Outline for Midterm #2
Math 4318, Spring 2011

I. Single Variable Functions: Differentiation

A. Definitions and first properties

1. We recalled a few definitions concerning limits form Analysis I.

2. A function $f : (a, b) \to \mathbb{R}$ is **differentiable** at a point $p \in (a, b)$ if the limit

$$\lim_{h \to 0} \frac{f(p+h) - f(p)}{h}$$

exists. If the limit exists we denote it by $f'(p)$ and call it the **derivative of f at p.**

3. Interpreting the derivative as a slope we could also define

$$f'(p) = \lim_{t \to p} \frac{f(t) - f(p)}{t - p}.$$

4. Theorem: If f is differentiable at p then f is continuous at p.

5. Theorem: Suppose f and g are differentiable at p. Then

i. The function $f \pm g$ is differentiable at p with

$$(f \pm g)'(p) = f'(p) \pm g'(p).$$

ii. The function fg is differentiable at p with

$$(fg)'(p) = f'(p)g(p) + f(p)g'(p).$$

iii. If $g(p) \neq 0$ then the function f/g is differentiable at p and

$$(f/g)'(p) = \frac{f'(p)g(p) - f(p)g'(p)}{g^2(p)}.$$

iv. If $h(x) = c$ for some constant $c \in \mathbb{R}$ then $h'(x) = 0$ for all x.

v. If $h(x) = x$ then $h'(x) = 1$ for all x.

vi. If f is differentiable at p and g is differentiable at $f(p)$ then $g \circ f$ is differentiable at p and

$$(g \circ f)'(p) = g'(f(p))f'(p).$$

6. Discussed the example

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

It is differentiable on $\mathbb{R}$ (including 0) but its derivative is not continuous at 0.

7. The **nth order derivative of f at p** is defined to be the derivative of the $(n-1)$st derivative of f at p, if it exists. Denote the nth derivative by $f^{(n)}(p)$. With this notation we take $f^{(0)}$ to denote f.

8. A function is **continuously differentiable of order r on the interval I** if $f^{(r)}$ exists and is continuous on I. (Note this implies $f^{(k)}$ exists and is continuous on I for all $k \leq r$.) We denote the set of continuously rth order differentiable functions on I by $C^r(I)$. The set $C^\infty(I)$ denotes functions whose derivative of all orders exist on I.

9. Clearly $C^r(I) \supset C^{r+1}(I)$. This inclusion is strict (that is for any r there are functions that are rth order continuously differentiable that are not continuously differentiable to order $r+1$). The first few examples are $f(x) = |x|, g(x) = x|x|, h(x) = |x|^3$...

B. The Mean Value Theorem

1. **Theorem (Mean value theorem):** If $f : [a, b] \to \mathbb{R}$ is continuous and f is differentiable on (a, b) then there is a point $c \in (a, b)$ such that
$$f(b) - f(a) = f'(c)(b - a).$$
2. **Theorem:** If $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$ and f has a local maximum of minimum at c then $f'(c) = 0$.
3. **Corollary:** If f is differentiable on (a, b) and there is some K such that $|f'(x)| leq K$ for all $x \in (a, b)$ then
$$|f(x) - f(y)| \leq K|x - y|$$
for all $x, y \in [a, b]$. (In particular f is Lipschitz.)
4. Suppose that f is differentiable on (a, b)
   i. If $f'(x) \geq 0$ for all x then f is increasing.
   ii. If $f'(x) \leq 0$ for all x then f is decreasing.
   iii. If $f'(x) > 0$ for all x then f is strictly increasing.
   iv. If $f'(x) < 0$ for all x then f is strictly decreasing.
5. **Corollary (Intermediate value theorem for derivatives):** Suppose that $f : [a, b] \to \mathbb{R}$ is differentiable. Given any λ between $f'(a)$ and $f'(b)$ there is some point $c \in (a, b)$ such that $f'(c) = \lambda$.
6. **Corollary (Inverse function theorem):** Suppose that $f : (a, b) \to \mathbb{R}$ is a differentiable function and $f'(x) \neq 0$ for all $x \in (a, b)$. Then f is a bijection onto its image. The inverse of f is continuous and
$$(f^{-1})'(y) = \frac{1}{f(x)}$$
where $y = f(x)$.
7. **Theorem:** Let f and g be two continuous functions on $[a, b]$ that are differentiable on (a, b). There is a point $c \in (a, b)$ such that
$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$
8. **Corollary (L'Hopital's Rule):** Let f, g be differentiable functions on (a, b). Suppose $f(x) \to 0$ and $g(x) \to 0$ as $x \to b$ and $g(x)$ and $g'(x)$ are not zero near b. If
$$\lim_{x \to b} \frac{f'(x)}{g'(x)} = L$$
then
$$\lim_{x \to b} \frac{f(x)}{g(x)} = L.$$
There are of course many other cases of L'Hopital's rule.

C. Taylor Polynomials

1. Let $f : [a, b] \to \mathbb{R}$ be a function and $c \in (a, b)$. If $f(c), f'(c), \ldots, f^{(n)}(c)$ exist then the **nth order Taylor polynomial of f at c** is
$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k.$$
2. **Theorem:** Let $f : [a, b] \to \mathbb{R}$ be a function such that $f', f'', \ldots, f^{(n+1)}$ all exists on (a, b). Let $P_n(x)$ be the nth order Taylor polynomial of f at c. Then for each $x \in [a, b]$ there is some point t between c and x such that
$$f(x) = P_n(x) + \frac{f^{(n+1)}(t)}{(n+1)!}(x - c)^{n+1}.$$

II. Single Variable Functions: Integration

A. Riemann integrability

1. A **partition** of an interval $[a, b]$ is a finite collection of points $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ such that $x_i < x_{i+1}$, $x_0 = a$ and $x_n = b$. The **intervals of a partition** are $I_i = [x_{i-1}, x_i]$ for $i = 1, \dots n$. The **lengths of a partition** are $\Delta x_i = x_i - x_{i-1}$. The **size of the partition** $\mathcal{P}$ is $\|\mathcal{P}\| = \max\{\Delta x_1, \dots \Delta x_n\}$.
2. A **tagged partition** $\mathcal{P}^t$ is a partition $\mathcal{P} = \{x_0, \dots, x_n\}$ together a choice of point t_i in each interval $[x_{i-1}, x_i]$.
3. If $f : [a, b] \to \mathbb{R}$ is a function and $\mathcal{P}$ is a tagged partition of $[a, b]$ then the **Riemann sum of f associated to $\mathcal{P}^t$** is
$$S(f, \mathcal{P}^t) = \sum_{i=1}^{n} f(t_i)\Delta x_i.$$
We say f is **Riemann integrable** if there is come number $I \in \mathbb{R}$ such that for every $\epsilon > 0$ there is some $\delta > 0$ such that for any tagged partition $\mathcal{P}^t$ of $[a, b]$ with $\|\mathcal{P}^t\| < \delta$ we have
$$|S(f, \mathcal{P}^t) - I| < \epsilon.$$
Let $\mathcal{R}([a, b])$ be the set of Riemann integrable functions on $[a, b]$.
4. **Lemma:** If $f \in \mathcal{R}([a, b])$ then the I in the definition above is uniquely determined.
5. If $f \in \mathcal{R}([a, b])$ then the **Riemann integral of f over** $[a, b]$ is the number I in the definition above. We denoted this number by
$$\int_a^b f(x)\,dx.$$
6. **Proposition:**
   - **i.** $\mathcal{R}([a, b])$ is a vector space.
   - **ii.** The map
   $$\mathcal{R}([a, b]) \to \mathbb{R} : f \mapsto \int_a^b f(x)\,dx$$
   is a linear map.
   - **iii.** The constant function $f(x) = k$ is in $\mathcal{R}([a, b])$ for any $[a, b]$ and
   $$\int_a^b k\,dx = k(b - a)$$
   - **iv.** If $f(x) \leq g(x)$ and $f, g \in \mathcal{R}([a, b])$ then
   $$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$
7. **Theorem:** If $f \in \mathcal{R}([a, b])$ then f is bounded on $[a, b]$.

B. Darboux integrability

1. Let $f : [a, b] \to \mathbb{R}$ be a bounded function and $\mathcal{P} = \{x_0, \dots, x_n\}$ a partition of $[a, b]$. For each $i = 1, \dots, n$ set
$$M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\} \quad \text{and} \quad m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}.$$
The **lower sum of f associated to the partition** $\mathcal{P}$ is
$$L(f, \mathcal{P}) = \sum_{i=1}^{n} m_i \Delta x_i$$

and the **upper sum** is

$$U(f,\mathcal{P}) = \sum_{i=1}^{n} M_i \Delta x_i.$$

The **lower integral of f over** $[a,b]$ is

$$\underline{\int_a^b} f(x)\,dx = \sup\{L(f,\mathcal{P}) : \mathcal{P} \text{ a partition of } [a,b]\}$$

and the **upper integral of f over** $[a,b]$ is

$$\overline{\int_a^b} f(x)\,dx = \int\{U(f,\mathcal{P}) : \mathcal{P} \text{ a partition of } [a,b]\}.$$

We say f is **Darboux integrable on** $[a,b]$ if

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx.$$

2. **Theorem:** A function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if it is Darboux integrable. And if it is integrable then

$$\int_a^b f(x)\,dx = \overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx.$$

3. We say a partition $\mathcal{P}'$ **refines** a partition $\mathcal{P}$ if $\mathcal{P} \subset \mathcal{P}'$.
4. **Lemma:** If $\mathcal{P}'$ refines $\mathcal{P}$ then

$$L(f,\mathcal{P}) \le L(f,\mathcal{P}') \le U(f,\mathcal{P}') \le U(f,\mathcal{P}).$$

5. **Lemma:** A bounded function $f : [a,b] \to \mathbb{R}$ is Darboux integrable if and only if for every $\epsilon > 0$ there is some partition $\mathcal{P}$ such that

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon.$$

6. **Corollary:** A bounded function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if for every $\epsilon > 0$ there is some partition $\mathcal{P}$ such that

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon.$$

7. **Corollary:** A continuous function $f : [a,b] \to \mathbb{R}$ is Riemann integrable.

C. Sets of measure zero and the Riemann-Lebesgue theorem

1. A set $S \subset \mathbb{R}$ has **measure zero** or is **a set of measure zero** if for every $\epsilon > 0$ there is a countable collection of intervals (a_i, b_i) such that $S \subset \cup(a_i, b_i)$ (that is the intervals are a cover of S) and

$$\sum (b_i - a_i) \le \epsilon.$$

2. **Lemma:**
   - **i.** A finite set has measure zero.
   - **ii.** A subset of a set of measure zero has measure zero.
   - **iii.** A countable union of sets of measure zero has measure zero.
   - **iv.** A countable set has measure zero.
   - **v.** The middle thirds Cantor set has measure zero.
3. A function $f : [a,b] \to \mathbb{R}$ is **continuous almost everywhere** if the set of points at which f is discontinuous has measure zero.
4. **Theorem (the Riemann-Lebesgue Theorem):** Function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if it is bounded and continuous almost everywhere.

5. **Corollary:** Every continuous and bounded piecewise continuous function on $[a, b]$ is Riemann integrable.
6. **Corollary:** If $f : [a, b] \to \mathbb{R}$ is Riemann integrable and $g : [c, d] \to \mathbb{R}$ is continuous with $f([a, b]) \subset [c, d]$ then $g \circ f$ is Riemann integrable.
7. **Corollary:** If $f \in \mathcal{R}([a, b])$ then $|f| \in \mathcal{R}([a, b])$ and

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx$$

8. **Corollary:** The product of Riemann integrable functions is Riemann integrable.
9. **Corollary:** If $a < c < b$ then $f \in \mathcal{R}([a, b])$ if and only if $f \in \mathcal{R}([a, c])$ and $f \in \mathcal{R}([c, b])$. Moreover,

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

10. **Corollary:** Suppose $f : [a.b] \to \mathbb{R}$ is non-negative (that is $f(x) \geq 0$) and integrable. Then $\int_a^b f(x)\, dx = 0$ implies that $f(x) = 0$ at every point x where f is continuous.

D. Fundamental theorem of calculus

1. **Theorem (Fundamental theorem of calculus I):** Let $f \in \mathcal{R}([a, b])$ and set

$$F(x) = \int_a^x f(t)\, dt$$

for all $x \in [a, b]$. Then the function F is continuous on $[a, b]$ and if f is continuous at c then F is differentiable at c and $F'(c) = f(c)$.
2. **Corollary:** Any continuous function $f : [a, b] \to \mathbb{R}$ has han anti-derivative (*i.e.* a function F such that $F'(x) = f(x)$ for all $x \in [a, b]$).
3. **Theorem (Fundamental theorem of calculus II):** If f is differentiable on $[a, b]$ and f' is integrable on $[a, b]$ then

$$\int_a^b f'(x)\, dx = f(b) - f(a).$$

4. **Theorem (Integration by parts):** Let f and g be functions on $[a, b]$ for which f' and g' are both integrable. Then

$$\int_a^b f'(x)g(x)\, dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x)\, dx.$$

5. **Theorem (Change of variables):** Let $\phi : [c, d] \to \mathbb{R}$ have continuous derivative and $f : [a, b] \to \mathbb{R}$ be continuous with $[a, b] \subset \phi([c, d])$. Then

$$\int_c^d f(\phi(t))\phi'(t)\, dt = \int_{\phi(c)}^{\phi(d)} f(x)\, dx.$$

E. Improper integrals

1. If $f : (a, b] \to \mathbb{R}$ is a function and $c \in (a, b]$ then if f is integrable on $[c, b]$ set

$$I_c = \int_c^b f(x)\, dx.$$

Define the **improper integral of f on** $[a, b]$ to be

$$\int_a^b f(x)\, dx = \lim_{c \to a^+} I_c$$

if the limit exists. (One can similarly define the improper integral of $f : [a, b) \to \mathbb{R}$.)

2. If $f : [a, b] \to \mathbb{R}$ is integrable then this definition agrees with the Riemann integral of f.

III. Sequences of functions and function spaces

A. Sequences of functions

1. If $S \subset \mathbb{R}$ and $\{f_n : S \to \mathbb{R}\}$ is a sequence of functions on f (that is if $\mathcal{F}(S, \mathbb{R})$ is the set of all functions from S to $\mathbb{R}$, then $\{f_n\}$ is a sequence in the set $\mathcal{F}(S, \mathbb{R})$), the we say the sequence **converges point wise to** f if for each $x \in S$ we have the sequence of numbers $\{f_n(x)\}$ converges to $f(x)$. That is for all $x \in S$ and $\epsilon > 0$ there is some N such that $|f(x) - f_n(x)| < \epsilon$ for all $n \geq N$.

2. We say that a sequence of functions $\{f_n : S \to \mathbb{R}\}$ **converges uniformly to** f **on** S if for all $\epsilon > 0$ there is a N such that $|f(x) - f_n(x)| < \epsilon$ for all $n \geq N$ and $x \in S$.

3. Lemma: The sequence $\{f_n : S \to \mathbb{R}\}$ converges uniformly to some function on S if and only if for all $\epsilon > 0$ there is an N such that $|f_n(x) - f_m(x)| < \epsilon$ for all $n, m \geq N$ and $x \in S$.

4. Theorem: If $f_n \to f$ uniformly on S and the f_n are continuous at $c \in S$ then f is continuous at c.

5. Theorem: If $\{f_n|\}$ is a sequence in $\mathcal{R}([a, b])$ and $f_n \to f$ uniformly on $[a, b]$ then $f \in \mathcal{R}(a, b])$ and

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

6. Theorem: Let $\{f_n : [a, b] \to \mathbb{R}\}$ be a sequence of functions. Suppose that

- **i.** there is some $c \in [a, b]$ such that $\{f_n(c)\}$ converges,
- **ii.** the functions f_n are all differentiable on $[a, b]$, and
- **iii.** $f_n' \to g$ uniformly on $[a, b]$ for some function g.

Then there is some function f such that $f_n \to f$ uniformly on $[a, b]$ and $f' = g$.

B. Metric space topology

1. Let X be a set. A **metric** on X is a function

$$d : X \times X \to \mathbb{R}$$

such that

- **i.** $d(p, q) \geq 0$ for all $p, q \in X$,
- **ii.** $d(p, q) = 0$ if and only if $p = q$
- **iii.** $d(p, q) = d(q, p)$, and
- **iv.** $d(p, q) \leq d(p, r) + d(r, q)$ for all $p, q, r \in X$. We think of $d(p, q)$ as being the distance between p and q. The pair (X, d) is called a **metric space**.

2. Example: Given a norm $\|\cdot\|$ on a vector space V one has a metric associated to the norm by setting

$$d(v, w) = \|v - w\|$$

for all $v, w \in V$. Thus if $\langle \cdot, \cdot \rangle$ is an inner product on V then $\|v\| = \sqrt{\langle v, v\rangle}$ is a norm and hence we get a metric associated to the inner product too.

3. Given a metric space (X, d) most all the concepts form Analysis I can be defined for (X, d) since they only depended on some notion of distance (where there we used the standard norm on $\mathbb{R}^n$). For example

- **i.** A set $U \subset X$ is **open** if for all $p \in U$ there is some $r > 0$ such that $B_r(p) \subset U$ where $B_r(p) = \{x \in X : d(x, p) < r\}$.
- **ii.** A **neighborhood** of a point $p \in X$ is an open set N containing p.
- **iii.** A subset $S \subset X$ has $p \in X$ as a **cluster point** if for all neighborhoods N of p we have $(N - \{p\}) \cap S \neq \emptyset$. (This is equivalent to saying for all $\epsilon > 0$ there is some $q \in S$ with $q \neq p$ and $d(q, p) < \epsilon$.)

iv. A set S in X is **closed** if it contains all of its cluster points.

v. A set S is bounded if there is some $p \in X$ and R such that $S \subset B_r(p)$.

vi. A sequence $\{x_n\}$ in X is **Cauchy** if for all $\epsilon > 0$ there is some N such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq N$.

Most all the other concepts can be rephrased in terms of d too.

4. A metric space is complete if every Cauchy sequence converges.

5. Theorem (from Analysis I): The space $(\mathbb{R}^n, d)$ is complete. (Here $d(x,y) = \sqrt{\sum (x_i - y_i)^2}$.)

6. Suppose V is a vector space and d is a metric on V that is complete. If d comes from a norm then (V, d) is called a **Banach space**. If d comes from an inner product then (V, d) is called a **Hilbert space**.

C. Function spaces

1. Let $D \subset \mathbb{R}$. We set

$$\mathcal{B}(D, \mathbb{R}) = \{\text{bounded functions } f : D \to \mathbb{R}\}$$

and for $f \in \mathcal{B}(D, \mathbb{R})$ let

$$\|f\|_\infty = \sup\{|f(x)| : x \in D\}.$$

We call this the **sup-norm** or **uniform norm** on $\mathcal{B}(D, \mathbb{R})$ and one easily checks that this is a norm.

2. Lemma: Let $\{f_n\}$ be a sequence in $\mathcal{B}(D, \mathbb{R})$. Then the following are equivalent

i. $f_n \to f$ uniformly on D.

ii. f_n converges to f in the sup-norm.

iii. $\|f_n - f\|_\infty \to 0$ as $n \to \infty$.

3. Theorem: The sequence $\{f_n\}$ in $\mathcal{B}(D, \mathbb{R})$ is Cauchy in the sup-norm if and only if it converges in the sup-norm. That is $(\mathcal{B}(D, \mathbb{R}), \|\cdot\|_\infty)$ is a Banach space.

4. Let

$$C^0_b\{D\} = \{f \in \mathcal{B}(D, \mathbb{R}) : f \text{ is continuous}\}.$$

(Notice that if D is compact then $C^0_b(D) = C^0(D)$.)

5. Theorem: $C^0_b(D)$ is a Banach space (in the sup-norm).

6. Theorem: The set $\mathcal{R}([a,b])$ (which is a subset of $\mathcal{B}([a,b], \mathbb{R})$) is a Banach space in the sup-norm and the function

$$I : \mathcal{R}([a,b]) \to \mathbb{R} : f \to \int_a^b f(x)\, dx$$

is continuous.

7. For a function in $f \in C^n([a,b])$ and any $k \leq n$ define

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \ldots + \|f^{(k)}\|_\infty.$$

It is easy to check this is a norm on $C^n([a,b])$.

8. Theorem: $(C^n([a,b]), \|\cdot\|_{C^n})$ is a Banach space.

D. Approximation of functions

1. Theorem (Weierstrass): Polynomials are dense in $C^0([a,b])$. (That is for every $f \in C^0([a,b])$ and $\epsilon > 0$ there is a polynomial p such that $\|f - p\| < \epsilon$.)

2. Given two functions $f, g : \mathbb{R} \to \mathbb{R}$ we define the **convolution** of f and g to be

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t)\ , dt,$$

if the integral is well defined (and it will be if, for example, one of the functions has compact support and the functions are integrable on compact intervals).

3. A sequence of functions $K_n(x)$ is called an **approximation to the identity** if

i. $K_n(x) \geq 0$ for all x and n,
ii. $\int K_n(x)\,dx = 1$ for all n, and
iii. $\int_{|x|\geq\epsilon} K_n(x)\,dx \to 0$ as $n \to \infty$ for any $\epsilon > 0$.

4. Lemma: If $\{K_n\}$ is an approximation to the identity and f is a compactly supported continuous function then $(f * K_n) \to f$ uniformly.

5. Lemma: If p is a polynomial and $f \in \mathcal{R}([a,b])$ has compact support then $p * f$ is a polynomial.

E. Fixed point theorems and differential equations

1. Theorem (Contraction mapping theorem): Let (M,d) be a complete metric space and $f : M \to M$ a contraction mapping (that is there is some $0 \leq k < 1$ such that $d(f(x), f(y)) \leq kd(x,y)$ for all $x, y \in M$). Then there is a unique fixed point $p \in M$ for f. (That is $f(p) = p$ and p is the only point with this property.)

2. Theorem (Existence ad uniqueness of solutions to ODEs): Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ be a function on a neighborhood D of (t_0, x_0) in $\mathbb{R}^2$. Assume that

i. f is Lipschitz in the x variable (that is there is some K such that $|f(t,x) - f(t,x')| \leq K|x - x'|$) and
ii. f is continuous on D.

Then there is some $\delta > 0$ and a unique C^1 solution to the initial value problem

$$\frac{dx}{dt} = f(t,x) \qquad x(t_0) = x_0$$

on $(t_0 - \delta, t_0 + \delta)$. (That is there is some continuously differentiable function $\gamma : (t_0 - \delta, t_0 + \delta) \to \mathbb{R}$ such that $\gamma'(t) = f(t, \gamma(t))$ and $\gamma(t_0) = x_0$.)

3. A function γ solves the initial value problem if and only if it is a fixed point of

$$\Phi(\gamma)(t) = x_0 + \int_{t_0}^{t} f(s, \gamma(s))\,ds.$$

F. Compactness in function spaces

1. A subset $S \subset \mathcal{F}(D, \mathbb{R}) = \{\text{functions from } D \text{ to } \mathbb{R}\}$ is called **equicontinuous** if for every $\epsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $x, y \in D$ with $|x - y| < \delta$ and $f \in S$. (Of course sequences are subsets so we can talk about equicontinuous sequences.)

2. Theorem: A subset $S \subset \mathcal{F}(D, \mathbb{R})$ is compact (in the sup-norm) if and only if it is closed, bounded and equicontinuous.

3. Recall that a set in a metric space is called compact if every open cover of the set has a finite sub-cover. This is equivalent to saying that every sequence in the set has a convergent sub-sequence.

4. Theorem (Arzelá-Ascoli Theorem): Let $\{f_n\}$ be a sequence of functions $f_n : D \to \mathbb{R}$. If

i. D is compact,
ii. $\{f_n\}$ is bounded in the sup-norm and
iii. $\{f_n\}$ is equicontinuous (not this implies that each f_n is continuous),

then there is a sub-seqence of $\{f_n\}$ that converges uniformly on D.

5. Theorem (Peano's Theorem): Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ be a function on a neighborhood D of (t_0, x_0) in $\mathbb{R}^2$. If f is continuous on D then there is some $\delta > 0$ and a unique C^1 solution to the initial value problem

$$\frac{dx}{dt} = f(t,x) \qquad x(t_0) = x_0$$

on $[t_0, t_0 + \delta]$.

6. Notice that when you drop the hypothesis of f being Lipschitz in the x variable you loose uniqueness of solutions. For example you can see that $x'(t) = \sqrt{x(t)}$ with the initial condition $x(0) = 0$ has infinitely many solutions. Indeed for $c \geq 0$ let $\gamma_c(t) = 0$ for $t \leq c$ and $\frac{1}{4}(t-c)^2$ for $t \geq c$ and we see that γ_c solves the initial value problem.

G. Series of functions

1. Given a sequence $\{g_k\}$ of functions on a set $D \subset \mathbb{R}$ we can associated the partial sums

$$s_n(x) = \sum_{k=0}^{n} g_k(x).$$

We say the **series** $\sum_{k=0}^{\infty} g_k$ **converges uniformly, respectively point-wise** to g if the sequence of partial sums $\{s_n\}$ converges uniformly, respectively point-wise, to g. We say the series **converges absolutely** if for each $x \in D$ the series of real numbers $\sum_{k=0}^{\infty} |g_k(x)|$ converges.

2. Theorem (Weierstrass M-test): Let $\{g_k\}$ be a sequence of functions on D and M_k be constants such that $|g_k(x)| \leq M_k$ for all $x \in D$. If $\sum_{k=0}^{\infty} M_k$ converges then $\sum_{k=0}^{\infty} g_k$ converges absolutely and uniformly on D.

3. Theorem: If $\sum_{k=0}^{\infty} g_k$ converges uniformly to g on $[a,b]$ and each $g_k \in \mathcal{R}([a,b])$ then $g \in \mathcal{R}([a,b])$ and

$$\int_a^b g(x)\,dx = \int_a^b \sum_{k=0}^{\infty} g_k(x)\,dx = \sum_{k=0}^{\infty} \int_a^b g_k(x)\,dx.$$

4. Suppose that $\sum_{k=0}^{\infty} g_k$ converges point-wise on $[a,b]$ and each of the g_k is differentiable. If $\sum_{k=0}^{\infty} g_k'$ converges uniformly on $[a,b]$ then

$$\left(\sum_{k=0}^{\infty} g_k(x)\right)' = \sum_{k=0}^{\infty} g_k'(x).$$

5. Theorem: There is a continuous function $f : \mathbb{R} \to \mathbb{R}$ that is nowhere differentiable!

6. Facts we did not prove: Lipschitz functions are differentiable almost everywhere. The set of nowhere differentiable functions is dense in the set of continuous functions.

H. Power series

1. A **power series** about the point c is a series of the form

$$\sum_{k=0}^{\infty} a_n(x-c)^n.$$

2. Recall from Analysis I:

i. The **radius of convergence** is defined to be $R = \frac{1}{\rho}$ where

$$\rho = \limsup_k |a_k|^{\frac{1}{k}}.$$

ii. If $\lim_{n\to infty} \frac{|a_n|}{|a_{n+1}|}$ exists then it is equal to the radius of convergence.

iii. For any $|x-c| < R$ the series $\sum_{k=0}^{\infty} a_n(x-c)^n$ converges absolutely and for $|x-c| > R$ the series diverges.

iv. A power series converges uniformly on any compact subset of $(c-R, c+R)$.

v. A power series defines a continuous function on $(c-R, c+R)$.

3. Theorem: If R is the radius of convergence of the power series

$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$$

Then f is differentiable of all orders on $(c-R, c+R)$ and

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)a_n(x-c)^{n-k}.$$

4. **Corollary:** Suppose $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ and $g(x) = \sum_{n=0}^{\infty} b_n(x-c)^n$ converge on some interval (a,b) containing c. If $f(x) = g(x)$ for all $x \in (a,b)$ then $a_n = b_n$ for all n.
5. A function f on an open interval is called **analytic** if it can be represented as a power series with non-zero radius of convergence, about each point in its domain.
6. **Theorem:** if $f : (a,b) \to \mathbb{R}$ is C^∞ and there is some constant M such that

$$|f^{(k)}(x)| \leq M$$

for all $x \in (a,b)$ then f is analytic.
7. Given a function $f : (a,b) \to \mathbb{R}$ that is infinitely differentiable, $c \in (a,b)$ and $\delta > 0$ such that $I_{c,\delta} = [c-\delta, c+\delta] \subset (a,b)$, we can set $M^k_{c,\delta} = \sup\{|f^{(k)}(x)| : x \in I_{c,\delta}\}$. We call

$$g_{c,\delta} = \limsup_{k\to\infty} (M^k_{c,\delta}/k!)^{1/k}$$

the **derivative growth rate of** f **δ-near** c. We say f has **locally bounded derivative growth rate in** (a,b) if for all $c \in (a,b)$ there is some δ so that $g_{c,\delta}$ is finite.
8. **Theorem:** An infinitely differentiable function $f : (a,b) \to \mathbb{R}$ is analytic if and only if it has locally bounded derivative growth rate.

IV. Derivatives in higher dimensions

A. Definitions and first properties

1. A map $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable at** $c \in A$ if there is a linear map

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$\lim_{x\to c} \frac{\|f(x) - (f(c) + L(x-c)\|}{\|x-c\|} = 0.$$

If such an L exists then we call it the **derivative of** f **at** c and denote it by $Df(c)$.
2. Alternate definition: L is the derivative of f at c if and only if for all $\epsilon > 0$ there is some $\delta > 0$ such that $\|x-c\| < \delta$ implies that $\|f(x) - (f(c) + L(x-c))\| \leq \epsilon\|x-c\|$.
3. **Lemma:** If there is a linear map L as in the definition above then it is uniquely determined by f and c.
4. **Theorem:** If $f : A \to \mathbb{R}^m$ is differentiable at $c \in A$ with A and open set in $\mathbb{R}^n$, then f is continuous at c.
5. Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a function on the open set A. We can write

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)).$$

The i^{th} **partial derivative of** f_j at $(x_1, \dots, x_n)$ is

$$\frac{\partial f_j}{\partial x_i} = \lim_{h\to 0} \frac{f_j(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

(if the limit exists.
6. **Theorem:** Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a function on the open set A. If f is differentiable on A then the partial derivatives $\frac{\partial f_j}{\partial x_i}$ exist for all i and j and in the standard basis for $\mathbb{R}^n$

and $\mathbb{R}^m$ the linear map $Df(x_1,\ldots,x_n)$ is given by the matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

This matrix is called the **Jacobian matrix of f at** $(x_1,\ldots,x_n)$.

7. If $f:\mathbb{R}^n\to\mathbb{R}$ is a differentiable function then $Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n}\end{bmatrix}$. That is it is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$. This is very similar to a familiar concept: the **gradient of** f. The gradient is a vector containing the partial derivatives of f

$$\nabla f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \vdots \\ \frac{\partial f_1}{\partial x_n} \end{bmatrix}.$$

While it is common to write this vector as a row vector, we will always write it as a column vector (actually all our vectors are column vectors, so when we think of a point $x\in\mathbb{R}^n$ we will think of it as a column vector so our matrix can act on it by matrix multiplication).

8. Theorem: Let $A\subset\mathbb{R}^m$ be an open set and $f:A\to\mathbb{R}^n$. Suppose that $f=(f_1,\ldots,f_m)$ and each partial derivative $\frac{\partial f_j}{\partial x_i}$ exists and is continuous near on A. Then f is differentiable on A.

9. Let $A\subset\mathbb{R}^m$ be an open set and $f:A\to\mathbb{R}^n$. For e a unit vector in $\mathbb{R}^n$ and $c\in A$ we define the **directional derivative of** f **at** c **in the direction of** e to be

$$f'(c,e)=\lim_{h\to 0}\frac{f(c+he)-f(c)}{h}$$

if the limit exists.

10. It is simple to check that $(Df(c))e = f'(c,e)$.

B. Chain rule and product rule

1. Theorem (Chain Rule): Let $A\subset\mathbb{R}^n$ and $B\subset\mathbb{R}^m$ be open sets. Suppose that $f:A\to\mathbb{R}^m$, $g:B\to\mathbb{R}^p$ and $f(A)\subset B$. If f is differentiable at $x_0\in A$ and g is differentiable at $f(x_0)\in B$ then $g\circ f$ is differentiable at x_0 and

$$D(g\circ f)(x_0) = \left(Dg(f(x_0))\right)\left(Df(x_0)\right).$$

2. Theorem (Product Rule): Let A be a subset of $\mathbb{R}^n$, $f:A\to\mathbb{R}^m$ and $g:A\to\mathbb{R}$ be differentiable functions. THen gf is differentiable and the derivative $D(gf)(x)$ is the linear function

$$D(gf)(x)(v) = \left(Dg(x)(v)\right)f(x) + g(x)\left(Df(x)(v)\right).$$

If $h:A\to\mathbb{R}^m$ is also differentiable then $h\bullet f$ is differentiable with derivative

$$D(h\bullet f)(x)(v) = \left(Dh(x)(v)\right)\bullet f(x) + h(x)\bullet\left(Df(x)(v)\right).$$

C. Mean Value Theorem

1. Theorem (Mean Value Theorem): Let $A\subset\mathbb{R}^n$ be an open set and $f:A\to\mathbb{R}^m$ a differentiable function. Suppose there is some M such that

$$\|Df(x)(v)\|\le M\|v\|$$

for all $v\in\mathbb{R}^n$ and $x\in A$. If the line segment connecting $p,q\in A$ is contained in A then

$$\|f(p)-f(q)\|\le M\|p-q\|.$$

2. **Corollary:** Suppose $A \subset \mathbb{R}^n$ is open and any two points in A can be connected by a straight line that is contained in A. If $f : A \to \mathbb{R}^m$ satisfies $Df(x) = 0$ for all $x \in A$ then f is constant.

D. Higher Derivatives

1. If $f : \mathbb{R}^n \to \mathbb{R}$ is a function then $\frac{\partial f}{\partial x_i}$ is another function $\mathbb{R}^n \to \mathbb{R}$ so we can take a partial derivative of it. The **second order partial derivative of f with respect to x_i then x_j** is $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})$. You can similarly define higher order partial derivatives.
2. **Theorem:** If f has continuous second order partial derivatives then
$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$
3. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable then the derivative of f at x, $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$, is a linear map. So for each $x \in \mathbb{R}^n$ the derivative gives a linear map $\mathbb{R}^n \to \mathbb{R}^m$. That is
$$Df : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) : x \mapsto Df(x),$$
where $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is the set of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. By Choosing a basis for $\mathbb{R}^n$ and $\mathbb{R}^m$ we can identify $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ with $m \times n$ matrices. The set of $m \times n$ matrices can be identified with $\mathbb{R}^{nm}$. So Df gives a map $Df : \mathbb{R}^n \to \mathbb{R}^{nm}$. The **second derivative of f at** x is the derivative of this map at x and is denoted $D^2 f(x) = D(Df)(x)$. (Notice that we don't really need to identify $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ with $\mathbb{R}^{nm}$ to define the derivative, since $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a vector space we can define the derivative as normal, but using the identification above might be physiologically more acceptable on the first pass since we have not talked about derivatives of functions to a general vector space). We can similarly define higher derivatives.
4. Notice that $D^2 f(x)$ is a linear map $\mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. So $D^2 f(x)(v)$ is a linear map $\mathbb{R}^n \to \mathbb{R}^m$. So $D^2 f(x)(v)(w)$ is an element of $\mathbb{R}^m$. Thus we see that $D^2 f(x)$ can be thought of as a bilinear map $D^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$, by writing $D^2 f(x)(v, w) = D^2 f(x)(v)(w)$.
5. If $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear map then by choosing a basis $e_1, \ldots, e_n$ for $\mathbb{R}^n$ we can set $b_{ij} = B(e_i, e_j)$. Then given two vectors $v = \sum v_i e_i$ and $w = \sum w_j e_j$ we see that
$$B(v, w) = \sum b_{ij} v_i w_j.$$
So we can think of
$$B(v, w) = v^t (b_{ij}) w$$
where (b_{ij}) is an $n \times n$ matrix.
6. **Theorem:** If $A \subset \mathbb{R}^n$ is an open set and $f : A \to \mathbb{R}$ is twice differentiable then in the standard basis for $|R^n$ we can write
$$D^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$
7. We say that $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is C^r, or r **times continuously differentiable**, if all the rth order partial derivatives exist and are continuous (which implies that $D^r f$ exists and is continuous).
8. In general if $f : \mathbb{R}^n \to \mathbb{R}$ is k times differentiable then in the standard basis $e_1, \ldots e_n$ for $\mathbb{R}^n$ we have
$$D^k f(x)(v_1, \ldots v_k) = \sum_{i_1, \ldots i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} (v_1)_{i_1} \ldots (v_k)_{i_k},$$

where $v_i = \sum (v_i)_j e_j$.

9. Theorem: Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}$ be a C^{r+1} function. If $x, y \in A$ and the line between x and y is in A then there is some c on the line between x and y such that

$$f(y) = \underbrace{\sum_{k=0}^{r} \frac{1}{k!}(D^k f(x))(\underbrace{y-x, \ldots, y-x}_{k \text{ times}})}_{r\text{th order Taylor polynomial}} + \frac{1}{(r+1)!} D^{r+1} f(c)(\underbrace{y-x, \ldots y-x}_{r+1 \text{ times}}).$$

E. Minima and Maxima

1. Theorem: Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}$ a differentiable function. If x_0 is a (local) maximum or minimum of f then $Df(x_0) = 0$ (that is x_0 is a critical point).

2. A bilinear form $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called **positive definite** if $B(x, x) > 0$ for all $x \neq 0$, it is **positive semi-definite** if $B(x, x) \geq 0$ for all x. We have similar definitions for negative definite and **negative semi-definite**.

3. Theorem: Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}$ be a C^2-function.

- **i.** If $x_0 \in A$ is a critical point then $D^2 f(x_0)$ positive definite implies that x_0 is a local minimum and $D^2 f(x_0)$ negative definite implies that x_0 is a local maximum.
- **ii.** If f has a local minimum at x_0 then $D^2 f(x_0)$ is positive semi-definite and if x_0 is a local maximum then $D^2 f(x_0)$ is negative semi-definite.

4. Lemma: If $\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ represents a bilinear form $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ then it is positive definite if and only if $a > 0$ and $ad - b^2 > 0$. It is negative definite if and only if $a < 0$ and $ad - b^2 > 0$.

F. Inverse Function Theorem

1. Theorem (Inverse Function Theorem): Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}^n$ a C^1-function. If $x_0 \in A$ and $Df(x_0)$ is invertible then there is a neighborhood U of x_0 and W of $f(x_0)$ such that $f|_U : U \to W$ is a bijection and $(f|_U)^{-1}$ is differentiable with derivative

$$D(f^{-1})(y) = (Df(x))^{-1}$$

where $f(x) = y \in W$.

V. Multivariable Integration

A. The integral

1. Let $A \subset \mathbb{R}^n$ be a bounded set and $f : A \to \mathbb{R}$ a bounded function. Let $A \subset B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ (we call something of the form of B a "rectangle"). A **partition** of B, $\mathcal{P}$, is a choice of partition of the intervals $[a_1, b_1], \ldots, [a_n, b_n]$. That is, $a_1 = x_0^1 < x_1^1 < \ldots < x_{m_1}^1 = b_1$ and $a_2 = x_0^2 < \ldots < x_{m_2}^2 = b_2$ and so on. This divides B into many (precisely $m_1 m_2 \ldots m_n$) smaller rectangles $B_{i_1,\ldots,i_n} = [x_{i_1-1}^1, x_{i_1}^1] \times \cdots \times [x_{i_n-1}^n - x_{i_n}^n]$. We define the **volume** of $B_{i_1,\ldots,i_n}$ to be

$$\text{vol}(B_{i_1,\ldots,i_n}) = (b_{i_1} - b_{i_1-1}) \ldots (b_{i_n} - b_{i_n-1})$$

and the **mesh** of $\mathcal{P}$ to be the longest length of an edge of one of the $B_{i_1,\ldots,i_n}$ and denote it by $m(\mathcal{P})$.

extend f to B by defining it to be 0 at points in $B - A$. We define the **lower sum of f for $\mathcal{P}$** to be

$$L(f, \mathcal{P}) = \sum_{i_1,\ldots,i_n} \inf\{f(x) : x \in B_{i_1,\ldots i_n}\} \text{vol}(B_{i_1,\ldots i_n})$$

and the **upper sum of f for $\mathcal{P}$** to be

$$U(f,\mathcal{P}) = \sum_{i_1,\dots,i_n} \sup\{f(x) : x \in B_{i_1,\dots i_n}\}\mathrm{vol}(B_{i_1,\dots i_n})$$

2. We say a partition $\mathcal{P}'$ **refines** $\mathcal{P}$ if each rectangle $\mathcal{P}'$ defines is contained in a rectangle of $\mathcal{P}$. It is easy to see the following.

i. If $\mathcal{P}'$ refines $\mathcal{P}$ then

$$L(f,\mathcal{P}) \le L(f,\mathcal{P}') \le U(f,\mathcal{P}') \le U(f,\mathcal{P}).$$

ii. If $\mathcal{P}$ and $\mathcal{Q}$ are any partitions of B then

$$L(f,\mathcal{P}) \le U(f,\mathcal{Q}).$$

Thus the lower sums are bounded below and the upper sums are bounded above.

3. Define the **upper integral of f over** A to be

$$\overline{\int_A} f = \inf\{U(f,\mathcal{P} : \mathcal{P} \text{ is a partition of } B\}$$

and the **lower integral of f over** A to be

$$\underline{\int_A} f = \sup\{L(f,\mathcal{P} : \mathcal{P} \text{ is a partition of } B\}.$$

We say f is **integrable over** A is $\overline{\int_A} f = \underline{\int_A} f$ and in this case we define the **integral** to be the common value and denote it $\int_A f$ or $\int_A f(x)\,dx$.

4. Theorem (Darboux's Theorem): Suppose A is a bounded set and contained in the rectangle B and $f : A \to \mathbb{R}$ is a bounded function. Then f is integrable over A with integral I if and only if for all $\epsilon > 0$ there is a $\delta > 0$ such that for any partition $\mathcal{P}$ with $m(\mathcal{P}) < \delta$ and points $x_{i_1,\dots,i_n} \in B_{i_i,\dots,i_n}$ we have

$$\left|\sum f(x_{i_1,\dots,i_n})\mathrm{vol}(B_{i_1,\dots,i_n}) - I\right| < \epsilon.$$

5. Theorem (Riemann's Criterion): If f and A are as in the last theorem then f is integrable if and only if for each $\epsilon > 0$ there is a partition $\mathcal{P}$ such that $0 \le U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon$.

B. Sets of measure zero and the Lebesgue Theorem

1. A set $S \subset \mathbb{R}^n$ has **measure zero** if for every $\epsilon > 0$ S can be covered by a countable collection of rectangles $\{R_i\}$ such that $\sum_i \mathrm{vol}(R_i) < \epsilon$.

2. Lemma:

i. Countable sets have measure zero.

ii. Subsets of measure zero have measure zero.

iii. Countable unions of sets of measure zero have measure zero.

iv. If V is a linear subspace of $\mathbb{R}^n$ of dimension less than n then V has measure zero in $\mathbb{R}^n$.

3. Theorem (Lebegue's Theorem): Suppose $A \subset R^n$ is a bounded set and $f : A \to \mathbb{R}$ is bounded. Extend f to all of $\mathbb{R}^n$ by setting it to zero for all $x \notin A$. Then f is integrable on A if and only if the set of discontinuities of the extended function f has measure zero (that is the extended function is continuous almost everywhere).

4. Theorem: If $f : A \to \mathbb{R}$ is integrable and $f(x) \ge 0$ for all $x \in A$ then $\int_A f = 0$ implies that $\{x \in A : f(x) \neq 0\}$ has measure zero.

C. Properties of the integral

1. **Theorem:** Let A and B be bounded subsets of $\mathbb{R}^n$, $f, g : A \to \mathbb{R}$ be bounded, integrable functions and $c \in \mathbb{R}$.
   i. Then $f \pm g$ is integrable and
$$\int_A f \pm g = \int_A f \pm \int_A g.$$
   ii. cf is integrable and
$$\int_A cf = c \int_A f$$
   iii. $|f|$ is integrable and
$$\left| \int_A f \right| \le \int_A |f|.$$
   iv. $f \le g$ then $\int_A f \le \int_A g$.
   v. If $f : A \cup B \to \mathbb{R}$ is a function, $A \cap B$ has measure zero and $f|_A$, $f|_B$ and $f|_{A\cap B}$ are integrable then f is integrable on $A \cup B$ and
$$\int_{A\cup B} f = \int_A f + \int_B f.$$
2. If $A \subset \mathbb{R}^n$ is a bounded set then set $\chi_A(x) = 1$ if $x \in A$ and zero otherwise. Then we say A **has volume** if χ_A is integrable on A and define the **volume of** A to be $\mathrm{vol}(A) = \int_A \chi_A$.
3. **Theorem (Mean Value Theorem):** If $f : A \to \mathbb{R}$ is continuous and A is compact, connected and has bounded volume then there is a point $x_0 \in A$ such that
$$\int_A f = f(x_0)\mathrm{vol}(A).$$

D. Fubini's Theorem

1. **Theorem (Fubini's Theorem):** Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$ be rectangles and $f : A \times B \to \mathbb{R}$ be integrable. For each $x \in A$ define $f_x : B \to \mathbb{R} : y \mapsto f(x, y)$ and for each $y \in B$ define $f_y : A \to \mathbb{R} : x \mapsto f(x, y)$. If f_x is integrable over B for each $x \in A$ then
$$\int_{A\times B} f = \int_A \left(\int_B f(x, y)\, dy \right) dx.$$
If f_y is integrable over A for each $y \in B$ then
$$\int_{A\times B} f = \int_B \left(\int_A f(x, y)\, dx \right) dy.$$