Notes for Partial Differential Equations.

Kuttler

December 12, 2003

Contents

Part I

Review Of Advanced Calculus

Continuous Functions Of One Variable

There is a theorem about the integral of a continuous function which requires the notion of uniform continuity. This is discussed in this section. Consider the function $f(x) = \frac{1}{x}$ for $x \in (0,1)$. This is a continuous function because, it is continuous at every point of $(0,1)$. However, for a given $\varepsilon > 0$, the δ needed in the ε, δ definition of continuity becomes very small as x gets close to 0. The notion of uniform continuity involves being able to choose a single δ which works on the whole domain of f. Here is the definition.

Definition 1.1 *Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Then f is uniformly continuous if for every $\varepsilon > 0$, there exists a δ* ***depending only on*** *ε such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \varepsilon$.*

It is an amazing fact that under certain conditions continuity implies uniform continuity.

Definition 1.2 *A set, $K \subseteq \mathbb{R}$ is sequentially compact if whenever $\{a_n\} \subseteq K$ is a sequence, there exists a subsequence, $\{a_{n_k}\}$ such that this subsequence converges to a point of K.*

The following theorem is part of the Heine Borel theorem.

Theorem 1.3 *Every closed interval, $[a,b]$ is sequentially compact.*

Proof: Let $\{x_n\} \subseteq [a,b] \equiv I_0$. Consider the two intervals $\left[a, \frac{a+b}{2}\right]$ and $\left[\frac{a+b}{2}, b\right]$ each of which has length $(b-a)/2$. At least one of these intervals contains x_n for infinitely many values of n. Call this interval I_1. Now do for I_1 what was done for I_0. Split it in half and let I_2 be the interval which contains x_n for infinitely many values of n. Continue this way obtaining a sequence of nested intervals $I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \cdots$ where the length of I_n is $(b-a)/2^n$. Now pick n_1 such that $x_{n_1} \in I_1$, n_2 such that $n_2 > n_1$ and $x_{n_2} \in I_2, n_3$ such that $n_3 > n_2$ and $x_{n_3} \in I_3$, etc. (This can be done because in each case the intervals contained x_n for infinitely many values of n.) By the nested interval lemma there exists a point, c contained in all these intervals. Furthermore,

$$|x_{n_k} - c| < (b-a)\, 2^{-k}$$

and so $\lim_{k\to\infty} x_{n_k} = c \in [a,b]$. This proves the theorem.

Theorem 1.4 *Let $f : K \to \mathbb{R}$ be continuous where K is a sequentially compact set in $\mathbb{R}$. Then f is uniformly continuous on K.*

Proof: If this is not true, there exists $\varepsilon > 0$ such that for every $\delta > 0$ there exists a pair of points, x_δ and y_δ such that even though $|x_\delta - y_\delta| < \delta$, $|f(x_\delta) - f(y_\delta)| \geq \varepsilon$. Taking a succession of values for δ equal to $1, 1/2, 1/3, \cdots$, and letting the exceptional pair of points for $\delta = 1/n$ be denoted by x_n and y_n,

$$|x_n - y_n| < \frac{1}{n}, |f(x_n) - f(y_n)| \geq \varepsilon.$$

Now since K is sequentially compact, there exists a subsequence, $\{x_{n_k}\}$ such that $x_{n_k} \to z \in K$. Now $n_k \geq k$ and so

$$|x_{n_k} - y_{n_k}| < \frac{1}{k}.$$

Consequently, $y_{n_k} \to z$ also. (x_{n_k} is like a person walking toward a certain point and y_{n_k} is like a dog on a leash which is constantly getting shorter. Obviously y_{n_k} must also move toward the point also. You should give a precise proof of what is needed here.) By continuity of f

$$0 = |f(z) - f(z)| = \lim_{k \to \infty} |f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon,$$

an obvious contradiction. Therefore, the theorem must be true.

The following corollary follows from this theorem and Theorem 1.3.

Corollary 1.5 *Suppose I is a closed interval, $I = [a, b]$ and $f : I \to \mathbb{R}$ is continuous. Then f is uniformly continuous.*

1.1 Exercises

1. A function, $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous or just Lipschitz for short if there exists a constant, K such that

$$|f(x) - f(y)| \leq K|x - y|$$

for all $x, y \in D$. Show every Lipschitz function is uniformly continuous.

2. If $|x_n - y_n| \to 0$ and $x_n \to z$, show that $y_n \to z$ also.

3. Consider $f : (1, \infty) \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$. Show f is uniformly continuous even though the set on which f is defined is not sequentially compact.

4. If f is uniformly continuous, does it follow that $|f|$ is also uniformly continuous? If $|f|$ is uniformly continuous does it follow that f is uniformly continuous? Answer the same questions with "uniformly continuous" replaced with "continuous". Explain why.

1.2 Theorems About Continuous Functions

In this section, proofs of some theorems which have not been proved yet are given.

Theorem 1.6 *The following assertions are valid*

1. *The function, $af + bg$ is continuous at x when f, g are continuous at $x \in D(f) \cap D(g)$ and $a, b \in \mathbb{R}$.*

2. *If and f and g are each real valued functions continuous at x, then fg is continuous at x. If, in addition to this, $g(x) \neq 0$, then f/g is continuous at x.*

3. *If f is continuous at x, $f(x) \in D(g) \subseteq \mathbb{R}$, and g is continuous at $f(x)$,then $g \circ f$ is continuous at x.*

4. *The function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = |x|$ is continuous.*

Proof: First consider 1.) Let $\varepsilon > 0$ be given. By assumption, there exist $\delta_1 > 0$ such that whenever $|x-y| < \delta_1$, it follows $|f(x) - f(y)| < \frac{\varepsilon}{2(|a|+|b|+1)}$ and there exists $\delta_2 > 0$ such that whenever $|x-y| < \delta_2$, it follows that $|g(x) - g(y)| < \frac{\varepsilon}{2(|a|+|b|+1)}$. Then let $0 < \delta \leq \min(\delta_1, \delta_2)$. If $|x-y| < \delta$, then everything happens at once. Therefore, using the triangle inequality

$$|af(x) + bf(x) - (ag(y) + bg(y))|$$

$$\begin{aligned} &\leq |a| |f(x) - f(y)| + |b| |g(x) - g(y)| \\ &< |a| \left(\frac{\varepsilon}{2(|a|+|b|+1)}\right) + |b| \left(\frac{\varepsilon}{2(|a|+|b|+1)}\right) < \varepsilon. \end{aligned}$$

Now consider 2.) There exists $\delta_1 > 0$ such that if $|y-x| < \delta_1$, then $|f(x) - f(y)| < 1$. Therefore, for such y,

$$|f(y)| < 1 + |f(x)|.$$

It follows that for such y,

$$|fg(x) - fg(y)| \leq |f(x)g(x) - g(x)f(y)| + |g(x)f(y) - f(y)g(y)|$$

$$\begin{aligned} &\leq |g(x)| |f(x) - f(y)| + |f(y)| |g(x) - g(y)| \\ &\leq (1 + |g(x)| + |f(y)|) [|g(x) - g(y)| + |f(x) - f(y)|]. \end{aligned}$$

Now let $\varepsilon > 0$ be given. There exists δ_2 such that if $|x-y| < \delta_2$, then

$$|g(x) - g(y)| < \frac{\varepsilon}{2(1 + |g(x)| + |f(y)|)},$$

and there exists δ_3 such that if $|x-y| < \delta_3$, then

$$|f(x) - f(y)| < \frac{\varepsilon}{2(1 + |g(x)| + |f(y)|)}$$

Now let $0 < \delta \leq \min(\delta_1, \delta_2, \delta_3)$. Then if $|x-y| < \delta$, all the above hold at once and so

$$|fg(x) - fg(y)| \leq$$

$$\begin{aligned} &(1 + |g(x)| + |f(y)|) [|g(x) - g(y)| + |f(x) - f(y)|] \\ &< (1 + |g(x)| + |f(y)|) \left(\frac{\varepsilon}{2(1 + |g(x)| + |f(y)|)} + \frac{\varepsilon}{2(1 + |g(x)| + |f(y)|)}\right) = \varepsilon. \end{aligned}$$

This proves the first part of 2.) To obtain the second part, let δ_1 be as described above and let $\delta_0 > 0$ be such that for $|x-y| < \delta_0$,

$$|g(x) - g(y)| < |g(x)|/2$$

and so by the triangle inequality,

$$-|g(x)|/2 \leq |g(y)| - |g(x)| \leq |g(x)|/2$$

which implies $|g(y)| \geq |g(x)|/2$, and $|g(y)| < 3|g(x)|/2$.

Then if $|x-y| < \min(\delta_0, \delta_1)$,

$$\left|\frac{f(x)}{g(x)} - \frac{f(y)}{g(y)}\right| = \left|\frac{f(x)g(y) - f(y)g(x)}{g(x)g(y)}\right|$$
$$\leq \frac{|f(x)g(y) - f(y)g(x)|}{\left(\frac{|g(x)|^2}{2}\right)}$$
$$= \frac{2|f(x)g(y) - f(y)g(x)|}{|g(x)|^2}$$

$$\leq \frac{2}{|g(x)|^2}\left[|f(x)g(y) - f(y)g(y) + f(y)g(y) - f(y)g(x)|\right]$$
$$\leq \frac{2}{|g(x)|^2}\left[|g(y)||f(x) - f(y)| + |f(y)||g(y) - g(x)|\right]$$
$$\leq \frac{2}{|g(x)|^2}\left[\frac{3}{2}|g(x)||f(x) - f(y)| + (1 + |f(x)|)|g(y) - g(x)|\right]$$
$$\leq \frac{2}{|g(x)|^2}(1 + 2|f(x)| + 2|g(x)|)\left[|f(x) - f(y)| + |g(y) - g(x)|\right]$$
$$\equiv M\left[|f(x) - f(y)| + |g(y) - g(x)|\right]$$

where M is defined by

$$M \equiv \frac{2}{|g(x)|^2}(1 + 2|f(x)| + 2|g(x)|)$$

Now let δ_2 be such that if $|x-y| < \delta_2$, then

$$|f(x) - f(y)| < \frac{\varepsilon}{2}M^{-1}$$

and let δ_3 be such that if $|x-y| < \delta_3$, then

$$|g(y) - g(x)| < \frac{\varepsilon}{2}M^{-1}.$$

Then if $0 < \delta \leq \min(\delta_0, \delta_1, \delta_2, \delta_3)$, and $|x-y| < \delta$, everything holds and

$$\left|\frac{f(x)}{g(x)} - \frac{f(y)}{g(y)}\right| \leq M\left[|f(x) - f(y)| + |g(y) - g(x)|\right]$$

$$< M\left[\frac{\varepsilon}{2}M^{-1} + \frac{\varepsilon}{2}M^{-1}\right] = \varepsilon.$$

This completes the proof of the second part of 2.)

Note that in these proofs no effort is made to find some sort of "best" δ. The problem is one which has a yes or a no answer. Either is it or it is not continuous.

Now consider 3.). If f is continuous at x, $f(x) \in D(g) \subseteq \mathbb{R}^p$, and g is continuous at $f(x)$, then $g \circ f$ is continuous at x. Let $\varepsilon > 0$ be given. Then there exists $\eta > 0$ such that if $|y - f(x)| < \eta$ and $y \in D(g)$, it follows that $|g(y) - g(f(x))| < \varepsilon$. From continuity of f at x, there exists $\delta > 0$ such that if $|x-z| < \delta$ and $z \in D(f)$, then $|f(z) - f(x)| < \eta$. Then if $|x-z| < \delta$ and $z \in D(g \circ f) \subseteq D(f)$, all the above hold and so

$$|g(f(z)) - g(f(x))| < \varepsilon.$$

This proves part 3.)

To verify part 4.), let $\varepsilon > 0$ be given and let $\delta = \varepsilon$. Then if $|x-y| < \delta$, the triangle inequality implies

$$\begin{aligned}|f(x) - f(y)| &= ||x| - |y|| \\ &\leq |x-y| < \delta = \varepsilon.\end{aligned}$$

This proves part 4.) and completes the proof of the theorem.

Next here is a proof of the intermediate value theorem.

Theorem 1.7 *Suppose $f : [a,b] \to \mathbb{R}$ is continuous and suppose $f(a) < c < f(b)$. Then there exists $x \in (a,b)$ such that $f(x) = c$.*

Proof: Let $d = \frac{a+b}{2}$ and consider the intervals $[a,d]$ and $[d,b]$. If $f(d) \geq c$, then on $[a,d]$, the function is $\leq c$ at one end point and $\geq c$ at the other. On the other hand, if $f(d) \leq c$, then on $[d,b]$ $f \geq 0$ at one end point and ≤ 0 at the other. Pick the interval on which f has values which are at least as large as c and values no larger than c. Now consider that interval, divide it in half as was done for the original interval and argue that on one of these smaller intervals, the function has values at least as large as c and values no larger than c. Continue in this way. Next apply the nested interval lemma to get x in all these intervals. In the n^{th} interval, let x_n, y_n be elements of this interval such that $f(x_n) \leq c, f(y_n) \geq c$. Now $|x_n - x| \leq (b-a) 2^{-n}$ and $|y_n - x| \leq (b-a) 2^{-n}$ and so $x_n \to x$ and $y_n \to x$. Therefore,

$$f(x) - c = \lim_{n\to\infty} (f(x_n) - c) \leq 0$$

while

$$f(x) - c = \lim_{n\to\infty} (f(y_n) - c) \geq 0.$$

Consequently $f(x) = c$ and this proves the theorem.

Lemma 1.8 *Let $\phi : [a,b] \to \mathbb{R}$ be a continuous function and suppose ϕ is $1-1$ on (a,b). Then ϕ is either strictly increasing or strictly decreasing on $[a,b]$.*

Proof: First it is shown that ϕ is either strictly increasing or strictly decreasing on (a,b).

If ϕ is not strictly decreasing on (a,b), then there exists $x_1 < y_1$, $x_1, y_1 \in (a,b)$ such that

$$(\phi(y_1) - \phi(x_1))(y_1 - x_1) > 0.$$

If for some other pair of points, $x_2 < y_2$ with $x_2, y_2 \in (a,b)$, the above inequality does not hold, then since ϕ is $1-1$,

$$(\phi(y_2) - \phi(x_2))(y_2 - x_2) < 0.$$

Let $x_t \equiv tx_1 + (1-t) x_2$ and $y_t \equiv ty_1 + (1-t) y_2$. Then $x_t < y_t$ for all $t \in [0,1]$ because

$$tx_1 \leq ty_1 \text{ and } (1-t) x_2 \leq (1-t) y_2$$

with strict inequality holding for at least one of these inequalities since not both t and $(1-t)$ can equal zero. Now define

$$h(t) \equiv (\phi(y_t) - \phi(x_t))(y_t - x_t).$$

Since h is continuous and $h(0) < 0$, while $h(1) > 0$, there exists $t \in (0,1)$ such that $h(t) = 0$. Therefore, both x_t and y_t are points of (a,b) and $\phi(y_t) - \phi(x_t) = 0$ contradicting the assumption that ϕ is one to one. It follows ϕ is either strictly increasing or strictly decreasing on (a,b).

This property of being either strictly increasing or strictly decreasing on (a,b) carries over to $[a,b]$ by the continuity of ϕ. Suppose ϕ is strictly increasing on (a,b), a similar argument holding for ϕ strictly decreasing on (a,b). If $x > a$, then pick $y \in (a,x)$ and from the above, $\phi(y) < \phi(x)$. Now by continuity of ϕ at a,

$$\phi(a) = \lim_{x\to a+} \phi(z) \leq \phi(y) < \phi(x).$$

Therefore, $\phi(a) < \phi(x)$ whenever $x \in (a,b)$. Similarly $\phi(b) > \phi(x)$ for all $x \in (a,b)$. This proves the lemma.

Corollary 1.9 *Let $f : (a,b) \to \mathbb{R}$ be one to one and continuous. Then $f(a,b)$ is an open interval, (c,d) and $f^{-1} : (c,d) \to (a,b)$ is continuous.*

Proof: Since f is either strictly increasing or strictly decreasing, it follows that $f(a,b)$ is an open interval, (c,d). Assume f is decreasing. Now let $x \in (a,b)$. Why is f^{-1} is continuous at $f(x)$? Since f is decreasing, if $f(x) < f(y)$, then $y \equiv f^{-1}(f(y)) < x \equiv f^{-1}(f(x))$ and so f^{-1} is also decreasing. Let $\varepsilon > 0$ be given. Let $\varepsilon > \eta > 0$ and $(x-\eta, x+\eta) \subseteq (a,b)$. Then $f(x) \in (f(x+\eta), f(x-\eta))$. Let $\delta = \min(f(x) - f(x+\eta), f(x-\eta) - f(x))$. Then if

$$|f(z) - f(x)| < \delta,$$

it follows

$$z \equiv f^{-1}(f(z)) \in (x-\eta, x+\eta) \subseteq (x-\varepsilon, x+\varepsilon)$$

so

$$\left|f^{-1}(f(z)) - x\right| = \left|f^{-1}(f(z)) - f^{-1}(f(x))\right| < \varepsilon.$$

This proves the theorem in the case where f is strictly decreasing. The case where f is increasing is similar.

The Integral

The integral originated in attempts to find areas of various shapes and the ideas involved in finding integrals are much older than the ideas related to finding derivatives. In fact, Archimedes[1] was finding areas of various curved shapes about 250 B.C.

2.1 Upper And Lower Sums

The Riemann integral pertains to bounded functions which are defined on a bounded interval. Let $[a,b]$ be a closed interval. A set of points in $[a,b]$, $\{x_0,\cdot\cdot\cdot,x_n\}$ is a partition if

$$a = x_0 < x_1 < \cdots < x_n = b.$$

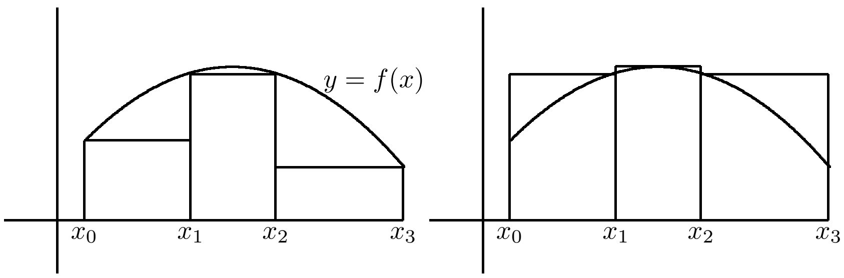
Such partitions are denoted by P or Q. For f a bounded function defined on $[a,b]$, let

$$M_i(f) \equiv \sup\{f(x) : x \in [x_{i-1}, x_i]\},$$
$$m_i(f) \equiv \inf\{f(x) : x \in [x_{i-1}, x_i]\}.$$

Also let $\Delta x_i \equiv x_i - x_{i-1}$. Then define upper and lower sums as

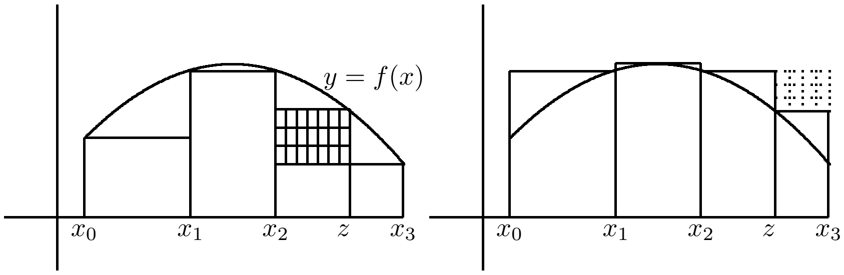
$$U(f,P) \equiv \sum_{i=1}^{n} M_i(f)\,\Delta x_i \text{ and } L(f,P) \equiv \sum_{i=1}^{n} m_i(f)\,\Delta x_i$$

respectively. The numbers, $M_i(f)$ and $m_i(f)$, are well defined real numbers because f is assumed to be bounded and $\mathbb{R}$ is complete. Thus the set $S = \{f(x) : x \in [x_{i-1}, x_i]\}$ is bounded above and below. In the following picture, the sum of the areas of the rectangles in the picture on the left is a lower sum for the function in the picture and the sum of the areas of the rectangles in the picture on the right is an upper sum for the same function which uses the same partition.



[1]Archimedes 287-212 B.C. found areas of curved regions by stuffing them with simple shapes which he knew the area of and taking a limit. He also made fundamental contributions to physics. The story is told about how he determined that a gold smith had cheated the king by giving him a crown which was not solid gold as had been claimed. He did this by finding the amount of water displaced by the crown and comparing with the amount of water it should have displaced if it had been solid gold.

What happens when you add in more points in a partition? The following pictures illustrate in the context of the above example. In this example a single additional point, labeled z has been added in.



Note how the lower sum got larger by the amount of the area in the shaded rectangle and the upper sum got smaller by the amount in the rectangle shaded by dots. In general this is the way it works and this is shown in the following lemma.

Lemma 2.1 *If $P \subseteq Q$ then*

$$U(f,Q) \leq U(f,P), \text{ and } L(f,P) \leq L(f,Q).$$

Proof: This is verified by adding in one point at a time. Thus let $P = \{x_0, \cdots, x_n\}$ and let $Q = \{x_0, \cdots, x_k, y, x_{k+1}, \cdots, x_n\}$. Thus exactly one point, y, is added between x_k and x_{k+1}. Now the term in the upper sum which corresponds to the interval $[x_k, x_{k+1}]$ in $U(f,P)$ is

$$\sup\{f(x) : x \in [x_k, x_{k+1}]\}(x_{k+1} - x_k) \tag{2.1}$$

and the term which corresponds to the interval $[x_k, x_{k+1}]$ in $U(f,Q)$ is

$$\sup\{f(x) : x \in [x_k, y]\}(y - x_k) + \sup\{f(x) : x \in [y, x_{k+1}]\}(x_{k+1} - y) \tag{2.2}$$

$$\equiv M_1(y - x_k) + M_2(x_{k+1} - y) \tag{2.3}$$

All the other terms in the two sums coincide. Now $\sup\{f(x) : x \in [x_k, x_{k+1}]\} \geq \max(M_1, M_2)$ and so the expression in (2.2) is no larger than

$$\sup\{f(x) : x \in [x_k, x_{k+1}]\}(x_{k+1} - y) + \sup\{f(x) : x \in [x_k, x_{k+1}]\}(y - x_k)$$

$$= \sup\{f(x) : x \in [x_k, x_{k+1}]\}(x_{k+1} - x_k),$$

the term corresponding to the interval, $[x_k, x_{k+1}]$ and $U(f,P)$. This proves the first part of the lemma pertaining to upper sums because if $Q \supseteq P$, one can obtain Q from P by adding in one point at a time and each time a point is added, the corresponding upper sum either gets smaller or stays the same. The second part is similar and is left as an exercise.

Lemma 2.2 *If P and Q are two partitions, then*

$$L(f,P) \leq U(f,Q).$$

Proof: By Lemma 2.1,

$$L(f,P) \leq L(f,P \cup Q) \leq U(f,P \cup Q) \leq U(f,Q).$$

Definition 2.3

$$\overline{I} \equiv \inf\{U(f,Q) \text{ where } Q \text{ is a partition}\}$$

$$\underline{I} \equiv \sup\{L(f,P) \text{ where } P \text{ is a partition}\}.$$

Note that $\underline{I}$ and $\overline{I}$ are well defined real numbers.

Theorem 2.4 $\underline{I} \leq \overline{I}$.

Proof: From Lemma 2.2,

$$\underline{I} = \sup\{L(f,P) \text{ where } P \text{ is a partition}\} \leq U(f,Q)$$

because $U(f,Q)$ is an upper bound to the set of all lower sums and so it is no smaller than the least upper bound. Therefore, since Q is arbitrary,

$$\begin{aligned}\underline{I} &= \sup\{L(f,P) \text{ where } P \text{ is a partition}\} \\ &\leq \inf\{U(f,Q) \text{ where } Q \text{ is a partition}\} \equiv \overline{I}\end{aligned}$$

where the inequality holds because it was just shown that $\underline{I}$ is a lower bound to the set of all upper sums and so it is no larger than the greatest lower bound of this set. This proves the theorem.

Definition 2.5 *A bounded function f is Riemann integrable, written as*

$$f \in R([a,b])$$

if

$$\underline{I} = \overline{I}$$

and in this case,

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

Thus, in words, the Riemann integral is the unique number which lies between all upper sums and all lower sums if there is such a unique number.

Recall the following Proposition which comes from the definitions.

Proposition 2.6 *Let S be a nonempty set and suppose* $\sup(S)$ *exists. Then for every $\delta > 0$,*

$$S \cap (\sup(S) - \delta, \sup(S)] \neq \emptyset.$$

If $\inf(S)$ *exists, then for every $\delta > 0$,*

$$S \cap [\inf(S), \inf(S) + \delta) \neq \emptyset.$$

This proposition implies the following theorem which is used to determine the question of Riemann integrability.

Theorem 2.7 *A bounded function f is Riemann integrable if and only if for all $\varepsilon > 0$, there exists a partition P such that*

$$U(f,P) - L(f,P) < \varepsilon. \tag{2.4}$$

Proof: First assume f is Riemann integrable. Then let P and Q be two partitions such that

$$U(f,Q) < \overline{I} + \varepsilon/2,\ L(f,P) > \underline{I} - \varepsilon/2.$$

Then since $\underline{I} = \overline{I}$,

$$U(f, Q \cup P) - L(f, P \cup Q) \leq U(f,Q) - L(f,P) < \overline{I} + \varepsilon/2 - (\underline{I} - \varepsilon/2) = \varepsilon.$$

Now suppose that for all $\varepsilon > 0$ there exists a partition such that (2.4) holds. Then for given ε and partition P corresponding to ε

$$\overline{I} - \underline{I} \leq U(f,P) - L(f,P) \leq \varepsilon.$$

Since ε is arbitrary, this shows $\underline{I} = \overline{I}$ and this proves the theorem.

The condition described in the theorem is called the Riemann criterion .

Not all bounded functions are Riemann integrable. For example, let

$$f(x) \equiv \begin{cases} 1 \text{ if } x \in \mathbb{Q} \\ 0 \text{ if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \tag{2.5}$$

Then if $[a,b] = [0,1]$ all upper sums for f equal 1 while all lower sums for f equal 0. Therefore the Riemann criterion is violated for $\varepsilon = 1/2$.

2.2 Exercises

1. Prove the second half of Lemma 2.1 about lower sums.

2. Verify that for f given in (2.5), the lower sums on the interval $[0,1]$ are all equal to zero while the upper sums are all equal to one.

3. Let $f(x) = 1 + x^2$ for $x \in [-1, 3]$ and let $P = \left\{-1, -\frac{1}{3}, 0, \frac{1}{2}, 1, 2\right\}$. Find $U(f,P)$ and $L(f,P)$.

4. Show that if $f \in R([a,b])$, there exists a partition, $\{x_0, \cdots, x_n\}$ such that for any $z_k \in [x_k, x_{k+1}]$,

$$\left| \int_a^b f(x)\, dx - \sum_{k=1}^n f(z_k)(x_k - x_{k-1}) \right| < \varepsilon$$

   This sum, $\sum_{k=1}^n f(z_k)(x_k - x_{k-1})$, is called a Riemann sum and this exercise shows that the integral can always be approximated by a Riemann sum.

5. Let $P = \left\{1, 1\frac{1}{4}, 1\frac{1}{2}, 1\frac{3}{4}, 2\right\}$. Find upper and lower sums for the function, $f(x) = \frac{1}{x}$ using this partition. What does this tell you about $\ln(2)$?

6. If $f \in R([a,b])$ and f is changed at finitely many points, show the new function is also in $R([a,b])$.

7. Define a "left sum" as

$$\sum_{k=1}^n f(x_{k-1})(x_k - x_{k-1})$$

   and a "right sum",

$$\sum_{k=1}^n f(x_k)(x_k - x_{k-1}).$$

Also suppose that all partitions have the property that $x_k - x_{k-1}$ equals a constant, $(b-a)/n$ so the points in the partition are equally spaced, and define the integral to be the number these right and left sums get close to as n gets larger and larger. Show that for f given in (2.5), $\int_0^x f(t)\ dt = 1$ if x is rational and $\int_0^x f(t)\ dt = 0$ if x is irrational. It turns out that the correct answer should always equal zero for that function, regardless of whether x is rational. This is shown in more advanced courses when the Lebesgue integral is studied. This illustrates why this method of defining the integral in terms of left and right sums is total nonsense.

2.3 Functions Of Riemann Integrable Functions

It is often necessary to consider functions of Riemann integrable functions and a natural question is whether these are Riemann integrable. The following theorem gives a partial answer to this question. This is not the most general theorem which will relate to this question but it will be enough for the needs of this book.

Theorem 2.8 *Let f, g be bounded functions and let $f([a,b]) \subseteq [c_1, d_1]$ and $g([a,b]) \subseteq [c_2, d_2]$. Let $H : [c_1, d_1] \times [c_2, d_2] \to \mathbb{R}$ satisfy,*

$$|H(a_1, b_1) - H(a_2, b_2)| \leq K[|a_1 - a_2| + |b_1 - b_2|]$$

for some constant K. Then if $f, g \in R([a,b])$ it follows that $H \circ (f, g) \in R([a,b])$.

Proof: In the following claim, $M_i(h)$ and $m_i(h)$ have the meanings assigned above with respect to some partition of $[a,b]$ for the function, h.

Claim: The following inequality holds.

$$|M_i(H \circ (f,g)) - m_i(H \circ (f,g))| \leq$$

$$K[|M_i(f) - m_i(f)| + |M_i(g) - m_i(g)|].$$

Proof of the claim: By the above proposition, there exist $x_1, x_2 \in [x_{i-1}, x_i]$ be such that

$$H(f(x_1), g(x_1)) + \eta > M_i(H \circ (f,g)),$$

and

$$H(f(x_2), g(x_2)) - \eta < m_i(H \circ (f,g)).$$

Then

$$|M_i(H \circ (f,g)) - m_i(H \circ (f,g))|$$

$$\begin{aligned}
&< 2\eta + |H(f(x_1), g(x_1)) - H(f(x_2), g(x_2))| \\
&< 2\eta + K[|f(x_1) - f(x_2)| + |g(x_1) - g(x_2)|] \\
&\leq 2\eta + K[|M_i(f) - m_i(f)| + |M_i(g) - m_i(g)|].
\end{aligned}$$

Since $\eta > 0$ is arbitrary, this proves the claim.

Now continuing with the proof of the theorem, let P be such that

$$\sum_{i=1}^{n} (M_i(f) - m_i(f)) \Delta x_i < \frac{\varepsilon}{2K},\ \sum_{i=1}^{n} (M_i(g) - m_i(g)) \Delta x_i < \frac{\varepsilon}{2K}.$$

Then from the claim,

$$\sum_{i=1}^{n} \left(M_i \left(H \circ (f, g)\right) - m_i \left(H \circ (f, g)\right)\right) \Delta x_i$$

$$< \sum_{i=1}^{n} K \left[\left|M_i (f) - m_i (f)\right| + \left|M_i (g) - m_i (g)\right|\right] \Delta x_i < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows $H \circ (f, g)$ satisfies the Riemann criterion and hence $H \circ (f, g)$ is Riemann integrable as claimed. This proves the theorem.

This theorem implies that if f, g are Riemann integrable, then so is $af + bg, |f|, f^2$, along with infinitely many other such continuous combinations of Riemann integrable functions. For example, to see that $|f|$ is Riemann integrable, let $H(a, b) = |a|$. Clearly this function satisfies the conditions of the above theorem and so $|f| = H(f, f) \in R([a, b])$ as claimed. The following theorem gives an example of many functions which are Riemann integrable.

Theorem 2.9 *Let $f : [a, b] \to \mathbb{R}$ be either increasing or decreasing on $[a, b]$. Then $f \in R([a, b])$.*

Proof: Let $\varepsilon > 0$ be given and let

$$x_i = a + i\left(\frac{b-a}{n}\right), \ i = 0, \cdots, n.$$

Then since f is increasing,

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{i=1}^{n} \left(f(x_i) - f(x_{i-1})\right)\left(\frac{b-a}{n}\right) \\ &= \left(f(b) - f(a)\right)\left(\frac{b-a}{n}\right) < \varepsilon \end{aligned}$$

whenever n is large enough. Thus the Riemann criterion is satisfied and so the function is Riemann integrable. The proof for decreasing f is similar.

Corollary 2.10 *Let $[a, b]$ be a bounded closed interval and let $\phi : [a, b] \to \mathbb{R}$ be Lipschitz continuous. Then $\phi \in R([a, b])$. Recall that a function, ϕ, is Lipschitz continuous if there is a constant, K, such that for all x, y,*

$$|\phi(x) - \phi(y)| < K|x - y|.$$

Proof: Let $f(x) = x$. Then by Theorem 2.9, f is Riemann integrable. Let $H(a, b) \equiv \phi(a)$. Then by Theorem 2.8 $H \circ (f, f) = \phi \circ f = \phi$ is also Riemann integrable. This proves the corollary.

In fact, it is enough to assume ϕ is continuous, although this is harder. This is the content of the next theorem which is where the difficult theorems about continuity and uniform continuity are used.

Theorem 2.11 *Suppose $f : [a, b] \to \mathbb{R}$ is continuous. Then $f \in R([a, b])$.*

Proof: By Corollary 1.5 on Page 8, f is uniformly continuous on $[a, b]$. Therefore, if $\varepsilon > 0$ is given, there exists a $\delta > 0$ such that if $|x_i - x_{i-1}| < \delta$, then $M_i - m_i < \frac{\varepsilon}{b-a}$. Let

$$P \equiv \{x_0, \cdots, x_n\}$$

be a partition with $|x_i - x_{i-1}| < \delta$. Then

$$U(f, P) - L(f, P) < \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

By the Riemann criterion, $f \in R([a, b])$. This proves the theorem.

2.4 Properties Of The Integral

The integral has many important algebraic properties. First here is a simple lemma.

Lemma 2.12 *Let S be a nonempty set which is bounded above and below. Then if $-S \equiv \{-x : x \in S\}$,*

$$\sup(-S) = -\inf(S) \tag{2.6}$$

and

$$\inf(-S) = -\sup(S). \tag{2.7}$$

Proof: Consider (2.6). Let $x \in S$. Then $-x \leq \sup(-S)$ and so $x \geq -\sup(-S)$. If follows that $-\sup(-S)$ is a lower bound for S and therefore, $-\sup(-S) \leq \inf(S)$. This implies $\sup(-S) \geq -\inf(S)$. Now let $-x \in -S$. Then $x \in S$ and so $x \geq \inf(S)$ which implies $-x \leq -\inf(S)$. Therefore, $-\inf(S)$ is an upper bound for $-S$ and so $-\inf(S) \geq \sup(-S)$. This shows (2.6). Formula (2.7) is similar and is left as an exercise.

In particular, the above lemma implies that for $M_i(f)$ and $m_i(f)$ defined above $M_i(-f) = -m_i(f)$, and $m_i(-f) = -M_i(f)$.

Lemma 2.13 *If $f \in R([a,b])$ then $-f \in R([a,b])$ and*

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

Proof: The first part of the conclusion of this lemma follows from Theorem 2.9 since the function $\phi(y) \equiv -y$ is Lipschitz continuous. Now choose P such that

$$\int_a^b -f(x)\,dx - L(-f,P) < \varepsilon.$$

Then since $m_i(-f) = -M_i(f)$,

$$\varepsilon > \int_a^b -f(x)\,dx - \sum_{i=1}^n m_i(-f)\,\Delta x_i = \int_a^b -f(x)\,dx + \sum_{i=1}^n M_i(f)\,\Delta x_i$$

which implies

$$\varepsilon > \int_a^b -f(x)\,dx + \sum_{i=1}^n M_i(f)\,\Delta x_i \geq \int_a^b -f(x)\,dx + \int_a^b f(x)\,dx.$$

Thus, since ε is arbitrary,

$$\int_a^b -f(x)\,dx \leq -\int_a^b f(x)\,dx$$

whenever $f \in R([a,b])$. It follows

$$\int_a^b -f(x)\,dx \leq -\int_a^b f(x)\,dx = -\int_a^b -(-f(x))\,dx \leq \int_a^b -f(x)\,dx$$

and this proves the lemma.

Theorem 2.14 *The integral is linear,*

$$\int_a^b (\alpha f + \beta g)(x)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx.$$

whenever $f, g \in R([a,b])$ and $\alpha, \beta \in \mathbb{R}$.

Proof: First note that by Theorem 2.8, $\alpha f + \beta g \in R([a,b])$. To begin with, consider the claim that if $f, g \in R([a,b])$ then

$$\int_a^b (f+g)(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx. \tag{2.8}$$

Let P_1,Q_1 be such that

$$U(f, Q_1) - L(f, Q_1) < \varepsilon/2,\ U(g, P_1) - L(g, P_1) < \varepsilon/2.$$

Then letting $P \equiv P_1 \cup Q_1$, Lemma 2.1 implies

$$U(f, P) - L(f, P) < \varepsilon/2, \text{ and } U(g, P) - U(g, P) < \varepsilon/2.$$

Next note that

$$m_i(f+g) \geq m_i(f) + m_i(g),\ M_i(f+g) \leq M_i(f) + M_i(g).$$

Therefore,

$$L(g+f, P) \geq L(f, P) + L(g, P),\ U(g+f, P) \leq U(f, P) + U(g, P).$$

For this partition,

$$\begin{aligned}\int_a^b (f+g)(x)\, dx &\in [L(f+g, P), U(f+g, P)] \\ &\subseteq [L(f, P) + L(g, P), U(f, P) + U(g, P)]\end{aligned}$$

and

$$\int_a^b f(x)\, dx + \int_a^b g(x)\, dx \in [L(f, P) + L(g, P), U(f, P) + U(g, P)].$$

Therefore,

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

$$U(f, P) + U(g, P) - (L(f, P) + L(g, P)) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This proves (2.8) since ε is arbitrary.

It remains to show that

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

Suppose first that $\alpha \geq 0$. Then

$$\int_a^b \alpha f(x)\, dx \equiv \sup\{L(\alpha f, P) : P \text{ is a partition}\} =$$

$$\alpha \sup\{L(f, P) : P \text{ is a partition}\} \equiv \alpha \int_a^b f(x)\, dx.$$

If $\alpha < 0$, then this and Lemma 2.13 imply

$$\int_a^b \alpha f(x)\ dx = \int_a^b (-\alpha)(-f(x))\ dx$$

$$= (-\alpha)\int_a^b (-f(x))\ dx = \alpha \int_a^b f(x)\ dx.$$

This proves the theorem.

Theorem 2.15 *If $f \in R([a,b])$ and $f \in R([b,c])$, then $f \in R([a,c])$ and*

$$\int_a^c f(x)\ dx = \int_a^b f(x)\ dx + \int_b^c f(x)\ dx. \tag{2.9}$$

Proof: Let P_1 be a partition of $[a,b]$ and P_2 be a partition of $[b,c]$ such that

$$U(f,P_i) - L(f,P_i) < \varepsilon/2,\ i = 1,2.$$

Let $P \equiv P_1 \cup P_2$. Then P is a partition of $[a,c]$ and

$$U(f,P) - L(f,P)$$

$$= U(f,P_1) - L(f,P_1) + U(f,P_2) - L(f,P_2) < \varepsilon/2 + \varepsilon/2 = \varepsilon. \tag{2.10}$$

Thus, $f \in R([a,c])$ by the Riemann criterion and also for this partition,

$$\begin{aligned}\int_a^b f(x)\ dx + \int_b^c f(x)\ dx &\in [L(f,P_1) + L(f,P_2), U(f,P_1) + U(f,P_2)] \\ &= [L(f,P), U(f,P)]\end{aligned}$$

and

$$\int_a^c f(x)\ dx \in [L(f,P), U(f,P)].$$

Hence by (2.10),

$$\left|\int_a^c f(x)\ dx - \left(\int_a^b f(x)\ dx + \int_b^c f(x)\ dx\right)\right| < U(f,P) - L(f,P) < \varepsilon$$

which shows that since ε is arbitrary, (2.9) holds. This proves the theorem.

Corollary 2.16 *Let $[a,b]$ be a closed and bounded interval and suppose that*

$$a = y_1 < y_2 \cdots < y_l = b$$

and that f is a bounded function defined on $[a,b]$ which has the property that f is either increasing on $[y_j, y_{j+1}]$ or decreasing on $[y_j, y_{j+1}]$ for $j = 1, \cdots, l-1$. Then $f \in R([a,b])$.

Proof: This follows from Theorem 2.15 and Theorem 2.9.
The symbol, $\int_a^b f(x)\ dx$ when $a > b$ has not yet been defined.

Definition 2.17 *Let $[a,b]$ be an interval and let $f \in R([a,b])$. Then*

$$\int_b^a f(x)\ dx \equiv -\int_a^b f(x)\ dx.$$

Note that with this definition,

$$\int_a^a f(x)\ dx = -\int_a^a f(x)\ dx$$

and so

$$\int_a^a f(x)\ dx = 0.$$

Theorem 2.18 *Assuming all the integrals make sense,*

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

Proof: This follows from Theorem 2.15 and Definition 2.17. For example, assume

$$c \in (a,b).$$

Then from Theorem 2.15,

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

and so by Definition 2.17,

$$\begin{aligned}\int_a^c f(x)\ dx &= \int_a^b f(x)\ dx - \int_c^b f(x)\ dx \\ &= \int_a^b f(x)\ dx + \int_b^c f(x)\ dx.\end{aligned}$$

The other cases are similar.

The following properties of the integral have either been established or they follow quickly from what has been shown so far.

$$\text{If } f \in R([a,b]) \text{ then if } c \in [a,b],\ f \in R([a,c]), \tag{2.11}$$

$$\int_a^b \alpha\, dx = \alpha(b-a), \tag{2.12}$$

$$\int_a^b (\alpha f + \beta g)(x)\ dx = \alpha \int_a^b f(x)\ dx + \beta \int_a^b g(x)\ dx, \tag{2.13}$$

$$\int_a^b f(x)\ dx + \int_b^c f(x)\ dx = \int_a^c f(x)\ dx, \tag{2.14}$$

$$\int_a^b f(x)\ dx \geq 0 \text{ if } f(x) \geq 0 \text{ and } a < b, \tag{2.15}$$

$$\left| \int_a^b f(x)\ dx \right| \leq \left| \int_a^b |f(x)|\ dx \right|. \tag{2.16}$$

The only one of these claims which may not be completely obvious is the last one. To show this one, note that

$$|f(x)| - f(x) \geq 0,\ |f(x)| + f(x) \geq 0.$$

Therefore, by (2.15) and (2.13), if $a < b$,

$$\int_a^b |f(x)|\ dx \geq \int_a^b f(x)\ dx$$

and

$$\int_a^b |f(x)|\ dx \geq -\int_a^b f(x)\ dx.$$

Therefore,

$$\int_a^b |f(x)|\ dx \geq \left| \int_a^b f(x)\ dx \right|.$$

If $b < a$ then the above inequality holds with a and b switched. This implies (2.16).

2.5 Fundamental Theorem Of Calculus

With these properties, it is easy to prove the fundamental theorem of calculus[2]. Let $f \in R([a,b])$. Then by (2.11) $f \in R([a,x])$ for each $x \in [a,b]$. The first version of the fundamental theorem of calculus is a statement about the derivative of the function

$$x \to \int_a^x f(t)\ dt.$$

Theorem 2.19 *Let $f \in R([a,b])$ and let*

$$F(x) \equiv \int_a^x f(t)\ dt.$$

Then if f is continuous at $x \in (a,b)$,

$$F'(x) = f(x).$$

Proof: Let $x \in (a,b)$ be a point of continuity of f and let h be small enough that $x + h \in [a,b]$. Then by using (2.14),

$$h^{-1}(F(x+h) - F(x)) = h^{-1} \int_x^{x+h} f(t)\ dt.$$

[2]This theorem is why Newton and Liebnitz are credited with inventing calculus. The integral had been around for thousands of years and the derivative was by their time well known. However the connection between these two ideas had not been fully made although Newton's predecessor, Isaac Barrow had made some progress in this direction.

Also, using (2.12),

$$f(x) = h^{-1} \int_x^{x+h} f(x) \, dt.$$

Therefore, by (2.16),

$$\left| h^{-1} (F(x+h) - F(x)) - f(x) \right| = \left| h^{-1} \int_x^{x+h} (f(t) - f(x)) \, dt \right|$$

$$\leq \left| h^{-1} \int_x^{x+h} |f(t) - f(x)| \, dt \right|.$$

Let $\varepsilon > 0$ and let $\delta > 0$ be small enough that if $|t - x| < \delta$, then

$$|f(t) - f(x)| < \varepsilon.$$

Therefore, if $|h| < \delta$, the above inequality and (2.12) shows that

$$\left| h^{-1} (F(x+h) - F(x)) - f(x) \right| \leq |h|^{-1} \varepsilon |h| = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows

$$\lim_{h \to 0} h^{-1} (F(x+h) - F(x)) = f(x)$$

and this proves the theorem.

Note this gives existence for the initial value problem,

$$F'(x) = f(x), \; F(a) = 0$$

whenever f is Riemann integrable and continuous.[3]

The next theorem is also called the fundamental theorem of calculus.

Theorem 2.20 *Let $f \in R([a,b])$ and suppose there exists an antiderivative for f, G, such that*

$$G'(x) = f(x)$$

for every point of (a,b) and G is continuous on $[a,b]$. Then

$$\int_a^b f(x) \, dx = G(b) - G(a). \tag{2.17}$$

Proof: Let $P = \{x_0, \cdots, x_n\}$ be a partition satisfying

$$U(f,P) - L(f,P) < \varepsilon.$$

Then

$$\begin{aligned} G(b) - G(a) &= G(x_n) - G(x_0) \\ &= \sum_{i=1}^n G(x_i) - G(x_{i-1}). \end{aligned}$$

[3]Of course it was proved that if f is continuous on a closed interval, $[a,b]$, then $f \in R([a,b])$ but this is a hard theorem using the difficult result about uniform continuity.

By the mean value theorem,

$$\begin{aligned} G(b) - G(a) &= \sum_{i=1}^{n} G'(z_i)(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} f(z_i)\,\Delta x_i \end{aligned}$$

where z_i is some point in $[x_{i-1}, x_i]$. It follows, since the above sum lies between the upper and lower sums, that

$$G(b) - G(a) \in [L(f,P), U(f,P)],$$

and also

$$\int_a^b f(x)\,dx \in [L(f,P), U(f,P)].$$

Therefore,

$$\left| G(b) - G(a) - \int_a^b f(x)\,dx \right| < U(f,P) - L(f,P) < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, (2.17) holds. This proves the theorem.

The following notation is often used in this context. Suppose F is an antiderivative of f as just described with F continuous on $[a,b]$ and $F' = f$ on (a,b). Then

$$\int_a^b f(x)\,dx = F(b) - F(a) \equiv F(x)\,|_a^b.$$

Definition 2.21 *Let f be a bounded function defined on a closed interval $[a,b]$ and let $P \equiv \{x_0, \cdots, x_n\}$ be a partition of the interval. Suppose $z_i \in [x_{i-1}, x_i]$ is chosen. Then the sum*

$$\sum_{i=1}^{n} f(z_i)(x_i - x_{i-1})$$

is known as a Riemann sum. Also,

$$||P|| \equiv \max\{|x_i - x_{i-1}| : i = 1, \cdots, n\}.$$

Proposition 2.22 *Suppose $f \in R([a,b])$. Then there exists a partition, $P \equiv \{x_0, \cdots, x_n\}$ with the property that for any choice of $z_k \in [x_{k-1}, x_k]$,*

$$\left| \int_a^b f(x)\,dx - \sum_{k=1}^{n} f(z_k)(x_k - x_{k-1}) \right| < \varepsilon.$$

Proof: Choose P such that $U(f,P) - L(f,P) < \varepsilon$ and then both $\int_a^b f(x)\,dx$ and $\sum_{k=1}^{n} f(z_k)(x_k - x_{k-1})$ are contained in $[L(f,P), U(f,P)]$ and so the claimed inequality must hold. This proves the proposition.

It is significant because it gives a way of approximating the integral.

The definition of Riemann integrability given in this chapter is also called Darboux integrability and the integral defined as the unique number which lies between all upper sums and all lower sums which is given in this chapter is called the Darboux integral . The definition of the Riemann integral in terms of Riemann sums is given next.

Definition 2.23 *A bounded function, f defined on $[a,b]$ is said to be Riemann integrable if there exists a number, I with the property that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if*

$$P \equiv \{x_0, x_1, \cdots, x_n\}$$

is any partition having $||P|| < \delta$, and $z_i \in [x_{i-1}, x_i]$,

$$\left| I - \sum_{i=1}^{n} f(z_i)(x_i - x_{i-1}) \right| < \varepsilon.$$

The number $\int_a^b f(x)\, dx$ is defined as I.

Thus, there are two definitions of the Riemann integral. It turns out they are equivalent which is the following theorem of of Darboux.

Theorem 2.24 *A bounded function defined on $[a,b]$ is Riemann integrable in the sense of Definition 2.23 if and only if it is integrable in the sense of Darboux. Furthermore the two integrals coincide.*

The proof of this theorem is left for the exercises in Problems 10 - 12. It isn't essential that you understand this theorem so if it does not interest you, leave it out. Note that it implies that given a Riemann integrable function f in either sense, it can be approximated by Riemann sums whenever $||P||$ is sufficiently small. Both versions of the integral are obsolete but entirely adequate for most applications and as a point of departure for a more up to date and satisfactory integral. The reason for using the Darboux approach to the integral is that all the existence theorems are easier to prove in this context.

2.6 Exercises

1. Let $F(x) = \int_{x^2}^{x^3} \frac{t^5+7}{t^7+87t^6+1}\, dt$. Find $F'(x)$.
2. Let $F(x) = \int_2^x \frac{1}{1+t^4}\, dt$. Sketch a graph of F and explain why it looks the way it does.
3. Let a and b be positive numbers and consider the function,
$$F(x) = \int_0^{ax} \frac{1}{a^2+t^2}\, dt + \int_b^{a/x} \frac{1}{a^2+t^2}\, dt.$$
Show that F is a constant.
4. Solve the following initial value problem from ordinary differential equations which is to find a function y such that
$$y'(x) = \frac{x^7+1}{x^6+97x^5+7},\ y(10) = 5.$$
5. If $F, G \in \int f(x)\, dx$ for all $x \in \mathbb{R}$, show $F(x) = G(x) + C$ for some constant, C. Use this to give a different proof of the fundamental theorem of calculus which has for its conclusion $\int_a^b f(t)\, dt = G(b) - G(a)$ where $G'(x) = f(x)$.
6. Suppose f is Riemann integrable on $[a,b]$ and continuous. (In fact continuous implies Riemann integrable.) Show there exists $c \in (a,b)$ such that
$$f(c) = \frac{1}{b-a} \int_a^b f(x)\, dx.$$
Hint: You might consider the function $F(x) \equiv \int_a^x f(t)\, dt$ and use the mean value theorem for derivatives and the fundamental theorem of calculus.

7. Suppose f and g are continuous functions on $[a,b]$ and that $g(x) \neq 0$ on (a,b). Show there exists $c \in (a,b)$ such that

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

Hint: Define $F(x) \equiv \int_a^x f(t)\, g(t)\, dt$ and let $G(x) \equiv \int_a^x g(t)\, dt$. Then use the Cauchy mean value theorem on these two functions.

8. Consider the function

$$f(x) \equiv \begin{cases} \sin\left(\frac{1}{x}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}.$$

Is f Riemann integrable? Explain why or why not.

9. Prove the second part of Theorem 2.9 about decreasing functions.

10. Suppose f is a bounded function defined on $[a,b]$ and $|f(x)| < M$ for all $x \in [a,b]$. Now let Q be a partition having n points, $\{x_0^*, \cdots, x_n^*\}$ and let P be any other partition. Show that

$$|U(f,P) - L(f,P)| \leq 2Mn\, ||P|| + |U(f,Q) - L(f,Q)|.$$

Hint: Write the sum for $U(f,P) - L(f,P)$ and split this sum into two sums, the sum of terms for which $[x_{i-1}, x_i]$ contains at least one point of Q, and terms for which $[x_{i-1}, x_i]$ does not contain any points of Q. In the latter case, $[x_{i-1}, x_i]$ must be contained in some interval, $\left[x_{k-1}^*, x_k^*\right]$. Therefore, the sum of these terms should be no larger than $|U(f,Q) - L(f,Q)|$.

11. $\uparrow$ If $\varepsilon > 0$ is given and f is a Darboux integrable function defined on $[a,b]$, show there exists $\delta > 0$ such that whenever $||P|| < \delta$, then

$$|U(f,P) - L(f,P)| < \varepsilon.$$

12. $\uparrow$ Prove Theorem 2.24.

Multivariable Calculus

3.1 Continuous Functions

What was done earlier for scalar functions is generalized here to include the case of a vector valued function.

Definition 3.1 *A function* $\mathbf{f} : D(\mathbf{f}) \subseteq \mathbb{R}^p \to \mathbb{R}^q$ *is continuous at* $\mathbf{x} \in D(\mathbf{f})$ *if for each* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that whenever* $\mathbf{y} \in D(\mathbf{f})$ *and*

$$|\mathbf{y} - \mathbf{x}| < \delta$$

it follows that

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \varepsilon.$$

$\mathbf{f}$ *is continuous if it is continuous at every point of* $D(\mathbf{f})$.

Note the total similarity to the scalar valued case.

3.1.1 Sufficient Conditions For Continuity

The next theorem is a fundamental result which will allow us to worry less about the ε δ definition of continuity.

Theorem 3.2 *The following assertions are valid.*

1. *The function,* $a\mathbf{f} + b\mathbf{g}$ *is continuous at* $\mathbf{x}$ *whenever* $\mathbf{f}$, $\mathbf{g}$ *are continuous at* $\mathbf{x} \in D(\mathbf{f}) \cap D(\mathbf{g})$ *and* $a, b \in \mathbb{R}$.
2. *If* $\mathbf{f}$ *is continuous at* $\mathbf{x}$, $\mathbf{f}(\mathbf{x}) \in D(\mathbf{g}) \subseteq \mathbb{R}^p$, *and* $\mathbf{g}$ *is continuous at* $\mathbf{f}(\mathbf{x})$ *,then* $\mathbf{g} \circ \mathbf{f}$ *is continuous at* $\mathbf{x}$.
3. *If* $\mathbf{f} = (f_1, \cdots, f_q) : D(\mathbf{f}) \to \mathbb{R}^q$, *then* $\mathbf{f}$ *is continuous if and only if each* f_k *is a continuous real valued function.*
4. *The function* $f : \mathbb{R}^p \to \mathbb{R}$, *given by* $f(\mathbf{x}) = |\mathbf{x}|$ *is continuous.*

The proof of this theorem is in the last section of this chapter. Its conclusions are not surprising. For example the first claim says that $(a\mathbf{f} + b\mathbf{g})(\mathbf{y})$ is close to $(a\mathbf{f} + b\mathbf{g})(\mathbf{x})$ when $\mathbf{y}$ is close to $\mathbf{x}$ provided the same can be said about $\mathbf{f}$ and $\mathbf{g}$. For the second claim, if $\mathbf{y}$ is close to $\mathbf{x}$, $\mathbf{f}(\mathbf{x})$ is close to $\mathbf{f}(\mathbf{y})$ and so by continuity of $\mathbf{g}$ at $\mathbf{f}(\mathbf{x})$, $\mathbf{g}(\mathbf{f}(\mathbf{y}))$ is close to $\mathbf{g}(\mathbf{f}(\mathbf{x}))$. To see the third claim is likely, note that closeness in $\mathbb{R}^p$ is the same as closeness in each coordinate. The fourth claim is immediate from the triangle inequality.

For functions defined on $\mathbb{R}^n$, there is a notion of polynomial just as there is for functions defined on $\mathbb{R}$.

Definition 3.3 *Let* α *be an* n *dimensional multi-index. This means*

$$\alpha = (\alpha_1, \cdots, \alpha_n)$$

where each α_i is a natural number or zero. Also, let

$$|\alpha| \equiv \sum_{i=1}^{n} |\alpha_i|$$

The symbol, $\mathbf{x}^\alpha$*,means*

$$\mathbf{x}^\alpha \equiv x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_3^{\alpha_n}.$$

An n dimensional polynomial of degree m is a function of the form

$$p(\mathbf{x}) = \sum_{|\alpha| \leq m} d_\alpha \mathbf{x}^\alpha.$$

where the d_α are real numbers.

The above theorem implies that polynomials are all continuous.

3.2 Exercises

1. Let $\mathbf{f}(t) = (t, \sin t)$. Show f is continuous at every point t.
2. Suppose $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|$ where K is a constant. Show that $\mathbf{f}$ is everywhere continuous. Functions satisfying such an inequality are called Lipschitz functions.
3. Suppose $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|^\alpha$ where K is a constant and $\alpha \in (0,1)$. Show that $\mathbf{f}$ is everywhere continuous.
4. Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is given by $f(\mathbf{x}) = 3x_1x_2 + 2x_3^2$. Use Theorem 3.2 to verify that f is continuous. **Hint:** You should first verify that the function, $\pi_k : \mathbb{R}^3 \to \mathbb{R}$ given by $\pi_k(\mathbf{x}) = x_k$ is a continuous function.
5. Generalize the previous problem to the case where $f : \mathbb{R}^q \to \mathbb{R}$ is a polynomial.
6. State and prove a theorem which involves quotients of functions encountered in the previous problem.

3.3 Limits Of A Function

As in the case of scalar valued functions of one variable, a concept closely related to continuity is that of the limit of a function. The notion of limit of a function makes sense at points, $\mathbf{x}$, which are limit points of $D(\mathbf{f})$ and this concept is defined next.

Definition 3.4 *Let $A \subseteq \mathbb{R}^m$ be a set. A point, $\mathbf{x}$, is a limit point of A if $B(\mathbf{x}, r)$ contains infinitely many points of A for every $r > 0$.*

Definition 3.5 *Let $\mathbf{f} : D(\mathbf{f}) \subseteq \mathbb{R}^p \to \mathbb{R}^q$ be a function and let $\mathbf{x}$ be a limit point of $D(\mathbf{f})$. Then*

$$\lim_{\mathbf{y} \to \mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{L}$$

if and only if the following condition holds. For all $\varepsilon > 0$ there exists $\delta > 0$ such that if

$$0 < |\mathbf{y} - \mathbf{x}| < \delta, \text{ and } \mathbf{y} \in D(\mathbf{f})$$

then,

$$|\mathbf{L} - \mathbf{f}(\mathbf{y})| < \varepsilon.$$

Theorem 3.6 *If* $\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{L}$ *and* $\lim_{y\to x} \mathbf{f}(\mathbf{y}) = \mathbf{L}_1$*, then* $\mathbf{L} = \mathbf{L}_1$.

Proof: Let $\varepsilon > 0$ be given. There exists $\delta > 0$ such that if $0 < |\mathbf{y}-\mathbf{x}| < \delta$ and $\mathbf{y} \in D(\mathbf{f})$, then

$$|\mathbf{f}(\mathbf{y}) - \mathbf{L}| < \varepsilon, \ |\mathbf{f}(\mathbf{y}) - \mathbf{L}_1| < \varepsilon.$$

Pick such a $\mathbf{y}$. There exists one because $\mathbf{x}$ is a limit point of $D(\mathbf{f})$. Then

$$|\mathbf{L} - \mathbf{L}_1| \leq |\mathbf{L} - \mathbf{f}(\mathbf{y})| + |\mathbf{f}(\mathbf{y}) - \mathbf{L}_1| < \varepsilon + \varepsilon = 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows $\mathbf{L} = \mathbf{L}_1$.

As in the case of functions of one variable, one can define what it means for $\lim_{\mathbf{y}\to\mathbf{x}} f(\mathbf{x}) = \pm\infty$.

Definition 3.7 *If* $f(\mathbf{x}) \in \mathbb{R}$, $\lim_{\mathbf{y}\to\mathbf{x}} f(\mathbf{x}) = \infty$ *if for every number* l*, there exists* $\delta > 0$ *such that whenever* $|\mathbf{y}-\mathbf{x}| < \delta$ *and* $\mathbf{y} \in D(\mathbf{f})$*, then* $f(\mathbf{x}) > l$.

The following theorem is just like the one variable version presented earlier.

Theorem 3.8 *Suppose* $\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{L}$ *and* $\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{g}(\mathbf{y}) = \mathbf{K}$ *where* $\mathbf{K}, \mathbf{L} \in \mathbb{R}^q$*. Then if* $a, b \in \mathbb{R}$,

$$\lim_{\mathbf{y}\to\mathbf{x}} (a\mathbf{f}(\mathbf{y}) + b\mathbf{g}(\mathbf{y})) = a\mathbf{L} + b\mathbf{K}, \tag{3.1}$$

$$\lim_{y\to x} \mathbf{f}\cdot\mathbf{g}(y) = \mathbf{L}\mathbf{K} \tag{3.2}$$

and if g *is scalar valued with* $\lim_{\mathbf{y}\to\mathbf{x}} g(\mathbf{y}) = K \neq 0$,

$$\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y})\, g(\mathbf{y}) = \mathbf{L}K. \tag{3.3}$$

Also, if $\mathbf{h}$ *is a continuous function defined near* $\mathbf{L}$*, then*

$$\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{h}\circ\mathbf{f}(\mathbf{y}) = \mathbf{h}(\mathbf{L}). \tag{3.4}$$

Suppose $\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{L}$*. If* $|\mathbf{f}(\mathbf{y}) - \mathbf{b}| \leq r$ *for all* $\mathbf{y}$ *sufficiently close to* $\mathbf{x}$*, then* $|\mathbf{L} - \mathbf{b}| \leq r$ *also.*

Proof: The proof of (3.1) is left for you. It is like a corresponding theorem for continuous functions. Now (3.2)is to be verified. Let $\varepsilon > 0$ be given. Then by the triangle inequality,

$$\begin{aligned} |\mathbf{f}\cdot\mathbf{g}(\mathbf{y}) - \mathbf{L}\cdot\mathbf{K}| &\leq |\mathbf{f}\mathbf{g}(\mathbf{y}) - \mathbf{f}(\mathbf{y})\cdot\mathbf{K}| + |\mathbf{f}(\mathbf{y})\cdot\mathbf{K} - \mathbf{L}\cdot\mathbf{K}| \\ &\leq |\mathbf{f}(\mathbf{y})|\,|\mathbf{g}(\mathbf{y}) - \mathbf{K}| + |\mathbf{K}|\,|\mathbf{f}(\mathbf{y}) - \mathbf{L}|. \end{aligned}$$

There exists δ_1 such that if $0 < |\mathbf{y}-\mathbf{x}| < \delta_1$ and $\mathbf{y} \in D(\mathbf{f})$, then

$$|\mathbf{f}(\mathbf{y}) - \mathbf{L}| < 1,$$

and so for such $\mathbf{y}$, the triangle inequality implies, $|\mathbf{f}(\mathbf{y})| < 1 + |\mathbf{L}|$. Therefore, for $0 < |\mathbf{y}-\mathbf{x}| < \delta_1$,

$$|\mathbf{f}\cdot\mathbf{g}(\mathbf{y}) - \mathbf{L}\cdot\mathbf{K}| \leq (1 + |\mathbf{K}| + |\mathbf{L}|)\,[|\mathbf{g}(\mathbf{y}) - \mathbf{K}| + |\mathbf{f}(\mathbf{y}) - \mathbf{L}|]. \tag{3.5}$$

Now let $0 < \delta_2$ be such that if $\mathbf{y} \in D(\mathbf{f})$ and $0 < |\mathbf{x}-\mathbf{y}| < \delta_2$,

$$|\mathbf{f}(\mathbf{y}) - \mathbf{L}| < \frac{\varepsilon}{2(1 + |\mathbf{K}| + |\mathbf{L}|)}, \ |\mathbf{g}(\mathbf{y}) - \mathbf{K}| < \frac{\varepsilon}{2(1 + |\mathbf{K}| + |\mathbf{L}|)}.$$

Then letting $0<\delta \leq \min (\delta_1, \delta_2)$, it follows from (3.5) that

$$|\mathbf{f} \cdot \mathbf{g}(\mathbf{y})-\mathbf{L} \cdot \mathbf{K}|<\varepsilon$$

and this proves (3.2).

The proof of (3.3) is left to you.

Consider (3.4). Since $\mathbf{h}$ is continuous near $\mathbf{L}$, it follows that for $\varepsilon>0$ given, there exists $\eta>0$ such that if $|\mathbf{y}-\mathbf{L}|<\eta$, then

$$|\mathbf{h}(\mathbf{y})-\mathbf{h}(\mathbf{L})|<\varepsilon$$

Now since $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})=\mathbf{L}$, there exists $\delta>0$ such that if $0<|\mathbf{y}-\mathbf{x}|<\delta$, then

$$|\mathbf{f}(\mathbf{y})-\mathbf{L}|<\eta .$$

Therefore, if $0<|\mathbf{y}-\mathbf{x}|<\delta$,

$$|\mathbf{h}(\mathbf{f}(\mathbf{y}))-\mathbf{h}(\mathbf{L})|<\varepsilon .$$

It only remains to verify the last assertion. Assume $|\mathbf{f}(\mathbf{y})-\mathbf{b}| \leq r$. It is required to show that $|\mathbf{L}-\mathbf{b}| \leq r$. If this is not true, then $|\mathbf{L}-\mathbf{b}|>r$. Consider $B(\mathbf{L},|\mathbf{L}-\mathbf{b}|-r)$. Since $\mathbf{L}$ is the limit of $\mathbf{f}$, it follows $\mathbf{f}(\mathbf{y}) \in B(\mathbf{L},|\mathbf{L}-\mathbf{b}|-r)$ whenever $\mathbf{y} \in D(\mathbf{f})$ is close enough to $\mathbf{x}$. Thus, by the triangle inequality,

$$|\mathbf{f}(\mathbf{y})-\mathbf{L}|<|\mathbf{L}-\mathbf{b}|-r$$

and so

$$\begin{aligned} r & <|\mathbf{L}-\mathbf{b}|-|\mathbf{f}(\mathbf{y})-\mathbf{L}| \leq||\mathbf{b}-\mathbf{L}|-|\mathbf{f}(\mathbf{y})-\mathbf{L}|| \\ & \leq|\mathbf{b}-\mathbf{f}(\mathbf{y})|, \end{aligned}$$

a contradiction to the assumption that $|\mathbf{b}-\mathbf{f}(\mathbf{y})| \leq r$.

Theorem 3.9 *For $\mathbf{f}: D(\mathbf{f}) \rightarrow \mathbb{R}^q$ and $\mathbf{x} \in D(\mathbf{f})$ a limit point of $D(\mathbf{f})$, $\mathbf{f}$ is continuous at $\mathbf{x}$ if and only if*

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})=\mathbf{f}(\mathbf{x}) .$$

Proof: First suppose $\mathbf{f}$ is continuous at $\mathbf{x}$ a limit point of $D(\mathbf{f})$. Then for every $\varepsilon>0$ there exists $\delta>0$ such that if $|\mathbf{y}-\mathbf{x}|<\delta$ and $\mathbf{y} \in D(\mathbf{f})$, then $|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})|<\varepsilon$. In particular, this holds if $0<|\mathbf{x}-\mathbf{y}|<\delta$ and this is just the definition of the limit. Hence $\mathbf{f}(\mathbf{x})=\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$.

Next suppose $\mathbf{x}$ is a limit point of $D(\mathbf{f})$ and $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})=\mathbf{f}(\mathbf{x})$. This means that if $\varepsilon>0$ there exists $\delta>0$ such that for $0<|\mathbf{x}-\mathbf{y}|<\delta$ and $\mathbf{y} \in D(\mathbf{f})$, it follows $|\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})|<\varepsilon$. However, if $\mathbf{y}=\mathbf{x}$, then $|\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})|=|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x})|=0$ and so whenever $\mathbf{y} \in D(\mathbf{f})$ and $|\mathbf{x}-\mathbf{y}|<\delta$, it follows $|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})|<\varepsilon$, showing $\mathbf{f}$ is continuous at $\mathbf{x}$.

The following theorem is important.

Theorem 3.10 *Suppose $\mathbf{f}: D(\mathbf{f}) \rightarrow \mathbb{R}^q$. Then for $\mathbf{x}$ a limit point of $D(\mathbf{f})$,*

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})=\mathbf{L} \tag{3.6}$$

if and only if

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} f_k(\mathbf{y})=L_k \tag{3.7}$$

where $\mathbf{f}(\mathbf{y}) \equiv(f_1(\mathbf{y}), \cdots, f_p(\mathbf{y}))$ and $\mathbf{L} \equiv(L_1, \cdots, L_p)$.

Proof: Suppose (3.6). Then letting $\varepsilon > 0$ be given there exists $\delta > 0$ such that if $0 < |\mathbf{y} - \mathbf{x}| < \delta$, it follows

$$|f_k(\mathbf{y}) - L_k| \leq |\mathbf{f}(\mathbf{y}) - \mathbf{L}| < \varepsilon$$

which verifies (3.7).

Now suppose (3.7) holds. Then letting $\varepsilon > 0$ be given, there exists δ_k such that if $0 < |\mathbf{y} - \mathbf{x}| < \delta_k$, then

$$|f_k(\mathbf{y}) - L_k| < \frac{\varepsilon}{\sqrt{p}}.$$

Let $0 < \delta < \min(\delta_1, \cdots, \delta_p)$. Then if $0 < |\mathbf{y} - \mathbf{x}| < \delta$, it follows

$$\begin{aligned} |\mathbf{f}(\mathbf{y}) - \mathbf{L}| &= \left(\sum_{k=1}^{p} |f_k(\mathbf{y}) - L_k|^2 \right)^{1/2} \\ &< \left(\sum_{k=1}^{p} \frac{\varepsilon^2}{p} \right)^{1/2} = \varepsilon. \end{aligned}$$

This proves the theorem.

This theorem shows it suffices to consider the components of a vector valued function when computing the limit.

Example 3.11 *Find* $\lim_{(x,y)\to(3,1)} \left(\frac{x^2-9}{x-3}, y \right)$.

It is clear that $\lim_{(x,y)\to(3,1)} \frac{x^2-9}{x-3} = 6$ and $\lim_{(x,y)\to(3,1)} y = 1$. Therefore, this limit equals $(6, 1)$.

Example 3.12 *Find* $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$.

First of all observe the domain of the function is $\mathbb{R}^2 \setminus \{(0,0)\}$, every point in $\mathbb{R}^2$ except the origin. Therefore, $(0,0)$ is a limit point of the domain of the function so it might make sense to take a limit. However, just as in the case of a function of one variable, the limit may not exist. In fact, this is the case here. To see this, take points on the line $y = 0$. At these points, the value of the function equals 0. Now consider points on the line $y = x$ where the value of the function equals $1/2$. Since arbitrarily close to $(0,0)$ there are points where the function equals $1/2$ and points where the function has the value 0, it follows there can be no limit. Just take $\varepsilon = 1/10$ for example. You can't be within $1/10$ of $1/2$ and also within $1/10$ of 0 at the same time.

Note it is necessary to rely on the definition of the limit much more than in the case of a function of one variable and it is the case there are no easy ways to do limit problems for functions of more than one variable. It is what it is and you will not deal with these concepts without agony.

3.4 Exercises

1. Find the following limits if possible
   (a) $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$
   (b) $\lim_{(x,y)\to(0,0)} \frac{x(x^2-y^2)}{(x^2+y^2)}$
   (c) $\lim_{(x,y)\to(0,0)} \frac{(x^2-y^4)^2}{(x^2+y^4)^2}$ **Hint:** Consider along $y = 0$ and along $x = y^2$.
   (d) $\lim_{(x,y)\to(0,0)} x \sin\left(\frac{1}{x^2+y^2} \right)$
   (e) $\lim_{(x,y)\to(1,2)} \frac{-2yx^2+8yx+34y+3y^3-18y^2+6x^2-13x-20-xy^2-x^3}{-y^2+4y-5-x^2+2x}$. **Hint:** It might help to write this in terms of the variables $(s,t) = (x-1, y-2)$.
2. In the definition of limit, why must $\mathbf{x}$ be a limit point of $D(\mathbf{f})$? **Hint:** If $\mathbf{x}$ were not a limit point of $D(\mathbf{f})$, show there exists $\delta > 0$ such that $B(\mathbf{x}, \delta)$ contains no points of $D(\mathbf{f})$ other than possibly $\mathbf{x}$ itself. Argue that 33.3 is a limit and that so is 22 and 7 and 11. In other words the concept is totally worthless.

3.5 The Limit Of A Sequence

As in the case of real numbers, one can consider the limit of a sequence of points in $\mathbb{R}^p$.

Definition 3.13 *A sequence* $\{\mathbf{a}_n\}_{n=1}^{\infty}$ *converges to* $\mathbf{a}$, *and write*

$$\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a} \text{ or } \mathbf{a}_n \to \mathbf{a}$$

if and only if for every $\varepsilon > 0$ *there exists* n_ε *such that whenever* $n \geq n_\varepsilon$,

$$|\mathbf{a}_n - \mathbf{a}| < \varepsilon.$$

In words the definition says that given any measure of closeness, ε, the terms of the sequence are eventually all this close to $\mathbf{a}$. There is absolutely no difference between this and the definition for sequences of numbers other than here bold face is used to indicate $\mathbf{a}_n$ and $\mathbf{a}$ are points in $\mathbb{R}^p$.

Theorem 3.14 *If* $\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a}$ *and* $\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a}_1$ *then* $\mathbf{a}_1 = \mathbf{a}$.

Proof: Suppose $\mathbf{a}_1 \neq \mathbf{a}$. Then let $0 < \varepsilon < |\mathbf{a}_1 - \mathbf{a}|/2$ in the definition of the limit. It follows there exists n_ε such that if $n \geq n_\varepsilon$, then $|\mathbf{a}_n - \mathbf{a}| < \varepsilon$ and $|\mathbf{a}_n - \mathbf{a}_1| < \varepsilon$. Therefore, for such n,

$$\begin{aligned} |\mathbf{a}_1 - \mathbf{a}| &\leq |\mathbf{a}_1 - \mathbf{a}_n| + |\mathbf{a}_n - \mathbf{a}| \\ &< \varepsilon + \varepsilon < |\mathbf{a}_1 - \mathbf{a}|/2 + |\mathbf{a}_1 - \mathbf{a}|/2 = |\mathbf{a}_1 - \mathbf{a}|, \end{aligned}$$

a contradiction.

As in the case of a vector valued function, it suffices to consider the components. This is the content of the next theorem.

Theorem 3.15 *Let* $\mathbf{a}_n = \left(a_1^n, \cdots, a_p^n\right) \in \mathbb{R}^p$. *Then* $\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a} \equiv (a_1, \cdots, a_p)$ *if and only if for each* $k = 1, \cdots, p$,

$$\lim_{n\to\infty} a_k^n = a_k. \tag{3.8}$$

Proof: First suppose $\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a}$. Then given $\varepsilon > 0$ there exists n_ε such that if $n > n_\varepsilon$, then

$$|a_k^n - a_k| \leq |\mathbf{a}_n - \mathbf{a}| < \varepsilon$$

which establishes (3.8).

Now suppose (3.8) holds for each k. Then letting $\varepsilon > 0$ be given there exist n_k such that if $n > n_k$,

$$|a_k^n - a_k| < \varepsilon/\sqrt{p}.$$

Therefore, letting $n_\varepsilon > \max(n_1, \cdots, n_p)$, it follows that for $n > n_\varepsilon$,

$$|\mathbf{a}_n - \mathbf{a}| = \left(\sum_{k=1}^{n} |a_k^n - a_k|^2\right)^{1/2} < \left(\sum_{k=1}^{n} \frac{\varepsilon^2}{p}\right)^{1/2} = \varepsilon,$$

showing that $\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a}$. This proves the theorem.

Example 3.16 *Let* $\mathbf{a}_n = \left(\frac{1}{n^2+1}, \frac{1}{n}\sin(n), \frac{n^2+3}{3n^2+5n}\right)$.

It suffices to consider the limits of the components according to the following theorem. Thus the limit is $(0, 0, 1/3)$.

Theorem 3.17 *Suppose* $\{\mathbf{a}_n\}$ *and* $\{\mathbf{b}_n\}$ *are sequences and that*

$$\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a} \text{ and } \lim_{n\to\infty} \mathbf{b}_n = \mathbf{b}.$$

Also suppose x *and* y *are real numbers. Then*

$$\lim_{n\to\infty} x\mathbf{a}_n + y\mathbf{b}_n = x\mathbf{a} + y\mathbf{b} \tag{3.9}$$

$$\lim_{n\to\infty} \mathbf{a}_n \cdot \mathbf{b}_n = \mathbf{a} \cdot \mathbf{b} \tag{3.10}$$

If $b_n \in \mathbb{R}$, *then*

$$\mathbf{a}_n b_n \to \mathbf{a} b.$$

Proof: The first of these claims is left for you to do. To do the second, let $\varepsilon > 0$ be given and choose n_1 such that if $n \geq n_1$ then

$$|\mathbf{a}_n - \mathbf{a}| < 1.$$

Then for such n, the triangle inequality and Cauchy Schwarz inequality imply

$$\begin{aligned} |\mathbf{a}_n \cdot \mathbf{b}_n - \mathbf{a} \cdot \mathbf{b}| &\leq |\mathbf{a}_n \cdot \mathbf{b}_n - \mathbf{a}_n \cdot \mathbf{b}| + |\mathbf{a}_n \cdot \mathbf{b} - \mathbf{a} \cdot \mathbf{b}| \\ &\leq |\mathbf{a}_n| \, |\mathbf{b}_n - \mathbf{b}| + |\mathbf{b}| \, |\mathbf{a}_n - \mathbf{a}| \\ &\leq (|\mathbf{a}| + 1) \, |\mathbf{b}_n - \mathbf{b}| + |\mathbf{b}| \, |\mathbf{a}_n - \mathbf{a}| \,. \end{aligned}$$

Now let n_2 be large enough that for $n \geq n_2$,

$$|\mathbf{b}_n - \mathbf{b}| < \frac{\varepsilon}{2(|\mathbf{a}| + 1)}, \text{ and } |\mathbf{a}_n - \mathbf{a}| < \frac{\varepsilon}{2(|\mathbf{b}| + 1)}.$$

Such a number exists because of the definition of limit. Therefore, let

$$n_\varepsilon > \max(n_1, n_2).$$

For $n \geq n_\varepsilon$,

$$\begin{aligned} |\mathbf{a}_n \cdot \mathbf{b}_n - \mathbf{a} \cdot \mathbf{b}| &\leq (|\mathbf{a}| + 1) \, |\mathbf{b}_n - \mathbf{b}| + |\mathbf{b}| \, |\mathbf{a}_n - \mathbf{a}| \\ &< (|\mathbf{a}| + 1) \frac{\varepsilon}{2(|\mathbf{a}| + 1)} + |\mathbf{b}| \frac{\varepsilon}{2(|\mathbf{b}| + 1)} \leq \varepsilon. \end{aligned}$$

This proves (3.9). The proof of (3.10) is entirely similar and is left for you.

3.5.1 Sequences And Completeness

Recall the definition of a Cauchy sequence.

Definition 3.18 $\{\mathbf{a}_n\}$ *is a Cauchy sequence if for all* $\varepsilon > 0$, *there exists* n_ε *such that whenever* $n, m \geq n_\varepsilon$,

$$|\mathbf{a}_n - \mathbf{a}_m| < \varepsilon.$$

A sequence is Cauchy means the terms are "bunching up to each other" as m, n get large.

Theorem 3.19 *Let* $\{\mathbf{a}_n\}_{n=1}^{\infty}$ *be a Cauchy sequence in* $\mathbb{R}^p$. *Then there exists* $\mathbf{a} \in \mathbb{R}^p$ *such that* $\mathbf{a}_n \to \mathbf{a}$.

Proof: Let $\mathbf{a}_n = \left(a_1^n, \cdots, a_p^n\right)$. Then

$$|a_k^n - a_k^m| \leq |\mathbf{a}_n - \mathbf{a}_m|$$

which shows for each $k = 1, \cdots, p$, it follows $\{a_k^n\}_{n=1}^{\infty}$ is a Cauchy sequence. By completeness of $\mathbb{R}$, it follows there exists a_k such that $\lim_{n\to\infty} a_k^n = a_k$. Letting $\mathbf{a} = (a_1, \cdots, a_p)$, it follows from Theorem 3.15 that

$$\lim_{n\to\infty} \mathbf{a}_n = \mathbf{a}.$$

This proves the theorem.

Theorem 3.20 *The set of terms in a Cauchy sequence in $\mathbb{R}^p$ is bounded in the sense that for all n, $|\mathbf{a}_n| < M$ for some $M < \infty$.*

Proof: Let $\varepsilon = 1$ in the definition of a Cauchy sequence and let $n > n_1$. Then from the definition,

$$|\mathbf{a}_n - \mathbf{a}_{n_1}| < 1.$$

It follows that for all $n > n_1$,

$$|\mathbf{a}_n| < 1 + |\mathbf{a}_{n_1}|.$$

Therefore, for all n,

$$|\mathbf{a}_n| \leq 1 + |\mathbf{a}_{n_1}| + \sum_{k=1}^{n_1} |\mathbf{a}_k|.$$

This proves the theorem.

Theorem 3.21 *If a sequence $\{\mathbf{a}_n\}$ in $\mathbb{R}^p$ converges, then the sequence is a Cauchy sequence.*

Proof: Let $\varepsilon > 0$ be given and suppose $\mathbf{a}_n \to \mathbf{a}$. Then from the definition of convergence, there exists n_ε such that if $n > n_\varepsilon$, it follows that

$$|\mathbf{a}_n - \mathbf{a}| < \frac{\varepsilon}{2}$$

Therefore, if $m, n \geq n_\varepsilon + 1$, it follows that

$$|\mathbf{a}_n - \mathbf{a}_m| \leq |\mathbf{a}_n - \mathbf{a}| + |\mathbf{a} - \mathbf{a}_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

showing that, since $\varepsilon > 0$ is arbitrary, $\{\mathbf{a}_n\}$ is a Cauchy sequence.

3.5.2 Continuity And The Limit Of A Sequence

Just as in the case of a function of one variable, there is a very useful way of thinking of continuity in terms of limits of sequences found in the following theorem. In words, it says a function is continuous if it takes convergent sequences to convergent sequences whenever possible.

Theorem 3.22 *A function $\mathbf{f} : D(\mathbf{f}) \to \mathbb{R}^q$ is continuous at $\mathbf{x} \in D(\mathbf{f})$ if and only if, whenever $\mathbf{x}_n \to \mathbf{x}$ with $\mathbf{x}_n \in D(\mathbf{f})$, it follows $\mathbf{f}(\mathbf{x}_n) \to \mathbf{f}(\mathbf{x})$.*

Proof: Suppose first that $\mathbf{f}$ is continuous at $\mathbf{x}$ and let $\mathbf{x}_n \rightarrow \mathbf{x}$. Let $\varepsilon > 0$ be given. By continuity, there exists $\delta > 0$ such that if $|\mathbf{y} - \mathbf{x}| < \delta$, then $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \varepsilon$. However, there exists n_δ such that if $n \geq n_\delta$, then $|\mathbf{x}_n - \mathbf{x}| < \delta$ and so for all n this large,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_n)| < \varepsilon$$

which shows $\mathbf{f}(\mathbf{x}_n) \rightarrow \mathbf{f}(\mathbf{x})$.

Now suppose the condition about taking convergent sequences to convergent sequences holds at $\mathbf{x}$. Suppose $\mathbf{f}$ fails to be continuous at $\mathbf{x}$. Then there exists $\varepsilon > 0$ and $\mathbf{x}_n \in D(f)$ such that $|\mathbf{x} - \mathbf{x}_n| < \frac{1}{n}$, yet

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_n)| \geq \varepsilon.$$

But this is clearly a contradiction because, although $\mathbf{x}_n \rightarrow \mathbf{x}$, $\mathbf{f}(\mathbf{x}_n)$ fails to converge to $\mathbf{f}(\mathbf{x})$. It follows $\mathbf{f}$ must be continuous after all. This proves the theorem.

3.6 Properties Of Continuous Functions

Functions of p variables have many of the same properties as functions of one variable. First there is a version of the extreme value theorem generalizing the one dimensional case.

Theorem 3.23 *Let C be closed and bounded and let $f : C \rightarrow \mathbb{R}$ be continuous. Then f achieves its maximum and its minimum on C. This means there exist, $\mathbf{x}_1, \mathbf{x}_2 \in C$ such that for all $\mathbf{x} \in C$,*

$$f(\mathbf{x}_1) \leq f(\mathbf{x}) \leq f(\mathbf{x}_2).$$

There is also the long technical theorem about sums and products of continuous functions. These theorems are proved in the next section.

Theorem 3.24 *The following assertions are valid*

1. *The function, $a\mathbf{f} + b\mathbf{g}$ is continuous at $\mathbf{x}$ when $\mathbf{f}$, $\mathbf{g}$ are continuous at $\mathbf{x} \in D(\mathbf{f}) \cap D(\mathbf{g})$ and $a, b \in \mathbb{R}$.*

2. *If and f and g are each real valued functions continuous at $\mathbf{x}$, then fg is continuous at $\mathbf{x}$. If, in addition to this, $g(\mathbf{x}) \neq 0$, then f/g is continuous at $\mathbf{x}$.*

3. *If $\mathbf{f}$ is continuous at $\mathbf{x}$, $\mathbf{f}(\mathbf{x}) \in D(\mathbf{g}) \subseteq \mathbb{R}^p$, and $\mathbf{g}$ is continuous at $\mathbf{f}(\mathbf{x})$,then $\mathbf{g} \circ \mathbf{f}$ is continuous at $\mathbf{x}$.*

4. *If $\mathbf{f} = (f_1, \cdots, f_q) : D(\mathbf{f}) \rightarrow \mathbb{R}^q$, then $\mathbf{f}$ is continuous if and only if each f_k is a continuous real valued function.*

5. *The function $f : \mathbb{R}^p \rightarrow \mathbb{R}$, given by $f(\mathbf{x}) = |\mathbf{x}|$ is continuous.*

3.7 Exercises

1. $\mathbf{f} : D \subseteq \mathbb{R}^p \rightarrow \mathbb{R}^q$ is Lipschitz continuous or just Lipschitz for short if there exists a constant, K such that

   $$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K |\mathbf{x} - \mathbf{y}|$$

   for all $\mathbf{x}, \mathbf{y} \in D$. Show every Lipschitz function is uniformly continuous which means that given $\varepsilon > 0$ there exists $\delta > 0$ independent of $\mathbf{x}$ such that if $|\mathbf{x} - \mathbf{y}| < \delta$, then $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \varepsilon$.

2. If $\mathbf{f}$ is uniformly continuous, does it follow that $|\mathbf{f}|$ is also uniformly continuous? If $|\mathbf{f}|$ is uniformly continuous does it follow that $\mathbf{f}$ is uniformly continuous? Answer the same questions with "uniformly continuous" replaced with "continuous". Explain why.

3.8 Proofs Of Theorems

This section contains the proofs of the theorems which were just stated without proof.

Theorem 3.25 *The following assertions are valid*

1. *The function,* $a\mathbf{f}+b\mathbf{g}$ *is continuous at* $\mathbf{x}$ *when* $\mathbf{f}$, $\mathbf{g}$ *are continuous at* $\mathbf{x}\in D(\mathbf{f})\cap D(\mathbf{g})$ *and* $a,b\in\mathbb{R}$.
2. *If and* f *and* g *are each real valued functions continuous at* $\mathbf{x}$, *then* fg *is continuous at* $\mathbf{x}$. *If, in addition to this,* $g(\mathbf{x})\neq 0$, *then* f/g *is continuous at* $\mathbf{x}$.
3. *If* $\mathbf{f}$ *is continuous at* $\mathbf{x}$, $\mathbf{f}(\mathbf{x})\in D(\mathbf{g})\subseteq\mathbb{R}^p$, *and* $\mathbf{g}$ *is continuous at* $\mathbf{f}(\mathbf{x})$ *,then* $\mathbf{g}\circ\mathbf{f}$ *is continuous at* $\mathbf{x}$.
4. *If* $\mathbf{f}=(f_1,\cdots,f_q):D(\mathbf{f})\to\mathbb{R}^q$, *then* $\mathbf{f}$ *is continuous if and only if each* f_k *is a continuous real valued function.*
5. *The function* $f:\mathbb{R}^p\to\mathbb{R}$, *given by* $f(\mathbf{x})=|\mathbf{x}|$ *is continuous.*

Proof: Begin with 1.) Let $\varepsilon>0$ be given. By assumption, there exist $\delta_1>0$ such that whenever $|\mathbf{x}-\mathbf{y}|<\delta_1$, it follows $|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})|<\frac{\varepsilon}{2(|a|+|b|+1)}$ and there exists $\delta_2>0$ such that whenever $|\mathbf{x}-\mathbf{y}|<\delta_2$, it follows that $|\mathbf{g}(\mathbf{x})-\mathbf{g}(\mathbf{y})|<\frac{\varepsilon}{2(|a|+|b|+1)}$. Then let $0<\delta\leq\min(\delta_1,\delta_2)$. If $|\mathbf{x}-\mathbf{y}|<\delta$, then everything happens at once. Therefore, using the triangle inequality

$$|a\mathbf{f}(\mathbf{x})+b\mathbf{f}(\mathbf{x})-(a\mathbf{g}(\mathbf{y})+b\mathbf{g}(\mathbf{y}))|$$

$$\begin{aligned}&\leq|a||\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})|+|b||\mathbf{g}(\mathbf{x})-\mathbf{g}(\mathbf{y})|\\&<|a|\left(\frac{\varepsilon}{2(|a|+|b|+1)}\right)+|b|\left(\frac{\varepsilon}{2(|a|+|b|+1)}\right)<\varepsilon.\end{aligned}$$

Now begin on 2.) There exists $\delta_1>0$ such that if $|\mathbf{y}-\mathbf{x}|<\delta_1$, then $|f(\mathbf{x})-f(\mathbf{y})|<1$. Therefore, for such $\mathbf{y}$,

$$|f(\mathbf{y})|<1+|f(\mathbf{x})|.$$

It follows that for such $\mathbf{y}$,

$$|fg(\mathbf{x})-fg(\mathbf{y})|\leq|f(\mathbf{x})g(\mathbf{x})-g(\mathbf{x})f(\mathbf{y})|+|g(\mathbf{x})f(\mathbf{y})-f(\mathbf{y})g(\mathbf{y})|$$

$$\begin{aligned}&\leq|g(\mathbf{x})||f(\mathbf{x})-f(\mathbf{y})|+|f(\mathbf{y})||g(\mathbf{x})-g(\mathbf{y})|\\&\leq(1+|g(\mathbf{x})|+|f(\mathbf{y})|)[|g(\mathbf{x})-g(\mathbf{y})|+|f(\mathbf{x})-f(\mathbf{y})|].\end{aligned}$$

Now let $\varepsilon>0$ be given. There exists δ_2 such that if $|\mathbf{x}-\mathbf{y}|<\delta_2$, then

$$|g(\mathbf{x})-g(\mathbf{y})|<\frac{\varepsilon}{2(1+|g(\mathbf{x})|+|f(\mathbf{y})|)},$$

and there exists δ_3 such that if $|\mathbf{x}-\mathbf{y}|<\delta_3$, then

$$|f(\mathbf{x})-f(\mathbf{y})|<\frac{\varepsilon}{2(1+|g(\mathbf{x})|+|f(\mathbf{y})|)}$$

Now let $0<\delta\leq\min(\delta_1,\delta_2,\delta_3)$. Then if $|\mathbf{x}-\mathbf{y}|<\delta$, all the above hold at once and

$$|fg(\mathbf{x})-fg(\mathbf{y})|\leq$$

$$\begin{aligned}&(1+|g(\mathbf{x})|+|f(\mathbf{y})|)\,[|g(\mathbf{x})-g(\mathbf{y})|+|f(\mathbf{x})-f(\mathbf{y})|]\\&<(1+|g(\mathbf{x})|+|f(\mathbf{y})|)\left(\frac{\varepsilon}{2(1+|g(\mathbf{x})|+|f(\mathbf{y})|)}+\frac{\varepsilon}{2(1+|g(\mathbf{x})|+|f(\mathbf{y})|)}\right)=\varepsilon.\end{aligned}$$

This proves the first part of 2.) To obtain the second part, let δ_1 be as described above and let $\delta_0>0$ be such that for $|\mathbf{x}-\mathbf{y}|<\delta_0$,

$$|g(\mathbf{x})-g(\mathbf{y})|<|g(\mathbf{x})|/2$$

and so by the triangle inequality,

$$-|g(\mathbf{x})|/2\le|g(\mathbf{y})|-|g(\mathbf{x})|\le|g(\mathbf{x})|/2$$

which implies $|g(\mathbf{y})|\ge|g(\mathbf{x})|/2$, and $|g(\mathbf{y})|<3|g(\mathbf{x})|/2.$

Then if $|\mathbf{x}-\mathbf{y}|<\min(\delta_0,\delta_1)$,

$$\begin{aligned}\left|\frac{f(\mathbf{x})}{g(\mathbf{x})}-\frac{f(\mathbf{y})}{g(\mathbf{y})}\right|&=\left|\frac{f(\mathbf{x})g(\mathbf{y})-f(\mathbf{y})g(\mathbf{x})}{g(\mathbf{x})g(\mathbf{y})}\right|\\&\le\frac{|f(\mathbf{x})g(\mathbf{y})-f(\mathbf{y})g(\mathbf{x})|}{\left(\frac{|g(\mathbf{x})|^2}{2}\right)}\\&=\frac{2|f(\mathbf{x})g(\mathbf{y})-f(\mathbf{y})g(\mathbf{x})|}{|g(\mathbf{x})|^2}\end{aligned}$$

$$\begin{aligned}&\le\frac{2}{|g(\mathbf{x})|^2}[|f(\mathbf{x})g(\mathbf{y})-f(\mathbf{y})g(\mathbf{y})+f(\mathbf{y})g(\mathbf{y})-f(\mathbf{y})g(\mathbf{x})|]\\&\le\frac{2}{|g(\mathbf{x})|^2}[|g(\mathbf{y})||f(\mathbf{x})-f(\mathbf{y})|+|f(\mathbf{y})||g(\mathbf{y})-g(\mathbf{x})|]\\&\le\frac{2}{|g(\mathbf{x})|^2}\left[\frac{3}{2}|\mathbf{g}(\mathbf{x})||f(\mathbf{x})-f(\mathbf{y})|+(1+|f(\mathbf{x})|)|g(\mathbf{y})-g(\mathbf{x})|\right]\\&\le\frac{2}{|g(\mathbf{x})|^2}(1+2|f(\mathbf{x})|+2|g(\mathbf{x})|)[|f(\mathbf{x})-f(\mathbf{y})|+|g(\mathbf{y})-g(\mathbf{x})|]\\&\equiv M[|f(\mathbf{x})-f(\mathbf{y})|+|g(\mathbf{y})-g(\mathbf{x})|]\end{aligned}$$

where

$$M\equiv\frac{2}{|g(\mathbf{x})|^2}(1+2|f(\mathbf{x})|+2|g(\mathbf{x})|)$$

Now let δ_2 be such that if $|\mathbf{x}-\mathbf{y}|<\delta_2$, then

$$|f(\mathbf{x})-f(\mathbf{y})|<\frac{\varepsilon}{2}M^{-1}$$

and let δ_3 be such that if $|\mathbf{x}-\mathbf{y}|<\delta_3$, then

$$|g(\mathbf{y})-g(\mathbf{x})|<\frac{\varepsilon}{2}M^{-1}.$$

Then if $0<\delta\le\min(\delta_0,\delta_1,\delta_2,\delta_3)$, and $|\mathbf{x}-\mathbf{y}|<\delta$, everything holds and

$$\left|\frac{f(\mathbf{x})}{g(\mathbf{x})}-\frac{f(\mathbf{y})}{g(\mathbf{y})}\right|\le M[|f(\mathbf{x})-f(\mathbf{y})|+|g(\mathbf{y})-g(\mathbf{x})|]$$

$$< M\left[\frac{\varepsilon}{2}M^{-1} + \frac{\varepsilon}{2}M^{-1}\right] = \varepsilon.$$

This completes the proof of the second part of 2.) Note that in these proofs no effort is made to find some sort of "best" δ. The problem is one which has a yes or a no answer. Either is it or it is not continuous.

Now begin on 3.). If $\mathbf{f}$ is continuous at $\mathbf{x}$, $\mathbf{f}(\mathbf{x}) \in D(\mathbf{g}) \subseteq \mathbb{R}^p$, and $\mathbf{g}$ is continuous at $\mathbf{f}(\mathbf{x})$,then $\mathbf{g}\circ\mathbf{f}$ is continuous at $\mathbf{x}$. Let $\varepsilon > 0$ be given. Then there exists $\eta > 0$ such that if $|\mathbf{y} - \mathbf{f}(\mathbf{x})| < \eta$ and $\mathbf{y} \in D(\mathbf{g})$, it follows that $|\mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{f}(\mathbf{x}))| < \varepsilon$. It follows from continuity of $\mathbf{f}$ at $\mathbf{x}$ that there exists $\delta > 0$ such that if $|\mathbf{x} - \mathbf{z}| < \delta$ and $\mathbf{z} \in D(\mathbf{f})$, then $|\mathbf{f}(\mathbf{z}) - \mathbf{f}(\mathbf{x})| < \eta$. Then if $|\mathbf{x} - \mathbf{z}| < \delta$ and $\mathbf{z} \in D(\mathbf{g}\circ\mathbf{f}) \subseteq D(\mathbf{f})$, all the above hold and so

$$|\mathbf{g}(\mathbf{f}(\mathbf{z})) - \mathbf{g}(\mathbf{f}(\mathbf{x}))| < \varepsilon.$$

This proves part 3.)

Part 4.) says: If $\mathbf{f} = (f_1, \cdots, f_q) : D(\mathbf{f}) \to \mathbb{R}^q$, then $\mathbf{f}$ is continuous if and only if each f_k is a continuous real valued function. Then

$$\begin{aligned} |f_k(\mathbf{x}) - f_k(\mathbf{y})| &\leq |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \\ &\equiv \left(\sum_{i=1}^{q} |f_i(\mathbf{x}) - f_i(\mathbf{y})|^2\right)^{1/2} \\ &\leq \sum_{i=1}^{q} |f_i(\mathbf{x}) - f_i(\mathbf{y})|. \end{aligned} \tag{3.11}$$

Suppose first that $\mathbf{f}$ is continuous at $\mathbf{x}$. Then there exists $\delta > 0$ such that if $|\mathbf{x} - \mathbf{y}| < \delta$, then $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \varepsilon$. The first part of the above inequality then shows that for each $k = 1, \cdots, q$, $|f_k(\mathbf{x}) - f_k(\mathbf{y})| < \varepsilon$. This shows the only if part. Now suppose each function, f_k is continuous. Then if $\varepsilon > 0$ is given, there exists $\delta_k > 0$ such that whenever $|\mathbf{x} - \mathbf{y}| < \delta_k$

$$|f_k(\mathbf{x}) - f_k(\mathbf{y})| < \varepsilon/q.$$

Now let $0 < \delta \leq \min(\delta_1, \cdots, \delta_q)$. For $|\mathbf{x} - \mathbf{y}| < \delta$, the above inequality holds for all k and so the last part of (3.11) implies

$$\begin{aligned} |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| &\leq \sum_{i=1}^{q} |f_i(\mathbf{x}) - f_i(\mathbf{y})| \\ &< \sum_{i=1}^{q} \frac{\varepsilon}{q} = \varepsilon. \end{aligned}$$

This proves part 4.)

To verify part 5.), let $\varepsilon > 0$ be given and let $\delta = \varepsilon$. Then if $|\mathbf{x} - \mathbf{y}| < \delta$, the triangle inequality implies

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &= ||\mathbf{x}| - |\mathbf{y}|| \\ &\leq |\mathbf{x} - \mathbf{y}| < \delta = \varepsilon. \end{aligned}$$

This proves part 5.) and completes the proof of the theorem.

Here is a multidimensional version of the nested interval lemma.

Lemma 3.26 *Let* $I_k = \prod_{i=1}^{p}\left[a_i^k, b_i^k\right] \equiv \left\{\mathbf{x} \in \mathbb{R}^p : x_i \in \left[a_i^k, b_i^k\right]\right\}$ *and suppose that for all* $k = 1, 2, \cdots$,

$$I_k \supseteq I_{k+1}.$$

Then there exists a point, $\mathbf{c} \in \mathbb{R}^p$ *which is an element of every* I_k.

Proof: Since $I_k \supseteq I_{k+1}$, it follows that for each $i = 1, \cdots, p$, $\left[a_i^k, b_i^k\right] \supseteq \left[a_i^{k+1}, b_i^{k+1}\right]$. This implies that for each i,

$$a_i^k \leq a_i^{k+1},\ b_i^k \geq b_i^{k+1}. \tag{3.12}$$

Consequently, if $k \leq l$,

$$a_i^l \leq a_i^l \leq b_i^l \leq b_i^k. \tag{3.13}$$

Now define

$$c_i \equiv \sup \left\{a_i^l : l = 1, 2, \cdots\right\}$$

By the first inequality in (3.12),

$$c_i = \sup \left\{a_i^l : l = k, k+1, \cdots\right\} \tag{3.14}$$

for each $k = 1, 2 \cdots$. Therefore, picking any k,(3.13) shows that b_i^k is an upper bound for the set, $\left\{a_i^l : l = k, k+1, \cdots\right\}$ and so it is at least as large as the least upper bound of this set which is the definition of c_i given in (3.14). Thus, for each i and each k,

$$a_i^k \leq c_i \leq b_i^k.$$

Defining $\mathbf{c} \equiv (c_1, \cdots, c_p)$, $\mathbf{c} \in I_k$ for all k. This proves the lemma.

The following definition is similar to that given earlier. It defines what is meant by a sequentially compact set in $\mathbb{R}^p$.

Definition 3.27 *A set, $K \subseteq \mathbb{R}^p$ is sequentially compact if and only if whenever $\{\mathbf{x}_n\}_{n=1}^{\infty}$ is a sequence of points in K, there exists a point, $\mathbf{x} \in K$ and a subsequence, $\{\mathbf{x}_{n_k}\}_{k=1}^{\infty}$ such that $\mathbf{x}_{n_k} \to \mathbf{x}$.*

It turns out the sequentially compact sets in $\mathbb{R}^p$are exactly those which are closed and bounded. Only half of this result will be needed in this book and this is proved next.

Theorem 3.28 *Let $C \subseteq \mathbb{R}^p$ be closed and bounded. Then C is sequentially compact.*

Proof: Let $\{\mathbf{a}_n\} \subseteq C$, let $C \subseteq \prod_{i=1}^p [a_i, b_i]$, and consider all sets of the form $\prod_{i=1}^p [c_i, d_i]$ where $[c_i, d_i]$ equals either $\left[a_i, \frac{a_i+b_i}{2}\right]$ or $[c_i, d_i] = \left[\frac{a_i+b_i}{2}, b_i\right]$. Thus there are 2^p of these sets because there are two choices for the i^{th} slot for $i = 1, \cdots, p$. Also, if $\mathbf{x}$ and $\mathbf{y}$ are two points in one of these sets,

$$|x_i - y_i| \leq 2^{-1} |b_i - a_i| .$$

Therefore, letting $D_0 = \left(\sum_{i=1}^p |b_i - a_i|^2\right)^{1/2}$,

$$\begin{aligned} |\mathbf{x} - \mathbf{y}| &= \left(\sum_{i=1}^p |x_i - y_i|^2\right)^{1/2} \\ &\leq 2^{-1} \left(\sum_{i=1}^p |b_i - a_i|^2\right)^{1/2} \equiv 2^{-1} D_0. \end{aligned}$$

In particular, since $\mathbf{d} \equiv (d_1, \cdots, d_p)$ and $\mathbf{c} \equiv (c_1, \cdots, c_p)$ are two such points,

$$D_1 \equiv \left(\sum_{i=1}^p |d_i - c_i|^2\right)^{1/2} \leq 2^{-1} D_0$$

Denote by $\{J_1, \cdots, J_{2^p}\}$ these sets determined above. Since the union of these sets equals all of I_0, it follows

$$C = \cup_{k=1}^{2^p} J_k \cap C.$$

Pick J_k such that $\mathbf{a}_n$ is contained in $J_k \cap C$ for infinitely many values of n. Let $I_1 \equiv J_k$. Now do to I_1 what was done to I_0 to obtain $I_2 \subseteq I_1$ and for any two points, $\mathbf{x}, \mathbf{y} \in I_2$

$$|\mathbf{x} - \mathbf{y}| \leq 2^{-1} D_1 \leq 2^{-2} D_0,$$

and $I_2 \cap C$ contains $\mathbf{a}_n$ for infinitely many values of n. Continue in this way obtaining sets, I_k such that $I_k \supseteq I_{k+1}$ and for any two points in $I_k, \mathbf{x}, \mathbf{y}$, it follows $|\mathbf{x} - \mathbf{y}| \leq 2^{-k} D_0$, and $I_k \cap C$ contains $\mathbf{a}_n$ for infinitely many values of n. By the nested interval lemma, there exists a point, $\mathbf{c}$ which is contained in each I_k.

Claim: $\mathbf{c} \in C$.

Proof of claim: Suppose $\mathbf{c} \notin C$. Since C is a closed set, there exists $r > 0$ such that $B(\mathbf{c}, r)$ is contained completely in $\mathbb{R}^p \setminus C$. In other words, $B(\mathbf{c}, r)$ contains no points of C. Let k be so large that $D_0 2^{-k} < r$. Then since $\mathbf{c} \in I_k$, and any two points of I_k are closer than $D_0 2^{-k}, I_k$ must be contained in $B(\mathbf{c}, r)$ and so has no points of C in it, contrary to the manner in which the I_k are defined in which I_k contains $\mathbf{a}_n$ for infinitely many values of n. Therefore, $\mathbf{c} \in C$ as claimed.

Now pick $\mathbf{a}_{n_1} \in I_1 \cap \{\mathbf{a}_n\}_{n=1}^{\infty}$. Having picked this, let $\mathbf{a}_{n_2} \in I_2 \cap \{\mathbf{a}_n\}_{n=1}^{\infty}$ with $n_2 > n_1$. Having picked these two, let $\mathbf{a}_{n_3} \in I_3 \cap \{\mathbf{a}_n\}_{n=1}^{\infty}$ with $n_3 > n_2$ and continue this way. The result is a subsequence of $\{\mathbf{a}_n\}_{n=1}^{\infty}$ which converges to $\mathbf{c} \in C$ because any two points in I_k are within $D_0 2^{-k}$ of each other. This proves the theorem.

Here is a proof of the extreme value theorem.

Theorem 3.29 *Let C be closed and bounded and let $f : C \to \mathbb{R}$ be continuous. Then f achieves its maximum and its minimum on C. This means there exist, $\mathbf{x}_1, \mathbf{x}_2 \in C$ such that for all $\mathbf{x} \in C$,*

$$f(\mathbf{x}_1) \leq f(\mathbf{x}) \leq f(\mathbf{x}_2).$$

Proof: Let $M = \sup\{f(\mathbf{x}) : \mathbf{x} \in C\}$. Recall this means $+\infty$ if f is not bounded above and it equals the least upper bound of these values of f if f is bounded above. Then there exists a sequence, $\{\mathbf{x}_n\}$ such that $f(\mathbf{x}_n) \to M$. Since C is sequentially compact, there exists a subsequence, $\mathbf{x}_{n_k}$, and a point, $\mathbf{x} \in C$ such that $\mathbf{x}_{n_k} \to \mathbf{x}$. But then since f is continuous at $\mathbf{x}$, it follows from Theorem 3.22 on Page 36 that $f(\mathbf{x}) = \lim_{k\to\infty} f(\mathbf{x}_{n_k}) = M$. This proves f achieves its maximum and also shows its maximum is less than ∞. Let $\mathbf{x}_2 = \mathbf{x}$. The case of a minimum is handled similarly.

Recall that a function is uniformly continuous if the following definition holds.

Definition 3.30 *Let $\mathbf{f} : D(\mathbf{f}) \to \mathbb{R}^q$. Then $\mathbf{f}$ is uniformly continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $|\mathbf{x} - \mathbf{y}| < \delta$, it follows $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \varepsilon$.*

Theorem 3.31 *Let $\mathbf{f} : C \to \mathbb{R}^q$ be continuous where C is a closed and bounded set in $\mathbb{R}^p$. Then $\mathbf{f}$ is uniformly continuous on C.*

Proof: If this is not so, there exists $\varepsilon > 0$ and pairs of points, $\mathbf{x}_n$ and $\mathbf{y}_n$ satisfying $|\mathbf{x}_n - \mathbf{y}_n| < 1/n$ but $|\mathbf{f}(\mathbf{x}_n) - \mathbf{f}(\mathbf{y}_n)| \geq \varepsilon$. Since C is sequentially compact, there exists $\mathbf{x} \in C$ and a subsequence, $\{\mathbf{x}_{n_k}\}$ satisfying $\mathbf{x}_{n_k} \to \mathbf{x}$. But $|\mathbf{x}_{n_k} - \mathbf{y}_{n_k}| < 1/k$ and so $\mathbf{y}_{n_k} \to \mathbf{x}$ also. Therefore, from Theorem 3.22 on Page 36,

$$\varepsilon \leq \lim_{k\to\infty} |\mathbf{f}(\mathbf{x}_{n_k}) - \mathbf{f}(\mathbf{y}_{n_k})| = |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x})| = 0,$$

a contradiction. This proves the theorem.

3.9 The Concept Of A Norm

To do calculus, you must have some concept of distance and this is often provided by a norm. In all this, $\mathbb{F}^n$ will denote either $\mathbb{R}^n$ or $\mathbb{C}^n$. You already know about norms in $\mathbb{R}^n$. The purpose of this section is to review and to prove the theorem that all norms are equivalent.

Definition 3.32 ***Norms*** *satisfy*

$$||\mathbf{x}|| \geq 0, \ \ ||\mathbf{x}|| = 0 \text{ if and only if } \mathbf{x} = 0,$$

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||,$$

$$||c\mathbf{x}|| = |c| \, ||\mathbf{x}||$$

whenever c is a scalar. A set, U in $\mathbb{F}^n$ is ***open*** *if for every $\mathbf{p} \in U$, there exists $\delta > 0$ such that*

$$B(p, \delta) \equiv \{\mathbf{x} : ||\mathbf{x} - \mathbf{p}|| < \delta\} \subseteq U.$$

This is often referred to by saying that every point of the set is an ***interior point****.*

To begin with here is a fundamental inequality called the **Cauchy Schwarz inequality** which is stated here in $\mathbb{C}^n$. First here is a simple lemma.

Lemma 3.33 *If $z \in \mathbb{C}$ there exists $\theta \in \mathbb{C}$ such that $\theta z = |z|$ and $|\theta| = 1$.*

Proof: Let $\theta = 1$ if $z = 0$ and otherwise, let $\theta = \frac{\overline{z}}{|z|}$. Recall that for $z = x + iy, \overline{z} = x - iy$.

Definition 3.34 *For $\mathbf{x} \in \mathbb{C}^n$,*

$$|\mathbf{x}| \equiv \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2}.$$

Theorem 3.35 *(Cauchy Schwarz)The following inequality holds for x_i and $y_i \in \mathbb{C}$.*

$$\left| \sum_{i=1}^{n} x_i \overline{y}_i \right| \leq \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2 \right)^{1/2}. \tag{3.15}$$

Proof: Let $\theta \in \mathbb{C}$ such that $|\theta| = 1$ and

$$\theta \sum_{i=1}^{n} x_i \overline{y}_i = \left| \sum_{i=1}^{n} x_i \overline{y}_i \right|$$

Thus

$$\theta \sum_{i=1}^{n} x_i \overline{y}_i = \sum_{i=1}^{n} x_i \overline{\left(\overline{\theta} y_i\right)} = \left| \sum_{i=1}^{n} x_i \overline{y}_i \right|.$$

Consider $p(t) \equiv \sum_{i=1}^{n} \left(x_i + t\overline{\theta} y_i\right) \left(\overline{x_i + t\overline{\theta} y_i}\right)$ where $t \in \mathbb{R}$.

$$\begin{aligned} 0 &\leq p(t) = \sum_{i=1}^{n} |x_i|^2 + 2t \operatorname{Re}\left(\theta \sum_{i=1}^{n} x_i \overline{y}_i \right) + t^2 \sum_{i=1}^{n} |y_i|^2 \\ &= |\mathbf{x}|^2 + 2t \left| \sum_{i=1}^{n} x_i \overline{y}_i \right| + t^2 |\mathbf{y}|^2 \end{aligned}$$

If $|\mathbf{y}| = 0$ then (3.15) is obviously true because both sides equal zero. Therefore, assume $|\mathbf{y}| \neq 0$ and then $p(t)$ is a polynomial of degree two whose graph opens up. Therefore, it either has no zeroes, two zeros or one repeated zero. If it has two zeros, the above inequality must be violated because in this case the graph must dip below the x axis. Therefore, it either has no zeros or exactly one. From the quadratic formula this happens exactly when

$$4\left|\sum_{i=1}^{n} x_i \overline{y}_i\right|^2 - 4|\mathbf{x}|^2 |\mathbf{y}|^2 \leq 0$$

and so

$$\left|\sum_{i=1}^{n} x_i \overline{y}_i\right| \leq |\mathbf{x}| |\mathbf{y}|$$

as claimed. This proves the inequality.

Theorem 3.36 *The norm $|\cdot|$ given in Definition 3.34 really is a norm. Also if $||\cdot||$ is any norm on $\mathbb{F}^n$. Then $||\cdot||$ is equivalent to $|\cdot|$. That is there exist constants, δ and Δ such that*

$$\delta ||\mathbf{x}|| \leq |\mathbf{x}| \leq \Delta ||\mathbf{x}|| . \tag{3.16}$$

Proof: All of the above properties of a norm are obvious except the second, the triangle inequality. To establish this inequality, use the Cauchy Schwarz inequality to write

$$\begin{aligned} |\mathbf{x}+\mathbf{y}|^2 &\equiv \sum_{i=1}^{n} |x_i + y_i|^2 \leq \sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} |y_i|^2 + 2\operatorname{Re}\sum_{i=1}^{n} x_i \overline{y}_i \\ &\leq |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2} \\ &= |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| = (|\mathbf{x}| + |\mathbf{y}|)^2 \end{aligned}$$

and this proves the second property above.

It remains to show the equivalence of the two norms. Letting $\{\mathbf{e}_k\}$ denote the usual basis vectors for $\mathbb{C}^n$, the Cauchy Schwarz inequality implies

$$\begin{aligned} ||\mathbf{x}|| &\equiv \left|\left|\sum_{i=1}^{n} x_i \mathbf{e}_i\right|\right| \leq \sum_{i=1}^{n} |x_i| \, ||\mathbf{e}_i|| \leq |\mathbf{x}| \left(\sum_{i=1}^{n} ||\mathbf{e}_i||^2\right)^{1/2} \\ &\equiv \delta^{-1} |\mathbf{x}| . \end{aligned}$$

This proves the first half of the inequality.

Suppose the second half of the inequality is not valid. Then there exists a sequence $\mathbf{x}^k \in \mathbb{F}^n$ such that

$$\left|\mathbf{x}^k\right| > k \left|\left|\mathbf{x}^k\right|\right|, \ k = 1, 2, \cdots.$$

Then define

$$\mathbf{y}^k \equiv \frac{\mathbf{x}^k}{|\mathbf{x}^k|}.$$

It follows

$$\left|\mathbf{y}^k\right| = 1, \quad \left|\mathbf{y}^k\right| > k \left|\left|\mathbf{y}^k\right|\right| \tag{3.17}$$

and the vector

$$\left(y_1^k, \cdot\cdot\cdot, y_n^k\right)$$

is a unit vector in $\mathbb{F}^n$. By the Heine Borel theorem from calculus, there exists a subsequence, still denoted by k such that

$$\left(y_1^k, \cdot\cdot\cdot, y_n^k\right) \rightarrow (y_1, \cdot\cdot\cdot, y_n) = \mathbf{y}$$

a unit vector. It follows from (3.17) and this that for

$$\mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{e}_i,$$

$$0 = \lim_{k\rightarrow\infty} \left|\left|\mathbf{y}^k\right|\right| = \lim_{k\rightarrow\infty} \left|\left|\sum_{i=1}^{n} y_i^k \mathbf{e}_i\right|\right| = \left|\left|\sum_{i=1}^{n} y_i \mathbf{e}_i\right|\right|$$

but not all the y_i equal zero because $\mathbf{y}$ is a unit vector. This contradicts the linear independence of $\{\mathbf{e}_1, \cdot\cdot\cdot, \mathbf{e}_n\}$ and proves the second half of the inequality.

Corollary 3.37 *Any two norms on $\mathbb{F}^n$ are equivalent. That is, if $||\cdot||$ and $|||\cdot|||$ are two norms on $\mathbb{F}^n$, then there exist positive constants, δ and Δ, independent of $\mathbf{x} \in X$ such that*

$$\delta\, |||\mathbf{x}||| \leq ||\mathbf{x}|| \leq \Delta\, |||\mathbf{x}|||\,.$$

Proof: By Theorem 3.36, there are positive constants $\delta_1, \Delta_1, \delta_2, \Delta_2$, all independent of $\mathbf{x} \in \mathbb{F}^n$ such that

$$\delta_2\, |||\mathbf{x}||| \leq |\mathbf{x}| \leq \Delta_2\, |||\mathbf{x}|||\,,$$

$$\delta_1\, ||\mathbf{x}|| \leq |\mathbf{x}| \leq \Delta_1\, ||\mathbf{x}||\,.$$

Then

$$\delta_2\, |||\mathbf{x}||| \leq |\mathbf{x}| \leq \Delta_1\, ||\mathbf{x}|| \leq \frac{\Delta_1}{\delta_1}\, |\mathbf{x}| \leq \frac{\Delta_1 \Delta_2}{\delta_1}\, |||\mathbf{x}|||$$

and so

$$\frac{\delta_2}{\Delta_1}\, |||\mathbf{x}||| \leq ||\mathbf{x}|| \leq \frac{\Delta_2}{\delta_1}\, |||\mathbf{x}|||$$

which proves the corollary.

3.10 The Operator Norm

Definition 3.38 *Let norms $||\cdot||_{\mathbb{F}^n}$ and $||\cdot||_{\mathbb{F}^m}$ be given on $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Then $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ denotes the space of linear transformations mapping $\mathbb{F}^n$ to $\mathbb{F}^m$. Recall that if you pick a basis on $\mathbb{F}^n$, $\{\mathbf{v}_1, \cdot\cdot\cdot, \mathbf{v}_n\}$ and $\mathbb{F}^m$, $\{\mathbf{w}_1, \cdot\cdot\cdot, \mathbf{w}_m\}$, a linear transformation, L determines a matrix in the following way. Let q_V and q_W be defined by*

$$q_V(\mathbf{x}) \equiv \sum_j x_j \mathbf{v}_j,\ q_W(\mathbf{y}) \equiv \sum_k y_k \mathbf{w}_k$$

Then the matrix, A is such that matrix multiplication makes the following diagram commute.

$$\begin{array}{ccccc} & & L & & \\ \{\mathbf{v}_1,\cdots,\mathbf{v}_n\} & V & \rightarrow & W & \{\mathbf{w}_1,\cdots,\mathbf{w}_m\} \\ & q_V\uparrow & \circ & \uparrow q_W & \\ & \mathbb{F}^n & \rightarrow & \mathbb{F}^m & \\ & & A & & \end{array} \tag{3.18}$$

$m\times n$ matrices. I will not attempt to make an issue of the difference between a matrix and a linear transformation because there is really no loss of generality in simply thinking of a matrix as a linear transformation using matrix multiplication. This is really what the above diagram says. Thus $A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ will mean that A is a linear transformation but I may also refer to it as a matrix. The ***operator norm*** *is defined by*

$$||A||\equiv\sup\{||Ax||_{\mathbb{F}^m}:||x||_{\mathbb{F}^n}\leq 1\}<\infty.$$

Theorem 3.39 *Denote by $||\cdot||$ the norm on either $\mathbb{F}^n$ or $\mathbb{F}^m$. The set of $m\times n$ matrices with this norm is a* ***complete normed linear space*** *of dimension nm with*

$$||A\mathbf{x}||\leq||A||\,||\mathbf{x}||\,.$$

Completeness *means that every Cauchy sequence converges.*

Proof: It is necessary to show the norm defined on $\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ really is a norm. Again the first and third properties listed above for norms are obvious. It remains to show the second and verify $||A||<\infty$. There exist constants $\delta,\Delta>0$ such that

$$\delta\,||\mathbf{x}||\leq|\mathbf{x}|\leq\Delta\,||\mathbf{x}||\,.$$

Then,

$$\begin{aligned}||A+B|| &\equiv \sup\{||(A+B)(\mathbf{x})||:||\mathbf{x}||\leq 1\}\\ &\leq \sup\{||A\mathbf{x}||:||\mathbf{x}||\leq 1\}+\sup\{||B\mathbf{x}||:||\mathbf{x}||\leq 1\}\\ &\equiv ||A||+||B||\,.\end{aligned}$$

Next consider the claim that $||A||<\infty$. This follows from

$$||A(\mathbf{x})||=\left|\left|A\left(\sum_{i=1}^n x_i\mathbf{e}_i\right)\right|\right|\leq\sum_{i=1}^n|x_i|\,||A(\mathbf{e}_i)||$$

$$\leq|\mathbf{x}|\left(\sum_{i=1}^n||A(\mathbf{e}_i)||^2\right)^{1/2}\leq\Delta\,||\mathbf{x}||\left(\sum_{i=1}^n||A(\mathbf{e}_i)||^2\right)^{1/2}<\infty.$$

Thus $||A||\leq\Delta\left(\sum_{i=1}^n||A(\mathbf{e}_i)||^2\right)^{1/2}$.

It is clear that a basis for $\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ consists of matrices of the form E_{ij} where E_{ij} consists of the $m\times n$ matrix having all zeros except for a 1 in the ij^{th} position. In effect, this considers $\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ as $\mathbb{F}^{nm}$. Think of the $m\times n$ matrix as a long vector folded up.

If $\mathbf{x}\neq\mathbf{0}$,

$$||A\mathbf{x}||\,\frac{1}{||\mathbf{x}||}=\left|\left|A\frac{\mathbf{x}}{||\mathbf{x}||}\right|\right|\leq||A|| \tag{3.19}$$

It only remains to verify completeness. Suppose then that $\{A_k\}$ is a Cauchy sequence in $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Then from (3.19) $\{A_k \mathbf{x}\}$ is a Cauchy sequence for each $\mathbf{x} \in \mathbb{F}^n$. This follows because

$$||A_k \mathbf{x} - A_l \mathbf{x}|| \leq ||A_k - A_l|| \, ||\mathbf{x}||$$

which converges to 0 as $k, l \to \infty$. Therefore, by completeness of $\mathbb{F}^n$, there exists $A\mathbf{x}$, the name of the thing to which the sequence, $\{A_k \mathbf{x}\}$ converges such that

$$\lim_{k \to \infty} A_k \mathbf{x} = A\mathbf{x}.$$

Then A is linear because

$$\begin{aligned} A(a\mathbf{x} + b\mathbf{y}) &\equiv \lim_{k\to\infty} A_k (a\mathbf{x} + b\mathbf{y}) \\ &= \lim_{k\to\infty} (aA_k\mathbf{x} + bA_k\mathbf{y}) \\ &= a \lim_{k\to\infty} A_k\mathbf{x} + b \lim_{k\to\infty} A_k\mathbf{y} \\ &= aA\mathbf{x} + bA\mathbf{y}. \end{aligned}$$

By the first part of this argument, $||A|| < \infty$ and so $A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. This proves the theorem.

It turns out that there are many ways of placing a norm on $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ and they are all equivalent. This follows because as noted above, you can think of $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ as $\mathbb{F}^{nm}$ and it was shown in Corollary 3.37 that any two norms on this space are equivalent. One popular norm is the following called the **Frobenius norm**. It is not an operator norm but instead is based on the idea of considering the $m \times n$ matrix as an element of $\mathbb{F}^{nm}$. Recall the trace of an $n \times n$ matrix, (a_{ij}) is just $\sum_j a_{jj}$. In other words, it is just the sum of the entries on the main diagonal.

Definition 3.40 *Define an inner product on $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ as follows.*

$$(A, B) \equiv tr(AB^*)$$

*where tr denotes the **trace**. Thus*

$$tr(A) \equiv \sum_i A_{ii},$$

the sum of the entries on the main diagonal. Then define $||A|| \equiv (A, A)^{1/2}$. It is obvious this is a norm from the argument above in Theorem 3.36 applied this time to $\mathbb{F}^{nm}$. This follows because

$$(A, B) \equiv tr(AB^*) \equiv \sum_i \sum_j a_{ij} \overline{b_{ij}}$$

There are many norms which are used on $\mathbb{C}^n$. The most common ones are listed below. By Corollary 3.37 they are all equivalent. This means that in any convergence question it does not make any difference which of these norms you use.

Definition 3.41 *Let $\mathbf{x} \in \mathbb{C}^n$. Then define for $p \geq 1$,*

$$||\mathbf{x}||_p \equiv \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

$$||\mathbf{x}||_1 \equiv \sum_{i=1}^n |x_i|,$$

$$||\mathbf{x}||_\infty \equiv \max\{|x_i|, i = 1, \cdots, n\},$$

$$||\mathbf{x}||_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$$

The last is the usual norm often referred to as the Euclidean norm.

It has already been shown that the last of the above norms is really a norm. It is easy to verify that $||\cdot||_1$ is a norm and also not hard to see that $||\cdot||_\infty$ is a norm. You should verify this. The norm, $||\cdot||_p$ is more difficult, however. The following inequality is called **Holder's inequality**. It is a generalization of the Cauchy Schwarz inequality. It is always assumed that $p > 1$ and p' is defined by the equation

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Proposition 3.42 *For* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$\sum_{i=1}^n |x_i| |y_i| \leq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \left(\sum_{i=1}^n |y_i|^{p'}\right)^{1/p'}$$

The proof will depend on the following lemma.

Lemma 3.43 *If* $a, b \geq 0$ *and* p' *is defined by* $\frac{1}{p} + \frac{1}{p'} = 1$, *then*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

Proof of the Proposition: If $\mathbf{x}$ or $\mathbf{y}$ equals the zero vector there is nothing to prove. Therefore, assume they are both nonzero. Let $A = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ and $B = \left(\sum_{i=1}^n |y_i|^{p'}\right)^{1/p'}$. Then using Lemma 3.43,

$$\begin{aligned} \sum_{i=1}^n \frac{|x_i|}{A} \frac{|y_i|}{B} &\leq \sum_{i=1}^n \left[\frac{1}{p}\left(\frac{|x_i|}{A}\right)^p + \frac{1}{p'}\left(\frac{|y_i|}{B}\right)^{p'}\right] \\ &= 1 \end{aligned}$$

and so

$$\sum_{i=1}^n |x_i| |y_i| \leq AB = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \left(\sum_{i=1}^n |y_i|^{p'}\right)^{1/p'}.$$

This proves the proposition.

Theorem 3.44 *The* p *norms do indeed satisfy the axioms of a norm.*

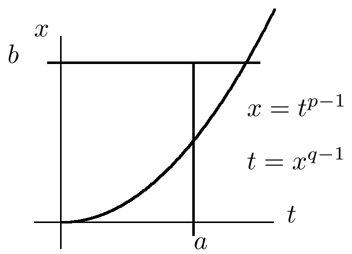
Proof: It is obvious that $||\cdot||_p$ does indeed satisfy most of the norm axioms. The only one that is not clear is the triangle inequality. To save notation write $||\cdot||$ in place of $||\cdot||_p$ in what follows. Note also that $\frac{p}{p'} = p - 1$. Then using the Holder inequality,

$$\begin{aligned}
||\mathbf{x}+\mathbf{y}||^p &= \sum_{i=1}^n |x_i+y_i|^p \\
&\leq \sum_{i=1}^n |x_i+y_i|^{p-1}|x_i| + \sum_{i=1}^n |x_i+y_i|^{p-1}|y_i| \\
&= \sum_{i=1}^n |x_i+y_i|^{\frac{p}{p'}}|x_i| + \sum_{i=1}^n |x_i+y_i|^{\frac{p}{p'}}|y_i| \\
&\leq \left(\sum_{i=1}^n |x_i+y_i|^p\right)^{1/p'} \left[\left(\sum_{i=1}^n |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |y_i|^p\right)^{1/p}\right] \\
&= ||\mathbf{x}+\mathbf{y}||^{p/p'} \left(||\mathbf{x}||_p + ||\mathbf{y}||_p\right)
\end{aligned}$$

so $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}||_p + ||\mathbf{y}||_p$. This proves the theorem.

It only remains to prove Lemma 3.43.

Proof of the lemma: Let $p' = q$ to save on notation and consider the following picture:



$$ab \leq \int_0^a t^{p-1}dt + \int_0^b x^{q-1}dx = \frac{a^p}{p} + \frac{b^q}{q}.$$

Note equality occurs when $a^p = b^q$.

Now $||A||_p$ is the operator norm of A taken with respect to $||\cdot||_p$.

Theorem 3.45 *The following holds.*

$$||A||_p \leq \left(\sum_k \left(\sum_j |A_{jk}|^p\right)^{q/p}\right)^{1/q}$$

Proof: Let $||\mathbf{x}||_p \leq 1$ and let $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n)$ where the $\mathbf{a}_k$ are the columns of A. Then

$$A\mathbf{x} = \left(\sum_k x_k \mathbf{a}_k\right)$$

and so by Holder's inequality,

$$\begin{aligned}
||A\mathbf{x}||_p &\equiv \left|\left|\sum_k x_k \mathbf{a}_k\right|\right|_p \leq \sum_k |x_k| \, ||\mathbf{a}_k||_p \\
&\leq \left(\sum_k |x_k|^p\right)^{1/p} \left(\sum_k ||\mathbf{a}_k||_p^q\right)^{1/q} \\
&\leq \left(\sum_k \left(\sum_j |A_{jk}|^p\right)^{q/p}\right)^{1/q}
\end{aligned}$$

and this shows $||A||_p \leq \left(\sum_k \left(\sum_j |A_{jk}|^p\right)^{q/p}\right)^{1/q}$ and proves the theorem.

3.11 The Frechet Derivative

Let U be an open set in $\mathbb{F}^n$, and let $\mathbf{f} : U \to \mathbb{F}^m$ be a function.

Definition 3.46 *A function* $\mathbf{g}$ *is* $o(\mathbf{v})$ *if*

$$\lim_{||\mathbf{v}|| \to 0} \frac{\mathbf{g}(\mathbf{v})}{||\mathbf{v}||} = \mathbf{0} \tag{3.20}$$

A function $\mathbf{f} : U \to \mathbb{F}^m$ *is differentiable at* $\mathbf{x} \in U$ *if there exists a linear transformation* $L \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ *such that*

$$\mathbf{f}(\mathbf{x} + \mathbf{v}) = \mathbf{f}(\mathbf{x}) + L\mathbf{v} + o(\mathbf{v})$$

This linear transformation L *is the definition of* $D\mathbf{f}(\mathbf{x})$*. This derivative is often called the Frechet derivative. .*

Note that it does not matter which norm is used in this definition because of Theorem 3.36 on Page 44 and Corollary 3.37. The definition means that the error,

$$\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x}) - L\mathbf{v}$$

converges to $\mathbf{0}$ faster than $||\mathbf{v}||$. Thus the above definition is equivalent to saying

$$\lim_{||\mathbf{v}|| \to 0} \frac{||\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x}) - L\mathbf{v}||}{||\mathbf{v}||} = 0. \tag{3.21}$$

Now it is clear this is just a generalization of the notion of the derivative of a function of one variable because in this more specialized situation,

$$\lim_{|v| \to 0} \frac{|f(x + v) - f(x) - f'(x) v|}{|v|} = 0,$$

due to the definition which says

$$f'(x) = \lim_{v \to 0} \frac{f(x + v) - f(x)}{v}.$$

For functions of n variables, you can't define the derivative as the limit of a difference quotient like you can for a function of one variable because you can't divide by a vector. That is why there is a need for a more general definition.

The term $o(\mathbf{v})$ is notation that is descriptive of the behavior in (3.20) and it is only this behavior that is of interest. Thus, if t and k are constants,

$$o(\mathbf{v}) = o(\mathbf{v}) + o(\mathbf{v}),\ o(t\mathbf{v}) = o(\mathbf{v}),\ ko(\mathbf{v}) = o(\mathbf{v})$$

and other similar observations hold. The sloppiness built in to this notation is useful because it ignores details which are not important. It may help to think of $o(\mathbf{v})$ as an adjective describing what is left over after approximating $\mathbf{f}(\mathbf{x}+\mathbf{v})$ by $\mathbf{f}(\mathbf{x}) + D\mathbf{f}(\mathbf{x})\mathbf{v}$.

Theorem 3.47 *The derivative is well defined.*

Proof: First note that for a fixed vector, $\mathbf{v}$, $o(t\mathbf{v}) = o(t)$. Now suppose both L_1 and L_2 work in the above definition. Then let $\mathbf{v}$ be any vector and let t be a real scalar which is chosen small enough that $t\mathbf{v} + \mathbf{x} \in U$. Then

$$\mathbf{f}(\mathbf{x}+t\mathbf{v}) = \mathbf{f}(\mathbf{x}) + L_1 t\mathbf{v} + o(t\mathbf{v}),\ \mathbf{f}(\mathbf{x}+t\mathbf{v}) = \mathbf{f}(\mathbf{x}) + L_2 t\mathbf{v} + o(t\mathbf{v}).$$

Therefore, subtracting these two yields $(L_2 - L_1)(t\mathbf{v}) = o(t\mathbf{v}) = o(t)$. Therefore, dividing by t yields $(L_2 - L_1)(\mathbf{v}) = \frac{o(t)}{t}$. Now let $t \to 0$ to conclude that $(L_2 - L_1)(\mathbf{v}) = 0$. Since this is true for all $\mathbf{v}$, it follows $L_2 = L_1$. This proves the theorem.

Lemma 3.48 *Let $\mathbf{f}$ be differentiable at $\mathbf{x}$. Then $\mathbf{f}$ is continuous at $\mathbf{x}$ and in fact, there exists $K > 0$ such that whenever $||\mathbf{v}||$ is small enough,*

$$||\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})|| \le K\,||\mathbf{v}||$$

Proof: From the definition of the derivative, $\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{x})\mathbf{v} + o(\mathbf{v})$. Let $||\mathbf{v}||$ be small enough that $\frac{o(||\mathbf{v}||)}{||\mathbf{v}||} < 1$ so that $||o(\mathbf{v})|| \le ||\mathbf{v}||$. Then for such $\mathbf{v}$,

$$\begin{aligned} ||\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})|| &\le ||D\mathbf{f}(\mathbf{x})\mathbf{v}|| + ||\mathbf{v}|| \\ &\le (||D\mathbf{f}(\mathbf{x})|| + 1)\,||\mathbf{v}|| \end{aligned}$$

This proves the lemma with $K = ||D\mathbf{f}(\mathbf{x})|| + 1$.

Theorem 3.49 *(The chain rule) Let U and V be open sets, $U \subseteq \mathbb{F}^n$ and $V \subseteq \mathbb{F}^m$. Suppose $\mathbf{f} : U \to V$ is differentiable at $\mathbf{x} \in U$ and suppose $\mathbf{g} : V \to \mathbb{F}^q$ is differentiable at $\mathbf{f}(\mathbf{x}) \in V$. Then $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{x}$ and*

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D(\mathbf{g}(\mathbf{f}(\mathbf{x})))\,D(\mathbf{f}(\mathbf{x})).$$

Proof: This follows from a computation. Let $B(\mathbf{x},r) \subseteq U$ and let r also be small enough that for $||\mathbf{v}|| \le r$, it follows that $\mathbf{f}(\mathbf{x}+\mathbf{v}) \in V$. Such an r exists because $\mathbf{f}$ is continuous at $\mathbf{x}$. For $||\mathbf{v}|| < r$, the definition of differentiability of $\mathbf{g}$ and $\mathbf{f}$ implies

$$\mathbf{g}(\mathbf{f}(\mathbf{x}+\mathbf{v})) - \mathbf{g}(\mathbf{f}(\mathbf{x})) =$$

$$\begin{aligned} & D\mathbf{g}(\mathbf{f}(\mathbf{x}))(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})) + o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})) \\ = \ & D\mathbf{g}(\mathbf{f}(\mathbf{x}))[D\mathbf{f}(\mathbf{x})\mathbf{v} + o(\mathbf{v})] + o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})) \\ = \ & D(\mathbf{g}(\mathbf{f}(\mathbf{x})))\,D(\mathbf{f}(\mathbf{x}))\mathbf{v} + o(\mathbf{v}) + o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})). \end{aligned} \tag{3.22}$$

It remains to show $o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})) = o(\mathbf{v})$.

By Lemma 3.48, with K given there, letting $\varepsilon > 0$, it follows that for $||\mathbf{v}||$ small enough,

$$||o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x}))|| \le (\varepsilon/K)\,||\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})|| \le (\varepsilon/K)\,K\,||\mathbf{v}|| = \varepsilon\,||\mathbf{v}||.$$

Since $\varepsilon > 0$ is arbitrary, this shows $o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x})) = o(\mathbf{v})$ because whenever $||\mathbf{v}||$ is small enough,

$$\frac{||o(\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x}))||}{||\mathbf{v}||} \leq \varepsilon.$$

By (3.22), this shows

$$\mathbf{g}(\mathbf{f}(\mathbf{x}+\mathbf{v})) - \mathbf{g}(\mathbf{f}(\mathbf{x})) = D(\mathbf{g}(\mathbf{f}(\mathbf{x})))\, D(\mathbf{f}(\mathbf{x}))\,\mathbf{v} + o(\mathbf{v})$$

which proves the theorem.

The derivative is a linear transformation. The matrix of this linear transformation taken with respect to the usual bases will be denoted by $J\mathbf{f}(\mathbf{x})$. Denote by π_i the mapping which takes a vector in $\mathbb{F}^m$ and delivers the i^{th} component of this vector. Also let $\mathbf{e}_i$ denote the vector of $\mathbb{F}^m$ which has a one in the i^{th} entry and zeroes elsewhere. Thus, if the components of $\mathbf{v}$ with respect to the standard basis vectors are v_i,

$$\mathbf{v} = \sum_i v_i \mathbf{e}_i$$

and so

$$\sum_i \sum_j J\mathbf{f}(\mathbf{x})_{ij}\, v_j \mathbf{e}_i = D\mathbf{f}(\mathbf{x})\,\mathbf{v}.$$

Doing π_i to both sides,

$$\sum_j J\mathbf{f}(\mathbf{x})_{ij}\, v_j = \pi_i(D\mathbf{f}(\mathbf{x})\,\mathbf{v}) \tag{3.23}$$

What are the entries of $J\mathbf{f}(x)$? Letting $\mathbf{f}(\mathbf{x}) = \sum_{i=1}^m f_i(\mathbf{x})\,\mathbf{e}_i$, it follows

$$f_i(\mathbf{x}+\mathbf{v}) - f_i(\mathbf{x}) = \pi_i(D\mathbf{f}(\mathbf{x})\,\mathbf{v}) + o(\mathbf{v}).$$

Thus, letting t be a small scalar, and replacing $\mathbf{v}$ with $t\mathbf{e}_i$,

$$f_i(\mathbf{x}+t\mathbf{e}_j) - f_i(\mathbf{x}) = t\pi_i(D\mathbf{f}(\mathbf{x})\,\mathbf{e}_j) + o(t).$$

Dividing by t, and letting $t \to 0$, $\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \pi_i(D\mathbf{f}(\mathbf{x})\,\mathbf{e}_j)$. This says the i^{th} component of $D\mathbf{f}(\mathbf{x})\,\mathbf{e}_j$ equals $\frac{\partial f_i(\mathbf{x})}{\partial x_j}$. Thus, from (3.23),

$$J\mathbf{f}(\mathbf{x})_{ij} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}. \tag{3.24}$$

This proves the following theorem

Theorem 3.50 *Let $\mathbf{f} : \mathbb{F}^n \to \mathbb{F}^m$ and suppose $\mathbf{f}$ is differentiable at $\mathbf{x}$. Then all the partial derivatives $\frac{\partial f_i(\mathbf{x})}{\partial x_j}$ exist and if $J\mathbf{f}(\mathbf{x})$ is the matrix of the linear transformation with respect to the standard basis vectors, then the ij^{th} entry is given by (3.24).*

In practice we tend to think in terms of the standard basis and identify the derivative with this $m \times n$ matrix. Thus, it is not uncommon to see people refer to the matrix, $\left(\frac{\partial f_i(\mathbf{x})}{\partial x_j}\right)$ as the linear transformation, $D\mathbf{f}(\mathbf{x})$.

What if all the partial derivatives of $\mathbf{f}$ exist? Does it follow that $\mathbf{f}$ is differentiable? Consider the following function.

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} \text{ if } (x,y) \neq (0,0) \\ 0 \text{ if } (x,y) = (0,0) \end{cases}.$$

Then from the definition of partial derivatives,

$$\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{0-0}{h}=0$$

and

$$\lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}=\lim_{h\to 0}\frac{0-0}{h}=0$$

However f is not even continuous at $(0,0)$ which may be seen by considering the behavior of the function along the line $y=x$ and along the line $x=0$. By Lemma 3.48 this implies f is not differentiable. Therefore, it is necessary to consider the correct definition of the derivative given above if you want to get a notion which generalizes the concept of the derivative of a function of one variable in such a way as to preserve continuity whenever the function is differentiable.

However, there are theorems which can be used to get differentiability of a function based on existence of the partial derivatives.

Definition 3.51 *When all the partial derivatives exist and are continuous the function is called a C^1 function.*

Because of the above which identifies the entries of $J\mathbf{f}$ with the partial derivatives, the following definition is equivalent to the above.

Definition 3.52 *Let $U\subseteq\mathbb{F}^n$ be an open set. Then $\mathbf{f}:U\to\mathbb{F}^m$ is $C^1(U)$ if $\mathbf{f}$ is differentiable and the mapping*

$$\mathbf{x}\to D\mathbf{f}(\mathbf{x}),$$

is continuous as a function from U to $\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$.

The following is an important abstract generalization of the familiar concept of partial derivative.

Definition 3.53 *Place a norm on $\mathbb{F}^n\times\mathbb{F}^m$ as follows.*

$$||(\mathbf{x},\mathbf{y})||\equiv\max\left(||\mathbf{x}||_{\mathbb{F}^n},||\mathbf{y}||_{\mathbb{F}^m}\right).$$

Now let $\mathbf{g}:U\subseteq\mathbb{F}^n\times\mathbb{F}^m\to\mathbb{F}^q$, where U is an open set in $\mathbb{F}^n\times\mathbb{F}^m$. Denote an element of $\mathbb{F}^n\times\mathbb{F}^m$ by $(\mathbf{x},\mathbf{y})$ where $\mathbf{x}\in\mathbb{F}^n$ and $\mathbf{y}\in\mathbb{F}^m$. Then the map $\mathbf{x}\to\mathbf{g}(\mathbf{x},\mathbf{y})$ is a function from the open set in X,

$$\{\mathbf{x}:(\mathbf{x},\mathbf{y})\in U\}$$

to $\mathbb{F}^q$. When this map is differentiable, its derivative is denoted by

$$D_1\mathbf{g}(\mathbf{x},\mathbf{y}),\ \text{or sometimes by } D_{\mathbf{x}}\mathbf{g}(\mathbf{x},\mathbf{y}).$$

Thus,

$$\mathbf{g}(\mathbf{x}+\mathbf{v},\mathbf{y})-\mathbf{g}(\mathbf{x},\mathbf{y})=D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}+o(\mathbf{v}).$$

A similar definition holds for the symbol $D_{\mathbf{y}}\mathbf{g}$ or $D_2\mathbf{g}$.

The following theorem will be very useful in much of what follows. It is a version of the mean value theorem.

Theorem 3.54 *Suppose U is an open subset of $\mathbb{F}^n$ and $\mathbf{f}:U\to\mathbb{F}^m$ has the property that $D\mathbf{f}(\mathbf{x})$ exists for all $\mathbf{x}$ in U and that, $\mathbf{x}+t(\mathbf{y}-\mathbf{x})\in U$ for all $t\in[0,1]$. (The line segment joining the two points lies in U.) Suppose also that for all points on this line segment,*

$$||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||\leq M.$$

Then

$$||\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})||\leq M\,||\mathbf{y}-\mathbf{x}||.$$

Proof: Let

$$S \equiv \{t \in [0,1] : \text{ for all } s \in [0,t],$$

$$||\mathbf{f}(\mathbf{x}+s(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})|| \leq (M+\varepsilon)s||\mathbf{y}-\mathbf{x}||\}.$$

Then $0 \in S$ and by continuity of $\mathbf{f}$, it follows that if $t \equiv \sup S$, then $t \in S$ and if $t < 1$,

$$||\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})|| = (M+\varepsilon)t||\mathbf{y}-\mathbf{x}||. \tag{3.25}$$

If $t < 1$, then there exists a sequence of positive numbers, $\{h_k\}_{k=1}^{\infty}$ converging to 0 such that

$$||\mathbf{f}(\mathbf{x}+(t+h_k)(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})|| > (M+\varepsilon)(t+h_k)||\mathbf{y}-\mathbf{x}||$$

which implies that

$$||\mathbf{f}(\mathbf{x}+(t+h_k)(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||$$

$$+||\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})|| > (M+\varepsilon)(t+h_k)||\mathbf{y}-\mathbf{x}||.$$

By (3.25), this inequality implies

$$||\mathbf{f}(\mathbf{x}+(t+h_k)(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))|| > (M+\varepsilon)h_k||\mathbf{y}-\mathbf{x}||$$

which yields upon dividing by h_k and taking the limit as $h_k \to 0$,

$$||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})|| \geq (M+\varepsilon)||\mathbf{y}-\mathbf{x}||.$$

Now by the definition of the norm of a linear operator,

$$M||\mathbf{y}-\mathbf{x}|| \geq ||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||\,||\mathbf{y}-\mathbf{x}|| \geq ||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})|| \geq (M+\varepsilon)||\mathbf{y}-\mathbf{x}||,$$

a contradiction. Therefore, $t = 1$ and so

$$||\mathbf{f}(\mathbf{x}+(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})|| \leq (M+\varepsilon)||\mathbf{y}-\mathbf{x}||.$$

Since $\varepsilon > 0$ is arbitrary, this proves the theorem.

The next theorem proves that if the partial derivatives exist and are continuous, then the function is differentiable.

Theorem 3.55 *Let $\mathbf{g}, U$ and Z be given as in Definition 3.53. Then $\mathbf{g}$ is $C^1(U)$ if and only if $D_1\mathbf{g}$ and $D_2\mathbf{g}$ both exist and are continuous on U. In this case,*

$$D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v}) = D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}.$$

Proof: Suppose first that $\mathbf{g} \in C^1(U)$. Then if $(\mathbf{x},\mathbf{y}) \in U$,

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-\mathbf{g}(\mathbf{x},\mathbf{y}) = D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0})+o(\mathbf{u}).$$

Therefore, $D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u} = D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0})$. Then

$$||(D_1\mathbf{g}(\mathbf{x},\mathbf{y})-D_1\mathbf{g}(\mathbf{x}',\mathbf{y}'))(\mathbf{u})|| =$$

$$||(D\mathbf{g}(\mathbf{x},\mathbf{y})-D\mathbf{g}(\mathbf{x}',\mathbf{y}'))(\mathbf{u},\mathbf{0})|| \leq$$

$$||D\mathbf{g}(\mathbf{x},\mathbf{y}) - D\mathbf{g}(\mathbf{x}',\mathbf{y}')||\,||(\mathbf{u},\mathbf{0})||.$$

Therefore,

$$||D_1\mathbf{g}(\mathbf{x},\mathbf{y}) - D_1\mathbf{g}(\mathbf{x}',\mathbf{y}')|| \leq ||D\mathbf{g}(\mathbf{x},\mathbf{y}) - D\mathbf{g}(\mathbf{x}',\mathbf{y}')||.$$

A similar argument applies for $D_2\mathbf{g}$ and this proves the continuity of the function, $(\mathbf{x},\mathbf{y}) \to D_i\mathbf{g}(\mathbf{x},\mathbf{y})$ for $i = 1, 2$. The formula follows from

$$\begin{aligned} D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v}) &= D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0}) + D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{0},\mathbf{v}) \\ &\equiv D_1\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{v}. \end{aligned}$$

Now suppose $D_1\mathbf{g}(\mathbf{x},\mathbf{y})$ and $D_2\mathbf{g}(\mathbf{x},\mathbf{y})$ exist and are continuous.

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})$$

$$+\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y})$$

$$= \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}) - \mathbf{g}(\mathbf{x},\mathbf{y}) + \mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}) +$$

$$[\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}) - (\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}))]$$

$$= D_1\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{u} + D_2\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{v} + o(\mathbf{v}) + o(\mathbf{u}) +$$

$$[\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}) - (\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}))]. \tag{3.26}$$

Let $\mathbf{h}(\mathbf{x},\mathbf{u}) \equiv \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})$. Then the expression in [ ] is of the form,

$$\mathbf{h}(\mathbf{x},\mathbf{u}) - \mathbf{h}(\mathbf{x},\mathbf{0}).$$

Also

$$D_2\mathbf{h}(\mathbf{x},\mathbf{u}) = D_1\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - D_1\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})$$

and so, by continuity of $(\mathbf{x},\mathbf{y}) \to D_1\mathbf{g}(\mathbf{x},\mathbf{y})$,

$$||D_2\mathbf{h}(\mathbf{x},\mathbf{u})|| < \varepsilon$$

whenever $||(\mathbf{u},\mathbf{v})||$ is small enough. By Theorem 3.54, there exists $\delta > 0$ such that if $||(\mathbf{u},\mathbf{v})|| < \delta$, the norm of the last term in (3.26) satisfies the inequality,

$$||\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}) - (\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v}) - \mathbf{g}(\mathbf{x},\mathbf{y}))|| < \varepsilon\,||\mathbf{u}||. \tag{3.27}$$

Therefore, this term is $o((\mathbf{u},\mathbf{v}))$. It follows from (3.27) and (3.26) that

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v}) =$$

$$\mathbf{g}(\mathbf{x},\mathbf{y}) + D_1\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{u} + D_2\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{v}+o(\mathbf{u}) + o(\mathbf{v}) + o((\mathbf{u},\mathbf{v}))$$

$$= \mathbf{g}(\mathbf{x},\mathbf{y}) + D_1\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{u} + D_2\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{v} + o((\mathbf{u},\mathbf{v}))$$

Showing that $D\mathbf{g}(\mathbf{x},\mathbf{y})$ exists and is given by

$$D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v}) = D_1\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{u} + D_2\mathbf{g}(\mathbf{x},\mathbf{y})\,\mathbf{v}.$$

The continuity of $(\mathbf{x},\mathbf{y}) \to D\mathbf{g}(\mathbf{x},\mathbf{y})$ follows from the continuity of $(\mathbf{x},\mathbf{y}) \to D_i\mathbf{g}(\mathbf{x},\mathbf{y})$. This proves the theorem.

Not surprisingly, it can be generalized to many more factors.

Definition 3.56 *For* $\mathbf{x} \in \prod_{i=1}^{n} \mathbb{F}^{r_i}$ *by*

$$||\mathbf{x}|| \equiv \max \{||\mathbf{x}_i||_i : i = 1, \cdots, n\}.$$

Now let $\mathbf{g} : U \subseteq \prod_{i=1}^{n} \mathbb{F}^{r_i} \to \mathbb{F}^q$, *where* U *is an open set. Then the map* $\mathbf{x}_i \to \mathbf{g}(\mathbf{x})$ *is a function from the open set in* $\mathbb{F}^{r_i}$,

$$\{\mathbf{x}_i : \mathbf{x} \in U\}$$

to $\mathbb{F}^q$. *When this map is differentiable, its derivative is denoted by* $D_i\mathbf{g}(\mathbf{x})$. *To aid in the notation, for* $\mathbf{v} \in X_i$, *let* $\theta_i \mathbf{v} \in \prod_{i=1}^{n} \mathbb{F}^{r_i}$ *be the vector* $(\mathbf{0}, \cdots, \mathbf{v}, \cdots, \mathbf{0})$ *where the* $\mathbf{v}$ *is in the* i^{th} *slot and for* $\mathbf{v} \in \prod_{i=1}^{n} \mathbb{F}^{r_i}$, *let* $\mathbf{v}_i$ *denote the entry in the* i^{th} *slot of* $\mathbf{v}$. *Thus by saying* $\mathbf{x}_i \to \mathbf{g}(\mathbf{x})$ *is differentiable is meant that for* $\mathbf{v} \in X_i$ *sufficiently small,*

$$\mathbf{g}(\mathbf{x} + \theta_i \mathbf{v}) - \mathbf{g}(\mathbf{x}) = D_i \mathbf{g}(\mathbf{x}) \mathbf{v} + o(\mathbf{v}).$$

Here is a generalization of Theorem 3.55.

Theorem 3.57 *Let* $\mathbf{g}, U, \prod_{i=1}^{n} \mathbb{F}^{r_i}$, *be given as in Definition 3.56. Then* $\mathbf{g}$ *is* $C^1(U)$ *if and only if* $D_i\mathbf{g}$ *exists and is continuous on* U *for each* i. *In this case,*

$$D\mathbf{g}(\mathbf{x})(\mathbf{v}) = \sum_k D_k \mathbf{g}(\mathbf{x}) \mathbf{v}_k \tag{3.28}$$

Proof: The only if part of the proof is left for you. Suppose then that $D_i\mathbf{g}$ exists and is continuous for each i. Note that $\sum_{j=1}^{k} \theta_j \mathbf{v}_j = (\mathbf{v}_1, \cdots, \mathbf{v}_k, \mathbf{0}, \cdots, \mathbf{0})$. Thus $\sum_{j=1}^{n} \theta_j \mathbf{v}_j = \mathbf{v}$ and define $\sum_{j=1}^{0} \theta_j \mathbf{v}_j \equiv \mathbf{0}$. Therefore,

$$\mathbf{g}(\mathbf{x} + \mathbf{v}) - \mathbf{g}(\mathbf{x}) = \sum_{k=1}^{n} \left[\mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k} \theta_j \mathbf{v}_j\right) - \mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k-1} \theta_j \mathbf{v}_j\right) \right] \tag{3.29}$$

Consider the terms in this sum.

$$\mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k} \theta_j \mathbf{v}_j\right) - \mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k-1} \theta_j \mathbf{v}_j\right) = \mathbf{g}(\mathbf{x} + \theta_k \mathbf{v}_k) - \mathbf{g}(\mathbf{x}) + \tag{3.30}$$

$$\left(\mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k} \theta_j \mathbf{v}_j\right) - \mathbf{g}(\mathbf{x} + \theta_k \mathbf{v}_k) \right) - \left(\mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k-1} \theta_j \mathbf{v}_j\right) - \mathbf{g}(\mathbf{x}) \right) \tag{3.31}$$

and the expression in (3.31) is of the form $\mathbf{h}(\mathbf{v}_k) - \mathbf{h}(\mathbf{0})$ where for small $\mathbf{w} \in X_k$,

$$\mathbf{h}(\mathbf{w}) \equiv \mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k-1} \theta_j \mathbf{v}_j + \theta_k \mathbf{w}\right) - \mathbf{g}(\mathbf{x} + \theta_k \mathbf{w}).$$

Therefore,

$$D\mathbf{h}(\mathbf{w}) = D_k \mathbf{g}\left(\mathbf{x} + \sum_{j=1}^{k-1} \theta_j \mathbf{v}_j + \theta_k \mathbf{w}\right) - D_k \mathbf{g}(\mathbf{x} + \theta_k \mathbf{w})$$

and by continuity, $||D\mathbf{h}(\mathbf{w})|| < \varepsilon$ provided $||\mathbf{v}||$ is small enough. Therefore, by Theorem 3.54, whenever $||\mathbf{v}||$ is small enough, $||\mathbf{h}(\theta_k \mathbf{v}_k) - \mathbf{h}(\mathbf{0})|| \leq \varepsilon ||\theta_k \mathbf{v}_k|| \leq \varepsilon ||\mathbf{v}||$ which shows that since ε is arbitrary, the expression in (3.31) is $o(\mathbf{v})$. Now in (3.30) $\mathbf{g}(\mathbf{x} + \theta_k \mathbf{v}_k) - \mathbf{g}(\mathbf{x}) = D_k \mathbf{g}(\mathbf{x}) \mathbf{v}_k + o(\mathbf{v}_k) = D_k \mathbf{g}(\mathbf{x}) \mathbf{v}_k + o(\mathbf{v})$. Therefore, referring to (3.29),

$$\mathbf{g}(\mathbf{x} + \mathbf{v}) - \mathbf{g}(\mathbf{x}) = \sum_{k=1}^{n} D_k \mathbf{g}(\mathbf{x}) \mathbf{v}_k + o(\mathbf{v})$$

which shows $D\mathbf{g}$ exists and equals the formula given in (3.28).

3.12 Higher Order Derivatives

If $f : U \subseteq \mathbb{F}^n \to \mathbb{F}^q$, then

$$\mathbf{x} \to D\mathbf{f}(\mathbf{x})$$

is a mapping from U to $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^q)$, a normed linear space.

Definition 3.58 *The following is the definition of the second derivative.*

$$D^2\mathbf{f}(\mathbf{x}) \equiv D(D\mathbf{f}(\mathbf{x})).$$

Thus,

$$D\mathbf{f}(\mathbf{x}+\mathbf{v}) - D\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})\mathbf{v} + o(\mathbf{v}).$$

This implies

$$D^2\mathbf{f}(\mathbf{x}) \in \mathcal{L}(\mathbb{F}^n, \mathcal{L}(\mathbb{F}^n, \mathbb{F}^q)),\ D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v}) \in \mathbb{F}^q,$$

and the map

$$(\mathbf{u}, \mathbf{v}) \to D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})$$

is a bilinear map having values in $\mathbb{F}^q$. The same pattern applies to taking higher order derivatives. Thus,

$$D^3\mathbf{f}(\mathbf{x}) \equiv D\left(D^2\mathbf{f}(\mathbf{x})\right)$$

and you can consider $D^3\mathbf{f}(\mathbf{x})$ as a trilinear map. Also, instead of writing

$$D^2 f(\mathbf{x})(\mathbf{u})(\mathbf{v}),$$

it is customary to write

$$D^2 f(\mathbf{x})(\mathbf{u}, \mathbf{v}).$$

$\mathbf{f}$ is said to be $C^k(U)$ if $\mathbf{f}$ and its first k derivatives are all continuous. For example, for $\mathbf{f}$ to be $C^2(U)$,

$$\mathbf{x} \to D^2\mathbf{f}(\mathbf{x})$$

would have to be continuous as a map from U to $\mathcal{L}(\mathbb{F}^n, \mathcal{L}(\mathbb{F}^n, \mathbb{F}^q))$. The following theorem deals with the question of symmetry of the map $D^2\mathbf{f}$ in the case where f is a real valued function.

Theorem 3.59 *Let U be an open subset of $\mathbb{F}^n$ and suppose $f : U \subseteq \mathbb{F}^n \to \mathbb{R}$ and $D^2 f(\mathbf{x})$ exists for all $\mathbf{x} \in U$ and $D^2 f$ is continuous at $\mathbf{x} \in U$. Then*

$$D^2 f(\mathbf{x})(\mathbf{u})(\mathbf{v}) = D^2 f(\mathbf{x})(\mathbf{v})(\mathbf{u}).$$

Proof: Let $B(\mathbf{x}, r) \subseteq U$ and let $t, s \in (0, r/2]$. Now define

$$\Delta(s,t) \equiv \frac{1}{st}\{\overbrace{f(\mathbf{x}+t\mathbf{u}+s\mathbf{v}) - f(\mathbf{x}+t\mathbf{u})}^{h(t)} - \overbrace{(f(\mathbf{x}+s\mathbf{v}) - f(\mathbf{x}))}^{h(0)}\}. \tag{3.32}$$

Let $h(t) = f(\mathbf{x}+s\mathbf{v}+t\mathbf{u}) - f(\mathbf{x}+t\mathbf{u})$. Therefore, by the mean value theorem from calculus,

$$\begin{aligned} \Delta(s,t) &= \frac{1}{st}(h(t) - h(0)) = \frac{1}{st}h'(\alpha t)\,t \\ &= \frac{1}{s}\left(Df(\mathbf{x}+s\mathbf{v}+\alpha t\mathbf{u})\,\mathbf{u} - Df(\mathbf{x}+\alpha t\mathbf{u})\,\mathbf{u}\right). \end{aligned}$$

Applying the mean value theorem again,

$$\Delta(s,t) = D^2 f(\mathbf{x}+\beta s\mathbf{v}+\alpha t\mathbf{u})(\mathbf{v})(\mathbf{u})$$

where $\alpha, \beta \in (0,1)$. If the terms $f(\mathbf{x}+t\mathbf{u})$ and $f(\mathbf{x}+s\mathbf{v})$ are interchanged in (3.32), $\Delta(s,t)$ is also unchanged and the above argument shows there exist $\gamma, \delta \in (0,1)$ such that

$$\Delta(s,t) = D^2 f(\mathbf{x}+\gamma s\mathbf{v}+\delta t\mathbf{u})(\mathbf{u})(\mathbf{v}).$$

Letting $(s,t) \to (0,0)$ and using the continuity of $D^2 f$ at $\mathbf{x}$,

$$\lim_{(s,t)\to(0,0)} \Delta(s,t) = D^2 f(\mathbf{x})(\mathbf{u})(\mathbf{v}) = D^2 f(\mathbf{x})(\mathbf{v})(\mathbf{u}).$$

Corollary 3.60 *Let U be an open subset of $\mathbb{F}^n$ and let $f : U \to \mathbb{R}$ be a function in $C^2(U)$. Then all mixed partial derivatives are equal.*

Proof: If $\mathbf{e}_i$ are the standard basis vectors, what is

$$D^2 f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j)?$$

To see what this is, use the definition to write

$$D^2 f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j) = t^{-1}s^{-1}D^2 f(\mathbf{x})(t\mathbf{e}_i)(s\mathbf{e}_j)$$

$$= t^{-1}s^{-1}(Df(\mathbf{x}+t\mathbf{e}_i) - Df(\mathbf{x}) + o(t))(s\mathbf{e}_j)$$

$$= t^{-1}s^{-1}(f(\mathbf{x}+t\mathbf{e}_i + s\mathbf{e}_j) - f(\mathbf{x}+t\mathbf{e}_i)$$

$$+o(s) - (f(\mathbf{x}+s\mathbf{e}_j) - f(\mathbf{x}) + o(s)) + o(t)s).$$

First let $s \to 0$ to get

$$t^{-1}\left(\frac{\partial f}{\partial x_j}(\mathbf{x}+t\mathbf{e}_i) - \frac{\partial f}{\partial x_j}(\mathbf{x}) + o(t)\right)$$

and then let $t \to 0$ to obtain

$$D^2 f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \tag{3.33}$$

Thus Theorem 3.59 asserts that the mixed partial derivatives are equal at $\mathbf{x}$ if they are defined near $\mathbf{x}$ and continuous at $\mathbf{x}$.

This theorem about equality of mixed partial derivatives turns out to be very important in paritial differential equations. It was first proved by Euler in the early to mid 1700's.

3.13 Implicit Function Theorem

The implicit function theorem is one of the greatest theorems in mathematics. To prove it, here is a useful lemma.

Lemma 3.61 *Let $A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ and suppose $||A|| \leq r < 1$. Then*

$$(I - A)^{-1} \text{ exists} \tag{3.34}$$

and

$$\left|\left|(I - A)^{-1}\right|\right| \leq (1 - r)^{-1}. \tag{3.35}$$

Furthermore, if

$$\mathcal{I} \equiv \left\{A \in \mathcal{L}(X, X) : A^{-1} \text{ exists}\right\}$$

the map $A \to A^{-1}$ is continuous on $\mathcal{I}$ and $\mathcal{I}$ is an open subset of $\mathcal{L}(X, X)$.

Proof: Consider $B_k \equiv \sum_{i=0}^{k} A^i$. Then if $N < l < k$,

$$||B_k - B_l|| \leq \sum_{i=N}^{k} \left|\left|A^i\right|\right| \leq \sum_{i=N}^{k} ||A||^i \leq \frac{r^N}{1 - r}.$$

It follows B_k is a Cauchy sequence and so it converges to $B \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$. Also,

$$(I - A) B_k = \sum_{i=0}^{k} A^i - \sum_{i=1}^{k+1} A^i = I - A^{k+1}$$

and similarly, $I - A^{k+1} = B_k (I - A)$ and so

$$I = \lim_{k \to \infty} (I - A) B_k = (I - A) B, \ I = \lim_{k \to \infty} B_k (I - A) = B (I - A).$$

Thus from the definition of the inverse, $(I - A)^{-1} = B = \sum_{i=0}^{\infty} A^i$. It follows

$$\left|\left|(I - A)^{-1}\right|\right| \leq \sum_{i=1}^{\infty} \left|\left|A^i\right|\right| \leq \sum_{i=0}^{\infty} ||A||^i = \frac{1}{1 - r}.$$

To verify the continuity of the inverse map, let $A \in \mathcal{I}$. Then

$$B = A\left(I - A^{-1}(A - B)\right)$$

and so if $\left|\left|A^{-1}(A - B)\right|\right| < 1$ it follows $B^{-1} = \left(I - A^{-1}(A - B)\right)^{-1} A^{-1}$ which shows $\mathcal{I}$ is open. Now for such B this close to A,

$$\left|\left|B^{-1} - A^{-1}\right|\right| = \left|\left|\left(I - A^{-1}(A - B)\right)^{-1} A^{-1} - A^{-1}\right|\right|$$

$$= \left|\left|\left(\left(I - A^{-1}(A - B)\right)^{-1} - I\right) A^{-1}\right|\right|$$

$$= \left|\left|\sum_{k=1}^{\infty} \left(A^{-1}(A - B)\right)^k A^{-1}\right|\right| \leq \sum_{k=1}^{\infty} \left|\left|A^{-1}(A - B)\right|\right|^k \left|\left|A^{-1}\right|\right|$$

$$= \frac{\left|\left|A^{-1}(A - B)\right|\right|}{1 - \left|\left|A^{-1}(A - B)\right|\right|} \left|\left|A^{-1}\right|\right|$$

which shows that if $||A - B||$ is small, so is $\left|\left|B^{-1} - A^{-1}\right|\right|$. This proves the lemma.

The next theorem is a very useful result in many areas. It will be used in this section to give a short proof of the implicit function theorem but it is also useful in studying differential equations and integral equations. It is sometimes called the uniform contraction principle.

Theorem 3.62 *Let (Y,ρ) and (X,d) be complete metric spaces and suppose for each $(x,y) \in X \times Y$, $T(x,y) \in X$ and satisfies*

$$d(T(x,y), T(x',y)) \leq rd(x,x') \tag{3.36}$$

where $0 < r < 1$ and also

$$d(T(x,y), T(x,y')) \leq M\rho(y,y'). \tag{3.37}$$

Then for each $y \in Y$ there exists a unique "fixed point" for $T(\cdot,y)$, $x \in X$, satisfying

$$T(x,y) = x \tag{3.38}$$

and also if $x(y)$ is this fixed point,

$$d(x(y), x(y')) \leq \frac{M}{1-r}\rho(y,y'). \tag{3.39}$$

Proof: First I show there exists a fixed point for the mapping, $T(\cdot,y)$. For a fixed y, let $g(x) \equiv T(x,y)$. Now pick any $x_0 \in X$ and consider the sequence,

$$x_1 = g(x_0),\ x_{k+1} = g(x_k).$$

Then by (3.36),

$$d(x_{k+1}, x_k) = d(g(x_k), g(x_{k-1})) \leq rd(x_k, x_{k-1}) \leq$$

$$r^2 d(x_{k-1}, x_{k-2}) \leq \cdots \leq r^k d(g(x_0), x_0).$$

Now by the triangle inequality,

$$d(x_{k+p}, x_k) \leq \sum_{i=1}^{p} d(x_{k+i}, x_{k+i-1})$$

$$\leq \sum_{i=1}^{p} r^{k+i-1} d(x_0, g(x_0)) \leq \frac{r^k d(x_0, g(x_0))}{1-r}.$$

Since $0 < r < 1$, this shows that $\{x_k\}_{k=1}^{\infty}$ is a Cauchy sequence. Therefore, it converges to a point in X, x. To see x is a fixed point,

$$x = \lim_{k\to\infty} x_k = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g(x).$$

This proves (3.38). To verify (3.39),

$$d(x(y), x(y')) = d(T(x(y), y), T(x(y'), y')) \leq$$

$$d(T(x(y), y), T(x(y), y')) + d(T(x(y), y'), T(x(y'), y'))$$

$$\leq M\rho(y,y') + rd(x(y), x(y')).$$

Thus $(1-r)\, d(x(y), x(y')) \leq M\rho(y,y')$. This also shows the fixed point for a given y is unique. This proves the theorem.

The implicit function theorem is one of the most important results in Analysis. It provides the theoretical justification for such procedures as implicit differentiation taught in Calculus courses and has surprising consequences in many other areas. It deals with the question of solving, $\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ for $\mathbf{x}$ in terms of $\mathbf{y}$ and how smooth the solution is. The proof I will give below will apply with no change to much more general situations.

Theorem 3.63 *(implicit function theorem) Suppose U is an open set in $\mathbb{F}^n \times \mathbb{F}^m$. Let $\mathbf{f} : U \to \mathbb{F}^n$ be in $C^1(U)$ and suppose*

$$\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0},\ D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n). \tag{3.40}$$

Then there exist positive constants, δ, η, such that for every $\mathbf{y} \in B(\mathbf{y}_0, \eta)$ there exists a unique $\mathbf{x}(\mathbf{y}) \in B(\mathbf{x}_0, \delta)$ such that

$$\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y}) = \mathbf{0}. \tag{3.41}$$

Furthermore, the mapping, $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ is in $C^1(B(\mathbf{y}_0, \eta))$.

Proof: Let $\mathbf{T}(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x} - D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} \mathbf{f}(\mathbf{x}, \mathbf{y})$. Therefore,

$$D_1\mathbf{T}(\mathbf{x}, \mathbf{y}) = I - D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} D_1\mathbf{f}(\mathbf{x}, \mathbf{y}). \tag{3.42}$$

by continuity of the derivative and Theorem 3.55, it follows that there exists $\delta > 0$ such that if $||(\mathbf{x} - \mathbf{x}_0, \mathbf{y} - \mathbf{y}_0)|| < \delta$, then

$$||D_1\mathbf{T}(\mathbf{x}, \mathbf{y})|| < \frac{1}{2},$$

$$\left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}\right|\right| ||D_2\mathbf{f}(\mathbf{x}, \mathbf{y})|| < M \tag{3.43}$$

where $M > \left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}\right|\right| ||D_2\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)||$. By Theorem 3.54, whenever $\mathbf{x}, \mathbf{x}' \in B(\mathbf{x}_0, \delta)$ and $\mathbf{y} \in B(\mathbf{y}_0, \delta)$,

$$||\mathbf{T}(\mathbf{x}, \mathbf{y}) - \mathbf{T}(\mathbf{x}', \mathbf{y})|| \leq \frac{1}{2} ||\mathbf{x} - \mathbf{x}'||. \tag{3.44}$$

Solving (3.42) for $D_1\mathbf{f}(\mathbf{x}, \mathbf{y})$,

$$D_1\mathbf{f}(\mathbf{x}, \mathbf{y}) = D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)(I - D_1\mathbf{T}(\mathbf{x}, \mathbf{y})).$$

By Lemma 3.61 and (3.43), $D_1\mathbf{f}(\mathbf{x}, \mathbf{y})^{-1}$ exists and

$$\left|\left|D_1\mathbf{f}(\mathbf{x}, \mathbf{y})^{-1}\right|\right| \leq 2\left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}\right|\right|. \tag{3.45}$$

Now restrict $\mathbf{y}$ some more. Let $0 < \eta < \min\left(\delta, \frac{\delta}{3M}\right)$. Then suppose $\mathbf{x} \in \overline{B(\mathbf{x}_0, \delta)}$ and $\mathbf{y} \in \overline{B(\mathbf{y}_0, \eta)}$. Consider $\mathbf{T}(\mathbf{x}, \mathbf{y}) - \mathbf{x}_0 \equiv \mathbf{g}(\mathbf{x}, \mathbf{y})$.

$$D_1\mathbf{g}(\mathbf{x}, \mathbf{y}) = I - D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} D_1\mathbf{f}(\mathbf{x}, \mathbf{y}) = D_1\mathbf{T}(\mathbf{x}, \mathbf{y}),$$

and

$$D_2\mathbf{g}(\mathbf{x}, \mathbf{y}) = -D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} D_2\mathbf{f}(\mathbf{x}, \mathbf{y}).$$

Thus by (3.43), (3.40) saying that $\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$, and Theorems 3.54 and (3.26), it follows that for such $(\mathbf{x}, \mathbf{y})$,

$$||\mathbf{T}(\mathbf{x}, \mathbf{y}) - \mathbf{x}_0|| = ||\mathbf{g}(\mathbf{x}, \mathbf{y})|| = ||\mathbf{g}(\mathbf{x}, \mathbf{y}) - \mathbf{g}(\mathbf{x}_0, \mathbf{y}_0)||$$

$$\leq \frac{1}{2} ||\mathbf{x} - \mathbf{x}_0|| + M ||\mathbf{y} - \mathbf{y}_0|| < \frac{\delta}{2} + \frac{\delta}{3} = \frac{5\delta}{6} < \delta. \tag{3.46}$$

Also for such $(\mathbf{x}, \mathbf{y}_i), i = 1, 2$, Theorem 3.54 and (3.43) yields

$$||\mathbf{T}(\mathbf{x}, \mathbf{y}_1) - \mathbf{T}(\mathbf{x}, \mathbf{y}_2)|| = \left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}(\mathbf{f}(\mathbf{x}, \mathbf{y}_2) - \mathbf{f}(\mathbf{x}, \mathbf{y}_1))\right|\right|$$

$$\leq M ||\mathbf{y}_2 - \mathbf{y}_1|| . \tag{3.47}$$

From now on assume $||\mathbf{x} - \mathbf{x}_0|| < \delta$ and $||\mathbf{y} - \mathbf{y}_0|| < \eta$ so that (3.47), (3.45), (3.46), (3.44), and (3.43) all hold. By (3.47), (3.44), (3.46), and the uniform contraction principle, Theorem 3.62 applied to $X \equiv \overline{B\left(\mathbf{x}_0, \frac{5\delta}{6}\right)}$ and $Y \equiv \overline{B(\mathbf{y}_0, \eta)}$ implies that for each $\mathbf{y} \in B(\mathbf{y}_0, \eta)$, there exists a unique $\mathbf{x}(\mathbf{y}) \in B(\mathbf{x}_0, \delta)$ (actually in $\overline{B\left(\mathbf{x}_0, \frac{5\delta}{6}\right)}$) such that $\mathbf{T}(\mathbf{x}(\mathbf{y}), \mathbf{y}) = \mathbf{x}(\mathbf{y})$ which is equivalent to

$$\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y}) = \mathbf{0}.$$

Furthermore,

$$||\mathbf{x}(\mathbf{y}) - \mathbf{x}(\mathbf{y}')|| \leq 2M ||\mathbf{y} - \mathbf{y}'|| . \tag{3.48}$$

This proves the implicit function theorem except for the verification that $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ is C^1. This is shown next. Letting $\mathbf{v}$ be sufficiently small, Theorem 3.55 and Theorem 3.54 imply

$$\mathbf{0} = \mathbf{f}(\mathbf{x}(\mathbf{y} + \mathbf{v}), \mathbf{y} + \mathbf{v}) - \mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y}) =$$

$$D_1\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y})(\mathbf{x}(\mathbf{y} + \mathbf{v}) - \mathbf{x}(\mathbf{y})) +$$

$$+D_2\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y})\mathbf{v} + o((\mathbf{x}(\mathbf{y} + \mathbf{v}) - \mathbf{x}(\mathbf{y}), \mathbf{v})) .$$

The last term in the above is $o(\mathbf{v})$ because of (3.48). Therefore, by (3.45), it is possible to solve the above equation for $\mathbf{x}(\mathbf{y} + \mathbf{v}) - \mathbf{x}(\mathbf{y})$ and obtain

$$\mathbf{x}(\mathbf{y} + \mathbf{v}) - \mathbf{x}(\mathbf{y}) = -D_1(\mathbf{x}(\mathbf{y}), \mathbf{y})^{-1} D_2\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y})\mathbf{v} + o(\mathbf{v})$$

Which shows that $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ is differentiable on $B(\mathbf{y}_0, \eta)$ and

$$D\mathbf{x}(\mathbf{y}) = -D_1(\mathbf{x}(\mathbf{y}), \mathbf{y})^{-1} D_2\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y}) .$$

Now it follows from the continuity of $D_2\mathbf{f}$, $D_1\mathbf{f}$, the inverse map, (3.48), and this formula for $D\mathbf{x}(\mathbf{y})$that $\mathbf{x}(\cdot)$ is $C^1(B(\mathbf{y}_0, \eta))$. This proves the theorem.

In practice, how do you verify the condition, $D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$?

$$\mathbf{f}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} f_1(x_1, \cdots, x_n, y_1, \cdots, y_n) \\ \vdots \\ f_n(x_1, \cdots, x_n, y_1, \cdots, y_n) \end{pmatrix}.$$

The matrix of the linear transformation, $D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)$ is then

$$\begin{pmatrix} \frac{\partial f_1(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_1} & \cdots & \frac{\partial f_1(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_1} & \cdots & \frac{\partial f_n(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_n} \end{pmatrix}$$

and from linear algebra, $D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)^{-1} \in \mathcal{L}(\mathbb{F}^n,\mathbb{F}^n)$ exactly when the above matrix has an inverse. In other words when

$$\det\begin{pmatrix} \frac{\partial f_1(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_1} & \cdots & \frac{\partial f_1(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_1} & \cdots & \frac{\partial f_n(x_1,\cdots,x_n,y_1,\cdots,y_n)}{\partial x_n} \end{pmatrix} \neq 0$$

at $(\mathbf{x}_0,\mathbf{y}_0)$. The above determinant is important enough that it is given special notation. Letting $\mathbf{z} = \mathbf{f}(\mathbf{x},\mathbf{y})$, the above determinant is often written as

$$\frac{\partial(z_1,\cdot\cdot\cdot,z_n)}{\partial(x_1,\cdot\cdot\cdot,x_n)}.$$

The next theorem is a very important special case of the implicit function theorem known as the inverse function theorem. Actually one can also obtain the implicit function theorem from the inverse function theorem. It is done this way in [14] and in [2].

Theorem 3.64 *(inverse function theorem) Let $\mathbf{x}_0 \in U \subseteq \mathbb{F}^n$ and let $\mathbf{f} : U \to \mathbb{F}^n$. Suppose*

$$\mathbf{f} \text{ is } C^1(U)\ , \quad \text{and} \quad D\mathbf{f}(\mathbf{x}_0)^{-1} \in \mathcal{L}(\mathbb{F}^n,\mathbb{F}^n). \tag{3.49}$$

Then there exist open sets, W, and V such that

$$\mathbf{x}_0 \in W \subseteq U, \tag{3.50}$$

$$\mathbf{f} : W \to V \text{ is one to one and onto}, \tag{3.51}$$

$$\mathbf{f}^{-1} \text{ is } C^1, \tag{3.52}$$

Proof: Apply the implicit function theorem to the function

$$\mathbf{F}(\mathbf{x},\mathbf{y}) \equiv \mathbf{f}(\mathbf{x}) - \mathbf{y}$$

where $\mathbf{y}_0 \equiv \mathbf{f}(\mathbf{x}_0)$. Thus the function $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ defined in that theorem is $\mathbf{f}^{-1}$. Now let

$$W \equiv B(\mathbf{x}_0,\delta) \cap \mathbf{f}^{-1}(B(\mathbf{y}_0,\eta))$$

and

$$V \equiv B(\mathbf{y}_0,\eta).$$

This proves the theorem.

Lemma 3.65 *Let*

$$O \equiv \{A \in \mathcal{L}(\mathbb{F}^n,\mathbb{F}^n) : A^{-1} \in \mathcal{L}(\mathbb{F}^n,\mathbb{F}^n)\} \tag{3.53}$$

and let

$$\mathfrak{I} : O \to \mathcal{L}(\mathbb{F}^n,\mathbb{F}^n),\ \mathfrak{I}A \equiv A^{-1}. \tag{3.54}$$

Then O is open and $\mathfrak{I}$ is in $C^m(O)$ for all $m = 1, 2, \cdots$ Also

$$D\mathfrak{I}(A)(B) = -\mathfrak{I}(A)(B)\mathfrak{I}(A). \tag{3.55}$$

Proof: Let $A \in O$ and let $B \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ with

$$||B|| \leq \frac{1}{2} \left|\left|A^{-1}\right|\right|^{-1}.$$

Then

$$\left|\left|A^{-1}B\right|\right| \leq \left|\left|A^{-1}\right|\right| ||B|| \leq \frac{1}{2}$$

and so by Lemma 3.61,

$$\left(I + A^{-1}B\right)^{-1} \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n).$$

Now

$$\begin{aligned} (A+B)\left(I + A^{-1}B\right)^{-1} A^{-1} &= A\left(I + A^{-1}B\right)\left(I + A^{-1}B\right)^{-1} A^{-1} \\ &= AA^{-1} = \text{id}, \end{aligned}$$

the identity map on X. Therefore,

$$(A+B)^{-1} = \left(I + A^{-1}B\right)^{-1} A^{-1} =$$

$$\sum_{n=0}^{\infty} (-1)^n \left(A^{-1}B\right)^n A^{-1} = \left[I - A^{-1}B + o(B)\right] A^{-1}$$

which shows that O is open and also,

$$\begin{aligned} \mathfrak{I}(A+B) - \mathfrak{I}(A) &= \sum_{n=0}^{\infty} (-1)^n \left(A^{-1}B\right)^n A^{-1} - A^{-1} \\ &= -A^{-1}BA^{-1} + o(B) \\ &= -\mathfrak{I}(A)(B)\mathfrak{I}(A) + o(B). \end{aligned}$$

This follows when you observe the higher order terms in B are $o(B)$. This shows (3.55). It follows from this that we can continue taking derivatives of $\mathfrak{I}$. For $||B_1||$ small,

$$-\left[D\mathfrak{I}(A+B_1)(B) - D\mathfrak{I}(A)(B)\right]$$

$$= \mathfrak{I}(A+B_1)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A)$$

$$\begin{aligned} = \; & \mathfrak{I}(A+B_1)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A+B_1) + \\ & \mathfrak{I}(A)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A) \end{aligned}$$

$$= \left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right](B)\mathfrak{I}(A+B_1)$$

$$+\mathfrak{I}(A)(B)\left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right]$$

$$\begin{aligned} = \; & \left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right](B)\left[A^{-1} - A^{-1}B_1A^{-1}\right] + \\ & \mathfrak{I}(A)(B)\left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right] \end{aligned}$$

$$= \mathfrak{I}(A)(B_1)\mathfrak{I}(A)(B)\mathfrak{I}(A) + \mathfrak{I}(A)(B)\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)$$

and so

$$D^2\mathfrak{I}(A)(B_1)(B) =$$

$$\mathfrak{I}(A)(B_1)\mathfrak{I}(A)(B)\mathfrak{I}(A) + \mathfrak{I}(A)(B)\mathfrak{I}(A)(B_1)\mathfrak{I}(A)$$

which shows $\mathfrak{I}$ is $C^2(O)$. Clearly we can continue in this way, which shows $\mathfrak{I}$ is in $C^m(O)$ for all $m = 1, 2, \cdots$

Corollary 3.66 *In the inverse or implicit function theorems, assume*

$$\mathbf{f} \in C^m(U), m \geq 1.$$

Then

$$\mathbf{f}^{-1} \in C^m(V)$$

in the case of the inverse function theorem. In the implicit function theorem, the function

$$\mathbf{y} \to \mathbf{x}(\mathbf{y})$$

is C^m.

Proof: We consider the case of the inverse function theorem.

$$D\mathbf{f}^{-1}(\mathbf{y}) = \mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right).$$

Now by Lemma 3.65, and the chain rule,

$$D^2\mathbf{f}^{-1}(\mathbf{y})(B) =$$

$$-\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)(B)\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)D^2\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)D\mathbf{f}^{-1}(\mathbf{y})$$

$$= -\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)(B)\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)\cdot$$

$$D^2\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right).$$

Continuing in this way we see that it is possible to continue taking derivatives up to order m. Similar reasoning applies in the case of the implicit function theorem. This proves the corollary.

3.13.1 The Method Of Lagrange Multipliers

As an application of the implicit function theorem, we consider the method of Lagrange multipliers from calculus. Recall the problem is to maximize or minimize a function subject to equality constraints. Let $f : U \to \mathbb{R}$ be a C^1 function and let

$$g_i(\mathbf{x}) = 0, \ i = 1, \cdots, m \tag{3.56}$$

be a collection of equality constraints with $m < n$. Now consider the system of nonlinear equations

$$\begin{aligned} f(\mathbf{x}) &= a \\ g_i(\mathbf{x}) &= 0, \ i = 1, \cdots, m. \end{aligned}$$

We say $\mathbf{x}_0$ is a local maximum if $f(\mathbf{x}_0) \geq f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{x}_0$ which also satisfies the constraints (3.56). A local minimum is defined similarly. Let $\mathbf{F} : U \times \mathbb{R} \to \mathbb{R}^{m+1}$ be defined by

$$\mathbf{F}(\mathbf{x},a) \equiv \begin{pmatrix} f(\mathbf{x}) - a \\ g_1(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{pmatrix}. \tag{3.57}$$

Now consider the $m+1 \times n$ Jacobian matrix,

$$\begin{pmatrix} f_{x_1}(\mathbf{x}_0) & \cdots & f_{x_n}(\mathbf{x}_0) \\ g_{1x_1}(\mathbf{x}_0) & \cdots & g_{1x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ g_{mx_1}(\mathbf{x}_0) & \cdots & g_{mx_n}(\mathbf{x}_0) \end{pmatrix}.$$

If this matrix has rank $m+1$ then some $m+1 \times m+1$ submatrix has nonzero determinant. It follows from the implicit function theorem that we can select $m+1$ variables, $x_{i_1}, \cdots, x_{i_{m+1}}$ such that the system

$$\mathbf{F}(\mathbf{x},a) = \mathbf{0} \tag{3.58}$$

specifies these $m+1$ variables as a function of the remaining $n-(m+1)$ variables and a in an open set of $\mathbb{R}^{n-m}$. Thus there is a solution $(\mathbf{x},a)$ to (3.58) for some $\mathbf{x}$ close to $\mathbf{x}_0$ whenever a is in some open interval. Therefore, $\mathbf{x}_0$ cannot be either a local minimum or a local maximum. It follows that if $\mathbf{x}_0$ is either a local maximum or a local minimum, then the above matrix must have rank less than $m+1$ which requires the rows to be linearly dependent. Thus, there exist m scalars,

$$\lambda_1, \cdots, \lambda_m,$$

and a scalar μ, not all zero such that

$$\mu \begin{pmatrix} f_{x_1}(\mathbf{x}_0) \\ \vdots \\ f_{x_n}(\mathbf{x}_0) \end{pmatrix} = \lambda_1 \begin{pmatrix} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{pmatrix} + \cdots + \lambda_m \begin{pmatrix} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{pmatrix}. \tag{3.59}$$

If the column vectors

$$\begin{pmatrix} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{pmatrix}, \cdots \begin{pmatrix} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{pmatrix} \tag{3.60}$$

are linearly independent, then, $\mu \neq 0$ and dividing by μ yields an expression of the form

$$\begin{pmatrix} f_{x_1}(\mathbf{x}_0) \\ \vdots \\ f_{x_n}(\mathbf{x}_0) \end{pmatrix} = \lambda_1 \begin{pmatrix} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{pmatrix} + \cdots + \lambda_m \begin{pmatrix} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{pmatrix} \tag{3.61}$$

at every point $\mathbf{x}_0$ which is either a local maximum or a local minimum. This proves the following theorem.

Theorem 3.67 *Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a C^1 function. Then if $\mathbf{x}_0 \in U$ is either a local maximum or local minimum of f subject to the constraints (3.56), then (3.59) must hold for some scalars $\mu, \lambda_1, \cdots, \lambda_m$ not all equal to zero. If the vectors in (3.60) are linearly independent, it follows that an equation of the form (3.61) holds.*

3.14 Taylor's Formula

First we recall the Taylor formula with the Lagrange form of the remainder. Since we will only need this on a specific interval, we will state it for this interval. See any good calculus book for a proof of this theorem.

Theorem 3.68 *Let $h : (-\delta, 1+\delta) \to \mathbb{R}$ have $m+1$ derivatives. Then there exists $t \in [0,1]$ such that*

$$h(1) = h(0) + \sum_{k=1}^{m} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m+1)}(t)}{(m+1)!}.$$

Now let $f : U \to \mathbb{R}$ where $U \subseteq X$ a normed linear space and suppose $f \in C^m(U)$. Let $\mathbf{x} \in U$ and let $r > 0$ be such that

$$B(\mathbf{x},r) \subseteq U.$$

Then for $||\mathbf{v}|| < r$ we consider

$$f(\mathbf{x}+t\mathbf{v}) - f(\mathbf{x}) \equiv h(t)$$

for $t \in [0,1]$. Then

$$h'(t) = Df(\mathbf{x}+t\mathbf{v})(\mathbf{v}), \; h''(t) = D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{v})(\mathbf{v})$$

and continuing in this way, we see that

$$h^{(k)}(t) = D^{(k)} f(\mathbf{x}+t\mathbf{v})(\mathbf{v})(\mathbf{v}) \cdots (\mathbf{v}) \equiv D^{(k)} f(\mathbf{x}+t\mathbf{v})\mathbf{v}^k.$$

It follows from Taylor's formula for a function of one variable that

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \sum_{k=1}^{m} \frac{D^{(k)} f(\mathbf{x})\mathbf{v}^k}{k!} + \frac{D^{(m+1)} f(\mathbf{x}+t\mathbf{v})\mathbf{v}^{m+1}}{(m+1)!}. \tag{3.62}$$

This proves the following theorem.

Theorem 3.69 *Let $f : U \to \mathbb{R}$ and let $f \in C^{m+1}(U)$. Then if*

$$B(\mathbf{x},r) \subseteq U,$$

and $||\mathbf{v}|| < r$, there exists $t \in (0,1)$ such that (3.62) holds.

Now we consider the case where $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$ is $C^2(U)$. Then from Taylor's theorem, if $\mathbf{v}$ is small enough, there exists $t \in (0,1)$ such that

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + Df(\mathbf{x})\mathbf{v} + \frac{D^2 f(\mathbf{x}+t\mathbf{v})\mathbf{v}^2}{2}.$$

Letting

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i,$$

where $\mathbf{e}_i$ are the usual basis vectors, the second derivative term reduces to

$$\frac{1}{2}\sum_{i,j} D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{e}_i)(\mathbf{e}_j) v_i v_j = \frac{1}{2}\sum_{i,j} H_{ij}(\mathbf{x}+t\mathbf{v}) v_i v_j$$

where

$$H_{ij}(\mathbf{x}+t\mathbf{v}) = D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{e}_i)(\mathbf{e}_j) = \frac{\partial^2 f(\mathbf{x}+t\mathbf{v})}{\partial x_j \partial x_i}.$$

Definition 3.70 *The matrix whose ij^{th} entry is $\frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$ is called the Hessian matrix.*

From Theorem 3.59, this is a symmetric matrix. By the continuity of the second partial derivative,

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + Df(\mathbf{x})\mathbf{v} + \frac{1}{2}\mathbf{v}^T H(\mathbf{x})\mathbf{v} +$$

$$\frac{1}{2}\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right). \tag{3.63}$$

where the last two terms involve ordinary matrix multiplication and

$$\mathbf{v}^T = (v_1 \cdots v_n)$$

for v_i the components of $\mathbf{v}$ relative to the standard basis.

Theorem 3.71 *In the above situation, suppose $Df(\mathbf{x}) = 0$. Then if $H(\mathbf{x})$ has all positive eigenvalues, $\mathbf{x}$ is a local minimum. If $H(\mathbf{x})$ has all negative eigenvalues, then $\mathbf{x}$ is a local maximum. If $H(\mathbf{x})$ has a positive eigenvalue, then there exists a direction in which f has a local minimum at $\mathbf{x}$, while if $H(\mathbf{x})$ has a negative eigenvalue, there exists a direction in which $H(\mathbf{x})$ has a local maximum at $\mathbf{x}$.*

Proof: Since $Df(\mathbf{x}) = 0$, formula (3.63) holds and by continuity of the second derivative, we know $H(\mathbf{x})$ is a symmetric matrix. Thus $H(\mathbf{x})$ has all real eigenvalues. Suppose first that $H(\mathbf{x})$ has all positive eigenvalues and that all are larger than $\delta^2 > 0$. Then $H(\mathbf{x})$ has an orthonormal basis of eigenvectors, $\{\mathbf{v}_i\}_{i=1}^n$ and if $\mathbf{u}$ is an arbitrary vector, we can write $\mathbf{u} = \sum_{j=1}^n u_j \mathbf{v}_j$ where $u_j = \mathbf{u} \cdot \mathbf{v}_j$. Thus

$$\mathbf{u}^T H(\mathbf{x})\mathbf{u} = \sum_{j=1}^n u_j \mathbf{v}_j^T H(\mathbf{x}) \sum_{j=1}^n u_j \mathbf{v}_j$$

$$= \sum_{j=1}^n u_j^2 \lambda_j \geq \delta^2 \sum_{j=1}^n u_j^2 = \delta^2 |\mathbf{u}|^2.$$

From (3.63) and the continuity of H, if $\mathbf{v}$ is small enough,

$$f(\mathbf{x}+\mathbf{v}) \geq f(\mathbf{x}) + \frac{1}{2}\delta^2 |\mathbf{v}|^2 - \frac{1}{4}\delta^2 |\mathbf{v}|^2 = f(\mathbf{x}) + \frac{\delta^2}{4}|\mathbf{v}|^2.$$

This shows the first claim of the theorem. The second claim follows from similar reasoning. Suppose $H(\mathbf{x})$ has a positive eigenvalue λ^2. Then let $\mathbf{v}$ be an eigenvector for this eigenvalue. Then from (3.63),

$$f(\mathbf{x}+t\mathbf{v}) = f(\mathbf{x}) + \frac{1}{2}t^2 \mathbf{v}^T H(\mathbf{x})\mathbf{v} +$$

$$\frac{1}{2}t^2\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right)$$

which implies

$$\begin{aligned} f(\mathbf{x}+t\mathbf{v}) &= f(\mathbf{x}) + \frac{1}{2}t^2\lambda^2 |\mathbf{v}|^2 + \frac{1}{2}t^2\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right) \\ &\geq f(\mathbf{x}) + \frac{1}{4}t^2\lambda^2 |\mathbf{v}|^2 \end{aligned}$$

whenever t is small enough. Thus in the direction $\mathbf{v}$ the function has a local minimum at $\mathbf{x}$. The assertion about the local maximum in some direction follows similarly. This proves the theorem.

This theorem is an analogue of the second derivative test for higher dimensions. As in one dimension, when there is a zero eigenvalue, it may be impossible to determine from the Hessian matrix what the local qualitative behavior of the function is. For example, consider

$$f_1(x,y) = x^4 + y^2,\ f_2(x,y) = -x^4 + y^2.$$

Then $Df_i(0,0) = \mathbf{0}$ and for both functions, the Hessian matrix evaluated at $(0,0)$ equals

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

but the behavior of the two functions is very different near the origin. The second has a saddle point while the first has a minimum there.

3.15 Weierstrass Approximation Theorem

In this section we give a proof of the important approximation theorem of Weierstrass about approximating an arbitrary continuous function with a polynomial. First of all, we consider what we mean by a polynomial in the case of functions of n variables.

Definition 3.72 *Let α be an n dimensional multi-index. This means*

$$\alpha = (\alpha_1, \cdots, \alpha_n)$$

where each α_i is a natural number or zero. Also, we let

$$|\alpha| \equiv \sum_{i=1}^{n} |\alpha_i|$$

When we write $\mathbf{x}^\alpha$, we mean

$$\mathbf{x}^\alpha \equiv x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_3^{\alpha_n}.$$

An n dimensional polynomial of degree m is a function of the form

$$p(\mathbf{x}) = \sum_{|\alpha| \leq m} d_\alpha \mathbf{x}^\alpha.$$

where the d_α are complex or real numbers.

Consider $\left(1-x^2\right)^k$ and you can see that for $x \in (-1,1)$, the graphs of these functions get thinner as k increases. Now define $\phi_k(x) \equiv A_k \left(1-x^2\right)^k$ where A_k is chosen in such a way that

$$\int_{-1}^{1} \phi_k(x)\, dx = 1.$$

Thus

$$A_k = \left(\int_{-1}^{1} \left(1-x^2\right)^k dx\right)^{-1}.$$

Therefore, ϕ_k has the same shape as the graph of $\left(1-x^2\right)^k$ except that the constant A_k will make ϕ_k much taller for large k because the area must always equal 1. Now

$$\begin{aligned}\int_{-1}^{1}\left(1-x^2\right)^k dx &= 2\int_0^1\left(1-x^2\right)^k dx \\ &\geq 2\int_0^1 x\left(1-x^2\right)^k dx = \frac{1}{k+1}\end{aligned}$$

and so $A_k \leq k+1$. Therefore, we can conclude that for $1 \geq |x| \geq r > 0$,

$$\phi_k(x) \leq \phi_k(r) \leq (k+1)\left(1-r^2\right)^k$$

and $\lim_{k\to\infty}\phi_k(r) = 0$.

For $\mathbf{x} \in \mathbb{R}^n$, we define $||\mathbf{x}||_\infty \equiv \max\{|x_i|, i = 1,\cdots,n\}$. We leave it as an exercise to verify that $||\cdot||_\infty$ is a norm on $\mathbb{R}^n$ and that if we define $P_r : \mathbb{R}^n \to \mathbb{R}^n$ for $r > 0$ by

$$P_r\mathbf{x} \equiv \begin{cases} \mathbf{x} \text{ if } ||\mathbf{x}||_\infty \leq r \\ \frac{\mathbf{x}}{||\mathbf{x}||_\infty} \text{ if } ||\mathbf{x}||_\infty > r, \end{cases}$$

it follows that P_r is continuous.

With this preparation, we are ready to prove the following lemma which will yield a proof of the Weierstrass theorem.

Lemma 3.73 *Let $f : \prod_{i=1}^n [-r,r] \to \mathbb{C}$ be continuous. Then for any $\varepsilon > 0$ there exists a polynomial, p such that*

$$||p-f||_\infty < \varepsilon.$$

Proof: For $\mathbf{x} = (x_1,\cdots,x_n) \in \prod_{i=1}^n [-r,r]$ define

$$p_k(\mathbf{x}) \equiv \int_{x_1-1}^{x_1+1}\cdots\int_{x_n-1}^{x_n+1}\overline{f}(\mathbf{y})\prod_{i=1}^n \phi_k(x_i-y_i)\,dy_1\cdots dy_n \tag{3.64}$$

where $\overline{f}$ is the continuous function defined on $\mathbb{R}^n$ by $\overline{f}(\mathbf{y}) \equiv f(P_r(\mathbf{y}))$. Note that in this definition, $\mathbf{y} \in \prod_{i=1}^n [-r-1, r+1]$ because $\mathbf{x} \in \prod_{i=1}^n [-r,r]$. We see this is a polynomial because $\prod_{i=1}^n \phi_k(x_i - y_i)$ is a polynomial in $x_1,\cdots,x_n$ having coefficients which are functions of $\mathbf{y}$. Therefore, the result of the iterated integration yields a polynomial in $x_1,\cdots,x_n$. If you are concerned about the existence of the iterated integral, note that in each iteration the process asks for the integral of a continuous function so there is no problem in writing this. Also, since $\int_{x-1}^{x+1}\phi_k(x-y)\,dy = 1$,

$$f(\mathbf{x}) = \int_{x_1-1}^{x_1+1}\cdots\int_{x_n-1}^{x_n+1} f(\mathbf{x})\prod_{i=1}^n \phi_k(x_i-y_i)\,dy_1\cdots dy_n. \tag{3.65}$$

Then we note that if $||\mathbf{z}||_\infty \geq r$, we have

$$\prod_{i=1}^n \phi_k(z_i) \leq \phi_k(r)^n \tag{3.66}$$

which converges to zero as $k \to \infty$. Now from (3.64) and (3.65) we find

$$|f(\mathbf{x}) - p(\mathbf{x})| \leq$$

$$\int_{x_1-1}^{x_1+1}\cdots\int_{x_n-1}^{x_n+1}\left|f(\mathbf{x}) - \overline{f}(\mathbf{y})\right|\prod_{i=1}^n \phi_k(x_i-y_i)\,dy_1\cdots dy_n. \tag{3.67}$$

Now since $\prod_{i=1}^{n}[-r-1,r+1]$ is compact, we can apply Theorem 3.31 on Page 42 and conclude $\overline{f}$ is uniformly continuous on this set and so if $\varepsilon > 0$ is given, there exists a $\delta > 0$ such that if $||\mathbf{x}-\mathbf{y}||_{\infty} < \delta$, then $\left|\overline{f}(\mathbf{x}) - \overline{f}(\mathbf{y})\right| < \varepsilon/2$. (Note $||\mathbf{x}||_{\infty}\sqrt{n} \geq |\mathbf{x}|$.) Using (3.66), the expression in (3.67) is dominated by

$$\int_{|x_1-y_1|\geq\delta}\cdots\int_{|x_n-y_n|\geq\delta}\left|f(\mathbf{x}) - \overline{f}(\mathbf{y})\right|\phi_k(\delta)^n\,dy_1\cdots dy_n +$$

$$\int_{x_1-\delta}^{x_1+\delta}\cdots\int_{x_n-\delta}^{x_n+\delta}(\varepsilon/2)\prod_{i=1}^{n}\phi_k(x_i-y_i)\,dy_1\cdots dy_n. \tag{3.68}$$

Since $\overline{f}$ is continuous, it is bounded on $\prod_{i=1}^{n}[-r-1,r+1]$ and so the first integral in (3.68) is dominated by an expression of the form $M\phi_k(\delta)^n$ where M does not depend on $\mathbf{x}\in\prod_{i=1}^{n}[-r,r]$ while the second is dominated by

$$\int_{x_1-1}^{x_1+1}\cdots\int_{x_n-1}^{x_n+1}\frac{\varepsilon}{2}\prod_{i=1}^{n}\phi_k(x_i-y_i)\,dy_1\cdots dy_n = \frac{\varepsilon}{2}.$$

Therefore, letting k be large enough, we have shown that $|f(\mathbf{x}) - p_k(\mathbf{x})| < \varepsilon$ for all $\mathbf{x}\in\prod_{i=1}^{n}[-r,r]$. This proves the lemma.

The Weierstrass theorem is as follows.

Theorem 3.74 *Let K be any compact subset of $\mathbb{R}^n$ and let $f: K\to\mathbb{C}$ be continuous. Then for every $\varepsilon > 0$, there exists a polynomial, p, such that $||p-f||_{\infty} < \varepsilon$, Where*

$$\begin{aligned} ||g||_{\infty} &\equiv \sup\{|g(\mathbf{x})| : \mathbf{x}\in K\} \\ &= \max\{|g(\mathbf{x})| : \mathbf{x}\in K\}. \end{aligned}$$

In words, this theorem states that you can get uniformly close to a given continuous function with a polynomial.

This is an easy theorem to prove from Lemma 3.73 and the Tietze extension theorem. However, I haven't presented the Tietze extension theorem and so I will only prove the following special case which is really just a different version of Lemma 3.73. This version is all that will be needed later.

Theorem 3.75 *Let $B = \overline{B(\mathbf{0},r)}$ and let $f: B\to\mathbb{C}$ be continuous. Then for every $\varepsilon > 0$, there exists a polynomial, p, such that $||p-f||_{\infty} < \varepsilon$, Where*

$$\begin{aligned} ||g||_{\infty} &\equiv \sup\{|g(\mathbf{x})| : \mathbf{x}\in K\} \\ &= \max\{|g(\mathbf{x})| : \mathbf{x}\in K\}. \end{aligned}$$

Proof of Theorem 3.75: Let

$$Q_r(\mathbf{x}) \equiv \begin{cases} \mathbf{x} \text{ if } |\mathbf{x}| \leq r \\ \frac{\mathbf{x}}{|\mathbf{x}|} \text{ if } |\mathbf{x}| > r \end{cases}$$

Here $|\cdot|$ refers to the usual Euclidean norm. It is not hard to verify that Q_r is continuous on $\mathbb{R}^n$. Now let F be the continuous function defined on $R\equiv\prod_{i=1}^{n}[-r,r]$ given by

$$F(\mathbf{x}) \equiv f(Q_r(\mathbf{x})).$$

Therefore, $F = f$ on B and now by Lemma 3.73, there exists a polynomial, p such that

$$||f-p||_{\infty,B} \leq ||F-p||_{\infty,R} < \varepsilon.$$

Here

$$||g||_{\infty,S} \equiv \sup\{|g(\mathbf{x})| : \mathbf{x}\in S\}.$$

3.16 Ascoli Arzela Theorem

Let K be a compact subset of $\mathbb{R}^n$ and consider the continuous real or complex valued functions defined on K, denoted here by $C(K)$. You can measure the distance between two of these functions as follows.

Definition 3.76 *For $f, g \in C(K)$, define*

$$||f - g|| \equiv \max \{|f(\mathbf{x}) - g(\mathbf{x})| : \mathbf{x} \in K\}.$$

The Ascoli Arzela theorem is a major result which tells which subsets of $C(K)$ are sequentially compact.

Definition 3.77 *Let $A \subseteq C(K)$. Then A is said to be uniformly equicontinuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $\mathbf{x}, \mathbf{y} \in K$ with $|\mathbf{x} - \mathbf{y}| < \delta$ and $f \in A$,*

$$|f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon.$$

The set, A is said to be uniformly bounded if for some $M < \infty$,

$$||f|| \leq M$$

for all $f \in A$.

Uniform equicontinuity is like saying that the whole set of functions, A, is uniformly continuous on K uniformly for $f \in A$. The version of the Ascoli Arzela theorem I will present here is the following.

Theorem 3.78 *Suppose K is a nonempty compact subset of $\mathbb{R}^n$ and $A \subseteq C(K)$ is uniformly bounded and uniformly equicontinuous. Then if $\{f_k\} \subseteq A$, there exists a function, $f \in C(K)$ and a subsequence, f_{k_l} such that*

$$\lim_{l \to \infty} ||f_{k_l} - f|| = 0.$$

To give a proof of this theorem, I will first prove some lemmas.

Lemma 3.79 *If K is a compact subset of $\mathbb{R}^n$, then there exists $D \equiv \{\mathbf{x}_k\}_{k=1}^{\infty} \subseteq K$ such that D is dense in K. Also, for every $\varepsilon > 0$ there exists a finite set of points, $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\} \subseteq K$, called an ε net such that*

$$\cup_{i=1}^{m} B(\mathbf{x}_i, \varepsilon) \supseteq K.$$

Proof: For $m \in \mathbb{N}$, pick $x_1^m \in K$. If every point of K is within $1/m$ of x_1^m, stop. Otherwise, pick

$$x_2^m \in K \setminus B(x_1^m, 1/m).$$

If every point of K contained in $B(x_1^m, 1/m) \cup B(x_2^m, 1/m)$, stop. Otherwise, pick

$$x_3^m \in K \setminus (B(x_1^m, 1/m) \cup B(x_2^m, 1/m)).$$

If every point of K is contained in $B(x_1^m, 1/m) \cup B(x_2^m, 1/m) \cup B(x_3^m, 1/m)$, stop. Otherwise, pick

$$x_4^m \in K \setminus (B(x_1^m, 1/m) \cup B(x_2^m, 1/m) \cup B(x_3^m, 1/m))$$

Continue this way until the process stops, say at $N(m)$. It must stop because if it didn't, there would be a convergent subsequence due to the compactness of K. Ultimately all terms of this convergent subsequence would be closer than $1/m$, violating the manner in which they are chosen. Then $D = \cup_{m=1}^{\infty} \cup_{k=1}^{N(m)} \{x_k^m\}$. This is countable because it is a countable union of countable sets. If $\mathbf{y} \in K$ and $\varepsilon > 0$, then for some m, $2/m < \varepsilon$ and so $B(\mathbf{y}, \varepsilon)$ must contain some point of $\{x_k^m\}$ since otherwise, the process stopped too soon. You could have picked $\mathbf{y}$. This proves the lemma.

Lemma 3.80 *Suppose D is defined above and $\{g_m\}$ is a sequence of functions of A having the property that for every $\mathbf{x}_k \in D$,*

$$\lim_{m\to\infty} g_m(\mathbf{x}_k) \text{ exists.}$$

Then there exists $g \in C(K)$ such that

$$\lim_{m\to\infty} ||g_m - g|| = 0.$$

Proof: Define g first on D.

$$g(\mathbf{x}_k) \equiv \lim_{m\to\infty} g_m(\mathbf{x}_k).$$

Next I show that $\{g_m\}$ converges at every point of K. Let $\mathbf{x} \in K$ and let $\varepsilon > 0$ be given. Choose $\mathbf{x}_k$ such that for all $f \in A$,

$$|f(\mathbf{x}_k) - f(\mathbf{x})| < \frac{\varepsilon}{3}.$$

I can do this by the equicontinuity. Now if p, q are large enough, say $p, q \geq M$,

$$|g_p(\mathbf{x}_k) - g_q(\mathbf{x}_k)| < \frac{\varepsilon}{3}.$$

Therefore, for $p, q \geq M$,

$$\begin{aligned} |g_p(\mathbf{x}) - g_q(\mathbf{x})| &\leq |g_p(\mathbf{x}) - g_p(\mathbf{x}_k)| + |g_p(\mathbf{x}_k) - g_q(\mathbf{x}_k)| + |g_q(\mathbf{x}_k) - g_q(\mathbf{x})| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

It follows that $\{g_m(\mathbf{x})\}$ is a Cauchy sequence having values in $\mathbb{R}$ or $\mathbb{C}$. Therefore, it converges. Let $g(\mathbf{x})$ be the name of the thing it converges to.

Let $\varepsilon > 0$ be given and pick $\delta > 0$ such that whenever $\mathbf{x}, \mathbf{y} \in K$ and $|\mathbf{x} - \mathbf{y}| < \delta$, it follows $|f(\mathbf{x}) - f(\mathbf{y})| < \frac{\varepsilon}{3}$ for all $f \in A$. Now let $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ be a δ net for K as in Lemma 3.79. Since there are only finitely many points in this δ net, it follows that there exists N such that for all $p, q \geq N$,

$$|g_q(\mathbf{x}_i) - g_p(\mathbf{x}_i)| < \frac{\varepsilon}{3}$$

for all $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\} \cdot$ Therefore, for arbitrary $\mathbf{x} \in K$, pick $\mathbf{x}_i \in \{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ such that $|\mathbf{x}_i - \mathbf{x}| < \delta$. Then

$$\begin{aligned} |g_q(\mathbf{x}) - g_p(\mathbf{x})| &\leq |g_q(\mathbf{x}) - g_q(\mathbf{x}_i)| + |g_q(\mathbf{x}_i) - g_p(\mathbf{x}_i)| + |g_p(\mathbf{x}_i) - g_p(\mathbf{x})| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since N does not depend on the choice of $\mathbf{x}$, it follows this sequence $\{g_m\}$ is uniformly Cauchy. That is, for every $\varepsilon > 0$, there exists N such that if $p, q \geq N$, then

$$||g_p - g_q|| < \varepsilon.$$

Next, I need to verify that the function, g is a continuous function. Let N be large enough that whenever $p, q \geq N$, the above holds. Then for all $\mathbf{x} \in K$,

$$|g(\mathbf{x}) - g_p(\mathbf{x})| \leq \frac{\varepsilon}{3} \tag{3.69}$$

whenever $p \geq N$. This follows from observing that for $p, q \geq N$,

$$|g_q(\mathbf{x}) - g_p(\mathbf{x})| < \frac{\varepsilon}{3}$$

and then taking the limit as $q \to \infty$ to obtain (3.69). Now let p satisfy (3.69) for all $\mathbf{x}$ whenever $p > N$. Also pick $\delta > 0$ such that if $|\mathbf{x} - \mathbf{y}| < \delta$, then

$$|g_p(\mathbf{x}) - g_p(\mathbf{y})| < \frac{\varepsilon}{3}.$$

Then if $|\mathbf{x} - \mathbf{y}| < \delta$,

$$\begin{aligned} |g(\mathbf{x}) - g(\mathbf{y})| &\leq |g(\mathbf{x}) - g_p(\mathbf{x})| + |g_p(\mathbf{x}) - g_p(\mathbf{y})| + |g_p(\mathbf{y}) - g(\mathbf{y})| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since ε was arbitrary, this shows that g is continuous.

It only remains to verify that $||g - g_k|| \to 0$. But this follows from (3.69). This proves the lemma.

With these lemmas, it is time to prove Theorem 3.78.

Proof of Theorem 3.78: Let $D = \{\mathbf{x}_k\}$ be the countable dense set of K gauranteed by Lemma 3.79 and let $\{(1,1), (1,2), (1,3), (1,4), (1,5), \cdots\}$ be a subsequence of $\mathbb{N}$ such that

$$\lim_{k\to\infty} f_{(1,k)}(\mathbf{x}_1) \text{ exists.}$$

Now let $\{(2,1), (2,2), (2,3), (2,4), (2,5), \cdots\}$ be a subsequence of $\{(1,1), (1,2), (1,3), (1,4), (1,5), \cdots\}$ which has the property that

$$\lim_{k\to\infty} f_{(2,k)}(\mathbf{x}_2) \text{ exists.}$$

Thus it is also the case that

$$f_{(2,k)}(\mathbf{x}_1) \text{ converges to } \lim_{k\to\infty} f_{(1,k)}(\mathbf{x}_1).$$

because every subsequence of a convergent sequence converges to the same thing as the convergent sequence. Continue this way and consider the array

$$\begin{array}{c} f_{(1,1)}, f_{(1,2)}, f_{(1,3)}, f_{(1,4)}, \cdots \text{ converges at } \mathbf{x}_1 \\ f_{(2,1)}, f_{(2,2)}, f_{(2,3)}, f_{(2,4)} \cdots \text{ converges at } \mathbf{x}_1 \text{ and } \mathbf{x}_2 \\ f_{(3,1)}, f_{(3,2)}, f_{(3,3)}, f_{(3,4)} \cdots \text{ converges at } \mathbf{x}_1,\ \mathbf{x}_2, \text{ and } \mathbf{x}_3 \\ \vdots \end{array}$$

Now let $g_k \equiv f_{(k,k)}$. Thus g_k is ultimately a subsequence of $\{f_{(m,k)}\}$ whenever $k > m$ and therefore, $\{g_k\}$ converges at each point of D. By Lemma 3.80 it follows there exists $g \in C(K)$ such that

$$\lim_{k\to\infty} ||g - g_k|| = 0.$$

This proves the Ascoli Arzela theorem.

Actually there is an if and only if version of it but the most useful case is what is presented here. The process used to get the subsequence in the proof is called the Cantor diagonalization procedure.

3.17 Systems Of Ordinary Differential Equations

3.17.1 The Banach Contraction Mapping Theorem

Let $A \subseteq \mathbb{R}^n$ and $BC(A;\mathbb{R}^n)$ denote the space of continuous $\mathbb{R}^n$ valued functions, f, which have the property that $\sup\{|f(x)| : x \in A\} < \infty$. We leave it to the reader to verify that $BC(A;\mathbb{R}^n)$ is a vector space whose vectors are the functions in $BC(A;\mathbb{R}^n)$. For $f \in BC(A;\mathbb{R}^n)$ let

$$||f|| = ||f||_\infty \equiv \sup\{|f(x)| : x \in A\}$$

where the norm in the parenthesis refers to the usual norm in $\mathbb{R}^n$. Then we can check that $||\cdot||$ satisfies the axioms of a norm. That is,

a.) $||f|| \geq 0$ and $||f|| = 0$ if and only if $f = 0$.

b.) $||af|| = |a| \, ||f||$ for all constants, a.

c.) $||f + g|| \leq ||f|| + ||g||$

The first two of these axioms are obvious. We consider the triangle inequality.

$$\begin{aligned} ||f+g|| &= \sup_{x \in A} \{|f(x) + g(x)|\} \\ &\leq \sup_{x \in A} \{|f(x)| + |g(x)|\} \\ &\leq ||f|| + ||g||, \end{aligned}$$

this last inequality holding because $||f|| + ||g|| \geq |f(x)| + |g(x)|$ for all $x \in A$.

Definition 3.81 *We say a normed linear space, $(X, ||\cdot||)$ is a Banach space if it is complete. Thus, whenever, $\{\mathbf{x}_n\}$ is a Cauchy sequence, there exists $\mathbf{x} \in X$ such that $\lim_{n\to\infty} ||\mathbf{x} - \mathbf{x}_n|| = 0$.*

The following proposition shows that $BC(A; \mathbb{R}^n)$ is an example of a Banach space.

Proposition 3.82 *$(BC(A; \mathbb{R}^n), || \, ||_\infty)$ is a Banach space.*

Proof: Suppose $\{f_r\}$ is a Cauchy sequence in $BC(A; \mathbb{R}^n)$. Then for each $x \in A$, $\{f_r(x)\}$ is a Cauchy sequence in $\mathbb{R}^n$ because

$$|f_k(x) - f_l(x)| \leq ||f_k - f_l||.$$

Let $\varepsilon > 0$ be given and choose N such that $p, q \geq N$ implies

$$||f_p - f_q|| < \varepsilon.$$

Then for each $x \in A$, if $p, q \geq N$, then

$$|f_q(x) - f_p(x)| < \varepsilon.$$

Thus $\{f_r(x)\}_{r=1}^{\infty}$ is a Cauchy sequence for each $x \in A$. Define $f(x)$ by

$$f(x) \equiv \lim_{r\to\infty} f_r(x).$$

Now letting $\varepsilon > 0$, choose N such that if $p, q \geq N$, then

$$||f_p - f_q|| < \varepsilon/3.$$

Therefore, for each $x \in A$, and $p, q \geq N$,

$$|f_p(x) - f_q(x)| < \varepsilon/3$$

so letting $q \to \infty$, it follows that for all $x \in A$,

$$|f_p(x) - f(x)| \leq \varepsilon/3. \tag{3.70}$$

I need to show that f is bounded and continuous and that $||f - f_p|| \to 0$ as $p \to \infty$. From (3.70),

$$|f(x)| \leq ||f_p|| + \varepsilon/3 < \infty$$

for all $x \in A$ and so f is bounded. It remains to verify that f is continuous at $x \in A$. Pick $p > N$ where for all such p, (3.70) holds. For $y \in A$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_p(x)| + |f_p(x) - f_p(y)| + |f_p(y) - f(y)| \\ &\leq \frac{2\varepsilon}{3} + |f_p(x) - f_p(y)|. \end{aligned}$$

Now f_p is continuous and so there exists $\delta > 0$ such that if $|y - x| < \delta$, then

$$|f_p(x) - f_p(y)| < \varepsilon/3.$$

Therefore, if $y \in A$ and $|y - x| < \delta$, it follows that

$$|f(x) - f(y)| < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

and so f is continuous. Now (3.70) shows that $||f_p - f|| \to 0$. This proves the proposition. .

Corollary 3.83 *Let K be a compact metric space and denote by $C(K; \mathbb{R}^n)$ the space of continuous functions having values in $\mathbb{R}^n$ which are defined on K. Then $C(K; \mathbb{R}^n)$ with the norm, $||\cdot||_\infty$ defined above is a Banach space.*

Proof: Since K is compact, it follows that $C(K; \mathbb{R}^n) = BC(K; \mathbb{R}^n)$ because for $f \in C(K; \mathbb{R}^n)$, $x \to ||f(x)||$ is a continuous function having values in $\mathbb{R}$ and therefore, achieves its maximum. Therefore, the above proposition gives the desired conclusion.

Thus $C(K; \mathbb{R}^n)$ and $BC(A; \mathbb{R}^n)$ are two examples of complete normed linear spaces. The following theorem is the contraction mapping theorem of Banach.

Theorem 3.84 *Suppose X is a complete normed linear space and $T : X \to X$ satisfies*

$$|Tx - Ty| \leq r|x - y|$$

where $r \in (0, 1)$. Then there exists a unique point x such that $Tx = x$.

Proof: Pick $x_0 \in X$ and consider $\{T^k(x_0)\}_{k=1}^\infty$. Then since T is a contraction map, as indicated, it follows easily that

$$\left|\left|T^{k+1}x_0 - T^k x_0\right|\right| \leq r^k ||Tx_0 - x_0||.$$

therefore, letting $q > p > N$,

$$||T^q x_0 - T^p x_0|| \leq \sum_{k=p+1}^{q-1} \left|\left|T^{k+1}x_0 - T^k x_0\right|\right| \leq \sum_{k=N}^{\infty} r^k |Tx_0 - x_0| \leq \frac{r^N}{(1-r)} |Tx_0 - x_0|.$$

Letting $\varepsilon > 0$ be given you can choose N large enough that

$$\frac{r^N}{(1-r)} |Tx_0 - x_0| < \varepsilon$$

and so for $p, q \geq N$,

$$||T^q x_0 - T^p x_0|| < \varepsilon$$

showing that $\{T^k(x_0)\}_{k=1}^\infty$ is a Cauchy sequence. Therefore, there exists x such that $T^k(x_0) \to x$. Therefore,

$$Tx = T \lim_{k\to\infty} T^k x_0 = \lim_{k\to\infty} T^{k+1} x_0 = x$$

showing that x is a fixed point. Suppose now that $Ty = y$. Then

$$||x - y|| = ||Tx - Ty|| \leq r\,||x - y||$$

which shows $y = x$ and so the fixed point is unique.

A mapping T which satisfies the hypothesis of the above theorem is called a contraction map.

The following is the fundamental theorem for existence of the initial value problem for ordinary differential equations.

Theorem 3.85 *Let $c \in [a, b]$ and suppose $\mathbf{f} : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and satisfies the Lipschitz condition,*

$$|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})| \leq K\,|\mathbf{x} - \mathbf{y}|, \tag{3.71}$$

and $\mathbf{x}_0 \in \mathbb{R}^n$ is given. Then there exists a unique solution on $[a, b]$ to the initial value problem,

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}),\ \mathbf{x}(c) = \mathbf{x}_0, \tag{3.72}$$

on $[a, b]$ and if $\mathbf{x}(t; \mathbf{x}_0)$ denotes this solution, there exists M depending only on K and $|b - a|$ such that for all $t \in [a, b]$,

$$|\mathbf{x}(t, \mathbf{x}_0) - \mathbf{x}(t, \mathbf{x}_0')| \leq M\,|\mathbf{x}_0 - \mathbf{x}_0'|. \tag{3.73}$$

Proof: This will use the norm

$$||\mathbf{x}||_\lambda \equiv \max\left\{|\mathbf{x}(t)|\,e^{-\lambda|t-c|} : t \in [a, b]\right\}. \tag{3.74}$$

You can see this norm is equivalent to the usual norm, $||\cdot||$ defined above. Let $0 < \delta < \min\left\{e^{-\lambda|t-c|} : t \in [a, b]\right\}$. Then for $\mathbf{x} \in C([a, b]; \mathbb{R}^n)$,

$$\delta\,||\mathbf{x}|| \leq ||\mathbf{x}||_\lambda \leq ||\mathbf{x}||. \tag{3.75}$$

Therefore, with respect to this new norm, $C([a, b]; \mathbb{R}^n)$ is a complete normed linear space. The solution to the initial value problem, (3.72), if there is one, must satisfy

$$\mathbf{x}(t, \mathbf{x}_0) = \mathbf{x}_0 + \int_c^t \mathbf{f}(r, \mathbf{x}(r, \mathbf{x}_0))\,dr. \tag{3.76}$$

Conversely, if $t \to \mathbf{x}(t, \mathbf{x}_0)$ solves the integral equation, (3.76), then it is a solution to the initial value problem, (3.72). This follows from the fundamental theorem of calculus and properties of integrals. Therefore, it suffices to look for solutions to (3.76). Define for $\mathbf{x} \in C([a, b]; \mathbb{R}^n)$,

$$T_{\mathbf{x}_0}(\mathbf{x})(t) \equiv \mathbf{x}_0 + \int_c^t \mathbf{f}(r, \mathbf{x}(r))\,dr.$$

I will show that if λ is large enough, then $T_{\mathbf{x}_0}$ is a contraction map on $C([a, b]; \mathbb{R}^n)$. Let $\mathbf{x}, \mathbf{y}$ be two elements of $C([a, b]; \mathbb{R}^n)$.

$$\begin{aligned}
||T_{\mathbf{x}_0}\mathbf{x} - T_{\mathbf{x}_0}\mathbf{y}||_\lambda &\leq \max\left\{\left|\int_c^t |\mathbf{f}(r, \mathbf{x}(r)) - \mathbf{f}(r, \mathbf{y}(r))|\,dr\right| e^{-\lambda|t-c|} : t \in [a, b]\right\} \\
&\leq \max\left\{\left|\int_c^t K\,|\mathbf{x}(r) - \mathbf{y}(r)|\,dr\right| e^{-\lambda|t-c|} : t \in [a, b]\right\} \\
&\leq K\,||\mathbf{x} - \mathbf{y}||_\lambda \max\left\{\left|\int_c^t e^{\lambda|r-c|}dr\right| e^{-\lambda|t-c|} : t \in [a, b]\right\}.
\end{aligned} \tag{3.77}$$

Now there are two cases to consider, the one where $t > c$ and the one where $t < c$.

First consider the one where $t > c$.

$$\left|\int_c^t e^{\lambda|r-c|}dr\right| e^{-\lambda|t-c|} = \int_c^t e^{\lambda(r-c)}dre^{-\lambda(t-c)} = \frac{1-e^{-\lambda(t-c)}}{\lambda} \leq \frac{1}{\lambda}.$$

Next consider the case where $t < c$. In this case,

$$\left|\int_c^t e^{\lambda|r-c|}dr\right| e^{-\lambda|t-c|} = \left(\int_t^c e^{\lambda(c-r)}dr\right) e^{-\lambda(c-t)} = \frac{-e^{\lambda(t-c)}+1}{\lambda} \leq \frac{1}{\lambda}.$$

Therefore, for every value of t,

$$||T_{\mathbf{x}_0}\mathbf{x} - T_{\mathbf{x}_0}\mathbf{y}||_\lambda \leq \frac{K}{\lambda}||\mathbf{x}-\mathbf{y}||_\lambda$$

so if $\lambda > K$, it follows $T_{\mathbf{x}_0}$ is a contraction map. Pick such a λ and let $r = K/\lambda$. Therefore, there exists a unique fixed point in $C([a,b];\mathbb{R}^n)$, $\mathbf{x}$ satisfying

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_c^t \mathbf{f}(r, \mathbf{x}(r))\, dr$$

and this is the unique solution to the initial value problem (3.72). This proves existence and uniqueness of the solution to the initial value problem. It remains to verify the estimate involving the initial condition.

$$||\mathbf{x}(\cdot,\mathbf{x}_0) - \mathbf{x}(\cdot,\mathbf{x}_0')||_\lambda = \left|\left|T_{\mathbf{x}_0}(\mathbf{x}(\cdot,\mathbf{x}_0)) - T_{\mathbf{x}_0'}(\mathbf{x}(\cdot,\mathbf{x}_0'))\right|\right|_\lambda \leq$$

$$\left|\left|T_{\mathbf{x}_0}(\mathbf{x}(\cdot,\mathbf{x}_0)) - T_{\mathbf{x}_0'}(\mathbf{x}(\cdot,\mathbf{x}_0))\right|\right|_\lambda + \left|\left|T_{\mathbf{x}_0'}(\mathbf{x}(\cdot,\mathbf{x}_0)) - T_{\mathbf{x}_0'}(\mathbf{x}(\cdot,\mathbf{x}_0'))\right|\right|_\lambda$$

$$\leq |\mathbf{x}_0 - \mathbf{x}_0'| + r\,||\mathbf{x}(\cdot,\mathbf{x}_0) - \mathbf{x}(\cdot,\mathbf{x}_0')||_\lambda$$

Therefore,

$$||\mathbf{x}(\cdot,\mathbf{x}_0) - \mathbf{x}(\cdot,\mathbf{x}_0')||_\lambda \leq \frac{1}{1-r}|\mathbf{x}_0 - \mathbf{x}_0'|.$$

From (3.75),

$$\begin{aligned} |\mathbf{x}(t,\mathbf{x}_0) - \mathbf{x}(t,\mathbf{x}_0')| &\leq ||\mathbf{x}(\cdot,\mathbf{x}_0) - \mathbf{x}(\cdot,\mathbf{x}_0')|| \\ &\leq \frac{1}{\delta}||\mathbf{x}(\cdot,\mathbf{x}_0) - \mathbf{x}(\cdot,\mathbf{x}_0')||_\lambda \\ &\leq \frac{1}{\delta(1-r)}|\mathbf{x}_0 - \mathbf{x}_0'|. \end{aligned}$$

This proves the theorem.

3.17.2 C^1 Surfaces And The Initial Value Problem

This is a nice theorem but it is not enough for partial differential equations. It is necessary to consider how the solutions depend on the initial data, $\mathbf{x}_0$. There are various ways to do this, the best being a direct appeal to the implicit function theorem in infinite dimensional spaces. However, there is a more pedestrian approach involving Gronwall's inequality which is presented next. The following lemma is usually referred to as Gronwall's inequality.

Lemma 3.86 *Suppose for $t > c$, $u(t) \leq C + \int_c^t ku(s)\,ds$ where u is a real valued continuous function and $k \geq 0$. Then for $t > c$,*

$$u(t) \leq Ce^{k(t-c)}.$$

Proof: Let $w(t) = \int_c^t u(s)\,ds$. Then the given inequality is of the form

$$w'(t) - kw(t) \leq C.$$

Now multiply both sides by $e^{-k(t-c)}$ to obtain

$$\left(we^{-k(t-c)}\right)' \leq Ce^{-k(t-c)}.$$

Now integrate both sides from c to t.

$$\begin{aligned} w(t)\,e^{-k(t-c)} - 0 &\leq C\int_c^t e^{-k(s-c)}ds \\ &= C\left(\frac{1-e^{-k(t-c)}}{k}\right) \end{aligned}$$

and so

$$w(t) \leq C\left(\frac{e^{k(t-c)}-1}{k}\right).$$

Therefore, from the original inequality satisfied by u,

$$u(t) \leq C + kw(t) = C + C\left(e^{k(t-c)} - 1\right) = Ce^{k(t-c)}.$$

This proves the lemma.

Not surprisingly, there is a version for $t < c$.

Lemma 3.87 *Suppose for $t < c$, $u(t) \leq C + \int_t^c ku(s)\,ds$ where u is a real valued continuous function and $k \geq 0$. Then for $t < c$,*

$$u(t) \leq Ce^{k(c-t)}.$$

Proof: Let $w(t) = \int_t^c u(s)\,ds$. Then the given inequality takes the form

$$-w'(t) \leq C + kw(t)$$

and so

$$w'(t) + kw(t) \geq -C.$$

Therefore,

$$\left(e^{k(t-c)}w(t)\right)' \geq -Ce^{k(t-c)}$$

and integrating from c to t while remembering that $t < c$,

$$e^{k(t-c)}w(t) \leq -C\int_c^t e^{k(s-c)}ds = -C\frac{e^{k(t-c)}-1}{k}.$$

Therefore,

$$w(t) \leq -C\frac{1-e^{-k(t-c)}}{k}.$$

Placing this in the original given inequality,

$$\begin{aligned} u(t) &\leq C + \int_t^c ku(s)\,ds \\ &= C + kw(t) \\ &\leq C - C\frac{1-e^{-k(t-c)}}{k}k \\ &= Ce^{k(c-t)} \end{aligned}$$

and this proves the lemma.

Corollary 3.88 *Suppose* $\mathbf{f}$ *satisfies the Lipschitz condition,*

$$|\mathbf{f}(r,\mathbf{x}) - \mathbf{f}(r,\mathbf{y})| \leq k|\mathbf{x}-\mathbf{y}|$$

and also that

$$\mathbf{u}(t) = \mathbf{u}_0 + \int_c^t \mathbf{f}(r,\mathbf{u}(r))\,dr$$

and

$$\mathbf{v}(t) = \mathbf{v}_0 + \int_c^t \mathbf{f}(r,\mathbf{v}(r))\,dr$$

for $t \in [a,b]$ *with* $c \in (a,b)$. *Then*

$$|\mathbf{v}(t) - \mathbf{u}(t)| \leq |\mathbf{v}_0 - \mathbf{u}_0| e^{k|t-c|}. \tag{3.78}$$

Proof: If $t > c$, then

$$|\mathbf{v}(t) - \mathbf{u}(t)| \leq |\mathbf{v}_0 - \mathbf{u}_0| + \int_c^t |\mathbf{v}(r) - \mathbf{u}(r)|\,dr$$

and so from Lemma 3.86 (3.78) holds. If $t \leq c$, the result follows from

$$|\mathbf{v}(t) - \mathbf{u}(t)| \leq |\mathbf{v}_0 - \mathbf{u}_0| + \int_t^c |\mathbf{v}(r) - \mathbf{u}(r)|\,dr$$

and Lemma 3.87.

In Theorem 3.85 it was assumed that $\mathbf{f}(t,\mathbf{x})$ was continuous, bounded, and satisfied a Lipschitz condition on the second argument. Suppose now that you are interested in the following problem in which $c \in [a,b]$.

$$\mathbf{x}_t(t,\mathbf{s}) = \mathbf{f}(t,\mathbf{x}(t,\mathbf{s})),\ \mathbf{x}(c,\mathbf{s}) = \mathbf{x}_0(\mathbf{s}) \tag{3.79}$$

Where $\mathbf{x}_0 \in C^1\left(\prod_{k=1}^m [c_k,d_k];\mathbb{R}^n\right)$. By $C^p\left(\prod_{k=1}^m [c_k,d_k];\mathbb{R}^n\right)$ I mean the set of functions which are restrictions to $\prod_{k=1}^m [c_k,d_k]$ of a function which is in $C^p(\mathbb{R}^m;\mathbb{R}^n)$, that is, one which is defined on all of $\mathbb{R}^m$ having values in $\mathbb{R}^n$ which is C^p.

Theorem 3.89 *Suppose* $\mathbf{x}_0 \in C^1\left(\prod_{k=1}^m [c_k,d_k];\mathbb{R}^n\right)$, $\mathbf{f}$ *is* C^1, $\mathbf{f}$ *and its first partial derivatives are bounded, and* $\mathbf{f}$ *satisfies a Lipschitz condition on the second argument as in Theorem 3.85. Letting* $\mathbf{x}(t,\mathbf{s})$ *denote the solution to the problem (3.79), it follows* $\mathbf{x}$ *is a* C^1 *function. That is all its partial derivative up to order 1 exist and are continuous on* $[a,b] \times \prod_{k=1}^m [c_k,d_k]$ *where you take the derivative from an appropriate side at the end points.*

Proof: It has been shown in Theorem 3.85 that

$$\mathbf{x}(t,\mathbf{s}) = \mathbf{x}_0(\mathbf{s}) + \int_c^t \mathbf{f}(r, \mathbf{x}(r,\mathbf{s}))\,dr. \tag{3.80}$$

Now consider the solution to

$$\mathbf{z}(t,\mathbf{s}) = \frac{\partial \mathbf{x}_0}{\partial s_k}(\mathbf{s}) + \int_c^t \mathbf{F}(r, \mathbf{z}(r,s))\,dr \tag{3.81}$$

where

$$\mathbf{F}(t,\mathbf{z}) \equiv D_2\mathbf{f}(t, \mathbf{x}(t,\mathbf{s}))\,\mathbf{z}. \tag{3.82}$$

For $\mathbf{x}(t,\mathbf{s})$ the solution to (3.79) for a fixed $\mathbf{s}$. Then F is continuous and Lipschitz continuous in the second argument with a single Lipschitz constant for all $\mathbf{s}$ because of the continuity of $\mathbf{x}$ and the assumption that $\mathbf{f}$ is C^1. By Theorem 3.85 it follows there exists a unique solution to (3.81). Furthermore, from this theorem, the solution, $\mathbf{z}$ is a continuous function of its arguments. This follows from the assumption that $\frac{\partial \mathbf{x}_0}{\partial s_k}$ is continuous and the estimate of Theorem 3.85 which gave a Lipschitz dependence on the initial condition. Now I want to show that $\mathbf{z}(t,\mathbf{s}) = \frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s})$. From (3.80) and (3.81),

$$\mathbf{x}(t,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(t,\mathbf{s}) - \mathbf{z}(t,\mathbf{s})\,h = \mathbf{x}_0(\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}_0(\mathbf{s}) - \frac{\partial \mathbf{x}_0}{\partial s_k}(\mathbf{s})\,h \tag{3.83}$$

$$+\int_c^t \left(\mathbf{f}(r, \mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k)) - \mathbf{f}(r,\mathbf{x}(r,\mathbf{s})) - D_2\mathbf{f}(t,\mathbf{x}(t,\mathbf{s}))\,\mathbf{z}(r,\mathbf{s})\,h\right)dr \tag{3.84}$$

Now consider the integrand of (3.84).

$$\mathbf{f}(r, \mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k)) - \mathbf{f}(r,\mathbf{x}(r,\mathbf{s})) - D_2\mathbf{f}(t,\mathbf{x}(t,\mathbf{s}))\,\mathbf{z}(r,\mathbf{s})\,h =$$

$$D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\left(\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s}) - \mathbf{z}(r,\mathbf{s})\,h\right) + o\left(\left|\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s})\right|\right).$$

The term, $o(|\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s})|)$ is actually $o(h)$ independent of r and $\mathbf{s}$. To see this, use the Lipschitz estimate of Theorem 3.85 along with the assumption that $\mathbf{x}_0$ is C^1 to conclude that for h small enough,

$$\left|o\left(\left|\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s})\right|\right)\right| \le \varepsilon\left|\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s})\right| \le C\varepsilon\,|h|$$

for some C which is independent of $\mathbf{s}$ and $r \in [a,b]$. Also the right side of (3.83) is $o(|h|)$ because of the assumption that $\mathbf{x}_0$ is C^1. Therefore, the equation (3.83) - (3.84) is of the form

$$\mathbf{x}(t,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(t,\mathbf{s}) - \mathbf{z}(t,\mathbf{s})\,h = o(|h|) + \int_c^t D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\left(\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(r,\mathbf{s}) - \mathbf{z}(r,\mathbf{s})\,h\right).$$

It follows from this and Gronwall's inequalities, Lemmas 3.86 and 3.87 that

$$\left|\mathbf{x}(t,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(t,\mathbf{s}) - \mathbf{z}(t,\mathbf{s})\,h\right| \le |o(|h|)|$$

and so $\mathbf{x}(t,\mathbf{s}+h\mathbf{e}_k) - \mathbf{x}(t,\mathbf{s}) - \mathbf{z}(t,\mathbf{s})\,h = o(|h|)$ showing that $\mathbf{z}(t,\mathbf{s}) = \frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s})$.

I have just shown that all partial derivatives of $\mathbf{x}(t,\mathbf{s})$ taken with respect to the components of $\mathbf{s}$ exist and are continuous. That $\mathbf{x}_t$ exists and is continuous follows from the continuity of $\mathbf{x}(t,\mathbf{s})$, the function, $\mathbf{f}$ and the differential equation satisfied by $\mathbf{x}$. Therefore, $\mathbf{x}$ is a C^1 function as claimed. This proves the theorem.

Corollary 3.90 *Suppose in addition to the hypotheses of Theorem 3.89 that $\mathbf{x}_0$ is the restriction to $\prod_{k=1}^{m}[c_k,d_k]$ of a function which is C^2, and not just C^1 and that $\mathbf{f}$ is C^2, not just C^1 and $\mathbf{f}$ and all its first and second order partial derivatives are bounded. Then $\mathbf{x}(t,\mathbf{s})$ is C^2. That is all its second partial derivatives exist and are continuous.*

Proof: As before,

$$\mathbf{x}(t,\mathbf{s}) = \mathbf{x}_0(\mathbf{s}) + \int_c^t \mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\,dr. \tag{3.85}$$

Theorem 3.89 shows that

$$\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}) = \frac{\partial \mathbf{x}}{\partial s_k}(\mathbf{s}) + \int_c^t D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})\,dr \tag{3.86}$$

and that each of these partial derivatives is continuous. Also

$$\mathbf{x}_{,t}(t,\mathbf{s}) = \mathbf{f}(t,\mathbf{x}(t,\mathbf{s})) \tag{3.87}$$

and so

$$\frac{\partial^2 \mathbf{x}}{\partial s_k \partial t}(t,\mathbf{s}) = D_2\mathbf{f}(t,\mathbf{x}(t,\mathbf{s}))\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}) \tag{3.88}$$

which shows that these sorts of second order partial derivatives exist and are continuous. Also from the fundamental theorem of calculus and (3.86),

$$\frac{\partial^2 \mathbf{x}}{\partial t \partial s_k}(t,\mathbf{s}) = D_2\mathbf{f}(t,\mathbf{x}(t,\mathbf{s}))\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s})$$

so these second order partial derivatives exist and are continuous also. Note this coincides with the mixed partial derivative taken in the other order as it must by the theorem on equality of mixed partial derivatives. What remains to check? The only case left to consider is $\frac{\partial^2 \mathbf{x}}{\partial s_l \partial s_k}$. It is handled much the same way as in the C^1 case.

Let $\mathbf{z}(t,\mathbf{s})$ be the solution to

$$\mathbf{z}(t,\mathbf{s}) = \frac{\partial^2 \mathbf{x}_0}{\partial s_l \partial s_k}(\mathbf{s}) +$$

$$\int_c^t \left[\frac{\partial}{\partial s_l}\left(D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})\right) + D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\,\mathbf{z}(r,\mathbf{s})\right]dr$$

Observe that $\frac{\partial}{\partial s_l}\left(D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})\right)$ is a known continuous function of r for each fixed $\mathbf{s}$. Therefore, as in the C^1 case, there exists a unique solution, $\mathbf{z}(t,\mathbf{s})$ to the above integral equation and this solution is continuous. I need to identify it with $\frac{\partial^2 \mathbf{x}}{\partial s_l \partial s_k}(t,\mathbf{s})$. Using the assumption that $\mathbf{x}_0$ is C^2, and referring to (3.86),

$$\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}+h\mathbf{e}_l) - \frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}) - \mathbf{z}(t,\mathbf{s})\,h = o(h) \tag{3.89}$$

$$+\int_c^t D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_l))\left(\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s}+h\mathbf{e}_l)\right) - D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s}) - \tag{3.90}$$

$$\left[\frac{\partial}{\partial s_l}\left(D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})\right)h + D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\,\mathbf{z}(r,\mathbf{s})\,h\right]dr. \tag{3.91}$$

This is a great and terrible mess. First I will massage the first term in (3.91). In doing so, recall that

$$o\left(\left|\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_k\right)-\mathbf{x}\left(r,\mathbf{s}\right)\right|\right)=o\left(|h|\right)$$

and it has the little o property uniformly in r and $\mathbf{s}$ because it has been established already that $\mathbf{x}$ is C^1 and all this is taking place on a compact set so there are upper bounds for absolute values of all partial derivatives of $\mathbf{x}$. Because of this observation,

$$\frac{\partial}{\partial s_l}\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)h=\frac{\partial}{\partial s_l}\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)h\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)$$

$$=\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)+o\left(|h|\right)$$

where $o\left(|h|\right)$ always refers to having the little o property uniformly in r and $\mathbf{s}$. Therefore, (3.89) - (3.91) equals

$$\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}\right)-\mathbf{z}\left(t,\mathbf{s}\right)h=o\left(h\right) \tag{3.92}$$

$$+\int_c^t D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)- \tag{3.93}$$

$$\left[\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)+D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\mathbf{z}\left(r,\mathbf{s}\right)h\right]dr. \tag{3.94}$$

which simplifies to

$$\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}\right)-\mathbf{z}\left(t,\mathbf{s}\right)h=o\left(h\right) \tag{3.95}$$

$$+\int_c^t D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\mathbf{z}\left(r,\mathbf{s}\right)hdr$$

and this equals

$$\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(t,\mathbf{s}\right)-\mathbf{z}\left(t,\mathbf{s}\right)h=o\left(h\right) \tag{3.96}$$

$$+\int_c^t D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)-\mathbf{z}\left(r,\mathbf{s}\right)h\right)dr \tag{3.97}$$

$$\int_c^t\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)dr. \tag{3.98}$$

Consider the integrand of the last integral.

$$\left(D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}+h\mathbf{e}_l\right)\right)-D_2\mathbf{f}\left(r,\mathbf{x}\left(r,\mathbf{s}\right)\right)\right)\left(\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}+h\mathbf{e}_l\right)-\frac{\partial \mathbf{x}}{\partial s_k}\left(r,\mathbf{s}\right)\right)=$$

$$[D_2(D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s})))(\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_l)-\mathbf{x}(r,\mathbf{s}))+o(|h|)]\left[\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s}+h\mathbf{e}_l)-\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})\right]$$

Since it is already known that $\mathbf{x}$ is C^1, it follows that $h^{-1}(\mathbf{x}(r,\mathbf{s}+h\mathbf{e}_l)-\mathbf{x}(r,\mathbf{s}))$ is bounded and now the continuity of $\frac{\partial \mathbf{x}}{\partial s_k}$ shows that this expression is $o(|h|)$. Furthermore, it has this property uniformly in r and $\mathbf{s}$ due to compactness of $[a,b]\times\prod_{k=1}^m[c_k,d_k]$. Therefore, (3.89) - (3.91) is of the form

$$\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}+h\mathbf{e}_l)-\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s})-\mathbf{z}(t,\mathbf{s})h=o(h) \tag{3.99}$$

$$+\int_c^t D_2\mathbf{f}(r,\mathbf{x}(r,\mathbf{s}))\left(\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s}+h\mathbf{e}_l)-\frac{\partial \mathbf{x}}{\partial s_k}(r,\mathbf{s})-\mathbf{z}(r,\mathbf{s})h\right)dr \tag{3.100}$$

and now the Gronwall inequalities, Lemmas 3.86 and 3.87 imply that

$$\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s}+h\mathbf{e}_l)-\frac{\partial \mathbf{x}}{\partial s_k}(t,\mathbf{s})-\mathbf{z}(t,\mathbf{s})h=o(h).$$

Therefore, $\mathbf{z}(t,\mathbf{s})=\frac{\partial^2\mathbf{x}}{\partial s_l\partial s_k}(t,\mathbf{s})$ and this completes showing that $\mathbf{x}$ is C^2. This proves the corollary.

I think you can see from this that if you assume $\mathbf{x}_0$ and $\mathbf{f}$ are both C^m and that all the partial derivatives of $\mathbf{f}$ up to order m are bounded, then the solution, $\mathbf{x}(t,\mathbf{s})$ will also be C^m.

The following corollary is a local result which depends on conditions which are easy to verify.

Corollary 3.91 *Let U be an open bounded and convex set and suppose $\mathbf{x}_0:\prod_{k=1}^m[c_k,d_k]\to U$ is a C^m function and let $\mathbf{f}\in C^m((c-T_1,c+T_1)\times V)$ where V is an open set containing $\overline{U}$. Then there exists $T>0$ such that there exists a unique solution, $\mathbf{x}$ to the initial value problem,*

$$\mathbf{x}_t(t,\mathbf{s})=\mathbf{f}(t,\mathbf{x}(t,\mathbf{s})),\ \mathbf{x}(c,\mathbf{s})=\mathbf{x}_0(s) \tag{3.101}$$

for each $s\in\prod_{k=1}^m[c_k,d_k]$ and $t\in[c-T,c+T]$. Furthermore, $\mathbf{x}$ is a C^m function for $(t,s)\in[c-T,c+T]\times\prod_{k=1}^m[c_k,d_k]$.

Proof: Let $d=\operatorname{dist}\left(\mathbf{x}_0\left(\prod_{k=1}^m[c_k,d_k]\right),U^C\right)$. Thus $d>0$ because $\mathbf{x}_0\left(\prod_{k=1}^m[c_k,d_k]\right)$ is a compact set. Now let ψ be a function which is infinitely differentiable which also satisfies $\psi=0$ on V^C, $\psi(\mathbf{x})\in[0,1]$ for all $\mathbf{x}$, and $\psi(\mathbf{x})=1$ for all $\mathbf{x}\in\overline{U}$.[1] Let $\mathbf{f}_1(t,\mathbf{x})\equiv\overline{\mathbf{f}}(t,\mathbf{x})\psi(\mathbf{x})$ where $\overline{\mathbf{f}}$ denotes the zero extension of $\mathbf{f}$ off V. Pick $T_2<T_1$. Then from the assumption that $\mathbf{f}$ is C^1 and U is convex, it follows $\mathbf{f}_1$ is Lipschitz in both arguments on $(c-T_1,c+T_1)\times U$. Therefore, there exists a unique solution to (3.101), $\mathbf{x}$, and $\mathbf{x}$ is a C^m function. The corollary is proved by showing that for all $\mathbf{s}\in\prod_{k=1}^m[c_k,d_k]$, $|\mathbf{x}(t,\mathbf{s})-\mathbf{x}_0(\mathbf{s})|<d$ whenever $t\in[c-T,c+T]$ for a sufficiently small T because if this is done, then for $(t,\mathbf{s})\in[c-T,c+T]\times\prod_{k=1}^m[c_k,d_k]$, the two functions, $\mathbf{f}(t,\mathbf{x}(t,\mathbf{s}))$ and $\mathbf{f}_1(t,\mathbf{x}(t,\mathbf{s}))$ are the same. But there exists a constant, C such that $|\mathbf{f}_1(t,\mathbf{x})|<C$ for all $(t,\mathbf{x})\in[c-T_1,c+T_1]\times\mathbb{R}^n$ and so for $0<T<d/C$, it follows that for $t\in[c-T,c+T]$ and $s\in[a,b]$,

$$\begin{aligned}|\mathbf{x}(t,\mathbf{s})-\mathbf{z}(\mathbf{s})| &\le \left|\int_c^t|\mathbf{f}_1(r,\mathbf{x}(r,\mathbf{s}))|\,dr\right|\\ &\le CT<d.\end{aligned}$$

This proves the corollary.

[1] The existence of such a function is pretty significant and will be discussed later in connection with the theory of the Lebesgue integral where it is most easily shown.

First Order PDE

4.1 Quasilinear First Order PDE

Definition 4.1 *A first order quasilinear PDE (partial differential equation) is one of the form*

$$\sum_{j=1}^{n} a_j(\mathbf{x}, u) u_{,j} = b(\mathbf{x}, u) \tag{4.1}$$

where here $u_{,j}$ denotes the partial derivative of u with respect to x_j and $\mathbf{x}$ is in an open subset of $\mathbb{R}^n$ and u is the unknown function. A Linear PDE is one which is of the form

$$\sum_{j=1}^{n} a_j(\mathbf{x}) u_{,j} = b(\mathbf{x}) u \tag{4.2}$$

It turns out that the subject splits conveniently into the case of quasilinear versus nonlinear. I will discuss the quasilinear case here. Suppose u is a solution. Then $(\mathbf{x}, u(\mathbf{x}))$ would give a n dimensional surface in $\mathbb{R}^{n+1}$. Suppose you want this surface to contain the $n-1$ dimensional set in $\mathbb{R}^n$, $(\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}))$ where $\mathbf{s} \in \prod_{i=1}^{n-1} [a_i, b_i]$. For example, in case $n = 2$, you would be looking for a 2 dimensional surface containing a curve. Then suppose $(\mathbf{x}(t, \mathbf{s}), z(t, \mathbf{s}))$ gives a parameterization of this surface for t near 0. If so,

$$u(\mathbf{x}(t, \mathbf{s})) = z(t, \mathbf{s})$$

and therefore,

$$z_{,t} - \sum_{i=1}^{n} u_{,i}(\mathbf{x}) x_{i,t} = 0.$$

From the partial differential equation, this will hold if

$$z_{,t}(t) = b(\mathbf{x}(t), z(t)), \ x_{i,t} = a_i(\mathbf{x}(t), z(t))$$

and you can insist the surface contains the given $n-1$ dimensional set by imposing the following initial conditions.

$$z(0, \mathbf{s}) = z_0(\mathbf{s}), \quad x_i(0, \mathbf{s}) = x_{0i}(\mathbf{s}).$$

By the theorems on dependence of solutions on initial data, it follows $\mathbf{x}$ is a C^1 function. Then if

$$\frac{\partial(x_1 \cdots x_n)}{\partial(t, s_1, \cdots s_{n-1})}(0, \mathbf{s}) \neq 0, \tag{4.3}$$

it follows from the inverse function theorem that the equations,

$$x_i = x_i(t, \mathbf{s}), i = 1, \cdots, n$$

can be locally inverted to obtain $t = t(\mathbf{x})$ and $s_k = s_k(\mathbf{x})$. Letting $z(t, \mathbf{s}) = u(\mathbf{x})$, it follows that

$$b(\mathbf{x},u) = b(\mathbf{x},z) = z_{,t} = \sum_j u_{,j}(\mathbf{x}) x_{j,t} = \sum_j u_{,j}(\mathbf{x}) a_j(\mathbf{x}, z) = \sum_j u_{,j}(\mathbf{x}) a_j(\mathbf{x}, u)$$

showing that the partial differential equation is satisfied in addition to having the surface contain the given $n-1$ dimensional set, $(\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}))$ where $\mathbf{s} \in \prod_{i=1}^{n-1} [a_i, b_i]$.

This shows that to obtain a solution of (4.1) which contains the $n-1$ dimensional set,

$$\left\{ (\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s})) : \mathbf{s} \in \prod_{i=1}^{n-1} [a_i, b_i] \right\}, \tag{4.4}$$

you should solve the initial value problem, the solutions of which are called characteristics:

$$x_{j,t} = a_j(\mathbf{x}, z), z_{,t} = b(\mathbf{x}, z), (\mathbf{x}(0, \mathbf{s}), z(0, \mathbf{s})) = (\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s})). \tag{4.5}$$

If a_j and b are all C^1 functions and if $(\mathbf{x}_0(\mathbf{s}), z_0(s))$ is C^1, and the condition, (4.3) holds, then there exists a unique solution to the PDE (4.1) given by $u(\mathbf{x}) = z(t(\mathbf{x}), \mathbf{s}(\mathbf{x}))$ for $\mathbf{x} \in (t, \mathbf{s})^{-1} \left([-T, T] \times \prod_{i=1}^{n-1} [a_i, b_i] \right)$ where T is some positive number. In other words there exists a local solution to the PDE which contains the given set described in (4.4). Consider condition (4.3). This condition is

$$\frac{\partial(x_1 \cdots x_n)}{\partial(t, s_1, \cdots s_{n-1})}(0, \mathbf{s}) \equiv \det \begin{pmatrix} \frac{\partial x_1}{\partial t}(0, \mathbf{s}) & \frac{\partial x_1}{\partial s^1}(0, \mathbf{s}) & \cdots & \frac{\partial x_1}{\partial s^{n-1}}(0, \mathbf{s}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial t}(0, \mathbf{s}) & \frac{\partial x_n}{\partial s^1}(0, \mathbf{s}) & \cdots & \frac{\partial x_n}{\partial s^{n-1}}(0, \mathbf{s}) \end{pmatrix} \neq 0.$$

From the given initial value problem, this reduces to

$$\det \begin{pmatrix} a_1((\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}))) & x_{01,s^1}(\mathbf{s}) & \cdots & x_{01,s^{n-1}}(\mathbf{s}) \\ \vdots & \vdots & & \vdots \\ a_n((\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}))) & x_{0n,s^1}(\mathbf{s}) & \cdots & x_{0n,s^{n-1}}(\mathbf{s}) \end{pmatrix} \neq 0$$

which is what is meant when it is said that the given set, $\left\{ (\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s})) : \mathbf{s} \in \prod_{i=1}^{n-1} [a_i, b_i] \right\}$ is not characteristic. The above discussion shows there exists a unique solution to this problem of a local solution to the PDE near a given non characteristic set. This problem is called the Cauchy problem. If the initial set is characteristic, you can see with a little effort that there may be more than one solution or maybe even no solution.

I emphasize again that if the functon, $\phi(\mathbf{x},z) \equiv z - u(\mathbf{x})$ is constant along characteristics, solutions to the system

$$x_{k,t} = a_k, z_{,t} = b,$$

then u is a solution to the PDE because

$$0 = \phi_{,t} = z_{,t} - \sum_k \frac{\partial u}{\partial x_k} \frac{\partial x_k}{\partial t} = b - \sum_k u_{,k} a_k.$$

Example 4.2 *Find the solution to the PDE* $(1 + x) u_x + y u_y = u$ *which satisfies* $u(x, x) = x$.

This is asking for the solution to this equation which contains the curve (s, s, s). Thus the equations for the characteristics are

$$x' = 1 + x,\ y' = y,\ z' = z,\ x(0, s) = s,\ y(0, s) = s,\ z(0, s) = s.$$

Hence $x(t, s) = -1 + (s + 1)e^t,\ y(t, s) = se^t,\ z(t, s) = se^t$. Now solving for s, t in terms of x, y, yields

$$t = \ln(x + 1 - y),\ s = \frac{y}{x + 1 - y}.$$

It follows the solution is

$$u(x, y) = z(t, s) = se^t = \frac{ye^{\ln(x+1-y)}}{x + 1 - y} = y.$$

You can see this works.

I hope you see that this method is no good in general because it requires you to find solutions to a system of possibly nonlinear ordinary differential equations and then to invert a system of nonlinear equations. You can't expect to be able to do these things using algebra. In other words you will be able to do the problems which have been cooked up to work out and that is about all. However, there are some interesting examples which do work.

Example 4.3 *Find the solution to the PDE* $uu_x + yu_y = x$ *which satisfies* $u(0, y) = 2y$.

Here the characteristic equations are

$$x' = z,\ y' = y,\ z' = x,\ x(0, s) = 0,\ y(0, s) = s,\ z(0, s) = 2s.$$

This is a pretty reasonable problem because it involves a linear system of equations.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}' = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, eigenvectors: $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \leftrightarrow -1, \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \leftrightarrow 1$The solution is

$$C_1(s) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{-t} + C_2(s) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^t + C_3(s) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} e^t$$

and it is necessary to choose the $C_k(s)$ such that the initial conditions hold. Thus it is necessary to solve

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 0 \\ s \\ 2s \end{pmatrix}$$

The solution is

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} -s \\ s \\ s \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} x(t, s) \\ y(t, s) \\ z(t, s) \end{pmatrix} = \begin{pmatrix} -se^{-t} + se^t \\ se^t \\ se^{-t} + se^t \end{pmatrix}$$

and the next step is to solve the system

$$x = -se^{-t} + se^{t}$$
$$y = se^{t}$$

for s and t in terms of x, y. This yields $t = -\ln\sqrt{1-\frac{x}{y}}$ and $s = y\sqrt{1-\frac{x}{y}}$. Therefore, the solution to the Cauchy problem is

$$u(x,y) = z(t,s) = y\sqrt{1-\frac{x}{y}}e^{\ln\sqrt{1-\frac{x}{y}}} + y\sqrt{1-\frac{x}{y}}e^{-\ln\sqrt{1-\frac{x}{y}}} = 2y - x.$$

You see this works. Just check it.

It should not be surprising that sometimes things just won't work out. Even in the case of ordinary differential equations it is sometimes necessary to write the solution implicitly. Suppose you have the equation,

$$a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u)$$

and you have somehow found two functions of three variables, ϕ and ψ which are constant along characteristics. Note that if you can find an explicit solution, $z = u(x,y)$, the function, $\phi(x,y,z) = z - u(x,y)$ works. Suppose also that $\nabla\phi \times \nabla\psi \neq 0$. This condition implies that the intersection of the two surfaces, $\phi(x,y,z) = \phi_0$ and $\psi(x,y,z) = \psi_0$ is locally a curve or is empty. This follows from the implicit function theorem. Now consider the implicitly defined surface, $F(\phi,\psi) = 0$. Consider the curve, $C = \{(\phi_0,\psi_0) : F(\phi_0,\psi_0) = 0\}$. Then the surface, $F(\phi,\psi) = 0$ can be written as

$$\cup_{(\phi_0,\psi_0)\in C}\,[\psi = \psi_0] \cap [\phi = \phi_0].$$

Consider the set, $[\psi = \psi_0] \cap [\phi = \phi_0]$. This is a curve due to the assumption that $\nabla\phi \times \nabla\psi \neq 0$. Now the characteristic curve through a point in this set must lie in both the surfaces $[\phi = \phi_0]$ and $[\psi = \psi_0]$ and so it must equal the intersection just written. Thus $F(\phi,\psi) = 0$ is a union of characteristic curves and so it defines a solution to the PDE implicitly. To illustrate, consider the previous example.

Example 4.4 *Find the solution to the PDE* $uu_x + yu_y = x$ *which satisfies* $u(0,y) = 2y$.

I want to find a couple of functions which are constant along characteristics. The characteristic curves are solutions of the ordinary differential equations,

$$x' = z, y' = y, z' = x.$$

Solving for dt this yields the system

$$\frac{dx}{z} = \frac{dy}{y} = \frac{dz}{x}.$$

Considering only the last two expressions,

$$xdx = zdz$$

and so along characteristics,

$$z^2 - x^2 = C_1.$$

Let the first function be $\phi(x,y,z) = z^2 - x^2$. This one is constant along characteristics. I need another such function. I just showed that along characteristics, $z^2 - x^2$ is a constant and so along characteristics, $z = \sqrt{C_1 + x^2}$. [1] Then going back to the equations,

$$\frac{dx}{\sqrt{C_1 + x^2}} = \frac{dy}{y}$$

[1] This process is really very sloppy. The idea is to mess around untill you get the two functions.

$$\int \frac{dx}{\sqrt{C_1 + x^2}} = \ln\left(x + \sqrt{(C_1 + x^2)}\right)$$

and so along characteristics,

$$\ln\left(x + \sqrt{(C_1 + x^2)}\right) - \ln y = C_2.$$

Thus another function is

$$\psi(x, y, z) = \ln\left(\frac{x + \sqrt{(C_1 + x^2)}}{y}\right).$$

From the above discussion, a solution will be

$$h\left(z^2 - x^2\right) = \ln\left(\frac{x + \sqrt{(C_1 + x^2)}}{y}\right)$$

assuming you can solve for ψ in terms of ϕ in the relation, $F(\phi, \psi) = 0$. Now it is just a matter of satisfying $u(0, y) = 2y$. Thus

$$h\left(4y^2\right) = \ln\left(\frac{\sqrt{C_1}}{y}\right)$$

and this tells what h should equal. Just let $s = 4y^2$ and so

$$h(s) = \ln\left(\frac{2\sqrt{C_1}}{\sqrt{s}}\right).$$

Consequently, the implicitly defined solution is

$$\ln\left(\frac{2\sqrt{C_1}}{\sqrt{z^2 - x^2}}\right) = \ln\left(\frac{x + \sqrt{(C_1 + x^2)}}{y}\right)$$

and so

$$\frac{2\sqrt{C_1}}{\sqrt{z^2 - x^2}} = \frac{x + \sqrt{(C_1 + x^2)}}{y}$$

Now recall that $C_1 = z^2 - x^2$. Then

$$2 = \frac{x + z}{y}$$

or in other words, $z = 2y - x$ which was found earlier.

Example 4.5 *Find an integrating factor for the ordinary differential equation,* $\left(x^2 + y^2\right) dx + xydy = 0$.

Recall that an integrating factor is a function of two variables which when you multiply the equation you get one which is exact. The function, u is an integrating factor exactly when

$$\left(\left(x^2 + y^2\right) u\right)_{,y} = (xyu)_{,x}.$$

This amounts to finding a solution to the following linear PDE

$$xyu_{,x} - \left(x^2 + y^2\right) u_{,y} = yu.$$

As above, the characteristics are solutions of

$$\frac{dx}{xy} = \frac{dy}{-\left(x^2+y^2\right)} = \frac{dz}{yz}.$$

Using the first and last terms, it follows $z-x = C_1$. Thus I will let $\phi = z-x$. Since ϕ is constant along characteristics, one solution is obtained by solving for z. Thus $u = z = x$. Now this should serve as an integrating factor. That is all that is needed.

From the above theory, you can see that the integrating factor approach is completely general from a theoretical point of view. That is, you can always reduce the given ODE to an exact ODE using an appropriate integrating factor. Of course the above statement is a lie because you won't be able to actually find the solution to the PDE in all cases. However, you do know it exists.

Example 4.6 *Find solutions to $uu_x + u_y = 0$, the inviscid Burger's equation.*

To find a lot of solutions, I will look for two functions which are constant along characteristics. The characteristic curves satisfy the differential equations

$$\frac{dx}{z} = \frac{dy}{1} = \frac{dz}{0}.$$

Thus z is constant with respect to t. One function is then $\phi(x,y,z) = z$. Then $dx = zdy$ and so $x = zy + C$ and so another function is $\psi(x,y,z) = zy - x$. Then a general solution would be $F(zy-x, z)$ and assuming F is such that the second variable can be solved for, this yields a general solution of the form

$$z = f(zy - x).$$

The solution then is defined implicitly by $u = f(uy-x)$. As an example, suppose $u(x,0) = \sin x$. Then

$$\sin x = f(-x)$$

and so $f(x) = -\sin x$. Thus the solution in this case would be $u = \sin(x - uy)$. Does it work?

$$uu_x + u_y = ?$$

From the chain rule,

$$u_x = \cos(x-uy)(1-yu_x),\ u_y = \cos(x-yu)(-u_y y - u).$$

Hence

$$\begin{aligned} u_x u + u_y &= u(\cos(x-uy)(1-yu_x)) + \cos(x-yu)(-u_y y - u) \\ &= u\cos(x-uy) - yuu_x - u_y y\cos(x-yu) - u\cos(x-yu) \\ &= -y\cos(x-yu)(u_x u + u_y) \end{aligned}$$

showing that $u_x u + u_y = 0$ as hoped.

4.2 Conservation Laws And Shocks

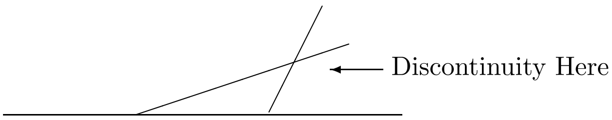
A PDE of the form $(G(u))_x + u_y = 0$ is called a conservation law. Typically, y refers to time and x refers to space. The inviscid Burger's equation is such an example with $G(u) = u^2$. The characteristic equations for such a conservation law are

$$\frac{dx}{dt} = G'(z),\ \frac{dy}{dt} = 1,\ \frac{dz}{dt} = 0.$$

Suppose you want a solution which contains the curve, $(s, 0, h(s))$. That is, you desire a solution to the conservation law which has $u(x, 0) = h(x)$. Then from the characteristic equations, you see easily that

$$z = h(s), \ y = t, \ x = G'(h(s))t + s.$$

In particular, z is a constant for a fixed value of s. Now consider $s_1 < s_2$ and suppose $G'(h(s_1)) > G'(h(s_2))$. Then it follows that on the two lines shown below z is constant, a different constant for each line.



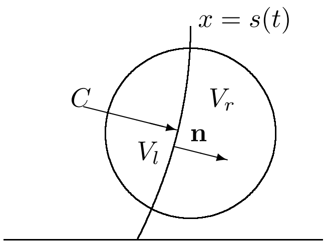
The problem is that the two lines intersect and so there are two values z is required to assume. There is a jump discontinuity in the solution. Of course this makes the partial differential equation inappropriate. This may give some idea of why up till now, all the solutions obtained have been local. However, there is another way to think of solutions to these equations which makes sense even though it might not make sense to talk about the partial derivatives. Let ϕ denote the restriction to $\mathbb{R} \times [0, \infty)$ of a function which is infinitely differentiable and equals zero outside some closed ball. The existence of such functions will be dealt with later. For now assume they exist. Then you multiply the differential equation by ϕ and do the integral,

$$\begin{aligned}
\int_{-\infty}^{\infty} \int_{0}^{\infty} ((G(u))_x + u_y)\, \phi dy dx &= \int_{0}^{\infty} \int_{-\infty}^{\infty} (G(u))_x\, \phi dx dy + \int_{-\infty}^{\infty} \int_{0}^{\infty} u_y \phi dy dx \\
&= -\int_{0}^{\infty} \int_{-\infty}^{\infty} G(u)\, \phi_x dx dy - \int_{-\infty}^{\infty} u(x, 0)\, \phi(x, 0)\, dx \\
-\int_{-\infty}^{\infty} \int_{0}^{\infty} u \phi_y dy dx &= 0
\end{aligned}$$

Since $u(x, 0) = h(x)$, this reduces to

$$\int_{0}^{\infty} \int_{-\infty}^{\infty} \left(G(u)\, \phi_x + u \phi_y\right) dx dy + \int_{-\infty}^{\infty} h(x)\, \phi(x, 0)\, dx = 0. \tag{4.6}$$

Now you notice that if you were to look for a solution, u which satisfied the above equation for all ϕ of the sort described, this would make perfect sense even if u wasn't continuous. Such solutions are sometimes called weak solutions or integral solutions. Now suppose you have a weak solution, u, that there is a smooth curve, $x = s(y)$ and to the left of it and right of it the function, u is C^1. Consider the following picture.



In this picture, $\mathbf{n} = (n_1, n_2)$ denotes the unit outer normal to V_l along the curve. Let $V = V_l \cup V_r$ and let ϕ be a test function which is infinitely differentiable and equals zero off some closed subset of V_l. Then the only thing with survives of the formula for the weak solution is

$$\int_{0}^{\infty} \int_{-\infty}^{\infty} \left(G(u)\, \phi_x + u \phi_y\right) dx dy = 0$$

and you can integrate by parts to obtain that

$$-\int_{V_l} (G(u)_x + u_y)\, \phi dx dy = 0.$$

Since this holds for arbitrary ϕ of the sort described, it follows that on V_l,

$$G(u)_x + u_y = 0$$

Similarly $G(u)_x + u_y = 0$ on V_r. Thus the partial differential equation is satisfied on $V_l \cup V_r$.

Now let ϕ be such a test function which is infinitely differentiable and vanishes off V. Thus ϕ might not vanish on the curve, $x = s(y)$ but it does vanish on the x axis. Then (4.6) reduces to

$$\int_{V_l} \left(G(u)\phi_x + u\phi_y\right) dA + \int_{V_r} \left(G(u)\phi_x + u\phi_y\right) dA = 0.$$

This equals

$$\int_{V_l} \Big((G(u)\phi)_x - G(u)_x \phi + (u\phi)_y - u_y\phi\Big) dA + \int_{V_r} \Big((G(u)\phi)_x - G(u)_x \phi + (u\phi)_y - u_y\phi\Big) dA$$

$$= \int_C G(u_l)\phi n_1 dl + \int_C u_l \phi n_2 dl - \int_C G(u_r)\phi n_1 dl - \int_C u_r \phi n_2 dl = 0$$

because of the divergence theorem and the above observation that u solves the partial differential equation on V_l and V_r. In this formula, u_l denotes the limit of u from the left and u_r denotes the limit of u from the right. Since ϕ is an arbitrary test function, it follows that along the curve, C,

$$(G(u_l) - G(u_r)) n_1 + (u_l - u_r) n_2 = 0$$

This is a very interesting relationship! From calculus,

$$\mathbf{n} = \left(\frac{1}{\sqrt{1 + s'(y)^2}}, \frac{-s'(y)}{\sqrt{1 + s'(y)^2}} \right)$$

and so at a point of C corresponding to y,

$$(G(u_l) - G(u_r)) \frac{1}{\sqrt{1 + s'(y)^2}} = (u_l - u_r) \left(\frac{s'(y)}{\sqrt{1 + s'(y)^2}} \right).$$

Written more simply, this says

$$[G(u)] = \frac{dx}{dy} [u] \tag{4.7}$$

where $[f] \equiv f_l - f_r$ denotes the jump. The formula (4.7) is called the Rankine Hugoniot jump condition.

The general theory of existence of weak solutions is pretty extensive. One good source for what has been done on these problems is the book by Smoller [19]. You can solve simple problems by using characteristics and that the solution should be constant along characteristics till they intersect a shock which can be located by the Rankine Hugoniot relation. When they cross, you can let the function have a jump. However, sometimes there is a region of the upper half plane which does not contain any characteristics. For example, suppose $u = u_-$ for $x < 0$ and $u = u_+$ for $x > 0$. Suppose also that $G'(u_-) < G'(u_+)$. Then there will be a wedge shaped region containing no characteristics. What should be the value of u in this region? Recall that the equation is $(G(u))_x + u_y = 0$. Now consider $u = \alpha\left(\frac{x}{y}\right)$ where α is chosen appropriately to solve the partial differential equation. Thus you need

$$G'\left(\alpha\left(\frac{x}{y}\right)\right) \alpha'\left(\frac{x}{y}\right) \frac{1}{y} + \alpha'\left(\frac{x}{y}\right) \left(\frac{-x}{y^2}\right) = 0.$$

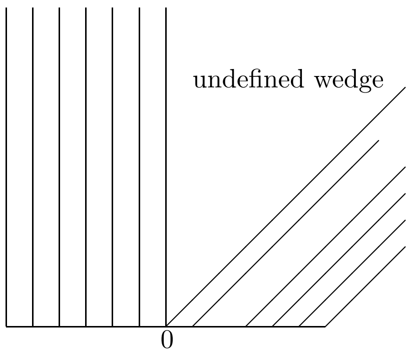
Cancelling $\alpha'\left(\frac{x}{y}\right)$, you get

$$G'\left(\alpha\left(\frac{x}{y}\right)\right) = \frac{x}{y}$$

and so it is appropriate to take $\alpha(r) = (G')^{-1}(r)$.

Example 4.7 *Burger's equation,* $\left(\frac{u^2}{2}\right)_x + u_y = 0$. *For initial condition, take* $u(x,0) = 0$ *if* $x < 0$ *and* $u(x,0) = 1$ *if* $x > 0$.

The characteristic equations are then $\frac{dx}{dt} = z$, $\frac{dy}{dt} = 1$, and $\frac{dz}{dt} = 0$. Therefore, z equals a constant and so $x(t) = zt + s$ while $y = t$. For $x < 0$, it follows $z = 0$ and $x = s$, while $y = t$. Thus these characteristics are straight vertical lines. For $x > 0$, $z = 1$ and so $x = t + s$ while $y = t$. Thus the characteristics are lines of slope 1 passing through the points of the positive x axis. In this case, no shocks form because there are no intersections of characteristics. The wedge between $x = 0$ and $y = x$ is left out, however. In this case, $G'(u) = u$ and so α above just equals the identity map. Therefore, you should let $u = x/y$ in the wedge and you will have a solution to the partial differential equation. In the second quadrant, let $u = 0$. In the first quadrant to the right of $y = x$, you have $u = 1$ and in the wedge, you have $u = x/y$.



Example 4.8 *In the example 4.7 find another weak solution.*

This is not hard to do. You could look for a shock solution. On the left side the value of u will be 0 and on the right it will have the value 1. Then by the Rankine Hugoniot equation,

$$-\frac{1}{2} = \frac{dx}{dy}(-1).$$

Thus you need to have $\frac{dx}{dy} = \frac{1}{2}$. Lets define $u(x,y) = 0$ if $x < \frac{y}{2}$ and $u(x,y) = 1$ if $x > \frac{y}{2}$. Is it a weak solution? In the two regions it works out fine and the Rankine Hugoniot relation holds so it must end up being a weak solution and it is very different than the one in Example 4.7. This shows very clearly that the weak solution to these conservation laws may not be unique!

Example 4.9 *Suppose* $G(u) = u(1-u)$ *and consider* $(G(u))_x + u_y = 0$ *with the initial condition,*

$$u(x,0) = \begin{cases} 1/2 \text{ if } x < 0 \\ 1 \text{ if } x > 0 \end{cases} = h(x).$$

Thus there is initially a shock. Lets find the location of this shock as a function of y.

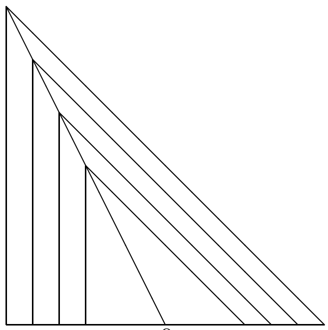
Recall the characteristic curve which goes through $(s, 0, h(s))$ is given by

$$z = h(s), \; y = t, \; x = G'(h(s))\, t + s$$

It follows from this that the characteristic curves through points on the negative x axis are vertical lines while those through points on the positive x axis are lines having slope equal to -1. To the right of the shock, the value of u would be 1 and to the left it would be $1/2$. Therefore, $[G(u)] = 1/4$ and $[u] = -1/2$. It follows from the Rankine Hugoniot relation,

$$[G(u)] = \frac{dx}{dy}[u]$$

that $x'(y) = -1/2$. Therefore, the shock would be a line through the origin which has slope -2. Thus $y = -2x$ gives the shock. Here is a picture.



4.3 Nonlinear First Order PDE

This is much harder. The PDE will be of the form

$$F(\mathbf{x}, u, \nabla u) = 0. \tag{4.8}$$

and I will look for solutions, u which are C^2. Furthermore, I will assume F is as smooth as it needs to be.

First I will show the appropriate generalization of the concept of characteristics. The idea of a characteristic is that F is constant along it. Let $\mathbf{p}$ denote ∇u and z denote u. Then in this case the characteristics will involve considering $\mathbf{x}$, $\mathbf{p}$, and z as solutions of ordinary differential equations. However, there must be a relationship between $\mathbf{p}$ and z if in the end, we take $u(\mathbf{x}(t)) = z(t)$ and $\mathbf{p} = \nabla u$. This requires

$$z'(t) = \sum_k u_{,k}(\mathbf{x}(t))\, x_k'(t) = \sum_k p_k(t)\, x_k'(t). \tag{4.9}$$

Now in terms of $\mathbf{p}$ and z, we need

$$F(\mathbf{x}, z, \mathbf{p}) = 0. \tag{4.10}$$

Also, $p_k = u_{,k}$ and so

$$p_k'(t) = \sum_i u_{,ki}(\mathbf{x}(t))\, x_i'(t) \tag{4.11}$$

What about $u_{,ki}$? This is where you refer to the PDE again in (4.8). Differentiate with respect to x_k

$$\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} u_{,k} + \sum_i \frac{\partial F}{\partial p_i} \frac{\partial p_i}{\partial x_k} = 0. \tag{4.12}$$

Now since u is C^2,

$$\frac{\partial p_i}{\partial x_k} = u_{,ik} = u_{,ki}$$

and so (4.12) is of the form

$$\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} u_{,k} + \sum_i \frac{\partial F}{\partial p_i} u_{,ki} = 0.$$

Therefore, letting

$$x_i' = \frac{\partial F}{\partial p_i}, \tag{4.13}$$

it follows from (4.12) and (4.11) that

$$\begin{aligned} p_k'(t) &= \sum_i u_{,ki}(\mathbf{x}(t))\, x_i'(t) = \sum_i u_{,ki}(\mathbf{x}(t)) \frac{\partial F}{\partial p_i} \\ &= -\left(\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} u_{,k} \right) = -\left(\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} p_k \right) \end{aligned} \tag{4.14}$$

This gives differential equations for $\mathbf{p}$ and $\mathbf{x}$. The remaining differential equation for z comes from (4.9) and (4.13). Thus,

$$z'(t) = \sum_k p_k(t)\, x_k'(t) = \sum_k p_k \frac{\partial F}{\partial p_k}. \tag{4.15}$$

You should notice that it is not just a curve in $\mathbb{R}^{n+1}$, $(\mathbf{x}(t), z(t))$, which is important in this context but a curve in $\mathbb{R}^{2n+1}$, $(\mathbf{x}(t), z(t), \mathbf{p}(t))$. This curve is referred to as a characteristic strip.

Definition 4.10 *Characteristic strips for the equation (4.10) are solutions of the differential equations, (4.13), (4.14), and (4.15).*

Lemma 4.11 *The function, $F(\mathbf{x}, z, \mathbf{p})$ is constant along any characteristic strip.*

Proof: Just differentiate with respect to t. Thus

$$\sum_i \frac{\partial F}{\partial x_i} x_i' + \frac{\partial F}{\partial z} z' + \sum_i \frac{\partial F}{\partial p_i} p_i' \tag{4.16}$$

$$= \sum_i \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial p_i} + \frac{\partial F}{\partial z} \sum_i p_i \frac{\partial F}{\partial p_i} - \sum_i \frac{\partial F}{\partial p_i} \left(\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial z} p_i \right) = 0. \tag{4.17}$$

Now let Γ be given by

$$\Gamma \equiv \left\{ (\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}), \mathbf{p}_0(\mathbf{s})) : \mathbf{s} \in \prod_{i=1}^{n-1} [a_i, b_i] \right\}$$

and I will assume the functions, $\mathbf{x}_0, z_0$, and $\mathbf{p}_0$ are all C^2. Then define $\mathbf{x}(t, \mathbf{s}), z(t, \mathbf{s})$, and $\mathbf{p}(t, \mathbf{s})$ as solutions to the following initial value problem.

$$\begin{aligned} x_{k,t} &= \frac{\partial F}{\partial p_k}, p_{k,t} = -\left(\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} p_k \right), z_{,t} = \sum_k p_k \frac{\partial F}{\partial p_k}, && (4.18) \\ x_k(0, \mathbf{s}) &= \mathbf{x}_0(s), p_k(0, \mathbf{s}) = p_{0k}(\mathbf{s}), z(0, \mathbf{s}) = z_0(\mathbf{s}). && (4.19) \end{aligned}$$

If we assume F and its partial derivatives are in C^2, then it will follow that $\mathbf{x}(t,\mathbf{s}), z(t,\mathbf{s})$, and $\mathbf{p}(t,\mathbf{s})$ will be C^2 functions of $\mathbf{s}$ and t and that these functions are defined on a set of the form

$$[-T,T]\times\prod_{i=1}^{n-1}[a_i,b_i]$$

It will always be assumed that F is smooth. Suppose now that

$$F(\mathbf{x}_0(\mathbf{s}),z_0(\mathbf{s}),\mathbf{p}_0(\mathbf{s}))=0. \tag{4.20}$$

Then it follows from Lemma 4.11 that for all $(t,\mathbf{s})\in[-T,T]\times\prod_{i=1}^{n-1}[a_i,b_i]$,

$$F(\mathbf{x}(t,\mathbf{s}),z(t,\mathbf{s}),\mathbf{p}(t,\mathbf{s}))=0.$$

Assume an appropriate condition which will allow the inversion of the system, $\mathbf{x}=\mathbf{x}(t,\mathbf{s})$ and solve for t and $\mathbf{s}$ in terms of $\mathbf{x}$. Thus assume

$$\frac{\partial(x_1,\cdots,x_n)}{\partial(t,s_1,s_2,\cdots,s_{n-1})}(0,\mathbf{s})\neq 0,\ \mathbf{s}\in\prod_{i=1}^{n-1}[a_i,b_i].$$

In other words,

$$\det\begin{pmatrix}\frac{\partial F(\mathbf{x}_0(\mathbf{s}),z_0(\mathbf{s}),\mathbf{p}_0(\mathbf{s}))}{\partial p_1} & \frac{\partial x_{01}}{\partial s^1}(\mathbf{s}) & \cdots & \frac{\partial x_{01}}{\partial s^{n-1}}(\mathbf{s})\\ \vdots & \vdots & & \vdots\\ \frac{\partial F(\mathbf{x}_0(\mathbf{s}),z_0(\mathbf{s}),\mathbf{p}_0(\mathbf{s}))}{\partial p_n} & \frac{\partial x_{0n}}{\partial s^1}(\mathbf{s}) & \cdots & \frac{\partial x_{0n}}{\partial s^{n-1}}(\mathbf{s})\end{pmatrix}\neq 0. \tag{4.21}$$

Then it follows that

$$\frac{\partial(x_1,\cdots,x_n)}{\partial(t,s_1,s_2,\cdots,s_{n-1})}(t,\mathbf{s})\neq 0$$

whenever t is small enough and so, let T be small enough that this happens on $[-T,T]\times\prod_{i=1}^{n-1}[a_i,b_i]$. It follows from the inverse function theorem that $\mathbf{x}=\mathbf{x}(t,\mathbf{s})$ defines t and $\mathbf{s}$ locally as C^2 functions of $\mathbf{x}$. Therefore, it is possible to write

$$u(\mathbf{x})=z(t,\mathbf{s}).$$

However, it is not clear that $u_{,i}$ equals p_i. In fact, there is no reason to suppose this at all. Therefore, something more must be required than (4.20). If $u_{,i}=p_i$, then as pointed out earlier

$$z_{,t}(t,\mathbf{s})=\sum_k p_k(t,\mathbf{s})\,x_{k,t}(t,\mathbf{s})$$

and this holds because of the differential equations. Thus

$$z_{,t}(t,\mathbf{s})=\sum_k p_k(t,\mathbf{s})\frac{\partial F}{\partial p_k}(t,\mathbf{s})=\sum_k p_k(t,\mathbf{s})\,x_{k,t}(t,\mathbf{s}).$$

However, it is also necessary that a similar formula must hold for differentiations with respect to s_k. In particular, it must be that

$$\frac{\partial z(t,\mathbf{s})}{\partial s_k}=\sum_i\frac{\partial u(\mathbf{x}(t,\mathbf{s}))}{\partial x_i}\frac{\partial x_i(t,\mathbf{s})}{\partial s_k}=\sum_i p_i(t,\mathbf{s})\frac{\partial x_i(t,\mathbf{s})}{\partial s_k}. \tag{4.22}$$

In the case where $t=0$ this reduces to the following strip condition.

$$\frac{\partial z_0(\mathbf{s})}{\partial s_k}=\sum_i p_{0i}(\mathbf{s})\frac{\partial x_{0i}(\mathbf{s})}{\partial s_k}. \tag{4.23}$$

Lemma 4.12 *Suppose the ordinary differential equations, (4.18) and (4.19) are valid and assume the strip conditions, (4.20) and (4.23). Then (4.22) holds for* $(t, \mathbf{s}) \in [-T, T] \times \prod_{i=1}^{n-1} [a_i, b_i]$ *for some* T *sufficiently small.*

Proof: First note the ordinary differential equations hold for $(t, \mathbf{s}) \in [-T, T] \times \prod_{i=1}^{n-1} [a_i, b_i]$ for some T sufficiently small. It remains only to verify the identity. Following John, [13], let

$$A(t, \mathbf{s}) = \frac{\partial z(t, \mathbf{s})}{\partial s_k} - \sum_i p_i(t, \mathbf{s}) \frac{\partial x_i(t, \mathbf{s})}{\partial s_k}.$$

By the assumption that (4.23) it follows that

$$A(0, \mathbf{s}) = 0.$$

Also,

$$\begin{aligned}
A_{,t} &= \frac{\partial^2 z(t, \mathbf{s})}{\partial s_k \partial t} - \sum_i \frac{\partial p_i(t, \mathbf{s})}{\partial t} \frac{\partial x_i(t, \mathbf{s})}{\partial s_k} - \sum_i p_i(t, \mathbf{s}) \frac{\partial^2 x_i}{\partial s_k \partial t} \\
&= z_{,ts_k} - \sum_i (p_i x_{i,t})_{,s_k} + \sum_i p_{i,s_k} x_{i,t} - \sum_i p_{i,t} x_{i,s_k} \\
&= \frac{\partial}{\partial s_k} \left(\overbrace{z_{,t} - \sum_i p_i x_{i,t}}^{=0} \right) + \sum_i p_{i,s_k} x_{i,t} - \sum_i p_{i,t} x_{i,s_k} \\
&= \sum_i p_{i,s_k} \frac{\partial F}{\partial p_i} + \sum_i \left(\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial z} p_i \right) x_{i,s_k} \\
&= \overbrace{\sum_i p_{i,s_k} \frac{\partial F}{\partial p_i} + \sum_i \frac{\partial F}{\partial x_i} x_{i,s_k} + \frac{\partial F}{\partial z} z_{,s_k}}^{=\frac{\partial F}{\partial s_k}} - \frac{\partial F}{\partial z} z_{,s_k} + \sum_i p_i \frac{\partial F}{\partial z} x_{i,s_k} \\
&= \frac{\partial F}{\partial s_k} - \frac{\partial F}{\partial z} z_{,s_k} + \sum_i p_i \frac{\partial F}{\partial z} x_{i,s_k}.
\end{aligned}$$

Now since

$$F(\mathbf{x}(t, \mathbf{s}), z(t, \mathbf{s}), \mathbf{p}(t, \mathbf{s})) = 0,$$

it follows that $\frac{\partial F}{\partial s_k} = 0$ and so

$$A_{,t} = -\frac{\partial F}{\partial z} (A(t, \mathbf{s})), \ A(0, \mathbf{s}) = 0.$$

Consequently, $A(t, \mathbf{s}) = 0$ and this proves the lemma.

Now it is possible to give the following existence theorem.

Theorem 4.13 *Suppose the hypotheses of Lemma 4.12,*

$$\begin{aligned}
x_{k,t} &= \frac{\partial F}{\partial p_k}, p_{k,t} = -\left(\frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial z} p_k \right), z_{,t} = \sum_k p_k \frac{\partial F}{\partial p_k}, && (4.24) \\
x_k(0, \mathbf{s}) &= \mathbf{x}_0(s), p_k(0, \mathbf{s}) = p_{0k}(\mathbf{s}), z(0, \mathbf{s}) = z_0(\mathbf{s}). && (4.25)
\end{aligned}$$

and the condition for solving $t, \mathbf{s}$ in terms of $\mathbf{x}$, (4.21),

$$\det\begin{pmatrix} \frac{\partial F(\mathbf{x}_0(\mathbf{s}),z_0(\mathbf{s}),\mathbf{p}_0(\mathbf{s}))}{\partial p_1} & \frac{\partial x_{01}}{\partial s^1}(\mathbf{s}) & \cdots & \frac{\partial x_{01}}{\partial s^{n-1}}(\mathbf{s}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial F(\mathbf{x}_0(\mathbf{s}),z_0(\mathbf{s}),\mathbf{p}_0(\mathbf{s}))}{\partial p_n} & \frac{\partial x_{0n}}{\partial s^1}(\mathbf{s}) & \cdots & \frac{\partial x_{0n}}{\partial s^{n-1}}(\mathbf{s}) \end{pmatrix} \neq 0 \tag{4.26}$$

along with the strip conditions

$$F(\mathbf{x}_0(\mathbf{s}), z_0(\mathbf{s}), \mathbf{p}_0(\mathbf{s})) = 0. \tag{4.27}$$

$$\frac{\partial z_0(\mathbf{s})}{\partial s_k} = \sum_i p_{0i}(\mathbf{s}) \frac{\partial x_{0i}(\mathbf{s})}{\partial s_k}. \tag{4.28}$$

Then for T small enough, it is possible to define

$$u(\mathbf{x}) \equiv z(t, \mathbf{s}) \equiv z(t(\mathbf{x}), \mathbf{s}(\mathbf{x}))$$

and for $(t,\mathbf{s}) \in [-T,T] \times \prod_{i=1}^{n-1}[a_i,b_i]$, $u_{x_i} = p_i$ and so there exists a local solution to the partial differential equation,

$$F(\mathbf{x}, u, \nabla u) = 0$$

which contains the set,

$$(\mathbf{x}_0(\mathbf{s}), z(\mathbf{s}), \mathbf{p}(\mathbf{s})),\ \mathbf{s} \in \prod_{i=1}^{n-1}[a_i,b_i].$$

Proof: Let T be small enough that

$$\frac{\partial(x_1, \cdots, x_n)}{\partial(t, s_1, s_2, \cdots, s_{n-1})}(t,\mathbf{s}) \neq 0 \tag{4.29}$$

for $(t,\mathbf{s}) \in [-T,T] \times \prod_{i=1}^{n-1}[a_i,b_i]$.

It only remains to verify that $p_i = u_{x_i}$. From Lemma 4.12 and the assumed ordinary differential equation solved by z,

$$z_{,s_k} = \sum_{i=1}^n p_i x_{i,s_k},\ z_{,t} = \sum_{i=1}^n p_i x_{i,t}.$$

But also from the definition of u in terms of z,

$$z_{,s_k} = \sum_{i=1}^n u_{,x_i} x_{i,s_k},\ z_{,t} = \sum_{i=1}^n u_{,x_i} x_{i,t}.$$

By (4.29) $p_i = u_{x_i}$. This proves the local existence theorem.

Example 4.14 *Letting $n = 2$, find a solution to to $u_x u_y = u$ which satisfies $u(0,y) = y^3$.*

In this case $F(\mathbf{x}, z, \mathbf{p}) = pq - z$. The first thing needed is to consider the given condition that $u(0,y) = y^3$ as part of a strip condition. Thus you must identify functions, $p_0(s)$ and $q_0(s)$ such that $\left(0, s, s^3, p_0(s), q_0(s)\right)$ satisfies the strip conditions

$$p_0(s) q_0(s) - s^3 = 0,\ \overbrace{3s^2}^{z_{0,s}} = p_0(s) \cdot \overbrace{0}^{x_{0,s}} + q_0(s) \cdot \overbrace{1}^{y_{0,s}}.$$

Hopefully you can find such functions although there is no gaurantee that this is the case. Also, there may be more than one possible choice for such functions. Here $q_0(s) = 3s^2$ and then when you see this is the case, the first equation indicates that $p_0(s) = \frac{1}{3}s$. Thus the initial strip is $\left(0, s, s^3, \frac{1}{3}s, 3s^2\right)$. Now it is time to write the ordinary differential equations for the characteristic strips in (4.24) and (4.25). These are

$$\frac{\partial x}{\partial t} = q,\ \frac{\partial y}{\partial t} = p,\ \frac{\partial z}{\partial t} = 2pq,\ \frac{\partial p}{\partial t} = p,\ \frac{\partial q}{\partial t} = q. \tag{4.30}$$

From the last two equations, it follows

$$p(t,s) = C(s)e^t,\ q(t,s) = D(s)e^t$$

where $C(s)$ and $D(s)$ must be chosen in such a way that the initial data are satisfied. Thus

$$\begin{aligned} p(0,s) &= p_0(s) = \frac{1}{3}s = C(s) \\ q(0,s) &= q_0(s) = 3s^2 = D(s). \end{aligned}$$

Thus $p(t,s) = \frac{1}{3}se^t,\ q(t,s) = 3s^2e^t$. Some progress has been made. Now the third equation in (4.30) becomes

$$\begin{aligned} \frac{\partial z}{\partial t} &= 2pq = 2\left(\frac{1}{3}se^t\right)\left(3s^2e^t\right) \\ &= 2s^3e^{2t} \end{aligned}$$

and so

$$z(t,s) = s^3e^{2t} + K(s)$$

where $K(s)$ needs to be chosen in such a way as to satisfy the initial condition. Therefore,

$$z(0,s) = s^3 + K(s) = s^3$$

and so $K(s) = 0$. Hence

$$z(t,s) = s^3e^{2t} \tag{4.31}$$

It remains to solve the initial value problems for x and y. Thus

$$\frac{\partial x}{\partial t} = q = 3s^2e^t$$

and so

$$x(t,s) = 3s^2e^t + L(s)$$

where $L(s)$ must be chosen in such a way as to satisfy the initial condition. Therefore,

$$x(0,s) = 3s^2 + L(s) = 0$$

and so $L(s) = -3s^2$ and

$$x(t,s) = 3s^2e^t - 3s^2. \tag{4.32}$$

The differential equation for y is

$$\frac{\partial y}{\partial t} = p = \frac{1}{3}se^t$$

and so $y(t,s) = \frac{1}{3}se^t + M(s)$ where $M(s)$ must be chosen to satisfy the initial condition. Thus

$$y(0,s) = \frac{1}{3}s + M(s) = s$$

and so

$$M(s) = \frac{2}{3}s.$$

Thus

$$y(t,s) = \frac{1}{3}se^t + \frac{2}{3}s. \tag{4.33}$$

Now you use (4.31) - (4.33) to find the solution $z(t,s) = u(x,y)$.

Listing these equations,

$$\begin{aligned} x &= 3s^2e^t - 3s^2 \\ 3y &= s\left(e^t + 2\right) \\ z &= s^3e^t \end{aligned}$$

The idea is to solve the first two for s and t in terms of x and y.

$$\begin{aligned} x &= 3s^2e^t - 3s^2 \\ 3y &= s\left(e^t + 2\right) \end{aligned}$$

After much affliction and suffering you find

$$t = \ln\left(\frac{2\left(6y - \sqrt{9y^2 - 4x}\right)}{3y + \sqrt{9y^2 - 4x}}\right),\ s = \frac{3y + \sqrt{9y^2 - 4x}}{6}.$$

Therefore, the solution is

$$\begin{aligned} u(x,y) &= z(t,s) = s^3e^{2t} \\ &= \left(\frac{3y + \sqrt{9y^2 - 4x}}{6}\right)^3 \left(\frac{2\left(6y - \sqrt{9y^2 - 4x}\right)}{3y + \sqrt{9y^2 - 4x}}\right)^2 \\ &= \frac{1}{54}\left(3y + \sqrt{(9y^2 - 4x)}\right)\left(6y - \sqrt{(9y^2 - 4x)}\right)^2 \end{aligned}$$

which is probably not the first thing you would have thought of.

Does it work?

$$u(0,y) = \frac{1}{54}\left(3y + \sqrt{9y^2}\right)\left(6y - \sqrt{9y^2}\right)^2 = y^3$$

so it does satisfy the given condition. What about the equation? Using the above formula,

$$\begin{aligned} u_x &= \frac{2}{3}y - \frac{1}{9}\sqrt{(9y^2 - 4x)} \\ u_y &= \frac{1}{6}\left(3y + \sqrt{(9y^2 - 4x)}\right)\left(6y - \sqrt{(9y^2 - 4x)}\right) \end{aligned}$$

And

$$\begin{aligned} u_x u_y - u &= \left(\frac{2}{3}y - \frac{1}{9}\sqrt{(9y^2 - 4x)}\right)\left(\frac{1}{6}\left(3y + \sqrt{(9y^2 - 4x)}\right)\left(6y - \sqrt{(9y^2 - 4x)}\right)\right) \\ &\quad -\frac{1}{54}\left(3y + \sqrt{(9y^2 - 4x)}\right)\left(6y - \sqrt{(9y^2 - 4x)}\right)^2 \\ &= 0 \end{aligned}$$

so it also satisfies the partial differential equation.

I hope you see that this was just lucky that the equations could be solved for t, s in terms of x, y using methods of algebra. In general, you can't do this at all. The inverse function theorem does not hold because of some algebraic trick. It tells you something exists but not how to find it.

4.3.1 Wave Propagation

Suppose you have a wave which propagates in two dimensions, the wave front being the level surface,

$$u(x, y) = t$$

where t is the time. Picking a point (x, y) on the edge of this wave front and supposing that the speed of the wave is $c(x, y)$,

$$\dot{x}^2 + \dot{y}^2 = c^2. \tag{4.34}$$

Also,

$$u(x(t), y(t)) = t$$

so

$$u_x \dot{x} + u_y \dot{y} = 1. \tag{4.35}$$

Also, the velocity would satisfy

$$(\dot{x}, \dot{y}) = k\nabla u \tag{4.36}$$

for some constant, k. Then from (4.35),

$$k\left(u_x^2 + u_y^2\right) = 1$$

and from (4.34) and (4.36),

$$k^2 u_x^2 + k^2 u_y^2 = c^2$$

and so $k = c^2$. Thus the partial differential equation satisfied by u would be

$$c^2\left(u_x^2 + u_y^2\right) = 1 \tag{4.37}$$

which is called the eikonal equation.

It has some very interesting properties.

Example 4.15 *Find a solution to (4.37) assuming c is a constant which contains the curve* $(x_0(s), y_0(s), z_0(s))$.

First you need to complete the curve into a strip. Thus you need to find $p_0(s)$ and $q_0(s)$ such that

$$(x_0(s), y_0(s), z_0(s), p_0(s), q_0(s))$$

satisfies the strip conditions.

$$c^2\left(p_0(s)^2 + q_0(s)^2\right) = 1, \tag{4.38}$$

$$z_0'(\mathbf{s}) = p_0(s)\, x_0'(s) + q_0(s)\, y_0'(s). \tag{4.39}$$

These equations are somewhat problematic. By the Cauchy Schwarz inequality applied to the second and then using the first, you must have

$$\begin{aligned} |z_0'(s)| &\leq \sqrt{p_0(s)^2 + q_0(s)^2}\sqrt{x_0'(s)^2 + y_0'(s)^2} \\ &\leq \frac{1}{c}\sqrt{x_0'(s)^2 + y_0'(s)^2} \end{aligned}$$

and so in order to solve (4.38) and (4.39) you must have

$$z_0'(s)^2 c^2 \leq x_0'(s)^2 + y_0'(s)^2.$$

When $z_0'(s)^2 c^2 < x_0'(s)^2 + y_0'(s)^2$ the initial curve is called space like. If the inequality is turned around, it is called time like. Suppose therefore, that

$$z_0'(s)^2 c^2 < x_0'(s)^2 + y_0'(s)^2.$$

The simplest case of this is when the initial curve lies in the xy plane and likely the most interesting case would be where this curve is a circle. Thus the curve would be of the form

$$(r\cos s, r\sin s, 0).$$

Then from (4.39) and (4.38),

$$0 = -p_0(s)\sin s + q_0(s)\cos s,\ c^2\left(p_0(s)^2 + q_0(s)^2\right) = 1$$

and this is solved if

$$p_0(s) = c^{-1}\cos s,\ q_0(s) = c^{-1}\sin s \tag{4.40}$$

or

$$p_0(s) = -c^{-1}\cos s,\ q_0(s) = -c^{-1}\sin s. \tag{4.41}$$

These lead to two different solutions to the problem. First consider (4.40). The characteristic equations are

$$\begin{aligned} \frac{\partial x}{\partial t} &= 2c^2 p,\ \frac{\partial y}{\partial t} = 2c^2 q,\ \frac{\partial z}{\partial t} = 2c^2\left(p^2 + q^2\right) \\ \frac{\partial p}{\partial t} &= 0,\ \frac{\partial q}{\partial t} = 0. \end{aligned}$$

Now from the initial conditions,

$$q = c^{-1}\sin s,\ p = c^{-1}\cos s,$$

$$x(t,s) = 2ct\cos s + r\cos s,\ y(t,s) = 2ct\sin s + r\sin s,\ z(t,s) = 2t.$$

Thus $(2ct+r)^2 = x^2 + y^2$ and so

$$t = \frac{\sqrt{x^2+y^2} - r}{2c}$$

which implies

$$u(x,y) = z(t,s) = 2\frac{\sqrt{x^2+y^2} - r}{2c} = \frac{\sqrt{x^2+y^2} - r}{c}.$$

Remember that the level surfaces, $u = t$ gave the wave front at time t. Therefore, if ρ is the distance of this wave front from the origin at time t, the above formula shows

$$\frac{\rho - r}{c} = t$$

and so

$$\frac{d\rho}{dt} = c$$

showing that the speed of the wave front equals c.

4.3.2 Complete Integrals

Here I will present another method for finding lots of solutions to $F(\mathbf{x}, z, \mathbf{p}) = 0$ involving something called a complete integral. I am following the treatment of this subject which is found in the partial differential equations book by Evans.

Definition 4.16 *Let $\mathbf{a} \in A$, an open set in $\mathbb{R}^n$ and let $\mathbf{x} \in U$, an open set in $\mathbb{R}^n$. Then $u(\mathbf{x},\mathbf{a})$ is called a complete integral of $F(\mathbf{x}, z, \mathbf{p}) = 0$ if for every $\mathbf{a}$, the function $\mathbf{x} \to \mathbf{u}(\mathbf{x},\mathbf{a})$ is a C^2 solution of the P.D.E. and*

$$\text{rank}\begin{pmatrix} u_{,a_1} & u_{,x_1a_1} & \cdots & u_{,x_1a_n} \\ \vdots & \vdots & & \vdots \\ u_{,a_n} & u_{,x_na_1} & \cdots & u_{,x_na_n} \end{pmatrix} = n. \tag{4.42}$$

Why the funny condition on the rank of the above $n \times (n+1)$ matrix? Consider the following system of nonlinear equations

$$\begin{gathered} u(\mathbf{x},\mathbf{a}) = z \\ u_{,x_1}(\mathbf{x},\mathbf{a}) = p_1 \\ \vdots \\ u_{,x_n}(\mathbf{x},\mathbf{a}) = p_n \end{gathered}$$

where $\mathbf{x}$ is fixed. The condition is exactly what is needed to pick n of the above equations and apply the inverse function theorem to determine locally a one to one and onto relationship between $\mathbf{a}$ and n of the variables $\{z, p_1, \cdots, p_n\}$. If there exists an open set, $V \subseteq \mathbb{R}^p$ for $p < n$ and a C^1 function, ψ such that for $\mathbf{b} \in \mathbb{R}^p, \psi(\mathbf{b}) = \mathbf{a}$, then from the application of the inverse function theorem just mentioned, there would be a one to one and onto C^1 function, ϕ such that $\phi \circ \psi$ would map V, an open set in $\mathbb{R}^p$ onto an open set in $\mathbb{R}^n$. But C^1 functions can't do this. Therefore, the components of $\mathbf{a}$ are all needed. You could not write $u(\mathbf{x},\mathbf{a}) = v(\mathbf{x},\mathbf{b})$ where $\mathbf{b} \in \mathbb{R}^p$ for $p < n$.

Definition 4.17 *Suppose* $\mathbf{x} \to u(\mathbf{x}, \mathbf{a})$ *is a solution to* $F(\mathbf{x}, z, \mathbf{p}) = 0$ *for each* $\mathbf{a} \in A$, *an open set in* $\mathbb{R}^m$. *Consider the* m *equations*

$$\frac{\partial u}{\partial a_k}(\mathbf{x}, \mathbf{a}) = 0, \ k = 1, \cdots, m \tag{4.43}$$

and suppose it is possible to solve for $\mathbf{a}$ *in terms of* $\mathbf{x}$ *in these equations. By the implicit function theorem this would occur if*

$$\frac{\partial (u_{,a_1} \cdots u_{,a_m})}{\partial (a_1, \cdots, a_m)} \neq 0$$

and u *is* C^2. *Then, writing* $\mathbf{a} = \mathbf{a}(\mathbf{x})$ *and substituting in to* $u(\mathbf{x}, \mathbf{a})$ *to obtain*

$$v(\mathbf{x}) \equiv u(\mathbf{x}, \mathbf{a}(\mathbf{x})),$$

v *is called an envelope of the functions,* $u(\mathbf{x}, \mathbf{a})$.

The interesting thing about the envelope is that it is a solution of the partial differential equation.

Theorem 4.18 *Let* $u(\mathbf{x}, \mathbf{a})$ *be as described in Definition 4.17 and let* $\mathbf{v}$ *be the envelope defined there. Then* v *is also a solution to the partial differential equation,* $F(\mathbf{x}, z, \mathbf{p}) = 0$.

Proof: Compute $v_{,x_k}$. By the chain rule,

$$v_{,x_k} = u_{,x_k} + \sum_i u_{,a_i} a_{i,x_k} = u_{,x_k}$$

because of the equations (4.43) which assure that $u_{,a_i} = 0$.

It follows that v solves the partial differential equation.

Now suppose $u(\mathbf{x}, \mathbf{a})$ is a complete integral for $F(\mathbf{x}, z, \mathbf{p}) = 0$ as in Definition 4.16. Let $\mathbf{a} = (\mathbf{a}', a_n)$. Then let $h(\mathbf{a}') = a_n$ where h is an arbitrary function. Let v_h be the envelope of $u(\mathbf{x}, \mathbf{a}', h(\mathbf{a}'))$. Then v_h is a solution to the partial differential equation which depends on the arbitrary function, h.

Example 4.19 *Find a complete integral for the eikonal equation,* $c^2(p^2 + q^2) = 1$ *and use it to obtain some solutions to the partial differential equation.*

Look for solutions which are of the form $u = X(x) + Y(y)$. Then plugging in to the PDE,

$$c^2\left((X')^2 + (Y')^2\right) = 1.$$

There are many ways to proceed from here. One way would be to separate the variables. Here is another. The equation says that (cX', cY') is a point on the unit circle. Therefore, there exists a such that

$$cY' = \sin a, \quad cX' = \cos a.$$

Therefore, $X(x) = \frac{1}{c}\cos(a)x + c_1$ and $Y(y) = \frac{1}{c}\sin(a)y + c_2$. Therefore, combining the two c_i into one constant, a complete integral would be

$$u = \frac{1}{c}\cos(a)x + \frac{1}{c}\sin(a)y + b.$$

Letting $b = h(a)$ the envelope of the functions, $\frac{1}{c}\cos(a)x + \frac{1}{c}\sin(a)y + h(a)$ would give solutions to the equation. Thus you need to solve

$$\frac{-1}{c}\sin(a)x + \frac{1}{c}\cos(a)y + h'(a) = 0$$

for a in terms of x and y. Lets take $h(a) = 0$ for simplicity. Then

$$\cos(a)\, y = \sin(a)\, x$$

and so $\tan(a) = y/x$ so $a = \arctan(y/x)$. Therefore, the solutions corresponding to this choice of h are

$$\begin{aligned} & \frac{1}{c}\cos(\arctan(y/x))\, x + \frac{1}{c}\sin(\arctan(y/x))\, y \\ = \quad & \frac{1}{c}\left(\frac{x^2}{\sqrt{x^2+y^2}} + \frac{y^2}{\sqrt{x^2+y^2}}\right) = \frac{1}{c}\sqrt{x^2+y^2}. \end{aligned}$$

By choosing other choices for h, you could obtain other solutions to the PDE.

The Laplace And Poisson Equation

5.1 The Divergence Theorem

The divergence theorem relates an integral over a set to one on the boundary of the set. It is also called Gauss's theorem.

Definition 5.1 *Let* $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$*. Denote by* $\widehat{\mathbf{x}}_k$ *the vector in* $\mathbb{R}^{n-1}$ *such that*

$$\widehat{\mathbf{x}}_k = (x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n).$$

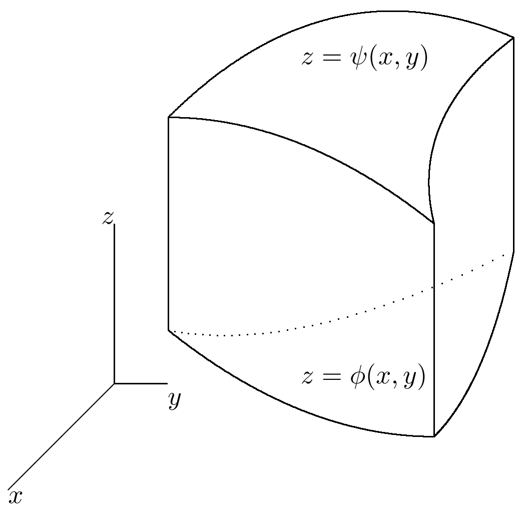
A subset, V *of* $\mathbb{R}^n$ *is called cylindrical in the* x_k *direction if it is of the form*

$$V = \{\mathbf{x} : \phi(\widehat{\mathbf{x}}_k) \leq x_k \leq \psi(\widehat{\mathbf{x}}_k) \ \text{for } \widehat{\mathbf{x}}_k \in D\}$$

Points on ∂D *are defined to be those for which every open ball in* $\mathbb{R}^{n-1}$ *contains points which are in* D *as well as points which are not in* D*. A similar definition holds for sets in* $\mathbb{R}^{n-1}$*. Thus if* V *is cylindrical in the* x_k *direction,*

$$\begin{aligned}
\partial V = & \{\mathbf{x} : \widehat{\mathbf{x}}_k \in \partial D \text{ and } x_k \in (\phi(\widehat{\mathbf{x}}_k), \psi(\widehat{\mathbf{x}}_k))\} \cup \\
& \{\mathbf{x} : \widehat{\mathbf{x}}_k \in D \text{ and } \mathbf{x} = (x_1, \cdots, x_{k-1}, \phi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n)\} \\
& \cup \{\mathbf{x} : \widehat{\mathbf{x}}_k \in D \text{ and } \mathbf{x} = (x_1, \cdots, x_{k-1}, \psi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n)\}.
\end{aligned}$$

The following picture illustrates the above definition in the case of V cylindrical in the z direction in the case of three dimensions.



Of course, many three dimensional sets are cylindrical in each of the coordinate directions. For example, a ball or a rectangle or a tetrahedron are all cylindrical in each direction. Similar sets are found in $\mathbb{R}^n$ but of course they would be harder to draw. The following lemma allows the exchange of the volume integral of a partial derivative for an area integral in which the derivative is replaced with multiplication by an appropriate component of the unit exterior normal.

Lemma 5.2 *Suppose V is cylindrical in the x_k direction and that ϕ and ψ are the functions in the above definition. Assume ϕ and ψ are C^1 functions and suppose F is a C^1 function defined on V. Also, let $\mathbf{n} = (n_1, \cdots, n_n)$ be the unit exterior normal to ∂V. Then*

$$\int_V \frac{\partial F}{\partial x_k}(\mathbf{x})\, dV = \int_{\partial V} F n_k \, dA.$$

Proof: From the fundamental theorem of calculus,

$$\int_V \frac{\partial F}{\partial x_k}(\mathbf{x})\, dV = \int_D \int_{\phi(\widehat{\mathbf{x}}_k)}^{\psi(\widehat{\mathbf{x}}_k)} \frac{\partial F}{\partial x_k}(\mathbf{x})\, dx_k \, d\widehat{x}_k$$

$$= \int \int_D \left[F\left(x_1, \cdots, x_{k-1}, \psi\left(\widehat{\mathbf{x}}_k\right), x_{k+1} \cdots, x_n\right) - F\left(x_1, \cdots, x_{k-1}, \phi\left(\widehat{\mathbf{x}}_k\right), x_{k+1} \cdots, x_n\right)\right] d\widehat{x}_k \tag{5.1}$$

Now the unit exterior normal on the the surface $(x_1, \cdots, x_{k-1}, \psi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n)$, $\widehat{\mathbf{x}}_k \in D$, referred to as the top surface is

$$\frac{1}{\sqrt{\sum_{i \neq k} \left(\psi_{x_i}\right)^2 + 1}} \left(-\psi_{x_1}, \cdots, -\psi_{x_{k-1}}, 1, -\psi_{x_{k+1}}, \cdots, -\psi_{x_n}\right).$$

This follows from the observation that the top surface is the level surface, $x_k - \psi(\widehat{\mathbf{x}}_k) = 0$ and so the gradient of this function of three variables is perpendicular to the level surface. It points in the correct direction because the x_k component is positive. Therefore, on the top surface,

$$n_k = \frac{1}{\sqrt{\sum_{i \neq k} \left(\psi_{x_i}\right)^2 + 1}}$$

Similarly, the unit outer normal to the surface on the bottom, the one of the form

$$(x_1, \cdots, x_{k-1}, \phi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n), \ \widehat{\mathbf{x}}_k \in D$$

is given by

$$\frac{1}{\sqrt{\sum_{i\neq k}(\phi_{x_i})^2+1}}\left(\phi_{x_1}, \cdots, \phi_{x_{k-1}}, -1, \phi_{x_{k+1}}, \cdots, \phi_{x_n}\right)$$

and so on the bottom surface,

$$n_z = \frac{-1}{\sqrt{\sum_{i\neq k}(\phi_{x_i})^2+1}}$$

Note that here the z component is negative because since it is the outer normal it must point down. On the lateral surface, the one where $\widehat{\mathbf{x}}_k \in \partial D$ and $x_k \in [\phi(\widehat{\mathbf{x}}_k), \psi(\widehat{\mathbf{x}}_k)]$, $n_k = 0$.

The area element on the top surface, denoted by T is $dA = \sqrt{\sum_{i\neq k}(\psi_{x_i})^2+1}\, d\widehat{x}_k$ while the area element on the bottom surface, denoted by B is $\sqrt{\sum_{i\neq k}(\phi_{x_i})^2+1}\, d\widehat{x}_k$. Therefore, the last expression in (5.1) is of the form,

$$\int_T F(x_1, \cdots, x_{k-1}, \psi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n) \frac{dA}{\sqrt{\sum_{i\neq k}(\psi_{x_i})^2+1}}$$

$$-\int_B F(x_1, \cdots, x_{k-1}, \phi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n) \frac{dA}{\sqrt{\sum_{i\neq k}(\phi_{x_i})^2+1}} =$$

$$\int_T F(x_1, \cdots, x_{k-1}, \psi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n)\, n_k dA$$

$$+\int_B F(x_1, \cdots, x_{k-1}, \phi(\widehat{\mathbf{x}}_k), x_{k+1} \cdots, x_n)\, n_k dA$$

$$+\int_{\text{Lateral surface}} F n_k\, dA,$$

the last term equaling zero because on the lateral surface, $n_k = 0$. Therefore, this reduces to $\int\int_{\partial V} F n_k\, dA$ as claimed.

Theorem 5.3 *Let V be cylindrical in each of the coordinate directions and let $\mathbf{F}$ be a C^1 vector field defined on V. Then*

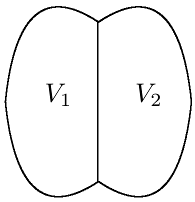
$$\int_V \nabla \cdot \mathbf{F}\, dV = \int_{\partial V} \mathbf{F} \cdot \mathbf{n}\, dA.$$

Proof: From the above lemma and corollary,

$$\begin{aligned}\int_V \nabla \cdot \mathbf{F}\, dV &= \int_V \sum_i \frac{\partial F_i}{\partial x_i}\, dV \\ &= \int_{\partial V} \sum_i F_i n_i\, dA \\ &= \int_{\partial V} \mathbf{F} \cdot \mathbf{n}\, dA.\end{aligned}$$

This proves the theorem.

The divergence theorem holds for much more general regions than this. Suppose for example you have a complicated region which is the union of finitely many disjoint regions of the sort just described which are cylindrical in each of the coordinate directions. Then the volume integral over the union of these would equal the sum of the integrals over the disjoint regions. If the boundaries of two of these regions intersect, then the area integrals will cancel out on the intersection because the unit exterior normals will point in opposite directions. Therefore, the sum of the integrals over the boundaries of these disjoint regions will reduce to an integral over the boundary of the union of these. Hence the divergence theorem will continue to hold. For example, consider the following picture. If the divergence theorem holds for each V_i in the following picture, then it holds for the union of these two.



There are much more general formulations for the divergence theorem than this. I have been tacitly assuming that the functions defining the top and bottoms are C^1 but this is not necessary. Lipshitz continuous is plenty. Also, it is not necessary to assume the set is a finite union of such sets which are cylindrical in all directions. The point is, the divergence theorem is really very good and you can use it with considerable confidence.

Definition 5.4 *Let u be a function defined on an open subset of $\mathbb{R}^n$. Then*

$$\Delta u \equiv \sum_i u_{x_i x_i}.$$

Δ is called the Laplacian. Also, for $\mathbf{n}$ the unit outer normal,

$$\frac{\partial F}{\partial n} \equiv \nabla F \cdot \mathbf{n}.$$

The following little result is now obvious and I leave the proof for you to do. It is called Green's identity.

Theorem 5.5 *Let U be an open set for which the divergence theorem holds and let $u, v \in C^2\left(\overline{U}\right)$. This means u, v are the restrictions to U of functions which are $C^2(\mathbb{R}^n)$. Then*

$$\int_U (v\Delta u - u\Delta v)\, dx = \int_{\partial U} \left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n} \right) dA$$

5.1.1 Balls

Recall, $B(\mathbf{x}, r)$ denotes the set of all $\mathbf{y} \in \mathbb{R}^n$ such that $|\mathbf{y} - \mathbf{x}| < r$. By the change of variables formula for multiple integrals or simple geometric reasoning, all balls of radius r have the same volume. Furthermore, simple reasoning or change of variables formula will show that the volume of the ball of radius r equals $\alpha_n r^n$ where α_n will denote the volume of the unit ball in $\mathbb{R}^n$. With the divergence theorem, it is now easy to give a simple relationship between the surface area of the ball of radius r and the volume. By the divergence theorem,

$$\int_{B(\mathbf{0},r)} \operatorname{div} \mathbf{x}\, dx = \int_{\partial B(\mathbf{0},r)} \mathbf{x} \cdot \frac{\mathbf{x}}{|\mathbf{x}|} dA$$

because the unit outward normal on $\partial B(\mathbf{0}, r)$ is $\frac{\mathbf{x}}{|\mathbf{x}|}$. Therefore,

$$n\alpha_n r^n = rA\left(\partial B(\mathbf{0}, r)\right)$$

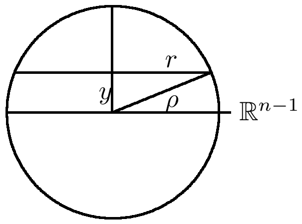
and so

$$A\left(\partial B\left(\mathbf{0},r\right)\right)=n\alpha_n r^{n-1}.$$

You recall the surface area of $S^2\equiv\left\{\mathbf{x}\in\mathbb{R}^3:|\mathbf{x}|=r\right\}$ is given by $4\pi r^2$ while the volume of the ball, $B\left(\mathbf{0},r\right)$ is $\frac{4}{3}\pi r^3$. This follows the above pattern. You just take the derivative with respect to the radius of the volume of the ball of radius r to get the area of the surface of this ball. Let ω_n denote the area of the sphere $S^{n-1}=\left\{\mathbf{x}\in\mathbb{R}^n:|\mathbf{x}|=1\right\}$. I just showed that

$$\omega_n=n\alpha_n.$$

I want to find α_n now and also to get a relationship between ω_n and ω_{n-1}. Consider the following picture of the ball of radius ρ seen on the side.



Taking slices at height y as shown and using that these slices have $n-1$ dimensional area equal to $\alpha_{n-1}r^{n-1}$, it follows

$$\alpha_n\rho^n=2\int_0^\rho\alpha_{n-1}\left(\rho^2-y^2\right)^{(n-1)/2}dy$$

In the integral, change variables, letting $y=\rho\cos\theta$. Then

$$\alpha_n\rho^n=2\rho^n\alpha_{n-1}\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta.$$

It follows that

$$\alpha_n=2\alpha_{n-1}\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta. \tag{5.2}$$

Consequently,

$$\omega_n=\frac{2n\omega_{n-1}}{n-1}\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta. \tag{5.3}$$

This is a little messier than I would like.

$$\begin{aligned}\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta &= -\cos\theta\sin^{n-1}\theta|_0^{\pi/2}+(n-1)\int_0^{\pi/2}\cos^2\theta\sin^{n-2}\theta\\ &= (n-1)\int_0^{\pi/2}\left(1-\sin^2\theta\right)\sin^{n-2}\left(\theta\right)d\theta\\ &= (n-1)\int_0^{\pi/2}\sin^{n-2}\left(\theta\right)d\theta-(n-1)\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta\end{aligned}$$

Hence

$$n\int_0^{\pi/2}\sin^n\left(\theta\right)d\theta=(n-1)\int_0^{\pi/2}\sin^{n-2}\left(\theta\right)d\theta \tag{5.4}$$

and so (5.3) is of the form

$$\omega_n = 2\omega_{n-1}\int_0^{\pi/2}\sin^{n-2}(\theta)\,d\theta. \tag{5.5}$$

So what is α_n explicitly? Clearly $\alpha_1 = 2$ and $\alpha_2 = \pi$.

Theorem 5.6 $\alpha_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$ *where* Γ *denotes the gamma function, defined for* $\alpha > 0$ *by*

$$\Gamma(\alpha) \equiv \int_0^\infty e^{-t}t^{\alpha-1}dt.$$

Proof: Recall that $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Now note the given formula holds if $n = 1$ because

$$\Gamma\left(\frac{1}{2}+1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

(I leave it as an exercise for you to verify that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.) Thus

$$\alpha_1 = 2 = \frac{\sqrt{\pi}}{\sqrt{\pi}/2}$$

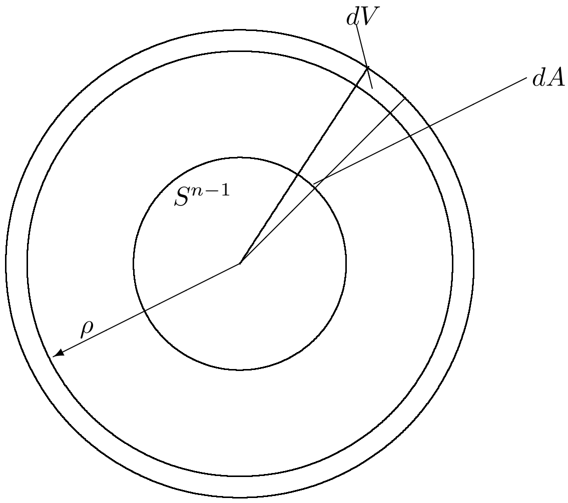
satisfying the formula. Now suppose this formula holds for $k \leq n$. Then from the induction hypothesis, (5.5), (5.4), (5.2) and (5.3),

$$\begin{aligned}
\alpha_{n+1} &= 2\alpha_n\int_0^{\pi/2}\sin^{n+1}(\theta)\,d\theta \\
&= 2\alpha_n\frac{n}{n+1}\int_0^{\pi/2}\sin^{n-1}(\theta)\,d\theta \\
&= 2\alpha_n\frac{n}{n+1}\frac{\alpha_{n-1}}{2\alpha_{n-2}} \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}\frac{n}{n+1}\pi^{1/2}\frac{\Gamma\left(\frac{n-2}{2}+1\right)}{\Gamma\left(\frac{n-1}{2}+1\right)} \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n-2}{2}+1\right)\left(\frac{n}{2}\right)}\frac{n}{n+1}\pi^{1/2}\frac{\Gamma\left(\frac{n-2}{2}+1\right)}{\Gamma\left(\frac{n-1}{2}+1\right)} \\
&= 2\pi^{(n+1)/2}\frac{1}{n+1}\frac{1}{\Gamma\left(\frac{n-1}{2}+1\right)} \\
&= \pi^{(n+1)/2}\frac{1}{\left(\frac{n+1}{2}\right)}\frac{1}{\Gamma\left(\frac{n-1}{2}+1\right)} \\
&= \pi^{(n+1)/2}\frac{1}{\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)} = \frac{\pi^{(n+1)/2}}{\Gamma\left(\frac{n+1}{2}+1\right)}.
\end{aligned}$$

This proves the theorem.

5.1.2 Polar Coordinates

The right way to discuss this is in the context of the Lebesgue integral where everything can be easily proved. I will give you a heuristic explanation of the technique of polar coordinates below. Consider the following picture.



From the picture, a Riemann sum for $\int_{B(\mathbf{0},r)} f(\mathbf{x})\,dx$ would involve summing things of the form

$$f(\rho\mathbf{w})\,dV = f(\rho\mathbf{w})\,\rho^{n-1}dAd\rho.$$

Thus

$$\int_{B(\mathbf{0},r)} f(\mathbf{x})\,dx = \int_0^r \int_{S^{n-1}} f(\rho\mathbf{w})\,\rho^{n-1}dAd\rho.$$

5.2 Poisson's Problem

The Poisson problem is to find u satisfying the two conditions

$$\Delta u = f, \text{ in } U, \; u = g \text{ on } \partial U. \tag{5.6}$$

Here U is an open bounded set for which the divergence theorem holds. When $f = 0$ this is called Laplace's equation and the boundary condition given is called a Dirichlet boundary condition. When $\Delta u = 0$, the function, u is said to be a harmonic function. When $f \neq 0$, it is called Poisson's equation. I will give a way of representing the solution to these problems. When this has been done, great and marvelous conclusions may be drawn about the solutions. Before doing anything else however, it is wise to prove a fundamental result called the weak maximum principle.

Theorem 5.7 *Suppose U is an open bounded set and*

$$u \in C^2(U) \cap C\left(\overline{U}\right)$$

and

$$\Delta u \geq 0 \textit{ in } U.$$

Then

$$\max\left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U}\right\} = \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\}.$$

Proof: Suppose not. Then there exists $\mathbf{x}_0 \in U$ such that

$$u(\mathbf{x}_0) > \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\}.$$

Consider $w_\varepsilon(\mathbf{x}) \equiv u(\mathbf{x}) + \varepsilon |\mathbf{x}|^2$. I claim that for small enough $\varepsilon > 0$, the function w also has this property. If not, there exists $\mathbf{x}_\varepsilon \in \partial U$ such that $w_\varepsilon(\mathbf{x}_\varepsilon) \geq w_\varepsilon(\mathbf{x})$ for all $\mathbf{x} \in U$. But since U is bounded, it follows the points, $\mathbf{x}_\varepsilon$ are in a compact set and so there exists a subsequence, still denoted by $\mathbf{x}_\varepsilon$ such that as $\varepsilon \to 0, \mathbf{x}_\varepsilon \to \mathbf{x}_1 \in \partial U$. But then for any $\mathbf{x} \in U$,

$$u(\mathbf{x}_0) \leq w_\varepsilon(\mathbf{x}_0) \leq w_\varepsilon(\mathbf{x}_\varepsilon)$$

and taking a limit as $\varepsilon \to 0$ yields

$$u(\mathbf{x}_0) \leq u(\mathbf{x}_1)$$

contrary to the property of $\mathbf{x}_0$ above. It follows that my claim is verified. Pick such an ε. Then w_ε assumes its maximum value in U say at $\mathbf{x}_2$. Then by the second derivative test,

$$\Delta w_\varepsilon(\mathbf{x}_2) = \Delta u(\mathbf{x}_2) + 2\varepsilon \leq 0$$

which requires $\Delta u(\mathbf{x}_2) \leq -2\varepsilon$, contrary to the assumption that $\Delta u \geq 0$. This proves the theorem.

The theorem makes it very easy to verify the following uniqueness result.

Corollary 5.8 *Suppose U is an open bounded set and*

$$u \in C^2(U) \cap C\left(\overline{U}\right)$$

and

$$\Delta u = 0 \text{ in } U, u = 0 \text{ on } \partial U.$$

Then $u = 0$.

Proof: From the weak maximum principle, $u \leq 0$. Now apply the weak maximum principle to $-u$ which satisfies the same conditions as u. Thus $-u \leq 0$ and so $u \geq 0$. Therefore, $u = 0$ as claimed.

Define

$$r_n(\mathbf{x}) \equiv \begin{cases} \ln|\mathbf{x}| \text{ if } n = 2 \\ \frac{1}{|\mathbf{x}|^{n-2}} \text{ if } n > 2 \end{cases}.$$

Then it is fairly routine to verify the following Lemma.

Lemma 5.9 *For r_n given above,*

$$\Delta r_n = 0.$$

Proof: I will verify the case where $n \geq 3$ and leave the other case for you.

$$D_{x_i}\left(\sum_{i=1}^n x_i^2\right)^{-(n-2)/2} = -(n-2)\, x_i \left(\sum_j x_j^2\right)^{-n/2}$$

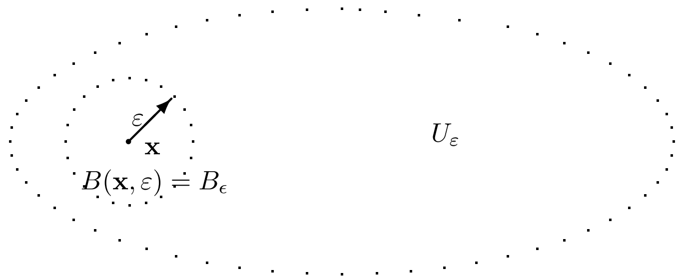
Therefore,

$$D_{x_i}\left(D_{x_i}(r_n)\right) = \left(\sum_j x_j^2\right)^{-(n+2)/2} (n-2)\left[n x_i^2 - \sum_{j=1}^n x_j^2\right].$$

It follows

$$\Delta r_n = \left(\sum_j x_j^2\right)^{\frac{-(n+2)}{2}} (n-2)\left(n\sum_{i=1}^n x_i^2 - \sum_{i=1}^n\sum_{j=1}^n x_j^2\right) = 0.$$

Now let U_ε be as indicated in the following picture. I have taken out a ball of radius ε which is centered at the point, $\mathbf{x} \in U$.



Then the divergence theorem will continue to hold for U_ε (why?) and so I can use Green's identity to write the following for $u, v \in C^2\left(\overline{U}\right)$.

$$\int_{U_\varepsilon} (u\Delta v - v\Delta u)\, dx = \int_{\partial U}\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) dA - \int_{\partial B_\varepsilon}\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) dA \tag{5.7}$$

Now, letting $\mathbf{x} \in U$, I will pick for v the function,

$$v(\mathbf{y}) \equiv r_n(\mathbf{y}-\mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y})$$

where $\psi^{\mathbf{x}}$ is a function which is chosen such that on ∂U,

$$\psi^{\mathbf{x}}(\mathbf{y}) = r_n(\mathbf{y}-\mathbf{x})$$

and $\psi^{\mathbf{x}}$ is in $C^2\left(\overline{U}\right)$ and also satisfies

$$\Delta\psi^{\mathbf{x}} = 0.$$

The existence of such a function is another issue. For now, assume such a function exists.[1] Then assuming such a function exists, (5.7) reduces to

$$-\int_{U_\varepsilon} v\Delta u dx = \int_{\partial U} u\frac{\partial v}{\partial n} dA - \int_{\partial B_\varepsilon}\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) dA. \tag{5.8}$$

The idea now is to let $\varepsilon \to 0$ and see what happens. Consider the term

$$\int_{\partial B_\varepsilon} v\frac{\partial u}{\partial n} dA.$$

[1] In fact, if the boundary of U is smooth enough, such a function will always exist, although this requires more work to show but this is not the point. The point is to explicitly find it and this will only be possible for certain simple choices of U.

The area is $O\left(\varepsilon^{n-1}\right)$ while the integrand is $O\left(\varepsilon^{-(n-2)}\right)$ in the case where $n \geq 3$. In the case where $n = 2$, the area is $O(\varepsilon)$ and the integrand is $O(|\ln|\varepsilon||)$. Now you know that $\lim_{\varepsilon \to 0} \varepsilon \ln |\varepsilon| = 0$ and so in the case $n = 2$, this term converges to 0 as $\varepsilon \to 0$. In the case that $n \geq 3$, it also converges to zero because in this case the integral is $O(\varepsilon)$.

Next consider the term

$$-\int_{\partial B_\varepsilon} u \frac{\partial v}{\partial n} dA = -\int_{\partial B_\varepsilon} u(\mathbf{y}) \left(\frac{\partial r_n}{\partial n} (\mathbf{y} - \mathbf{x}) - \frac{\partial \psi^{\mathbf{x}}}{\partial n} (\mathbf{y}) \right) dA.$$

This term does not disappear as $\varepsilon \to 0$. First note that since $\psi^{\mathbf{x}}$ has bounded derivatives,

$$\lim_{\varepsilon \to 0} -\int_{\partial B_\varepsilon} u(\mathbf{y}) \left(\frac{\partial r_n}{\partial n} (\mathbf{y} - \mathbf{x}) - \frac{\partial \psi^{\mathbf{x}}}{\partial n} (\mathbf{y}) \right) dA = \lim_{\varepsilon \to 0} \left(-\int_{\partial B_\varepsilon} u(\mathbf{y}) \frac{\partial r_n}{\partial n} (\mathbf{y} - \mathbf{x}) \, dA \right) \tag{5.9}$$

and so it is just this last item which is of concern.

First consider the case that $n = 2$. In this case,

$$\nabla r_2(\mathbf{y}) = \left(\frac{y_1}{|\mathbf{y}|^2}, \frac{y_2}{|\mathbf{y}|^2} \right)$$

Also, on ∂B_ε, the exterior unit normal, $\mathbf{n}$, equals

$$\frac{1}{\varepsilon} (y_1 - x_1, y_2 - x_2).$$

It follows that on ∂B_ε,

$$\frac{\partial r_2}{\partial n} (\mathbf{y} - \mathbf{x}) = \frac{1}{\varepsilon} (y_1 - x_1, y_2 - x_2) \cdot \left(\frac{y_1 - x_1}{|\mathbf{y} - \mathbf{x}|^2}, \frac{y_2 - x_2}{|\mathbf{y} - \mathbf{x}|^2} \right) = \frac{1}{\varepsilon}.$$

Therefore, this term in (5.9) converges to

$$-u(\mathbf{x}) \, 2\pi. \tag{5.10}$$

Next consider the case where $n \geq 3$. In this case,

$$\nabla r_n(\mathbf{y}) = -(n-2) \left(\frac{y_1}{|\mathbf{y}|^n}, \cdots, \frac{y_n}{|\mathbf{y}|} \right)$$

and the unit outer normal, $\mathbf{n}$, equals

$$\frac{1}{\varepsilon} (y_1 - x_1, \cdots, y_n - x_n).$$

Therefore,

$$\frac{\partial r_n}{\partial n} (\mathbf{y} - \mathbf{x}) = -\frac{(n-2)}{\varepsilon} \frac{|\mathbf{y} - \mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} = \frac{-(n-2)}{\varepsilon^{n-1}}.$$

Letting ω_n denote the $n-1$ dimensional surface area of the unit sphere, S^{n-1}, it follows that the last term in (5.9) converges to

$$u(\mathbf{x}) (n-2) \, \omega_n \tag{5.11}$$

Finally consider the integral,

$$\int_{B_\varepsilon} v \Delta u dx.$$

$$\begin{aligned}\int_{B_\varepsilon}|v\Delta u|\,dx &\leq C\int_{B_\varepsilon}|r_n(\mathbf{y}-\mathbf{x})-\psi^{\mathbf{x}}(\mathbf{y})|\,dy\\ &\leq C\int_{B_\varepsilon}|r_n(\mathbf{y}-\mathbf{x})|\,dy+O(\varepsilon^n)\end{aligned}$$

Using polar coordinates to evaluate this improper integral in the case where $n\geq 3$,

$$\begin{aligned}C\int_{B_\varepsilon}|r_n(\mathbf{y}-\mathbf{x})|\,dx &= C\int_0^\varepsilon\int_{S^{n-1}}\frac{1}{\rho^{n-2}}\rho^{n-1}dAd\rho\\ &= C\int_0^\varepsilon\int_{S^{n-1}}\rho dAd\rho\end{aligned}$$

which converges to 0 as $\varepsilon\to 0$. In the case where $n=2$

$$C\int_{B_\varepsilon}|r_n(\mathbf{y}-\mathbf{x})|\,dx=C\int_0^\varepsilon\int_{S^{n-1}}\ln(\rho)\,\rho dAd\rho$$

which also converges to 0 as $\varepsilon\to 0$. Therefore, returning to (5.8) and using the above limits, yields in the case where $n\geq 3$,

$$-\int_U v\Delta udx=\int_{\partial U}u\frac{\partial v}{\partial n}dA+u(\mathbf{x})(n-2)\,\omega_n, \tag{5.12}$$

and in the case where $n=2$,

$$-\int_U v\Delta udx=\int_{\partial U}u\frac{\partial v}{\partial n}dA-u(\mathbf{x})\,2\pi. \tag{5.13}$$

These two formulas show that it is possible to represent the solutions to Poisson's problem provided the function, $\psi^{\mathbf{x}}$ can be determined. I will show you can determine this function in the case that $U=B(\mathbf{0},r)$.

5.2.1 Poisson's Problem For A Ball

Lemma 5.10 *When* $|\mathbf{y}|=r$ *and* $\mathbf{x}\neq\mathbf{0}$,

$$\left|\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}\right|=|\mathbf{x}-\mathbf{y}|,$$

and for $|\mathbf{x}|,|\mathbf{y}|<r,\mathbf{x}\neq\mathbf{0}$,

$$\left|\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}\right|\neq 0.$$

Proof: Suppose first that $|\mathbf{y}|=r$. Then

$$\begin{aligned}\left|\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}\right|^2 &= \left(\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}\right)\cdot\left(\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}\right)\\ &= \frac{|\mathbf{x}|^2}{r^2}|\mathbf{y}|^2-2\mathbf{y}\cdot\mathbf{x}+r^2\frac{|\mathbf{x}|^2}{|\mathbf{x}|^2}\\ &= |\mathbf{x}|^2-2\mathbf{x}\cdot\mathbf{y}+|\mathbf{y}|^2=|\mathbf{x}-\mathbf{y}|^2.\end{aligned}$$

This proves the first claim. Next suppose $|\mathbf{x}|,|\mathbf{y}|<r$ and suppose, contrary to what is claimed, that

$$\frac{\mathbf{y}|\mathbf{x}|}{r}-\frac{r\mathbf{x}}{|\mathbf{x}|}=\mathbf{0}.$$

Then

$$\mathbf{y}\left|\mathbf{x}\right|^{2}=r^{2}\mathbf{x}$$

and so $\left|\mathbf{y}\right|\left|\mathbf{x}\right|^{2}=r^{2}\left|\mathbf{x}\right|$ which implies

$$\left|\mathbf{y}\right|\left|\mathbf{x}\right|=r^{2}$$

contrary to the assumption that $\left|\mathbf{x}\right|,\left|\mathbf{y}\right|<r$.

Let

$$\psi^{\mathbf{x}}\left(\mathbf{y}\right)\equiv\begin{cases}\left|\frac{\mathbf{y}\left|\mathbf{x}\right|}{r}-\frac{r\mathbf{x}}{\left|\mathbf{x}\right|}\right|^{-(n-2)}, & r^{-(n-2)}\text{ for }\mathbf{x}=\mathbf{0}\text{ if }n\geq 3\\ \ln\left|\frac{\mathbf{y}\left|\mathbf{x}\right|}{r}-\frac{r\mathbf{x}}{\left|\mathbf{x}\right|}\right|, & \ln\left(r\right)\text{ if }\mathbf{x}=\mathbf{0}\text{ if }n=2\end{cases}$$

Note that

$$\lim_{\mathbf{x}\to\mathbf{0}}\left|\frac{\mathbf{y}\left|\mathbf{x}\right|}{r}-\frac{r\mathbf{x}}{\left|\mathbf{x}\right|}\right|=r.$$

Then $\psi^{\mathbf{x}}\left(\mathbf{y}\right)=r_{n}\left(\mathbf{y}-\mathbf{x}\right)$ if $\left|\mathbf{y}\right|=r$, and $\Delta\psi^{\mathbf{x}}=0$. This last claim is obviously true if $\mathbf{x}\neq\mathbf{0}$. If $\mathbf{x}=\mathbf{0}$, then $\psi^{\mathbf{0}}\left(\mathbf{y}\right)$ equals a constant and so it is also obvious in this case that $\Delta\psi^{\mathbf{x}}=0$. The following lemma is easy to obtain.

Lemma 5.11 *Let*

$$f\left(\mathbf{y}\right)=\begin{cases}\left|\mathbf{y}-\mathbf{x}\right|^{-(n-2)} & \text{if } n\geq 3\\ \ln\left|\mathbf{y}-\mathbf{x}\right| & \text{if } n=2\end{cases}.$$

Then

$$\nabla f\left(\mathbf{y}\right)=\begin{cases}\frac{-(n-2)(\mathbf{y}-\mathbf{x})}{\left|\mathbf{y}-\mathbf{x}\right|^{n}} & \text{if } n\geq 3\\ \frac{\mathbf{y}-\mathbf{x}}{\left|\mathbf{y}-\mathbf{x}\right|^{2}} & \text{if } n=2\end{cases}.$$

Also, the outer normal on $\partial B\left(\mathbf{0},r\right)$ is $\mathbf{y}/r$.

From Lemma 5.11 it follows easily that for $v\left(\mathbf{y}\right)=r_{n}\left(\mathbf{y}-\mathbf{x}\right)-\psi^{\mathbf{x}}\left(\mathbf{y}\right)$ and $\mathbf{y}\in\partial B\left(\mathbf{0},r\right)$, then for $n\geq 3$,

$$\begin{aligned}\frac{\partial v}{\partial n} &= \frac{\mathbf{y}}{r}\cdot\left[\frac{-\left(n-2\right)\left(\mathbf{y}-\mathbf{x}\right)}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}+\left(\frac{\left|\mathbf{x}\right|}{r}\right)^{-(n-2)}\left(n-2\right)\frac{\left(\mathbf{y}-\frac{r^{2}}{\left|\mathbf{x}\right|^{2}}\mathbf{x}\right)}{\left|\mathbf{y}-\frac{r^{2}}{\left|\mathbf{x}\right|^{2}}\mathbf{x}\right|^{n}}\right]\\ &= \frac{-\left(n-2\right)}{r}\frac{\left(r^{2}-\mathbf{y}\cdot\mathbf{x}\right)}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}+\frac{\frac{\left|\mathbf{x}\right|^{2}}{r^{2}}}{\left(\frac{\left|\mathbf{x}\right|}{r}\right)^{n}}\frac{\left(n-2\right)}{r}\frac{\left(r^{2}-\frac{r^{2}}{\left|\mathbf{x}\right|^{2}}\mathbf{x}\cdot\mathbf{y}\right)}{\left|\mathbf{y}-\frac{r^{2}}{\left|\mathbf{x}\right|^{2}}\mathbf{x}\right|^{n}}\\ &= \frac{-\left(n-2\right)}{r}\frac{\left(r^{2}-\mathbf{y}\cdot\mathbf{x}\right)}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}+\frac{\left(n-2\right)}{r}\frac{\left(\frac{\left|\mathbf{x}\right|^{2}}{r^{2}}r^{2}-\mathbf{x}\cdot\mathbf{y}\right)}{\left|\frac{\left|\mathbf{x}\right|}{r}\mathbf{y}-\frac{r}{\left|\mathbf{x}\right|}\mathbf{x}\right|^{n}}\end{aligned}$$

which by Lemma 5.10 equals

$$\begin{aligned}&\frac{-\left(n-2\right)}{r}\frac{\left(r^{2}-\mathbf{y}\cdot\mathbf{x}\right)}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}+\frac{\left(n-2\right)}{r}\frac{\left(\frac{\left|\mathbf{x}\right|^{2}}{r^{2}}r^{2}-\mathbf{x}\cdot\mathbf{y}\right)}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}\\ =&\frac{-\left(n-2\right)}{r}\frac{r^{2}}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}+\frac{\left(n-2\right)}{r}\frac{\left|\mathbf{x}\right|^{2}}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}\\ =&\frac{\left(n-2\right)}{r}\frac{\left|\mathbf{x}\right|^{2}-r^{2}}{\left|\mathbf{y}-\mathbf{x}\right|^{n}}.\end{aligned}$$

In the case where $n = 2$, and $|\mathbf{y}| = r$, then Lemma 5.10 implies

$$\begin{aligned}
\frac{\partial v}{\partial n} &= \frac{\mathbf{y}}{r} \cdot \left[\frac{(\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2} - \left(\frac{|\mathbf{x}|}{r}\right) \frac{\left(\frac{\mathbf{y}|\mathbf{x}|}{r} - \frac{r\mathbf{x}}{|\mathbf{x}|}\right)}{\left|\frac{\mathbf{y}|\mathbf{x}|}{r} - \frac{r\mathbf{x}}{|\mathbf{x}|}\right|^2} \right] \\
&= \frac{\mathbf{y}}{r} \cdot \left[\frac{(\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2} - \frac{\left(\frac{\mathbf{y}|\mathbf{x}|^2}{r^2} - \mathbf{x}\right)}{|\mathbf{y}-\mathbf{x}|^2} \right] \\
&= \frac{1}{r} \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^2}.
\end{aligned}$$

Referring to (5.12) and (5.13), we would hope a solution, u to Poisson's problem satisfies for $n \geq 3$

$$-\int_U (r_n(\mathbf{y}-\mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y})) f(\mathbf{y})\, dy = \int_{\partial U} g(\mathbf{y}) \left(\frac{(n-2)}{r} \frac{|\mathbf{x}|^2 - r^2}{|\mathbf{y}-\mathbf{x}|^n} \right) dA(y) + u(\mathbf{x})(n-2)\omega_n.$$

Thus

$$u(\mathbf{x}) = \frac{1}{\omega_n (n-2)} \cdot$$

$$\left[\int_U (\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})) f(\mathbf{y})\, dy + \int_{\partial U} g(\mathbf{y}) \left(\frac{(n-2)}{r} \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n} \right) dA(y) \right]. \tag{5.14}$$

In the case where $n = 2$,

$$-\int_U (r_2(\mathbf{y}-\mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y})) f(\mathbf{y})\, dx = \int_{\partial U} g(\mathbf{y}) \left(\frac{1}{r} \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^2} \right) dA(\mathbf{y}) - u(\mathbf{x})\, 2\pi$$

and so in this case,

$$u(\mathbf{x}) = \frac{1}{2\pi} \left[\int_U (r_2(\mathbf{y}-\mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y})) f(\mathbf{y})\, dx + \int_{\partial U} g(\mathbf{y}) \left(\frac{1}{r} \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^2} \right) dA(\mathbf{y}) \right]. \tag{5.15}$$

5.2.2 Does It Work In Case $f = 0$?

It turns out these formulas work better than you might expect. In particular, they work in the case where g is only continuous. In deriving these formulas, more was assumed on the function than this. In particular, it would have been the case that g was equal to the restriction of a function in $C^2(\mathbb{R}^n)$ to $\partial B(\mathbf{0}, r)$. The problem considered here is

$$\Delta u = 0 \text{ in } U, \ u = g \text{ on } \partial U$$

From (5.14) it follows that if u solves the above problem, known as the Dirichlet problem, then

$$u(\mathbf{x}) = \frac{r^2 - |\mathbf{x}|^2}{\omega_n r} \int_{\partial U} g(\mathbf{y}) \frac{1}{|\mathbf{y}-\mathbf{x}|^n} dA(\mathbf{y}).$$

I have shown this in case $u \in C^2(\overline{U})$ which is more specific than to say $u \in C^2(U) \cap C(\overline{U})$. Nevertheless, it is enough to give the following lemma.

Lemma 5.12 *The following holds for $n \geq 3$.*

$$1 = \int_{\partial U} \frac{r^2 - |\mathbf{x}|^2}{r\omega_n |\mathbf{y} - \mathbf{x}|^n} dA(y).$$

For $n = 2$,

$$1 = \int_{\partial U} \frac{1}{2\pi r} \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^2} dA(\mathbf{y}).$$

Proof: Consider the problem

$$\Delta u = 0 \text{ in } U, \; u = 1 \text{ on } \partial U.$$

I know a solution to this problem which is in $C^2\left(\overline{U}\right)$, namely $u \equiv 1$. Therefore, by Corollary 5.8 this is the only solution and since it is in $C^2\left(\overline{U}\right)$, it follows from (5.14) that in case $n \geq 3$,

$$1 = u(\mathbf{x}) = \int_{\partial U} \frac{r^2 - |\mathbf{x}|^2}{r\omega_n |\mathbf{y} - \mathbf{x}|^n} dA(y)$$

and in case $n = 2$, the other formula claimed above holds.

Theorem 5.13 *Let $U = B(\mathbf{0}, r)$ and let $g \in C(\partial U)$. Then there exists a unique solution $u \in C^2(U) \cap C\left(\overline{U}\right)$ to the problem*

$$\Delta u = 0 \text{ in } U, \; u = g \text{ on } \partial U.$$

This solution is given by the formula,

$$u(\mathbf{x}) = \frac{1}{\omega_n r} \int_{\partial U} g(\mathbf{y}) \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) \tag{5.16}$$

for every $n \geq 2$. Here $\omega_2 = 2\pi$.

Proof: That $\Delta u = 0$ in U follows from the observation that the difference quotients used to compute the partial derivatives converge uniformly in $\mathbf{y} \in \partial U$ for any given $\mathbf{x} \in U$. To see this note that for $\mathbf{y} \in \partial U$, the partial derivatives of the expression,

$$\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n}$$

taken with respect to x_k are uniformly bounded and continuous. In fact, this is true of all partial derivatives. Therefore you can take the differential operator inside the integral and write

$$\Delta_x \frac{1}{\omega_n r} \int_{\partial U} g(\mathbf{y}) \frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) = \frac{1}{\omega_n r} \int_{\partial U} g(\mathbf{y}) \Delta_x \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} \right) dA(\mathbf{y}) = 0.$$

It only remains to verify that it achieves the desired boundary condition. Let $\mathbf{x}_0 \in \partial U$. From Lemma 5.12,

$$|g(\mathbf{x}_0) - u(\mathbf{x})| \leq \frac{1}{\omega_n r} \int_{\partial U} |g(\mathbf{y}) - g(\mathbf{x}_0)| \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} \right) dA(\mathbf{y}) \tag{5.17}$$

$$\leq \frac{1}{\omega_n r} \int_{[|\mathbf{y} - \mathbf{x}_0| < \delta]} |g(\mathbf{y}) - g(\mathbf{x}_0)| \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} \right) dA(\mathbf{y}) + \tag{5.18}$$

$$\frac{1}{\omega_n r} \int_{[|\mathbf{y} - \mathbf{x}_0| \geq \delta]} |g(\mathbf{y}) - g(\mathbf{x}_0)| \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y} - \mathbf{x}|^n} \right) dA(\mathbf{y}) \tag{5.19}$$

where δ is a positive number. Letting $\varepsilon > 0$ be given, choose δ small enough that if $|\mathbf{y} - \mathbf{x}_0| < \delta$, then $|g(\mathbf{y}) - g(\mathbf{x}_0)| < \frac{\varepsilon}{2}$. Then for such δ,

$$\begin{aligned}
\frac{1}{\omega_n r}\int_{[|\mathbf{y}-\mathbf{x}_0|<\delta]} |g(\mathbf{y}) - g(\mathbf{x}_0)| \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n}\right) dA(\mathbf{y}) &\leq \frac{1}{\omega_n r}\int_{[|\mathbf{y}-\mathbf{x}_0|<\delta]} \frac{\varepsilon}{2}\left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n}\right) dA(\mathbf{y}) \\
&\leq \frac{1}{\omega_n r}\int_{\partial U} \frac{\varepsilon}{2}\left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n}\right) dA(\mathbf{y}) = \frac{\varepsilon}{2}.
\end{aligned}$$

Denoting by M the maximum value of g on ∂U, the integral in (5.19) is dominated by

$$\begin{aligned}
\frac{2M}{\omega_n r}\int_{[|\mathbf{y}-\mathbf{x}_0|\geq\delta]} \left(\frac{r^2 - |\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n}\right) dA(\mathbf{y}) &\leq \frac{2M}{\omega_n r}\int_{[|\mathbf{y}-\mathbf{x}_0|\geq\delta]} \left(\frac{r^2 - |\mathbf{x}|^2}{[|\mathbf{y}-\mathbf{x}_0| - |\mathbf{x}-\mathbf{x}_0|]^n}\right) dA(\mathbf{y}) \\
&\leq \frac{2M}{\omega_n r}\int_{[|\mathbf{y}-\mathbf{x}_0|\geq\delta]} \left(\frac{r^2 - |\mathbf{x}|^2}{\left[\delta - \frac{\delta}{2}\right]^n}\right) dA(\mathbf{y}) \\
&\leq \frac{2M}{\omega_n r}\left(\frac{2}{\delta}\right)^n \int_{\partial U} \left(r^2 - |\mathbf{x}|^2\right) dA(\mathbf{y})
\end{aligned}$$

If $|\mathbf{x} - \mathbf{x}_0|$ is sufficiently small. Then taking $|\mathbf{x} - \mathbf{x}_0|$ still smaller, if necessary, this last expression is less than $\varepsilon/2$ because $|\mathbf{x}_0| = r$ and so $\lim_{\mathbf{x}\to\mathbf{x}_0}\left(r^2 - |\mathbf{x}|^2\right) = 0$. This proves $\lim_{\mathbf{x}\to\mathbf{x}_0} u(\mathbf{x}) = g(\mathbf{x}_0)$ and this proves the existence part of this theorem. The uniqueness part follows from Corollary 5.8.

Actually, I could have said a little more about the boundary values in Theorem 5.13. Since g is continuous on ∂U, it follows g is uniformly continuous and so the above proof shows that actually $\lim_{\mathbf{x}\to\mathbf{x}_0} u(\mathbf{x}) = g(\mathbf{x}_0)$ uniformly for $\mathbf{x}_0 \in \partial U$.

Not surprisingly, it is not necessary to have the ball centered at $\mathbf{0}$ for the above to work.

Corollary 5.14 *Let $U = B(\mathbf{x}_0, r)$ and let $g \in C(\partial U)$. Then there exists a unique solution $u \in C^2(U) \cap C\left(\overline{U}\right)$ to the problem*

$$\Delta u = 0 \text{ in } U, \ u = g \text{ on } \partial U.$$

This solution is given by the formula,

$$u(\mathbf{x}) = \frac{1}{\omega_n r}\int_{\partial U} g(\mathbf{y}) \frac{r^2 - |\mathbf{x}-\mathbf{x}_0|^2}{|\mathbf{y}-\mathbf{x}|^n} dA(\mathbf{y}) \tag{5.20}$$

for every $n \geq 2$. Here $\omega_2 = 2\pi$.

This corollary implies the following.

Corollary 5.15 *Let u be a harmonic function defined on an open set, $U \subseteq \mathbb{R}^n$ and let $\overline{B(\mathbf{x}_0, r)} \subseteq U$. Then*

$$u(\mathbf{x}_0) = \frac{1}{\omega_n r^{n-1}}\int_{\partial B(\mathbf{x}_0, r)} u(\mathbf{y})\, dA$$

The representation formula, (5.16) is called Poisson's integral formula. I have now shown it works better than you had a right to expect for the Laplace equation. What happens when $f \neq 0$?

5.2.3 The Case Where $f \neq 0$, Poisson's Equation

I will verify the results for the case $n \geq 3$. The case $n = 2$ is entirely similar.

Lemma 5.16 *Let $f \in C\left(\overline{U}\right)$ or in $L^p(U)$ for $p > n/2$* [2]*. Then for each $\mathbf{x} \in U$, and $\mathbf{x}_0 \in \partial U$,*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{1}{\omega_n (n-2)} \int_U \left(\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y} - \mathbf{x})\right) f(\mathbf{y})\, dy = 0.$$

Proof:
Claim:

$$\lim_{\delta \to 0} \int_{B(\mathbf{x}_0, \delta)} \psi^{\mathbf{x}}(\mathbf{y}) |f(\mathbf{y})|\, dy = 0, \ \lim_{\delta \to 0} \int_{B(\mathbf{x}_0, \delta)} r_n(\mathbf{y} - \mathbf{x}) |f(\mathbf{y})|\, dy = 0.$$

Proof of the claim: There is nothing much to show if $\mathbf{x} = \mathbf{0}$ so suppose $\mathbf{x} \neq \mathbf{0}$.

$$\begin{aligned} \int_{B(\mathbf{x}_0, \delta)} \psi^{\mathbf{x}}(\mathbf{y}) |f(\mathbf{y})|\, dy &= \int_{B(\mathbf{0}, \delta)} r_n \left(\frac{(\mathbf{x}_0 + \mathbf{z}) |\mathbf{x}|}{r} - \frac{r\mathbf{x}}{|\mathbf{x}|} \right) |f(\mathbf{x}_0 + \mathbf{z})|\, dz \\ &= \int_0^\delta \int_{S^{n-1}} r_n \left(\frac{(\mathbf{x}_0 + \rho \mathbf{w}) |\mathbf{x}|}{r} - \frac{r\mathbf{x}}{|\mathbf{x}|} \right) |f(\mathbf{x}_0 + \rho \mathbf{w})| \rho^{n-1} d\sigma d\rho \end{aligned}$$

Now from the formula for r_n, there exists $\delta_0 > 0$ such that for $\rho \in [0, \delta_0]$,

$$r_n \left(\frac{(\mathbf{x}_0 + \rho \mathbf{w}) |\mathbf{x}|}{r} - \frac{r\mathbf{x}}{|\mathbf{x}|} \right) \rho^{n-2}$$

is bounded. Therefore,

$$\int_{B(\mathbf{x}_0, \delta)} \psi^{\mathbf{x}}(\mathbf{y}) |f(\mathbf{y})|\, dy \leq C \int_0^\delta \int_{S^{n-1}} |f(\mathbf{x}_0 + \rho \mathbf{w})| \rho\, d\sigma d\rho.$$

If f is continuous, this is dominated by an expression of the form

$$C' \int_0^\delta \int_{S^{n-1}} \rho\, d\sigma d\rho$$

which converges to 0 as $\delta \to 0$. If $f \in L^p(U)$, then by Holder's inequality, for $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned} \int_0^\delta \int_{S^{n-1}} |f(\mathbf{x}_0 + \rho \mathbf{w})| \rho\, d\sigma d\rho &= \int_0^\delta \int_{S^{n-1}} |f(\mathbf{x}_0 + \rho \mathbf{w})| \rho^{2-n} \rho^{n-1} d\sigma d\rho \\ &\leq \left(\int_0^\delta \int_{S^{n-1}} |f(\mathbf{x}_0 + \rho \mathbf{w})|^p \rho^{n-1} d\sigma d\rho \right)^{1/p} \cdot \\ &\quad \left(\int_0^\delta \int_{S^{n-1}} \left(\rho^{2-n}\right)^q \rho^{n-1} d\sigma d\rho \right)^{1/q} \\ &\leq C \|f\|_{L^p(U)}. \end{aligned}$$

Similar reasoning shows that

$$\lim_{\delta \to 0} \int_{B(\mathbf{x}_0, \delta)} r_n(\mathbf{y} - \mathbf{x}) |f(\mathbf{y})|\, dy = 0.$$

[2] If you don't know what this is, ignore it. Just do the part where f is continuous.

This proves the claim.

Let $\varepsilon > 0$ be given and choose $\delta > 0$ such that $r/2 > \delta > 0$ and small enough that

$$\int_{B(\mathbf{x}_0,\delta)} |f(\mathbf{y})| \, |\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})| \, dy < \varepsilon.$$

If $|\mathbf{x} - \mathbf{x}_0|$ is small enough, both

$$|\psi^{\mathbf{x}}(\mathbf{y})| \text{ and } r_n(\mathbf{y}-\mathbf{x})$$

are larger than $\delta/2$ for all $\mathbf{y} \in U \setminus B(\mathbf{x}_0,\delta)$. Therefore,

$$|\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})|$$

converges uniformly to 0 for $\mathbf{y} \in U \setminus B(\mathbf{x}_0,\delta)$. It follows

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \int_{U\setminus B(\mathbf{x}_0,\delta)} f(\mathbf{y}) \left(\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})\right) dy = 0.$$

This proves the lemma.

The following lemma follows from this one and Theorem 5.13.

Lemma 5.17 *Let $f \in C(\overline{U})$ or in $L^p(U)$ for $p > n/2$ and let $g \in C(\partial U)$. Then if u is given by (5.14) in the case where $n \geq 3$ or by (5.15) in the case where $n = 2$, then if $\mathbf{x}_0 \in \partial U$,*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} u(\mathbf{x}) = g(\mathbf{x}_0).$$

Not surprisingly, you can relax the condition that $g \in C(\partial U)$ but I won't do so here.

The next question is about the partial differential equation satisfied by u for u given by (5.14) in the case where $n \geq 3$ or by (5.15) for $n = 2$. This is going to introduce a new idea. I will just sketch the main ideas and leave you to work out the details, most of which have already been considered in a similar context.

Definition 5.18 *Let U be an open subset of $\mathbb{R}^n$. $C_c^\infty(U)$ is the vector space of all infinitely differentiable functions which equal zero for all $\mathbf{x}$ outside of some compact set contained in U. Similarly, $C_c^m(U)$ is the vector space of all functions which are m times continuously differentiable and whose support is a compact subset of U.*

Example 5.19 *Let $U = B(\mathbf{z}, 2r)$*

$$\psi(\mathbf{x}) = \begin{cases} \exp\left[\left(|\mathbf{x}-\mathbf{z}|^2 - r^2\right)^{-1}\right] & \text{if } |\mathbf{x}-\mathbf{z}| < r, \\ 0 \text{ if } |\mathbf{x}-\mathbf{z}| \geq r. \end{cases}$$

Then a little work shows $\psi \in C_c^\infty(U)$. This is left for you verify. The following also is easily obtained.

Lemma 5.20 *Let U be any open set. Then $C_c^\infty(U) \neq \emptyset$.*

Proof: Pick $\mathbf{z} \in U$ and let r be small enough that $B(\mathbf{z}, 2r) \subseteq U$. Then let $\psi \in C_c^\infty(B(\mathbf{z}, 2r)) \subseteq C_c^\infty(U)$ be the function of the above example.

Let $\phi \in C_c^\infty(U)$ and let $\mathbf{x} \in U$. Let U_ε denote the open set which has $\overline{B(\mathbf{y},\varepsilon)}$ deleted from it, much as was done earlier. In what follows I will denote with a subscript of x things for which $\mathbf{x}$ is the variable. Then denoting by $G(\mathbf{y},\mathbf{x})$ the expression $\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})$, it is easy to verify that $\Delta_x G(\mathbf{y},\mathbf{x}) = 0$ and so by Fubini's theorem,

$$\int_U \frac{1}{\omega_n(n-2)} \left[\int_U \left(\psi^{\mathbf{x}}(\mathbf{y}) - r_n(\mathbf{y}-\mathbf{x})\right) f(\mathbf{y}) \, dy\right] \Delta_x \phi(\mathbf{x}) \, dx$$

$$
\begin{aligned}
&= \lim_{\varepsilon\to 0}\int_{U_\varepsilon}\frac{1}{\omega_n(n-2)}\left[\int_U(\psi^{\mathbf{x}}(\mathbf{y})-r_n(\mathbf{y}-\mathbf{x}))f(\mathbf{y})\,dy\right]\Delta_x\phi(\mathbf{x})\,dx\\
&= \lim_{\varepsilon\to 0}\int_U\left(\int_{U_\varepsilon}\frac{1}{\omega_n(n-2)}(\psi^{\mathbf{x}}(\mathbf{y})-r_n(\mathbf{y}-\mathbf{x}))\Delta_x\phi(\mathbf{x})\,dx\right)f(\mathbf{y})\,dy\\
&= \lim_{\varepsilon\to 0}\frac{1}{\omega_n(n-2)}\int_U f(\mathbf{y})\left[-\int_{\partial B(\mathbf{y},\varepsilon)}\left(G\frac{\partial\phi}{\partial n_x}-\phi\frac{\partial G}{\partial n_x}\right)dA(x)\right]dy\\
&= \lim_{\varepsilon\to 0}\frac{1}{\omega_n(n-2)}\int_U f(\mathbf{y})\int_{\partial B(\mathbf{y},\varepsilon)}\phi\frac{\partial G}{\partial n_x}dA(x)\,dy\\
&= \lim_{\varepsilon\to 0}\frac{1}{\omega_n(n-2)}\int_U f(\mathbf{y})\int_{\partial B(\mathbf{y},\varepsilon)}\phi\frac{\partial r_n}{\partial n_x}(\mathbf{x}-\mathbf{y})\,dA(x)\,dy\\
&= \lim_{\varepsilon\to 0}\frac{1}{\omega_n}\int_U f(\mathbf{y})\int_{\partial B(\mathbf{y},\varepsilon)}\phi\frac{1}{\varepsilon^{n-1}}dA(x)\,dy = \int_U f(\mathbf{y})\,\phi(\mathbf{y})\,dy.
\end{aligned}
$$

Similar but easier reasoning shows that

$$\int_U\left(\frac{1}{\omega_n r}\int_{\partial U}g(\mathbf{y})\frac{r^2-|\mathbf{x}|^2}{|\mathbf{y}-\mathbf{x}|^n}dA(\mathbf{y})\right)\Delta_x\phi(\mathbf{x})\,dx=0.$$

Therefore, if $n\geq 3$, and u is given by (5.14), then whenever $\phi\in C_c^\infty(U)$,

$$\int_U u\Delta\phi dx=\int_U f\phi dx. \tag{5.21}$$

The same result holds for $n=2$.

Definition 5.21 *$\Delta u=f$ on U in the weak sense or in the sense of distributions if for all $\phi\in C_c^\infty(U)$, (5.21) holds.*

This with Lemma 5.17 proves the following major theorem.

Theorem 5.22 *Let $f\in C(\overline{U})$ or in $L^p(U)$ for $p>n/2$ and let $g\in C(\partial U)$. Then if u is given by (5.14) in the case where $n\geq 3$ or by (5.15) in the case where $n=2$, then u solves the differential equation of the Poisson problem in the sense of distributions along with the boundary conditions.*

5.3 The Half Plane

Everything which was done for a ball will work for a half plane, H which is of the form,

$$H\equiv\{\mathbf{x}\in\mathbb{R}^n : x_n>0\}.$$

I will only consider the Laplace equation with Dirichlet condition,

$$\Delta u=0,\ \text{in } H,\ u=g \text{ on } \partial H=\mathbb{R}^{n-1}.$$

The same things will work except in this case, it is problematic to base the derivation on an appeal to the divergence theorem or Green's formula because this was not established for unbounded sets. However, I will pretend there is no problem with this issue in coming up with the formula. Then it will be easy to verify that it works.

For $\mathbf{x}=(x_1,\cdots,x_{n-1},x_n)\in\mathbb{R}^n$, define

$$\mathbf{x}^*\equiv(x_1,\cdots,x_{n-1},-x_n).$$

For $\mathbf{x} \in H$, let

$$\psi^{\mathbf{x}}(\mathbf{y}) \equiv r_n(\mathbf{y} - \mathbf{x}^*)$$

and define

$$v(\mathbf{x}, \mathbf{y}) \equiv r_n(\mathbf{y} - \mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y}).$$

Then $\Delta v = 0$ and for $\mathbf{y} \in \partial H = \mathbb{R}^{n-1}$,

$$v(\mathbf{x}, \mathbf{y}) = 0.$$

As before, you might expect that for $n \geq 3$,

$$-\int_H v \Delta u dx = \int_{\partial H} u \frac{\partial v}{\partial n} dA + u(\mathbf{x})(n-2)\omega_n, \tag{5.22}$$

and in the case where $n = 2$,

$$-\int_H v \Delta u dx = \int_{\partial H} u \frac{\partial v}{\partial n} dA - u(\mathbf{x}) 2\pi. \tag{5.23}$$

and so for the Laplace equation with Dirichlet conditions, for $n \geq 3$,

$$u(\mathbf{x}) = \frac{-1}{(n-2)\omega_n} \int_{\partial H} g \frac{\partial v}{\partial n} dA, \tag{5.24}$$

and in the case where $n = 2$,

$$u(\mathbf{x}) = \frac{1}{2\pi} \int_{\partial H} g \frac{\partial v}{\partial n} dA. \tag{5.25}$$

It is of course necessary to find $\frac{\partial v}{\partial n}$. The unit outer normal in this case is just $-\mathbf{e}_n$. and when you work it out using the obvious fact that $|\mathbf{y} - \mathbf{x}| = |\mathbf{y} - \mathbf{x}^*|$, you get

$$\frac{-2x_n(n-2)}{|\mathbf{y} - \mathbf{x}|^n}.$$

Therefore, in the case where $n \geq 3$,

$$u(\mathbf{x}) = \frac{1}{\omega_n} \int_{\mathbb{R}^{n-1}} g(\mathbf{y}') \frac{2x_n}{|\mathbf{y}' - \mathbf{x}'|^n} dy' \tag{5.26}$$

where for $\mathbf{x} = (x_1, \cdots, x_{n-1} x_n)$, $\mathbf{x}' = (x_1, \cdots, x_{n-1})$. In case $n = 2$, the same formula results. Everything else will work just as well as with a ball but there is one loose end. Why is

$$\frac{1}{\omega_n} \int_{\partial H} \frac{2x_n}{|\mathbf{y} - \mathbf{x}|^n} dA = 1?$$

This was the key result in showing the boundary conditions held. This equals

$$\frac{2x_n}{\omega_n} \int_{\partial H} \frac{1}{\left(x_n^2 + |\mathbf{y}' - \mathbf{x}'|^2\right)^{n/2}} dA(\mathbf{y}) = \frac{2x_n}{\omega_n} \int_{\mathbb{R}^{n-1}} \frac{1}{\left(x_n^2 + |\mathbf{y}' - \mathbf{x}'|^2\right)^{n/2}} dy'$$

$$
\begin{aligned}
&= \frac{2x_n}{\omega_n}\int_{\mathbb{R}^{n-1}} \frac{1}{\left(x_n^2+|\mathbf{y}'|^2\right)^{n/2}}dy' \\
&= \frac{2x_n}{\omega_n}\int_0^{\infty}\int_{S^{n-1}} \frac{\rho^{n-2}}{\left(x_n^2+\rho^2\right)^{n/2}}dAd\rho \\
&= \frac{2x_n\omega_{n-1}}{\omega_n}\int_0^{\infty} \frac{\rho^{n-2}}{\left(x_n^2+\rho^2\right)^{n/2}}d\rho \\
&= \frac{2\omega_{n-1}}{\omega_n}\int_0^{\infty} \frac{u^{n-2}}{\left(1+u^2\right)^{n/2}}d\rho \\
&= \frac{2\omega_{n-1}}{\omega_n}\int_0^{\pi/2} \frac{(\tan\theta)^{n-2}}{(\sec\theta)^n}\sec^2(\theta)\,d\theta \\
&= \frac{2\omega_{n-1}}{\omega_n}\int_0^{\pi/2} \sin^{n-2}(\theta)\,d\theta. \qquad (5.27)
\end{aligned}
$$

Now recall (5.5) on Page 112 which said

$$\omega_n = 2\omega_{n-1}\int_0^{\pi/2}\sin^{n-2}(\theta)\,d\theta. \qquad (5.28)$$

This formula implies (5.27) equals 1. With this information, you can see that this solves the Laplace equation with the Dirichlet condition.

You should consider when the integral in (5.26) even makes sense. This will be the case if g is any bounded continuous function. This may be verified using polar coordinates. The following theorem is the final result.

Theorem 5.23 *Let g be a bounded continuous function defined on ∂H. Then there exists a solution to the Laplace equation with Dirichlet boundary conditions,*

$$\Delta u = 0 \text{ in } H, \text{ and } u = g \text{ on } \partial H$$

which satisfies $u \in C^2(H)\cap C\left(\overline{H}\right)$ and is given by

$$u(\mathbf{x}) = \frac{1}{\omega_n}\int_{\mathbb{R}^{n-1}} g(\mathbf{y}')\frac{2x_n}{|\mathbf{y}'-\mathbf{x}'|^n}dy'.$$

Note there is no uniqueness assertion made. This is because solutions are no longer unique. Consider $u(\mathbf{x}) \equiv 0$ and $u(\mathbf{x}) = x_n$.

5.4 Properties Of Harmonic Functions

Consider the problem for $g \in C(\partial U)$.

$$\Delta u = 0 \text{ in } U,\ u = g \text{ on } \partial U.$$

When $U = B(\mathbf{x}_0, r)$, it has now been shown there exists a unique solution to the above problem satisfying $u \in C^2(U)\cap C\left(\overline{U}\right)$ and it is given by the formula

$$u(\mathbf{x}) = \frac{r^2-|\mathbf{x}-\mathbf{x}_0|^2}{\omega_n r}\int_{\partial B(\mathbf{x}_0,r)}\frac{g(\mathbf{y})}{|\mathbf{y}-\mathbf{x}|^n}dA(y) \qquad (5.29)$$

It was also noted that this formula implies the mean value property for harmonic functions,

$$u(\mathbf{x}_0) = \frac{1}{\omega_n r^{n-1}}\int_{\partial B(\mathbf{x}_0,r)} u(\mathbf{y})\,dA(y). \qquad (5.30)$$

The mean value property can also be formulated in terms of an integral taken over $B(\mathbf{x}_0, r)$.

Lemma 5.24 *Let u be harmonic and C^2 on an open set, V and let $\overline{B(\mathbf{x}_0, r)} \subseteq V$. Then*

$$u(\mathbf{x}_0) = \frac{1}{m_n(B(\mathbf{x}_0, r))} \int_{B(\mathbf{x}_0,r)} u(\mathbf{y})\, dy$$

where here $m_n(B(\mathbf{x}_0, r))$ denotes the volume of the ball.

Proof: From the method of polar coordinates and the mean value property given in (5.30), along with the observation that $m_n(B(\mathbf{x}_0, r)) = \frac{\omega_n}{n} r^n$,

$$\begin{aligned}
\int_{B(\mathbf{x}_0,r)} u(\mathbf{y})\, dy &= \int_0^r \int_{S^{n-1}} u(\mathbf{x}_0 + \mathbf{y}) \rho^{n-1} dA(\mathbf{y})\, d\rho \\
&= \int_0^r \int_{\partial B(\mathbf{0},\rho)} u(\mathbf{x}_0 + \mathbf{y})\, dA(y)\, d\rho \\
&= u(\mathbf{x}_0) \int_0^r \omega_n \rho^{n-1} d\rho = u(\mathbf{x}_0) \frac{\omega_n}{n} r^n = u(\mathbf{x}_0)\, m_n(B(\mathbf{x}_0, r)).
\end{aligned}$$

This proves the lemma.

There is a very interesting theorem which says roughly that the values of a nonnegative harmonic function are all comparable. It is known as Harnack's inequality.

Theorem 5.25 *Let U be an open set and let $u \in C^2(U)$ be a nonnegative harmonic function. Also let U_1 be a connected open set which is bounded and satisfies $\overline{U_1} \subseteq U$. Then there exists a constant, C, depending only on U_1 such that*

$$\max\left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U_1}\right\} \leq C \min\left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U_1}\right\}$$

Proof: There is a positive distance between $\overline{U_1}$ and U^C because of compactness of $\overline{U_1}$. Therefore there exists $r > 0$ such that whenever $\mathbf{x} \in \overline{U_1}$, $B(\mathbf{x}, 2r) \subseteq U$. Then consider $\mathbf{x} \in \overline{U_1}$ and let $|\mathbf{x} - \mathbf{y}| < r$. Then from Lemma 5.24

$$\begin{aligned}
u(\mathbf{x}) &= \frac{1}{m_n(B(\mathbf{x}, 2r))} \int_{B(\mathbf{x},2r)} u(\mathbf{z})\, dz \\
&= \frac{1}{2^n m_n(B(\mathbf{x}, r))} \int_{B(\mathbf{x},2r)} u(\mathbf{z})\, dz \\
&\geq \frac{1}{2^n m_n(B(\mathbf{y}, r))} \int_{B(\mathbf{y},r)} u(\mathbf{z})\, dz = \frac{1}{2^n} u(\mathbf{y}).
\end{aligned}$$

The fact that $u \geq 0$ is used in going to the last line. Since $\overline{U_1}$ is compact, there exist finitely many balls having centers in $\overline{U_1}$, $\{B(\mathbf{x}_i, r)\}_{i=1}^m$ such that

$$\overline{U_1} \subseteq \cup_{i=1}^m B(\mathbf{x}_i, r/2).$$

Furthermore each of these balls must have nonempty intersection with at least one of the others because if not, it would follow that $\overline{U_1}$ would not be connected. Letting $\mathbf{x}, \mathbf{y} \in U_1$, there must be a sequence of these balls, $B_1, B_2, \cdots, B_k$ such that $\mathbf{x} \in B_1, \mathbf{y} \in B_k$, and $B_i \cap B_{i+1} \neq \emptyset$ for $i = 1, 2, \cdots, k-1$. Therefore, picking a point, $\mathbf{z}_{i+1} \in B_i \cap B_{i+1}$, the above estimate implies

$$u(\mathbf{x}) \geq \frac{1}{2^n} u(\mathbf{z}_2),\ u(\mathbf{z}_2) \geq \frac{1}{2^n} u(\mathbf{z}_3),\ u(\mathbf{z}_3) \geq \frac{1}{2^n} u(\mathbf{z}_4), \cdots,\ u(\mathbf{z}_k) \geq \frac{1}{2^n} u(\mathbf{y}).$$

Therefore,

$$u(\mathbf{x}) \geq \left(\frac{1}{2^n}\right)^k u(\mathbf{y}) \geq \left(\frac{1}{2^n}\right)^m u(\mathbf{y}).$$

Therefore, for all $\mathbf{x} \in \overline{U_1}$,

$$\sup \{u(\mathbf{y}) : \mathbf{y} \in U_1\} \leq (2^n)^m u(\mathbf{x})$$

and so

$$\max \left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U_1}\right\} = \sup \{u(\mathbf{y}) : \mathbf{y} \in U_1\}$$

$$\leq (2^n)^m \inf \{u(\mathbf{x}) : \mathbf{x} \in U_1\} = (2^n)^m \min \left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U_1}\right\}.$$

This proves the inequality.

The next theorem comes from the representation formula for harmonic functions given above.

Theorem 5.26 *Let U be an open set and suppose $u \in C^2(U)$ and u is harmonic. Then in fact, $u \in C^\infty(U)$. That is, u possesses all partial derivatives and they are all continuous.*

Proof: Let $B(\mathbf{x}_0,r) \subseteq U$. I will show that $u \in C^\infty(B(\mathbf{x}_0,r))$. From (5.29), it follows that for $\mathbf{x} \in B(\mathbf{x}_0,r)$,

$$\frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{\omega_n r} \int_{\partial B(\mathbf{x}_0,r)} \frac{u(\mathbf{y})}{|\mathbf{y} - \mathbf{x}|^n} dA(y) = u(\mathbf{x}).$$

It is obvious that $\mathbf{x} \to \frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{\omega_n r}$ is infinitely differentiable. Therefore, consider

$$\mathbf{x} \to \int_{\partial B(\mathbf{x}_0,r)} \frac{u(\mathbf{y})}{|\mathbf{y} - \mathbf{x}|^n} dA(y). \tag{5.31}$$

Take $\mathbf{x} \in B(\mathbf{x}_0, r)$ and consider a difference quotient for $t \neq 0$.

$$\left(\int_{\partial B(\mathbf{x}_0,r)} u(\mathbf{y}) \frac{1}{t} \left(\frac{1}{|\mathbf{y} - (\mathbf{x} + t\mathbf{e}_k)|^n} - \frac{1}{|\mathbf{y} - \mathbf{x}|^n}\right) dA(y)\right)$$

Then by the mean value theorem, the term

$$\frac{1}{t} \left(\frac{1}{|\mathbf{y} - (\mathbf{x} + t\mathbf{e}_k)|^n} - \frac{1}{|\mathbf{y} - \mathbf{x}|^n}\right)$$

equals

$$-n \left|\mathbf{x} + t\theta(t)\mathbf{e}_k - \mathbf{y}\right|^{-(n+2)} (x_k + \theta(t) t - y_k)$$

and as $t \to 0$, this converges uniformly for $\mathbf{y} \in \partial B(\mathbf{x}_0, r)$ to

$$-n \left|\mathbf{x} - \mathbf{y}\right|^{-(n+2)} (x_k - y_k).$$

This uniform convergence implies you can take a partial derivative of the function of $\mathbf{x}$ given in (5.31) obtaining the partial derivative with respect to x_k equals

$$\int_{\partial B(\mathbf{x}_0,r)} \frac{-n(x_k - y_k) u(\mathbf{y})}{|\mathbf{y} - \mathbf{x}|^{n+2}} dA(y).$$

Now exactly the same reasoning applies to this function of $\mathbf{x}$ yielding a similar formula. The continuity of the integrand as a function of $\mathbf{x}$ implies continuity of the partial derivatives. The idea is there is never any problem because $\mathbf{y} \in \partial B(\mathbf{x}_0, r)$ and $\mathbf{x}$ is a given point not on this boundary. This proves the theorem.

Liouville's theorem is a famous result in complex variables which asserts that an entire bounded function is constant. A similar result holds for harmonic functions.

Theorem 5.27 *(Liouville's theorem) Suppose u is harmonic on $\mathbb{R}^n$ and is bounded. Then u is constant.*

Proof: From the Poisson formula

$$\frac{r^2 - |\mathbf{x}|^2}{\omega_n r} \int_{\partial B(\mathbf{0},r)} \frac{u(\mathbf{y})}{|\mathbf{y}-\mathbf{x}|^n} dA(y) = u(\mathbf{x}).$$

Now from the discussion above,

$$\frac{\partial u(\mathbf{x})}{\partial x_k} = \frac{-2x_k}{\omega_n r} \int_{\partial B(\mathbf{x}_0,r)} \frac{u(\mathbf{y})}{|\mathbf{y}-\mathbf{x}|^n} dA(y) + \frac{r^2 - |\mathbf{x}|^2}{\omega_n r} \int_{\partial B(\mathbf{0},r)} \frac{u(\mathbf{y})(y_k - x_k)}{|\mathbf{y}-\mathbf{x}|^{n+2}} dA(y)$$

Therefore, letting $|u(\mathbf{y})| \leq M$ for all $\mathbf{y} \in \mathbb{R}^n$,

$$\begin{aligned} \left|\frac{\partial u(\mathbf{x})}{\partial x_k}\right| &\leq \frac{2|\mathbf{x}|}{\omega_n r} \int_{\partial B(\mathbf{x}_0,r)} \frac{M}{(r-|\mathbf{x}|)^n} dA(y) + \frac{\left(r^2 - |\mathbf{x}|^2\right) M}{\omega_n r} \int_{\partial B(\mathbf{0},r)} \frac{1}{(r-|\mathbf{x}|)^{n+1}} dA(y) \\ &= \frac{2|\mathbf{x}|}{\omega_n r} \frac{M}{(r-|\mathbf{x}|)^n} \omega_n r^{n-1} + \frac{\left(r^2 - |\mathbf{x}|^2\right) M}{\omega_n r} \frac{1}{(r-|\mathbf{x}|)^{n+1}} \omega_n r^{n-1} \end{aligned}$$

and these terms converge to 0 as $r \to \infty$. Since the inequality holds for all $r > |\mathbf{x}|$, it follows $\frac{\partial u(\mathbf{x})}{\partial x_k} = 0$. Similarly all the other partial derivatives equal zero as well and so u is a constant. This proves the theorem.

5.5 Laplace's Equation For General Sets

Here I will consider the Laplace equation with Dirichlet boundary conditions on a general bounded open set, U. Thus the problem of interest is

$$\Delta u = 0 \text{ on } U, \text{ and } u = g \text{ on } \partial U.$$

I will be presenting Perron's method for this problem. This method is based on exploiting properties of subharmonic functions which are functions satisfying the following definition.

Definition 5.28 *Let U be an open set and let u be a function defined on U. Then u is subharmonic if it is continuous and for all $\mathbf{x} \in U$,*

$$u(\mathbf{x}) \leq \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x},r)} u(\mathbf{y}) \, dA \tag{5.32}$$

whenever r is small enough.

Compare with Corollary 5.15.

5.5.1 Properties Of Subharmonic Functions

The first property is a maximum principle. Compare to Theorem 5.7.

Theorem 5.29 *Suppose U is a bounded open set and u is subharmonic on U and continuous on $\overline{U}$. Then*

$$\max\left\{u(\mathbf{y}) : \mathbf{y} \in \overline{U}\right\} = \max\left\{u(\mathbf{y}) : \mathbf{y} \in \partial U\right\}.$$

Proof: Suppose $\mathbf{x} \in U$ and $u(\mathbf{x}) = \max\left\{u(\mathbf{y}) : \mathbf{y} \in \overline{U}\right\} \equiv M$. Let V denote the connected component of U which contains $\mathbf{x}$. Then since u is subharmonic on V, it follows that for all small $r > 0$, $u(\mathbf{y}) = M$ for all $\mathbf{y} \in \partial B(\mathbf{x}, r)$. Therefore, there exists some $r_0 > 0$ such that $u(\mathbf{y}) = M$ for all $\mathbf{y} \in B(\mathbf{x}, r_0)$ and this shows $\{\mathbf{x} \in V : u(\mathbf{x}) = M\}$ is an open subset of V. However, since u is continuous, it is also a closed subset of V. Therefore, since V is connected,

$$\{\mathbf{x} \in V : u(\mathbf{x}) = M\} = V$$

and so by continuity of u, it must be the case that $u(\mathbf{y}) = M$ for all $\mathbf{y} \in \partial V \subseteq \partial U$. This proves the theorem because $M = u(\mathbf{y})$ for some $\mathbf{y} \in \partial U$.

As a simple corollary, the proof of the above theorem shows the following startling result.

Corollary 5.30 *Suppose U is a connected open set and that u is subharmonic on U. Then either*

$$u(\mathbf{x}) < \sup\{u(\mathbf{y}) : \mathbf{y} \in U\}$$

for all $\mathbf{x} \in U$ or

$$u(\mathbf{x}) \equiv \sup\{u(\mathbf{y}) : \mathbf{y} \in U\}$$

for all $\mathbf{x} \in U$.

The next result indicates that the maximum of any finite list of subharmonic functions is also subharmonic.

Lemma 5.31 *Let U be an open set and let $u_1, u_2, \cdots, u_p$ be subharmonic functions defined on U. Then letting*

$$v \equiv \max(u_1, u_2, \cdots, u_p),$$

it follows that v is also subharmonic.

Proof: Let $\mathbf{x} \in U$. Then whenever r is small enough to satisfy the subharmonicity condition for each u_i.

$$\begin{aligned} v(\mathbf{x}) &= \max(u_1(\mathbf{x}), u_2(\mathbf{x}), \cdots, u_p(\mathbf{x})) \\ &\leq \max\left(\frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x},r)} u_1(\mathbf{y})\, dA(y), \cdots, \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x},r)} u_p(\mathbf{y})\, dA(y)\right) \\ &\leq \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x},r)} \max(u_1, u_2, \cdots, u_p)(\mathbf{y})\, dA(y) = \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x},r)} v(\mathbf{y})\, dA(y). \end{aligned}$$

This proves the lemma.

The next lemma concerns modifying a subharmonic function on an open ball in such a way as to make the new function harmonic on the ball. Recall Corollary 5.14 which I will list here for convenience.

Corollary 5.32 *Let $U = B(\mathbf{x}_0, r)$ and let $g \in C(\partial U)$. Then there exists a unique solution $u \in C^2(U) \cap C\left(\overline{U}\right)$ to the problem*

$$\Delta u = 0 \text{ in } U, \ u = g \text{ on } \partial U.$$

This solution is given by the formula,

$$u(\mathbf{x}) = \frac{1}{\omega_n r} \int_{\partial U} g(\mathbf{y}) \frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) \tag{5.33}$$

for every $n \geq 2$. Here $\omega_2 = 2\pi$.

Definition 5.33 *Let U be an open set and let u be subharmonic on U. Then for $\overline{B(\mathbf{x}_0,r)} \subseteq U$ define*

$$u_{\mathbf{x}_0,r}(\mathbf{x}) \equiv \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \notin B(\mathbf{x}_0,r) \\ \frac{1}{\omega_n r}\int_{\partial B(\mathbf{x}_0,r)} u(\mathbf{y}) \frac{r^2-|\mathbf{x}-\mathbf{x}_0|^2}{|\mathbf{y}-\mathbf{x}|^n} dA(\mathbf{y}) & \text{if } \mathbf{x} \in B(\mathbf{x}_0,r) \end{cases}$$

Thus $u_{\mathbf{x}_0,r}$ is harmonic on $B(\mathbf{x}_0,r)$, and equals to u off $B(\mathbf{x}_0,r)$. The wonderful thing about this is that $u_{\mathbf{x}_0,r}$ is still subharmonic on all of U. Also note that from Corollary 5.15 on Page 121 every harmonic function is subharmonic.

Lemma 5.34 *Let U be an open set and $\overline{B(\mathbf{x}_0,r)} \subseteq U$ as in the above definition. Then $u_{\mathbf{x}_0,r}$ is subharmonic on U and $u \leq u_{\mathbf{x}_0,r}$.*

Proof: First I show that $u \leq u_{\mathbf{x}_0,r}$. This follows from the maximum principle. Here is why. The function $u - u_{\mathbf{x}_0,r}$ is subharmonic on $B(\mathbf{x}_0,r)$ and equals zero on $\partial B(\mathbf{x}_0,r)$. Here is why: For $\mathbf{z} \in B(\mathbf{x}_0,r)$,

$$u(\mathbf{z}) - u_{\mathbf{x}_0 r}(\mathbf{z}) = u(\mathbf{z}) - \frac{1}{\omega \rho^{n-1}} \int_{\partial B(\mathbf{z},\rho)} u_{\mathbf{x}_0,r}(\mathbf{y})\, dA(\mathbf{y})$$

for all ρ small enough. This is by the mean value property of harmonic functions and the observation that $u_{\mathbf{x}_0 r}$ is harmonic on $B(\mathbf{x}_0,r)$. Therefore, from the fact that u is subharmonic,

$$u(\mathbf{z}) - u_{\mathbf{x}_0 r}(\mathbf{z}) \leq \frac{1}{\omega \rho^{n-1}} \int_{\partial B(\mathbf{z},\rho)} (u(\mathbf{y}) -_{\mathbf{x}_0,r}(\mathbf{y}))\, dA(\mathbf{y})$$

Therefore, for all $\mathbf{x} \in B(\mathbf{x}_0,r)$,

$$u(\mathbf{x}) - u_{\mathbf{x}_0,r}(\mathbf{x}) \leq 0.$$

The two functions are equal off $B(\mathbf{x}_0,r)$.

The condition for being subharmonic is clearly satisfied at every point, $\mathbf{x} \notin \overline{B(\mathbf{x}_0,r)}$. It is also satisfied at every point of $B(\mathbf{x}_0,r)$ thanks to the mean value property, Corollary 5.15 on Page 121. It is only at the points of $\partial B(\mathbf{x}_0,r)$ where the condition needs to be checked. Let $\mathbf{z} \in \partial B(\mathbf{x}_0,r)$. Then since u is given to be subharmonic, it follows that for all r small enough,

$$\begin{aligned} u_{\mathbf{x}_0,r}(\mathbf{z}) &= u(\mathbf{z}) \leq \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x}_0,r)} u(\mathbf{y})\, dA \\ &\leq \frac{1}{\omega_n r^{n-1}} \int_{\partial B(\mathbf{x}_0,r)} u_{\mathbf{x}_0,r}(\mathbf{y})\, dA. \end{aligned}$$

This proves the lemma.

Definition 5.35 *For U a bounded open set and $g \in C(\partial U)$, define*

$$w_g(\mathbf{x}) \equiv \sup\{u(\mathbf{x}) : u \in S_g\}$$

where S_g consists of those functions u which are subharmonic with $u(\mathbf{y}) \leq g(\mathbf{y})$ for all $\mathbf{y} \in \partial U$ and $u(\mathbf{y}) \geq \min\{g(\mathbf{y}) : \mathbf{y} \in \partial U\} \equiv m$.

Note that $S_g \neq \emptyset$ because $u(\mathbf{x}) \equiv m$ is a member of S_g. Also all functions in S_g have values between m and $\max\{g(\mathbf{y}) : \mathbf{y} \in \partial U\}$. The fundamental result is the following absolutely amazing incredible result.

Proposition 5.36 *Let U be a bounded open set and let $g \in C(\partial U)$. Then $w_g \in S_g$ and in addition to this, w_g is harmonic.*

Proof: Let $\overline{B(\mathbf{x}_0, 2r)} \subseteq U$ and let $\{\mathbf{x}_k\}_{k=1}^{\infty}$ denote a countable dense subset of $\overline{B(\mathbf{x}_0, r)}$. Let $\{u_{1k}\}$ denote a sequence of functions of S_g with the property that

$$\lim_{k \to \infty} u_{1k}(\mathbf{x}_1) = w_g(\mathbf{x}_1).$$

By Lemma 5.34, it can be assumed each u_{1k} is a harmonic function in $B(\mathbf{x}_0, 2r)$ since otherwise, you could use the process of replacing u with $u_{\mathbf{x}_0, 2r}$. Similarly, for each l, there exists a sequence of harmonic functions in S_g, $\{u_{lk}\}$ with the property that

$$\lim_{k \to \infty} u_{lk}(\mathbf{x}_l) = w_g(\mathbf{x}_l).$$

Now define

$$w_k = (\max(u_{1k}, \cdots, u_{kk}))_{\mathbf{x}_0, 2r}.$$

Then each $w_k \in S_g$, each w_k is harmonic in $B(\mathbf{x}_0, 2r)$, and for each $\mathbf{x}_l$,

$$\lim_{k \to \infty} w_k(\mathbf{x}_l) = w_g(\mathbf{x}_l).$$

For $\mathbf{x} \in \overline{B(\mathbf{x}_0, r)}$

$$w_k(\mathbf{x}) = \frac{1}{\omega_n 2r} \int_{\partial B(\mathbf{x}_0, 2r)} w_k(\mathbf{y}) \frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) \tag{5.34}$$

and so there exists a constant, C which is independent of k such that for all $i = 1, 2, \cdots, n$ and $\mathbf{x} \in \overline{B(\mathbf{x}_0, r)}$,

$$\left| \frac{\partial w_k(\mathbf{x})}{\partial x_i} \right| \leq C$$

Therefore, this set of functions, $\{w_k\}$ is equicontinuous on $\overline{B(\mathbf{x}_0, r)}$ as well as being uniformly bounded and so by the Ascoli Arzela theorem, it has a subsequence which converges uniformly on $\overline{B(\mathbf{x}_0, r)}$ to a continuous function I will denote by w which has the property that for all k,

$$w(\mathbf{x}_k) = w_g(\mathbf{x}_k) \tag{5.35}$$

Also since each w_k is harmonic,

$$w_k(\mathbf{x}) = \frac{1}{\omega_n r} \int_{\partial B(\mathbf{x}_0, r)} w_k(\mathbf{y}) \frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) \tag{5.36}$$

Passing to the limit in (5.36) using the uniform convergence, it follows

$$w(\mathbf{x}) = \frac{1}{\omega_n r} \int_{\partial B(\mathbf{x}_0, r)} w(\mathbf{y}) \frac{r^2 - |\mathbf{x} - \mathbf{x}_0|^2}{|\mathbf{y} - \mathbf{x}|^n} dA(\mathbf{y}) \tag{5.37}$$

which shows that w is also harmonic. I have shown that $w = w_g$ on a dense set. Also, it follows that $w(\mathbf{x}) \leq w_g(\mathbf{x})$ for all $\mathbf{x} \in \overline{B(\mathbf{x}_0, r)}$. It remains to verify these two functions are in fact equal.

Claim: w_g is lower semicontinuous on U.

Proof of claim: Suppose $\mathbf{z}_k \to \mathbf{z}$. I need to verify that

$$\liminf_{k \to \infty} w_g(\mathbf{z}_k) \geq w_g(\mathbf{z}).$$

Let $\varepsilon > 0$ be given and pick $u \in S_g$ such that $w_g(\mathbf{z}) - \varepsilon < u(\mathbf{z})$. Then

$$w_g(\mathbf{z}) - \varepsilon < u(\mathbf{z}) = \lim \inf_{k \to \infty} u(\mathbf{z}_k) \leq \lim \inf_{k \to \infty} w_g(\mathbf{z}_k).$$

Since ε is arbitrary, this proves the claim.

Using the claim, let $\mathbf{x} \in \overline{B(\mathbf{x}_0, r)}$ and pick $\mathbf{x}_{k_l} \to \mathbf{x}$ where $\{\mathbf{x}_{k_l}\}$ is a subsequence of the dense set, $\{\mathbf{x}_k\}$. Then

$$w_g(\mathbf{x}) \geq w(\mathbf{x}) = \lim \inf_{l \to \infty} w(\mathbf{x}_{k_l}) = \lim \inf_{l \to \infty} w_g(\mathbf{x}_{k_l}) \geq w_g(\mathbf{x}).$$

This proves $w = w_g$ and since w is harmonic, so is w_g. This proves the proposition.

It remains to consider whether the boundary values are assumed. This requires an additional assumption on the set, U. It is a remarkably mild assumption, however.

Definition 5.37 *A bounded open set, U has the barrier condition at $\mathbf{z} \in \partial U$, if there exists a function, $b_{\mathbf{z}}$ called a barrier function which has the property that $b_{\mathbf{z}}$ is subharmonic on U, $b_{\mathbf{z}}(\mathbf{z}) = 0$, and for all $\mathbf{x} \in \partial U \setminus \{\mathbf{z}\}, b_{\mathbf{z}}(\mathbf{x}) < 0$.*

The main result is the following remarkable theorem.

Theorem 5.38 *Let U be a bounded open set which has the barrier condition at $\mathbf{z} \in \partial U$ and let $g \in C(\partial U)$. Then the function, w_g, defined above is in $C^2(U)$ and satisfies*

$$\Delta w_g = 0 \text{ in } U,$$

$$\lim_{\mathbf{x} \to \mathbf{z}} w_g(\mathbf{x}) = g(\mathbf{z}).$$

Proof: From Proposition 5.36 it follows $\Delta w_g = 0$. Let $\mathbf{z} \in \partial U$ and let $b_{\mathbf{z}}$ be the barrier function at $\mathbf{z}$. Then letting $\varepsilon > 0$ be given, the function

$$u_-(\mathbf{x}) \equiv \max(g(\mathbf{z}) - \varepsilon + K b_{\mathbf{z}}(\mathbf{x}), m)$$

is subharmonic for all $K > 0$.

Claim: For K large enough, $g(\mathbf{z}) - \varepsilon + K b_{\mathbf{z}}(\mathbf{x}) \leq g(\mathbf{x})$ for all $\mathbf{x} \in \partial U$.

Proof of claim: Let $\delta > 0$ and let $B_\delta = \max\{b_{\mathbf{z}}(\mathbf{x}) : \mathbf{x} \in \partial U \setminus B(\mathbf{z}, \delta)\}$. Then $B_\delta < 0$ by assumption and the compactness of $\partial U \setminus B(\mathbf{z}, \delta)$. Choose $\delta > 0$ small enough that if $|\mathbf{x} - \mathbf{z}| < \delta$, then $g(\mathbf{x}) - g(\mathbf{z}) + \varepsilon > 0$. Then for $|\mathbf{x} - \mathbf{z}| < \delta$,

$$b_{\mathbf{z}}(\mathbf{x}) \leq \frac{g(\mathbf{x}) - g(\mathbf{z}) + \varepsilon}{K}$$

for any choice of positive K. Now choose K large enough that $B_\delta < \frac{g(\mathbf{x}) - g(\mathbf{z}) + \varepsilon}{K}$ for all $\mathbf{x} \in \partial U$. This can be done because $B_\delta < 0$. It follows the above inequality holds for all $\mathbf{x} \in \partial U$. This proves the claim.

Let K be large enough that the conclusion of the above claim holds. Then, for all $\mathbf{x}$, $u_-(\mathbf{x}) \leq g(\mathbf{x})$ for all $\mathbf{x} \in \partial U$ and so $u_- \in S_g$ which implies $u_- \leq w_g$ and so

$$g(\mathbf{z}) - \varepsilon + K b_{\mathbf{z}}(\mathbf{x}) \leq w_g(\mathbf{x}). \tag{5.38}$$

This is a very nice inequality and I would like to say

$$\lim_{\mathbf{x} \to \mathbf{z}} g(\mathbf{z}) - \varepsilon + K b_{\mathbf{z}}(\mathbf{x}) = g(\mathbf{z}) - \varepsilon \leq \lim \inf_{\mathbf{x} \to \mathbf{z}} w_g(\mathbf{x}) \leq \lim \sup_{\mathbf{x} \to \mathbf{z}} w_g(\mathbf{x}) = w_g(\mathbf{z}) \leq g(\mathbf{z})$$

but this would be wrong because I do not know that w_g is continuous at a boundary point. I only have shown that it is harmonic in U. Therefore, a little more is required. Let

$$u_+(\mathbf{x}) \equiv g(\mathbf{z}) + \varepsilon - K b_{\mathbf{z}}(\mathbf{x}).$$

Then $-u_+$ is subharmonic and also if K is large enough, it follows from reasoning similar to that of the above claim that

$$-u_+(\mathbf{x}) = -g(\mathbf{z}) - \varepsilon + Kb_{\mathbf{z}}(\mathbf{x}) \leq -g(\mathbf{x})$$

on ∂U. Therefore, letting $u \in S_g, u - u_+$ is a subharmonic function which satisfies for $\mathbf{x} \in \partial U$,

$$u(\mathbf{x}) - u_+(\mathbf{x}) \leq g(\mathbf{x}) - g(\mathbf{x}) = 0.$$

Consequently, the maximum principle implies $u \leq u_+$ and so since this holds for every $u \in S_g$, it follows

$$w_g(\mathbf{x}) \leq u_+(\mathbf{x}) = g(\mathbf{z}) + \varepsilon - Kb_{\mathbf{z}}(\mathbf{x}).$$

It follows that

$$g(\mathbf{z}) - \varepsilon + Kb_{\mathbf{z}}(\mathbf{x}) \leq w_g(\mathbf{x}) \leq g(\mathbf{z}) + \varepsilon - Kb_{\mathbf{z}}(\mathbf{x})$$

and so,

$$g(\mathbf{z}) - \varepsilon \leq \liminf_{\mathbf{x}\to\mathbf{z}} w_g(\mathbf{x}) \leq \limsup_{\mathbf{x}\to\mathbf{z}} w_g(\mathbf{x}) \leq g(\mathbf{z}) + \varepsilon.$$

Since ε is arbitrary, this shows

$$\lim_{\mathbf{x}\to\mathbf{z}} w_g(\mathbf{x}) = g(\mathbf{z}).$$

This proves the theorem.

5.5.2 Poisson's Problem Again

Corollary 5.39 *Let U be a bounded open set which has the barrier condition and let $f \in C\left(\overline{U}\right), g \in C(\partial U)$. Then there exists at most one solution, $u \in C^2(U) \cap C\left(\overline{U}\right)$ to Poisson's problem. If there is a solution, then it is of the form*

$$u(\mathbf{x}) = \frac{-1}{(n-2)\,\omega_n}\left[\int_U G(\mathbf{x},\mathbf{y})\, f(\mathbf{y})\, dy + \int_{\partial U} g(\mathbf{y}) \frac{\partial G}{\partial n_y}(\mathbf{x},\mathbf{y})\, dA(\mathbf{y})\right], \text{ if } n \geq 3, \tag{5.39}$$

$$u(\mathbf{x}) = \frac{1}{2\pi}\left[\int_{\partial U} g(\mathbf{y}) \frac{\partial G}{\partial n_y}(\mathbf{x},\mathbf{y})\, dA + \int_U G(\mathbf{x},\mathbf{y})\, f(\mathbf{y})\, dx\right], \text{ if } n = 2 \tag{5.40}$$

for $G(\mathbf{x},\mathbf{y}) = r_n(\mathbf{y}-\mathbf{x}) - \psi^{\mathbf{x}}(\mathbf{y})$ where $\psi^{\mathbf{x}}$ is a function which satisfies $\psi^{\mathbf{x}} \in C^2(U) \cap C\left(\overline{U}\right)$

$$\Delta\psi^{\mathbf{x}} = 0,\ \psi^{\mathbf{x}}(\mathbf{y}) = r_n(\mathbf{x}-\mathbf{y}) \text{ for } \mathbf{y} \in \partial U.$$

Furthermore, if u is given by the above representations, then u is a weak solution to Poisson's problem.

Proof: Uniqueness follows from Corollary 5.8 on Page 114. If u_1 and u_2 both solve the Poisson problem, then their difference, w satisfies

$$\Delta w = 0, \text{ in } U,\ w = 0 \text{ on } \partial U.$$

The same arguments used earlier show that the representations in (5.39) and (5.40) both yield a weak solution to Poisson's problem.

The function, G in the above representation is called Green's function. Much more can be said about the Green's function.

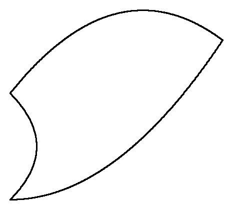
How can you recognize that a bounded open set, U has the barrier condition? One way would be to check the following condition.

Condition 5.40 *For each* $\mathbf{z} \in \partial U$, *there exists* $\mathbf{x_z} \notin \overline{U}$ *such that* $|\mathbf{x_z} - \mathbf{z}| < |\mathbf{x_z} - \mathbf{y}|$ *for every* $\mathbf{y} \in \partial U \setminus \{\mathbf{z}\}$.

Proposition 5.41 *Suppose Condition 5.40 holds. Then* U *satisfies the barrier condition.*

Proof: For $n \geq 3$, let $b_{\mathbf{z}}(\mathbf{y}) \equiv r_n(\mathbf{y} - \mathbf{x_z}) - r_n(\mathbf{z} - \mathbf{x_z})$. Then $b_{\mathbf{z}}(\mathbf{z}) = 0$ and if $\mathbf{y} \in \partial U$ with $\mathbf{y} \neq \mathbf{z}$, then clearly $b_{\mathbf{z}}(\mathbf{y}) < 0$. For $n = 2$, let $b_{\mathbf{z}}(\mathbf{y}) = -\ln|\mathbf{y} - \mathbf{x_z}| + \ln|\mathbf{z} - \mathbf{x_z}|$. This works out the same way.

Here is a picture of a domain which satisfies the barrier condition.



In fact, you have to have a fairly pathological example in order to find something which does not satisfy the barrier condition. You might try to think of some examples. Think of $B(\mathbf{0}, 1) \setminus \{z \text{ axis}\}$ for example. The points on the z axis which are in $B(\mathbf{0}, 1)$ become boundary points of this new set. Thus this set can't satisfy the above condition. Could this set have the barrier property?

Maximum Principles

6.1 Elliptic Equations

Definition 6.1 *Let the functions, a_{ij} be continuous and suppose $a_{ij}(\mathbf{x}) = a_{ji}(\mathbf{x})$ and satisfy*

$$\sum_{ij} a_{ij}(\mathbf{x}) \xi_i \xi_j \geq \delta^2 |\boldsymbol{\xi}|^2, \ \delta > 0. \tag{6.1}$$

Then an elliptic operator is one which is of the form

$$Lu(\mathbf{x}) = \sum_{ij} a_{ij}(\mathbf{x}) u_{,ij}(\mathbf{x}) + \sum_i b_i(\mathbf{x}) u_{,i} + c(\mathbf{x}) u(\mathbf{x}) \tag{6.2}$$

where b_i and c_i are also continuous functions. An elliptic equation will be one which is of the form

$$Lu = f$$

where L is given in (6.2).

6.2 Maximum Principles For Elliptic Problems

There are two maximum principles for elliptic equations which are of major importance, the weak maximum principle and the strong maximum principle. For much more on maximum principles than presented here you should see the book by Protter and Weinberger, [15].

6.2.1 Weak Maximum Principle

The weak maximum principle is as follows.

Theorem 6.2 *Let U be a bounded open set and suppose $u \in C^2(U) \cap C(\overline{U})$. Also suppose that $c(\mathbf{x}) = 0$ in the above definition of L and*

$$Lu(\mathbf{x}) \geq 0 \tag{6.3}$$

for all $\mathbf{x} \in U$. Then

$$\max\{u(\mathbf{x}) : \mathbf{x} \in \overline{U}\} = \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\}. \tag{6.4}$$

Proof: Suppose the conclusion is not true. Then there exists $\mathbf{x}_0 \in U$ such that $u(\mathbf{x}_0) > \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\}$. Now consider

$$v_\varepsilon(\mathbf{x}) \equiv u(\mathbf{x}) + \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}}$$

where $\mathbf{z} \notin \overline{U}$. Also suppose that R is large enough that $B(\mathbf{z}, R) \supseteq U$. I claim that for ε small enough, v_ε has its maximum at a point of U. To see this, let

$$\delta \equiv u(\mathbf{x}_0) - \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\}.$$

Then

$$\begin{aligned} v_\varepsilon(\mathbf{x}_0) - \max\{v_\varepsilon(\mathbf{x}) : \mathbf{x} \in \partial U\} &\geq u(\mathbf{x}_0) - \varepsilon e^{\frac{R^2}{\alpha}} \\ &\quad - \max\{u(\mathbf{x}) : \mathbf{x} \in \partial U\} \\ &= \delta - \varepsilon e^{\frac{R^2}{\alpha}} > 0 \end{aligned}$$

if

$$\varepsilon < (\delta/2)\, e^{-\frac{R^2}{\alpha}}.$$

Claim: Let $w(\mathbf{x}) = \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}}$. Then $L(w) > 0$ if α is small enough.

Proof of claim:

$$w_{,i} = \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}} \left(2\frac{(x_i - z_j)}{\alpha}\right).$$

Then

$$\begin{aligned} w_{,ij} &= \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}} \left[\left(2\frac{(x_i - z_i)}{\alpha}\right)\left(2\frac{(x_j - z_j)}{\alpha}\right) + \delta_{ij}\frac{2}{\alpha}\right] \\ &= \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}} \left[\frac{4}{\alpha^2}(x_i - z_i)(x_j - z_j) + \delta_{ij}\frac{2}{\alpha}\right]. \end{aligned}$$

It follows

$$\begin{aligned} L(w) &= \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}} \left\{\sum_{ij} a_{ij}\left[\frac{4}{\alpha^2}(x_i - z_i)(x_j - z_j) + \delta_{ij}\frac{2}{\alpha}\right] + \sum_i b_i\left(2\frac{(x_i - z_j)}{\alpha}\right)\right\} \\ &\geq \varepsilon e^{\frac{|\mathbf{x}-\mathbf{z}|^2}{\alpha}} \left\{\frac{4}{\alpha^2}\delta^2 |\mathbf{x}-\mathbf{z}|^2 + \operatorname{trace}((a_{ij}))\frac{2}{\alpha} + \sum_i b_i\left(2\frac{(x_i - z_j)}{\alpha}\right)\right\}. \end{aligned}$$

Now since $\mathbf{z} \notin \overline{U}$, $|\mathbf{x}-\mathbf{z}|$ is bounded away from zero and so for small enough α, the above expression is larger than 0. This establishes the claim.

Pick such an $\varepsilon > 0$ described above and to save on notation refer to v_ε more simply as v from now on. Let $\mathbf{q}$ be a point of $\overline{U}$ where the maximum value of v is achieved. Thus

$$v(\mathbf{q}) \equiv \max\left\{v(\mathbf{x}) : \mathbf{x} \in \overline{U}\right\} > \max\{v(\mathbf{x}) : \mathbf{x} \in \partial U\}.$$

Now change coordinates letting $v(\mathbf{y}) = v(\mathbf{x})$ where for Q an orthogonal matrix,

$$x_i = \sum_j Q_{ij} y_j,\ \sum_i Q_{ij} x_i = y_j$$

for Q an orthogonal matrix. That is, $Q^T Q = I$. Then

$$\frac{\partial v}{\partial x_i} = \sum_k \frac{\partial v}{\partial y_k}\frac{\partial y_k}{\partial x_i} = \sum_k \frac{\partial v}{\partial y_k} Q_{ik}$$

and so also

$$\frac{\partial^2 v}{\partial x_j \partial x_i}(\mathbf{x}) = \sum_{k,l} \frac{\partial^2 v}{\partial y_l \partial y_k}(\mathbf{y}) Q_{jl} Q_{ik}.$$

Now from the assumption that $Lu \geq 0$ in U,

$$0 \leq Lu(\mathbf{q}) = Lv(\mathbf{q}) - L(w(\mathbf{q}))$$

and so

$$Lv(\mathbf{q}) \geq L(w(\mathbf{q})) > 0.$$

Since $\mathbf{q}$ is a local maximum, the partial derivatives of v all vanish at $\mathbf{q}$ and so letting $\mathbf{q} = Q\mathbf{y}_\mathbf{q}$

$$0 < Lv(\mathbf{q}) = \sum_{ij} a_{ij}(\mathbf{q}) \frac{\partial^2 v}{\partial x_j \partial x_i}(\mathbf{q}) = \sum_{ij} a_{ij}(\mathbf{q}) \sum_{k,l} \frac{\partial^2 v}{\partial y_l \partial y_k}(\mathbf{y}_\mathbf{q}) Q_{jl} Q_{ik}. \tag{6.5}$$

Let Q be an orthogonal matrix with the property that

$$Q^T(a_{ij}(\mathbf{q}))Q = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

where the λ_i are the eigenvalues of the symmetric matrix, $(a_{ij}(\mathbf{q}))$. By condition (6.1) all these eigenvalues are positive, each larger than δ. Then the right side of (6.5) reduces to $\sum_k \lambda_k \frac{\partial^2 v}{(\partial y_k)^2}(\mathbf{y}_\mathbf{q})$ and so from (6.5),

$$0 < \sum_k \lambda_k \frac{\partial^2 v}{\partial y_k^2}(\mathbf{y}_\mathbf{q}) \tag{6.6}$$

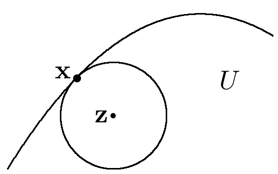
but $\mathbf{y} \to v(\mathbf{y})$ has a local maximum at $\mathbf{y}_\mathbf{q}$ and so by the second derivative test,

$$\frac{\partial^2 v}{\partial y_k^2}(\mathbf{y}_\mathbf{q}) \leq 0$$

which shows the right side of (6.6) is no larger than 0, a contradiction. This proves the theorem.

6.2.2 Strong Maximum Principle

Definition 6.3 *Let U be an open set. Then U has the interior ball condition at $\mathbf{x} \in \partial U$ if there exists $\mathbf{z} \in U$ and $r > 0$ such that $B(\mathbf{z}, r) \subseteq U$ and $\mathbf{x} \in \partial B(\mathbf{z}, r)$.*



The following lemma of Hopf is the main idea.

Lemma 6.4 *Let U be a bounded open set and suppose $\mathbf{x}_0 \in \partial U$ and U has the interior ball condition at $\mathbf{x}_0$ with the ball being $B(\mathbf{z}, r)$. Also let $c(\mathbf{x}) = 0$ in the definition of L, (6.2) and suppose $u \in C^2(U) \cap C^1(\overline{U})$ satisfies*

$$Lu \geq 0 \text{ in } U. \tag{6.7}$$

Then if $u(\mathbf{x}_0) = \max\left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U}\right\}$ and $u(\mathbf{x}) < u(\mathbf{x}_0)$ for $\mathbf{x} \in U$, it follows

$$\frac{\partial u}{\partial n}(\mathbf{x}_0) > 0 \tag{6.8}$$

where $\mathbf{n}$ is the exterior unit normal to the ball at the point $\mathbf{x}_0$.

Proof: Let $B(\mathbf{z}, r)$ be the interior ball for the interior ball condition such that $\mathbf{x}_0 \in \partial B(\mathbf{z}, r)$. Let

$$v(\mathbf{x}) \equiv e^{\lambda|\mathbf{x}-\mathbf{z}|^2} - e^{\lambda r^2}. \tag{6.9}$$

Thus $v(\mathbf{x}_0) = 0$. Now consider $L(v)$.

$$v_{,i} = e^{\lambda|\mathbf{x}-\mathbf{z}|^2} 2\lambda(x_i - z_i),\ v_{,ij} = e^{\lambda|\mathbf{x}-\mathbf{z}|^2} 4\lambda^2 (x_j - z_j)(x_i - z_i) + e^{\lambda|\mathbf{x}-\mathbf{z}|^2} 2\lambda\delta_{ij}. \tag{6.10}$$

For convenience, since $c(\mathbf{x}) = 0$,

$$Lu(\mathbf{x}) = \sum_{ij} a_{ij}(\mathbf{x}) u_{,ij}(\mathbf{x}) + \sum_i b_i(\mathbf{x}) u_{,i} \tag{6.11}$$

Therefore, Lv is given by

$$\begin{aligned}
Lv &= e^{\lambda|\mathbf{x}-\mathbf{z}|^2} \sum_{ij} 4\lambda^2 a_{ij}(\mathbf{x})(x_j - z_j)(x_i - z_i) + 2\lambda\delta_{ij}a_{ij}(\mathbf{x}) + e^{\lambda|\mathbf{x}-\mathbf{z}|^2} \sum_i b_i(\mathbf{x}) 2\lambda(x_i - z_i) \\
&= e^{\lambda|\mathbf{x}-\mathbf{z}|^2} \left(\sum_{ij} 4\lambda^2 a_{ij}(\mathbf{x})(x_j - z_j)(x_i - z_i) + 2\lambda \operatorname{trace}(A(\mathbf{x})) + \sum_i b_i(\mathbf{x}) 2\lambda(x_i - z_i) \right) \\
&\geq e^{\lambda|\mathbf{x}-\mathbf{z}|^2} \left(4\lambda^2\delta^2 |\mathbf{x}-\mathbf{z}|^2 + 2\lambda \operatorname{trace}(a_{ij}(\mathbf{x})) + 2\lambda \sum_i b_i(\mathbf{x})(x_i - z_i) \right)
\end{aligned}$$

Now consider the open set, A which consists of the points between $B(\mathbf{z}, r/2)$ and $B(\mathbf{z}, r)$. Thus on A, $|\mathbf{x}-\mathbf{z}| \geq r/2$ and so $Lv \geq 0$ provided $|\lambda|$ is large enough. In this argument, λ will be a negative number having absolute value large enough that this condition holds. Also note that from (6.10) and $\lambda < 0$, it follows that

$$\frac{\partial v}{\partial n}(\mathbf{x}_0) < 0.$$

Consider the function, $w(\mathbf{x}) = u(\mathbf{x}) + v(\mathbf{x})$. Then $Lw \geq 0$ in A. From the definition of v in (6.9) it follows $w = u$ on $\partial B(\mathbf{z}, r)$. Therefore, $w - u(\mathbf{x}_0) \leq 0$ on $\partial B(\mathbf{z}, r)$. If $|\lambda|$ is large enough, $|v|$ is very small on $\partial B(\mathbf{z}, r/2)$ and so it can be assumed that for $x \in \partial B(\mathbf{z}, r/2)$,

$$w(\mathbf{x}) - u(\mathbf{x}_0) \equiv u(\mathbf{x}) + v(\mathbf{x}) - u(\mathbf{x}_0) < 0$$

also. By the weak maximum principle, it follows that $w - u(\mathbf{x}_0) \leq 0$ on A. But this function equals 0 at $\mathbf{x}_0$ and so letting $\mathbf{n}$ be the unit outer normal to $B(\mathbf{z}, r)$ at $\mathbf{x}_0$, it follows

$$\frac{\partial(w - u(\mathbf{x}_0))}{\partial n}(\mathbf{x}_0) \geq 0.$$

Thus

$$\frac{\partial u}{\partial n}(\mathbf{x}_0) \geq -\frac{\partial v}{\partial n}(\mathbf{x}_0) > 0$$

and this proves the lemma.

The strong maximum principle is as follows.

Theorem 6.5 *Let U be bounded, open and connected and suppose $c(\mathbf{x}) \equiv 0$ in the definition of L. Also suppose $u \in C^2(U) \cap C(\overline{U})$ satisfies $Lu \geq 0$ in U. Let*

$$M \equiv \max\left\{u(\mathbf{x}) : \mathbf{x} \in \overline{U}\right\}.$$

Then if $u(\mathbf{x}) = M$ for some $\mathbf{x} \in U$, it follows $u(\mathbf{x}) = M$ for all $\mathbf{x} \in U$.

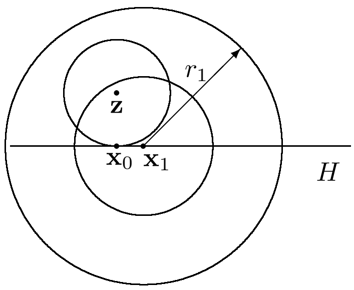
Proof: Let $H \equiv \left\{\mathbf{x} \in \overline{U} : u(\mathbf{x}) = M\right\}$. Then H is a closed set because u is continuous. Suppose $U \setminus H \neq \emptyset$. If for every $\mathbf{x} \in H \cap U$, there exists $B(\mathbf{x}, r_{\mathbf{x}})$ such that $B(\mathbf{x}, r_{\mathbf{x}}) \cap U \subseteq H$. Then

$$(\cup_{\mathbf{x} \in H \cap U} B(\mathbf{x}, r_{\mathbf{x}}) \cap U) \cup (U \setminus H) = U$$

and so U would be the union of disjoint nonempty open sets and hence not connected. Therefore, there exists $\mathbf{x}_1 \in H \cap U$ such that for all $r > 0$, $B(\mathbf{x}_1, r)$ contains points of $U \setminus H$. Pick r_1 such that $\overline{B(\mathbf{x}_1, r_1)} \subseteq U$ and choose $\mathbf{z} \in (U \setminus H) \cap B(\mathbf{x}_1, r_1/2)$. Thus

$$\operatorname{dist}(\mathbf{z}, H) < \frac{r_1}{2} < \operatorname{dist}(\mathbf{z}, \partial B(\mathbf{x}_1, r_1)) < \operatorname{dist}(\mathbf{z}, \partial U).$$

Pick $r > 0$ such that $\overline{B(\mathbf{z}, r)} \cap H \neq \emptyset$ but $B(\mathbf{z}, r) \cap H = \emptyset$. Thus $r \leq r_1/2$ because $\mathbf{z}$ is closer to $\mathbf{x}_1$ than $r_1/2$ and $\mathbf{x}_1 \in H$. Letting $\mathbf{x}_0 \in \overline{B(\mathbf{z}, r)} \cap H$, it follows $\mathbf{x}_0 \in U \cap H$ and $U \setminus H$ satisfies the interior ball condition at $\mathbf{x}_0$. The situation is illustrated in the following picture.



Therefore, $\nabla u(\mathbf{x}_0) = \mathbf{0}$ because $\mathbf{x}_0$ is an interior point at which the maximum is achieved. But then by Hopf's lemma

$$0 < \frac{\partial u}{\partial n}(\mathbf{x}_0) = \nabla u(\mathbf{x}_0) \cdot \mathbf{n} = 0,$$

a contradiction. Hence $U \setminus H = \emptyset$ and this proves the theorem.

6.3 Maximum Principles For Parabolic Problems

Definition 6.6 *A partial differential equation is called parabolic if it is of the form*

$$Lu = u_t$$

where L is the operator defined in (6.1) - (6.2).

There are maximum principles for parabolic problems just as there are for elliptic problems. In what follows, U will be an open bounded set in $\mathbb{R}^n$ and $U_T \equiv U \times (0, T)$. Also, $\Gamma_T \equiv U \times \{0\} \cup \partial U \times [0, T]$.

6.3.1 The Weak Parabolic Maximum Principle

The following theorem is the parabolic weak maximum principle.

Theorem 6.7 *Suppose* $\mathbf{x} \to \mathbf{u}(\mathbf{x},t)$ *is in* $C^2(U_T) \cap C(\overline{U_T})$ *and* $t \to u(\mathbf{x},t)$ *is in* $C^1([0,T])$. *Also suppose that in the definition of* L *given in (6.1) - (6.2),* $c = 0$ *and that* a_{ij} *and* b_i *are all continuous on* $\overline{U_T}$ *and that*

$$Lu \geq u_t$$

on U_T. *Then*

$$\max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\} = \max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\}.$$

Proof: Let $M = \max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\}$ and suppose, contrary to the conclusion of the theorem that

$$\max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\} > \max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\}$$

Then for some $(\mathbf{x}_0, t_0)$ satisfying $\mathbf{x}_0 \in U$ and $t_0 > 0$,

$$u(\mathbf{x}_0, t_0) = M.$$

Consider $v(\mathbf{x},t) \equiv u(\mathbf{x},t) - \varepsilon t$ for $t \in [0,T]$. Then if $\varepsilon > 0$ is small enough, v_ε also has the property that

$$\max\left\{v_\varepsilon(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\} > \max\left\{v_\varepsilon(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\}.$$

To see this, let $\delta = u(\mathbf{x}_0, t_0) - \max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\}$. Then

$$\begin{aligned} v_\varepsilon(\mathbf{x}_0, t_0) - \max\left\{v_\varepsilon(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\} &\geq u(\mathbf{x}_0, t_0) - \varepsilon T - \max\left\{u(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\} - \varepsilon T \\ &\geq \delta - 2\varepsilon T \end{aligned}$$

so you simply take $\varepsilon < \delta/2T$. Pick $\varepsilon > 0$ this small and denote v_ε as v from now on.

Letting $v(\mathbf{x}_1, t_1) = \max\left\{v(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\}$, it follows $\mathbf{x}_1 \in U$ and $t_1 > 0$. Therefore, $v_t(\mathbf{x}_1, t_1) \geq 0$. Also at this point,

$$Lv = Lu \geq u_t = v_t + \varepsilon \geq \varepsilon > 0.$$

Let

$$U' \equiv \left\{\mathbf{x} : Lv(\mathbf{x},t_1) > 0\right\}.$$

Then U' is open and by the weak maximum principle for elliptic problems, it follows that since $\mathbf{x}_1 \in U'$,

$$\max\left\{v(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\} = \max\left\{v(\mathbf{x},t_1) : \mathbf{x} \in \overline{U'}\right\} = \max\left\{v(\mathbf{x},t_1) : \mathbf{x} \in \partial U'\right\}.$$

If $\mathbf{x}_2 \in \partial U'$ is such that $v(\mathbf{x}_2, t_1) = \max\left\{v(\mathbf{x},t_1) : \mathbf{x} \in \overline{U'}\right\}$ which equals $\max\left\{v(\mathbf{x},t) : \mathbf{x} \in \overline{U_T}\right\}$, then it follows that $\mathbf{x}_2 \notin U$ because if it were, the same argument just given would show that $\mathbf{x}_2$ is not really a boundary point of U'. It would be a point of U'. Therefore, $\mathbf{x}_2 \in \partial U$ and so $v(\mathbf{x}_2, t_1) = \max\left\{v(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\}$ contrary to the above observation that, since ε was small enough,

$$\max\left\{v(\mathbf{x},t) : (\mathbf{x},t) \in \overline{U_T}\right\} > \max\left\{v(\mathbf{x},t) : (\mathbf{x},t) \in \Gamma_T\right\}.$$

This proves the theorem.

The following Lemma whose proof is just like the proof of Theorem 6.7 will be used in the proof of the strong maximum principle.

Lemma 6.8 *Let W be a bounded open set in $\mathbb{R}^{n+1}$ and suppose $\mathbf{x} \to u(\mathbf{x}, t)$ is in $C^2(W) \cap C(\overline{W})$ and $t \to u(\mathbf{x},t)$ is in $C^1(W)$ and $u \in C(\overline{W})$. Also suppose that in the definition of L given in (6.1) - (6.2), $c = 0$ and that a_{ij} and b_i are all continuous on $\overline{U_T}$ and that*

$$Lu \geq u_t$$

on W. Then

$$\max\left\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\} = \max\left\{u(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\}.$$

Proof: Let $M = \max\left\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\}$ and suppose, contrary to the conclusion of the theorem that

$$\max\left\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\} > \max\left\{u(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\}$$

Then for some $(\mathbf{x}_0, t_0) \in W$,

$$u(\mathbf{x}_0, t_0) = M.$$

Consider $v(\mathbf{x}, t) \equiv u(\mathbf{x}, t) - \varepsilon t$. Then if $\varepsilon > 0$ is small enough, v_ε also has the property that

$$\max\left\{v_\varepsilon(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\} > \max\left\{v_\varepsilon(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\}.$$

To see this, let $\delta = u(\mathbf{x}_0, t_0) - \max\left\{u(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\}$. Then letting $W \subseteq B(\mathbf{0}, K)$,

$$\begin{aligned} v_\varepsilon(\mathbf{x}_0, t_0) - \max\left\{v_\varepsilon(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\} &\geq u(\mathbf{x}_0, t_0) - \varepsilon K - \max\left\{u(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\} - \varepsilon K \\ &\geq \delta - 2\varepsilon K \end{aligned}$$

so you simply take $\varepsilon < \delta/2K$. Pick $\varepsilon > 0$ this small and denote v_ε as v from now on.

Letting $v(\mathbf{x}_1, t_1) = \max\left\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\}$, it follows $(\mathbf{x}_1, t_1) \in W$. Therefore, $v_t(\mathbf{x}_1, t_1) = 0$. Also at this point,

$$Lv = Lu \geq u_t = v_t + \varepsilon = \varepsilon > 0.$$

Let

$$W_{t_1} \equiv \left\{\mathbf{x} : Lv(\mathbf{x},t_1) > 0\right\}.$$

Then W_{t_1} is open in $\mathbb{R}^n$ and by the weak maximum principle for elliptic problems, it follows that since $\mathbf{x}_1 \in W_{t_1}$,

$$\max\left\{v(\mathbf{x}, t) : (\mathbf{x},t) \in \overline{W}\right\} = \max\left\{v(\mathbf{x}, t_1) : \mathbf{x} \in \overline{W_{t_1}}\right\} = \max\left\{v(\mathbf{x}, t_1) : \mathbf{x} \in \partial W_{t_1}\right\},$$

the boundary of W_{t_1} taken in $\mathbb{R}^n$. If $\mathbf{x}_2 \in \partial W_{t_1}$ is such that $v(\mathbf{x}_2, t_1) = \max\left\{v(\mathbf{x}, t_1) : \mathbf{x} \in \overline{W_{t_1}}\right\}$ which equals $\max\left\{v(\mathbf{x}, t) : \mathbf{x} \in \overline{W}\right\}$, then it follows that $(\mathbf{x}_2, t_1) \notin W$ because if it were, the same argument just given would show that $\mathbf{x}_2$ is not really a boundary point of W_{t_1}. Therefore, $(\mathbf{x}_2, t_1) \in \partial W$ and $v(\mathbf{x}_2, t_1) = \max\left\{v(\mathbf{x}, t) : (\mathbf{x},t) \in \overline{W}\right\}$ contrary to the above observation that, since ε was small enough,

$$\max\left\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{W}\right\} > \max\left\{v(\mathbf{x}, t) : (\mathbf{x},t) \in \partial W\right\}.$$

This proves the lemma.

6.3.2 The Strong Parabolic Maximum Principle

Let U_T and Γ_T be given above.

Theorem 6.9 *Suppose that in the definition of L given in (6.1) - (6.2), $c = 0$ and that a_{ij} and b_i are all continuous on $\overline{U_T}$ and that*

$$Lu \geq u_t$$

on U_T and that U is an open bounded connected set. Suppose also $\mathbf{x} \rightarrow \mathbf{u}(\mathbf{x}, t)$ is in $C^2(U_T)$ and $t \rightarrow u(\mathbf{x}, t)$ is in $C^1(U_T)$ while $u \in C\left(\overline{U_T}\right)$. Let

$$M \equiv \max\left\{u(\mathbf{x}, t) : (\mathbf{x},t) \in \overline{U_T}\right\}.$$

Then if for some $(\mathbf{x}_0, t_0) \in U_T$ with $t_0 > 0$, $u(\mathbf{x}_0, t_0) = M$, it follows that $u(\mathbf{x}, t) = M$ for all $(\mathbf{x},t) \in U \times [0, t_0]$.

The proof is accomplished through the use of the following three lemmas.

Lemma 6.10 *Let $0 \leq t_0 < t_1 < T$ and suppose $u < M$ in $U \times (t_0, t_1)$. Then $u < M$ on $U \times \{t_1\}$.*

Proof: Suppose not. Then there exists $\mathbf{x}_1 \in U$ and $u(\mathbf{x}_1, t_1) = M$. Then define

$$v(\mathbf{x}, t) \equiv \exp\left(-|\mathbf{x} - \mathbf{x}_1|^2 - \alpha(t - t_1)\right) - 1.$$

Then

$$Lv - v_t = \exp\left(-|\mathbf{x} - \mathbf{x}_1|^2 - \alpha(t - t_1)\right) \cdot$$

$$\left[4\sum_{ij} a_{ij}(x_i - x_{1i})(x_j - x_{1j}) - 2\left(\sum_i a_{ii} + b_i(x_i - x_{1i})\right) + \alpha\right]$$

which is greater than 0 for all $(\mathbf{x},t) \in U_T$ if α is large enough. Always let α be this large.

Now if r is small enough,

$$\overline{B((\mathbf{x}_1, t_1), r)} \subseteq U_T.$$

Consider the paraboloid,

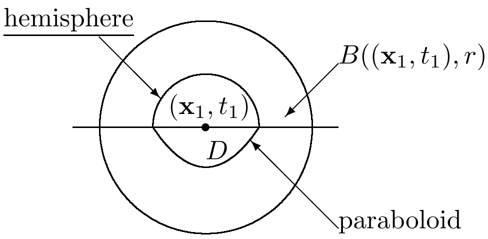
$$-|\mathbf{x} - \mathbf{x}_1|^2 - \alpha(t - t_1) = -\frac{1}{\alpha}$$

which is of the form

$$t = t_1 - \frac{1}{\alpha^2} + \frac{1}{\alpha}|\mathbf{x} - \mathbf{x}_1|^2.$$

Then if α is large enough, the intersection of this paraboloid with the plane $\{(\mathbf{x}, t_1) : \mathbf{x} \in \mathbb{R}^n\}$ is a sphere having center at $(\mathbf{x}_1, t_1)$ with radius equal to r_1 where $r_1 < r$ given above and the vertex of this paraboloid is higher than t_0. Then on this paraboloid, $v = e^{-(1/\alpha)} - 1 < 0$.

Now let D denote the open set containing $(\mathbf{x}_1, t_1)$ which lies between this paraboloid and the hemisphere centered at $(\mathbf{x}_1, t_1)$ having radius r_1 as shown in the following picture.



Thus on ∂D

$$L(u+\varepsilon v-M)-(u+\varepsilon v-M)_t=$$

$$L(u+\varepsilon v-M)-u_t-\varepsilon v_t>Lu-u_t+\varepsilon Lv-\varepsilon v_t>0.$$

On the top part of ∂D, $v<0$ and so $u+\varepsilon v-M<0$. On the bottom part of ∂D it was just noted that $v<0$ and so $u+\varepsilon v-M<0$ on ∂D. Now by Lemma 6.8 it follows that $u+\varepsilon v-M<0$ in D. However, this function equals zero at $(\mathbf{x}_1,t_1)$ which is a contradiction. This proves the lemma.

The next lemma indicates that u equals M only at the top or bottom of balls in which $u<M$. More precisely,

Lemma 6.11 *If $u(\mathbf{x}_0,t_0)=M$ where $(\mathbf{x}_0,t_0)$ on $\partial B((\mathbf{z},\tau),r)$ where $\overline{B((\mathbf{z},\tau),r)}\subseteq U_T$ but $u(\mathbf{x},t)<M$ for all $(\mathbf{x},t)\in B((\mathbf{z},\tau),r)$ and for all other $(\mathbf{x},t)\in\partial B((\mathbf{z},\tau),r)$ then $\mathbf{x}_0=\mathbf{z}$. The conclusion remains unchanged if the condition that $u(\mathbf{x},t)<M$ for all other $(\mathbf{x},t)\in\partial B((\mathbf{z},\tau),r)$ is dropped.*

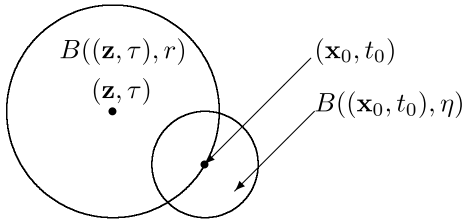
Proof: Suppose this is not true. That is, suppose $(\mathbf{x}_0,t_0)\in\partial B((\mathbf{z},\tau),r)$ and $u(\mathbf{x}_0,t_0)=M$ but for every other point in $\overline{B((\mathbf{z},\tau),r)}$, $u<M$ and yet $\mathbf{x}_0\neq\mathbf{z}$. Define the function,

$$v(\mathbf{x},t)\equiv\exp\left(-\alpha|\mathbf{x}-\mathbf{z}|^2-\alpha|t-\tau|^2\right)-\exp\left(-\alpha r^2\right)$$

Thus $v>0$ in $B((\mathbf{z},\tau),r)$, equal to zero on $\partial B((\mathbf{z},\tau),r)$, and less than 0 off $B((\mathbf{z},\tau),r)$. Now let η be small enough that

$$\overline{B((\mathbf{x}_0,t_0),\eta)}\subseteq U_T.$$

The following picture is representative of the two balls.



Then

$$Lv-v_t=\exp\left(-\alpha|\mathbf{x}-\mathbf{z}|^2-\alpha|t-\tau|^2\right)\cdot$$

$$\left(\sum_{ij}4\alpha^2a_{ij}(x_i-z_i)(x_j-z_j)-2\alpha\sum_i a_{ii}-2\alpha\sum_i b_i(x_i-z_i)+2\alpha(t-\tau)\right)$$

$$\geq \exp\left(-\alpha |\mathbf{x}-\mathbf{z}|^2 - \alpha |t-\tau|^2\right)\left[4\alpha^2\delta^2 |\mathbf{x}-\mathbf{z}|^2 - 2\alpha \operatorname{trace}((a_{ij})) - 2\alpha \sum_i b_i (x_i - z_i) + 2\alpha (t-\tau)\right]$$

Choosing η small enough, $|\mathbf{x}-\mathbf{z}|$ is bounded away from 0 for $\mathbf{x} \in B((\mathbf{x}_0, t_0), \eta)$ and so the term, $4\alpha^2\delta^2 |\mathbf{x}-\mathbf{z}|^2$ dominates all the others in $[\cdot]$ if α is chosen very large. Therefore, for all large enough α,

$$Lv - v_t > 0 \text{ on } B((\mathbf{x}_0, t_0), \eta).$$

Now $u - M < 0$ on $\partial B((\mathbf{x}_0, t_0), \eta) \cap \overline{B((\mathbf{z}, \tau), r)}$ which is a compact set and so $u - M$ is negative and bounded away from 0 on this set. Therefore, if σ is a sufficiently small positive number,

$$u - M + \sigma v < 0$$

on $\partial B((\mathbf{x}_0, t_0), \eta) \cap \overline{B((\mathbf{z}, \tau), r)}$. On $\partial B((\mathbf{x}_0, t_0), \eta) \cap B((\mathbf{z}, \tau), r)^C$, $v < 0$ and $u - M \leq 0$ so the above function is negative on all of $\partial B((\mathbf{x}_0, t_0), \eta)$. Since $\partial B((\mathbf{x}_0, t_0), \eta)$ is a compact set, the above function is negative and bounded away from zero on $\partial B((\mathbf{x}_0, t_0), \eta)$. Also,

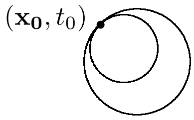
$$L(u - M + \sigma v) - (u - M + \sigma v)_t = Lu + \sigma Lv - u_t - \sigma v_t > 0.$$

But now Lemma 6.8 applies to $B((\mathbf{x}_0, t_0), \eta)$ and it follows that for all $(\mathbf{x}, t) \in B((\mathbf{x}_0, t_0), \eta)$,

$$(u - M + \sigma v)(\mathbf{x}, t) < 0$$

contrary to the fact that $(u - M + \sigma v)(\mathbf{x}_0, t_0) = 0$.

It only remains to verify the claim that the conclusion of the lemma still holds if the condition that $u(\mathbf{x}, t) < M$ for all other $(\mathbf{x}, t) \in \partial B((\mathbf{z}, \tau), r)$ is dropped. Drop this condition then and note that if $\mathbf{z} \neq \mathbf{x}_0$, the same is true for $\mathbf{x}_0$ and the center of the smaller ball shown in the following picture.



This smaller ball satisfies the condition which was dropped and so a contradiction is obtained. This proves the lemma.

The next lemma is where the connectedness of U is used. Up till now, this has not been required.

Lemma 6.12 *Suppose* $u(\mathbf{x}_0, t_0) < M$ *where* $\mathbf{x}_0 \in U$ *and* $t_0 \in (0, T)$. *Then* $u(\mathbf{x}, t_0) < M$ *for all* $\mathbf{x} \in U$.

Proof: Suppose not. Then letting $H \equiv \{\mathbf{x} \in U : u(\mathbf{x}, t_0) = M\}$, it follows that H is nonempty. Since u is continuous, it follows that H is closed. Therefore, since U is connected, H must not be open because if it were $U = H \cup (U \setminus H)$, a disjoint union of relatively open sets. Thus there exists $\mathbf{x}_1 \in H$ such that $\mathbf{x}_1$ is not an interior point of H. Consider rays from $\mathbf{x}_1$ of the form $\mathbf{x}_1 + s\mathbf{v}$ where

$$0 < |\mathbf{v}| < \operatorname{dist}\left(\mathbf{x}_1, U^C\right) \equiv \inf\left\{|\mathbf{x}_1 - \mathbf{z}| : \mathbf{z} \in U^C\right\}.$$

I claim that for some $\mathbf{v}$, the ray just described has the property that for $s \in (0, 1]$, $u(\mathbf{x}_1 + s\mathbf{v}, t_0) < M$. If this were not so, then the points of the form $\mathbf{x}_1 + s\mathbf{v}$ for $s \in (0, 1]$ and $0 < |\mathbf{v}| < \operatorname{dist}\left(\mathbf{x}_1, U^C\right)$ along with the single point, $\mathbf{x}_1$ would yield an open set containing $\mathbf{x}_1$ which is contained in H, contrary to the assumption that $\mathbf{x}_1$ is not an interior point. Pick such a vector, $\mathbf{v}$ and let $\mathbf{x}_2 = \mathbf{x}_1 + \mathbf{v}$. Thus, from the construction, the line segment joining $\mathbf{x}_1$ to $\mathbf{x}_2$ is contained in U and every point of this line segment is not in H except $\mathbf{x}_1$ which is in H.

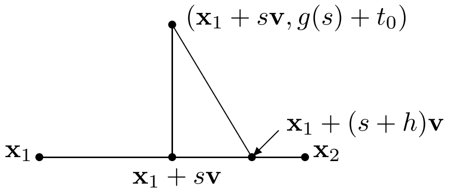
Let $H_1 \equiv \left\{(\mathbf{x}, t) \in \overline{U_T} : u(\mathbf{x}, t) = M\right\}$ and define

$$d(\mathbf{x}) \equiv \operatorname{dist}((\mathbf{x}, t_0), H_1) \equiv \inf\{|(\mathbf{x}, t_0) - \mathbf{h}| : \mathbf{h} \in H_1\}.$$

Let $\mathbf{v} \equiv \mathbf{x}_2 - \mathbf{x}_1$ so that $\mathbf{x}_1 + s\mathbf{v}$, $s \in [0,1]$ is a parametrization of the line segment joining the points $\mathbf{x}_1$ and $\mathbf{x}_2$. Let $g(s) \equiv d(\mathbf{x}_1 + s\mathbf{v})$. From Lemma 6.11 the point of H_1 which is closest to $\mathbf{x}_1 + s\mathbf{v}$ is

$$(\mathbf{x}_1 + s\mathbf{v}, t_0 \pm d(\mathbf{x}_1 + s\mathbf{v})) = (\mathbf{x}_1 + s\mathbf{v}, t_0 \pm g(s)).$$

Consider the following picture.



In this picture the point at the top represents the closest point of H_1 to $\mathbf{x}_1 + s\mathbf{v}$ and so the point of H_1 which is closest to $\mathbf{x}_1 + (s+h)\mathbf{v}$ is at least as close as this one.

It follows

$$g(s+h) \leq \sqrt{h^2 |\mathbf{v}|^2 + g(s)^2}.$$

by similar reasoning,

$$g(s) \leq \sqrt{h^2 |\mathbf{v}|^2 + g(s+h)^2}.$$

Therefore, for $h > 0$,

$$\frac{g(s+h) - \sqrt{h^2 |\mathbf{v}|^2 + g(s+h)^2}}{h} \leq \frac{g(s+h) - g(s)}{h} \leq \frac{g(s) - \sqrt{h^2 |\mathbf{v}|^2 + g(s)^2}}{h}$$

and for $h < 0$, the inequalities just get turned around. Because of the construction above, $g(s) > 0$ whenever $s > 0$. Now if $a > 0$, it is routine to verify that

$$\left| \frac{a - \sqrt{h^2 |\mathbf{v}|^2 + a^2}}{h} \right| \leq \frac{|h| |\mathbf{v}|}{a + \sqrt{h^2 |\mathbf{v}|^2 + a^2}}$$

and so the limit as $h \to 0$ of this expression equals zero. By the squeezing theorem of calculus, it follows that for all $s \in (0,1)$, $g'(s) = 0$. Therefore, by the mean value theorem, g must be a constant which contradicts the fact that $g(0) = M$ and $g(1) < M$. This proves the lemma.

With these lemmas, it is now not hard to prove the strong maximum principle.

Proof of Theorem 6.9: I have shown that for $t \in (0,T)$, either $u(\mathbf{x},t) = M$ for all $\mathbf{x} \in U$ or $u(\mathbf{x},t) < M$ for all $\mathbf{x} \in U$ (Lemma 6.12). Suppose then that $u(\mathbf{x}_0, t_0) = M$ for some $\mathbf{x}_0 \in U$ and $t_0 \in (0,T]$. Let

$$G \equiv \{t \in (0,t_0) : u(\mathbf{x},t) < M \text{ for all } \mathbf{x} \in U\}.$$

Then G is open because of continuity of u. If $u(\mathbf{x},t) < M$, this will be true for t' near t and so by Lemma 6.12, $t' \in G$. Therefore, $G = \cup_{i=1}^{\infty} (a_i, b_i)$ where the open intervals are connected components of G. If G is nonempty, then some of these are nonempty. But by Lemma 6.10, if (a_i, b_i) is one of these which is nonempty, then $b_i \in G$ as well which means that (a_i, b_i) was not really a connected component. I could get a larger connected open interval contained in G which is of the form $(a_i, b_i + \delta)$ for small enough δ. It must be the case that G is empty and so $u(\mathbf{x},t) = M$ for all $\mathbf{x} \in U$ and $t \leq t_0$. This proves the theorem.

Bibliography

[1] **Adams R.** *Sobolev Spaces*, Academic Press, New York, San Francisco, London, 1975.

[2] **Apostol, T. M.,** *Mathematical Analysis,* Addison Wesley Publishing Co., 1969.

[3] **Apostol, T. M.,** *Calculus second edition,* Wiley, 1967.

[4] **Bergh J. and Löfström J**. *Interpolation Spaces,* Springer Verlag 1976.

[5] **Dunford N.** and **Schwartz J.T.** *Linear Operators,* Interscience Publishers, a division of John Wiley and Sons, New York, part 1 1958, part 2 1963, part 3 1971.

[6] **Evans L.C.** and **Gariepy,** *Measure Theory and Fine Properties of Functions,* CRC Press, 1992.

[7] **Evans L.C.** *Partial Differential Equations,* Berkeley Mathematics Lecture Notes. 1993.

[8] **Federer H.,** *Geometric Measure Theory,* Springer-Verlag, New York, 1969.

[9] **Gagliardo, E.**, Properieta di alcune classi di funzioni in piu variabili, *Ricerche Mat.* 7 (1958), 102-137.

[10] **Grisvard, P.** *Elliptic problems in nonsmooth domains,* Pittman 1985.

[11] **Hewitt E.** and **Stromberg K.** *Real and Abstract Analysis,* Springer-Verlag, New York, 1965.

[12] **Hörmander, Lars** *Linear Partial Differential Operators,* Springer Verlag, *1976.*

[13] **John, Fritz,** *Partial Differential Equations,* Fourth edition, Springer Verlag, 1982.

[14] **Kuttler K.L.,** *Modern Analysis* CRC Press 1998.

[15] **Protter, M. and Weinberger, H.**, *Maximum Principles in Differential Equations*, Springer Verlag, New Youk, 1984.

[16] **Rudin W.** *Real and Complex Analysis,* third edition, McGraw-Hill, 1987.

[17] **Rudin W.** *Functional Analysis,* second edition, McGraw-Hill, 1991.

[18] **Smart D.R.** *Fixed point theorems* Cambridge University Press, 1974.

[19] **Smoller Joel,** Shock Waves and Reaction Diffusion Equations, Springer, 1980.

[20] **Stein E.** *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, Princeton, N. J., 1970.

[21] **Varga R. S.** *Functional Analysis and Approximation Theory in Numerical Analysis,* SIAM 1984.

[22] **Yosida K.** *Functional Analysis,* Springer-Verlag, New York, 1978.

Index