inverse function theorem:

We proved the inverse function theorem assuming that $f'(a) = I$. We shall now deduce the general case. We start with some auxiliary observations.

Let A be a 2×2 matrix and consider the linear transformation of R^2 to itself, $\varphi(x) = Ax$. This is a continuous map, simply because each coordinate of $\varphi(x)$ is a linear combinations of coordinates of x. If A is non-singular, then it has an inverse A^{-1} and the map $x \mapsto A^{-1}x$ is likewise continuous.

When A is non-singular, then for every closed set S, the set

$$A(S) = \{Ax : x \in S\}$$

is closed. Indeed, if $y_n \in A(S)$ and $y_n \to y$, then $A^{-1}(y_n) \in S$ and $A^{-1}(y_n) \to A^{-1}(y)$ and S being closed we conclude $A^{-1}(y) \in S$ which shows that $y = A(A^{-1}y) \in A(S)$. Since closed sets are precisely complements of open sets, we conclude that for any open set V, the set $A(V)$ is open.

Returning to the inverse function theorem, let $f : \Omega \to R^2$ be C^1 function with $f'(a) = A$, non-singular. Then define the function $g(x) = A^{-1}f(x)$ on Ω. This is again a C^1 function and $g'(a) = A^{-1}A = I$. Thus there is are open sets U, V such that $a \in U$, $b = g(a) \in V$, g is one-to-one on U onto V, the inverse map $h : V \to U$ is C^1, and $h'(y) = [g'(x)]^{-1}$ for $y \in V$ and $g(x) = y$.

the set $W = A(V)$ is an open set. Easy to see that $f : U \to W$ is one-one on U onto W, indeed $f(x) = Ag(x)$. The inverse map $\xi : W \to U$ is given by $\xi(z) = h(A^{-1}z)$ and is hence composition of C^1 maps. Thus it is again C^1. Chain rule now verifies the formula for the derivative of the inverse map ξ.

Integration:

We shall now proceed to imitate the concept of lower sums and upper sums and the concept of integral for function of two variables. Of course, this is not just for the sake of imitation. Just as finding areas motivated us towards integration last semester, finding volumes is one of the reasons for integrating functions of two variables, We do not spend time but you should understand why we are doing all this.

Basically, we would like to start with a bounded function defined on a bounded region; partition the region into small sets; for each set T in the partition wish to calculate $a(T)M(T)$ where $a(T)$ is the area of T and $M(T)$ is the supremum of the function in that set T. Sum of all these will then give us the upper sum. Similarly we define the lower sum.

There is one problem in implementing above plan. How do we know the area of T in the calculation above? We follow the maxim: solve the simplest problem first. From high school we are familiar with rectangles and their areas. so we decide to partition into rectangles. But if your original region is not a rectangle, then you can not partition it into rectangles.

Thus we consider, as a first step, bounded functions defined on rectangle and also consider partitions into rectangles with sides parallel to the axes.

Till further announcement, rectangle always means usual rectangle with sides parallel to the two standard axes.

Let $S = [a, b] \times [c, d]$ be a (closed) rectangle and $f : S \to R$ be a bounded function. Recall a partition of the interval $[a, b]$ is a finite sequence of points

$$a = a_0 < a_1 < a_2 < \cdots < a_k = b$$

or equivalently the intervals

$$[a_0, a_1], \quad [a_1, a_2], \quad \cdots \quad , [a_{k-1}, a_k].$$

If we also have a partition of $[c, d]$

$$c = c_0 < c_1 < c_2 < \cdots < c_l = d$$

then we define product partition of the rectangle S as the collection of the non-overlapping rectangles,

$$S_{ij} = [a_i, a_{i+1}] \times [b_j, b_{j+1}]; \quad 0 \le i \le k-1; 0 \le j \le l-1.$$

Such partitions of S are called product partitions. It is denoted simply as $\Pi = \{S_{ij}\}$.

Let us agree on meanings to some phrases. if we have a rectangle $T = [\alpha, \beta] \times [\gamma, \delta]$ we say that points

$$\{(x, y) : \alpha < x < \beta;\ \gamma < y < \delta\}$$

are interior points of the rectangle T. In other words it is precisely the set $(\alpha, \beta) \times (\gamma, \delta)$. The remaining points of T are called boundary points of T. In other words boundary consists of

$$\{\alpha\} \times [\gamma, \delta] \ \cup \ \{\beta\} \times [\gamma, \delta] \ \cup \ [\alpha, \beta] \times \{\gamma\} \ \cup \ [\alpha, \beta] \times \{\delta\}.$$

Two rectangles are non-overlapping if any point common to both of them is boundary point of both the rectangles. Thus, as sets they may not be disjoint, but they have no common interior points. Any two different rectangles above are non-overlapping.

For a rectangle T as above, its area is $(\beta - \alpha) \times (\delta - \gamma)$. This is also denoted by $|T|$.

Let us now consider a product partition as above of S. Denote

$$M_{ij} = \sup\{f(x) : \ x \in S_{ij}\}, \quad m_{ij} = \inf\{f(x) : \ x \in S_{ij}\}.$$

$$U(f, \Pi) \ = \ \sum_{i,j} M_{ij}|S_{ij}| \quad L(f, \Pi) \ = \ \sum_{i,j} m_{ij}|S_{ij}|.$$

$U(f, \Pi)$ is called the upper sum for the partition Π and $L(f, \Pi)$ is called the lower sum.

Say that f is **integrable** if the supremum of all lower sums equals infimum of the upper sums (over all product partitions). We denote this common value by

$$\int_S f; \quad \int_S f(x, y) d(x, y), \qquad \int_{[a,b]\times[c,d]} f(x, y) d(x, y).$$

this is called the integral or double integral of f.

Here are standard facts whose proofs are exactly same (?) as in the one dimensional case.

Theorem: $L(f, \Pi) \leq U(f, \Pi)$.

Recall that a partition η of the interval $[a, b]$ is finer than a partition π if every point that appears in π also appears in η. In other words η is obtained by adding more (possibly zero) points to π. Thus a finer partition cuts the interval into more pieces or into finer pieces. A product partition Π_2 is finer than Π_1 if the corresponding partitions on each side are finer for Π_2.

Theorem: If Π_2 is finer than Π_1, then

$$L(f, \Pi_1) \leq L(f, \Pi_2); \qquad U(f, \Pi_1) \geq U(f, \Pi_2).$$

In other words as the partition becomes finer the upper sums reduce while the lower sums increase.

Since we are considering only product partitions the theorem above was stated for product partitions, but it is true for any partitions, one finer than the other.

Theorem: For any partitions Π_1 and Π_2, $L(f, \Pi_1) \leq U(f, \Pi_2)$.

Theorem: f is integrable iff for any $\epsilon > 0$, there is a product partition Π such that $U(f, \Pi) - L(f, \Pi) < \epsilon$.

Proof: f is integrable iff sup of lower sums equals inf of upper sums. this is same as saying that for every $\epsilon > 0$, there is a lower sum and an upper sum which are ϵ-close. Since lower sums increase and upper sums decrease, this is same as saying that there is one partition for which the lower and upper sums differ by at most ϵ.

Theorem: Every continuous function is integrable.

The proof is along the expected lines. Given $\epsilon > 0$, use uniform continuity to get a product partition so that within each rectangle of the partition the values of the function differ by at most $\epsilon / Aa(S)$. Remember $a(S) = (b-a)(d-c)$, area of S,

We have been considering product partitions. We can consider any partition into rectangles, in fact, any reasonable partition. But to see that this also leads to the same answer, we have to wait. Right now, we can at least observe one thing. Any partition into rectangles parallel to the axes can also be considered, we arrive at the same answer.

Let us for this para, by a partition mean partition into rectangles with sides parallel to the axes. Superficially it appears if you allow all possible, not necessarily product, partitions you have many possibilities. But the crucial fact is the following.

Theorem: given any partition (rectangles, sides parallel to the axes) Π there is a finer partition Π_1 which is product partition.

In fact you only have to consider all the corners of all the rectangles of the partition Π, take their first coordinates as the partition of the side $[a, b]$; take all the second coordinates of all the corners as the partition of $[c, d]$ and consider Π_1 to be the product of these partitions of the two sides.

Let A be the set of all lower sums corresponding to only product partitions and B be the set of all lower sums corresponding to all partitions (into rectangles with sides parallel to the axes). Thus $A \subset B$, so that $\sup A \leq \sup B$.

Since finer partitions (whether product or not) have larger lower sums, the theorem above tells that for every number in B there is something larger than that in A. Thus $\sup B \leq \sup A$. This shows that A and B have the same sup.

Thus whether you make lower sums with product partitions or you make lower sums with all possible partitions (rectangles with sides parallel to the axes) you get the same sup. Similar remark holds for the upper sums. As a consequence if you define integrability using these partitions you get nothing new.

Theorem: If f and g are (bounded) integrable then so is $f + g$ and cF where c is any number. Further,

$$\int (f + g) = \int f + \int g; \quad \int cf = c \int f.$$

First observe the following. If f and g are two functions (real valued) on a set T with supremums $M(f)$ and $M(g)$, then for any point $x \in T$

$$(f + g)(x) = f(x) + g(x) \leq M(f) + M(g).$$

Thus $M(f + g) \leq M(f) + M(g)$. Similar remark applies for the infimums. This leads to the followng. For any partition Π

$$L(f) + L(g) \leq L(f + g) \leq U(f + g) \leq U(f) + U(g).$$

Since f and g are integrable, take partitions Π_n such that

$$L(f, \Pi_n) \to \int f; \quad U(f, \Pi_n) \to \int f;$$

$$L(g,\Pi_n) \to \int g,\ U(g,\Pi_n) \to \int g.$$

Then the earlier display shows that

$$L(f+g,\Pi_n) \to \int f + \int g; \quad U(f+g,\Pi_n) \to \int f \int g.$$

Tis shows that $f+g$ is integrable and $\int(f+g) = \int f + \int g$.

Similar and simpler argument shows $\int cf = c \int f$.

Let S be union of non-overlapping rectangles $S_1, \cdots S_k$ all having sides parallel to the axes. Then
Theorem: f is integrable on S iff it is integrable on each S_i and then

$$\int_S f = \sum \int_{S_i} f.$$

Strictly speaking, we should not say f is integrable on S_i, we should say restriction of f to S_i is integrable on S_i. Similarly, we should be writing

$$\int_{S_i} f_i$$

where f_i is restriction of f to S_i. But we shall not be so strict.

Proof: If f is integrable on S then you can take a partition Π such that $U(f,S,\Pi) - L(f,S,\Pi) < \epsilon$. By taking a larger partition, if necessary, we assume that each S_i is union of sets in the partition. If this is not already so, you only need to put-in the x coordinates of all the corner points of all the S_i in the partition of $[a,b]$ on the x-axis, and similarly put-in the y coordinates of all the corner points of all the rectangles S_i to get a partition of $[c,d]$. clearly for each i, those sets in Π that are contained in S_i is a product partition Π_i of S_i and

$$U(f,S_i,\Pi_i) - L(f,S_i,\Pi_i) \le U(f,S,\Pi) - L(f,S,\Pi) \le \epsilon$$

showing that f is integrable on each S_i.

Conversely, if f is integrable on each S_i then you take product partition of η_i of S_i for each i so that

$$U(f,S_i,\eta_i) - L(f,S_i,\eta_i) < \epsilon/k.$$

Putting all the x coordinates of all the corners of all the partition rectangles together we get a partition of $[a,b]$ and and putting together all the

y-coordinates of all the corners of all the partition rectangles together we get a partition of $[c, d]$ and thus we get product partition Π of S. If we restrict this Π to S_i we get a product partition Π_i which is finer than the η_i we started with on S_i. Thus

$$U(f, S, \Pi) - L(f, S, \Pi) = \sum_1^k [U(f, S_i, \Pi_i) \leq \epsilon.$$

This shows that f is integrable on S.

In fact this last display can be strengthened into

$$U(f, S, \Pi) = \sum U(f, S_i, \Pi_i); \quad L(f, S, \Pi) = \sum L(f, S_i, \Pi_i)$$

Small sets:

In case of functions of one variable, we did not answer the question: which functions (bounded function on a bounded interval) are integrable. This was because most of the functions that we come across at the elementary level have finitely many discontinuities and we have shown that such functions are integrable. This was enough for life to go on.

Unfortunately, in higher dimensions, we do need to tackle this problem head on. The reason is the following. Afterall, we need to integrate functions which areanot necessarily defined on rectangles, even if it ias defined on rectangles, then the rectangle need not have sides parallel tot eh axes. Thus partitions that we have been considering will be not enough. Of course, we would not complicate too much. If f is afunction given on a bounded set Ω, we just enclose Ω in a rectangle S, extend f to all of Ω by defining its value to be zero for points of $S - \Omega$ to be zero.

But then even if we started with a nice continuous function on Ω, this extensionn is rarely continuous on S and we need some assurance that we have not destroyed continuity too much and this extended function on S is indeed integrable. If moreover this value does not depend on which rectangle you choose to enclose Ω, then we are justified to regard integral of this extended function on the rectangle S as integral of f on Ω.

A set is small if it can be fit into any bag, no matter how small is the bag. More precisely, a set $A \subset R^2$ is small if given any $\epsilon > 0$, we can get finitely many or countable many rectangles $S_1, S_2, \cdots$ such that

$$\sum a(S_i) < \epsilon; \qquad A \subset \cup_i S_i.$$

It does not matter whether we take closed rectangles or open rectangles in the above definition. The reason is the following. If you can do with open rectangles $(a_i, b_i) \times (c_i, d_i)$ then you can, without changing areas consider theirclosed rectangles $[a_i, b_i] \times [c_i, d_i]$. Conversely, if you can do with closed recatngles, such a simple minded argument of removing boundary will not work because those open rectangles may not cover all of the set A. But this is achieved as follows. Let $\epsilon > 0$ be given. Get closed rectangles with total area smaller than $\epsilon/2$ covering A. Increase each of the rectangles a little bit so that the area of i-th rectangle is increased by only $\epsilon/2^{i+2}$ an make it closed. These will do.

Clearly every single point set is small. Even a countable set is a small set because you can use$\epsilon/2^i$ argument. In fact the samev $\epsilon/2^i$ allows you to show that union of countably many small sets is small. Also clear is that subset of a small set is small. It is a nice exercise to show that this concept agrees with your intuition by showing the following. A rectangle $[a, b] \times [c, d]$ with $a < b$, $c < d$ is indeed not small where as its boundary is small.

Here is then the relevance of small sets to our problem.

Theorem: Let f be a bounded function on a bounded rectangle $S = [a, b] \times [c, d]$. then f is integrable iff its set of discontinuity points form a small set.

We shall prove this theorem. But we first need some preliminaries.

Oscillation:

so we have to finally understand discotinuity points. Of course we did discuss a little about left limits, right limits and discontinuity points last semester. But now we need to get a quantitative feeling for discontinuity.

Let $f : S \subset R^2 \to R$ and $a \in S$. We want to understand how far away from continuity is f at the point a. for $\delta > 0$, let us define

$$O(a, \delta) = \sup\{f(x) : x \in [a - \delta, a + \delta]\} - \inf\{f(x) : x \in [a - \delta, a + \delta]\}.$$

Here we have used a notation, $[a - \delta, a + \delta]$ is not an interval but is the rectangle $[a_1 - \delta, a_1 + \delta] \times [a_2 - \delta, a_2 + \delta]$. This is suggestive if understood carefully, otherwise it would be confusing.

In case the function is continuous at a then definition of continuity tells us that, given any $\epsilon > 0$ we can choose a $\delta > 0$ so that $O(a, \delta) < 2\epsilon$. Observe that $O(a, \delta)$ decreases as $\delta \downarrow 0$. Let us define

$$O(a) = \lim_{\delta \downarrow 0} O(\delta).$$

This is called oscillation of the function at the point a. Here is its importance.

Theorem: f is continuous at a iff $O(a) = 0$.

In fact if the function is continuous we have already seen that given $\epsilon > 0$ there is $\delta > 0$ such that $O(a, \delta) < 2\epsilon$ and this holds for all smaller δ too. This shows that $O(a) = 0$. Conversely, if $O(a) = 0$ then given $\epsilon > 0$, there is a $\delta > 0$ so that $O(a, \delta) < \epsilon$. In particular $|(x) - f(a)| < \epsilon$ whenever $|x - a| < \delta$ showing continuity of f at a.

Theorem: For any $\epsilon > 0$, the set $\{a \in S : O(a) \geq \epsilon\}$ is closed in S. That is, if $a_n \to a$ and all these points a_n and a are in S and if each a_n is in this set then so is a.

Note that if $O(a) < \epsilon$ then there is $\delta > 0$ so that $O(a, \delta) < \epsilon$. But then for every point $b \in (a - \delta, a + \delta)$ we have small square around b contained in $(a - \delta, a + \delta)$ which shows that $O(b) < \epsilon$. In other words none of the a_n are in this square ahowing a_n does not converge to a; a contradiction.

In defining oscillation we have used squares. It is alright to use recatngles, but then $O(\delta)$ will be indexed by two numbers $O(\delta_1, \delta_2)$, lengths of the two sides. You should take limit as both $\delta_1 \to 0$ and $\delta_2 \to 0$. You get nothing new, simply because every such rectangle contains a square.

You can also take instead of square around a, a disc around $B(a, \delta) = \{x : ||x - a|| \leq \delta\}$ and define $O(a, \delta)$. Again since every square contains a disc and conversely, you get nothing new.

As you may have noticed, we have been taking closed square or closed disc. But you can take open squares or open discs too. Since closed squares contain open squares and open square contains a closed square, we get nothing new.

compact sets:

A closed bounded subset $K \subset R^2$ is simply called a compact set. Compact sets ahve the following nice property: if some one puts into many bags, we can fit it in finitely many of those bags.

Theorem: Let K be a compact set and you have a collection of open rectangles which cover K. This means that every point of K is in one of these rectangles. Then you can select finitely many of the given rectangles which also cover K.

The argument is standard and was seen several times. First put the set K in a big closed bounded rectangle S. Let $\mathcal{U}$ be a collection of open rectangles covering K for which the conclusion is false. Cut the rectangle into four parts at the mid points of the sides, one of these parts can not be covered by finitely many sets from $\mathcal{U}$. Take one such part S_1. Cut this into four pieces as above and pick a part S_2 so that the part of K in S_2 can not be covered by finitely many sets from $\mathcal{U}$. Thus we get a sequence of closed rectangles (S_n) which are decreasing.n fact lengths of sides are hlf length of the previous one; the part of K in S_n can not be covered by finitely many sets from $\mathcal{U}$. Cantor intersection theorem gives you a point a common to all the rectangles which will be in S because S is closed. This point is in some open rectangkle T of the family $\mathcal{U}$
and hence some $S_n \subset T$ contradicting the choice of S_n.

back to integration:

suppose that $f : S \to R$ bounded function on a rectangle $S = [a, b] \times [c, d]$. If $O(a) < \epsilon$ for all $a \in S$, then there is a product partition Π such that $U(\Pi) - L)\Pi) < \epsilon a(S)$.

Proof. For every point $a \in S$ there is a open rectangle T_a such that $O(f, T_a) < \epsilon$. In stead of showing δ we are showing the rectangle T_a. These open rectangles cover S and use finitely many of them to cover. Take their intersection with S to obtain finitely many rectangles$T_1, T_2, \cdots, T_k$ which cover S and in each of them the sup minus inf of the fumction is $< \epsilon$.. Now do the usual thing. Take first coordinates of the corners of these rectangles and the second coordinates of the corners to obtain partition of the sides of S Take product partition. Within each rectangle fo the partition, we have sup minus inf of the function is $< \epsilon$. Thus for this partition Π, we have $U(\Pi) - \text{Ł}(\Pi) < \epsilon a(S)$.

We shall now prove the theorem on integrability. Let f be a closed bounded rectangle and $f : S \to R$ be bounded, say $|f| \leq M$. Let us assume that the set D of points where f is not continuous is a small set. Fix $\epsilon > 0$. We wish to exhibit a partition so that $U - L < \epsilon$. The idea is the following. The set $D_\epsilon = \{x : O(x) \geq \epsilon/2a(S).\}$ is a closed and bounded set and hence is compact. Put each of these points in a small open rectangle of area $< \epsilon/2Ma(S)$. Take finitely many of these which cover D_ϵ. Let this part of S be denoted S_1. Thios finite collection already gives a partition of S_1 On the remaing part $S_2 = S - S_1$ the oscillation is small and we can make U_L smaller than $\epsilon/2$ by proper choice of partition, using previous result. The previous theorem is for a rectangle and our S_2 is unlikely to be a rectangle. We need to carefully argue and this we shall do later.

This is simple and will be precisely executed later..