

MAT257: Real Analysis II

QiLin Xue

Fall 2021

Contents

1	Differentiation	2
1.1	Inverse Function Theorem	2
2	Integration	6
2.1	Integrability, Measure 0, Content 0, Integrability and Continuity	6
3	Fubini's Theorem	7

1 Differentiation

1.1 Inverse Function Theorem

Theorem: Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable in an open set containing a and $\det f'(a) \neq 0$. Then there is an open set V containing a and an open set W containing $f(a)$ such that $f : V \rightarrow W$ has a continuous inverse $f^{-1} : W \rightarrow V$ which is differentiable and for all $y \in W$ satisfies

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1} \quad (1)$$

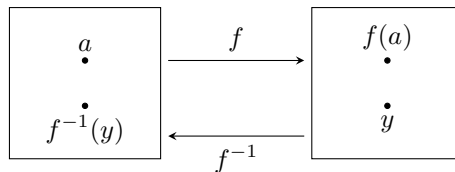
Motivation: In 1D calculus, if a function has a derivative $f'(x) > 0$ at some point $x = a$, then around this point the function is monotone and has an inverse, per the intermediate value theorem. We want something similar for multiple dimensions, however there is no equivalent intermediate value theorem.

We will motivate our proof with the following steps

1. Prove the last step, as it is the easiest.
2. WLOG, make the simplifying assumption that $f'(a) = I$.
3. Define “all-scale fidelity” to describe how vectors in an open neighbourhood around a are *nearly* preserved. Then show that f has all-scale fidelity on some neighbourhood U around a .
4. Given $y \in W$. We can construct a sequence (x_i) such that $\{f(x_i)\}$ is a Cauchy sequence which converges to y .
5. Show that there exists an $x \in V$ such that $f(x) = y$, by invoking continuity. Said differently, $f|_V : V \rightarrow W$ is onto.
6. Show that $f|_V : V \rightarrow W$ is injective (1-1).
7. We have shown that f^{-1} exists. We now need to show that f^{-1} is continuous.
8. Show that f^{-1} is differentiable at the point $f(a) = b$.
9. Show that f^{-1} is differentiable at a point near b .
10. Show that $f^{-1}(y)$ is continuously differentiable near b .

Let us perform these steps:

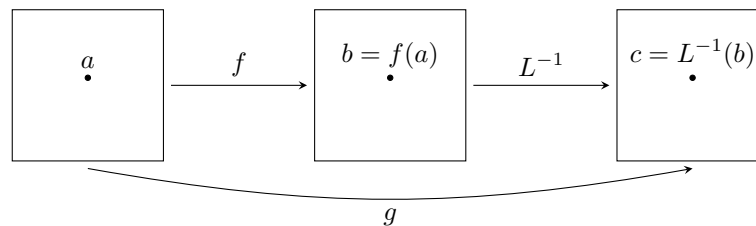
1. Consider the following setup



Recall that $f \circ f^{-1} = I$, so differentiating and applying the chain rule we can write

$$f'(f^{-1}(y)) \cdot (f^{-1})'(y) = I \implies (f^{-1})'(y) = [f'(f^{-1}(y))]^{-1} \quad (2)$$

2. We are allowed to write $f'(a) = I$ since every invertible matrix is just the identity with a change of basis. More concretely, consider another composition represented below, where $L = f'(a)$:



By the chain rule, we have $g'(a) = L^{-1} \circ f'(a)$ since L^{-1} is a linear transformation. Note that $f'(a) = L$, so $g'(a) = I$ is the identity.

If the IFT was true for functions whose differential is I , then it's true for g , so there exists g^{-1} . Also, $f^{-1} = g^{-1} \circ L^{-1}$. Therefore, if g^{-1} is continuously differentiable, then f^{-1} would also be continuously differentiable. Thus, it is sufficient to only look at the case where the differential is the identity.

3. Consider a small neighbourhood around a . Intuitively, we should expect vectors and the image of the vectors (which need not start at a) to look roughly the same.

Specifically, f has **all-scale-fidelity** on some neighbourhood U of a with *fidelity factor* $\frac{1}{257}$. This means for all $x_1, x_2 \in U$, we have

$$|(x_2 - x_1) - (f(x_2) - f(x_1))| \leq \frac{1}{257} |x_2 - x_1| \quad (3)$$

Proof. From a previous theorem, if we have a function $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ whose derivative $|g'| < M$ is bounded in some open set, then we can write

$$|g(x_1) - g(x_2)| < n^2 M |x_1 - x_2|. \quad (4)$$

Now consider $g(x) = f(x) - x$. The derivative is $g' = f' - I$ so $g'(a) = 0$. Therefore, there exists an open rectangle U of a where we can say $|g'(x)| \leq \frac{1}{257n^2}$.

By the previous theorem, for any $x_1, x_2 \in U$, we have

$$|g(x_1) - g(x_2)| \leq \frac{1}{257} |x_1 - x_2| \quad (5)$$

However, the LHS is just

$$\begin{aligned} |g(x_1) - g(x_2)| &= |f(x_1) - x_1 - (f(x_2) - x_2)| \\ &= |(x_2 - x_1) - (f(x_2) - f(x_1))| \end{aligned}$$

□

4. Given $y \in W$, we wish to find an $x \in V$ such that $f(x) = y$. We will travel this direction in W , but we may “miss” by a bit. We can then repeat this process, with each step we travel in the direction of y .

Put this formally, let $W = B_{r/2}(b)$. Given $y \in W$, we claim that there exists an $x \in B_r(a)$ such that $f(x) = y$. Indeed,

$$x_1 = a + (y - b) \quad (6)$$

$$x_2 = x_1 + (y - f(x_1)) \quad (7)$$

$$x_3 = x_2 + (y - f(x_2)) \quad (8)$$

$$x_n = x_{n-1} + (y - f(x_{n-1})) \quad (9)$$

But the difference between any two consecutive terms is just the LHS of the all-scale fidelity

$$|x_n - x_{n-1}| = |(x_{n-1} - x_{n-2}) - (f(x_{n-1}) - f(x_{n-2}))| \quad (10)$$

$$\leq \frac{1}{257} |x_{n-1} - x_{n-2}| \quad (11)$$

$$\leq \frac{1}{257^{n-1}} |x_1 - x_0| \quad (12)$$

$$\leq \frac{1}{257^{n-1}} |y - b| \quad (13)$$

$$\leq \frac{1}{257^{n-1}} \frac{r}{2} \quad (14)$$

We now need to show that each x_i is within the ball of radius r around a (since this is only when all-scale fidelity is defined). It can be shown via induction that $|x_n - a| \leq r$.

Finally, we show that (x_n) is a Cauchy-Sequence. We can immediately show this by noting that

$$|x_n - x_m| \leq \frac{1}{257^m} r \quad (15)$$

so (x_n) is cauchy.

5. While we have shown that $f(x_n)$ converges to y , we have not yet shown that this is possible, i.e. what if there is a discontinuity? We can invoke the continuity of f to show that there does exist such an x_n . We have

$$|f(x_n) - y| = |x_{n+1} - x_n| \leq \frac{1}{257^n} \frac{r}{2} \rightarrow 0 \quad (16)$$

so from continuity, there exists an x such that

$$|f(x) - y| = \lim_{n \rightarrow \infty} |f(x_n) - y| = 0. \quad (17)$$

Therefore, $f(x) = y$. We can now define $V = F^{-1}(W)$ and now

$$F|_V : V \rightarrow W \quad (18)$$

is onto and 1-1.

6. We have constructed x in one such way. How do we know that if we use a different procedure, we find a different x ?

Assume that $f(x_1) = f(x_2)$ where $x_{1,2} \in B_r(a)$. Then by ASF,

$$|(x_1 - x_2) - (f(x_1) - f(x_2))| \leq \frac{1}{257} |x_1 - x_2| \quad (19)$$

$$|x_1 - x_2| \leq \frac{1}{257} |x_1 - x_2| \quad (20)$$

which is true if and only if $x_1 = x_2$.

7. It might seem that continuity for f^{-1} is cheap. After all, the difference between two vectors in both the input and image space is roughly the same. However, the mistake is written in terms of $|x_1 - x_2|$, so this reasoning becomes circular. We need to reformulate the ASF principle such that the mistake is in terms of $|y_1 - y_2|$.

To simplify things, let $\alpha = x_1 - x_2$ and $\beta = f(x_1) - f(x_2)$. Then ASF says that

$$|\alpha - \beta| \leq \frac{1}{257} |\alpha| \quad (21)$$

But $\alpha = \beta + \alpha - \beta$. By the triangle inequality, this becomes

$$|\alpha - \beta| \leq \frac{1}{257} (|\beta| + |\alpha - \beta|) \quad (22)$$

$$\frac{256}{257} |\alpha - \beta| \leq \frac{1}{257} |\beta| \quad (23)$$

$$|\alpha - \beta| \leq \frac{1}{256} |\beta| \quad (24)$$

8. Let us first show that f^{-1} is differentiable at the point b . We can write

$$f^{-1}(b + h) = f^{-1}(b) + I \cdot h + e(h) \quad (25)$$

We want to show that $e(h)$ is tiny. We want to rearrange this in a form such that we can apply all scale fidelity. Let $b + h = y_2$ and let $x_2 = f^{-1}(y_2)$. Let $b = y_1$ and $a = x_1$, so the above just becomes

$$x_2 = x_1 + y_2 - y_1 + e(h) \quad (26)$$

However the error once rearranged becomes

$$|e(h)| = |(x_2 - x_1) - (y_2 - y_1)| \leq \frac{1}{256} |y_2 - y_1| \quad (27)$$

Since $y_2 - y_1 = h$, we end up with

$$\frac{|e(h)|}{|h|} \leq \frac{1}{256} \quad (28)$$

Now we have a problem: we want to show that this approaches zero. This condition is not good enough. However, this constant was chosen arbitrarily, so we can make $\frac{|e(h)|}{|h|}$ as small as possible.

9. Similarly, the choice of b was also arbitrary. If the conditions for the IFT hold at a , then they have to hold in an open set near a (due to continuity). Therefore, we can rewrite the entire proof by considering points near b and not just b .
10. To show f^{-1} is continuously differentiable, we can use the chain rule:

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1} \quad (29)$$

We can conclude that $f^{-1}(y)$ is continuous in y and $f'(x)$ is continuous in x . Therefore, $M \mapsto M^{-1}$ is a continuous operation on matrices. Specifically, it is a function that maps $\mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$. This is not everywhere defined. But where defined, the inverse is continuous by Cramer's Law. There is an explicit formula for the inverse, so this map is continuous.

2 Integration

2.1 Integrability, Measure 0, Content 0, Integrability and Continuity

Definition: A is of measure 0 means that for $\forall \epsilon > 0$, there exists open (alternatively closed) rectangles $(R_i)_{i=1}^{\infty}$ such that

1. $A \subset \bigcup R_i$
2. $\sum \text{vol}(R_i) < \epsilon$.

For example, finite & countable sets, along with $\mathbb{R} \subset \mathbb{R}^2$ are of measure 0.

Definition: A set X is called **countably infinite** if there is a surjective function F such that $F : \mathbb{N} \rightarrow X$

A few facts that follows:

1. Finite sets are countable. (Typically, we exclude this from the definition.)
2. Subsets of countable sets are countable. (Proof: Elements in a countable set can be enumerated. Simply select a new enumeration.)
3. A finite/countable union of countable sets is countable. If A_i is countable for all i , then $\bigcup A_i$ is countable.

Proof. Consider

$$\begin{array}{llll} A_1 : & a_{11} & a_{12} & a_{13} & \cdots \\ A_2 : & a_{21} & a_{22} & a_{23} & \cdots \\ A_3 : & a_{31} & a_{32} & a_{33} & \cdots \end{array}$$

To enumerate the union, we look at the diagonals, i.e.

$$\bigcup A = \{a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}, \dots\} \quad (30)$$

□

Another fact that immediately follows is that the set of integers is countable, since $\mathbb{Z} = (-\mathbb{N}) \cup \{0\} \cup \mathbb{N}$

4. If A, B are countable, then $A \times B$ is countable.

Proof. We can prove this similar to how a countable union of sets is countable. Alternatively, we can write

$$A \times B = \bigcup_{b \in B} A \times \{b\} \quad (31)$$

And each set $A \times \{b\}$ is countable.

□

A fact that follows is that the set of rationals is countable, since $\mathbb{Q} \subset \mathbb{Z} \times \mathbb{Z}$.

It may seem a bit suspicious since up to this point, everything is countable. However, there are sets that are uncountable!

Theorem: \mathbb{R} is not countable. And hence, irrational numbers are not countable, i.e. there are “more” reals than naturals, more irrationals than rationals.

Proof. Assume that \mathbb{R} is countable, that is (a_i) is an enumeration of the real numbers. Let x be a real number whose k^{th} decimal digit is different from the k^{th} decimal digit of a_k . Note that x cannot be any of the a_k s, hence $\{a_k\} \neq \mathbb{R}$. □

We can now state and prove a few statements about measure-0.

1. If A is measure-0 and $B \subset A$, then B is measure-0.
2. A countable union of measure-0 sets is measure-0.

Proof. Suppose $\forall i$, A_i is of measure 0, so given $\epsilon > 0$, we can cover A_i with countably many rectangles whose $\sum \text{vols} < \frac{\epsilon}{2^i}$. Take all the rectangles above together, as a countable collection of countable sets, this collection of rectangles is countable, and

$$\sum \text{vols} < \sum_{i=1}^{\infty} \frac{\epsilon}{2^i} = \epsilon. \quad (32)$$

Finally this collection covers

$$\bigcup A_i = A \quad (33)$$

so A is measure-0. \square

Note that we have stated before that $\mathbb{R} \subset \mathbb{R}^2$

Warning: Countable sets are measure-0, but the converse is not true! For example, $\mathbb{R} \subset \mathbb{R}^2$ has measure-0, but is not countable.

Another Example: Let \mathcal{C} be the cantor set $\subset [0, 1]$ which is uncountable, yet of measure-0 in \mathbb{R} . Let

$$\mathcal{C} = \{0.C_1C_2C_3C_4 \cdots : C_i \in \{0, 2\}\} \quad (34)$$

\mathcal{C} is uncountable for the same reason as \mathbb{R} . Define C_k to be the union of 2^k intervals of length $\frac{1}{3^k}$. Therefore,

$$\mathcal{C} \subset C_k \quad (35)$$

where C_k itself is a union of intervals of total length $2^k \cdot \frac{1}{3^k} = \left(\frac{2}{3}\right)^k$, which approaches 0.

Therefore, \mathcal{C} is measure-0. As an aside, each C_k is compact, so $\mathcal{C} = \bigcap C_k$ is therefore also compact.

Theorem: $[a, b] \subset \mathbb{R}$ is *not* measure 0. In fact, $\mathbb{R} \subset \mathbb{R}^n$ is not measure 0.

Definition: $A \subset \mathbb{R}^n$ is said to be of content-0 if $\forall \epsilon > 0$ it is contained in a finite union of rectangles whose sum of volumes is smaller than ϵ .

Note that $\mathbb{Z} \in \mathbb{R}$ is not of content-0.

3 Fubini's Theorem

Currently, we have the tools to define the integral, but we don't have the tools to compute the integral yet. We will start off with a loose example.

Example 1: Integrate xy on $[0, 1]_x \times [0, 1]_y$. Fubini's theorem loosely tells us that we can fix x , and then fix y . Namely,

$$\int_{[0,1] \times [0,1]} xy \, dx \, dy = \int_0^1 x \int_0^1 y \, dy \, dx = \frac{1}{4} \quad (36)$$

We will now formally state it.

Theorem: (Tempting but Incorrect) Let $A \subset \mathbb{R}^n_x$ and $B \subset \mathbb{R}^m_y$ be rectangles, set $R = A \times B \subset \mathbb{R}^{n+m}$. Let

$$F : R \rightarrow \mathbb{R} \quad (37)$$

be an integrable function. Let

$$g(x) = \int_B f(x, y) \, dy. \quad (38)$$

Then

$$\int_R F = \int_A g \, dx \quad (39)$$

Note that this is incorrect for general functions f , but is true if f is continuous.

Warning: Note that we cannot define $g(x) = \int_B f(x, y) \, dy$ and $\int_A f = \int_A g \, dx$ since while f is integrable over $A \times B$, it is not necessarily integrable over B . For example, suppose we have a function defined as

$$f(x, y) = \begin{cases} 0 & x < 0.5 \\ 1 & x > 0.5 \\ 1_{\mathbb{Q}} & x = 0.5 \end{cases} \quad (40)$$

where $1_{\mathbb{Q}}$ is the Dirichlet function, defined to be 1 if rational and 0 otherwise. f will be integrable in the region $[0, 1] \times [0, 1]$ but is not integrable if we restrict it to the line $x = 0.5$. This is because the set of discontinuities is of measure 0 in \mathbb{R}^2 but is of measure 1 in \mathbb{R} .

While it may be tempting to write the theorem as

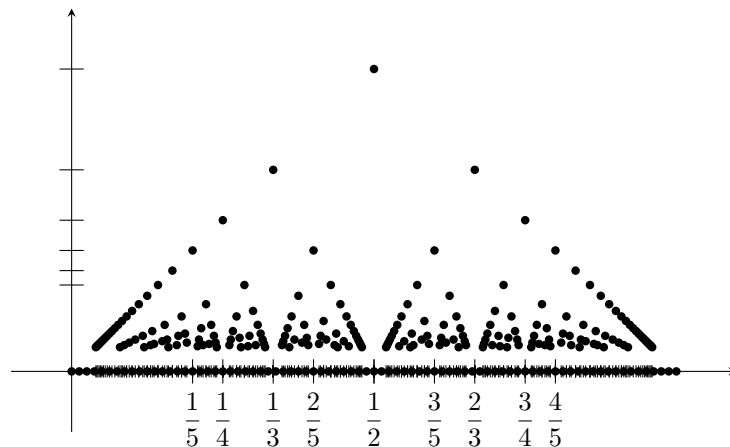
$$g(x) = \begin{cases} \int_B f(x, y) \, dy & f(x, -) \text{ is integrable} \\ 17 & \text{otherwise.} \end{cases} \quad (41)$$

and define

$$\int_R f = \int_A g \quad (42)$$

which might solve the problem of discontinuities. However, this is still wrong. Here is a counter example.

Consider $h(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \\ 0 & \text{otherwise} \end{cases}$ defined on $[0, 1]$, known as Thomae's function, which looks like the below



Note that $h(x)$ is discontinuous on \mathbb{Q} but continuous on \mathbb{Q}^C . Since \mathbb{Q} is of measure 0, h is integrable. The integral is

$$\int_0^1 h(x) = 0 \quad (43)$$

and we can prove this by cropping the function about some arbitrary $y = c$. Since there are only a finite amount of points above this line, we can “chop” them off. Now we extend this to two variables. Consider

$$f(x, y) = \begin{cases} \frac{1}{q} & x, y \in \mathbb{Q}, x = \frac{p}{q} \\ 0 & \text{otherwise} \end{cases} \quad (44)$$

defined on the set $[0, 1] \times [0, 1]$. If we try to compute g using the second incorrect attempt, we get

$$g(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ 17 & x \in \mathbb{Q} \end{cases} \quad (45)$$

but this is a “bed of nails” function with respect to x , and is not integrable, and thus we cannot expect an equality.

Note that the choice for $g(x)$ might sound stupid. After all, we can choose 0 instead of 17, to remove the problematic values. However, we can just shift $f(x, y)$ to create another counterexample.

Let us now state the correct theorem,

Theorem: (Correct Fubini's Theorem) Let $A \subset \mathbb{R}_x^n$ and $B \subset \mathbb{R}_y^m$ be rectangles, set $R = A \times B \subset \mathbb{R}^{n+m}$. Let $f : R \rightarrow \mathbb{R}$ be an integrable function and let

$$\underline{g}(x) = \int_L f(x, y) dy = L(f(x, -)) = \sup \text{ lower sums for } f(x, -) \quad (46)$$

$$\bar{g}(x) = \int_U f(x, y) dy = U(f(x, -)) = \inf \text{ upper sums} \quad (47)$$

Then \underline{g} and \bar{g} are integrable and

$$\int_R f dx dy = \int_A \underline{g} dx = \int_A \bar{g} dx \quad (48)$$

Let us go back to our previous counterexample. Now,

$$\underline{g}(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ 0 & x \in \mathbb{Q} \end{cases} \quad (49)$$

so $\underline{g}(x) = 0$ and $\int \underline{g} = \int 0 = 0$. On the other hand,

$$\bar{g}(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ \frac{1}{q} & x \in \mathbb{Q} \end{cases} \quad (50)$$

which is just $h(x)$ from earlier, which we have computed the integral already to be 0. Now that we have worked through examples, but we have not yet proved the theorem yet.

Before we do so, note that if f is continuous, all that is a non-issue

$$g(x) = \int_B f(x, y) dy = \bar{g}(x) = \underline{g}(x) \quad (51)$$

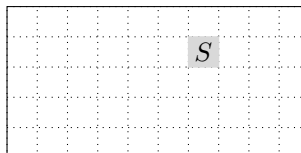
for all x , so the naive Fubini's Theorem holds. Likewise, if $f(x, -)$ is integrable except for finitely many x 's, then the second attempt we made holds.

Proof. We sketch out the proof as follows:

1. Bound $L(f, P) \leq L(\underline{g}, P_A)$ and $U(f, P) \leq U(\bar{g}, P_A)$.
2. Show that \bar{g} and \underline{g} is integrable.

We will now carry out these steps:

1. Consider a partition P of $R \subset \mathbb{R}^{n+m}$, which is illustrated below. We can restrict our attention to the first n and the last m coordinates. We can always write it as $P_A \times P_B$, where P_A is a partition of A and P_B is a partition of B .



If $S \in P$, we can write $S = S' \times S''$ where $S' \in P_A$ and $S'' \in P_B$.

Given this partition $P = P_A \times P_B$ of R , we have

$$L(f, P) = \sum_{S \in P} V(S) \cdot \inf_{(x,y) \in S} f(x, y) \quad (52)$$

$$= \sum_{\substack{S' \in P_A \\ S'' \in P_B}} V(S') V(S'') \cdot \inf_{x \in S'} \inf_{y \in S''} f(x, y) \quad (53)$$

$$= \sum_{S' \in P_A} V(S') \sum_{S'' \in P_B} V(S'') \inf_{x \in S'} \inf_{y \in S''} f(x, y) \quad (54)$$

We are aiming to pull the infimums outside. We can do this with the following lemma.

Lemma 1: Let $h_k : X \rightarrow \mathbb{R}_1$. Then

$$\sum_k \inf h_k(x) \leq \inf \sum_k h_k(x) \quad (55)$$

Proof. Note that $\inf_x h_k(x) \leq h_k(y)$ for all y given any k . This means that we can sum both sides

$$\sum_k \inf h_k(x) \leq \sum_k h_k(y) \quad (56)$$

The left hand side is a constant but the right hand side is a function of y . Since this inequality is true for all y , we can just pick the y to minimize the right hand side:

$$\sum_k \inf_x h_k(x) \leq \inf_y \sum_k h_k(y) \quad (57)$$

which is just the lemma with different variable names. □

Using this lemma, we can bound $L(f, P)$ by

$$L(f, P) \leq \sum_{S' \in P_A} V(S') \inf_{x \in S'} \underbrace{\sum_{S'' \in P_B} V(S'') \inf_{y \in S''} f(x, y)}_{L(f(x, -), P_B)} \quad (58)$$

$$\leq \sum_{S' \in P_A} V(S') \inf_{x \in S'} \underline{g}(x) \quad (59)$$

$$= L(\underline{g}, P_A) \quad (60)$$

Similarly, we can do the same thing with supremums to get

$$L(f, P) \leq L(\underline{g}, P_A) \quad U(\bar{g}, P_A) \leq U(f, P) \quad (61)$$

2. Note that both $L(\bar{g}, P_A)$ and $U(\underline{g}, P_A)$ are bounded by both $L(\underline{g}, P_A)$ (on the lower end) and $U(\bar{g}, P_A)$ (on the upper end). Let us restrict our attention to

$$L(\underline{g}, P_A) \leq U(\underline{g}, P_A). \quad (62)$$

This means that \underline{g} is integrable. Similarly, \bar{g} is integrable.

Now assume $\epsilon > 0$ and P was chosen such that $U(f, P) - L(f, P) < \epsilon$ (which is possible since f is integrable), then

$$U(\underline{g}, P_A) - L(\underline{g}, P_A) \leq \epsilon \quad (63)$$

and

$$U(\bar{g}, P_A) - L(\bar{g}, P_A) \leq \epsilon \quad (64)$$

so \underline{g} and \bar{g} are integrable on A .

3. From the inequalities earlier, we can write

$$L(f, P) \leq \int_A \bar{g}, \int_A \underline{g} \leq U(f, P) \quad (65)$$

which is true for every P . This means that we can take the infimum and supremum over all partitions, to get

$$\int f \leq \int \bar{g}, \int \underline{g} \leq \int f \quad (66)$$

so $\int f = \int \bar{g} = \int \underline{g}$ and we are done.

□