|| \leq |x - y| \) for \( x, y \in \mathbb{R} \).
Chapter 3

The Limit of a Sequence

3.1 Definition of limit.

Definition 3.1.1. The number $L$ is the limit of the sequence $\{a_n\}$ iff given $\varepsilon > 0$, we have $a_n \approx L$ for $n \gg 1$:

$$\forall \varepsilon > 0, \exists N \text{ such that } n \geq N \implies |a_n - L| < \varepsilon.$$

That is, $a_n$ is a good approximation to $L$ for large $n$. And it can be made increasingly better by taking $n$ larger.

To prove a limit: typically, find a rule for $N$ in terms of $\varepsilon$.

Example 3.1.1. Prove $\lim_{n \to \infty} \frac{n-1}{n+1} = 1$.

Solution. Given $\varepsilon > 0$, we need to show $\frac{n-1}{n+1} \approx 1$ for $n \gg 1$. That is, we need

$$\left| \frac{n-1}{n+1} - 1 \right| = \left| \frac{n-1 - n + 1}{n+1} \right| = \frac{2}{n+1} < \varepsilon.$$

Observe: $\frac{2}{n+1} < \varepsilon \iff \frac{2}{\varepsilon} < n+1$. So for $N = \frac{2}{\varepsilon} - 1$, the estimate will hold for $n \geq N$. \qed

Example 3.1.2. Prove $\lim_{n \to \infty} (\sqrt{n+1} - \sqrt{n}) = 0$. 

\footnote{April 18, 2007}
Proof. Use the identity \( A - B = \frac{A^2 - B^2}{A + B} \) to get

\[
\left| \sqrt{n + 1} - \sqrt{n} \right| = \frac{1}{\sqrt{n + 1} + \sqrt{n}} < \frac{1}{2\sqrt{n}}.
\]

Then since

\[
\frac{1}{2\sqrt{n}} < \varepsilon \iff \frac{1}{4\varepsilon^2} < n,
\]
we will have the required estimate for \( n \geq N = \frac{1}{4\varepsilon^2}. \)

\[\square\]

Theorem 3.1.2. \( a_n \to 0 \iff |a_n| \to 0. \)

Proof. HW.

\[\square\]

3.2 The uniqueness of limits. The \( K - \varepsilon \) principle.

Theorem 3.2.1 (Uniqueness of limits). A sequence \( a_n \) has at most one limit.

Proof. We must show that \( (a_n \to L) \) and \( (a_n \to L') \implies L = L' \). Assume that both \( L \) and \( L' \) are limits of \( a_n \) and suppose, by way of contradiction, that \( L \neq L' \). Then we may choose \( \varepsilon = \frac{1}{2}|L - L'| \), so that \( \varepsilon > 0 \) and \( 2\varepsilon = |L - L'| \). From the assumptions, we have \( a_n \approx \varepsilon L \) and \( a_n \approx \varepsilon L' \), for sufficiently large \( n \). Thus,

\[
L \approx \varepsilon a_n \approx \varepsilon L' \implies L \approx_{2\varepsilon} L' \iff |L - L'| < 2\varepsilon = |L - L'|. \quad \nexists \ a \neq a. \quad \square
\]

Theorem 3.2.2 (Incr Seq Thm). \( \{a_n\} \) is increasing, \( \lim a_n = L \implies a_n \leq L, \forall n. \)

Proof. We show the contrapositive (WHAT IS IT?), so suppose \( a_k > L \) for some \( k \). Now:

1. If \( \{a_n\} \) is increasing, then \( |a_n - L| \geq |a_k - L| > 0 \) for \( n \geq k \), so \( \lim a_n \neq L \).

2. If \( \lim a_n = L \), then for \( \varepsilon = (a_k - L) \), we can find \( N \) such that \( a_n \approx \varepsilon L \) for \( n \geq N \).

But then \( a_n < a_k \) and \( n \geq k \), for all these \( n \).

\[\square\]

Theorem 3.2.3. \( \{a_n\} \) is decreasing and \( \lim a_n = L \implies a_n \geq L \) for all \( n \).

Example:

Theorem 3.2.4. \( \lim a_n = a \) and \( \lim b_n = b \implies \lim(a_n + b_n) = a + b. \)
Proof. Given \( \varepsilon > 0 \), we can find \( N_1, N_2 \) such that
\[
\begin{align*}
    n \geq N_1 & \implies |a_n - a| < \varepsilon, \quad \text{and} \quad n \geq N_2 \implies |b_n - b| < \varepsilon. \\
\end{align*}
\]
Then let \( N = \max(N_1, N_2) \) and
\[
\begin{align*}
    n \geq N & \implies |(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \\
            & \leq |a_n - a| + |b_n - b| < \varepsilon + \varepsilon = 2\varepsilon. \quad \square
\end{align*}
\]

Theorem 3.2.5 (The \( K-\varepsilon \) principle.). \( \{a_n\} \) is a sequence and for any \( \varepsilon > 0 \), it is true that \( a_n \approx_{K\varepsilon} L \) for \( n \gg 1 \), where \( K > 0 \) is a fixed constant. Then \( \lim a_n = L \).

3.3 Infinite limits.

Definition 3.3.1. \( \lim a_n = \infty \) means that for any \( M > 0 \), we have \( a_n > M \) for \( n \gg 1 \). Then \( a_n \) tends to infinity: \( a_n \to \infty \).

Example 3.3.1. \( \{\log n\} \) tends to \( \infty \).
Since \( \log x \) is monotone increasing, it is order-preserving:
\[
a < b \implies \log a < \log b.
\]
Given \( M > 0 \), we then have
\[
n > e^M \implies \log n > \log e^M = M.
\]

3.4 An important limit.

Theorem 3.4.1 (The limit of \( a^n \)).
\[
\lim_{n \to \infty} = \begin{cases} 
\infty, & a > 1, \\
1, & a = 1, \\
0, & |a| < 1.
\end{cases}
\]
Proof. Case $a > 1$: let $M > 0$ be given. Since $a = 1 + x$ for some $x > 0$,

$$a^n = (1 + x)^n = 1 + nx + \cdots + \binom{n}{k}x^k + \cdots + x^n.$$  

All the terms on the right are positive, so

$$a^n > 1 + nx > M \iff n > (M - 1)/x.$$

So choose $N = (M - 1)/x$. Case $a = 1$: obvious.

Case $|a| < 1$: let $\varepsilon > 0$ be given. Then

$$|a| < 1 \implies \frac{1}{|a|} > 1 \implies \left(\frac{1}{|a|}\right)^n > \frac{1}{\varepsilon} \implies |a|^n < \varepsilon. \quad \square$$

It is also useful to note that for $c > 0$,

$$0 < a < 1 \implies 0 < ac < c \implies 0 < a(a^n) < a^n \implies a^{n+1} < a^n.$$  

So the limit is monotone decreasing for small $a$. Also,

$$a > 1 \implies ac > c \implies a(a^n) > a^n \implies a^{n+1} > a^n,$$

so the limit is monotone increasing for large $a$.

Example 3.4.1. Let $f_1(x) = x, f_2 = x^2, ... f_n(x) = x^n$ be a sequence of functions, each defined on $[0,1]$. Define a new function as the pointwise limit:

$$f(x) := \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} x^n$$  

$$= \begin{cases} 
0, & 0 \leq x < 1, \\
1, & a = 1. 
\end{cases}$$

Each of the functions $f_n(x) = x^n$ is continuous, but the limit $f(x)$ is not continuous!
3.5 Writing limit proofs.

Read on your own ... each student will get something different.

\[ A < n(n + 1) \quad \text{rather than} \quad A < n(n + 1) \]
\[ = n^2 + n \quad < n^2 + n \]

Grading: note §3.5

3.6 Some limits involving integrals.

Example 3.6.1. Let \( a_n := \int_0^1 (x^2 + 2)^n \, dx \). Show that \( \lim a_n = \infty \).

Solution. Estimate the integrand from below:

\[ x^2 + 2 \geq 2, \forall x \]
\[ (x^2 + 2)^n \geq 2^n, \forall x, n \]
\[ \int_0^1 (x^2 + 2)^n \, dx \geq \int_0^1 2^n \, dx = 2^n. \quad \text{to be shown ...} \]

Since \( 2^n \to \infty \), \( a_n \to \infty \): given \( M > 0 \),
\[ n > \log_2 M \implies \int_0^1 (x^2 + 2)^n \, dx \geq 2^n \geq M. \]

Question 2. Let \( a_n := \int_0^1 (x + 2)^n \, dx \). If you are to show that \( \lim a_n = \infty \), how would the argument differ from the above?

Example 3.6.2. Show \( \int_0^1 (x^2 + 1)^n \, dx \to \infty \).

Solution. The previous argument gives \( (x^2 + 1)^n \geq 1^n = 1 \), useless.

Since \( f(x) = x^2 + 1 \) is increasing with \( f(0.1) = 1.01 \),

\[ x^2 + 1 \geq 1.01 > 1, \text{ for } 0.1 \leq x \leq 1 \]
\[ (x^2 + 1)^n \geq (1.01)^n, \text{ for } 0.1 \leq x \leq 1 \]
\[ \int_{0.1}^1 (x^2 + 1)^n \, dx \geq \int_{0.1}^1 (1.01)^n \, dx = \frac{9}{10}(1.01)^n \to \infty. \]
3.7 Another limit involving integrals.

Example 3.7.1. Let \( a_n := \int_0^{\pi/2} \sin^n x \, dx \). Determine \( \lim a_n \).

Solution. Observe:

\[
0 \leq x \leq \frac{\pi}{2} \implies 0 \leq \sin x \leq 1 \implies 0 \leq \sin^n x \leq 1.
\]

Then \( f_n(x) = \sin^n x \to 0 \), for every value of \( x \) except \( x = \frac{\pi}{2} \), since \( f_n\left(\frac{\pi}{2}\right) = 1^n \to 1 \). We expect \( \lim \int f_n = \int \lim f_n = \int 0 = 0 \) (NOT generally true!)

Given \( \varepsilon > 0 \), we will show the area under \( \sin^n x \) is less than \( 2\varepsilon \) on this interval, for \( n \gg 1 \). Divide the interval at the point \( a := \frac{\pi}{2} - \varepsilon \).

\[
\text{left-hand area} < L = \text{area of flat rectangle} = a \sin^n a \\
\text{right-hand area} < R = \text{area of tall rectangle} = \varepsilon.
\]

Then \( \sin a < 1 \implies \sin^n a < \frac{\varepsilon}{a} \implies a \sin^n a < \varepsilon \) for \( n \gg 1 \).

\[
\int_0^{\pi/2} \sin^n x \, dx = \text{total area} < L + R < 2\varepsilon, \quad n \gg 1.
\]

Exercises: 3.2.3, 3.3.2, 3.4.1, 3.4.3, 3.7.1  Rec: #3.3.1, 3.3.3, 3.4.2,

3.4.4 Problems: 3-1,  Rec: 3-4  Due: Feb. 5

1. Show that \( e_n \to 0 \iff |e_n| \to 0 \).
Chapter 4

Error Term Analysis

4.1 The error term

Previously: \( a_n \to L \).

Now: how fast does \( a_n \to L \)?

Example 4.1.1.

\[
a_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} \to \log 2
\]

\[
b_n = \frac{2}{1 \cdot 3} + \frac{2}{3 \cdot 3^2} + \frac{2}{5 \cdot 3^5} + \cdots + \frac{2}{(2n-1)3^{2n-1}} \to \log 2
\]

But \( a_{100} = a_{99} - \frac{1}{100} \) is still changing the second decimal place. By contrast, \( b_3 \) is already accurate to three decimal places.

Example 4.1.2. For a converging sequence \( a_n \to L \), the error term is \( e_n = a_n - L \).

Theorem 4.1.1. Let \( a_n = L + e_n \). Then \( a_n \to L \iff e_n \to 0 \).

Note: \( e_n \to 0 \iff |e_n| \to 0 \), so it’s fine to define the error term as \( e_n = L - a_n \).
4.2 Geometric series error term.

Recall: for \( |r| < 1 \), we have

\[
1 + r + r^2 + \cdots + r^n + \cdots = \sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}.
\]

Theorem 4.2.1. Take \( a_n := 1 + r + r^2 + \cdots + r^n = \sum_{k=0}^{n} r^k \). Then \( \lim a_n = \frac{1}{1 - r} \).

Proof.

\[
a_n = 1 + r + r^2 + \cdots + r^n
\]

\[
r a_n = r + r^2 + r^3 + \cdots + r^{n+1}
\]

\[
a_n - r a_n = 1 - r + r^2 + r^3 - \cdots - r^n + r^{n+1}
\]

\[
a_n (1 - r) = 1 - r^{n+1}
\]

\[
a_n = \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r} - \frac{r^{n+1}}{1 - r}
\]

The error term is \( e_n = \frac{r^{n+1}}{1 - r} \to 0 \) by Thm. 3.4.1. So \( a_n \to L \) by Thm. 4.1.1. \( \square \)

Example 4.2.1. Show \( b_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} \to \log 2 \).

Solution. Put \( r = -u \) into the geometric series and then integrate:

\[
1 - u + u^2 - u^3 + \cdots + (-1)^{n-1} u^{n-1} = \frac{1}{1 + u} - (-1)^n \frac{u^n}{1 + u} \quad u \neq 1
\]

\[
1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1} u^{n-1} = \log 2 \pm \int_0^1 \frac{u^n}{1 + u} \, du.
\]

(Since \( \int_0^1 u^k \, du = \left[ \frac{u^{k+1}}{k+1} \right]_0^1 = \frac{1}{k+1} \).) The error term is

\[
e_n = \int_0^1 \frac{u^n}{1 + u} \, du \leq \int_0^1 u^n \, dx = \frac{1}{n + 1} \to 0. \square
\]

4.3 Newton’s method

Let \( \alpha \in \mathbb{R} \). Then Newton’s method can (often) find a sequence \( a_n \to \alpha \), where \( a_n \) is
defined in terms of \( a_{n-1} \).
How to find a zero of a function \( f(x) \)? Pick \( a_0 \) nearby. Then

\[
f'(a_n) = \frac{f(a_n)}{a_n - a_{n+1}} \implies a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}.
\]

**Example 4.3.1.** Find a sequence \( a_n \to \sqrt{2} \).

*Solution.* We need a zero of \( f(x) = x^2 - 2 \), so

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

Use Thm. 4.1.1: the error term is \( e_n = a_n - \sqrt{2} \), so

\[
e_{n+1} = a_{n+1} - \sqrt{2} = \frac{1}{2} \left( a_n + \frac{2}{a_n} \right) - \sqrt{2} = \frac{1}{2} \left( e_n + \sqrt{2} + \frac{2}{e_n + \sqrt{2}} \right) - \sqrt{2} = \frac{e_n^2}{2(\sqrt{2} + e_n)} \leq \frac{e_n^2}{2(\sqrt{2} - |e_n|)}.
\]

So if we pick \( a_0 \) within \( \varepsilon = \frac{1}{2} \) of \( \sqrt{2} \), then \(|e_0| < \frac{1}{2}\), and

\[
|e_n| < \frac{1}{2} \implies |e_{n+1}| \leq \frac{e_n^2}{2(\sqrt{2} - |e_n|)} \leq \frac{e_n^2}{2(1 - \frac{1}{2})} \leq e_n^2.
\]

So \(|e_0| < \frac{1}{2} \implies |e_n| \to 0\), by Thm. 3.4.1.

Why \( \varepsilon = \frac{1}{2} \)?

### 4.4 The Fibonacci numbers

The Fibonacci sequence \( 0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots \) is defined by

\[
F_0 = 0, \ F_1 = 1, \ F_{n+2} = F_{n+1} + F_n.
\]
Consider the sequence $a_n$ of rational numbers

$$1 \ 1 \ 2 \ 3 \ 5 \ \frac{1}{1} \ \frac{1}{2} \ \frac{1}{3} \ \frac{1}{5} \ \frac{8}{5} \ \ldots.$$ 

This sequence satisfies

$$a_{n+1} = \frac{1}{a_n + 1}.$$ 

Let us assume that $M = \lim a_n$. Then $M$ should satisfy

$$M = \frac{1}{M + 1} \implies M^2 + M - 1 = 0$$ 

$$\implies M = \frac{\sqrt{5} - 1}{2} \approx 0.618 \ldots$$

The error term is $e_n = a_n - M$, so

$$e_{n+1} = a_{n+1} - M = \frac{1}{a_n + 1} - M$$

by (*)

$$= \frac{1}{e_n + M + 1} - M$$

$$= \frac{1 - M - M^2 - Me_n}{e_n + M + 1}$$

$$= -\frac{M}{e_n + M + 1} e_n$$

by (**) 

$$= -\frac{\sqrt{5} - 1}{2e_n + \sqrt{5} + 1} e_n$$

by (***)

$$|e_{n+1}| \leq \frac{\sqrt{5} - 1}{1 + \sqrt{5} - 2|e_n|} e_n$$

$$\leq \frac{2e_n + \sqrt{5} + 1}{2}$$

$$\geq \sqrt{5} - 1 - 2|e_n|$$

$$= \frac{3 - 1}{1 + 2 - 2|e_n|} e_n$$

$$< \frac{2}{3 - 2|e_n|} e_n$$

$$< 2 \sqrt{5} < 3$$

Now suppose $|e_n| < \varepsilon = \frac{1}{2}$. Then

$$|e_{n+1}| < \frac{2}{3 - 2|e_n|} e_n < \frac{2}{3 - 1} e_n.$$
4.4 The Fibonacci numbers

Hmm ... this doesn’t help. Try $\varepsilon = \frac{1}{10}$. Then

$$|e_{n+1}| < \frac{2}{3-2|e_n|} e_n < \frac{2}{\frac{30}{10} - \frac{10}{10}} e_n = \frac{20}{28}e_n = \frac{5}{7}e_n.$$  

Then $|e_n| \leq \frac{5}{7}(\frac{5}{7}e_{n-1}) = \cdots = (\frac{5}{7})^{n+1}e_0 \to 0$.

Exercises: 4.1.1, 4.3.3  Recommended: #4.2.1, 4.4.1

Problems: 4-1  Recommended: #4-2

Due: Feb. 5

1. Show that $e_n \to 0 \iff |e_n| \to 0$. 

Chapter 5

The Limit Theorems

5.1 Limits of sums, products, quotients

Theorem 5.1.1. Let $a_n \to L$ and $b_n \to M$, where $L, M \in \mathbb{R}$.

(i) $\forall r, s \in \mathbb{R}, ra_n + sb_n \to rL + sM$. (linearity)

(ii) $a_nb_n \to LM$. (multiplicativity)

(iii) $b_n/a_n \to M/L$ if $L \neq 0$ and $a_n \neq 0, \forall n$.

(Just like the proof of $a_n + b_n$ is increasing.)

(i). Given $k \in \mathbb{N}$, we can find $N_1, N_2$ such that

$$n \geq N_1 \implies |a_n - L| < \frac{1}{k}, \quad \text{and} \quad n \geq N_2 \implies |b_n - M| < \frac{1}{k}.$$ 

Then let $N = \max(N_1, N_2)$. For $n \geq N$, we have

$$|(a_n + b_n) - (L + M)| = |(a_n - L) + (b_n - M)|$$

$$\leq |a_n - L| + |b_n - M| \quad \Delta \text{ ineq}$$

$$< \frac{1}{k} + \frac{1}{k} = \frac{2}{k}.$$ 

$\square$

\footnote{April 18, 2007}
(ii). (diff from book)

Given \( k \in \mathbb{N} \), we can again find \( N_1, N_2 \) such that

\[
n \geq N_1 \implies |a_n - L| < \frac{1}{k}, \quad \text{and} \quad n \geq N_2 \implies |b_n - M| < \frac{1}{k}.
\]

Then let \( N = \max(N_1, N_2) \) and

\[
|a_n b_n - LM| = |a_n b_n - a_n M + a_n M - LM| \\
\leq |a_n b_n - a_n M| + |a_n M - LM| \quad \Delta \text{ ineq} \\
\leq |a_n||b_n - M| + |M| |a_n - L| \quad \text{mult law for } | \cdot | \\
< |a_n| \frac{1}{k} + |M| \frac{1}{k} \quad \text{estimates for } n \geq N \\
< (J + |M|) \frac{1}{k} \quad \text{convergent sequences are bounded, Prob. 3-4}
\]

(iii). First, show \( \frac{1}{a_n} \to \frac{1}{L} \).

\[
\left| \frac{1}{a_n} - \frac{1}{L} \right| = \left| \frac{L - a_n}{a_n L} \right| = \frac{|L - a_n|}{|a_n||L|}
\]

So choose \( N \) such that

\[
n \geq N \implies |a_n - L| < \frac{1}{2} |L| \\
\implies |a_n| > \frac{1}{2} |L| \\
\implies \frac{|L - a_n|}{|a_n||L|} < \frac{2 |L - a_n|}{|L||L|^2/2} < \frac{2 |L|^2}{|L|^2/2} = n \gg 1. \quad \square
\]

Example 5.1.1.

\[
\lim_{n \to \infty} \frac{n^2 - 3n}{n^3 - 2n - 1} = \lim_{n \to \infty} \frac{1/n - 3/n^2}{1 - 2/n^2 - 1/n^3} \\
= \lim_{n \to \infty} \frac{1/n - 3/n^2}{1 - 2/n^2 - 1/n^3} \\
= \lim_{n \to \infty} \frac{1/n - 3/n^2}{1 - 2 \lim_{n \to \infty} 1/n^2 - \lim_{n \to \infty} 1/n^3} \quad (i) \\
= \frac{0 - 3 \cdot 0}{1 - 2 \cdot 0 - 0} = 0
\]
5.2 Comparison Theorems

**Question 3.** Prove that \( \lim a_n^2 = 0 \implies \lim a_n = 0 \).

**Solution.** The contrapositive is \( \lim a_n^2 \neq 0 \implies \lim a_n \neq 0 \), so suppose \( \lim a_n = L \neq 0 \). Then

\[
\lim a_n^2 = \lim a_n \cdot \lim a_n = L^2 \neq 0 \]

LIES! LIES! LIES! Cannot assume \( \lim a_n \) exists.

**Theorem 5.1.2.**

1. If \( a_n \to \infty \) and \( \{b_n\} \) bounded below, then \( a_n + b_n \to \infty \).

2. If \( a_n \to \infty \) and \( b_n \geq c > 0 \) for \( n \gg 1 \), then \( a_nb_n \to \infty \).

3. If \( a_n \to \infty \), then \( 1/a_n \to 0 \).

4. If \( a_n > 0, n \gg 1, \) and \( a_n \to 0 \), then \( 1/a_n \to \infty \).

**Example 5.1.2.** Examples of (2), above:

\[
\frac{n^2 - 3}{n+1} = (n - 3) \frac{n+3}{n+1} \to \infty
\]

\[
\frac{n^2 - 3}{n^2 + 1} = (n - 3) \frac{n+3}{n^2 + 1} \to 1
\]

5.2 Comparison Theorems

**Theorem 5.2.1.** Let \( x_n \to x \) and \( y_n \to y \). Then \( x_n \leq y_n \implies x \leq y \).

**Proof.** Use contradiction: suppose not. Then \( x_n \leq y_n \) but \( x > y \). Then \( |x - y| > 0 \), so we can use this for \( \varepsilon \). Find \( N \) such that

\[
|x_n - x| < \frac{|x - y|}{2} \quad \text{and} \quad |y_n - y| < \frac{|x - y|}{2} \quad \text{for} \quad n \geq N.
\]

For each \( x_n, y_n \) past the \( N^{th} \), \( y_n < x + \frac{|x-y|}{2} < x_n \). \( \square \)

**Corollary 5.2.2.** (Squeeze Thm) If \( x_n \to L, \ y_n \to L \) and \( x_n \leq z_n \leq y_n \) for some sequences \( \{x_n\}, \{y_n\}, \{z_n\} \), then \( \lim z_n = L \).

NOTE: for both, even if \( x_n < y_n \), can only conclude \( x \leq y \).

**Example 5.2.1.** \( \cos n/n^2 \).

DID THE PREV TWO THEOREMS WORK FOR \( x_n \to \infty \)?

No: \( |x - y| \) would not be defined.
Theorem 5.2.3. Let \( x_n \to \infty \) and \( x_n \leq y_n \). Then \( y_n \to \infty \).

Proof. Homework. \( \square \)

Example 5.2.2. \( a > 1 \implies a^n \to \infty \).

\[
1 < x \implies x = 1 + k, k > 0
\]

\[
x^n > 1 + nk
\]

\(
\text{binomial thm}
\)

\[
x^n \to \infty
\]

Using the Squeeze Thm with integrals.

Example 5.2.3. \( \lim_{n \to \infty} \frac{\log n!}{n \log n} = 1 \).

Solution. First, note that

\[
\log n! = \log(1 \cdot 2 \cdot 3 \ldots n) = \log 1 + \log 2 + \log 3 + \ldots \log n
\]

\[
\leq \log n + \log n + \log n + \ldots \log n = n \log n.
\]

and also that

\[
\log n! = \log 1 + \log 2 + \log 3 + \ldots \log n
\]

\[
\geq \int_1^n \log x \, dx = [x \log x - x]_1^n = n \log n - n + 1.
\]

Therefore, we can Squeeze

\[
n \log n - n + 1 \leq \log n! \leq n \log n
\]

\[
1 - \frac{1}{\log n} + \frac{1}{n \log n} \leq \frac{\log n!}{n \log n} \leq 1.
\]

Interpretation: \( \log n! \simeq \log n^n \).

5.3 Location theorems

Theorem 5.3.1 (Limit location theorem). If \( \{a_n\} \) is convergent, then

\( i \) \( a_n \leq M, n \gg 1 \implies \lim a_n \leq M, \) and
(ii) $a_n \geq M, n \gg 1 \implies \lim a_n \geq M$.

(i). By hyp, let $a_n \to L$. We need to show $L \leq M$. For any $k \in \mathbb{N}$, we have

$$a_n \approx_L L, n \gg 1 \implies L - \frac{1}{k} < a_n < L + \frac{1}{k} \implies L - \frac{1}{k} \leq M.$$  

Letting $k \to \infty$, we get $L \leq M$ by Comparison Thm. \qed

**Corollary 5.3.2.** If $\{a_n\}, \{b_n\}$ convergent, then $a_n \leq b_n, n \gg 1 \implies \lim a_n \leq \lim b_n$.

**Proof.** Apply prev to $a_n - b_n \leq 0$. \qed

**Theorem 5.3.3** (Sequence location theorem). If $\{a_n\}$ is convergent, then

(i) $\lim a_n < M \implies a_n < M, n \gg 1$, and

(ii) $\lim a_n > M \implies a_n > M, n \gg 1$.

(i). By hyp, let $a_n \to L$, so $L < M$.

Choose $\varepsilon = |L - M|/2$. Then for $n \gg 1$, each $a_n$ is within $|L - M|/2$ of $L$. \qed

IMPORTANT: be careful with $<$ and $\leq$ here.

**Example 5.3.1.** Thm. 5.3.1(ii) cannot be written $a_n > M, n \gg 1 \implies \lim a_n > M$.

Counterexample: $\frac{1}{n} > 0$ but $\frac{1}{n} \to 0 \not\to 0$.

Thm. 5.3.3(ii) cannot be written $\lim a_n \geq M \implies a_n \geq M, n \gg 1$.

Counterexample: $\frac{1}{n} > 0$ but $\frac{1}{n} \to 0 \not\to 0$.

### 5.4 Subsequences

**Definition 5.4.1.** If $\{a_n\}$ is a sequence, then a subsequence is a new sequence obtained from the original by deleting some (possibly infinitely many) terms, but keeping the order intact. The subsequence is denoted $\{a_{n_k}\}$.

**Example 5.4.1.** The sequence $\{(−1)^n(1 + \frac{1}{n})\}$ has the monotone decreasing subsequence $1 + \frac{1}{2^n}$ obtained by taking every second term. (SKETCH) Infinitely many deletions.

$a_n = \begin{cases} -2, & n \text{ odd} \\ \frac{3}{2}, & n = 1 \\ \frac{4}{3}, & n = 2 \\ \frac{5}{4}, & n = 3 \\ \vdots & \text{and so on} \end{cases}$
The Limit Theorems

$$n = 1, 2, 3, 4, 5, 6, \ldots$$

$$n_1 = 2, n_2 = 4, n_3 = 6, \ldots$$

$$a_{n_1} = a_2 = \frac{3}{2}, a_{n_2} = a_4 = \frac{5}{4}, a_{n_3} = a_6 = \frac{7}{6}.$$

Note: $$n_k \geq k$$ and $$n_1 < n_2 < \ldots$$

**Example 5.4.2.** The sequence \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\} has subsequence \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\} obtained by deleting the first term. A single deletion. \{\frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \ldots\} is not a subsequence. \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\} is not a subsequence.

**Theorem 5.4.2** (Subsequence Thm). \(a_n \to L\) iff \(a_{n_k} \to L\), for every subsequence \(\{a_{n_k}\}\).

**Proof.** (⇒) By hyp, we can find \(N\) such that \(a_n \approx L\) for \(n \geq N\). Suppose we take any subsequence \(\{a_{n_k}\}\). Since the indices are strictly increasing, \(n_1 < n_2 < \ldots\), we have \(n_k \geq N\) for \(k \geq 1\). Then \(a_{n_k} \approx L\) for \(k \geq 1\).

(⇐) Choose the subsequence obtained by deleting no terms from the original. □

**USE:** Subsequences are good for showing that a limit does not exist.

If you can find subsequences tending to two different limits, the original sequence doesn’t converge.

**Example 5.4.3.** Consider

$$1 + 1,$$

$$1 + \frac{1}{2}, 2 + \frac{1}{3},$$

$$1 + \frac{1}{4}, 2 + \frac{1}{5}, 3 + \frac{1}{6},$$

$$1 + \frac{1}{7}, 2 + \frac{1}{8}, 3 + \frac{1}{9}, 4 + \frac{1}{10}, \ldots$$

For any \(n \in \mathbb{N}\), this sequence has a subsequence tending to \(n\).

Obviously, the sequence doesn’t converge.

**QUESTION:** How to make a sequence with \(\mathbb{Q}\) as the set of limit points?
Example 5.4.4. \( \lim \sin n \) doesn’t exist. (note: no \( \pi \))

Solution. We will find two subsequences and show that they cannot have the same limit.

Note that \( \sin x > \sqrt{2}/2 \) for \( x \in (\pi/4, 3\pi/4) \), and hence for any \( m \in \mathbb{N} \) we have

\[
x \in (2k\pi + \frac{\pi}{4}, 2k\pi + \frac{3\pi}{4}) \implies \sin x > \sqrt{2}/2.
\]

(SKETCH). Each of these intervals has length \( \frac{\pi}{2} > 1 \), so there is an integer in each one. Let \( n_k \) be an integer in \( (2k\pi + \frac{\pi}{4}, 2k\pi + \frac{3\pi}{4}) \). Then \( \sin n_k > \sqrt{2}/2 \). Similarly,

\[
x \in ((2k+1)\pi + \frac{\pi}{4}, (2k+1)\pi + \frac{3\pi}{4}) \implies \sin x < -\sqrt{2}/2.
\]

Let \( m_k \) be an integer in \( (2k\pi + \frac{\pi}{4}, 2k\pi + \frac{3\pi}{4}) \). Then \( \sin m_k < -\sqrt{2}/2 \). Since \( |\sin n_k - \sin m_k| > \sqrt{2} \), they cannot tend to the same limit. \( \square \)

5.5 Two common mistakes

First mistake: trouble with inequalities.

Example 5.5.1. \( a_n \to 0, b_n \) bounded \( \implies a_nb_n \to 0 \).

Note: can’t use Product Thm, since don’t know that \( b_n \) has a limit. Attempt:

\[
L \leq b_n \leq M \quad 0 \leq |b_n| \leq K \\
a_nL \leq a_nb_n \leq a_nM \quad 0 \leq |a_n||b_n| \leq |a_n|K \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0
\]

What if \( a_n \) is sometimes negative? Solution: use \( |\cdot| \) instead.

Second mistake: repeating previous results. Whenever possible, cite a previous theorem to make life easier.

Exercises: #5.1.4, 5.2.4, 5.3.6, 5.4.2  Recommended: #5.1.5, 5.2.3, 5.3.2, 5.3.4, 5.3.5

Problems: 5-1, 5-2, 5-3  Recommended: #
1. If \( x_n \to \infty \) and \( x_n \leq y_n \), then \( y_n \to \infty \).

2. \( x_n \to x \) iff every neighbourhood of \( x \) of the form \((x - \varepsilon, x + \varepsilon), \varepsilon > 0\) contains all but finitely many points \( x_n \).

3. Can you have a sequence \( \{a_n\} \) which, for any given rational number \( p \in \mathbb{Q} \) has a subsequence \( a_{n_k} \to p \)? Construct one or prove it is impossible.
Chapter 6

The Completeness Property

6.1 Introduction. Nested intervals.

Definition 6.1.1. Let \( A_1, A_2, \ldots \) be a sequence of sets in \( \mathbb{R} \). This sequence is nested iff \( A_1 \supseteq A_2 \supseteq \ldots \).

Example 6.1.1. If this is a sequence of intervals \( A_n = [a_n, b_n] = \{x : 0 \leq x \leq 1\} \), then this means \( a_n \leq a_{n+1} \leq b_{n+1} \leq b_n, \forall n \).

Note: \( \bigcap_{n=1}^{\infty} A_n = \{x : x \in A_n, \forall n\} \).

Theorem 6.1.2 (Nested Intervals Thm). Suppose that \( A_n = [a_n, b_n] \) is a nested sequence of intervals with \( \lim(a_n - b_n) = 0 \). Then \( \bigcap_{n=1}^{\infty} A_n = \{L\} \). Also, \( a_n \to L \) and \( b_n \to L \).

Proof. There are 5 steps.

(i) \( a_n \leq b_n \) for any \( n, m \). Suppose \( a_n > b_m \). Then

\[
\begin{align*}
    n > m \implies b_n &\leq b_m < a_n \uparrow \\
    n < m \implies b_m &< a_n \leq a_m \uparrow
\end{align*}
\]

(ii) \( \{a_n\} \) is increasing and convergent, so let \( L = \lim a_n \).

\( \{a_n\} \) is increasing by nestedness, and bounded by (i), so converges by completeness.

---

\(^1\)April 18, 2007
(iii) \( \forall n, a_n \leq L \leq b_n \).

Part (ii) shows \( a_n \leq L \) (this is exactly a thm: \( a_n \not\to L \implies a_n \leq L \))
\[
 a_n \leq b_n \implies L \leq b_m \text{ by Limit Location Thm.}
\]

(iv) \( L = \lim b_n \) and \( L \) is the only number common to all intervals.

We add the two convergent sequences \( \{a_n\} \) and \( \{b_n - a_n\} \) to get
\[
\lim b_n = \lim(a_n + (b_n - a_n)) = \lim a_n + \lim(b_n - a_n) = L + 0 = L.
\]

This shows that the intersection can contain only \( L \), for suppose it also contained some number slightly larger than \( L \): call it \( L + \varepsilon, \varepsilon > 0 \). Then \( b_n \to L \) implies that \( b_n < L + \varepsilon \) for \( n \gg 1 \), so \( L + \varepsilon \) cannot be in the intersection. The same is true for any number slightly less than \( L \).

\[\square\]

**Example 6.1.2.** \( a_n = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n} \) converges.

**Proof.** Let \( a_0 = 0 \). Due to the alternating sign (which begins positive),
\[
|a_n - a_{n+1}| = \frac{1}{n} \to 0 \implies a_{2n} \leq a_{2n+2} \leq a_{2n+3} \leq a_{2n+1}.
\]

Then we have a decreasing nested sequence of intervals
\[
[a_0, a_1] \supseteq [a_2, a_3] \supseteq \ldots,
\]
so \( \bigcap_{n=0}^{\infty} [a_{2n}, a_{2n+1}] = \{L\} \) and \( a_n \to L \).

\[\square\]

### 6.2 Cluster points

Consider the sequence \( 1, 2, \frac{1}{3}, 4, \frac{1}{5}, 6, \frac{1}{7}, \ldots \). This contains two subsequences,
\[
1, 2, 4, 6, \cdots \to \infty
\]
\[
1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \cdots \to 0,
\]
so it clearly doesn’t converge. But it has a subsequence which does converge.
Definition 6.2.1. $K \in \mathbb{R}$ is a cluster point (or a limit point) of $\{a_n\}$ iff

$$\forall \varepsilon > 0, a_n \approx_{\varepsilon} K \text{ for infinitely many } n.$$  

Example 6.2.1. $(-1)^n$ has two cluster points: $\pm 1$. 

$(-1)^n(1 + \frac{1}{n})$ has the same two cluster points.

Neither of these sequences has limit.

Theorem 6.2.2 (Cluster Point Thm). $K$ is a cluster point of $\{a_n\}$ iff $K$ is the limit of a subsequence of $\{a_n\}$.

Proof. ($\Rightarrow$) Assuming that $K$ is a cluster point, we will construct a subsequence converging to it. Given $\varepsilon > 0$, we can find $a_n \approx_{\varepsilon} K$, so:

for $\varepsilon = 1$, choose $a_{n_1} \approx_{1} K$

for $\varepsilon = \frac{1}{2}$, choose $a_{n_2} \approx_{\frac{1}{2}} K$ and $n_2 > n_1$

...\n
for $\varepsilon = \frac{1}{j}$, choose $a_{n_j} \approx_{\frac{1}{j}} K$ and $n_j > n_{j-1}$

Then $a_{n_k} \to K$.

($\Leftarrow$) Given $\varepsilon > 0$, find $J$ such that $j \geq J \implies |a_{n_j} - K| < \varepsilon$.

Then $a_n \approx_{\varepsilon} K$ for $n \in \{n_J, n_{J+1}, \ldots\}$. $\square$

Together, the Subsequence Thm and Cluster Point Thm give an easy way to prove that a sequence doesn’t converge:

Any sequence with more than one cluster point does not converge.

6.3 Bolzano-Weierstrass Theorem

Theorem 6.3.1 (Bolzano-Weierstrass). A bounded sequence in $\mathbb{R}$ has a convergent subsequence.

Proof. Suppose $\{x_n\}$ is bounded, so that

$$a_0 \leq x_n \leq b_0, \forall n.$$
By Cluster Pt Thm, suffices to find a cluster point of the sequence. Apply the bisection method (aka divide-and-conquer): let \( c \) be the midpoint of \([a_0, b_0]\). Then at least one of \([a, c]\) or \([c, b_0]\) contains infinitely many points \( x_n \); call it \([a_1, b_1]\). (Choose the first one, if both have infinitely many.) Continuing, we get a nested sequence

\[
[a_0, b_0] \supseteq [a_1, b_1] \supseteq \ldots \supseteq [a_m, b_m] \supseteq \ldots
\]

Since

\[
|b_n - a_n| = \frac{|b_0 - a_0|}{2^n} \to 0,
\]

the Nested Intervals Thm gives

\[
\exists L \in \bigcap [a_n, b_n].
\]

Claim: \( L \) is a cluster point of \( \{x_n\} \). Given \( \varepsilon > 0 \), choose \( N \) large enough that \( |b_n - a_n| < \varepsilon \).

Then \([a_n, b_n] \subseteq (L - \varepsilon, L + \varepsilon)\) and contains infinitely many of the \( x_n \).

### 6.4 Cauchy sequences

**Definition 6.4.1.** A sequence \( \{a_n\} \) in \( \mathbb{R} \) is a **Cauchy sequence** iff

\[
\forall \varepsilon > 0, a_m \approx_\varepsilon a_n, \text{ for } m, n \gg 1.
\]

This means \( |a_m - a_n| \xrightarrow{m,n \to \infty} 0 \), or

\[
\forall \varepsilon > 0, \exists N_\varepsilon \text{ such that } m, n \geq N \implies |a_m - a_n| < \varepsilon.
\]

**Example 6.4.1** (Nonexample). \( 1, 2, 2 \frac{1}{2}, 3, 3 \frac{1}{3}, 3 \frac{2}{3}, 4, \ldots \)

**Definition 6.4.2** (Alternative). A sequence \( \{a_n\} \) in \( \mathbb{Q} \) is a **Cauchy sequence** iff

\[
\forall \varepsilon > 0, \exists (a, b) \text{ such that } |b - a| < \varepsilon \text{ and } \{x_N, x_{N+1}, x_{N+1}, \ldots \} \subseteq (a, b), \text{ for some } N.
\]

**Example 6.4.2.** The tail of the nonCauchy sequence has no upper bound, so any interval containing it is of the form \([x, \infty)\).

The definition of Cauchy sequence makes no claim about convergence (to \( L \in \mathbb{R} \), e.g.)!

How to know when a Cauchy sequence converges?
Theorem 6.4.3 (Cauchy Criterion). A sequence in \( \mathbb{R} \) has a limit \( \iff \) it is Cauchy.

Proof. (\( \Rightarrow \)) Homework (6.4.1)

(\( \Leftarrow \)) Let \( \{a_n\} \) be Cauchy. Then

(i) \( \{a_n\} \) is bounded.

Find \( N \) such that

\[
m, n \geq N \implies a_n \approx \varepsilon \ a_m
\]

\[
n \geq N \implies a_n \approx \varepsilon \ a_N.
\]

Then for \( n \geq N \),

\[
a_N - \varepsilon \leq a_n \leq a_N + \varepsilon.
\]

(ii) \( \{a_n\} \) has a convergent subsequence \( \{a_{n,i}\} \), by Bolzano-Weierstrass.

(iii) Define \( L := \lim_{n \to \infty} a_{n,i} \). Then

\[
|a_n - L| = |a_n - a_{n,i} + a_{n,i} - L|
\]

\[
\leq |a_n - a_{n,i}| + |a_{n,i} - L| \quad \Delta \text{ ineq.}
\]

Since Cauchy, can ensure \( |a_n - a_{n,i}| < \varepsilon \), for \( n, i > \).

Since subsequence converges, can ensure \( |a_{n,i} - L| < \varepsilon \), for \( i > \). So \( a_n \to L \) \( \square \)

Example 6.4.3. Prove the convergence of the sequence of Fibonacci fractions

\[
a_1 = 1, a_{n+1} = \frac{1}{a_n + 1}.
\]

Solution. Want to estimate \( |a_m - a_n| \). Wlog, let \( m > n \). Consider

\[
|a_n - a_{n+1}| = \left| \frac{1}{a_{n-1} + 1} - \frac{1}{a_n + 1} \right| = \frac{|a_{n-1} - a_n|}{(a_n + 1)(a_{n-1} + 1)} = c_n |a_{n-1} - a_n|,
\]

where \( c_n := 1/(a_n + 1)(a_{n-1} + 1) \). If we could show \( 0 \leq c_n \leq C \), then we’d have a bound

\[
|a_n - a_{n+1}| \leq C |a_{n-1} - a_n| \leq C^2 |a_{n-2} - a_{n-1}| \leq \cdots \leq C^{n-1} |a_1 - a_2|.
\]

Then we could sum up the terms \( |a_n - a_{n+1}| \) as a geometric series. Looking at the seq
one would guess \( a_n \geq \frac{1}{2} \), implying
\[
\frac{1}{c_n} = (a_n + 1)(a_{n-1} + 1) \geq \frac{3}{2} \cdot \frac{3}{2} > 2 \implies 0 \leq c_n \leq \frac{1}{2}.
\]

Claim: \( a_n \geq \frac{1}{2} \).

Proof of Claim. By induction: The basis step is \( a_1 = 1 \geq \frac{1}{2} \) and the induction step is
\[
\frac{1}{2} \leq a_n \leq 1 \implies \frac{3}{2} \leq a_n + 1 \leq 2, \quad + 1
\]
\[
\implies \frac{1}{2} \leq \frac{1}{a_n + 1} \leq \frac{2}{3}, \quad \le \text{ rules}
\]
\[
\implies \frac{1}{2} \leq a_{n+1} \leq \frac{2}{3} \leq 1, \quad \text{defn of } a_{n+1}. \quad \square
\]

Thus we have \( |a_n - a_{n+1}| < 1/2^n \), which implies
\[
|a_n - a_m| \leq |a_n - a_{n+1}| + |a_{n+1} - a_{n+2}| + \cdots + |a_{m-1} - a_m|
\]
\[
< \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{m-1}}
\]
\[
= \frac{1}{2^n}(1 + \frac{1}{2} + \frac{1}{4} + \cdots)
\]
\[
= \frac{1}{2^{n-1}} < \varepsilon, \quad \text{for } m > n \gg 1.
\]

This shows the sequence is Cauchy, so converges by Cauchy Criterion. \( \square \)

### 6.5 Completeness Property for sets

**Definition 6.5.1.** Let \( S \subseteq \mathbb{R} \). An upper bound for \( S \) is a number \( b \) such that \( x \in S \implies x \leq b \). \( S \) is said to be bounded above iff \( S \) has an upper bound. \( b \) is a sharp upper bound for \( S \) if no number less than \( b \) is an upper bound, i.e., if \( b \) is best possible.

**Definition 6.5.2.** \( m \in \mathbb{R} \) is the maximum of \( S \) iff \( m \) is an upper bound of \( S \) and \( m \in S \).

**Example 6.5.1.** A bounded set may not have a maximum: \((0, 1)\) is bounded above by 1, and by no number less than 1. But \( 1 \notin (0, 1) \).

NOTE: the max of a set must be contained in the set. What about when the set doesn’t contained the element that “ought” to be the max? There is a workaround:

**Definition 6.5.3.** The supremum of \( S \) is a sharp upper bound for \( S \).
6.5 Completeness Property for sets

(i) \( \beta \) is an upper bound for \( S \); \( x \in S \implies x \leq \beta \).

(ii) \( \beta \) is the least upper bound; \( b \) is an upper bound for \( S \implies \beta \leq b \).

A (nonempty) bounded set may not have a maximum, but it WILL have a supremum: this is a characterizing property of \( \mathbb{R} \).

**Theorem 6.5.4** (Completeness of \( \mathbb{R} \)). If \( S \subseteq \mathbb{R} \) is nonempty and bounded above, then \( \sup S \) exists in \( \mathbb{R} \).

**Proof.** Two steps: use bisection and nested intervals to locate a candidate for \( \sup S \), then prove it using Limit Location.

1. Since \( S \) is bounded above, let \( b_0 \) be an upper bound.

   Since \( S \) is nonempty, let \( a_0 \in S \). Then expect \( \sup S \) to be somewhere in \([a_0, b_0]\).

   Bisect \( S \); let \( c \) be the midpoint of \([a_0, b_0]\). Then expect \( \sup S \) to be in \([a_0, c]\) or \([c, b_0]\).

   \[
   [a_1, b_1] := \begin{cases} 
   [a_0, c], & \text{c is an upper bound for } S, \\
   [c, b_0], & \text{else.}
   \end{cases}
   \]

   Iterating the bisection procedure gives a sequence of nested intervals \([a_n, b_n]\) such that

   (i) \([a_n, b_n]\) contains a point of \( S \),

   (ii) \( b_n \) is an upper bound of \( S \), and

   (iii) \(|b_n - a_n| \to 0\).

   By the Nested Intervals Thm, \( \exists! \beta \in \bigcap_{n=1}^{\infty} [a_n, b_n] \), and

   \[\lim a_n = \lim b_n = \beta.\]

2. \( \beta = \sup S \).

   (i) \( \beta \) is an upper bound.

   \[
   x \in S \implies x \leq b_n, \forall n, \quad S \leq b_n
   \]

   \[\implies x \leq \lim b_n = \beta, \quad \text{Limit Loc Thm.}\]
(ii) $\beta$ is the least upper bound. Let $b$ be any other upper bound of $S$. Then

$$a_n \leq b, \forall n \implies \beta = \lim a_n \leq b.$$

Definition 6.5.5. Let $S \subseteq \mathbb{R}$. An lower bound for $S$ is a number $b$ such that $x \in S \implies x \geq b$. $S$ is said to be bounded below iff $S$ has an lower bound. $b$ is a sharp lower bound for $S$ if no number greater than $b$ is an lower bound, i.e., if $b$ is best possible.

Definition 6.5.6. $m \in \mathbb{R}$ is the minimum of $S$ iff $m$ is an lower bound of $S$ and $m \in S$.

Definition 6.5.7. The infimum of $S$ is a sharp lower bound for $S$.

(i) $\beta$ is an lower bound for $S$; $x \in S \implies x \leq \beta$.

(ii) $\beta$ is the greatest lower bound; $b$ is an lower bound for $S \implies \beta \geq b$.

Theorem 6.5.8. If $-S = \{-x : x \in S\}$, then $\inf S = -\sup(-S)$.

Exercises: #6.4.2, 6.5.3(aceg) Recommended: #6.2.2, 6.4.1

Problems: #6-2 Recommended: #6-3

Due: Feb.

1. Prove the equivalence of the two definitions of Cauchy sequence.

2. If $\{x_n\}$ is Cauchy in $\mathbb{R}$ and some subsequence $\{x_{n_k}\}$ converges to $x \in \mathbb{R}$, then prove the full sequence $\{x_n\}$ also converges to $x$. 
Chapter 7

Infinite Series

7.1 Series and sequences

1

Definition 7.1.1. An (infinite) series is a sum of a sequence \( \{a_k\} \):

\[
\sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + \ldots.
\]

To make it clear that the terms of the sequence \( \{a_k\} \) are added in order, define

\[
\sum_{k=0}^{\infty} a_k = \lim_{n \to \infty} \sum_{k=0}^{n} a_k = \lim s_n,
\]

where \( s_n := \sum_{k=0}^{n} a_k \). The series \( \sum a_k \) converges or diverges as the sequence \( s_n \) does.

Thus, a series is any sequence which can be written in a certain simple recursive form:

\[
s_n = s_{n-1} + f(n).
\]

Example 7.1.1. Geometric series: \( 1 + r + r^2 + \cdots = \sum_{k=0}^{\infty} r^k \) is the limit of

\[
s_n = 1 + r + r^2 + \cdots + r^n = s_{n-1} + r^n.
\]
Example 7.1.2. Harmonic series: \(1 + \frac{1}{2} + \frac{1}{3} + \cdots = \sum_{k=1}^{\infty} \frac{1}{k}\) is the limit of

\[s_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = s_{n-1} + \frac{1}{n}.
\]

Definition 7.1.2. A telescoping series is one that can be written in the form

\[\sum_{k=0}^{\infty} (a_{k+1} - a_k).
\]

A telescoping series has partial sums

\[s_n = (a_1 - a_0) + (a_2 - a_1) + (a_3 - a_2) + \cdots + (a_{n-1} - a_{n-2}) + (a_n - a_{n-1}) = a_n - a_0.
\]

So the sum can be found as

\[\sum_{k=0}^{\infty} (a_{k+1} - a_k) = \lim_{s_n} s_n = \lim_{a_n} a_n - a_0.
\]

Given any sequence, this provides a way to write a series which has as partial sums, the terms of the original sequence:

1. Start with any sequence \(\{x_n\}\).

2. Define \(a_0 = x_0\) and for \(n \geq 1\), let \(a_n := x_n - x_{n-1}\).

3. Then \(x_n\) is the \(n^{th}\) partial sum

\[\sum_{k=0}^{n} a_k = x_0 + (x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1}) = x_n.
\]

Example 7.1.3 (Euler’s gamma). Let \(s_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log(n+1), n \geq 1\).

To make this into a series, let

\[a_n = s_n - s_{n-1} = \frac{1}{n} - \log(n+1) + \log n = \frac{1}{n} - \log \frac{n+1}{n}, n \geq 1
\]

\[\Rightarrow \sum_{n=1}^{\infty} \left(\frac{1}{n} - \log \frac{n+1}{n}\right) = \gamma.
\]

Note that

\[\int_{n}^{n+1} \frac{dx}{x} = [\log x]_{n+1}^{n+1} = \log(n+1) - \log n,
\]
so we have
\[ \gamma = \lim_{N \to \infty} \left( \sum_{n=1}^{\infty} \frac{1}{n} - \int_{1}^{\infty} \frac{1}{x} \, dx \right) = \sum_{n=1}^{\infty} \frac{1}{n} - \int_{1}^{\infty} \frac{1}{x} \, dx. \]

7.2 Elementary convergence tests

Almost always impossible to find the actual sum of a series (exceptions: geometric, telescoping, Fourier), so more important to know if converges.

**Theorem 7.2.1.** \( \sum a_n \) converges \( \implies \) \( a_n \to 0 \).

**Proof.** Let \( s_n \) be a partial sum of the series and \( S = \lim s_n \). Then

\[
\begin{align*}
s_n &= s_{n-1} + a_n \\
a_n &= s_n - s_{n-1} \\
\lim a_n &= \lim(s_n - s_{n-1}) = \lim s_n - \lim s_{n-1} = S - S = 0.
\end{align*}
\]

**NOTE:**

1. The converse is FALSE:

\[
\frac{1}{n} \to 0 \text{ but } \sum \frac{1}{n} \text{ diverges.}
\]

2. Most often used as contrapositive: \( a_n \not\to 0 \implies \sum a_n \text{ diverges.} \)

\[
\sum_{n=1}^{\infty} (-1)^n, \quad \sum_{n=1}^{\infty} \left( \frac{n}{n+1} \right)^n
\]

**Theorem 7.2.2** (Tail-convergence). \( \sum_{n=N_0}^{\infty} a_n \) converges \( \iff \) \( \sum_{n=0}^{\infty} a_n \) converges \( \iff \sum_{n=N}^{\infty} a_n, \forall N. \)

**Proof.** Idea: \( \lim s_n = \lim s_{n+N} \).

**Theorem 7.2.3** (Cauchy Criterion for series). \( \sum a_n \) converges iff

\[
\forall \varepsilon > 0, \quad m, n \gg 1 \implies \left| \sum_{k=n}^{m} a_k \right| < \varepsilon.
\]
Theorem 7.2.4 (Linearity). \( \forall p, q \in \mathbb{R}, \) if \( \sum a_n \) and \( \sum b_n \) converge, then \( \sum (pa_n + qb_n) \) converges and
\[
\sum (pa_n + qb_n) = p \sum a_n + q \sum b_n.
\]

Proof. \( \lim (ps_n + qt_n) = p \lim s_n + q \lim t_n. \)

So we have \( \sum a_n + \sum b_n = \sum (a_n + b_n) \) and \( \sum ca_n = c \sum a_n. \) However,
\[
\sum a_n b_n \neq \sum a_n \sum b_n, \text{ because } 1 + a_1 b_2 + a_2 b_2 + \ldots a_n b_n \neq (1 + a_1 + \cdots + a_n)(1 + b_2 + \cdots + b_n).
\]

(The more cross terms on right.)

Theorem 7.2.5 (Increasing & bounded). If \( 0 \leq a_n, \forall n, \) then \( \sum a_n \) converges iff the partial sums are bounded.

Proof. \( (\Rightarrow) \lim s_n \) exists \( \implies \{s_n\} \) bounded.

\( (\Leftarrow) s_n = s_{n-1} + a_n \geq s_{n-1}, \) so monotone. Then \( \{s_n\} \) bounded implies \( \{s_n\} \) convergent by completeness. \( \square \)

Theorem 7.2.6 ((Direct) Comparison Thm). If \( 0 \leq a_n \leq b_n, \forall n, \) then
\[
\sum b_n \text{ converges } \implies \sum a_n \text{ converges.}
\]

In this case, \( \sum a_n \leq \sum b_n. \)

The converse of this statement is also quite helpful:
\[
0 \leq a_n \leq b_n, \sum a_n \text{ diverges } \implies \sum b_n \text{ diverges.}
\]

Proof. Define the partial sums
\[
s_n = \sum_{k=1}^{n} a_k \text{ and } t_n = \sum_{k=1}^{n} b_k.
\]

By hyp, \( T := \lim t_n \) exists, and it is an upper bound for \( \{t_n\} \) by the Incr Seq Thm, so
\[
0 \leq a_n \leq b_n \implies s_n \leq t_n \leq T,
\]
7.3 Series with negative terms

which implies \( \sum a_n \) converges, by the Incr & bounded Thm. Finally, \( S = \sum a_n \leq T \) by Limit Location Thm. \( \square \)

IMPORTANT:

completeness of \( \mathbb{R} \) \( \Rightarrow \) increasing & bounded thm \( \Rightarrow \) Comparison Thms.

7.3 Series with negative terms

The above Comparison Thm is only for positive-term series.

**Definition 7.3.1.** \( \sum a_n \) is *absolutely convergent* iff \( \sum |a_n| \) converges.
\( \sum a_n \) is *conditionally convergent* iff \( \sum |a_n| \) diverges but \( \sum a_n \) converges.

**Example 7.3.1.**

1. For a positive-term series, convergence \( \equiv \) absolute convergence.

2. \( \sum \frac{(-1)^n}{n^2} \) and \( \sum \frac{(-1)^n}{n!} \) are absolutely convergent, since \( \sum \frac{1}{n^2} \) and \( \sum \frac{1}{n!} \) are convergent.

3. \( \sum \frac{(-1)^n}{n} \) is conditionally convergent, since the harmonic series diverges.

**Example 7.3.2.** If \( f : \mathbb{R} \to \mathbb{R} \) is any function, then define

\[
    f^+(x) := \max\{f(x), 0\}, \quad \text{and} \quad f^-(x) := -\min\{f(x), 0\}.
\]

Then \( f^+(x) \geq 0 \) and \( f^-(x) \geq 0 \), \( \forall x \). Also:

\[
    f(x) = f^+(x) - f^-(x) \quad \text{and} \quad |f(x)| = f^+(x) + f^-(x).
\]

A sequence is just a function \( a : \mathbb{N} \to \mathbb{R} \) where we usually write \( a_n \) for \( a(n) \), but it can be similarly decomposed

\[
    a_n^+ := \max\{a_n, 0\}, \quad a_n^- := -\min\{a_n, 0\},
\]

so that both \( \{a_n^+\} \) and \( \{a_n^-\} \) are positive sequences and

\[
    a_n = a_n^+ - a_n^- \quad \text{and} \quad |a_n| = a_n^+ + a_n^-.
\]
(SKETCH EXAMPLE:) \(\sum (-1)^n n\).

**Theorem 7.3.2** (Absolute convergence thm). \(\sum |a_n|\) converges \(\implies \sum a_n\) converges.

*Proof 1.* Split the series into positive and negative components as above: \(\{a_n^+\}\) and \(\{a_n^-\}\).

Then \(0 \leq a_n^+ \leq |a_n| \implies \sum a_n^+ \leq \sum |a_n|\), by Comp Thm, and same for \(a_n^-\). Linearity Thm lets us add two convergent series. \(\square\)

*Proof 2.* Apply the Cauchy criterion to \(\sum_{k=n}^{m} |a_k| \leq \sum_{k=n}^{m} |a_k|\).

Completeness is what implies the Comparison Thm, so it also implies this Thm. Strangely, this property is *equivalent* to completeness! First, need a couple of definitions.

**Definition 7.3.3.** A *vector space* is a set \(X\) where any two element of \(X\) can be added, or multiplied by a number in \(\mathbb{R}\). (There are more details, but this is all we’ll need.)

**Definition 7.3.4.** A *norm* on a vector space is a function from \(X\) to \(\mathbb{R}\) that satisfies

(i) \(\|x\| \geq 0, \|x\| = 0 \iff x = 0\).

(ii) \(\|ax\| = |a| \cdot \|x\|, \forall a \in \mathbb{R}\).

(iii) \(\|x - z\| \leq \|x - y\| + \|y - z\|, \forall x, y, z \in X\).

NOTE: the scalars in these two definitions can be replaced by \(\mathbb{Q}\), \(\mathbb{C}\), or any other field.

**Example 7.3.3.**
\(\mathbb{R}^n\) with \(\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}\).

\(M_n(\mathbb{R})\) with \(\|A\| = \sum_{i=1}^{n} |a_{ij}|\).

the continuous functions on an interval \(C(I)\) with \(\|f\| = \sup_{x \in I} |f(x)|\).

the continuous functions on an interval \(C(I)\) with \(\|f\|_1 = \int_I |f(x)| \, dx\).

the continuous functions on an interval \(C(I)\) with \(\|f\|_2 = (\int_I |f(x)|^2 \, dx)^{1/2}\).

Now we can show that in any normed vector space, completeness (defined as convergence of Cauchy sequences) is equivalent to summability of absolutely convergent series.
Theorem 7.3.5. Suppose we have a vector space \((X, \| \cdot \|)\). Then \(X\) is complete iff every absolutely convergent series in \(X\) converges.

Proof. \((\Rightarrow)\) Suppose that every Cauchy sequence in \(X\) converges and that \(\sum_{k=1}^{\infty} \| x_k \|\) converges. Must show that \(\sum_{k=1}^{\infty} x_k\) converges.

Show that the sequence of partial sums is Cauchy, hence converges.

Let \(s_n = \sum_{k=1}^{n} x_k\). Then for \(n > m\), we have

\[
\left\| s_n - s_m \right\| = \left\| \sum_{k=1}^{n} x_k - \sum_{k=1}^{m} x_k \right\| = \left\| \sum_{k=m+1}^{n} x_k \right\|
\leq \sum_{k=m+1}^{n} \| x_k \| \text{ by } \Delta \text{ ineq}
< \varepsilon, \text{ for } m > 1,
\]

since \(\sum_{k=1}^{\infty} \| x_k \|\) converges, \(\sum_{k=N}^{\infty} \| x_k \| \xrightarrow{N \to \infty} 0\) by Tail-Conv Thm.

\((\Leftarrow)\) Suppose that \(\sum_{k=1}^{\infty} \| x_k \|\) converges \(\implies\) \(\sum_{k=1}^{\infty} x_k\) converges. Use this to show that any Cauchy sequence converges.

Let \(\{x_n\}\) be Cauchy. Then

\[
\forall \varepsilon > 0, \exists N, \text{ such that } m, n \geq N \implies \| x_n - x_m \| < \varepsilon, \text{ or}
\forall j \in \mathbb{N}, \exists n_j, \text{ such that } m, n \geq n_j \implies \| x_n - x_m \| < \frac{1}{2^j}.
\]

So we can find a subsequence \(\{x_{n_j}\}\), choosing \(n_1 < n_2 < \ldots\). Define

\[
y_1 = x_{n_1},
y_j = x_{n_j} - x_{n_{j-1}}, j > 1.
\]

Then \(\sum_{j=1}^{k} y_j = x_{n_k}\) (by telescoping), and

\[
\sum_{j=1}^{\infty} \| y_j \| \leq \| y_1 \| + \sum_{j=1}^{\infty} \frac{1}{2^j} = \| y_1 \| + 1 < \infty.
\]

So \(\lim x_{n_k} = \sum y_j\) exists, i.e., \(x_{n_j} \to x \in X\). Since \(\{x_n\}\) is Cauchy, it must also converge to the same limit (REC HW from §6); for \(m, n \geq N\), have

\[
\| x_n - x \| = \| x_n - x_{n_k} + x_{n_k} - x \| \leq \| x_n - x_{n_k} \| + \| x_{n_k} - x \| < 2\varepsilon, N \gg 1.
\]
Recap: \((\Rightarrow)\) comes by writing a series as a sequence, \((\Leftarrow)\) comes by writing a sequence as a series, using the telescoping trick.

### 7.4 Ratio and Root tests

**Theorem 7.4.1** (Ratio test). Suppose \(a_n \neq 0, n \gg 1\), and \(\lim \left| \frac{a_{n+1}}{a_n} \right| = L\). Then

\[
L < 1 \implies \sum |a_n| \text{ converges}, L > 1 \implies \sum |a_n| \text{ diverges}.
\]

*If \(L = 1\) or \(L\) doesn’t exist, the test tells nothing.*

**Proof.** Compare \(\sum |a_n|\) to geometric series.

**case (1) \(0 \leq L < 1\).** Then pick \(M\) such that \(L < M < 1\).

(SKETCH WHY \(M\) is necessary for the following ineq).

By Seq Loc Thm,

\[
\left| \frac{a_{n+1}}{a_n} \right| \to L \implies \left| \frac{a_{n+1}}{a_n} \right| < M,
\]

for all \(n\) larger than some \(N\). Then we get a recursion relation

\[
\left| \frac{a_{n+1}}{a_n} \right| < M \implies |a_{n+1}| < |a_n|M.
\]

Applying this to \(a_{N+k}\) and iterating,

\[
|a_{N+k}| < |a_{N+k-1}|M < |a_{N+k-2}|M^2 < \cdots < |a_N|M^k
\]

\[
\sum_{k=0}^{\infty} |a_{N+k}| < \sum_{k=0}^{\infty} |a_N|M^k = |a_N| \sum_{k=0}^{\infty} M^k
\]

RHS converges by geom, so Comp Thm gives convergence of LHS.

Then Tail-Conv Thm gives convergence of \(\sum |a_n|\)

**case (2) \(L > 1\).** Exercise 7.4.2.

\(\square\)

**Theorem 7.4.2** (Root test). Suppose \(\lim |a_n|^{1/n} = L\). Then

\[
L < 1 \implies \sum |a_n| \text{ converges},
\]

\[
L > 1 \implies \sum a_n \text{ diverges}.
\]
7.5 Integral, p-series, and asymptotic comparison

If \( L = 1 \) or doesn’t exist, then test tells nothing.

Proof. By cases, like for Ratio Test.

The Ratio Test is easier and more common, but the Root Test is applicable more broadly (Prob 7-4).

7.5 Integral, p-series, and asymptotic comparison

Theorem 7.5.1 (Integral Test). Suppose \( f(x) \geq 0 \) and decreasing for \( x \geq N \in \mathbb{N} \). Then \( \sum f(n) \) converges iff \( \int_{N}^{\infty} f(x) \, dx \) is finite.

Proof. Define the area \( A_n := \int_{n}^{n+1} f(x) \, dx \). The area of the rectangle under the graph is \( f(n+1) \times |n - (n+1)| = f(n+1) \). Thus,

\[
0 \leq f(n+1) \leq A_n, \text{ for } n \geq N.
\]

Shifting one unit to the right, the region under the graph is contained in the rectangle, so

\[
0 \leq A_n \leq f(n), \text{ for } n \geq N.
\]

This gives

\[
\sum_{k=N+1}^{N+n} f(k) \leq \int_{N+1}^{N+n} f(x) \, dx \leq \sum_{k=N}^{N+n-1} f(k) \leq \int_{N}^{N+n-1} f(x) \, dx
\]

and so the two sequences converge or diverge together.

Theorem 7.5.2 (p-series). \( \sum \frac{1}{n^p} \) converges iff \( p > 1 \).

Proof. For \( p \geq 0, \frac{1}{n^p} \) is increasing, so apply the integral test to

\[
\int_{1}^{\infty} \frac{1}{x^p} \, dx = \begin{cases} 
\lim_{r \to \infty} \frac{r^{1-p} - 1}{1-p}, & p \neq 1, \\
\lim_{r \to \infty} \log r, & p = 1.
\end{cases}
\]

For \( p = 1, \log r \to \infty \) and both diverge.

For \( p > 1, r^{1-p} \to 0 \), so both converge.

For \( 0 \leq p < 1, r^{1-p} \to \infty \), so both diverge.

Finally, consider \( p < 0 \) and put \( q = -p > 0 \). Then \( \sum n^q \) diverges by \( n^{th} \) term test.
Theorem 7.5.3 (Asymptotic comparison test). If \( \lim \frac{|a_n|}{|b_n|} = 1 \), then

\[
\sum |a_n| \text{ converges} \iff \sum |a_n| \text{ converges}.
\]

Proof. Homework 7.4.4.

\[\square\]

### 7.6 Alternating series test

Theorem 7.6.1. If \( \{a_n\} \) is positive and strictly decreasing, then \( \sum (-1)^n a_n \) converges.

Proof. Since the signs alternate, we get

\[s_{2k} = s_{2k-1} + a_{2k}, \quad s_{2k+1} = s_{2k} - a_{2k+1},\]

so \( s_{2k+1} = s_{2k} - a_{2k+1} < s_{2k} \).

Also, \( s_{2k} = s_{2k-1} + a_{2k} \implies s_{2k-1} < s_{2k} \).

Also, \( s_{2k+1} = s_{2k} - a_{2k+1} = s_{2k-1} + (a_{2k} - a_{2k+1}) \implies s_{2k-1} < s_{2k+1} \).

Thus, we have a nested sequence \([s_{2k-1}, s_{2k}]\) with \( |s_{2k-1} - s_{2k}| = |a_{2k}| \to 0 \). Thus, \( \{s_n\} \) is a Cauchy sequence.

\[\square\]

Corollary 7.6.2. For an alternating series, \( e_n = |s_n - S| < a_{n+1} \), where \( S = \sum (-1)^n a_n \).

Proof. HW: since \( S \in [s_{2k-1}, s_{2k}] \), apply inequalities from prev proof.

\[\square\]

Theorem 7.6.3 (Cauchy’s subsequence test). Suppose \( a_1 \geq a_2 \geq \cdots \geq 0 \). Then

\[
\sum_{n=1}^{\infty} a_n \text{ converges} \iff \sum_{k=0}^{\infty} 2^k a_{2^k} \text{ converges}.
\]

\[a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, \ldots\]

\[a_1, 2a_2, 4a_4, 8a_8, \ldots\]

Proof. Since positive-term, enough to show boundedness of partial sums. Define

\[s_n := a_1 + a_2 + \cdots + a_n,\]

\[t_k := a_1 + 2a_2 + \cdots + 2^k a_{2^k}.\]
For $n < 2^k$,
\[ s_n \leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \cdots + (a_{2^k} + \cdots + a_{2^k+1-1}) \]
\[ \leq a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k} = t_k. \]
This shows $s_n \leq t_k$. Meanwhile, for $n > 2^k$,
\[ s_n \geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \cdots + (a_{2^k-1+1} + \cdots + a_{2^k}) \]
\[ \geq a_1 + a_2 + 2a_4 + 4a_8 + \cdots + 2^{k-1} a_{2^k} \]
\[ \geq \frac{1}{2} a_1 + a_2 + 2a_4 + 4a_8 + \cdots + 2^{k-1} a_{2^k} = \frac{1}{2} t_k. \]
This shows $2s_n \geq t_k$, so that
\[ s_n \leq t_k \leq 2s_n. \]
Thus \( \{s_n\} \) and \( \{t_k\} \) are both bounded or both unbounded.

### 7.7 Rearrangements

**Theorem 7.7.1.** If $\sum a_n$ is absolutely convergent, then any rearrangement of it is also convergent, and has the same sum.

**Proof.** First, suppose $a_n \geq 0$. Let $\sum a_{n'}$ denote a rearrangement of the original series; the partial sums are $s_n$ and $s_n'$. Fix $\varepsilon > 0$ and show $|s_n - s_n'| < \varepsilon$.

The hypothesis means that (by Tail-conv) for some $N$,
\[ \sum_{k=N}^{\infty} |a_k| < \varepsilon. \]
By going far enough in the rearranged series, can ensure that
\[ \{a_1, a_2, \ldots, a_N\} \subseteq \{a_1', a_2', \ldots, a_p'\}, \]
so that
\[ \sum_{k=p+1}^{\infty} |a_k'| \leq \sum_{k=N}^{\infty} |a_k| < \varepsilon. \]
Theorem 7.7.2. If $\sum a_n$ is conditionally convergent, then for any $x \in \mathbb{R}$, there is a rearrangement which sums to $x$ (or even diverges to $\pm \infty$).

Nonproof. To be conditionally convergent, must have infinitely many positive and negative terms, so separate into $\sum a^+_n$ and $\sum a^-_n$. For $x \geq 0$, form rearrangement as follows:

1. Add positive terms until sum exceeds $x$. Stop as soon as $\sum_{j=1}^J a^+_j \geq x$.
2. Add negative terms until $x$ exceeds sum. Stop as soon as $\sum_{j=1}^J a^+_j - \sum_{k=1}^K a^-_k \leq x$.
3. Repeat.

Since $a^+_n \to \infty$ and $a^-_n \to \infty$, neither step (i) nor step (ii) can ever go on for infinitely many steps. Since $a_n$ is conditionally convergent, $a_n \to 0$ and the procedure generates a nested sequence of intervals.

\[\Box\]

7.8 Multiplication of Series

Theorem 7.8.1 (Cauchy Product). Suppose that $\sum a_n = A$ and $\sum b_n = B$, and at least one of them converges absolutely. Define $c_n = \sum_{k=0}^n a_kb_{n-k}$. Then $\sum c_n = AB$.

Proof. Wlog, suppose it is $\sum a_n$ that converges absolutely, and define

\[A_n := \sum_{k=0}^n a_n, \quad B_n := \sum_{k=0}^n b_n, \quad C_n := \sum_{k=0}^n c_n, \quad \beta_n = B_n - B.\]

Then we use an error term estimate:

\[C_n = a_0b_0 + (a_0b_1 + a_1b_0) + \cdots + (a_0b_n + \cdots + a_nb_0)\]
\[= a_0B_n + a_1B_{n-1} + \cdots + a_nB_0\]
\[= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \cdots + a_n(B + \beta_0)\]
\[= A_nB + a_0\beta_n + \cdots + a_n\beta_0.\]

So let $e_n := a_0\beta_n + \cdots + a_n\beta_0$ and it will suffice to show $e_n \to 0$.

To use the absolute convergence of $\sum a_n$, let $\alpha := \sum |a_n|$.

Fix $\varepsilon > 0$. Since $\sum b_n$ converges, choose $N$ such that $n \geq N \implies |\beta_n| < \varepsilon$, so that

\[|e_n| \leq |a_0\beta_n + \cdots a_{n-N-1}\beta_{N+1}| + |\beta_Na_{n-N} + \cdots + a_n\beta_0|\]
\[ \leq \varepsilon \alpha + |\beta_N a_{n-N} + \cdots + a_n \beta_0| \]

Now since \( a_n \to 0 \), for \( N \) fixed and \( n \gg 1 \), we can make \( |\beta_N a_{n-N} + \cdots + a_n \beta_0| < \varepsilon \). Then \( |e_n| < (\alpha + 1)\varepsilon \).

**Theorem 7.8.2** (Dirichlet’s test). Suppose that the partial sums \( A_n = \sum_{i=1}^{n} a_i \) form a bounded sequence, and suppose there is a sequence \( \{b_i\} \) with \( b_i \geq b_{i+1} \) and \( b_i \to 0 \). Then show \( \sum a_i b_i \) converges.

**Exercises:** #7.2.2, 7.4.2, 7.4.4  **Recommended:** #7.2.3, 7.3.1, 7.4.1,

**7.4.3**

**Problems:** #7-4  **Recommended:** #7-5, 7-7

**Due:** Mar.

1. If \( \sum a_n \) converges and \( \{b_n\} \) is monotonic and bounded, then \( \sum a_n b_n \) converges.

2. Give an example to show that you can have \( \sum x_n \) diverge and \( \sum y_n \) diverge, but \( \sum x_n y_n \) converges.

3.  
   (a) Show that if \( \sum a_n \) converges absolutely, then \( \sum a_n^2 \) does, too. Is this true without the hypothesis of absolute convergence?
   
   (b) If \( \sum a_n \) converges and \( a_n \geq 0 \), what can be said about \( \sum \sqrt{a_n} \)?

4.  
   (a) (Dirichlet’s Test) Suppose that the partial sums \( A_n = \sum_{i=1}^{n} a_i \) form a bounded sequence, and suppose there is a sequence \( \{b_i\} \) with \( b_i \geq b_{i+1} \) and \( b_i \to 0 \). Then show \( \sum a_i b_i \) converges. (Hint: \( \left| \sum_{i=p}^{q} a_i b_i \right| = \left| \sum_{i=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p \right| \))

   (b) Use Dirichlet’s test to prove the alternating series test.

5. For an alternating series, \( e_n = |s_n - S| < a_{n+1} \), where \( S = \sum (-1)^n a_n \).
Chapter 8

Power Series

8.1 Intro, radius of convergence

Definition 8.1.1. A power series is a series of the form \( \sum a_n x^n \), where \( x \) is a variable. The \( n^{th} \) term of a power series is \( a_n x^n \) (rather than just \( a_n \)).

\( \sum a_n x^n \) is a family of series, one for each value of \( x \). We are interested in the subfamily corresponding to

\[ A = \{ x \in \mathbb{R} : \sum |a_n x^n| \text{ converges} \} . \]

Then we can define a function

\[ f : A \rightarrow \mathbb{R}, \quad f(x) = \sum a_n x^n . \]

Example 8.1.1. Where does \( \sum_{n=1}^{\infty} \frac{x^{2n}}{2n} \) converge?

Solution.

\[ \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{x^{2n+1}}{x^{2n}} \right| \cdot \frac{2^{n+1}}{2^n} = \frac{n}{2(n+1)} |x|^2 \rightarrow \frac{|x|^2}{2}, \forall x \in \mathbb{R} . \]

Ratio test implies convergence if this final quantity is \(< 1\), so find where this is true:

\[ \frac{|x|^2}{2} < 1 \iff |x| < \sqrt{2} . \]

1 April 18, 2007
So converges for \(-\sqrt{2} < x < \sqrt{2}\) and diverges for \(|x| > \sqrt{2}\).

(What happens for \(x = \pm\sqrt{2}\)? Evaluate:)

\[
\sum_{n=1}^{\infty} \frac{(\pm\sqrt{2})^{2n}}{2^n n} = \sum_{n=1}^{\infty} \frac{2^n}{2^n n} = \sum_{n=1}^{\infty} \frac{1}{n}.
\]

This behavior is typical:

**Theorem 8.1.2.** For any power series \(\sum a_n x^n\), \(\exists R \geq 0\) such that

\[
\sum a_n x^n \text{ converges absolutely for } |x| < R,
\]

\[
\sum a_n x^n \text{ diverges for } |x| > R.
\]

*R is the radius of convergence of the power series.* (Note: may have \(R = 0\).) By convention, \(R = \infty\) iff the series converges \(\forall x \in \mathbb{R}\).

**Proof.** (I) First step: show convergence for \(x = c\) implies convergence for \(|x| < |c|\).

This show the domain of convergence is an interval \((-c, c)\).

For \(c = 0\), it is trivial, so assume \(c > 0\).

If \(\sum a_n c^n\) converges, then \(a_n c^n \to 0\), so \(|a_n c^n| \leq M\) for some fixed \(M > 0\). Then for \(|x| < c\),

\[
|a_n x^n| = |a_n c^n| \left| \frac{x}{c} \right|^n \leq M \left| \frac{x}{c} \right|^n.
\]

Since \(\sum M \left| \frac{x}{c} \right|^n\) is a geometric series with \(\left| \frac{x}{c} \right| < 1\), it converges and by Comparison, \(\sum a_n x^n\) converges absolutely.

(II) Second step: define \(R\) as the sup of the \(c\)'s in the First step, and show that it separates convergent from divergent.

\[
A := \{ x \in \mathbb{R} : \sum |a_n x^n| \text{ converges} \} = (-c, c).
\]

If \(A = \mathbb{R}\) then let \(R = \infty\). Otherwise, \(\exists b \notin A\), hence also \(-b \notin A\). Then \(|b|\) is an upper bound for \(A\), because \(c \in A \implies (-c, c) \subseteq A\), by first part.

So \(\sup A\) exists; define \(R = \sup A\).

(III) \(|x| < R \implies \sum |a_n x^n| \text{ converges.}\) This is because \(|x| < R \implies x \in (-c, c) \subseteq A\).
8.2 Convergence at endpoints, Abel summation

Convergence of power series can be difficult to determine at \( R, -R \).

**Definition 8.2.1.** *Abel summation.* Suppose that a power series \( \sum a_n x^n \) converges to a continuous function \( f \) on \((-1, 1)\). If \( f \) is defined and continuous at \( x = 1 \), then say the series is Abel-summable to \( f(1) \), even if the series diverges at \( x = 1 \):

\[
\sum a_n x^n = f(x) \implies \sum a_n = f(1).
\]

This can be used to find the value of uncooperative numerical series.

**Example 8.2.1.** The divergent series \( \sum (-1)^n = 1 - 1 + 1 - 1 + 1 - \ldots \) is Abel-summable to \( \frac{1}{2} \):

\[
1 - x + x^2 - x^3 + \cdots = \sum_{n=0}^{\infty} x^n = \frac{1}{1 + x}, \quad |x| < 1.
\]

This function is continuous at \( x = 1 \), so the Abel sum of \( \sum (-1)^n \) is \( \frac{1}{2} \).

8.3 Linearity of power series

**Theorem 8.3.1.** If \( f(x) = \sum a_n x^n \) and \( g(x) = \sum b_n x^n \), then for any \( p, q \in \mathbb{R} \),

\[
pf(x) + qg(x) = \sum (pa_n + qb_n)x^n
\]

is valid on the common domain of convergence of \( f, g \).

**Proof.** Since this is true for every fixed value of \( x \) by prev thm, done. \( \square \)

**Example 8.3.1.**

\[
1 + x + x^2 + x^3 + \ldots = \frac{1}{1 - x}, \quad |x| < 1
\]
\[ 1 - x + x^2 - x^3 + \ldots = \frac{1}{1 + x}, \quad |x| < 1 \]
\[ 2(1 + x^2 + x^4 + \ldots) = \frac{1}{1 - x} + \frac{1}{1 + x} = \frac{2}{1 - x^2} \]
\[ 1 + x^2 + x^4 + \ldots = \frac{1}{1 - x^2} = \sum (x^2)^n. \]

### 8.4 Multiplication of power series

**Q:** In what order to take the terms of an infinite FOIL-expansion? (No “L”!)

**A:** Group terms by powers of \( x \) (convention).

**Theorem 8.4.1** (Cauchy product). If \( f(x) = \sum a_n x^n \) and \( g(x) = \sum b_n x^n \), then
\[
f(x)g(x) = \sum_{n=1}^{\infty} \left( \sum_{i+j=n} a_i b_j \right) x^n = \sum_{n=1}^{\infty} \left( \sum_{i=0}^{n} a_i b_{n-i} \right) x^n
\]
on their common domain of convergence.

**Proof.** The result holds at any fixed \( x \in (-R, R) \) by Cauchy Prod thm for numerical series.

**Exercises:** #8.1.1(beh), 8.3.1  
**Recommended:** #8.2.2

**Problems:** #8-1  
**Recommended:** #

**Due:** Mar.

1. For how many points \( x \in \mathbb{R} \) can a power series converge conditionally?
Chapter 9

Functions of One Variable

9.1 Functions

Definition 9.1.1. A function from a set $D$ to a set $R$ is a subset $f \subseteq D \times R$ for which each element $d \in D$ appears in exactly one element $(d, \cdot) \in f$. Write $f : D \to R$. If $(x, y) \in f$, then we usually write $f(x) = y$.

$D$ is the domain of the function; the subset of $R$ of elements $x$ for which the function is defined.

$R$ is the range; a subset of $R$ which contains all the points $f(x)$. Generally, assume $R = \mathbb{R}$.

Definition 9.1.2. A function is a rule of assignment $x \mapsto f(x)$, where for each $x$ in the domain, $f(x)$ is a unique and well-defined element of the range.

$f(x) = y$ means "$f$ maps $x \in D$ to $f(x) \in R$".

NOTE: a function whose domain is $\mathbb{N}$ is generally called a sequence, and $a : \mathbb{N} \to \mathbb{R}$ is denoted by $a_n := a(n)$.

Definition 9.1.3. The image of $f$ is the subset

$$\text{Im } f := \{y \in R : \exists x \in D, f(x) = y\} \subseteq R.$$
The function $f$ is surjective or onto iff $\text{Im } f = \mathbb{R}$, that is,

$$\forall y \in \mathbb{R}, \exists x \in D, f(x) = y.$$ 

**Definition 9.1.4.** For $f : D \to \mathbb{R}$, the preimage of $B \subseteq \mathbb{R}$ is the subset

$$f^{-1}(B) := \{x : f(x) \in B\} \subseteq D.$$ 

**Example 9.1.1.** The preimage of $[0, 1]$ under $f(x) = x^2$ is $[-1, 1]$.
The preimage of $[-1, 1]$ under $f(x) = \sin x$ is $\mathbb{R}$.
The preimage of $\{1\}$ under $f(x) = \sin x$ is $\{2k\pi\}_{k \in \mathbb{Z}}$.
The preimage of $[0, 1]$ under $\log x$ is $[1, e]$.

**NOTE:** one can discuss the preimage of any function but the preimage is not necessarily a function. In fact, the preimage $f^{-1}$ is a function iff $f$ is both injective and surjective.

**Definition 9.1.5.** A function $f$ is injective or one-to-one iff no two distinct points in $D$ get mapped onto the same point in $R$, i.e.

$$f(x) = f(y) \implies x = y.$$ 

**Example 9.1.2.** $f(x) = x^2$ is injective on $(0, \infty)$ but not on $\mathbb{R}$.
$f(x) = \frac{1}{x}$ is injective on $\mathbb{R} \setminus \{0\}$.

**Theorem 9.1.6.** TFAE:

1. $f$ is invertible.
2. $f$ is bijective.
3. $f^{-1}$ is a function.
4. $\exists g$ such that $g \circ f = \text{id}_D$ and $f \circ g = \text{id}_R$.

Suppose $f^{-1}$ exists. Then a point $(x, y)$ is in the graph of $f$ iff $(y, x)$ is in the graph of $f^{-1}$.

**NOTE:** an easy way to see that $f$ is injective is to prove it is strictly increasing (or decreasing) on its domain.

**NOTE:** if $f$ is continuous, an easy way to see that $f$ is surjective onto $[a, b]$ is to find $x$ such that $f(x) = a$ and $y$ such that $f(y) = b$. 

Theorem 9.1.7. A function which is continuous and strictly increasing (or decreasing) is invertible.

Theorem 9.1.8. If $f$ is injective, then its inverse can be defined on its image.

9.2 Algebraic operations on functions

Definition 9.2.1.

\[
\begin{align*}
  f + g + (f + g)(x) & := f(x) + g(x) \\
  fg + (f g)(x) & := f(x) g(x) \\
  c f + (c f)(x) & := c f(x) \\
  g \circ f + (g \circ f)(x) & := g(f(x))
\end{align*}
\]

Definition 9.2.2. Translation: for $a > 0$,

- $f(x + a)$ left-shift by $a$
- $f(x - a)$ right-shift by $a$.

Change of scale: for $a > 1$,

- $f(x/a)$ horizontal expansion by factor $a$
- $f(ax)$ horizontal contraction by factor $a$.

Vertical expansion/contraction is $cf$.

9.3 Properties of functions

All properties like increasing, decreasing, monotone, etc, of sequences apply directly to functions: replace $n = a, n + 1 = b$, where $b > a$ throughout the defns.

Symmetries of functions:

Definition 9.3.1. $f$ is even iff $f(-x) = f(x), \forall x \in D$. $f$ is odd iff $f(-x) = -f(x), \forall x \in D$. 
Theorem 9.3.2. An even function times an even function is even. An odd function times an even function is odd. What about an odd function times an odd function? (Compare with the rules for integers.)

Proof. HW

Theorem 9.3.3. Suppose the domain of $f$ is symmetric about 0. Then $f$ has a unique representation as the sum of an even and an odd function:

$$f(x) = E(x) + O(x), \quad E(x) \text{ even and } O(x) \text{ odd.}$$

Proof. HW 9.3.1

Definition 9.3.4. $f$ is periodic iff $\exists T > 0$ such that $f(x + T) = f(x), \forall x \in D$. The smallest such $T$ is called the period of $f$, if it exists.

Example 9.3.1. A function which is even an monotone must be constant.

9.4 Elementary functions

1. Rational functions $\frac{p(x)}{q(x)}$, where $p, q$ are polynomials.

2. The six trigonometric functions $\sin x, \cos x, \ldots$, and their inverses.

3. The exponential $e^x$ and its inverse $\log x$.

4. The algebraic functions $y = y(x)$ which satisfy some

$$y^n + a_{n-1}(x)y^{n-1} + \ldots + a_1(x)y + a_0(x) = 0,$$

where $a_k(x)$ are rational functions. (eg, $\sqrt{x}$)

Exercises: #9.2.1,9.2.2,9.3.4,  Recommended: #9.3.5
Problems: #9-1,9-2  Recommended: #
Due: Mar.

1. An even function times an even function is even. An odd function times an even function is odd. What about an odd function times an odd function? (Compare with the rules for integers.)
Chapter 10

Local and Global Behavior

10.1 Intervals. Estimating functions

Intervals are “building blocks” of sets, especially domains.

\[(a, b) \text{ open} \quad (a, \infty) \text{ open} \quad (-\infty, b) \text{ open}\]
\[[a, b] \text{ closed} \quad [a, \infty] \text{ wrong}\]
\[[a, b) \text{ neither} \quad (a, b] \text{ neither} \]

Definition 10.1.1. A \( \delta \)-neighbourhood of \( a \) is \((a - \delta, a + \delta)\), and TFAE

\[x \in (a - \delta, a + \delta), \quad a - \delta < x < a + \delta, \quad |x - a| < \delta, \quad x \approx_\delta a.\]

Definition 10.1.2. \( B \) is an upper bound for \( f \) on an interval \( I \) iff \( B \) is an upper bound of \( f(I) \). Similarly,

\[\sup_{I} f(x) = \sup \{f(x) : x \in I\}\]
\[\max_{I} f(x) = \max \{f(x) : x \in I\}\]

Similarly for inf, min.

NOTE: NONE of these need exist!

\(^1\)April 18, 2007
Theorem 10.1.3 (Completeness for functions). Suppose $f$ is bounded above on $I$. Then $\sup_I f(x)$ exists.

NOTE: there may not actually exists $c \in I$ for which $f(c) = \sup_I f(x)$!

Boundedness/estimates: $|f(x)| \leq B$.
This mean $f$ is bounded above by the constant function $B(x) \equiv B$.
More generally, use a bound like $|f(x)| \leq g(x)$.

NOTE: $|f(x)g(x)| = |f(x)||g(x)|$, $|f(x) + g(x)| \leq |f(x)| + |g(x)|$.

USE: apply to integrals.

Theorem 10.1.4. If $f < g$ on $I$ and the integrals exist, then for any $a, b \in I$,

$$\int_a^b f(x) \, dx < \int_a^b g(x) \, dx, \quad a < b.$$ 

Proof. coming ...

Corollary 10.1.5. If $f$ is bounded on $[a, b]$ and its integral exists, then

$$\int_a^b f(x) \, dx \leq M(b-a).$$ 

Proof. Use $g(x) \equiv M$, where $M$ is a bound for $f$.

Example 10.1.1. Show that $\text{erf } x = \int_0^x e^{-t^2/2} \, dt$ is bounded above on the interval $[0, \infty)$.

Solution. We have an upper bound for the (positive) integrand given by

$$t \leq t^2 \implies e^{-t^2/2} \leq e^{-t/2},$$

however this is only true for $t \geq 1$. But suffices to only consider this domain!

$$\text{erf } x = \int_0^x e^{-t^2/2} \, dt$$

$$= \int_0^1 e^{-t^2/2} \, dt + \int_1^x e^{-t^2/2} \, dt$$

$$\leq M + \int_1^x e^{-t/2} \, dt$$
10.2 Approximating functions

\[ M + \frac{1}{2} e^{-1/2} \leq \leq M + 2e^{-1/2}, \quad x \geq 1. \]

More generally:

**Theorem 10.1.6.** \( f \approx \varepsilon \, g \) on \([a, b]\) \( \implies \int_a^b f(x) \, dx \approx \varepsilon (b-a) \int_a^b g(x) \, dx. \)

**Proof.** HW; apply the definition of \( \approx \varepsilon \) and the corollary above.

10.2 Approximating functions

**Definition 10.2.1.** Similar to prev, TFAE for \( x \in I \)

\[ g(x) - \delta < f(x) < g(x) + \delta, \quad |f(x) - g(x)| < \delta, \quad f(x) \approx \delta \, g(x), \]

\[ f(x) = g(x) + e(x), \quad |e(x)| < \varepsilon. \]

**Example 10.2.1.** To find a \( \delta \)-neighbourhood of 0 where \( \sin x \approx \varepsilon \, x \), \( \varepsilon = .001 \), apply the corollary to Alternating Series test to the series expn

\[ \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \]

\[ |\sin x - x| \leq \frac{x^3}{3!}, \quad 0 < x < 1 \]

\[ < 0.001, \quad x^3 < .006, \quad \text{so } x < 0.18. \]

Since \( \sin x \) is symmetric about 0 (odd fn), let \( \delta = 0.18 \).

10.3 Local behavior

The local behavior of \( f \) is what can be seen by studying \( f \) in a neighbourhood of \( x \).

**Definition 10.3.1.** \( f \) is locally increasing at \( x \) means \( f \) is increasing on \( I = (x - \delta, x + \delta) \), for some \( \delta \).

\( f \) is locally bounded at \( x \) means \( f \) is bounded on \( I = (x - \delta, x + \delta) \), for some \( \delta \).
Example 10.3.1. \( \sin x \) is locally increasing at every \( 2\pi n, n \in \mathbb{Z} \).

NOTE: \( f \) need not actually be defined at \( x \).

Example 10.3.2. \( \frac{1}{x} \) is locally bounded at any \( x \neq 0 \). However, \( \frac{1}{x} \) is not bounded near 0; it’s unbounded in any neighbourhood of 0.

\( \sin \frac{1}{x} \) is not monotonic near 0.

Theorem 10.3.2. \( f, g \) locally bounded near \( x \) \( \implies \) so is \( f + g \).

Proof. Straightforward: just use \( \delta = \min\{\delta_f, \delta_g\} \).

Definition 10.3.3. A property of \( f \) is true for \( x \gg 1 \) or “for \( x \) near \( \infty \)” or “at \( \infty \)” if it holds on some interval \((a, \infty)\).

Example 10.3.3. \( \frac{1}{x} \) and \( e^{-x} \) are functions that “vanish at \( \infty \)”, i.e., satisfy \( |f(x)| < \varepsilon \) on \((N, \infty)\) for sufficiently large \( N \). Every nonconstant polynomial is unbounded near \( \infty \).

10.4 Local and global properties

Definition 10.4.1. \( f \) is locally bounded on \( I = (a, b) \) iff it is locally bounded at any point in \( I \).

Example 10.4.1. \( \frac{1}{x} \) is locally bounded on \((0, \infty)\) but not bounded. NOTE: must use increasingly \( \delta \) for \( a \) near 0.

Definition 10.4.2. \( f \) is locally increasing on \( I \) iff it is locally increasing on every interval in its domain.

Example 10.4.2. \( \tan x \) is locally increasing but not increasing on \( \mathbb{R} \setminus \{(n + \frac{1}{2})\pi\}_{n \in \mathbb{Z}} \).

Most important examples: continuity and differentiability. These are local but not pointwise properties: need some neighbourhood of the point in question.

Example 10.4.3. \( f \) is positive at \( c \). This is a pointwise property, and it satisfied iff \( f(c) > 0 \). Don’t need a neighbourhood.

Exercises: #10.1.7, 10.1.9, 10.3.2  Recommended: #10.1.1, 10.3.1

Problems: #10-2  Recommended: #10-3

Due: Mar.

1. If \( f \) and \( g \) are bounded on \( I \), show that \( f + g \) and \( fg \) are bounded on \( I \).
Chapter 11

Continuity and limits

11.1 Continuous functions

IDEA: if $x$ and $y$ are related by a function $f(x) = y$, then we want to say that $y$ varies continuously with respect to $x$ iff small changes in $x$ produce small changes in $y$; no jumping!

Example 11.1.1. Define the Heaviside function

$$h(x) = \begin{cases} 
0, & x \leq 0 \\
1, & x > 0.
\end{cases}$$

Then for $x$ in any small neighbourhood of $0$, varying $x$ by $\delta$ can produce a sudden jump of distance $1$; cannot make this jump less than, e.g., $\varepsilon = \frac{1}{2}$ by restricting to smaller $\delta$.

Definition 11.1.1. $f$ is continuous at $c$ iff it is defined at $c$ and

$$\forall \varepsilon > 0, \ 0 \approx c \implies f(x) \approx f(c)$$

$$\forall \varepsilon > 0, \exists \delta, \ x \approx_\delta c \implies f(x) \approx f(c)$$

$$\forall \varepsilon > 0, \exists \delta, \ |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.$$
Definition 11.1.2. \( f \) has a limit from the left at \( x \) (or is left-continuous at \( x \)) iff

\[
\forall \varepsilon > 0, \quad x \approx \delta, x < c \implies f(x) \approx \varepsilon f(c)
\]

\[
\forall \varepsilon > 0, \exists \delta, \quad 0 < c - x < \delta \implies |f(x) - f(c)| < \varepsilon.
\]

Write \( f(c-) := \lim_{x \to c-} f(x) \). Similarly for \( f(c+) := \lim_{x \to c+} f(x) \):

\[
\forall \varepsilon > 0, \exists \delta, \quad 0 < x - c < \delta \implies |f(x) - f(c)| < \varepsilon.
\]

So clearly \( f \) is continuous at \( c \) iff \( f(c-) = f(c+) \) (where of course both limits must exist).

Definition 11.1.3. \( f \) is continuous on \( I \) iff \( I \) is an interval (possibly infinite) and \( f \) is continuous at every point \( c \in I \).

Definition 11.1.4. \( f \) is Lipschitz (or strongly continuous) iff it satisfies a Lipschitz condition:

\[
\forall x, y \quad |f(x) - f(y)| \leq C|x - y|, \quad C > 0.
\]

In this case, \( C \) is the Lipschitz constant of \( f \).

Theorem 11.1.5. \( f \) is Lipschitz \( \implies \) \( f \) is continuous.

Proof. Let \( \delta = \frac{\varepsilon}{C} \).

Example 11.1.2. Show \( \sin x \) is Lipschitz continuous with constant \( C = 1 \).

Solution. Let \( P_0 \) be a point on the unit circle, corresponding to the arc \( AP_0 \) of length \( c \). Then \( P_0 = (\cos c, \sin c) \) in coordinates. We need to see that for \( x \) near \( c \), \( P = (\cos x, \sin x) \) is near \( P_0 \). The difference between the heights of the two triangles is

\[
|\sin x - \sin c| = |PR| \leq |PP_0| = |x - c|,
\]

since the line \( PR \) is the shortest curve from \( P \) to the horizontal line \( P_0R \) (\( PP_0 \) is the arc along the circle).

Theorem 11.1.6. \( |\int f(t) \, dt| \leq \int |f(t)| \, dt \).

Theorem 11.1.7. \( \int f(t) \, dt + \int g(t) \, dt = \int (f(t) + g(t)) \, dt \).
Example 11.1.3. Show that the integral \( \int_0^\pi \frac{\sin xt}{t} \, dt \) depends continuously on \( x \), and is in fact Lipschitz continuous with constant \( \pi \).

Solution. We prove \( f(x) := \int_0^\pi \frac{\sin xt}{t} \, dt \) is a Lipschitz continuous function:

\[
|f(x) - f(y)| = \left| \int_0^\pi \frac{\sin xt}{t} \, dt - \int_0^\pi \frac{\sin yt}{t} \, dt \right| \\
= \left| \int_0^\pi \frac{\sin xt - \sin yt}{t} \, dt \right| \\
\leq \int_0^\pi \left| \frac{\sin xt - \sin yt}{t} \right| \, dt \\
\leq \int_0^\pi \left| \frac{x-t}{t} \right| \, dt \\
= |x - y| : \int_0^\pi 1 \, dt \\
= \pi |x - y|.
\]

11.1.1 Discontinuities

Definition 11.1.8. \( f \) has a simple (or removable) discontinuity at \( c \) iff (re)defining \( f(c) \) could make \( f \) continuous at \( c \). In this case, \( f(c-) = f(c+) \), but \( f(c) \) is something else.

Example 11.1.4. \( f(x) = \frac{x^2-9}{x-3} \).

Definition 11.1.9. \( f \) has a jump discontinuity at \( c \) iff \( f(c-) \neq f(c+) \) (in which case \( f(c) \) is immaterial).

Example 11.1.5. \( f(x) = \frac{x}{|x|} \).

Definition 11.1.10. \( f \) has an infinite discontinuity at \( c \) iff \( f(c-) = \pm\infty \) or \( f(c+) = \pm\infty \).

Example 11.1.6. \( f(x) = \frac{1}{x} \).

Definition 11.1.11. \( f \) has an essential singularity at \( c \) iff it is not one of the other kinds.

Example 11.1.7. \( f(x) = \sin \frac{1}{x} \).

11.2 Limits of functions

If \( f(c) \) is not defined, we may still be able to talk about what it “ought” to be.

Definition 11.2.1. \( f(x) \) has the limit \( L \) as \( x \to c \) iff

\[
\forall \varepsilon > 0, \ x \approx_\delta c, x \neq c \implies f(x) \approx_\varepsilon L
\]
∀ε > 0, ∃δ, 0 < |x − c| < δ ⇒ |f(x) − L| < ε.

Example 11.2.1. \( \lim_{x \to 0} x \sin \frac{1}{x} = 0. \)

Solution. Fix ε > 0. Then for 0 < |x| < ε, \( |x \sin \frac{1}{x}| = |x| \cdot |\sin \frac{1}{x}| \leq |x| < \varepsilon. \)

Example 11.2.2. \( \lim_{x \to 1^-} \sqrt{1 - x^2} = 0. \)

Solution. The function is not defined for \( x > 1 \). Fix ε > 0. Then for \( 0 < 1 - x < \frac{\varepsilon}{2} \),

\[
\sqrt{1 - x^2} = \sqrt{1 - x \sqrt{1 + x}} < \sqrt{2 \sqrt{1 + x}} < \varepsilon.
\]

Definition 11.2.2. \( f \) has a limit at \( \infty \) iff given ε > 0, we have \( f(x) \approx_{\varepsilon} L \) for \( x \gg 1 \).

Example 11.2.3. \( \lim_{1+x^2} = 0. \)

Solution. Fix ε > 0. Then for \( x > \varepsilon^{-1/2}, 1 + x^2 > \frac{1}{\varepsilon} \Rightarrow \frac{1}{1+x^2} < \varepsilon. \)

Definition 11.2.3. If \( f \) is defined in \( (c - \delta, c + \delta) \) except at \( c \), for some \( \delta > 0 \). Then \( \lim_{x \to c} f(x) = \infty \) iff given \( N \in \mathbb{N} \),

\[
\exists \delta_0, \ x \approx_{\delta_0} c \Rightarrow f(x) > N.
\]

Example 11.2.4. \( \lim_{x \to 0} \frac{1}{x^2} = \infty. \)

Solution. Fix \( N \in \mathbb{N} \). For any \( x \neq 0 \) with \( |x| < \frac{1}{\sqrt{N}} \), we have \( x^2 < \frac{1}{N} \Rightarrow \frac{1}{x^2} > N. \)

11.3 Limit theorems for functions

11.4 Limits and continuity

Continuous functions are useful because they preserve limits:

\[
\lim_{x \to c} f(x) = f(\lim_{x \to c} x) = f(c).
\]

The reason is that continuous functions map convergent sequences to convergent sequences.
11.4 Limits and continuity

**Theorem 11.4.1.** \( f \) is continuous at \( x \) iff \( \lim_{x \to c} f(x) = f(c) \).

**Proof.** We must show the following two lines are equivalent:

\[
\forall \varepsilon > 0, \ x \approx \delta, x \neq c \implies f(x) \approx \varepsilon f(c) \\
\forall \varepsilon > 0, \ x \approx c \implies f(x) \approx \varepsilon f(c).
\]

Since \( x = c \implies f(x) = f(c) \), this is trivial. \( \square \)

**Theorem 11.4.2** (Sequential continuity). \( \lim_{t \to x} f(t) = L \) iff \( \lim_{n \to \infty} f(x_n) = L \) for every \( \{x_n\} \) with \( x_n \to x \).

**Proof.** (\( \Rightarrow \)) Choose a sequence \( x_n \to x \), and fix \( \varepsilon > 0 \). Since \( f \) is continuous, \( \exists \delta > 0 \) for which

\[
|x - t| < \delta \implies |f(t) - L| < \varepsilon.
\]

Also, there is \( N \) such that

\[
n \geq N \implies |x_n - x| < \delta.
\]

Thus, \( n \geq N \implies |f(x_n) - L| < \varepsilon \).

(\( \Leftarrow \)) Contrapositive: suppose it is false that \( \lim_{t \to x} f(t) = L \). Then:

\[
\exists \varepsilon > 0, \forall \delta > 0, \exists t, \ |x - t| < \delta \text{ and } |f(t) - L| \geq \varepsilon \\
\exists \varepsilon > 0, \forall n \in \mathbb{N}, \exists t_n, \ |x - t_n| < \frac{1}{n} \text{ and } |f(t_n) - L| \geq \varepsilon
\]

This produces \( t_n \to x \) for which it is false that \( \lim_{n \to \infty} f(x_n) = L \). \( \square \)

By this theorem, our entire "limit toolkit" transfers to functions: the comparison theorems, linearity, products & quotients, error form, \( K - \varepsilon \) principle, Squeeze Theorem, Limit Location, Function Location (instead of Sequence Location).

**Corollary 11.4.3.** \( f \) is continuous at \( x \), iff \( \lim_{n \to \infty} f(x_n) = f(x) \) for every \( \{x_n\} \) with \( x_n \to x \).

In other words, for a continuous function \( f \),

\[
x_n \to x \implies f(x_n) \to f(x),
\]

so that continuous functions map convergent sequences to convergent sequences.
Some of the basic limit theorems for functions (the new “function limit toolkit”) can be extended for continuous functions.

**Theorem 11.4.4** (Positivity). *If* \( f \) *is continuous at* \( c \) *and* \( f(c) > 0 \), *then* \( f(x) > 0 \) *for* \( x \approx_x c \).

**Proof.** Let \( x_n \to c \). Then \( f(x_n) \to f(c) \). Since \( f(c) > 0 \), the Sequence Location Theorem gives \( f(x_n) > 0 \) for \( n \gg 1 \).

This proof amounts to saying:

Since \( \lim_{x \to c} f(x) = f(c) > 0 \), the Function Location Theorem gives \( f(x) > 0 \) for \( x \approx_x c \).

**Theorem 11.4.5.** If \( f, g \) are continuous, then so are \( f + g \), \( f \cdot g \), and (if \( g \neq 0 \)) \( f/g \).

**Proof.** Continuous functions preserve sequences, and limits are linear & multiplicative for sequences.

**Theorem 11.4.6.** Let \( x = g(t), c = g(b) \). If \( g(t) \) is continuous at \( b \) and \( f(x) \) is continuous at \( c \), then \( f \circ g(t) = f(g(t)) \) is continuous at \( b \).

**Proof.** Given \( \varepsilon > 0 \), \( \exists \delta > 0 \) such that

\[
\begin{align*}
f(x) &\approx_{x} f(c) \quad \text{for} \ x \approx_{x} c \\
g(t) &\approx_{t} g(b) \quad \text{for} \ t \approx_{t} b
\end{align*}
\]

continuity of \( f \)

continuity of \( g \).

Then \( t \approx_{t} b \implies x = g(t) \approx_{t} g(b) = c \implies f(x) \approx_{x} f(c) \).

**Example 11.4.1.** \( f(x) = \cos \frac{1}{x} \) has an essential discontinuity at 0.

**Solution.** HW: Consider the sequences

\[
\begin{align*}x_n & := \frac{1}{2n\pi} & f(x_n) &= 1 \\
y_n & := \frac{1}{(2n+1)\pi} & f(y_n) &= -1
\end{align*}
\]

By Sequential Continuity, \( f(0+) \) cannot exist.
Theorem 11.4.7 (Pasting Lemma). Let \( f \) be continuous on \([a, b]\) and \( g \) be continuous on \([b, c]\). If \( f(b) = g(b) \), then

\[
    h(x) := \begin{cases} 
        f(x) & x \in [a, b] \\
        g(x) & x \in [b, c] 
    \end{cases}
\]

is continuous on \([a, c]\).

Proof. HW

Since \( f, g \) are continuous on their domains, only remains to check continuity at \( b \). Choose \( \delta_1 \) so that

\[
    0 < b - x < \delta_1 \implies |f(x) - f(b)| < \varepsilon,
\]

and choose \( \delta_2 \) so that

\[
    0 < x - b < \delta_2 \implies |g(x) - g(b)| < \varepsilon.
\]

Define \( \delta := \min\{\delta_1, \delta_2\} \). Then

\[
    x \approx_\delta b \implies f(x) \approx_\varepsilon f(b) = g(b). \tag*{\qed}
\]

Exercises: \#11.1.4, 11.2.2, 11.3.3, 11.4.2, 11.5.1, 11.5.4 \hspace{0.5cm} \text{Recommended:} \#11.3.5, 11.3.6, 11.4.4, 11.5.5, 11.5.6

Problems: \#11-1, 11-2 \hspace{0.5cm} \text{Recommended:} \#11-3

Due: Mar.
Math 311  

Continuity and limits
Chapter 12

Intermediate Value Theorem

12.1 Existence of zeros

Theorem 12.1.1 (Bolzano’s Thm). Let $f$ be continuous on $[a, b]$ with $f(a) < 0 < f(b)$. Then there is a point $c \in (a, b)$ where $f(c) = 0$.

Proof. Define

$$C := \{x \in [a, b] : f(x) \leq 0\}.$$ 

Since $a \in C$ and $b$ is an upper bound for $C$, we may define

$$c := \sup C.$$ 

Suppose $f(c) > 0$. Then by the Positivity Theorem, we could find $\varepsilon > 0$ with $f(x) > 0$ on $(c - \varepsilon, c)$, so that $c$ would not be the smallest upper bound (e.g. $\frac{c - \varepsilon}{2}$ would be smaller).

Suppose $f(c) < 0$. Then by the Negativity Theorem, we could find $\varepsilon > 0$ with $f(x) < 0$ on $(c, c + \varepsilon)$, so that $c$ would not be an upper bound.

Theorem 12.1.2 (IVT). Let $f$ be continuous on $[a, b]$ and $f(a) \leq f(b)$. Then for $k \in \mathbb{R}$,

$$f(a) \leq k \leq f(b) \implies \exists c \in [a, b], f(c) = k.$$
Proof. If \( f(a) = f(b) \), then with \( k = a \) or \( k = b \) we are done, so assume wlog \( f(a) < k < f(b) \). Then \( f(a) - k < 0 < f(b) - k \), so apply Bolzano’s Thm to get \( f(c) - k = 0 \). 

\[ \square \]

12.2 Applications of Bolzanno

Example 12.2.1. A polynomial of odd degree has a real zero.

Proof. Consider \( x^{2k+1} = x(x^2)^k \). Then

\[ x < 0 \implies x(x^2)^k < 0, \quad x > 0 \implies x(x^2)^k > 0. \]

Apply this to the polynomial \( a_0 + a_1 x + \ldots + a_n x^n \), assuming \( a_n \neq 0 \) (so it has degree \( n = 2k + 1 \)). The lesser terms will not matter for \( |x| > 1 \):

\[
|a_{n-1}x^{n-1} + \ldots + a_0| \leq |a_{n-1}|x^{n-1} + \ldots + |a_0| \\
\leq |a_n x^n| \left( \frac{|a_{n-1}|}{|a_n x|} + \frac{|a_{n-2}|}{|a_n x^2|} + \ldots + \frac{|a_0|}{|a_n x^n|} \right) \\
< |a_n x^n|.
\]

Then the polynomial satisfies the conditions for Bolzano’s Thm. 

\[ \square \]

Theorem 12.2.1 (Intersection Principle). (a) The solutions of \( f(x) = g(x) \) are the values of \( x \) for which the graphs intersect.

(b) If \( f, g \) are continuous on \( [a, b] \) and \( f(a) \leq g(a) \) but \( f(b) \geq g(b) \), then the graphs intersect at some \( c \in [a, b] \).

Proof of (a). A point in the graph of \( f \) looks like \((x, f(x)) \in \mathbb{R}^2\). So a point lies on each graph iff \((c, f(c)) = (c, g(c)) \) in \( \mathbb{R}^2 \).

\[ \square \]

Proof of (b). Apply Bolzano’s Theorem to the continuous function \( f(x) - g(x) \).

\[ \square \]

12.3 Monotonicity and the IVP

Definition 12.3.1. A function \( f \) has the Intermediate Value Property on \([a, b]\) iff it is defined on \([a, b]\) and takes on all values between \( f(a) \) and \( f(b) \) as \( x \) varies between \( a \) and \( b \).
12.4 Inverse functions

**Theorem 12.3.2** (Continuity for monotone functions). *If the function* \( f \) *is strictly increasing and has the IVP on* \([a, b]\), *then it is continuous on* \([a, b]\).

**Proof.** Wlog, let \( f \) be increasing. Let \( c \in (a, b) \); show \( f \) continuous at \( c \). Fix \( \varepsilon > 0 \) and find \( x_1, x_2 \) such that
\[
f(x_1) = f(c) - \varepsilon \quad \text{and} \quad f(x_2) = f(c) + \varepsilon.
\]
Since \( f \) is strictly increasing, a small enough \( \varepsilon \) will ensure \( f(x_1) > f(a) \) and \( f(x_2) < f(b) \). Then IVP ensures these two points exist, and strictly increasing gives
\[
a < x_1 < c < x_2 < b \quad \text{and} \quad f(a) < f(x_1) < f(c) < f(x_2) < f(b),
\]
so that these points \( x_1, x_2 \) are unique. Define \( \delta := \min\{|c - x_1|, |c - x_2|\} \). Then we have \((c - \delta, c + \delta) \subseteq (x_1, x_2) \) and so
\[
x \approx \delta c \implies f(x) \approx \varepsilon f(c)
\]
\[
|x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.
\]

**NOTE:** it is clear how the proof goes through for a strictly decreasing function, but it also extends to a piecewise strictly monotone function, e.g. \( \sin x \). So to prove \( \sin x \) is continuous, you could apply this theorem and it would only remain to check that \( \sin x \) is continuous at the points \( x = (2k + 1)\pi/2 \), which follows immediately from the Pasting Lemma.

12.4 Inverse functions

**Theorem 12.4.1** (Inverse of increasing functions). *If* \( y = f(x) \) *is continuous and strictly increasing on* \([a, b]\), *then it has an inverse function* \( x = g(y) \) *which is also continuous and strictly increasing.*

**Proof.** (1) \( g \) is defined on \([f(a), f(b)]\).

Fix any point \( y \) in \([f(a), f(b)]\). By IVT, there is an \( x \in [a, b] \) with \( f(x) = y \), and this point is unique because \( f \) is strictly increasing. Thus, \( g(y) := x \) is well-defined.
(2) \( g \) is strictly increasing.

Let \( x_1 := g(y_1) \) and \( x_2 := g(y_2) \). Since \( f \) is strictly increasing,

\[
\begin{align*}
    x_1 \leq x_2 & \implies f(x_1) \leq f(x_2) \\
    g(y_1) \leq g(y_2) & \implies y_1 \leq y_2 \\
    y_1 > y_2 & \implies g(y_1) > g(y_2).
\end{align*}
\]

(3) \( g \) is continuous.

We’ve seen \( g \) is monotone. To see \( g \) has IVP, note that

\[
    a \leq x \leq b \implies g(y) = x,
\]

where \( y := f(x) \). Then done by prev thm.

\[\Box\]

NOTE: of course, this theorem is also true with “increasing” replaced by “decreasing” throughout, \textit{mutatis mutandis}.

Exercises: #12.1.1, 12.2.1 (make a map), 12.4.1  Recommended: #12

Problems: #12-2, 12-7  Recommended: #12

Due: Mar.
Chapter 13

Continuity and Compact Intervals

13.1 Compact intervals

Definition 13.1.1. A set $S \subseteq \mathbb{R}$ is sequentially compact iff every sequence in $S$ has a subsequence converging to a point of $S$.

Given an interval $(a, b)$, how could this fail?

$$b - \frac{1}{n} \in (a, b) \text{ for } n \gg 1, \quad b - \frac{1}{n} \to b, \text{ but } b \not\in (a, b).$$

Fix: use a closed interval instead! That fixes this flaw. But $[a, \infty)$ is a closed interval, so consider $\{a + n\}$.

Fix: also require boundedness.

Definition 13.1.2. A interval in $\mathbb{R}$ is compact iff it is closed and bounded. More generally, a set is compact if it is obtained from taking finite unions or arbitrary intersections of compact intervals.

Example 13.1.1. The discrete set $\{0, 1, 2\}$ is compact, since it can be written as a union of finitely many (ie, three) singleton sets. Each of these is compact, since, eg,

$$\{0\} = \bigcap_{n=1}^{\infty} [0, \frac{1}{n}].$$

\footnote{April 18, 2007}
Note: obviously, cannot take infinite unions, or else

\[ [a, \infty) = \bigcup_{n=1}^{\infty} [a + n, a + n + 1] \]

would be a compact interval which is not bounded.

**Theorem 13.1.3** (Sequential Compactness Thm). *A compact interval \([a, b]\) is sequentially compact.*

**Proof.** Let \(\{x_n\} \subseteq [a, b]\). Then \(a \leq x_n \leq b\), so the sequence is bounded and B-W thm applies to give a convergent subsequence \(x_{n_i} \to c\). Then the Limit Location thm gives

\[ a \leq x_{n_i} \leq b \implies a \leq \lim_{i \to \infty} x_{n_i} \leq b. \]

NOTE: the following converse also holds: if \(S \subseteq \mathbb{R}\) is sequentially compact, then it is also compact. However, proving this requires developing the full definition of compactness in terms of open covers, etc, so we just take it on faith.

**Theorem 13.1.4.** *If \(S \subseteq \mathbb{R}\) is a sequentially compact set, then \(S\) is compact.*

### 13.2 Bounded continuous functions

**Theorem 13.2.1** (Boundedness Thm). *If \(f\) is continuous on a compact interval \(I\), then \(f\) is bounded on \(I\).*

**Proof.** Show \(f\) is bounded above via contrapositive:

\[ f \text{ has no upper bound on } I \implies f \text{ not continuous on } I. \]

Define a sequence in \(I\) by choosing \(x_n\) such that \(f(x_n) > n, \forall n \in \mathbb{N}\). Since \(f\) is unbounded, this is possible. Since \(I\) is compact, sequential compactness gives a subsequence \(\{x_{n_i}\}\) which converges, so that \(x_{n_i} \to c \in I\). Further,

\[ f(x_n) \to \infty \implies f(x_{n_i}) \to \infty. \]
So \( f \) cannot be continuous, or else we’d have

\[
 f(c) = \lim_{i \to \infty} f(x_n_i) = \infty, \nabla
\]

The proof that \( f \) must be bounded below follows, \textit{mutatis mutandis}.

\[ \square \]

\textbf{NOTE:} compactness allows us infer a global property (boundedness) from a local property (continuity).

\textbf{NOTE:} compactness is necessary. Consider \( x^2 \) on \([0, \infty)\) or \( \frac{1}{x} \) on \((0, 1)\).

\section*{13.3 Extrema of continuous functions}

\textbf{Theorem 13.3.1} (Maximum Thm). \textit{Let \( f \) be continuous on the compact interval \( I \). Then \( f \) attains its maximum and minimum on \( I \): \exists \alpha, \beta \in I \) such that

\[
 f(\alpha) = \inf_{x \in I} f(x), \quad \text{and} \quad f(\beta) = \sup_{x \in I} f(x).
\]

Proof. Boundedness Thm shows \( f \) is bounded above, so completeness gives existence of

\[
 M := \sup_{x \in I} f(x).
\]

So \( f(x) \leq M \) for every \( x \in I \), and \( M \) is the smallest number for which this inequality is true. Thus, we can pick

\[
 x_n \in I \quad \text{with} \quad f(x_n) \in [M - \frac{1}{n}, M] \subseteq I, n \gg 1.
\]

By Sequential Compactness Thm, \( \{x_n\} \) has a convergent subsequence with limit \( \beta := \lim_{i \to \infty} x_{n_i} \in I \). Then the Squeeze Thm gives

\[
 M - \frac{1}{n} \leq f(x_{n_i}) \leq M \quad \implies \quad M \leq \lim_{i \to \infty} f(x_{n_i}) \leq M,
\]

and continuity of \( f \) gives \( f(\beta) = \lim_{i \to \infty} f(x_{n_i}) = M \).

\[ \square \]

\textbf{NOTE:} here, compactness allows us to infer a global property (having a max) from a local property (continuity).

\textbf{NOTE:} compactness is necessary. Else \( x \) has no max or min on \((0, 1)\) and no max on \([0, \infty)\).
13.4 The mapping viewpoint

Consider \( f : D \to \mathbb{R} \) as a map of \( D \). We are interested in the image \( f(D) = \{ f(x) : x \in D \} \).

The continuous image of a compact set is compact.

**Theorem 13.4.1** (Continuous mapping thm). If \( f \) is defined and continuous on a compact interval \( I \), then \( f(I) \) is a compact interval.

**Proof.** By Maximum Thm, there are \( \alpha, \beta \) with

\[
 f(\alpha) = m := \inf_{x \in I} f(x) \quad \text{and} \quad f(\beta) = M := \sup_{x \in I} f(x).
\]

We show \( f(I) = [m, M] \) by double inclusion.

\((\subseteq)\) By the definition of \( m, M \),

\[
 x \in I \implies m \leq f(x) \leq M \implies f(x) \in [m, M].
\]

\((\supseteq)\) Let \( y \in [m, M] \). Since \( f \) is continuous, IVT gives

\[
 f(\alpha) \leq y \leq f(\beta) \implies \exists x \in [\alpha, \beta], f(x) = y.
\]

But \( f(x) = y \) means \( y \in f(I) \). \( \square \)

**NOTE:** the proof of this theorem used Max Thm and IVT. However, the Contin Mapping Thm also implies these two (HW 13.4.1).

13.5 Uniform continuity

**Definition 13.5.1.** \( f \) is continuous on \( I \) iff

\[
 \forall \varepsilon > 0, \forall x \in I, \exists \delta > 0, \quad y \approx_{\delta} x \implies f(y) \approx_{\varepsilon} f(x).
\]

This is a local property: it is verified by considering \( f \) on small neighbourhoods of points in \( I \). Here, \( \delta \) depends on \( x \) and \( \varepsilon \), since \( \delta \) is found after these are fixed.

**Definition 13.5.2.** \( f \) is uniformly continuous on \( I \) iff

\[
 \forall \varepsilon > 0, \exists \delta > 0, \forall x \in I, \quad y \approx_{\delta} x \implies f(y) \approx_{\varepsilon} f(x).
\]
This is a global property: the $\delta$ is independent of $x$ and therefore must work for all $x$ in $I$ simultaneously. Here, $\delta$ depends on $\varepsilon$ only.

NOTE: it is nonsense to ask if $f$ is uniformly continuous at a point.

NOTE: it is nonsense to say “$f$ is uniformly continuous”. You must specify the domain, e.g., $f$ is uniformly continuous on $I$.

Example 13.5.1. $f(x) = \frac{1}{x}$ is uniformly continuous on $(\frac{1}{n}, \infty)$ but not on $(0, \infty)$.

Theorem 13.5.3. $f$ is uniformly continuous on $I$ implies that $f$ is continuous on $I$.

Proof. Too easy to assign for rec HW. \qed

Theorem 13.5.4. On a compact interval, continuity implies uniform continuity.

Proof. Assume the domain $I$ is a compact interval, and prove the contrapositive. Suppose $f$ is not uniformly continuous. Then

$$\exists \varepsilon_0, \forall \delta, \exists x, x \approx \delta y \text{ and } f(x) \neq \varepsilon f(y),$$

Applying this to $\delta_n = \frac{1}{n}$, we can find an $x_n, y_n$ for each $\delta_n$. That is, we can construct sequences $\{x_n\}, \{y_n\}$ where for each $n$,

$$x_n \approx \delta_n y_n \text{ and } |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

By Sequential Compactness, obtain a subsequence $x_{n_i} \to c \in I$ and let $\{y_{n_i}\}$ be the corresponding subsequence (same indices as in $\{x_{n_i}\}$).

By construction, $(x_{n_i} - y_{n_i}) \to 0$. By Linearity of Limits, $y_{n_i} \to c$, too.

Now $f$ is not continuous at $x = c$, since for any $\delta > 0$, can find $\frac{1}{n} < \delta$, and then

$$|x - y| < \frac{1}{n} \text{ and } |f(x) - f(y)| \geq \varepsilon_0. \qed$$

NOTE: here, compactness allows us to infer a global property (uniform continuity) from a local property (continuity).

Required: Ex #13.1.1, 13.2.1, 13.3.1, 13.4.1, 13.5.6, Prob 13-1, 13-3, 13-6

Recommended: Ex #13.1.2, 13.3.3, 13.5.2, Prob 13-2, 13-4

1. In addition to #13-6, show that $\frac{1}{x}$ is uniformly continuous on $(\frac{1}{n}, \infty)$, for $n \in \mathbb{N}$. 
Chapter 14

Differentiation: Local Properties

14.1 The derivative

Definition 14.1.1. Let \( f \) be defined for \( x \approx a \). The derivative of \( f \) at \( a \) is the limit

\[
f'(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h},
\]

provided it exists. In this case, we say \( f \) is differentiable at \( a \).

Note: let \( x = a + h \) to see the equivalence of these two formulas.

\[ \begin{array}{ll}
\text{derivative} & f'(a) \text{ slope of tangent line to graph at } a \\
\text{difference quotient} & \frac{f(x) - f(a)}{x - a} \text{ slope of secant through } a \\
\frac{dy}{dx} & \text{ROC of } y \text{ w/r } x \text{ when } x = a \\
& \frac{\Delta y}{\Delta x} \text{ avg ROC of } y \text{ over } [a, a + \Delta x]
\end{array} \]

\( f'(a) \) is a number.

Now think of \( f'(a) \) as the value of a function \( f'(x) \) when evaluated at \( x = a \).

Definition 14.1.2. \( f \) is differentiable on an open interval \( I \) if the limit \( f'(x) \) exists for every \( x \in I \). The function so defined is \( f' \), the derivative of \( f \).

Definition 14.1.3. If \( f' \) is continuous on \( I \) then we say \( f \) is continuously differentiable and write \( f \) is \( C^1 \) or \( f \in C^1(I) \). Similarly, if \( f'' := (f')' \) is continuous, we say \( f \in C^2 \), etc.

\[ ^1 \text{April 18, 2007} \]
Another way to say $f$ is continuous is $f \in C^0$.

**Definition 14.1.4.** The right-hand derivative of $f$ is

$$f'(a+) := \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a},$$

and the left-hand derivative of $f$ is

$$f'(a-) := \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a},$$

whenever the limits exist.

**Example 14.1.1.** $f(x) = |x|$ is continuous everywhere, but not differentiable at 0. Reason: $f'(0-) \neq f'(0+)$.

**Example 14.1.2.** $f(x) = \sqrt{x}$ is differentiable.

Domain is only $[0, \infty)$, so understood that “differentiable” means $f'(x)$ exists on $(0, \infty)$ and $f'(0+)$ exists.

**Theorem 14.1.5.** $f$ is differentiable at $x_0$ implies $f$ is continuous at $x_0$.

**Proof.** $f(t) - f(x) = \frac{f(t) - f(x)}{t - x} \cdot (t - x) \to f'(x) \cdot 0 = 0$. \qed

In fact, $f$ must be Lipschitz.

**Definition 14.1.6.** $f$ is differentiable at $x_0$ iff

$$\forall \varepsilon > 0, \exists \delta > 0, \quad 0 < |x - x_0| < \delta \quad \Rightarrow \quad \left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| < \varepsilon,$$

in which case $f'(x_0) = L$.

Since $x \neq x_0$, multiply the inequality to obtain

**Definition 14.1.7.** $f$ is differentiable at $x_0$ iff

$$\forall m > 0, \exists n > 0, \quad 0 < |x - x_0| < \frac{1}{n} \quad \Rightarrow \quad |f(x) - (f(x_0) + f'(x_0)(x - x_0))| < \frac{|x - x_0|}{m}.$$

So $f \approx g$ for $g(x) = f(x_0) + f'(x_0)(x - x_0)$.

**Definition 14.1.8.** If $f_n(x) \xrightarrow{g(x)} \infty$, then $f$ “blows up” faster than $g$. If $f_n(x) \xrightarrow{g(x)} 0$, then $g$ “blows up” faster than $f$; write $f(x) = o(g(x))$.

Write $f(x) = O(g(x))$ iff $\frac{f(x)}{g(x)} \leq b < \infty$ as $x \to x_0$. 

Then “$f$ is differentiable” means $f(x) - g(x) = o(|x - x_0|)$, where $g$ is the affine approximation to $f$: $g(x) = f(x_0)+f'(x_0)(x-x_0)$.

### 14.2 Differentiation formulas

Notation for derivatives: read it in the book!

Meanwhile, it is useful to think of differentiation as an operator $D$, i.e., a function on functions:

$$D : \{\text{functions}\} \to \{\text{functions}\}, \quad D : f \mapsto f', \quad D(f) = f'. $$

Solving a differential equation, like

$$3f''(x) - x^2f'(x) - 4f(x) = g(x)$$

then amounts to inverting an operator:

$$(3D^2 - x^2D - 4)f = g$$

$$f = (3D^2 - x^2D - 4)^{-1}g.$$ 

So it will be helpful to study properties of the operator $(3D^2 - x^2D - 4)$. This amounts to studying properties of the derivative.

**Theorem 14.2.1** (Differentiation algebra). Let $f, g$ be differentiable on an interval $I$ on which $g \neq 0$. Then

(i) [Linearity] $D(af + bg) = aD(f) + bD(g)$, $\forall a,b \in \mathbb{R}$.

(ii) [Product Rule] $D(fg) = D(f)g + fD(g)$.

(iii) [Quotient Rule] $D(f/g) = (D(f)g - fD(g))/g^2$.

*Proof of (i).* HW: follows immediately from limit defn of derivatives.

*Proof of (ii).* Let $h = fg$ so that

$$h(t) - h(x) = f(t)g(t) - f(t)g(x) + f(t)g(x) - f(x)g(x)$$
\[
\frac{h(t) - h(x)}{t - x} = \frac{f(t)[g(t) - g(x)]}{t - x} + \frac{g(x)[f(t) - f(x)]}{t - x}.
\]

Proof of (iii). HW: Let \( h = f/g \) so that
\[
\frac{h(t) - h(x)}{t - x} = \frac{1}{g(t)g(x)} \left[ g(x)f(t) - f(x)g(t) - f(x)g(x) \right].
\]

Theorem 14.2.2 (Chain Rule). If \( f \) is differentiable at \( x \) and \( g \) is differentiable at \( f(x) \), then \( g \circ f \) is differentiable at \( x \) with \( (g \circ f)'(x) = g'(f(x))f'(x) \).

Proof. Let \( y = f(x) \) and \( s = f(t) \) and define \( h(t) = g(f(t)) \). By the defn of derivative,
\[
\begin{align*}
  f(t) - f(x) &= (t - x)[f'(x) + o(1)] \quad \text{as } t \to x \\
  g(s) - g(y) &= (s - y)[g'(y) + o(1)] \quad \text{as } s \to y.
\end{align*}
\]
So:
\[
\begin{align*}
  h(t) - h(x) &= g(f(t)) - g(f(x)) \\
  &= (s - y)[g'(y) + o(1)] \\
  &= [f(t) - f(x)][g'(f(x)) + o(1)] \\
  &= [(t - x)[f'(x) + o(1)]][g'(f(x)) + o(1)] \\
  \frac{h(t) - h(x)}{t - x} &= [f'(x) + o(1)][g'(f(x)) + o(1)].
\end{align*}
\]
Note: \( s \to y \) as \( t \to x \), by continuity of \( f \).

Theorem 14.2.3 (Baby Inverse Fn Thm). \( f : (a, b) \to (c, d) \) is \( C^1 \) and \( f'(x) > 0 \) on \( (a, b) \). Then \( f \) is invertible and is \( C^1 \) with \( (f^{-1})'(y) = 1/f'(x) \), if \( y = f(x) \).

Proof. Use the chain rule to differentiate both sides of the identity \( f^{-1}(f(x)) = x \).

14.3 Derivatives and local properties

Already: differentiability implies continuity.

Theorem 14.3.1 (Continuity of derivatives). \( f \) is differentiable on \( (a, b) \). Then \( f' \) assumes every value between \( f'(s) \) and \( f'(t) \), for \( a < s < t < b \).
Proof. (SKIP?)
Let \( \lambda \in (f'(s), f'(t)) \), and define \( g(t) = f(t) - \lambda t \) so that

\[
g'(a) = f(a) - \lambda < 0 \quad \Rightarrow \quad g(t_1) < g(a) \text{ for some } a < t_1 < b, \text{ and}
\]
\[
g'(b) = f(b) - \lambda > 0 \quad \Rightarrow \quad g(t_2) < g(b) \text{ for some } a < t_2 < b.
\]

Then \( g \) attains its min on \([a, b]\) at some point \( x \) such that \( a < x < b \). It follows that \( g'(x) = 0 \), hence \( f'(x) = \lambda \).

\[\]

Corollary 14.3.2. If \( f \) is differentiable on \([a, b]\), then \( f' \) cannot have any simple or jump discontinuities on \([a, b]\).

Theorem 14.3.3. Suppose \( f \in C^1(I) \), \( I \) open.

1. \( f \) is locally increasing on \( I \) \( \iff \) \( f'(x) \geq 0 \).

2. \( f \) is locally decreasing on \( I \) \( \iff \) \( f'(x) \leq 0 \).

Proof. Let \( a \in I \). Then \( f \) locally increasing at \( a \) means that for \( 0 < x - a < \delta \),

\[
f(x) \geq f(a) \iff \frac{f(x) - f(a)}{x - a} \geq 0
\]
\[
\iff f'(a+) \geq 0.
\]

Since \( f' \) exists by hyp, we have \( f'(a) = f'(a+) \).

For part (ii), apply part (i) to the increasing function \(-f(x)\) to obtain \(-f'(x) \geq 0\).

Definition 14.3.4. Let \( f \) be defined on the open interval \( I \). Then

1. \( c \in I \) is a local maximizer (local maximum point) of \( f \) iff \( f(c) \geq f(x) \) for \( x \approx_c c \).

2. \( c \in I \) is a local minimizer (local minimum point) of \( f \) iff \( f(c) \leq f(x) \) for \( x \approx_c c \).

NOTE: a local extremizer must be an interior point; endpoints don’t count!

Theorem 14.3.5. Suppose \( f \in C^1(I) \), \( I \) open. If \( a \in I \) is a local extremizer, then \( f'(a) = 0 \).

Proof. Choose \( \delta \) such that \( a < a - \delta < a < a + \delta < b \). Then for \( a - \delta < t < a \), we have

\[
\frac{f(t) - f(a)}{t - a} \geq 0.
\]

Letting \( t \to a \), get \( f'(a) \geq 0 \). Similarly for the other inequality.
Definition 14.3.6. $c$ is a critical point of $f$ iff $f'(x) = 0$.

Theorem 14.3.7. If $c$ is a local extremizer of $f$, then $c$ is a critical point of $f$.

We would like the converse, but it isn’t true: $f(x) = x^3$ has a critical point at $x = 0$, which is not a local extremizer. When is a critical point an extremizer?

Theorem 14.3.8 (Isolation principle). If we can find an open $I$ on which $f \in C^1$ and

(i) $I$ contains exactly one critical point $c$ of $f$, and

(ii) the Max Thm indicates that $f$ has an extremizer on $I$,

then $f$ has a unique extremum on $I$ at $c$.

Example 14.3.1. Let $f(x) = xe^{-x}$. Find & classify the extremizers.

Solution. The function is differentiable on $I = (-\infty, \infty)$. Since

$$f'(x) = e^{-x}(1-x),$$

we have

$$f'(x) = 0 \iff x = 1.$$ (SKETCH)

From the sketch, see $f(0) = 0$, $f(1) = \frac{1}{e}$, $f(2) < \frac{1}{e}$.

Hence $f$ has a local max inside of $[0, 2]$.

Hence $f$ has a unique local max at $x = 1$. \hfill \Box

Required: Ex #14.1.2, 14.2.4, 14.3.2, Prob 14-2, 14-3, 14-4

Recommended: Ex #14.1.4, 14.1.5, 14.2.5, Prob 14-5, 14-6

1. Prove from the limit definition that $D(x^n) = nx^{n-1}$ for $n = 0, 1, 2, 3, \ldots$. Hint: multiply out $(x - a) \sum_{k=0}^{n} x^k a^{n-k}$ and avoid induction.
Chapter 15

Differentiation: Global Properties

15.1 The Mean-Value Theorem

Theorem 15.1.1 (Rolle’s). If \( f \) is continuous on \([a, b]\) and differentiable on the interior and \( f(a) = f(b) \), then \( \exists x \in (a, b) \) such that \( f'(x) = 0 \).

Proof. If \( f \) is constant on the interval, we are done, so wlog let \( h(t) > h(a) \) somewhere. Then \( f \) attains its max at some point \( x \in (a, b) \) (since \([a, b]\) is compact), and a prev thm gives \( f'(x) = 0 \). □

Theorem 15.1.2 (Cauchy mean value thm). \( f, g \) are continuous on \([a, b]\) and differentiable on the interior. Then \( \exists x \in (a, b) \) for which

\[
[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).
\]

Proof. For \( a \leq t \leq b \), define \( h(t) := [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t), \) so that \( h \) is continuous on \([a, b]\) and differentiable on the interior and

\[
h(a) = f(b)g(a) - f(a)g(b) = h(b).
\]

By Rolle’s Thm, get \( h'(x) = 0 \) for some \( x \). □

Interp: \( x = f(s) \), \( y = g(t) \) (SKETCH).

\(^1\)April 18, 2007
Corollary 15.1.3 ("The" Mean Value Thm). If \( f \) continuous on \([a, b]\) and differentiable on the interior, then \( \exists x \in (a, b) \) for which

\[
f(b) - f(a) = (b - a)f'(x).
\]

Proof. Let \( g(x) = x \) in prev. \( \square \)

Recall: “\( f \) is differentiable” means

\[
f(t) \approx f(s) + f'(s)(t-s) = g(t),
\]

in the sense that \( f(t) - g(t) = o(|t-s|) \). The MVT says that there is some point \( x \) nearby (i.e., \( x \in (s, t) \)) for which this becomes an actual equality:

\[
f(t) = f(s) + f'(x)(t-s).
\]

Theorem 15.1.4. \( f \) is differentiable on \((a, b)\).

1. If \( f'(x) \geq 0, \forall x \in (a, b) \), then \( f \) is increasing.

2. If \( f'(x) \leq 0, \forall x \in (a, b) \), then \( f \) is decreasing.

3. If \( f'(x) = 0, \forall x \in (a, b) \), then \( f \) is constant.

Proof. These can all be read off from the equation

\[
f(t) - f(s) = (t-s)f'(x),
\]

which is always valid for some \( x \in (s, t) \). \( \square \)

15.4 L’Hôpital’s rule for indeterminate forms

We have seen that the differentiability of \( f(x) \) at \( x = c \) is equivalent to the fact that \( f(c) + f'(c)(x - c) \) for \( x \approx c \), or more precisely,

\[
f(x) = f(c) + f'(c)(x - c) + o(|x-c|).
\]

We use this to give a conceptually clear proof of L’Hôpital’s rule.
15.4 L'Hôpital's rule for indeterminate forms

Theorem 15.4.1. Let \( f, g \in C^1(A) \) where \( A \) is a neighbourhood of \( a \). If \( \lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0 \) and \( g'(x) \neq 0 \) for \( x \in A \setminus \{a\} \), then if the RHS limit exists,

\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.
\]

Proof. We only need to take the limit of

\[
\frac{f(x)}{g(x)} = \frac{f(c) + f'(c)(x - c) + o(|x - c|)}{g(c) + g'(c)(x - c) + o(|x - c|)} \quad \text{defn of diff}
\]

\[
= \frac{f'(c)(x - c) + o(|x - c|)}{g'(c)(x - c) + o(|x - c|)} \quad \text{continuity}
\]

\[
= \frac{f'(c) + o(1)}{g'(c) + o(1)} \quad \text{cancellation.} \]

Theorem 15.4.2. Let \( f, g \in C^1(A) \) where \( A \) is a neighbourhood of \( a \). If \( \lim_{x \to a} f(x) = \lim_{x \to a} g(x) = \infty \) and \( g'(x) \neq 0 \) for \( x > 1 \), then if the RHS limit exists,

\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.
\]

Proof. Fix \( \varepsilon > 0 \) and wlog, let \( g > 0 \) for \( x > 1 \). When the RHS limit finite, then \( \frac{f'(x)}{g'(x)} \xrightarrow[x \to a]{} L \), and for our \( \varepsilon \) we can choose \( \delta \) such that

\[
x \approx \delta a \implies \frac{f'(x)}{g'(x)} \approx_{\varepsilon} L.
\]

In fact, for for \( x, t \approx \delta a \) and some \( c \) between \( x \) and \( t \),

\[
\frac{f'(c)}{g'(c)} = \frac{f(x) - f(t)}{g(x) - g(t)} \approx_{\varepsilon} L
\]

\[
L - \varepsilon < \frac{f(x) - f(t)}{g(x) - g(t)} < L + \varepsilon. \quad (*)
\]

Hold \( t \) fixed for now. For small enough \( \delta_1 \),

\[
x \approx_{\delta_1} a \implies g(x) > 0, \ g(x) > g(t).
\]

We need to isolate \( f(x)/g(x) \) and get a bound for it. To this end, note that \( \frac{g(x) - g(t)}{g(x)} > 0 \), so we can multiply the inequality \((*)\) above to obtain

\[
(L - \varepsilon) \frac{g(x) - g(t)}{g(x)} < \frac{f(x) - f(t)}{g(x)} < (L + \varepsilon) \frac{g(x) - g(t)}{g(x)}
\]

\[
\frac{(L - \varepsilon)g(x) - g(t)}{g(x)} < \frac{f(x) - f(t)}{g(x)} < \frac{(L + \varepsilon)g(x) - g(t)}{g(x)}
\]
\[
(L - \varepsilon) \left(1 - \frac{g(t)}{g(x)}\right) + \frac{f(t)}{g(x)} < \frac{f(x)}{g(x)} < (L + \varepsilon) \left(1 - \frac{g(t)}{g(x)}\right) + \frac{f(t)}{g(x)}
\]

\[
L - \varepsilon - \frac{(L - \varepsilon)g(t) - f(t)}{g(x)} < \frac{f(x)}{g(x)} < L + \varepsilon - \frac{(L + \varepsilon)g(t) - f(t)}{g(x)}
\]

Choose \(\delta_2 \leq \delta_1\) so that

\[
x \approx \delta_2 a \quad \Rightarrow \quad \frac{(L - \varepsilon)g(t) - f(t)}{g(x)}, \quad \frac{(L + \varepsilon)g(t) - f(t)}{g(x)} < \varepsilon.
\]

Then for \(x \approx \delta_2 a\), we have

\[
L - 2\varepsilon < \frac{f(x)}{g(x)} < L + 2\varepsilon.
\]

When the RHS is \(\infty\), the proof is similar to the prev. \(\Box\)

**Required:** Ex #15.1.3, 15.2.1, 15.2.2, 15.4.4 Prob 15-1, 15-2

**Recommended:** Ex #15.1.1(a), 15.1.4, Prob 15-3, 15-5

1. Let \(f : \mathbb{R} \to \mathbb{R}\) satisfy \(|f(x) - f(y)| \leq (x - y)^2\) for all \(x, y \in \mathbb{R}\). Prove \(f\) is constant.
Chapter 16

Linearization and Convexity

16.1 Linearization

Recall: “f is differentiable” means

\[ f(t) \approx g(t) = f(s) + f'(s)(t - s), \]

in the sense that \( f(t) - g(t) = o(|t - s|) \). Higher derivatives give better approximation.

**Theorem 16.1.1** (Linearization Error Term). Suppose \( f \in C^2(I) \) and \( a \in I \). For each \( x \in I \), there is a point \( c \) between \( x \) and \( a \) for which

\[ f(x) = f(a) + f'(a)(x - a) + \frac{f''(c)}{2}(x - a)^2. \]

This is a special case of Taylor’s Thm \((n = 2)\), which we will prove shortly. We will see that a differentiable function can be approximated by its derivatives, and that sometimes it can even be written as a power series. In either case, a polynomial makes a decent local approximation.

Note: this case is often used for approximation/optimization with multivariable systems, since it makes sense for matrices:

\[ f(x) \approx f(c) + \nabla f(c)(x - c) + \frac{1}{2}(x - c)^T H(f)(x - c). \]

\(^1\)April 18, 2007
16.2 Applications to convexity

**Theorem 16.2.1** (2\textsuperscript{nd} Deriv Test). Suppose \( f \in C^2(I), a \in I, \) and \( f'(a) = 0. \)

\[
\begin{align*}
  f''(a) > 0 & \implies f(x) \text{ has a strict local min at } a \\
  f''(a) < 0 & \implies f(x) \text{ has a strict local max at } a.
\end{align*}
\]

**Proof.** Applying the Linearization Error Term, we get a \( c \) for which

\[
f(x) = f(a) + f'(a)(x - a) + \frac{f''(c)}{2}(x - a)^2
\]

If \( f''(a) > 0, \) then \( f''(x) > 0 \) for \( x \approx a \) and hence \( f''(c) > 0 \) \( \implies f(x) > f(a). \)

At a critical point, the graph of \( f \) has a horizontal tangent, and \( f'' \) tells if the graph lies above or below the tangent.

At other points, the tangent may not be horizontal, but one can still use \( f'' \) to see if the graph is above or below the tangent.

**Definition 16.2.2.** \( f \) is convex on \( I \) iff for \( 0 \leq t \leq 1, \) we have

\[
f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y), \quad \forall x, y.
\]

Note that any point between \( x \) and \( y \) is written \((1 - t)x + ty, \) where \( 0 \leq t \leq 1. \) So this inequality just states that the function value in the interior of \((x, y)\) cannot be greater than the corresponding point on the secant line from \((x, f(x))\) to \((y, f(y)).\)

**Theorem 16.2.3.** \( f \) is convex on \( I \) \( \iff \) for every \( x \in [a, b] \subseteq I, \)

\[
\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a}.
\]

This theorem just states that for a convex function, secant lines from a fixed starting point have increasing slopes.

**Theorem 16.2.4.** Let \( f \in C^1(I). \) Then \( f \) is convex \( \iff \) \( f(x) \geq f(c) + f'(c)(x - c). \)

**Proof.** HW 16.2.2.

**Theorem 16.2.5** (2\textsuperscript{nd} Deriv Test). If \( f \in C^2(I), \) then \( f'' \geq 0 \) \( \implies f \) is convex on \( I. \)
Proof. From prev, we have
\[ f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2, \]
so \( f''(c) \geq 0 \) gives the inequality.

**Theorem 16.2.6** (1st Deriv Test). Let \( f \in C^1(I) \). Then \( f \) convex on \( I \) \( \iff \) \( f'(x) \) is increasing on \( I \).

**Proof.** \((\Rightarrow)\) Pick \( a < b \) in \( I \). Then
\[
\begin{align*}
    f(b) &\geq f(a) + f'(a)(b-a) \iff f'(a) \leq \frac{f(b) - f(a)}{b-a} \\
    f(a) &\leq f(b) + f'(a)(b-a) \iff f'(b) \geq \frac{f(b) - f(a)}{b-a}.
\end{align*}
\]

**Definition 16.2.7.** Let \( f \) be a function on \((a, b)\) and pick \( c \in (a, b)\). The line \( y(x) = m(x-c) + f(c) \) is called a supporting line at \( c \) iff it always lies below the graph of \( f \):
\[ f(x) \geq m(x-c) + f(c). \]

**Required:** Ex #16.1(a), 16.1.2, 16.2.2, 16.2.4 Prob 16-1, 16-2

**Recommended:** Ex #16.2.1, Prob 16

1. Prove that for \( f \in C^1 \), \( f \) is convex iff \( f(x) \geq f(c) + f'(c)(x-c) \).

2. Prove that \( f \) is convex iff every point \( x \) in the domain of \( f \) has a supporting line.
Chapter 17

Taylor Approximation

17.1 Taylor polynomials

Definition 17.1.1. If \( f \in C^n \) and \( f^{(n)} = (f^{(n-1)})' \), the \( n^{th} \) Taylor polynomial of \( f \) at \( a \) is

\[
T_n(a, x) = f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + \cdots + \frac{1}{n!} f^{(n)}(a)(x - a)^n.
\]

Definition 17.1.2. Two functions \( f, g \in C^n(I) \) have \( n^{th} \)-order agreement at \( a \) iff \( f^{(k)}(a) = g^{(k)}(a) \) for \( k = 0, 1, \ldots, n \).

Theorem 17.1.3. If \( f^{(n)} \) exists at \( a \), then \( T_n(a, x) \) is the unique polynomial of degree \( n \) in powers of \( (x - a) \) having \( n^{th} \)-order agreement with \( f(x) \) at \( a \).

Proof. Suppose \( p(x) = c_0 + c_1(x - a) + \cdots + c_n(x - a)^n \) is a polynomial in \( (x - a) \). After \( k \) differentiations, the terms \( c_0, \ldots, c_{k-1} \) vanish:

\[
p^{(k)}(x) = k!c_k + \text{(terms with } (x - a) \text{ as a factor)}.
\]

So \( p^{(k)}(a) = k!c_k \). If \( f(x) \) and \( p(x) \) have \( n^{th} \)-order agreement,

\[
f^{(k)}(a) = k!c_k \implies c_k = \frac{f^{(k)}(a)}{k!}, \forall k = 1, \ldots, n.
\]

1\footnote{April 18, 2007}
17.2 Taylor’s theorem with Lagrange remainder

Theorem 17.2.1 (Taylor’s Thm). Suppose \( f \in C^{n+1}(I) \) for some open interval \( I \) containing \( a \) and \( x \). Then for some \( c \) between \( a \) and \( x \),

\[
f(x) = f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2 + \cdots + \frac{1}{n!} f^{(n)}(a)(x-a)^n + R_n(x),
\]

\[
R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-a)^{n+1}.
\]

**NOTE:** \( c \) depends on \( x \), so RHS of \( f \) is not a polynomial; \( f^{(n+1)}(c) \) is not a constant.

Proof. We show the theorem holds at \( x = b \). Let \( P \) be defined by

\[
P(x) = T_n(a,x) + C(x-a)^{n+1},
\]

where \( C \) is chosen so that \( f(b) = P(b) \), i.e., \( C = \frac{f(b)-T_n(a,b)}{(b-a)^n} \). Let

\[
g(x) = f(x) - P(x)
\]

so that \( g(a) = g(b) = 0 \). Must show \( f^{(n+1)}(c) = (n+1)!C \) for some \( c \in (a,b) \). Since \( g^{(n+1)}(x) = f^{(n+1)}(x) - (n+1)!C \), suffices to find a zero of \( g^{(n+1)} \) on \( (a,b) \).

Since \( T_n^{(k)}(a) = f^{(k)} \) for \( k = 0, \ldots, n \), we have

\[
g(a) = g'(a) = \cdots = g^{(n+1)}(x) = 0.
\]

Applying the MVT \( n \) times to the derivatives of \( g \),

\[
g(x) = g'(x_1)(x-a)
\]

\[
g'(x_1) = g''(x_2)(x_2-x_1)
\]

\[
\vdots
\]

\[
g^{(n)}(x_n) = g'(x_{n+1})(x_n-a),
\]

each time using \( g^{(k)}(x_k) = 0 \) and the fact that \( f \in C^n \implies g \in C^n \). Then \( g^{(n+1)}(x_n) = 0 \) for \( x_n \in (a,b) \).

Corollary 17.2.2. \( f = T_n + o(|x-a|^n) \) as \( x \to a \). 

\(\square\)
17.3 Estimating error in Taylor’s approximation

The expression for the error term in \( f(x) \approx T_n(x) \) is exact:

\[
R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} (x - a)^{n+1}
\]

but who knows where \( c \in (x, a) \) really is? Use \( R_n \) to estimate the error instead.

**Example 17.3.1.** The error in the third-order approximation to \( e^x \) at 0 is \( R_3(x) = \frac{x^4}{4!} \).

17.4 Taylor series

**Definition 17.4.1.** The Taylor series at 0 of a function \( f \in C^n(I) \) is

\[
\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \]

**Definition 17.4.2.** \( f \) is analytic at 0 if \( f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \) on \((-R, R)\), where \( R \) is the radius of convergence of the Taylor series.

Since \( f(x) = T_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1} \) for some \( c \in (-x, x) \), this is equivalent to requiring

\[
\frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1} \to 0.
\]

It can be very hard to show that the Taylor series actually converges to the function it’s supposed to converge to. In the complex case, things are easier.

\[
f \text{ is analytic } \iff f \in C^\infty \iff f \text{ satisfies C-R eqns.}
\]

Elementary functions can also sometimes be checked explicitly.

**Example 17.4.1.** \( e^x = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!} \).

**Solution.** Given a (fixed) \( x \in \mathbb{R} \), must show the remainder is small for \( n \gg 1 \). Choose \( N \) such that \(|x| < \frac{N}{2}\). Then we have

\[
|R_n(x)| = \left| \frac{e^c}{n!} x^n \right| \leq e^{|x|} \frac{|x|^n}{n!} \to e^{|x|} \cdot 0 = 0.
\]
by Exercise 3.4.2.

Required: Ex #17.1.1, 17.2.1, 17.3.3, 17.4.1 Prob 17-1
Recommended: Ex #17.3.4 Prob 17
Chapter 18

Integrability

18.1 Introduction. Partitions.

Let \( f(x) \) be a function defined on \([a, b]\). We want to define its integral \( \int_a^b f(x) \, dx \).

Definition 18.1.1. A partition \( P \) of the interval \([a, b]\) is a finite set of points \( \{a = x_0, x_1, x_2, \ldots, x_n = b\} \), where \( x_i < x_{i+1} \).

The mesh of a partition is the size of the largest subinterval:

\[
|P| := \max_i (x_i - x_{i-1}).
\]

An \( n \)-partition is one containing \( n \) subintervals.

18.2 Integrability

Definition 18.2.1. On each subinterval \([x_i, x_{i+1}]\), of the partition \( P \), define

\[
M_i = \sup \{ f(x) : x_{i-1} \leq x \leq x_i \}
\]

\[
m_i = \inf \{ f(x) : x_{i-1} \leq x \leq x_i \}
\]
$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1})$

$L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1})$

$U(f, P)$ is the upper sum of $f$ on the partition $P$, and $L(f, P)$ is the lower sum of $f$ on the partition $P$.

**Definition 18.2.2.** If $\sup_P L(f, P) = \inf_P U(f, P)$, then the integral of $f$ on $[a, b]$ is defined to be the common value, denoted $\int_a^b f(x) \, dx$. We say $f$ is (Riemann-) integrable on $[a, b]$ and write $f \in \mathcal{R}[a, b]$.

**Definition 18.2.3.** The partition $P'$ is a refinement of $P$ iff $P \subseteq P'$.

Note that adding points to the partition has two effects:

1. the lengths of the subintervals decreases, and
2. $L(f, P)$ increases and $U(f, P)$ decreases.

**Lemma 18.2.4** (Upper and lower sums). For $P \subseteq P'$, $L(f, P') \geq L(f, P)$ and $U(f, P) \leq U(f, P')$.

**Proof.** Do the case where $P' = P \cup \{y\}$ first; then the sums only change on the one subinterval. General case by repetition. \qed

**Corollary 18.2.5.** For $f : [a, b] \to \mathbb{R}$, any lower sum is less than any upper sum.

**Proof.** Let $P_1, P_2$ be any two partitions of the interval. Let $P'$ be the common refinement. Then $L(f, P_1) \leq L(f, P') \leq U(f, P') \leq U(f, P_2)$. \qed

Suppose we have a nested sequence of partitions $P_1 \subseteq P_2 \subseteq \ldots$. Then $\{L(f, P_i)\}$ and $\{U(f, P_i)\}$ are monotonic sequences, each bounded by any element of the other. So both converge. $f$ is integrable when they converge to the same value.

**Definition 18.2.6.** The oscillation of $f$ is

$$Osc(f, P) = U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1})$$
We now show that \( \int_a^b f(x) \, dx \) exists iff \( \text{Osc}(f, P_n) \to 0 \), for some nested sequence of partitions \( \{P_n\} \).

**Theorem 18.2.7** (Oscillation). \( f \in \mathcal{R}[a, b] \) iff \( \forall \varepsilon > 0, \exists P \text{ such that } \text{Osc}(f, P) < \varepsilon \).

**Proof.** (\( \Rightarrow \)) Let \( f \in \mathcal{R} \) and fix \( \varepsilon > 0 \). Then there are partitions \( P_1, P_2 \) such that

\[
U(f, P_1) - \int f < \varepsilon \quad \text{and} \quad \int f - L(f, P_2) < \varepsilon.
\]

Let \( P = P_1 \cup P_2 \) be the common refinement. Then

\[
U(f, P) \leq U(f, P_1) < \int f + \varepsilon < L(f, P_2) + 2\varepsilon \leq L(f, P) + 2\varepsilon,
\]

so that \( U(f, P) - L(f, P) < 2\varepsilon \).

(\( \Leftarrow \)) Apply the inequality \( U(f, P) - L(f, P) < \varepsilon \) to

\[
L(f, P) \leq \sup P L(f, P) \leq \inf P U(f, P) \leq U(f, P)
\]

to get \( 0 < \inf_P U(f, P) - \sup_P L(f, P) < \varepsilon \). Since this is true for any \( \varepsilon > 0 \), done.

**Corollary 18.2.8.** If \( f \in \mathcal{R}[a, b] \), then for any sequence \( \{P_i\} \) of partitions of \([a, b]\) with \( |P_i| \to 0 \),

\[
\lim_{i \to \infty} U(f, P_i) = \int_a^b f(x) \, dx \quad \text{and} \quad \lim_{i \to \infty} L(f, P_i) = \int_a^b f(x) \, dx.
\]

**Proof.** Fix \( \varepsilon > 0 \). Since \( f \) is integrable, the Oscillation theorem gives \( N \) such that

\[
i \geq N \implies U(f, P_i) - L(f, P_i) < \varepsilon.
\]

Since we always have \( U(f, P_i) < \int f < L(f, P_i) \), it is clear that \( U(f, P_i) - \int f < \varepsilon \) and \( \int f - L(f, P_i) < \varepsilon \).

Previously, we saw that an upper bound \( x \) of \( \{a_n\} \) is actually the supremum iff given \( \varepsilon > 0 \), we can always find an \( a_n \) with \( |a_n - x| < \varepsilon \). If we know that \( \{a_n\} \) is monotonic, then \( a_n \to x \) iff we can always find an \( a_n \) with \( |a_n - x| < \varepsilon \).

In the current situation, the definition of \( \int_a^b f(x) \, dx \) in terms of sups & infs means that \( f \in \mathcal{R}[a, b] \) iff given \( \varepsilon > 0 \), we can find a partition \( P \) for which \( \text{Osc}(f, P) = U(f, P) - \int_a^b f(x) \, dx \).
$L(f, p) < \varepsilon$. Since a nested sequence of partitions gives monotonic sequences $\{U(f, P_n)\}$ and $\{L(f, P_n)\}$, this is equivalent to the condition

$$\lim_{n \to \infty} U(f, P_n) = \lim_{n \to \infty} L(f, P_n).$$

### 18.3 Integrability of monotone or continuous $f$

Next, two sufficient conditions for integrability: continuity and monotonicity.

**Theorem 18.3.1.** $f$ monotonic on $[a, b] \implies f \in \mathcal{R}[a, b]$.

**Proof.** Suppose $f$ is monotonic increasing (the proof is analogous in the other case) so that on any partition,

$$M_i = f(x_i), \quad m_i = f(x_{i-1}), \quad i = 1, \ldots, n.$$  

For any $n$, choose a partition by dividing $[a, b]$ into $n$ equal subintervals; then $x_i - x_{i-1} = (b - a)/n$ for each $i = 1, \ldots, n$. Then can get

$$Osc(f, P) = U(f, P) - L(f, P) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \frac{b - a}{n}$$

$$= \frac{b - a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))$$

$$\leq \frac{b - a}{n} \cdot [f(b) - f(a)]$$

$$< \varepsilon$$

for large enough $n$. Then apply Oscillation thm. \qed

**Theorem 18.3.2.** If $f$ is continuous on $[a, b]$ then $f \in \mathcal{R}[a, b]$.

**Proof.** Fix $\varepsilon > 0$ and choose $\gamma > 0$ such that $0 < (b - a)\gamma < \varepsilon$. Since $f$ is uniformly continuous on $[a, b]$, there is $\delta > 0$ such that $|f(x) - f(t)| < \gamma$ whenever

$$|x - t| < \delta, \quad x, t \in [a, b].$$
18.4 Basic properties of integrable functions

If $P$ is any partition of $[a, b]$ such that $|x_i - x_{i-1}| < \delta, \forall i$, then

$$|x - t| < \delta \implies |f(x) - f(t)| < \gamma \implies M_i - m_i < \gamma.$$ 

Therefore,

$$\text{Osc}(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \gamma \sum_{i=1}^{n} (x_i - x_{i-1}) = \gamma|b - a| < \varepsilon.$$

By Oscillation thm, $f \in \mathcal{R}$.

18.4 Basic properties of integrable functions

**Theorem 18.4.1.** Let $f \in \mathcal{R}[a, b]$ and suppose $m \leq f \leq M$. If $\varphi$ is continuous on $[m, M]$ and $h(x) := \varphi(f(x))$ for $x \in [a, b]$, then $h \in \mathcal{R}[a, b]$.

**Proof.** Fix $\varepsilon > 0$. Since $\varphi$ is uniformly continuous on $[m, M]$, find $0 < \delta < \varepsilon$ such that

$$|s - t| < \delta \implies |\varphi(s) - \varphi(t)| < \varepsilon, \quad s, t \in [m, m].$$

Since $f \in \mathcal{R}$, find $P = \{x_0, \ldots, x_n\}$ such that

$$\text{Osc}(f, P) < \delta^2.$$ 

$M_i, m_i$ are extrema of $f$, $M'_i, m'_i$ are for $h$. Subdivide the set of indices $\{1, \ldots, n\}$ into two classes:

$$i \in A \iff M_i - m_i < \delta,$$

$$i \in B \iff M_i - m_i \geq \delta.$$ 

For $i \in A$, have $M'_i - m'_i \leq \varepsilon$ by choice of $\delta$. For $i \in B$, have $M'_i - m'_i \leq 2\sup_{m \leq t \leq M} |\varphi(t)|$. 

By prev bound of $\delta^2$,

$$\delta \sum_{i \in B} (x_i - x_{i-1}) \leq \sum_{i \in B} (M_i - m_i)(x_i - x_{i-1}) < \delta^2$$

$$\sum_{i \in B} (x_i - x_{i-1}) < \delta.$$ 

Then

$$Osc(h, P) = \sum_{i \in A} (M'_i - m'_i)(x_i - x_{i-1}) + \sum_{i \in B} (M'_i - m'_i)(x_i - x_{i-1})$$

$$\leq \varepsilon(b - a) + 2\delta \sup |\varphi(t)|$$

$$< \varepsilon(b - a + 2 \sup |\varphi(t)|).$$

We will skip the rest of this section, as it is essentially repeated in the next chapter.

**Required:** Ex #18.1.2, 18.2.2, 18.3.2, 18.3.4

**Recommended:** Ex #18.3.1, 18.3.3

1.
Chapter 19

The Riemann Integral

§19.1 Refinement of partitions: we have already covered it.

§19.2 Definition of the Riemann integral: we have already covered it. ¹

19.3 Riemann sums

Definition 19.3.1. For $f \in R[a, b]$, a Riemann sum for $f(x)$ over $P$ is any sum of the form

$$S_f(P) = \sum_{i=1}^{n} f(x'_i)(x_i - x_{i-1}), \quad \text{where } x'_i \in (x_{i-1}, x_i).$$

So the upper and lower sums are just special Riemann sums, and we always have

$$m_i \leq f(x'_i) \leq M \implies L(f, P) \leq S_f(P) \leq U(f, P).$$

Theorem 19.3.2 (Riemann sums). Let $f \in R[a, b]$ and suppose $\{P_k\}$ is a sequence of partitions of $[a, b]$ such that $|P_k| \to 0$. Then

$$\lim_{k \to \infty} S_f(P_k) = \int_{a}^{b} f(x) \, dx.$$  

Proof. This is immediate from the Squeeze Thm applied to the inequalities:

$$L(f, P_k) \leq S_f(P_k) \leq U(f, P_k).$$  

¹April 18, 2007
19.4 Basic properties of the integral

Theorem 19.4.1. Let \( f, g \in \mathcal{R}[a,b] \).

(i) (Linearity) \( f + g \in \mathcal{R}[a,b] \) and \( cf \in \mathcal{R}[a,b] \), \( \forall c \in \mathbb{R} \), with \( \int (f + g) \, dx = \int f \, dx + \int g \, dx \)
and \( \int cf \, dx = c \int f \, dx \).

(ii) \( \max\{f, g\}, \min\{f, g\} \in \mathcal{R}[a,b] \).

(iii) \( fg \in \mathcal{R}[a,b] \).

(iv) \( f \leq g \) on \([a,b]\) \implies \( \int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx \).

(v) \( f \in \mathcal{R}[a,b] \) and \( \int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx \).

(vi) If \( |f(x)| \leq M \) on \([a,b]\), then \( \int_a^b f(x) \, dx \leq M(b - a) \).

Proof of (i). Fix \( \varepsilon > 0 \) and suppose \( f = f_1 + f_2 \). Find partitions \( P_1, P_2 \) such that

\[
Osc(f_1, P_1) < \varepsilon, \quad \text{and} \quad Osc(f_2, P_2) < \varepsilon
\]

Let \( P = P_1 \cup P_2 \) be the common refinement. Then these inequalities are still true, and

\[
L(f_1, P) + L(f_2, P) \leq L(f, P) \leq U(f, P) \leq U(f_1, P) + U(f_2, P),
\]

which implies \( Osc(f, P) < 2\varepsilon \). Hence, \( f \in \mathcal{R} \), and

\[
U(f_j, P) < \int f_j(x) \, dx + \varepsilon,
\]

which implies (by the long inequality above) that

\[
\int f \, dx \leq U(f, P) < \int f_1 \, dx + \int f_2 \, dx + 2\varepsilon.
\]

Since \( \varepsilon \) was arbitrary, this gives \( \int f \, dx \leq \int f_1 \, dx + \int f_2 \, dx \). Similarly for the other inequality. For \( cf \), use \( \sum cM_i(x_i - x_{i-1}) = c \sum M_i(x_i - x_{i-1}) \), etc. \( \square \)

Proof of (ii). Let \( h(x) := \max\{f(x), g(x)\} \). Since \( f(x), g(x) \leq h(x) \),

\[
Osc(h, P) \leq Osc(f, P) + Osc(g, P) < \varepsilon.
\]

\( \square \)
19.5 Interval addition property

Proof of (iii). Use $\varphi(t) = t^2$ to get $f^2 \in \mathcal{R}$ (compositions of integrable functions are integrable), then observe $4fg = (f + g)^2 - (f - g)^2$. 

Proof of (iv). HW. Use $g - f \geq 0$ and part (i). 

Proof of (v). Use $\varphi(t) = |t|$ to get $|f| \in \mathcal{R}$, then choose $c = \pm 1$ to make $c \int f \geq 0$ and observe 

$$cf \leq |f| \implies \left| \int f \right| = c \int f = \int cf \leq \int |f|.$$ 

Proof of (vi). HW. Use (iv) and show $\int_a^b 1 \, dx = b - a$. 

19.5 Interval addition property

Theorem 19.5.1 (Interval addition). For $a < c < b$, $\int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx$. 

Proof. Note that you can always refine the partition by adding $c$. Once $c$ is in the partition, taking the sup over values of $f$ on a subinterval of $[a, c]$ is completely independent of whatever $f$ does on subintervals of $[c, b]$. 

Definition 19.5.2. We define 

$$\int_a^a f(x) \, dx = 0, \forall a \quad \text{and} \quad \int_b^a f(x) \, dx = -\int_a^b f(x) \, dx, \forall a, b.$$ 

Corollary 19.5.3. For any $a, b, c$, $\int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx$. 

19.6 Piecewise properties

Definition 19.6.1. A property of $f$ holds piecewise on $[a, b]$ if there is a partition of the interval such that the property holds for $f$ on each open subinterval $(x_{i-1}, x_i)$. 

Example 19.6.1. $f(x) = \tan x$ is piecewise continuous and piecewise monotone on $\mathbb{R}$. The Heaviside function is piecewise constant. (Hence also piecewise continuous and piecewise monotone.) 

Theorem 19.6.2 (Finite discrepancies). Suppose $f, g \in \mathcal{R}[a, b]$ and suppose $f(x) = g(x)$ for all but finitely many values of $x \in [a, b]$. Then $\int_a^b f(x) \, dx = \int_a^b g(x) \, dx$. 

Proof. First, suppose \( f \) differs from \( g \) only for \( x = c \). Then define a sequence of partitions \( \{ P_n \} \) which each contain the subinterval \((p_n, q_n) := (c - \frac{1}{n}, c + \frac{1}{n})\). Since \( f = g \) outside this subinterval, suffices to note that

\[
\lim_{n \to \infty} f(c)(q_n - p_n) = \lim_{n \to \infty} g(c)(q_n - p_n) = 0.
\]

By the Riemann Sum Thm, this case is complete. For the general case, just use induction.

\[\square\]

**Theorem 19.6.3.** If \( f \) is bounded and either piecewise continuous or piecewise monotone on \([a, b]\) (with respect to a partition \( P = \{x_0, x_1, \ldots, x_n\}\)), then \( f \in \mathcal{R}[a, b] \) and

\[
\int_a^b f(x) \, dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) \, dx.
\]

**Proof.** If \( f \) is monotone, suffices to apply the Interval Addition property \( n \) times.

Otherwise, we build a clever partition by working around the “bad points” of \( f \). Away from the bad points, we use continuity to get integrability. At the bad points, we use the bound on \( f \) (and a “horizontal squeeze” to estimate the oscillation.

Let \( \alpha_1 < \alpha_2 < \cdots < \alpha_n \) be the discontinuities of \( f \) on \([a, b]\). Let \( I_k = (\alpha_k - \frac{\epsilon}{2n}, \alpha_k + \frac{\epsilon}{2n}) \) be an open interval about \( \alpha_k \). Then the total measure of these intervals is

\[
\left| \bigcup_{k=1}^n I_k \right| \leq \sum_{k=1}^n |I_k| = \sum_{k=1}^n \frac{\epsilon}{n} = \epsilon.
\]

If we remove the intervals \( I_k \) from \([a, b]\), the remaining set \( K = [a, b] \setminus \bigcup I_k \) is compact, and so \( f \) is uniformly continuous on \( K \). Choose \( \delta > 0 \) such that

\[
|s - t| < \delta \implies |f(s) - f(t)| < \epsilon.
\]

Now let \( P = \{a = x_0, x_1, \ldots, x_m = b\} \) be any partition of \([a, b]\) such that

(a) \( P \) contains the endpoints \( \{u_1, \ldots, u_n, v_1, \ldots, v_n\} \) of the “removed” intervals \( I_k \).

(b) \( P \) contains no point \( x \) which lies inside an interval \( I_k \).

(c) Whenever \( x_i \) is not one of the \( u_k \), then \( x_i - x_{i-1} < \delta \).

Put \( M = \sup |f(x)| \). Then \( M_k - m_k \leq 2M \). In fact, we have \( M_i - m_i \leq \epsilon \) unless \( x_{i-1} \)
19.6 Piecewise properties

Figure 19.1: The graph of \( t = s^{p-1} \)

is one of the \( u_k \).

\[
Osc(f, P) = U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1})
= \sum_{\text{contin}} (M_i - m_i)(x_i - x_{i-1}) + \sum_k (M_k - m_k)(v_k - u_k)
\leq \varepsilon \sum_{\text{contin}} (x_i - x_{i-1}) + 2M \sum_k (v_k - u_k)
= \varepsilon(b - a) + 2M\varepsilon.
\]

By Oscillation thm and \( K-\varepsilon \) (with \( K = (b - a) + 2M \)), \( f \in \mathbb{R} \).

**Theorem 19.6.4** (Young’s Inequality). For conjugate exponents \( p, q \) (\( \frac{1}{p} + \frac{1}{q} = 1 \), \( p, q > 1 \) and \( a, b \geq 0 \),

\[
ab \leq \frac{a^p}{p} + \frac{b^q}{q}.
\]

**Proof.** Consider the graph of \( t = s^{p-1} \): Since

\[
\frac{1}{p} + \frac{1}{q} = 1 \implies \frac{1}{p} = 1 - \frac{1}{q} = \frac{q-1}{q} \implies p - 1 = \frac{1}{q-1},
\]

this is also the graph of \( s = t^{q-1} \).

Now (1) = \( \int_0^a s^{p-1} = \left[ \frac{a^p}{p} \right]_0^a = \frac{a^p}{p} \), and (2) = \( \int_0^b t^{q-1} = \left[ \frac{b^q}{q} \right]_0^b = \frac{b^q}{q} \).

Thus the area of the entire shaded region is \( (1) + (2) = \frac{a^p}{p} + \frac{b^q}{q} \), which is clearly always larger than the box of area \( ab \).
Theorem 19.6.5 (Hölder’s Inequality). For \( f, g \in \mathcal{R}[a, b] \) and \( \frac{1}{p} + \frac{1}{q} = 1 \), \( p, q > 1 \),

\[
\left| \int_a^b f(x)g(x) \, dx \right| \leq \left( \int_a^b |f(x)|^p \, dx \right)^{1/p} \left( \int_a^b |g(x)|^q \, dx \right)^{1/q}.
\]

Proof. Put \( A = \left( \int_a^b |f(x)|^p \, dx \right)^{1/p} \), \( B = \left( \int_a^b |g(x)|^q \, dx \right)^{1/q} \). Note: \( A, B \neq 0 \) or else trivial. Then let \( a = \frac{|f(x)|}{A}, b = \frac{|g(x)|}{B} \) and apply Young’s:

\[
ab = \frac{|f(x)g(x)|}{AB} \leq \frac{|f(x)|^p}{pA^p} + \frac{|g(x)|^q}{qB^q} = \frac{a^p}{p} + \frac{b^q}{q}.
\]

but \( A^p = \int |f|^p \, dx \) and \( B^q = \int |g|^q \, dx \), so this is

\[
\frac{1}{AB} \int_a^b |f(x)g(x)| \, dx \leq \frac{A^p}{pA^p} + \frac{B^q}{qB^q} = \frac{1}{p} + \frac{1}{q} = 1.
\]

\[
\int_a^b |f(x)g(x)| \, dx \leq \left( \int_a^b |f(x)|^p \, dx \right)^{1/p} \left( \int_a^b |g(x)|^q \, dx \right)^{1/q}.
\]

For \( p = q = 2 \), this is called the Cauchy-Schwartz Inequality.

Recall from Chap 16:
The line \( y = m(x-c) + g(c) \) is a supporting line to \( g \) at \( c \) iff it always lies below the graph of \( g \),

\[
g(x) \geq m(x-c) + g(c),
\]

and \( g \) is convex iff every point \( x \) in the domain of \( g \) has a supporting line.

Theorem 19.6.6 (Jensen’s Inequality). Let \( g \) be convex on \( \mathbb{R} \) and let \( f \in s\mathcal{R}[a, b] \). Then

\[
\int g(f(t)) \, dt \geq g \left( \int f(t) \, dt \right).
\]

In Chapter 20, we will see that \( e^x \) is convex.

Theorem 19.6.7. For \( f \in \mathcal{R}[a, b], \int e^{f(t)} \, dt \geq e^{\int f(t) \, dt} \).

**Required: Ex #18.4.1, 18.4.2, 19.2.1, 19.3.1, 19.4.2, 19.6.3 Prob**

**Recommended: Ex #19.2.3, 19.3.2, 19.4.3, 19.4.4, 19.5.1 Prob 19-2, 19-3**

1. Prove Jensen’s Inequality. **Hint:** Let \( \alpha = \int f(t) \, dt \) and pick a supporting line \( y(x) = m(x-\alpha) + \varphi(\alpha) \) at \( \alpha \). Deduce that \( \prod b_n^a \leq \sum a_nb_n \) for \( \sum a_n = 1; a_n, b_n \geq 0. \)
Chapter 20

Derivatives and Integrals

20.1 First fundamental theorem of calculus

"The fundamental theorem(s) of calculus" (next two thms) shows that integration and differentiation are almost inverse operations.

**Definition 20.1.1.** A primitive of \( f \) (or antiderivative) is a function \( F \) such that \( f = F' \).

**Theorem 20.1.2.** Any two primitives of \( f \) differ only by a constant.

*Proof.* Let \( F, G \) both be primitives of \( f \). Then

\[
(F - G)' = F' - G' = f - f = 0 \quad \Rightarrow \quad F - G = c, \text{ for some } c \in \mathbb{R}.
\]

Putting this minor result together with the next two shows that

\[
D(I(f)) = f, \quad \text{but} \quad I(D(f)) = f + c,
\]

so integration and differentiation are almost inverse operations.

**Theorem 20.1.3** (Integration of derivative, FToC1). If \( f \in \mathcal{R}[a,b] \) and \( f \) has a primitive \( F \) which is differentiable on \( [a,b] \), then \( \int_a^b f(x) \, dx = F(b) - F(a) \).

---

1 April 18, 2007
Proof. Fix $\varepsilon > 0$ and choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that $\text{Osc}(f, P) < \varepsilon$.

By the MVT, get $t_i \in [x_{i-1}, x_i]$ such that

$$f(t_i)(x_i - x_{i-1}) = F(x_i) - F(x_{i-1}), \quad i = 1, \ldots, n$$

$$\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} F(x_i) - F(x_{i-1}) = F(b) - F(a).$$

Since $L(f, P) \leq \sum f(t_i)(x_i - x_{i-1}) \leq U(f, P),$

$\text{Osc}(f, P) < \varepsilon \implies \left| \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - \int_{a}^{b} f(x) \, dx \right| < \varepsilon. \quad \Box$

20.2 Second fundamental theorem of calculus

Theorem 20.2.1 (Differentiation of integral, FToC2). If $f \in \mathcal{R}[a, b]$, then $F(x) = \int_{a}^{x} f(t) \, dt \in C^0[a, b]$ and $F'(c) = f(c)$ if $f$ is continuous at $c$.

Proof. Since $f \in \mathcal{R}$, $|f(t)| \leq M$ for $a \leq t \leq b$. For $a \leq x < y \leq b$,

$$|F(y) - F(x)| = \left| \int_{x}^{y} f(t) \, dt \right| \leq \int_{x}^{y} |f(t)| \, dt \leq M(y - x).$$

This Lipschitz condition gives uniform continuity of $F$ on $[a, b]$.

If $f$ is continuous at $c$, then given $\varepsilon > 0$, have $\delta > 0$ such that

$$|t - c| < \delta \implies |f(t) - f(c)| < \varepsilon, \forall t \in [a, b].$$

If we choose $s < t \in [a, b]$ such that $c - \delta < s \leq c \leq t < c + \delta$, then

$$\left| \frac{F(t) - F(s)}{t - s} - f(c) \right| = \left| \frac{1}{t - s} \int_{s}^{t} [f(u) - f(c)] \, du \right| < \varepsilon$$

shows that $F'(c) = f(c). \quad \Box$

NOTE: $f \in C^0[a, b] \implies F \in C^1[a, b]$.

Theorem 20.2.2 (Existence and uniqueness for ODE). Let $f \in C^0(I)$ and $a \in I$. Then
the initial value problem
\[
\begin{align*}
y' &= f(x), \\
y(a) &= b,
\end{align*}
\]
has the unique solution \( y = F(x) \), where \( F(x) = b + \int_a^x f(t) \, dt \).

**Proof.** For existence, note that the given function satisfies the given IVP, by the FToC2. For uniqueness, suppose \( G(x) \) also satisfies the IVP. Then \( F' = G' \), so \( F(x) = G(x) + c \). Then the initial condition gives \( G(a) = b = F(a) \), so \( c = 0 \). \( \square \)

**Corollary 20.2.3** (FToC2 implies FToC1). Let \( F(x) \) have the continuous derivative \( f(x) \) on \([a, b]\). Then \( \int_a^b f(t) \, dt = F(b) - F(a) \).

**Proof.** Let \( G(x) := \int_a^x f(t) \, dt \), so that
\[
G'(x) = f(x) = F'(x), \quad x \in [a, b]
\]
by FToC2 and the uniqueness of primitives. Then by Uniqueness for ODE, we get \( G(x) = F(x) + c \) so that
\[
\int_a^x f(t) \, dt = F(x) + c
\]
for some \( c \in \mathbb{R} \). Then setting \( x = a \) gives \( c = -F(a) \), and setting \( x = b \) gives the result. \( \square \)

### 20.3 Other relations between integrals and derivatives

**Theorem 20.3.1** (Integration by parts). If \( f, g \in C^1[a, b] \), then
\[
\int_a^b f(x)g'(x) \, dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x) \, dx.
\]

**Proof.** Put \( h(x) = f(x)g(x) \), so \( h, h' \in \mathcal{R} \) by Integral properties thm. Use the Integration of derivative thm on \( h' \). \( \square \)

NOTE: IBP is product rule in reverse, just like CoV is chain rule in reverse.
Theorem 20.3.2 (Change of variable). If \( g \in C^1[a, b] \), \( g \) is increasing, and \( f \in \mathcal{R}[g(a), g(b)] \), then \( f \circ g \in \mathcal{R}[a, b] \) and
\[
\int_a^b f(g(x))g'(x) \, dx = \int_{g(a)}^{g(b)} f(y) \, dy.
\]

**Proof.** Let \( F \) be a primitive of \( f \), so \( F' = f \). Then the chain rule gives
\[
f(g(x))g'(x) = (F(g(x)))'
\]
\[
= F'(g(b)) - F'(g(a))
\]
\[
= \int_{g(a)}^{g(b)} f(y) \, dy,
\]
where the last two equalities come by FToC1. \( \square \)

### Logarithm and exponential

**Definition 20.4.1.** The **natural logarithm** function is \( \log x := \int_1^x \frac{dt}{t} \), for \( x > 0 \).

**Theorem 20.4.2.**

(i) \( \log x \) is differentiable on \( \mathbb{R}^+ \) and \( (\log x)' = \frac{1}{x} \).

(ii) \( \log x \) is strictly increasing and concave.

(iii) \( \log xy = \log x + \log y \) and \( \log(1/x) = -\log x \).

(iv) \( \log x^r = r \log x \).

(v) \( \lim_{x \to \infty} \log x = \infty \) and \( \lim_{x \to 0^+} \log x = -\infty \).

(vi) \( \log : (0, \infty) \to \mathbb{R} \) is a bijection, i.e., it is invertible.

(vii) There is a unique number \( e \) such that \( \log e = 1 \).

**Proof of (i).** Immediate from FToC. \( \square \)

**Proof of (ii).** It is clear that \( \log x \) is increasing because \( \frac{1}{x} > 0 \) for \( x > 0 \). It is concave because the second derivative is \( -\frac{1}{x^2} < 0 \). \( \square \)

**Proof of (iii).** \( \log(xy) = \int_1^{xy} \frac{dt}{t} = \int_1^x \frac{dt}{t} + \int_x^{xy} \frac{dt}{t} = \log x + \int_1^y \frac{dt}{t} = \log x + \log y \) by change of variable: \( t = xu \). Then \( 0 = \log 1 = \log(x^{1/x}) = \log x + \log \frac{1}{x} \) shows \( \log \frac{1}{x} = -\log x \). \( \square \)
Proof of (iv). If \( r \in \mathbb{Z} \), apply the previous part. If \( r = \frac{1}{n} \),

\[
\log x = \log(x^{1/n})^n = n \log x^{1/n} \implies \frac{1}{n} \log x = \log x^{1/n}.
\]

Now for \( r \in \mathbb{Q} \), it follows from (iii) and what we’ve just shown. So let \( x \in \mathbb{R} \). Then for any sequence of rationals \( r_n \to r \),

\[
\log x^r = \log x^{\lim_{n \to \infty} r_n} = \lim_{n \to \infty} \log x^{r_n} \quad \text{continuity of } \log x
\]

\[
= \lim_{n \to \infty} n r_n \log x \quad \text{first part; } r_n \in \mathbb{Q}
\]

\[
= r \log x.
\]

Proof of (v). Consider the sequence \( \log x^n \). By previous part, this is \( n \log x \to \infty \). Since \( \log x \) is strictly increasing, this suffices to give the first limit. The second limit comes the same way.

Proof of (vi). Immediate from strictly increasing; surjective comes from the previous part.

Proof of (vii). Immediate from (vi).

Definition 20.4.3. The exponential function \( \exp x \) is defined as the inverse of \( \log x \), as proven to exist in the previous thm.

Theorem 20.4.4. (i) \( \exp : \mathbb{R} \to \mathbb{R}^+ \) is convex and differentiable with \( (\exp x)' = \exp x \).

(ii) \( \exp(x + y) = \exp x \cdot \exp y \), and \( \exp(-x) = \frac{1}{\exp x} \), and \( \exp(rx) = (\exp x)^r \).

(iii) \( \exp 0 = 1, \exp 1 = e, \exp r = e^r \).

(iv) \( a^x = e^{x \log a}, a > 0 \).

Proof of (i)-(ii). Immediate from Inverse function thm (for the first two) and the rule for \( \log x \), e.g.,

\[
\log \exp(x + y) = x + y = \log \exp x + \log \exp y = \log(\exp x \cdot \exp y)
\]

\[
\exp(x + y) = \exp x \cdot \exp y.
\]
Proof of (iii). The first two are immediate from results for \( \log x \). To see \( e^r = \exp r \), proceed in the same fashion as for \( \log x \): true for integers by prev, true for rationals, then use continuity to extend to reals.

Proof of (iv). \( e^{x \log a} = \exp(x \log a) = \exp \log a^x = a^x \).

20.5 Stirling’s Formula

**Theorem 20.5.1** (Stirling’s Formula). \( \lim_{n \to \infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1 \).

**Proof.** Later.

This is usually written \( n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \) and indicates that the RHS provides an approximation to \( n! \).

20.6 Growth rate of functions

Already done.

Required: Ex #20.2.4, 20.3.2, 20.3.3, 20.3.4, 20.6.2 Prob 20-1, 20-3

Recommended: Ex #20.3.1, 20.4.2, 20.4.3, 20.6.1 Prob 20-2
Chapter 21

Improper Integrals

21.1 Basic Definitions

Definition 21.1.1. Define the improper integral

\[ \int_a^\infty f(t) \, dt := \lim_{N \to \infty} \int_a^N f(t) \, dt. \]

If \( \lim_{x \to b^-} f(t) = \pm \infty \) then we define the improper integral

\[ \int_a^b f(t) \, dt := \lim_{u \to b^-} \int_a^u f(t) \, dt, \]

or similarly if \( \lim_{x \to a^+} f(t) = \pm \infty \).

The improper integral is said to converge or exist iff the limit exists; otherwise, diverges.

NOTE: if \( f \) is integrable on \([a, \infty)\), it means that \( f \in \mathcal{R}[a, b] \) for any \( b > a \). To say \( f \in \mathcal{R}[a, \infty) \) means both that \( f \) is integrable on \([a, \infty)\) AND that the improper integral converges.

Example 21.1.1. The standard example:

\[ \int_1^\infty \frac{1}{x^p} \, dx \]

converges iff \( p > 1 \).
We know that both diverge for \( p = 1 \) by the integral test applied to \( \sum \frac{1}{n} \) (using CoV \( u = \frac{1}{x} \) for \( \int_0^1 \)). Then the other divergence follows by Comparison. For convergence,

\[
\int_1^\infty \frac{1}{x^p} \, dx = \lim_{N \to \infty} \int_1^N x^{-p} \, dx = \lim_{N \to \infty} \left[ \frac{x^{1-p}}{1-p} \right]_1^N = \lim_{N \to \infty} \left( \frac{N^{1-p} - 1}{1-p} \right) = \frac{1}{p-1}
\]

**Definition 21.1.2.** The **Cauchy Principal Value** is a doubly improper integral, where the limits are taken simultaneously.

**Example 21.1.2.** Consider \( \int_{\mathbb{R}} \frac{t}{1 + t^2} \, dt \)

**Solution.** Using CPV, this can be evaluated

\[
\int_{-\infty}^{\infty} \frac{t}{1 + t^2} \, dt = \lim_{R \to \infty} \int_{-R}^{R} \frac{t}{1 + t^2} \, dt = \frac{1}{2} \lim_{R \to \infty} \left( \log(1 + R^2) - \log(1 + (-R)^2) \right)
\]

If this were broken into separate integrals,

\[
\int_{-\infty}^{\infty} \frac{t}{1 + t^2} \, dt = \lim_{P \to \infty} \int_{0}^{P} \frac{t}{1 + t^2} \, dt + \lim_{Q \to \infty} \int_{-Q}^{0} \frac{t}{1 + t^2} \, dt = \infty - \infty,
\]

and we could not evaluate.

\[\square\]

### 21.2 Comparison theorems

Analogues of the results for series, and proved similarly.

**Theorem 21.2.1** (Tail-convergence). If \( f \) is integrable on (any compact subinterval of) \([a, \infty)\), then

\[ f \in \mathcal{R}[a, \infty) \iff f \in \mathcal{R}[b, \infty), \quad \forall b > a. \]

**Theorem 21.2.2.**

1. If \( f \) is increasing for \( x \gg 1 \) and \( \lim_{x \to \infty} f(x) = L \), then \( f(x) \leq L \) for \( x \gg 1 \).

2. If \( f \) is increasing and if \( f(x) \leq B \) for \( x \gg 1 \), then \( \lim_{x \to \infty} f(x) \) exists and \( \lim_{x \to \infty} f(x) \leq B \).
Theorem 21.2.3 (Comparison for Improper Integrals). If $0 \leq f(x) \leq g(x)$ for $f, g$ integrable on $[a, b]$ (where $b$ may be $\infty$), then $g \in sR[a, \infty)$ implies
\[
\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx.
\]

Proof. On any finite interval, we have $\int_a^R f(t) \, dt \leq \int_a^R g(t) \, dt$, and since $g \geq 0$, the integral can only increase as $R \to \infty$. Now apply the previous theorem to the function $F(R) = \int_a^R f(t) \, dt \leq \int_a^\infty g(t) \, dt$. \hfill \Box

We already saw this example:

Example 21.2.1. Show that erf $x = \int_0^x e^{-t^2/2} \, dt$ is bounded above on the interval $[0, \infty)$. (So now we know this means the improper integral $\lim_{x \to \infty} \text{erf} \, x$ exists. )

Solution. We have an upper bound for the (positive) integrand given by
\[
t \leq t^2 \implies e^{-t^2/2} \leq e^{-t/2},
\]
however this is only true for $t \geq 1$. But suffices to only consider this domain!
\[
\int_0^\infty e^{-t^2/2} \, dt = \int_0^1 e^{-t^2/2} \, dt + \int_1^\infty e^{-t^2/2} \, dt \leq M + \lim_{x \to \infty} \int_1^x e^{-t/2} \, dt \leq M + 2e^{-1/2}. \hfill \Box
\]

Theorem 21.2.4 (Asymptotic Comparison). Suppose $f, g$ are integrable on $[a, \infty)$ and $f \sim g$ as $x \to \infty$. Then
\[
\int_a^\infty f(t) \, dt \text{ converges } \iff \int_a^\infty g(t) \, dt \text{ converges.}
\]

Example 21.2.2. Does $\int_0^\infty \frac{dx}{\sqrt{1 + x^3}}$ converge?

Solution. Both endpoints are improper. For $f(x) = \frac{1}{\sqrt{1 + x^3}}$,
\[
f(x) \sim \frac{1}{\sqrt{x}}, \: x \approx 0^+ \quad \text{and} \quad \int_0^1 \frac{dx}{x^{1/2}} \text{ converges},
\]
\[
f(x) \sim \frac{1}{x^2}, \: x \gg 1 \quad \text{and} \quad \int_1^\infty \frac{dx}{x^2} \text{ converges},
\]
so it is convergent. \hfill \Box
21.3 The Gamma function

Definition 21.3.1. The Gamma function is

\[ \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, dt, \quad x > 0. \]

Theorem 21.3.2. The Gamma function has the following properties.

1. \( \Gamma(n+1) = n! \).

   Proof. Induct on \( n \). Basis: \( \int_0^\infty e^{-t} \, dt = 1 = 0! \). Then IBP:
   \[ \int_0^R t^n e^{-t} \, dt = -R^n e^{-R} + n \int_0^R t^{n-1} e^{-t} \, dt \xrightarrow{R \to \infty} 0 + n\Gamma(n). \]

2. \( \Gamma(x) \) is defined for any \( x > 0 \); \( \lim_{x \to 0^+} \Gamma(x) = \lim_{x \to \infty} \Gamma(x) = \infty \).

   Proof. Each of \( \int_1^\infty x e^{-t} \, dt \) and \( \int_1^\infty x e^{-t} \, dt \) converges by comparison. Then \( \Gamma(x) \xrightarrow{x \to \infty} \infty \) because \( n! \xrightarrow{n \to \infty} \infty \). The other limit is HW 21.3.1.

3. \( \Gamma(x+1) = x\Gamma(x) \).

   Proof. Note that in the proof of (i) we didn’t actually use \( n \in \mathbb{N} \).

4. \( \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \).

   Proof. For \( 0 < a < b < \infty \), we have
   \[ \int_a^b \frac{e^{-t}}{\sqrt{t}} \, dt = \int_0^{\sqrt{\infty}} s^{-2} e^{-s^2} \, ds \]
   \[ = \sqrt{\pi} \]
   \[ \xrightarrow{a \to 0, b \to \infty} 2 \int_0^\infty e^{-s^2} \, ds \]

5. \( \Gamma \in C^\infty(0, \infty) \). Also, \( \Gamma'(1) = \gamma \).

   Proof. Later; we’d need to differentiate under the integral.

6. \( \Gamma(x) \) is convex and so is \( \log \Gamma(x) \).
21.4 Absolute and conditional convergence

Proof. To see $\Gamma(x)$ is convex, use the positivity of $\Gamma'(x) = \int_0^\infty t^{x-1} \log^2 t \, dt$. Hölder’s Inequality gives

$$\Gamma\left(\frac{x}{p} + \frac{y}{q}\right) \leq \Gamma(x)^{1/p} \Gamma(y)^{1/q} \quad \square$$

Theorem 21.3.3. If $f > 0$ on $(0, \infty)$ and (i) $f(x + 1) = xf(x)$, (ii) $f(1) = 1$, and (iii) $\log f(x)$ is convex. Then $f(x) = \Gamma(x)$.

21.4 Absolute and conditional convergence

Definition 21.4.1. $\int_a^\infty f(x) \, dx$ converges absolutely iff $\int_a^\infty |f(x)| \, dx$ converges; it converges conditionally iff it converges but not absolutely.

Theorem 21.4.2. If $f$ is integrable on $[a, \infty)$ and $|f| \in R[a, \infty)$, then $f \in R[a, \infty)$.

Proof. Write $f$ as the difference of two nonnegative functions $f = f^+ - f^-$,

$$f^+(x) = \frac{1}{2}(|f(x)| + f(x)) = \max\{0, f(x)\}$$
$$f^-(x) = \frac{1}{2}(|f(x)| - f(x)) = \max\{0, -f(x)\}.$$

Then $0 \leq f^+(x), f^-(x) \leq |f(x)|$, so the convergence of $\int |f|$ gives the convergence of

$$\int f(x) \, dx = \int f^+(x) \, dx - \int f^-(x) \, dx$$

by Comparison Test. \quad \square

Example 21.4.1. $\int_0^\infty \frac{\sin x}{x} \, dx$ converges conditionally.

Solution. Since $f(x) = \frac{\sin x}{x} \to 1$ as $x \to 0$, the integrand is bounded and continuous on $(0, 1)$ and thus has a finite integral. So we restrict to $\int_1^\infty f(x) \, dx$. Then

$$\int_1^R \frac{\sin x}{x} \, dx = \left[-\frac{\cos x}{x}\right]_1^R - \int_1^R \frac{\cos x}{x^2} \, dx$$
$$\leq M + \int_1^R \left|\frac{\cos x}{x^2}\right| \, dx$$
$$\leq M + \int_1^R \frac{1}{x^2} \, dx,$$

which is finite. To see that the integral does not converge absolutely, do HW 21.4.2. \quad \square
Required: Ex #21.2.1(begh), 21.2.4, 21.3.1, 21.4.2 Prob 21-4, 21-5
Recommended: Ex #21.1.3, 21.1.2, 21.2.2, 21.2.3 Prob 21

1.
Chapter 22

Sequences and Series of Functions

22.1 Pointwise and uniform convergence

What does it mean to say a sequence of functions \{f_n\} converges? I.e., how to define when \(\lim f_n(x) = f(x)\)? There are different (nonequivalent) ways to define such a limit.

What does it mean to say a sum of functions \{f_n\} converges? I.e., how to define \(\sum f_n(x) = f(x)\)? We have seen power series, where \(f_n(x) = a_n x^n\), but what about other kinds of functions?

We want to know when things are valid, like for \(f(x) = \sum f_n(x)\),

\[
\frac{d}{dx} \sum f_n(x) = \sum \frac{d}{dx} f_n(x) \\
\int f(x) \, dx = \sum \int f_n(x) \, dx,
\]

Or even if it is valid to compute things like

\[
\Gamma'(x) = \int_0^\infty \frac{\partial}{\partial x} (t^{x-1} e^{-t}) \, dt = \int_0^\infty t^{x-1} \frac{\log t}{e^t} \, dt.
\]

These operations all involve interchanging the order of limits; series, integrals and derivatives are all defined in terms of limits.

**Definition 22.1.1.** Let \{f_n\} be a sequence of functions all defined on some common
domain $I$. Then $f_n$ converges pointwise iff $\lim_{n} f_n(x)$ exists for every $x \in I$. In this case, we can define the limit function by

$$f(x) := \lim_{n} f_n(x),$$

and write $f_n \xrightarrow{pw} f$. The defn is equivalent to:

$$\forall \varepsilon > 0, \forall x \in I, \exists N, \quad n \geq N \implies f_n(x) \approx f(x).$$

**Example 22.1.1.** Let $f_n(x) = \frac{x}{x+n}$ on $\mathbb{R}$. Then

$$\lim_{x \to \infty} \lim_{n \to \infty} f_n(x) = \lim_{x \to \infty} 0 = 0,$$

$$\lim_{n \to \infty} \lim_{x \to \infty} f_n(x) = \lim_{n \to \infty} 1 = 1.$$

**Example 22.1.2.** Let $f_n(x) = x^n$ on $I = [0, 1]$. Then $\{f_n\}$ converges pointwise and

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

So a sequence of continuous functions can converge pointwise to something which is not continuous!

Even worse:

**Example 22.1.3.** Let $f_k(x) = \lim_{n \to \infty} (\cos k!x\pi)^{2n}$. Then whenever $k!x$ is an integer, $f_k(x) = 1$. If $x = p/q$ is rational, then for $k \geq q$, $f_n(x) = 1$. If $k!x$ is not an integer (for example, if $x$ is irrational), then $f_k(x) = 0$. We have an everywhere discontinuous limit function

$$f(x) = \lim_{k \to \infty} \lim_{n \to \infty} (\cos k!x\pi)^{2n} = \begin{cases} 0, & x \in \mathbb{R} \setminus \mathbb{Q}, \\ 1, & x \in \mathbb{Q}. \end{cases}$$

**Example 22.1.4.** Let $f_n(x) = \frac{x^2}{(1+x^2)^n}$ on $\mathbb{R}$ and consider

$$f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n} = \begin{cases} 0, & x = 0, \\ 1 + x^2, & x \neq 0, \end{cases}$$

since the series is geometric for $x \neq 0$. So a series of continuous functions can converge
pointwise to something which is not continuous! (Not even integrable!)

**Example 22.1.5.** Let \( f_n(x) = \frac{\sin(nx)}{\sqrt{n}} \) on \( \mathbb{R} \). Then

\[
 f(x) = \lim_{n \to \infty} f_n(x) = 0, \quad \forall x \in \mathbb{R},
\]

so \( f'(x) = 0 \). On the other hand, \( f'_n(x) = \sqrt{n} \cos(nx) \) so that \( \lim_{n \to \infty} f'_n(x) \neq f'(x) \).

**Example 22.1.6.** Let \( f_n(x) = n^2x(1-x^2)^n \) on \([0,1]\). Then \( \lim_{n \to \infty} f_n(x) = 0 \) for any \( x \in [0,1] \). Thus, trivially \( \int_0^1 \lim_{n \to \infty} f_n(x) \, dx = 0 \). However,

\[
 \lim_{n \to \infty} \int_0^1 f_n(x) \, dx = \lim_{n \to \infty} \frac{n^2}{2n + 2} = \infty.
\]

CONCLUSION: pointwise convergence sucks.

**Definition 22.1.2.** Let \( \{f_n\} \) be a sequence of functions with a common domain \( I \). The sequence converges uniformly to \( f \) on \( I \) iff there exists some function \( f \) for which

\[
 \forall \varepsilon > 0, \exists N, \quad n \geq N \implies f_n(x) \approx_{\varepsilon} f(x), \forall x \in I.
\]

We write \( f_n \xrightarrow{\text{unif}} f \).

NOTE: \( \forall x \) appears at the end: \( N \) does not depend on \( x \). This is the "uniform nature" of the convergence; \( N \) works globally for all of \( I \).

NOTE: uniform convergence implies pointwise convergence.

**Theorem 22.1.3.** Suppose \( f(x) = \lim_n f_n(x) \) pointwise. Then

\[
 f_n \xrightarrow{\text{unif}} f \iff \sup_{x \in I} |f_n(x) - f(x)| - n \to 0.
\]

Proof. \( |f_n(x) - f(x)| < \varepsilon, \forall x \) is equivalent to the condition \( \sup_{x \in I} |f_n(x) - f(x)| < \varepsilon \).}

**Example 22.1.7.** \( x^n \) does not converge uniformly on \([0,1]\).

For \( f(x) \equiv 0 \), \( \sup_{x \in I} |f_n(x) - f(x)| = 1 \to 0 \).

More directly, choose \( \varepsilon = \frac{1}{2} \). For any fixed \( n \), one can find \( x \approx 1^- \) such that \( x^n > \frac{1}{2} = \varepsilon \).

**Definition 22.1.4.** \( \sum f_n \) converges pointwise or uniformly iff the corresponding sequence of partial sums converges pointwise or uniformly.
Example 22.1.8. \( \sum \frac{x^n}{n} \) converges uniformly to \( e^x \) on any compact interval \([-R, R]\), but not on \( \mathbb{R} \).

Since \( |c| \leq R \implies 0 < e^c \leq e^R \), we have

\[
\left| e^x - \left( 1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} \right) \right| \leq \frac{e^R e^{n+1}}{(n+1)!} \to 0.
\]

To see that the convergence is not uniform on \( \mathbb{R} \), note that for any given (fixed) \( n \),

\[
n = 2k \implies \lim_{x \to -\infty} s_n(x) = \infty
\]

\[
n = 2k + 1 \implies \lim_{x \to -\infty} s_n(x) = -\infty,
\]

whereas \( \lim_{x \to -\infty} e^x = 0 \). Hence the sup is unbounded for any \( n \) and cannot go to 0.

22.2 Criteria for uniform convergence

Theorem 22.2.1 (Cauchy Criterion). \( \{f_n\} \) converges uniformly on \( I \) iff

\[
\forall \varepsilon > 0, \exists N \quad m, n \geq N \implies |f_n(x) - f_m(x)| < \varepsilon, \forall x.
\]

Proof. HW. (\( \Rightarrow \)), use \( \Delta \) ineq. (\( \Leftarrow \)), use pointwise Cauchy Crit to obtain the limit \( f \). \( \square \)

Theorem 22.2.2 (Weierstrass M-test). Let \( \{f_n\} \) be defined on \( I \) and satisfy \( |f_n(x)| \leq M_n \). If \( \sum M_n \) converges, then \( \sum f_n(x) \) converges uniformly on \( I \).

Proof. Fix \( \varepsilon > 0 \). Then

\[
\left| \sum_{i=n}^{m} f_i(x) \right| \leq \sum_{i=n}^{m} |f_i(x)| \leq \sum_{i=n}^{m} M_n < \varepsilon,
\]

for \( n, m \gg 1 \), because \( \sum M_n \) converges. The result follows from the previous thm. \( \square \)

Example 22.2.1. \( \sum \frac{\cos nx}{n} \) converges uniformly on \( \mathbb{R} \).

Note that \( \left| \frac{\cos nx}{n} \right| \leq \frac{1}{n^2} \) and \( \sum \frac{1}{n^2} \) converges.

In fact, \( \sum \frac{\cos f_n(x)}{n^2} \) converges uniformly on \( \mathbb{R} \) for any arbitrary \( f_n(x) \).

Theorem 22.2.3 (Unif convergence of power series). If \( \sum a_n x^n \) has radius of convergence \( R \), then the series converges uniformly on \([-L, L]\) whenever \( 0 \leq L < R \).
22.3 Continuity and uniform convergence

Proof. We know $\sum a_n x^n$ converges absolutely for $|x| \leq L < R$, so apply the Weierstrass M-test with $|a_n x^n| \leq |a_n| L^n = M_n$. \qed

22.3 Continuity and uniform convergence

Theorem 22.3.1. A uniform limit of continuous functions is continuous.

Proof. Suppose we have $f_n \xrightarrow{\text{unif}} f$, where each $f_n \in C^0(I)$. Then NTS $f$ is continuous at an arbitrary point $c \in I$. Given $\varepsilon > 0$, use unif convergence to pick $N$ such that

$$n \geq N \implies f_n(x) \approx_\varepsilon f(x).$$

Then since $f_n$ is continuous at $c$,

$$x \approx_\delta c \implies f_n(x) \approx_\varepsilon f_n(c).$$

Combine the two to obtain

$$f(x) \approx_\varepsilon f_n(x) \approx_\varepsilon f_n(c) \approx_\varepsilon f(c) \implies f(x) \approx_3 f(c).$$

In $\Delta$-ineq form, we used estimates on the RHS of

$$|f(x) - f(c)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(c)| + |f_n(c) - f(c)|.$$ 

Corollary 22.3.2. If $\sum f_n(x)$ converges uniformly on $I$, then it converges to a continuous function. In particular, a power series is continuous inside its interval of convergence.

Proof. We’ll prove that it’s differentiable in a moment, so wait until then. \qed

22.4 Term-by-term integration

Theorem 22.4.1 (Integration of a uniform limit). Let $f_n \xrightarrow{\text{unif}} f$, where each $f_n \in \mathcal{R}[a,b]$. Then $f \in \mathcal{R}[a,b]$ and $\lim_{n \to \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx$. 
Proof. Put $\varepsilon_n := \sup_{a \leq x \leq b} |f_n(x) - f(x)|$ so that $f_n - \varepsilon_n \leq f \leq f_n + \varepsilon_n$. Then the upper and lower sums satisfy

$$\int_{a}^{b} (f_n - \varepsilon_n) \, dx \leq \int_{a}^{b} f \, dx \leq \int_{a}^{b} (f_n + \varepsilon_n) \, dx$$

and thus $0 \leq \text{Osc}(f, P) \leq 2\varepsilon_n(b-a) \xrightarrow{n \to \infty} 0$. Thus $f \in \mathcal{R}[a, b]$. Now (*) becomes

$$\int_{a}^{b} f_n \, dx \leq \int_{a}^{b} f \, dx \leq \int_{a}^{b} f_n \, dx,$$

which gives

$$\left| \int_{a}^{b} f_n \, dx - \int_{a}^{b} f \, dx \right| \leq 2\varepsilon_n(b-a) \xrightarrow{n \to \infty} 0.$$

Theorem 22.4.2 (Term-by-term integration of a series). If $\sum f_k(x)$ converges uniformly on $[a, b]$ and each $f_k \in \mathcal{R}[a, b]$, then $\int_{a}^{b} f \, dx = \sum \int_{a}^{b} f_k(x) \, dx$.

Proof.

$$\int_{a}^{b} f \, dx = \int_{a}^{b} \left( \sum_{k=0}^{\infty} f_k(x) \right) \, dx = \lim_{n \to \infty} \int_{a}^{b} \sum_{k=0}^{n} f_k(x) \, dx = \lim_{n \to \infty} \sum_{k=0}^{n} \int_{a}^{b} f_k(x) \, dx.$$
22.5 Term-by-term differentiation

$b < \infty$, and suppose $f_n \xrightarrow{\text{unif}} f$ on every compact subset of $(0, \infty)$. If $g \in \mathcal{R}[0, \infty)$, then

$$|f_n| \leq g \implies \lim_{n \to \infty} \int_0^\infty f_n(x) \, dx = \int_0^\infty f(x) \, dx.$$  

Proof. HW.

Theorem 22.4.4 (Stirling’s Formula). $\lim_{n \to \infty} \frac{\Gamma(x+1)}{(x/e)^x \sqrt{2\pi x}} = 1$.

Proof. HW.

22.5 Term-by-term differentiation

Example 22.5.1. Recall $f_n(x) = \frac{\sin nx}{\sqrt{n}}$. This converges uniformly to $f \equiv 0$, but $f'_n(x) \not\to f'(x)$! Not even uniform convergence can save us now! Need stronger hypothesis.

Theorem 22.5.1. Let $f_n \in C^1(I)$, $f_n \xrightarrow{\text{pw}} f$ and $f'_n \xrightarrow{\text{unif}} g$. Then $f \in C^1(I)$ and $f'(x) = g(x)$.

Proof. Fix a point $a \in I$. Then FToC1 gives

$$f_n(x) - f_n(a) = \int_a^x f'_n(t) \, dt \xrightarrow{n \to \infty} \int_a^x g(t) \, dt.$$  

However, we also have $f_n(x) - f_n(a) \to f(x) - f(a)$, so apply FToC2 to

$$f(x) - f(a) = \int_a^x g(t) \, dt$$

to see that $f \in C^1(I)$ with $f'(x) = g(x)$.

This can be strengthened:

Theorem 22.5.2. Suppose $f_n \in C^1(I)$ and $\{f'_n\}$ converges uniformly. If $\{f_n(c)\}$ converges for some $c \in I$, then $f_n \xrightarrow{\text{unif}} f \in C^1(I)$ and $\lim_{n \to \infty} f'_n(x) = f'(x)$.

Proof. Not for the faint of heart.

Corollary 22.5.3. Let $f_k \in C^1(I)$. If $\sum f_k$ converges pointwise and $\sum f'_k$ converges uniformly, then $f(x) := \sum f_k(x) \in C^1(I)$ and $f'(x) = \sum f'_k(x)$. 

Proof. Let \( s_n(x) := \sum_{k=0}^{n} f_k(x) \). Then \( s'_n(x) = \sum_{k=0}^{n} f'_k(x) \) \( \xrightarrow{\mathrm{unif}} f' \), \( s_n \in C^1(I) \), and \( s_n \xrightarrow{\mathrm{pw}} f \), so apply the previous thm.

Example 22.5.2 (Sawtooth function). Recall the uniformly convergent series

\[
f(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\cos(2k+1)x}{(2k+1)^2} = x, \quad 0 \leq x \leq \pi.
\]

Differentiating term-by-term,

\[
f'(x) = \frac{x^2}{2} - \frac{4}{\pi} \left( \sin x + \frac{\sin 3x}{3^2} + \frac{\sin 5x}{5^2} + \cdots \right).
\]

To establish this equality, we’d need uniform convergence of the series on the right. Unfortunately, this doesn’t converge uniformly on \( \mathbb{R} \). If it did, prev thm would give \( f' \in C^0(\mathbb{R}) \), but the sawtooth is clearly nondifferentiable for \( x = k\pi \).

It turns out that it does converge uniformly on \((k\pi, (k+1)\pi)\). Moral: convergence of Fourier series can be subtle.

### 22.6 Power series and analyticity

**Theorem 22.6.1.** Suppose \( f(x) = \sum_{n=0}^{\infty} a_n x^n \) converges on some open interval \( I \). Then \( f \in C^\infty(I) \) and the derivative can be found term-by-term:

\[
f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1}, x \in I.
\]

Before proving this, we need a few results.

**Lemma 22.6.2 (Abel’s Lemma).** Let \( b_n \geq 0 \) be a decreasing sequence and let \( \sum a_n \) be a series whose partial sums are bounded: \( |a_0 + a_1 + \cdots + a_n| \leq A \). Then for all \( n \in \mathbb{N} \),

\[
|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \leq Ab_1.
\]

**Proof.** Use the summation by parts identity:

\[
\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p,
\]
with the partial sums \( A_N := \sum_{n=1}^{N} a_n \) and \( |A_N| \leq A \), so \( A_0 = 0 \) and
\[
\sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N-1} A_n (b_n - b_{n+1}) + A_N b_N - A_0 b_1 \\
\leq A \sum_{n=1}^{N-1} (b_n - b_{n+1}) + A b_N \\
\leq A (b_1 - b_N) + A b_N \\
= A b_1.
\]

If \( A \) were an upper bound on the partial sums of \( \sum |a_n| \) then we could just use \( \Delta \)-ineq.

Abel’s lemma is a workaround for dealing with conditional convergence in this situation.

**Theorem 22.6.3** (Abel’s Theorem). Let \( f(x) = \sum_{n=0}^{\infty} c_k x^k \) converge at the point \( x = R \). Then the series converges uniformly on \([0, R]\).

Proof. Fix \( \varepsilon > 0 \). Since
\[
f(x) = \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{\infty} (c_k R^k) \left( \frac{x}{R} \right)^k,
\]
apply Abel’s Lemma with \( a_k = c_k R^k \) and \( b_k = \left( \frac{x}{R} \right)^k \). Note that \( \sum_{k=0}^{\infty} c_k R^k \) converges by hypothesis, so we can pick \( N \) for which
\[
n, m \geq N \implies \left| \sum_{k=n}^{m} c_k R^k \right| < \varepsilon.
\]

Using \( \varepsilon \) as a bound on the partial sums of \( \sum_{j=0}^{\infty} c_{k+j} R^{k+j} \) and noting that \( x < R \implies \left( \frac{x}{R} \right)^{k+j} \) is decreasing, the Lemma gives
\[
\left| \sum_{j=1}^{n} (c_{k+j} R^{k+j}) \left( \frac{x}{R} \right)^{k+j} \right| < 2\varepsilon \left( \frac{x}{R} \right)^{k+1}.
\]
Thus the Cauchy Criterion for uniform convergence of a series is satisfied, by \( K-\varepsilon \).

Of course, a similar result holds for \( x = -R \). Consequently, we have an easy corollary:

**Corollary 22.6.4.** If a power series converges pointwise on \((-R, R)\), then it converges uniformly on any compact interval \( K \subseteq (-R, R) \).
Theorem 22.6.5. If a power series \( \sum_{n=0}^{\infty} a_n x^n \) converges on \( I = (-R,R) \), then the differentiated power series \( \sum_{n=0}^{\infty} na_n x^{n-1} \) also converges on \( I \).

Proof. First, from the Ratio Test, we know that \( \lim \left| \frac{b_{n+1}}{b_n} \right| = r < 1 \) implies \( \sum b_n \) converges. Then if \( 0 < s < 1 \),

\[
\left| \frac{(n+1)s^n}{ns^{n-1}} \right| = \frac{n+1}{n}s \rightarrow s \in (0,1) \Rightarrow \sum ns^{n-1} \text{ converges.}
\]

Thus \( ns^{n-1} \) is bounded. Now choose \( t \) to satisfy \( |x| < t < R \) and observe that

\[
|na_n x^{n-1}| = \frac{1}{t} \cdot n \cdot \left| \frac{x}{t} \right|^{n-1} \cdot |a_n t^n|.
\]

Using \( s = \left| \frac{x}{t} \right| \) in the first part, we obtain a bound \( |ns^{n-1}| \leq B \), so that

\[
\left| \sum_{n=0}^{\infty} na_n x^{n-1} \right| \leq \sum_{n=0}^{\infty} |na_n x^{n-1}|
= \sum_{n=0}^{\infty} \frac{1}{t} \cdot n \cdot \left| \frac{x}{t} \right|^{n-1} \cdot |a_n t^n|
\leq \frac{B}{t} \sum_{n=0}^{\infty} |a_n t^n|,
\]

which converges, since \( t \in I \).

Consequently, the convergence of the differentiated series is uniform on any compact \( K \subseteq I \). Returning to the first theorem of the section, this gives

Theorem 22.6.6. Suppose \( f(x) = \sum_{n=0}^{\infty} a_n x^n \) converges on some open interval \( I \). Then \( f \in C^\infty(I) \) and the derivative can be found term-by-term:

\[
f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1}, x \in I.
\]

Example 22.6.1. Find a closed-form expression for

\[
f(x) = \frac{x^2}{1 \cdot 2} + \frac{x^3}{2 \cdot 3} + \cdots = \sum_{n=0}^{\infty} \frac{x^{n+2}}{(n-1) \cdot n}.
\]

Solution. Ratio test gives convergence on \( I = (-1,1) \). Differentiating twice,

\[
f''(x) = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.
\]
Then integrating this twice, get
\[ f(x) = x + (1 - x) \log(1 - x), \quad |x| < 1, \]
using the constants of integration found from \( f(0) = 0, f'(0) = 0. \)

**Corollary 22.6.7.** On its interval of convergence, the Taylor series of \( f(x) = \sum_{n=0}^{\infty} a_n x^n \) is just itself.

**Proof.** Applying the previous theorem, see that after \( n \) differentiations, get
\[ f^{(n)}(x) = n! a_n + c_1 x + c_2 x^2 + \ldots, \quad c_i \in \mathbb{R}. \]
Thus evaluating at \( x = 0 \) gives \( a_n = f^{(n)}(0)/n! \).

**Corollary 22.6.8.** If \( \sum_{n=0}^{\infty} a_n x^n \) has radius of convergence \( R > 0 \) and \( \sum_{n=0}^{\infty} a_n x^n = 0 \) for \( |x| < R \), then \( a_n = 0, \forall n. \)

**Proof.** Consider the Taylor series of \( f(x) \equiv 0. \)

**Corollary 22.6.9.** If \( f(x) = \sum_{n=0}^{\infty} a_n x^n \) has radius of convergence \( R > 0 \) and \( f(x) = \sum_{n=0}^{\infty} b_n x^n \) for \( x \in (-R, R) \), then \( a_n = b_n, \forall n. \)

**Proof.** Apply prev to \( \sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n - b_n) x^n. \)

**Example 22.6.2 (Series solutions to ODE.).** Find a solution to the IVP
\[
\begin{align*}
  y' + xy &= 0, \\
  y(0) &= 1.
\end{align*}
\]

**Solution.** Assume that \( y \) has a series representation: \( y = \sum a_n x^n, 0 \leq |x| < R \). We “just” need to find the \( a_n \). Differentiating term-by-term,
\[
y' + xy = \sum_{n=1}^{\infty} n a_n x^{n-1} + \sum_{n=0}^{\infty} a_n x^{n+1}
= \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n + \sum_{n=1}^{\infty} a_{n-1} x^n
= a_1 + \sum_{n=1}^{\infty} ((n+1) a_{n+1} + a_{n-1}) x^n = 0.
\]
This shows $a_1 = 0$. Now by prev corollary, just need to solve

$$(n + 1)a_{n+1} + a_{n-1} = 0 \quad \Rightarrow \quad a_{n+2} = -\frac{a_n}{n + 2}, \, n \geq 0.$$ 

We start with $y(0) = a_0 = 1$. Then

$$a_1 = a_3 = a_5 = \cdots = a_{2k+1} = 0,$$
$$a_2 = -\frac{a_0}{0 + 2} = -\frac{1}{2},$$
$$a_4 = -\frac{a_2}{2 + 2} = \left(-\frac{1}{2}\right)^2 \left(-\frac{1}{4}\right),$$
$$a_6 = -\frac{a_4}{4 + 2} = \left(-\frac{1}{2}\right) \left(-\frac{1}{4}\right) \left(-\frac{1}{6}\right)$$
$$\vdots$$
$$a_{2n} = \frac{(-1)^n}{2^n n!}.$$ 

The resulting series is

$$y(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} x^{2n} = \sum_{n=0}^{\infty} \left(-\frac{x^2}{2}\right)^n \frac{1}{n!} = e^{-x^2/2},$$

and Ratio Test shows it converges for all $x \in \mathbb{R}$. 

**Definition 22.6.10.** A function $f$ is said to be *analytic* (at 0) iff

$$\lim_{n \to \infty} \sum_{k=0}^{n} a_k x^k = f(x),$$

that is, if the Taylor series of $f$ converges $f$.

**Example 22.6.3.** We spent a lot of time proving that an analytic function is infinitely differentiable, but the converse is false. Define

$$f(x) = \begin{cases} 
    e^{-1/x^2}, & x \geq 0 \\
    0, & x \leq 0.
\end{cases}$$

Then $f \in C^\infty(\mathbb{R})$ with $f^{(n)}(0) = 0$, so the Taylor series of $f$ is $\sum 0x^n = 0 \neq f$. 

Required: Ex #22.1.1(ac), 22.2.2(d), 22.3.3, 22.4.1, 22.4.4, 22.6.2 Prob
22.6 Power series and analyticity

22-1, 22-2, 22-4

Recommended: Ex #22.1.2, 22.2.5, 22.3.1, 22.4.3, 22.6.3, 22.6.4