Important theorems and their proofs

This is a list of some important theorems and proofs from IB Analysis & Topology. Please note that this list is not exhaustive and there may be a few inaccuracies.

Contents

List of Theorems

1 Uniform convergence and uniform continuity

Theorem 1.1: Switching summation and differentiation

Let (f_n) be a sequence of continuously differentiable functions on $[a, b]$. Assume further that

(i) $\sum_{n=1}^{\infty} f_n'(x)$ converges uniformly on $[a, b]$.

(ii) There exists $c \in [a, b]$ such that $\sum_{n=1}^{\infty} f_n(c)$ converges.

Then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to a continuously differentiable function f on $[a, b]$ and $f'(x) = \sum_{n=1}^{\infty} f_n'(x)$ for all $x \in [a, b]$.

Proof 1.1

Main idea: let $g(x) = \sum_{n=1}^{\infty} f_n'(x)$ and solve $f' = g$ with initial conditions $f(c) = \sum_{n=1}^{\infty} f_n(c)$.

- g is integrable so let $\lambda = \sum_{n=1}^{\infty} f_n(c)$ and define

$$f(x) = \lambda + \int_c^x g(t)\mathrm{d}t$$

  Then f is differentiable with $f'(x) = g$ and $f(c) = \lambda$.
- We also know

$$f_k(x) = f_k(c) + \int_c^x f_k'(t)\mathrm{d}t$$

- Now use these to bound $|f(x) - \sum_{k=1}^{n} f_k(x)|$ and conclude $\sum_{k=1}^{n} f_k(x) \to f(x)$ uniformly.

Theorem 1.2: General Principle of Uniform Convergence

If (f_n) is a uniformly Cauchy sequence of scalar functions on a set S, then (f_n) converges uniformly to some function f on S.

Proof 1.2

Main idea : we know that if it converges uniformly, then it converges pointwise. So we just want to show it converges uniformly to its pointwise limit.

- The existence of a pointwise limit follows from the General Principle of Convergence [IA Analysis]. For each $x \in S$ let $f(x)$ be the limit of $f_n(x)$.
- Now we just bound $|f_m(x) - f_n(x)| < \varepsilon$ by Cauchy convergence and let $m \to \infty$.

Theorem 1.3: Weierstrass M-test

Let (f_n) be a sequence of scalar functions on a set S. Let (M_n) be a sequence of non-negative real numbers such that $\sum_{n=1}^{\infty} M_n$ converges. Further suppose $|f_n(x)| \leq M_n$ for all $n \in \mathbb{N}$ and all $x \in S$. Then $\sum_{n=1}^{\infty} f_n$ converges uniformly on S.

Proof 1.3

Main idea: Since $\sum_{k=1}^{n} M_k$ is a Cauchy sequence, we can show $\sum_{k=1}^{n} f_k(x)$ is uniformly Cauchy and hence converges uniformly by the GPUC.

Theorem 1.4: Uniform convergence of power series

Let $\sum_{n=0}^{\infty} c_n(z-a)^n$ be a power series with radius of convergence R. Then for any $r \in (0, R)$ the power series $\sum_{n=0}^{\infty} c_n(z-a)^n$ converges uniformly on $D(a,r) \subseteq \mathbb{C}$.

Proof 1.4

Main idea: take some $w \in D(a,R) \setminus D(a,r)$ and bound $|c_n(w-a)^n| \leq M$ for every n. Then apply the Weierstrass M-test after comparing with a geometric series.

- Fix $w \in D(a,R) \setminus D(a,r)$. Then $\sum_{n=0}^{\infty} c_n(w-a)^n$ converges so $c_n(w-a)^n \to 0$ and we can find $M > 0$ such that $|c_n(w-a)^n| \leq M$ for all $n \in \mathbb{N}$.
- Set $\rho = \frac{r}{|w-a|} < 1$. Then for all $z \in D(a,r)$

$$|c_n(z-a)^n| = |c_n(w-a)^n| \cdot \left| \frac{(z-a)^n}{(w-a)^n} \right| \leq M \frac{r^n}{|w-a|^n} = M\rho^n$$

- Note $\sum_{n=0}^{\infty} M\rho^n$ converges so the result follows by the Weierstrass M-test.

Theorem 1.5: Continuous functions on a closed & bounded interval are uniformly continuous

Let f be a scalar function on a closed & bounded interval $[a,b]$. If f is continuous on $[a,b]$ then f is uniformly continuous on $[a,b]$.

Proof 1.5

Main idea: Argue by contradiction: if it were not, then for some ε and for any δ we could always find x, y a distance δ apart with $|f(x) - f(y)| \geq \varepsilon$. By taking $\delta = 1/n$ for each $n \in \mathbb{N}$ this gives x_n and y_n which contradict Bolzano-Weierstrass.

2 Completeness and the Contraction Mapping Theorem

Theorem 2.1: Contraction Mapping Theorem

Let M be a non-empty, complete metric space and let $f : M \to M$ be a contraction mapping. Then f has a unique fixed point $z \in M$.

Proof 2.1

Main idea: once we have shown uniqueness, then we construct a fixed point by considering an "infinite iteration" $f(f(\dots f(x_0)\dots))$ starting from some $x_0 \in M$. Fix $\lambda < 1$ such that $= d(f(x) - f(y)) \leq \lambda d(x, y)$ for all $x, y \in M$.

- Uniqueness: suppose z, w were both fixed points of f. Then $d(f(z), f(w)) < \lambda d(z, w) = \lambda d(f(z), f(w))$ which is a contradiction.
- Existence: fix $x_0 \in M$ and define $x_n = f(x_{n-1})$ for $n \in \mathbb{N}$. Then for any $n \in \mathbb{N}$

  $$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \lambda d(x_{n-1}, x_n) \leq \lambda^n d(x_0, x_1)$$

  Then $d(x_m, x_n) \leq \sum_{k=m}^{n-1} d(x_k, x_{k+1})$ so using the above bound we can bound this by a geometric series to find $d(x_m, x_n) \leq \frac{\lambda^m}{1-\lambda} d(x_0, x_1)$.
- Note that $\frac{\lambda^m}{1-\lambda} d(x_0, x_1) \to 0$ as $m \to \infty$ so (x_n) is Cauchy. Since M is complete $x_n \to z$ for some $z \in M$. By continuity of f this means $x_{n+1} = f(x_n) \to f(z)$.

Theorem 2.2: Lindelöf-Picard

Given $n \in \mathbb{N}$, $a, b, R, \in \mathbb{R}$ with $a < b$ and $R > 0$, $y_0 \in \mathbb{R}^n$ and a continuous function

$$\phi : [a, b] \times B_R(y_0) \to \mathbb{R}^n$$

Assume that ϕ is K-Lipschitz in its second variable for some $K > 0$. Then there exists $\varepsilon > 0$ such that for any $t_0 \in [a, b]$ the initial value problem

$$f'(t) = \phi(t, f(t)), f(t_0) = y_0$$

has a unique solution on $[c, d] = [t_0 - \varepsilon, t_0 + \varepsilon] \cap [a, b]$.

Proof 2.2

Main idea: We want to say that $f(t) = y_0 + \int_{t_0}^{t} \phi(s, f(s))\mathrm{d}s$ is the solution.

- ϕ is bounded since it takes values on a closed and bounded subset of $\mathbb{R}^{n+1}$. Hence let C be the supremum of ϕ on its domain. We will show that $\varepsilon = \min(\frac{R}{C}, \frac{1}{2K})$ works[a].
- Let $[c, d] = [t_0 - \varepsilon, t_0 + \varepsilon]$. Since $B_R(y_0)$ is complete, it follows that $M = C([c, d], B_R(y_0))$ is complete in the uniform metric D. Now define $T : M \to M$ by
$$T(g)(t) = y_0 + \int_{t_0}^{t} \phi(s, g(s))\mathrm{d}t$$
- Note that T is well defined since $T(g)$ is differentiable and takes values in $B_R(y_0)$ as
$$||T(g)(t) - y_0|| \leq |t - t_0| \sup_{s\in[c,d]} ||\phi(s, g(s))|| = \varepsilon C \leq R$$
- Now f is a solution of the IVP if and only if $f \in M$ and $T(f) = f$. Now it is just required to show T is a contradction mapping which we can do since
$$||T(g)(t) - T(h)(t)|| \leq |t - t_0|KD(g, h) \leq \varepsilon KD(g, h)$$
so taking a supremum of this over $t \in [c, d]$ we have $D(Tg, Th) \leq \varepsilon KD(g, h) \leq \frac{1}{2}D(g, h)$.

[a] I suppose this is worth remembering.

3 Compactness

Theorem 3.1: The Topological Inverse Function Theorem

Let $f : X \to Y$ be a continuous bijection between topological spaces. If X is compact and Y is Hausdorff, then f is an open map, and hence a homeomorphism.

Proof 3.1

Main idea: We want to use the fact that closed subsets of compact sets are compact, and that compact subsets of Hausdorff sets are closed.

- Fix an open subset $U \subseteq X$ and let $K = X \setminus U$. Then K is closed, and since X is compact, K is compact.
- Now f is continuous so $f(K)$ is a compact subset of Y and so $f(K)$ is closed in Y.
- Finally note that since f is a bijection we have $f(K) = f(X \setminus U) = f(Y) \setminus f(X)$ and so $f(X)$ is open.

Theorem 3.2: Tychonov's Theorem

The product of compact topological spaces is compact in the product topology.

Proof 3.2

Main idea: as per usual, take an open cover and show there is a finite subcover.

- Let $\mathcal{W}$ be an open cover of $X \times Y$. Define
$$\mathcal{U} = \{U \times V : U \text{ open in } X, V \text{ open in } Y \text{ and } \exists W \in \mathcal{W} : U \times V \subseteq W\}$$
- It is easy to see that $\mathcal{U}$ is an open cover for $X \times Y$. We now show that it is sufficient to show $\mathcal{U}$ has a finite subcover. Indeed if $(U_i)_{i=1}^n$ and $(V_i)_{i=1}^n$ are elements of $\mathcal{U}$ such that $X \times Y = \bigcup_{i=1}^n U_i \times V_i$, we can take $W_i \in \mathcal{W}$ with $U_i \times V_i \subseteq W_i$ for each i. Then $X \times Y = \bigcup_{i=1}^n W_i$ so $(W_i)_{i=1}^n$ is a finite subcover of $\mathcal{W}$.
- Now fix $x \in X$. Then $\{x\} \times Y$ is a continuous image of Y (take $y \mapsto (x, y)$) and so it is compact. Since $\mathcal{U}$ covers $\{x\} \times Y$, there is a finite subfamily of $\mathcal{U}$ which covers it. So we can take $n_x \in \mathbb{N}$ and open subsets $U_{x,1}, U_{x,2}, \ldots, U_{x,n_x}$ of X and $V_{x,1}, V_{x,2}, \ldots, V_{x,n_x}$ of Y such that $U_{x,i} \times V_{x,i} \in \mathcal{U}$ for each i and $\{x\} \times Y \subseteq \bigcup_{i=1}^{n_x} U_{x,i} \times V_{x,i}$.
- Wlog assume $x \in U_{x,i}$ for each i. Then $U_x = \bigcap_{i=1}^{n_x}$ is an open set in X containing x. Furthermore $\bigcap_{i=1}^{n_x} U_{x,i} \times V_{x,i} \supseteq U_x \times Y$.
- Now carry out the above process for each $x \in X$ and obtain the open cover $\{U_x : x \in X\}$ for X. Then X is compact so we can take a finite subset $F \subseteq X$ such that $X = \bigcup_{x \in F} U_x$. Hence
$$X \times Y = \left(\bigcup_{x \in F} U_x \right) \times Y \subseteq \bigcup_{x \in F} \bigcup_{i=1}^{n_x} U_{x,i} \times V_{x,i}$$
and so we have our finite subcover of $\mathcal{U}$.

Theorem 3.3: Compactness is equivalent to sequential compactness in a metric space

In a metric space (M, d) the following are equivalent

(i) M is compact.

(ii) M is sequentially compact.

(iii) M is complete and totally bounded.

Proof 3.3

Main idea: we will show $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$.

- First we show $(a) \Rightarrow (b)$. Let (x_n) be a sequence in M. For each n define $T_n = \{x_k : k > n\}$. Note that if (x_n) has a convergent subsequence then $x \in \bigcap_{n\in\mathbb{N}} \mathrm{Cl}(T_n)$.

- Note that $\bigcap_{n\in\mathbb{N}} \mathrm{Cl}(T_n)$ is non-empty (since if it were, $\bigcup_{n\in\mathbb{N}} M \setminus \mathrm{Cl}(T_n) = M$ and by compactness $M = M \setminus \mathrm{Cl}(T_n)$ for some n). Now fix $x \in \bigcap_{n\in\mathbb{N}} \mathrm{Cl}(T_n)$. We will show that a subsequence of (x_n) converges to x. Since $x \in \mathrm{Cl}(T_1)$ we can take $k_1 > 1$ with $x_{k_1} \in D_1(x)$. Now $x \in \mathrm{Cl}(T_{k_1})$ so we can take $k_2 > k_1$ with $x_{k_2} \in D_{1/2}(x)$. Continue this to inductively obtain a subsequence (x_{k_n}) of (x_n) such that $d(x_n, x) < \frac{1}{n}$ and so $x_{k_n} \to x$.

- Now we show $(b) \Rightarrow (c)$. First show that M is complete. Let (x_n) be Cauchy in M. By sequential compactness, there is a subsequence (x_{k_n}) which converges to some $x \in M$ and so $x_n \to x$.

- Suppose M were not totally bounded. Then there is some $\varepsilon > 0$ such that there is no finite ε-net for M. Then construct a sequence (x_n) as follows. Pick any $x_1 \in M$ and let $x_n \in M \setminus \bigcup_{k=1}^{n-1} B_\varepsilon(x_k)$ (we can't 'run out' of elements since there is no ε-net). Now $d(x_m, x_n) > \varepsilon$ for all $m \neq n$ and so it has no Cauchy subsequence which contradicts sequential compactness.

- Finally we show $(c) \Rightarrow (a)$. Let $\mathcal{U}$ be an open cover for M and suppose M is not finitely covered by $\mathcal{U}$. Now construct a nested sequence $A_0 \supseteq A_1 \supseteq \dots$ of non-empty closed subsets of M such that $\operatorname{diam} A_n \to 0$ and A_n cannot be finitely covered by $\mathcal{U}$ for every n. Let $A_0 = M$ and define A_n in terms of A_{n-1} as follows: take $K \in \mathbb{N}$ and non-empty closed subsets $B_1, \dots, B_K$ of A_{n-1} such that $A_{n-1} = \bigcup_{k=1}^{K} B_k$ and $\operatorname{diam} B_k \leq \frac{1}{n}$ for $1 \leq k \leq K$. Now there must be some k such that B_k cannot be finitely covered by $\mathcal{U}$ (or else A_{n-1} would be finitely covered by $\mathcal{U}$). Then let $A_n = B_k$.

- Now for each n choose $x_n \in A_n$ arbitrarily. It is simple to show that (x_n) is Cauchy and so converges by completness of M. Since the A_n are nested and closed we have $x \in A_n$ for every n. $\mathcal{U}$ is an open cover for M so there is some $U \in \mathcal{U}$ with $x \in U$ and hence some $r > 0$ with $D_r(x) \subseteq U$. Now pick $n \in \mathbb{N}$ with $\operatorname{diam} A_n < r$ and so $A_n \subseteq D_r(x) \subseteq U$. This contradicts the fact that A_n cannot be finitely covered by $\mathcal{U}$.

4 Differentiation

Theorem 4.1: The Chain Rule

Let U be an open subset of $\mathbb{R}^m$ and V be an open subset of $\mathbb{R}^n$. We are given functions $f : U \to \mathbb{R}^n$ with $f(U) \subseteq V$ and $g : V \to \mathbb{R}^p$. Let $a \in U$ and $b = f(a)$. If f is differentiable at a, and g is differentiable at $b = f(a)$, then $g \circ f$ is differentiable at a, and $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$.

Proof 4.1

Main idea: since we think of differentiation as a linear approximation, we can simply apply this approximation twice, and show it is suitably 'good'.

- Let $S = f'(a)$ and $T = g'(b)$. Then for suitable error functions ε and ζ

$$f(a+h) = f(a) + S(h) + ||h||\varepsilon(h)$$
$$g(b+k) = g(b) + T(k) + ||k||\zeta(k)$$

- Now

$$g(f(a+h)) = g(f(a) + S(h) + ||h||\varepsilon(h))$$

so we 'expand out' and use linearity of T to show this is equal to $g(f(a)) + g'(f(a)) \circ f'(a)(h) + ||h||T(\varepsilon(h)) + ||k||\zeta(k)$ where $k = S(h) + ||h||\varepsilon(h)$.

- Finally we show that $||h||T(\varepsilon(h)) + ||k||\zeta(k)$ is $o(h)$, which is simple by bounding suitably.

Theorem 4.2: Partial Derivatives imply Differentiable

We are given an open subset $U \subseteq \mathbb{R}^m$, a function $f : U \to \mathbb{R}^n$ and $a \in U$. If there exists $r > 0$ such that $D_r(a) \subseteq U$ and for each $i \in \{1, \ldots, m\}$, the partial derivative $D_i f(x)$ exists for all $x \in D_r(a)$ and the map $x \mapsto D_i f(x) : D_r(a) \to \mathbb{R}^n$ is continuous at a, then f is differentiable at a.

Proof 4.2

Main idea: if the derivative exists, we know what it must be since we have all the partial derivatives. So we just need to show this works. To do this we pass to a function which is real valued with real domain so we can apply the mean value theorem.

- Note that we can wlog take $n = 1$ by considering the components of f in $\mathbb{R}^n$. Take $a = (a_1, \ldots, a_m)$. If f is differentiable at a, then $f'(a)(h_1, \ldots, h_m) = \sum_{i=1}^m h_i D_i f(a)$ so we want to show

$$f(a+h) = f(a) + \sum_{i=1}^{m} h_i D_i f(a) + o(||(h_1, \ldots, h_m)||)$$

- Now note

$$\begin{aligned}
&f(a+h) - f(a) \\
&= f(a+h) - f(a_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) \\
&+ f(a_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) - f(a_1, a_2, a_3 + h_3, \ldots, a_m + h_m) \\
&\ \vdots \\
&+ f(a_1, a_2, a_3, \ldots, a_{m-1}, a_m + h_m) - f(a_1, a_2, \ldots, a_m)
\end{aligned}$$

- Note that by the MVT we have

$$\begin{aligned}
&f(a_1 + h_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) - f(a_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) \\
&= D_1 f(a_1 + t_1 h_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) h_1
\end{aligned}$$

  for some $t_1 \in (0, 1)$.

- The other terms are dealt with analagously to get

$$\begin{aligned}
&f(a+h) - f(a) \\
&= D_1 f(a_1 + t_1 h_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) h_1 \\
&+ D_2 f(a_1, a_2 + t_2 h_2, a_3 + h_3, \ldots, a_m + h_m) h_2 \\
&\ \vdots \\
&+ D_m f(a_1, \ldots, a_m + t_m h_m) h_m
\end{aligned}$$

- Now we have the bound

$$\begin{aligned}
&||f(a+h) - f(a) - \sum_{i=1}^{m} h_i D_i f(a)|| \\
&\leq \{|D_1 f(a_1 + t_1 h_1, a_2 + h_2, a_3 + h_3, \ldots, a_m + h_m) - D_1 f(a_1, a_2, \ldots, a_m)| \\
&\quad + \ldots + |D_m f(a_1, \ldots, a_m + t_m h_m) - D_m f(a_1, \ldots, a_m)|\} \, ||h||
\end{aligned}$$

  and by continuity of the D_i at a, this is $o(h)$ as $h \to 0$.

Theorem 4.3: The Mean Value Inequality

Let U be an open subset of $\mathbb{R}^m$, let $f : U \to \mathbb{R}^n$ be a function that is differentiable at every $z \in U$ and let $a, b \in U$. Assume that the line segment $[a, b]$ joining a and b is contained in U and that for some M we have $||f'(z)|| \leq M$ for all $z \in [a, b]$. Then

$$||f(b) - f(a)|| \leq M||b - a||$$

Proof 4.3

Main idea: we consider a suitable real valued function with real domain so we can apply the usual Mean Value Theorem.

- Consider our line $\gamma(t) = a + tu$ where $u = b - a$. Then by our assumption, $\gamma([0, 1]) \subseteq U$ and so the composition $f \circ \gamma$ is defined on $[0, 1]$.
- Now consider the function $\varphi : \gamma^{-1}(U) \to \mathbb{R}^m$ by $\varphi(t) = \langle f(a + tu), v\rangle$ where $v = f(b) - f(a)$.
- Noting that $\varphi(1) - \varphi(0) = ||f(b) - f(a)||^2$ we can apply the Mean Value Theorem to φ (since it is a composition of a differentiable and linear function and hence differentiable).
- Applying this gives

  $$\varphi(1) - \varphi(0) = \langle f'(a+\theta u)(u), v\rangle \leq ||f'(a+\theta u)(u)||||v|| \leq M||b-a||||f(b)-f(a)||$$

Theorem 4.4: The Inverse Function Theorem

Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ be a C^1-function on U. Let $a \in U$ and assume that $f'(a)$ is invertible. Then there are subsets U, V of $\mathbb{R}^n$ such that $a \in V \subseteq U$ and $f|_V : V \to W$ is a bijection with C^1 inverse g and $g'(y) = f'(g(y))^{-1}$ for every $y \in W$.

Proof 4.4

Main idea: we reduce to the case where $a = f(a) = 0$ and $f'(a) = I$. Then we apply the Mean Value Inequality and Contraction Mapping Theorem.

- Let $T = f'(a)$ and define $h(x) = T^{-1}(f(x+a) - f(a))$. It is easy to show $h(0) = 0$, $h'(0) = I$ and h has domain $U - a$. Furthermore it is sufficient to prove the theorem for h since $f(x) = T(h(x-a)) + f(a)$.

- Now we may assume $f(0) = 0$ and $f'(0) = I$. Since f' is continuous, we can choose $r > 0$ such that $B_r(0) \subseteq U$ and $||f'(x) - I|| \leq 1/2$ for all $x \in B_r(0)$. Now define $p : U \to \mathbb{R}^n$ by $p(x) = f(x) - x$. Then p is differentiable with $p'(x) = f'(x) - I$ for all $x \in U$. Hence $p'(x) \leq 1/2$ for all $x \in B_r(0)$. Now for any $x, y \in B_r(0)$ the line segment $(1-t)x + ty$ is at most distance r from the origin. So by the Mean Value Inequality (applied to p with $M = 1/2$)

$$||f(x)-f(y)|| = ||p(x)+x-(p(y)+y)|| \geq ||x-y|| - ||p(x)-p(y)|| \geq \frac{1}{2}||x-y||$$

- Let $s = r/2$. We will show that $f(D_r(0)) \supseteq D_s(0)$. Fix some $w \in D_s(0)$ and for $x \in B_r(0)$ let $q(x) = w - p(x)$. Then since $p(0) = 0$ we have

$$||q(x)|| \leq ||w|| + ||p(x) - p(0)|| \leq ||w|| + \frac{1}{2}||x - 0|| < 2s = r$$

  So q is a contraction mapping onto $B_r(0)$. By the contraction mapping theorem q has a unique fixed point x, i.e there is a unique x with $f(x) = w$.

- Set $W = D_s(0)$ and $V = f^{-1}(D_s(0)) \cap D_r(0)$. We will show these work as the V and W as stated in the theorem. By the above, $f|_V : V \to W$ is a bijection. Let g be its inverse. Then given $u, v \in W$ let $x = g(u)$ and $y = g(v)$. Then we had

$$||f(x) - f(y)|| \geq \frac{1}{2}||x - y|| \implies ||x - y|| \geq \frac{1}{2}||g(u) - g(y)||$$

  and so g is a 2-Lipschitz map and hence continuous.

- The final step is **non-examinable** so for the purposes of this document will not be included.

Theorem 4.5: Symmetry of mixed partial derivatives

We are given an open set $U \subseteq \mathbb{R}^m$, a function $f : U \to \mathbb{R}^n$ and a point $a \in U$. Assume f is twice differentiable on an open set V with $a \in V \subseteq U$. Assume $f'' : V \to \text{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ is continuous at a. Then

$$f''(a)(u, v) = f''(a)(v, u)$$

Proof 4.5

Main idea: we reduce to $n = 1$ and then consider a suitible function to apply the Mean Value Theorem to.

- First show that we can take $n = 1$ wlog by noting $[D_u f(x)]_j = f_j'(x)(u) = D_u f_j(x)$.
- Now define φ on a neighbourhood of $(0, 0) \in \mathbb{R}^2$ by

$$\varphi(s,t) = f(a + su + tv) - f(a + tv) - f(a + su) + f(a)$$

- Now for $s, t \in \mathbb{R}$ fixed consider $\psi(y) = f(a + yu + tv) - f(a + yu)$ (so $\varphi(s,t) = \psi(s) - \psi(0)$). Then by the Mean Value Theorem

$$\varphi(s,t) = s\psi'(\alpha s) = s(D_u f(a + \alpha sy + tv)) - D_u f(a + \alpha su))$$

for some $\alpha(s,t) \in (0,1)$.

- Now apply the Mean Value Theorem to $y \mapsto D_u f(a + \alpha su + yv)$ to get

$$\varphi(s,t) = stD_v D_u f(a + \alpha su + \beta tv) = stf''(a + \alpha su + \beta tv)(v,u)$$

for some $\beta(s,t) \in (0,1)$.

- Doing this for every (s,t) and using continuity of f'' at a we have

$$\frac{\varphi(s,t)}{st} = f''(a + \alpha su + \beta tv)(v,u) \to f''(a)(v,u) \text{ as } (s,t) \to (0,0)$$

- Now repeat the above instead with $\psi(y) = f(a + su + yv) - f(a + yv)$ to get

$$\frac{\varphi(s,t)}{st} \to f''(a)(u,v) \text{ as } (s,t) \to (0,0)$$