Introduction to Real Analysis
Spring 2014 Lecture Notes

Vern I. Paulsen

April 22, 2014

Contents

Chapter 1

Sequences and Series of Functions

In this chapter we introduce different notions of convergence for sequence and series of functions and then examine how integrals and derivatives behave upon taking limits of functions in these various senses. We then apply these results to power series and Fourier series.

Definition 1.1. *Given a set* X, *a metric space* (Y, ρ), *functions* $f_n : X \to Y$, $n \in \mathbb{N}$ *and* $f : X \to Y$, *we say that the sequence of functions* $\{f_n\}$ **converges pointwise to** f *provided that for each* $x \in X$, *the sequence of points* $\{f_n(x)\}$ *converges to the point* $f(x)$ *in the metric of* Y. *That is, provided that* $\lim_n \rho(f_n(x), f(x)) = 0$. *When this occurs we write,* $f_n \xrightarrow{ptw} f$.

Note that the statement that $\{f_n\}$ converges pointwise to f is equivalent to the requirement that for each $\epsilon > 0$ and $x \in X$, there is a N_x so that when $n > N_x$ we have that $\rho(f_n(x), f(x)) < \epsilon$. That is, since pointwise convergence only requires convergence at each point in X, the value that we take for N could depend on the individual point x as well as on ϵ. When the value for N can be picked depending only on ϵ and independent of the point x, then we call the convergence *uniform.* The precise definition follows.

Definition 1.2. *Given a set* X, *a metric space* (Y, ρ), *functions* $f_n : X \to Y$, $n \in \mathbb{N}$ *and* $f : X \to Y$, *we say that the sequence of functions* $\{f_n\}$ **converges uniformly to** f *provided that for each* $\epsilon > 0$, *there is* N, *so that when* $n > N$, *then for every* $x \in X$, *we have* $\rho(f_n(x), f(x)) < \epsilon$. *When this occurs we write,* $f_n \xrightarrow{u} f$.

Note that uniform convergence always implies pointwise convergence, since it is the stronger condition that N is independent of the point x.

Just as pointwise convergence is the requirement that $\lim_n \rho(f_n(x), f(x)) = 0$, uniform convergence is equivalent to the requirement that the sequence of numbers $s_n = \sup\{\rho(f_n(x), f(x)) : x \in X\}$ converges to 0. That is, provided that

$$\lim_n [\sup\{\rho(f_n(x), f(x)) : x \in X\}] = 0.$$

We prove this below.

Proposition 1.3. *Let X be a set, let (Y, ρ) a metric space, let $f_n : X \to Y$, $n \in \mathbb{N}$ and $f : X \to Y$ be functions and set*

$$s_n = \sup\{\rho(f_n(x), f(x)) : x \in X\}.$$

Then $f_n \xrightarrow{u} f$ if and only if $\lim_{n\to\infty} s_n = 0$.

Proof. First assume that $f_n \xrightarrow{u} f$. Let $\epsilon > 0$ be given. By the definition of uniform convergence, there is a N so that when $n > N$, for every $x \in X$, we have $\rho(f_n(x), f(x)) < \epsilon/2$. But this implies that for $n > N$, we have

$$s_n = \sup\{\rho(f_n(x), f(x)) : x \in X\} \leq \epsilon/2 < \epsilon.$$

Thus, for $n > N$ we have that $|s_n - 0| < \epsilon$ and since ϵ was arbitrary, this proves that $\lim_{n\to\infty} s_n = 0$.

Conversely, assume that $\lim_{n\to\infty} s_n = 0$. Then given $\epsilon > 0$, there is a N so that $n > N$ implies that $0 \leq s_n < \epsilon$ whihc implies that for every $n > N$ and for every $x \in X$, we have $\rho(f_n(x), f(x)) < \epsilon$. This proves that $f_n \xrightarrow{u} f$. □

Note that for both of these definitions, we did not need for X to be a metric space, although in many of the interesting examples, we will also have X a metric space.

Also note that if $f_n \xrightarrow{\text{u}} f$, then $f_n \xrightarrow{\text{ptw}} f$.

Example 1.4. *Let $X = Y = [0, 1]$ be endowed with the usual metric, let $f_n(x) = x^n$ and let $f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$. Then it is easy to see that $f_n \xrightarrow{ptw} f$. To decide if the convergence is also uniform, we compute*

$$s_n = \sup\{\rho(f_n(x), f(x)) : 0 \leq x \leq 1\} = \sup\{|x^n - f(x)| : 0 \leq x \leq 1\} = \sup\{x^n : 0 \leq x < 1\} = 1.$$

Since, $\lim_n \sup\{\rho(f_n(x), f(x)) : 0 \leq x \leq 1\} = 1 \neq 0$, *we conclude that* $\{f_n\}$ *does not converge uniformly to f.*

Thus although uniform convergence implies pointwise convergence, we have that pointwise convergence does not imply uniform convergence. For this reason we say that uniform convergence is a **stronger** convergence than pointwise convergence.

In this example, each of the functions f_n is continuous, but f is clearly not continuous. Thus, this example shows that the pointwise limit of continuous functions need **not** be a continuous function. Later we will see that uniform limits of continuous functions are again continuous.

Example 1.5. *This is a slight modification of the first example. Let* $X = [0, B], Y = [0, 1]$ *with* $0 < B < 1$, *let* $f_n(x) = x^n$ *and let* $f(x) = 0$. *Now*

$$s_n = \sup\{|f_n(x) - f(x)| : x \in X\} = \sup\{x^n : 0 \le x \le B\} = B^n.$$

Since, $\lim_{n\to\infty} B^n = 0$, *we have that* $f_n \xrightarrow{u} f$.

Example 1.6. *Let* $X = Y = \mathbb{R}$ *with the usual metric, let* $f_n(x) = \frac{x}{n}$ *and let* $f(x) = 0$. *Then* $f_n \xrightarrow{ptw} f$ *but* $\{f_n\}$ *does not converge to* f *uniformly.*

Example 1.7. *Let* $X = Y = [-\pi, +\pi]$ *with the usual metric, let* $f_n(x) = \frac{sin(nx)}{n}$ *and let* $f(x) = 0$. *Since* $\sup\{|f_n(x) - f(x)| : -\pi \le x \le +\pi\} = \frac{1}{n}$ *we have that* $f_n \xrightarrow{u} f$, *but* $f_n'(0) = 1$, *while* $f'(0) = 0$.

Thus, even uniform convergence of functions does not guarantee convergence of their derivatives.

Problem 1.8. *Let* $X = Y = [0, 1]$ *with the usual metric, let* $f_n : X \to Y$ *be defined by* $f_n(x) = x^{1/n}$ *and let*

$$f(x) = \begin{cases} 0 & \textit{when } x = 0 \\ 1 & \textit{when } x \neq 0 \end{cases}.$$

Prove that $f_n \xrightarrow{ptw} f$ *and that* $\{f_n\}$ *does not converge to* f *uniformly.*

Problem 1.9. *Let* $X \subset \mathbb{R}$ *be compact, let* $f_n : X \to \mathbb{R}$ *be defined by* $f_n(x) = \frac{x}{n}$ *and let* $f(x) = 0$ *be the 0 function. Prove that* $f_n \xrightarrow{u} f$.

Problem 1.10. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be continuous at* 0. *Set* $g_n(x) = f(\frac{x}{n})$ *and let* g *be the constant function that is equal to* $f(0)$.

1. *Prove that* $g_n \xrightarrow{ptw} g$ *as functions on* $\mathbb{R}$.
2. *Give an example of a function* f *such that the convergence is not uniform as functions on* $\mathbb{R}$.
3. *Prove that if* $X \subseteq \mathbb{R}$ *is any compact subset of* $\mathbb{R}$, *then* $g_n \xrightarrow{u} g$ *as functions on* X.

1.1 Behavior of Riemann Integrals with Limits

We now look at how Riemann integration behaves under pointwise and uniform convergence.

Example 1.11. *Let* $X = Y = [0,1]$ *with the usual metric. Recall that the rationals are countable and let* $\{r_n\}_{n\in\mathbb{N}}$ *be an enumeration of the rational numbers in* $[0,1]$*. We set*

$$f_n(x) = \begin{cases} 1 & x = r_k, 1 \le k \le n \\ 0 & \textit{otherwise} \end{cases} \quad \textit{and } f(x) = \begin{cases} 1 & x \textit{ rational} \\ 0 & x \textit{ irrational} \end{cases}.$$

It is easily checked that $f_n \xrightarrow{ptw} f$*, but for every* n*,* $\sup\{|f_n(x) - f(x)| : 0 \le x \le 1\} = 1$ *and so the sequence does not converge uniformly to* f*. Now verify that each function* f_n *is Riemann integrable on* $[0,1]$ *with* $\int_0^1 f_n(x)dx = 0$*, but* f *is not Riemann integrable.*

Thus, we see that if $\{f_n\}$ is a sequence of Riemann integrable fuctions that converges pointwise to a function f, then f might not even be Riemann integrable! In this case there is no chance that the limit of the integrals is the integral of the limit.

The next example shows that even when the limit function is Riemann integrable, pointwise convergence is not enough to guarantee that the limit of the integrals is the integral of the limit.

Example 1.12. *Let* $X = Y = [0,1]$*, set*

$$f_n(x) = \begin{cases} 2n^2x & 0 \le x \le \frac{1}{2n} \\ 2n - 2n^2x & \frac{1}{2n} < x \le \frac{1}{n} \\ 0 & \frac{1}{n} < x \le 1 \end{cases}$$

and let $f(x) = 0$*. Then* $f_n \xrightarrow{ptw} f$*, all of these functions are Riemann integrable on* $[0,1]$*, but* $\int_0^1 f_n(x)dx = 1/2$ *for all* n*, while* $\int_0^1 f(x)dx = 0$*.*

These last two examples show that the formula,

$$\lim_n \int_a^b f_n(x)dx = \int_a^b \lim_n f_n(x)dx,$$

is **not valid** when the limit of the functions is meant in the pointwise sense.

We now prove that all of these problems "go away" when one considers uniform convergence instead.

Lemma 1.13. *Let $\alpha : [a,b] \to \mathbb{R}$ be an increasing function and let $f, g : [a,b] \to \mathbb{R}$ be two bounded functions. If $\delta = \sup\{|f(x) - g(x)| : a \leq x \leq b\}$, then for any partition $\mathcal{P} = \{x_0, ..., x_n\}$ of $[a,b]$ e have that $U(f,\mathcal{P},\alpha) \leq U(g,\mathcal{P},\alpha) + \delta(\alpha(b) - \alpha(a))$ and $L(f,\mathcal{P},\alpha) \geq L(g,\mathcal{P},\alpha) - \delta(\alpha(b) - \alpha(a))$.*

Proof. For each x, we have that $g(x) - \delta \leq f(x) \leq g(x) + \delta$. Hence, for each subinterval of the partition we have that $m_i(g) - \delta = \inf\{g(x) : x_{i-1} \leq x \leq x_i\} - \delta \leq \inf\{f(x) : x_{i-1} \leq x \leq x_i\} = m_i(f)$ and $M_i(f) = \sup\{f(x) : x_{i-1} \leq x \leq x_i\} \leq \sup\{g(x) : x_{i-1} \leq x \leq x_i\} + \delta = M_i(g) + \delta$.

Thus, $U(f,\mathcal{P},\alpha) = \sum_{i=1}^n M_i(f)(\alpha(x_i) - \alpha(x_{i-1})) \leq \sum_{i=1}^n (M_i(g)+\delta)(\alpha(x_i) - \alpha(x_{i-1})) = U(g,\mathcal{P},\alpha) + \delta(\alpha(b) - \alpha(a))$.

The proof of the other inequality is similar. □

Theorem 1.14. *Let $\alpha : [a,b] \to \mathbb{R}$ be an increasing function on $[a,b]$ and let $\{f_n\}$ be a sequence of functions that are Riemann-Stieltjes integrable on $[a,b]$ with respect to α and converge uniformly to a function f on $[a,b]$. Then f is Riemann-Stieltjes integrable on $[a,b]$ with respect to α and*

$$\lim_n \int_a^b f_n d\alpha = \int_a^b f d\alpha.$$

Proof. We first prove that f satisfies the Riemann-Stieltjes integrability criterion on $[a,b]$. The case that $\alpha(a) = \alpha(b)$ is trivial, so we assume that $\alpha(b) - \alpha(a) > 0$.

Given $\epsilon > 0$, set $\delta = \frac{\epsilon}{4(\alpha(b)-\alpha(a))}$ and choose an integer N so that for $n \geq N$ we have that $\sup\{|f(x) - f_n(x)| : a \leq x \leq b\} < \delta$.

Since f_N is Riemann-Stieltjes integrable, we may choose a partition $\mathcal{P}$, so that $U(f_N,\mathcal{P},\alpha) - L(f_N,\mathcal{P},\alpha) < \frac{\epsilon}{2}$. Applying the lemma, we have that

$$\begin{aligned} &U(f,\mathcal{P},\alpha) - L(f,\mathcal{P},\alpha) \leq \\ &\quad [U(f_N,\mathcal{P},\alpha) + \delta(\alpha(b) - \alpha(a))] - [L(f_N,\mathcal{P},\alpha) - \delta(\alpha(b) - \alpha(a))] = \\ &\qquad [U(f_N,\mathcal{P},\alpha) - L(f_N,\mathcal{P},\alpha)] + 2\delta(\alpha(b) - \alpha(a)) < \frac{\epsilon}{2} + \frac{\epsilon}{2}. \end{aligned}$$

Thus, f satisfies the Riemann-Stieltjes integrability criterion.

Also, for any $n \geq N$, applying the lemma again, we have that

$$\begin{aligned} &\int_a^b f d\alpha = \inf\{U(f,\mathcal{P},\alpha) : \mathcal{P}\} \leq \\ &\qquad \inf\{U(f_n,\mathcal{P},\alpha) + \delta(\alpha(b) - \alpha(a)) : \mathcal{P}\} = \int_a^b f_n d\alpha + \frac{\epsilon}{4}, \end{aligned}$$

while

$$\int_a^b f d\alpha = \sup\{L(f, \mathcal{P}, \alpha) : \mathcal{P}\} \geq$$

$$\sup\{L(f_n, \mathcal{P}, \alpha) - \delta(\alpha(b) - \alpha(a)) : \mathcal{P}\} = \int_a^b f_n d\alpha - \frac{\epsilon}{4}.$$

These two inequalities show that for $n \geq N$, we have that

$$|\int_a^b f d\alpha - \int_a^b f_n d\alpha| \leq \frac{\epsilon}{4} < \epsilon.$$

Since ϵ was arbitrary, we have that

$$\lim_n \int_a^b f_n d\alpha = \int_a^b f d\alpha.$$

□

Problem 1.15. *Prove that the functions $\{f_n\}$ in Example 1.7 do converge pointwise to f, prove that each f_n is Riemann integrable with Riemann integral equal to 0 and prove that f is not Riemann integrable.*

Problem 1.16. *Prove that the functions $\{f_n\}$ of Example 1.8 do converge pointwise to f and prove that each f_n is Riemann integrable with integral equal to 1/2.*

1.2 Uniform Convergence and Continuity

We have already seen that pointwise limits of continuous functions need not be continuous. In this section we will show that uniform limits preserve continuity. Some of these results have been seen in a different form in Chapter I.4.

Theorem 1.17. *Let (X, d) and (Y, ρ) be metric spaces, let $f_n : X \to Y, n \in \mathbb{N}$ be a sequence of functions from X to Y that converge uniformly to a function $f : X \to Y$ and let $x_0 \in X$. If f_n is continuous at x_0 for all $n \in \mathbb{N}$, then f is continuous at x_0.*

Proof. Given $\epsilon > 0$, we may pick N so that for $n > N$, we have that $\sup\{\rho(f_n(x), f(x)) : x \in X\} < \epsilon/3$. Fix any $n > N$, and using the fact that f_n is continuous at x_0, we may pick $\delta > 0$, so that $d(x, x_0) < \delta$ implies that $\rho(f_n(x), f_n(x_0)) < \epsilon/3$.

Hence, for $d(x, x_0) < \delta$, by applying the triangle inequality twice, we have that

$$\rho(f(x), f(x_0)) \leq \rho(f(x), f_n(x)) + \rho(f_n(x), f(x_0)) \leq \\ \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f(x_0)) < \epsilon/3 + \epsilon/3 + \epsilon/3,$$

and we conclude that f is continuous at x_0. □

Corollary 1.18. *Let (X, d) and (Y, ρ) be metric spaces, let $f_n : X \to Y, n \in \mathbb{N}$ be a sequence of functions from X to Y that converge uniformly to a function $f : X \to Y$. If f_n is continuous for all $n \in \mathbb{N}$, then f is continuous.*

We now consider the set of all continuous functions between two metric spaces and show that when the domain space is compact, then this set can be endowed with a metric.

Lemma 1.19. *Let (X, d) and (Y, ρ) be metric spaces and let $f, g : X \to Y$, be continuous functions. Then the function $h : X \to \mathbb{R}$ defined by $h(x) = \rho(f(x), g(x))$ is continuous.*

Proof. Fix a point $x_0 \in X$, and we will show that h is continuous at x_0. Given $\epsilon > 0$, since f and g are continuous at x_0, there exists $\delta_1 > 0$ and $\delta_2 > 0$, so that $d(x, x_0) < \delta_1$ implies that $\rho(f(x), f(x_0)) < \epsilon/2$, while $d(x, x_0) < \delta_2$ implies that $\rho(g(x), g(x_0)) < \epsilon/2$. Let $\delta = \min\{\delta_1, \delta_2\}$, then for $d(x, x_0) < \delta$, and using the reverse triangle inequality we have that

$$|h(x) - h(x_0)| = |\rho(f(x), g(x)) - \rho(f(x_0), g(x_0)| \leq \\ |\rho(f(x), g(x)) - \rho(f(x_0), g(x))| + |\rho(f(x_0), g(x)) - \rho(f(x_0), g(x_0))| \leq \\ \rho(f(x), f(x_0)) + \rho(g(x), g(x_0)) < \epsilon/2 + \epsilon/2.$$

Since $\epsilon > 0$ was arbitrary, this proves that h is continuous at x_0. □

Definition 1.20. *Given two metric spaces (X, d) and (Y, ρ), we let $C(X; Y)$ denote the set of all continuous functions from X to Y. Given $f, g \in C(X; Y)$ we set*

$$\gamma(f, g) = \sup\{\rho(f(x), g(x)) : x \in X\}.$$

When this supremum is unbounded, we set $\gamma(f, g) = +\infty$.

Theorem 1.21. *Let (X, d) and (Y, ρ) be metric spaces with X compact. Then*

1. *γ defines a metric on $C(X; Y)$,*

2. *a sequence* $\{f_n\} \subseteq C(X;Y)$ *converges to* $f \in C(X;Y)$ *in the metric* γ *if and only if* $f_n \overset{u}{\to} f$,

3. *if* (Y, ρ) *is complete, then* $(C(X;Y), \gamma)$ *is also a complete metric space.*

Proof. Given $f, g \in C(X;Y)$, by the above lemma the function $x \to \rho(f(x), g(x))$ is a continuous function from the compact metric space X to $\mathbb{R}$. Hence, it is bounded and so $\gamma(f,g) \geq 0$ is a finite real number. Note that $\gamma(f,g) = 0$ if and only if $\rho(f(x), g(x)) = 0$ for every $x \in X$ which is if and only if $f(x) = g(x)$ for every $x \in X$.

Clearly, $\gamma(f,g) = \gamma(g,f)$. So all that remains to show that γ is a metric is to verify the triangle inequality. To this end let $f, g, h \in C(X;Y)$. Then

$$\gamma(f,g) = \sup\{\rho(f(x), g(x)) : x \in X\} \leq \\ \sup\{\rho(f(x), h(x)) + \rho(h(x), g(x)) : x \in X\} \leq \gamma(f,h) + \gamma(h,g).$$

Thus, γ is a metric and 1) is proven.

Next we prove the second statement. To see this we refer back to Proposition 1.3. Note that the s_n of that proposition is $s_n = \rho(f_n, f)$. Thus, $\{f_n\}$ converges to f in the metric ρ iff $\lim_{n\to\infty} \rho(f_n, f) = 0$ iff $\lim_{n\to\infty} s_n = 0$. By Proposition 1.3, this last statement is equivalent to requiring that $f_n \overset{u}{\to} f$.

To prove the third statement, we let $\{f_n\} \subseteq C(X;Y)$ be a sequence that is Cauchy in the metric γ, and we must prove that there exists a function $f \in C(X;Y)$, so that $\{f_n\}$ converges to f in the metric γ.

Since for each fixed $x \in X$, we have that $\rho(f_n(x), f_m(x)) \leq \gamma(f_n, f_m)$, we see that the sequence of points in Y, given by $\{f_n(x)\}$ is also Cauchy. Since (Y, ρ) is complete, there is a point that this sequence converges to and we set $f(x) = \lim_n f_n(x)$. Since $x \in X$, was an arbitrary point, we see that we can define a function, $f : X \to Y$, by setting for each point $x \in X$, $f(x) = \lim_n f_n(x)$.

If we can prove that the sequence $\{f_n\}$ converges uniformly to f, then by the above theorem, f will be a continuous function from X to Y and by 2) we will have that $\{f_n\}$ converges to f in the metric γ, so our proof will be complete.

Given $\epsilon > 0$, since our original sequence of functions is Cauchy, we may pick N, so that for any $n, m > N$, we have that $\gamma(f_n, f_m) < \epsilon/2$. Now fix $x_0 \in X$, and $n > N$, then we have that

$$\rho(f(x_0), f_n(x_0)) = \lim_m \rho(f_m(x_0), f_n(x_0)) \leq \epsilon/2.$$

Since x_0 was arbitrary, we have that

$$\sup\{\rho(f(x), f_n(x)) : x \in X\} \leq \epsilon/2 < \epsilon,$$

for every $n > N$. This proves that the convergence is uniform. Thus, $f \in C(X;Y)$ and we see that this last inequality also shows that $\gamma(f, f_n) < \epsilon$ for $n > N$. Thus the sequence $\{f_n\}$ converges to the continuous function f in the metric γ and our proof is complete. □

When $Y = \mathbb{R}$ with the usual metric, then we simplify notation by setting $C(X) = C(X;\mathbb{R})$. We can now see that Theorem I.4.7 is a special case of the above theorem.

When $Y = \mathbb{R}^k$ with the usual Euclidean metric, then we can think of a function in $\vec{f} \in C(X;\mathbb{R}^k)$ as a vector-valued function $\vec{f}(x) = (f_1(x), ..., f_k(x))$ where each $f_i : X \to \mathbb{R}$. Using Theorem I.3.18, we see that $\vec{f}$ is continuous if and only if each $f_i \in C(X)$.

Problem 1.22. *Let $\alpha : [a,b] \to \mathbb{R}$ be an increasing function and let $\mathcal{R}([a,b],\alpha)$ denote the set of functions on [a,b] that are Riemann-Stieltjes integrable with respect to α. Given $f, g \in \mathcal{R}([a,b],\alpha)$ define $\gamma(f,g) = \sup\{|f(x)-g(x)| : a \leq x \leq b\}$. Prove that γ defines a metric on $\mathcal{R}([a,b],\alpha)$ and that $(\mathcal{R}([a,b],\alpha),\gamma)$ is a complete metric space.*

Problem 1.23. *Let $\vec{f_n} = (f_{1,n}, ..., f_{k,n}) \in C(X;\mathbb{R}^k)$ be a sequence of functions. Prove that $\vec{f_n}$ converges uniformly to $\vec{g} = (g_1, ..., g_k)$ if and only if for each $i, 1 \leq i \leq k$, we have that $f_{i,n}$ converges uniformly to g_i.*

Problem 1.24. *Let $f_n : [0,1] \to \mathbb{R}$ be defined by $f_n(x) = \frac{nx}{1+n^2x^2}$. Use standard results from calculus to prove that $f_n \xrightarrow{ptw} 0$, where 0 denotes the function that is identically equal to 0 and that $\lim_{n\to\infty} \int_0^1 f_n(x)dx = 0$. Prove or disprove that $f_n \xrightarrow{u} 0$.*

1.3 Uniform Convergence and Derivatives

In this section we look at one key theorem about the behaviour of derivatives under the taking of uniform limits. Given a function $f : (a,b) \to \mathbb{R}$ we will say that f is **differentiable on (a,b)** provided that it is differentiable at each point $x, a < x < b$.

Theorem 1.25. *Let $\{f_n\}$ be a sequence of differentiable functions on (a,b). If $f_n \xrightarrow{ptw} f$ and $f_n' \xrightarrow{u} g$ on (a,b) then f is differentiable on (a,b), $f' = g$, and $f_n \xrightarrow{u} f$.*

Proof. Fix $c, a < c < b$ and define

$$h(x) = \begin{cases} \frac{f(x)-f(c)}{x-c}, & x \neq c \\ g(c), & x = c \end{cases}.$$

If we can show that h is continuous at c, then that will show that $f'(c)$ exists and is equal to $g(c)$. If we let

$$h_n(x) = \begin{cases} \frac{f_n(x)-f_n(c)}{x-c}, & x \neq c \\ f_n'(c), & x = c \end{cases},$$

then since f_n is differentiable, h_n is continuous at c. We will prove that h is continuous at c by showing that the sequence $\{h_n\}$ converges uniformly to h.

Given $\epsilon > 0$, we may choose N_1 so that for $n > N_1$, we have that $\sup\{|f_n'(x) - g(x)| : a < x < b\} < \epsilon/3$. Then for any $m, n > N_1$, and any x we will have that $|f_m'(x) - f_n'(x)| \leq |f_m'(x) - g(x)| + |g(x) - f_n'(x)| < 2\epsilon/3$.

Thus, applying the Mean Value Theorem, for any $m, n > N_1$ and $x \neq c$ we will have a point x_1 between x and c so that $|h_m(x) - h_n(x)| = |\frac{(f_m(x)-f_n(x))-(f_m(c)-f_n(c))}{x-c}| = |(f_m - f_n)'(x_1)| < 2\epsilon/3$. Thus, for $x \neq c$, we have that $|h(x) - h_n(x)| = \lim_{m\to+\infty} |h_m(x) - h_n(x)| \leq 2\epsilon/3 < \epsilon$.

On the other hand, at the point c, we have that $h(c) = g(c) = \lim_n f_n'(c) = \lim_n h_n(c)$. So we may choose N_2 so that for $n > N_2$, we have that $|h(c) - h_n(c)| < \epsilon$. Setting $N = \max\{N_1, N_2\}$, we have that for $n > N$, and any $a < x < b$, $|h(x) - h_n(x)| < \epsilon$. This proves that the sequence $\{h_n\}$ converges uniformly to h, which completes the proof that $f' = g$.

To see that $f_n \xrightarrow{u} f$, given any $\delta > 0$, we may choose N_1 so that for $n > N_1$, $|h_n(x) - h(x)| < \delta$ for all $x, a < x < b$. Substituting in the definitions of these functions we see that

$$|(f_n(x) - f_n(c)) - (f(x) - f(c))| < \delta|x - c|,$$

and hence,

$$|f_n(x) - f(x)| < \delta|x - c| + |f_n(c) - f(c)|.$$

Now given $\epsilon > 0$, choose δ so that $\delta \max\{|b - c|, |c - a|\} = \epsilon/2$ and take the corresponding N_1. Also, since $f_n \xrightarrow{\text{ptw}} f$, we may choose N_2, so that for $n > N_2$, we have $|f_n(c) - f(c)| < \epsilon/2$.

Finally, taking $N_3 = \max\{N_1, N_2\}$, we have that for any $n > N_3$, and any $x, a < x < b$, we have that

$$|f_n(x) - f(x)| < \delta|x - c| + |f_n(c) - f(c)| < \epsilon/2 + \epsilon/2.$$

□

The following problem generalizes Theorem 1.25.

Problem 1.26. *Let $\{f_n\}$ be a sequence of differentiable functions on (a, b). If $\{f_n'\}$ converges uniformly to a function g on (a, b) and for some given point $x_0, a < x_0 < b$, the sequence of real numbers, $\{f_n(x_0)\}$ converges, then prove that there exists a differentiable function f on (a, b) such that $f_n \xrightarrow{u} f$ with $f' = g$. [Hint: Use the Mean Value Theorem to show $\{f_n(x)\}$ is Cauchy for any $a < x < b$.]*

1.4 Series of Functions

Recall that given real numbers, $\{a_k\}_{k\in\mathbb{N}}$ we can form the **infinite series**

$$\sum_{n=1}^{+\infty} a_k.$$

We say that this series *converges* to a real number s provided that: when we form the sequence of **partial sums,**

$$s_n = \sum_{k=1}^{n} a_k,$$

then $\lim_n s_n = s$.

When each $a_k \geq 0$, then the partial sums s_n are an increasing sequence of non-negative reals. Recall that in this case either the sequence $\{s_n\}$ is bounded above and $\lim_n s_n = \sup\{s_n : n \in \mathbb{N}\}$ or the sequence $\{s_n\}$ is not bounded above, in which case we write $\lim_n s_n = +\infty$ and we will also write $\sum_{k=1}^{+\infty} a_k = +\infty$.

Now suppose that we have a set X and functions, $f_k : X \to \mathbb{R}, k \in \mathbb{N}$. We wish to define convergence of a series of functions,

$$\sum_{k=1}^{+\infty} f_k.$$

To do this note that for each $x \in X$, we have real numbers $\{f_k(x)\}_{k\in\mathbb{N}}$. If for each $x \in X$ the series,

$$\sum_{k=1}^{+\infty} f_k(x)$$

converges, then we may define a function $s : X \to \mathbb{R}$ by setting

$$s(x) = \sum_{k=1}^{+\infty} f_k(x).$$

In the case we say that $\sum_{k=1}^{+\infty} f_k$ **converges** to the function s. Note that this is equivalent to saying that the sequence of partial sum functions,

$$s_n = \sum_{k=1}^{n} f_k,$$

converges pointwise to s. For this reason, we will sometimes, for emphasis, say that the series **converges pointwise** to s and write

$$s = ptw - \sum_{k=1}^{+\infty} f_k.$$

When the sequence $\{s_n\}$ of partial sum functions converges uniformly to s then we say that the series $\sum_{k=1}^{+\infty} f_k$ **converges uniformly** to s. To indicate this stronger convergence, we will write

$$s = u - \sum_{k=1}^{+\infty} f_k.$$

Thus, when we write $s = \sum_{k=1}^{+\infty} f_k$ we are only asserting that the partial sums converge pointwise. It is important to remember that the "default" notion of convergence for series of functions is pointwise convergence of the partial sums.

However, as we saw in the earlier sections, unifrom convergence has many important properties, so we will often be interested in when we have this stronger form of convergence.

Example 1.27. *Let $X = [0, 1]$, set $f_1(x) = x$ and for $n > 1$, set $f_n(x) = x^n - x^{n-1}$. Then the series "telescopes" and for $n > 1$, we have that the partial sums satisfy $s_n(x) = x^n$. Thus, by our earlier examples, we see that $s_n \xrightarrow{ptw} s$, where $s(x) = \begin{cases} 0 & 0 \le x < 1 \\ 1 & x = 1 \end{cases}$ and the convergence of these partial sums is* **not** *uniform. Thus, we have*

$$s = \sum_{k=1}^{+\infty} f_k,$$

but

$$s \neq u - \sum_{k=1}^{+\infty} f_k.$$

One of the best ways to determine if a series of functions converges uniformly is given by **Weierstrass' M-test.**

Theorem 1.28 (Weierstrass' M-test)**.** *Let X be a set, let $f_k : X \to \mathbb{R}$, $k \in \mathbb{N}$ be a sequence of functions and let $M_k \in \mathbb{R}$ be a sequence satisfying $\sup\{|f_k(x)| : x \in X\} \leq M_k$. If $\sum_{k=1}^{+\infty} M_k$ is finite, then there exists a function $s : X \to \mathbb{R}$ such that $s = u - \sum_{k=1}^{+\infty} f_k$.*

Proof. Set $A = \sum_{k=1}^{+\infty} M_k$. Since each $M_k \geq 0$, we have that the partial sums of the series $\sum_{k=1}^{+\infty} M_k$ are increasing to A. Thus, given $\epsilon > 0$, if we choose an integer N so that $0 \leq A - \sum_{k=1}^{N} M_k < \epsilon$, then $\sum_{k=N+1}^{+\infty} M_k < \epsilon$.

If we set $s_n(x) = \sum_{k=1}^{n} f_k(x)$, then for any $m > n > N$, we have that

$$|s_m(x) - s_n(x)| = |\sum_{k=n+1}^{m} f_k(x)| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k < \epsilon.$$

This inequality shows that, for each $x \in X$, the sequence of real numbers $\{s_n(x)\}$ is Cauchy and so has a limit, $s(x)$. Moreover, $s(x) = \sum_{k=1}^{+\infty} f_k(x)$.

Finally, since this inequality is independent of x, we have that

$$|s(x) - s_n(x)| = \lim_m |s_m(x) - s_n(x)| \leq \epsilon,$$

for any $n > N$. Since $\epsilon > 0$ was arbitrary and since N is independent of the point $x \in X$, this shows that the sequence of partial sums $\{s_n\}$ converges uniformly to s.

Thus, we have that $s = u - \sum_{k=1}^{+\infty} f_k$. ☐

Example 1.29. *Let $f_k : \mathbb{R} \to \mathbb{R}$ be defined by $f_k(x) = \frac{sin(kx)}{k^3}$. Since $|f_k(x)| \leq \frac{1}{k^3}$ and the series $\sum_{k=1}^{+\infty} \frac{1}{k^3}$ converges(by the integral test), by the Weierstrass M-test, we have that the partial sums*

$$s_n(x) = \sum_{k=1}^{n} f_k(x)$$

converge uniformly to the function $s(x) = \sum_{k=1}^{+\infty} f_k(x)$, i.e., $s = u - \sum_{k=1}^{+\infty} f_k$. Since each f_k is continuous, we have that s is continuous by Corollary 1.18.

Moreover, note that

$$s_n'(x) = \sum_{k=1}^{n} \frac{cos(kx)}{k^2}$$

and that $|\frac{cos(kx)}{k^2}| \leq \frac{1}{k^2}$ *which is another convergent series. Applying Weierstrass' M-test again, we have that* $s_n' \xrightarrow{u} g$ *where*

$$g(x) = u - \sum_{k=1}^{+\infty} \frac{cos(kx)}{k^2}.$$

Thus, g *is continuous, and by Theorem 1.25,* $s'(x) = g(x)$*, in other words,*

$$s' = u - \sum_{k=1}^{+\infty} f_k'$$

and we get that the "derivative of the sum is the sum of the derivatives" in this case.

Note that we run into difficulties if we try to differentiate again, we have no way yet to decide if the series,

$$\sum_{k=1}^{+\infty} f_k''(x) = \sum_{k=1}^{+\infty} \frac{-sin(kx)}{k}$$

even converges pointwise.

Problem 1.30. *Let* $f_n(x) = \frac{x}{1+n^3x^2}$*. Prove that the series* $\sum_{n=1}^{+\infty} f_n(x)$ *converges uniformly on* $\mathbb{R}$*.*

1.5 An Increasing Function with a Dense Set of Discontinuities

The Weierstrass M-test can be used to construct many functions with interesting and surprising properties. Here we construct a function that is increasing, continuous at every irrational number and discontinuous at every rational number.

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is said to have a **jump discontinuity** at c provided that both the one-sided limits as x approaches c exist, but the three numbers $f(c), \lim_{x\to c^-} f(x)$ and $\lim_{x\to c^+} f(x)$ do not all have the same value. A simple example of such a function is the "jump" function

$$J_c(x) = \begin{cases} 0 & x < c \\ 1 & c \le x \end{cases}.$$

Also, recall that the rational numbers are countable and let $\{r_n\}_{n\in\mathbb{N}}$ be an enumeration of the rationals. We now define a rather "bizarre" function as follows. Set

$$B(x) = \sum_{k=1}^{+\infty} \frac{1}{2^k} J_{r_k}(x).$$

Since the functions, $f_k(x) = \frac{1}{2^k} J_{r_k}(x)$ satisfy

$$\sup\{|f_k(x)| : x \in \mathbb{R}\} = \frac{1}{2^k}$$

and $\sum_{k=1}^{+\infty} \frac{1}{2^k} = 1$ by the Weierstrass M-test we see that $B = u - \sum_{k=1}^{+\infty} f_k$. We summarize some of the properties of the function B below. First, we will need a result many have probably seen.

Proposition 1.31. *Let $g : \mathbb{R} \to \mathbb{R}$ be an increasing function and let $c \in \mathbb{R}$. Then both of the one-sided limits exist at c with $\lim_{x\to c^-} g(x) = \sup\{g(x) : x < c\}$ and $\lim_{x\to c^+} g(x) = \inf\{g(x) : c < x\}$. Consequently, an increasing function can only have jump discontinuities.*

Proof. We only do the case of the limit from the left, the other case is similar. Since g is increasing, for $x < c$, we have $g(x) \le g(c)$. Thus, $g(c)$ is an upper bound for the set $\{g(x) : x < c\}$. This shows that the supremum exists and is not $+\infty$. Set $L^- = \sup\{g(x) : x < c\}$. Now given $\epsilon > 0$, $L^- - \epsilon$ is no longer an upper bound, so there exists an x_1 with $L^- - \epsilon < g(x_1) \le L^-$. Let $\delta = c - x_1$, then for any x with $c - \delta < x < c$, we have $x_1 < x < c$ and hence, $L^- - \epsilon < g(x_1) \le g(x) \le L^-$. Hence, $c - \delta < x < c$ implies that $|L^- - g(x)| < \epsilon$. This proves that $\lim_{x\to c^-} g(x) = L^-$. □

Proposition 1.32. *Let B be the function defined above. Then:*

1. *B is strictly increasing, that is, if $x < y$, then $B(x) < B(y)$,*
2. *B is continuous at every irrational number,*
3. *B has a jump discontinuity at every rational number with $\lim_{x\to r_n^-} B(x) = B(r_n) - \frac{1}{2^n}$ and $\lim_{x\to r_n^+} B(x) = B(r_n)$,*

4. $\lim_{x\to-\infty} B(x) = 0$,

5. $\lim_{x\to+\infty} B(x) = 1$.

Proof. To see the first statement, note that $x \le y$ implies that $f_k(x) \le f_k(y)$ for every k. Hence, $B(x) = \sum_{k=1}^{+\infty} f_k(x) \le \sum_{k=1}^{+\infty} f_k(y) = B(y)$. On the other hand, if $x < y$, then there is a rational between them, so there exists n with $x < r_n < y$. Thus, $f_n(x) = 0$, while $f_n(y) = \frac{1}{2^n}$. This implies that $B(x) + \frac{1}{2^n} \le B(y)$, and the first statement follows.

To see the second, let $\mathcal{I}$ denote the set of irrational numbers. Then each function f_k is continuous on $\mathcal{I}$. Since B is the uniform sum of continuous functions on $\mathcal{I}$, B is continuous on $\mathcal{I}$.

Since B is increasing, the two one-sided limits exist and are given by the suprema and infima as in the above proposition. Given any $\epsilon > 0$, choose $K > n$ so that $\sum_{k=K+1}^{+\infty} \frac{1}{2^k} < \epsilon$ and let $\delta = \min\{|r_n - r_k| : 1 \le k \le K, \text{ and } k \ne n\}$. Then for any x with $r_n < x < r_n + \delta$ we have that $f_k(r_n) = f_k(x)$ for $1 \le k \le K$. Hence,

$$|B(r_n) - B(x)| = B(x) - B(r_n) = \sum_{k=K+1}^{+\infty} f_k(x) - f_k(r_n) \le \sum_{k=K+1}^{+\infty} f_k(x) < \epsilon.$$

This shows that $\lim_{x\to r_n^+} B(X) = B(r_n)$.

On the other hand, if $r_n - \delta < x < r_n$, then $f_k(x) = f_k(r_n), 1 \le k \le K, k \ne n$, while $f_n(x) = 0$ and $f_n(r_n) = \frac{1}{2^n}$. Hence,

$$|(B(r_n) - \frac{1}{2^n}) - B(x)| = \sum_{k=K+1}^{+\infty} (f_k(r_n) - f_k(x)) \le \sum_{k=K+1}^{+\infty} \frac{1}{2^k} < \epsilon.$$

This proves that $\lim_{x\to r_n^-} B(x) = B(r_n) - \frac{1}{2^n}$.

To prove statements 4 and 5, given $\epsilon > 0$, let again $\sum_{k=K+1}^{+\infty} \frac{1}{2^k} < \epsilon$ and set $m = \min\{r_k : 1 \le k \le K\}$ and $M = \max\{r_k : 1 \le k \le K\}$. Then for $x < m$, we have that $0 \le B(x) = \sum_{k=K+1}^{+\infty} f_k(x) \le \sum_{k=K+1}^{+\infty} \frac{1}{2^k} < \epsilon$, which proves that $\lim_{x\to-\infty} B(x) = 0$. While for $x > M$, we have that $1 - \epsilon < \sum_{k=1}^{K} \frac{1}{2^k} = \sum_{k=1}^{K} f_k(x) \le \sum_{k=1}^{+\infty} f_k(x) \le \sum_{k=1}^{+\infty} \frac{1}{2^k} = 1$, which proves that $\lim_{x\to+\infty} B(X) = 1$. □

Since the function B is increasing, it is also possible to do Riemann-Stieltjes integration with respect to B. Because of the series representation of B it is possible to write explicit formulas for this integration. We do the case when the interval has irrational endpoints to keep things simple.

Recall that continuous functions are always Riemann-Stieltjes integrable. We will also need a well-known estimate, whose proof we leave as an exercise:

Lemma 1.33. *Let* $\alpha : [a, b] \to \mathbb{R}$ *be increasing and let* $f : [a, b] \to \mathbb{R}$ *be Riemann-Stieltjes integrable with respect to* α *with* $M = \sup\{|f(x)|; a \leq x \leq b\}$. *Then*

$$|\int_a^b f d\alpha| \leq M(\alpha(b) - \alpha(a)).$$

Proposition 1.34. *Let* $a < b$ *be irrational numbers and let* $f : [a, b] \to \mathbb{R}$ *be continuous and set* $f(x) = 0$ *for* $x \notin [a, b]$. *Then*

$$\int_a^b f dB = \sum_{k=1}^{+\infty} \frac{f(r_k)}{2^k}.$$

Proof. Let $B_K = \sum_{k=1}^{K} f_k$ and $R_K = \sum_{k=K+1}^{+\infty} f_k$. Then $B = B_K + R_K$ and B_K and R_K are both increasing functions. By Proposition I.5.21, we have that

$$\int_a^b f dB = \int_a^b f dB_K + \int_a^b f dR_K =$$
$$\sum_{k=1}^{K} \int_a^b f df_k + \int_a^b f dR_K =$$
$$\sum_{k=1}^{K} \frac{f(r_k)}{2^k} + \int_a^b f dR_K,$$

by Theorem I.5.23.

Now we want to show that $\int_a^b f dR_K \to 0$ as $K \to +\infty$. To see this let $M = \sup\{|f(x)| : a \leq x \leq b\}$ and using the lemma, we have that

$$|\int_a^b f dR_K| \leq M(R_k(b) - R_k(a)) \leq M \sum_{k=K+1}^{+\infty} \frac{1}{2^k} = \frac{M}{2^K},$$

which clearly tends to 0 as $K \to +\infty$.

Thus, we have that

$$\int_a^b f dB = \lim_{K \to +\infty} \sum_{k=1}^{K} \frac{f(r_k)}{2^k},$$

and the result follows. ∎

With a little care one can improve on Proposition 1.22. If a function f has a jump discontinuity at c, then we call

$$\delta = |\lim_{x\to c^-} f(x) - \lim_{x\to c^+} f(x)|,$$

the **size** of the jump discontinuity.

Problem 1.35. *Let $f : [a,b] \to \mathbb{R}$ be an increasing function. Prove that if f has jump discontinuities of size at least δ at points $\{c_1, ..., c_k\}$, with $a < c_1 < ... < c_k < b$ then*

$$k \leq \frac{f(b) - f(a)}{\delta}.$$

Problem 1.36. *Use the result of the last problem to show that an increasing function $f : [a,b] \to \mathbb{R}$ can have at most countably many discontinuities. Prove that an increasing function $f : \mathbb{R} \to \mathbb{R}$ can have at most countably many discontinuities.*

Problem 1.37. *Prove Lemma 1.33.*

1.6 A Space Filling Curve

Given a continuous function $f : \mathbb{R} \to \mathbb{R}^2$ we often imagine its graph as the path traced out by a moving point and in calculus we are often interested in computing the length of the path traveled. We also have a naive concept of how dimension behaves. We think of the real line as being "one-dimensional" and imagine that the continuous image of a one-dimensional set should also be "one-dimensional".

In this section we construct a continuous function $f : [0,1] \to [0,1] \times [0,1]$ that is onto, which shows how ungraphable a continuous function can be and also shows that our "primitive" notions of dimension require some concrete hypotheses to make them precise. The construction that we present is due to I.J. Schoenberg.

First we define a continuous function $\phi : [0,2] \to [0,1]$ by setting

$$\phi(t) = \begin{cases} 0, & 0 \leq t \leq 1/3, \\ 3t - 1, & 1/3 \leq t \leq 2/3, \\ 1, & 2/3 \leq t \leq 4/3, \\ -3t + 5, & 4/3 \leq t \leq 5/3, \\ 3t - 1, & 5/3 \leq t \leq 2. \end{cases}$$

We then extend ϕ to a periodic function on all of $\mathbb{R}$(still denoted by ϕ) setting $\phi(t+2) = \phi(t)$. Since $\phi(0) = \phi(2)$, this periodic extension is continuous.

Now define functions by,

$$f_1(t) = \sum_{k=1}^{+\infty} \frac{\phi(3^{2k-2}t)}{2^k} \text{ and } f_2(t) = \sum_{k=1}^{+\infty} \frac{\phi(3^{2k-1}t)}{2^k}.$$

Since ϕ is bounded by 1, each of the functions in these series is bounded by $M_k = 2^{-k}$ and consequently, they converge uniformly to define continuous functions f_1, f_2 on $\mathbb{R}$. Moreover, since $|\phi(t)| \leq 1$, we see that $|f_i(t)| \leq 1$.

Thus, we may define a continuous function $f : \mathbb{R} \to [0,1] \times [0,1]$ by setting $f(t) = (f_1(t), f_2(t))$.

We will now prove that f maps $[0,1]$ *onto* the unit square $[0,1] \times [0,1]$. To see this, pick any point (a, b) with $0 \leq a, b \leq 1$. It is most convenient if we use their binary expansion. That is, we choose sequences $a_n, b_n \in \{0,1\}$, such that

$$a = \sum_{n=1}^{+\infty} \frac{a_n}{2^n} \text{ and } b = \sum_{n=1}^{+\infty} \frac{b_n}{2^n}.$$

Set $c = 2\sum_{n=1}^{+\infty} \frac{c_n}{3^n}$ where $c_{2n-1} = a_n$ and $c_{2n} = b_n$. Clearly, $0 \leq c \leq 2\sum_{n=1}^{+\infty} \frac{1}{3^n} = 1$.

We claim that $f_1(c) = a$ and $f_2(c) = b$.

Note that

$$3^k c = 2\sum_{n=1}^{k} \frac{c_n}{3^{n-k}} + 2\sum_{n=k+1}^{+\infty} \frac{c_n}{3^{n-k}}.$$

Since $n-k$ is negative for $1 \leq n \leq k$, the first sum is an even integer. Let d_k denote the second term, then $\phi(3^k c) = \phi(d_k)$.

Now if $c_{k+1} = 0$, then $0 \leq d_k = 2\sum_{n=k+2}^{+\infty} \frac{c_n}{3^{n-k}} \leq 2\sum_{m=2}^{+\infty} \frac{1}{3^m} = 1/3$. Hence, $\phi(3^k c) = \phi(d_k) = 0 = c_{k+1}$.

When $c_{k+1} = 1$, a similar calculation shows that $2/3 \leq d_k \leq 1$, and hence, $\phi(3^k c) = \phi(d_k) = 1$.

Thus, in either case, $\phi(3^k c) = c_{k+1}$, and we have that

$$f_1(c) = \sum_{k=1}^{+\infty} \frac{\phi(3^{2k-2}C)}{2^k} = \sum_{k=1}^{+\infty} \frac{c_{2k-1}}{2^k} = a,$$

while

$$f_2(c) = \sum_{k=1}^{+\infty} \frac{\phi(3^{2k-1}c)}{2^k} = \sum_{k=1}^{+\infty} \frac{c_{2k}}{2^k} = b.$$

Problem 1.38. *Show that when* $c_{k+1} = 1$, *that* $2/3 \leq d_k \leq 1$.

Problem 1.39. *Use the above result to construct a continuous function* $f : [0, 1] \to [0, 1] \times [0, 1] \times [0, 1]$ *that is onto.*

1.7 Power Series

Given $x_0 \in \mathbb{R}$ by a **power series centered at** x_0 we mean an expression of the form

$$\sum_{k=0}^{+\infty} a_k (x - x_0)^k,$$

where $a_k \in \mathbb{R}$ are constants and for convenience of notation we let $(x - x_0)^0$ denote the function that is constantly equal to 1. Thus, a power series is just a particular type of series of functions, where each function $f_k(x) = a_k(x - x_0)^k$ is a function from $\mathbb{R}$ to $\mathbb{R}$.

In this section we study convergence of power series, behavior of their deriviatives, integrals and the behavior of sums and products of power series.

Note that the above power series always converges to a_0 when $x = x_0$. Thus every power series always converges at this one point.

Most theorems about power series rely on the *root test* for series, so we recall it here. Since the convergence or divergence of a series doesn't depend on the first few terms, we will write our series as $\sum_k a_k$.

Theorem 1.40 (Root Test)**.** *Given a series* $\sum_k a_k$, *set*

$$A = \limsup_k |a_k|^{1/k}.$$

1. *If* $A < 1$, *then* $\sum_k a_k$ *converges absolutely.*
2. *If* $A > 1$, *then the series* $\sum_k a_k$ *diverges.*
3. *There exist series with* $A = 1$ *which converge and series with* $A = 1$ *which diverge, i.e.,* $A = 1$ *is inconclusive.*

Proof. If $A < 1$, then pick r with $A < r < 1$. By the definition of the limsup, there is a K so that for every $k \geq K$, we have that $|a_k|^{1/k} < r$. Hence, $|a_k| < r^k$ for $k \geq K$. But the series $\sum_{k=K}^{+\infty} r^k$ converges and hence $\sum_{k=K}^{+\infty} |a_k|$ converges by the comparison test. Thus, $\sum a_k$ converges absolutely.

If $A > 1$, then pick r with $A > r > 1$. Again by the definition of the limsup, for every integer K, there must exist $k > K$, such that $|a_k|^{1/k} > r$.

This implies that $|a_k| > r^k > 1$. Thus, there are infinitely many terms in the series that are greater than 1. This implies that $\lim_k a_k \neq 0$, and hence the series diverges.

Finally, $\sum 1/k$ and $\sum 1/k^2$ are series for which $A = 1$, one diverges and the other converges. □

Given a power series as above we set $A = \limsup_{n\to+\infty} |a_n|^{1/n}$, where we allow the possibility $A = +\infty$ to indicate that the limit is unbounded.

We set

$$R = \begin{cases} +\infty & A = 0 \\ 1/A & 0 < A < +\infty \\ 0 & A = +\infty \end{cases},$$

and we call R the **radius of convergence of the power series.** The following result explains this definition.

Recall that a series of numbers $\sum_{k=0}^{+\infty} b_k$ is said to **converge absolutely** provided that the series $\sum_{k=0}^{+\infty} |b_k|$ converges and that absolute convergence of a series implies convergence of the series.

Theorem 1.41. *Let $\sum_{k=0}^{+\infty} a_k(x - x_0)^k$ be a power series and let R be the radius of convergence. Then*

1. *for each fixed x satisfying $|x - x_0| < R$, the series converges absolutely to define a continuous function $s(x)$ on the set $|x - x_0| < R$,*
2. *for each x satisfying $|x - x_0| > R$, the series diverges,*
3. *for each $r < R$, the series converges uniformly on the closed interval $[x_0 - r, x_0 + r]$ to the function s.*

Proof. If we fix x with $|x - x_0| = b < R$, then we have that

$$\limsup_{k\to+\infty} |a_k(x - x_0)^k|^{1/k} = \limsup_{k\to+\infty} b|a_k|^{1/k} = bA < 1.$$

Thus, by the root test, the series of numbers $\sum_{k=0}^{+\infty} a_k(x - x_0)^k$ converges absolutely. This proves the series converges to define a function on the set $|x - x_0| < R$.

If $b = |x - x_0| > R$, then $\limsup_{k\to+\infty} |a_k(x - x_0)^k|^{1/k} = bA > 1$. By Theorem 1.40.2, the series diverges.

To prove the third statement, we apply Weierstrass' M-test. On the closed interval $[x_0 - r, x_0 + r]$ we have that the function $f_k(x) = a_k(x - x_0)^k$ satisfies $|f_k(x)| \leq |a_k|r^k = M_k$. Applying the root test to the series

$\sum_{k=0}^{+\infty} M_k$, we see that $\limsup_{k\to+\infty} M_k^{1/k} = r(\limsup_{k\to+\infty} |a_k|^{1/k}) = rA < 1$. Thus, by the root test, the series $\sum_{k=0}^{+\infty} M_k$ converges and so by the Weierstrass' M-test the series of functions converges uniformly on its domain.

Now we show that s is continuous on $|x - x_0| < R$. Given any x_1, with $|x_1 - x_0| < R$, choose r with $|x_1 - x_0| < r < R$. Since the convergence is uniform on $|x - x_0| \le r$, and each function $a_k(x - x_0)^k$ is continuous, $s(x)$ is continuous on $[x_0 - r, x_0 + r]$ and hence at x_1. This finishes the proof of the theorem. □

Corollary 1.42. *Let $\sum_{k=0}^{+\infty} a_k(x - x_0)^k$ be a power series. If this series converges for a real number x_1, with $|x_1 - x_0| = R_1$, then $R_1 \le R$.*

Example 1.43. *Set $a_0 = 1$ and $a_k = k^k$ for $k \ge 1$. Then $A = +\infty$, $R = 0$ and the power series $\sum_{k=0}^{+\infty} a_k x^k$ only converges for $x = 0$.*

Example 1.44 (The Geometric Series)**.** *Set $a_k = 1$ for all k, then $A = 1 = R$, and the series $\sum_{k=0}^{+\infty} x^k$ converges for $-1 < x < +1$. This series does not converge at either of the endpoints, $x = +1$ or $x = -1$.*

Since $\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$, we see that the partial sums of this series, $s_N(x) = \sum_{k=0}^{N} x^k$ converge to the function $\frac{1}{1-x}$ uniformly on any interval of the form $[-r, +r], r < 1$.

Another method that is often useful for estimating radii of convergence is the ratio test. The following result shows that it always gives a lower estimate on the radius of convergence. We then give an example to show that it in fact does not always give the actual radius of convergence.

Proposition 1.45. *Let $B = \limsup_{k\to+\infty} \frac{|a_{k+1}|}{|a_k|}$, then $\limsup_{k\to+\infty} |a_k|^{1/k} \le B$. Hence, $1/B \le 1/A = R$.*

Proof. Let $\delta > 0$, since $B = \inf_m \sup\{\frac{|a_{k+1}|}{|a_k|} : k \ge m\}$, there exists an m, so that $\sup\{\frac{|a_{k+1}|}{|a_k|} : k \ge m\} \le B + \delta$. This implies that

$$|a_{m+j}| \le (B + \delta)^j |a_m|.$$

Hence, $A = \limsup_{j\to+\infty} |a_{m+j}|^{1/(m+j)} \le \limsup_{j\to+\infty} (B+\delta)^{\frac{j}{j+m}} |a_m|^{1/(m+j)} = B + \delta$.

Since δ was arbitrary, $A \le B$. □

The above result also shows that the root test is a more powerful test, in the following sense. If $B \le 1$, then $A \le 1$. So any time that the ratio test shows that a series converges, the root test would also show that it converges. However, since $A \leq B$ it is possible that $A < 1 < B$ in which case the series would converge, but you would not see that with the ratio.

The advantage of the ratio test is that often the limit of the ratios is easier to compute.

Example 1.46. *Let* $a_k = \begin{cases} 2^k, & k \text{ even} \\ 3^k, & k \text{ odd} \end{cases}$ *and set* $f(x) = \sum_{k=0}^{+\infty} a_k x^k$*. Since* $\limsup_{k\to+\infty} |a_k|^{1/k} = 3$*, this power series has radius of convergence* $1/3$*. However,* $\limsup_{k\to+\infty} \frac{|a_{k+1}|}{|a_k|} = +\infty$*.*

Thus, the ratio test only allows you to deduce that the series converges for $x = 0$, while we know that it converges for any $|x| < 1/3$.

Example 1.47 (The Exponential Series)**.** *Let* $a_k = \frac{1}{k!}$ *where we set* $0! = 1$*. Then* $B = \limsup_{k\to+\infty} \frac{a_{k+1}}{a_k} = 0$*. Hence by Proposition 1.45,* $A = 0$*, too and the series* $\sum_{k=0}^{+\infty} \frac{x^k}{k!}$ *has radius of convergence* $R = +\infty$*. The function that it converges to is denoted* e^x *or* $exp(x)$*. On every interval* $[-r, +r]$ *we have that the partial sums of this series converge uniformly to* e^x*.*

The number

$$e = \sum_{k=0}^{+\infty} \frac{1}{k!}$$

plays an important role in mathematics, so it is often useful to know how accurately the partial sums of the series approximate e.

Proposition 1.48. *Let* $s_n = \sum_{k=0}^{n} \frac{1}{k!}$*, then* $0 < e - s_n < \frac{1}{(n+1)!}\frac{n+2}{n+1}$*.*

Proof. We have that

$$e - s_n = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots = \frac{1}{(n+1)!}[1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots] <$$
$$\frac{1}{(n+1)!}[1 + \frac{1}{(n+2)} + \frac{1}{(n+2)^2} + \cdots] = \frac{1}{(n+1)!}\frac{n+2}{n+1}.$$

□

Theorem 1.49. *The number e is irrational.*

Proof. Suppose instead that $e = \frac{p}{q}$, with p, q positive integers. Then we have that $0 < (q!)[e - s_q] < \frac{q!}{(q+1)!}\frac{q+2}{q+1} = \frac{q+2}{(q+1)^2} < 1$. Since $e = \frac{p}{q}$, we have that $(q!)e$ is an integer. Also, $(q!)s_q = \sum_{k=0}^{q} \frac{q!}{k!}$ is a sum of integers. Hence, $(q!)[e - s_q]$ is an integer strictly between 0 and 1, contradiction. □

Problem 1.50. *Prove that on the interval* $(-1/3, +1/3)$ *the power series of Example 1.46 converges to the function*

$$\frac{1}{1-4x^2} + \frac{3x}{1-9x^2}.$$

Problem 1.51. *Prove that if* $a_k \in \mathbb{R}$ *and* $\sum_{k=0}^{+\infty} a_k$ *converges absolutely, then* $\sum_{k=0}^{+\infty} a_k x^k$ *converges uniformly on* $[-1, +1]$.

Problem 1.52. *Prove that if* $a_k \in \mathbb{R}$ *and* $\sum_{k=0}^{+\infty} a_k$ *converges, then* $\sum_{k=0}^{+\infty} a_k x^k$ *converges uniformly on* $[-r, +1]$ *for any* $0 \le r < 1$.

1.8 Operations on Power Series

In this section we examine how power series behave under sums, products, differentiation and integration.

Lemma 1.53. *For any* $0 < A$ *we have that* $\lim_{k\to+\infty} A^{1/k} = 1$ *and* $\lim_{k\to+\infty}(k+1)^{1/k} = 1$.

Proof. We first prove the second statement. Write $(k+1)^{1/k} = 1 + b_k$ for some $b_k > 0$. Then we have that $k+1 = (1+b_k)^k = 1 + kb_k + \frac{k^2}{2}b_k^2 + ...$, where all the remaining terms are non-negative. Hence, $1 + \frac{k^2}{2}b_k^2 \le 1 + k$ and it follows that $b_k^2 \le \frac{2}{k}$. Thus, $b_k^2 \to 0$ and $b_k \to 0$, from which it follows that $(k+1)^{1/k} \to 1$.

Now let $1 \le A$, then for k large enough we have that $A \le k+1$ and hence, $1 \le A^{1/k} \le (k+1)^{1/k}$. Thus, by the "Squeeze Theorem" and the fact $\lim_k (k+1)^{1/k} = 1$, we have that $\lim_k A^{1/k} = 1$. Now if $0 < A < 1$, then $\frac{1}{A} \ge 1$ and hence $1 = \lim_k (1/A)^{1/k} = \lim_k \frac{1}{A^{1/k}}$. Hence, $\lim_k A^{1/k} = \frac{1}{1} = 1$. □

Theorem 1.54. *Let the power series* $f(x) = \sum_{k=0}^{+\infty} a_k(x-x_0)^k$ *have radius of convergence* $R > 0$. *Then*

1. *the power series* $g(x) = \sum_{k=0}^{+\infty} a_{k+1}(k+1)(x-x_0)^k$ *has radius of convergence at least* R,

2. *the function* f *is differentiable on the interval* $|x-x_0| < R$, *and* $f'(x) = g(x)$,

3. *the power series* $F(x) = \sum_{k=1}^{+\infty} \frac{a_{k-1}}{k}(x-x_0)^k$ *has radius of convergence at least* R *and* $F'(x) = f(x)$.

Proof. Let $A = \limsup_{k\to+\infty} |a_k|^{1/k}$. We only consider the case that $0 < A < +\infty$, and leave the other cases to the reader. Given any $\delta > 0$, we may choose K large enough so that $\sup\{|a_k|^{1/k} : k \geq K\} \leq A + \delta$. We have that $\sup\{|a_{k+1}(k+1)|^{1/k} : k \geq K\} = \sup\{[|a_{k+1}|^{1/(k+1)}]^{(k+1)/k}(k+1)^{1/k} : k \geq K\} \leq \sup\{(A+\delta)^{(k+1)/k}(k+1)^{1/k} : k \geq K\}$. Now by the above lemma we may choose K large enough so that for any $k \geq K$, the above product is as close to $A + \delta$ as we may choose. Since $\delta > 0$, was arbitrary, we have that

$$\limsup_{k\to+\infty} |a_{k+1}(k+1)|^{1/k} \leq \limsup_{k\to+\infty} |a_k|^{1/k},$$

and so the power series for g has a greater radius of convergence than f.

The second statement follows from the first by considering the partial sums of $s_n(x) = \sum_{k=0}^{n} a_k(x-x_0)^k$ of f. On any given closed interval $[x_0 - r, x_0 + r]$, with $r < R$, these functions converge uniformly to f and their derivatives, s_n' are the partial sums of g which converge uniformly to g. Hence, applying Theorem 1.20, we have that $f' = g$.

The proof that the power series for F has radius of convergence at least R is similar and then the fact that $F' = f$ follows by applying the first two results to the power series for F. □

Corollary 1.55. *Let $f(x) = \sum_{k=0}^{+\infty} a_k(x-x_0)^k$ have radius of convergence R. Then on $|x-x_0| < R$, f is infinitely differentiable, the derivatives of f can be obtained by formally differentiating the power series for f term-by-term and these new power series will have radius of convergence R.*

Proof. By the previous theorem, each derivative will have radius of convergence at least R and so it can again be differentiated. The only thing that needs to be checked is that they all have radius of convergence exactly R. To see this suppose that the power series for f' had radius of convergence $R_1 > R$. Then by the third part of the theorem, the integral of f' would also have radius of convergence at least R_1. But this function differs from f by at most a constant and so f would have radius of convergence R_1. Contradiction. Hence the radius of convergence of every derivative has exactly the same radius of convergence as f. □

The following is elementary so we omit its proof.

Proposition 1.56. *Let $f(x) = \sum_{k=0}^{+\infty} a_k(x-x_0)^k$ and $g(x) = \sum_{k=0}^{+\infty} b_k(x-x_0)^k$ both have radius of convergence at least R. Then $\sum_{k=0}^{+\infty}(a_k+b_k)(x-x_0)^k$ has radius of convergence at least R and converges to $f(x)+g(x)$ on this domain.*

Products of power series are more interesting. Note that if we have two polynomials $f(x) = \sum_{k=0}^{N} a_k x^k$ and $g(x) = \sum_{k=0}^{M} b_k x^k$ and we think of them as power series by setting all of their higher order coefficients equal to 0, then gathering terms we see that

$$f(x)g(x) = \sum_{k=0} c_k x^x$$

where $c_0 = a_0 b_0, c_1 = a_0 b_1 + a_1 b_0$, and $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. This motivates the following results:

Theorem 1.57 (Cauchy Product Theorem). *Assume that $\sum_{k=0}^{+\infty} a_k$ converges absolutely to A and $\sum_{k=0}^{+\infty} b_k$ converges absolutely to B, and set $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Then $\sum_{k=0}^{+\infty} c_k$ converges absolutely to AB.*

Proof. Set $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$ and $C_n = \sum_{k=0}^{n} c_k$. Also let $P = \sum_{k=0}^{+\infty} |a_k|$ and let $Q = \sum_{k=0}^{+\infty} |b_k|$.

Given a finite set $S \subseteq \mathbb{N} \times \mathbb{N}$, we shall adopt the notation, $\sum_S a_k b_j$ to denote the sum over all pairs (k, j) that belong to S.

First to see the absolute convergence, note that

$$\sum_{n=0}^{N} |c_n| = \sum_{n=0}^{N} |\sum_{k=0}^{n} a_k b_{n-k}| \leq$$

$$\sum_{n=0}^{N} \sum_{k=0}^{n} |a_k||b_{n-k}| = \sum_{S_N} |a_k||b_j|,$$

where $S_N = \{(k, j) : 0 \leq k + j \leq N\}$.

Now since $S_N \subseteq R_N = \{(k, j) : 0 \leq k \leq N, 0 \leq j \leq N\}$, we have that

$$\sum_{n=0}^{N} |c_n| \leq \sum_{R_N} |a_k||b_j| = (\sum_{k=0}^{N} |a_k|)(\sum_{j=0}^{N} |b_j|).$$

Thus, for every N, we have that

$$\sum_{n=0}^{N} |c_n| \leq (\sum_{k=0}^{+\infty} |a_k|)(\sum_{j=0}^{+\infty} |b_j|) = PQ$$

and so $\sum c_n$ is absolutely convergent.

Next, note that $A_N B_N - C_N = \sum_{R_N \setminus S_N} a_k b_j$. Since $\lim_N A_N B_N = AB$, to prove that $\sum_{n=0}^{+\infty} c_n = AB$, it will be enough to show that $\lim_N (A_N B_N - C_N) = 0$.

To this end given $\epsilon > 0$, choose L so that $\sum_{j=L}^{+\infty} |b_j| < \epsilon/2P$ and $\sum_{k=L}^{+\infty} |a_k| < \epsilon/2Q$.

Now if $N \geq 2L$, then $R_N \backslash S_N \subseteq \{(k,j) : 0 \leq k \leq N, L \leq j \leq N\} \cup \{(k,j) : L \leq k \leq N, 0 \leq j \leq N\}$. Hence,

$$|A_N B_N - C_N| \leq \sum_{R_N \backslash S_N} |a_k||b_j| \leq \sum_{k=0}^{N}\sum_{j=L}^{N} |a_k||b_j| + \sum_{k=L}^{N}\sum_{j=0}^{N} |a_k||b_j| < P(\epsilon/2P) + Q(\epsilon/2Q) = \epsilon.$$

This inequality shows that $A_N B_N - C_N \to 0$ as $N \to +\infty$ and so the proof is complete. □

As an application of this result we prove a familiar fact about the exponential function.

Proposition 1.58. *Let $a, b \in \mathbb{R}$. Then $exp(a) \cdot exp(b) = exp(a+b)$.*

Proof. Let $a_k = \frac{a^k}{k!}$ and let $b_k = \frac{b^k}{k!}$ so that $exp(a) = \sum_{k=0}^{+\infty} a_k$, $exp(b) = \sum_{k=0}^{+\infty} b_k$ and both series converge absolutely. Thus, by the Cauchy Product theorem, if we set $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, then $exp(a)exp(b) = \sum_{n=0}^{+\infty} c_n$.

However, using the binomial formula,

$$c_n = \frac{1}{n!}\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^k b^{n-k} = \frac{1}{n!}\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \frac{1}{n!}(a+b)^n.$$

Hence, $\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} \frac{(a+b)^n}{n!} = exp(a+b)$. □

We can now prove our main theorem about products of power series.

Theorem 1.59. *Let $f(x) = \sum_{k=0}^{+\infty} a_k (x-x_0)^k$ and $g(x) = \sum_{k=0}^{+\infty} b_k (x-x_0)^k$ be two power series both with radius of convergence at least R and set $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Then the power series $h(x) = \sum_{k=0}^{+\infty} c_k (x-x_0)^k$ has radius of convergence at least R and for any $x, |x-x_0| < R$, we have that $h(x) = f(x)g(x)$.*

Proof. For each fixed x, with $|x - x_0| < R$, we have that both series of numbers converge absolutely. Hence, by the above result their Cauchy product converges absolutely. But when we fix x, the k-th term of each series is

$\hat{a_k} = a_k(x - x_0)^k$ and $\hat{b_k} = b_k(x - x_0)^k$. Thus, the n-th term of the Cauchy product is

$$\hat{c_n} = \sum_{k=0}^{n} \hat{a}_k \hat{b}_{n-k} = \sum_{k=0}^{n} (a_k(x - x_0)^k)(b_{n-k}(x - x_x)^{n-k}) = c_n(x - x_0)^n.$$

Hence, by the Cauchy product theorem for each x with $|x - x_)| < R$, the series $\sum_{n=0}^{+\infty} c_n(x - x_0)^n$ converges absolutely to the number $f(x)g(x)$. Thus, the product formula holds and by Corollary this implies that the series $\sum_{n=0}^{+\infty} c_n(x - x_0)^n$ has radius of convergence at least R. □

Corollary 1.60. *Let* $c_n = \sum_{k=0}^{n} a_k b_{n-k}$*, then*

$$\limsup_n |c_n|^{1/n} \leq \max\{\limsup_n |a_n|^{1/n}, \limsup_n |b_n|^{1/n}\}.$$

Proof. We only do the case where $0 < A < +\infty$ and $0 < B < +\infty$. Let $A = \limsup_n |a_n|^{1/n}$ and $B = \limsup_n |b_n|^{1/n}$. Then the power series $\sum_{k=0}^{+\infty} a_k x^k$ and $\sum_{k=0}^{+\infty} b_k x^k$ both have radius of convergence $R = \min\{1/A, 1/B\}$, and hence the series $\sum_{k=0}^{+\infty} c_k x^k$ converges absolutely for $|x| < R$. By the root test this implies that $\limsup_n |c_n x^n|^{1/n} < 1$, for $|x| < R$. Hence, $\limsup_n |c_n|^{1/n} \leq 1/R = \max\{A, B\}$. □

The following result tells us a great deal about convergence of a power series at an endpoint of its interval of convergence.

Theorem 1.61. *Let* $f(x) = \sum_{n=0}^{+\infty} a_n(x - x_0)^n$ *have radius of convergence at least* R*. If* $s = \sum_{n=0}^{+\infty} a_n R^n$ *converges, then*

$$\lim_{x \to (x_0+R)^-} f(x) = s.$$

Thus, if one sets $f(x_0 + R) = s$*, then* f *is defined and continuous on the interval* $(x_0 - R, x_0 + R]$*.*

The case for the other endpoint of the interval is similar and is left to the exercises.

Proof. The continuity statement follows from the limit result. We only prove the case $R = 1$ and $x_0 = 0$, the general case follows easily from this case by substitution. Thus, we must prove that $\lim_{x \to 1^-} f(x) = s$.

For $n \geq 0$, let $s_n = \sum_{k=0}^{n} a_k$ denote the partial sums of the series and for convenience of notation set $s_{-1} = 0$. Then we will have that $a_n = s_n - s_{n-1}$ for $n \geq 0$.

Since the sequence of partial sums is convergent, it is bounded, say, $|s_n| \le M$ and we have $\limsup_n |s_n|^{1/n} \le \limsup_n M^{1/n} = 1$. Thus, the power series $g(x) = \sum_{n=0}^{+\infty} s_n x^n$ has radius of convergence at least 1.

We have

$$\sum_{n=0}^{K} a_n x^n = \sum_{n=0}^{K} (s_n - s_{n-1}) x^n = \sum_{n=0}^{K} s_n x^n - \sum_{n=1}^{K} s_{n-1} x^n = \sum_{n=0}^{K} s_n x^n - \sum_{n=0}^{K-1} s_n x^{n+1} = s_K x^K + (1-x) \sum_{n=0}^{K-1} s_n x^n.$$

Thus, for $|x| < 1$, we have

$$f(x) = \lim_{K \to +\infty} \sum_{n=0}^{K} a_n x^n = \lim_{K \to +\infty} s_K x^K + (1-x) \sum_{n=0}^{K-1} s_n x^n = (1-x) g(x),$$

since $s_K x^K \to 0$, for $|x| < 1$.

Note that for $|x| < 1$, we have $(1-x)\sum_{n=0}^{+\infty} x^n = 1$.

Given $\epsilon > 0$, choose N so that $n > N$ implies that $|s - s_n| < \epsilon/2$. Let $C_N = \sum_{n=0}^{N} |s - s_n|$, and set $\delta = \frac{\epsilon}{2C_N}$. Then for $1 - \delta < x < 1$, we have that

$$|f(x) - s| = |(1-x) \sum_{n=0}^{+\infty} s_n x^n - s(1-x) \sum_{n=0}^{+\infty} x^n| =$$
$$|1-x| \cdot |\sum_{n=0}^{+\infty} (s_n - s) x^n| \le |1-x| \sum_{n=0}^{N} |s_n - s| x^n + |1-x| \sum_{n=N+1}^{+\infty} \frac{\epsilon}{2} x^n \le$$
$$\delta \sum_{n=0}^{N} |s - s_n| + |1-x| \frac{\epsilon}{2} \frac{x^{N+1}}{1-x} \le \delta C_N + \frac{\epsilon}{2} x^{N+1} < \epsilon.$$

This proves that $\lim_{x \to 1^-} f(x) = s$, which completes the proof. □

Problem 1.62. *Use the power series for* $\frac{1}{(1-x)}$ *to find a power series that converges for* $|x| < 1$ *to* $\frac{1}{(1-x)^2}$*, two different ways. First, using differentiation and secondly using the Cauchy product. Cite theorems to justify your calculations.*

Problem 1.63. *Use the fact that* $(ln(x))' = \frac{1}{x}$*, differentiation and the formula,* $\frac{1}{x} = \frac{1}{1-(1-x)}$*, to find a power series for* $ln(x)$ *that converges for* $|x - 1| < 1$*. Justify your calculations with theorems.*

Problem 1.64. *Give a careful proof that the series* $f(x) = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} x^{2k}$ *and* $g(x) = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$ *both have radius of convergence* $+\infty$.

Problem 1.65. *Let* f, g *be the functions given by the above power series. Prove that* $f'(x) = -g(x)$, *that* $g'(x) = f(x)$ *and that* $f^2(x) + g^2(x) = 1$ *for every* $x \in \mathbb{R}$.

Problem 1.66. *Complete the proof of Theorem 1.61 by showing that the case* $R = 1$ *and* $x_0 = 0$ *implies the general case.*

Problem 1.67. *Let* $f(x) = \sum_{n=0}^{+\infty} a_n (x - x_0)^n$ *be a power series with radius of convergence at least* R. *Prove that if* $s = \sum_{n=0}^{+\infty} a_n (-R)^n$ *converges, then* $\lim_{x \to (x_0 - R)^+} f(x) = s$. *Conclude that setting* $f(x_0 - R) = s$ *defines a functions that is continuous on* $[x_0 - R, x_0 + R)$.

1.9 Taylor Series

In this section we take a careful look at Taylor polynomials and convergence of Taylor series.

When f is differentiable on an interval and the function f' is also differentiable, then we write f'' for the derivative of f'. This notation becomes tedious for more than a couple of derivatives, so we adopt the notation, $f^{(n)}$ for the function that is the result of performing the derivative operation n times. In particular, $f^{(1)} = f'$ and $f^{(0)} = f$.

We will say that a function is differentiable on a closed interval $[a, b]$ provided that it is differentiable at each point in (a, b) and that at the two endpoints, we have the one-sided limits exist.

Recall that if a function has a derivative, then it is continuous. So when we say that f is n-times differentiable on an interval $[a, b]$ this means that $f, f^{(1)}, ..., f^{(n-1)}$ exist and are continuous on $[a, b]$ and that the derivative of the function $f^{(n-1)}$ exists at each point of $[a, b]$, but we make no claims about its continuity.

Theorem 1.68 (Taylor's Theorem)**.** *Let* $f : [a, b] \to \mathbb{R}$ *be* $(n-1)$*-times differentiable on* $[a, b]$, *let* $f^{(n)}$ *exist on* (a, b), *let* $x_0 \in [a, b]$, *and set*

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!} (t - x_0)^k.$$

Then for any $x \in [a, b]$, $x \neq x_0$, *there exists* c *between* x *and* x_0 *such that* $f(x) = P(x) + \frac{f^{(n)}(c)}{n!} (x - x_0)^n$.

Note that when $n = 1$ this is the Mean Value Theorem. The requirement that $x \neq x_0$ guarantees that $f^{(n)}$ exists for every t between x and x_0. The polynomial P appearing in the statement of the theorem is called the **(n-1)-th Taylor polynomial centered at** x_0.

Proof. Set $M = \frac{f(x)-P(x)}{(x-x_0)^n}$ so that we need to show that there exists c between x and x_0 with $n!M = f^{(n)}(c)$. Let

$$g(t) = f(t) - P(t) - M(t-x_0)^n, a \leq t \leq b,$$

so that $g^{(n)}(t) = f^{(n)}(t) - n!M, a < t < b$. The proof wil be complete if we show that there is c between x and x_0 with $g^{(n)}(c) = 0$.

Since $P^{(k)}(x_0) = f^{(k)}(x_0), 0 \leq k \leq n-1$, we have that $g^{(k)}(x_0) = 0, 0 \leq k \leq n-1$. Now, $g(x_0) = g(x) = 0$, so by the MVT there exists x_1 between x_0 and x with $g'(x_1) = 0$. Thus, $g'(x_0) = g'(x_1) = 0$ and so applying the MVT again, there exists x_2 between x_0 and x_1 with $g^{(2)}(x_2) = 0$. Continuing in this fashion we obtain points x_k with $g^{(k)}(x_k) = 0$, for $1 \leq k \leq n$. Setting $c = x_n$ completes the proof. □

Corollary 1.69. *Let $0 < R \leq +\infty$ and let f be a real-valued function that is infinitely differentiable for $|x - x_0| < R$. If for each $r < R$, we have that*

$$\lim_{n\to+\infty}(\sup\{\frac{|f^{(n)}(t)|}{n!} : |t-x_0| \leq r\})r^n = 0,$$

then the Taylor series

$$\sum_{n=0}^{+\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

has radius of convergence at least R and converges to $f(x)$ for $|x - x_0| < R$.

Proof. Given any x, with $|x - x_0| < R$, pick $r < R$ with $|x - x_0| < r$. By Taylors Theorem,

$$|f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k| \leq \sup\{\frac{f^{(n)}(t)}{n!} : |t-x_0| \leq r\}r^n,$$

so by the above estimate, the partial sums of the series converge to $f(x)$. □

The following famous example is a function that is infinitely differentiable, its Taylor series centered at 0 has infinite radius of convergence, yet the Taylor series is only equal to the function at the point 0. To avoid too many calculations, we only sketch the details.

We defined a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Proposition 1.70. *The function f is infinitely differentiable on $\mathbb{R}$, $f^{(n)}(0) = 0$ for every n, the Taylor series for f centered at 0 has infinite radius of convergence, but it converges to $f(x)$ only for $x = 0$.*

Proof. Once we verify the formula for the derivative, we will have that all the coefficients of the Taylor series are 0's and hence the Taylor series converges everywhere to the function that is identically 0. But $f(x) = 0$ only when $x = 0$. Thus, all that remains is to verify the formula for the derivatives.

We compute $f'(0) = \lim_{x\to 0} \frac{e^{-1/x^2} - 0}{x - 0} = \lim_{x\to 0} \frac{1/x}{e^{1/x^2}}$. If we substitute $y = 1/x^2$ then as $x \to 0$, we have $y \to +\infty$. Using the series expression for e^y we see that for $y \geq 0$, we have $y^{n+1}/(n+1)! \leq e^y$. From this it follows that $\frac{y^n}{e^y} \leq \frac{(n+1)!}{y}$ and so $\lim_{y\to+\infty} \frac{y^n}{e^y} = 0$ for any $n \in \mathbb{N}$. Consequently, $\lim_{y\to+\infty} \frac{y^r}{e^y} = 0$ for any real number r.

This shows that $f'(0) = \lim_{x\to 0} \frac{\pm\sqrt{y}}{e^y} = 0$.

To compute higher derivatives, note that for $x \neq 0$, $f'(x) = 2x^{-3}e^{-1/x^2} = 2\frac{\pm y^{3/2}}{e^y}$. Hence,

$$f^{(2)}(0) = \lim_{x\to 0} \frac{f'(x) - 0}{x - 0} = \lim_{x\to 0} 2\frac{y^2}{e^y} = 0.$$

The remaining calculations are similar. Inductively, we show that $f^{(n)}(0) = 0$ and that for $x \neq 0$, $f^{(n)}(x) = p(1/x)e^{-1/x^2}$ where p is some polynomial. Consequently,

$$f^{(n+1)}(0) = \lim_{x\to 0} \frac{\pm\sqrt{y}p(\pm\sqrt{y})}{e^y} = 0.$$

□

Problem 1.71. *Use Corollary 1.59 to prove that the Taylor series centered at 0 for cos(x) and sin(x) converge to these functions on the whole real line and show that these Taylor series are the power series f and g of Problem 1.64.*

Earlier, we found a power series in $(x - 1)$ that converged to $ln(x)$ for $|x - 1| < 1$. In the following problems we will see that this is indeed the Taylor series for $ln(x)$ centered at 1, and see what the above results say about convergence of this series.

Problem 1.72. *Compute the Taylor series for* $ln(x)$ *centered at* 1 *and use Corollary 1.59 to prove that the Taylor series centered at* 1 *for* $ln(x)$ *converges to* $ln(x)$ *for* $|x-1| < 1/2$.

Problem 1.73. *Use the estimate in Taylor's theorem to prove that for* $1 < x < 2$, *the Taylor series for* $ln(x)$ *centered at* 1 *converges to* $ln(x)$.

Problem 1.74. *Find the Taylor series for* $arctan(x) = tan^{-1}(x)$ *centered at* 0 *and prove that it converges to* $arctan(x)$ *on* $(-1, +1)$. *Use the fact that* $arctan(1/\sqrt{3}) = \pi/6$ *to give an infinite series representation for* π.

Problem 1.75. *Use Theorem 1.61, the fact that* $arctan(x)$ *is continuous at* $+1$ *and that* $arctan(+1) = \pi/4$ *to give another infinite series representation for* π.

1.10 Polynomial Approximation

In this section we prove that every continuous function on an interval can be aproximated uniformly by a polynomial. First we need a prelimary calculation.

Let $Q_n(x) = c_n(1-x^2)^n$ where c_n is chosen so that $\int_{-1}^{+1} Q_n(x)dx = 1$. Note that $g(x) = (1-x^2)^n - 1 + nx^2 \geq 0$ for $0 \leq x \leq 1$. To see this differentiate, check that the derivative is positive on this interval and is positive at $x = 0$. Since $c_n^{-1} = \int_{-1}^{+1}(1-x^2)^n dx = 2\int_0^1(1-x^2)^n dx \geq 2\int_0^{\sqrt{1/n}}(1-x^2)^n dx \geq 2\int_0^{\sqrt{1/n}}(1-nx^2)dx = \frac{4}{3\sqrt{n}}$ we get that $c_n \leq \frac{3\sqrt{n}}{4}$. Also, since Q_n is decreasing on [0,1], we have that for $0 < \delta \leq |x| \leq 1$,

$$Q_n(x) \leq \sqrt{n}(1-x^2)^n \leq \sqrt{n}(1-\delta^2)^n \to 0.$$

Thus, $Q_n \xrightarrow{\text{u}} 0$ on $\delta \leq |x| \leq 1$.

Theorem 1.76 (Weierstrass' Approximation Theorem)**.** *If* f *is a continuous function on [a,b], then there exists a sequence of polynomials* p_n, *such that* $p_n \xrightarrow{u} f$ *on [a,b].*

Proof. Suppose we prove the theorem for [0,1] and f is continuous on [a,b]. Setting $\tilde{f}(t) = f(a+(b-a)t)$, we obtain a continuous function, which we can approximate uniformly by polynomials $\tilde{p}_n$ on [0,1]. Setting $p_n(x) = \tilde{p}_n(\frac{x-a}{b-a})$ we obtain a sequence of polynomials that approximate f uniformly. Thus, it will be sufficient to consider the case of [0,1].

If we let $g(x) = f(x) - f(0) - xf(1)$, then $g(0) = g(1) = 0$ and f is equal to g plus a polynomial. Thus, if we can approximate g by a sequence of polynomials, then we can approximate f by a sequence of polynomials. Hence, it will be enough to prove the theorem under the extra assumption that $f(0) = f(1) = 0$.

Since $f(0 = f(1) = 0$, if we set $f(x) = 0$ for $x < 0$ and for $x > 1$, then this extends f to a continuous function on the whole real line.

For $0 \le x \le 1$, set

$$p_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_0^1 f(s)Q_n(s-x)ds,$$

where $s = x + t$, which shows that p_n is a polynomial. Since f is uniformly continuous on [0,1], given $\epsilon > 0$, we may choose $\delta > 0$, so that $|x - y| < \delta$, $0 \le x, y \le 1$, implies that $|f(x) - f(y)| < \epsilon/2$. Because f is identically 0 outside of this interval, this same ϵ-δ relation holds for any $x, y \in \mathbb{R}$.

Let $M = \sup\{|f(t)| : 0 \le t \le 1\}$ and recall that $Q_n(x) \ge 0$. Using the extended values of f we have that $p_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_{-1}^{+1} f(x+t)Q_n(t)dt$. Hence,

$$|p_n(x)-f(x)| = |\int_{-1}^{+1} (f(x+t)-f(x))Q_n(t)dt| \le \int_{-1}^{+1} |f(x+t)-f(x)|Q_n(t)dt \le$$
$$2M\int_{-1}^{-\delta} Q_n(t)dt + \int_{-\delta}^{+\delta} \epsilon/2Q_n(t)dt + 2M\int_{+\delta}^{1} Q_n(t)dt.$$

The middle of these three integrals is less than $\epsilon/2$ and since Q_n tends uniformly to 0 on the intervals of the first and third integrals, we may pick N so that for $n > N$, the other two integrals are at most $\epsilon/(4M)$. Thus, for $n > N$, we will have that $|p_n(x) - f(x)| < \epsilon$ for all $0 \le x \le 1$. □

Corollary 1.77. *For any $A > 0$, there is a sequence of polynomials $\{p_n\}$ with $p_n(0) = 0$ that converge uniformly on $[-A, +A]$ to $|x|$.*

Proof. By Weierstrass' theorem there is a sequence of polynomials q_n such that $q_n \xrightarrow{u} f$ on $[-A, +A]$ where $f(x) = |x|$. Then we have that $q_n(0) \to f(0) = 0$. Let $p_n(t) = q_n(t) - q_n(0)$, so that $p_n(0) = 0$.

Given $\epsilon > 0$, there is an N such that $n > N$ implies $|q_n(t) - f(t)| < \epsilon/2$ for all $-A \le t \le +A$. Then we have that $|p_n(t) - f(t)| = |q_n(t) - q_n(0) - f(t) + f(0)| \le |q_n(t) - f(t)| + |q_n(0) - f(0)| < \epsilon/2 + \epsilon/2 = \epsilon$ for all $n > N$ and all $-A \le t \le +A$. Hence, $p_n \xrightarrow{textu}$. □

The above proof is an example of a "convolution" proof. Given continuous functions f, g on [-1,1], we can set

$$f * g(x) = \int_0^1 f(s)g(s-x)ds.$$

This behaves like a product in the sense that $(f_1 + f_2) * g = f_1 * g + f_2 * g$, $f*(g_1+g_2) = f*g_1+f*g_2$ and for any constant c, $(cf)*g = f*(cg) = c(f*g)$. In the above proof we showed that,

$$f * Q_n \xrightarrow{u} f.$$

Thus, the functions Q_n behave "approximately" as an identity for the "convolution product". The concepts of convolution and approximate identities, play an important role in signal processing and harmonic analysis.

There is another proof of the Weierstrass approximation theorem due to Bernstein that gives the approximating polynomials by "sampling" the function f. Although we will not prove it here, we wish to state the theorem. This type of "sampling theorem" is also important in signal processing and in the area known as "approximation theory".

Definition 1.78. *Given a function $f : [0,1] \to \mathbb{R}$ then the* **n-th Bernstein polynomial of f** *is the polynomial,*

$$B_n(x) = \sum_{k=0}^{n} f(\frac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k}.$$

Theorem 1.79 (Bernstein's Approximation Theorem)**.** *If $f : [0,1] \to \mathbb{R}$ is continuous, then $B_n \xrightarrow{u} f$.*

Problem 1.80. *Let f be continuous on [a,b] and let $M = \sup\{|f(t)| : a \leq t \leq b\}$. Prove that there exists a sequence of polynomials $\{p_n\}$ that converges uniformly to f and satisfy $\sup\{|p_n(t)| : a \leq t \leq b\} \leq M$, for all n.*

Problem 1.81. *Let f be continuous on [a,b] and let c be a point, $a \leq c \leq b$. Prove that there exists a sequence of polynomials $\{p_n\}$ satisfying $p_n(c) = f(c)$ that converges uniformly to f on [a,b].*

Problem 1.82. *Use induction and the previous problem to prove that if f is continuous on [a,b] and $\{x_1, ..., x_K\}$ is a finite collection of points in [a,b], then there exists a sequence of polynomials $\{p_n\}$ satisfying $p_n(x_i) = f(x_i)$ for $1 \leq i \leq K$ that converges uniformly to f.*

1.11 The Stone-Weierstrass Theorem

The Stone-Weierstrass theorem is a generalization of Weierstrass' theorem that applies to more general settings. We shall actually need this general form later for the study of Fourier series. first some definitions.

Definition 1.83. *Let E be a set and let $\mathcal{A}$ be a collection of functions from E to $\mathbb{R}$. We call $\mathcal{A}$ an* **algebra of functions** *provided that*

- $f, g \in \mathcal{A}$, *implies that* $f + g \in \mathcal{A}$,
- $f, g \in \mathcal{A}$, *implies that* $fg \in \mathcal{A}$,
- $f \in \mathcal{A}$ *and* $r \in \mathbb{R}$ *implies that* $rf \in \mathcal{A}$.

Notice that the first and third properties are the defining properties of a vector space. Some examples of algebras of functions that we know are the contnuous real-valued functions on a metric space, the Riemann integrable functions on an interval.

Also, the set of polynomials satisfies the above three properties, so they may be regarded as an algebra of functions on any subset $E \subseteq \mathbb{R}$.

Definition 1.84. *A collection $\mathcal{A}$ of real-valued functions on a set E is said to* **separate points,** *provided that given any $x_1, x_2 \in E$, with $x_1 \neq x_2$, there is $f \in \mathcal{A}$ with $f(x_1) \neq f(x_2)$. the collection is said to* **vanish at no point,** *provided that for each $x_1 \in E$, there is $f \in \mathcal{A}$ with $f(x_1) \neq 0$.*

Proposition 1.85. *Let (X, d) be a metric space and let $C(X)$ denote the set of continuous functions from X to $\mathbb{R}$. Then $C(X)$ is an algebra of functions that separates points and vanishes at no point.*

Proof. We have already noticed that it is an algebra of function, since sums and products of continuous functions are continuous. Also, recall that if we fix any point $x_1 \in X$, then we proved last semester that $f(x) = d(x, x_1)$ is continuous. Hence, if $x_2 \neq x_1$, then $f(x_2) \neq 0$ while $f(x_1) = 0$. Thus, $f(x_1) \neq f(x_2)$, and so $C(X)$ separates points. Finally, if we set $g(x) = d(x, x_1) + 1$, then $g(x_1) \neq 0$. So $C(X)$ vanishes at no point. □

Also since every continuous function on [a,b] is Riemann integrable on [a,b], the algebra of Riemann integrable functions also separates points and vanishes at no point.

Finally, by just considering first degree polynomials, we can see that the algebra of polynomials separates points and vanishes at no point.

For some examples that fail these properties. Suppose that we consider the set of polynomials in x^2, then $p(-a) = p(+a)$ and so thes polynomials do not separate points when viewed as functions on $[-1. + 1]$, but they vanish at no point. Also if we consider the polynomials with constant term equal to 0, then these are an algebra of functions on $[-1, +1]$, they separate points on $[-1, +1]$, but every function in this set vanishes at 0.

Here is the main theorem:

Theorem 1.86 (Stone-Weierstrass Theorem)**.** *Let (K, d) be a compact metric space and let $\mathcal{A}$ be an algebra of continuous real valued functions on K. If $\mathcal{A}$ separates points on K and vanishes at no point, then given any $f \in C(K)$ there is a sequence of functions $\{f_n\}$ in $\mathcal{A}$, such that $f_n \xrightarrow{u} f$.*

When every function in $C(K)$ is the uniform limit of a sequence of functions in $\mathcal{A}$, then we say that **$\mathcal{A}$ is uniformly dense in $C(K)$**. Thus, the Stone-Weierstrass Theorem is often summarized as saying that if an algebra of real-valued functions on a compact metric space, separates points and vanishes at no point, then it is uniformly dense in $C(K)$.

This theorem is also true for any compact, Hausdorff topological space(a concept that we haven't yet encountered). There is also a version for complex-valued functions. The only extra condition that is needed in the complex case is that if $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$.

To prove this result, it will be convenient to first prove some lemmas.

Lemma 1.87. *Let $\mathcal{A}$ be an algebra of functions on E that separates points and vanishes at no point. If $x_1 \neq x_2$ are points in E and $c_1, c_2 \in \mathbb{R}$, then there exists $f \in \mathcal{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.*

Proof. Since $\mathcal{A}$ separates points there exists $g \in \mathcal{A}$ with $g(x_1) \neq g(x_2)$. Since $\mathcal{A}$ vanishes at no point, there exists $h_i \in \mathcal{A}, i = 1, 2$, with $h_i(x_i) \neq 0, i = 1, 2$.

Set

$$f(x) = c_1 \frac{g(x) - g(x_2)}{g(x_1) - g(x_2)} \frac{h_1(x)}{h_1(x_1)} + c_2 \frac{g(x) - g(x_1)}{g(x_2) - g(x_1)} \frac{h_2(x)}{h_2(x_2)},$$

then $f \in \mathcal{A}$ and $f(x_i) = c_i, i = 1, 2$. □

Lemma 1.88. *Let $\mathcal{A}$ be an algebra of functions on a set E and let $\mathcal{B}$ be the set of functions that are uniform limits of functions in $\mathcal{A}$. Then $\mathcal{B}$ is an algebra of functions on E and if $\{g_n\}$ is a sequence in $\mathcal{B}$ and $g_n \xrightarrow{u} g$, then $g \in \mathcal{B}$.*

Proof. If $f, g \in \mathcal{B}$ and $r \in \mathbb{R}$, then there are sequences $\{f_n\}$ and $\{g_n\}$ in $\mathcal{A}$ with $f_n \stackrel{u}{\to} f$ and $g_n \stackrel{u}{\to} g$. We have that $f_n + g_n \stackrel{f}{\to} +g$, so that $f + g \in \mathcal{B}$. Similarly, $f_n g_N \stackrel{u}{\to} fg$ and $rf_n \stackrel{u}{\to} rf$, so $fg \in \mathcal{B}$ and $rf \in \mathcal{B}$. Hence $\mathcal{B}$ is an algebra.

Finally, if $g_n \in \mathcal{B}$ and $g_n \stackrel{u}{\to} g$, then we may pick $f_n \in \mathcal{A}$, with $\sup\{|g_n(t) - f_n(t)| : t \in E\} < 1/n$. It then follows that $f_n \stackrel{u}{\to} h$ and so $g \in \mathcal{B}$. □

We are now prepared to prove the Stone-Weierstrass Theorem.

Proof. For $\mathcal{B}$ as defined in the last Lemma, we need to prove that $\mathcal{B} = C(K)$.

Let $f \in \mathcal{A}$. Since $f : K \to \mathbb{R}$ is continuous and K is compact, there is an $A > 0$, so that $-A \le f(x) \le +A$. By Corollary 1.66 there is a sequence of polynomials $\{p_n\}$ so that $p_n \stackrel{u}{\to} |t|$ on $[-A, +A]$ and $p_n(0) = 0$. Since these polynomials have no constant term and since $\mathcal{A}$ is an algebra, it follows that $p_n(f) \in \mathcal{A}$.

Also for $x \in K$, we have $|p_n(f(x)) - |f(x)|| \le \sup\{|p_n(t) - |t|| : -A \le t \le +A\}$. Hence, $p_n(f) \stackrel{u}{\to} |f|$ and we have that $|f| \in \mathcal{B}$. By the above lemma, if $f \in \mathcal{B}$, then $|f| \in \mathcal{B}$.

Given any two functions f, g on X, we let $min\{f, g\}$ be the function defined by $\min\{f, g\}(x) = \min\{f(x), g(x)\}$. We define $\max\{f, g\}$ similarly. Since for any two real numbers, $\min\{a, b\} = \frac{a+b}{2} - \frac{|a-b|}{2}$ and $\max\{a, b\} = \frac{a+b}{2} + \frac{|a-b|}{2}$, we see that for $f, g \in \mathcal{B}$,

$$\max\{f, g\} = \frac{f+g}{2} + \frac{|f-g|}{2} \in \mathcal{B} \text{ and } \min\{f, g\} = \frac{f+g}{2} - \frac{|f-g|}{2} \in \mathcal{B}.$$

Note that if $f_1, f_2, f_3 \in \mathcal{B}$, then $\max\{f_1, f_2, f_3\} = \max\{f_1, \max\{f_2, f_3\}\} \in \mathcal{B}$. Thus, by induction if $f_1, ..., f_n \in \mathcal{B}$, then $\max\{f_1, ..., f_n\} \in \mathcal{B}$, and similarly, $\min\{f_1, ..., f_n\} \in \mathcal{B}$.

Given $f \in C(K)$, a point $x \in K$ and $\epsilon > 0$, we claim that there is a function $g_x \in \mathcal{B}$, with $g_x(x) = f(x)$ and $g_x(t) > f(t) - \epsilon$.

To see why this is true, first given $y \ne x$, by the first Lemma, we can find $h_y \in \mathcal{A}$ with $h_y(x) = f(x)$ and $h_y(y) = f(y)$. Since $h_y - f$ is continuous and $(h_y - f)(y) = 0$, we may find a small ball of radius say $\delta_y > 0$, so that when $d(t, y) < \delta_y$, then $|h_y(t) - f(t)| < \epsilon$, which implies that $h_y(t) > f(t) - \epsilon$ for $t \in B(y; \delta_y)$. Now the set of all balls $\{B(y; \delta_y) : y \in K\}$ is an open cover of K so we may pick a finite subcover, i.e., $K = \cup_{i=1}^n B(y_i; \delta_{y_i})$. Since each $h_{y_i} \in \mathcal{A} \subseteq \mathcal{B}$, we have that $g_x = \max\{h_{y_1}, ... h_{y_n}\} \in \mathcal{B}$. But given any $t \in K$, there is an i so that $t \in B(y_i; \delta_{y_i})$, and so $g_x(t) \ge h_{y_i}(t) > f(t) - \epsilon$.

Now in a similar way, since each $g_x(x) = f(x)$, we may choose $r_x > 0$, so that $d(x, t) < r_x$ implies that $|g_x(t) - f(t)| < \epsilon$ and so $g_x(t) < f(t) + \epsilon$.

Again $\{B(x; r_x) : x \in K\}$ is an open cover of K and so we may choose a finite subcover, $K = \cup_{j=1}^m B(x_j; r_{x_j})$.

Set $h = \min\{g_{x_1}, ..., g_{x_n}\} \in \mathcal{B}$. Given any $t \in K$, there is j so that $t \in B(x_j, r_{x_j})$ and hence, $h(t) \leq g_{x_j}(t) < f(t) + \epsilon$. On the other hand, there must be an l, so that $h(t) = \min\{g_{x_1}(t), ..., g_{x_n}(t)\} = g_{x_l}(t) > f(t) - \epsilon$.

These two inequalities show that we have produced a $h \in \mathcal{B}$, with $|h(t) - f(t)| < \epsilon$ for all $t \in K$. Taking $\epsilon = 1/n$, and the corresponding $h_n \in \mathcal{B}$, gives us a sequence $\{h_n\}$ in $\mathcal{B}$ so that $h_n \overset{u}{\to} f$. By the second lemma, we can instead choose $f_n \in \mathcal{A}$, with $f_n \overset{u}{\to} f$. This completes the proof of the theorem. □

Recall that when we are given a polynomial $p(x_1, ..., x_n)$ in n variables, then we regard p as a function on $\mathbb{R}^n$.

Problem 1.89. *Let $K \subseteq \mathbb{R}^n$ be a compact set and let $\mathcal{A}$ be the algebra of polynomials in n variables, regarded as functions on K. Prove that $\mathcal{A}$ is uniformly dense in $C(K)$.*

Problem 1.90. *Let $\mathcal{A}$ be an algebra of functions on E that separates points and vanishes at no point. Let $x_1, ..., x_n$ be n distinct points in E and let $c_1, ..., c_n$ be real numbers. Prove that there is a function in $f \in \mathcal{A}$ such that $f(x_i) = c_i$, $i = 1, ..., n$.*

Problem 1.91. *Let $T = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, a compact set and let $\phi : [-\pi, +\pi] \to T$ be the continuous function defined by $\phi(t) = (cos(t), sin(t))$. If $h \in C(T)$ and $f(t) = h \circ \phi(t)$, then $f \in c([-\pi, +\pi])$ and $f(-\pi) = f(+\pi)$. Prove the converse, that is, prove that if $f \in C([-\pi, +\pi])$, with $f(-\pi) = f(+\pi)$, then there exists a continuous function $h \in C(T)$, with $f = h \circ \phi$.[HINT: The proof of this problem does not use the Stone-Weierstrass theorem.]*

1.12 Fourier Series

In addition to the Taylor series, another famous way to approximate functions is by means of their Fourier series. These series play an important role in signal processing and in many applications to finding solutions of differential equations.

Definition 1.92. *A* **trigonometric polynomial** *is any finite sum of the form $f(x) = a_0 + \sum_{n=1}^{N}(a_n cos(nx) + b_n sin(nx))$.*

It is convenient to write a_0 for the constant coefficient since $cos(0x) = 1$. Also, since $sin(0x) = 0$, we could write $f(x) = \sum_{n=0}^{N}(a_n cos(nx)+b_n sin(nx))$ and nothing would be changed. This sometimes makes it easier in formulas, since we can avoid the "special case" of $n = 0$.

The idea of Fourier series is to try and represent more general functions as "trigonometric series" that is as limits of infinite sums of trigonometric polynomials. The first question that we shall address, is to determine what the coefficients should be if one wants to represent a function as such a series.

If one uses the trigonometric identities,

$cos(a)cos(b) = 1/2[cos(a-b)+cos(a+b)]$ and $sin(a)sin(b) = 1/2[cos(a-b)-cos(a+b)]$,

then one obtains the following integral identities for integers m and n:

- $\int_{-\pi}^{+\pi} cos(nx)sin(mx)dx = 0$, for all m, n,
- $\int_{-\pi}^{+\pi} cos(nx)cos(mx)dx = \begin{cases} 0, & m \neq n \\ \pi, & m = n, \end{cases}$
- $\int_{-\pi}^{+\pi} sin(nx)sin(mx)dx = \begin{cases} 0, & m \neq n \\ \pi, & m = n \end{cases}.$

Because of these formulas, if we have a trig polynomial f as in the above Definition, then we can recover the constants by integration, namely,

- $a_0 = 1/2\pi \int_{-\pi}^{+\pi} f(x)1dx$,
- $a_n = 1/\pi \int_{-\pi}^{+\pi} f(x)cos(nx)dx$, for $1 \leq n \leq N$,
- $b_n = 1/\pi \int_{-\pi}^{+\pi} f(x)sin(nx)dx$, for $1 \leq n \leq N$.

Moreover, if $n > N$, then each of these integrals would be 0.

Thus, if we "imagine" a more general function f corresponding to an infinite series of cosine and sine functions, then these formulas tell us the values of the coefficients of that series. This leads to the following definition.

Definition 1.93. *Let $f : [-\pi, +\pi] \to \mathbb{R}$ be a Riemann integrable function. Then we set*

- $a_0 = 1/2\pi \int_{-\pi}^{+\pi} f(x)1dx$,
- $a_n = 1/\pi \int_{-\pi}^{+\pi} f(x)cos(nx)dx$, *for* $n \in \mathbb{N}$,

- $b_n = 1/\pi \int_{-\pi}^{+\pi} f(x)sin(nx)dx$, *for* $n \in \mathbb{N}$,

and we call these numbers the **Fourier coefficients of** f. *The infinite series,*

$$a_0 + \sum_{n=1}^{+\infty}(a_n cos(nx) + b_n sin(nx))$$

is called the **Fourier series for** f. *We set*

$$s_N(f,x) = a_0 + \sum_{n=0}^{N}(a_n cos(nx) + b_n sin(nx),$$

and call this the N*-th partial sum of the series.*

Note that each of the functions in the above formulas is Riemann integrable, since the product of Riemann integrable functions is again Riemann integrable.

One problem that we wish to discuss is to what extent the Fourier series converges to the function f. Note that if we start with a Riemann integrable function f and we change its value at a single point x_0 to get a new function g, then g will still be Riemann integrable and all the Fourier coefficients of g will be equal to those of f. Hence, $s_N(f,x) = s_N(g,x)$ for all N and x. So if the Fourier series does converge to a value at the point x_0, then there is no way that it could give both $f(x_0)$ and $g(x_0)$. Thus, in general, we cannot expect the Fourier series for f to converge pointwise to f.

Problem 1.94. *Let* $h(x) = \begin{cases} 0, & -\pi \le x \le 0 \\ 1, & 0 < x \le +\pi \end{cases}$. *Compute the Fourier coefficients for* h *and* $s_N(h,x)$.

Problem 1.95. *Let* $f(x) = |x|$. *Compute the Fourier coefficients for* f *and* $s_N(f,x)$.

1.13 Orthonormal Sets of Functions

Many of the results that we shall need about trigonometric polynomials are true because they are an orthonormal set of functions. So we first prove the results in this more general setting.

Definition 1.96. *A set of* $\mathcal{I}$ *of Riemann integrable functions on [a,b] is called an* **orthogonal family** *provided that for any* $f, g \in \mathcal{I}$, *with* $f \neq g$,

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

If, in addition, $\int_a^b f(x)^2 dx = 1$ *for every* $f \in \mathcal{I}$*, then we call the set an* **orthonormal family.**

By the formulas in the last section, we see that

$$\mathcal{I} = \{1, cos(nx), sin(nx) : n \in \mathbb{N}\}$$

is an orthogonal set of functions on $[-\pi, +\pi]$ and that

$$\mathcal{I} = \{\frac{1}{\sqrt{2\pi}}, \frac{cos(nx)}{\sqrt{\pi}}, \frac{sin(nx)}{\sqrt{\pi}} : n \in \mathbb{N}\}$$

is an orthonormal set of functions on $[-\pi, +\pi]$.

We will see that there are some advantages to working with orthonormal sets of functions.

Definition 1.97. *Let* f *be a Riemann integrable function on [a,b] and let* $\mathcal{I} = \{\phi_\alpha\}_{\alpha \in A}$ *where* A *is some index set be an orthonormal set of functions, then we call*

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

the **(generalized) Fourier coefficients of f.**

If $f(x) = a_0 + \sum_{n=1}^N (a_n cos(nx) + b_n sin(nx))$, then when we use the orthonormal set of functions, $\mathcal{I} = \{\frac{1}{\sqrt{2\pi}}, \frac{cos(nx)}{\sqrt{\pi}}, \frac{sin(nx)}{\sqrt{\pi}} : n \in \mathbb{N}\}$, we see that the generalized Fourier coefficients of f with respect to this set are $\hat{a_0} = \int_{-\pi}^{+\pi} f(x)\frac{1}{\sqrt{2\pi}} dx = a_0\sqrt{2\pi}$, and similarly, $\hat{a_n} = a_n\sqrt{\pi}$, and $\hat{b_n} = b_n\sqrt{\pi}$. Thus,

$$s_N(f;x) = a_0 + \sum_{n=1}^N (a_n cos(nx) + b_n sin(nx)) = \hat{a_0}(\frac{1}{\sqrt{2\pi}}) + \sum_{n=1}^N (\hat{a_n}(\frac{cos(nx)}{\sqrt{\pi}}) + \hat{b_n}(\frac{sin(nx)}{\sqrt{\pi}})),$$

and we see that the partial sums of the Fourier series for f are exactly the sums that one obtains by summing the generalized Fourier coefficients with respect to this orthonormal family.

We will see that orthonormal sets make the problem of finding the best aproximation to a function in something called the L^2-norm easy.

Definition 1.98. *Let* f *be a Riemann integrable function on [a,b], then the* L^2*-norm of* f *is the number,*

$$\|f\| = \sqrt{\int_a^b f(x)^2 dx}.$$

Note that many functions have an L^2-norm of 0. For example, any function that is non-zero at only finitely many points. Exactly which functions have L^2-norm of 0 is examined more closely in courses on measure theory.

Theorem 1.99 (Theorem on Best Approximation)**.** *Let f be a Riemann integrable function on [a,b], let $\{\phi_1, ..., \phi_N\}$ be a finite set of orthonormal functions on [a,b], let $c_n = \int_a^b f(x)\phi_n(x)dx$ be the generalized Fourier coefficients of f, set $s_N(x) = \sum_{n=1}^N c_n\phi_n(x)$, let $b_1, ..., b_N$ be arbitrary real numbers and set $t_N(x) = \sum_{n=1}^N b_n\phi_n(x)$. Then*

- $\|f - s_N\| \leq \|f - t_N\|$,
- $\|f - s_N\| = \|f - t_N\|$ *if and only if* $c_j = b_j$ *for all* $1 \leq j \leq N$.

In summary, this theorem says that if one takes the vector space spanned by the functions $\{\phi_1, ..., \phi_N\}$ then among all functions in that vector space, the function given by using the generalized Fourier coefficients is the function that is closest to f in the L^2-norm and it is the unique function in that subspace that is closest to f. Applying this theorem to the Fourier series, e see that $S_n(f; x)$ is the unique function in the$(2N + 1)$-dimensional vector space spanned by $\{1, cos(nx), sin(nx) : 1 \leq n \leq N\}$ that is closest to f in the L^2-norm.

Before proving the theorem, we prove some useful formulas.

Lemma 1.100. *Let $\{\phi_1, ..., \phi_N\}$ be an orthonormal set of functions on [a,b], and let $t_N = \sum_{n=1}^N b_n\phi_n$, then*

$$\|t_N\|^2 = \sum_{n=1}^N b_n^2.$$

Proof. We have that

$$\|t_N\|^2 = \int_a^b t_N(x)^2 = \sum_{i,j=1}^N \int_a^b b_i b_j \phi_i(x)\phi_j(x)dx = \sum_{j=1}^N b_j^2.$$

□

We are now ready to prove the theorem.

Proof. For any real numbers $b_1, ..., b_N$ we have that

$$\|f - t_N\|^2 = \int_a^b (f(x)^2 - 2f(x)t_N(x) + t_N(x)^2)dx =$$

$$\|f\|^2 - 2\sum_{n=1}^{N}\int_a^b f(x)b_n\phi_n(x)dx + \sum_{n=1}^{N} b_n^2 =$$

$$\|f\|^2 - 2\sum_{n=1}^{N} c_n b_n + \sum_{n=1}^{N} b_n^2 = \|f\|^2 - \sum_{n=1}^{N} c_n^2 + \sum_{n=1}^{N}(c_n^2 - 2c_n b_n + b_n^2) =$$

$$\|f\|^2 - \sum_{n=1}^{N} c_n^2 + \sum_{n=1}^{N}(c_n - b_n)^2.$$

Thus, when $b_n = c_n$ for all n,

$$\|f - s_N\|^2 = \|f\|^2 - \sum_{n=1}^{N} c_n^2,$$

and for any $t_N \neq s_N$,

$$\|f - t_N\|^2 = \|f - s_N\|^2 + \sum_{n=1}^{N}(c_n - b_n)^2 > \|f - s_N\|^2.$$

□

This theorem has two important consequences.

Corollary 1.101 (Bessel's Inequality)**.** *Let $\{\phi_n : n \in \mathbb{N}\}$ be an orthonormal set of functions on [a,b], let f be Riemann integrable on [a,b] and let $c_n -$ $\int_a^b f(x)\phi_n(x)dx$ be the generalized Fourier coefficients of f. Then*

$$\sum_{n=1}^{+\infty} |c_n|^2 \leq \|f\|^2.$$

In particular, the series $\sum_{n=1}^{+\infty} |c_n|^2$ of squares of Fourier coefficients is always summable for every Riemann integrable function.

Proof. In the proof of the above theorem, we saw that

$$\sum_{n=1}^{N} c_n^2 = \|f\|^2 - \|f - s_N\|^2 \leq \|f\|^2.$$

□

Corollary 1.102 (Parseval's Equation)**.** *Let $\{\phi_n : n \in \mathbb{N}\}$ be an orthonormal set of functions on [a,b], let f be Riemann integrable on [a,b], let $c_n - \int_a^b f(x)\phi_n(x)dx$ be the generalized Fourier coefficients of f and let $s_N(x) = \sum_{n=1}^N c_n\phi_n(x)$. Then* $\lim_{N\to+\infty} \|f - s_N\| = 0$ *if and only if* $\sum_{n=1}^{+\infty} c_n^2 = \|f\|^2$.

Proof. In the proof of the above theorem we saw that

$$\|f - s_N\|^2 = \|f\|^2 - \sum_{n=1}^{N} c_n^2.$$

Thus, the left hand side has limit zero if and only if the right hand side has limit zero. □

The statement, $\|f\|^2 = \sum_{n=1}^{+\infty} c_n^2$ is generally referred to as **Parseval's identity** and the above corollary is often summarized as saying that Parseval's identity holds if and only if $\lim_{N\to+\infty} \|f - s_N\| = 0$.

1.14 Fourier Series, Continued

We now wish to prove that Parseval's identity holds for the Fourier series. We will first need some preliminary theorems. These theorems are both important approximation results. The first tells us that in the L^2-norm, Riemann integrable functins can be approximated by continuous functions. The second says that 2π-periodic continuous functions can be approximated uniformly by trigonometric polynomials.

Theorem 1.103. *Let f be Riemann integrable on $[a, b]$, then for every $\epsilon > 0$ there exists a continuous function g such that*

$$\|f - g\|_2 = (\int_a^b |f(t) - g(t)|^2 dt)^{1/2} < \epsilon.$$

Proof. Let $M = \sup\{|f(t)| : a \le t \le b\}$ and choose a partition, $\mathcal{P} = \{a = x_0 < ... < x_n = b\}$ such that

$$U(f;\mathcal{P}) - L(f;\mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\epsilon^2}{2M},$$

where as usual $m_i = \inf\{f(t) : x_{i-1} \le t \le x_i\}$ and $M_i = \sup\{f(t) : x_{i-1} \le t \le x_i\}$.

Define a function $g : [a, b] \to \mathbb{R}$ by setting

$$g(t) = \frac{x_i - t}{x_i - x_{i-1}} f(x_{i-1}) + \frac{t - x_{i-1}}{x_i - x_{i-1}} f(x_i), \text{ for } x_{i-1} \leq t \leq x_i,$$

so that g is linear on each subinterval $[x_{i-1}, x_i]$ and $g(x_i) = f(x_i)$ for each i. Since the formulas for g agree at the endpoint of each subinterval, g is continuous on $[a, b]$.

Also, since on each subinterval g is linear and agrees with f at the endpoints, we have that $m_i \leq \min\{f(x_{i-1}), f(x_i)\} \leq g(t) \leq \max\{f(x_{i-1}), f(x_i)\} \leq M_i$, for $x_{i-1} \leq t \leq x_i$. This implies that $|f(t) - g(t)| \leq M_i - m_i$, for $x_{i-1} \leq t \leq x_i$.

Hence,

$$\begin{aligned}\|f - g\|_2^2 = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t) - g(t)|^2 dt &\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (M_i - m_i)^2 dt \\ &\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} 2M(M_i - m_i) dt = 2M(U(f : \mathcal{P}) - L(f; \mathcal{P})) < \epsilon^2\end{aligned}$$

and the result follows. □

Before our next result we need a lemma.

Lemma 1.104. *Every function of the form $cos^n(t), sin^m(t)$, and $cos^k(t)sin^j(t)$, for $n, m, k, j \in \mathbb{N}$ is a trigonometric polynomial.*

Proof. Using the trigonometric identities, $cos(at)cos(bt) = 1/2[cos(at + bt) + cos(at - bt)]$, $sin(at)sin(bt) = 1/2[cos(at - bt) - cos(at + bt)]$ and $cos(at)sin(bt) = 1/2[sin(a + b) - sin(a - b)]$ we can iteratively reduce products of trig functions into sums that are one degree lower. □

Theorem 1.105. *Let $f : [-\pi, +\pi] \to \mathbb{R}$ be continuous with $f(-\pi) = f(+\pi)$ and let $\epsilon > 0$. Then there is a trigonometric polynomial g, such that $|f(t) - g(t)| < \epsilon$ for $-\pi \leq t \leq +\pi$.*

Proof. Let $T = \{(x, y) : x^2 + y^2 = 1\}$. By Problem 1.84 There is a continuous function $h : T \to \mathbb{R}$ such that $f(t) = h(cos(t), sin(t))$. Since the polynomials in x and y are an algebra of functions on T that separate points and vanish at no point and since T is compact, by the Stone-Weierstrass theorem there is a polynomial $p(x, y)$ such that $|h(x, y) - p(x, y)| < \epsilon$ for every $(x, y) \in T$.

Let $g(t) = p(cos(t), sin(t))$. By the Lemma, g is a trigonometric polynomial and

$$\sup\{|f(t) - g(t)| : -\pi \leq t \leq +\pi\} \leq \sup\{|h(x, y) - p(x, y)| : x^2 + y^2 = 1\}.$$

□

Two important inequalities.

Proposition 1.106. *Let f, g be Riemann integrable on $[a, b]$ then*

1. *Cauchy-Schwarz inequality,* $|\int_a^b f(t)g(t)dt| \leq \|f\|_2\|g\|_2$,
2. *Minkowski inequality,* $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$.

Proof. First check that when $a, c \geq 0$ then the polynomial $at^2 + bt + c \geq 0$ for all t if and only if it only has one or zero roots. This is because its graph is a parabola that opens upwards. But it has one or zero roots if and only if $b^2 - 4ac \leq 0$.

Now since

$$0 \leq \int_a^b (tf(x) + g(x))^2 dx = \|f\|_2^2 t^2 + 2t \int_a^b f(x)g(x)dx + \|g\|_2^2,$$

we have that

$$4(\int_a^b f(x)g(x)dx)^2 \leq 4\|f\|_2^2\|g\|_2^2$$

and Schwarz's inequality follows by taking square roots.

To see Minkowski's inequality, we note that

$$\|f+g\|_2^2 = \int_a^b (f(x)+g(x))^2 dx = \int_a^b f(x)^2 dx + 2\int_a^b f(x)g(x)dx + \int_a^b g(x)^2 dx$$
$$\leq \|f\|_2^2 + 2\|f\|_2\|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2,$$

and the result follows by taking square roots. □

We leave the proof of the following lemma to the exercises.

Lemma 1.107. *Let g be a continuous function on $[a, b]$ and let $\epsilon > 0$ be given. Then there exists a continuous function g_1 on $[a, b]$ such $g_1(a) = g_1(b) = 0$ and $\|g - g_1\|_2 < \epsilon$.*

Now for the main theorem on Fourier series.

Theorem 1.108. *Let f be a Riemann integrable function on $[-\pi, +\pi]$ and let $s_N(f;x) = a_0 + \sum_{n=1}^{N}(a_n cos(nx) + b_n sin(nx))$ denote the partial sums of its Fourier series. Then:*

1. $\lim_{N\to+\infty} \|f - s_N\|_2 = 0$,
2. $\|f\|_2^2 = \int_{-\pi}^{+\pi} f(x)^2 = \pi[2a_0^2 + \sum_{n=1}^{+\infty}(a_n^2 + b_n^2)]$.

Proof. Recall that the functions $\{\frac{1}{\sqrt{2\pi}}, \frac{cos(nx)}{\sqrt{\pi}}, \frac{sin(nx)}{\sqrt{\pi}} : n \geq 1\}$ form an orthonormal family and by the Theorem on Best Approximation,

$$s_N(f;x) = (\sqrt{2\pi}a_0)(\frac{1}{\sqrt{2\pi}}) + \sum_{n=1}^{N}(\sqrt{\pi}a_n)(\frac{cos(nx)}{\sqrt{\pi}}) + (\sqrt{\pi}b_n)(\frac{sin(nx)}{\sqrt{\pi}}) =$$
$$\hat{a_0}(\frac{1}{\sqrt{2\pi}}) + \sum_{n=1}^{N}(\hat{a_n}\frac{cos(nx)}{\sqrt{\pi}} + \hat{b_n}\frac{sin(nx)}{\sqrt{\pi}}),$$

is the unique function in the span of these functions that is closest to f.

From this fact we see that $\|f - s_N\|_2 \geq \|f - s_{N+1}\|_2$ for all N, since as the vector space grows larger the distance to the nearest function must decrease.

Now we prove that the limit in (1) converges to zero. Given $\epsilon > 0$, by theorem 1.96 we may pick a continuous function g on $[-\pi, +\pi]$ so that $\|f - g\|_2 < \epsilon/3$. By Lemma 1.100, we may pick a continuous function g_1 with $g_1(-\pi) = g_1(+\pi) = 0$, such that $\|g - g_1\|_2 < \epsilon/3$. Now by Theorem 1.98, we can find a trigonometric polynomial p such that $|g_1(t) - p(t)| < \frac{\epsilon}{\sqrt{18\pi}}$.

This last inequality implies that

$$\|g_1 - p\|_2^2 = \int_{-\pi}^{+\pi}(g(t) - p(t))^2 dt < \int_{-\pi}^{+\pi}\frac{\epsilon^2}{18\pi} = \frac{\epsilon^2}{9},$$

so that $\|g_1 - p\|_2 < \epsilon/3$.

Now p is a trigonometric polynomial of some degree M and by Minkowski, we have that

$$\|f - p\|_2 = \|(f - g) + (g - g_1) + (g_1 - p)\|_2 < \epsilon.$$

By the Theorem on Best Approximation, $\|f - s_M\|_2 \leq \|f - p\|_2 < \epsilon$ and hence for any $N \geq M$, we have $\|f - s_N\| < \epsilon$ and so (1) follows.

Now since $\|f - s_N\|_2 \to 0$, by Parseval we have that

$$\|f\|_2^2 = \hat{a_0}^2 + \sum_{n=1}^{+\infty}(\hat{a_n}^2 + \hat{b_n}^2) = \pi(2a_0^2 + \sum_{n=1}^{+\infty}(a_n^2 + b_n^2),$$

using the identities relating the "hatted" coefficients to their hatless cousins.
□

Problem 1.109. *Prove Lemma 1.100.*

Problem 1.110. *Use the above theorem together with problems 1.87 and 1.88 to give some infinite series for* π.

Problem 1.111. *Prove that the set of trigonometric polynomials is an algebra of functions on* $[-\pi, +\pi]$ *that vanishes at no point. Also prove that if* $-\pi \leq x_1 < x_2 < +\pi$, *then there exists a trigonometric polynomial* p *with* $p(x_1) \neq p(x_2)$.

1.15 Equicontinuity and the Arzela-Ascoli Theorem

The Heine-Borel theorem tells us that every bounded sequence of numbers has a convergent subsequence and that every bounded sequence of vectors in $\mathbb{R}^k$ has a convergent subsequence. This theorem also gave us a characterization of compact sets in $\mathbb{R}^k$, namely, a set in $\mathbb{R}^k$ is compact if and only if it is closed and bounded.

In this section we consider the corresponding problem for sets of continuous real-valued functions whose domain is a compact set. The result is, formally, very similar to the Heine-Borel result with one extra ingredient, the concept of *equicontinuity.*

Definition 1.112. *Let* (X, d) *and* (Y, ρ) *be metric spaces. A set* $\mathcal{F}$ *of functions from* X *to* Y *is called* **equicontinuous** *provided that for every* $\epsilon > 0$, *there is* $\delta > 0$, *so that for every* $f \in \mathcal{F}$ *and every* $x_1, x_2 \in X$ *with* $d(x_1, x_2) < \delta$, *we have* $\rho(f(x_1), f(x_2) < \epsilon$.

Note that the definition implies that every function in $\mathcal{F}$ is uniformly continuous and that the same values of ϵ-δ work for every function in $\mathcal{F}$.

Here is an example of an equicontinuous family. Let $X = Y = \mathbb{R}$, fix a number M and let $\mathcal{F}$ denote the set of all differentiable functions on $\mathbb{R}$ with the property that $|f'(x)| \leq M$ for all x. Given any $\epsilon > 0$, if we set $\delta = \epsilon/M$, then when $|x_1 - x_2| < \delta$, by the Mean Vaule Theorem, we will have that $|f(x_1) - f(x_2)| = |f'(c)(x_1 - x_2)| \leq M|x_1 - x_2| < M\delta = \epsilon$.

Definition 1.113. *Let* E *be a set and let* $\mathcal{F}$ *be a set of real-valued functions on* E. *We say that* $\mathcal{F}$ *is* **pointwise bounded** *provided that for each* $x \in E$, *there is a constant* M_x *such that for every* $f \in \mathcal{F}$, $|f(x)| \leq M_x$.

The following result uses the idea of Cantor's diagonalization process that we encountered last semester.

Proposition 1.114. *Let $E = \{x_i\}_{i\in\mathbb{N}}$ be a countable set and let $\{f_n\}$ be a sequence of real-valued functions on E that is pointwise bounded. Then there is a subsequence $\{f_{n_k}\}$ such that for every $x \in E$, the sequence $\{f_{n_k}(x)\}$ converges.*

Proof. Recall that a subsequence of a convergent sequence still converges to the same point. Since $\{f_n\}$ is pointwise bounded, the sequence $\{f_n(x_1)\}$ of real numbers is bounded and so we may choose a subsequence so that $\{f_{n_k}(x_1)\}$ is convergent(as $k \to +\infty$). Now because the sequence(indexed by k), $\{f_{n_k}(x_2)\}$ is bounded we may choose a subsubsequence(indexed by say j) so that $\{f_{n_{k_j}}(x_2)\}$ converges as $j \to +\infty$. This subsubsequence, would have the property that it converges for the points x_1 and x_2.

Because the language of "subsubsequences" is a bit confusing, it is best to remember that a subsequence is an infinite subset $S_1 \subseteq \mathbb{N}$ and then we are listing the numbers, $S_1 = \{n_1 < n_2 < ...\}$ in their natural order. So when we speak about a subsubsequence we are just taking an infinite subset $S_2 \subseteq S_1$.

Thus, we see that we can define infinite sets $\mathbb{N} \supseteq S_1 \supseteq S_2 \supseteq ...$ so that the set $S_J = \{n_1 < n_2 < ...\}$ has the property that $\{f_{n_k}(x_j)\}$ converges for $j = 1, ..., J$. Now if the intersection $\cap S_j$ of all these sets was non-empty and infinite, then this would define a subsequence with the property that $\{f_{n_k}(x_j)\}$ converged for every j. The problem is that the intersection could be empty or finite.

However, using Cantor's diagonalization idea, we define a subsequence by keeping the k-th number in S_k. Note that this is a strictly increasing sequence of numbers. The sequence of numbers that we define in this manner, then has the property that $\{n_1, n_2, ...\} \subseteq S_1$, $\{n_2, n_3, ...\} \subseteq S_2$, and $\{n_k, n_{k+1}, ...\} \subseteq S_k$, for all k. Because the behavior of the first few terms of a sequence doesn't affect its convergence, we see that this subsequence has the property that $\{f_{n_k}(x_j)\}$ converges as $k \to +\infty$ for every j. □

We now can state the main thoerem of this section.

Theorem 1.115 (Arzela-Ascoli)**.** *Let (K, d) be a compact metric space and let $f_n : K \to \mathbb{R}$, $n \in \mathbb{N}$ be a sequence of continuous real-valued functions on K that is equicontinuous and pointwise bounded. Then there is a subsequence $\{f_{n_k}\}$ and a continuous function $f : K \to \mathbb{R}$ such that $f_{n_k} \xrightarrow{u} f$.*

Proof. Recall that every compact metric space has a countable dense set. Let $E = \{x_j\}$ be a countable dense set in K.

By the previous proposition, we may choose a subsequence $\{f_{n_k}\}$ such that $\lim_k f_{n_k}(x_j)$ exists for every j.

We claim that for every $x \in K$, the sequence $\{f_{n_k}(x)\}$ is Cauchy. To see this, given $\epsilon > 0$, by equicontinuity, we may pick $\delta > 0$, so that $d(x,y) < \delta$ implies for every n that $|f_n(x) - f_n(y)| < \epsilon/3$. Now pick j so that $d(x,x_j) < \delta$ and pick I so that $i,k > I$ implies that $|f_{n_k}(x_j) - f_{n_i}(x_j)| < \epsilon/3$. Then for $i,k > I$ we have that

$$|f_{n_k}(x) - f_{n_i}(x)| \le |f_{n_k}(x) - f_{n_k}(x_j)| + |f_{n_k}(x_j) - f_{n_i}(x_j)| + |f_{n_i}(x_j) - f_{n_i}(x)| < \epsilon.$$

Now since this sequence is Cauchy and $\mathbb{R}$ is complete, we may define a function $f : K \to \mathbb{R}$ by setting $f(x) = \lim_k f_{n_k}(x)$.

We now claim that $\{f_{n_k}\}$ converges uniformly to f. To see this, given $\epsilon > 0$, let δ be as before. Using the fact that K is compact, we may choose $\{x_1, ..., x_p\}$ so that $K = \cup_{j=1}^{p} B(x_j; \delta)$.

Now for each $j = 1, ..., p$ we may pick I_j so that $i, k > I_j$ implies that $|f_{n_k}(x_j) - f_{n_i}(x_j)| < \epsilon/3$. Let $I = \max\{I_1, ..., I_p\}$.

Then for $k > I$ and any $x \in K$, we have that $x \in B(x_j; \delta)$ for some value of j, $1 \le j \le p$ and hence,

$$|f(x) - f_{n_k}(x)| = \lim_i |f_{n_i}(x) - f_{n_k}(x)| \le$$
$$\limsup_i |f_{n_i}(x) - f_{n_i}(x_j)| + |f_{n_i}(x_j) - f_{n_k}(x_j)| \le \epsilon/3 + \epsilon/3 < \epsilon.$$

□

Corollary 1.116 (Arzela-Ascoli)**.** *Let (K,d) be a compact metric space and let $C(K)$ denote the set of continuous functions from K to $\mathbb{R}$ endowed with the uniform metric $\gamma(f,g) = \sup\{|f(x) - g(x)| : x \in K\}$. Then $\mathcal{F} \subseteq C(K)$ is a compact subset of $(C(K), \gamma)$ if and only if $\mathcal{F}$ is closed, equicontinuous and pointwise bounded.*

Proof. Recall that a sequence converges in the metric γ if and only if it converges uniformly. If $\mathcal{F}$ is closed, equicontinuous and pointwise bounded and $\{f_n\}$ is a sequence in $\mathcal{F}$, then by the Arzela-Ascoli theorem, there is a subsequence $\{f_{n_k}\}$ that converges uniformly, and hence in the metric γ to a continuous function f. Since $\mathcal{F}$ is closed, $f \in \mathcal{F}$. Thus, $\mathcal{F}$ is sequentially compact, which we have shown is the same as compact.

Conversely, assume that $\mathcal{F}$ is compact, then $\mathcal{F}$ is closed. Given $\epsilon > 0$, since $\mathcal{F}$ is compact we may choose finitely many functions in $\mathcal{F}$ such that

$\mathcal{F} \subseteq \cup_{n=1}^{N} B(f_n, \epsilon/3)$. Since K is compact, each function f_n is uniformly continuous and so there exists $\delta_n > 0$ so that $d(x,y) < \delta_n$ implies that $|f_n(x) - f_n(y)| < \epsilon/3$. Let $\delta = \min\{\delta_1, ..., \delta_n\} > 0$.

Given any $f \in \mathcal{F}$ there is a j so that $\gamma(f, f_j) < \epsilon/3$. Hence, if $d(x,y) < \delta$, then

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < \epsilon.$$

This proves that $\mathcal{F}$ is equicontinuous.

To see that $\mathcal{F}$ is pointwise bounded, let $\{f_1, ..., f_n\}$ be chosen as above for the value $\epsilon = 1$. Let $M_j = \sup\{|f_j(x)| : x \in K\}$ which is finite since K is compact and let $M = \max\{M_1, ..., M_n\}$. Given any $f \in \mathcal{F}$ there is j so that $\gamma(f, f_j) < 1$ and hence, $|f(x)| \leq |f(x) - f_j(x)| + |f_j(x)| \leq 1 + M_j \leq M + 1$. Thus, the functions in $\mathcal{F}$ are pointwise bounded and even uniformly bounded. □

Problem 1.117. *Let $\mathcal{F}$ be a set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ that is equicontinuous. Prove that if $\sup\{|f(0)| : f \in \mathcal{F}\} < +\infty$, then $\mathcal{F}$ is pointwise bounded.*

Problem 1.118. *Give an example of a metric space (X, d) and an equicontinuous set of continuous real-valued functions on X that is pointwise bounded at one point but not pointwise bounded.*

Chapter 2

Multivariable Differential Calculus

We shall always endow $\mathbb{R}^n$ with the Euclidean metric. Given a set $S \subseteq \mathbb{R}^n$ a function $\vec{f} : S \to \mathbb{R}^m$ has the form $\vec{f}(\vec{x}) = (f_1(\vec{x}), ..., f_m(\vec{x}))$ where we call the functions $f_j : S \to \mathbb{R}$ the **component functions.** In this chapter we study the theory of differentiablity of such functions. We begin with directional derivatives.

Definition 2.1. *Let $S \subseteq \mathbb{R}^n$, let $\vec{x_0} \in S$ be an interior point let $\vec{f} : S \to \mathbb{R}^m$ and let $\vec{u} = (u_1, ..., u_n) \in \mathbb{R}^n$. We say that $\vec{f}$ has a* **directional derivative in the direction $\vec{u}$ at $\vec{x_0}$** *provided that the following vector-valued limit exists,*

$$\lim_{t \to 0} \frac{\vec{f}(\vec{x_0} + t\vec{u}) - \vec{f}(\vec{x_0})}{t}.$$

When this limit exists, we denote it by $\vec{f'}_{\vec{u}}(\vec{x_0})$.

Many books make the requirement that $\vec{u}$ is of unit length. We don't make that requirement here.

Note that when $\vec{u} = \vec{0}$, then the numerator in the above limit is always $\vec{0}$, hence the limit always exists and

$$\vec{f'_{\vec{0}}}(\vec{x_0}) = \vec{0}.$$

For this reason we wil generally be interested in the case $\vec{u} \neq \vec{0}$.

Note that when $\vec{f} : S \to \mathbb{R}^m$, then provided that it exists, $\vec{f'_{\vec{u}}}(\vec{x_0}) \in \mathbb{R}^m$. Also, recall that the limit of a vector-valued function exists if and only if the limit of each component function exists. Thus, we see that $\vec{f'}_{\vec{u}}(\vec{x_0})$ exists if and only if for each j, $1 \leq j \leq m$, the real limit

$$b_j = \lim_{t \to 0} \frac{f_j(\vec{x_0} + t\vec{u}) - f_j(\vec{x_0})}{t}$$

exists and in this case $\vec{f'_{\vec{u}}}(\vec{x_0}) = (b_1, ..., b_m)$.

Let $m = 1$, so that $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ and let $\vec{x_0} = (a_1, ..., a_n)$. For $i, 1 \leq i \leq n$, we let $\vec{e_i}$ denote the "standard" basis vectors, that is the vector that is 1 in the i-th entry and 0 in all other entries. In this case we have that

$$f'_{\vec{e_i}}(\vec{x_0}) = \lim_{t \to 0} \frac{f(\vec{x_0} + t\vec{e_i}) - f(\vec{x_0})}{t} = \lim_{t \to 0} \frac{f(a_1, ..., a_{i-1}, a_i + t, a_{i+1}, ..., a_n) - f(a_1, ..., a_n)}{t} = \frac{\partial f}{\partial x_i}(\vec{x_0}),$$

the **partial derivative of f in the i-th direction.**

For this reason, when $m \geq 1$, and $\vec{f} : S \to \mathbb{R}^m$, we will often use the notation,

$$\vec{f'}_{\vec{e_i}}(\vec{x_0}) = \frac{\partial \vec{f}}{\partial x_i},$$

and we still call this vector in $\mathbb{R}^m$ the **partial derivative of $\vec{f}$ in the i-th direction.**

Some examples are in order.

Example 2.2. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by*

$$f(x_1, x_2) = \begin{cases} \frac{x_1^2 x_2}{x_1^4 + x_2^2}, & (x_1, x_2) \neq (0,0) \\ 0, & (x_1, x_2) = (0,0) \end{cases}.$$

Given $\vec{u} = (u_1, u_2) \neq (0,0)$, we have that

$$f'_{\vec{u}}(0,0) = \lim_{t \to 0} \frac{f(tu_1, tu_2) - f(0,0)}{t} = \lim_{t \to 0} \frac{\frac{t^3 u_1^2 u_2}{t^4 u_1^4 + t^2 u_2^2} - 0}{t} = \lim_{t \to 0} \frac{u_1^2 u_2}{t^2 u_1^4 + u_2^2} = \begin{cases} 0, & u_2 = 0 \\ \frac{u_1^2}{u_2}, & u_2 \neq 0 \end{cases}.$$

Thus, we see that the partial derivative of f exists at $(0,0)$ for every direction $\vec{u}$.

However, we will now show that f is not even continuous at $(0,0)$. To see that f is not continuous, consider

$$\lim_{t\to 0} f(t,t^2) = \lim_{t\to 0} \frac{t^4}{t^4+t^4} = 1/2 \neq f(0,0),$$

even though $\lim_{t\to 0}(t,t^2) = (0,0)$. Hence, f is not continuous at $(0,0)$.

Thus unlike the one variable case, where existence of the derivative guarantees continuity, **existence of the directional derivatives at a point in every direction does not even guarantee that the function is continuous.**

The next example shows that it is possible for a function to be continuous at $(0,0)$ have both partial derivatives exist, yet the directional derivatives in other directions do not exist.

Example 2.3. *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be defined by*

$$f(x_1,x_2) = \begin{cases} x_1 - x_2, & \textit{when } x_1 \geq 0 \textit{ or } x_2 \geq 0 \\ 0, & x_1 < 0 \textit{ and } x_2 < 0 \end{cases}.$$

It is easily checked that this function is continuous at $(0,0)$ and that $\frac{\partial f}{\partial x_1}(0,0) = 1$ and $\frac{\partial f}{\partial x_2}(0,0) = -1$.

If $\vec{u} = (a,b)$ lies in the 2nd or 4th quadrants, i.e., $ab < 0$, then

$$f'_{\vec{u}}(0,0) = \lim_{t\to 0} \frac{f(ta,tb)}{t} = \lim_{t\to 0} \frac{t(a-b)}{t} = a - b.$$

But if $\vec{u} = (a,b)$ with $a > 0$ and $b > 0$, then

$$f'_{\vec{u}}(0,0) = \lim_{t\to 0} \frac{f(ta,tb)}{t}.$$

When $t \to 0^-$, this numerator is 0 and so this one-sided limit exists and is equal to 0. But

$$\lim_{t\to 0^+} \frac{f(ta,tb)}{t} = a - b,$$

and so the two one-sided limits are equal if and only if $a = b$. Thus, for $a > 0, b > 0$, $f'_{\vec{u}}(0,0)$ exists only for $a = b$.

Thus, we see that the existence of both partial derivatives does not guarantee the existence of directional derivatives in all directions. In fact, we see that the set of directions for which directional derivatives exist may be a very surprising set.

2.1 The Total Derivative

To avoid all the problems associated with just partial derivatives, we use the total derivative which is actually a linear map. This is motivated by the tangent line approximation in one variable.

Definition 2.4. *Let $E \subseteq \mathbb{R}^n$ and let $\vec{x_0}$ be an interior point of E and let $\vec{f} : E \to \mathbb{R}^m$. We say that $\vec{f}$ is* **differentiable at** $\vec{x_0}$ *and has* **total derivative** L *where $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map provided that*

$$\lim_{\vec{h}\to\vec{0}} \frac{\|\vec{f}(\vec{x_0}+\vec{h}) - \vec{f}(\vec{x_0}) - L(\vec{h})\|_2}{\|\vec{h}\|_2} = 0.$$

When the total derivative exists, we set $\vec{f}'(\vec{x_0}) = L$.

We will denote the vector appearing in the numerator of the above limit by $E_f(\vec{h})$, that is,

$$E_f(\vec{h}) = \vec{f}(\vec{x_0}+\vec{h}) - \vec{f}(\vec{x_0}) - L(\vec{h}).$$

In this notation the condition for differentiability can be written as

$$\lim_{\vec{h}\to\vec{0}} \frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} = 0.$$

We shall often use the following reformulation.

Proposition 2.5. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$. Then $\vec{f}$ is differentiable at $\vec{x_0}$ if and only if for every $\epsilon > 0$, there exists $r > 0$ so that $\|\vec{h}\|_2 < r$ implies that $\|E_f(\vec{h})\|_2 \leq \epsilon\|\vec{h}\|_2$.*

Note that in the reformulation we have moved the $\|\vec{h}\|_2$ from the denominator to the right hand side of the inequality and we are now allowing $\vec{h} = \vec{0}$.

Proof. Requiring the limit to be 0 is equivalent to requiring that for every $\epsilon > 0$, there is an $r > 0$, so that when $\|\vec{h}\|_2 < r$ and $\vec{h} \neq \vec{0}$, we have

$$\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} < \epsilon$$

or

$$\|E_f(\vec{h})\|_2 < \epsilon\|\vec{h}\|_2.$$

Hence, $\|E_f(\vec{h})\|_2 \leq \epsilon\|\vec{h}\|_2$, for $\|\vec{h}\|_2 < r$ and $\vec{h} \neq \vec{0}$. Now that $\|\vec{h}\|_2$ is no longer in the denominator, we can allow $\vec{h} = \vec{0}$. When $\vec{h} = \vec{0}$, both sides of the above inequality are 0, so the less than or equal to holds there as well.

Conversely, if we only have the condition with the less than or equal to condition, then given $\epsilon > 0$, we may pick an $r > 0$, so that for $\|\vec{h}\|_2 < r$, we have that $\|E_f(\vec{h})\|_2 \leq \frac{\epsilon}{2}\|\vec{h}\|_2$. Hence, for every $\|\vec{h}\|_2 < r$ and $\vec{h} \neq \vec{0}$, we have that

$$\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} \leq \frac{\epsilon}{2} < \epsilon.$$

□

Notice that the condition that $\vec{f}$ be differentiable at $\vec{x_0}$ is that, not only does $E_f(\vec{h})$ tend to 0 as $\vec{h}$ tends to $\vec{0}$ but it does it so rapidly that it still tends to 0 when divided by $\|\vec{h}\|_2$.

Recall that every linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ is given by multiplication by an $m \times n$ matrix $A = (a_{i,j})$. Because matrix multiplication is best visualized by considering vectors as column vectors and column vectors are hard to include in text, we will write a vector $\vec{h} \in \mathbb{R}^n$ as $\vec{h} = (h_1, ..., h_n)^t$ where the superscript t denotes transpose. Thus,

$$L(\vec{h}) = A \cdot \vec{h} = (\sum_{j=1}^{n} a_{1,j}h_j, ..., \sum_{j=1}^{n} a_{n,j}h_j)^t.$$

We will often write $L = L_A$ when we want to emphasize that L is the linear map arising as multiplication by the matrix A. Other times we will simply identify the linear map and the matrix and write $L = A$.

We now look at implications of a function being differentiable.

Theorem 2.6. *Let $E \subseteq \mathbb{R}^n$ and let $\vec{x_0}$ be an interior point of E and let $\vec{f} : E \to \mathbb{R}^m$. If $\vec{f}$ is differentiable at $\vec{x_0}$ with $f'(\vec{x_0}) = L$, then the directional derivatives also exist for all directions and*

$$\vec{f}'_{\vec{u}}(\vec{x_0}) = L(\vec{u}).$$

Proof. The result is obvious when $\vec{u} = \vec{0}$, so we only consider $\vec{u} \neq \vec{0}$. We have

$$\vec{f}'_{\vec{u}}(\vec{x_0}) = \lim_{t\to 0} \frac{\vec{f}(\vec{x_0} + t\vec{u}) - \vec{f}(\vec{x_0})}{t} =$$

$$\lim_{t\to 0} \frac{\vec{f}(\vec{x_0} + t\vec{u}) - \vec{f}(\vec{x_0}) - L(t\vec{u})}{t} + \frac{L(t\vec{u})}{t} = L(\vec{u}) + \lim_{t\to 0} \frac{\vec{f}(\vec{x_0} + t\vec{u}) - \vec{f}(\vec{x_0}) - L(t\vec{u})}{t} =$$

$$L(\vec{u}) + \lim_{t\to 0} \frac{E_f(t\vec{u})}{t}.$$

When $\vec{u} \neq \vec{0}$, then

$$\|\frac{E_f(t\vec{u})}{t}\|_2 = \frac{\|E_f(t\vec{u})\|_2}{\|t\vec{u}\|_2}\|\vec{u}\|_2 \to 0,$$

by the differentiability condition. Thus,

$$\frac{E_f(t\vec{u})}{t} \to \vec{0}$$

as $t \to 0$ and the result follows. □

Theorem 2.7. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$ and let $f_i : E \to \mathbb{R}, 1 \leq i \leq m$ denote the components of $\vec{f}$. If $\vec{f}$ is differentiable at $\vec{x_0}$, then $\frac{\partial f_i(\vec{x_0})}{\partial x_j}$ exists for all i and j and*

$$\vec{f}'(\vec{x_0}) = \Big(\frac{\partial f_i(\vec{x_0})}{\partial x_j}\Big).$$

Proof. Let $\vec{e_j}$ denote the unit vector in the j-th direction. Then

$$(\frac{\partial f_1}{\partial x_j}(\vec{x_0}), ..., \frac{\partial f_m}{\partial x_j}(\vec{x_0}))^t = \frac{\partial \vec{f}}{\partial x_j}(\vec{x_0}) = \vec{f}'_{\vec{e_j}}(\vec{x_0}) = \vec{f}'(\vec{x_0})(\vec{e_j}).$$

Now note that if one multiplies a matrix times the vector $\vec{e_j}$, then the resulting vector is the j-th column of the matrix. the result follows. □

Some special cases of the above notation are useful. First we look at the case where the range i one dimensional ad then when the domain is one dimensional.

Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $f : E \to \mathbb{R}$ be differentiable at $\vec{x_0}$. Then $f'(\vec{x_0})$ is the $1 \times n$ matrix

$$\Big(\frac{\partial f}{\partial x_1}(\vec{x_0}), ..., \frac{\partial f}{\partial x_n}(\vec{x_0})\Big).$$

In calculus, this (row) vector is usually deoted by ∇f. Thus, in our notation for $f : \mathbb{R}^n \to \mathbb{R}$ we have that when f is differentiable,

$$f' = \nabla f.$$

We now look at the case when the domain is one dimensional. Let $E \subseteq \mathbb{R}$, let x_0 be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$ and let $f_i : E \to \mathbb{R}, 1 \leq i \leq m$ denote the components of $\vec{f}$. If $\vec{f}$ is differentiable at $\vec{x_0}$ then

$\vec{f}'(x_0) = \left(\frac{\partial f_i}{\partial x_1}(x_0)\right)$ is an $m \times 1$ matrix. Notice that since there is really only one variable, we usually omit the ∂ notation, use ordinary derivatives signs and this (column) vector, in calculus, is denoted by $\frac{d\vec{f}}{dx}(x_0)$. Thus, in this case,

$$\vec{f}'(x_0) = \Big(\frac{df_1}{dx}(x_0), ..., \frac{df_m}{dx}(x_0)\Big)^t = \frac{d\vec{f}}{dx}(x_0).$$

Proposition 2.8. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$ and let $f_i : E \to \mathbb{R}, 1 \leq i \leq m$ denote the components of $\vec{f}$. Then $\vec{f}$ is differentiable at $\vec{x_0}$ if and only if f_i is differentiable at$\vec{x_0}$ for all $i = 1, ..., m$.*

Proof. Note that $E_f(\vec{h})$ is a vector whose components are $E_i(\vec{h}), i = 1, ..., m$. Thus, we have that $|E_{f_i}(\vec{h})| \leq \|E_f(\vec{h})||_2$, for $i = 1, ..., m$. Conversely,

$$\|E_f(\vec{h})\|_2^2 \leq |E_{f_1}(\vec{h})|^2 + \cdots + |E_{f_m}(\vec{h})|^2.$$

If $\vec{f}$ is differentiable at $\vec{x_0}$ and given $\epsilon > 0$, we pick $r > 0$, so that $\frac{\|E_f(\vec{h})\|_2}{\|h\|_2} < \epsilon$ for $\|\vec{h}\|_2 < r, \vec{h} \neq \vec{0}$, then $\frac{|E_{f_i}(\vec{h})|}{\|\vec{h}\|_2} < \epsilon$ for $\|\vec{h}\|_2 < r, \vec{h} \neq \vec{0}$.

Conversely, if each f_i is differentiable at $\vec{x_0}$, then given $\epsilon > 0$, we may pick $r_i > 0$ so that $\frac{|E_{f_i}(\vec{h})|}{\|\vec{h}\|_2} < \frac{\epsilon}{\sqrt{m}}$, for $\|\vec{h}\|_2 < r_i, \vec{h} \neq \vec{0}$. If we let $r = \min\{r_1, ..., r_m\}$, then for $\|\vec{h}\|_2 < r, \vec{h} \neq \vec{0}$, we have

$$\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} = \frac{\sqrt{|E_{f_1}(\vec{h})|^2 + \cdots + |E_{f_m}(\vec{h})|^2}}{\|\vec{h}\|_2} =$$
$$\sqrt{\Big(\frac{|E_{f_1}(\vec{h})|}{\|\vec{h}\|_2}\Big)^2 + \cdots + \Big(\frac{|E_{f_m}(\vec{h})|}{\|\vec{h}\|_2}\Big)^2} < \sqrt{(\frac{\epsilon}{\sqrt{m}})^2 + \cdots + (\frac{\epsilon}{\sqrt{m}})^2} = \epsilon.$$

Hence, $\vec{f}$ is differentiable at $\vec{x_0}$. □

Problem 2.9. *Let $g, h : \mathbb{R} \to \mathbb{R}$. Prove that if g is differentiable at a(in the usual sense) and h is differentiable at b(in the usual sense), then $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x_1, x_2) = g(x_1)h(x_2)$ is differentiable at $\vec{x_0} = (a, b)$.*

Problem 2.10. *Use the above problem to give an example of a function $f : \mathbb{R}^2 \to \mathbb{R}$ that is differentiable at $(0, 0)$, $\frac{\partial f}{\partial x_j}, j = 1, 2$ both exist everywhere, but the partial derivatives are not continuous at $(0, 0)$.*

Problem 2.11. *Let $E \subseteq \mathbb{R}$, let x_0 be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$ and let $f_i : E \to \mathbb{R}, 1 \le i \le m$ denote the components of $\vec{f}$. Prove that if $\frac{df_i}{dx}$ exists at x_0 for all $1 \le i \le m$, then $\vec{f}$ satisfies the definition of differentiable at x_0.*

Problem 2.12. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$ and let $f_i : E \to \mathbb{R}, 1 \le i \le m$ denote the components of $\vec{f}$. Fix a (row) vector $\vec{a} = (a_1, ..., a_m) \in \mathbb{R}^m$ and let $g : E \to \mathbb{R}$ be the function $g = a_1 f_1 + \cdots + a_m f_m$, i.e., $g = \vec{a} \cdot \vec{f}$. Prove that if $\vec{f}$ is differentiable at $\vec{x_0}$, then g is differentiable at $\vec{x_0}$ and*

$$g'(\vec{x_0}) = \vec{a} \cdot \vec{f}'(\vec{x_0}),$$

where this last product is the product of a $1 \times m$ matrix and an $m \times n$ matrix.

2.2 Differentiability and Continuity

We saw earlier that the existence of partial derivatives is not enough to imply continuity. Now we will prove that a differentiable function is continuous.

Definition 2.13. *Let $A = (a_{i,j})$ be an $m \times n$ matrix. Then we set*

$$\|A\|_2 = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2}.$$

For example, if we I_n denote the $n \times n$ identity matrix, i.e., the matrix with 1's for its diagonal entries and 0's elsewhere, then $\|I_n\|_2 = \sqrt{n}$.

Proposition 2.14. *Let $A = (a_{i,j})$ be an $m \times n$ matrix and let $\vec{h} = (h_1, ..., h_n)^t \in \mathbb{R}^n$, then*

$$\|A(\vec{h})\|_2 \le \|A\|_2 \|\vec{h}\|_2.$$

Proof. Using the Schwarz inequality, we have that

$$\|A(\vec{h})\|_2^2 = \|(\sum_{j=1}^{n} a_{1,j} h_j, ..., \sum_{j=1}^{n} a_{m,j} h_j)\|_2^2 =$$
$$\sum_{i=1}^{m} \Big[\sum_{j=1}^{n} a_{i,j} h_j \Big]^2 \le \sum_{i=1}^{m} \Big[(\sum_{j=1}^{n} a_{i,j}^2)(\sum_{j=1}^{n} h_j^2) \Big] =$$
$$\sum_{i=1}^{m} \Big[(\sum_{j=1}^{n} a_{i,j}^2) \|\vec{h}\|_2^2 \Big] = \|A\|_2^2 \|\vec{h}\|_2^2.$$

□

Theorem 2.15. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{f} : E \to \mathbb{R}^m$. If $\vec{f}$ is differentiable at $\vec{x_0}$, then $\vec{f}$ is continuous at $\vec{x_0}$.*

Proof. Let $\vec{f}'(\vec{x_0}) = A$. Since $\vec{f}$ is differentiable at $\vec{x_0}$ for $\epsilon_1 = 1$, there exists $r_1 > 0$, such that $\|\vec{h}\|_2 < r_1, \vec{h} \neq \vec{0}$ implies that $\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} < 1$. Hence, $\|E_f(\vec{h}\|_2 \leq \|\vec{h}\|_2$ for all $\|\vec{h}\|_2 < r_1$.

Given $\epsilon > 0$, set $r_2 = \frac{\epsilon}{\|A\|_2+1}$, and let $r = \min\{r_1, r_2\}$. Then for $\|\vec{h}\|_2 < r$, we have that

$$\|\vec{f}(\vec{x_0} + \vec{h}) - \vec{f}(\vec{x_0})\|_2 = \|A(\vec{h}) + E_f(\vec{h})\|_2 \leq \\ \|A(\vec{h})\|_2 + \|E_f(\vec{h})\|_2 \leq \|A\|_2\|\vec{h}\|_2 + \|\vec{h}\|_2 = (\|A\|_2 + 1)\|\vec{h}\|_2 < \epsilon.$$

Hence, $\vec{f}$ is continuous at $\vec{x_0}$. □

Since the function of Example 2.2 is not continuous at $(0, 0)$, it is not differentiable at $(0, 0)$. Thus, the function of Example 2.2 is a **function that has directional derivatives in every direction, but is not differentiable.**

2.3 The Chain Rule

In calculus we often learn many multivariable versions of the chain rule depending on the situation. But the most concise form involves composition of linear maps and matrix products.

Recall that given linear maps $L_B : \mathbb{R}^n \to \mathbb{R}^m$ and $L_A : \mathbb{R}^m \to \mathbb{R}^p$, their composition is the linear map $L_A \circ L_B : \mathbb{R}^n \to \mathbb{R}^p$, and it is given by multiplication by the matrix that is the product of the matrices A and B, where the (i,j)-entry of the matrix product is given by

$$AB = \Big(\sum_{k=1}^{m} a_{i,k} b_{k,j}\Big).$$

In particular, $L_A \circ L_B = L_{AB}$.

Theorem 2.16. *Let $E \subseteq \mathbb{R}^n$, let $\vec{x_0}$ be an interior point of E, let $\vec{g} : E \to \mathbb{R}^m$ be differentiable at $\vec{x_0}$ with $\vec{g}'(\vec{x_0}) = B$ and let $Y \subseteq \mathbb{R}^m$ with $\vec{g}(E) \subseteq Y$, let $\vec{y_0} = \vec{g}(\vec{x_0})$ be an interior point of Y and let $\vec{f} : Y \to \mathbb{R}^p$ be differentiable at $\vec{y_0}$ with $\vec{f}'(\vec{y_0}) = A$. Then the function $\vec{f} \circ \vec{g} : E \to \mathbb{R}^p$ is differentiable at $\vec{x_0}$ and*

$$(\vec{f} \circ \vec{g})'(\vec{x_0}) = \vec{f}'(\vec{g}(\vec{x_0})) \circ \vec{g}'(\vec{x_0}) = AB.$$

Proof. Let $\epsilon > 0$. Using the differentiablility of $\vec{g}$ at $\vec{x_0}$ we may pick $r_1 > 0$, so that $\|\vec{h}\|_2 < r_1$ implies that $\|E_g(\vec{h})\|_2 \le \|\vec{h}\|_2$. Using the differentiablity of $\vec{f}$ at $\vec{y_0}$ we may choose $r_2 > 0$, so that for $\|\vec{k}\|_2 < r_2$, we have that $\|E_f(\vec{k})\| \le \frac{\epsilon}{2(\|B\|_2+1)}\|\vec{k}\|_2$. Finally, using the differentiablilty of $\vec{g}$ at $\vec{x_0}$ again, we may choose $r_3 > 0$, so that for $\|\vec{h}\|_2 < r_3$, we have that $\|E_g(\vec{h})\|_2 \le \frac{\epsilon}{2(\|A\|_2+1)}\|\vec{h}\|_2$.

Let $r = \min\{r_1, \frac{r_2}{\|B\|_2+1}, r_3\}$. Then for $\|\vec{h}\|_2 < r$, we have that

$$\|B(\vec{h}) + E_g(\vec{h})\|_2 \le \|B\|_2\|\vec{h}\|_2 + \|E_g(\vec{h})\|_2 \le (\|B\|_2 + 1)\|\vec{h}\|_2 < r_2,$$

and hence,

$$\begin{gathered}\|\vec{f}\circ\vec{g}(\vec{x_0}+\vec{h}) - \vec{f}\circ\vec{g}(\vec{x_0}) - AB(\vec{h})\|_2 = \|\vec{f}\big(\vec{g}(\vec{x_0}+\vec{h})\big) - \vec{f}\big(\vec{g}(\vec{x_0})\big) - AB(\vec{h})\|_2 = \\ \|\vec{f}\big(\vec{g}(\vec{x_0}) + B(\vec{h}) + E_g(\vec{h})\big) - \vec{f}\big(\vec{g}(\vec{x_0})\big) - A\big(B(\vec{h}) + E_g(\vec{h})\big) + A\big(E_g(\vec{h})\big)\|_2 = \\ \|E_f\big(B(\vec{h}) + E_g(\vec{h})\big) + A\big(E_g(\vec{h})\big)\|_2 \le \|E_f\big(B(\vec{h}) + E_g(\vec{h})\big)\|_2 + \|A\big(E_g(\vec{h})\big)\|_2 \le \\ \frac{\epsilon}{2(\|B\|_2+1)}\|B(\vec{h}) + E_g(\vec{h})\|_2 + \|A\|_2\|\|E_g(\vec{h})\|_2 \le \\ \frac{\epsilon}{2(\|B\|_2+1)}(\|B\|_2+1)\|\vec{h}\|_2 + \|A\|_2\frac{\epsilon}{2(\|A\|_2+1)}\|\vec{h}\|_2 \le \epsilon\|\vec{h}\|_2.\end{gathered}$$

This proves that the derivative exists and is equal to AB. □

We now look at a couple of applications of the Chain Rule. First notice that if we fix $\vec{x_0}, \vec{u} \in \mathbb{R}^n$ and define $\vec{g} : \mathbb{R} \to \mathbb{R}^n$, by $\vec{g}(t) = \vec{x_0} + t\vec{u}$, then $\vec{g}$ is differentiable by Problem 2.11 and

$$\vec{g}'(0) = \frac{d\vec{g}}{dt}(0) = \vec{u}.$$

Hence, if $E \subseteq \mathbb{R}^n$, $\vec{x_0}$ is an interior point of E and $\vec{f} : E \to \mathbb{R}^m$ is differentiable at $\vec{x_0}$, then by the Chain Rule

$$\vec{f} \circ \vec{g}(t) = \vec{f}(\vec{x_0} + t\vec{u}),$$

is differentiable at 0, and

$$\vec{f}'_{\vec{u}}(\vec{x_0}) = \Big(\vec{f} \circ \vec{g}\Big)'(0) = \vec{f}'(\vec{x_0})\vec{u}.$$

Thus, we see that Theorem 2.6 is also a consequence of the Chain Rule.

2.4 Multivariable Mean Value Theorem

Recall that the one-variable Mean Value theorem says that if $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists $a < t_0 < b$ such that $f(b) - f(a) = f'(t_0)(b - a)$. Now suppose that we have $\vec{f} : [a, b] \to \mathbb{R}^m$ with component functions f_i that is differentiable on (a, b) and continuous on $[a, b]$. A direct extension of the mean value theorem would be for there to exist $a < t_0 < b$, so that $\vec{f}(b) - \vec{f}(a) = \vec{f}'(t_0)(b - a)$. However, looking at each component, this would imply that for this one value t_0, we had $f_i(b) - f_i(a) = f_i'(t_0)(b - a)$, for all $i = 1, ..., m$. Thus, a single numer t_0 would be working, simultaneously, as the point for every one of the functions, $f_1, ..., f_m$. This is just not true, in fact it is quite easy to find differentiable functions $f_1, f_2 : [0, 1] \to \mathbb{R}$ so that there are unique points $0 < t_1, t_2 < 1$ with $f_1(1) - f_1(0) = f_1'(t_1)(1 - 0)$ and $f_2(1) - f_2(0) = f_2'(t_2)(1 - 0)$, but $t_1 \neq t_2$.

Thus, the most obvious multivariable generalization of the mean value theorem is not true even for just vector-valued functions. But there is a multivariable mean value theorem that is nearly as useful as the one variable version.

Definition 2.17. *A subset $C \subseteq \mathbb{R}^n$ is called* **convex** *if whenever $\vec{x_2}, \vec{x_1} \in C$, then every point on the line segment that joins them*

$$\vec{x_1} + t(\vec{x_2} - \vec{x_1}), 0 \leq t \leq 1,$$

is also in C.

For some examples of convex sets, note that every ball $B(\vec{x_0}; r)$ is convex, and the set $C = \{(x_1, ..., x_n) \in \mathbb{R}^n : x_i > 0 \forall i\}$ is convex.

Theorem 2.18 (Multivariable Mean Value Theorem)**.** *Let $C \subseteq \mathbb{R}^n$ be an open convex set, let $\vec{f} : C \to \mathbb{R}^m$ and let $f_i : C \to \mathbb{R}, 1 \leq i \leq m$ denote the components of $\vec{f}$. If $\vec{f}$ is differentiable at each point in C, $\vec{a} = (a_1, ..., a_m) \in \mathbb{R}^m$ and $\vec{x_2}, \vec{x_1} \in C$, then there exists $\vec{z} \in C$ such that*

$$\vec{a} \cdot (\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})) = \vec{a} \cdot [\vec{f}'(\vec{z})(\vec{x_2} - \vec{x_1})].$$

Proof. Since the line segment $\vec{x_1} + t(\vec{x_2} - \vec{x_1}), 0 \leq t \leq 1$ joining $\vec{x_1}$ to $\vec{x_2}$ stays inside C, we may define $g : [0, 1] \to \mathbb{R}$ by

$$g(t) = a_1 f_1(\vec{x_1} + t(\vec{x_2} - \vec{x_1})) + \cdots + a_m f_m(\vec{x_1} + t(\vec{x_2} - \vec{x_1})).$$

By the chain rule g is differentiable on $(0,1)$ and g is continuous on $[0,1]$ and so there is a point $0 < t_0 < 1$ with

$$\vec{a}\cdot\vec{f}(\vec{x_2}) - \vec{a}\cdot\vec{f}(\vec{x_1}) = g(1) - g(0) = g'(t_0) =$$
$$a_1 f_1'(\vec{x_1} + t_0(\vec{x_2} - \vec{x_1}))\cdot(\vec{x_2} - \vec{x_1}) + \cdots + a_m f_m'(\vec{x_1} + t_0(\vec{x_2} - \vec{x_1}))\cdot(\vec{x_2} - \vec{x_1}) =$$
$$\vec{a}\cdot[f'(\vec{x_1} + t_0(\vec{x_2} - \vec{x_1}))(\vec{x_2} - \vec{x_1})] = \vec{a}\cdot[\vec{f}'(\vec{z})(\vec{x_2} - \vec{x_1})],$$

where $\vec{z} = \vec{x_1} + t_0(\vec{x_2} - \vec{x_1}) \in C$. □

For the next result it will be better to have some better estimates.

Definition 2.19. *Given an $m \times n$ matrix A we define the* **norm** *of the matrix by*

$$\|A\| = \sup\{\frac{\|A\vec{h}\|_2}{\|h\|_2} : \vec{h} \neq \vec{0}\},$$

where the supremum is taken over all non-zero vectors $\vec{h} \in \mathbb{R}^n$.

Note that we have the inequality $\|A\vec{h}\|_2 \leq \|A\|\|\vec{h}\|_2$, and another way to view the number $\|A\|$ is that it is the smallest constant K satisfying, $\|A\vec{h}\|_2 \leq K\|\vec{h}\|_2$. In fact, it is not hard to show that

$$\|A\| = \inf\{K : \|A\vec{h}\|_2 \leq K\|\vec{h}\|_2 \forall \vec{h}\}.$$

Since we have shown earlier that $\|A\vec{h}\|_2 \leq \|A\|_2\|\vec{h}\|_2$, we have that

$$\|A\| \leq \|A\|_2.$$

If we consider the $n \times n$ identity matrix I_n, then we have that $I_n\vec{h} = \vec{h}$, and so $\|I_n\| = 1$, but $\|I_n\|_2 = \sqrt{n}$.

Theorem 2.20. *Let $C \subseteq \mathbb{R}^n$ be an open convex set and let $\vec{f} : C \to \mathbb{R}^m$. If $\vec{f}$ is differentiable at every point in C and $\|\vec{f}'(\vec{z})\| \leq M$ for every $\vec{z} \in C$, then $\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2 \leq M\|\vec{x_2} - \vec{x_1}\|_2$ for any $\vec{x_2}, \vec{x_1} \in C$.*

Proof. Let $\vec{a} = \vec{f}(\vec{x_2} - \vec{f}(\vec{x_1})$ and as before define $g : [0,1] \to \mathbb{R}$ by $g(t) = \vec{a}\cdot\vec{f}(\vec{x_1} + t(\vec{x_2} - \vec{x_1}))$. We have

$$\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2^2 = \vec{a}\cdot\vec{f}(\vec{x_2}) - \vec{a}\cdot\vec{f}(\vec{x_1}) = g(1) - g(0) =$$
$$g'(t_0) = \vec{a}\cdot[\vec{f}'(\vec{x_1} + t_0(\vec{x_2} - \vec{x_1}))(\vec{x_2} - \vec{x_1})] =$$
$$\vec{a}\cdot[\vec{f}'(\vec{z})(\vec{x_2} - \vec{x_1})].$$

Now applying the Schwarz inequality,

$$\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2^2 \leq \vec{a} \cdot [\vec{f}'(\vec{z})(\vec{x_2} - \vec{x_1})] \leq \|\vec{a}\|_2 \|\vec{f}'(\vec{z})(\vec{x_2} - \vec{x_1})\|_2 \leq \|\vec{a}\|_2 \|\vec{f}'(\vec{z})\| \|\vec{x_2} - \vec{x_1}\|_2 \leq M\|\vec{a}\|_2 \|\vec{x_2} - \vec{x_1}\|_2 = M\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2 \|\vec{x_2} - \vec{x_1}\|_2,$$

and after cancelling $\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2$ from both sides of the equation, the result follows. □

Theorem 2.21. *Let $U \subseteq \mathbb{R}^n$ be an open connected set and let $\vec{f} : U \to \mathbb{R}^m$. If $\vec{f}$ is differentiable at each point of U and $\vec{f}'(\vec{x}) = 0$ for all $\vec{x} \in U$, then $\vec{f}$ is constant.*

Proof. Pick a point $\vec{b} \in U$ and let $\vec{c} = \vec{f}(\vec{b})$. We will prove that $\vec{f}(\vec{x}) = \vec{c}$ for all $\vec{x} \in U$. To this end let $V = \{\vec{x} \in U : \vec{f}(\vec{x}) = \vec{c}\}$. We have that $\vec{b} \in V$, so V is non-empty. We will prove that V is an open and closed in the connected metric space (U, d_2), and hence $V = U$ and the theorem is proved.

It remains to verify the claim. First, if $\vec{x_0} \in V$, then since $\vec{x_0} \in U$, there is an $r > 0$, so that $B(\vec{x_0}; r) \subseteq U$. Now if $\vec{x_1}, \vec{x_2} \in B(\vec{x_0}; r)$, then by our last theorem with $M = 0$, we have $\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2 \leq 0\|\vec{x_2} - \vec{x_1}\|_2$. Thus, $\vec{f}(\vec{x_2}) = \vec{f}(\vec{x_1})$. □

2.5 Continuously Differentiable Functions

In this section we give a sufficient condition for a function to be differentiable.

Definition 2.22. *Let $U \subseteq \mathbb{R}^n$, let $\vec{f} : U \to \mathbb{R}^m$ and let $f_i, i = 1, ..., m$ be the component functions. We say that $\vec{f}$ is $\mathcal{C}^1$* **on** *U provided that $\frac{\partial f_i}{\partial x_j}$ exist for $1 \leq i \leq m$ and $1 \leq j \leq n$ and are continuous on U.*

Proposition 2.23. *Let $U \subseteq \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}$ be $\mathcal{C}^1$ on U. Then f is differentiable on U.*

Proof. For clarity of exposition we only do the case $n = 3$. Let $\vec{x_0} = (a, b, c) \in U$. To prove that f is differentiable at $\vec{x_0}$ we must show that for $\vec{h} = (h_1, h_2, h_3)$, we have

$$\lim_{\vec{h} \to \vec{0}} \frac{|f(\vec{x_0} + \vec{h}) - f(\vec{x_0}) - \frac{\partial f(\vec{x_0})}{\partial x_1} h_1 - \frac{\partial f(\vec{x_0})}{\partial x_2} h_2 - \frac{\partial f(\vec{x_0})}{\partial x_3} h_3|}{\|\vec{h}\|_2} = 0.$$

We regroup the numerator and then apply the one variable mean value theorem three times to get

$$\begin{aligned}
f(\vec{x_0}+\vec{h}) - f(\vec{x_0}) - \frac{\partial f(\vec{x_0})}{\partial x_1}h_1 - \frac{\partial f(\vec{x_0})}{\partial x_2}h_2 - \frac{\partial f(\vec{x_0})}{\partial x_3}h_3 = \\
f((a+h_1, b+h_2, c+h_3)) - f((a+h_1, b+h_2, c)) - \frac{\partial f(\vec{x_0})}{\partial x_3}h_3 \\
+ f((a+h_1, b+h_2, c)) - f((a+h_1, b, c)) - \frac{\partial f(\vec{x_0})}{\partial x_2}h_2 \\
+ f((a+h_1, b, c)) - f((a, b, c)) - \frac{\partial f((a,b,c))}{\partial x_1}h_1 \\
= \frac{\partial f((a+h_1, b+h_2, t_3)}{\partial x_3}h_3 - \frac{\partial f((a,b,c))}{\partial x_3}h_3 \\
+ \frac{\partial f((a+h_1, t_2, , c))}{\partial x_2}h_2 - \frac{\partial f((a,b,c))}{\partial x_2}h_2 \\
+ \frac{\partial f((t_1, b, c))}{\partial x_1}h_1 - \frac{\partial f((a,b,c))}{\partial x_1}h_1 \\
= \Big(\frac{\partial f((a+h_1, b+h_2, t_3)}{\partial x_3} - \frac{\partial f((a,b,c))}{\partial x_3}\Big)h_3 \\
+ \Big(\frac{\partial f((a+h_1, t_2, , c))}{\partial x_2} - \frac{\partial f((a,b,c))}{\partial x_2}\Big)h_2 \\
+ \Big(\frac{\partial f((t_1, b, c))}{\partial x_1} - \frac{\partial f((a,b,c))}{\partial x_1}\Big)h_1, \quad (2.1)
\end{aligned}$$

where t_3 is between c and $c+h_3$, t_2 is between b and $b+h_2$, and t_1 is between a and $a+h_1$.

Thus, the distance between each of the points (t_1, b, c), $(a+h_1, t_2, c)$, and $(a+h_1, b+h_2, t_3)$ and (a, b, c) is less than $\sqrt{h_1^2+h_2^2+h_3^2} = \|\vec{h}\|_2$. By the continuity of the partial derivatives, given $\epsilon > 0$ there is an $r > 0$, so that when $\|\vec{h}\|_2 < r$, we have that the differences between each pair of partials in the parentheses above is at most $\epsilon_1 = \epsilon/\sqrt{3}$.

Thus, if we apply the Schwarz inequality to (2.1), we get that the above quantity is less than

$$\sqrt{\epsilon_1^2+\epsilon_1^2+\epsilon_1^2}\sqrt{h_1^2+h_2^2+h_3^2} = \epsilon\|\vec{h}\|_2.$$

Thus, for $\|\vec{h}\|_2 < r$, we have that the above fraction is at most ϵ. □

Theorem 2.24. *Let $U \subseteq \mathbb{R}^n$ be open and let $\vec{f} : U \to \mathbb{R}^m$ be $\mathcal{C}^1$ on U. Then $\vec{f}$ is differentiable on U.*

Proof. By the above proposition each component function $f_i : U \to \mathbb{R}$ is differentiable. By Proposition 2.8, the fact that each component function is differentiable implies that $\vec{f}$ is differentiable. □

Since $\vec{f}' = (\frac{\partial f_i}{\partial x_j})$ and each entry of this matrix is continuous, we see that $\vec{f}'$ is continuous.

2.6 The Inverse Function Theorem

Loosely stated the Inverse Functions Theorem says that if $U \subseteq \mathbb{R}^n$, $\vec{f} : U \to \mathbb{R}^n$ is $\mathcal{C}^1$ and the matrix $\vec{f}'(\vec{x_0})$ is invertible, then that is enough to guarantee that in a small neighborhood of $\vec{x_0}$, the function $\vec{f}$ is one-to-one, so that an inverse function $\vec{g}$ exists and that inverse function will also be $\mathcal{C}^1$ on some neighborhood of $\vec{y_0} = \vec{f}(\vec{x_0})$, with $\vec{g}'(\vec{y_0}) = \big(\vec{f}'(\vec{x_0})\big)^{-1}$.

Before we can prove the inverse function theorem, we will need a few preliminary results about matrices and their norms.

Lemma 2.25. *Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Then $\|AB\| \leq \|A\|\|B\|$.*

Proof. Given any $\vec{h} \in \mathbb{R}^p$, we have that

$$\|(AB)\vec{h}\|_2 = \|A(B\vec{h})\|_2 \leq \|A\|\|B\vec{h}\|_2 \leq \|A\|\|B\|\|\vec{h}\|_2,$$

and the result follows. □

Lemma 2.26. *If A is an $n \times n$ matrix that is invertible, then for any $\vec{h} \in \mathbb{R}^n$, we have*

$$\frac{\|\vec{h}\|_2}{\|A^{-1}\|} \leq \|A\vec{h}\|_2.$$

Proof. We have that

$$\|\vec{h}\|_2 = \|A^{-1}(A\vec{h})\|_2 \leq \|A^{-1}\|\|A\vec{h}\|_2,$$

and the result follows by dividing this inequality by $\|A^{-1}\|$. □

Lemma 2.27 (The Reverse Triangle Inequality)**.** *Let $\vec{x}, \vec{y} \in \mathbb{R}^n$, then*

$$|\|\vec{x}\|_2 - \|\vec{y}\|_2| \leq \|\vec{x} + \vec{y}\|_2.$$

Proof. By Minkowski's inequality,

$$\|\vec{x}\|_2 = \|(\vec{x}+\vec{y}) - \vec{y}\|_2 \le \|\vec{x}+\vec{y}\|_2 + \|\vec{y}\|_2,$$

which implies that $\|\vec{x}\|_2 - \|\vec{y}\|_2 \le \|\vec{x}+\vec{y}\|_2$. Reversing the roles of $\vec{x}$ and $\vec{y}$, yields that $\|\vec{y}\|_2 - \|\vec{x}\|_2 \le \|\vec{x}+\vec{y}\|_2$. These two inequalities yield the result. □

The following result uses norms of matrices(a metric concept) to guarantee invertibility(a linear algebra concept)! Roughly, it says that if a matrix is close enough to an invertible matrix, then it is also invertible.

Theorem 2.28. *Let A and B be $n \times n$ matrices. If A is invertible and $\|A - B\| < \frac{1}{\|A^{-1}\|}$, then B is invertible.*

Proof. Suppose that $B\vec{x} = \vec{0}$ with $\vec{x} \neq \vec{0}$. Then e have that $A\vec{x} = (A - B)\vec{x}$, and hence, $\vec{x} = A^{-1}(A - B)\vec{x}$.

Computing the norm of both sides, yields

$$\|\vec{x}\|_2 = \|A^{-1}(A-B)\vec{x}\|_2 \le \|A^{-1}\|\|A-B\|\|\vec{x}\|_2 < \|\vec{x}\|_2.$$

Thus, $\|\vec{x}\|_2 < \|\vec{x}\|_2$! A contradiction. Hence, $B\vec{x} = \vec{0}$ implies that $\vec{x} = \vec{0}$, and so B is one-to-one, and from linear algebra, this means that B is also onto, and so invertible. □

Theorem 2.29 (Inverse Function Theorem)**.** *Let $U \subseteq \mathbb{R}^n$ be open, let $\vec{x_0} \in U$, and let $\vec{f} : U \to \mathbb{R}^n$ be $\mathcal{C}^1$ on U. If $\vec{f}'(\vec{x_0})$ is invertible, then:*

1. *there exists $r > 0$ such that $\vec{f}$ is one-to-one on $B(\vec{x_0}; r)$,*
2. *$f'(\vec{x})$ is invertible for $\vec{x} \in B(\vec{x_0}; r)$,*
3. *$V = \vec{f}(B(\vec{x_0}; r))$ is open,*
4. *if $\vec{g} = \vec{f^{-1}}$, denotes the inverse function, then $\vec{g} : V \to B(\vec{x_0}; r)$ is $\mathcal{C}^1$ on V and $\vec{g}'(\vec{y}) = \left(\vec{f}'(\vec{g}(\vec{y}))\right)^{-1}$.*

Proof. Let $A = \vec{f}'(\vec{x_0})$ and fix a number λ, $0 < \lambda < 1$. Note that

$$\|\vec{f}'(\vec{x}) - A\|^2 \le \|\vec{f}'(\vec{x}) - A\|_2^2 = \sum_{i,j=1}^{n} \Big(\frac{\partial f_i(\vec{x})}{\partial x_j} - \frac{\partial f_i(\vec{x_0})}{\partial x_j}\Big)^2.$$

Thus, using the continuity of all the partial derivatives, we may choose $r > 0$ so that $\|\vec{x} - \vec{x_0}\|_2 < r$ implies that

$$\|\vec{f}'(\vec{x}) - \vec{f}'(\vec{x_0})\| \leq \|\vec{f}'(\vec{x}) - A\|_2 < \frac{\lambda}{\|A^{-1}\|} < \frac{1}{\|A^{-1}\|}.$$

Hence, $f'(x)$ is invertible by Theorem 2.28.

Let $\vec{f_1}(\vec{x}) = \vec{f}(\vec{x}) - A(\vec{x} - \vec{x_0})$. Then for $\vec{x} \in B(\vec{x_0}; r)$ we have that

$$\|\vec{f_1}(\vec{x})\| = \|\vec{f}(\vec{x}) - A\| < \frac{\lambda}{\|A^{-1}\|}.$$

Using Theorem 2.20, we have that for $\vec{x_2}, \vec{x_1} \in B(\vec{x_0}; r)$,

$$\|\vec{f_1}(\vec{x_2}) - \vec{f_1}(\vec{x_1})\|_2 \leq \frac{\lambda}{\|A^{-1}\|}\|\vec{x_2} - \vec{x_1}\|_2. \tag{2.2}$$

Hence, for any $\vec{x_2}, \vec{x_1} \in B(\vec{x_0}; r)$ we have

$$\begin{aligned} \|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2 = \|\vec{f_1}(\vec{x_2}) - \vec{f_1}(\vec{x_1}) + A(\vec{x_2} - \vec{x_1})\|_2 \geq \\ \|A(\vec{x_2} - \vec{x_1})\|_2 - \|\vec{f_1}(\vec{x_2}) - \vec{f_1}(\vec{x_1})\|_2 \geq \frac{\|\vec{x_2} - \vec{x_1}\|_2}{\|A^{-1}\|} - \frac{\lambda}{\|A^{-1}\|}\|\vec{x_2} - \vec{x_1}\|_2 \geq \\ \frac{(1-\lambda)}{\|A^{-1}\|}\|\vec{x_2} - \vec{x_1}\|_2. \end{aligned} \tag{2.3}$$

Note that when $\vec{x_2} \neq \vec{x_1}$ then the last term in this inequality is strictly positive, and so $\|\vec{f}(\vec{x_2}) - \vec{f}(\vec{x_1})\|_2 \neq 0$, which implies that $\vec{f}(\vec{x_2}) \neq \vec{f}(\vec{x_1})$. Thus, $\vec{f}$ is one-to-one on $B(\vec{x_0}; r)$ and we have proven (1).

We now prove that $V = \vec{f}(B(\vec{x_0}; r))$ is open. To this end pick $\vec{y_1} \in V$, so that $\vec{y_1} = \vec{f}(\vec{x_1})$ for some $\vec{x_1} \in B(\vec{x_0}; r)$. Choose an $r_1 > 0$ sufficiently small that the closed ball,

$$B^- = B(\vec{x_1}; r_1)^- \subseteq B(\vec{x_0}; r).$$

We claim that if $\|\vec{y} - \vec{y_1}\| < \frac{(1-\lambda)r_1}{\|A^{-1}\|}$, then there is $\vec{x} \in B^-$ with $\vec{f}(\vec{x}) = \vec{y}$. This will prove that $B(\vec{y_1}; \frac{(1-\lambda)r_1}{\|A^{-1}\|} \subseteq V$ and so V is open.

We now prove the claim. To this end we define a function $\phi : B^- \to \underline{\mathbb{R}}^n$ by setting

$$\phi(\vec{x}) = \vec{x} + A^{-1}(\vec{y} - \vec{f}(\vec{x})).$$

We have that

$$\phi'(\vec{x}) = I_n - A^{-1}(\vec{f}'(\vec{x})) = A^{-1}(A - \vec{f}'(\vec{x})).$$

Hence,

$$\|\phi'(\vec{x})\| \leq \|A^{-1}\|\|A - \vec{f}'(\vec{x})\| \leq \lambda,$$

for any $\vec{x} \in B(\vec{x_0}; r)$, by the way that we chose r. Applying Theorem 2.20 again, we have that $\|\phi(\vec{x_3}) - \phi(\vec{x_2})\|_2 \leq \lambda\|\vec{x_3} - \vec{x_2}\|_2$. In particular ϕ is a contraction mapping.

We now wish to show that $\phi : B^- \to B^-$. First note that

$$\|\phi(\vec{x_1}) - \vec{x_1}\| = \|A^{-1}(\vec{y} - \vec{f}(\vec{x_1}))\| = \|A^{-1}(\vec{y} - \vec{y_1})\| < (1-\lambda)r_1.$$

Thus, for any $\vec{x} \in B^-$, we have that

$$\|\phi(\vec{x}) - \vec{x_1}\| \leq \|\phi(\vec{x}) - \phi(\vec{x_1})\|_2 + \|\phi(\vec{x_1}) - \vec{x_1}\|_2 \leq \\ \lambda\|\vec{x} - \vec{x_1}\|_2 + (1-\lambda)r_1 \leq \lambda r_1 + (1-\lambda)r_1 = r_1,$$

and so $\phi(\vec{x}) \in B^-$.

Since B^- is closed and bounded, it is a compact set by the Heine-Borel theorem. Thus, ϕ is a contraction mapping of a compact set back into itself. By the Contraction Mapping Theorem, ϕ has a unique fixed point. Thus, there is a point $\vec{x} \in B^-$, with

$$\vec{x} = \phi(\vec{x}) = \vec{x} + A^{-1}(\vec{y} - \vec{f}(\vec{x}).$$

This implies that $\vec{0} = A^{-1}(\vec{y} - \vec{f}(\vec{x}))$ and hence, $\vec{y} = \vec{f}(\vec{x})$.

This completes the proof of the claim and hence V is open.

We now prove (3). Pick $\vec{y} \in V$ and $\vec{y} + \vec{k} \in V$. Then there exist $\vec{x}$ and $\vec{x} + \vec{h}$ in $B(\vec{x_0}; r)$ with $\vec{y} = \vec{f}(\vec{x})$ and $\vec{y} + \vec{k} = \vec{f}(\vec{x} + \vec{h})$. By inequality (2.3), we have that

$$\|\vec{k}\|_2 = \|\vec{f}(\vec{x} + \vec{h}) - \vec{f}(\vec{x})\|_2 \geq \frac{(1-\lambda)}{\|A^{-1}\|}\|\vec{h}\|_2.$$

Thus, if $\vec{k} \to \vec{0}$, then $\vec{h} \to \vec{0}$.

Let $T = \big(\vec{f}'(\vec{x})\big)^{-1}$, then we have that

$$\frac{\|\vec{g}(\vec{y} + \vec{k}) - \vec{g}(\vec{y}) - T(\vec{k})\|_2}{\|\vec{k}\|_2} = \frac{\|(\vec{x} + \vec{h}) - \vec{x} - T\big(\vec{f}(\vec{x} + \vec{h}) - \vec{f}(\vec{x})\big)\|_2}{\|\vec{k}\|_2}$$

$$= \frac{\|\vec{h} + T(\vec{f}(\vec{x})) - T\big(\vec{f}(\vec{x} + \vec{h})\big)\|_2}{\|\vec{k}\|_2} = \frac{\|T\Big(\vec{f}'(\vec{x})\big(\vec{h}\big) + \vec{f}(\vec{x}) - \vec{f}(\vec{x} + \vec{h})\Big)\|_2}{\|\vec{k}\|_2} \leq$$

$$\frac{\|T\|\|E_f(\vec{h})\|_2}{\|\vec{k}\|_2} \leq \frac{\|T\|\|E_f(\vec{h})\|_2}{(1-\lambda)\|\vec{h}\|_2/\|A^{-1}\|} = \\ \frac{\|A^{-1}\|\|T\|}{1-\lambda}\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2}.$$

As $\vec{k} \to \vec{0}$, we have $\vec{h} \to \vec{0}$, and so by the differentiability of $\vec{f}$, as $\vec{k} \to \vec{0}$, we also have that $\frac{\|E_f(\vec{h})\|_2}{\|\vec{h}\|_2} \to 0$. Thus, given $\epsilon > 0$ we may choose $\|\vec{k}\|_2$ sufficiently small so that

$$\frac{\|E_f(\vec{h}\|_2}{\|\vec{h}\|_2} < \frac{(1-\lambda)\epsilon}{\|A^{-1}\|\|T\|}.$$

This proves that $\frac{\|E_g(\vec{k})\|_2}{\|\vec{k}\|_2} \to 0$ as $\vec{k} \to \vec{0}$, and so $\vec{g}$ is differentiable at $\vec{y}$ with

$$\vec{g}'(\vec{y}) = T = \left(\vec{f}'(\vec{x})\right)^{-1} = \left(\vec{f}'(\vec{g}(\vec{y}))\right)^{-1}.$$

Finally, to see that $\vec{g}$ is $\mathcal{C}^1$ on V, we note that since $\vec{g}$ is differentiable at each point in V it is continuous at each point in V.

Cramer's rule for the inverse of a matrix expresses each entry in the inverse as the quotient of one of the entries and a determinant of a submatrix. Since determinants are made up of sums and products of entries, if each component of a matrix is a continuous function, then each subdeterminant is a continuous function. When the matrix is invertible, then none of these subdeterminants can be equal to 0.

Hence, we see by Cramer's rule that if we are given a matrix of continuous functions and we know that the matrix is invertible, then each entry of the inverse matrix is also a continuous function

Thus, since $\vec{g}$ is continuous and $\vec{f}$ is $\mathcal{C}^1$, each entry of $\vec{f}'(\vec{g}(\vec{y}))$ is a continuous function of $\vec{y}$. By the Cramer's rule reasoning, each entry of $\vec{g}'(\vec{y}) = \left(\vec{f}'(\vec{g}(\vec{y}))\right)^{-1}$ is also a continuous function of $\vec{y}$. But the entries of $\vec{g}'(\vec{y})$ are the partial derivatives, $\frac{\partial g_i(\vec{y})}{\partial y_j}$. Thus, all the partials are continuous functions and so $\vec{g}$ is $\mathcal{C}^1$ on V. This completes the proof of the theorem. □

Problem 2.30. *Let* $\vec{f} : \mathbb{R}^3 \to \mathbb{R}^3$ *have component functions* $f_1(x_1, x_2, x_3) = 3x_1 + 2x_2 + x_3$, $f_2(x_1, x_2, x_3) = 2x_1x_2 + x_3^2$, $f_3(x_1, x_2, x_3) = 4x_1x_2x_3$. *Note that* $\vec{f}(1, 2, 0) = (7, 4, 0)$. *Use the inverse function theorem to compute a linear approximation to the vector* $\vec{x}$ *that solves* $\vec{f}(\vec{x}) = (7.1, 4.2, 0.3)$.

As a quick summary, we will often say that the inverse function theorem says that if $\vec{f}(\vec{x}) = \vec{y}$ is $\mathcal{C}^1$ and $\vec{f}'(\vec{x_0})$ is invertible, then locally we can solve for $\vec{x}$ as a function of $\vec{y}$.

2.7 The Multi-Variable Newton and Quasi-Newton Methods

We start with the motivation, which is very much like the 1-variable case. Suppose that $\vec{f} : \mathbb{R}^n \to \mathbb{R}^n$ and we wish to solve $\vec{f}(\vec{x}) = \vec{0}$. If we are given an approximate solution $\vec{x}_1$ and we wish to improve it a "better" approximate solution, $\vec{x}_2$, then we write

$$\vec{f}(\vec{x}_2) = \vec{f}(\vec{x}_1 + (\vec{x}_2 - \vec{x}_1)) \approx \vec{f}(\vec{x}_1) + \vec{f}'(\vec{x}_1)(\vec{x}_2 - \vec{x}_1).$$

Setting $\vec{f}(\vec{x}_2) = \vec{0}$ and solving for $\vec{x}_2$ yields

$$\vec{x}_2 = \vec{x}_1 - \left(\vec{f}'(\vec{x}_1)\right)^{-1}\vec{f}(\vec{x}_1).$$

Thus, **Newton's Method** is to start with an initial guess for the solution $\vec{x}_1$ and form the sequence,

$$\vec{x}_{n+1} = \vec{x}_n - \left(\vec{f}'(\vec{x}_n)\right)^{-1}\vec{f}(\vec{x}_n).$$

Of course, to form this sequence we will need that the matrix $\vec{f}'(\vec{x}_n)$ is invertible for all n.

Notice that at each step in this method, to compute $\vec{x}_{n+1}$ we need to compute the inverse of the matrix $\vec{f}'(\vec{x}_n)$. This is time consuming.

The **Quasi-Newton Method** is simpler to implement because it uses $\vec{f}'(\vec{x}_1)$ to approximate the remaining derivatives. Thus, this method only requires computing the inverse of a single matrix, namely, the inverse of the derivative at the initial point. Thus, the quasi-Newton method, is to start with an initial "guess" $\vec{x}_1$ and form the sequence

$$\vec{x}_{n+1} = \vec{x}_n - \left(\vec{f}'(\vec{x}_1)\right)^{-1}\vec{f}(\vec{x}_n).$$

Of course for both of these algorithms, success depends a lot on how good the initial "guess" is.

Theorem 2.31 (Quasi-Newton Method)**.** *Let $U \subseteq \mathbb{R}^n$ be an open set, let $\vec{f} : U \to \mathbb{R}^n$ and assume that there is $\vec{z} \in U$ such that $\vec{f}(\vec{z}) = \vec{0}$. If $\vec{f}$ is $\mathcal{C}^1$ on U and the matrix $\vec{f}'(\vec{z})$ is invertible, then there exists an $r > 0$, so that for any $x_1 \in B(\vec{z}; r)$ the sequence $\{\vec{x}_n\}$ generated by the Quasi-Newton Method converges to $\vec{z}$.*

Proof. First, since $\vec{f}$ is $\mathcal{C}^1$ and $\vec{f}'(\vec{z})$ is invertible, we may pick r_1 such that for $\vec{x} \in B(\vec{z}; r_1)$ the matrix $\vec{f}'(\vec{x})$ is invertible. Also, since all of the entries of this matrix are continuous functions of $\vec{x}$ by applying Cramer's Rule again, we see that the entries of $\left(\vec{f}'(\vec{x})\right)^{-1}$ and the norm of this matrix will also vary continuously with $\vec{x}$. Thus, we may find an $0 < r_2 \leq r_1$ such that for $\vec{x} \in B(\vec{z}; r_2)$ we have a bound, say,

$$\|\left(\vec{f}'(\vec{x})\right)^{-1}\| \leq M.$$

Finally, fix some number λ, $0 < \lambda < 1$ and again using the continuity of the partial derivatives, we may pick an $r_3 > 0$ so that when $\vec{x}, \vec{y} \in B(\vec{z}; r_3)$ then $\|\vec{f}'(\vec{x}) - \vec{f}'(\vec{y})\| < M^{-1}\lambda$.

Set $r = \min\{r_1, r_2, r_3\}$. We claim that for any intial $\vec{x}_1 \in B(\vec{z}; r)$ the sequence generated by the quasi-Newton algorithm converges to $\vec{z}$.

To simplify notation let $A = \vec{f}'(\vec{x}_1)$, so that

$$x_{n+1} = x_n - A^{-1}\vec{f}(\vec{x}_n).$$

We must show that this sequence converges to $\vec{z}$.

Now define $\vec{g}(\vec{x}) = x - A^{-1}\vec{f}(\vec{x})$, so that

$$\vec{x}_{n+1} = \vec{g}(\vec{x}_n).$$

Note that $\vec{g}(\vec{x}) = \vec{x}$ if and only if $\vec{f}(\vec{x}) = \vec{0}$.

Then

$$\vec{g}'(\vec{x}) = I - A^{-1}\vec{f}'(\vec{x}) = A^{-1}[A - \vec{f}'(\vec{x})].$$

Hence, for any $\vec{x} \in B(\vec{z}; r)$,

$$\|\vec{g}'(\vec{x})\| \leq \|A^{-1}\| \, \|A - \vec{f}'(\vec{x})\| \leq MM^{-1}\lambda = \lambda < 1.$$

Thus, by the corollary to the Multivariable Mean Value Theorem, for $\vec{x}, \vec{y} \in B(\vec{z}; r)$ we have that

$$\|\vec{g}(\vec{x}) - \vec{g}(\vec{y})\|_2 \leq \lambda\|\vec{x} - \vec{y}\|_2,$$

that is $\vec{g}$ is a strict contraction. Also,

$$\|\vec{g}(\vec{x}) - \vec{z}\|_2 = \|\vec{g}(\vec{x}) - \vec{g}(\vec{z})\|_2 \leq \lambda\|\vec{x} - \vec{z}\|_2 < r.$$

So $\vec{g}$ maps $B(\vec{z}; r)$ back into itself. Finally, as in the proof of the contraction mapping theorem, we have

$$\|\vec{x}_{n+1} - \vec{z}\|_2 = \|\vec{g}(\vec{x}_n) - \vec{g}(\vec{z})\|_2 \leq \lambda\|\vec{x}_n - \vec{z}\|_2.$$

Thus, by induction, we have that

$$\|\vec{x}_{n+1} - \vec{z}\|_2 \leq \lambda^n\|\vec{x}_1 - \vec{z}\|_2 \to 0,$$

and the proof is complete. □

Remark 2.32. *Note that the proof also shows that* $\vec{z}$ *is the unique point in* $B(\vec{z}; r)$ *that satisfies* $\vec{f}(\vec{x}) = \vec{0}$. *To see this note that if* $\vec{x}_1$ *was any point satisfying* $\vec{f}(\vec{x}_1) = \vec{0}$, *then using this point for our initial guess, we would have that* $\vec{x}_n = \vec{x}_1$ *fo all* n, *while at the same time,* $\vec{x}_n \to \vec{z}$.

Problem 2.33. *Let* $\vec{f}(x, y) = (x^2 - y, y^2 + 3x - 1)$. *Prove that there is some point* $\vec{z} = (a, b)$ *with* $0 \le a, b \le 1$ *satisfying* $\vec{f}(\vec{z}) = (0, 0)$. *Pick* $\vec{x}_1 = (1, 1)$ *and compute* $\vec{x}_2, \vec{x}_3$ *for both the quasi-Newton method and Newton method.*

2.8 The Implicit Function Theorem

Given a set of n equations in $n + m$ unknowns, we would like to be able to express n of the unknowns, called the *dependent variables*, in terms of m of the unknowns, called either the *free variables* or *independent variables.* The implicit function theorem gives us conditions under which we can do this locally and gives us a formula for computing the partial derivatives of the dependent variables in terms of the independent variables.

For example, from calculus we know that on the sphere given by the equation

$$x^2 + y^2 + z^2 = 9,$$

we can define the dependent variable z as a function of x and y, except when $z = 0$. Our calculus approach would be to write

$$\frac{\partial}{\partial x}[x^2 + y^2 + z^2] = \frac{\partial}{\partial x}[9]$$

and hence,

$$2x + 0 + 2z\frac{\partial z}{\partial x} = 0,$$

implying

$$\frac{\partial z}{\partial x} = -\frac{x}{z}.$$

The reason that we get 0 for the second term in this sum, was that x and y were both considered as independent variables and hence, $\frac{\partial y}{\partial x} = 0$. Similarly, $\frac{\partial z}{\partial y} = -\frac{y}{z}$.

Thus, at the point $(2, 1, -2)$, we would have that $\frac{\partial z}{\partial x} = 1$, $\frac{\partial z}{\partial y} = 1/2$. What these numbers tell us is that for nearby points $(2 + h_1, 1 + h_2, z)$ on the sphere we know that a good linear approximation to z is

$$z \cong -2 + 1h_1 + (1/2)h_2.$$

In this case we can see directly why $z = 0$ is problematical. For $z > 0$, we can actually sove this equation explicitly and $z = +\sqrt{9 - x^2 - y^2}$, while for $z < 0$, $z = -\sqrt{9 - x^2 - y^2}$. Thus, in any neighborhood of a point of the form $(x_0, y_0, 0)$ for each nearby (x, y) there are two nearby z's that satisfy the equation and so z is not uniquely determined as a function of (x, y).. In this case, we can actually picture this quite easily, since $(x_0, y_0,)$ is a point on the equator of this sphere and in any neighborhood of such a point, for a given nearby (x, y) there is a $z > 0$ "above" the equator and $z < 0$ "below" the equator.

For a second example, we consider the pair of equations,

$$f_1(x, y, z) = x^2 + y^2 + z^2 = 9 \text{ and } f_2(x, y, z) = 3x + 4y + 2z = 6,$$

which represent the intersection of the sphere with a plane yielding a circle on the sphere. This "one-dimensional" object should have only one "free variable". So in this case we might like to regard z and y as dependent variables and compute their derivatives with respect to the independent variable x.

Computing $\frac{d}{dx}$ of the pair of equations, yields

$$2x + 2y\frac{dy}{dx} + 2z\frac{dz}{dx} = 0 \text{ and } 3 + 4\frac{dy}{dx} + 2\frac{dz}{dx} = 0,$$

which can be expressed in matrix vector form as

$$\begin{pmatrix} 2x \\ 3 \end{pmatrix} + \begin{pmatrix} 2y & 2z \\ 4 & 2 \end{pmatrix} \begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Note that the first vector is $\frac{d\vec{f}}{dx}$ the matrix is $\vec{f}'$ if we only regard $\vec{f}$ as a function of y, z and the second vector is $\frac{d}{dx}(y, z)^t$.

At the point $(2, 1, -2)$ this matrix is invertible and we obtain,

$$\begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix} = -\begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \frac{-1}{20}\begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 3 \end{pmatrix} = \frac{-1}{20}\begin{pmatrix} 20 \\ -10 \end{pmatrix} = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}.$$

What these numbers tell us is that if near $(2, 1, -2)$ we wish to find a point on this circle on the sphere of the form $(2 + h, y, z)$ then a good linear approximation is given by

$$y \cong 1 - h \text{ and } z \cong -2 + h/2.$$

Note that these calculations tell us a great deal about this circle that would be difficult to see directly. For example, the above matrix is not invertible when $y - 4z = 0$. So perhaps if we tried to explicitly solve the above pair of equations for y and z as functions of x near a point of the form $(x_0, 4b, b)$ then we would find that there were multiple solutions near such points and so y and z cannot be expressed as functions, or perhaps the tangent line to the circle is vertical in either the y-direction or z-direction and so no derivative exists.

In the general case of the implicit function theorem, we will want to examine n equations $f_1, ..., f_n$ in $m + n$ variables and regard n of these variables as depending on m independent variables. To keep our notation close to what we encounter in calculus, we shall write $\vec{y} = (y_1, .., y_n)$ for the dependent variables and $\vec{x} = (x_1, ...x_m)$ for the independent variables. Our goal is to consider the equations,

$$f_1(\vec{x}, \vec{y}) = c_1, ..., f_n(\vec{x}, \vec{y}) = c_n,$$

where $c_1, ..., c_n$ are constants and for a given point $(\vec{x_0}, \vec{y_0})$ that satisfies these equations to compute

$$\frac{\partial y_i}{\partial x_j}, 1 \leq i \leq n, 1 \leq j \leq m.$$

If we write this in vector notation, with $\vec{f} = (f_1, ..., f_n)$, then we have that $\vec{f} : \mathbb{R}^{m+n} \to \mathbb{R}^n$, with a typical vector in $\mathbb{R}^{m+n}$ written as $(\vec{x}, \vec{y})$, with $\vec{x} \in \mathbb{R}^m$, and $\vec{y} \in \mathbb{R}^n$. Thus, our equation becomes $\vec{f}(\vec{x}, \vec{y}) = \vec{c}$ and we know that $\vec{f}(\vec{x_0}, \vec{y_0}) = \vec{c}$.

Note that the derivative of $\vec{f}$ will be an $n \times (m + n)$ matrix. The first m columns of this matrix will involve derivatives with respect to the x variables and the next n columns will be derivatives with respect to the y variables. Thus,

$$\vec{f}'(\vec{x}, \vec{y}) = \Big[\frac{\partial \vec{f}}{\partial x_1}, ..., \frac{\partial \vec{f}}{\partial x_m}, \frac{\partial \vec{f}}{\partial y_1}, ..., \frac{\partial \vec{f}}{\partial y_n}\Big].$$

To simplify notation further, we will write

$$\vec{f}'(\vec{x}, \vec{y}) = \Big[\frac{\partial \vec{f}}{\partial \vec{x}}, \frac{\partial \vec{f}}{\partial \vec{y}}\Big],$$

where

$$\frac{\partial \vec{f}}{\partial \vec{x}}$$

is the $n \times m$ matrix consisting of the first m columns of $\vec{f}'$ and

$$\frac{\partial \vec{f}}{\partial \vec{y}}$$

is the $n \times n$ matrix consisting of the last n columns of $\vec{f}'$. In this notation, our goal is to find the $n \times m$ matrix

$$\frac{\partial \vec{y}}{\partial \vec{x}} = \Big(\frac{\partial y_i}{\partial x_j}\Big).$$

This notation really carries the day, because what we will show is that, in matrix notation using matrix products,

$$\frac{\partial \vec{f}}{\partial \vec{x}} + \frac{\partial \vec{f}}{\partial \vec{y}}\frac{\partial \vec{y}}{\partial \vec{x}} = 0.$$

Hence, when the $n \times n$ matrix $\frac{\partial \vec{f}}{\partial \vec{y}}$ is invertible, we expect to find that

$$\frac{\partial \vec{y}}{\partial \vec{x}} = -\Big(\frac{\partial \vec{f}}{\partial \vec{y}}\Big)^{-1}\frac{\partial \vec{f}}{\partial \vec{x}}.$$

We are now prepared to prove the implicit function theorem. First two lemmas about matrices.

Lemma 2.34. *Let* $A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$ *be a* $(n_1 + n_2) \times (p_1 + p_2)$ *matrix partitioned into submatrices* $A_{i,j}$ *of size* $n_i \times p_j$ *and let* $B = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix}$ *be a* $(p_1 + p_2) \times (m_1 + m_2)$ *matrix partioned into submatrices* $B_{i,j}$ *of size* $p_i \times m_j$. *Then*

$$A \cdot B = \begin{pmatrix} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} & A_{1,1}B_{1,2} + A_{1,2}B_{2,2} \\ A_{2,1}B_{1,1} + A_{2,2}B_{2,1} & A_{2,1}B_{1,2} + A_{2,2}B_{2,2} \end{pmatrix}.$$

Lemma 2.35. *Let* $A = \begin{pmatrix} A_{1,1} & 0 \\ A_{2,1} & A_{2,2} \end{pmatrix}$ *be a* $(n_1 + n_2) \times (n_1 + n_2)$ *matrix partitioned into submatrices of size* $n_i \times n_j$. *If* $A_{1,1}$ *and* $A_{2,2}$ *are invertible, then* A *is invertible and*

$$A^{-1} = \begin{pmatrix} A_{1,1}^{-1} & 0 \\ -A_{2,2}^{-1}A_{2,1}A_{1,1}^{-1} & A_{2,2}^{-1} \end{pmatrix}.$$

Theorem 2.36 (Implicit Function Theorem)**.** *Let* $U \subseteq \mathbb{R}^{m+n}$ *and write points in* U *as* $(\vec{x}, \vec{y})$ *with* $\vec{x} \in \mathbb{R}^m$, $\vec{y} \in \mathbb{R}^n$, *let* $\vec{f} : \mathbb{R}^{m+n} \to \mathbb{R}^n$ *be* $\mathcal{C}^1$ *on* U, *let* $(\vec{x_0}, \vec{y_0}) \in U$ *with* $\vec{f}(\vec{x_0}, \vec{y_0}) = \vec{c}$. *If the* $n \times n$ *matrix*

$$\Big(\frac{\partial \vec{f}}{\partial \vec{y}}(\vec{x_0}, \vec{y_0})\Big)$$

is invertible, then there is an $r > 0$ *and a* $\mathcal{C}^1$ *function* $\vec{g} : B(\vec{x_0}; r) \to \mathbb{R}^n$ *with* $\vec{g}(\vec{x_0}) = \vec{y_0}$ *such that*

$$\vec{f}(\vec{x}, \vec{g}(\vec{x})) = \vec{c}$$

and

$$\vec{g}'(\vec{x}) = -\Big(\frac{\partial \vec{f}}{\partial \vec{y}}(\vec{x}, \vec{g}(\vec{x}))\Big)^{-1}\Big(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}, \vec{g}(\vec{x}))\Big),$$

for all $\vec{x} \in B(\vec{x_0}; r)$.

Proof. Let $A_{2,1}$ be the $n \times m$ matrix $\Big(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x_0}, \vec{y_0})\Big)$ and let $A_{2,2}$ be the $n \times n$ matrix $\Big(\frac{\partial \vec{f}}{\vec{y}}(\vec{x_0}, \vec{y_0})\Big)$. Define $\vec{F} : U \to \mathbb{R}^{m+n}$ by

$$\vec{F}(\vec{x}, \vec{y}) = \begin{pmatrix} \vec{x} \\ \vec{f}(\vec{x}, \vec{y}) \end{pmatrix}.$$

Then $\vec{F}$ is $\mathcal{C}^1$ on U and its derivative is an $(m+n) \times (m+n)$ matrix, writing this matrix in block form we have

$$\vec{F'}(\vec{x}, \vec{y}) = \begin{pmatrix} I_m & 0 \\ \Big(\frac{\partial \vec{f}}{\partial \vec{x}}\Big) & \Big(\frac{\partial \vec{f}}{\partial \vec{y}}\Big) \end{pmatrix}.$$

In particular,

$$\vec{F'}(\vec{x_0}, \vec{y_0}) = \begin{pmatrix} I_m & 0 \\ A_{2,1} & A_{2,2} \end{pmatrix}.$$

By Lemma 2.32 this matrix is invertible, and so we may apply the inverse function theorem to conclude that there exists $r_1 > 0$ so that $\vec{F}$ is on-to-one on $B = B((\vec{x_0}, \vec{y_0}); r_1) \subseteq \mathbb{R}^{m+n}$, $V = \vec{F}(B)$ is open and the inverse of $\vec{F}$, $\vec{G} : V \to B$ is $\mathcal{C}^1$ on V with the formula for the derivative of $\vec{G}$ as given in the theorem.

Note that since $\vec{F}(\vec{x}, \vec{y}) = (\vec{x}, \vec{f}(\vec{x}, \vec{y}))$, we have that if $(\vec{x}, \vec{z}) \in V$, then

$$\vec{G}(\vec{x}, \vec{z}) = (\vec{x}, \vec{g_1}(\vec{x}, \vec{z})),$$

for some function $\vec{g_1} : V \to \mathbb{R}^n$. Moreover, $\vec{g_1}$ will also be $\mathcal{C}^1$ on V since it is just the last n components of $\vec{G}$.

Since V is open and $\vec{F}(\vec{x_0}, \vec{y_0}) = (\vec{x_0}, \vec{c}) \in V$, there is $r > 0$, so that $\|\vec{x} - \vec{x_0}\|_2 < r$ implies that $(\vec{x}, \vec{c}) \in V$. Thus, on $B(\vec{x_0}; r) \subseteq \mathbb{R}^m$, we may define a function

$$\vec{g} : B(\vec{x_0}; r) \to \mathbb{R}^n \text{ by } \vec{g}(\vec{x}) = \vec{g_1}(\vec{x}, \vec{c}).$$

Since $\vec{F} \circ \vec{G}(\vec{x}, \vec{z}) = (\vec{x}, \vec{z})$, we have for $\vec{x} \in B(\vec{x_0}; r)$,

$$\Big(\vec{x}, \vec{c}\Big) = \vec{F} \circ \vec{G}(\vec{x}, \vec{c}) = \vec{F}(\vec{x}, \vec{g}(\vec{x})) = \Big(\vec{x}, \vec{f}\big(\vec{x}, \vec{g}(\vec{x})\big)\Big),$$

and it follows that

$$\vec{f}\big(\vec{x}, \vec{g}(\vec{x})\big) = \vec{c}.$$

Finally, to find the derivative formula, first note that

$$\vec{g'}(\vec{x}) = \big(\frac{\partial \vec{g_1}}{\partial x}(\vec{x}, \vec{c})\big),$$

and since $\vec{F} \circ \vec{G}(\vec{x}, \vec{z}) = (\vec{x}, \vec{z})$ by the chain rule

$$\begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix} = \vec{F'}\big(\vec{G}(\vec{x}, \vec{c})\big)\vec{G'}(\vec{x}, \vec{c}) =$$
$$\begin{pmatrix} I_m & 0 \\ \Big(\frac{\partial \vec{f}}{\partial x}(\vec{x}, \vec{g_1}(\vec{x}, \vec{c}))\Big) & \Big(\frac{\partial \vec{f}}{\partial \vec{y}}(\vec{x}, \vec{g_1}(\vec{x}, \vec{c}))\Big) \end{pmatrix} \begin{pmatrix} I_m & 0 \\ \Big(\frac{\partial \vec{g_1}}{\partial \vec{x}}(\vec{x}, \vec{c})\Big) & \Big(\frac{\partial \vec{g_1}}{\partial \vec{z}}(\vec{x}, \vec{c})\Big) \end{pmatrix}$$

Examining the (2,1)-entries of these block matrices, yields that

$$0 = \Big(\frac{\partial \vec{f}}{\partial x}\big(\vec{x}, \vec{g_1}(\vec{x}, \vec{c})\big)\Big) \cdot I_m + \Big(\frac{\partial \vec{f}}{\partial \vec{y}}\big(\vec{x}, \vec{g_1}(\vec{x}, \vec{c})\big)\Big) \cdot \Big(\frac{\partial \vec{g_1}}{\partial \vec{x}}(\vec{x}, \vec{c})\Big) =$$
$$\Big(\frac{\partial \vec{f}}{\partial x}\big(\vec{x}, \vec{g}(\vec{x})\big)\Big) + \Big(\frac{\partial \vec{f}}{\partial \vec{y}}\big(\vec{x}, \vec{g}(\vec{x})\big)\Big) \cdot \vec{g'}(\vec{x}).$$

Solving for $\vec{g'}$ yields

$$\vec{g'}(\vec{x}) = -\Big(\frac{\partial \vec{f}}{\partial \vec{y}}(\vec{x}, \vec{g}(\vec{x}))\Big)^{-1} \cdot \Big(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}, \vec{g}(\vec{x}))\Big).$$

□

Using our simplified notation, the function $\vec{g}$ is "solving" for the "y" variables, so that $y_i = g_i(\vec{x})$. Thus, this last equation(ignoring the variables) is

$$\big(\frac{\partial y_i}{\partial x_j}\big) = \big(\frac{\partial g_i}{\partial x_j}\big) = \vec{g'} = -\big(\frac{\partial \vec{f}}{\partial \vec{y}}\big)^{-1} \cdot \big(\frac{\partial \vec{f}}{\partial \vec{x}}\big).$$

Problem 2.37. *Consider the equations*

$$u^5 + xv^2 - y + w = 0,$$
$$v^5 + yu^2 - x + w = 0,$$
$$w^4 + y^5 - x^4 - 1 = 0$$

which are satisfied by $x = 1, y = 1, u = 1, v = 1, w = -1$. *Show that the hypotheses of the implicit function are met allowing us to express* $u = g_1(x, y), v = g_2(x, y), w = g_3(x, y)$ *in a neighborhood of* $(x_0, y_0) = (1, 1)$.

- *Use the implicit function theorem to compute the 6 partial derivatives of the g's with respect to* x *and* y *at* $(1, 1)$.

- *Use these partials to give approximate values of* u, v, w *that satisfy these equations for* $x = 1.1$ *and* $y = 0.8$.

Problem 2.38. *Consider the equations,*

$$xu^2 + yv^2 + xy - 9 = 0,$$
$$xv^2 + yu^2 - xy - 7 = 0.$$

Assuming that a point (x_0, y_0, u_0, v_0) *satisfies these equations, what conditions on* x_0, y_0 *are needed to guarantee that the hypostheses of the implicit function theorem are met allowing us to express* u, v *as functions of* x, y *near* x_0, y_0*.? Repeat the exercise, for expressing* x, y *as functions of* u, v.

2.9 Local Extrema and the Second Derivative Test

Definition 2.39. *Let* $U \subseteq \mathbb{R}^n$ *and let* $f : U \to \mathbb{R}$. *A point* $\vec{x} \in U$ *is called a* **local maximum** *if there is an* $r > 0$ *so that if* $\|\vec{y} - \vec{x}\| < r$ *then* $f(\vec{x}) \geq f(\vec{y})$ *and a* **strict local maximum** *if there is an* $r > 0$ *so that* $f(\vec{x}) > f(\vec{y})$ *for* $0 < \|\vec{x} - \vec{y}\| < r$. *Similarly,* $\vec{x}$ *is called a* **local minimum***(respectively,* **strict local minimum***) if* $f(\vec{x}) \leq f(\vec{y})$ *for* $\|\vec{x} - \vec{y}\| < r$ *(respectively,* $f(\vec{x}) < f(\vec{y})$ *for* $0 < \|\vec{x} - \vec{y}\| < r$*). The point* $\vec{x}$ *is called a* **local extrema***(respectively,* **strict local extrema***) if it is either a local maximum or local minimum(respectively, strict local maximum or strict local minimum).*

Proposition 2.40 (1st Derivative Test)**.** *Let* $U \subseteq \mathbb{R}^n$ *be open, let* $f : U \to \mathbb{R}$ *and let* $\vec{x} \in U$ *be a local extrema. If* f *is differentiable at* $\vec{x}$ *then* $f'(\vec{x}) = \vec{0}$.

Proof. Say that $f(\vec{x}) \geq f(\vec{y})$ for all $\vec{y} \in B(\vec{x}; r)$. Let $\vec{u}$ be any unit vector and define $g : (-r, +r) \to \mathbb{R}^n$ by $g(t) = \vec{x} + t\vec{u}$ and $h : (-r, +r) \to \mathbb{R}$ by $h(t) = f(g(t))$. Then h has a local max at $t = 0$ and so by the one-variable theorem and chain rule,

$$0 = h'(0) = f'(g(0)) \cdot g'(0) = f'(\vec{x}) \cdot \vec{u}.$$

This shows that $f'(\vec{x})$ is a vector whose dot product with every unit vector is 0. Hence, $f'(\vec{x}) = \vec{0}$. □

We now wish to prove the analogue of the one-variable second derivative test. For this we will need a few new concepts.

Definition 2.41. *Let $P = (p_{i,j}) = P^t$ be an $n \times n$ matrix. Then P is called* **positive definite** *provided that for every non-zero vector $\vec{v} = (v_1, ..., v_n)$ we have that*

$$(P\vec{v}) \cdot \vec{v} = \sum_{i,j=1}^{n} p_{i,j} v_i v_j > 0.$$

One needs to be careful when using other books. Some only demand that the " $>$ " be a "$\geq$" and use the phrase "strictly positive definite" instead.

There are many theorems in linear algebra characterizing these matrices. The following will be useful for us.

Given $P = (p_{i,j})$ an $n \times n$ matrix for each k, $1 \leq k \leq n$, we let P_k be the $k \times k$ matrix

$$P_k = (p_{i,j})_{i,j=1}^{k}.$$

Theorem 2.42. *Let $P = P^t$ be an $n \times n$ matrix. Then P is positive definite if and only if*

$$det(P_k) > 0, \ \text{for } 1 \leq k \leq n.$$

Recall the idea of higher order partial derivatives. The **order** is the total number of partial derivatives that one takes.

Definition 2.43. *Let $U \subseteq \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}$. We say that f is $\mathcal{C}^k$* **on** *U if all the partial derivatives up to order k exist and are continuous on U.*

Theorem 2.44 (Equality of Mixed Partials)**.** *Let $U \subseteq \mathbb{R}^n$ and let $f : U \to \mathbb{R}$. If f is $\mathcal{C}^2$ on U, then for any $1 \leq i, j \leq n$,*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Proof. It is enough to consider the case $n = 2$ and to simplify we'll label the variables x and y. Suppose that there is a point where they are not equal, then one is strictly larger and by continuity of the partials, the larger one stays strictly larger on an open set. So say that $\frac{\partial^2 f}{\partial x \partial y} > \frac{\partial^2 f}{\partial y \partial x}$ on an open set $V \subseteq U$.

Pick a closed rectangle contained in V, $[a, b] \times [c, d] \subseteq V$. Then we have that

$$\int_a^b \int_c^d \frac{\partial^2 f}{\partial x \partial y} dy dx > \int_a^b \int_c^d \frac{\partial^2 f}{\partial y \partial x} dy dx.$$

Starting with the 2nd integral we have that

$$\int_a^b \int_c^d \frac{\partial^2 f}{\partial y \partial x}(x, y) dy dx = \int_a^b \Big(\int_c^d \frac{\partial}{\partial y}[\frac{\partial f}{\partial x}(x, y)] dy \Big) dx =$$
$$\int_a^b \Big(\frac{\partial f}{\partial x}(x, d) - \frac{\partial f}{\partial x}(x, c) \Big) dx = \big(f(b, d) - f(a, d) \big) - \big(f(b, c) - f(a, c) \big)$$

However, after first reversing the order of integration, one gets the same answer for the first integral.

Contradicting the strict inequality. □

Definition 2.45. *Let $U \subseteq \mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be twice differentiable. Then the $n \times n$ matrix*

$$H(f) = \Big(\frac{\partial^2 f}{\partial x_i \partial x_j} \Big)$$

is called the **Hessian of f.**

By the last result if f is $\mathcal{C}^2$ on U then the $H(f) = H(f)^t$.

Theorem 2.46 (Second Derivative Test)**.** *Let $U \subseteq \mathbb{R}^n$ be open, let $\vec{x}_0 \in U$ and let $f : U \to \mathbb{R}$. If*

- $f'(\vec{x}_0) = \vec{0}$,
- *f is $\mathcal{C}^2$ on a neighborhood of $\vec{x}_0$,*
- *$H(f)(\vec{x}_0)$ is positive definite,*

then $\vec{x}_0$ is a strict local minimum.

Remark 2.47. *To test for a local maximum, we just apply the above theorem to $-f$. Thus, $\vec{x}_0$ will be a local maximum if the first two conditions are met and $-H(f)(\vec{x}_0)$ is positive definite.*

We need two lemmas.

Lemma 2.48. *Let $h : (-r, +r) \to \mathbb{R}$ be twice differentiable on $(-r, +r)$ with $h'(0) = 0$. If $h''(s) > 0$ for all $|s| < r$, then $h(0) < h(t)$ for all $0 < |t| < r$.*

Proof. We apply Taylor's theorem to write

$$h(t) = h(0) + h'(0)t + \frac{h''(s)}{2}t^2 = h(0) + \frac{h''(s)}{2}t^2,$$

for some s between 0 and t. By our hypotheses the 2nd term is positive and so $h(t) > h(0)$. □

Lemma 2.49. *Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ be $\mathcal{C}^2$ on U, and let $\vec{x}_0 \in U$. If $H(f)(\vec{x}_0)$ is positive definite, then there is $r > 0$ such that $H(f)(\vec{x})$ is positive definite for $\vec{x} \in B(\vec{x}_0; r)$.*

Proof. To check if a matrix A is positive definite it is enough to check if $(A\vec{u}) \cdot \vec{u} > 0$ for every unit vector $\vec{u}$. But since the set of unit vectors is a compact subset and this is clearly a continuous function there will exist a minimum value $m > 0$ so that $(A\vec{u}) \cdot \vec{u} \geq m$. Now, we may pick a distance δ so that if B with $|a_{i,j} - b_{i,j}| < \delta$, then $(B\vec{u}) \cdot \vec{u} \geq m/2 > 0$.

Since f is $\mathcal{C}^2$ the entries of $H(f)$ all vary continuously and so we may pick an $r > 0$ so that if $\|\vec{x} - \vec{x}_0\| < r$ then

$$\Big| \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) \Big| < \delta.$$

□

Proof. Pick an r as given by the last lemma so that for $\|\vec{x} - \vec{x}_0\| < r$ we have that $H(f)(\vec{x})$ is positive definite.

Fix a unit vector $\vec{u} = (u_1, ..., u_n)$ define $g : (-r, +r) \to \mathbb{R}^n$ by $g(t) = \vec{x}_0 + t\vec{u}$ and let $h : (-r, +r) \to \mathbb{R}$ be $h(t) = f(g(t))$ as earlier. We have that

$$h'(t) = f'(g(t)) \cdot \vec{u} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(g(t))u_j.$$

Now when we apply the chain rule to differentiate each function in this sum we see that

$$h''(t) = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(g(t))u_i u_j \Big) = [H(f)(g(t))\vec{u}] \cdot \vec{u}.$$

Hence, $h''(t) > 0$ for all $|t| < r$.

By the one-variable second derivative test we have that

$$f(\vec{x}_0 + t\vec{u}) > f(\vec{x}_0),$$

for $0 < |t| < r$. Since this holds for every direction $\vec{u}$ we have that

$$f(\vec{x}) > f(\vec{x}_0)$$

whenever $\|\vec{x} - \vec{x}_0\| < r$. □

Definition 2.50. *A point $\vec{x}_0$ such that $f'(\vec{x}_0) = \vec{0}$ but neither $H(f)(\vec{x}_0)$ nor $-H(f)(\vec{x}_0)$ is positive definite is called a* **saddle point.**

Problem 2.51. *For $f(x,y,z) = (x^3-3x+4)(y^2+1)(z^2+1)$ find and classify all points where the derivative vanishes.*

Problem 2.52. *Let $f(x,y,z) = x^2 + (x-1)^2(y-2)^2 + (z-1)^2y$. Find and classify all the points where the deriviative vanishes.*

2.10 Taylor Series in Several Variables

First we introduce **multi-index notation,** which simplifies many formulas. By a multi-index in dimension n we mean a tuple $I = (i_1, ..., i_n)$ where each i_k is a non-negative integer. We set

- $|I| = i_1 + \cdots + i_n$,
- $x^I = x_1^{i_1} \cdots x_n^{i_n}$, which is a monomial in the variables $x_1, ..., x_n$,
- $I! = (i_1!) \cdots (i_n!)$.
- given a function$f$ of n variables,

$$\frac{\partial^{|I|} f}{\partial x^I} = \frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_n}}{\partial x_n^{i_n}} f$$

By a polynomial in n variables, we mean any finite sum of the form

$$p(x_1, ..., x_n) = p(\vec{x}) = \sum a_I x^I,$$

where a_I are real numbers. The **total degree** or more shortly, **degree** of such a polynomial is the maximum of $|I|$ over all non-zero terms. We can write the most general polynomial in n variables of degree at most M as

$$\sum_{|I|\leq M} a_I x^I.$$

Note that given a polynomial, $p(\vec{x}) = \sum a_I x^I$, we obtain the coefficient of x^I by

$$a_I = \frac{1}{I!}\frac{\partial^{|I|}p}{\partial x^I}(\vec{0}).$$

Thus, given an open set $U \subseteq \mathbb{R}^n$ a point $x_0 \in U$ and an infinitely differentiable function $f : U \to \mathbb{R}$ we define its **Taylor series centered at** x_0 to be

$$\sum_{|I|\geq 0} \frac{1}{I!}\Big(\frac{\partial^{|I|}f}{\partial x^I}(x_0)\Big)(x-x_0)^I,$$

where if $x_0 = (a_1, ..., a_n)$ and $x = (x_1, ..., x_n)$ then

$$(x-x_0)^I = (x_1 - a_1)^{i_1} \cdots (x_n - a_n)^{i_n}.$$

The **Taylor polynomial of degree** M is defined to be

$$\sum_{|I|\leq M} \frac{1}{I!}\Big(\frac{\partial^{|I|}f}{\partial x^I}(x_0)\Big)(x-x_0)^I.$$

The following theorem is called the **multi-variable Taylor Theorem with Lagrange form of the remainder.**

Theorem 2.53. *Let $r > 0$, let $\vec{x}_0 \in \mathbb{R}^n$ and let $f : B(\vec{x}_0; r) \to \mathbb{R}$ be $\mathcal{C}^{M+1}$ on $B(\vec{x}_0; r)$. Then for each $\vec{x} \in B(\vec{x}_0; r)$ there is $\vec{z}$ on the line segment joining $\vec{x}_0$ and $\vec{x}$ such that*

$$f(\vec{x}) = \sum_{|I|\leq M} \frac{1}{I!}\Big(\frac{\partial^{|I|}f}{\partial x^I}(\vec{x}_0)\Big)(x-x_0)^I + \sum_{|I|=M+1} \frac{1}{I!}\Big(\frac{\partial^{M+1}f}{\partial x^I}(\vec{z})\Big)(x-x_0)^I.$$

Proof. Set

$$g(t) = f(\vec{x}_0 + t(\vec{x} - \vec{x}_0)),$$

so that $g(0) = f(\vec{x}_0)$ and $g(1) = f(\vec{x})$. Also g is defined for $|t| < R = \frac{r}{\|\vec{x}-\vec{x}_0\|}$ where $R > 1$ and has $M + 1$ derivatives on the interval $(-R, +R)$.

Thus, by the one variable Taylor theorem,

$$f(\vec{x}) = g(1) = \sum_{j=0}^{M} \frac{g^{(j)}(0)}{j!} 1^j + \frac{g^{(M+1)}(s)}{(M+1)!},$$

for some $0 < s < 1$.

The remainder of the proof is done by showing that

$$g^{(j)}(t) = \sum_{|I|=j} \frac{j!}{I!} \Big(\frac{\partial^j f}{\partial x^I} (\vec{x}_0 + t(\vec{x} - \vec{x}_0)) \Big) (x - x_0)^I.$$

This last fact is a messy counting argument. We only illustrate it in the case that $n = 2$ and $j = 3$. Set $\vec{x}_0 = (a_1, a_2)$. So first notice that by the chain rule

$$g'(t) = \sum_{k=1}^{2} \Big(\frac{\partial f}{\partial x_k} (a_1 + t(x_1 - a_1), a_2 + t(x_2 - a_2)) \Big) (x_k - a_k).$$

Repeating the chain rule three times yields,

$$g^{(3)}(t) = \sum_{i,j,k=1}^{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} (a_1 + t(x_1 - a_1), a_2 + t(x_2 - a_2)) \Big) (x_i - a_i)(x_j - a_j)(x_k - a_k).$$

Now comes the counting part. For example if $I = (2,1)$ then $(\vec{x} - \vec{x}_0)^I = (x_1 - a_1)^2 (x_2 - a_2)^1$ and we need to count how many times this term appears in the above sum. Well we have 3 factors and we need to pick 2 of them to be $(x_1 - a_1)$ and 1 of them to be $(x_2 - a_2)$ we can do this exactly

$$\binom{3}{2} = \frac{3!}{(2!)(1!)} = \frac{j!}{I!}$$

ways!

In the general case our sum for the jth deriviative will involve all possible products of j terms involving the factors $(x_1 - a_1), ..., (x_n - a_n)$ and to compute the coefficient of $(x - x_0)^I$ we need to compute how many of these products will be equal to this term. So among the j factors we need to pick $(x_1 - a_1)$ exactly i_1 times, $(x_2 - a_2)$ exactly i_2 times, etc. The answer to this combinatorial problem is exactly

$$\frac{j!}{(i_1!) \cdots (i_n!)} = \frac{j!}{I!}.$$

□

Problem 2.54. *For $f(x, y) = cos(x^2y)$ compute the Taylor polynomial of degree 2 centered at $\vec{x}_0 = (0, 0)$ and at $\vec{x}_0 = (0, 1)$.*

Problem 2.55. *Repeat the above problem for $f(x, y) = e^x + sin(y)$.*