Math 521 – Fall 2008

521

July 28, 2009

The text references in these notes are to R.C. Buck, *Advanced Calculus*, 3rd edition, Waveland Press. In these notes we shall often use the abbreviations

$\Longrightarrow$ for "implies",
$\Longleftrightarrow$ for "if and only if",
$\exists$ for "there exists",
$\forall$ for "for all".

These abbreviations help clarify the logical structure of the definitions and theorems. The following standard notations are also used:

$\mathbf{Z}_+$	the set of positive integers (**natural numbers**).
$\mathbf{N}$	the set of nonnegative integers.
$\mathbf{Z}$	the set of integers.
$\mathbf{Q}$	the set of rational numbers.
$\mathbf{R}$	the set of real numbers.
$\mathbf{C}$	the set of complex numbers.

The usual **interval notation** is used, e.g.

$$[a,b) := \{x \in \mathbf{R} : a \le x < b\}, \qquad (-\infty, b] := \{x \in \mathbf{R} : x \le b\}, \qquad \text{etc.}$$

We write := to indicate that two objects are equal by definition. We also signal definitions by writing *iff* instead of *if and only if.*

Throughout $\mathbf{R}^n$ denotes the vector space of all $n$tuples of real numbers, so $\mathbf{R}^1 = \mathbf{R}$, $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$, $\mathbf{R}^3 = \mathbf{R} \times \mathbf{R} \times \mathbf{R}$, etc. (Buck uses the term n **space** as a synonym for $\mathbf{R}^n$.)

1 Sets, Functions, and Maps

1. The notation $x \in A$ means that the point x is an element of the set A. The notation $A \subseteq B$ means that A is a subset of B, i.e. for all x we have $x \in A \implies x \in B$. By definition, two sets are **equal** if each is a subset of the other:

$$A = B \iff A \subseteq B \text{ and } B \subseteq A.$$

The notation $\{x : P(x)\}$ denotes the set of all x for which the property $P(x)$ is true. The notation $\{x \in A : P(x)\}$ denotes the set of all $x \in A$ for which the property $P(x)$ is true. If A and B are sets, then the sets

$$A \cup B := \{x : x \in A \text{ or } x \in B\}, \qquad A \cap B := \{x : x \in A \text{ and } x \in B\},$$

are called respectively the **union** and **intersection** of A and B. The **empty set** is denoted $\emptyset$: $x \notin \emptyset$ for all x. Two sets are **disjoint** iff they have no elements in common, i.e iff $A \cap B = \emptyset$. (See Buck page 5.) The set

$$B \setminus A := \{y \in B : y \notin A\}$$

is called the **complement** of A in B. (See page 30.) The set

$$A \times B := \{(x, y) : x \in A \text{ and } y \in B\}$$

of all ordered pairs (x, y) with $x \in A$ and $y \in B$ is called the **Cartesian product** of A and B.

2. An **indexed family of sets** is a function which assigns a set A_i to each element i of a set I. The set I is called the **index set** of the family and the family is usually denoted $\{A_i\}_{i \in I}$. The **union** and **intersection** of the indexed family are defined by

$$x \in \bigcup_{i \in I} A_i \iff \exists i \in I \;\; x \in A_i.$$

$$x \in \bigcap_{i \in I} A_i \iff \forall i \in I \;\; x \in A_i.$$

The notation $\exists i \in I$ is an abbreviation for "there exists $i \in I$", and the notation $\forall i \in I$ is an abbreviation for "for all $i \in I$". In these definitions the set I can be infinite. For finite sets I we recover the earlier definitions, e.g. for $I = \{1, 2\}$ we have

$$\bigcup_{i\in\{1,2\}} A_i = A_1 \cup A_2, \qquad \bigcap_{i\in\{1,2\}} A_i = A_1 \cap A_2.$$

This illustrates the logical principle that $\exists$ is like an “infinite or” and $\forall$ is like an “infinite and”. From logic we know that

$$\text{not } \forall \iff \exists \text{ not}, \qquad \text{and} \qquad \text{not } \exists \iff \forall \text{ not},$$

so for any set B and any indexed family $\{A_i\}_{i\in I}$ we have

$$B \setminus \bigcup_{i\in I} A_i = \bigcap_{i\in I} (B \setminus A_i), \qquad B \setminus \bigcap_{i\in I} A_i = \bigcup_{i\in I} (B \setminus A_i).$$

In particular, for $I = \{1, 2\}$ we have

$$B \setminus (A_1 \cup A_2) = (B \setminus A_1) \cap (B \setminus A_2), \quad B \setminus (A_1 \cap A_2) = (B \setminus A_1) \cup (B \setminus A_2).$$

Also by logic we have set theoretic distributive laws

$$B \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I} (B \cap A_i), \qquad B \cup \bigcap_{i\in I} A_i = \bigcap_{i\in I} (B \cup A_i).$$

In particular, for $I = \{1, 2\}$ we have

$$B \cup (A_1 \cap A_2) = (B \cup A_1) \cap (B \cup A_2), \quad B \cap (A_1 \cup A_2) = (B \cap A_1) \cup (B \cap A_2).$$

3. A **function** is a rule which assigns a value $f(x)$ to every point x from a set called the **domain** of the function. The set

$$\operatorname{graph}(f) := \{(x, y) : y = f(x)\}$$

of all pairs (x, y) such that $y = f(x)$ is called the **graph** of the function f.

4. Let X and Y be sets. We say that f is a **map** from X to Y and write $f : X \to Y$ when f is a function which assigns a point $y = f(x) \in Y$ to each point $x \in X$. Two maps $f : X \to Y$ and $f' : X' \to Y'$ are said to be **equal** when $X = X'$, $Y = Y'$, and $f(x) = f'(x)$ for all $x \in X$. Thus if f and f' equal maps, then $\operatorname{graph}(f) = \operatorname{graph}(f')$ but not conversely (because $Y = Y'$ is part of the definition of equality for maps). (See Buck page 23).

5. When $A \subseteq X$, $B \subseteq Y$, and $f : X \to Y$, the sets

$$\begin{aligned} f(A) &:= \{y \in Y : \exists x \in A \text{ such that } y = f(x)\}, \\ f^{-1}(B) &:= \{x \in X : f(x) \in B\}, \end{aligned}$$

are called respectively the **image** of A by f and the **pre-image** (or **inverse image**) of B by f. (See Buck page 76.)

Problem 6. Let $f : X \to Y$, $\{A_i\}_{i \in I}$ be a family of subsets of X, and $\{B_i\}_{i \in I}$ be a family of subsets of Y. Say which of the following inclusions are (always) true. Prove each true inclusion and give a counterexample for each false one.

$$\begin{array}{ll} f^{-1}\left(\bigcup_{i \in I} B_i\right) \subseteq \bigcup_{i \in I} f^{-1}(B_i)? & \bigcup_{i \in I} f^{-1}(B_i) \subseteq f^{-1}\left(\bigcup_{i \in I} B_i\right)? \\ f^{-1}\left(\bigcap_{i \in I} B_i\right) \subseteq \bigcap_{i \in I} f^{-1}(B_i)? & \bigcap_{i \in I} f^{-1}(B_i) \subseteq f^{-1}\left(\bigcap_{i \in I} B_i\right)? \\ f\left(\bigcup_{i \in I} A_i\right) \subseteq \bigcup_{i \in I} f(A_i)? & \bigcup_{i \in I} f(A_i) \subseteq f\left(\bigcup_{i \in I} A_i\right)? \\ f\left(\bigcap_{i \in I} A_i\right) \subseteq \bigcap_{i \in I} f(A_i)? & \bigcap_{i \in I} f(A_i) \subseteq f\left(\bigcap_{i \in I} A_i\right)? \end{array}$$

7. If $f : X \to Y$ and $g : Y \to Z$, then the **composition** of f and g is the map $g \circ f : X \to Z$ defined by

$$(g \circ f)(x) = g(f(x))$$

for $x \in X$. For any set X the **identity map** of X is the map $\mathrm{id}_X : X \to X$ defined by

$$\mathrm{id}_X(x) = x$$

for $x \in X$. Clearly

$$\mathrm{id}_Y \circ f = f \text{ and } f \circ \mathrm{id}_X = f$$

for $f : X \to Y$.

8. A map $g : Y \to X$ is said to be a **left inverse** for the map $f : X \to Y$ iff $g \circ f = \mathrm{id}_X$, i.e. iff $g(f(x)) = x$ for all $x \in X$. A map $g : Y \to X$ is said to be a **right inverse** for the map $f : X \to Y$ iff $f \circ g = \mathrm{id}_Y$, i.e. iff $f(g(y)) = y$ for all $y \in Y$. A map $g : Y \to X$ is said to be a (two sided) **inverse** to the map $f : X \to Y$ iff it is both a left inverse and a right inverse to f. If g is a left inverse to f and g' is a right inverse to f then $g = g'$. (Proof: $g = g \circ \mathrm{id}_X = g \circ f \circ g' = \mathrm{id}_Y \circ g' = g'$.) In this case there is a unique inverse and it is denoted f^{-1}. So if $f : X \to Y$ has an inverse $f^{-1} : Y \to X$, then

$$y = f(x) \iff x = f^{-1}(y)$$

for $x \in X$ and $y \in Y$.

9. A map $f : X \to Y$ is said to be **one-one** iff

$$\forall x_1, x_2 \in X \; [f(x_1) = f(x_2) \implies x_1 = x_2]$$

and it is said to be **onto** iff $Y = f(X)$, i.e.

$$\forall y \in Y \; \exists x \in X \; y = f(x)$$

Thus

(1) A map is one-one if and only if it has a left inverse;

(2) A map is onto if and only if it has a right inverse;

(3) A map is one-one and onto if and only if it has an inverse.

Remark 10. The 'only if' part of item (2) is called the **Axiom of Choice**. It was once controversial because one can imagine a situation where one can prove that a map f is onto but where one can not give an explicit formula for a right inverse.

Remark 11. The above assertions (1-3) are false if continuity (defined later) is required: There is a continuous one-one and onto map whose inverse is not continuous (and hence continuous one-one map which does not have a continuous left inverse, and a continuous onto map which does not have a continuous right inverse). See Example 64 below.

Example 12. Consider the four maps

$$f_1 : \mathbf{R} \to \mathbf{R}, \quad f_2 : [0, \infty) \to \mathbf{R}, \quad f_3 : \mathbf{R} \to [0, \infty), \quad f_4 : [0, \infty) \to [0, \infty)$$

defined by $f_i(x) = x^2$. Then f_1 is not one-one and not onto, f_2 is one-one but not onto, f_3 is onto but not one-one, and f_4 is one-one onto. Any map $g_2\mathbf{R} \to [0, \infty)$ such that $g_2(y) = \sqrt{y}$ for $y \geq 0$ is a left inverse to f_2, and any map $g_3 : [0, \infty) \to \mathbf{R}$ such that $g_3(y) = \pm\sqrt{y}$ (the ± can depend on y) is a right inverse to f_3 The inverse map to f_4 is $f_4^{-1}(y) = \sqrt{y}$.

2 Axioms for the Real Numbers

We state here the axioms for the real number system $\mathbf{R}$. We shall accept these axioms without proof but it can proved (from more general axioms) that there is an essentially unique structure satisfying them.

13. Algebraic Axioms. The set $\mathbf{R}$ of **real numbers** is equipped with two operations

$$\mathbf{R} \times \mathbf{R} \to \mathbf{R} : (a, b) \mapsto a + b, \qquad \mathbf{R} \times \mathbf{R} \to \mathbf{R} : (a, b) \mapsto a \cdot b$$

such that the usual laws of grade school arithmetic hold:

(Commutative Laws) $a + b = b + a$ and $a \cdot b = b \cdot a$.

(Associative Laws) $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(Distributive Law) $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

(Zero,One) There are (necessarily unique) distinct elements $0, 1 \in \mathbf{R}$ such that $a + 0 = a$ and $a \cdot 1 = a$ for all $a \in \mathbf{R}$.

(Inverses) For every $a \in \mathbf{R}$ there is a (necessarily unique) element $-a$ such that $a + (-a) = 0$. For every $a \in \mathbf{R} \setminus \{0\}$ there is a (necessarily unique) element a^{-1} such that $a \cdot a^{-1} = 1$.

The standard notations from high school algebra are used: in particular, $ab := a \cdot b$, $a - b := a + (-b)$, and $a/b = a \cdot b^{-1}$.

14. Order Axioms. The set $\mathbf{R}$ has an order relation denoted $a < b$ satisfying the following laws for all $a, b, c \in \mathbf{R}$:

(Trichotomy) Exactly one of the alternatives $a < b$, $a = b$, $b < a$, holds.

(Addition) $a < b \implies a + c < b + c$.

(Multiplication) $0 < a, b \implies 0 < ab$.

The other order notations are defined as usual, i.e. $a < b \iff b > a$ and $a \leq b \iff b \geq a \iff$ either $a < b$ or $a = b$.

Remark 15. All the rules of algebra used in College Algebra (Math 112) follow from the Algebraic Axioms 13 and Order Axioms 14. For example, $(a+b)^2 = a^2 + 2ab + b^2$, $a^2 \geq 0$, etc.

Definition 16. The set S of real numbers is said to be **bounded above** iff there is a number $b \in \mathbf{R}$ such that $x \leq b$ for all $x \in S$; the number b is then called an **upper bound** for S. A number $b \in \mathbf{R}$ is called a **least upper bound** for S iff it is an upper bound for S and $b \leq b'$ for every other upper bound b' for S. Similarly the set S is said to be **bounded below** iff there is an number $a \in \mathbf{R}$ such that $a \leq x$ for all $x \in S$; the element b is then called a **lower bound** for S. An element $a \in \mathbf{R}$ is called a **greatest lower bound** iff it is an lower bound for S and $a' \leq a$ for every other lower bound a' for S. The words **infimum** and *greatest lower bound* are synonymous as are the words **supremum** and *least upper bound.* The least upperbound of the set S will be denoted $\sup(S)$ and the greatest lower bound of the set S will be denoted $\inf(S)$.

17. Completeness Axiom. Every set S of real numbers which is bounded above has a least upper bound, i.e.

$$\text{if } x \leq b \text{ for all } x \in S, \text{ then } x \leq \sup(S) \leq b \text{ for all } x \in S.$$

Because multiplication by -1 reverses the order it is the same to say that every set which is bounded below has a greatest lower bound. Thus

$$\text{if } a \leq x \text{ for all } x \in S, \text{ then } a \leq \inf(S) \leq x \text{ for all } x \in S.$$

Theorem 18 (Archimedean Property)**.** *There is neither an infinite real number nor an infinitesimal real number. More precisely,*

(1) *There is no real number which is larger than every integer.*

(2) *For every positive real number $\varepsilon > 0$ there is a positive integer n such that $1/n < \varepsilon$.*

Problem 19. Prove Theorem 18. Hint: If $\omega > n$ for every integer n what about $\omega - 1$?

Problem 20. Let $\mathcal{R}$ denote the set of real valued rational functions, i.e. $f \in \mathcal{R}$ iff $f(x) = p(x)/q(x)$ where $p(x)$ and $q(x)$ are polynomials with real coefficients (and $q(x)$ is not the zero polynomial). For $f, g \in \mathcal{R}$ define an order relation by the condition that $f > g$ iff there exists an M such that $f(x) > g(x)$ for all $x > M$. Then the set $\mathcal{R}$ satisfies the algebraic axioms and order axioms given above. View $\mathbf{R}$ (and hence $\mathbf{Z}$) as a subset of $\mathcal{R}$ by identifying the real number c with the constant function whose value is always c. Exhibit (in the lingo of Theorem 18) an infinite element $f \in \mathcal{R}$ and an infinitesimal element $g \in \mathcal{R}$.

3 Distance

21. The **distance** $d(p, q)$ between two points $p = (x_1, x_2, \ldots, x_n)$ and $q = (y_1, y_2, \ldots, y_n)$ in $\mathbf{R}^n$ is defined by

$$d(p, q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

The distance $d(v, 0)$ from a vector $v \in \mathbf{R}^n$ to the origin is called the **norm** of v denoted by $|v|$ so

$$d(p, q) := |p - q|.$$

The norm satisfies the following laws for $v, w \in \mathbf{R}^n$:

(Zero Norm) $|v| = 0 \iff v = 0$,

(Homogeneity) $|av| = a\,|v|$ if $a > 0$,

(Symmetry) $|-v| = |v|$,

(Triangle Inequality) $v + w \leq |v| + |w|$.

(The Zero Norm Law holds because a sum of squares vanishes only if each summand vanishes and the Triangle Inequality is proved in Corollary on page 14 of Buck.) The laws for the norm imply that the distance function satisfies the following:

(Zero Distance) $d(p, q) = 0 \iff p = q$,

(Symmetry) $d(p, q) = d(q, p)$,

(Triangle Inequality) $d(p, r) \leq d(p, q) + d(q, r)$.

These are proved by reading $v = p - q$ and $w = q - r$ in the corresponding law for the norm.

Remark 22. In Chapters I-V Buck writes $|p - q|$ instead of $d(p, q)$ but uses the notation $d(p, q)$ starting in Chapter VI in a more general setting. (See Buck page 304.)

Definition 23. For $p \in \mathbf{R}^n$ and $\delta > 0$ the set

$$B(p, \delta) := \{q \in \mathbf{R}^n : d(p, q) < \delta\}$$

is called the **open ball** centered at p with radius δ. Note that when $n = 1$ the open ball is an open interval: for $a \in \mathbf{R}$

$$B(x, \delta) = \{x \in \mathbf{R} : |x - a| < \delta\} = \{x \in \mathbf{R} : a - \delta < x < a + \delta\} = (a - \delta, a + \delta).$$

,

4 Topological Terminology

The following definitions are from section I.5 of Buck. In all these definitions the term *set* means *subset of* $\mathbf{R}^n$.

Definition 24. A set $U \subseteq \mathbf{R}^n$ is **open** iff for every $p \in S$ there exists a $\delta > 0$ such that $B(p, \delta) \subseteq U$. The **interior** of a set $S \subseteq \mathbf{R}^n$ is the set of all $p \in S$ such that $B(p, \delta) \subseteq S$ for some $\delta > 0$. Thus a set is open if and only if it equals its interior and the interior of a set is the largest open set contained in S. A **neighborhood** of a point p is a set containing p in its interior. The **exterior** of a set S is the interior of its complement $\mathbf{R}^n \setminus S$. The **boundary** of a set S is the set of points which are neither exterior or interior to S. It is denoted by $\mathrm{bdry}(S)$. A set $S \subseteq \mathbf{R}^n$ is **closed** iff its complement $\mathbf{R}^n \setminus S$ is open. The **closure** of the set S is the set

$$\bar{S} := S \cup \mathrm{bdry}(S).$$

It is the smallest closed set containing S.

Definition 25. A point p is a **cluster point** of a set S iff every neighborhood of p contains infinitely many points of S. The terms **limit point** and **accumulation point** are synonymous with *cluster point*, but for Buck the term limit point has a slightly differnt meaning when applied to a sequence. See Definition 34 and Remark 36 below. A point $p \in S$ is an **isolated** point of S iff some neighborhood of p contains no other point of S.

Definition 26. A set S is **disconnected** iff there are disjoint open sets U and V such that $S \subseteq U \cup V$ and both $S \cap U$ and $S \cap V$ are nonempty. A set is **connected** iff it is not disconnected.

Theorem 27. *A subset $S \subseteq \mathbf{R}$ of the real line is connected if and only if S is an interval, i.e. $[a, b] \subseteq S$ whenever $a, b \in S$.*

Proof. We prove only if. Assume that there exist $a, b \in S$ with $[a, b] \not\subseteq S$. Then there is a $c \in [a, b]$ with $c \notin S$. Let $U = (-\infty, c)$ and $V = (c, \infty)$. The point c lies in the open interval (a, b) as $a, b \in S$ so $a \in U$ and $b \in V$. Hence both $S \cap U$ and $S \cap V$ are nonempty and clearly $S \subseteq U \cup V$ (as $c \notin S$). Hence the open sets U and V separate S so S is disconnected as required.

We prove if. Assume that S is disconnected, i.e. that there exist open sets $U, V \subseteq \mathbf{R}$ with $S \subseteq U \cup V$, $S \cap U \neq \emptyset$, $S \cap V \neq \emptyset$, and $U \cap V = \emptyset$. Choose $a \in S \cap U$ and $b \in S \cap V$. Then $a \neq b$ as $U \cap V = \emptyset$. Assume without loss of generality that $a < b$. (The case $b < a$ is the same.) We must show that $[a, b] \not\subseteq S$, and for this it is enough (as $S \subseteq U \cup V$) to show that $[a, b] \not\subseteq U \cup V$.

The set $T := \{x \in [a, b] : [a, x) \subseteq U\}$ is nonempty (as it contains a) and bounded above (as b is an upperbound) so it has a supremum c. Clearly $[a, c) \subseteq T$. Also $a < c$ (as $(a - \varepsilon, a + \varepsilon) \subseteq U$ for sufficiently small $\varepsilon > 0$) and $c < b$ (as $(b - \varepsilon, b + \varepsilon) \subseteq V \subseteq \mathbf{R} \setminus U$ for sufficiently small $\varepsilon > 0$). If $c \in U$, then $(c - \varepsilon, c + \varepsilon) \subseteq U$ for sufficiently small $\varepsilon > 0$ so $[a, c + \varepsilon) = [a, c) \cup (a - \varepsilon, a + \varepsilon) \subset U$ so $[a, c + \varepsilon) \subseteq T$ contradicting the fact that c is an upperbound for T. If $c \in V$, then $(c - \varepsilon, c + \varepsilon) \subseteq V$ for sufficiently small $\varepsilon > 0$ so $[a, c - \varepsilon) \cap V = \emptyset$ contradicting the fact that c is the least upperbound of T. Hence $c \notin U \cup V$ as required. ☐

Definition 28. A set S is **bounded** iff it is contained in some large ball, i.e. iff there exists $M > 0$ such that $|p| < M$ for all $p \in S$. Thus a set or real numbers is bounded if and only if it is bounded above and bounded below. (See Definition 16.)

5 Limits

29. Let p_0 be a cluster point of a set S and F be a function defined on S (but possibly not at p_0). The notation

$$\lim_{p \to p_0} F(p) = L$$

means that for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $p \in S \cap B(p_0, \delta) \setminus \{p_0\}$ we have $|f(p) - L| < \varepsilon$, i.e. iff

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall p \in S \; [0 < |p - p_0| < \delta \implies |f(p) - L| < \varepsilon].$$

When $p_0 \in S$ and p_0 is a cluster point of S we have that a function f defined on S is continuous at p_0 (see Definition 48 below) if and only if

$$\lim_{p \to p_0} f(p) = f(p_0)$$

(and the function is trivially continuous at a point $p_0 \in S$ which is not a cluster point of S). However, the limit notation is usually used in situations where (p_0 is a cluster point of S but) $p_0 \notin S$. For example, the **derivative** of a real valued function $f : I \to \mathbf{R}$ defined on an open interval $I \subseteq \mathbf{R}$ is defined by

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

The ratio in the limit is undefined when $x = x_0$ but is defined for nearby values of x.

30. For a real valued function f defined on a subset of $\mathbf{R}$ we can extend the definition of the notation $\lim_{x \to a} F(x) = L$ to include the cases where $a = \pm\infty$ and/or $L = \pm\infty$ as follows. Let

$$\hat{\mathbf{R}} := \{-\infty\} \cup \mathbf{R} \cup \{\infty\}$$

consist of the set of real numbers together with two additional points which we think of as located at infinity. The set $\hat{\mathbf{R}}$ is sometimes called the set of **extended real numbers**. For $b \in \hat{\mathbf{R}}$, a set $U \subseteq \mathbf{R}$ is called **neighborhood** of b iff

either $b \in \mathbf{R}$ and U contains an open interval $(b - \rho, b + \rho)$ for some $\rho > 0$,

or else $b = \infty$ and U contains an open interval (M, ∞) for some $M > 0$,

or else $b = -\infty$ and U contains an open interval $(-\infty, -M)$ for some $M > 0$.

Because $B(p, \rho) = (b - \rho, b + \rho)$ this definition agrees with the definition in paragraph 24. A point $a \in \hat{\mathbf{R}}$ is called a **cluster point** of a subset $S \subseteq \mathbf{R}$

iff every neighborhood of b intersect S in an infinite set. If $f : S \to \mathbf{R}$ and a is a cluster point of S, then the notation

$$\lim_{x\to a} F(x) = L$$

means that for every neighborhood V of L there is a neighborhood U of a such that $f(U \cap S) \subseteq V$.

6 Sequences

31. A **sequence** is a function defined on a subset of the integers. (Usually this subset is the set $\mathbf{Z}_+ := \{n \in \mathbf{Z} : n > 0\}$ of positive integers.) It is customary to denote the value of a sequence at an integer n with a subscript rather than with parentheses and to denote a sequence with a notation like $\{p_n\}_n$ or $\{p_n\}_{n\in\mathbf{Z}_+}$. When $n_1 < n_2 < n_3 < \cdots$ is an increasing sequence of positive integers, the sequence $\{p_{n_k}\}_k$ is called a **subsequence** of the sequence $\{p_n\}_n$.

Definition 32. The sequence $\{p_n\}_n$ of points of $\mathbf{R}^m$ is said to **converge** to the point $p \in \mathbf{R}^m$ iff

$$\lim_{n\to\infty} p_n = p$$

i.e. iff for every neighborhood U of p there exists $N > 0$ such that $n \geq N \implies p_n \in U$. This is sometimes abbreviated as $p_n \to p$ as $n \to \infty$. We say a sequence **converges** or is **convergent** iff it converges to p for some $p \in \mathbf{R}^m$. A sequence is said to **diverge** when it does not converge. (A sequence in $\mathbf{R}$ whose limit is infinite is also said to diverge.)

Theorem 33. *A set S is closed if and only if it is closed under limits of sequences, i.e. whenever $\{p_n\}_n$ is a sequence of points of S and $\lim_{n\to\infty} p_n = p$ we have $p \in S$. (Theorem 5 page 40.)*

Definition 34. A point p is a **limit point** of a sequence $\{p_n\}$ iff some subsequence $\{p_{n_k}\}_k$ converges to p.

Example 35. The limit points of the sequence $p_n = (-1)^n$ are 1 and -1. The limit points of the sequence $q_n = (-1)^n + 1/n$ are also 1 and -1.

Remark 36. Every sequence $\{p_n\}_n$ determines a set $\{p_n : n \in \mathbf{Z}_+\}$. Buck calls this set the **trace** of the sequence, but that terminology is uncommon. The trace can be finite, for example the trace of the sequence $p_n = (-1)^n$ is the two element set $\{-1, 1\}$. If the trace of a sequence is finite then there must be at least one constant subsequence and the common value is a limit point of the sequence. By definition only an infinite set has a cluster point.

Theorem 37 (Bolzano-Weierstrass)**.** *Every bounded infinite subset of* $\mathbf{R}^n$ *has an cluster point.*

Corollary 38. *Every bounded sequence in* $\mathbf{R}^n$ *has a limit point. (Buck Theorem 22 page 62.)*

Definition 39. A sequence $\{a_n\}_n$ of real numbers is said to be **increasing** iff $a_1 \leq a_2 \leq \cdots$, is said to be **decreasing** iff $a_1 \geq a_2 \geq \cdots$, and is called **monotonic** iff it is either increasing or decreasing.

Theorem 40. *A bounded monotonic sequence is convergent. (Buck page 47.)*

Proof. Assume the sequence $\{a_n\}$ is increasing and bounded above. Let $a = \sup\{a_n\}$. Then $a \geq a_n$ for all n (as a is an upper bound) but for $\varepsilon > 0$ $a_N > a - \varepsilon$ for some N (as a is the least upper bound). Hence $a_n > a - \varepsilon$ for $n \geq N$ (as the sequence is increasing). □

41. For any sequence $\{a_k\}_k$ of real numbers we have $\{a_k : k \geq n\} \subseteq \{a_k : k \geq m\}$ for $m < n$. If the sequence $\{a_k\}_k$ is bounded above, then the sequence $s_n := \sup\{a_k : k \geq n\}$ is decreasing. The limit of the latter sequence is denoted

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \sup\{a_k : k \geq n\}.$$

Similarly for a sequence which is bounded below,

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \inf\{a_k : k \geq n\}.$$

Definition 42. A sequence $\{p_n\}$ is called **Cauchy** iff

$$\lim_{m,n\to\infty} |p_n - p_m| = 0$$

i.e. iff for every $\varepsilon > 0$ there exists $N > 0$ such that $n, m \geq N \implies |p_n - p_m| < \varepsilon$.

Theorem 43 (Cauchy Convergence Criterion)**.** *A sequence in* $\mathbf{R}^n$ *converges if and only if it is a Cauchy sequence. (Buck Theorem 23 and its corollary on pages 62-63.)*

7 Compact Sets

Definition 44. An **open cover** of a set S is a collection $\{U_i\}_{i\in I}$ of open sets such that $S \subseteq \bigcup_{i\in I} U_i$. The subset S is **compact** iff every compact cover $\{U_i\}_{i\in I}$ of S has finite subcover, i.e. there are indices $i_1, i_2, \dots, i_n \in I$ such that $S \subseteq U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_n}$.

Theorem 45. *The closed interval $[a, b]$ is compact. (Buck Theorem 24 page 65.)*

Theorem 46 (Heine Borel)**.** *A subset of $\mathbf{R}^n$ is compact if and only if it is closed and bounded. (Buck Theorem 25 page 65.)*

Remark 47. Call a set S **sequentially compact** iff for every sequence $\{p_n\}_n$ of points in S there is a point $p \in S$ and a subsequence $\{p_{n_k}\}_k$ which converges to p. Combining Theorems 37 and 46 we see that a set is sequentially compact if and only if it is compact if and only if it is closed and bounded. It follows

8 Continuity

Let $f : D \to \mathbf{R}^m$ where $D \subseteq \mathbf{R}^n$.

Definition 48. The function f is said to be **continuous** at a point $p \in D$ iff for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $q \in D$

$$|q - p| < \varepsilon \implies |f(q) - f(p)| < \varepsilon.$$

Theorem 49. *The function f is continuous at $p \in D$ if and only if for every sequence $\{p_n\}$ of points in D we have*

$$\lim_{n\to\infty} p_n = p \implies \lim_{n\to\infty} f(p_n) = f(p). \tag{1}$$

Proof. We prove 'only if'. Assume f is continuous at p. Choose sequence $\{p_n\}$ of points in D. Assume

$$\lim_{n\to\infty} p_n = p. \tag{2}$$

Choose $\varepsilon > 0$. Because f is assumed to be continuous at p there is a $\delta > 0$ such that for all $q \in D$

$$|q - p| < \delta \implies |f(q) - f(p)| < \varepsilon. \tag{3}$$

By (2) there is an N such that $|p_n - p| < \delta$ for $n > N$. Hence by (3) $f(p_n) - f(q)| < \varepsilon$ for $n > N$. This proves

$$\lim_{n\to\infty} f(p_n) = f(p). \tag{4}$$

as required.

We prove 'if'. Assume that f is not continuous at $p \in D$. Then there is an $\varepsilon > 0$ such that for every $\delta > 0$ there is a $q \in D$ such that

$$|q - p| < \delta \text{ but } |f(q) - f(p)| \geq \varepsilon.$$

In particular, for each $n \in \mathbf{Z}_+$ there is a q_n such that

$$|q_n - p| < \frac{1}{n} \text{ but } |f(q_n) - f(p)| \geq \varepsilon.$$

But then (2) holds but (4) fails. This proves that (1) is false as required. □

Definition 50. The function f is said to be **continuous** iff it is continuous at every point of D iff

$$\forall p \in D\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall q \in D \left[|q - p| < \varepsilon \implies |f(q) - f(p)| < \varepsilon \right].$$

The function f is said to be **uniformly continuous** iff

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall p \in D\ \forall q \in D \left[|q - p| < \varepsilon \implies |f(q) - f(p)| < \varepsilon \right].$$

(For continuity $\delta = \delta(p, \varepsilon)$; for uniform continuity $\delta = \delta(\varepsilon)$.)

51. The function f is said to be **Lipschitz** iff there is a constant M such that

$$|f(p) - f(q)| \leq M|p - q|$$

for all $p, q \in D$. A Lipschitz function is uniformly continuous. (Proof: $\delta = \varepsilon/M$.)

Problem 52. Let $f(x) = x^p$. Show that f is Lipschitz on every compact interval $[a, b] \subseteq (0, \infty)$. For which values of p is f uniformly continuous on $(0, \infty)$? Hint: Use the Mean Value Theorem from calculus. (See Theorem 73 below.) Theorem 53 below may also help.

Theorem 53. *Assume that D is compact and f is continuous. The f is uniformly continuous.*

Proof. Choose $\varepsilon > 0$. Then, because f is continuous, for every $p \in D$ there is a $\delta = \delta(p) > 0$ such that

$$|q - p| < \delta(p) \implies |f(q) - f(p)| < \frac{\varepsilon}{2}.$$

Let $U_p := B(p, \delta(p)/2)$. The sets U_p cover D (since $p \in U_p$), i.e. $D \subseteq \bigcup_p U_p$. As D is compact, finitely many of these sets cover D, i.e.

$$D \subseteq U_{p_1} \cup U_{p_2} \cup \cdots \cup U_{p_n}. \tag{5}$$

Define

$$\delta := \tfrac{1}{2} \min\{\delta(p_1), \delta(p_2), \ldots, \delta(p_n)\}.$$

Choose $p, q \in D$. Assume $|q - p| < \delta$. By (5) we have that $p \in U_{p_k}$ for some k. Hence

$$|p - p_k| < \delta(p_k)/2. \tag{6}$$

But $\delta \leq \delta(p_k)/2$ by its definition so

$$|q - p_k| \leq |q - p| + |p - p_k| \leq \delta + \frac{\delta(p_k)}{2} \leq \frac{\delta(p_k)}{2} + \frac{\delta(p_k)}{2} = \delta(p_k). \tag{7}$$

Hence, by the definition of $\delta(\cdot)$ and Equations (6) and (7) we have

$$|f(p) - f(q)| \leq |f(p) - f(p_k)| + |f(q) - f(p_k)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as required. □

Remark 54. A proof using the Bolzano-Weierstrass theorem instead on the Heine-Borel theorem is given in Buck page 85. The proof given here is like the proof sketched in Exercises 11 and 12 in Buck page 89.

Theorem 55. *Assume $U \subseteq \mathbf{R}^n$ is open. Then the map $f : U \to \mathbf{R}^m$ is continuous if and only if the preimage of every open subset of $\mathbf{R}^m$ is an open subset of $\mathbf{R}^n$, i.e. if and only if whenever $V \subseteq \mathbf{R}^m$ is open, the set $f^{-1}(V)$ is also open. (Buck Theorem 3 page 76.)*

Theorem 56. *Assume that $D \subseteq \mathbf{R}^n$, $E \subseteq \mathbf{R}^m$, and that $f : D \to E$ and $g : E \to \mathbf{R}^\ell$ are continuous. Then $g \circ f : D \to \mathbf{R}^\ell$ is continuous. (Buck Theorem 5 page 78.)*

Theorem 57. *The continuous image of a compact set is compact: If $f : D \to \mathbf{R}^m$ is continuous and D is compact, then $f(D)$ is compact. (Buck Theorem 13 on page 93.)*

Corollary 58. *The continuous image of compact set is bounded. (Buck Theorem 10 on page 90.)*

Theorem 59. *If $f : D \to \mathbf{R}$ is continuous and D is compact, then f assumes its maximum on D, i.e. there exists $p \in D$ such that $f(q) \leq f(p)$ for all $q \in D$. Similarly for the minimum. (Buck Theorem 11 page 91.)*

Theorem 60. *The continuous image of a connected set is connected: If $f : D \to \mathbf{R}^m$ is continuous and D is connected, then $f(D)$ is connected. (Buck Theorem 15 on page 94.)*

Corollary 61 (Intermediate Value Theorem)**.** *Assume that S is connected and that $f : S \to \mathbf{R}$ is continuous. Suppose that $a, b \in f(S)$ and that $a < c < b$. Then $c \in f(S)$. (Buck Theorem 14 on page 93.)*

Remark 62. The Intermediate Value Theorem from calculus is a special case. It says that if $f : [\alpha, \beta] \to \mathbf{R}$ is a real valued continuous function on the closed interval $[\alpha, \beta] \subseteq \mathbf{R}$, $\{a, b\} = \{f(\alpha), f(\beta)\}$, and $a \leq c \leq b$, then the equation $f(x) = c$ has a solution $x \in [\alpha, \beta]$.

Theorem 63. *Let $f : D \to E$ be continuous, one-one, and onto, and assume D (and hence by Theorem 57 also E) is compact. Then $f^{-1} : E \to D$ is continuous.*

Proof. Choose a convergent sequence $\{q_n\}_n$ in E and a let

$$q := \lim_{n\to\infty} q_n$$

be its limit. We will show that

$$f^{-1}(q) = \lim_{n\to\infty} f^{-1}(q_n); \tag{\#}$$

the Theorem will then follow by Theorem 49. By Bolzano Weierstrass and Heine Borel there is a convergent subsequence $\{f^{-1}(q_{n_k})\}_k$. Let

$$p := \lim_{k\to\infty} f^{-1}(q_{n_k})$$

be its limit. Now f is assumed to be continuous so

$$f(p) = f\left(\lim_{k\to\infty} f^{-1}((q_{n_k})\right) = \lim_{k\to\infty} f(f^{-1}((q_{n_k})) = \lim_{k\to\infty} q_{n_k} = q.$$

But $f(p) = q \implies p = f^{-1}(q)$ so $f^{-1}(q) = \lim_{k\to\infty} f^{-1}((q_{n_k})$. If (#) fails, then there is a neighborhood U of $f^{-1}(q)$ such that for every N there exists $n > N$ with $f^{-1}(q_n) \notin U$, i.e. there is a subsequence $f^{-1}(q_{m_j})$ with $f^{-1}(q_{m_j}) \notin U$. As before choose a further subsequence (again denoted $f^{-1}(q_{m_j})$ which converges and let

$$p' := \lim_{j\to\infty} f^{-1}(q_{m_j})$$

denote the limit. Then $p' \notin U$ (else we would have $f^{-1}(q_{m_j}) \in U$ for sufficiently large j) so $p' \neq p$. But as before

$$f(p') = f\left(\lim_{j\to\infty} f^{-1}(q_{m_j})\right) = \lim_{j\to\infty} f(f^{-1}(q_{m_j})) = \lim_{j\to\infty} q_{m_j} = q.$$

But now $f(p) = q = f(p')$ which contradicts the fact that f is one-one. $\square$

Example 64. Let $S := \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 1\}$ denote the unit circle in $\mathbf{R}^2$ and define $f : [0, 2\pi) \to S$ by $f(\theta) = (\cos\theta, \sin\theta)$. Then f is one-one onto and continuous but f^{-1} is not continuous. The interval $[0, 2\pi)$ is not compact. These example shows both that a continuous one-one onto map whose inverse is not continuous. (See Remark 11.)

65. A real valued function $f : U \to \mathbf{R}$ defined on a subset U of the real numbers $\mathbf{R}$ is called **increasing** iff $x_1 \leq x_2 \implies f(x_1) \leq f(x_2)$ for all $x_1, x_2 \in U$, **decreasing** iff $x_1 \leq x_2 \implies f(x_1) \geq f(x_2)$ for all $x_1, x_2 \in U$, **monotonic** iff it is either increasing or decreasing. The function f is called **strictly increasing** iff $x_1 < x_2 \implies f(x_1) < f(x_2)$ for all $x_1, x_2 \in U$, **strictly decreasing** iff $x_1 < x_2 \implies f(x_1) > f(x_2)$ for all $x_1, x_2 \in U$, **strictly monotonic** iff it is either strictly increasing or strictly decreasing.

Theorem 66. *A continuous function $f : I \to \mathbf{R}$ defined on an interval $I \subseteq \mathbf{R}$ is one-one if and only if it is strictly monotonic. When these equivalent conditions hold, the image $J = f(I)$ is again an interval and the inverse function is continuous.*

Problem 67. Prove Theorem 66. (This theorem is proved in Buck Theorem 18 page 96 and Theorem 25 page 114, but Buck assumes that the intervals are closed and bounded. This assumption can be removed.)

Theorem 68. *Let $S \subseteq \mathbf{R}^n$ be and $f : S \to \mathbf{R}^m$ be uniformly continuous. Then the function f can be continuously extended to the closure $\bar{S}$ of S. i.e. there is a continuous function $F : \bar{S} \to \mathbf{R}^m$ such that $F(p) = f(p)$ for $p \in S$. (Buck Theorem 25 on page 109.)*

Example 69. The function $f : (0, 1] \to \mathbf{R}$ defined by $f(x) = \sin(1/x)$ cannot be extended to a continuous function on the closure $[0, 1]$ of $(0, 1]$.

Problem 70. Fix a positive number $a \in \mathbf{R}$. The purpose of this problem is to define the exponential a^x for $x \in \mathbf{R}$. Define $a^0 := 1$ and for n a positive integer define

$$a^n := \underbrace{a \cdot a \cdots a}_{n}, \qquad a^{-n} := 1/a^n.$$

Then

(1) Prove that for every nonzero integer n there is a unique solution $b > 0$ to the equation $b^n = a$. Define $a^{1/n}$ to be this unique solution, i.e. $a^{1/n} = b \iff b^n = a$.

(2) For a rational number q define a^q by $a^q = (a^m)^{1/n}$ where $q = m/n$. Prove that this definition is independent of the choice of the integers m and n such that $q = m/n$.

(3) Prove that there is a unique continuous function

$$\mathbf{R} \to (0, \infty) : x \mapsto a^x$$

such that $a^x = a^q$ when x is a rational number q.

In your proof make clear which theorems from these notes you are appealing to. Also make your proof self contained so that a person who doesn't have access to the statement of the problem can follow it. (You needn't provide proofs for the theorems you use, but do provide references to them.) In your proof of (3) you may use the inequality

$$|a^p - a^q| \leq M|p - q|$$

which is true when $a > 1$, N is a positive integer, $1 \leq p, q \leq N+1$, and

$$M := \left(\sum_{k=1}^{N} \frac{1}{k} \right) a^N.$$

You need not prove this inequality but use calculus to show where it comes from. Hint: What is the definition of $\ln x$, e^x, and a^x used in calculus? The natural logarithm function $\ln x$ is usually defined as an integral. How do you bound an integral by a sum? How does the Mean Value Theorem (Buck Theorem 3 page 118) give inequalities like this?

9 Derivatives

Definition 71. The function $f : I \to \mathbf{R}$ is said to be **differentiable** at the point $x_0 \in I$ iff the limit

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

exists; we say that f is differentiable on a set iff it is differentiable at each point x_0 in the set. The function f' is called the **derivative** of f.

Theorem 72. *A differentiable function is continuous.*

Theorem 73 (Mean Value Theorem)**.** *Suppose that f is differentiable of (a, b) and continuous on $[a, b]$. The there exists $c \in (a, b)$ such that*

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

Corollary 74. *Assume that f is differentiable on I. Then the derivative f' vanishes identically on I if and only if f is constant on I.*

10 The Integral

75. A **partition** of the closed interval $[a, b]$ is an increasing finite sequence $\{x_k\{_{0 \leq k \leq n}$ with $x_0 = a$ and $x_n = b$, For any bounded function f defined on $[a, b]$ and any partition $P = \{x_k\}_{0 \leq k \leq n}$ of $[a, b]$ define the **upper sum** by

$$\bar{S}(f, P) := \sum_{k=1}^{n} \bar{f}_k (x_k - x_{k-1}), \qquad \bar{f}_k := \sup\{f(x) : x_{k-1} \leq x \leq x_k\},$$

and the lower sum by

$$\underline{S}(f,P) := \sum_{k=1}^{n} \underline{f}_k(x_k, x_{k-1}), \qquad \underline{f}_k := \inf\{f(x) : x_{k-1} \le x \le x_k\}.$$

Theorem 76. *Assume that* $f : [a,b] \to \mathbf{R}$ *is continuous. Then there is a unique number* $\int_a^b f$ *called the* **definite integral** *of* f *on the interval* $[a,b]$ *such that the inequality*

$$\underline{S}(f,P) \le \int_a^b f \le \bar{S}(f,P)$$

holds for every partition P *of the interval* $[a,b]$.

Theorem 77. *The definite integral satisfies the following properties.*

(Constant) $\displaystyle\int_a^b 1 = (b-a).$

(Linearity). $\displaystyle\int_a^b (f+g) = \int_a^b f + \int_a^b g, \qquad \int_a^b cf = c\int_a^b f$ *for* $c \in \mathbf{R}$.

(Additivity). $\displaystyle\int_a^b f + \int_b^c f = \int_a^c f.$

(Order Preservation). *If* $f(x) \le g(x)$ *for all* $x \in [a,b]$, *then* $\displaystyle\int_a^b f \le \int_a^b g.$

(Triangle Inequality). $\displaystyle\left|\int_a^b f\right| \le \int_a^b |f|.$

Theorem 78 (Fundamental Theorem of Calculus). *Assume that* I *is an open interval,* $a \in I$, *and that* $F(x)$ *is defined by*

$$F(x) := \int_a^x f.$$

The F *is differentiable on* $I \cap (a,\infty)$ *and its derivative is* f. *Hence*

$$\int_a^b f = F(b) - F(a)$$

for $b \ge a$. *The number* $\int_a^b f$ *is called the* **integral** *of* f *on the interval* $[a,b]$.

Remark 79. Henceforth we use the more traditional notation $\int_a^b f(x)\,dx$ for the integral. The reader is reminded that the variable x in this expression is a **dummy variable**, i.e.

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt.$$

11 Taylor's Formula

Theorem 80 (Taylor's Formula – Lagrange Form)**.** *Let I be an interval, $a \in I$, and $f : I \to \mathbf{R}^m$ be of class C^{n+1}. Then*

$$f(x) = P_n(a,x) + R_n(a,x)$$

for any $x \in I$ where

$$P_n(a,x) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k, \qquad R_n(a,x) := \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt.$$

The polynomial $P_n(a,x)$ is called the **Taylor polynomial** *of degree n at a and $R_n(a,x)$ is called the nth* **Taylor remainder***.*

Proof. By induction on n. For $n = 0$ this is the Fundamental Theorem of Calculus. We assume the formula for n and integrate by parts:

$$\begin{aligned} v &= -\frac{(x-t)^{n+1}}{(n+1)!}, & u &= f^{(n+1)}(t), \\ dv &= \frac{(x-t)^n}{n!}\,dt, & du &= f^{(n+2)}(t)\,dt, \end{aligned}$$

$$\begin{aligned} R_n(a,x) &= \int_a^x u\,dv \\ &= uv\Big|_a^x - \int_a^x v\,du \\ &= \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,dt \\ &= \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + R_{n+1}(a,x). \end{aligned}$$

Adding $P_n(a,c)$ to both sides gives $f(x) = P_{n+1}(a,x) + R_{n+1}(a,x)$. $\square$

Corollary 81. *If* $|f^{(n+1)}(t)| \leq M$ *for* $t \in I$ *then*

$$|R_n(a,x)| \leq \frac{M|x-a|^{n+1}}{(n+1)!}$$

for $x \in I$.

Proof. Assume that $x > a$. (The case $x < a$ is similar.)

$$\begin{aligned} |R_n(a,x)| &= \left| \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \, dt \right| \\ &\leq \int_a^x \left| \frac{f^{(n+1)}(t)}{n!}(x-t)^n \right| \, dt \\ &\leq \int_a^x M \frac{(x-t)^n}{n!} = M \frac{(x-a)^{n+1}}{(n+1)!}. \end{aligned}$$ □

Remark 82. In Math 222 it is shown that that there is a number c between a and x such that

$$R_n(a,x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

This form of the remainder has the advantage that it is easy to remember: the remainder is the next term in the series with $f^{(n+1)}(a)$ replaced by $f^{(n+1)}(c)$. However, this version of the theorem only holds when f is real valued, i.e. when $m = 1$. For $m > 1$ there will be a different value of c for each component of f.

12 Series

83. A sequence determines a series and a series determines a sequence. More precisely, a sequence $\{a_k\}_k$ determines a **series** whose **partial sums** are

$$S_n = \sum_{k=1}^{n} a_k := a_1 + a_2 + \cdots + a_n,$$

and the **terms** of the series may be recovered from the sequence of partial sums via the formula

$$a_n = S_n - S_{n-1}, \qquad a_1 = S_1.$$

Convergence of the series is synonymous with convergence of the sequence of partial sums:

$$\sum_{k=1}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

Theorem 84. *If the series $\sum_k a_k$ converges, then the nth term converges to zero. (Buck Theorem 2 page 230.)*

Example 85. The nth partial sum

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + \cdots + x^n$$

of the **geometric series** is easy to compute:

$$(1-x)\sum_{k=0}^{n} x^k = \sum_{k=0}^{n} (x^k - x^{k+1}) = 1 - x^{n+1}$$

(as the sum telescopes) so dividing by $(1-x)$ gives

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$$

Hence if $|x| < 1$ we have the formula

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

for the infinite sum.

Theorem 86 (Cauchy Convergence Criterion for Series)**.** *A series $\sum_k a_k$ converges if and only if*

$$\lim_{m,n\to\infty} \sum_{k=m+1}^{n} a_k = 0$$

(Buck Theorem 3 page 265.)

Proof. This follows immediately from the formula

$$\sum_{k=m+1}^{n} a_k = S_n - S_m, \qquad S_n := \sum_{k=1}^{n} a_k, \qquad \sum_{k=m+1}^{n} a_k := a_{m+1} + \cdots + a_n.$$

and Theorem 43 (the Cauchy Convergence Criterion for Sequences) □

87. The notation

$$\sum_{k=1}^{\infty} a_k = \infty$$

means that for every $M > 0$ there exists an integer $N > 0$ such that $\sum_{k=1}^{n} a_k > M$ for $n > N$. If $a_k \geq 0$ for all k then the sequence of partial sums is monotonic increasing so by Theorem 40 either the limit $\sum_{k=1}^{n} a_k$ exists (i.e. the sequence of partial sums is bounded) or $\sum_{k=1}^{\infty} = \infty$ (i.e. the sequence of partial sums is unbounded).

Definition 88. The series $\sum_k a_k$ is said to **converge absolutely**

$$\sum_{k=1}^{\infty} |a_k| < \infty.$$

A series which converges but does not converge absolutely is said to **converge conditionally**.

Theorem 89. *A series which converges absolutely converges.*

Proof. This is an immediate consequence of the inequality

$$\left| \sum_{k=m+1}^{n} a_k \right| \leq \sum_{k=m+1}^{n} |a_k|$$

and Theorem 86. □

Theorem 90 (Comparison Test)**.** *If* $0 \leq a_k \leq b_k$ *and the series* $\sum_k b_k$ *converges, then the series* $\sum_k a_k$ *does. (Buck Theorem 5 page 231.)*

Theorem 91 (Integral Test)**.** *Assume that* $a_k = f(k)$ *where* $f : [1, \infty) \to [0, \infty)$ *is monotonic decreasing. Then the improper integral* $\int_1^{\infty} f(x)\, dx < \infty$ *converges if and only if the sum* $\sum_{k=1}^{\infty} a_k < \infty$ *does. (Buck Theorem 10 page 233.)*

Theorem 92 (Root Test)**.** *Let*

$$r = \limsup_{k \to \infty} |a_k|^{1/k}$$

Then the series $\sum_k a_k$ *converges absolutely if* $r < 1$ *and diverges (does not converge) if* $r > 1$*. If* $a_k = 1/k$*, the* $r = 1$ *but* $\sum_k a_k = \infty$*. If* $a_k = 1/k^2$*, the* $r = 1$ *and* $\sum_k a_k < \infty$*. (Buck Theorem 9 page 232.)*

Remark 93. The proofs the convergence tests tell us how to estimate the error, i.e. the difference between a partial sum and the infinite sum. For example, by the Integral Test, the series $\sum_k 1/k^2$ converges and

$$\left|\sum_{k=1}^{\infty}\frac{1}{k^2}-\sum_{k=1}^{n}\frac{1}{k^2}\right| = \left|\sum_{k=n+1}^{\infty}\frac{1}{k^2}\right| \le \int_n^\infty \frac{dx}{x^2} = \frac{1}{n}.$$

Similarly if $r < 1$ in the Root Test and $r < \rho < 1$ there is an N such that $|a_k| < \rho^k$ for $k > N$ and hence

$$\left|\sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n} a_k\right| = \sum_{k=n+1}^{\infty} |a_k| \le \sum_{k=n+1}^{\infty} \rho^k = \frac{\rho^{n+1}}{1-\rho} \le \frac{\rho^N}{1-\rho}$$

for $n > N$.

Problem 94. Show that for $p > 1$ the series $\sum_{k=1}^{\infty} k^{-p}$ converges and that the estimate

$$\left|\sum_{k=1}^{\infty} k^{-p} - \sum_{k=1}^{n} k^{-p}\right| \le \frac{(n+1)^{1-p}}{p-1}$$

holds for the difference (error) between the nth partial sum and the limit.

Problem 95. (Alternating harmonic series) Show that the series $\sum_{k=1}^{\infty}(-1)^k/k$ converges conditionally (but not absolutely) to $-\ln 2$. Hint: Use Taylor's Theorem. See Buck section 3.5 page 147. (Warning: There is a mistake in the formula (3-37) for $R_n(x)$.) You will need to show that the absolute value of the integrand is $\le 1/n$.

Definition 96. A series $\sum_{k=1}^{\infty} b_k$ is said to be a **rearrangement** of the series $\sum_{k=1}^{\infty} a_k$ iff there is a one-one correspondence $\sigma : \mathbf{Z}_+ \to \mathbf{Z}_+$ such that $b_k = a_{\sigma(k)}$.

Theorem 97. **(1)** *Any rearrangement of an absolutely convergent series converges absolutely to the same limit.*

(2) *Assume that the series* $\sum_{k=1}^{n} a_k$ *converges conditionally and* $L \in \hat{\mathbf{R}}$ *(see paragraph 30). Then there is a rearrangement* $\sum_{k=1}^{\infty} b_k$ *of the series* $\sum_{k=1}^{n} a_k$ *such that*

$$\sum_{k=1}^{\infty} b_k = L.$$

Proof. Part (1) is Theorem 13 page 239 of Buck. Part (2) is proved on pages 238-9 of Buck in the special case where $a_n = (-1)^n/n$ and $L = 10$; the general argument is much the same but uses Theorem 84. $\square$

13 Uniform Convergence

Definition 98. A sequence $\{f_n : U \to V\}$ of functions is said to **converge pointwise** to the function $f : U \to \mathbf{R}$ iff $\lim_{n\to\infty} f_n(p) = f(p)$ for all $p \in D$, i.e. iff

$$\forall p \in D\, \forall \varepsilon > 0\, \exists N\, \forall n\, \big[n > N \implies |f_n(p) - f(p)| < \varepsilon\big].$$

A sequence $\{f_n : U \to V\}_n$ of functions is said to **converge uniformly** to the function $f : U \to V$ iff $\lim_{n\to\infty} \sup_{p\in U} |f_n(p) - f(p)| = 0$, i.e. iff

$$\forall \varepsilon > 0\, \exists N\, \forall p \in U\, \forall n\big[n > N \implies |f_n(p) - f(p)| < \varepsilon\big].$$

For a sequence $\{u_k : D \to \mathbf{R}^m\}_k$ of functions the series $\sum_k u_k$ of functions is said to converge pointwise or uniformly iff the sequence $f_n = \sum_{k=0}^n u_k$ of partial sum does.

Example 99. The sequence $f_n : [0, 1] \to \mathbf{R}$ defined by $f_n(x) = x^n$ converges pointwise but not uniformly to the function

$$f(x) = \begin{cases} 0 & \text{for } 0 \le x < 1 \\ 1 & \text{for } x = 1. \end{cases}$$

Theorem 100. *If the sequence $\{f_n : U \to V\}_n$ of functions converges uniformly to the function $f : U \to V$ and each f_n is continuous, then the limit f is also continuous. (Buck Theorem 3 page 266.)*

Theorem 101 (Weierstrass Comparison Test)**.** *Assume that the functions $u_k : U \to \mathbf{R}^m$ satisfy $|u_k(p)| \le M_k$ where $\sum_k M_k < \infty$. Then the series $\sum_k u_k$ converges uniformly. (Buck Theorem 2 page 266.)*

Theorem 102. *Assume that the sequence of continuous functions $\{f_n : [a, b] \to \mathbf{R}\}_n$ converges uniformly to a function f. Then the limit of the integrals is the integral of the limit, i.e.*

$$\lim_{n\to\infty} \int_a^b f(x)\, dx = \int_a^b f(x)\, dx.$$

Corollary 103. *Let I be an open interval and $\{f_n : I \to \mathbf{R}\}_n$ be a sequence of differentiable functions. Assume that the sequence $\{f_n\}_n$ converges uniformly to $f : I \to \mathbf{R}$ and that the sequence $\{f'_n\}_n$ of derivatives also converges uniformly. Then the limit f is differentiable and the limit of the derivatives is the derivative of the limit, i.e.*

$$\lim_{n\to\infty} f'_n(x) = f'(x)$$

for $x \in I$.

Problem 104. Assume that the sequence $\{b_n\}_n$ is eventually bounded by the sequence $\{n^{-p}\}_n$ i.e there is an N such that

$$|b_n| < n^{-p}$$

for $n > N$.

(1) Show that, if $p > 1$, the series

$$f(x) := \sum_{n=1}^{\infty} b_n \sin nx \tag{104-1}$$

converges uniformly.

(2) Show that, if $p > 1$, then

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx)\, dx. \tag{104-2}$$

(3) Show that, if $p > 2$, then f is differentiable and that

$$f'(x) := \sum_{n=1}^{\infty} n b_n \cos nx.$$

Hint: See Problem 94 above. You may use any of the theorems stated above but state which theorems you are using and verify that the hypotheses of the theorems are satisfied.

Problem 105. Continue the notation of Problem 104. Show that the series

$$u(t,x) := \sum_{n=1}^{\infty} e^{-n^2 t} b_n \sin nx \tag{105-1}$$

converges uniformly on $[0,\infty) \times [0,\pi]$ if $p > 1$ and that the limit satisfies the partial differential equation[1]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{105-2}$$

on the open set $(0,\infty) \times (0,\pi)$. Show also that u is continuous on the closed set $[0,\infty) \times [0,\pi]$, that it satisfies the initial condition

$$u(0,x) = f(x) \tag{105-3}$$

and the boundary condition

$$u(t,0) = u(t,\pi) = 0. \tag{105-4}$$

(You may use any of the theorems stated above but state which theorems you are using and verify that the hypotheses of the theorems are satisfied. Hint: $e^{-n^2 t}$ is very small if $t > 0$ and n large.) It looks like this exercise proves that the solution of the partial differential equation (105-2) subject to the intial condition (105-3) and the boundary condition (105-4) is given by (105-1) where the coeffcients are defined by (104-2). Is there anything missing for a rigorous proof?

14 Power Series

106. A series of form $\sum_{k=0}^{\infty} c_k (x-a)^k$ is called a **power series** centered at a. The **radius of convergence** of the power series is the number R defined by

$$\frac{1}{R} := \limsup_{k\to\infty} |c_k|^{1/k}.$$

(If the $\limsup$ is infinite, then $R := 0$ and if the $\limsup$ is zero, then $R := \infty$.)

---

[1] This PDE is called the **Heat Equation**.

Problem 107. (A formula for the radius of convergence). Assume that the coefficients c_k are nonzero. Show that

$$\limsup_{k\to\infty} |c_k|^{1/k} = \lim_{k\to\infty} \frac{|c_{k+1}|}{|c_k|}$$

if the limit on the right exists.

Theorem 108. *Let* $\sum_{k=0}^{\infty} c_k(x-a)^k$ *and* R *be its radius of convergence. Then the series converges if* $|x-a| < R$ *and diverges if* $|x-a| > R$*. More precisely,*

(i) *If* $0 < r < R$ *the series converges uniformly the interval* $[a-r, a+r]$*.*

(ii) *If* $|x-a| > R$ *then the* n*th term of the series is unbounded and hence does not converge to zero (so the series does not converge by Theorem 84).*

(Buck Theorem 14 page 240.)

Theorem 109. *Let* $\sum_{k=0}^{\infty} c_k(x-a)^k$ *be a power series and* R *be its radius of convergence. Denote the sum by*

$$f(x) := \sum_{k=0}^{\infty} c_k(x-a)^k$$

for $|x-a| < R$*. Then* f *is differentiable on the interval* $(a-R, a+R)$*, its derivative is given by term-wise differentiation, i.e.*

$$f'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1},$$

and the radius convergence of this last power series is also R*.*

Corollary 110 (Taylor Series)**.** *Continue the hypotheses of Theorem 109. Then* f *is infinitely differentiable on the interval* $(a-R, a+R)$*, the* j*th derivative* $f^{(j)}$ *of* f *is*

$$f^{(j)}(x) = \sum_{k=j}^{\infty} k(k-1)\cdots(k-j+1)c_k(x-a)^{k-j} = \sum_{k=j}^{\infty} \frac{k!c_k}{j!}(x-a)^{k-j},$$

so that $c_j = f^{(j)}(a)/j!$*, i.e.*

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(a)}{j!}(x-a)^j.$$