# CONTENTS

1. **Stone’s Theorem And Partitions Of Unity**
   - 1.1 Partitions Of Unity And Stone’s Theorem .................................. 83
   - 1.2 An Extension Theorem, Retracts .................................................. 87
   - 1.3 Something Which Is Not A Retract ............................................. 92
   - 1.4 Exercises ..................................................................................... 97

2. **The Bochner Integral**
   - 2.1 Strong And Weak Measurability ............................................... 99
   - 2.2 The Essential Bochner Integral ................................................... 107
   - 2.3 The Spaces $L^p(\Omega; X)$ ......................................................... 112
   - 2.4 Measurable Representatives ....................................................... 116
   - 2.5 Vector Measures ........................................................................ 118
   - 2.6 The Riesz Representation Theorem ............................................ 123
   - 2.7 Pointwise Behavior Of Weakly Convergent Sequences .................. 129
   - 2.8 Test Functions ............................................................................ 130
   - 2.9 Exercises .................................................................................... 131

3. **The Derivative**
   - 3.1 Limits Of A Function .................................................................. 135
   - 3.2 Basic Definitions ........................................................................ 139
   - 3.3 The Chain Rule ........................................................................... 141
   - 3.4 The Derivative Of A Compact Mapping ........................................ 141
   - 3.5 The Matrix Of The Derivative ..................................................... 142
   - 3.6 A Mean Value Inequality .............................................................. 144
   - 3.7 Higher Order Derivatives ............................................................. 150
   - 3.8 The Derivative And The Cartesian Product ................................... 151
   - 3.9 Mixed Partial Derivatives ............................................................. 155
   - 3.10 Implicit Function Theorem ......................................................... 157
   - 3.11 More Derivatives ....................................................................... 164
   - 3.12 Lyapunov Schmidt Procedure ................................................... 165
   - 3.13 Analytic Functions .................................................................... 169
   - 3.14 Ordinary Differential Equations ................................................ 176
   - 3.15 Tensor Products
     - 3.15.1 The Norm In Tensor Product Space ...................................... 182
     - 3.15.2 The Taylor Formula And Tensors .......................................... 185
   - 3.16 Exercises ................................................................................... 186

4. **Fixed Point Theorems And More**
   - 4.1 The Schauder Fixed Point Theorem ............................................. 191
   - 4.2 The Kakutani Fixed Point Theorem ............................................. 197
   - 4.3 A Variational Principle of Ekeland ............................................... 200
     - 4.3.1 Caristie Fixed Point Theorem ............................................... 203
     - 4.3.2 A Density Result .................................................................. 205
   - 4.4 Critical Points ............................................................................. 207
     - 4.4.1 Mountain Pass Theorem In Hilbert Space .............................. 207
## CONTENTS

7.4.2 A Locally Lipschitz Selection, Pseudogradients 212
7.4.3 Mountain Pass Theorem In Banach Space 219
7.5 The min max Theorem 223
7.6 Exercises 228

## 8 Degree Theory, An Introduction 235
8.1 Sard's Lemma 236
8.2 Preliminary Results 238
8.3 Definitions And Elementary Properties
  8.3.1 The Degree For $C^\infty (\Omega; R^n)$ 243
  8.3.2 Definition Of The Degree For Continuous Functions 247
8.4 Borsuk's Theorem 251
8.5 Applications 254
8.6 The Product Formula 258
8.7 A Function With Values In Smaller Dimensions 269
8.8 The Leray Schauder Degree 274
8.9 Exercises 284

## 9 Nonlinear Operators In Banach Space 291
9.1 Some Nonlinear Single Valued Operators 292
9.2 Duality Maps 301
9.3 Penalization And Projection Operators 308
9.4 Set-Valued Maps, Pseudomonotone Operators 311
9.5 Generalized Gradients 319
9.6 Maximal Monotone Operators
  9.6.1 Maximal Monotone Operators 323
  9.6.2 Surjectivity Theorems 331
  9.6.3 Approximation Theorems 341
  9.6.4 Sum Of Maximal Monotone Operators 349
  9.6.5 Convex Functions, An Example 354
9.7 Perturbation Theorems 368
9.8 Exercises 380

## 10 Maximal Monotone Operators, Hilbert Space 381
10.1 Lipschitz Functions 381
10.2 Basic Theory 384
10.3 Evolution Inclusions 391
10.4 Subgradients
  10.4.1 General Results 397
  10.4.2 Hilbert Space 409
10.5 A Perturbation Theorem 415
10.6 An Evolution Inclusion 417
10.7 A More Complicated Perturbation Theorem 420
10.8 An Evolution Inclusion 422
## CONTENTS

### 11 The Yankov von Neumann Aumann theorem 431
- 11.1 Multifunctions And Their Measurability 442
  - 11.1.1 The General Case 442
  - 11.1.2 A Special Case Which Is Easier 445

### A Covering Theorems 447

### B Banach Spaces 453
- B.1 Introduction 453
- B.2 Theorems Based On Baire Category 453
  - B.2.1 Baire Category Theorem 453
  - B.2.2 Uniform Boundedness Theorem 457
  - B.2.3 Open Mapping Theorem 458
  - B.2.4 Closed Graph Theorem 460
- B.3 Quotient Spaces 462
- B.4 Operators With Closed Range 464
- B.5 Closed Subspaces 471

### C Hilbert Spaces 475
- C.1 Basic Theory 475
- C.2 Approximations In Hilbert Space 481
- C.3 Orthonormal Sets 484

### D The Hausdorff Maximal Theorem 487
- D.1 The Hamel Basis 491
- D.2 Exercises 492
Chapter 1

Metric Spaces And General Topological Spaces

1.1 Introduction

This first part of the book features fundamental topological ideas, many of which are necessary to do a decent job in presenting nonlinear analysis. In addition, much of it is part of the topic of nonlinear analysis. It is assumed the reader has had a typical advanced calculus course including functions of many variables. The middle section has the essential ideas in linear functional analysis which are sometimes needed. Other distinctive topics are covered in the exercises. The emphasis is on real spaces throughout.

1.2 Metric Space

It is assumed the reader knows about metric spaces already. However, this chapter will include the most important basic properties including some which were not considered earlier.

**Definition 1.2.1** A metric space is a set, $X$ and a function $d : X \times X \to [0, \infty)$ which satisfies the following properties.

\[
\begin{align*}
d(x, y) &= d(y, x) \\
d(x, y) &\geq 0 \text{ and } d(x, y) = 0 \text{ if and only if } x = y \\
d(x, y) &\leq d(x, z) + d(z, y).
\end{align*}
\]

You can check that $\mathbb{R}^n$ and $\mathbb{C}^n$ are metric spaces with $d(x, y) = |x - y|$. However, there are many others. The definitions of open and closed sets are the same for a metric space as they are for $\mathbb{R}^n$.

**Definition 1.2.2** A set, $U$ in a metric space is open if whenever $x \in U$, there exists $r > 0$ such that $B(x, r) \subseteq U$. As before, $B(x, r) \equiv \{y : d(x, y) < r\}$. Closed sets are
8 CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

those whose complements are open. A point \( p \) is a limit point of a set, \( S \) if for every \( r > 0, B(p, r) \) contains infinitely many points of \( S \). A sequence, \( \{x_n\} \) converges to a point \( x \) if for every \( \varepsilon > 0 \) there exists \( N \) such that if \( n \geq N, \) then \( d(x, x_n) < \varepsilon. \) \( \{x_n\} \) is a Cauchy sequence if for every \( \varepsilon > 0 \) there exists \( N \) such that if \( m, n \geq N, \) then \( d(x_n, x_m) < \varepsilon. \)

**Lemma 1.2.3** In a metric space, \( X \) every ball, \( B(x, r) \) is open. A set is closed if and only if it contains all its limit points. If \( p \) is a limit point of \( S, \) then there exists a sequence of distinct points of \( S, \{x_n\} \) such that \( \lim_{n \to \infty} x_n = p. \)

**Proof:** Let \( z \in B(x, r). \) Let \( \delta = r - d(z, x). \) Then if \( w \in B(z, \delta), \)

\[
d(w, x) \leq d(z, x) + d(z, w) < d(z, x) + r - d(x, z) = r.
\]

Therefore, \( B(z, \delta) \subseteq B(x, r) \) and this shows \( B(x, r) \) is open.

The properties of balls are presented in the following theorem.

**Theorem 1.2.4** Suppose \((X, d)\) is a metric space. Then the sets \( \{B(x, r) : r > 0, \ x \in X\} \) satisfy

\[
\bigcup \{B(x, r) : r > 0, \ x \in X\} = X
\]

(1.2.1)

If \( p \in B(x_1, r_1) \cap B(z, r_2) \), there exists \( r > 0 \) such that

\[
B(p, r) \subseteq B(x_1, r_1) \cap B(z, r_2).
\]

(1.2.2)

**Proof:** Observe that the union of these balls includes the whole space, \( X \) so \( (X) \) is obvious. Consider \( B(p, r) \subseteq B(x_1, r_1) \cap B(z, r_2). \) Consider

\[
r \equiv \min(r_1 - d(x, p), r_2 - d(z, p))
\]

and suppose \( y \in B(p, r). \) Then

\[
d(y, x) \leq d(y, p) + d(p, x) < r_1 - d(x, p) + d(x, p) = r_1
\]

and so \( B(p, r) \subseteq B(x_1, r_1). \) By similar reasoning, \( B(p, r) \subseteq B(z, r_2). \)

Let \( K \) be a closed set. This means \( K^C \equiv X \setminus K \) is an open set. Let \( p \) be a limit point of \( K. \) If \( p \in K^C, \) then since \( K^C \) is open, there exists \( B(p, r) \subseteq K^C. \) But this contradicts \( p \) being a limit point because there are no points of \( K \) in this ball. Hence all limit points of \( K \) must be in \( K. \)

Suppose next that \( K \) contains its limit points. Is \( K^C \) open? Let \( p \in K^C. \) Then \( p \) is not a limit point of \( K. \) Therefore, there exists \( B(p, r) \) which contains at most finitely many points of \( K. \) Since \( p \notin K, \) it follows that by making \( r \) smaller if necessary, \( B(p, r) \) contains no points of \( K. \) That is \( B(p, r) \subseteq K^C \) showing \( K^C \) is open. Therefore, \( K \) is closed.

Suppose now that \( p \) is a limit point of \( S. \) Let \( x_1 \in (S \setminus \{p\}) \cap B(p, 1). \) If \( x_1, \ldots, x_k \) have been chosen, let

\[
r_{k+1} \equiv \min \left\{ d(p, x_i), i = 1, \ldots, k, \frac{1}{k + 1} \right\}.
\]

Let \( x_{k+1} \in (S \setminus \{p\}) \cap B(p, r_{k+1}). \)
Lemma 1.2.5 If \( \{x_n\} \) is a Cauchy sequence in a metric space, \( X \) and if some subsequence, \( \{x_{n_k}\} \) converges to \( x \), then \( \{x_n\} \) converges to \( x \). Also if a sequence converges, then it is a Cauchy sequence.

Proof: Note first that \( n_k \geq k \) because in a subsequence, the indices, \( n_1, n_2, \ldots \) are strictly increasing. Let \( \varepsilon > 0 \) be given and let \( N \) be such that for \( k > N, d(x, x_{n_k}) < \varepsilon/2 \) and for \( m, n \geq N, d(x_m, x_n) < \varepsilon/2 \). Pick \( k > n \). Then if \( n > N \),

\[
d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\]

Finally, suppose \( \lim_{n \to \infty} x_n = x \). Then there exists \( N \) such that if \( n > N \), then \( d(x_n, x) < \varepsilon/2 \). It follows that for \( m, n > N \),

\[
d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\]

A useful idea is the idea of distance from a point to a set.

Definition 1.2.6 Let \( (X, d) \) be a metric space and let \( S \) be a nonempty set in \( X \). Then

\[
\text{dist} (x, S) \equiv \inf \{ d(x, y) : y \in S \}.
\]

The following lemma is the fundamental result.

Lemma 1.2.7 The function, \( x \to \text{dist} (x, S) \) is continuous and in fact satisfies

\[
|\text{dist} (x, S) - \text{dist} (y, S)| \leq d(x, y).
\]

Proof: Suppose \( \text{dist} (x, y) \) is as least as large as \( \text{dist} (y, S) \). Then pick \( z \in S \) such that \( d(y, z) \leq \text{dist} (y, S) + \varepsilon \). Then

\[
|\text{dist} (x, S) - \text{dist} (y, S)| = \text{dist} (x, S) - \text{dist} (y, S) \\
\leq d(x, z) - (d(y, z) - \varepsilon) \\
= d(x, z) - d(y, z) + \varepsilon \\
\leq d(x, y) + d(y, z) - d(y, z) + \varepsilon \\
= d(x, y) + \varepsilon.
\]

Since \( \varepsilon > 0 \) is arbitrary, this proves the lemma.

1.3 Compactness In Metric Space

Many existence theorems in analysis depend on some set being compact. Therefore, it is important to be able to identify compact sets. The purpose of this section is to describe compact sets in a metric space.
Definition 1.3.1 Let \( A \) be a subset of \( X \). \( A \) is compact if whenever \( A \) is contained in the union of a set of open sets, there exists finitely many of these open sets whose union contains \( A \). (Every open cover admits a finite subcover.) \( A \) is “sequentially compact” means every sequence has a convergent subsequence converging to an element of \( A \).

In a metric space compact is not the same as closed and bounded!

Example 1.3.2 Let \( X \) be any infinite set and define \( d(x, y) = 1 \) if \( x \neq y \) while \( d(x, y) = 0 \) if \( x = y \).

You should verify the details that this is a metric space because it satisfies the axioms of a metric. The set \( X \) is closed and bounded because its complement is \( \emptyset \) which is clearly open because every point of \( \emptyset \) is an interior point. (There are none.) Also \( X \) is bounded because \( X = B(x, 2) \). However, \( X \) is clearly not compact because \( \{ B(x, 1/2) : x \in X \} \) is a collection of open sets whose union contains \( X \) but since they are all disjoint and nonempty, there is no finite subset of these whose union contains \( X \). In fact \( B(x, 1/2) = \{ x \} \).

From this example it is clear something more than closed and bounded is needed.

If you are not familiar with the issues just discussed, ignore them and continue.

Definition 1.3.3 In any metric space, a set \( E \) is totally bounded if for every \( \varepsilon > 0 \) there exists a finite set of points \( \{ x_1, \cdots, x_n \} \) such that \( E \subseteq \bigcup_{i=1}^{n} B(x_i, \varepsilon) \).

This finite set of points is called an \( \varepsilon \) net.

The following proposition tells which sets in a metric space are compact. First here is an interesting lemma.

Lemma 1.3.4 Let \( X \) be a metric space and suppose \( D \) is a countable dense subset of \( X \). In other words, it is being assumed \( X \) is a separable metric space. Consider the open sets of the form \( B(d, r) \) where \( r \) is a positive rational number and \( d \in D \). Denote this countable collection of open sets by \( \mathcal{B} \). Then every open set is the union of sets of \( \mathcal{B} \). Furthermore, if \( \mathcal{C} \) is any collection of open sets, there exists a countable subset, \( \{ U_n \} \subseteq \mathcal{C} \) such that \( \bigcup_{n=1}^{\infty} U_n = \bigcup \mathcal{C} \).

Proof: Let \( U \) be an open set and let \( x \in U \). Let \( B(x, \delta) \subseteq U \). Then by density of \( D \), there exists \( d \in D \cap B(x, \delta/4) \). Now pick \( r \in \mathbb{Q} \cap (\delta/4, 3\delta/4) \) and consider \( B(d, r) \). Clearly, \( B(d, r) \) contains the point \( x \) because \( r > \delta/4 \). Is \( B(d, r) \subseteq B(x, \delta) \)? if so, this proves the lemma because \( x \) was an arbitrary point of \( U \). Suppose \( z \in B(d, r) \).

Then

\[
d(z, x) \leq d(z, d) + d(d, x) < r + \frac{\delta}{4} + \frac{\delta}{4} = \frac{3\delta}{4} + \frac{\delta}{4} = \delta
\]

Now let \( \mathcal{C} \) be any collection of open sets. Each set in this collection is the union of countably many sets of \( \mathcal{B} \). Let \( \mathcal{B}' \) denote the sets of \( \mathcal{B} \) which are contained in some set of \( \mathcal{C} \). Thus \( \bigcup \mathcal{B}' = \bigcup \mathcal{C} \). Then for each \( B \in \mathcal{B}' \), pick \( U_B \in \mathcal{C} \) such that \( B \subseteq U_B \). Then \( \{ U_B : B \in \mathcal{B}' \} \) is a countable collection of sets of \( \mathcal{C} \) whose union equals \( \bigcup \mathcal{C} \). Therefore, this proves the lemma.
1.3. COMPACTNESS IN METRIC SPACE

**Proposition 1.3.5** Let \((X,d)\) be a metric space. Then the following are equivalent.

\[(X,d)\text{ is compact}, \quad (1.3.3)\]
\[(X,d)\text{ is sequentially compact}, \quad (1.3.4)\]
\[(X,d)\text{ is complete and totally bounded}. \quad (1.3.5)\]

**Proof:** Suppose \([\text{2.3.2}]) and let \(\{x_k\}\) be a sequence. Suppose \(\{x_k\}\) has no convergent subsequence. If this is so, then no value of the sequence is repeated more than finitely many times. Also \(\{x_k\}\) has no limit point because if it did, there would exist a subsequence which converges. To see this, suppose \(p\) is a limit point of \(\{x_k\}\). Then in \(B(p,1)\) there are infinitely many points of \(\{x_k\}\). Pick one called \(x_{k_1}\). Now if \(x_{k_1}, x_{k_2}, \ldots, x_{k_n}\) have been picked with \(x_{k_i} \in B(p,1/i)\), consider \(B(p,1/(n+1))\). There are infinitely many points of \(\{x_k\}\) in this ball also. Pick \(x_{k_{n+1}}\) such that \(k_{n+1} > k_n\). Then \(\{x_{k_n}\}_{n=1}^{\infty}\) is a subsequence which converges to \(p\) and it is assumed this does not happen. Thus \(\{x_k\}\) has no limit points. It follows the set

\[C_n = \cup \{x_k : k \geq n\}\]

is a closed set because it has no limit points and if

\[U_n = C_n^C,\]

then

\[X = \cup_{n=1}^{\infty} U_n,\]

but there is no finite subcovering, because no value of the sequence is repeated more than finitely many times. This contradicts compactness of \((X,d)\). Note \(x_k\) is not in \(U_n\) whenever \(k > n\). Thus \([\text{2.3.2}]) implies \([\text{2.3.4}])

Now suppose \([\text{2.3.4}]) and let \(\{x_n\}\) be a Cauchy sequence. Is \(\{x_n\}\) convergent? By sequential compactness \(x_{n_k} \to x\) for some subsequence. By Lemma \([\text{2.2.2}]\) it follows that \(\{x_n\}\) also converges to \(x\) showing that \((X,d)\) is complete. If \((X,d)\) is not totally bounded, then there exists \(\varepsilon > 0\) for which there is no \(\varepsilon\) net. Hence there exists a sequence \(\{x_k\}\) with \(d(x_k, x_l) \geq \varepsilon\) for all \(l \neq k\). By Lemma \([\text{2.2.2}]\) again, this contradicts \([\text{2.3.4}]) because no subsequence can be a Cauchy sequence and so no subsequence can converge. This shows \([\text{2.3.4}]) implies \([\text{2.3.5}])

Now suppose \([\text{2.3.5}]). What about \([\text{2.3.5}])? Let \(\{p_n\}\) be a sequence and let \(\{x_i^n\}_{i=1}^{m_n}\) be a \(2^{-n}\) net for \(n = 1, 2, \ldots\). Let

\[B_n \equiv B \left(x_i^n, 2^{-n}\right)\]

be such that \(B_n\) contains \(p_k\) for infinitely many values of \(k\) and \(B_n \cap B_{n+1} \neq \emptyset\). To do this, suppose \(B_n\) contains \(p_k\) for infinitely many values of \(k\). Then one of the sets which intersect \(B_n\), \(B(x_{i+1}, 2^{-(n+1)})\) must contain \(p_k\) for infinitely many values of \(k\) because all these indices of points from \(\{p_n\}\) contained in \(B_n\) must be accounted for in one of finitely many sets, \(B(x_{i+1}, 2^{-(n+1)})\). Thus there exists a strictly increasing sequence of integers, \(n_k\) such that

\[p_{n_k} \in B_k.\]
CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

Then if \( k \geq l \),

\[
d(p_n^k, p_n^l) \leq \sum_{i=l}^{k-1} d(p_{n+1}^i, p_n^i) < \sum_{i=l}^{k-1} 2^{-(i-1)} < 2^{-(l-2)},
\]

Consequently \( \{p_n^k\} \) is a Cauchy sequence. Hence it converges because the metric space is complete. This proves 1.3.4.

Now suppose 1.3.4 and 1.3.5 which have now been shown to be equivalent. Let \( D_n \) be a \( n^{-1} \) net for \( n = 1, 2, \cdots \) and let

\[ D = \bigcup_{n=1}^{\infty} D_n. \]

Thus \( D \) is a countable dense subset of \((X, d)\).

Now let \( C \) be any set of open sets such that \( \bigcup C \supseteq X \). By Lemma 1.3.4, there exists a countable subset of \( C \),

\[ \bar{C} = \{U_n\}_{n=1}^{\infty} \]

such that \( \bigcup \bar{C} = \bigcup C \). If \( C \) admits no finite subcover, then neither does \( \bar{C} \) and there exists \( p_n \in X \setminus \bigcup_{k=1}^{n} U_k \). Then since \( X \) is sequentially compact, there is a subsequence \( \{p_n^k\} \) such that \( \{p_n^k\} \) converges. Say

\[ p = \lim_{k \to \infty} p_n^k. \]

All but finitely many points of \( \{p_n^k\} \) are in \( X \setminus \bigcup_{k=1}^{n} U_k \). Therefore \( p \in X \setminus \bigcup_{k=1}^{n} U_k \) for each \( n \). Hence

\[ p \notin \bigcup_{k=1}^{\infty} U_k \]

contradicting the construction of \( \{U_n\}_{n=1}^{\infty} \) which required that \( \bigcup_{n=1}^{\infty} U_n \supseteq X \). Hence \( X \) is compact. \( \blacksquare \)

Consider \( \mathbb{R}^n \). In this setting totally bounded and bounded are the same. This will yield a proof of the Heine Borel theorem from advanced calculus.

**Lemma 1.3.6** A subset of \( \mathbb{R}^n \) is totally bounded if and only if it is bounded.

**Proof:** Let \( A \) be totally bounded. Is it bounded? Let \( x_1, \cdots, x_p \) be a 1 net for \( A \). Now consider the ball \( B(0, r+1) \) where \( r \geq \max \{|x_i| : i = 1, \cdots, p\} \). If \( z \in A \), then \( z \in B(x_j, 1) \) for some \( j \) and so by the triangle inequality,

\[ |z - 0| \leq |z - x_j| + |x_j| < 1 + r. \]

Thus \( A \subseteq B(0, r+1) \) and so \( A \) is bounded.

Now suppose \( A \) is bounded and suppose \( A \) is not totally bounded. Then there exists \( \varepsilon > 0 \) such that there is no \( \varepsilon \) net for \( A \). Therefore, there exists a sequence of
points \( \{a_i\} \) with \( |a_i - a_j| \geq \varepsilon \) if \( i \neq j \). Since \( A \) is bounded, there exists \( r > 0 \) such that
\[
A \subseteq [\neg r, r]^n.
\]
\((x \in [\neg r, r]^n)\) means \( x_i \in [\neg r, r) \) for each \( i \). Now define \( S \) to be all cubes of the form
\[
\prod_{k=1}^{n}[a_k, b_k]
\]
where
\[
a_k = \neg r + i2^{-p}r, \quad b_k = \neg r + (i + 1)2^{-p}r,
\]
for \( i \in \{0, 1, \cdots, 2^{p+1} - 1\} \). Thus \( S \) is a collection of \( (2^{p+1})^n \) non overlapping cubes whose union equals \( [\neg r, r]^n \) and whose diameters are all equal to \( 2^{-p}r \sqrt{n} \). Now choose \( p \) large enough that the diameter of these cubes is less than \( \varepsilon \). This yields a contradiction because one of the cubes must contain infinitely many points of \( \{a_i\} \).

The next theorem is called the Heine Borel theorem and it characterizes the compact sets in \( \mathbb{R}^n \).

**Theorem 1.3.7** A subset of \( \mathbb{R}^n \) is compact if and only if it is closed and bounded.

**Proof:** Since a set in \( \mathbb{R}^n \) is totally bounded if and only if it is bounded, this theorem follows from Proposition 1.3.5 and the observation that a subset of \( \mathbb{R}^n \) is closed if and only if it is complete.

### 1.4 Some Applications Of Compactness

The following corollary is an important existence theorem which depends on compactness.

**Corollary 1.4.1** Let \( X \) be a compact metric space and let \( f : X \rightarrow \mathbb{R} \) be continuous. Then \( \max \{ f(x) : x \in X \} \) and \( \min \{ f(x) : x \in X \} \) both exist.

**Proof:** First it is shown \( f(X) \) is compact. Suppose \( C \) is a set of open sets whose union contains \( f(X) \). Then since \( f \) is continuous \( f^{-1}(U) \) is open for all \( U \in C \). Therefore, \( \{ f^{-1}(U) : U \in C \} \) is a collection of open sets whose union contains \( X \). Since \( X \) is compact, it follows finitely many of these, \( \{ f^{-1}(U_1), \cdots, f^{-1}(U_p) \} \) contains \( X \) in their union. Therefore, \( f(X) \subseteq \bigcup_{k=1}^{p}U_k \) showing \( f(X) \) is compact as claimed.

Now since \( f(X) \) is compact, Theorem 1.3.7 implies \( f(X) \) is closed and bounded. Therefore, it contains its inf and its sup. Thus \( f \) achieves both a maximum and a minimum.

**Definition 1.4.2** Let \( X, Y \) be metric spaces and \( f : X \rightarrow Y \) a function. \( f \) is uniformly continuous if for all \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that whenever \( x_1 \) and \( x_2 \) are two points of \( X \) satisfying \( d(x_1, x_2) < \delta \), it follows that \( d(f(x_1), f(x_2)) < \varepsilon \).
CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

A very important theorem is the following.

**Theorem 1.4.3** Suppose \( f : X \to Y \) is continuous and \( X \) is compact. Then \( f \) is uniformly continuous.

**Proof:** Suppose this is not true and that \( f \) is continuous but not uniformly continuous. Then there exists \( \varepsilon > 0 \) such that for all \( \delta > 0 \) there exist points, \( p_\delta \) and \( q_\delta \) such that \( d(p_\delta, q_\delta) < \delta \) and yet \( d(f(p_\delta), f(q_\delta)) \geq \varepsilon \). Let \( p_n \) and \( q_n \) be the points which go with \( \delta = 1/n \). By Proposition 1.3.5, \( \{p_n\} \) has a convergent subsequence, \( \{p_{n_k}\} \) converging to a point, \( x \in X \). Since \( d(p_n, q_n) < 1/n \), it follows that \( q_{n_k} \to x \) also. Therefore,

\[
\varepsilon \leq d(f(p_{n_k}), f(q_{n_k})) \leq d(f(p_{n_k}), f(x)) + d(f(x), f(q_{n_k}))
\]

but by continuity of \( f \), both \( d(f(p_{n_k}), f(x)) \) and \( d(f(x), f(q_{n_k})) \) converge to 0 as \( k \to \infty \) contradicting the above inequality. \( \blacksquare \)

Another important property of compact sets in a metric space concerns the finite intersection property.

**Definition 1.4.4** If every finite subset of a collection of sets has nonempty intersection, the collection has the finite intersection property.

**Theorem 1.4.5** Suppose \( F \) is a collection of compact sets in a metric space, \( X \) which has the finite intersection property. Then there exists a point in their intersection. \((\bigcap F \neq \emptyset).\)

**Proof:** First I show each compact set is closed. Let \( K \) be a nonempty compact set and suppose \( p \notin K \). Then for each \( x \in K \), let \( V_x = B(x, d(p, x) / 3) \) and \( U_x = B(p, d(p, x) / 3) \) so that \( U_x \) and \( V_x \) have empty intersection. Then since \( V \) is compact, there are finitely many \( V_x \) which cover \( K \) say \( V_{x_1}, \ldots, V_{x_n} \). Then let \( U = \bigcap_{i=1}^n U_{x_i} \). It follows \( p \in U \) and \( U \) has empty intersection with \( K \). In fact \( U \) has empty intersection with \( \bigcup_{i=1}^n V_{x_i} \). Since \( U \) is an open set and \( p \in K \), it follows \( K^C \) is an open set.

Consider now the claim about the intersection. If this were not so,

\[
\bigcup \{F^C : F \in F\} = X
\]

and so, in particular, picking some \( F_0 \in F \),

\[
\{F^C : F \in F\}
\]

would be an open cover of \( F_0 \). Since \( F_0 \) is compact, some finite subcover, \( F_1^C, \ldots, F_m^C \) exists. But then

\[
F_0 \subseteq \bigcup_{k=1}^m F_k^C
\]

which means \( \bigcap_{k=0}^m F_k = \emptyset \), contrary to the finite intersection property. To see this, note that if \( x \in F_0 \), then it must fail to be in some \( F_k \) and so it is not in \( \bigcap_{k=0}^m F_k \). Since this is true for every \( x \) it follows \( \bigcap_{k=0}^m F_k = \emptyset \).
1.5. ASCOLI ARZELA THEOREM

Theorem 1.4.6 Let $X_i$ be a compact metric space with metric $d_i$. Then $\prod_{i=1}^{m} X_i$ is also a compact metric space with respect to the metric, $d(x,y) \equiv \max_i (d_i(x_i,y_i))$.

Proof: This is most easily seen from sequential compactness. Let $\{x^k\}_{k=1}^{\infty}$ be a sequence of points in $\prod_{i=1}^{m} X_i$. Consider the $i^{th}$ component of $x^k$, $x^k_i$. It follows $\{x^k_i\}$ is a sequence of points in $X_i$ and so it has a convergent subsequence. Compactness of $X_1$ implies there exists a subsequence of $x^k$, denoted by $\{x^{k_1}\}$ such that

$$\lim_{k_1 \to \infty} x^{k_1}_1 \to x_1 \in X_1.$$ 

Now there exists a further subsequence, denoted by $\{x^{k_2}\}$ such that in addition to this, $x^{k_2}_2 \to x_2 \in X_2$. After taking $m$ such subsequences, there exists a subsequence, $\{x^l\}$ such that $\lim_{l \to \infty} x^l_i = x_i \in X_i$ for each $i$. Therefore, letting $x \equiv (x_1, \ldots, x_m)$, $x^l \to x$ in $\prod_{i=1}^{m} X_i$.

1.5 Ascoli Arzela Theorem

Definition 1.5.1 Let $(X,d)$ be a complete metric space. Then it is said to be locally compact if $B(x,r)$ is compact for each $r > 0$.

Thus if you have a locally compact metric space, then if $\{a_n\}$ is a bounded sequence, it must have a convergent subsequence.

Let $K$ be a compact subset of $\mathbb{R}^n$ and consider the continuous functions which have values in a locally compact metric space, $(X,d)$ where $d$ denotes the metric on $X$. Denote this space as $C(K,X)$.

Definition 1.5.2 For $f,g \in C(K,X)$, where $K$ is a compact subset of $\mathbb{R}^n$ and $X$ is a locally compact complete metric space define

$$\rho_K (f,g) \equiv \sup \{ d(f(x),g(x)) : x \in K \}.$$ 

Then $\rho_K$ provides a distance which makes $C(K,X)$ into a metric space.

The Ascoli Arzela theorem is a major result which tells which subsets of $C(K,X)$ are sequentially compact.

Definition 1.5.3 Let $A \subseteq C(K,X)$ for $K$ a compact subset of $\mathbb{R}^n$. Then $A$ is said to be uniformly equicontinuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $x,y \in K$ with $|x - y| < \delta$ and $f \in A$,

$$d(f(x), f(y)) < \varepsilon.$$ 

The set, $A$ is said to be uniformly bounded if for some $M < \infty$, and $a \in X$,

$$f(x) \in B(a,M)$$

for all $f \in A$ and $x \in K$. 

Uniform equicontinuity is like saying that the whole set of functions, \( A \), is uniformly continuous on \( K \) uniformly for \( f \in A \). The version of the Ascoli Arzela theorem I will present here is the following.

**Theorem 1.5.4** Suppose \( K \) is a nonempty compact subset of \( \mathbb{R}^n \) and \( A \subseteq C(K, X) \) is uniformly bounded and uniformly equicontinuous. Then if \( \{ f_k \} \subseteq A \), there exists a function, \( f \in C(K, X) \) and a subsequence, \( f_{k_l} \) such that

\[
\lim_{l \to \infty} \rho_K (f_{k_l}, f) = 0.
\]

To give a proof of this theorem, I will first prove some lemmas.

**Lemma 1.5.5** If \( K \) is a compact subset of \( \mathbb{R}^n \), then there exists \( D = \{ x_k \}_{k=1}^\infty \subseteq K \) such that \( D \) is dense in \( K \). Also, for every \( \varepsilon > 0 \) there exists a finite set of points, \( \{ x_1, \ldots, x_m \} \subseteq K \), called an \( \varepsilon \) net such that

\[
\bigcup_{i=1}^m B(x_i, \varepsilon) \supseteq K.
\]

**Proof:** For \( m \in \mathbb{N} \), pick \( x_1^m \in K \). If every point of \( K \) is within \( 1/m \) of \( x_1^m \), stop. Otherwise, pick \( x_2^m \in K \setminus B(x_1^m, 1/m) \).

If every point of \( K \) contained in \( B(x_1^m, 1/m) \cup B(x_2^m, 1/m) \), stop. Otherwise, pick \( x_3^m \in K \setminus (B(x_1^m, 1/m) \cup B(x_2^m, 1/m)) \).

If every point of \( K \) is contained in \( B(x_1^m, 1/m) \cup B(x_2^m, 1/m) \cup B(x_3^m, 1/m) \), stop. Otherwise, pick

\[
x_4^m \in K \setminus (B(x_1^m, 1/m) \cup B(x_2^m, 1/m) \cup B(x_3^m, 1/m))
\]

Continue this way until the process stops, say at \( N(m) \). It must stop because if it didn’t, there would be a convergent subsequence due to the compactness of \( K \). Ultimately all terms of this convergent subsequence would be closer than \( 1/m \), violating the manner in which they are chosen. Then \( D = \bigcup_{m=1}^\infty \bigcup_{k=1}^{N(m)} \{ x_k^m \} \). This is countable because it is a countable union of countable sets. If \( y \in K \) and \( \varepsilon > 0 \), then for some \( m \), \( 2/m < \varepsilon \) and so \( B(y, \varepsilon) \) must contain some point of \( \{ x_k^m \} \) since otherwise, the process stopped too soon. You could have picked \( y \).

**Lemma 1.5.6** Suppose \( D \) is defined above and \( \{ g_m \} \) is a sequence of functions of \( A \) having the property that for every \( x_k \in D \),

\[
\lim_{m \to \infty} g_m(x_k) \text{ exists.}
\]

Then there exists \( g \in C(K, X) \) such that

\[
\lim_{m \to \infty} \rho(g_m, g) = 0.
\]
1.5. ASCOLI ARZELA THEOREM

**Proof:** Define \( g \) first on \( D \).

\[
g(x_k) \equiv \lim_{m \to \infty} g_m(x_k).
\]

Next I show that \( \{ g_m \} \) converges at every point of \( K \). Let \( x \in K \) and let \( \varepsilon > 0 \) be given. Choose \( x_k \) such that for all \( f \in A \),

\[
d(f(x_k), f(x)) < \frac{\varepsilon}{3}.
\]

I can do this by the equicontinuity. Now if \( p, q \) are large enough, say \( p, q \geq M \),

\[
d(g_p(x_k), g_q(x_k)) < \frac{\varepsilon}{3}.
\]

Therefore, for \( p, q \geq M \),

\[
d(g_p(x), g_q(x)) \leq d(g_p(x), g_p(x_k)) + d(g_p(x_k), g_q(x_k)) + d(g_q(x_k), g_q(x))
\]

\[
< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\]

It follows that \( \{ g_m(x) \} \) is a Cauchy sequence having values \( X \). Therefore, it converges. Let \( g(x) \) be the name of the thing it converges to.

Let \( \varepsilon > 0 \) be given and pick \( \delta > 0 \) such that whenever \( x, y \in K \) and \( |x - y| < \delta \), it follows \( d(f(x), f(y)) < \frac{\varepsilon}{3} \) for all \( f \in A \). Now let \( \{ x_1, \ldots, x_m \} \) be a \( \delta \) net for \( K \) as in Lemma 1.5.5. Since there are only finitely many points in this \( \delta \) net, it follows that there exists \( N \) such that for all \( p, q \geq N \),

\[
d(g_q(x_i), g_p(x_i)) < \frac{\varepsilon}{3}
\]

for all \( \{ x_1, \ldots, x_m \} \). Therefore, for arbitrary \( x \in K \), pick \( x_i \in \{ x_1, \ldots, x_m \} \) such that \( |x_i - x| < \delta \). Then

\[
d(g_q(x), g_p(x)) \leq d(g_q(x), g_q(x_i)) + d(g_q(x_i), g_p(x_i)) + d(g_p(x_i), g_p(x))
\]

\[
< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\]

Since \( N \) does not depend on the choice of \( x \), it follows this sequence \( \{ g_m \} \) is uniformly Cauchy. That is, for every \( \varepsilon > 0 \), there exists \( N \) such that if \( p, q \geq N \), then

\[
\rho(g_p, g_q) < \varepsilon.
\]

Next, I need to verify that the function, \( g \) is a continuous function. Let \( N \) be large enough that whenever \( p, q \geq N \), the above holds. Then for all \( x \in K \),

\[
d(g(x), g_p(x)) \leq \frac{\varepsilon}{3}
\]

whenever \( p \geq N \). This follows from observing that for \( p, q \geq N \),

\[
d(g_q(x), g_p(x)) < \frac{\varepsilon}{3}
\]
and then taking the limit as \( q \to \infty \) to obtain 1.5.6. In passing to the limit, you can use the following simple claim.

**Claim:** In a metric space, if \( a_n \to a \), then \( d(a_n, b) \to d(a, b) \).

**Proof of the claim:** You note that by the triangle inequality, 
\[
d(a_n, b) - d(a, b) \leq d(a_n, a) \text{ and } d(a, b) - d(a_n, b) \leq d(a_n, a)
\]
and so
\[
|d(a_n, b) - d(a, b)| \leq d(a_n, a).
\]

Now let \( p \) satisfy 1.5.6 for all \( x \) whenever \( p > N \). Also pick \( \delta > 0 \) such that if \( |x - y| < \delta \), then
\[
d(g_p(x), g_p(y)) < \varepsilon.
\]
Then if \( |x - y| < \delta \),
\[
d(g(x), g(y)) \leq d(g(x), g_p(x)) + d(g_p(x), g_p(y)) + d(g_p(y), g(y)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\]
Since \( \varepsilon \) was arbitrary, this shows that \( g \) is continuous.

It only remains to verify that \( \rho(g, g_k) \to 0 \). But this follows from 1.5.6. \( \blacksquare \)

**Proof of Theorem 1.5.4:** Let \( D = \{x_k\} \) be the countable dense set of \( K \) guaranteed by Lemma 1.5.5 and let \( \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), \ldots\} \) be a subsequence of \( \mathbb{N} \) such that
\[
\lim_{k \to \infty} f_{(1,k)}(x_1) \text{ exists.}
\]
This is where the local compactness of \( X \) is being used. Now let \( \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), \ldots\} \) be a subsequence of \( \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), \ldots\} \) which has the property that
\[
\lim_{k \to \infty} f_{(2,k)}(x_2) \text{ exists.}
\]
Thus it is also the case that
\[
f_{(2,k)}(x_1) \text{ converges to } \lim_{k \to \infty} f_{(1,k)}(x_1).
\]
because every subsequence of a convergent sequence converges to the same thing as the convergent sequence. Continue this way and consider the array
\[
f_{(1,1)}, f_{(1,2)}, f_{(1,3)}, f_{(1,4)}, \ldots \text{ converges at } x_1
\]
\[
f_{(2,1)}, f_{(2,2)}, f_{(2,3)}, f_{(2,4)}, \ldots \text{ converges at } x_1 \text{ and } x_2
\]
\[
f_{(3,1)}, f_{(3,2)}, f_{(3,3)}, f_{(3,4)}, \ldots \text{ converges at } x_1, x_2, \text{ and } x_3
\]
\[
\vdots
\]
1.6. ANOTHER GENERAL VERSION

Now let \( g_k \equiv f_{(k,k)} \). Thus \( g_k \) is ultimately a subsequence of \( \{ f_{(m,k)} \} \) whenever \( k > m \) and therefore, \( \{ g_k \} \) converges at each point of \( D \). By Lemma 1.5.6 it follows there exists \( g \in C(K;X) \) such that

\[
\lim_{k \to \infty} \rho(g, g_k) = 0. \quad \blacksquare
\]

Actually there is an if and only if version of it but the most useful case is what is presented here. The process used to get the subsequence in the proof is called the Cantor diagonalization procedure.

1.6 Another General Version

This will use the characterization of compact metric spaces to give a proof of a general version of the Arzella Ascoli theorem. See Naylor and Sell \[17\] which is where I saw this general formulation.

Definition 1.6.1 Let \((X,d_X)\) be a compact metric space. Let \((Y,d_Y)\) be another complete metric space. Then \(C(X,Y)\) will denote the continuous functions which map \(X\) to \(Y\). Then \(\rho\) is a metric on \(C(X,Y)\) defined by

\[
\rho(f,g) = \sup_{x \in X} d_Y(f(x),g(x)).
\]

Theorem 1.6.2 \((C(X,Y),\rho)\) is a complete metric space.

Proof: It is first necessary to show that \(\rho\) is well defined. In this argument, I will just write \(d\) rather than \(d_X\) or \(d_Y\). To show this, note that

\[
x \to d(f(x),g(x))
\]

is a continuous function because \(f, g\) are continuous and

\[
|d(f(x),g(x)) - d(f(y),g(y))| \leq d(f(x),f(y)) + d(g(x),g(y))
\]

This follows from the triangle inequality. Say \(d(f(x),g(x)) \geq d(f(y),g(y))\). Otherwise just replace \(x\) with \(y\) and repeat the argument. Then in this case, it reduces to the claim that

\[
d(f(x),g(x)) \leq d(f(x),f(y)) + d(g(x),g(y)) + d(f(y),g(y))
\]

However, by the triangle inequality, the right side of the above is at least as large as

\[
d(f(x),f(y)) + d(g(x),f(y)) \geq d(f(x),g(x)).
\]

It follows that \(\rho(f,g)\) is just the maximum of a continuous function defined on a compact set.
Clearly $\rho(f,g) = \rho(g,f)$ and
\[ \rho(f,g) + \rho(g,h) = \sup_{x \in X} d(f(x),g(x)) + \sup_{x \in X} d(g(x),h(x)) \geq \sup_{x \in X} (d(f(x),g(x)) + d(g(x),h(x))) \geq \sup_{x \in X} d(f(x),h(x)) = \rho(f,h) \]
so the triangle inequality holds.

It remains to check completeness. Let $\{f_n\}$ be a Cauchy sequence. Then from the definition, $\{f_n(x)\}$ is a Cauchy sequence in $Y$ and so it converges to something called $f(x)$. I have to verify that $x \to f(x)$ is continuous.
\[ d(f(x), f(\hat{x})) \leq d(f(x), f_n(x)) + d(f_n(x), f(\hat{x})) = \lim_{m \to \infty} d(f_m(x), f_n(x)) + d(f_n(x), f(\hat{x})) \leq 2\lim_{m \to \infty} \sup_{x \in X} \rho(f_m, f_n) + d(f(x), f(\hat{x})) \leq \frac{2\varepsilon}{3} + d(f(x), f(\hat{x})) \]
whenever $n$ is large enough thanks to $\{f_n\}$ being a Cauchy sequence. Fix such an $n$. Now by continuity of $f_n$, there exists a $\delta > 0$ such that if $d(x, \hat{x}) < \delta$, then $d(f_n(x), f(\hat{x})) < \varepsilon/3$ and so for this choice of $\delta$, $d(f(x), f(\hat{x})) < \varepsilon$ which shows that $f$ is indeed continuous. Then
\[ d(f(x), f_n(x)) = \lim_{m \to \infty} d(f_m(x), f_n(x)) \leq \lim_{m \to \infty} \sup_{x \in X} \rho(f_m, f_n) < \varepsilon \]
provided $n$ is large enough. Then for such $n$
\[ \rho(f, f_n) = \sup_{x} d(f(x), f_n(x)) \leq \lim_{m \to \infty} \sup_{x} \rho(f_m, f_n) < \varepsilon \]
and so the Cauchy sequence converges to $f$. ■

Here is a useful lemma.

**Lemma 1.6.3** Let $S$ be a totally bounded subset of $(X, d)$ a metric space. Then $\overline{S}$ is also totally bounded.

**Proof:** Suppose not. Then there exists a sequence $\{p_n\} \subseteq \overline{S}$ such that $d(p_m, p_n) \geq \varepsilon$ for all $m \neq n$. Now let $q_n \in B(p_n, \frac{\varepsilon}{8}) \cap S$. Then it follows that
\[ \frac{\varepsilon}{8} + d(q_n, q_m) + \frac{\varepsilon}{8} \geq d(p_n, q_n) + d(q_n, q_m) + d(q_m, p_m) \geq d(p_n, q_m) \geq \varepsilon \]
and so $d(q_n, q_m) > \frac{\varepsilon}{2}$. This contradicts total boundedness of $S$. ■

Next, here is an important definition.
1.6. ANOTHER GENERAL VERSION

Definition 1.6.4 Let $\mathcal{A} \subseteq C(X,Y)$ where $(X,d_X)$ and $(Y,d_Y)$ are metric spaces. Thus $\mathcal{A}$ is a set of continuous functions mapping $X$ to $Y$. Then $\mathcal{A}$ is said to be equicontinuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $d_X(x_1,x_2) < \delta$ then for all $f \in \mathcal{A}$, $d_Y(f(x_1),f(x_2)) < \varepsilon$. (This is uniform continuity which is uniform in $\mathcal{A}$.) $\mathcal{A}$ is said to be pointwise compact if $\{f(x) : f \in \mathcal{A}\}$ has compact closure in $Y$.

Here is the Ascoli–Arzela theorem.

Theorem 1.6.5 Let $(X,d_X)$ be a compact metric space and let $(Y,d_Y)$ be a complete metric space. Thus $(C(X,Y),\rho)$ is a complete metric space. Let $\mathcal{A} \subseteq C(X,Y)$ be pointwise compact and equicontinuous. Then $\mathcal{A}$ is compact. Here the closure is taken in $(C(X,Y),\rho)$. The converse also holds.

Proof: The more useful direction is that the two conditions imply compactness of $\mathcal{A}$. I prove this first. Since $\mathcal{A}$ is a closed subset of a complete space, it follows that $\mathcal{A}$ will be compact if it is totally bounded. In showing this, it follows from Lemma 1.6.2 that it suffices to verify that $\mathcal{A}$ is totally bounded. Suppose this is not so. Then there exists $\varepsilon > 0$ and a sequence of points of $\mathcal{A}$, $\{f_n\}$ such that $\rho(f_n,f_m) \geq \varepsilon$ whenever $n \neq m$.

By equicontinuity, there exists $\delta > 0$ such that if $d(x,y) < \delta$, then $d(f(x),f(y)) < \varepsilon/8$ for all $f \in \mathcal{A}$. Let $\{x_i\}_{i=1}^{\infty}$ be a $\delta$ net for $X$. Since there are only finitely many $x_i$, it follows from pointwise compactness that there exists a subsequence, still denoted by $\{f_n\}$ which converges at each $x_i$. Then for $x \in B(x_i,\delta)$,

\[
d(f_n(x),f_m(x)) \leq d(f_n(x),f_n(x_i)) + d(f_n(x_i),f_m(x_i)) + d(f_m(x_i),f_m(x)) < \frac{\varepsilon}{8} + d(f_n(x_i),f_m(x_i)) + \frac{\varepsilon}{8} < \frac{\varepsilon}{2}
\]

provided $m,n \geq N_i$. Then let $N > \max\{N_i,i \leq p\}$. If $m,n > N$, then for arbitrary $x \in X$,

\[
d(f_n(x),f_m(x)) \leq \frac{\varepsilon}{2} \text{ so } \rho(f_n,f_m) \leq \frac{\varepsilon}{2}
\]

contrary to the condition above which had $\rho(f_n,f_m) \geq \varepsilon$ for all $n \neq m$. It follows that $\mathcal{A}$ and hence $\overline{\mathcal{A}}$ is totally bounded. This proves the more important direction.

Next suppose $\mathcal{A}$ is compact. Why must $\mathcal{A}$ be pointwise compact and equicontinuous? If it fails to be pointwise compact, then there exists $x \in X$ such that $\{f(x) : f \in \mathcal{A}\}$ is not contained in a compact set of $Y$. Thus there exists $\varepsilon > 0$ and a sequence of functions in $\mathcal{A}\{f_n\}$ such that $d(f_n(x),f_m(x)) \geq \varepsilon$. But this implies $\rho(f_m,f_n) \geq \varepsilon$ and so $\overline{\mathcal{A}}$ fails to be totally bounded, a contradiction. Thus $\mathcal{A}$ must be pointwise compact. Now why must it be equicontinuous? If it is not, then for each $n \in \mathbb{N}$ there exists $\varepsilon > 0$ and $x_n,y_n \in X$ such that $d(x_n,y_n) < 1/n$ but for some $f_n \in \mathcal{A}$, $d(f_n(x_n),f_n(y_n)) \geq \varepsilon$. However, by compactness, there exists a subsequence $\{f_{n_k}\}$ such that $\lim_{k \to \infty} \rho(f_{n_k},f) = 0$ and also that $x_{n_k},y_{n_k} \to x \in X$. Hence

\[
\varepsilon \leq d(f_{n_k}(x_{n_k}),f_{n_k}(y_{n_k})) \leq d(f_{n_k}(x_{n_k}),f(x_{n_k})) + d(f(x_{n_k}),f(y_{n_k})) + d(f(y_{n_k}),f_{n_k}(y_{n_k}))
\]
$$\leq \rho(f_{n_k}, f) + d(f(x_{n_k}), f(y_{n_k})) + \rho(f, f_{n_k})$$

and now this is a contradiction because each term on the right converges to 0. The middle term converges to 0 because $f(x_{n_k}), f(y_{n_k}) \to f(x)$. $\blacksquare$

### 1.7 The Tietze Extension Theorem

It turns out that if $H$ is a closed subset of a metric space, $(X, d)$ and if $f : H \to [a, b]$ is continuous, then there exists $g$ defined on all of $X$ such that $g = f$ on $H$ and $g$ is continuous. This is called the Tietze extension theorem. First it is well to recall continuity in the context of metric space.

**Definition 1.7.1** Let $(X, d)$ be a metric space and suppose $f : X \to Y$ is a function where $(Y, \rho)$ is also a metric space. For example, $Y = \mathbb{R}$. Then $f$ is continuous at $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(z)) < \varepsilon$ whenever $d(x, z) < \delta$. As is usual in such definitions, $f$ is said to be continuous if it is continuous at every point of $X$.

The following lemma gives an important example of a continuous real valued function defined on a metric space, $(X, d)$.

**Lemma 1.7.2** Let $(X, d)$ be a metric space and let $S \subseteq X$ be a nonempty subset. Define

$$\operatorname{dist}(x, S) \equiv \inf \{d(x, y) : y \in S\}.$$  \hspace{1cm} (1.7.7)

Then $x \to \operatorname{dist}(x, S)$ is a continuous function satisfying the inequality,

$$|\operatorname{dist}(x, S) - \operatorname{dist}(y, S)| \leq d(x, y).$$

**Proof:** The continuity of $x \to \operatorname{dist}(x, S)$ is obvious if the inequality is established. So let $x, y \in X$. Without loss of generality, assume $\operatorname{dist}(x, S) \geq \operatorname{dist}(y, S)$ and pick $z \in S$ such that $d(y, z) - \varepsilon < \operatorname{dist}(y, S)$. Then

$$|\operatorname{dist}(x, S) - \operatorname{dist}(y, S)| = \operatorname{dist}(x, S) - \operatorname{dist}(y, S) \leq d(x, z) - (d(y, z) - \varepsilon) \leq d(z, y) + d(x, y) - d(y, z) + \varepsilon = d(x, y) + \varepsilon.$$

Since $\varepsilon$ is arbitrary, this proves (1.7.7).

**Lemma 1.7.3** Let $H, K$ be two nonempty disjoint closed subsets of a metric space, $(X, d)$. Then there exists a continuous function, $g : X \to [-1, 1]$ such that $g(H) = -1/3$, $g(K) = 1/3$, $g(X) \subseteq [-1/3, 1/3]$.

**Proof:** Let

$$f(x) = \frac{\operatorname{dist}(x, H)}{\operatorname{dist}(x, H) + \operatorname{dist}(x, K)}.$$  

The denominator is never equal to zero because if $\operatorname{dist}(x, H) = 0$, then $x \in H$ because $H$ is closed. (To see this, pick $h_k \in B(x, 1/k) \cap H$. Then $h_k \to x$ and
1.7. THE TIETZE EXTENSION THEOREM

since $H$ is closed, $x \in H.$) Similarly, if dist$(x, K) = 0$, then $x \in K$ and so the denominator is never zero as claimed. Hence, by Lemma 1.7.2, $f$ is continuous and from its definition, $f = 0$ on $H$ and $f = 1$ on $K$. Now let $g(x) \equiv \frac{3}{4} \left(f(x) - \frac{1}{2}\right)$. Then $g$ has the desired properties.

**Definition 1.7.4** For $f$ a real or complex valued bounded continuous function defined on a metric space, $M$

$$||f||_M \equiv \sup \{|f(x)| : x \in M\}.$$ 

**Lemma 1.7.5** Suppose $M$ is a closed set in $X$ where $(X, d)$ is a metric space and suppose $f : M \to [-1, 1]$ is continuous at every point of $M$. Then there exists a function, $g$ which is defined and continuous on all of $X$ such that $||f - g||_M < \frac{2}{3}.$

**Proof:** Let $H = f^{-1}([-1, -1/3]), K = f^{-1}([1/3, 1])$. Thus $H$ and $K$ are disjoint closed subsets of $M$. Suppose first $H, K$ are both nonempty. Then by Lemma 1.7.2 there exists $g$ such that $g$ is a continuous function defined on all of $X$ and $g(H) = -1/3, g(K) = 1/3$, and $g(X) \subseteq [-1/3, 1/3]$. It follows $||f - g||_M < 2/3$. If $H = \emptyset$, then $f$ has all its values in $[-1/3, 1]$ and so letting $g \equiv 1/3$, the desired condition is obtained. If $K = \emptyset$, let $g \equiv -1/3.$

**Lemma 1.7.6** Suppose $M$ is a closed set in $X$ where $(X, d)$ is a metric space and suppose $f : M \to [-1, 1]$ is continuous at every point of $M$. Then there exists a function, $g$ which is defined and continuous on all of $X$ such that $g = f$ on $M$ and $g$ has its values in $[-1, 1]$.

**Proof:** Let $g_1$ be such that $g_1(X) \subseteq [-1/3, 1/3]$ and $||f - g_1||_M \leq \frac{2}{3}.$ Suppose $g_1, \ldots, g_m$ have been chosen such that $g_j(X) \subseteq [-1/3, 1/3]$ and

$$\left\| f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right\|_M < \left(\frac{2}{3}\right)^m.$$ (1.7.8)

Then

$$\left\| \left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right) \right\|_M \leq 1$$

and so $\left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right)$ can play the role of $f$ in the first step of the proof. Therefore, there exists $g_{m+1}$ defined and continuous on all of $X$ such that its values are in $[-1/3, 1/3]$ and

$$\left\| \left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right) - g_{m+1} \right\|_M \leq \frac{2}{3}.$$ 

Hence

$$\left\| f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right\|_M - \left(\frac{2}{3}\right)^m g_{m+1} \right\|_M \leq \left(\frac{2}{3}\right)^{m+1}.$$
It follows there exists a sequence, \( \{ g_i \} \) such that each has its values in \([-1/3, 1/3]\) and for every \( m \) holds. Then let 

\[
g(x) = \sum_{i=1}^{\infty} \left( \frac{2}{3} \right)^{i-1} g_i(x).
\]

It follows 

\[
|g(x)| \leq \sum_{i=1}^{\infty} \left( \frac{2}{3} \right)^{i-1} g_i(x) \leq \sum_{i=1}^{m} \left( \frac{2}{3} \right)^{i-1} \frac{1}{3} \leq 1
\]

and since convergence is uniform, \( g \) must be continuous. The estimate \( \text{Lemma } 1.7.8 \) implies \( f = g \) on \( M \).

The following is the Tietze extension theorem.

**Theorem 1.7.7** Let \( M \) be a closed nonempty subset of a metric space \((X, d)\) and let \( f : M \to [a, b] \) is continuous at every point of \( M \). Then there exists a function, \( g \) continuous on all of \( X \) which coincides with \( f \) on \( M \) such that \( g(X) \subseteq [a, b] \).

**Proof:** Let \( f_1(x) = 1 + \frac{2}{b-a} (f(x) - b) \). Then \( f_1 \) satisfies the conditions of Lemma \( 1.7.6 \) and so there exists \( g_1 : X \to [-1,1] \) such that \( g \) is continuous on \( X \) and equals \( f_1 \) on \( M \). Let \( g(x) = (g_1(x) - 1) \left( \frac{b-a}{2} \right) + b \). This works.

### 1.8 Some Simple Fixed Point Theorems

The following is of more interest in the case of normed vector spaces, but there is no harm in stating it in this more general setting. You should verify that the functions described in the following definition are all continuous.

**Definition 1.8.1** Let \( f : X \to Y \) where \( (X, d) \) and \( (Y, \rho) \) are metric spaces. Then \( f \) is said to be Lipschitz continuous if for every \( x, \hat{x} \in X \), \( \rho(f(x), f(\hat{x})) \leq rd(x, \hat{x}) \). The function is called a contraction map if \( r < 1 \).

The big theorem about contraction maps is the following.

**Theorem 1.8.2** Let \( f : (X, d) \to (X, d) \) be a contraction map and let \( (X, d) \) be a complete metric space. Thus Cauchy sequences converge and also \( d(f(x), f(\hat{x})) \leq rd(x, \hat{x}) \) where \( r < 1 \). Then \( f \) has a unique fixed point. This is a point \( x \in X \) such that \( f(x) = x \). Also, if \( x_0 \) is any point of \( X \), then 

\[
d(x, x_0) \leq \frac{d(x_0, f(x_0))}{1 - r}
\]

Also, for each \( n \), 

\[
d(f^n(x_0), x_0) \leq \frac{d(x_0, f(x_0))}{1 - r},
\]

and \( x = \lim_{n \to \infty} f^n(x_0) \).
Proof: Pick $x_0 \in X$ and consider the sequence of iterates of the map,

$$x_0, f(x_0), f^2(x_0), \ldots.$$  

We argue that this is a Cauchy sequence. For $m < n$, it follows from the triangle inequality,

$$d(f^m(x_0), f^n(x_0)) \leq \sum_{k=m}^{n-1} d(f^{k+1}(x_0), f^{k}(x_0)) \leq \sum_{k=m}^{\infty} r^k d(f(x_0), x_0)$$

The reason for this last is as follows.

$$d(f^2(x_0), f(x_0)) \leq r d(f(x_0), x_0)$$
$$d(f^3(x_0), f^2(x_0)) \leq r d(f^2(x_0), f(x_0)) \leq r^2 d(f(x_0), x_0)$$

and so forth. Therefore,

$$d(f^n(x_0), f^m(x_0)) \leq d(f(x_0), x_0) \frac{r^m}{1 - r}$$

which shows that this is indeed a Cauchy sequence. Therefore, there exists $x$ such that

$$\lim_{n \to \infty} f^n(x_0) = x$$

By continuity,

$$f(x) = f\left(\lim_{n \to \infty} f^n(x_0)\right) = \lim_{n \to \infty} f^{n+1}(x_0) = x.$$  

Also note that this estimate yields

$$d(x_0, f^n(x_0)) \leq \frac{d(x_0, f(x_0))}{1 - r}$$

Now $d(x_0, x) \leq d(x_0, f^n(x_0)) + d(f^n(x_0), x)$ and so

$$d(x_0, x) - d(f^n(x_0), x) \leq \frac{d(x_0, f(x_0))}{1 - r}$$

Letting $n \to \infty$, it follows that

$$d(x_0, x) \leq \frac{d(x_0, f(x_0))}{1 - r}$$

It only remains to verify that there is only one fixed point. Suppose then that $x, x'$ are two. Then

$$d(x, x') = d(f(x), f(x')) \leq r d(x', x)$$

and so $d(x, x') = 0$ because $r < 1$. 

The above is the usual formulation of this important theorem, but we actually proved a better result.
Corollary 1.8.3 Let $B$ be a closed subset of the complete metric space $(X, d)$ and let $f : B \to X$ be a contraction map

$$d(f(x), f(\hat{x})) \leq rd(x, \hat{x}), \quad r < 1.$$ 

Also suppose there exists $x_0 \in B$ such that the sequence of iterates $\{f^n(x_0)\}_{n=1}^{\infty}$ remains in $B$. Then $f$ has a unique fixed point in $B$ which is the limit of the sequence of iterates. This is a point $x \in B$ such that $f(x) = x$. In the case that $B = B(x_0, \delta)$, the sequence of iterates satisfies the inequality

$$d(f^n(x_0), x_0) \leq \frac{d(x_0, f(x_0))}{1 - r}$$ 

and so it will remain in $B$ if

$$\frac{d(x_0, f(x_0))}{1 - r} < \delta.$$ 

Proof: By assumption, the sequence of iterates stays in $B$. Then, as in the proof of the preceding theorem, for $m < n$, it follows from the triangle inequality,

$$d(f^m(x_0) , f^n(x_0)) \leq \sum_{k=m}^{n-1} d(f^{k+1}(x_0), f^k(x_0)) \leq \sum_{k=m}^{n-1} r^k d(x_0, f(x_0)) = \frac{r^m}{1 - r} d(f(x_0), x_0).$$

Hence the sequence of iterates is Cauchy and must converge to a point $x$ in $X$. However, $B$ is closed and so it must be the case that $x \in B$. Then as before,

$$x = \lim_{n \to \infty} f^n(x_0) = \lim_{n \to \infty} f^{n+1}(x_0) = f \left( \lim_{n \to \infty} f^n(x_0) \right) = f(x).$$

As to the sequence of iterates remaining in $B$ where $B$ is a ball as described, the inequality above in the case where $m = 0$ yields

$$d(x_0, f^m(x_0)) \leq \frac{1}{1 - r} d(f(x_0), x_0)$$

and so, if the right side is less than $\delta$, then the iterates remain in $B$. As to the fixed point being unique, it is as before. If $x, x'$ are both fixed points in $B$, then $d(x, x') = d(f(x), f(x')) \leq rd(x, x')$ and so $x = x'$. 

Sometimes you have the contraction depending on a parameter $\lambda$. Then there is a principle of uniform contractions.

Corollary 1.8.4 Suppose $f : X \times \Lambda \to X$ where $\Lambda$ is a metric space and $X$ is a complete metric space. Suppose $f$ satisfies

1. $d(f(x, \lambda), f(y, \lambda)) \leq rd(x, y)$ for each $\lambda \in \Lambda$. 

2. \( \lambda \rightarrow f(x, \lambda) \) is continuous as a map from \( \Lambda \) to \( X \).

Then if \( x(\lambda) \) is the fixed point, it follows that \( \lambda \rightarrow x(\lambda) \) is continuous.

**Proof:** Pick \( x_0 \in X \) and consider the above sequence of iterates, \( \{f^n(x, \lambda)\} \). Let \( \rho \) be the metric on \( \Lambda \). Then there is a fixed point and if \( x(\lambda) \) is this unique fixed point,

\[
d(x(\lambda), x_0) \leq \frac{d(f(x_0, \lambda), x_0)}{1 - r}
\]

In particular, you could start with \( x_0 = x(\mu) \) and conclude that

\[
d(x(\lambda), x(\mu)) \leq \frac{d(f(x(\mu), \lambda), x(\mu))}{1 - r}
\]

\[
\leq \frac{d(f(x(\mu), \lambda), f(x(\mu), \mu))}{1 - r} + \frac{d(f(x(\mu), \mu), x(\mu))}{1 - r}
\]

\[
= \frac{d(f(x(\mu), \lambda), f(x(\mu), \mu))}{1 - r}
\]

Now by continuity of \( \lambda \rightarrow f(x, \lambda) \), it follows that if \( \rho(\lambda, \mu) \) is small enough, the above is no larger than

\[
\varepsilon \frac{(1 - r)}{1 - r} = \varepsilon
\]

Hence, if \( \rho(\lambda, \mu) \) is small enough, we have

\[
d(x(\lambda), x(\mu)) < \varepsilon. \quad \blacksquare
\]

This is called the uniform contraction principle.

The contraction mapping theorem has an extremely useful generalization. In order to get a unique fixed point, it suffices to have some power of \( f \) a contraction map.

**Theorem 1.8.5** Let \( f : (X, d) \rightarrow (X, d) \) have the property that for some \( n \in \mathbb{N} \), \( f^n \) is a contraction map and let \( (X, d) \) be a complete metric space. Then there is a unique fixed point for \( f \). As in the earlier theorem the sequence of iterates \( \{f^n(x_0)\}_{n=1}^\infty \) also converges to the fixed point.

**Proof:** From Theorem 1.8.2 there is a unique fixed point for \( f^n \). Thus

\[
f^n(x) = x
\]

Then

\[
f^n(f(x)) = f^{n+1}(x) = f(x)
\]

By uniqueness, \( f(x) = x \).
Now consider the sequence of iterates. Suppose it fails to converge to $x$. Then there is $\varepsilon > 0$ and a subsequence $n_k$ such that
\[
    d(f^{n_k}(x_0), x) \geq \varepsilon
\]
Now $n_k = p_k n + r_k$ where $r_k$ is one of the numbers $\{0, 1, 2, \ldots, n - 1\}$. It follows that there exists one of these numbers which is repeated infinitely often. Call it $r$ and let the further subsequence continue to be denoted as $n_k$. Thus
\[
    d\left(f^{p_k n + r}(x_0), x\right) \geq \varepsilon
\]
In other words,
\[
    d\left(f^{p_k n + r}(x_0), x\right) \geq \varepsilon
\]
However, from Theorem 1.8.2, as $k \to \infty$, $f^{p_n n}(f^r(x_0)) \to x$ which contradicts the above inequality. Hence the sequence of iterates converges to $x$, as it did for $f$ a contraction map. ■

**Definition 1.8.6** Let $f : (X, d) \to (Y, \rho)$ be a function. Then it is said to be uniformly continuous on $X$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $x, \hat{x}$ are two points of $X$ with $d(x, \hat{x}) < \delta$, it follows that $\rho(f(x), f(\hat{x})) < \varepsilon$.

Note the difference between this and continuity. With continuity, the $\delta$ could depend on $x$ but here it works for any pair of points in $X$.

**Lemma 1.8.7** Suppose $x_n \to x$ and $y_n \to y$. Then $d(x_n, y_n) \to d(x, y)$.

**Proof:** Consider the following.
\[
    d(x, y) \leq d(x, x_n) + d(x, y) \leq d(x, x_n) + d(x, y_n) + d(y_n, y)
\]
so
\[
    d(x, y) - d(x, y_n) \leq d(x, x_n) + d(y_n, y)
\]
Similarly
\[
    d(x_n, y_n) - d(x, y) \leq d(x, x_n) + d(y_n, y)
\]
and so
\[
    |d(x_n, y_n) - d(x, y)| \leq d(x, x_n) + d(y_n, y)
\]
and the right side converges to 0 as $n \to \infty$. ■

There is a remarkable result concerning compactness and uniform continuity.

**Theorem 1.8.8** Let $f : (X, d) \to (Y, \rho)$ be a continuous function and let $K$ be a compact subset of $X$. Then the restriction of $f$ to $K$ is uniformly continuous.

**Proof:** First of all, $K$ is a metric space and $f$ restricted to $K$ is continuous. Now suppose it fails to be uniformly continuous. Then there exists $\varepsilon > 0$ and pairs of points $x_n, \hat{x}_n$ such that $d(f(x_n), \hat{x}_n) < 1/n$ but $\rho(f(x_n), f(\hat{x}_n)) \geq \varepsilon$. Since $K$ is
compact, it is sequentially compact and so there exists a subsequence, still denoted as \(\{x_n\}\) such that \(x_n \to x \in K\). Then also \(\hat{x}_n \to x\) also and so
\[
\rho(f(x), f(x)) = \lim_{n \to \infty} \rho(f(x_n), f(\hat{x}_n)) \geq \varepsilon
\]
which is a contradiction. Note the use of Lemma 1.8.7 in the equal sign.

Next is to consider the meaning of convergence of sequences of functions. There are two main ways of convergence of interest here, pointwise and uniform convergence.

**Definition 1.8.9** Let \(f_n: X \to Y\) where \((X, d), (Y, \rho)\) are two metric spaces. Then \(\{f_n\}\) is said to converge pointwise to a function \(f: X \to Y\) if for every \(x \in X\),
\[
\lim_{n \to \infty} f_n(x) = f(x)
\]
\(\{f_n\}\) is said to converge uniformly if for all \(\varepsilon > 0\), there exists \(N\) such that if \(n \geq N\), then
\[
\sup_{x \in X} \rho(f_n(x), f(x)) < \varepsilon
\]

Here is a well known example illustrating the difference between pointwise and uniform convergence.

**Example 1.8.10** Let \(f_n(x) = x^n\) on the metric space \([0, 1]\). Then this function converges pointwise to
\[
f(x) = \begin{cases} 
0 & \text{on } [0, 1) \\
1 & \text{at } 1 
\end{cases}
\]
but it does not converge uniformly on this interval to \(f\).

Note how the target function \(f\) in the above example is not continuous even though each function in the sequence is. The nice thing about uniform convergence is that it takes continuity of the functions in the sequence and imparts it to the target function. It does this for both continuity at a single point and uniform continuity. Thus uniform convergence is a very superior thing.

**Theorem 1.8.11** Let \(f_n : X \to Y\) where \((X, d), (Y, \rho)\) are two metric spaces and suppose each \(f_n\) is continuous at \(x \in X\) and also that \(f_n\) converges uniformly to \(f\) on \(X\). Then \(f\) is also continuous at \(x\). In addition to this, if each \(f_n\) is uniformly continuous on \(X\), then the same is true for \(f\).

**Proof:** Let \(\varepsilon > 0\) be given. Then
\[
\rho(f(x), f(\hat{x})) \leq \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(\hat{x})) + \rho(f_n(\hat{x}), f(\hat{x}))
\]
By uniform convergence, there exists \(N\) such that both \(\rho(f(x), f_n(x))\) and \(\rho(f_n(\hat{x}), f(\hat{x}))\) are less than \(\varepsilon/3\) provided \(n \geq N\). Thus picking such an \(n\),
\[
\rho(f(x), f(\hat{x})) \leq \frac{2\varepsilon}{3} + \rho(f_n(x), f_n(\hat{x}))
\]
Now from the continuity of $f_n$, there exists $\delta > 0$ such that if $d(x, \hat{x}) < \delta$, then $\rho(f_n(x), f_n(\hat{x})) < \varepsilon/3$. Hence, if $d(x, \hat{x}) < \delta$, then
\[
\rho(f(x), f(\hat{x})) \leq \frac{2\varepsilon}{3} + \rho(f_n(x), f_n(\hat{x})) < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\]
Hence, $f$ is continuous at $x$.

Next consider uniform continuity. It follows from the uniform convergence that if $x, \hat{x}$ are any two points of $X$, then if $n \geq N$, then, picking such an $n$,
\[
\rho(f(x), f(\hat{x})) \leq \frac{2\varepsilon}{3} + \rho(f_n(x), f_n(\hat{x}))
\]
By uniform continuity of $f_n$ there exists $\delta$ such that if $d(x, \hat{x}) < \delta$, then the term on the right in the above is less than $\varepsilon/3$. Hence if $d(x, \hat{x}) < \delta$, then $\rho(f(x), f(\hat{x})) < \varepsilon$ and so $f$ is uniformly continuous as claimed.

1.9 Banach Spaces

In metric space, you maybe can’t add the elements of the space. An important special case of a metric space is a Banach space, described in the following definition. Most of what is done later on will be in the context of Banach space.

**Definition 1.9.1** A vector space $V$ with field of $\{\text{Banach spaces}\}$ scalars either $\mathbb{R}$ or $\mathbb{C}$ is called a normed linear space or normed vector space if there is a norm defined on the vector space $\|\cdot\| : V \to [0, \infty)$ which is something which satisfies the following axioms.

- $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$
- $\|ax\| = |a| \|x\|$ for every scalar $a$
- $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in V$

Then one can define a distance on this as follows.
\[
d(x, y) \equiv \|x - y\|
\]
A normed linear space is called a Banach space if, with respect to this distance, it is a complete metric space.

**Observation 1.9.2** The distance is a good enough metric.

To see this, note that it is obvious that $d(x, y) = d(y, x)$ and $d(x, y) \geq 0$ and equals 0 if and only if $x = y$. The main thing to check is the triangle inequality.
\[
d(x, y) + d(y, z) = \|x - y\| + \|y - z\| \geq \|x - y + y - z\| = \|x - z\| = d(x, z).
\]
This will be discussed more in the exercises.
1.10  General Topological Spaces

It turns out that metric spaces are not sufficiently general for some applications. This section is a brief introduction to general topology. In making this generalization, the properties of balls which are the conclusion of Theorem 1.2.4 on Page 8 are stated as axioms for a subset of the power set of a given set which will be known as a basis for the topology. More can be found in [38] and the references listed there.

**Definition 1.10.1** Let $X$ be a nonempty set and suppose $\mathcal{B} \subseteq \mathcal{P}(X)$. Then $\mathcal{B}$ is a basis for a topology if it satisfies the following axioms.

1.) Whenever $p \in A \cap B$ for $A, B \in \mathcal{B}$, it follows there exists $C \in \mathcal{B}$ such that $p \in C \subseteq A \cap B$.
2.) $\bigcup \mathcal{B} = X$.

Then a subset, $U$, of $X$ is an open set if for every point, $x \in U$, there exists $B \in \mathcal{B}$ such that $x \in B \subseteq U$. Thus the open sets are exactly those which can be obtained as a union of sets of $\mathcal{B}$. Denote these subsets of $X$ by the symbol $\tau$ and refer to $\tau$ as the topology or the set of open sets.

Note that this is simply the analog of saying a set is open exactly when every point is an interior point.

**Proposition 1.10.2** Let $X$ be a set and let $\mathcal{B}$ be a basis for a topology as defined above and let $\tau$ be the set of open sets determined by $\mathcal{B}$. Then

\begin{align*}
\emptyset & \in \tau, \quad X \in \tau, \quad \text{(1.10.9)} \\
\text{If } C \subseteq \tau, \text{ then } \bigcup C & \in \tau \quad \text{(1.10.10)} \\
\text{If } A, B \in \tau, \text{ then } A \cap B & \in \tau. \quad \text{(1.10.11)}
\end{align*}

**Proof:** If $p \in \emptyset$ then there exists $B \in \mathcal{B}$ such that $p \in B \subseteq \emptyset$ because there are no points in $\emptyset$. Therefore, $\emptyset \in \tau$. Now if $p \in X$, then by part 2.) of Definition 1.10.1 $p \in B \subseteq X$ for some $B \in \mathcal{B}$ and so $X \in \tau$.

If $C \subseteq \tau$, and if $p \in \bigcup C$, then there exists a set, $B \in \mathcal{C}$ such that $p \in B$. However, $B$ is itself a union of sets from $\mathcal{B}$ and so there exists $C \in \mathcal{B}$ such that $p \in C \subseteq B \subseteq \bigcup C$. This verifies (1.10.10).

Finally, if $A, B \in \tau$ and $p \in A \cap B$, then since $A$ and $B$ are themselves unions of sets of $\mathcal{B}$, it follows there exists $A_1, B_1 \in \mathcal{B}$ such that $A_1 \subseteq A, B_1 \subseteq B$, and $p \in A_1 \cap B_1$. Therefore, by 1.) of Definition 1.10.1 there exists $C \in \mathcal{B}$ such that $p \in C \subseteq A_1 \cap B_1 \subseteq A \cap B$, showing that $A \cap B \in \tau$ as claimed. Of course if $A \cap B = \emptyset$, then $A \cap B \in \tau$. This proves the proposition.

**Definition 1.10.3** A set $X$ together with such a collection of its subsets satisfying (1.10.9)-(1.10.11) is called a topological space. $\tau$ is called the topology or set of open sets of $X$.

**Definition 1.10.4** A topological space is said to be Hausdorff if whenever $p$ and $q$ are distinct points of $X$, there exist disjoint open sets $U, V$ such that $p \in U, q \in V$. In other words points can be separated with open sets.
CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

32

Definition 1.10.5 A subset of a topological space is said to be closed if its complement is open. Let \( p \) be a point of \( X \) and let \( E \subseteq X \). Then \( p \) is said to be a limit point of \( E \) if every open set containing \( p \) contains a point of \( E \) distinct from \( p \).

Note that if the topological space is Hausdorff, then this definition is equivalent to requiring that every open set containing \( p \) contains infinitely many points from \( E \). Why?

Theorem 1.10.6 A subset, \( E \), of \( X \) is closed if and only if it contains all its limit points.

Proof: Suppose first that \( E \) is closed and let \( x \) be a limit point of \( E \). Is \( x \in E \)? If \( x \not\in E \), then \( E^C \) is an open set containing \( x \) which contains no points of \( E \), a contradiction. Thus \( x \in E \).

Now suppose \( E \) contains all its limit points. Is the complement of \( E \) open? If \( x \not\in E^C \), then \( x \) is not a limit point of \( E \) because \( E \) has all its limit points and so there exists an open set, \( U \) containing \( x \) such that \( U \) contains no point of \( E \) other than \( x \). Since \( x \not\in E \), it follows that \( x \in U \subseteq E^C \) which implies \( E^C \) is an open set because this shows \( E^C \) is the union of open sets.

Theorem 1.10.7 If \( (X, \tau) \) is a Hausdorff space and if \( p \in X \), then \( \{p\} \) is a closed set.

Proof: If \( x \neq p \), there exist open sets \( U \) and \( V \) such that \( x \in U, p \in V \) and \( U \cap V = \emptyset \). Therefore, \( \{p\}^C \) is an open set so \( \{p\} \) is closed.

Note that the Hausdorff axiom was stronger than needed in order to draw the conclusion of the last theorem. In fact it would have been enough to assume that if \( x \neq y \), then there exists an open set containing \( x \) which does not intersect \( y \).

Definition 1.10.8 A topological space \( (X, \tau) \) is said to be regular if whenever \( C \) is a closed set and \( p \) is a point not in \( C \), there exist disjoint open sets \( U \) and \( V \) such that \( p \in U, C \subseteq V \). Thus a closed set can be separated from a point not in the closed set by two disjoint open sets.

\[ U \quad p \quad \text{Regular} \quad V \]

\[ U \quad p \quad \text{Hausdorff} \quad V \]

\[ U \quad p \quad \text{Regular} \quad V \]
1.10. GENERAL TOPOLOGICAL SPACES

**Definition 1.10.9** The topological space, \((X, \tau)\) is said to be normal if whenever \(C\) and \(K\) are disjoint closed sets, there exist disjoint open sets \(U\) and \(V\) such that \(C \subseteq U, K \subseteq V\). Thus any two disjoint closed sets can be separated with open sets.

**Definition 1.10.10** Let \(E\) be a subset of \(X\). \(\overline{E}\) is defined to be the smallest closed set containing \(E\).

**Lemma 1.10.11** The above definition is well defined.

**Proof:** Let \(C\) denote all the closed sets which contain \(E\). Then \(C\) is nonempty because \(X \in C\).

\[
(\cap \{A : A \in C\})^C = \cup \{A^C : A \in C\},
\]

an open set which shows that \(\cap C\) is a closed set and is the smallest closed set which contains \(E\).

**Theorem 1.10.12** \(\overline{E} = E \cup \{\text{limit points of } E\}\).

**Proof:** Let \(x \in \overline{E}\) and suppose that \(x \notin E\). If \(x\) is not a limit point either, then there exists an open set, \(U\), containing \(x\) which does not intersect \(E\). But then \(U^C\) is a closed set which contains \(E\) which does not contain \(x\), contrary to the definition that \(\overline{E}\) is the intersection of all closed sets containing \(E\). Therefore, \(x\) must be a limit point of \(E\) after all.

Now \(E \subseteq \overline{E}\) so suppose \(x\) is a limit point of \(E\). Is \(x \in \overline{E}\)? If \(H\) is a closed set containing \(E\), which does not contain \(x\), then \(H^C\) is an open set containing \(x\) which contains no points of \(E\) other than \(x\) negating the assumption that \(x\) is a limit point of \(E\).

The following is the definition of continuity in terms of general topological spaces. It is really just a generalization of the \(\varepsilon - \delta\) definition of continuity given in calculus.

**Definition 1.10.13** Let \((X, \tau)\) and \((Y, \eta)\) be two topological spaces and let \(f : X \to Y\). \(f\) is continuous at \(x \in X\) if whenever \(V\) is an open set of \(Y\) containing \(f(x)\), there exists an open set \(U \in \tau\) such that \(x \in U\) and \(f(U) \subseteq V\). \(f\) is continuous if \(f^{-1}(V) \in \tau\) whenever \(V \in \eta\).

You should prove the following.

**Proposition 1.10.14** In the situation of Definition 1.10.13 \(f\) is continuous if and only if \(f\) is continuous at every point of \(X\).
Definition 1.10.15 Let \((X_i, \tau_i)\) be topological spaces. \(\prod_{i=1}^{n} X_i\) is the Cartesian product. Define a product topology as follows. Let \(B = \prod_{i=1}^{n} A_i\) where \(A_i \in \tau_i\). Then \(B\) is a basis for the product topology.

Theorem 1.10.16 The set \(B\) of Definition 1.10.15 is a basis for a topology.

Proof: Suppose \(x \in \prod_{i=1}^{n} A_i \cap \prod_{i=1}^{n} B_i\) where \(A_i\) and \(B_i\) are open sets. Say \(x = (x_1, \ldots, x_n)\).

Then \(x_i \in A_i \cap B_i\) for each \(i\). Therefore, \(x \in \prod_{i=1}^{n} A_i \cap B_i \in B\) and \(\prod_{i=1}^{n} A_i \cap B_i \subseteq \prod_{i=1}^{n} A_i\).

The definition of compactness is also considered for a general topological space. This is given next.

Definition 1.10.17 A subset, \(E\), of a topological space \((X, \tau)\) is said to be compact if whenever \(C \subseteq \tau\) and \(E \subseteq \bigcup C\), there exists a finite subset of \(C\), \(\{U_1 \cdot \cdot \cdot U_m\}\), such that \(E \subseteq \bigcup_{i=1}^{m} U_i\). (Every open covering admits a finite subcovering.) \(E\) is precompact if \(\overline{E}\) is compact. A topological space is called locally compact if it has a basis \(B\), with the property that \(\overline{B}\) is compact for each \(B \in B\).

In general topological spaces there may be no concept of “bounded”. Even if there is, closed and bounded is not necessarily the same as compactness. However, in any Hausdorff space every compact set must be a closed set.

Theorem 1.10.18 If \((X, \tau)\) is a Hausdorff space, then every compact subset must also be a closed set.

Proof: Suppose \(p \notin K\). For each \(x \in X\), there exist open sets, \(U_x\) and \(V_x\) such that \(x \in U_x\), \(p \in V_x\), and 
\[U_x \cap V_x = \emptyset.\]

If \(K\) is assumed to be compact, there are finitely many of these sets, \(U_{x_1}, \ldots, U_{x_m}\) which cover \(K\). Then let \(V \equiv \cap_{i=1}^{m} V_{x_i}\). It follows that \(V\) is an open set containing \(p\) which has empty intersection with each of the \(U_{x_i}\). Consequently, \(V\) contains no points of \(K\) and is therefore not a limit point of \(K\).

A useful construction when dealing with locally compact Hausdorff spaces is the notion of the one point compactification of the space.

Definition 1.10.19 Suppose \((X, \tau)\) is a locally compact Hausdorff space. Then let \(\overline{X} \equiv X \cup \{\infty\}\) where \(\infty\) is just the name of some point which is not in \(X\) which is called the point at infinity. A basis for the topology \(\overline{\tau}\) for \(\overline{X}\) is 
\[\tau \cup \{K^C\ \text{where} \ K \ \text{is a compact subset of} \ X\}.\]

The complement is taken with respect to \(\overline{X}\) and so the open sets, \(K^C\) are basic open sets which contain \(\infty\).
The reason this is called a compactification is contained in the next lemma.

**Lemma 1.10.20** If \((X,\tau)\) is a locally compact Hausdorff space, then \((\tilde{X},\tilde{\tau})\) is a compact Hausdorff space. Also if \(U\) is an open set of \(\tilde{\tau}\), then \(U \setminus \{\infty\}\) is an open set of \(\tau\).

**Proof:** Since \((X,\tau)\) is a locally compact Hausdorff space, it follows \((\tilde{X},\tilde{\tau})\) is a Hausdorff topological space. The only case which needs checking is the one of \(p \in X\) and \(\infty\). Since \((X,\tau)\) is locally compact, there exists an open set of \(\tau\), \(U\) having compact closure which contains \(p\). Then \(p \in U\) and \(\infty\in \overline{U}\) and these are disjoint open sets containing the points, \(p\) and \(\infty\) respectively. Now let \(C\) be an open cover of \(\tilde{X}\) with sets from \(\tilde{\tau}\). Then \(\infty\) must be in some set, \(U_{\infty}\) from \(C\), which must contain a set of the form \(K^C\) where \(K\) is a compact subset of \(X\). Then there exist sets from \(C, U_1, \cdots, U_r\), which cover \(K\). Therefore, a finite subcover of \(\tilde{X}\) is \(U_1, \cdots, U_r, U_{\infty}\).

To see the last claim, suppose \(U\) contains \(\infty\) since otherwise there is nothing to show. Notice that if \(C\) is a compact set, then \(X \setminus C\) is an open set. Therefore, if \(x \in U \setminus \{\infty\}\), and if \(\tilde{X} \setminus C\) is a basic open set contained in \(U\) containing \(\infty\), then if \(x\) is in this basic open set of \(\tilde{X}\), it is also in the open set \(X \setminus C\subseteq U \setminus \{\infty\}\). If \(x\) is not in any basic open set of the form \(\tilde{X} \setminus C\) then \(x\) is contained in an open set of \(\tau\) which is contained in \(U \setminus \{\infty\}\). Thus \(U \setminus \{\infty\}\) is indeed open in \(\tau\).

**Definition 1.10.21** If every finite subfamily of a collection of sets has nonempty intersection, the collection has the finite intersection property.

**Theorem 1.10.22** Let \(K\) be a set whose elements are compact subsets of a Hausdorff topological space, \((X,\tau)\). Suppose \(K\) has the finite intersection property. Then \(\emptyset \neq \cap K\).

**Proof:** Suppose to the contrary that \(\emptyset = \cap K\). Then consider

\[ C \equiv \{K^C : K \in K\}. \]

It follows \(C\) is an open cover of \(K_0\) where \(K_0\) is any particular element of \(K\). But then there are finitely many \(K \in K, K_1, \cdots, K_r\) such that \(K_0 \subseteq \cup_{i=1}^r K_i^C\) implying that \(\cap_{i=1}^r K_i = \emptyset\), contradicting the finite intersection property.

**Lemma 1.10.23** Let \((X,\tau)\) be a topological space and let \(B\) be a basis for \(\tau\). Then \(K\) is compact if and only if every open cover of basic open sets admits a finite subcover.

**Proof:** Suppose first that \(X\) is compact. Then if \(C\) is an open cover consisting of basic open sets, it follows it admits a finite subcover because these are open sets in \(C\).

Next suppose that every basic open cover admits a finite subcover and let \(C\) be an open cover of \(X\). Then define \(C\) to be the collection of basic open sets which are
CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

contained in some set of \( C \). It follows \( \mathcal{C} \) is a basic open cover of \( X \) and so it admits a finite subcover, \( \{U_1, \cdots, U_p\} \). Now each \( U_i \) is contained in an open set of \( \mathcal{C} \). Let \( O_i \) be a set of \( \mathcal{C} \) which contains \( U_i \). Then \( \{O_1, \cdots, O_p\} \) is an open cover of \( X \). \( \blacksquare \)

In fact, much more can be said than Lemma 1.10.23. However, this is all which I will present here. The case of a subbasis is presented a little later.

One other thing should be noted. If you have a topological space, \((X, \tau)\) and if \( Y \subseteq X \), then you can consider this as a topological space also by letting the topology equal sets of the form \( U \cap Y \) such that \( U \in \tau \). Then denoting this new topology as \( \sigma \), you have \((Y, \sigma)\) is a topological space. You can verify that \( \sigma \) satisfies all the necessary properties of a topology. It is called the relative topology.

1.11 Connected Sets

Stated informally, connected sets are those which are in one piece. More precisely,

**Definition 1.11.1** A set, \( S \) in a general topological space is separated if there exist sets, \( A, B \) such that

\[
S = A \cup B, \quad A, B \neq \emptyset, \quad \overline{A \cap B} = \overline{B \cap A} = \emptyset.
\]

In this case, the sets \( A \) and \( B \) are said to separate \( S \). A set is connected if it is not separated.

One of the most important theorems about connected sets is the following.

**Theorem 1.11.2** Suppose \( U \) and \( V \) are connected sets having nonempty intersection. Then \( U \cup V \) is also connected.

**Proof:** Suppose \( U \cup V = A \cup B \) where \( \overline{A \cap B} = \overline{B \cap A} = \emptyset \). Consider the sets, \( A \cap U \) and \( B \cap U \). Since

\[
(\overline{A \cap U}) \cap (B \cap U) = (A \cap U) \cap (\overline{B \cap U}) = \emptyset,
\]

It follows one of these sets must be empty since otherwise, \( U \) would be separated. It follows that \( U \) is contained in either \( A \) or \( B \). Similarly, \( V \) must be contained in either \( A \) or \( B \). Since \( U \) and \( V \) have nonempty intersection, it follows that both \( V \) and \( U \) are contained in one of the sets, \( A, B \). Therefore, the other must be empty and this shows \( U \cup V \) cannot be separated and is therefore, connected.
The intersection of connected sets is not necessarily connected as is shown by
the following picture.

\[
\begin{array}{c}
\text{U} \\
\text{V}
\end{array}
\]

**Theorem 1.11.3** Let \( f : X \to Y \) be continuous where \( X \) and \( Y \) are topological
spaces and \( X \) is connected. Then \( f(X) \) is also connected.

**Proof:** To do this you show \( f(X) \) is not separated. Suppose to the contrary
that \( f(X) = A \cup B \) where \( A \) and \( B \) separate \( f(X) \). Then consider the sets, \( f^{-1}(A) \)
and \( f^{-1}(B) \). If \( z \in f^{-1}(B) \), then \( f(z) \in B \) and so \( f(z) \) is not a limit point
of \( A \). Therefore, there exists an open set, \( U \) containing \( f(z) \) such that \( U \cap A = \emptyset \).
But then, the continuity of \( f \) implies that \( f^{-1}(U) \) is an open set containing \( z \) such
that \( f^{-1}(U) \cap f^{-1}(A) = \emptyset \). Therefore, \( f^{-1}(B) \) contains no limit points of \( f^{-1}(A) \).
Similar reasoning implies \( f^{-1}(A) \) contains no limit points of \( f^{-1}(B) \). It follows
that \( X \) is separated by \( f^{-1}(A) \) and \( f^{-1}(B) \), contradicting the assumption that \( X \)
was connected.

An arbitrary set can be written as a union of maximal connected sets called
connected components. This is the concept of the next definition.

**Definition 1.11.4** Let \( S \) be a set and let \( p \in S \). Denote by \( C_p \) the union of all
connected subsets of \( S \) which contain \( p \). This is called the connected component
determined by \( p \).

**Theorem 1.11.5** Let \( C_p \) be a connected component of a set \( S \) in a general topological
space. Then \( C_p \) is a connected set and if \( C_p \cap C_q \neq \emptyset \), then \( C_p = C_q \).

**Proof:** Let \( C \) denote the connected subsets of \( S \) which contain \( p \). If \( C_p = A \cup B \)
where

\[
\overline{A} \cap B = \overline{B} \cap A = \emptyset,
\]

then \( p \) is in one of \( A \) or \( B \). Suppose without loss of generality \( p \in A \). Then every
set of \( C \) must also be contained in \( A \) also since otherwise, as in Theorem 1.11.3, the
set would be separated. But this implies \( B \) is empty. Therefore, \( C_p \) is connected.
From this, and Theorem 1.11.4, the second assertion of the theorem is proved.

This shows the connected components of a set are equivalence classes and par-
tition the set.
A set, \( I \) is an interval in \( \mathbb{R} \) if and only if whenever \( x, y \in I \) then \((x, y) \subseteq I\). The following theorem is about the connected sets in \( \mathbb{R} \).

**Theorem 1.11.6** A set, \( C \) in \( \mathbb{R} \) is connected if and only if \( C \) is an interval.

**Proof:** Let \( C \) be connected. If \( C \) consists of a single point, \( p \), there is nothing to prove. The interval is just \([p, p]\). Suppose \( p < q \) and \( p, q \in C \). You need to show \((p, q) \subseteq C\).

Let \( x \in (p, q) \setminus C \) let \( C \cap (−∞, x) = A \), and \( C \cap (x, ∞) = B \). Then \( C = A \cup B \) and the sets, \( A \) and \( B \) separate \( C \) contrary to the assumption that \( C \) is connected.

Conversely, let \( I \) be an interval. Suppose \( I \) is separated by \( A \) and \( B \). Pick \( x \in A \) and \( y \in B \). Suppose without loss of generality that \( x < y \). Now define the set, \( S = \{ t \in [x, y] : [x, t] \subseteq A \} \) and let \( l \) be the least upper bound of \( S \). Then \( l \in A \) so \( l \notin B \) which implies \( l \in A \).

But if \( l \notin B \), then for some \( \delta > 0 \),

\[ (l, l + \delta) \cap B = \emptyset \]

contradicting the definition of \( l \) as an upper bound for \( S \). Therefore, \( l \in B \) which implies \( l \notin A \) after all, a contradiction. It follows \( I \) must be connected.

The following theorem is a very useful description of the open sets in \( \mathbb{R} \).

**Theorem 1.11.7** Let \( U \) be an open set in \( \mathbb{R} \). Then there exist countably many disjoint open sets, \( \{(a_i, b_i)\}_{i=1}^{∞} \) such that \( U = \bigcup_{i=1}^{∞} (a_i, b_i) \).

**Proof:** Let \( p \in U \) and let \( z \in C_p \), the connected component determined by \( p \). Since \( U \) is open, there exists, \( \delta > 0 \) such that \((z − \delta, z + \delta) \subseteq U \). It follows from Theorem 1.11.6 that

\[ (z − \delta, z + \delta) \subseteq C_p \]

This shows \( C_p \) is open. By Theorem 1.11.6, this shows \( C_p \) is an open interval, \((a, b)\) where \( a, b \in [−∞, ∞] \). There are therefore at most countably many of these connected components because each must contain a rational number and the rational numbers are countable. Denote by \( \{(a_i, b_i)\}_{i=1}^{∞} \) the set of these connected components.

**Definition 1.11.8** A topological space, \( E \) is arcwise connected if for any two points, \( p, q \in E \), there exists a closed interval, \([a, b]\) and a continuous function, \( γ : [a, b] → E \) such that \( γ(a) = p \) and \( γ(b) = q \). \( E \) is locally connected if it has a basis of connected open sets. \( E \) is locally arcwise connected if it has a basis of arcwise connected open sets.
An example of an arcwise connected topological space would be the any subset of \( \mathbb{R}^n \) which is the continuous image of an interval. Locally connected is not the same as connected. A well known example is the following.

\[
\left\{ \left( x, \sin \frac{1}{x} \right) : x \in (0, 1] \right\} \cup \{ (0, y) : y \in [-1, 1] \}
\]

(11.11.12)

You can verify that this set of points considered as a metric space with the metric from \( \mathbb{R}^2 \) is not locally connected or arcwise connected but is connected.

**Proposition 1.11.9** If a topological space is arcwise connected, then it is connected.

**Proof:** Let \( X \) be an arcwise connected space and suppose it is separated. Then \( X = A \cup B \) where \( A, B \) are two separated sets. Pick \( p \in A \) and \( q \in B \). Since \( X \) is given to be arcwise connected, there must exist a continuous function \( \gamma : [a, b] \to X \) such that \( \gamma(a) = p \) and \( \gamma(b) = q \). But then we would have \( \gamma([a, b]) = (\gamma([a, b]) \cap A) \cup (\gamma([a, b]) \cap B) \) and the two sets, \( \gamma([a, b]) \cap A \) and \( \gamma([a, b]) \cap B \) are separated thus showing that \( \gamma([a, b]) \) is separated and contradicting Theorem 1.11.6 and Theorem 1.11.3. It follows that \( X \) must be connected as claimed.

**Theorem 1.11.10** Let \( U \) be an open subset of a locally arcwise connected topological space, \( X \). Then \( U \) is arcwise connected if and only if \( U \) is connected. Also the connected components of an open set in such a space are open sets, hence arcwise connected.

**Proof:** By Proposition 1.11.9 it is only necessary to verify that if \( U \) is connected and open in the context of this theorem, then \( U \) is arcwise connected. Pick \( p \in U \). Say \( x \in U \) satisfies \( \mathcal{P} \) if there exists a continuous function, \( \gamma : [a, b] \to U \) such that \( \gamma(a) = p \) and \( \gamma(b) = x \).

\[
A \equiv \{ x \in U \text{ such that } x \text{ satisfies } \mathcal{P} \}
\]

If \( x \in A \), there exists, according to the assumption that \( X \) is locally arcwise connected, an open set, \( V \), containing \( x \) and contained in \( U \) which is arcwise connected. Thus letting \( y \in V \), there exist intervals, \( [a, b] \) and \( [c, d] \) and continuous functions having values in \( U \), \( \gamma, \eta \) such that \( \gamma(a) = p, \gamma(b) = x, \eta(c) = x \), and \( \eta(d) = y \). Then let \( \gamma_1 : [a, b + d - c] \to U \) be defined as

\[
\gamma_1(t) = \begin{cases} 
\gamma(t) & \text{if } t \in [a, b] \\
\eta(t + c - b) & \text{if } t \in [b, b + d - c]
\end{cases}
\]

Then it is clear that \( \gamma_1 \) is a continuous function mapping \( p \) to \( y \) and showing that \( V \subseteq A \). Therefore, \( A \) is open. \( A \neq \emptyset \) because there is an open set, \( V \) containing \( p \) which is contained in \( U \) and is arcwise connected.

Now consider \( B \equiv U \setminus A \). This is also open. If \( B \) is not open, there exists a point \( z \in B \) such that every open set containing \( z \) is not contained in \( B \). Therefore,
letting $V$ be one of the basic open sets chosen such that $z \in V \subseteq U$, there exist points of $A$ contained in $V$. But then, a repeat of the above argument shows $z \in A$ also. Hence $B$ is open and so if $B \neq \emptyset$, then $U = B \cup A$ and so $U$ is separated by the two sets, $B$ and $A$ contradicting the assumption that $U$ is connected.

It remains to verify the connected components are open. Let $z \in C_p$ where $C_p$ is the connected component determined by $p$. Then picking $V$ an arcwise connected open set which contains $z$ and is contained in $U$, $C_p \cup V$ is connected and contained in $U$ and so it must also be contained in $C_p$.

As an application, consider the following corollary.

**Corollary 1.11.11** Let $f : \Omega \to \mathbb{Z}$ be continuous where $\Omega$ is a connected open set. Then $f$ must be a constant.

**Proof:** Suppose not. Then it achieves two different values, $k$ and $l \neq k$. Then $\Omega = f^{-1}(l) \cup f^{-1}\left(\{m \in \mathbb{Z} : m \neq l\}\right)$ and there are disjoint nonempty open sets which separate $\Omega$. To see they are open, note

$$f^{-1}\left(\{m \in \mathbb{Z} : m \neq l\}\right) = f^{-1}\left(\bigcup_{m \neq l} \left(m - \frac{1}{6}, m + \frac{1}{6}\right)\right)$$

which is the inverse image of an open set.

### 1.12 The Tychonoff Theorem

Sometimes it is necessary to consider infinite Cartesian products of topological spaces. When you have finitely many topological spaces in the product and each is compact, it can be shown that the Cartesian product is compact with the product topology. It turns out that the same thing holds for infinite products but you have to be careful how you define the topology. The first thing likely to come to mind by analogy with finite products is not the right way to do it.

#### 1.12.1 Partially Ordered Sets

**Definition 1.12.1** Let $\mathcal{F}$ be a nonempty set. $\mathcal{F}$ is called a partially ordered set if there is a relation, denoted here by $\leq$, such that

$x \leq x$ for all $x \in \mathcal{F}$.

If $x \leq y$ and $y \leq z$ then $x \leq z$.

$\mathcal{C} \subseteq \mathcal{F}$ is said to be a chain if every two elements of $\mathcal{C}$ are related. This means that if $x, y \in \mathcal{C}$, then either $x \leq y$ or $y \leq x$. Sometimes a chain is called a totally ordered set. $\mathcal{C}$ is said to be a maximal chain if whenever $\mathcal{D}$ is a chain containing $\mathcal{C}$, $\mathcal{D} = \mathcal{C}$.

The most common example of a partially ordered set is the power set of a given set with $\subseteq$ being the relation. It is also helpful to visualize partially ordered sets as trees. Two points on the tree are related if they are on the same branch of
the tree and one is higher than the other. Thus two points on different branches
would not be related although they might both be larger than some point on the
trunk. You might think of many other things which are best considered as partially
ordered sets. Think of food for example. You might find it difficult to determine
which of two favorite pies you like better although you may be able to say very
easily that you would prefer either pie to a dish of lard topped with whipped cream
and mustard. The following theorem is equivalent to the axiom of choice. For a
discussion of this, see the appendix on the subject.

A major result is the following theorem.

**Theorem 1.12.2 (Hausdorff maximal principle)** Let $F$ be a nonempty partially
ordered set. Then there exists a maximal chain.

### 1.12.2 Alexander Sub-basis Theorem

The main tool in the study of products of compact topological spaces is the Alexan-
der sub-basis theorem which is presented next. Recall a set is compact if every basic
open cover admits a finite subcover. This was pretty easy to prove. However, there
is a much smaller set of open sets called a subbasis which has this property. The
proof of this result is much harder.

**Definition 1.12.3** $S \subseteq \tau$ is called a subbasis for the topology $\tau$ if the set $B$ of finite
intersections of sets of $S$ is a basis for the topology, $\tau$.

**Theorem 1.12.4** Let $(X, \tau)$ be a topological space and let $S \subseteq \tau$ be a subbasis for
$\tau$. Then if $H \subseteq X$, $H$ is compact if and only if every open cover of $H$ consisting
entirely of sets of $S$ admits a finite subcover.

**Proof:** The only if part is obvious because the subbasic sets are themselves open.

If every basic open cover admits a finite subcover then the set in question is
compact. Suppose then that $H$ is a subset of $X$ having the property that subbasic
open covers admit finite subcovers. Is $H$ compact? Assume this is not so. Then
what was just observed about basic covers implies there exists a basic open cover
of $H$, $O$, which admits no finite subcover. Let $F$ be defined as

$$\{O : O \text{ is a basic open cover of } H \text{ which admits no finite subcover}\}.$$

The assumption is that $F$ is nonempty. Partially order $F$ by set inclusion and use
the Hausdorff maximal principle to obtain a maximal chain, $C$, of such open covers
and let

$$D = \cup C.$$

If $D$ admits a finite subcover, then since $C$ is a chain and the finite subcover has only
finitely many sets, some element of $C$ would also admit a finite subcover, contrary
to the definition of $F$. Therefore, $D$ admits no finite subcover. If $D'$ properly
contains $D$ and $D'$ is a basic open cover of $H$, then $D'$ has a finite subcover of $H$
since otherwise, \( C \) would fail to be a maximal chain, being properly contained in \( C \cup \{D'\} \). Every set of \( D \) is of the form

\[
U = \bigcap_{i=1}^{m} B_i, \quad B_i \in S
\]

because they are all basic open sets. If it is the case that for all \( U \in D \) one of the \( B_i \) is found in \( D \), then replace each such \( U \) with the subbasic set from \( D \) containing it. But then this would be a subbasic open cover of \( H \) which by assumption would admit a finite subcover contrary to the properties of \( D \). Therefore, one of the sets of \( D \), denoted by \( U \), has the property that

\[
U = \bigcap_{i=1}^{m} B_i, \quad B_i \in S
\]

and no \( B_i \) is in \( D \). Thus \( D \cup \{B_i\} \) admits a finite subcover, for each of the above \( B_i \) because it is strictly larger than \( D \). Let this finite subcover corresponding to \( B_i \) be denoted by

\[
V_i^1, \ldots, V_i^{m_i}, B_i
\]

Consider

\[
\{U, V_j^1, j = 1, \ldots, m_i, i = 1, \ldots, m\}.
\]

If \( p \in H \setminus \bigcup \{V_j^1\} \), then \( p \in B_i \) for each \( i \) and so \( p \in U \). This is therefore a finite subcover of \( D \) contradicting the properties of \( D \). Therefore, \( F \) must be empty and this proves the theorem. \( \blacksquare \)

**Definition 1.12.5** Let \( I \) be a set and suppose for each \( i \in I \), \((X_i, \tau_i)\) is a nonempty topological space. The Cartesian product of the \( X_i \), denoted by \( \prod_{i \in I} X_i \), consists of the set of all choice functions defined on \( I \) which select a single element of each \( X_i \). Thus \( f \in \prod_{i \in I} X_i \) means for every \( i \in I \), \( f(i) \in X_i \). The axiom of choice says \( \prod_{i \in I} X_i \) is nonempty. Let

\[
P_j(A) \equiv \prod_{i \in I} B_i
\]

where \( B_i \equiv X_i \) if \( i \neq j \) and \( B_j \equiv A \). A subbasis for a topology on the product space consists of all sets \( P_j(A) \) where \( A \in \tau_j \). (These sets have an open set from the topology of \( X_j \) in the \( j \)th slot and the whole space in the other slots.) Thus a basis consists of finite intersections of these sets. Note that the intersection of two of these basic sets is another basic set and their union yields \( \prod_{i \in I} X_i \). Therefore, they satisfy the condition needed for a collection of sets to serve as a basis for a topology. This topology is called the product topology and is denoted by \( \prod \tau_i \).

**Proposition 1.12.6** The product topology is the smallest topology \( \tau \) for \( X \equiv \prod_{i \in I} X_i \) such that each \( \pi_i \) is continuous. Here \( \pi_i \) is defined in the following manner. For \( x \in X \), \( \pi_i(x) \equiv x_i \). Thus \( \pi_i \) delivers the \( i \)th entry of \( x \).

**Proof:** Let \( \tau_p \) denote the product topology. If each \( \pi_i \) is continuous, this means \( \pi_i^{-1}(A) = P_i(A) \in \tau \). By definition, \( \tau \) must consist of all unions of finite intersections of the sets \( P_i(A) \). It must include all such sets to be a topology and
it can’t contain any more sets if it is to be the smallest topology such that each $\pi_i$ is continuous. This is a description of $\tau_p$ and so $\tau = \tau_p$. ■

This gives an alternative way of defining the product topology. It is just the smallest topology for which the coordinate maps are all continuous. It is tempting to define a basis for a topology to be sets of the form $\prod_{i \in I} A_i$ where $A_i$ is open in $X_i$. This is not the same thing at all. Note that the basis in the product topology just described has at most finitely many slots filled with an open set which is not the whole space. The thing just mentioned in which every slot may be filled by a proper open set is called the box topology and there exist people who are interested in it. It is much too big to be interesting here.

The Alexander subbasis theorem is used to prove the Tychonoff theorem which says that if each $X_i$ is a compact topological space, then in the product topology, $\prod_{i \in I} X_i$ is also compact.

**Theorem 1.12.7** If $(X_i, \tau_i)$ is compact, then so is $(\prod_{i \in I} X_i, \tau)$ where $\tau$ is the product topology.

**Proof:** By the Alexander subbasis theorem, the theorem will be proved if every subbasic open cover admits a finite subcover. Therefore, let $O$ be a subbasic open cover of $X \equiv \prod_{i \in I} X_i$. Let

$$O_j = \{Q \in O : \pi_i Q = X_i \text{ for } i \neq j\}$$

$$\pi_j O_j = \{\pi_j Q : Q \in O_j\}$$

Thus $O_j$ are those sets which might have a proper open subset of $X_j$ in the $j^{th}$ position. If each $\pi_j O_j$ fails to cover $X_j$, then there exists

$$f \in \prod_{j \in I} X_j \setminus \cup_{j \in I} \pi_j O_j$$

Now $f$ is contained in some open set from $O$ which must be in some $O_i$. Hence $\pi_j f = f(j) \in \cup \pi_j O_j$ but this does not happen. Hence for some $j$, $\pi_j O_j$ must cover $X_j$.

$$X_j = \cup \pi_j O_j$$

and so by compactness of $X_j$, there exist $A_1, \ldots, A_m$, sets in $\tau_j$ such that $X_j \subseteq \cup_{k=1}^m A_k$ and letting $\pi_j U_k = A_k$ for $U_k \in O_j$, $\{U_k\}_{k=1}^m$ covers $\prod_{i \in I} X_i$. By the Alexander subbasis theorem this proves $\prod_{i \in I} X_i$ is compact. ■

### 1.13 Exercises

1. Let $d(x, y) = |x - y|$ for $x, y \in \mathbb{R}$. Show that this is a metric on $\mathbb{R}$.

2. Now consider $\mathbb{R}^n$. Let $||x||_\infty \equiv \max \{|x_i|, i = 1, \ldots, n\}$. Define

$$d(x, y) \equiv ||x - y||_\infty.$$

Show that this is a metric on $\mathbb{R}^n$. In the case of $n = 2$, describe the ball $B(0, r)$. **Hint:** First show that $||x + y|| \leq ||x|| + ||y||$. 


3. Let $C([0,T])$ denote the space of functions which are continuous on $[0,T]$.

Define

$$\|f\| = \sup_{t \in [0,T]} |f(t)| = \max_{t \in [0,T]} |f(t)|$$

Verify the following. $\|f + g\| \leq \|f\| + \|g\|$. Then use to show that $d(f, g) = \|f - g\|$ is a metric and that with this metric, $(C([0,T]), d)$ is a metric space.

4. Recall that $[a, b]$ is compact. This was done in single variable advanced calculus. It comes from the least upper bound property for completeness. That is, every sequence has a convergent subsequence. Also recall that a sequence of numbers $\{x_n\}$ is a Cauchy sequence means that for every $\varepsilon > 0$ there exists $N$ such that if $m, n > N$, then $|x_n - x_m| < \varepsilon$. First show that every Cauchy sequence is bounded. Next, using the compactness of closed intervals, show that every Cauchy sequence has a convergent subsequence. Thus $\mathbb{R}$ with the usual metric just described is complete because every Cauchy sequence converges. This is showing that the least upper bound property implies that every Cauchy sequence converges.

5. Using the result of the above problem, show that $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a complete metric space. That is, every Cauchy sequence converges. Here $d(x, y) \equiv \|x - y\|_\infty$.

6. Suppose you had $(X_i, d_i)$ is a metric space. Now consider the product space

$$X = \prod_{i=1}^{n} X_i$$

with $d(x, y) = \max \{d_i(x_i, y_i), i = 1 \cdots n\}$. Would this be a metric space? If so, prove that this is the case.

Does triangle inequality hold? **Hint:** For each $i$,

$$d_i(x_i, z_i) \leq d_i(x_i, y_i) + d_i(y_i, z_i) \leq d(x, y) + d(y, z)$$

Now take max of the two ends.

7. In the above example, if each $(X_i, d_i)$ is complete, explain why $(X, d)$ is also complete.

8. Show that $C([0,T])$ is a complete metric space. That is, show that if $\{f_n\}$ is a Cauchy sequence, then there exists $f \in C([0,T])$ such that $\lim_{n \to \infty} d(f, f_n) = \lim_{n \to \infty} \|f - f_n\| = 0$. **Hint:** First, you know that $\{f_n(t)\}$ is a Cauchy sequence for each $t$. Why? Now let $f(t)$ be the name of the thing to which $f_n(t)$ converges. Recall why the uniform convergence implies $t \to f(t)$ is continuous. Give the proof. It was done in single variable advanced calculus. Review and write down proof. Also show that $\|f - f_n\| \to 0$. 
9. Let $X$ be a nonempty set of points. Say it has infinitely many points. Define $d(x, y) = 1$ if $x \neq y$ and $d(x, y) = 0$ if $x = y$. Show that this is a metric. Show that in $(X, d)$ every point is open and closed. In fact, show that every set is open and every set is closed. Is this a complete metric space? Explain why. Describe the open balls.

10. Show that the union of any set of open sets is an open set. Show the intersection of any set of closed sets is closed. Let $A$ be a nonempty subset of a metric space $(X, d)$. Then the closure of $A$, written as $\overline{A}$ is defined to be the intersection of all closed sets which contain $A$. Show that $A = A \cup A'$. That is, to find the closure, you just take the set and include all limit points of the set.

11. Let $A'$ denote the set of limit points of $A$, a nonempty subset of a metric space $(X, d)$. Show that $A'$ is closed.

12. A theorem was proved which gave three equivalent descriptions of compactness of a metric space. One of them said the following: A metric space is compact if and only if it is complete and totally bounded. Suppose $(X, d)$ is a complete metric space and $K \subseteq X$. Then $(K, d)$ is also clearly a metric space having the same metric as $X$. Show that $(K, d)$ is compact if and only if it is closed and totally bounded. Note the similarity with the Heine Borel theorem on $\mathbb{R}$. Show that on $\mathbb{R}$, every bounded set is also totally bounded. Thus the earlier Heine Borel theorem for $\mathbb{R}$ is obtained.

13. Suppose $(X_i, d_i)$ is a compact metric space. Then the Cartesian product is also a metric space. That is $(\prod_{i=1}^n X_i, d)$ is a metric space where $d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max\{d_i(x_i, y_i)\}$. Show that $(\prod_{i=1}^n X_i, d)$ is compact. Recall the Heine Borel theorem for $\mathbb{R}$. Explain why $\prod_{i=1}^n [a_i, b_i]$ is compact in $\mathbb{R}^n$ with the distance given by $d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max\{|x_i - y_i|\}$. **Hint:** It suffices to show that $(\prod_{i=1}^n X_i, d)$ is sequentially compact. Let $(x^n)_{m=1}^\infty$ be a sequence. Then $(x^n_1)_{m=1}^\infty$ is a sequence in $X_1$. Therefore, it has a subsequence $(x^n_{1k_1})_{k_1=1}^\infty$ which converges to a point $x_1 \in X_1$. Now consider $(x^n_{2k_2})_{k_2=1}^\infty$ the second components. It has a subsequence denoted as $k_2$ such that $(x^n_{2k_2})_{k_2=1}^\infty$ converges to a point $x_2 \in X_2$. Explain why $\lim_{k_2 \to \infty} x^n_{2k_2} = x_2 \in X_2$. Continue doing this $n$ times. Explain why $\lim_{k_n \to \infty} x^n_{1k_n} = x_l \in X_l$ for each $l$. Then explain why this is the same as saying $\lim_{k_n \to \infty} x^{k_n} = x$ in $(\prod_{i=1}^n X_i, d)$.

14. If you have a metric space $(X, d)$ and a compact subset of $(X, d)$ $K$, suppose that $L$ is a closed subset of $K$. Explain why $L$ must also be compact. **Hint:** Go right to the definition. Take an open covering of $L$ and consider this along with the open set $L^C$ to obtain an open covering of $K$. Now use compactness.
of $K$. Use this to explain why every closed and bounded set in $\mathbb{R}^n$ is compact. Here the distance is given by $d(x, y) \equiv \max_{1 \leq i \leq n} \{|x_i - y_i|\}$.

15. Show that compactness is a topological property in the following sense. If $(X, d), (Y, \rho)$ are both metric spaces and $f : X \rightarrow Y$ has the property that $f$ is one to one, onto, and continuous, and also $f^{-1}$ is one to one onto and continuous, then the two metric spaces are compact or not compact together. That is one is compact if and only if the other is.

16. Consider $\mathbb{R}$ the real numbers. Define a distance in the following way.

$$\rho(x, y) \equiv |\arctan(x) - \arctan(y)|$$

Show this is a good enough distance and that the open sets which come from this distance are the same as the open sets which come from the usual distance $d(x, y) = |x - y|$. Explain why this yields that the identity mapping $f(x) = x$ is continuous with continuous inverse as a map from $(\mathbb{R}, d)$ to $(\mathbb{R}, \rho)$. To do this, you show that an open ball taken with respect to one of these is also open with respect to the other. However, $(\mathbb{R}, \rho)$ is not a complete metric space while $(\mathbb{R}, d)$ is. Thus, unlike compactness. Completeness is not a topological property. **Hint:** To show the lack of completeness of $(\mathbb{R}, \rho)$, consider $x_n = n$. Show it is a Cauchy sequence with respect to $\rho$.

17. It is useful to define the following distance function. Let $(X, d)$ be a metric space and $S \subseteq X, S \neq \emptyset$. Then $\operatorname{dist}(x, S) \equiv \inf\{d(x, y) : y \in S\}$. Show that this always satisfies

$$|\operatorname{dist}(x, S) - \operatorname{dist}(z, S)| \leq d(x, z)$$

This is a really neat result.

18. If $K$ is a compact subset of $(X, d)$ and $y \notin K$, show that there always exists $x \in K$ such that $d(x, y) = \operatorname{dist}(y, K)$. Give an example in $\mathbb{R}^2$ to show two disjoint closed sets such that the distance between them is 0.

19. You know that if $f : X \rightarrow X$ for $X$ a complete metric space, then if

$$d(f(x), f(y)) < rd(x, y)$$

it follows that $f$ has a unique fixed point theorem. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$f(t) = t + (1 + e^t)^{-1}$$

Show that $|f(t) - f(s)| < |t - s|$, but $f$ has no fixed point.

20. Let $X$ be a complete metric space. Let $U_n$ be a dense open set. That is, $\overline{U_n} = X$. Show that $D \equiv \cap_{n=1}^{\infty} U_n$ is dense.
1.13. EXERCISES

21. If \((X, d)\) is a metric space, show that there is a bounded metric \(\rho\) such that the open sets for \((X, d)\) are the same as those for \((X, \rho)\).

22. Let \((X, d)\) be a metric space where \(d\) is a bounded metric. Let \(C\) denote the collection of closed subsets of \(X\). For \(A, B \in C\), define

\[
\rho(A, B) \equiv \inf \{ \delta > 0 : A_\delta \supseteq B \text{ and } B_\delta \supseteq A \}
\]

where for a set \(S\),

\[
S_\delta \equiv \{ x : \text{ dist}(x, S) \equiv \inf \{ d(x, s) : s \in S \} \leq \delta \}.
\]

Show \(x \to \text{ dist}(x, S)\) is continuous and that therefore, \(S_\delta\) is a closed set containing \(S\). Also show that \(\rho\) is a metric on \(C\). This is called the Hausdorff metric.

23. Suppose \((X, d)\) is a compact metric space. Show \((C, \rho)\) is a complete metric space. **Hint:** Show first that if \(W_n \downarrow W\) where \(W_n\) is closed, then \(\rho(W_n, W) \to 0\). Now let \(\{A_n\}\) be a Cauchy sequence in \(C\). Then if \(\varepsilon > 0\) there exists \(N\) such that when \(m, n \geq N\), then \(\rho(A_n, A_m) < \varepsilon\). Therefore, for each \(n \geq N\),

\[
(A_n)_\varepsilon \supseteq \bigcup_{k=n}^\infty A_k.
\]

Let \(A \equiv \cap_{n=1}^\infty \bigcup_{k=n}^\infty A_k\). By the first part, there exists \(N_1 > N\) such that for \(n \geq N_1\),

\[
\rho\left(\bigcup_{k=n}^\infty A_k, A\right) < \varepsilon, \quad \text{and} \quad (A_n)_\varepsilon \supseteq \bigcup_{k=n}^\infty A_k.
\]

Therefore, for such \(n\), \(A_\varepsilon \supseteq W_n \supseteq A_n\) and \((W_n)_\varepsilon \supseteq (A_n)_\varepsilon \supseteq A\) because

\[
(A_n)_\varepsilon \supseteq \bigcup_{k=n}^\infty A_k \supseteq A.
\]

24. Let \(X\) be a compact metric space. Show \((C, \rho)\) is compact. **Hint:** Let \(D_n\) be a \(2^{-n}\) net for \(X\). Let \(K_n\) denote finite unions of sets of the form \(B(p, 2^{-n})\) where \(p \in D_n\). Show \(K_n\) is a \(2^{-(n-1)}\) net for \((C, \rho)\).

25. Consider the space of \(m \times n\) matrices. Let \(\|A\| = \sqrt{\text{trace}(A^*A)}\) and let \(\|A - B\| = \text{dist}(A, B)\). Show this is a norm on the space of \(m \times n\) matrices.

26. An \(n \times n\) matrix \(U\) is unitary if \(U^*U = I\). Show that any sequence of unitary matrices has a converging subsequence which converges to a unitary matrix with respect to the norm \(\|A\| = (\text{trace}(A^*A))^{1/2}\).

27. A function \(f : X \to \mathcal{P}(Y)\) (Power set of \(Y\), set of all subsets of \(Y\)) where \((X, d), (Y, \rho)\) are metric spaces is called upper semicontinuous if whenever \(x_n \to x\) and \(U\) is an open set containing \(f(x)\), it follows that \(f(x_n) \in U\) for all \(n\) large enough. Now consider the space of \(n \times n\) matrices. The distance between two of these \(A, B\) will be \(\|A - B\|\) where \(\|A\|^2 = \text{trace}(A^*A)\). Let \(f(A)\) denote the set of eigenvalues of \(A\). Show that \(f\) is upper semicontinuous.
28. A separable metric space \((X,d)\) is one for which there is a countable dense subset \(D = \{d_k\}_{k=1}^{\infty}\). This means that if \(B(x,r)\) is an arbitrary ball, then \(B(x,r) \cap D \neq \emptyset\). Show that in any separable metric space \((X,d)\), there exists a countable set \(B\) of balls such that every open subset of \((X,d)\) is the union of these balls. This \(B\) is called a countable basis. A complete separable metric space is called a Polish space.

29. Show that every separable metric space has the Lindelöf property. This means that if you have any set \(S\) and an open cover \(C\) of \(S\), then there exists a countable subset of \(C\) which also covers \(S\).

30. Let \((X_i,d_i)\) be a complete separable metric space, a Polish space \(i \in \mathbb{N}\). Consider the infinite product, \(X \equiv \prod_{i=1}^{\infty} X_i\) with the following metric. Denoting \(x,y\) as two points in \(X\),

\[
d(x,y) = \sum_{i=1}^{\infty} 2^{-i} \frac{d_i(x_i,y_i)}{1 + d_i(x_i,y_i)}
\]

Show \((X,d)\) is a Polish space. Also show that it has the same topology as \(X \equiv \prod_{i=1}^{\infty} X_i\) with the product topology \(\tau\).

31. Let

\[
X = \prod_{i=1}^{\infty} X_i
\]

where each \((X_i,d_i)\) is a compact metric space. Define a distance as in the above problem. Show that \((X,d)\) is a compact metric space directly without using the Alexander subbasis theorem or the general Tychonoff theorem. Instead, use a Cantor diagonalization argument.

32. A Banach space \(X\) is a complete normed vector space. It dual space, defined as the set of continuous linear \(F\) valued mappings is denoted as \(X'\). Here \(F\) is \(\mathbb{R}\) or \(\mathbb{C}\). Then show that a linear map \(f\) from \(X\) to \(F\), more generally to another Banach space, is continuous if and only if any of the following equivalent conditions hold.

(a) \(f\) is continuous at 0

(b) \(f\) is bounded, \(\|f(x)\| \leq C\|x\|\)

(c) \(f\) is continuous at every \(x\).

For \(f \in X', \|f\| \equiv \sup_{\|x\| \leq 1} |f(x)|\). Then show that this defines a norm and with respect to this norm, \(\|f(x)\| \leq \|f\| \|x\|\), \(X'\) is a Banach space. When \(f : X \to Y\) where \(Y\) is another Banach space and \(f\) is linear, then the same three equivalences hold for \(f\) to be continuous. Verify that this is also true. The set of such continuous linear functions is denoted as \(\mathcal{L}(X,Y)\). The notation \(\langle x^*, x \rangle\) will often be used to denote \(x^*(x)\) where here \(x^* \in X'\) and \(x \in X\). This will be especially the case in the material on nonlinear operators.
33. Let there be a set \( S \subseteq \mathcal{L}(X,Y) \) where \( Y \) is a normed linear space and \( X \) is a Banach space. Thus \( f \in S \) means it is linear and \( \sup_{\|x\| \leq 1} \|f(x)\| \equiv \|f\|. \)

(a) Let \( U_n \equiv \{ x \in X : \|f(x)\| > n \|x\| \text{ for some } f \in S \} \). Show that \( U_n \) is an open set.

(b) Using Problem 24 show that if each \( U_n \) is dense, then there exists a dense subset of \( X \) such that for \( x \) in this set \( \sup_{f \in S} \|f(x)\| = \infty \)

(c) The alternative is that some \( U_n \) is not dense. Show that this implies there exists \( M \) such that \( \|f\| < M \) for all \( f \in S \).

**Weak and Weak * Topologies Compactness**

34. Let \( X \) be a Banach space. Then on \( X' \) one can define a topology known as the weak * topology as follows. A subbasis for this topology will be sets of the form \( B_x(f,\delta) \) defined as \( \{ g \in X' : |g(x) - f(x)| < \delta \} \). Verify that this is a subbasis for a topology which has a basis of the form

\[
B_{x_1,\ldots,x_n}(f,\delta) \equiv \left\{ g \in X' : \max_{i=1\ldots n} \{|g(x_i) - f(x_i)|\} < \delta \right\}
\]

Now let \( \overline{B}(0,1) \) be the closed unit ball in \( X' \). Suppose \( X \) is separable. Verify that in this case, the relative topology on this closed unit ball is the same as would be obtained if the subbasis were defined only in terms of \( x \in D \) where \( D \) is a countable dense subset of \( X \). **Hint:** Say \( g \in B_x(f,\delta) \) and show there exists \( d \in D, 0 < r < 1 \) such that

\[
|g(d) - f(x)| < r\delta, \quad \|d - x\| < \frac{1}{2} \left( \delta - \frac{\delta (1 + r)}{2} \right)
\]

Then show that \( B_d(g, \frac{1}{2}\delta (1 - r)) \subseteq B_x(f,\delta) \). The first set is one of those defined in terms of \( D \) and this shows that an arbitrary basic set is the union of these specialized basic sets. Let \( \tau \) be the weak* topology on \( \overline{B}(0,1) \) and let \( \tau_D \) be this new one. Note that \( B_{x_1,\ldots,x_n}(f,\delta) = \cap_{i=1}^n B_{x_i}(f,\delta) \) and so if you have \( g \) in this basic open set of \( \tau \), you could obtain \( g \) is in an open set from \( \tau_D \) which is contained in \( B_{x_1,\ldots,x_n}(f,\delta) \). Thus every open set in \( \tau \) is open in \( \tau_D \) so \( \tau \subseteq \tau_D \) but the other inclusion is obvious.

35. The elements of \( \overline{B}(0,1) \equiv \overline{B} \) are functions defined on \( X \). For \( f \in \overline{B}, f(x) \in \overline{B}(0,\|x\|) \). These functions in \( \overline{B} \) happen to be linear functions of course. However, you could consider the infinite product space

\[
\prod_{x \in X} \overline{B}(0,\|x\|) \subseteq \prod_{x \in X} \mathbb{F}
\]

with the product topology. This is also the set of functions defined on \( X \) such that \( f(x) \in \overline{B}(0,\|x\|) \). By Tychonoff’s theorem, the set \( \prod_{x \in X} \overline{B}(0,\|x\|) \) is compact in \( \prod_{x \in X} \mathbb{F} \) with respect to the product topology. Then \( \overline{B} \subseteq \overline{B}(0,1) \)
Letting

The next few problems are an alternate treatment of this important topic. Let

34

and

that when

36.

CHAPTER 1. METRIC SPACES AND GENERAL TOPOLOGICAL SPACES

\( \prod_{x \in X} B(0, \|x\|) \) and the weak * topology on \( \tilde{B} \) is the same as the product topology on \( \prod_{x \in X} \mathbb{F} \). Now I claim that \( \tilde{B} \) is actually a closed subset of the compact set \( \prod_{x \in X} B(0, \|x\|) \) with respect to the product topology. It is clearly contained in this compact set. Suppose \( h \notin \tilde{B} \) but \( h \in \prod_{x \in X} B(0, \|x\|) \). Then the only way this can happen is for \( h \) to fail to be linear. Hence there exist scalars \( a, b \) and vectors of \( X, x, y \) such that

\[ h(ax + by) \neq ah(x) + bh(y) \]

Consider \( B_{x,y,ax+by}(h, 1/n) \). If for all \( n \) this basic set contains some \( f_n \in B \), then \( f_n(x) \to h(x), f_n(ax + by) \to h(ax + by) \) and \( f_n(y) \to h(y) \). Then, since \( f_n \) is linear,

\[ h(ax + by) = \lim_{n \to \infty} f_n(ax + by) = \lim_{n \to \infty} (af_n(x) + bf_n(y)) = ah(x) + bh(y) \]

which is a contradiction. Hence for some \( n \), \( B_{x,y,ax+by}(h, 1/n) \) fails to intersect \( \tilde{B} \) showing that \( \tilde{B} \) is indeed closed. Since it is a closed subset of a compact set, it is compact in the product topology which coincides with the weak * topology. This is the Banach Alaoglu theorem. It is really Tychonoff’s theorem. The case of most interest here is the one where \( X \) is separable. Show using Problems 36 and 37 that when \( X \) is separable, \( \tilde{B} \) is not just compact, but also weak * sequentially compact. Note that from the definition of weak * convergence, \( f_n \) converges weak * to \( f \) if and only if \( \lim_{n \to \infty} f_n(x) = f(x) \) for all \( x \in X \). Verify that this is the case also.

36. Letting \( X \) be a Banach space, you can consider \( X \) as being contained in \( X'' \) as follows. For \( f \in X' \), let \( x(f) \equiv f(x) \). Then this convention does define \( x \in X'' \) because it is linear and continuous. Show this. Why do we care? When every element of \( X'' \) is obtained this way, the space is called a reflexive Banach space. Then the weak * topology on \( X \subseteq X'' \) is obtained from subbasic sets of the form

\[ B_f(x, \delta) \equiv \{ y \in X : |f(x) - f(y)| = |x(f) - y(f)| < \delta \} \]

This is called the weak topology on \( X \). When \( X = X'' \), it follows from the above problems that bounded sets in \( X \) are weakly sequentially compact provided \( X' \) is also separable. (This assumption is not needed. For the applications of most interest, it is clear anyway. ) This is part of the Eberlein Smulian theorem. Fill in the details. You need to explain why \( x \in X \) in this definition is also a continuous linear functional defined on \( X' \). For your information, it is not hard to generalize to the case where \( X \) is not separable. You show that it is possible to reduce to the case of a separable Banach space. We will denote the identification of \( x \in X \) with \( x \in X'' \) by a mapping called \( J \). Thus \( J : X \to X'' \) is given as above.

A different approach

37. The next few problems are an alternate treatment of this important topic. Let

\( X \) be a real Banach space (It works just as well for complex Banach spaces.)
1.13. EXERCISES

Consider the product space

\[ P \equiv \prod_{x \in X} \mathbb{R} \]

These are the functions mapping \( X \) to \( \mathbb{R} \). The set of linear functions defined on \( X \) denoted as \( L \), is a closed subspace of \( P \). The weak * topology on \( L \) is just the relative topology obtained from the product topology on \( P \). Thus a subbasis for the weak * topology consists of \( \pi^{-1}_{x} (B(f(x), r)) \cap L \) where \( B(f(x), r) \) is an open ball in \( \mathbb{R} \) and \( \pi_{x} \) is the projection map from \( P \) to \( x \).

Note that for \( f \in P, \pi_{x} f = f(x) \). An open set in the subbasis for the weak * topology on \( L \) just described consists of \( \{ g \in P : |g(x) - f(x)| < r \} \cap L \). We denote this set as \( B_{x}(f, r) \).

Now let \( B(0,1) \) be the closed unit ball in \( X' \). This means that \( \|f\| \leq 1 \) and \( f \) is linear. Then

\[ B \equiv B(0,1) \subseteq \prod_{x \in X} B(0,\|x\|) \equiv K \]

where in the product \( B(0,\|x\|) \) is the closed ball of radius \( \|x\| \) in \( \mathbb{R} (\mathbb{C}) \). Show why \( L \) is a closed subset of \( P \) and then show why \( B(0,1) \) is a closed subset of \( \prod_{x \in X} B(0,\|x\|) \) in the product topology. Then explain why \( B(0,1) \), the closed unit ball in \( X' \) is weak * compact. This is the Banach Alaoglu theorem.

38. In the context of the above problem, suppose \( D \) is a countable dense subset of \( X \). Then you could consider a subbasis for a topology on \( X' \) to be sets of the form

\[ \{ g \in P : |g(x) - f(x)| < r \} \cap L \equiv B_{x}(f, r), \ x \in D \]

Show that these special sets yield the same topology on \( B \) where \( B = B(0,\|x\|) \subseteq X' \). Then show that this weak * topology on \( B \) is actually a metric space.

39. Let \( D_{0} \) be a countable dense subset of \( X \) in the above problem. Let \( D \) be the span of \( D_{0} \) using only rational scalars. Show that for \( L \) those functions which are linear on \( D \) with field of scalars the rational numbers,

\[ \prod_{x \in D} B(0,\|x\|) \cap L \subseteq P \equiv \prod_{x \in D} \mathbb{R} \]

which is a compact set by Tychonoff’s theorem. The product space on the right is also a metric space because \( D \) is countable. Verify that \( L \) is a closed subspace of \( P \) with respect to the product topology. Then tell why \( \prod_{x \in D} B(0,\|x\|) \cap L \) is a compact set in the product topology. Next show that if \( \{f_{n}\} \) is a bounded sequence in \( X' \), there is a subsequence \( \{f_{n_{k}}\} \) and an \( f \in X' \) such that \( f_{n_{k}}(x) \to f(x) \) for all \( x \in X \). Thus this proves the following theorem: If \( \{f_{n}\} \) is a bounded sequence in \( X' \) and if \( X \) is separable, then there is a subsequence \( \{f_{n_{k}}\} \) and a function \( f \in X' \) such that for all \( x \in X, f_{n_{k}}(x) \to f(x) \).
40. Letting $X$ be a Banach space, you can consider $X$ as being contained in $X''$ as follows. For $f \in X'$, let $x(f) \equiv f(x)$. Then this convention does define $x \in X''$ because it is linear and continuous. Show this. Why do we care? When every element of $X''$ is obtained this way, the space is called a reflexive Banach space. The weak topology on $X$ is really just the weak * topology on $X''$ restricted to $X$. When $X = X''$ so that every element of $X''$ is obtained as just described and $X'$ is separable, show that if $\{x_n\}$ is any bounded sequence in $X$, there exists a subsequence $\{x_{n_k}\}$ and $x \in X$ such that for all $f \in X', f(x_{n_k}) \equiv x_{n_k}(f) \to x(f) \equiv f(x)$. This is part of the Eberlein Smulian theorem. You don’t have to assume $X'$ is separable but this is shown later. In most examples of interest it is separable.

41. Let $(\mathbb{R}^n, \mathcal{F}, \mu)$ be a measure space where $\mu$ is a regular complete $\sigma$ finite Borel measure. Explain why for $\Omega$ a Borel subset of $\mathbb{R}^n$, the unit ball in $L^\infty(\Omega)$ is weak * sequentially compact. Meaning that if you have a bounded set $\{f_n\}$ in $L^\infty(\Omega)$, there will be a subsequence $\{f_{n_k}\}$ and a function $f \in L^\infty(\Omega)$ such that for all $g \in L^1(\Omega)$, $\int_\Omega f_{n_k} g d\mu \to \int_\Omega f g d\mu$. 
Chapter 2

Hahn Banach Theorem, Convexity

The Hahn-Banach theorem has nothing to do with topology. However, it is one of the major results in functional analysis. I am presenting the essentials here and giving applications in the exercises. It is very surprising that this major theorem has a lot to do with the separation of convex sets. This theorem is an essential tool in nonlinear analysis.

Before presenting this theorem, here are some preliminaries about partially ordered sets.

Theorem 2.0.1 (Hausdorff Maximal Principle) Let $F$ be a nonempty partially ordered set. Then there exists a maximal chain.

2.1 Gauge Functions And Hahn Banach Theorem

Definition 2.1.1 Let $X$ be a real vector space $\rho: X \to \mathbb{R}$ is called a gauge function if

\begin{align*}
\rho(x + y) &\leq \rho(x) + \rho(y), \\
\rho(ax) &= a\rho(x) \text{ if } a \geq 0. \tag{2.1.1}
\end{align*}

Suppose $M$ is a subspace of $X$ and $z \notin M$. Suppose also that $f$ is a linear real-valued function having the property that $f(x) \leq \rho(x)$ for all $x \in M$. Consider the problem of extending $f$ to $M \oplus \mathbb{R}z$ such that if $F$ is the extended function, $F(y) \leq \rho(y)$ for all $y \in M \oplus \mathbb{R}z$ and $F$ is linear. Since $F$ is to be linear, it suffices to determine how to define $F(z)$. Letting $a > 0$, it is required to define $F(z)$ such that the following hold for all $x, y \in M$.

\[
F(x) + aF(z) = F(x + az) \leq \rho(x + az),
\]
CHAPTER 2. HAHN BANACH THEOREM, CONVEXITY

\[
\widehat{F}(y) - aF(z) = F(y - az) \leq \rho(y - az). \tag{2.1.2}
\]

Now if these inequalities hold for all \( y/a \), they hold for all \( y \) because \( M \) is given to be a subspace. Therefore, multiplying by \( a^{-1} \) implies that what is needed is to choose \( F(z) \) such that for all \( x,y \in M \),

\[
f(x) + F(z) \leq \rho(x + z), \quad f(y) - \rho(y - z) \leq F(z)
\]

and that if \( F(z) \) can be chosen in this way, this will satisfy (2.1.2) for all \( x,y \) and the problem of extending \( f \) will be solved. Hence it is necessary to choose \( F(z) \) such that for all \( x,y \in M \),

\[
f(y) - \rho(y - z) \leq F(z) \leq \rho(x + z) - f(x). \tag{2.1.3}
\]

Is there any such number between \( f(y) - \rho(y - z) \) and \( \rho(x + z) - f(x) \) for every pair \( x,y \in M \)? This is where \( f(x) \leq \rho(x) \) on \( M \) and that \( f \) is linear is used. For \( x,y \in M \),

\[
\rho(x + z) - f(x) - [f(y) - \rho(y - z)]
\]

\[
= \rho(x + z) + \rho(y - z) - (f(x) + f(y))
\]

\[
\geq \rho(x + y) - f(x + y) \geq 0.
\]

Therefore there exists a number between

\[
\sup \{f(y) - \rho(y - z) : y \in M\}
\]

and

\[
\inf \{\rho(x + z) - f(x) : x \in M\}
\]

Choose \( F(z) \) to satisfy (2.1.3). This has proved the following lemma.

**Lemma 2.1.2** Let \( M \) be a subspace of \( X \), a real linear space, and let \( \rho \) be a gauge function on \( X \). Suppose \( f : M \to \mathbb{R} \) is linear, \( z \notin M \), and \( f(x) \leq \rho(x) \) for all \( x \in M \). Then \( f \) can be extended to \( M \oplus \mathbb{R}z \) such that, if \( F \) is the extended function, \( F \) is linear and \( F(x) \leq \rho(x) \) for all \( x \in M \oplus \mathbb{R}z \).

With this lemma, the Hahn Banach theorem can be proved.

**Theorem 2.1.3** (Hahn Banach theorem) Let \( X \) be a real vector space, let \( M \) be a subspace of \( X \), let \( f : M \to \mathbb{R} \) be linear, let \( \rho \) be a gauge function on \( X \), and suppose \( f(x) \leq \rho(x) \) for all \( x \in M \). Then there exists a linear function, \( F : X \to \mathbb{R} \), such that

a.) \( F(x) = f(x) \) for all \( x \in M \)

b.) \( F(x) \leq \rho(x) \) for all \( x \in X \).
2.2. SEPARATION THEOREMS

Proof: Let \( F = \{(V,g) : V \supseteq M, V \text{ is a subspace of } X, g : V \to \mathbb{R} \text{ is linear}, g(x) = f(x) \text{ for all } x \in M, \text{ and } g(x) \leq \rho(x) \text{ for } x \in V \} \). Then \((M,f) \in F \) so \( F \neq \emptyset \). Define a partial order by the following rule.

\[(V,g) \leq (W,h) \]

means

\[ V \subseteq W \text{ and } h(x) = g(x) \text{ if } x \in V. \]

By Theorem 2.0.1, there exists a maximal chain, \( C \subseteq F \). Let \( Y = \bigcup \{(V,g) : (V,g) \in C \} \)
and let \( h : Y \to \mathbb{R} \) be defined by \( h(x) = g(x) \) where \( x \in V \) and \((V,g) \in C \). This is well defined because if \( x \in V_1 \) and \( V_2 \) where \((V_1,g_1) \) and \((V_2,g_2) \) are both in the chain, then since \( C \) is a chain, the two element related. Therefore, \( g_1(x) = g_2(x) \).

Also \( h \) is linear because if \( ax + by \in Y \), then \( x \in V_1 \) and \( y \in V_2 \) where \((V_1,g_1) \) and \((V_2,g_2) \) are elements of \( C \). Therefore, letting \( V \) denote the larger of the two \( V_i \), and \( g \) be the function that goes with \( V \), it follows \( ax + by \in V \) where \((V,g) \in C \). Therefore,

\[ h(ax + by) = g(ax + by) = ag(x) + bg(y) = ah(x) + bh(y). \]

Also, \( h(x) = g(x) \leq \rho(x) \) for any \( x \in Y \) because for such \( x \), \( x \in V \) where \((V,g) \in C \).

Is \( Y = X \)? If not, \( Y \neq X \) because for such \( x \), \( x \in V \) where \((V,g) \in C \).

2.2 Separation Theorems

Here \( X \) will be a real Banach space. A set \( K \), is said to be convex if whenever \( x, y \in K \),

\[ \lambda x + (1 - \lambda) y \in K \]

for all \( \lambda \in [0,1] \).

Definition 2.2.1 Let \( U \) be an open convex set containing \( 0 \) and define

\[ m(x) = \inf \{ t > 0 : x/t \in U \}. \]

This is called a Minkowski functional.

Proposition 2.2.2 Let \( X \) be a real Banach space, \( U \) is an open convex set containing \( 0 \). Then \( m \) is defined on \( X \) and satisfies

\[ m(x + y) \leq m(x) + m(y) \]  \hspace{1cm} (2.2.4)

\[ m(\lambda x) = \lambda m(x) \text{ if } \lambda > 0. \]  \hspace{1cm} (2.2.5)

Thus, \( m \) is a gauge function on \( X \).
**Proof:** Let \( x \in X \) be arbitrary. There exists

\[
0 \in B (0, r) \subseteq U.
\]

Then if \( x \) is arbitrary, you have \( \frac{x}{2} \in B (0, r) \) whenever \( t \) is large enough. In fact, you could take \( \frac{1}{t} = 2^{-1} \| x \|^{-1} r. \) Therefore, \( m (x) \leq \frac{2 \| x \|}{r} \). Thus \( m (x) \) is defined for any \( x. \)

Let \( x/t \in U, y/s \in U \). Then since \( U \) is convex,

\[
\frac{x + y}{t + s} = \left( \frac{t}{t + s} \right) \left( \frac{x}{t} \right) + \left( \frac{s}{t + s} \right) \left( \frac{y}{s} \right) \in U.
\]

It follows that

\[
m (x + y) \leq t + s.
\]

Choosing \( s, t \) such that \( t - \varepsilon < m (x) \) and \( s - \varepsilon < m (y), \)

\[
m (x + y) \leq m (x) + m (y) + 2 \varepsilon.
\]

Since \( \varepsilon \) is arbitrary, this shows \( 2.2.4 \). It remains to show \( 2.2.5 \). Let \( x/t \in U. \) Then if \( \lambda > 0, \)

\[
\frac{\lambda x}{\lambda t} \in U
\]

and so \( m (\lambda x) \leq \lambda t. \) Thus \( m (\lambda x) \leq \lambda m (x) \) for all \( \lambda > 0. \) Hence

\[
m (x) = m (\lambda^{-1} \lambda x) \leq \lambda^{-1} m (\lambda x) \leq \lambda^{-1} \lambda m (x) = m (x)
\]

and so

\[
\lambda m (x) = m (\lambda x). \quad \square
\]

**Lemma 2.2.3** Let \( U \) be an open convex set containing \( 0 \) and let \( q \notin U. \) Then there exists \( f \in X' \) such that

\[
f (q) > f (x)
\]

for all \( x \in U. \)

**Proof:** Let \( m \) be the Minkowski functional just defined and let

\[
F (cq) = cm (q)
\]

for \( c \in \mathbb{R}. \) If \( c > 0 \) then

\[
F (cq) = m (cq)
\]

while if \( c \leq 0, \)

\[
F (cq) = cm (q) \leq 0 \leq m (cq).
\]

By the Hahn Banach theorem, \( F \) has an extension \( g, \) defined on all of \( X \) satisfying

\[
g (x + y) = g (x) + g (y), \quad g (cx) = cg (x)
\]
2.2. SEPARATION THEOREMS

for all \( c \in \mathbb{R} \), and
\[
g(x) \leq m(x).
\]
Thus, \( g(-x) \leq m(-x) \) and so
\[
-m(-x) \leq g(x) \leq m(x), \quad -g(x) = g(-x) \leq m(-x)
\]
It follows as in the Proposition \[2.2.2\] that
\[
m(x) \leq C \| x \|
\]
and hence \( g \) is continuous by Problem \[32\] on Page \[48\].

**Corollary 2.2.4** Let \( U \) be an open nonempty convex set and let \( q \notin U \). Then there exists \( f \in X' \) such that
\[
f(q) > f(x)
\]
for all \( x \in U \).

**Proof:** Let \( u_0 \in U \) and consider \( \hat{U} = U - u_0 \). Then \( 0 \in \hat{U} \) and \( q - u_0 \notin \hat{U} \). By Lemma \[2.2.3\], there exists \( f \in X' \) such that
\[
f(q - u_0) > f(x - u_0)
\]
for all \( x \in U \). Thus \( f(q) > f(x) \) for all \( x \in U \). \( \blacksquare \)

**Theorem 2.2.5** Let \( K \) be closed and convex in a real Banach space and let \( p \notin K \). Then there exists a real number, \( c \), and \( f \in X' \) such that
\[
f(p) > c > f(k)
\]
for all \( k \in K \).

**Proof:** Since \( K \) is closed, and \( p \notin K \), there exists \( r > 0 \) such that
\[
K \cap B(p, r) = \emptyset.
\]
Thus
\[
p \notin K + B(0, r)
\]
Pick \( k_0 \in K \) and let
\[
U = K + B(0, r) - k_0, \quad q = p - k_0.
\]
It follows that \( U \) is an open convex set containing \( 0 \) and \( q \notin U \). This is because \( p \notin K + B(0, r) \). Therefore, by Lemma \[2.2.3\], there exists \( f \in X' \) such that
\[
f(p-k_0) = f(q) > f(k+e-k_0)
\]
so
\[
f(p) > f(k) + f(e)
\]
(2.2.6)
for all \( k \in K \) and \( e \in B(0,r) \). If \( f(e) = 0 \) for all \( e \in B(0,r) \), then \( f = 0 \) and Lemma 2.2.6 could not hold. Therefore, \( f(\hat{e}) > 0 \) for some \( \hat{e} \in B(0,r) \) and so,

\[
 f(p) > f(k) + f(\hat{e})
\]

for all \( k \in K \). Let \( c_1 \equiv \sup \{ f(k) : k \in K \} \). Then for all \( k \in K \),

\[
 f(p) \geq c_1 + f(\hat{e}) > c_1 + \frac{f(\hat{e})}{2} > f(k).
\]

Let \( c = c_1 + \frac{f(\hat{e})}{2} \).

\[ K \]

\{ x : f(x) = c \} \quad \bullet p

2.3 Convex Functions

The following tells what is meant by a convex function.

**Definition 2.3.1** Let \( \phi : X \to (-\infty, \infty] \) be a function. It is convex if \( \phi(tx + (1 - t)y) \leq t\phi(x) + (1 - t)\phi(y) \) for all \( t \in [0, 1] \). It is proper if there exists \( x \) such that \( \phi(x) < \infty \). The set of all such \( x \) is called the domain of \( \phi \), denoted by \( D(\phi) \).

In this section, we consider the special case where the Banach space is just \( \mathbb{R}^n \). However, first is a general result valid on any normed linear space which is really surprising. It says that if \( \phi \) is bounded on an open set, then it is continuous on this open set. This also generalizes to locally convex topological vector space.

**Lemma 2.3.2** Suppose \( \phi \) is convex, and proper. Also suppose

\[
 \phi(y) < a
\]

for all \( y \in U \) where \( U \) is some open set containing \( x \). Then \( \phi \) is continuous at \( x \). In fact, if \( B(x,r) \subseteq U \), then for \( \varepsilon \in (0,1/2) \), and \( y \in \overline{B(x,\varepsilon r)} \),

\[
 2\varepsilon \phi(x) - 2\varepsilon a \leq \phi(y) - \phi(x) \leq 2\varepsilon a - 2\varepsilon \phi(x). 
\]

**Proof:** Let the open set be \( U \) and let \( x \in B(x,r) \subseteq U \). Let \( \varepsilon \in (0,1/2) \). If \( y \in \overline{B(x,\varepsilon r)} \), then

\[
 y = 2\varepsilon \left( \frac{y-x}{2\varepsilon} + x \right) + (1 - 2\varepsilon) x.
\]

and

\[
 y = 2\varepsilon \left( \frac{y-x}{2\varepsilon} + x \right) + (1 - 2\varepsilon) x.
\]
Therefore, if \( y \in B(x, \varepsilon r) \),
\[
\phi(y) \leq 2\varepsilon \phi \left( \frac{y - x}{2\varepsilon} + x \right) + (1 - 2\varepsilon) \phi(x) \tag{2.3.8}
\]
\[
\leq 2\varepsilon a + (1 - 2\varepsilon) \phi(x).
\]

Now it is also true that
\[
y = (1 + 2\varepsilon)x - 2\varepsilon \left( x + \frac{x - y}{2\varepsilon} \right)
\]
and so
\[
x = \frac{y}{1 + 2\varepsilon} + \frac{2\varepsilon}{1 + 2\varepsilon} \left( x + \frac{x - y}{2\varepsilon} \right)
\]
and, therefore,
\[
\phi(x) \leq \frac{\phi(y)}{1 + 2\varepsilon} + \frac{2\varepsilon}{1 + 2\varepsilon} \phi \left( x + \frac{x - y}{2\varepsilon} \right) \tag{2.3.9}
\]
\[
\leq \frac{\phi(y)}{1 + 2\varepsilon} + \frac{2\varepsilon a}{1 + 2\varepsilon}.
\]

By (2.3.8) and (2.3.9),
\[
(1 + 2\varepsilon) \phi(x) - 2a \varepsilon \leq \phi(y) \leq (1 - 2\varepsilon) \phi(x) + 2\varepsilon a
\]
and so
\[
2\varepsilon \phi(x) - 2a \varepsilon < \phi(y) - \phi(x) \leq 2a \varepsilon - 2\varepsilon \phi(x).
\]

Since \( \varepsilon \) is arbitrary, this proves (2.3.7) and the continuity of \( \phi \) at \( x \). \( \blacksquare \)

**Corollary 2.3.3** If \( \phi \) is convex and proper, defined on \( \mathbb{R}^n \), then \( \phi \) is continuous on the interior of \( D(\phi) \).

**Proof:** Let \( x \in \text{interior of } (D(\phi)) \). Then \( x \in B(x,r) \subseteq D(\phi) \) for some \( r > 0 \). Define \( v_i \equiv \delta e_i \) for \( i = 1, \ldots, n \), and \( v_0 \equiv 0 \). Let
\[
S \equiv \left\{ \sum_{i=0}^{n} \lambda_i v_i : \sum_{i=0}^{n} \lambda_i = 1 \text{ and } \lambda_i > 0 \text{ for all } i \right\}.
\]

Then if you have
\[
\sum_{i=0}^{n} \lambda_i v_i, \sum_{i=0}^{n} \mu_i v_i
\]
two points of \( S \), then the distance between them is
\[
\left| \sum_{i=0}^{n} (\lambda_i - \mu_i) v_i \right| \leq \left( \sum_{i=0}^{n} (\lambda_i - \mu_i)^2 \right)^{1/2} \left( \sum_{i=0}^{n} \delta_i^2 \right)^{1/2} \leq \delta (n + 1)
\]
Then $S$ is an open set of diameter less than $n\delta$ which contains the point

\[ \sum_{i=0}^{n} \frac{1}{1+n} v_i \equiv p. \]

Let $U \equiv x + (S - p)$. Then $U$ is an open set containing $x$ which has diameter less than $n\delta$ and so $x \in U \subseteq B(x,r)$ whenever $\delta$ is small enough. Also $x - p + v_i \in B(x,r)$ for each $i$ whenever $\delta$ is small enough. (Both $v_i$ and $p$ are multiples of $\delta$.) Let $\delta$ be sufficiently small. Then it follows that

\[
\phi \left( x - p + \sum_{i=0}^{n} \lambda_i v_i \right) = \phi \left( \sum_{i=0}^{n} \lambda_i (x - p + v_i) \right) \\
\leq \sum_{i=0}^{n} \lambda_i \phi (x - p + v_i) \\
\leq \max \left\{ \phi (x - p + v_i) \right\}_{i=0}^{n}.
\]

Therefore, $\phi$ is continuous by Lemma 2.3.2.

In particular, if $\phi$ is defined and finite on all of $\mathbb{R}^n$, then it is everywhere continuous.

### 2.4 Conjugate Functions

In this section, is a description of a way to obtain convex lower semicontinuous functions from functions which are not necessarily so. First is a definition. Let $X$ be a Banach space.

**Definition 2.4.1** A function $\phi : X \to (-\infty, \infty]$ is said to be lower semicontinuous, l.s.c. if $\text{epi}(\phi)$ is closed in $X \times (-\infty, \infty]$ where

\[ \text{epi}(\phi) \equiv \{(x, \alpha) : \alpha \geq \phi(x)\} \]

To say this is closed in $X \times (-\infty, \infty]$ means that if $(x_n, \alpha_n) \in \text{epi}(\phi)$ and $(x_n, \alpha_n) \to (x, \alpha)$, then $\alpha \geq \phi(x)$. If $\alpha$ is $\infty$, convergence of $\alpha_n$ to $\alpha$ is the usual thing. For every $m$, eventually, for $n$ large enough, $\alpha_n > m$.

**Definition 2.4.2** Let $\phi : X \to (-\infty, \infty]$ be some function, not necessarily convex but satisfying $\phi(y) < \infty$ for some $y \in X$. Define $\phi^* : X' \to (-\infty, \infty]$ by

\[ \phi^*(x^*) \equiv \sup \{ x^*(y) - \phi(y) : y \in X \}. \]

This function, $\phi^*$, defined above, is called the conjugate function of $\phi$ or the polar of $\phi$. Note that $\phi^*(x^*) > -\infty$ because $\phi(y) < \infty$ for some $y$.

**Theorem 2.4.3** Let $X$ be a real Banach space. Then $\phi^*$ is convex and l.s.c.
2.4. CONJUGATE FUNCTIONS

Proof: Let $\lambda \in [0,1]$. Then

\[
\phi^* (\lambda x^* + (1 - \lambda) y^*) = \sup \{ (\lambda x^* + (1 - \lambda) y^*) (y) - \phi (y) : y \in X \}
\]

\[
\sup \{ \lambda (x^* (y) - \phi (y)) + (1 - \lambda) (y^* (y) - \phi (y)) : y \in X \}
\]

\[
\leq \lambda \phi^* (x^*) + (1 - \lambda) \phi^* (y^*).
\]

It remains to show the function is l.s.c. Consider $f_y (x^*) \equiv x^* (y) - \phi (y)$. Then $f_y$ is obviously convex. Also to say that $(x, \alpha) \in \text{epi} (\phi^*)$ is to say that $\alpha \geq x^* (y) - \phi (y)$ for all $y$. Thus

\[
\text{epi} (\phi^*) = \cap_{y \in X} \text{epi} (f_y).
\]

Therefore, if $\text{epi} (f_y)$ is closed, this will prove the theorem. If $(x^*, a) \notin \text{epi} (f_y)$, then $a < x^* (y) - \phi (y)$ and, by continuity, for $b$ close enough to $a$ and $y^*$ close enough to $x^*$ then

\[
b < y^* (y) - \phi (y), \ (y^*, b) \notin \text{epi} (f_y)
\]

Thus $\text{epi} (f_y)$ is closed.

This theorem holds with no change in the proof if $X$ is only a locally convex topological vector space and $X'$ is given the weak * topology.

Definition 2.4.4 We define $\phi^{**}$ on $X$ by

\[
\phi^{**} (x) \equiv \sup \{ x^* (x) - \phi^* (x^*) : x^* \in X' \}.
\]

The following lemma comes from separation theorems. First is a simple observation.

Observation 2.4.5 $f \in (X \times \mathbb{R})'$ if and only if there exists $x^* \in X'$ and $\alpha \in \mathbb{R}$ such that $f (x, \lambda) = x^* (x) + \lambda \alpha$. To get $x^*$, you can simply define $x^* (x) \equiv f (x, 0)$ and to get $\alpha$ you just let $\alpha \lambda \equiv f (0, \lambda)$. Why does such an $\alpha$ exist? You know that $f (0, a \lambda + b \delta) = af (0, \alpha) + b f (0, \delta)$ and so in fact $\lambda \rightarrow f (0, \lambda)$ satisfies the Cauchy functional equation $g (x + y) = g (x) + g (y)$ and is continuous so there is only one thing it can be and that is $f (0, \lambda) = \alpha \lambda$ for some $\alpha$.

This picture illustrates the conclusion of the following lemma.

\[
\begin{cases}
epi(\phi) \\
(x_0, \beta) \\
\beta + (z^*, y - x_0) + \delta < \phi(y)
\end{cases}
\]

Lemma 2.4.6 Let $\phi : X \rightarrow (-\infty, \infty]$ be convex and lower semicontinuous and $\phi (x) < \infty$ for some $x$. (proper). Then if $\beta < \phi (x_0)$ so that $(x_0, \beta)$ is not in $\text{epi} (\phi)$, it follows that there exists $\delta > 0$ and $z^* \in X'$ such that for all $y$,

\[
z^* (y - x_0) + \beta + \delta < \phi (y), \ \text{all } y \in X
\]
CHAPTER 2. HAHN BANACH THEOREM, CONVEXITY

Proof: Let $C = \text{epi} (\phi) \cap (X \times \mathbb{R})$. Then $C$ is a closed convex nonempty set and it does not contain the point $(x_0, \beta)$. Let $\hat{\beta} > \beta$ be slightly larger so that also $(x_0, \hat{\beta}) \notin C$. Thus there exists $y^* \in X'$ and $\alpha \in \mathbb{R}$ such that for some $\hat{c}$, and all $y \in X$,

$$y^* (x_0) + \alpha \hat{\beta} > \hat{c} > y^* (y) + \alpha \phi (y)$$

for all $y \in X$. Now you can’t have $\alpha \geq 0$ because

$$\alpha \left( \hat{\beta} - \phi (y) \right) > y^* (y) - x_0$$

and you can let $y = x_0$ to have

$$\alpha < 0$$

Hence $\alpha < 0$ and so, dividing by it yields that for all $y \in X$,

$$x^* (x_0) + \hat{\beta} < c < x^* (y) + \phi (y)$$

where $x^* = y^*/\alpha, \hat{c}/\alpha \equiv c$. Then

$$(-x^*) (y - x_0) + \beta + \left( \hat{\beta} - \beta \right) < c - x^* (y) < \phi (y)$$

$$(-x^*) (y - x_0) + \beta + \delta < \phi (y), \delta \equiv \hat{\beta} - \beta$$

Let $z^* = -x^*$.

Theorem 2.4.7 $\phi^{**} (x) \leq \phi (x)$ for all $x$ and if $\phi$ is convex and l.s.c., $\phi^{**} (x) = \phi (x)$ for all $x \in X$.

Proof:

$$\phi^{**} (x) \equiv \sup \left\{ x^* (x) - \sup \{ x^* (y) - \phi (y) : y \in X \} : x^* \in X' \right\}$$

$$\leq \sup \{ x^* (x) - (x^* (x) - \phi (x)) \} = \phi (x).$$

Next suppose $\phi$ is convex and l.s.c. If $\phi^{**} (x_0) < \phi (x_0)$, then using Lemma 2.4.6, there exists $x_0^*, \delta > 0$ such that for all $y \in X$,

$$(x_0^*) (y - x_0) + \phi^{**} (x_0) + \delta < \phi (y)$$

$$x_0^* (y) - \phi (y) + \delta < x_0^* (x_0) - \phi^{**} (x_0)$$

Thus, since this holds for all $y$,

$$\phi^* (x_0^*) + \delta \leq x_0^* (x_0) - \phi^{**} (x_0)$$

$$\phi^{**} (x_0) + \delta \leq x_0^* (x_0) - \phi^* (x_0^*)$$
Then
\[ \phi^{**}(x_0) \equiv \sup \{ x^* (x_0) - \phi^* (x^*), x^* \in X' \} \geq x^*_0 (x_0) - \phi^* (x^*_0) \geq \phi^{**}(x_0) + \delta \]
a contradiction. \[\blacksquare\]

Note that the above shows that epi (\(\phi^{**}\)) \(\supseteq\) epi (\(\phi\)). Also note that epi (\(\phi^{**}\)) is closed. In general, when you do a * to something, you get the epigraph closed.

The following corollary is descriptive of the situation just discussed. It says that to find epi (\(\phi^{**}\)) it suffices to take the intersection of all closed convex sets which contain epi (\(\phi\)).

**Corollary 2.4.8** epi (\(\phi^{**}\)) is the smallest closed convex set containing epi (\(\phi\)).

**Proof:** epi (\(\phi^{**}\)) \(\supseteq\) epi (\(\phi\)) from Theorem 2.4.7. Also epi (\(\phi^{**}\)) is closed by the proof of Theorem 2.4.3. Suppose epi (\(\phi\)) \(\subseteq\) K \(\subseteq\) epi (\(\phi^{**}\)) and K is convex and closed. Let
\[ \psi (x) \equiv \min \{a : (x, a) \in K\} \]
\(\{a : (x, a) \in K\}\) is a closed subset of \((\infty, \infty]\) so the minimum exists.) \(\psi\) is also a convex function with epi (\(\psi\)) = K. To see \(\psi\) is convex, let \(\lambda \in [0, 1]\). Then, by the convexity of K,
\[ \lambda (x, \psi (x)) + (1 - \lambda) (y, \psi (y)) \]
\[= (\lambda x + (1 - \lambda) y, \lambda \psi (x) + (1 - \lambda) \psi (y)) \in K.\]
It follows from the definition of \(\psi\) that
\[ \psi (\lambda x + (1 - \lambda) y) \leq \lambda \psi (x) + (1 - \lambda) \psi (y).\]
Then
\[ \phi^{**} \leq \psi \leq \phi \]
and so from the definitions,
\[ \phi^{***} \leq \psi^* \leq \phi^* \]
That is, putting a * on the functions on either side of an inequality turns around the inequality. Thus from the definitions and Theorem 2.4.7,
\[ \phi^{**} = \phi^{****} \leq \psi^{**} = \psi \leq \phi^{**}.\]
Therefore, \(\psi = \phi^{**}\) and epi (\(\phi^{**}\)) is the smallest closed convex set containing epi (\(\phi\)) as claimed. \[\blacksquare\]

**Notation 2.4.9** For \(x^* \in X'\) and \(x \in X\), it is often the case that people write \(\langle x^*, x \rangle\) to indicate \(x^* (x)\).
CHAPTER 2. HAHN BANACH THEOREM, CONVEXITY

2.5 Exercises

1. Let \( J : X \to X'' \) be defined by \( Jx(f) \equiv f(x) \). Show that \( \|Jx\|_{X''} = \|x\|_X \).

2. A fundamental consideration is the notion of uniform convexity of the norm. A norm on a Banach space \( X \) is said to be uniformly convex if the following condition holds. If

\[
\|x_n\|,\|y_n\| \leq 1 \quad \text{and} \quad \|x_n + y_n\| \to 2, \quad \text{then} \quad \|x_n - y_n\| \to 0.
\]

an amazing result which is of great importance is the following. If a Banach space has a uniformly convex norm and if \( x_n \to x \) weakly and \( \|x_n\| \to \|x\| \), then \( \|x_n - x\| \to 0 \). Thus weak convergence and convergence of the norms implies strong convergence.

3. Let \( A \in \mathcal{L}(X,Y) \). That is, \( A \) is a continuous linear mapping from \( X \) to \( Y \) where \( X,Y \) are Banach spaces. We define \( A^* : Y' \to X' \) by

\[
A^∗y^*(x) = y^*(Ax)
\]

Show that \( A^* \) is continuous, linear, and that \( \|A^*\| = \|A\| \). This is called the adjoint map.

4. We say \( \phi : X \to (-\infty, \infty] \) is convex if \( \phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y) \).

It is lower semicontinuous if

\[
\text{epi} \phi \equiv \{(x,\alpha) : \alpha \geq \phi(x)\}
\]

is closed in the topological space \( X \times \mathbb{R} \) with the product topology. Also assume that \( \phi \) is proper which means that \( \phi(x) < \infty \) for some \( x \). The set of such \( x \) is called the effective domain of \( \phi, D(\phi) \). Show that if \( \beta < \phi(x_0) \) so that \((x_0,\beta) \notin \text{epi} \phi\), then there exists \( \gamma \in \mathbb{R} \) and \( x^* \in X' \) such that for all \( x \in X \),

\[
\phi(x) > \gamma + \langle x^*, x \rangle, \quad \gamma + \langle x^*, x_0 \rangle > \beta
\]

You might want to draw a picture to illustrate this claim. Go over the details of this. It is proved in the chapter.

5. Suppose \( K \) is a convex set in a Banach space \( X \). Let \( K_0 \) denote its interior consisting of the union of all open subsets of \( K \). Show that \( K_0 \) is also convex.

Hint: Consider \( x, y \in K_0 \) so there is \( r > 0 \) such that \( B(x,r), B(y,r) \) are both in \( K_0 \). Consider

\[
U \equiv \bigcup_{t \in [0,1]} tB(x,r) + (1-t)B(y,r)
\]

and argue that this is an open set in \( K \) which contains \( tx + (1-t)y \).
6. Suppose $K$ is a convex set in a Banach space $X$ which has nonempty interior. Let $K_0$ denote its interior consisting of the union of all open subsets of $K$. Show that

$$K = K_0$$

That is, the closure of $K$ is the same as the closure of the interior of $K$.

**Hint:** Consider this picture. Show that the cone $C$ is open. This cone is of the form $C = \bigcup_{t \in [0,1), x \in B(x,r)} ((1-t) \hat{x} + ty)$. Note how it does not include $y$, the point of the cone. However, $y \in K$.

7. Suppose $\phi : X \to (-\infty, \infty]$ is convex, lower semicontinuous, and proper. Let $G_0$ denote the interior of $\text{epi} (\phi) \cap X \times \mathbb{R}$. Suppose also that for some $x_0 \in D(\phi)$, whenever $x_n \to x_0$,

$$\phi(x_0) \geq \lim inf_{n \to \infty} \phi(x_n).$$

Then $G_0 \neq \emptyset$. In fact if $\alpha > \phi(x_0)$, show $(x_0, \alpha) \in G_0$. Recall that $D(\phi) \equiv \{ x \in X : \phi(x) < \infty \}$. **Hint:** If this is not so, then there would be $(x_n, \alpha_n) \to (x, \alpha)$ but $\alpha_n \leq \phi(x_n)$. Now get a contradiction.

8. Suppose $G_0$ is the interior of $G \equiv \text{epi}(\phi) \cap (X \times \mathbb{R})$ and that $G_0 \neq \emptyset$ and let $x \in D(\phi)$. Show that there exists $y^* \in X'$ such that for all $y \in X$,

$$y^*(y-x) = \langle y^*, y-x \rangle \leq \phi(y) - \phi(x).$$

This $y^*$ is called a subgradient. We say $y^* \in \partial\phi(x)$ if the above inequality holds for all $y \in X$. Note that from Problem 6, if there is any point $x_0 \in D(\phi)$ for which $\phi(x_0) \geq \lim inf_{n \to \infty} \phi(x_n)$ whenever $x_n \to x_0$, then it follows that $\partial\phi(x) \neq \emptyset$ for all $x \in D(\phi)$. Also, you can show that the condition that $\text{epi} \phi$ is closed is equivalent to saying that whenever $x_n \to x, \phi(x_0) \leq \lim inf_{n \to \infty} \phi(x_n)$. Thus this is an incredible result. Knowledge of a limit
property at a single point of $D(\phi)$ is sufficient to yield the existence of a subgradient at every point of $D(\phi)$. Also from Lemma 2.3.2, if $\phi$ is bounded on any nonempty open set, this condition holds. **Hint:** You should be using Problems 6, 8, and Corollary 2.2.4.

9. Suppose $Y$ is a closed subspace of $X$ a real Banach space, and suppose that $x_0 \notin Y$. Show that there exists $f \in X'$ such that $f(x_0) \neq 0$ but $f(Y) = 0$.

10. Let $X$ be a Banach space. When $J : X \to X''$ is onto, the space $X$ is called reflexive. If $Y$ is a closed subspace of $X$, and $X$ is reflexive, show that $Y$ is also reflexive. **Hint:** You might consider the following diagram.

\[
\begin{align*}
Y'' & \xrightarrow{i}^* X'' \\
Y' & \xrightarrow{i \text{ onto}} X' \\
Y & \to X
\end{align*}
\]

11. Suppose $\phi : X \to (-\infty, \infty]$ is convex and lower semicontinuous. Recall that this meant that $G \equiv \text{epi}(\phi) \cap X \times \mathbb{R}$ is a closed convex subset of $X \times \mathbb{R}$. Show, using separation theorems that $G$ is also closed in the weak topology of $X \times \mathbb{R}$. Thus $G$ is weakly lower semicontinuous.

12. You can also separate two disjoint convex sets. Let $A$ and $B$ be disjoint, convex and nonempty sets with $B$ open. Then there exists $f \in X'$ such that

$$f(a) < f(b)$$

for all $a \in A$ and $b \in B$. **Hint:** Note that $B - A + a_0 - b_0$ is open and contains 0. Now apply the separation theorem Lemma 2.2.3.

13. Suppose you have a Banach space $X$ which has a countable subset $D_0$ such that every weakly open subset of $X$ has a point of $D_0$. Now let $D$ consist of all linear combinations of $D_0$ where the scalars are rational. Show that $D$ must also be strongly dense.

14. Problem 36 on Page 50 was the Eberlein Smulian theorem for the case of $X$ reflexive and $X'$ separable. That is, bounded sequences in $X$ have weakly convergent subsequences. Show that no assumption of separability of $X'$ needs to be made. **Hint:** Consider the following diagram.

\[
\begin{align*}
B^{**} & \xrightarrow{i**} Y'' & \xrightarrow{j}^1 X'' \\
\text{weakly separable} & B^* & \xrightarrow{i^* \text{ onto}} Y' \\
\text{separable} & B & \to Y & \to X
\end{align*}
\]

In this diagram, you have a sequence $\{x_n\}$ in the closed unit ball of $X$. (It suffices to assume this.) The space $Y$ is the closure of the span of the $\{x_n\}$.
Therefore, it is separable. Apply the Banach Alaoglu theorem to \( B^* \) to conclude it is weak \(^*\) compact. Then since \( Y \) is separable, the weak \(^*\) topology on \( B^* \) comes from a suitable metric, Problem 31 on Page 43 or 38 on Page 45. Therefore, \( B^* \) and hence \( Y' \) is a separable metric space with a countable dense set \( D \). This is in the weak topology. By the above problem, this set \( D \) is also dense in \( Y' \) so \( Y' \) is separable. It follows that there is a subsequence \( \{x_{n_k}\} \) which converges weakly to some \( x \) in \( Y \), this by Problem 40 on Page 52. Finish the argument.

15. A sequence of \( \{x_n\} \) is said to be weakly bounded if for every \( f \in X' \), \( \sup_n |f(x_n)| < \infty \). Show that if a sequence is weakly bounded, then it is actually bounded with respect to the norm. **Hint:** Use Problem 33 on Page 45 and what was shown above that \( J \) preserves norms.

16. The Riesz representation theorem says that if you have \( (\Omega, \mathcal{F}, \mu) \) a measure space, then \( (L^p(\Omega))^\prime = L^{p'}(\Omega) \) in the sense that if \( g \in (L^p(\Omega))^\prime \) then there is a unique \( \hat{g} \in L^{p'}(\Omega) \) such that \( g(f) = \int_\Omega \hat{g} f d\mu \). You can consider real spaces here. Show that \( L^p \) is a reflexive Banach space, \( p > 1 \). You know it is a Banach space from usual considerations. Show it is reflexive.
Chapter 3

Approximation Of Functions

3.1 Introduction

This is on the Stone Weierstrass theorem, one of the major results on approximation. It gives a very nice example of the usefulness of the definition of a locally compact Hausdorff space. This theorem is just a very useful one to have on hand. It will be nice to have when we discuss the Brouwer degree for example. First we give a simple theorem on approximation by polynomials. Then follows a profound generalization due to Stone.

3.2 The Bernstein Polynomials

This short chapter is on the important Weierstrass approximation theorem. It is about approximating an arbitrary continuous function uniformly by a polynomial. It will be assumed first that the polynomials are real valued with real coefficients. First here is some notation.

The following estimate will be the basis for the Weierstrass approximation theorem. It is actually a statement about the variance of a binomial random variable.

**Lemma 3.2.1** The following estimate holds for \( x \in [0, 1] \).

\[
\sum_{k=0}^{m} \binom{m}{k} (k - mx)^2 x^k (1 - x)^{m-k} \leq \frac{1}{4} m
\]

**Proof:** By the Binomial theorem,

\[
\sum_{k=0}^{m} \binom{m}{k} \left(e^t x\right)^k (1 - x)^{m-k} = (1 - x + e^t x)^m. \tag{3.2.1}
\]

Differentiating both sides with respect to \( t \) and then evaluating at \( t = 0 \) yields

\[
\sum_{k=0}^{m} \binom{m}{k} k x^k (1 - x)^{m-k} = m x.
\]
Now doing two derivatives of \( e^t \) with respect to \( t \) yields

\[
\sum_{k=0}^{m} \binom{m}{k} k^2 (e^t)^k (1-x)^{m-k} = m(m-1)(1-x+e^t x)^{m-2} e^{2t}x^2 \\
+ m(1-x+e^t x)^{m-1} xe^t.
\]

Evaluating this at \( t = 0 \),

\[
\sum_{k=0}^{m} \binom{m}{k} k^2 (x)^k (1-x)^{m-k} = m(m-1)x^2 + mx.
\]

Therefore,

\[
\sum_{k=0}^{m} \binom{m}{k} (k-mx)^2 x^k (1-x)^{m-k} = m(m-1)x^2 + mx - 2m^2x^2 + m^2x^2 \\
= m(x-x^2) \leq \frac{1}{4}m. \blacksquare
\]

**Theorem 3.2.2** Let \( f \) be continuous on \( [a,b] \). Then there exists a sequence of polynomials converging uniformly to \( f \) on \( [a,b] \).

**Proof:** First assume \( [a,b] = [0,1] \). Then let

\[
p_n(x) \equiv \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.
\]

Then

\[
|f(x) - p_n(x)| \leq \sum_{k=0}^{n} \binom{n}{k} \left|f\left(\frac{k}{n}\right) - f(x)\right| x^k (1-x)^{n-k}
\]

Let \( \delta > 0 \) be such that if \( |x-y| < \delta \), then \( |f(x) - f(y)| < \varepsilon /2 \).

\[
|f(x) - p_n(x)| \leq \sum_{k:(k/n)-x<\delta} \binom{n}{k} \left|f\left(\frac{k}{n}\right) - f(x)\right| x^k (1-x)^{n-k} \\
+ \sum_{k:(k/n)-x\geq\delta} \binom{n}{k} \left|f\left(\frac{k}{n}\right) - f(x)\right| x^k (1-x)^{n-k}
\]

and so

\[
|f(x) - p_n(x)| \leq \sum_{k=0}^{n} \binom{n}{k} \frac{\varepsilon}{2} x^k (1-x)^{n-k} \\
+ 2\|f\|_\infty \sum_{k:(k/n)-x\geq\delta} \binom{n}{k} x^k (1-x)^{n-k}
\]
3.3. STONE WEIERSTRASS THEOREM

where \( \|f\|_\infty \equiv \max \{ |f(x)| : x \in [0, 1] \} \). The second sum is over \( k \) such that \((k - nx)^2 \geq n^2 \delta^2\). Thus

\[
|f(x) - p_n(x)| \leq \frac{\varepsilon}{2} + \frac{1}{n^2 \delta^2} \|f\|_\infty \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) (k - nx)^2 \frac{(1 - x)^{n-k}}{4^n} \leq \frac{\varepsilon}{2} + \frac{1}{n^2 \delta^2} \|f\|_\infty \frac{1}{2}^n
\]

Thus whenever \( n \) is large enough, \( \|f - p_n\|_\infty \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \) and so convergence is uniform.

As to an arbitrary interval, let \( l : [0, 1] \rightarrow [a, b] \) be one to one onto and linear. Then if \( f \) is continuous on \([a, b]\), let \( g = f \circ l \). Then \( g \) is continuous on \([0, 1]\) and so there are polynomials \( p_n \) converging uniformly to \( g \) on \([0, 1]\). Then \( p_n \circ l^{-1} \) is a sequence of polynomials which converges uniformly to \( f \) on \([a, b]\).

Here is some notation which tells what is meant by polynomials as functions of many variables.

**Definition 3.2.3** \( \alpha = (\alpha_1, \cdots, \alpha_n) \) for \( \alpha_1 \cdots \alpha_n \) positive integers is called a multi-index. For \( \alpha \) a multi-index, \(|\alpha| \equiv \alpha_1 + \cdots + \alpha_n \) and if \( x \in \mathbb{R}^n \),

\[
x = (x_1, \cdots, x_n),
\]

and \( f \) a function, define

\[
x^\alpha \equiv x_1^{\alpha_1}x_2^{\alpha_2} \cdots x_n^{\alpha_n}.
\]

A polynomial in \( n \) variables of degree \( m \) is a function of the form

\[
p(x) = \sum_{|\alpha| \leq m} a_\alpha x^\alpha.
\]

Here \( \alpha \) is a multi-index as just described.

One can prove the following theorem which will be left for the exercises. It will come from the general Stone Weierstrass theorem presented in the next section.

**Theorem 3.2.4** Let \( K \) be a compact set in \( \mathbb{R}^n \) and let \( f \) be a continuous function defined on \( K \). Then there exists a sequence of polynomials \( \{p_m\} \) converging uniformly to \( f \) on \( K \).

3.3 Stone Weierstrass Theorem

3.3.1 The Case Of Compact Sets

There is a profound generalization of the Weierstrass approximation theorem due to Stone.

**Definition 3.3.1** \( \mathcal{A} \) is an algebra of functions if \( \mathcal{A} \) is a vector space and if whenever \( f, g \in \mathcal{A} \) then \( fg \in \mathcal{A} \).
The next result is the key to the profound generalization of the Weierstrass theorem due to Stone in which an interval will be replaced by a compact or locally compact set and polynomials will be replaced with elements of an algebra satisfying certain axioms.

**Corollary 3.3.2** On the interval \([-M, M]\), there exist polynomials \(p_n\) such that
\[
p_n(0) = 0
\]
and
\[
\lim_{n \to \infty} ||p_n - \cdot||_\infty = 0.
\]

**Proof:** By Theorem 3.2.2 there exists a sequence of polynomials, \(\{\tilde{p}_n\}\) such that \(\tilde{p}_n \to ||\cdot||\) uniformly. Then let \(p_n(t) \equiv \tilde{p}_n(t) - \tilde{p}_n(0)\).

**Definition 3.3.3** An algebra of functions, \(A\) defined on \(A\), annihilates no point of \(A\) if for all \(x \in A\), there exists \(g \in A\) such that \(g(x) \neq 0\). The algebra separates points if whenever \(x_1 \neq x_2\), then there exists \(g \in A\) such that \(g(x_1) \neq g(x_2)\).

The following generalization is known as the Stone Weierstrass approximation theorem.

**Theorem 3.3.4** Let \(A\) be a compact topological space and let \(A \subseteq C(A; \mathbb{R})\) be an algebra of functions which separates points and annihilates no point. Then \(A\) is dense in \(C(A; \mathbb{R})\).

**Proof:** First here is a lemma.

**Lemma 3.3.5** Let \(c_1\) and \(c_2\) be two real numbers and let \(x_1 \neq x_2\) be two points of \(A\). Then there exists a function \(f_{x_1x_2}\) such that
\[
f_{x_1x_2}(x_1) = c_1, \quad f_{x_1x_2}(x_2) = c_2.
\]

**Proof of the lemma:** Let \(g \in A\) satisfy
\[
g(x_1) \neq g(x_2).
\]
Such a \(g\) exists because the algebra separates points. Since the algebra annihilates no point, there exist functions \(h\) and \(k\) such that
\[
h(x_1) \neq 0, \quad k(x_2) \neq 0.
\]
Then let
\[
u \equiv gh - g(x_2) h, \quad v \equiv gk - g(x_1) k.
\]
It follows that \(u(x_1) \neq 0\) and \(u(x_2) = 0\) while \(v(x_2) \neq 0\) and \(v(x_1) = 0\). Let
\[
f_{x_1x_2} \equiv \frac{c_1 u}{u(x_1)} + \frac{c_2 v}{v(x_2)}.
\]
Now continue the proof of Theorem 3.3.4.

First note that $\mathcal{A}$ satisfies the same axioms as $\mathcal{A}$ but in addition to these axioms, $\mathcal{A}$ is closed. The closure of $\mathcal{A}$ is taken with respect to the usual norm on $C(A)$,

$$\|f\|_\infty = \max \{|f(x)| : x \in A\}.$$ 

Suppose $f \in \mathcal{A}$ and suppose $M$ is large enough that $\|f\|_\infty < M$.

Using Corollary 3.3.2, let $p_n$ be a sequence of polynomials such that

$$\|p_n - f\|_\infty \to 0, \quad p_n(0) = 0.$$ 

It follows that $p_n \circ f \in \mathcal{A}$ and so $|f| \in \mathcal{A}$ whenever $f \in \mathcal{A}$. Also note that

$$\max(f, g) = \frac{|f - g| + (f + g)}{2}$$

$$\min(f, g) = \frac{(f + g) - |f - g|}{2}.$$ 

Therefore, this shows that if $f, g \in \mathcal{A}$ then

$$\max(f, g), \min(f, g) \in \mathcal{A}.$$ 

By induction, if $f_i, i = 1, 2, \ldots, m$ are in $\mathcal{A}$ then

$$\max(f_i, i = 1, 2, \ldots, m), \quad \min(f_i, i = 1, 2, \ldots, m) \in \mathcal{A}.$$ 

Now let $h \in C(A; \mathbb{R})$ and let $x \in A$. Use Lemma 3.3.5 to obtain $f_{xy}$, a function of $\mathcal{A}$ which agrees with $h$ at $x$ and $y$. Letting $\varepsilon > 0$, there exists an open set $U(y)$ containing $y$ such that

$$f_{xy}(z) > h(z) - \varepsilon \quad \text{if} \quad z \in U(y).$$ 

Since $A$ is compact, let $U(y_1), \ldots, U(y_l)$ cover $A$. Let

$$f_x \equiv \max(f_{xy_1}, f_{xy_2}, \ldots, f_{xy_l}).$$ 

Then $f_x \in \mathcal{A}$ and

$$f_x(z) > h(z) - \varepsilon$$ 

for all $z \in A$ and $f_x(x) = h(x)$. This implies that for each $x \in A$ there exists an open set $V(x)$ containing $x$ such that for $z \in V(x)$,

$$f_x(z) < h(z) + \varepsilon.$$ 

Let $V(x_1), \ldots, V(x_m)$ cover $A$ and let

$$f \equiv \min(f_{x_1}, \ldots, f_{x_m}).$$
Therefore, 
\[ f(z) < h(z) + \varepsilon \]
for all \( z \in A \) and since \( f_x(z) > h(z) - \varepsilon \) for all \( z \in A \), it follows
\[ f(z) > h(z) - \varepsilon \]
also and so
\[ |f(z) - h(z)| < \varepsilon \]
for all \( z \). Since \( \varepsilon \) is arbitrary, this shows \( h \in \overline{A} \) and proves \( \overline{A} = C_0(A;\mathbb{R}) \).

3.3.2 The Case Of Locally Compact Sets

Next we give a generalization to locally compact Hausdorff space. Of course the thing of most interest is \( \mathbb{R}^n \).

**Definition 3.3.6** Let \((X,\tau)\) be a locally compact Hausdorff space. \( C_0(X) \) denotes the space of real or complex valued continuous functions defined on \( X \) with the property that if \( f \in C_0(X) \), then for each \( \varepsilon > 0 \) there exists a compact set \( K \) such that \( |f(x)| < \varepsilon \) for all \( x \notin K \). Define
\[ \|f\|_{\infty} = \sup \{|f(x)| : x \in X\}. \]

**Lemma 3.3.7** For \((X,\tau)\) a locally compact Hausdorff space with the above norm, \( C_0(X) \) is a complete space.

**Proof:** Let \((\overline{X},\overline{\tau})\) be the one point compactification described in Lemma 1.10.20.

\[ D \equiv \{ f \in C(\overline{X}) : f(\infty) = 0 \}. \]

Then \( D \) is a closed subspace of \( C(\overline{X}) \). For \( f \in C_0(X) \),
\[ \tilde{f}(x) \equiv \begin{cases} f(x) & \text{if } x \in X \\ 0 & \text{if } x = \infty \end{cases} \]

and let \( \theta : C_0(X) \to D \) be given by \( \theta f = \tilde{f} \). Then \( \theta \) is one to one and onto and also satisfies \( \|f\|_{\infty} = \||\theta f||_{\infty} \). Now \( D \) is complete because it is a closed subspace of a complete space and so \( C_0(X) \) with \( ||\cdot||_{\infty} \) is also complete. 

The above refers to functions which have values in \( \mathbb{C} \) but the same proof works for functions which have values in any complete normed linear space.

In the case where the functions in \( C_0(X) \) all have real values, I will denote the resulting space by \( C_0(X;\mathbb{R}) \) with similar meanings in other cases.

With this lemma, the generalization of the Stone Weierstrass theorem to locally compact sets is as follows.
Theorem 3.3.8 Let \( \mathcal{A} \) be an algebra of functions in \( C_0(X; \mathbb{R}) \) where \((X, \tau)\) is a locally compact Hausdorff space which separates the points and annihilates no point. Then \( \mathcal{A} \) is dense in \( C_0(X; \mathbb{R}) \).

**Proof:** Let \( \left( \hat{X}, \hat{\tau} \right) \) be the one point compactification as described in Lemma 1.10.20. Let \( \hat{\mathcal{A}} \) denote all finite linear combinations of the form

\[
\left\{ \sum_{i=1}^{n} c_i \tilde{f}_i + c_0 : f \in \mathcal{A}, \ c_i \in \mathbb{R} \right\}
\]

where for \( f \in C_0(X; \mathbb{R}) \),

\[
\tilde{f}(x) \equiv \begin{cases} f(x) & \text{if } x \in X \\ 0 & \text{if } x = \infty \end{cases}
\]

Then \( \hat{\mathcal{A}} \) is obviously an algebra of functions in \( C \left( \hat{X}; \mathbb{R} \right) \). It separates points because this is true of \( \mathcal{A} \). Similarly, it annihilates no point because of the inclusion of \( c_0 \) an arbitrary element of \( \mathbb{R} \) in the definition above. Therefore from Theorem 3.3.4, \( \hat{\mathcal{A}} \) is dense in \( C \left( \hat{X}; \mathbb{R} \right) \). Letting \( f \in C_0(X; \mathbb{R}) \), it follows \( \tilde{f} \in C \left( \hat{X}; \mathbb{R} \right) \) and so there exists a sequence \( \{h_n\} \subseteq \hat{\mathcal{A}} \) such that \( h_n \) converges uniformly to \( f \). Now \( h_n \) is of the form \( \sum_{i=1}^{n} c_i \tilde{f}_i + c_0 \) and since \( \tilde{f}(\infty) = 0 \), you can take each \( c_0^n = 0 \) and so this has shown the existence of a sequence of functions in \( \mathcal{A} \) such that it converges uniformly to \( f \).

3.3.3 The Case Of Complex Valued Functions

What about the general case where \( C_0(X) \) consists of complex valued functions and the field of scalars is \( \mathbb{C} \) rather than \( \mathbb{R} \)? The following is the version of the Stone Weierstrass theorem which applies to this case. You have to assume that for \( f \in \mathcal{A} \) it follows \( \overline{f} \in \mathcal{A} \). Such an algebra is called self adjoint.

**Theorem 3.3.9** Suppose \( \mathcal{A} \) is an algebra of functions in \( C_0(X) \), where \( X \) is a locally compact Hausdorff space, which separates the points, annihilates no point, and has the property that if \( f \in \mathcal{A} \), then \( \overline{f} \in \mathcal{A} \). Then \( \mathcal{A} \) is dense in \( C_0(X) \).

**Proof:** Let \( \text{Re} \mathcal{A} \equiv \{ \text{Re} f : f \in \mathcal{A} \} \), \( \text{Im} \mathcal{A} \equiv \{ \text{Im} f : f \in \mathcal{A} \} \). First I will show that \( \mathcal{A} = \text{Re} \mathcal{A} + i \text{ Im} \mathcal{A} = \text{Im} \mathcal{A} + i \text{ Re} \mathcal{A} \). Let \( f \in \mathcal{A} \). Then

\[
f = \frac{1}{2} \left( f + \overline{f} \right) + \frac{1}{2} \left( f - \overline{f} \right) = \text{Re} f + i \text{ Im} f \in \text{Re} \mathcal{A} + i \text{ Im} \mathcal{A}
\]

and so \( \mathcal{A} \subseteq \text{Re} \mathcal{A} + i \text{ Im} \mathcal{A} \). Also

\[
f = \frac{1}{2i} \left( if + i\overline{f} \right) - \frac{i}{2} \left( if + i\overline{f} \right) = \text{Im} (if) + i \text{ Re} (if) \in \text{Im} \mathcal{A} + i \text{ Re} \mathcal{A}
\]
This proves one half of the desired equality. Now suppose \( h \in \Re \mathcal{A} + i \Im \mathcal{A} \). Then \( h = \Re g_1 + i \Im g_2 \) where \( g_i \in \mathcal{A} \). Then since \( \Re g_1 = \frac{1}{2} (g_1 + \overline{g_1}) \), it follows \( \Re g_1 \in \mathcal{A} \). Similarly \( \Im g_2 \in \mathcal{A} \). Therefore, \( h \in \mathcal{A} \). The case where \( h \in \Im \mathcal{A} + i \Re \mathcal{A} \) is similar. This establishes the desired equality.

Now \( \Re \mathcal{A} \) and \( \Im \mathcal{A} \) are both real algebras. I will show this now. First consider \( \Im \mathcal{A} \). It is obvious this is a real vector space. It only remains to verify that the product of two functions in \( \Im \mathcal{A} \) is in \( \Im \mathcal{A} \). Note that from the first part, \( \Re \mathcal{A}, \Im \mathcal{A} \) are both subsets of \( \mathcal{A} \) because, for example, if \( u \in \Im \mathcal{A} \) then \( u + 0 \in \Im \mathcal{A} + i \Re \mathcal{A} = \mathcal{A} \). Therefore, if \( v, w \in \Im \mathcal{A} \), both \( iv \) and \( ivw \) are in \( \mathcal{A} \) and so \( \Im (ivw) = vw \) and \( ivw \in \mathcal{A} \). Similarly, \( \Re \mathcal{A} \) is an algebra.

Both \( \Re \mathcal{A} \) and \( \Im \mathcal{A} \) must separate the points. Here is why: If \( x_1 \neq x_2 \), then there exists \( f \in \mathcal{A} \) such that \( f (x_1) \neq f (x_2) \). If \( \Im f (x_1) \neq \Im f (x_2) \), this shows there is a function in \( \Im \mathcal{A}, \Im f \) which separates these two points. If \( \Im f \) fails to separate the two points, then \( \Re f \) must separate the points and so you could consider \( \Im (if) \) to get a function in \( \Im \mathcal{A} \) which separates these points. This shows \( \Im \mathcal{A} \) separates the points. Similarly \( \Re \mathcal{A} \) separates the points.

Neither \( \Re \mathcal{A} \) nor \( \Im \mathcal{A} \) annihilate any point. This is easy to see because if \( x \) is a point there exists \( f \in \mathcal{A} \) such that \( f (x) \neq 0 \). Thus either \( \Re f (x) \neq 0 \) or \( \Im f (x) \neq 0 \). If \( \Im f (x) \neq 0 \), this shows this point is not annihilated by \( \Im \mathcal{A} \). If \( \Im f (x) = 0 \), consider \( \Im (if) (x) = \Re f (x) \neq 0 \). Similarly, \( \Re \mathcal{A} \) does not annihilate any point.

It follows from Theorem \( \Re \mathcal{A} \) and \( \Im \mathcal{A} \) are dense in the real valued functions of \( C_0 (X) \). Let \( f \in C_0 (X) \). Then there exists \( \{h_n\} \subseteq \Re \mathcal{A} \) and \( \{g_n\} \subseteq \Im \mathcal{A} \) such that \( h_n \rightarrow \Re f \) uniformly and \( g_n \rightarrow \Im f \) uniformly. Therefore, \( h_n + ig_n \in \mathcal{A} \) and it converges to \( f \) uniformly. \( \blacksquare \)

### 3.4 The Holder Spaces

We consider these spaces as spaces of functions defined on an interval \([0, 1]\) although one could have \([0, T]\) just as easily. A slightly more general version is in the exercises. They are a very interesting example of spaces which are not separable.

**Definition 3.4.1** Let \( p > 1 \). Then \( f \in C^{1/p} ([0, 1]) \) means that \( f \in C ([0, 1]) \) and also

\[
\rho_p (f) \equiv \sup \{ \frac{|f (x) - f (y)|}{|x - y|^{1/p}} : x, y \in X, x \neq y \} < \infty
\]

Then the norm is defined as \( \|f\|_{C([0, 1])} + \rho_p (f) \equiv \|f\|_1^{1/p} \).

We leave it as an exercise to verify that \( C^{1/p} ([0, 1]) \) is a complete normed linear space.

Let \( p > 1 \). Then \( C^{1/p} ([0, 1]) \) is not separable. Define uncountably many functions, one for each \( \varepsilon \) where \( \varepsilon \) is a sequence of \(-1\) and \(1\). Thus \( \varepsilon_k \in \{-1, 1\} \). Thus \( \varepsilon \neq \varepsilon' \) if the two sequences differ in at least one slot, one giving \( 1 \) and the other
3.4. THE HOLDER SPACES

equaling $-1$. Now define

$$f_{\varepsilon}(t) \equiv \sum_{k=1}^{\infty} \varepsilon_k 2^{-k/p} \sin(2^k \pi t)$$

Then this is $1/p$ Holder. Let $s < t$.

$$|f_{\varepsilon}(t) - f_{\varepsilon}(s)| \leq \sum_{k \leq |\log_2 (t-s)|} |2^{-k/p} \sin(2^k \pi t) - 2^{-k/p} \sin(2^k \pi s)|
+ \sum_{k > |\log_2 (t-s)|} |2^{-k/p} \sin(2^k \pi t) - 2^{-k/p} \sin(2^k \pi s)|$$

If $t = 1$ and $s = 0$, there is really nothing to show because then the difference equals 0. There is also nothing to show if $t = s$. From now on, $0 < t - s < 1$. Let $k_0$ be the largest integer which is less than or equal to $|\log_2 (t-s)| = \log_2 (t-s)$. Note that $-\log (t-s) > 0$ because $0 < t - s < 1$. Then

$$|f_{\varepsilon}(t) - f_{\varepsilon}(s)| \leq \sum_{k \leq k_0} |2^{-k/p} \sin(2^k \pi t) - 2^{-k/p} \sin(2^k \pi s)|
+ \sum_{k > k_0} |2^{-k/p} \sin(2^k \pi t) - 2^{-k/p} \sin(2^k \pi s)|$$

Now $k_0 \leq -\log_2 (t-s) < k_0 + 1$ and so $-k_0 \geq \log_2 (t-s) \geq -(k_0 + 1)$. Hence

$$2^{-k_0} \geq |t-s| \geq 2^{-k_0} 2^{-1}$$

and so

$$2^{-k_0/p} \geq |t-s|^{1/p} \geq 2^{-k_0/p} 2^{-1/p}$$

Using this in the sums,

$$|f_{\varepsilon}(t) - f_{\varepsilon}(s)| \leq |t-s| C_p + \sum_{k > k_0} 2^{-k/p} 2^{k_0/p} 2^{-k_0/p} 2^{-1/p}$$

$$\leq |t-s| C_p + \sum_{k > k_0} 2^{-k/p} 2^{k_0/p} \left(2^{1/p} |t-s|^{1/p} \right)^2$$

$$\leq |t-s| C_p + \sum_{k > k_0} 2^{-(k-k_0)/p} \left(2^{1/p} |t-s|^{1/p} \right)^2$$

$$\leq |t-s| C_p + \left(2^{1+1/p} \right) \sum_{k=1}^{\infty} 2^{-k/p} |t-s|^{1/p}$$

$$= |t-s| C_p + D_p |t-s|^{1/p} \leq C_p |t-s|^{1/p} + D_p |t-s|^{1/p}$$
Thus \( f_\epsilon \) is indeed \( 1/p \) Holder continuous.

Now consider \( \epsilon \neq \epsilon' \). Suppose the first discrepancy in the two sequences occurs with \( \epsilon_j \). Thus one is 1 and the other is \(-1\). Let \( t = \frac{i+1}{2^{j+1}}, s = \frac{i}{2^{j+1}} \)

\[
|f_\epsilon (t) - f_\epsilon (s) - (f_\epsilon' (t) - f_\epsilon' (s))| = \left| \sum_{k=j}^{\infty} \epsilon_k 2^{-k/p} \sin \left( \frac{2k\pi}{2^{j+1}} \right) - \sum_{k=j}^{\infty} \epsilon'_k 2^{-k/p} \sin \left( \frac{2k\pi}{2^{j+1}} \right) \right|
\]

Now consider \( k > j \)

\[
\sin \left( \frac{2^j \pi}{2^{j+1}} \right) = \sin \left( \frac{\pi}{2^{j+1}} \right) = 0
\]

for some integer \( m \). Thus the whole mess reduces to

\[
(\epsilon_j - \epsilon'_j) 2^{-j/p} \sin \left( \frac{2^j \pi (i + 1)}{2^{j+1}} \right) - (\epsilon_j - \epsilon'_j) 2^{-j/p} \sin \left( \frac{\pi i}{2^{j+1}} \right) = 2^{-j/p}
\]

In particular, \( |t - s| = \frac{1}{2^{j+1}} \) so \( 2^{1/p} |t - s|^{1/p} = 2^{-j/p} \)

\[
|f_\epsilon (t) - f_\epsilon (s) - (f_\epsilon' (t) - f_\epsilon' (s))| = 2 \left( 2^{1/p} \right) |t - s|^{1/p}
\]

which shows that

\[
\sup_{0 \leq s < t \leq 1} \frac{|f_\epsilon (t) - f_\epsilon' (t) - (f_\epsilon (s) - f_\epsilon' (s))|}{|t - s|^{1/p}} \geq 2^{1/p} (2)
\]

Thus there exists a set of uncountably many functions in \( C^{1/p} ([0, T]) \) and for any two of them \( f, g \), you get

\[
\|f - g\|_{C^{1/p}([0, 1])} > 2
\]

so \( C^{1/p} ([0, 1]) \) is not separable.

### 3.5 Exercises

1. Let \((X, \tau), (Y, \eta)\) be topological spaces and let \( A \subseteq X \) be compact. Then if \( f : X \to Y \) is continuous, show that \( f (A) \) is also compact.

2. ↑ In the context of Problem 1, suppose \( \mathbb{R} = Y \) where the usual topology is placed on \( \mathbb{R} \). Show \( f \) achieves its maximum and minimum on \( A \).
3. Let $V$ be an open set in $\mathbb{R}^n$. Show there is an increasing sequence of compact sets, $K_m$, such that $V = \bigcup_{m=1}^{\infty} K_m$. \textbf{Hint:} Let

$$C_m = \left\{ x \in \mathbb{R}^n : \text{dist} (x, V^C) \geq \frac{1}{m} \right\}$$

where

$$\text{dist}(x, S) \equiv \inf \{|y - x| \text{ such that } y \in S\}.$$ 

Consider $K_m \equiv C_m \cap B(0, m)$.

4. Let $\alpha \in [0, 1]$. Define, for $X$ a compact subset of $\mathbb{R}^p$,

$$C^\alpha (X; \mathbb{R}^n) \equiv \{ f \in C(X; \mathbb{R}^n) : \rho_\alpha (f) + ||f|| \equiv ||f||_\alpha < \infty \}$$

where

$$||f|| \equiv \sup\{|f(x)| : x \in X\}$$

and

$$\rho_\alpha (f) \equiv \sup\{|f(x) - f(y)|/|x - y|^\alpha : x, y \in X, x \neq y\}.$$ 

Show that $(C^\alpha (X; \mathbb{R}^n), ||\cdot||_\alpha)$ is a complete normed linear space.

5. Let $\{f_n\}_{n=1}^{\infty} \subseteq C^\alpha (X; \mathbb{R}^n)$ where $X$ is a compact subset of $\mathbb{R}^p$ and suppose

$$||f_n||_\alpha \leq M$$

for all $n$. Show there exists a subsequence, $n_k$, such that $f_{n_k}$ converges in $C(X; \mathbb{R}^n)$. The given sequence is called precompact when this happens. (This also shows the embedding of $C^\alpha (X; \mathbb{R}^n)$ into $C(X; \mathbb{R}^n)$ is a compact embedding.) Note that it is likely the case that $C^\alpha (X; \mathbb{R}^n)$ is not separable although it embeds continuously into a nice separable space. In fact, $C^\alpha ([0, T]; \mathbb{R}^n)$ can be shown to not be separable. See Definition 3.4.1 and the discussion which follows it.

6. Use the general Stone Weierstrass approximation theorem to prove Theorem 3.2.4.

7. Define $G_1$ to be the functions of the form $p(x) e^{-a|x|^2}$ where $a > 0$ is rational and $p(x)$ is a polynomial having all rational coefficients, $a_n$ being “rational” if it is of the form $a + ib$ for $a, b \in \mathbb{Q}$. Let $G$ be all finite sums of functions in $G_1$. Thus $G$ is an algebra of functions which has the property that if $f \in G$ then $\int f \in G$. Show that if $f$ is an arbitrary function in $C_0(\mathbb{R}^n)$, then there exists a sequence $\{f_n\}$ of functions in $G$ which converges uniformly to $f$.

8. Let $G$ consist of linear combinations of multiples of functions $t \to e^{-st}$ for $s > M$. Show this is dense in $C_0([0, \infty))$. Now suppose $f$ is a continuous function satisfying $|f(t)| < Ce^{-t}$. Suppose also that for all $s$ larger than $M$,

$$\int_0^\infty f(t) e^{-st} dt = 0$$
Show that then \( f(t) = 0 \) for all \( t \). Generalize this to the case where it is only known that \( f(t) \) has exponential growth. That is, for some \( \lambda, |f(t)| < Ce^{\lambda t} \).

This little problem justifies the method of Laplace transforms. You recall that you identified functions when their Fourier transforms were the same for large \( s \). However, they did not tell you why this is valid.

9. This problem will help to understand that a certain kind of function exists. Infinitely differentiable functions are not necessarily analytic.

\[
f(x) = \begin{cases} 
e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}
\]

show that \( f \) is infinitely differentiable. Note that you only need to be concerned with what happens at 0. There is no question elsewhere. This is a little fussy but is not too hard.

10. ↑Let \( f(x) \) be as given above. Now let

\[
\hat{f}(x) = \begin{cases} f(x) & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}
\]

Show that \( \hat{f}(x) \) is also infinitely differentiable. Now consider let \( r > 0 \) and define

\[
g(x) = \hat{f}(-(x-r)) \hat{f}(x+r).
\]

show that \( g \) is infinitely differentiable and vanishes for \( |x| \geq r \). Let \( \psi(x) = \prod_{k=1}^{p} g(x_k) \). For \( U = B(0,2r) \) with the norm given by \( ||x|| = \max \{|x_k|, k \leq p\} \), show that \( \psi \in C^\infty(U) \).

11. ↑Using the above problem, let \( \psi \in C^\infty(B(0,1)) \). Also let \( \psi \geq 0 \) as in the above problem. Show there exists \( \psi \geq 0 \) such that \( \psi \in C^\infty(B(0,1)) \) and \( \int \psi dm_p = 1 \). Now define

\[
\psi_n(x) = n^p \psi(nx)
\]

Show that \( \psi_n \) equals zero off a compact subset of \( B(0,\frac{1}{n}) \) and \( \int \psi_n dm_p = 1 \). We say that \( \text{spt}(\psi_n) \subseteq B(0,\frac{1}{n}) \). \( \text{spt}(f) \) is defined as the closure of the set on which \( f \) is not equal to 0. Such a sequence of functions as just defined \( \{\psi_n\} \) where \( \int \psi_n dm_p = 1 \) and \( \psi_n \geq 0 \) and \( \text{spt}(\psi_n) \subseteq B(0,\frac{1}{n}) \) is called a mollifier.

12. ↑Let \( \psi_n \) be a mollifier and let \( U \) be a bounded open set in \( \mathbb{R}^n \). Then \( f \in C(U) \) means that \( f \) is the restriction to \( U \) of a continuous function defined on \( \mathbb{R}^n \).

The convolution is

\[
f * \psi_n(x) = \int_{\mathbb{R}^n} f(x-y) \psi_n(y) dy
\]

Show that this function is infinitely differentiable and equals

\[
\int_{\mathbb{R}^n} f(y) \psi_n(x-y) dy
\]
also show that \( f \ast \psi_n \) converges uniformly to \( f \) on \( \bar{U} \). Note that if \( f \) has compact support, then so does \( f \ast \psi_n \). Also one sometimes considers \( \{ \psi_\varepsilon \} \) such that \( \varepsilon \to 0 \), \( \text{spt}(\psi_\varepsilon) \subseteq B(0, \varepsilon) \), \( \psi_\varepsilon(x) \geq 0 \), \( \int_{\mathbb{R}^n} \psi_\varepsilon(x) \, dx = 1 \). Then \( \varepsilon \to 0 \) corresponds to \( n \to \infty \).

13. Let \( U \) be a bounded open set. Show that there exists an infinitely differentiable function \( f : \bar{U} \to [0, 1] \) such that \( f = 1 \) on \( \bar{U} \) and \( f \) equals zero off some bounded open set. **Hint:** Consider \( W = U + B(0, 1) \) and \( W = W + B(0, 1) \). Now let \( \psi \) be infinitely differentiable having compact support in \( B(0, 1/2) \) and consider convolving \( X_{\bar{W}} \) with \( \psi \) where also \( \int \psi = 1 \).

14. If \( f : \bar{U} \to \mathbb{R} \) is continuous where \( U \) is a bounded open set in \( \mathbb{R}^n \), show that there exists an infinitely differentiable function \( g \) defined on \( \mathbb{R}^n \) being nonzero only in some bounded open set such that

\[
\|f - g\| \equiv \max_{x \in \bar{U}} |f(x) - g(x)| < \varepsilon.
\]

15. Do the above problem without using the Stone-Weierstrass theorem.
Chapter 4

Stone’s Theorem And Partitions Of Unity

This section is devoted to Stone’s theorem which says that a metric space is paracompact, defined below. See [45] for this which is where I read it. First is the definition of what is meant by a refinement.

Definition 4.0.1 Let $S$ be a topological space. We say that a collection of sets $\mathcal{D}$ is a refinement of an open cover $\mathcal{S}$, if every set of $\mathcal{D}$ is contained in some set of $\mathcal{S}$. An open refinement would be one in which all sets are open, with a similar convention holding for the term “closed refinement”.

Definition 4.0.2 We say that a collection of sets $\mathcal{D}$, is locally finite if for all $p \in S$, there exists $V$ an open set containing $p$ such that $V$ has nonempty intersection with only finitely many sets of $\mathcal{D}$.

Definition 4.0.3 We say $S$ is paracompact if it is Hausdorff and for every open cover $\mathcal{S}$, there exists an open refinement $\mathcal{D}$ such that $\mathcal{D}$ is locally finite and $\mathcal{D}$ covers $S$.

Theorem 4.0.4 If $\mathcal{D}$ is locally finite then

$$\bigcup\{D : D \in \mathcal{D}\} = \overline{\bigcup\{D : D \in \mathcal{D}\}}.$$

Proof: It is clear the left side is a subset of the right. Let $p$ be a limit point of $\bigcup\{D : D \in \mathcal{D}\}$.

and let $p \in V$, an open set intersecting only finitely many sets of $\mathcal{D}$, $D_1...D_n$. If $p$ is not in any of $\overline{D_i}$ then $p \in W$ where $W$ is some open set which contains no points of $\bigcup_{i=1}^{n} D_i$. Then $V \cap W$ contains no points of any set of $\mathcal{D}$ and this contradicts the assumption that $p$ is a limit point of $\bigcup\{D : D \in \mathcal{D}\}$. 

83
Thus \( p \in D_i \) for some \( i \).

We say \( \mathcal{S} \subseteq \mathcal{P}(S) \) is countably locally finite if

\[
\mathcal{S} = \bigcup_{n=1}^{\infty} \mathcal{S}_n
\]

and each \( \mathcal{S}_n \) is locally finite. The following theorem appeared in the 1950’s. It will be used to prove Stone’s theorem.

**Theorem 4.0.5** Let \( S \) be a regular topological space. (If \( p \in U \) open, then there exists an open set \( V \) such that \( p \in \overline{V} \subseteq U \).) The following are equivalent

1.) Every open covering of \( S \) has a refinement that is open, covers \( S \) and is countably locally finite.

2.) Every open covering of \( S \) has a refinement that is locally finite and covers \( S \). (The sets in refinement maybe not open.)

3.) Every open covering of \( S \) has a refinement that is closed, locally finite, and covers \( S \). (Sets in refinement are closed.)

4.) Every open covering of \( S \) has a refinement that is open, locally finite, and covers \( S \). (Sets in refinement are open.)

**Proof:**

1.) \( \Rightarrow \) 2.)

Let \( \mathcal{S} \) be an open cover of \( S \) and let \( \mathcal{B} \) be an open countably locally finite refinement

\[
\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{B}_n
\]

where \( \mathcal{B}_n \) is an open refinement of \( \mathcal{S} \) and \( \mathcal{B}_n \) is locally finite. For \( B \in \mathcal{B}_n \), let

\[
E_n(B) = B \setminus \bigcup_{k<n} (\cup \{B : B \in \mathcal{B}_k\}).
\]

Thus, in words, \( E_n(B) \) consists of points in \( B \) which are not in any set from any \( \mathcal{B}_k \) for \( k < n \).

**Claim:** \( \{E_n(B) : n \in \mathbb{N}, B \in \mathcal{B}_n\} \) is locally finite.

**Proof of the claim:** Let \( p \in S \). Then \( p \in B_0 \in \mathcal{B}_n \) for some \( n \). Let \( V \) be open, \( p \in V \), and \( V \) intersects only finitely many sets of \( \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_n \). Then consider \( B_0 \cap V \). If \( m > n \),

\[
(B_0 \cap V) \cap E_m(B) \subseteq \left[ \bigcup_{k<m} (\cup \{B : B \in \mathcal{B}_k\}) \right]^c \subseteq B_0^c.
\]

In words, \( E_m(B) \) has nothing in it from any of the \( \mathcal{B}_k \) for \( k < m \). In particular, it has nothing in it from \( B_0 \). Thus \( (B_0 \cap V) \cap E_m(B) = \emptyset \) for \( m > n \). Thus \( p \in B_0 \cap V \) which intersects only finitely many sets of \( \mathcal{S} \), no more than those intersected by \( V \). This establishes the claim.

**Claim:** \( \{E_n(B) : n \in \mathbb{N}, B \in \mathcal{B}_n\} \) covers \( S \).

**Proof:** Let \( p \in S \) and let \( n = \min\{k \in \mathbb{N} : p \in B \text{ for some } B \in \mathcal{B}_k\} \). Let \( p \in B \in \mathcal{B}_n \). Then \( p \in E_n(B) \).
The two claims show that 1.) $\Rightarrow$ 2.).

2.) $\Rightarrow$ 3.)

Let $\mathcal{S}$ be an open cover and let

$$
\mathcal{G} \equiv \{ U : U \text{ is open and } \overline{U} \subseteq V \in \mathcal{S} \text{ for some } V \in \mathcal{S} \}.
$$

Then since $S$ is regular, $\mathcal{G}$ covers $S$. (If $p \in S$, then $p \in U \subseteq \overline{U} \subseteq V \in \mathcal{S}$.) By 2.), $\mathcal{G}$ has a locally finite refinement $\mathcal{C}$, covering $S$. Consider

$$
\{ \mathcal{E} : E \in \mathcal{C} \}.
$$

This collection of closed sets covers $S$ and is locally finite because if $p \in S$, there exists $V$, $p \in V$, and $V$ has nonempty intersections with only finitely many elements of $\mathcal{C}$, say $E_1, \cdots, E_n$. If $\overline{\mathcal{E}} \cap V \neq \emptyset$, then $E \cap V \neq \emptyset$ and so $V$ intersects only $\overline{E_1}, \cdots, \overline{E_n}$. This shows 2.) $\Rightarrow$ 3.).

3.) $\Rightarrow$ 4.) Here is a table of symbols with a short summary of their meaning.

<table>
<thead>
<tr>
<th>Open covering</th>
<th>Locally finite refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{S}$ original covering</td>
<td>$\mathcal{B}$ by 3. can be closed refinement</td>
</tr>
<tr>
<td>$\mathcal{F}$ open interectors</td>
<td>$\mathcal{C}$ closed refinement</td>
</tr>
</tbody>
</table>

Let $\mathcal{S}$ be an open cover and let $\mathcal{B}$ be a locally finite refinement which covers $S$. By 3.) we can take $\mathcal{B}$ to be a closed refinement but this is not important here. Let

$$
\mathcal{F} \equiv \{ U : U \text{ is open and } U \text{ intersects only finitely many sets of } \mathcal{B} \}.
$$

Then $\mathcal{F}$ covers $S$ because $\mathcal{B}$ is locally finite. If $p \in S$, then there exists an open set $U$ containing $p$ which intersects only finitely many sets of $\mathcal{B}$. Thus $p \in U \in \mathcal{F}$. By 3., $\mathcal{F}$ has a locally finite closed refinement $\mathcal{C}$, which covers $S$. Define for $B \in \mathcal{B}$

$$
\mathcal{C} (B) \equiv \{ C \in \mathcal{C} : C \cap B = \emptyset \}
$$

Thus these closed sets $C$ do not intersect $B$ and so $B$ is in their complement. We use $\mathcal{C} (B)$ to fatten up $B$. Let

$$
E (B) \equiv \overline{\bigcup \{ C : C \in \mathcal{C} (B) \}}^C.
$$

In words, $E (B)$ is the complement of the union of all closed sets of $\mathcal{C}$ which do not intersect $B$. Thus $E (B) \supseteq B$, and has fattened up $B$. Then since $\mathcal{C} (B)$ is locally finite, $E (B)$ is an open set by Theorem [4.0.4]. Now let $F (B)$ be defined such that for $B \in \mathcal{B}$,

$$
B \subseteq F (B) \in \mathcal{S}
$$

(by definition $B$ is in some set of $\mathcal{S}$), and let

$$
\mathcal{L} = \{ E (B) \cap F (B) : B \in \mathcal{B} \}
$$

The intersection with $F (B)$ is to ensure that $\mathcal{L}$ is a refinement of $\mathcal{S}$. The important thing to notice is that if $C \in \mathcal{C}$ intersects $E (B)$, then it must also intersect
B. If not, you could include it in the list of closed sets which do not intersect \( B \) and whose complement is \( E(B) \). Thus \( E(B) \) would be too large.

**Claim:** \( \mathcal{L} \) covers \( S \).

This claim is obvious because if \( p \in S \) then \( p \in B \) for some \( B \in \mathfrak{B} \). Hence

\[
p \in E(B) \cap F(B) \in \mathcal{L}.
\]

**Claim:** \( \mathcal{L} \) is locally finite and a refinement of \( \mathfrak{S} \).

**Proof:** It is clear \( \mathcal{L} \) is a refinement of \( \mathfrak{S} \) because every set of \( \mathcal{L} \) is a subset of a set of \( \mathfrak{S} \), \( F(B) \). Let \( p \in S \). There exists an open set \( W \), such that \( p \in W \) and \( W \) intersects only \( C_1, \ldots, C_n \), elements of \( \mathfrak{C} \). Hence \( W \subseteq \bigcup_{i=1}^n C_i \) since \( \mathfrak{C} \) covers \( S \).

But \( C_i \) is contained in a set \( U_i \in \mathfrak{F} \) which intersects only finitely many sets of \( \mathfrak{B} \). Thus each \( C_i \) intersects only finitely many \( B \in \mathfrak{B} \) and so each \( C_i \) intersects only finitely many of the sets, \( E(B) \). (If it intersects \( E(B) \), then it intersects \( B \).) Thus \( W \) intersects only finitely many of the \( E(B) \), hence finitely many of the \( E(B) \cap F(B) \). It follows that \( \mathcal{L} \) is locally finite.

It is obvious that 4.\( \Rightarrow \) 1.).

The following theorem is Stone’s theorem.

**Theorem 4.0.6** If \( S \) is a metric space then \( S \) is paracompact (Every open cover has a locally finite open refinement also an open cover.)

**Proof:** Let \( \mathfrak{G} \) be an open cover. Well order \( \mathfrak{G} \). For \( B \in \mathfrak{G} \),

\[
B_n \equiv \{ x \in B : \text{dist} (x, B^C) < \frac{1}{2^n} \}, \; n = 1, 2, \ldots
\]

Thus \( B_n \) is contained in \( B \) but approximates it up to \( 2^{-n} \). Let

\[
E_n(B) = B_n \setminus \cup \{ D : D \prec B \text{ and } D \neq B \}
\]

where \( \prec \) denotes the well order. If \( B, D \in \mathfrak{G} \), then one is first in the well order. Let \( D \prec B \). Then from the construction, \( E_n(B) \subseteq D^C \) and \( E_n(D) \) is further than \( 1/2^n \) from \( D^C \). Hence, assuming neither set is empty,

\[
\text{dist} (E_n(B), E_n(D)) \geq 2^{-n}
\]
4.1. PARTITIONS OF UNITY AND STONE’S THEOREM

for all $B, D \in \mathcal{S}$. Fatten up $E_n(B)$ as follows.

$$\overline{E_n(B)} \equiv \bigcup \{B(x, 8^{-n}) : x \in E_n(B)\}.$$ 

Thus $\overline{E_n(B)} \subseteq B$ and

$$\text{dist} \left( \overline{E_n(B)}, \overline{E_n(D)} \right) \geq \frac{1}{2^n} - 2 \left(\frac{1}{8}\right)^n \equiv \delta_n > 0.$$ 

It follows that the collection of open sets

$$\{\overline{E_n(B)} : B \in \mathcal{S}\} \equiv \mathfrak{B}_n$$

is locally finite. In fact, $B \left(p, \frac{\delta_n}{2}\right)$ cannot intersect more than one of them. In addition to this,

$$S \subseteq \cup \{\overline{E_n(B)} : n \in \mathbb{N}, B \in \mathcal{S}\}$$

because if $p \in S$, let $B$ be the first set in $\mathcal{S}$ to contain $p$. Then $p \in E_n(B)$ for $n$ large enough because it will not be in anything deleted. Thus this is an open countably locally finite refinement. Thus 1.) in the above theorem is satisfied.

4.1 Partitions Of Unity And Stone’s Theorem

First observe that if $S$ is a nonempty set, then $\text{dist} \left( x, S \right)$ satisfies $|\text{dist} \left( x, S \right) - \text{dist} \left( y, S \right)| \leq d(x, y)$. To see this,

$$|\text{dist} \left( x, S \right) - \text{dist} \left( y, S \right)| \leq d(x, y)$$

To see this, say $\text{dist} \left( x, S \right)$ is the larger of the two. Then there exists $z \in S$ such that

$$\text{dist} \left( y, S \right) \geq d(y, z) + \varepsilon$$

It follows that

$$|\text{dist} \left( x, S \right) - \text{dist} \left( y, S \right)|$$

$$= \text{dist} \left( x, S \right) - \text{dist} \left( y, S \right)$$

$$\leq \text{dist} \left( x, S \right) - (d(y, z) + \varepsilon)$$

$$\leq d(x, z) - d(y, z) + \varepsilon$$

$$\leq d(x, y) + d(y, z) - d(y, z) + \varepsilon = d(x, y) + \varepsilon$$

Since $\varepsilon > 0$ is arbitrary, this shows the desired conclusion.

**Theorem 4.1.1** Let $S$ be a metric space and let $\mathcal{S}$ be any open cover of $S$. Then there exists a set $\mathfrak{F}$, an open refinement of $\mathcal{S}$, and functions $\{\phi_F : F \in \mathfrak{F}\}$ such that

$$\phi_F : S \rightarrow [0, 1]$$
CHAPTER 4. STONE’S THEOREM AND PARTITIONS OF UNITY

\( \phi_F \) is continuous
\( \phi_F (x) \) equals 0 for all but finitely many \( F \in \mathcal{F} \)
\[ \sum \{ \phi_F (x) : F \in \mathcal{F} \} = 1 \text{ for all } x \in S. \]

Each \( \phi_F \) is locally Lipschitz continuous which means that for each \( z \) there is an open set \( W \) containing \( z \) for which, if \( x, y \in W \), then there is a constant \( K \) such that
\[ |\phi_F (x) - \phi_F (y)| \leq Kd(x, y) \]

**Proof:** By Stone’s theorem, there exists a locally finite refinement \( \mathcal{F} \) covering \( S \). For \( F \in \mathcal{F} \)
\[ g_F (x) \equiv \text{dist} (x, F^C) \]
Let
\[ \phi_F (x) \equiv \left( \sum \{ g_F (x) : F \in \mathcal{F} \} \right)^{-1} g_F (x). \]
Now
\[ \sum \{ g_F (x) : F \in \mathcal{F} \} \]
is a continuous function because if \( x \in S \), then there exists an open set \( W \) with \( x \in W \) and \( W \) has nonempty intersection with only finitely many sets of \( F \in \mathcal{F} \). Then for \( y \in W \),
\[ \sum \{ g_F (y) : F \in \mathcal{F} \} = \sum_{i=1}^{n} g_{F_i} (y). \]
Since \( \mathcal{F} \) is a cover of \( S \),
\[ \sum \{ g_F (x) : F \in \mathcal{F} \} \neq 0 \]
for any \( x \in S \). Hence \( \phi_F \) is continuous. This also shows \( \phi_F (x) = 0 \) for all but finitely many \( F \in \mathcal{F} \). It is obvious that
\[ \sum \{ \phi_F (x) : F \in \mathcal{F} \} = 1 \]
from the definition.

Let \( z \in S \). Then there is an open set \( W \) containing \( z \) such that \( W \) has nonempty intersection with only finitely many \( F \in \mathcal{F} \). Thus for \( y, x \in W \),
\[ |\phi_{F_j} (x) - \phi_{F_j} (y)| \leq \left| \frac{\sum_{i=1}^{n} g_{F_i} (y) - g_{F_j} (y) \sum_{i=1}^{n} g_{F_i} (x)}{\sum_{i=1}^{n} g_{F_i} (x)} \right| \]
If \( F \) is not one of these \( F_i \), then \( g_F (x) = \phi_F (x) = \phi_F (y) = g_F (y) = 0 \). Thus there is nothing to show for these. It suffices to consider the ones above. Restricting \( W \) if necessary, we can assume that for \( x \in W \),
\[ \sum_{F} g_F (x) = \sum_{i=1}^{n} g_{F_i} (x) > \delta > 0, g_{F_j} (x) < \Delta < \infty, j \leq n \]
4.2. AN EXTENSION THEOREM, RETRACTS

Then, simplifying the above, and letting $x, y \in W$, for each $j \leq n$,

$$|\phi_{F_j}(x) - \phi_{F_j}(y)| \leq \frac{1}{\delta^2} \left| g_{F_j}(x) \sum_F g_F(y) - g_{F_j}(y) \sum_F g_F(x) \right| + \frac{1}{\delta^2} \sum_{i=1}^{n} |g_{F_i}(y) - g_{F_i}(x)|$$

$$\leq \frac{1}{\delta^2} \Delta |g_{F_j}(x) - g_{F_j}(y)| + \frac{1}{\delta^2} \sum_{i=1}^{n} |g_{F_i}(y) - g_{F_i}(x)|$$

$$\leq \frac{\Delta}{\delta^2} d(x, y) + \frac{\Delta}{\delta^2} nd(x, y) = (n + 1) \frac{\Delta}{\delta^2} d(x, y)$$

Thus on this set $W$ containing $z$, all $\phi_F$ are Lipschitz continuous with Lipschitz constant $(n + 1) \frac{\Delta}{\delta^2}$.

The functions described above are called a partition of unity subordinate to the open cover $\mathcal{S}$. A useful observation is contained in the following corollary.

**Corollary 4.1.2** Let $S$ be a metric space and let $\mathcal{S}$ be any open cover of $S$. Then there exists a set $\mathcal{F}$, an open refinement of $\mathcal{S}$, and functions $\{\phi_F : F \in \mathcal{F}\}$ such that

$$\phi_F : S \to [0, 1]$$

$$\phi_F$$ is continuous

$$\phi_F(x)$$ equals 0 for all but finitely many $F \in \mathcal{F}$

$$\sum_{F \in \mathcal{F}} \phi_F(x) = 1$$ for all $x \in S$.

Each $\phi_F$ is Lipschitz continuous. If $U \in \mathcal{S}$ and $H$ is a closed subset of $U$, the partition of unity can be chosen such that each $\phi_F = 0$ on $H$ except for one which equals 1 on $H$.

**Proof:** Just change your open cover to consist of $U$ and $V \setminus H$ for each $V \in \mathcal{S}$. Then every function but one equals 0 on $H$ and so exactly one of them equals 1 on $H$.

---

4.2 An Extension Theorem, Retracts

**Lemma 4.2.1** Let $A$ be a closed set in a metric space and let $x_n \notin A$, $x_n \to a_0 \in A$ and $a_n \in A$ such that $d(a_n, x_n) < 6 \text{ dist}(x_n, A)$. Then $a_n \to a_0$.

**Proof:** By assumption,

$$d(a_n, a_0) \leq d(a_n, x_n) + d(x_n, a_0) < 6 \text{ dist}(x_n, A) + d(x_n, a_0) \leq 6d(x_n, a_0) + d(x_n, a_0) = 7d(x_n, a_0)$$

and this converges to 0.
Note that there was nothing magic about 6 in the above. Another number would work as well.

In the proof of the following theorem, you get a covering of $A^C$ with open balls $B$ such that for each of these balls, there exists $a \in A$ such that for all $x \in B$, $\|x - a\| \leq 6 \text{dist}(x, A)$. The 6 is not important. Any other constant with this property would work. Then you use Stone’s theorem.

A Banach space is a normed vector space which is also a complete metric space where the metric comes from the norm.

$$d(x, y) = \|x - y\|$$

Thus you can add things in a Banach space. Much more will be considered about Banach spaces a little later.

**Definition 4.2.2** A Banach space is a complete normed linear space. If you have a subset $B$ of a Banach space, then $\text{conv}(B)$ denotes the smallest closed convex set which contains $B$. It can be obtained by taking the intersection of all closed convex sets containing $B$. Recall that a set $C$ is convex if whenever $x, y \in C$, then so is $\lambda x + (1 - \lambda) y$ for all $\lambda \in [0, 1]$. Note how this makes sense in a vector space but maybe not in a general metric space.

In the following theorem, we have in mind both $X$ and $Y$ are Banach spaces, but this is not needed in the proof. All that is needed is that $X$ is a metric space and $Y$ a normed linear space or possibly something more general in which it makes sense to do addition and scalar multiplication.

**Theorem 4.2.3** Let $A$ be a closed subset of a metric space $X$ and let $F : A \to Y$, $Y$ a normed linear space. Then there exists an extension of $F$ denoted as $\hat{F}$ such that $\hat{F}$ is defined on all of $X$ and agrees with $F$ on $A$. It has values in $\text{conv}(F(A))$, the convex hull of $F(A)$.

**Proof:** For each $c \notin A$, let $B_c$ be a ball contained in $A^C$ centered at $c$ where distance of $c$ to $A$ is at least $\text{diam}(B_c)$.
4.2. AN EXTENSION THEOREM, RETRACTS

So for \( x \in B_c \), what about \( \text{dist} (x, A) \)? How does it compare with \( \text{dist} (c, A) \)?

\[
\text{dist} (c, A) \leq d (c, x) + \text{dist} (x, A) \\
\leq \frac{1}{2} \text{diam} (B_c) + \text{dist} (x, A) \\
\leq \frac{1}{2} \text{dist} (c, A) + \text{dist} (x, A)
\]

so

\[
\text{dist} (c, A) \leq 2 \text{dist} (x, A)
\]

Now the following is also valid. Letting \( x \in B_c \) be arbitrary, it follows from the assumption on the diameter that there exists \( a_0 \in A \) such that \( d (c, a_0) < 2 \text{dist} (c, A) \).

Then

\[
d (x, a_0) \leq \sup_{y \in B_c} d (y, a_0) \leq \sup_{y \in B_c} (d (y, c) + d (c, a_0)) \leq \frac{\text{diam} (B_c)}{2} + 2 \text{dist} (c, A)
\]

\[
\leq \frac{\text{dist} (c, A)}{2} + 2 \text{dist} (c, A) < 3 \text{dist} (c, A)
\]

(4.2.1)

It follows from \( \text{conv} \)

\[
d (x, a_0) \leq 3 \text{dist} (c, A) \leq 6 \text{dist} (x, A)
\]

Thus for any \( x \in B_c \), there is an \( a_0 \in A \) such that \( d (x, a_0) \) is bounded by a fixed multiple of the distance from \( x \) to \( A \).

By Stone’s theorem, there is a locally finite open refinement \( \mathcal{R} \). These are open sets each of which is contained in one of the balls just mentioned such that each of these balls is the union of sets of \( \mathcal{R} \). Thus \( \mathcal{R} \) is a locally finite cover of \( A^C \). Since \( x \in A^C \) is in one of those balls, it was just shown that there exists \( a_R \in A \) such that for all \( x \in R \in \mathcal{R} \) we have \( d (x, a_R) \leq 6 \text{dist} (x, A) \). Of course there may be more than one because \( R \) might be contained in more than one of those special balls. One \( a_R \) is chosen for each \( R \in \mathcal{R} \).

Now let \( \phi_R (x) \equiv \text{dist} (x, R^C) \). Then let

\[
\hat{F} (x) \equiv \begin{cases} F (x) & \text{for } x \in A \\
\sum_{R \in \mathcal{R}} F (a_R) \frac{\phi_R (x)}{\sum_{R \in \mathcal{R}} \phi_R (x)} & \text{for } x \notin A \end{cases}
\]

The sum in the bottom is always finite because the covering is locally finite. Also, this sum is never 0 because \( \mathcal{R} \) is a covering. Also \( \hat{F} \) has values in \( \text{conv} \left( F (K) \right) \).

It only remains to verify that \( \hat{F} \) is continuous. It is clearly so on the interior of \( A \) thanks to continuity of \( F \). It is also clearly continuous on \( A^C \) because the functions \( \phi_R \) are continuous. So it suffices to consider \( x_n \to a \in \partial A \subseteq A \) where \( x_n \notin A \) and see whether \( F (a) = \lim_{n \to \infty} \hat{F} (x_n) \).

Suppose this does not happen. Then there is a sequence converging to some \( a \in \partial A \) and \( \varepsilon > 0 \) such that

\[
\varepsilon \leq \left\| \hat{F} (a) - \hat{F} (x_n) \right\| \text{ all } n
\]
For \( x_n \in R \), it was shown above that 
\[ d(x_n, a_{R_n}) \leq 6 \ \text{dist}(x_n, A). \]
By the above Lemma \[ \text{Lemma } 4.2.1 \], it follows that 
\[ a_{R_n} \to a \] and so 
\[ F(a_{R_n}) \to F(a). \]

By local finiteness of the cover, each \( x_n \) involves only finitely many \( R \). Thus, in this limit process, there are countably many \( R \) involved \( \{R_j\}_{j=1}^{\infty} \). Thus one can apply Fatou’s lemma.

\[ \varepsilon \leq \lim_{n \to \infty} \inf_{n \to \infty} \left\| \hat{F}(a) - \hat{F}(x_n) \right\| \]
\[ \leq \sum_{j=1}^{\infty} \lim_{n \to \infty} \inf_{n \to \infty} \left\| F(a_{R_j,n}) - F(a) \right\| \frac{\phi_{R_j}(x_{R_j,n})}{\sum_{j=1}^{\infty} \phi_{R_j}(x_{R_j,n})} \]
\[ \leq \sum_{j=1}^{\infty} \lim_{n \to \infty} \inf_{n \to \infty} \left\| F(a_{R_j,n}) - F(a) \right\| = 0 \]

The last step is needed because you lose local finiteness as you approach \( \partial A \). Note that the only thing needed was that \( X \) is a metric space. The addition takes place in \( Y \) so it needs to be a vector space. Did it need to be complete? No, this was not used. Nor was completeness of \( X \) used. The main interest here is in Banach spaces, but the result is more general than that.

It also appears that \( \hat{F} \) is locally Lipschitz on \( A^C \).

**Definition 4.2.4** Let \( S \) be a subset of \( X \), a Banach space. Then it is a retract if there exists a continuous function \( R : X \to S \) such that \( Rs = s \) for all \( s \in S \). This \( R \) is a retraction. More generally, \( S \subseteq T \) is called a retract of \( T \) if there is a continuous \( R : T \to S \) such that \( Rs = s \) for all \( s \in S \).

**Theorem 4.2.5** Let \( K \) be closed and convex subset of \( X \) a Banach space. Then \( K \) is a retract.

**Proof:** By Theorem \[ \text{Lemma } 4.2.4 \], there is a continuous function \( \hat{I} \) extending \( I \) to all of \( X \). Then also \( \hat{I} \) has values in \( \text{conv}(IK) = \text{conv}(K) = K \). Hence \( \hat{I} \) is a continuous function which does what is needed. It maps everything into \( K \) and keeps the points of \( K \) unchanged. ■

Sometimes people call the set a retraction also or the function which does the job a retraction. This seems like strange thing to call it because a retraction is the act of repudiating something you said earlier. Nevertheless, I will call it that. Note that if \( S \) is a retract of the whole metric space \( X \), then it must be a retract of every set which contains \( S \).

### 4.3 Something Which Is Not A Retract

The next lemma is a fundamental result which will be used to develop the Brouwer degree. It will also be used to give a short proof of the Brouwer fixed point theorem.
4.3. **SOMETHING WHICH IS NOT A RETRACT** in the exercises. This major fixed point theorem is probably the most fundamental theorem in nonlinear analysis. The proof outlined in the exercises is from [21].

**Lemma 4.3.1** Let \( g : U \to \mathbb{R}^n \) be \( C^2 \) where \( U \) is an open subset of \( \mathbb{R}^n \). Then

\[
\sum_{j=1}^{n} \text{cof} (Dg)_{ij,j} = 0,
\]

where here \( (Dg)_{ij} \equiv g_{i,j} \equiv \frac{\partial g_i}{\partial x_j} \). Also, \( \text{cof} (Dg)_{ij} = \frac{\partial \det(Dg)}{\partial g_{i,j}} \).

**Proof:** From the cofactor expansion theorem,

\[
\det (Dg) = \sum_{i=1}^{n} g_{i,j} \text{cof} (Dg)_{ij}
\]

and so

\[
\frac{\partial \det (Dg)}{\partial g_{i,j}} = \text{cof} (Dg)_{ij}
\] (4.3.2)

which shows the last claim of the lemma. Also

\[
\delta_{kj} \det (Dg) = \sum_{i} g_{i,k} (\text{cof} (Dg))_{ij}
\] (4.3.3)

because if \( k \neq j \) this is just the cofactor expansion of the determinant of a matrix in which the \( k^{th} \) and \( j^{th} \) columns are equal. Differentiate \( 4.3.3 \) with respect to \( x_j \) and sum on \( j \). This yields

\[
\sum_{r,s,j} \delta_{kj} \frac{\partial (\det Dg)}{\partial g_{r,s}} g_{r,sj} = \sum_{ij} g_{i,kj} (\text{cof} (Dg))_{ij} + \sum_{ij} g_{i,k} \text{cof} (Dg)_{ij,j}.
\]

Hence, using \( \delta_{kj} = 0 \) if \( j \neq k \) and \( 4.3.2 \),

\[
\sum_{rs} (\text{cof} (Dg))_{rs} g_{r,sk} = \sum_{rs} g_{r,ks} (\text{cof} (Dg))_{rs} + \sum_{ij} g_{i,k} \text{cof} (Dg)_{ij,j}.
\]

Subtracting the first sum on the right from both sides and using the equality of mixed partials,

\[
\sum_{i} g_{i,k} \left( \sum_{j} (\text{cof} (Dg))_{ij,j} \right) = 0.
\]

If \( \det (g_{i,k}) \neq 0 \) so that \( (g_{i,k}) \) is invertible, this shows \( \sum_{j} (\text{cof} (Dg))_{ij,j} = 0 \). If \( \det (Dg) = 0 \), let

\[
g_k = g + \varepsilon_k I
\]

where \( \varepsilon_k \to 0 \) and \( \det (Dg + \varepsilon_k I) \equiv \det (Dg_k) \neq 0 \). Then

\[
\sum_{j} (\text{cof} (Dg))_{ij,j} = \lim_{k \to \infty} \sum_{j} (\text{cof} (Dg_k))_{ij,j} = 0 \]
Definition 4.3.2 Let $h$ be a function defined on an open set, $U \subseteq \mathbb{R}^n$. Then $h \in C^k(U)$ if there exists a function $g$ defined on an open set, $W$ containing $U$ such that $g = h$ on $U$ and $g$ is $C^k(W)$.

Lemma 4.3.3 There does not exist $h \in C^2(B(0,R))$ such that $h : B(0,R) \to \partial B(0,R)$ which also has the property that $h(x) = x$ for all $x \in \partial B(0,R)$. That is, there is no retraction of $B(0,R)$ to $\partial B(0,R)$.

Proof: Suppose such an $h$ exists. Let $\lambda \in [0,1]$ and let $p_\lambda(x) \equiv x + \lambda(h(x) - x)$. This function, $p_\lambda$ is a homotopy of the identity map and the retraction, $h$. Let

$$I(\lambda) = \int_{B(0,R)} \det(Dp_\lambda(x)) \, dx.$$ 

Then using the dominated convergence theorem,

$$I'(\lambda) = \int_{B(0,R)} \sum_{i,j} \frac{\partial \det(Dp_\lambda(x))}{\partial p_{\lambda ij}} \frac{\partial p_{\lambda ij}}{\partial \lambda} \, dx$$

$$= \int_{B(0,R)} \sum_{i} \sum_{j} \frac{\partial \det(Dp_\lambda(x))}{\partial p_{\lambda ij}} (h_i(x) - x_i)_j \, dx$$

$$= \int_{B(0,R)} \sum_{i} \sum_{j} \text{cof}(Dp_\lambda(x))_{ij} (h_i(x) - x_i)_j \, dx$$

Now by assumption, $h_i(x) = x_i$ on $\partial B(0,R)$ and so one can integrate by parts and write

$$I'(\lambda) = - \sum_{i} \int_{B(0,R)} \sum_{j} \text{cof}(Dp_\lambda(x))_{ij} (h_i(x) - x_i) \, dx = 0.$$ 

Therefore, $I(\lambda)$ equals a constant. However,

$$I(0) = m_n(B(0,R)) > 0$$

but

$$I(1) = \int_{B(0,R)} \det(Dh(x)) \, dm_n = \int_{\partial B(0,R)} \#(y) \, dm_n = 0$$

because from polar coordinates or other elementary reasoning,

$$m_n(\partial B(0,1)) = 0.$$ 

The last formula uses the change of variables formula for functions which are not one to one. In this formula, $\#(y)$ equals the number of $x$ such that $h(x) = y$. 
4.3. SOMETHING WHICH IS NOT A RETRACT

To see this is so in case you have not seen this, note that $h$ is $C^1$ and so the inverse function theorem from advanced calculus applies. Thus

$$
\int_{B(0,R)} \det (Dh(x)) \, dm_n = \int_{\{\det(Dh(x)) > 0\}} \det (Dh(x)) \, dm_n \\
+ \int_{\{\det(Dh(x)) < 0\}} \det (Dh(x)) \, dm_n
$$

Thus $h$ is locally one to one on the two open sets $[\det (Dh(x)) > 0], [\det (Dh(x)) < 0]$

Now use inverse function theorem and change of variables for one to one $h$ to verify that both of these integrals equal 0. You cover $[\det (Dh(x)) > 0]$ with countably many balls on which $h$ is one to one and then use change of variables for each of these integrals over $[\det (Dh(x)) > 0]$ intersected with this ball.

The following is the Brouwer fixed point theorem for $C^2$ maps.

**Lemma 4.3.4** If $h \in C^2 \left(\overline{B(0,R)}\right)$ and $h : \overline{B(0,R)} \to \overline{B(0,R)}$, then $h$ has a fixed point, $x$ such that $h(x) = x$.

**Proof:** Suppose the lemma is not true. Then for all $x$, $|x - h(x)| \neq 0$. Then define

$$
g(x) = h(x) + \frac{x - h(x)}{|x - h(x)|} t(x)
$$

where $t(x)$ is nonnegative and is chosen such that $g(x) \in \partial B(0,R)$. This mapping is illustrated in the following picture.

If $x \to t(x)$ is $C^2$ near $\overline{B(0,R)}$, it will follow $g$ is a $C^2$ retraction onto $\partial B(0,R)$ contrary to Lemma 4.3.3. Thus $t(x)$ is the nonnegative solution to

$$
H(x, t) = |h(x)|^2 + 2 \left(h(x), \frac{x - h(x)}{|x - h(x)|}\right) t + t^2 = R^2
$$

(4.3.4)

Then

$$
H_t(x, t) = 2 \left(h(x), \frac{x - h(x)}{|x - h(x)|}\right) + 2t.
$$
Let \( f \) and so the restriction of \( x \). Then it has a fixed point \( \), which cannot happen unless \( \frac{\pi}{2} \). But this cannot happen because the angle between \( h(x) \) and \( x - h(x) \) cannot be \( \pi/2 \). Alternatively, if the above equals zero, you would need

\[
( h(x), x - h(x) ) = 0.
\]

But this cannot happen because the angle between \( h(x) \) and \( x - h(x) \) cannot be \( \pi/2 \). Alternatively, if the above equals zero, you would need

\[
( h(x), x - h(x) ) = | h(x) |^2 = R^2
\]

which cannot happen unless \( x = h(x) \) which is assumed not to happen. Therefore, \( x \to t(x) \) is \( C^2 \) near \( B(0,R) \) and so \( g(x) \) given above contradicts Lemma \[\text{[4.3.4]}\] 

Then the Brouwer fixed point theorem is as follows.

**Theorem 4.3.5** Let \( f : B(0,R) \to B(0,R) \) be continuous, this being a ball in \( \mathbb{R}^p \). Then it has a fixed point \( x \in B(0,R) \) such that \( f(x) = x \).

**Proof:** You can extend \( f \) to assume it is defined on all of \( \mathbb{R}^p \), \( f(\mathbb{R}^p) \subseteq B(0,R) \), the convex hull of \( B(0,R) \). Then letting \( \{ \psi_n \} \) be a mollifier, let \( f_n \equiv f \ast \psi_n \). Thus

\[
| f_n(x) | = \left| \int_{\mathbb{R}^p} f(t) \psi_n(x - t) \, dt \right|
\leq \int_{\mathbb{R}^p} | f(t) | \psi_n(x - t) \, dt
\leq R \int_{\mathbb{R}^p} \psi_n(x - t) \, dt = R
\]

and so the restriction of \( f_n \) to \( B(0,R) \) is \( C^2 \left( B(0,R) \right) \). Therefore, there exists \( x_n \in B(0,R) \) such that \( f_n(x_n) = x_n \). The functions \( f_n \) converge uniformly to \( f \) on \( B(0,R) \).

\[
| f(x) - f_n(x) | = \left| \int_{B(0,\frac{1}{n})} (f(x) - f(x - t)) \psi_n(t) \, dt \right|
\leq \int_{B(0,\frac{1}{n})} | f(x) - f(x - t) | \psi_n(t) \, dt < \varepsilon
\]
provided \( n \) is large enough, this for every \( x \in \overline{B(0,R)} \), this by uniform continuity of \( f \) on \( \overline{B(0,R+1)} \). There exists a subsequence, still called \( \{x_n\} \) which converges to \( x \in \overline{B(0,R)} \). Then using the uniform convergence of \( f_n \) to \( f \),

\[
f(x) = \lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} f_n(x_n) = \lim_{n \to \infty} x_n = x \quad \square
\]

**Definition 4.3.6** A nonempty topological space \( A \) is said to have the fixed point property if every continuous mapping \( f : A \to A \) has a fixed point.

### 4.4 Exercises

1. Suppose you have a Banach space \( X \) and a set \( A \subseteq X \). Suppose \( A \) is a retract of \( B \) where \( B \) has the fixed point property. By this is meant that \( A \subseteq B \) and there is a continuous function \( f : B \to A \) such that \( f \) equals the identity on \( A \). Show that it follows that then \( A \) also has the fixed point property.

2. Show that the fixed point property is a topological property. That is, if you have \( A, B \) two topological spaces and there is a continuous one to one onto mapping \( f : A \to B \) which has continuous inverse, then the two topological spaces either both have the fixed point property or neither one does.

3. The Brouwer fixed point theorem says that every closed ball in \( \mathbb{R}^n \) centered at 0 has the fixed point property. Show that it follows that every bounded convex closed set in \( \mathbb{R}^n \) has the fixed point property. **Hint:** You know that the closed convex set is a retract of \( \mathbb{R}^n \). Now if it is also a bounded set, then you could enclose it in \( B(0,r) \) for some large enough \( r \).

4. Convex closed sets in \( \mathbb{R}^n \) are retracts. Are there other examples of retracts not considered by Theorem 4.2.3?

5. In \( \mathbb{R}^2 \), consider an annulus, \( \{x : 1 \leq |x| \leq 2\} \). Show that this set does not have the fixed point property. Could it be a retract of \( \mathbb{R}^2 \)?

6. Does \( \{x \in \mathbb{R}^n : |x| = 1\} \) have the fixed point property? If \( n = 2 \), it is certainly the case that it does not. Consider a rotation through an angle of \( \pi/6 \). Letting \( R \) denote such a rotation, consider \( (x, y, z) \to (R(x, y), -z) \).

7. Suppose you have a closed subset \( H \) of \( X \) a metric space and suppose also that \( \mathcal{C} \) is an open cover of \( H \). Show there is another open cover \( \mathcal{C}' \) such that the closure of each open set in \( \mathcal{C} \) is contained in some set of \( \mathcal{C}' \). Do this in such a way that the new open cover is locally finite if \( \mathcal{C} \) is. **Hint:** You might want to use the fact that metric space is normal.

8. If \( H \) is a closed nonempty subset of \( \mathbb{R}^n \) and \( \mathcal{C} \) is an open cover of \( H \), show that there is a refined open cover such that each of the new open sets are bounded. In the partition of unity result obtained above, applied to \( H \) show that the functions in the partition of unity can be assumed to be infinitely differentiable with compact support.
9. Check that the conclusion of Theorem 4.2.3 applies for $X$ just a metric space. Then apply it to give another proof of the Tietze extension theorem.

10. There are many different proofs of the Brouwer fixed point theorem. Let $l$ be a line segment. Label one end with $A$ and the other end $B$. Now partition the segment into $n$ little pieces and label each of these partition points with either $A$ or $B$. Show there is an odd number of little segments with one end labeled with $A$ and the other labeled with $B$. If $f : l \to l$ is continuous, use the fact it is uniformly continuous and this little labeling result to give a proof for the Brouwer fixed point theorem for a one dimensional segment. Next consider a triangle. Label the vertices with $A, B, C$ and on the edge determined by $A, B$ label each vertex with either $A$ or $B$ and similarly with the segments $BC$ and $AC$. Then subdivide this triangle into little triangles, $T_1, \ldots, T_m$ in such a way that any pair of these little triangles intersects either along an entire edge or a vertex label the vertices of these little triangles on the side joining $A$ and $B$ either $A$ or $B$. Do something similar with the other two sides of the triangle. Now label the unlabeled vertices of these little triangles with either $A, B, C$ in any way. Show there is an odd number of little triangles having their vertices labeled as $A, B, C$. Use this to show the Brouwer fixed point theorem for any triangle. This approach generalizes to higher dimensions and you will see how this would take place if you are successful in going this far. This is an outline of the Sperner’s lemma approach to the Brouwer fixed point theorem. Are there other sets besides compact convex sets which have the fixed point property?

11. There is something interesting about fixed points. Suppose you have that $h_k : B \to B$ for $B$ a compact set and each $h_k$ has a fixed point. Suppose also that $h_k$ converges to $h$ uniformly on $B$. Then $h$ also has a fixed point. Verify this.

12. This problem is for those who have had some linear functional analysis. The Brouwer fixed point theorem is a finite dimensional creature. Consider a separable Hilbert space $H$ with a complete orthonormal basis $\{e_k\}_{k=1}^{\infty}$. Then define the following map. For $x = \sum_{i=1}^{\infty} x_i e_i$, define $L(\sum_{i=1}^{\infty} x_i e_i) \equiv \sum_{i=1}^{\infty} x_i e_{i+1}$. Now let $f(x) \equiv \frac{1}{2} (1 - \|x\|_H) e_1 + Lx$. Verify that $f : B(0, 1) \to B(0, 1)$ is continuous and yet it has no fixed point. This example is in 29.
Chapter 5

The Bochner Integral

5.1 Strong And Weak Measurability

This really belongs in the part of the book devoted to linear functional analysis. I am including it here because this important topic is almost universally neglected in a typical linear functional analysis course and it provides the basis for doing much of nonlinear analysis, especially that part pertaining to evolution equations.

In this chapter \((\Omega, \mathcal{S}, \mu)\) will be a \(\sigma\)-finite measure space and \(X\) will be a Banach space which contains the values of either a function or a measure. The Banach space will be either a real or a complex Banach space but the field of scalars does not matter and so it is denoted by \(F\) with the understanding that \(F = \mathbb{C}\) unless otherwise stated. The theory presented here includes the case where \(X = \mathbb{R}^n\) or \(\mathbb{C}^n\) but it does not include the situation where \(f\) could have values in a space like \([0, \infty]\). To begin with here is a definition.

**Definition 5.1.1** A function, \(x : \Omega \to X\), for \(X\) a Banach space, is a simple function if it is of the form

\[
x(s) = \sum_{i=1}^{n} a_i \chi_{B_i}(s)
\]

where \(B_i \in \mathcal{S}\) and \(\mu(B_i) < \infty\) for each \(i\). A function \(x\) from \(\Omega\) to \(X\) is said to be strongly measurable if there exists a sequence of simple functions \(\{x_n\}\) converging pointwise to \(x\). The function \(x\) is said to be weakly measurable if, for each \(f \in X'\),

\[
f \circ x
\]

is a scalar valued measurable function.

Earlier, a function was measurable if inverse images of open sets were measurable. Something similar holds here. The difference is that another condition needs to hold.
CHAPTER 5. THE BOCHNER INTEGRAL

Theorem 5.1.2 $x$ is strongly measurable if and only if $x^{-1}(U)$ is measurable for all $U$ open in $X$ and $x(\Omega)$ is separable.

Proof: Suppose first $x^{-1}(U)$ is measurable for all $U$ open in $X$ and $x(\Omega)$ is separable. Let $\{a_n\}_{n=1}^\infty$ be the dense subset of $x(\Omega)$. It follows $x^{-1}(B)$ is measurable for all $B$ Borel because

\[ \{ B : x^{-1}(B) \text{ is measurable} \} \]

is a $\sigma$ algebra containing the open sets. Let

\[ U^n_k \equiv \{ z \in X : \|z - a_k\| \leq \min\{\|z - a_l\|_{l=1}^n\} \}. \]

In words, $U^n_k$ is the set of points of $X$ which are as close to $a_k$ as they are to any of the $a_l$ for $l \leq n$.

\[ B^n_k \equiv x^{-1}(U^n_k), \quad D^n_k \equiv B^n_k \setminus (\cup_{l=1}^{k-1} B^n_l), \quad D^n_1 \equiv B^n_1, \]

and

\[ x_n(s) \equiv \sum_{k=1}^{n} a_k X_{D^n_k}(s). \]

Thus $x_n(s)$ is a closest approximation to $x(s)$ from $\{a_k\}_{k=1}^{n}$ and so $x_n(s) \to x(s)$ because $\{a_n\}_{n=1}^\infty$ is dense in $x(\Omega)$. Furthermore, $x_n$ is measurable because each $D^n_k$ is measurable.

Since $(\Omega, S, \mu)$ is $\sigma$ finite, there exists $\Omega_n \uparrow \Omega$ with $\mu(\Omega_n) < \infty$. Let

\[ y_n(s) \equiv X_{\Omega_n}(s) x_n(s). \]

Then $y_n(s) \to x(s)$ for each $s$ because for any $s, s \in \Omega_n$ if $n$ is large enough. Also $y_n$ is a simple function because it equals 0 off a set of finite measure.

Now suppose that $x$ is strongly measurable. Then some sequence of simple functions, $\{x_n\}$, converges pointwise to $x$. Then $x_n^{-1}(W)$ is measurable for every open set $W$ because it is just a finite union of measurable sets. Thus, $x_n^{-1}(W)$ is measurable for every Borel set $W$. This follows by considering

\[ \{ W : x_n^{-1}(W) \text{ is measurable} \} \]

and observing this is a $\sigma$ algebra which contains the open sets. Since $X$ is a metric space, it follows that if $U$ is an open set in $X$, there exists a sequence of open sets, $\{V_n\}$ which satisfies

\[ \nabla_n \subseteq U, \quad \nabla_n \subseteq V_{n+1}, \quad U = \cup_{n=1}^\infty V_n. \]

Then

\[ x^{-1}(V_m) \subseteq \bigcup_{n<\infty} \bigcap_{k\geq n} x^{-1}(V_m) \subseteq x^{-1}(\nabla_m). \]
5.1. STRONG AND WEAK MEASURABILITY

This implies

\[ x^{-1}(U) = \bigcup_{m<\infty} x^{-1}(V_m) \]

\[ \subseteq \bigcup_{m<\infty} \bigcup_{n<\infty} \bigcap_{k\geq n} x^{-1}(V_m) \subseteq \bigcup_{m<\infty} x^{-1}(\overline{V}_m) \subseteq x^{-1}(U). \]

Since

\[ x^{-1}(U) = \bigcup_{m<\infty} \bigcup_{n<\infty} \bigcap_{k\geq n} x^{-1}(V_m), \]

it follows that \( x^{-1}(U) \) is measurable for every open \( U \). It remains to show \( x(\Omega) \) is separable. Let

\[ D \equiv \text{all values of the simple functions } x_n \]

which converge to \( x \) pointwise. Then \( D \) is clearly countable and dense in \( \overline{D} \), a set which contains \( x(\Omega) \).

Claim: \( x(\Omega) \) is separable.

Proof of claim: For \( n \in \mathbb{N} \), let \( B_n \equiv \{ B(d,r) : 0 < r < \frac{1}{n}, \ r \text{ rational}, \ d \in D \} \).

Thus \( B_n \) is countable. Let \( z \in \overline{D} \). Consider \( B \left( z, \frac{1}{n} \right) \). Then there exists \( d \in D \cap B \left( z, \frac{1}{\sqrt{n}} \right) \). Now pick \( r \in \mathbb{Q} \cap \left( \frac{1}{\sqrt{n}}, \frac{1}{n} \right) \) so that \( B(d,r) \in B_n \). Now \( z \in B(d,r) \) and so this shows that \( x(\Omega) \subseteq \overline{D} \subseteq \bigcup B_n \) for each \( n \). Now let \( B'_n \) denote those sets of \( B_n \) which have nonempty intersection with \( x(\Omega) \). Say \( B'_n = \{ B'_k \}_{n,k=1}^{\infty} \). By the axiom of choice, there exists \( x'_n \in B'_k \cap x(\Omega) \). Then if \( z \in x(\Omega) \), \( z \) is contained in some set of \( B'_n \) which also contains a point of \( \{ x'_n \}_{n,k=1}^{\infty} \). Therefore, \( z \) is at least as close as \( 2/n \) to some point of \( \{ x'_n \}_{n,k=1}^{\infty} \) which shows \( \{ x'_n \}_{n,k=1}^{\infty} \) is a countable dense subset of \( x(\Omega) \). Therefore \( x(\Omega) \) is separable. ■

The last part also shows that a subset of a separable metric space is also separable. Therefore, the following simple corollary is obtained.

Corollary 5.1.3 If \( X \) is a separable Banach space then \( x \) is strongly measurable if and only if \( x^{-1}(U) \) is measurable for all \( U \) open in \( X \).

The next lemma is interesting for its own sake. Roughly it says that if a Banach space is separable, then the unit ball in the dual space is weak * separable. This will be used to prove Pettis’s theorem, one of the major theorems in this subject which relates weak measurability to strong measurability.

Lemma 5.1.4 If \( X \) is a separable Banach space with \( B' \) the closed unit ball in \( X' \), then there exists a sequence \( \{ f_n \}_{n=1}^{\infty} \equiv D' \subseteq B' \) with the property that for every \( x \in X \),

\[ ||x|| = \sup_{f \in D'} |f(x)|. \]

If \( H \) is a dense subset of \( X' \) then \( D' \) may be chosen to be contained in \( H \).
Proof: Let \( \{a_k\} \) be a countable dense set in \( X \), and consider the mapping
\[
\phi_n : B' \to \mathbb{F}^n
\]
given by
\[
\phi_n (f) \equiv (f(a_1), \ldots, f(a_n)).
\]

Then \( \phi_n (B') \) is contained in a compact subset of \( \mathbb{F}^n \) because \( |f(a_k)| \leq ||a_k|| \). Therefore, there exists a countable dense subset of \( \phi_n (B') \), \( \{ \phi_n (f_k) \}_{k=1}^\infty \). Then pick \( h^k_j \in H \cap B' \) such that \( \lim_{j \to \infty} ||f_k - h^k_j|| = 0 \). Then \( \{ \phi_n (h^k_j), k,j \} \) must also be dense in \( \phi_n (B') \). Let \( D'_n = \{ h^k_j, k,j \} \). Define
\[
D' \equiv \bigcup_{k=1}^\infty D'_k.
\]

Note that for each \( x \in X \), there exists \( f_x \in B' \) such that \( f_x(x) = ||x|| \). From the construction,
\[
||a_m|| = \sup \{ ||f(a_m)|| : f \in D' \}
\]
because \( f_{a_m}(a_m) \) is the limit of numbers \( f(a_m) \) for \( f \in D'_m \subset D' \). Therefore, for \( x \) arbitrary,
\[
||x|| \leq ||x - a_m|| + ||a_m|| = \sup \{ ||f(a_m)|| : f \in D' \} + ||x - a_m||
\]
\[
\leq \sup \{ ||f(a_m - x) + f(x)|| : f \in D' \} + ||x - a_m||
\]
\[
\leq \sup \{ ||f(x)|| : f \in D' \} + 2||x - a_m|| \leq ||x|| + 2||x - a_m||.
\]

Since \( a_m \) is arbitrary and the \( \{a_m\}_{m=1}^\infty \) are dense, this establishes the claim of the lemma. \( \blacksquare \)

The next theorem is one of the most important results in the subject. It is due to Pettis and appeared in 1938.

**Theorem 5.1.5** If \( x \) has values in a separable Banach space \( X \). Then \( x \) is weakly measurable if and only if \( x \) is strongly measurable.

**Proof:** It is necessary to show \( x^{-1}(U) \) is measurable whenever \( U \) is open. Since every open set is a countable union of balls, it suffices to show \( x^{-1}(B(a,r)) \) is measurable for any ball, \( B(a,r) \). Since every open ball is the countable union of closed balls, it suffices to verify \( x^{-1}(\overline{B(a,r)}) \) is measurable. From Lemma \ref{lemma}.

\[
x^{-1}(\overline{B(a,r)}) = \{ s : ||x(s) - a|| \leq r \}
\]
\[
= \left\{ s : \sup_{f \in D'} |f(x(s) - a)| \leq r \right\}
\]
\[
= \cap_{f \in D'} \{ s : |f(x(s) - a)| \leq r \}
\]
\[
= \cap_{f \in D'} \{ s : |f(x(s)) - f(a)| \leq r \}
\]
\[
= \cap_{f \in D'} (f \circ x)^{-1} \overline{B(f(a),r)}
\]
which equals a countable union of measurable sets because it is assumed that \( f \circ x \) is measurable for all \( f \in X' \).

Next suppose \( x \) is strongly measurable. Then there exists a sequence of simple functions \( x_n \) which converges to \( x \) pointwise. Hence for all \( f \in X' \), \( f \circ x_n \) is measurable and \( f \circ x_n \to f \circ x \) pointwise. Thus \( x \) is weakly measurable. ■

The same method of proof yields the following interesting corollary.

**Corollary 5.1.6** Let \( X \) be a separable Banach space and let \( B(X) \) denote the \( \sigma \) algebra of Borel sets. Let \( H \) be a dense subset of \( X' \). Then \( B(X) = \sigma(H) \equiv \mathcal{F} \), the smallest \( \sigma \) algebra of subsets of \( X \) which has the property that every function, \( x^* \in H \) is measurable.

**Proof:** First I need to show \( \mathcal{F} \) contains open balls because then \( \mathcal{F} \) will contain the open sets and hence the Borel sets. As noted above, it suffices to show \( \mathcal{F} \) contains closed balls. Let \( D' \) be those functionals in \( B' \) defined in Lemma 5.1.4.

\[
\{ x : ||x - a|| \leq r \} = \left\{ x : \sup_{x^* \in D'} |x^*(x - a)| \leq r \right\} = \bigcap_{x^* \in D'} \{ x : |x^*(x - a)| \leq r \} = \bigcap_{x^* \in D'} \{ x : |x^*(x)| - |x^*(a)| \leq r \} = \bigcap_{x^* \in D'} x^* - 1 \left( B\left(x^*(a), r \right) \right) \in \sigma(H)
\]

which is measurable because this is a countable intersection of measurable sets. Thus \( \mathcal{F} \) contains open sets so \( \sigma(H) \equiv \mathcal{F} \supseteq B(X) \).

To show the other direction for the inclusion, note that each \( x^* \) is \( B(X) \) measurable because \( x^* - 1 \) (open set) = open set. Therefore, \( B(X) \supseteq \sigma(H) \). ■

It is important to verify the limit of strongly measurable functions is itself strongly measurable. This happens under very general conditions. Suppose \( X \) is any separable metric space and let \( \tau \) denote the open sets of \( X \). Then it is routine to see that \( \tau \) has a countable basis, \( \mathcal{B} \).

(5.1.1)

Whenever \( U \in \mathcal{B} \), there exists a sequence of open sets, \( \{V_m\}_{m=1}^\infty \), such that

\[
\cdots V_m \subseteq \overline{V}_m \subseteq V_{m+1} \subseteq \cdots \cup_{m=1}^\infty V_m.
\]

(5.1.2)

**Theorem 5.1.7** Let \( f_n \) and \( f \) be functions mapping \( \Omega \) to \( X \) where \( \mathcal{F} \) is a \( \sigma \) algebra of measurable sets of \( \Omega \) and \( (X, \tau) \) is a topological space satisfying \( \tau \mathcal{B} \). Then if \( f_n \) is measurable, and \( f(\omega) = \lim_{n \to \infty} f_n(\omega) \), it follows that \( f \) is also measurable. (Pointwise limits of measurable functions are measurable.)

**Proof:** Let \( \mathcal{B} \) be the countable basis of \( \tau \mathcal{B} \) and let \( U \in \mathcal{B} \). Let \( \{V_m\} \) be the sequence of \( \mathcal{B} \). Since \( f \) is the pointwise limit of \( f_n \),

\[
f^{-1}(V_m) \subseteq \{ \omega : f_k(\omega) \in V_m \text{ for all } k \text{ large enough} \} \subseteq f^{-1}(\overline{V}_m).
\]
Therefore,
\[ f^{-1}(U) = \bigcup_{m=1}^{\infty} f^{-1}(V_m) \subseteq \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} f^{-1}(V_m) \]
\[ \subseteq \bigcup_{m=1}^{\infty} f^{-1}(V_m) = f^{-1}(U). \]
It follows \( f^{-1}(U) \in F \) because it equals the expression in the middle which is measurable. Now let \( W \in \tau \). Since \( B \) is countable, \( W = \bigcup_{n=1}^{\infty} U_n \) for some sets \( U_n \in B \). Hence
\[ f^{-1}(W) = \bigcup_{n=1}^{\infty} f^{-1}(U_n) \in F. \]

Note that the same conclusion would hold for any topological space with the property that for any open set \( U \), it has such a sequence of \( V_k \) attached to it as in 5.1.2.

**Corollary 5.1.8** \( x \) is strongly measurable if and only if \( x(\Omega) \) is separable and \( x \) is weakly measurable.

**Proof:** Strong measurability clearly implies weak measurability. If \( x_n(\omega) \to x(\omega) \) where \( x_n \) is simple, then \( f(x_n(\omega)) \to f(x(\omega)) \) for all \( f \in X' \). Hence \( f \circ x \) is measurable by Theorem 5.1.7 because it is the limit of a sequence of measurable functions. Let \( D \) denote the set of all values of \( x_n \). Then \( \overline{D} \) is a separable set containing \( x(\Omega) \). Thus \( \overline{D} \) is a separable metric space. Therefore \( x(\Omega) \) is separable also by the last part of the proof of Theorem 5.1.2.

Now suppose \( D \) is a countable dense subset of \( x(\Omega) \) and \( x \) is weakly measurable. Let \( Z \) be the subset consisting of all finite linear combinations of \( D \) with the scalars coming from the set of rational points of \( F \). Thus, \( Z \) is countable. Letting \( Y = \overline{Z} \), \( Y \) is a separable Banach space containing \( x(\Omega) \). If \( f \in Y' \), \( f \) can be extended to an element of \( X' \) by the Hahn Banach theorem. Therefore, \( x \) is a weakly measurable \( Y \) valued function. Now use Theorem 5.1.7 to conclude \( x \) is strongly measurable.

Weakly measurable as defined above means \( s \to x^*(x(\omega)) \) is measurable for every \( x^* \in X' \). The next lemma ties this weak measurability to the usual version of measurability in which a function is measurable when inverse images of open sets are measurable.

**Lemma 5.1.9** Let \( X \) be a Banach space and let \( x : (\Omega, F) \to K \subseteq X \) where \( K \) is weakly compact and \( X' \) is separable. Then \( x \) is weakly measurable if and only if \( x^{-1}(U) \in F \) whenever \( U \) is a weakly open set.

**Proof:** By Problem 34 on Page 49 and Problem 36 on Page 50, there exists a metric \( d \), such that the metric space topology with respect to \( d \) coincides with the weak topology. Since \( K \) is compact, it follows that \( K \) is also separable. Hence it is completely separable and so there exists a countable basis of open sets \( B \) for the weak topology on \( K \). It follows that if \( U \) is any weakly open set, covered by basic sets of the form \( B_A(x, r) \) where \( A \) is a finite subset of \( X' \), there exists a countable collection of these sets of the form \( B_A(x, r) \) which covers \( U \).
5.1. STRONG AND WEAK MEASURABILITY

Suppose now that $x$ is weakly measurable. To show $x^{-1}(U) \in \mathcal{F}$ whenever $U$ is weakly open, it suffices to verify $x^{-1}(B_A(z,r)) \in \mathcal{F}$ for any set, $B_A(z,r)$. Let $A = \{x_1^*, \ldots, x_m^*\}$. Then

$$x^{-1}(B_A(z,r)) = \{ s \in \Omega : \rho_A(x(s) - z) < r \}
\equiv \left\{ s \in \Omega : \max_{x^* \in A} |x^*(x(s) - z)| < r \right\}
\equiv \bigcup_{i=1}^m \left\{ s \in \Omega : |x_i^*(x(s)) - z| < r \right\}
\equiv \bigcup_{i=1}^m \left\{ s \in \Omega : |x_i^*(x(s)) - x_i^*(z)| < r \right\}
$$

which is measurable because each $x_i^* \circ x$ is given to be measurable.

Next suppose $x^{-1}(U) \in \mathcal{F}$ whenever $U$ is weakly open. Then in particular this holds when $U = B_{x^*}(z,r)$ for arbitrary $x^*$. Hence

$$\{ s \in \Omega : x(s) \in B_{x^*}(z,r) \} \in \mathcal{F}.$$  

But this says the same as

$$\{ s \in \Omega : |x^*(x(s)) - x^*(z)| < r \} \in \mathcal{F}$$

Since $x^*(z)$ can be a completely arbitrary element of $\mathcal{F}$, it follows $x^* \circ x$ is an $\mathcal{F}$ valued measurable function. In other words, $x$ is weakly measurable according to the former definition. ■

One can also define weak $\ast$ measurability and prove a theorem just like the Pettis theorem above. The next lemma is the analogue of Lemma 5.1.4.

**Lemma 5.1.10** Let $B$ be the closed unit ball in $X$. If $X'$ is separable, there exists a sequence $\{x_m\}_{m=1}^\infty \equiv D \subseteq B$ with the property that for all $y^* \in X'$,

$$||y^*|| = \sup_{x \in D} |y^*(x)|.$$  

**Proof:** Let

$$\{x_k^*\}_{k=1}^\infty$$

be the dense subspace of $X'$. Define $\phi_n : B \to \mathbb{F}^n$ by

$$\phi_n(x) \equiv (x_1^*(x), \ldots, x_n^*(x)).$$

Then $|x_k^*(x)| \leq ||x_k^*||$ and so $\phi_n(B)$ is contained in a compact subset of $\mathbb{F}^n$. Therefore, there exists a countable set, $D_n \subseteq B$ such that $\phi_n(D_n)$ is dense in $\phi_n(B)$. Let

$$D \equiv \bigcup_{n=1}^\infty D_n.$$  

It remains to verify this works. Let $y^* \in X'$. Then there exists $y$ such that

$$|y^*(y)| > ||y^*|| - \varepsilon.$$
By density, there exists one of the $x^*_k$ from the countable dense subset of $X'$ such that also

$$|x^*_k(y)| > ||y^*|| - \varepsilon, \ |x^*_k - y^*| < \varepsilon.$$ 

Now $x^*_k(y) \in \phi_k(B)$ and so there exists $x \in D_k \subseteq D$ such that

$$|x^*_k(x)| > ||y^*|| - \varepsilon.$$ 

Then since $||x^*_k - y^*|| < \varepsilon$, this implies

$$|y^*(x)| \geq ||y^*|| - 2\varepsilon.$$ 

Since $\varepsilon > 0$ is arbitrary,

$$||y^*|| \leq \sup_{x \in D} |y^*(x)| \leq ||y^*|| \quad \blacksquare$$

The next theorem is another version of the Pettis theorem. First here is a definition.

**Definition 5.1.11** A function $y$ having values in $X'$ is weak $\ast$ measurable, when for each $x \in X$, $y(\cdot)(x)$ is a measurable scalar valued function.

**Theorem 5.1.12** If $X'$ is separable and $y : \Omega \rightarrow X'$ is weak $\ast$ measurable, then $y$ is strongly measurable.

**Proof:** It is necessary to show $y^{-1}(B(a^*, r))$ is measurable. This will suffice because the separability of $X'$ implies every open set is the countable union of such balls of the form $B(a^*, r)$. It also suffices to verify inverse images of closed balls are measurable because every open ball is the countable union of closed balls. From Lemma 5.1.10,

$$y^{-1}(B(a^*, r)) = \{s : ||y(s) - a^*|| \leq r\}$$

$$= \left\{s : \sup_{x \in D} |(y(s) - a^*)(x)| \leq r\right\}$$

$$= \left\{s : \sup_{x \in D} |y(s)(x) - a^*(x)| \leq r\right\}$$

$$= \cap_{x \in D} y(\cdot)(x)^{-1}(B(a^*(x), r))$$

which is a countable intersection of measurable sets by hypothesis. \(\blacksquare\)

The following are interesting consequences of the theory developed so far and are of interest independent of the theory of integration of vector valued functions.

**Theorem 5.1.13** If $X'$ is separable, then so is $X$. 

---

106 \textit{CHAPTER 5. THE BOCHNER INTEGRAL}
5.2. THE ESSENTIAL BOCHNER INTEGRAL

Proof: Let $D = \{x_m\} \subseteq B$, the unit ball of $X$, be the sequence promised by Lemma 5.1.10. Let $V$ be all finite linear combinations of elements of $\{x_m\}$ with rational scalars. Thus $V$ is a separable subspace of $X$. The claim is that $V = X$. If not, there exists $x_0 \in X \setminus V$.

But by the Hahn Banach theorem there exists $x_0^* \in X'$ satisfying $x_0^* (x_0) \neq 0$, but $x_0^* (v) = 0$ for every $v \in V$. Hence

$$||x_0^*|| = \sup_{x \in D} |x_0^*(x)| = 0,$$

a contradiction.

Corollary 5.1.14 If $X$ is reflexive, then $X$ is separable if and only if $X'$ is separable.

Proof: From the above theorem, if $X'$ is separable, then so is $X$. Now suppose $X$ is separable with a dense subset equal to $D$. Then since $X$ is reflexive, $J(D)$ is dense in $X''$ where $J$ is the James map satisfying $Jx (x^*) = x^* (x)$. Then since $X''$ is separable, it follows from the above theorem that $X'$ is also separable.

5.2 The Essential Bochner Integral

Definition 5.2.1 Let $a_k \in X$, a Banach space and let

$$x(s) = \sum_{k=1}^{n} a_k \chi_{E_k}(s)$$

(5.2.3)

where for each $k$, $E_k$ is measurable and $\mu(E_k) < \infty$. Then define

$$\int_{\Omega} x(s) \, d\mu = \sum_{k=1}^{n} a_k \mu(E_k).$$

Proposition 5.2.2 Definition 5.2.1 is well defined.

Proof: It suffices to verify that if

$$\sum_{k=1}^{n} a_k \chi_{E_k}(s) = 0,$$

then

$$\sum_{k=1}^{n} a_k \mu(E_k) = 0.$$

Let $f \in X'$. Then

$$f \left( \sum_{k=1}^{n} a_k \chi_{E_k}(s) \right) = \sum_{k=1}^{n} f(a_k) \chi_{E_k}(s) = 0.$$
and, therefore,

$$0 = \int_{\Omega} \left( \sum_{k=1}^{n} f(a_k) \chi_{E_k}(s) \right) \, d\mu = \sum_{k=1}^{n} f(a_k) \mu(E_k) = f \left( \sum_{k=1}^{n} a_k \mu(E_k) \right).$$

Since $f \in X'$ is arbitrary, and $X'$ separates the points of $X$, it follows that

$$\sum_{k=1}^{n} a_k \mu(E_k) = 0$$

as claimed. This proves the proposition.

It follows easily from this proposition that $\int_{\Omega} d\mu$ is well defined and linear on simple functions.

**Definition 5.2.3** A strongly measurable function $x$ is Bochner integrable if there exists a sequence of simple functions $x_n$ converging to $x$ pointwise and satisfying

$$\int_{\Omega} ||x_n(s) - x_m(s)|| \, d\mu \to 0 \text{ as } m, n \to \infty. \quad (5.2.4)$$

If $x$ is Bochner integrable, define

$$\int_{\Omega} x(s) \, d\mu \equiv \lim_{n \to \infty} \int_{\Omega} x_n(s) \, d\mu. \quad (5.2.5)$$

**Theorem 5.2.4** The Bochner integral is well defined and if $x$ is Bochner integrable and $f \in X'$,

$$f \left( \int_{\Omega} x(s) \, d\mu \right) = \int_{\Omega} f(x(s)) \, d\mu \quad (5.2.6)$$

and

$$\left\| \int_{\Omega} x(s) \, d\mu \right\| \leq \int_{\Omega} ||x|| \, d\mu. \quad (5.2.7)$$

Also, the Bochner integral is linear. That is, if $a, b$ are scalars and $x, y$ are two Bochner integrable functions, then

$$\int_{\Omega} (ax(s) + by(s)) \, d\mu = a \int_{\Omega} x(s) \, d\mu + b \int_{\Omega} y(s) \, d\mu \quad (5.2.8)$$

**Proof:** First it is shown that the triangle inequality holds on simple functions and that the limit in (5.2.5) exists. Thus, if $x$ is given by (5.2.3) with the $E_k$ disjoint,

$$\left\| \int_{\Omega} x(s) \, d\mu \right\|$$

$$= \left\| \int_{\Omega} \sum_{k=1}^{n} a_k \chi_{E_k}(s) \, d\mu \right\| = \left\| \sum_{k=1}^{n} a_k \mu(E_k) \right\|$$

$$\leq \sum_{k=1}^{n} ||a_k|| \mu(E_k) = \int_{\Omega} \sum_{k=1}^{n} ||a_k|| \chi_{E_k}(s) \, d\mu = \int_{\Omega} ||x(s)|| \, d\mu.$$
which shows the triangle inequality holds on simple functions. This implies
\[
\left| \int_\Omega x_n(s) \, d\mu - \int_\Omega x_m(s) \, d\mu \right| = \left| \int_\Omega (x_n(s) - x_m(s)) \, d\mu \right| \\
\leq \int_\Omega \|x_n(s) - x_m(s)\| \, d\mu
\]
which verifies the existence of the limit in (5.2.4). This completes the first part of the argument.

Next it is shown the integral does not depend on the choice of the sequence satisfying (5.2.4) so that the integral is well defined. Suppose \(y_n, x_n\) both satisfy (5.2.4) and converge to \(x\) pointwise. By Fatou’s lemma,
\[
\left| \int_\Omega y_n \, d\mu - \int_\Omega x \, d\mu \right| \leq \int_\Omega \|y_n - x\| \, d\mu + \int_\Omega \|x - x_m\| \, d\mu
\]
\[
\leq \lim \inf_{k \to \infty} \int_\Omega \|y_n - y_k\| \, d\mu + \lim \inf_{k \to \infty} \int_\Omega \|x_k - x_m\|
\leq \varepsilon/2 + \varepsilon/2
\]
if \(m\) and \(n\) are chosen large enough. Since \(\varepsilon\) is arbitrary, this shows the limit is the same for both sequences and demonstrates the Bochner integral is well defined.

It remains to verify the triangle inequality on Bochner integral functions and the claim about passing a continuous linear functional inside the integral. Let \(x\) be Bochner integrable and let \(x_n\) be a sequence which satisfies the conditions of the definition. Define
\[
y_n(s) = \begin{cases} x_n(s) & \text{if } \|x_n(s)\| \leq 2\|x(s)\|, \\ 0 & \text{if } \|x_n(s)\| > 2\|x(s)\|. \end{cases}
\]
(5.2.9)
Thus
\[y_n(s) = x_n(s) \chi_{\|x_n\| \leq 2\|x\|}(s)\] .
If \(x(s) = 0\) then \(y_n(s) = 0\) for all \(n\). If \(\|x(s)\| > 0\) then for all \(n\) large enough,
\[y_n(s) = x_n(s)\]
Thus, \(y_n(s) \to x(s)\) and
\[\|y_n(s)\| \leq 2\|x(s)\|.
\]
(5.2.10)
By Fatou’s lemma,
\[
\int_\Omega \|x\| \, d\mu \leq \lim \inf_{n \to \infty} \int_\Omega \|x_n\| \, d\mu.
\]
(5.2.11)
Also from (5.2.4) and the triangle inequality on simple functions, \(\left\{ \int_\Omega \|x_n\| \, d\mu \right\}_{n=1}^{\infty}\) is a Cauchy sequence and so it must be bounded. Therefore, by (5.2.11) and the dominated convergence theorem,
\[
0 = \lim_{n,m \to \infty} \int_\Omega \|y_n - y_m\| \, d\mu
\]
(5.2.12)
and it follows $x_n$ can be replaced with $y_n$ in Definition 5.2.3.

From Definition 5.2.1,

$$f \left( \int_{\Omega} y_n \, d\mu \right) = \int_{\Omega} f(y_n) \, d\mu.$$  

Thus,

$$f \left( \int_{\Omega} x \, d\mu \right) = \lim_{n \to \infty} f \left( \int_{\Omega} y_n \, d\mu \right) = \lim_{n \to \infty} \int_{\Omega} f(y_n) \, d\mu = \int_{\Omega} f(x) \, d\mu,$$

the last equation holding from the dominated convergence theorem and 5.2.10 and 5.2.11. This shows 5.2.6. To verify 5.2.7, 

$$\int_{\Omega} x(s) \, d\mu = \lim_{n \to \infty} \int_{\Omega} y_n(s) \, d\mu \leq \lim_{n \to \infty} \int_{\Omega} ||y_n(s)|| \, d\mu = \int_{\Omega} ||x(s)|| \, d\mu,$$

where the last equation follows from the dominated convergence theorem and 5.2.10, 5.2.11.

It remains to verify 5.2.8. Let $f \in X'$. Then from 5.2.10

$$f \left( \int_{\Omega} (ax(s) + by(s)) \, d\mu \right) = \int_{\Omega} (af(x(s)) + bf(y(s))) \, d\mu = a \int_{\Omega} f(x(s)) \, d\mu + b \int_{\Omega} f(y(s)) \, d\mu = f \left( a \int_{\Omega} x(s) \, d\mu + b \int_{\Omega} y(s) \, d\mu \right).$$

Since $X'$ separates the points of $X$, it follows

$$\int_{\Omega} (ax(s) + by(s)) \, d\mu = a \int_{\Omega} x(s) \, d\mu + b \int_{\Omega} y(s) \, d\mu$$

and this proves 5.2.8. This proves the theorem.

**Theorem 5.2.5** An $X$ valued function, $x$, is Bochner integrable if and only if $x$ is strongly measurable and

$$\int_{\Omega} ||x(s)|| \, d\mu < \infty. \quad (5.2.13)$$

In this case there exists a sequence of simple functions $\{y_n\}$ satisfying 5.2.4, $y_n(s)$ converging pointwise to $x(s)$,

$$||y_n(s)|| \leq 2 ||x(s)|| \quad (5.2.14)$$

and

$$\lim_{n \to \infty} \int_{\Omega} ||x(s) - y_n(s)|| \, d\mu = 0. \quad (5.2.15)$$
5.2. THE ESSENTIAL BOCHNER INTEGRAL

Proof: Suppose $x$ is strongly measurable and condition $5.2.13$ holds. Since $x$ is strongly measurable, there exists a sequence of simple functions, $\{x_n\}$ converging pointwise to $x$. As before, let

$$y_n(s) = \begin{cases} x_n(s) & \text{if } ||x_n(s)|| \leq 2||x(s)||, \\ 0 & \text{if } ||x_n(s)|| > 2||x(s)||. \end{cases}$$

Then $5.2.14$ holds for $y_n$ and $y_n(s) \to x(s)$. Also

$$0 = \lim_{m,n \to \infty} \int_\Omega ||y_n(s) - y_m(s)|| \, d\mu$$

since otherwise, there would exist $\varepsilon > 0$ and $N_\varepsilon \to \infty$ as $\varepsilon \to 0$ and $n_\varepsilon, m_\varepsilon > N_\varepsilon$ such that

$$\int_\Omega ||y_{n_\varepsilon}(s) - y_{m_\varepsilon}(s)|| \, d\mu \geq \varepsilon.$$

But then taking a limit as $\varepsilon \to 0$ and using the dominated convergence theorem and $5.2.14$ and $5.2.13$, this would imply $0 \geq \varepsilon$. Therefore, $x$ is Bochner integrable. $5.2.15$ follows from the dominated convergence theorem and $5.2.14$.

Now suppose $x$ is Bochner integrable. Then it is strongly measurable and there exists a sequence of simple functions $\{x_n\}$ such that $x_n(s)$ converges pointwise to $x$ and

$$\lim_{m,n \to \infty} \int_\Omega ||x_n(s) - x_m(s)|| \, d\mu = 0.$$

Therefore, as before, since $\{\int_\Omega x_n \, d\mu\}_{n=1}^{\infty}$ is a Cauchy sequence, it follows

$$\left\{ \int_\Omega ||x_n|| \, d\mu \right\}_{n=1}^{\infty}$$

is also a Cauchy sequence because

$$\left| \int_\Omega ||x_n|| \, d\mu - \int_\Omega ||x_m|| \, d\mu \right| \leq \int_\Omega (||x_n|| - ||x_m||) \, d\mu \leq \int_\Omega ||x_n - x_m|| \, d\mu.$$

Thus

$$\int_\Omega ||x|| \, d\mu \leq \lim_{n \to \infty} \inf \int_\Omega ||x_n|| \, d\mu < \infty.$$

Using $5.2.16$ it follows $y_n$ satisfies $5.2.14$, converges pointwise to $x$ and then from the dominated convergence theorem $5.2.15$ holds. This proves the theorem.

Here is a simple corollary.

**Corollary 5.2.6** Let an $X$ valued function $x$ be Bochner integrable and let $L \in \mathcal{L}(X,Y)$ where $Y$ is another Banach space. Then $Lx$ is a $Y$ valued Bochner integrable function and

$$L \left( \int_\Omega x(s) \, d\mu \right) = \int_\Omega Lx(s) \, d\mu.$$
Proof: From Theorem 5.2.5 there is a sequence of simple functions \( \{ y_n \} \) having the properties listed in that theorem. Then consider \( \{ Ly_n \} \) which converges pointwise to \( Lx \). Since \( L \) is continuous and linear,

\[
\int_\Omega ||Ly_n - Lx||_Y d\mu \leq ||L|| \int_\Omega ||y_n - x||_X d\mu
\]

which converges to 0. This implies

\[
\lim_{m,n \to \infty} \int_\Omega ||Ly_n - Ly_m|| d\mu = 0
\]

and so by definition \( Lx \) is Bochner integrable. Also

\[
\int_\Omega x(s) d\mu = \lim_{n \to \infty} \int_\Omega y_n(s) d\mu
\]

\[
\int_\Omega Ly(s) d\mu = \lim_{n \to \infty} \int_\Omega Ly_n(s) d\mu
\]

\[
= \lim_{n \to \infty} L \int_\Omega y_n(s) d\mu
\]

\[
||L \left( \int_\Omega x(s) d\mu \right) - \int_\Omega Lx(s) d\mu ||_Y
\]

\[
\leq ||L \left( \int_\Omega x(s) d\mu \right) - L \int_\Omega y_n(s) d\mu ||_Y
\]

\[
+ \left| \int_\Omega Ly_n(s) d\mu - \int_\Omega Lx(s) d\mu \right|_Y < \varepsilon/2 + \varepsilon/2 = \varepsilon
\]

whenever \( n \) large enough. This proves the corollary.

5.3 The Spaces \( L^p(\Omega; X) \)

Definition 5.3.1 \( x \in L^p(\Omega; X) \) for \( p \in [1, \infty) \) if \( x \) is strongly measurable and

\[
\int_\Omega ||x(s)||^p d\mu < \infty
\]

Also

\[
||x||_{L^p(\Omega; X)} \equiv ||x||_p \equiv \left( \int_\Omega ||x(s)||^p d\mu \right)^{1/p}
\]

(5.3.17)

As in the case of scalar valued functions, two functions in \( L^p(\Omega; X) \) are considered equal if they are equal a.e. With this convention, and using the same arguments found in the presentation of scalar valued functions it is clear that \( L^p(\Omega; X) \) is a normed linear space with the norm given by (5.3.17). In fact, \( L^p(\Omega; X) \) is a Banach space. This is the main contribution of the next theorem.
Lemma 5.3.2 If \( x_n \) is a Cauchy sequence in \( L^p(\Omega; X) \) satisfying
\[
\sum_{n=1}^{\infty} ||x_{n+1} - x_n||_p < \infty,
\]
then there exists \( x \in L^p(\Omega; X) \) such that \( x_n(s) \to x(s) \) a.e. and
\[
||x - x_n||_p \to 0.
\]

Proof: Let
\[
g_N(s) = \sum_{n=1}^{N} ||x_{n+1}(s) - x_n(s)||_X.
\]
Then by the triangle inequality,
\[
\left( \int_{\Omega} g_N(s)^p \, d\mu \right)^{1/p} \leq \sum_{n=1}^{N} \left( \int_{\Omega} ||x_{n+1}(s) - x_n(s)||^p \, d\mu \right)^{1/p} \leq \sum_{n=1}^{\infty} ||x_{n+1} - x_n||_p < \infty.
\]

Let
\[
g(s) = \lim_{N \to \infty} g_N(s) = \sum_{n=1}^{\infty} ||x_{n+1}(s) - x_n(s)||_X.
\]
By the monotone convergence theorem,
\[
\left( \int_{\Omega} g(s)^p \, d\mu \right)^{1/p} = \lim_{N \to \infty} \left( \int_{\Omega} g_N(s)^p \, d\mu \right)^{1/p} < \infty.
\]
Therefore, there exists a set of measure 0, \( E \), such that for \( s \notin E \), \( g(s) < \infty \). Hence, for \( s \notin E \),
\[
\lim_{N \to \infty} x_{N+1}(s)
\]
exists because
\[
x_{N+1}(s) = x_{N+1}(s) - x_1(s) + x_1(s) = \sum_{n=1}^{N} (x_{n+1}(s) - x_n(s)) + x_1(s).
\]
Thus, if \( N > M \), and \( s \) is a point where \( g(s) < \infty \),
\[
||x_{N+1}(s) - x_{M+1}(s)||_X \leq \sum_{n=M+1}^{N} ||x_{n+1}(s) - x_n(s)||_X \leq \sum_{n=M+1}^{\infty} ||x_{n+1}(s) - x_n(s)||_X
\]
which shows that \( \{x_{N+1}(s)\}_{N=1}^{\infty} \) is a Cauchy sequence. Now let

\[
x(s) \equiv \begin{cases} 
\lim_{N \to \infty} x_N(s) & \text{if } s \notin E, \\
0 & \text{if } s \in E.
\end{cases}
\]

By Theorem 5.1.2, \( x_n(\Omega) \) is separable for each \( n \). Therefore, \( x(\Omega) \) is also separable. Also, if \( f \in X' \), then

\[
f(x(s)) = \lim_{N \to \infty} f(x_N(s))
\]

if \( s \notin E \) and \( f(x(s)) = 0 \) if \( s \in E \). Therefore, \( f \circ x \) is measurable because it is the limit of the measurable functions,

\[
f \circ x_N X_{E^c}.
\]

Since \( x \) is weakly measurable and \( x(\Omega) \) is separable, Corollary 5.1.8 shows that \( x \) is strongly measurable. By Fatou’s lemma,

\[
\int_{\Omega} ||x(s) - x_N(s)||^p \, d\mu \leq \liminf_{M \to \infty} \int_{\Omega} ||x_M(s) - x_N(s)||^p \, d\mu.
\]

But if \( N \) and \( M \) are large enough with \( M > N \),

\[
\left( \int_{\Omega} ||x_M(s) - x_N(s)||^p \, d\mu \right)^{1/p} \leq \sum_{n=N}^{M} ||x_{n+1} - x_n||_p \\
\leq \sum_{n=N}^{\infty} ||x_{n+1} - x_n||_p < \varepsilon
\]

and this shows, since \( \varepsilon \) is arbitrary, that

\[
\lim_{N \to \infty} \int_{\Omega} ||x(s) - x_N(s)||^p \, d\mu = 0.
\]

It remains to show \( x \in L^p(\Omega; X) \). This follows from the above and the triangle inequality. Thus, for \( N \) large enough,

\[
\left( \int_{\Omega} ||x(s)||^p \, d\mu \right)^{1/p} \\
\leq \left( \int_{\Omega} ||x_N(s)||^p \, d\mu \right)^{1/p} + \left( \int_{\Omega} ||x(s) - x_N(s)||^p \, d\mu \right)^{1/p} \\
\leq \left( \int_{\Omega} ||x_N(s)||^p \, d\mu \right)^{1/p} + \varepsilon < \infty.
\]

This proves the lemma.

**Theorem 5.3.3** \( L^p(\Omega; X) \) is complete. Also every Cauchy sequence has a subsequence which converges pointwise.
5.3. THE SPACES $L^p(\Omega; X)$

**Proof:** If $\{x_n\}$ is Cauchy in $L^p(\Omega; X)$, extract a subsequence $\{x_{n_k}\}$ satisfying
\[ \left\|x_{n_{k+1}} - x_{n_k}\right\|_p \leq 2^{-k} \]
and apply Lemma 5.3.2. The pointwise convergence of this subsequence was established in the proof of this lemma. This proves the theorem because if a subsequence of a Cauchy sequence converges, then the Cauchy sequence must also converge.

**Observation 5.3.4** If the measure space is Lebesgue measure then you have continuity of translation in $L^p(\mathbb{R}^n; X)$ in the usual way. More generally, for $\mu$ a Radon measure on $\Omega$ a locally compact Hausdorff space, $C_c(\Omega; X)$ is dense in $L^p(\Omega; X)$. Here $C_c(\Omega; X)$ is the space of continuous $X$ valued functions which have compact support in $\Omega$. The proof of this little observation follows immediately from approximating with simple functions and then applying the appropriate considerations to the simple functions.

Clearly Fatou’s lemma and the monotone convergence theorem make no sense for functions with values in a Banach space but the dominated convergence theorem holds in this setting.

**Theorem 5.3.5** If $x$ is strongly measurable and $x_n(s) \rightarrow x(s)$ a.e. with
\[ \|x_n(s)\| \leq g(s) \text{ a.e.} \]
where $g \in L^1(\Omega)$, then $x$ is Bochner integrable and
\[ \int_{\Omega} x(s) \, d\mu = \lim_{n \rightarrow \infty} \int_{\Omega} x_n(s) \, d\mu. \]

**Proof:** $\|x_n(s) - x(s)\| \leq 2g(s)$ a.e. so by the usual dominated convergence theorem,
\[ 0 = \lim_{n \rightarrow \infty} \int_{\Omega} \|x_n(s) - x(s)\| \, d\mu. \]
Also,
\[ \int_{\Omega} \|x_n(s) - x_m(s)\| \, d\mu \]
\[ \leq \int_{\Omega} \|x_n(s) - x(s)\| \, d\mu + \int_{\Omega} \|x_m(s) - x(s)\| \, d\mu, \]
and so $\{x_n\}$ is a Cauchy sequence in $L^1(\Omega; X)$. Therefore, by Theorem 5.3.5, there exists $y \in L^1(\Omega; X)$ and a subsequence $x_{n'}$ satisfying
\[ x_{n'}(s) \rightarrow y(s) \text{ a.e. and in } L^1(\Omega; X). \]
But $x(s) = \lim_{n' \rightarrow \infty} x_{n'}(s)$ a.e. and so $x(s) = y(s)$ a.e. Hence
\[ \int_{\Omega} \|x(s)\| \, d\mu = \int_{\Omega} \|y(s)\| \, d\mu < \infty. \]
which shows that \( x \) is Bochner integrable. Finally, since the integral is linear,
\[
\left| \int_{\Omega} x(s) \, d\mu - \int_{\Omega} x_n(s) \, d\mu \right| = \left| \int_{\Omega} (x(s) - x_n(s)) \, d\mu \right| \\
\leq \int_{\Omega} \| x_n(s) - x(s) \| \, d\mu,
\]
and this last integral converges to 0. This proves the theorem.

### 5.4 Measurable Representatives

In this section consider the special case where \( X = L^1(B, \nu) \) where \((B, \mathcal{F}, \nu)\) is a \( \sigma \) finite measure space and \( x \in L^1(\Omega; X) \). Thus for each \( s \in \Omega \), \( x(s) \in L^1(B, \nu) \). In general, the map
\[
(s, t) \rightarrow x(s)(t)
\]
will not be product measurable, but one can obtain a measurable representative. This is important because it allows the use of Fubini’s theorem on the measurable representative.

By Theorem 5.2.5, there exists a sequence of simple functions, \( \{x_n\} \), of the form
\[
x_n(s) = \sum_{k=1}^{m} a_k \chi_{E_k}(s)
\]
where \( a_k \in L^1(B, \nu) \) which satisfy the conditions of Definition 5.2.3 and
\[
\|x_n - x_m\|_{L^1(\Omega; L^1(B))} \to 0 \text{ as } m, n \to \infty
\]
(5.4.19)
For such a simple function, you can assume the \( E_k \) are disjoint and then
\[
\|x_n\|_{L^1(\Omega; L^1(B))} = \sum_{k=1}^{m} \|a_k\|_{L^1(B)} \mu(E_k) = \sum_{k=1}^{m} \int_{B} |a_k| \, d\nu(E_k)
\]
\[
= \int_{\Omega} \int_{B} |a_k(t)| \, d\nu(t) \chi_{E_k}(s) \, d\mu(s)
\]
\[
= \int_{\Omega} \int_{B} |x_n| \, d\nu d\mu
\]
Also, each \( x_n \) is product measurable. Thus from 5.4.13
\[
\|x_n - x_m\|_{L^1(\Omega; L^1(B))} = \int_{\Omega} \int_{B} |x_n - x_m| \, d\nu d\mu
\]
which shows that \( \{x_n\} \) is a Cauchy sequence in \( L^1(\Omega \times B, \mu \times \lambda) \). Then there exists \( y \in L^1(\Omega \times B, \mu \times \lambda) \) and a subsequence still called \( \{x_n\} \) such that
\[
\lim_{n \to \infty} \int_{\Omega} \int_{B} |x_n - y| \, d\nu d\mu = \lim_{n \to \infty} \int_{\Omega} \|x_n - y\|_{L^1(B)} \, d\mu = \|x_n - y\|_{L^1(\Omega; L^1(B))} = 0.
\]
Now consider 5.4.19 Since \( \lim_{m \to \infty} x_m(s) = x(s) \) in \( L^1(B) \), it follows from Fatou’s lemma that
\[
\|x_n - x\|_{L^1(\Omega, L^1(B))} \leq \liminf_{m \to \infty} \|x_n - x_m\|_{L^1(\Omega, L^1(B))} < \varepsilon
\]
for all \( n \) large enough. Hence
\[
\lim_{n \to \infty} \|x_n - x\|_{L^1(\Omega, L^1(B))} = 0
\]
and so
\[
x(s) = y(s) \quad \text{in} \quad L^1(B) \quad \mu \text{ a.e. } s
\]
In particular, for a.e. \( s \), it follows that
\[
x(s)(t) = y(s, t) \quad \text{for a.e. } t.
\]

Now \( \int_{\Omega} x(s) \, d\mu \in X = L^1(B, \nu) \) so it makes sense to ask for \( (\int_{\Omega} x(s) \, d\mu)(t) \), at least \( \mu \) a.e. \( t \). To find what this is, note
\[
\left\| \int_{\Omega} x_n(s) \, d\mu - \int_{\Omega} x(s) \, d\mu \right\|_X \leq \int_{\Omega} \|x_n(s) - x(s)\|_X \, d\mu.
\]
Therefore, since the right side converges to 0,
\[
\lim_{n \to \infty} \left\| \int_{\Omega} x_n(s) \, d\mu - \int_{\Omega} x(s) \, d\mu \right\|_X = 0
\]
But
\[
\left( \int_{\Omega} x_n(s) \, d\mu \right)(t) = \int_{\Omega} x_n(s, t) \, d\mu \text{ a.e. } t.
\]
Therefore
\[
\lim_{n \to \infty} \int_{B} \left| \int_{\Omega} x_n(s, t) \, d\mu - \left( \int_{\Omega} x(s) \, d\mu \right)(t) \right| \, d\nu = 0. \tag{5.4.20}
\]
Also, since \( x_n \to y \) in \( L^1(\Omega \times B) \),
\[
0 = \lim_{n \to \infty} \int_{B} \int_{\Omega} |x_n(s, t) - y(s, t)| \, d\mu \, d\nu \geq \lim_{n \to \infty} \int_{B} \left| \int_{\Omega} x_n(s, t) \, d\mu - \int_{\Omega} y(s, t) \, d\mu \right| \, d\nu. \tag{5.4.21}
\]
From 5.4.20 and 5.4.21
\[
\int_{\Omega} y(s, t) \, d\mu = \left( \int_{\Omega} x(s) \, d\mu \right)(t) \text{ a.e. } t.
\]
This proves the following theorem.
Theorem 5.4.1  Let \( X = L^1(B) \) where \((B, \mathcal{F}, \nu)\) is a \(\sigma\) finite measure space and let \( x \in L^1(\Omega; X) \). Then there exists a measurable representative, \( y \in L^1(\Omega \times B) \), such that
\[
x(s) = y(s, \cdot) \quad \text{a.e. } s \in \Omega, \text{ the equation in } L^1(B),
\]
and
\[
\int_{\Omega} y(s, t) \, d\mu = \left( \int_{\Omega} x(s) \, d\mu \right)(t) \quad \text{a.e. } t.
\]

5.5 Vector Measures

There is also a concept of vector measures.

Definition 5.5.1  Let \((\Omega, S)\) be a set and a \(\sigma\) algebra of subsets of \(\Omega\). A mapping
\[
F : S \rightarrow X
\]
is said to be a vector measure if
\[
F(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} F(E_i)
\]
whenever \(\{E_i\}_{i=1}^{\infty}\) is a sequence of disjoint elements of \(S\). For \(F\) a vector measure,
\[
|F|(A) \equiv \sup \{ \sum_{F \in \pi(A)} ||\mu(F)|| : \pi \text{ is a partition of } A \}.
\]

This is the same definition that was given in the case where \(F\) would have values in \(\mathbb{C}\), the only difference being the fact that now \(F\) has values in a general Banach space \(X\) as the vector space of values of the vector measure. Recall that a partition of \(A\) is a finite set, \(\{F_1, \ldots, F_n\} \subseteq S\) such that \(\bigcup_{i=1}^{n} F_i = A\). The same theorem about \(|F|\) proved in the case of complex valued measures holds in this context with the same proof. For completeness, it is included here.

Theorem 5.5.2  If \(|F|(\Omega) < \infty\), then \(|F|\) is a measure on \(S\).

Proof:  Let \(E_1\) and \(E_2\) be sets of \(S\) such that \(E_1 \cap E_2 = \emptyset\) and let \(\{A_1, \ldots, A_{n_i}\} = \pi(E_i)\), a partition of \(E_i\) which is chosen such that
\[
|F|(E_i) - \varepsilon < \sum_{j=1}^{n_i} ||F(A_j)|| \quad i = 1, 2.
\]

Consider the sets which are contained in either of \(\pi(E_1)\) or \(\pi(E_2)\), it follows this collection of sets is a partition of \(E_1 \cup E_2\) which is denoted here by \(\pi(E_1 \cup E_2)\). Then by the above inequality and the definition of total variation,
\[
|F|(E_1 \cup E_2) \geq \sum_{F \in \pi(E_1 \cup E_2)} ||F(F)|| > |F|(E_1) + |F|(E_2) - 2\varepsilon,
\]
which shows that since $\varepsilon > 0$ was arbitrary,

$$|F|(E_1 \cup E_2) \geq |F|(E_1) + |F|(E_2). \quad (5.5.22)$$

Let $\{E_j\}_{j=1}^\infty$ be a sequence of disjoint sets of $\mathcal{S}$ and let $E_\infty = \bigcup_{j=1}^\infty E_j$. Then by the definition of total variation there exists a partition of $E_\infty$, $\pi(E_\infty) = \{A_1, \ldots, A_n\}$ such that

$$|F|(E_\infty) - \varepsilon < \sum_{i=1}^n \|F(A_i)\|.$$

Also,

$$A_i = \bigcup_{j=1}^\infty A_i \cap E_j$$

and so by the triangle inequality, $\|F(A_i)\| \leq \sum_{j=1}^\infty \|F(A_i \cap E_j)\|$. Therefore, by the above,

$$|F|(E_\infty) - \varepsilon < \sum_{i=1}^n \sum_{j=1}^\infty \|F(A_i \cap E_j)\| \leq \sum_{j=1}^\infty \sum_{i=1}^n \|F(A_i \cap E_j)\| \leq \sum_{j=1}^\infty |F|(E_j)$$

because $\{A_i \cap E_j\}_{i=1}^n$ is a partition of $E_j$.

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

$$|F|(\bigcup_{j=1}^\infty E_j) \leq \sum_{j=1}^\infty |F|(E_j).$$

Also, (5.5.22) implies that whenever the $E_i$ are disjoint, $|F|(\bigcup_{j=1}^n E_j) \geq \sum_{j=1}^n |F|(E_j)$. Therefore,

$$\sum_{j=1}^\infty |F|(E_j) \geq |F|(\bigcup_{j=1}^\infty E_j) \geq |F|(\bigcup_{j=1}^n E_j) \geq \sum_{j=1}^n |F|(E_j).$$

Since $n$ is arbitrary,

$$|F|(\bigcup_{j=1}^\infty E_j) = \sum_{j=1}^\infty |F|(E_j)$$

which shows that $|F|$ is a measure as claimed. This proves the theorem.

**Definition 5.5.3** A Banach space is said to have the Radon Nikodym property if whenever

$$(\Omega, \mathcal{S}, \mu)$$

is a finite measure space
$F : S \to X$ is a vector measure with $|F|(\Omega) < \infty$

$F \ll \mu$

then one may conclude there exists $g \in L^1(\Omega; X)$ such that

$$F(E) = \int_E g(s) \, d\mu$$

for all $E \in S$.

Some Banach spaces have the Radon Nikodym property and some don’t. No attempt is made to give a complete answer to the question of which Banach spaces have this property but the next theorem gives examples of many spaces which do.

**Theorem 5.5.4** Suppose $X'$ is a separable dual space. Then $X'$ has the Radon Nikodym property.

**Proof:** Let $F \ll \mu$ and let $|F|(\Omega) < \infty$ for $F : S \to X'$, a vector measure. Pick $x \in X$ and consider the map

$$E \to F(E)(x)$$

for $E \in S$. This defines a complex measure which is absolutely continuous with respect to $|F|$. Therefore, by the Radon Nikodym theorem, there exists $f_x \in L^1(\Omega, |F|)$ such that

$$F(E)(x) = \int_E f_x(s) \, d|F|.$$  \hfill (5.5.23)

**Claim:** $|f_x(s)| \leq ||x||$ for $|F|$ a.e. $s$.

**Proof of claim:** Consider the closed ball in $F$, $\overline{B(0, ||x||)}$ and let $B \equiv B(p, r)$ be an open ball contained in its complement. Let $f_x^{-1}(B) \equiv E \in S$. I want to argue that $|F|(E) = 0$ so suppose $|F|(E) > 0$. then

$$|F|(E) ||x|| \geq ||F(E)|| ||x|| \geq |F(E)(x)|$$

and so from (5.5.23),

$$\frac{1}{|F|(E)} \left| \int_E f_x(s) \, d|F| \right| \leq ||x||.$$  \hfill (5.5.24)

But on $E, |f_x(s) - p| < r$ and so

$$\frac{1}{|F|(E)} \left| \int_E f_x(s) \, d|F| - p \right| < r$$

which contradicts (5.5.23) because $B(p, r)$ was given to have empty intersection with $\overline{B(0, ||x||)}$. Therefore, $|F|(E) = 0$ as hoped. Now $\mathbb{F} \setminus \overline{B(0, ||x||)}$ can be covered by countably many such balls and so $|F|\left(\mathbb{F} \setminus \overline{B(0, ||x||)}\right) = 0$.

Denote the exceptional set of measure zero by $N_x$. By Theorem 5.1.13, X is separable. Letting $D$ be a dense, countable subset of $X$, define

$$N_1 \equiv \bigcup_{x \in D} N_x.$$
Thus 
\[ |F| (N_1) = 0. \]

For any \( E \in \mathcal{S}, x, y \in D, \) and \( a, b \in \mathbb{F}, \)
\[
\int_E f_{ax+by} (s) \, d|F| = F (E) (ax + by) = aF (E) (x) + bF (E) (y)
\]
\[
= \int_E (af_x (s) + bf_y (s)) \, d|F|. \tag{5.5.25}
\]

Since \( f \) holds for all \( E \in \mathcal{S}, \) it follows
\[
f_{ax+by} (s) = af_x (s) + bf_y (s)
\]
for \( |F| \) a.e. \( s \) and \( x, y \in D. \) Let \( \tilde{D} \) consist of all finite linear combinations of the form \( \sum_{i=1}^{m} a_i x_i \) where \( a_i \) is a rational point of \( \mathbb{F} \) and \( x_i \in D. \) If
\[
\sum_{i=1}^{m} a_i x_i \in \tilde{D},
\]
the above argument implies
\[
f_{\sum_{i=1}^{m} a_i x_i} (s) = \sum_{i=1}^{m} a_i f_{x_i} (s) \text{ a.e.}
\]

Since \( \tilde{D} \) is countable, there exists a set, \( N_2, \) with
\[
|F| (N_2) = 0
\]
such that for \( s \notin N_2,
\[
f_{\sum_{i=1}^{m} a_i x_i} (s) = \sum_{i=1}^{m} a_i f_{x_i} (s) \tag{5.5.26}
\]
whenever \( \sum_{i=1}^{m} a_i x_i \in \tilde{D}. \) Let
\[
N = N_1 \cup N_2
\]
and let
\[
h_x (s) \equiv \mathcal{X}_{N_c} (s) f_x (s)
\]
for all \( x \in \tilde{D}. \) Now for \( x \in X \) define
\[
h_x (s) \equiv \lim_{x' \to x} \{ \tilde{h}_{x'} (s) : x' \in \tilde{D} \}.
\]
This is well defined because if \( x' \) and \( y' \) are elements of \( \tilde{D}, \) the above claim and \( \text{Ref. 24} \) imply
\[
\left| \tilde{h}_{x'} (s) - \tilde{h}_{y'} (s) \right| = \left| \tilde{h}_{(x'-y')} (s) \right| \leq \| x' - y' \|.
\]
Using the dominated convergence theorem may be applied to conclude that for $x_n \to x$, with $x_n \in \tilde{D}$,
\[
\int_{E} h_{x} (s) \, d|F| = \lim_{n \to \infty} \int_{E} \hat{h}_{x_n} (s) \, d|F| = \lim_{n \to \infty} F (E) (x_n) = F (E) (x).  \tag{5.5.27}
\]
It follows from the density of $\tilde{D}$ that for all $x, y \in X$ and $a, b \in F$,
\[
|h_{x} (s)| \leq ||x||, \quad h_{ax+by} (s) = ah_{x} (s) + bh_{y} (s),  \tag{5.5.28}
\]
for all $s$ because if $s \in N$, both sides of the equation in 5.5.28 equal 0.

Let $\theta (s)$ be given by
\[
\theta (s) (x) = h_{x} (s).
\]
By 5.5.28 it follows that $\theta (s) \in X'$ for each $s$. Also
\[
\theta (s) (x) = h_{x} (s) \in L^{1} (\Omega)
\]
so $\theta (\cdot)$ is weak * measurable. Since $X'$ is separable, Theorem 5.1.12 implies that $\theta$ is strongly measurable. Furthermore, by 5.5.28,
\[
||\theta (s)|| \equiv \sup_{||x|| \leq 1} |\theta (s) (x)| \leq \sup_{||x|| \leq 1} |h_{x} (s)| \leq 1.
\]
Therefore,
\[
\int_{\Omega} ||\theta (s)|| \, d|F| < \infty
\]
so $\theta \in L^{1} (\Omega; X')$. By 5.2.10, if $E \in S$,
\[
\int_{E} h_{x} (s) \, d|F| = \int_{E} \theta (s) (x) \, d|F| = \left( \int_{E} \theta (s) \, d|F| \right) (x).  \tag{5.5.29}
\]
From 5.5.27 and 5.5.29,
\[
\left( \int_{E} \theta (s) \, d|F| \right) (x) = F (E) (x)
\]
for all $x \in X$ and therefore,
\[
\int_{E} \theta (s) \, d|F| = F (E).\]
Finally, since $F \ll \mu, |F| \ll \mu$ also and so there exists $k \in L^{1} (\Omega)$ such that
\[
|F| (E) = \int_{E} k(s) \, d\mu
\]
for all $E \in S$, by the Radon Nikodým Theorem. It follows
\[
F (E) = \int_{E} \theta (s) \, d|F| = \int_{E} \theta (s) k(s) \, d\mu.
\]
Letting $g (s) = \theta (s) k (s)$, this has proved the theorem.
Corollary 5.5.5 Any separable reflexive Banach space has the Radon Nikodym property.

It is not necessary to assume separability in the above corollary. For the proof of a more general result, consult Vector Measures by Diestal and Uhl, [14].

5.6 The Riesz Representation Theorem

The Riesz representation theorem for the spaces $L^p(\Omega; X)$ holds under certain conditions. The proof follows the proofs given earlier for scalar valued functions.

Definition 5.6.1 If $X$ and $Y$ are two Banach spaces, $X$ is isometric to $Y$ if there exists $\theta \in L(X,Y)$ such that

$$||\theta x||_Y = ||x||_X.$$  

This will be written as $X \cong Y$. The map $\theta$ is called an isometry.

The next theorem says that $L^p'(\Omega; X')$ is always isometric to a subspace of $(L^p(\Omega; X))'$ for any Banach space, $X$.

Theorem 5.6.2 Let $X$ be any Banach space and let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space. Let $p \geq 1$ and let $1/p + 1/p' = 1$. (If $p = 1$, $p' \equiv \infty$.) Then $L^p'(\Omega; X')$ is isometric to a subspace of $(L^p(\Omega; X))'$. Also, for $g \in L^p'(\Omega; X')$,

$$\sup_{||f||_p \leq 1} \left| \int_{\Omega} g(s)(f(s)) \, d\mu \right| = ||g||_{p'}.$$  

Proof: First observe that for $f \in L^p(\Omega; X)$ and $g \in L^p'(\Omega; X')$,

$$s \rightarrow g(s)(f(s))$$

is a function in $L^1(\Omega)$. (To obtain measurability, write $f$ as a limit of simple functions. Holder’s inequality then yields the function is in $L^1(\Omega)$.) Define

$$\theta : L^p'(\Omega; X') \rightarrow (L^p(\Omega; X))'$$

by

$$\theta g(f) = \int_{\Omega} g(s)(f(s)) \, d\mu.$$  

Holder’s inequality implies

$$||\theta g|| \leq ||g||_{p'}$$  

and it is also clear that $\theta$ is linear. Next it is required to show

$$||\theta g|| = ||g||.$$
This will first be verified for simple functions. Let

\[ g(s) = \sum_{i=1}^{m} c_i \mathcal{X}_{E_i}(s) \]

where \( c_i \in \mathcal{X}' \), the \( E_i \) are disjoint and

\[ \bigcup_{i=1}^{m} E_i = \Omega. \]

Then \( ||g|| \in L^{p'}(\Omega) \). Let \( \varepsilon > 0 \) be given. By the scalar Riesz representation theorem, there exists \( h \in L^{p}(\Omega) \) such that

\[ ||h||_{L^1(\Omega)} = 1 \]

and

\[ \int_{\Omega} ||g(s)||_{X'} h(s) \, d\mu \geq ||g||_{L^{p'}(\Omega; X')} - \varepsilon. \]

Now let \( d_i \) be chosen such that

\[ c_i(d_i) \geq ||c_i||_{X'} - \varepsilon / ||h||_{L^1(\Omega)} \]

and \( ||d_i||_{X} \leq 1 \). Let

\[ f(s) \equiv \sum_{i=1}^{m} d_i h(s) \mathcal{X}_{E_i}(s). \]

Thus \( f \in L^p(\Omega; X) \) and \( ||f||_{L^p(\Omega; X)} \leq 1 \). This follows from

\[ ||f||_p^p = \int_{\Omega} \sum_{i=1}^{m} ||d_i||_X^p |h(s)|^p \mathcal{X}_{E_i}(s) \, d\mu \]

\[ = \sum_{i=1}^{m} \left( \int_{E_i} |h(s)|^p \, d\mu \right) ||d_i||_X^p \leq \int_{\Omega} |h|^p \, d\mu = 1. \]

Also

\[ ||\theta g|| \geq ||\theta g(f)|| = \left| \int_{\Omega} g(s) (f(s)) \, d\mu \right| \geq \left| \int_{\Omega} \sum_{i=1}^{m} \left( ||c_i||_{X'} - \varepsilon / ||h||_{L^1(\Omega)} \right) h(s) \mathcal{X}_{E_i}(s) \, d\mu \right| \]

\[ \geq \left| \int_{\Omega} ||g(s)||_{X'} h(s) \, d\mu \right| - \varepsilon \left| \int_{\Omega} h(s) / ||h||_{L^1(\Omega)} \, d\mu \right| \]

\[ \geq ||g||_{L^{p'}(\Omega; X')} - 2\varepsilon. \]

Since \( \varepsilon \) was arbitrary,

\[ ||\theta g|| \geq ||g|| \] (5.6.31)

and from 5.6.31 this shows equality holds in 5.6.31 whenever \( g \) is a simple function.

In general, let \( g \in L^{p'}(\Omega; X') \) and let \( g_n \) be a sequence of simple functions converging to \( g \) in \( L^{p'}(\Omega; X') \). Then

\[ ||\theta g|| = \lim_{n \to \infty} ||\theta g_n|| = \lim_{n \to \infty} ||g_n|| = ||g||. \]

This proves the theorem and shows \( \theta \) is the desired isometry.
5.6. THE RIESZ REPRESENTATION THEOREM

Theorem 5.6.3 If $X$ is a Banach space and $X'$ has the Radon Nikodym property, then if $(\Omega, S, \mu)$ is a finite measure space,

$$(L^p(\Omega; X))' \cong L^{p'}(\Omega; X')$$

and in fact the mapping $\theta$ of Theorem 5.6.2 is onto.

Proof: Let $l \in (L^p(\Omega; X))'$ and define $F(E) \in X'$ by

$$F(E)(x) \equiv l(\mathcal{X}_E(\cdot) x).$$

Lemma 5.6.4 $F$ defined above is a vector measure with values in $X'$ and $|F|(\Omega) < \infty$.

Proof of the lemma: Clearly $F(E)$ is linear. Also

$$||F(E)|| = \sup_{||x|| \leq 1} ||F(E)(x)||$$

$$\leq ||l|| \sup_{||x|| \leq 1} ||\mathcal{X}_E(\cdot) x||_{L^p(\Omega; X)} \leq ||l|| \mu(E)^{1/p}.$$ 

Let $\{E_i\}_{i=1}^\infty$ be a sequence of disjoint elements of $S$ and let $E = \cup_{n<\infty} E_n$.

$$\left| F(E)(x) - \sum_{k=1}^n F(E_k)(x) \right| = \left| l(\mathcal{X}_E(\cdot) x) - \sum_{i=1}^n l(\mathcal{X}_{E_i}(\cdot) x) \right|$$

$$\leq ||l|| \left| \mathcal{X}_E(\cdot) x - \sum_{i=1}^n \mathcal{X}_{E_i}(\cdot) x \right|_{L^p(\Omega; X)}$$

$$\leq ||l|| \mu \left( \bigcup_{k>n} E_k \right)^{1/p} ||x||.$$ 

Since $\mu(\Omega) < \infty$,

$$\lim_{n \to \infty} \mu \left( \bigcup_{k>n} E_k \right)^{1/p} = 0$$

and so inequality (5.6.32) shows that

$$\lim_{n \to \infty} \left| F(E) - \sum_{k=1}^n F(E_k) \right| = 0.$$ 

To show $|F|(\Omega) < \infty$, let $\varepsilon > 0$ be given, let $\{H_1, \cdots, H_n\}$ be a partition of $\Omega$, and let $||x_i|| \leq 1$ be chosen in such a way that

$$F(H_i)(x_i) > ||F(H_i)|| - \varepsilon/n.$$
Thus

\[-\varepsilon + \sum_{i=1}^{n} ||F(H_i)|| < \sum_{i=1}^{n} l(X_{H_i} (\cdot) x_i) \leq ||l|| \sum_{i=1}^{n} X_{H_i} (\cdot) x_i \left\| \right\|_{L^p(\Omega; X)} \]

\[\leq ||l|| \left( \int_{\Omega} \sum_{i=1}^{n} X_{H_i} (s) \, d\mu \right)^{1/p} = ||l|| \mu(\Omega)^{1/p}.\]

Since \(\varepsilon > 0\) was arbitrary,

\[\sum_{i=1}^{n} ||F(H_i)|| < ||l|| \mu(\Omega)^{1/p}.\]

Since the partition was arbitrary, this shows \(|F| (\Omega) \leq ||l|| \mu(\Omega)^{1/p}\) and this proves the lemma.

Continuing with the proof of Theorem 5.6.3, note that

\[F \ll \mu.\]

Since \(X'\) has the Radon Nikodym property, there exists \(g \in L^1 (\Omega; X')\) such that

\[F (E) = \int_{E} g (s) \, d\mu.\]

Also, from the definition of \(F (E)\),

\[l \left( \sum_{i=1}^{n} x_i X_{E_i} (\cdot) \right) = \sum_{i=1}^{n} l(X_{E_i} (\cdot) x_i) \]

\[= \sum_{i=1}^{n} F (E_i) (x_i) = \sum_{i=1}^{n} \int_{E_i} g (s) (x_i) \, d\mu. \quad (5.6.33)\]

It follows from (5.6.33) that whenever \(h\) is a simple function,

\[l (h) = \int_{\Omega} g (s) (h (s)) \, d\mu. \quad (5.6.34)\]

Let

\[G_n = \{ s : ||g (s)||_{X'} \leq n \}\]

and let

\[j : L^p (G_n; X) \to L^p (\Omega; X)\]

be given by

\[jh (s) = \begin{cases} h (s) & \text{if } s \in G_n, \\ 0 & \text{if } s \notin G_n. \end{cases} \]
5.6. THE RIESZ REPRESENTATION THEOREM

Letting $h$ be a simple function in $L^p(G_n; X)$,

$$j^*l(h) = l(jh) = \int_{G_n} g(s)(h(s)) \, d\mu.$$  \hspace{1cm} (5.6.35)

Since the simple functions are dense in $L^p(G_n; X)$, and $g \in L^{p'}(G_n; X')$, it follows 

5.6.35 holds for all $h \in L^p(G_n; X)$. By Theorem 5.6.2,

$$||g||_{L^{p'}(G_n; X')} = ||j^*l||_{(L^p(G_n; X))'} \leq ||l||_{(L^p(\Omega; X))'}.$$  

By the monotone convergence theorem,

$$||g||_{L^{p'}(\Omega; X')} = \lim_{n \to \infty} ||g||_{L^{p'}(G_n; X')} \leq ||l||_{(L^p(\Omega; X))'}.$$  

Therefore $g \in L^{p'}(\Omega; X')$ and since simple functions are dense in $L^p(\Omega; X)$, 5.6.35 holds for all $h \in L^p(\Omega; X)$. Thus $l = \theta g$ and the theorem is proved because, by Theorem 5.6.2, $||l|| = ||g||$ and the mapping $\theta$ is onto because $l$ was arbitrary.

As in the scalar case, everything generalizes to the case of $\sigma$ finite measure spaces. The proof is almost identical.

**Lemma 5.6.5** Let $(\Omega, S, \mu)$ be a $\sigma$ finite measure space and let $X$ be a Banach space such that $X'$ has the Radon Nikodym property. Then there exists a measurable function, $r$ such that $r(x) > 0$ for all $x$, such that $|r(x)| < M$ for all $x$, and $\int r \, d\mu < \infty$. For

$$\Lambda \in (L^p(\Omega; X))', \ p \geq 1,$$

there exists a unique $h \in L^{p'}(\Omega; X')$, $L^\infty(\Omega; X')$ if $p = 1$ such that

$$\Lambda f = \int h(f) \, d\mu.$$  

Also $||h|| = ||\Lambda||$. ($||h|| = ||h||_{p'}$ if $p > 1$, $||h||_\infty$ if $p = 1$). Here

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

**Proof:** First suppose $r$ exists as described. Also, to save on notation and to emphasize the similarity with the scalar case, denote the norm in the various spaces by $|.|$. Define a new measure $\tilde{\mu}$, according to the rule

$$\tilde{\mu}(E) = \int_E r \, d\mu. \hspace{1cm} (5.6.36)$$

Thus $\tilde{\mu}$ is a finite measure on $S$. Now define a mapping, $\eta : L^p(\Omega; X, \mu) \to L^p(\Omega; X, \tilde{\mu})$ by

$$\eta f = r^{-\frac{1}{p}} f.$$  

Then

$$||\eta f||_{L^p(\tilde{\mu})}^p = \int |r^{-\frac{1}{p}} f|^p \, d\tilde{\mu} = ||f||_{L^p(\mu)}^p.$$
and so $\eta$ is one to one and in fact preserves norms. I claim that also $\eta$ is onto. To see this, let $g \in L^p(\Omega; X, \mu)$ and consider the function, $r^{\frac{1}{p'}} g$. Then

$$\int |r^{\frac{1}{p'}} g|^p \, d\mu = \int |g|^p \, r \, d\mu = \int |g|^p \, d\bar{\mu} < \infty$$

Thus $r^{\frac{1}{p'}} g \in L^p(\Omega; X, \mu)$ and $\eta \left( r^{\frac{1}{p'}} g \right) = g$ showing that $\eta$ is onto as claimed. Thus $\eta$ is one to one, onto, and preserves norms. Consider the diagram below which is descriptive of the situation in which $\eta^*$ must be one to one and onto.

Then for $\Lambda \in (L^p(\mu))'$, there exists a unique $\tilde{\Lambda} \in L^p(\bar{\mu})'$ such that $\eta^* \tilde{\Lambda} = \Lambda$, $\|\tilde{\Lambda}\| = \|\Lambda\|$. By the Riesz representation theorem for finite measure spaces, there exists a unique $h \in L^{p'}(\bar{\mu}) \equiv L^{p'}(\Omega; X', \bar{\mu})$ which represents $\tilde{\Lambda}$ in the manner described in the Riesz representation theorem. Thus $\|h\|_{L^{p'}(\bar{\mu})} = \|\tilde{\Lambda}\| = \|\Lambda\|$ and for all $f \in L^p(\mu)$,

$$\Lambda(f) = \eta^* \tilde{\Lambda}(f) = \int h(\eta f) \, d\bar{\mu} = \int r h \left( r^{-\frac{1}{p'}} f \right) \, d\mu$$

Now

$$\int |r^{\frac{1}{p'}} h|^p' \, d\mu = \int |h|^p' \, r \, d\mu = ||h||_{L^{p'}(\bar{\mu})}^p < \infty.$$ 

Thus $\|r^{\frac{1}{p'}} h\|_{L^{p'}(\mu)} = \|h\|_{L^{p'}(\bar{\mu})} = \|\tilde{\Lambda}\| = \|\Lambda\|$ and represents $\Lambda$ in the appropriate way. If $p = 1$, then $1/p' \equiv 0$. Now consider the existence of $r$. Since the measure space is $\sigma$ finite, there exist $\{\Omega_n\}$ disjoint, each having positive measure and their union equals $\Omega$. Then define

$$r(\omega) \equiv \sum_{n=1}^{\infty} \frac{1}{n^2} \mu(\Omega_n)^{-1} \chi_{\Omega_n}(\omega)$$

This proves the Lemma.

**Theorem 5.6.6 (Riesz representation theorem)** Let $(\Omega, S, \mu)$ be $\sigma$ finite and let $X'$ have the Radon Nikodym property. Then for

$$\Lambda \in (L^p(\Omega; X, \mu))', \ p \geq 1$$
there exists a unique $h \in L^q(\Omega, X', \mu)$, $L^\infty(\Omega, X', \mu)$ if $p = 1$ such that
\[ \Lambda f = \int h(f) \, d\mu. \]
Also $||h|| = ||\Lambda||. (||h||_p = ||h||_q$ if $p > 1$, $||h||_\infty$ if $p = 1$). Here
\[ \frac{1}{p} + \frac{1}{q} = 1. \]

Proof: The above lemma gives the existence part of the conclusion of the theorem. Uniqueness is done as before.

Corollary 5.6.7 If $X'$ is separable, then for $(\Omega, S, \mu)$ a $\sigma$ finite measure space,
\[ (L^p(\Omega; X))' \cong L^{p'}(\Omega; X'). \]

Corollary 5.6.8 If $X$ is separable and reflexive, then for $(\Omega, S, \mu)$ a $\sigma$ finite measure space,
\[ (L^p(\Omega; X))' \cong L^{p'}(\Omega; X'). \]

Corollary 5.6.9 If $X$ is separable and reflexive and $(\Omega, S, \mu)$ a $\sigma$ finite measure space, then if $p \in (1, \infty)$, then $L^p(\Omega; X)$ is reflexive.

Proof: This is just like the scalar valued case.

5.7 Pointwise Behavior Of Weakly Convergent Sequences

There is an interesting little result which relates to weak limits in $L^2(\Gamma, E)$ for $E$ a Banach space. I am not sure where to put this thing but think that this would be a good place for it. It obviously generalizes to $L^p$ spaces.

Proposition 5.7.1 Let $E$ be a Banach space and let $\{u_n\}$ be a sequence in $L^2(\Gamma, E)$ and let $G(x)$ be a weakly compact set in $E$, and $u_n(x) \in G(x)$ a.e. for each $n$. Let
\[ \limsup \{u_n(x)\} \]
be the set of all weak limits of subsequences of $\{u_n(x)\}$ and let $H(x)$ be the closure of the convex hull of $\limsup \{u_n(x)\}$. Then if $u_n \to u$ weakly in $L^2(\Gamma, E)$, then $u(x) \in H(x)$ for a.e. $x$.

Proof: Let $H = \{w \in L^2(\Gamma, E) : w(x) \in H(x)$ a.e. $\}$. Then $H$ is convex. If you have $w_i \in H$, then since each $H(x)$ is convex, it follows that $\lambda w_1(x) + (1 - \lambda) w_2(x) \in H$ for a.e. $x$ and $\lambda \in [0, 1]$. Is $H$ closed? Suppose you have $w_n \in H$ and $w_n \to w$ in $L^2(\Gamma, E)$. Then there is a subsequence such that pointwise convergence happens a.e. and so since $H$ is closed, you have $w(x) \in H$ for a.e. $x$. Hence $H$ is also weakly closed in $L^2(\Gamma, H)$. Thus if $u$ is the weak limit of $\{u_n\}$ in $L^2(\Gamma, E)$, it must be the case that $u(x) \in H(x)$ a.e. (1)

As a case of this which might be pretty interesting, suppose $G(x)$ is not just weakly compact but also convex. Then $H(x) = G(x)$ and you can say that $u(x) \in H(x)$ a.e. whenever it is a weak limit in $L^2(\Gamma, E)$ of functions $u_n$ for which $u_n(x) \in G(x)$. 
5.8 Test Functions

Here is a fundamental lemma which is of great use in weak solutions of nonlinear evolution equations.

**Lemma 5.8.1** Suppose \( g \in L^1([a,b];X) \) where \( X \) is a Banach space. Then if \( \int_a^b g(t) \phi(t) \, dt = 0 \) for all \( \phi \in C_\infty^c(a,b) \), then \( g(t) = 0 \) a.e.

**Proof:** Let \( S \) be a measurable subset of \((a,b)\) and let \( K \subseteq S \subseteq V \subseteq (a,b) \) where \( K \) is compact, \( V \) is open and \( m \left( V \setminus K \right) < \varepsilon \). Let \( K \prec h \prec V \) as in the proof of the Riesz representation theorem for positive linear functionals. Enlarging \( K \) slightly and convolving with a mollifier, it can be assumed \( h \in C_\infty^c(a,b) \). Then

\[
\left| \int_a^b X_S(t) g(t) \, dt \right| = \left| \int_a^b (X_S(t) - h(t)) g(t) \, dt \right| \\
\leq \int_a^b |X_S(t) - h(t)| \| g(t) \| \, dt \\
\leq \int_{V \setminus K} \| g(t) \| \, dt.
\]

Now let \( K_n \subseteq S \subseteq V_n \) with \( m \left( V_n \setminus K_n \right) < 2^{-n} \). Then from the above,

\[
\left| \int_a^b X_S(t) g(t) \, dt \right| \leq \int_a^b X_{V_n \setminus K_n}(t) \| g(t) \| \, dt
\]

and the integrand of the last integral converges to 0 a.e. as \( n \to \infty \) because \( \sum_n m \left( V_n \setminus K_n \right) < \infty \). By the dominated convergence theorem, this last integral converges to 0. Therefore, whenever \( S \subseteq (a,b) \),

\[
\int_a^b X_S(t) g(t) \, dt = 0.
\]

Since the endpoints have measure zero, it also follows that for any measurable \( S \), the above equation holds.

Now \( g \in L^1([a,b];X) \) and so it is measurable. Therefore, \( g([a,b]) \) is separable. Let \( D \) be a countable dense subset and let \( E \) denote the set of linear combinations of the form \( \sum a_i d_i \) where \( a_i \) is a rational point of \( \mathbb{F} \) and \( d_i \in D \). Thus \( E \) is countable. Denote by \( Y \) the closure of \( E \) in \( X \). Thus \( Y \) is a separable closed subspace of \( X \) which contains all the values of \( g \).

Now let \( S_n = g^{-1} \left( B(y_n, \|y_n\|/2) \right) \) where \( E = \{y_n\}_{n=1}^{\infty} \). Therefore, \( \cup_n S_n = g^{-1} \left( X \setminus \{0\} \right) \). This follows because if \( x \in Y \) and \( x \neq 0 \), then in \( B \left( x, \frac{\|x\|}{4} \right) \) there is a point of \( E, y_n \). Therefore, \( \|y_n\| > \frac{3}{4} \|x\| \) and so \( \frac{\|y_n\|}{2} > \frac{3}{8} \|x\| > \|x\|/4 \) so \( x \in B \left( y_n, \|y_n\|/2 \right) \). It follows that if each \( S_n \) has measure zero, then \( g(t) = 0 \) for a.e.
t. Suppose then that for some $n$, the set, $S_n$ has positive measure. Then from what was shown above,

$$\|y_n\| = \left\| \frac{1}{m(S_n)} \int_{S_n} g(t) \, dt - y_n \right\| = \left\| \frac{1}{m(S_n)} \int_{S_n} g(t) \, dt - y_n \right\|$$

$$\leq \frac{1}{m(S_n)} \int_{S_n} \|g(t) - y_n\| \, dt \leq \frac{1}{m(S_n)} \int_{S_n} \|y_n\|/2 \, dt = \|y_n\|/2$$

and so $y_n = 0$ which implies $S_n = \emptyset$, a contradiction to $m(S_n) > 0$. This contradiction shows each $S_n$ has measure zero and so as just explained, $g(t) = 0$ a.e.

5.9 Exercises

1. Show $L^1(\mathbb{R})$ is not reflexive. **Hint:** $L^1(\mathbb{R})$ is separable. What about $L^\infty(\mathbb{R})$?

2. If $f \in L^1(\mathbb{R}^n; X)$ for $X$ a Banach space, does the usual fundamental theorem of calculus work? That is, can you say $\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(t) \, dm = f(x)$ a.e.?

3. Suppose $\{f_n\}$ are measurable $X$ valued functions with respect to $(\Omega, \mathcal{F}, \mu)$ where $\mu(\Omega) < \infty$ and suppose $f_n(\omega) \to f(\omega)$ in $X$ for each $\omega$. Here $X$ is a Banach space. Show that for every $\varepsilon > 0$ there exists a measurable set of measure less than $\varepsilon, E$ such that the convergence of $f_n$ to $f$ is uniform on $E^C$. This is Egoroff’s theorem for Banach space valued functions.

4. A set of functions $S$ in $L^1(\Omega, X)$ where $(\Omega, \mathcal{F}, \mu)$ is a finite measure space is said to be uniformly integrable if for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $A \in \mathcal{F}$ and $\mu(A) < \delta$, then

$$\int_A \|f\| \, d\mu \leq \varepsilon$$

for all $f \in S$.

Show that if you have a set of functions $S$ for which they are bounded in $L^p(\Omega, X)$ for some $p > 1$, then $S$ is uniformly integrable.

5. Does the Vitali convergence theorem hold for Bochner integrable functions? If so, give a statement of the appropriate theorem and a proof. Recall what it usually says. If you have uniformly integrable functions on a finite measure space which converge pointwise, then $\int_{\Omega} f_n \, d\mu \to \int_{\Omega} f \, d\mu$.

6. For a finite measure space, show that the Vitali convergence theorem implies the dominated convergence theorem.

7. Give an example where the Vitali convergence theorem yields convergence of the integrals but the dominated convergence theorem fails to do so.
8. Suppose \( f \in L^1 (a, b; X) \) and for all \( \phi \in C^\infty_c (a, b) \), \( \int_a^b f (t) \phi' (t) \, dt = 0 \). Then there exists a constant, \( a \in X \) such that \( f (t) = a \) a.e. \textbf{Hint:} Let

\[
\psi_\phi (x) = \int_a^x [\phi (t) - \left( \int_a^b \phi (y) \, dy \right) \phi_0 (t)] \, dt, \quad \phi_0 \in C^\infty_c (a, b), \quad \int_a^b \phi_0 (x) \, dx = 1
\]

Then explain why \( \psi_\phi \in C^\infty_c (a, b) \), \( \psi'_\phi = \phi - \left( \int_a^b \phi (y) \, dy \right) \phi_0 \). Then use the assumption on \( \psi_\phi \). Next use Lemma 5.8.1. Verify that

\[
f (y) = \int_a^b f (t) \phi_0 (t) \, dt \text{ a.e.} \quad y
\]

9. Let \( f \in L^1 ([a, b], X) \). Then we say that the weak derivative of \( f \) is in \( L^1 ([a, b], X) \) if there is a function denoted as \( f' \in L^1 ([a, b], X) \) such that for all \( \phi \in C^\infty_c (a, b) \),

\[
- \int_a^b f (t) \phi' (t) \, dt = \int_a^b f' (t) \phi (t) \, dt
\]

Show that this definition is well defined. Next, using the above problems, show that if \( f, f' \in L^1 ([a, b], X) \), it follows that there is a continuous function, denoted by \( t \to \hat{f} (t) \) such that \( \hat{f} (t) = f (t) \) a.e. \( t \) and

\[
\hat{f} (t) = \hat{f} (a) + \int_0^t f' (s) \, ds
\]

Thus, unlike the classical definition of the derivative, when a function and its derivative are both in \( L^1 \), it has a representative \( \hat{f} \) which equals the function a.e. such that \( \hat{f} \) can be recovered from its derivative. Recall the well known example of this not working out which is based on the Cantor function which you should see in a real analysis course. This function had zero derivative a.e. and yet it climbed from 0 to 1 on the unit interval. Thus one could not recover it from integrating its classical derivative.

10. Say you have \( f, f' \) are in \( L^1 ([0, T], X) \). Thus there is a representative for \( f \) such that

\[
f (t) = f (0) + \int_0^t f' (s) \, ds
\]

Show that

\[
\|f (t)\|_X \leq C \left( \|f\|_{L^1([0,T],X)} + \|f'\|_{L^1([0,T],X)} \right)
\]

If \( f, f' \) are in \( L^p ([0, T], X) \) show a similar inequality holds. \textbf{Hint:} The case of \( p > 1 \) follows right away from the above. Just explain why.
11. Let $E \subseteq W \subseteq X$ where the injection map is continuous from $W$ to $X$ and compact from $E$ to $W$. Then for every $\varepsilon > 0$ there exists a constant, $C_\varepsilon$ such that for all $u \in E$,

$$\|u\|_W \leq \varepsilon \|u\|_E + C_\varepsilon \|u\|_X$$

Show that this is the case. The hypotheses mean that these $E, W, X$ are Banach spaces and the statement that the injection map is compact means that if you have a bounded set in $E$, then it is precompact in $W$. **Hint:** Suppose not. Then for some $\varepsilon > 0$, there are $u_n$

$$\|u_n\|_W \geq \varepsilon \|u_n\|_E + n \|u_n\|_X$$

Now divide both sides by $\|u_n\|_E$ and let $v_n = u_n / \|u_n\|_E$.

$$\|v_n\|_W \geq \varepsilon + n \|v_n\|_X$$

Then you have $\{v_n\}$ is bounded in $E$. Now use the compactness of the embedding. Obtain a contradiction. This is called Erling’s lemma and it is a useful result in nonlinear problems. It often comes to the rescue when you are trying to get an estimate of some sort.

12. Suppose as in the above that $E \subseteq W \subseteq X$ where the injection map is continuous from $W$ to $X$ and compact from $E$ to $W$. Let

$$S = \left\{ u \text{ such that } \|u(t)\|_E \leq R \text{ for all } t \in [a, b], \right. \left. \text{ and } \|u(s) - u(t)\|_X \leq R |t - s|^{1/q} \right\}, q > 1$$

Show that $S \subseteq C([0, T])$. Note that you know it is bounded into $E$ and there is a Holder condition into $X$. This suffices to verify that the function is actually continuous into $W$. **Hint:** You should use the previous problem. Also show that the Holder condition is valid if the weak derivatives are bounded in $L^q(a,b,X)$

13. Let $D$ be a countable dense subset of $[a,b]$. Show that if $\{u_n\} \subseteq S$ then there exists a subsequence still denoted as $\{u_k\}$ which has the property that it converges in $W$ at each point of $D$. Next show that this subsequence actually converges in $W$ at every point of $[a,b]$. This will use the Holder condition. **Hint:** This is a Cantor diagonal argument as in usual proof of Ascoli Arzela theorem follows with a routine estimate.

14. In the situation of the above problem, show that the subsequence converges uniformly to $u$ in $W$ and that $u \in C([a,b], W)$. This is an infinite dimensional version of the Ascoli Arzela theorem which is due to Simon. It is incredibly useful for dealing with nonlinearities in evolution equations and inclusions. This has shown that the $S$ in Problem is precompact in $C([a,b], W)$. This means that a subsequence exhibits strong uniform convergence in this space. This is enough to overwhelm all sorts of bad nonlinearities provided they involve lower order terms. However, to appreciate this result, you need to have a good study of Sobolev spaces or some other kind of function space where there are nice compact embedding theorems available.
15. Let $E \subseteq W \subseteq X$ where the injection map is continuous from $W$ to $X$ and compact from $E$ to $W$. Let $p \geq 1$, let $q > 1$, and define

$$S \equiv \{ u \in L^p([a, b]; E) : \text{for some } C, \|u(t) - u(s)\|_X \leq C |t - s|^{1/q}$$

and $\|u\|_{L^p([a, b]; E)} \leq R$. 

Thus $S$ is bounded in $L^p([a, b]; E)$ and Holder continuous into $X$. Pick $k$ large. Let $a = t_0 < t_1 < \cdots < t_k = b$, $t_i - t_{i-1} = (b - a)/k$. Let $\{u_n\} \subseteq S$. Define the step functions

$$\overline{u}_n(t) \equiv \sum_{i=1}^k \overline{u}_{n,i} \chi_{[t_{i-1}, t_i)}(t), \quad \overline{u}_{n,i} = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} u_n(s) \, ds.$$ 

Show that if $k$ is large enough, and $\eta > 0$,

$$\|u_n - \overline{u}_n\|_{L^p([a, b]; W)} < \frac{\eta}{4}$$

\textbf{Hint:} By Erling’s lemma,

$$\|u_n(t) - u_n(s)\|_W^p \leq \varepsilon \|u_n(t) - u_n(s)\|_E^p + C_\varepsilon \|u_n(t) - u_n(s)\|_X^p$$

$$\leq 2^{p-1} \varepsilon \left( \|u_n(t)\|_W^p + \|u_n(s)\|_W^p \right) + C_\varepsilon |t - s|^{p/q}$$

From Jensen’s inequality that

$$\|u_n(t) - \overline{u}_n(t)\|_W^p = \sum_{i=1}^k \left\| \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} (u_n(t) - u_n(s)) \, ds \right\|_W^p \chi_{[t_{i-1}, t_i)}(t)$$

$$\leq \sum_{i=1}^k \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \|u_n(t) - u_n(s)\|_W^p \, ds \chi_{[t_{i-1}, t_i)}(t)$$

Now you make use of the above inequality.

16. Show that $S$ is precompact where $S$ is the set of the previous problem. This result that $S$ is precompact in $L^p([a, b], W)$ is due to Temann or Lions. It was used to study weak solutions to the initial boundary value problems for higher dimensional Navier Stokes equations for incompressible fluids.
Chapter 6

The Derivative

6.1 Limits Of A Function

As in the case of scalar valued functions of one variable, a concept closely related to continuity is that of the limit of a function. The notion of limit of a function makes sense at points $x$, which are limit points of $D(f)$ and this concept is defined next. In all that follows ($V, \|\cdot\|$) and ($W, \|\cdot\|$) are two normed linear spaces. Recall the definition of limit point first.

**Definition 6.1.1** Let $A \subseteq W$ be a set. A point $x$, is a limit point of $A$ if $B(x, r)$ contains infinitely many points of $A$ for every $r > 0$.

**Definition 6.1.2** Let $f: D(f) \subseteq V \to W$ be a function and let $x$ be a limit point of $D(f)$. Then

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

if and only if the following condition holds. For all $\varepsilon > 0$ there exists $\delta > 0$ such that if $0 < \|y - x\| < \delta$, and $y \in D(f)$ then,

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

**Theorem 6.1.3** If $\lim_{y \to x} f(y) = L$ and $\lim_{y \to z} f(y) = L_1$, then $L = L_1$.

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

$$\|f(y) - L\| < \varepsilon, \|f(y) - L_1\| < \varepsilon.$$ 

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

$$\|L - L_1\| \leq \|L - f(y)\| + \|f(y) - L_1\| < \varepsilon + \varepsilon = 2\varepsilon.$$ 

Since $\varepsilon > 0$ was arbitrary, this shows $L = L_1$. □

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

135
Definition 6.1.4 If \( f(x) \in \mathbb{R} \), \( \lim_{y \to x} f(x) = \infty \) if for every number \( l \), there exists \( \delta > 0 \) such that whenever \( \|y - x\| < \delta \) and \( y \in D(f) \), then \( f(x) > l \). \( \lim_{y \to x} f(x) = -\infty \) if for every number \( l \), there exists \( \delta > 0 \) such that whenever \( \|y - x\| < \delta \) and \( y \in D(f) \), then \( f(x) < l \).

The following theorem is just like the one variable version of calculus.

Theorem 6.1.5 Suppose \( f : D(f) \subseteq V \to \mathbb{F}^n \). Then for \( x \) a limit point of \( D(f) \),

\[
\lim_{y \to x} f(y) = L \tag{6.1.1}
\]

if and only if

\[
\lim_{y \to x} f_k(y) = L_k \tag{6.1.2}
\]

where \( f(y) = (f_1(y), \ldots, f_p(y)) \) and \( L = (L_1, \ldots, L_p) \).

Suppose here that \( f \) has values in \( W \), a normed linear space and

\[
\lim_{y \to x} f(y) = L, \quad \lim_{y \to x} g(y) = K
\]

where \( K, L \in W \). Then if \( a, b \in \mathbb{F} \),

\[
\lim_{y \to x} (af(y) + bg(y)) = aL + bK, \tag{6.1.3}
\]

If \( W \) is an inner product space,

\[
\lim_{y \to x} (f,g)(y) = (L,K) \tag{6.1.4}
\]

If \( g \) is scalar valued with \( \lim_{y \to x} g(y) = K \),

\[
\lim_{y \to x} f(y)g(y) = LK. \tag{6.1.5}
\]

Also, if \( h \) is a continuous function defined near \( L \), then

\[
\lim_{y \to x} h \circ f(y) = h(L). \tag{6.1.6}
\]

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

Proof: Suppose \( \lim_{y \to x} f_k(y) = L_k \). Then letting \( \varepsilon > 0 \) be given there exists \( \delta > 0 \) such that if \( 0 < \|y-x\| < \delta \), it follows

\[
|f_k(y) - L_k| \leq \|f(y) - L\| < \varepsilon
\]

which verifies \( \lim_{y \to x} f_k(y) = L_k \).

Now suppose \( \lim_{y \to x} f(y) = L \) holds. Then letting \( \varepsilon > 0 \) be given, there exists \( \delta_k \) such that if \( 0 < \|y-x\| < \delta_k \), then

\[
|f_k(y) - L_k| < \varepsilon.
\]
6.1. LIMITS OF A FUNCTION

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

$$\|f(y) - L\|_{\infty} < \varepsilon$$

Any other norm on $\mathbb{F}^m$ would work out the same way because the norms are all equivalent.

Each of the remaining assertions follows immediately from the coordinate descriptions of the various expressions and the first part. However, I will give a different argument for these.

The proof of 6.1.3 is left for you. Now 6.1.4 is to be verified. Let $\varepsilon > 0$ be given. Then by the triangle inequality,

$$|(f,g) (y) - (L,K)| \leq \|(f,g) (y) - (f (y), K)\| + \|(f (y), K) - (L,K)\|$$

$$\leq \|f (y)\| \|g (y) - K\| + \|K\| \|f (y) - L\|.$$ 

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

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

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

$$|(f,g) (y) - (L,K)| \leq (1 + \|K\| + \|L\|) \|g (y) - K\| + \|f (y) - L\|.$$ 

(6.1.7)

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

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

Then letting $0 < \delta \leq \min (\delta_1, \delta_2)$, it follows from (6.1.7) that

$$|(f,g) (y) - (L,K)| < \varepsilon$$

and this proves 6.1.4.

The proof of 6.1.5 is left to you.

Consider 6.1.6. Since $h$ is continuous near $L$, it follows that for $\varepsilon > 0$ given, there exists $\eta > 0$ such that if $\|y-L\| < \eta$, then

$$\|h (y) - h (L)\| < \varepsilon$$

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

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

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

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

It only remains to verify the last assertion. Assume $\|f (y) - b\| \leq r$. It is required to show that $\|L-b\| \leq r$. If this is not true, then $\|L-b\| > r$. Consider
Find is continuous at $f(y_0)$ and so whenever $y \in D(f)$ is close enough to $x$. Thus, by the triangle inequality,

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

and so

$$r < \|L-b\| - \|f(y) - L\| \leq \|b-L\| - \|f(y) - L\|$$

$$\leq \|b-f(y)\|,$$

a contradiction to the assumption that $\|b-f(y)\| \leq r$. \[\square\]

The relation between continuity and limits is as follows.

**Theorem 6.1.6** For $f : D(f) \to W$ and $x \in D(f)$ a limit point of $D(f)$, $f$ is continuous at $x$ if and only if

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

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

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

**Example 6.1.7** Find $\lim_{(x,y) \to (3,1)} \left( \frac{x^2-y}{x^2-y^2}, y \right)$.

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

**Example 6.1.8** Find $\lim_{(x,y) \to (0,0)} \frac{xy}{x^2+y^2}$.

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

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

The concept of derivative generalizes right away to functions defined on a normed linear space. However, no attempt will be made to consider derivatives from one side or another. This is because there isn’t a well defined side. However, it is certainly the case that there are more general notions which include such things. I will present a fairly general notion of the derivative of a function which is defined on a normed vector space which has values in a normed vector space.

In what follows, \( X, Y \) will denote normed vector spaces. Recall that \( \mathcal{L}(X,Y) \) will denote the bounded linear transformations from \( X \) to \( Y \).

Let \( U \) be an open set in \( X \), and let \( f : U \to Y \) be a function.

**Definition 6.2.1** A function \( g \) is \( o(v) \) if

\[
\lim_{||v|| \to 0} \frac{g(v)}{||v||} = 0 \tag{6.2.8}
\]

A function \( f : U \to Y \) is differentiable at \( x \in U \) if there exists a linear transformation \( L \in \mathcal{L}(X,Y) \) such that

\[
f(x + v) = f(x) + Lv + o(v)
\]

This linear transformation \( L \) is the definition of \( Df(x) \). This derivative is often called the Frechet derivative.

In finite dimensions, the question whether a given function is differentiable is independent of the norm used on the finite dimensional vector space. That is, a function is differentiable with one norm if and only if it is differentiable with another norm. This is because all norms are equivalent on a finite dimensional space.

The definition (6.2.8) means the error,

\[
f(x + v) - f(x) - Lv
\]

converges to 0 faster than \( ||v|| \). Thus the above definition is equivalent to saying

\[
\lim_{||v|| \to 0} \frac{||f(x + v) - f(x) - Lv||}{||v||} = 0 \tag{6.2.9}
\]

or equivalently,

\[
\lim_{y \to x} \frac{||f(y) - f(x) - Df(x)(y - x)||}{||y - x||} = 0. \tag{6.2.10}
\]

The symbol \( o(v) \) should be thought of as an adjective. Thus, if \( t \) and \( k \) are constants,

\[
o(v) = o(v) + o(v), \quad o(tv) = o(v), \quad ko(v) = o(v)
\]

and other similar observations hold.

**Theorem 6.2.2** The derivative is well defined.
Proof: First note that for a fixed vector \( v \), \( o(tv) = o(t) \). This is because
\[
\lim_{t \to 0} \frac{o(tv)}{|t|} = \lim_{t \to 0} \frac{|v|}{||tv||} o(tv) = 0
\]
Now suppose both \( L_1 \) and \( L_2 \) work in the above definition. Then let \( v \) be any vector and let \( t \) be a real scalar which is chosen small enough that \( tv + x \in U \). Then
\[
f(x + tv) = f(x) + L_1 tv + o(tv), \quad f(x + tv) = f(x) + L_2 tv + o(tv).
\]
Therefore, subtracting these two yields \( (L_2 - L_1)(tv) = o(tv) = o(t) \). Therefore, dividing by \( t \) yields \( (L_2 - L_1)(v) = \frac{o(t)}{t} \). Now let \( t \to 0 \) to conclude that \( (L_2 - L_1)(v) = 0 \). Since this is true for all \( v \), it follows \( L_2 = L_1 \). ■

Lemma 6.2.3 Let \( f \) be differentiable at \( x \). Then \( f \) is continuous at \( x \) and in fact, there exists \( K > 0 \) such that whenever \( ||v|| \) is small enough,
\[
||f(x + v) - f(x)|| \leq K ||v||
\]
Also if \( f \) is differentiable at \( x \), then
\[
o(||f(x + v) - f(x)||) = o(v)
\]
Proof: From the definition of the derivative,
\[
f(x + v) - f(x) = Df(x)v + o(v).
\]
Let \( ||v|| \) be small enough that \( \frac{o(||v||)}{||v||} < 1 \) so that \( ||o(v)|| \leq ||v|| \). Then for such \( v \),
\[
||f(x + v) - f(x)|| \leq ||Df(x)v|| + ||v|| \leq (||Df(x)|| + 1)||v||
\]
This proves the lemma with \( K = ||Df(x)|| + 1 \). Recall the operator norm discussed in Definition 3.2.8.

The last assertion is implied by the first as follows. Define
\[
h(v) = \begin{cases} 
\frac{o(||f(x + v) - f(x)||)}{||f(x + v) - f(x)||} & \text{if } ||f(x + v) - f(x)|| \neq 0 \\
0 & \text{if } ||f(x + v) - f(x)|| = 0 
\end{cases}
\]
Then \( \lim_{||v|| \to 0} h(v) = 0 \) from continuity of \( f \) at \( x \) which is implied by the first part. Also from the above estimate,
\[
\left\| o\left(\frac{||f(x + v) - f(x)||}{||v||}\right) \right\| ||v|| = h(v) ||f(x + v) - f(x)|| ||v|| \leq h(v) (||Df(x)|| + 1)
\]
This establishes the second claim. ■

Here \( ||Df(x)|| \) is the operator norm of the linear transformation \( Df(x) \).
6.3 The Chain Rule

With the above lemma, it is easy to prove the chain rule.

**Theorem 6.3.1 (The chain rule)** Let $U$ and $V$ be open sets $U \subseteq X$ and $V \subseteq Y$. Suppose $f : U \to V$ is differentiable at $x \in U$ and suppose $g : V \to \mathbb{R}^q$ is differentiable at $f(x) \in V$. Then $g \circ f$ is differentiable at $x$ and

$$D(g \circ f)(x) = Dg(f(x)) \cdot Df(x).$$

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

$$g(f(x + v)) - g(f(x)) =$$

$$Dg(f(x)) (f(x + v) - f(x)) + o(f(x + v) - f(x))$$

$$= Dg(f(x)) [Df(x)v + o(v)] + o(f(x + v) - f(x))$$

$$= Dg(f(x)) Df(x)v + o(f(x + v) - f(x))$$

$$= Dg(f(x)) Df(x)v + o(v) \quad (6.3.11)$$

By Lemma 6.3.1. From the definition of the derivative $D(g \circ f)(x)$ exists and equals $Dg(f(x)) \cdot Df(x)$. ■

6.4 The Derivative Of A Compact Mapping

Here is a little definition about compact mappings. It turns out that if you have a differentiable mapping which is also compact, then the derivative must also be compact.

**Definition 6.4.1** Let $C \in \mathcal{L}(X,Y)$. It is said to be compact if it takes bounded sets to precompact sets. If $f$ is a function defined on an open subset $U$ of $X$, then $f$ is called compact if $f$ (bounded set) = (precompact).

**Theorem 6.4.2** Let $f : U \subseteq X \to Y$ where $f$ takes bounded sets to precompact sets. Then $Df(x)$ also takes bounded sets in $X$ to precompact sets in $Y$.

**Proof:** If this is not so, then there exists a bounded set $B$ in $X$ and for some $\varepsilon > 0$ a sequence of points $Df(x) b_n$ such that all these points are further apart than $\varepsilon$. Without loss of generality, one can assume $B = B(0, r)$, a ball. In fact, one can assume that $r > 0$ is as small as desired because if $Df(x) B(0, r)$ is precompact, then so is $Df(x) B(0, R)$, $R > r$. Just get an $\varepsilon = \frac{\varepsilon}{n}$ net $\{Df(x) x_n\}_{n=1}^N$ for $Df(x) B(0, r)$ and consider $\{\frac{r}{n} Df(x) x_n\}_{n=1}^N$. Such a net $\cup_n B(Df(x) x_n, \varepsilon \frac{r}{n})$ covers $Df(x) B(0, r)$ and so $\cup_n B(Df(x) x_n, \varepsilon)$ covers $Df(x) B(0, r)$. 


CHAPTER 6. THE DERIVATIVE

Choose \( r \) very small so that \( r < \varepsilon/4 \) and

\[
f(x + x_n) - f(x) = Df(x)x_n + o(x_n), \quad \|o(x_n)\| < \|x_n\|
\]

and there are infinitely many \( Df(x)x_n \) further apart than \( \varepsilon, x_n \in B(0, r) \). Then consider \( B(x, r) \) and \( \{f(x + x_n)\}_{n=1}^{\infty} \).

\[
\|f(x + x_n) - f(x + x_m)\| \geq \|Df(x)x_n - Df(x)x_m\| - \|o(x_n) - o(x_m)\|
\]

\[
\geq \varepsilon - 2\varepsilon/4 = \varepsilon/2
\]

contradicting the assertion that \( f \) takes bounded sets to precompact sets. ■

6.5 The Matrix Of The Derivative

The case of most interest here is the only one I will discuss. It is the case where \( X = \mathbb{R}^n \) and \( Y = \mathbb{R}^m \), the function being defined on an open subset of \( \mathbb{R}^n \). Of course this all generalizes to arbitrary vector spaces and one considers the matrix taken with respect to various bases. As above, \( f \) will be defined and differentiable on an open set \( U \subseteq \mathbb{R}^n \).

The matrix of \( Df(x) \) is the matrix having the \( i \)th column equal to \( Df(x)e_i \) and so it is only necessary to compute this. Let \( t \) be a small real number such that both

\[
\frac{f(x + te_i) - f(x) - Df(x)(te_i)}{t} = o(t)
\]

Therefore,

\[
\frac{f(x + te_i) - f(x)}{t} = Df(x)(e_i) + o(t)
\]

The limit exists on the right and so it exists on the left also. Thus

\[
\frac{\partial f(x)}{\partial x_i} \equiv \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t} = Df(x)(e_i)
\]

and so the matrix of the derivative is just the matrix which has the \( i \)th column equal to the \( i \)th partial derivative of \( f \). Note that this shows that whenever \( f \) is differentiable, it follows that the partial derivatives all exist. It does not go the other way however as discussed later.

**Theorem 6.5.1** Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) and suppose \( f \) is differentiable at \( x \). Then all the partial derivatives \( \frac{\partial f(x)}{\partial x_j} \) exist and if \( Jf(x) \) is the matrix of the linear transformation, \( Df(x) \) with respect to the standard basis vectors, then the \( ij \)th entry is given by \( \frac{\partial f_i(x)}{\partial x_j} \) also denoted as \( f_{i,j} \) or \( f_{i,x_j} \). It is the matrix whose \( i \)th column is

\[
\frac{\partial f(x)}{\partial x_i} \equiv \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}.
\]
Of course there is a generalization of this idea called the directional derivative.

**Definition 6.5.2** In general, the symbol

\[ D_v f(x) \]

is defined by

\[ \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \]

where \( t \in \mathbb{F} \). In case \(|v| = 1\) and the norm is the standard Euclidean norm, this is called the directional derivative. More generally, with no restriction on the size of \( v \) and in any linear space, it is called the Gateaux derivative. \( f \) is said to be Gateaux differentiable at \( x \) if there exists \( D_v f(x) \) such that

\[ \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} = D_v f(x) \]

where \( v \to D_v f(x) \) is linear. Thus we say it is Gateaux differentiable if the Gateaux derivative exists for each \( v \) and \( v \to D_v f(x) \) is linear.\[1\]

What if all the partial derivatives of \( f \) exist? Does it follow that \( f \) is differentiable? Consider the following function, \( f : \mathbb{R}^2 \to \mathbb{R} \),

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

Then from the definition of partial derivatives,

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

and

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

However \( f \) is not even continuous at \((0, 0)\) which may be seen by considering the behavior of the function along the line \( y = x \) and along the line \( x = 0 \). By Lemma \[2\] this implies \( f \) is not differentiable. Therefore, it is necessary to consider the correct definition of the derivative given above if you want to get a notion which generalizes the concept of the derivative of a function of one variable in such a way as to preserve continuity whenever the function is differentiable.

---

\[1\] René Gateaux was one of the many young French men killed in world war I. This derivative is named after him, but it developed naturally from ideas used in the calculus of variations which were due to Euler and Lagrange back in the 1700’s.
6.6 A Mean Value Inequality

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

**Lemma 6.6.1** Let \( Y \) be a normed vector space and suppose \( h : [0,1] \to Y \) is differentiable and satisfies
\[
\|h'(t)\| \leq M.
\]
Then
\[
\|h(1) - h(0)\| \leq M.
\]

**Proof:** Let \( \varepsilon > 0 \) be given and let
\[
S \equiv \{ t \in [0,1] : \text{ for all } s \in [0,t], \|h(s) - h(0)\| \leq (M+\varepsilon) s \}
\]
Then \( 0 \in S. \) Let \( t = \sup S. \) Then by continuity of \( h \) it follows
\[
\|h(t) - h(0)\| = (M+\varepsilon) t \quad (6.6.12)
\]
Suppose \( t < 1. \) Then there exist positive numbers, \( h_k \) decreasing to 0 such that
\[
\|h(t+h_k) - h(0)\| > (M+\varepsilon) (t+h_k)
\]
and now it follows from \( 6.6.12 \) and the triangle inequality that
\[
\|h(t+h_k) - h(t)\| + \|h(t) - h(0)\| = \|h(t+h_k) - h(t)\| + (M+\varepsilon) t > (M+\varepsilon) (t+h_k)
\]
and so
\[
\|h(t+h_k) - h(t)\| > (M+\varepsilon) h_k
\]
Now dividing by \( h_k \) and letting \( k \to \infty \)
\[
\|h'(t)\| \geq M + \varepsilon,
\]
a contradiction. Thus \( t = 1. \)

**Theorem 6.6.2** Suppose \( U \) is an open subset of \( X \) and \( f : U \to Y \) has the property that \( Df(x) \) exists for all \( x \) in \( U \) and that, \( x + t(y-x) \in U \) for all \( t \in [0,1] \). (The line segment joining the two points lies in \( U \).) Suppose also that for all points on this line segment,
\[
\|Df(x+t(y-x))\| \leq M.
\]
Then
\[
\|f(y) - f(x)\| \leq M \|y - x\|.
\]
Proof: Let \( h(t) \equiv f(x + t(y - x)) \).

Then by the chain rule,

\[
 h'(t) = Df(x + t(y - x))(y - x)
\]

and so

\[
 ||h'(t)|| = ||Df(x + t(y - x))(y - x)|| 
\leq M ||y - x||
\]

by Lemma 6.6.1.

\[
 ||h(1) - h(0)|| = ||f(y) - f(x)|| \leq M ||y - x||. \]

Here is a little result which will help to tie the case of \( \mathbb{R}^n \) in to the abstract theory presented for arbitrary spaces.

**Theorem 6.6.3** Let \( X \) be a normed vector space having basis \( \{v_1, \ldots, v_n\} \) and let \( Y \) be another normed vector space having basis \( \{w_1, \ldots, w_m\} \). Let \( U \) be an open set in \( X \) and let \( f: U \to Y \) have the property that the Gateaux derivatives,

\[
 D_{v_k}f(x) \equiv \lim_{t \to 0} \frac{f(x + tv_k) - f(x)}{t}
\]

exist and are continuous functions of \( x \). Then \( Df(x) \) exists and

\[
 Df(x)v = \sum_{k=1}^{n} D_{v_k}f(x)a_k
\]

where

\[
 v = \sum_{k=1}^{n} a_kv_k.
\]

Furthermore, \( x \to Df(x) \) is continuous; that is

\[
 \lim_{y \to x} ||Df(y) - Df(x)|| = 0.
\]

**Proof:** Let \( v = \sum_{k=1}^{n} a_kv_k \). Then

\[
 f(x + v) - f(x) = f(x + \sum_{k=1}^{n} a_kv_k) - f(x).
\]

Then letting \( \sum_{k=1}^{0} a_kv_k = 0 \), \( f(x + v) - f(x) \) is given by

\[
 \sum_{k=1}^{n} \left[ f(x + \sum_{j=1}^{k} a_jv_j) - f(x + \sum_{j=1}^{k-1} a_jv_j) \right]
\]
\[
= \sum_{k=1}^{n} [f(x + a_kv_k) - f(x)] + \\
\sum_{k=1}^{n} \left[ \left( f(x + a_jv_j) - f(x + a_kv_k) \right) - \left( f(x + \sum_{j=1}^{k-1} a_jv_j) - f(x) \right) \right]
\]

Consider the \( k \)th term in (6.6.13). Let
\[
h(t) \equiv f\left(x + \sum_{j=1}^{k-1} a_jv_j + ta_kv_k \right) - f(x + ta_kv_k)
\]
for \( t \in [0,1] \). Then
\[
h'(t) = a_k \lim_{h \to 0} \frac{1}{a_kh} \left( f\left(x + \sum_{j=1}^{k-1} a_jv_j + (t + h) a_kv_k \right) - f(x + (t + h) a_kv_k) \right)
\]
\[
- \left( f\left(x + \sum_{j=1}^{k-1} a_jv_j + ta_kv_k \right) - f(x + ta_kv_k) \right)
\]
and this equals
\[
\left( Dv_k f\left(x + \sum_{j=1}^{k-1} a_jv_j + ta_kv_k \right) - Dv_k f(x + ta_kv_k) \right) a_k \quad (6.6.14)
\]

Now without loss of generality, it can be assumed that the norm on \( X \) is given by
\[
||v|| \equiv \max \left\{ |a_k| : v = \sum_{j=1}^{n} a_kv_k \right\}
\]
because this is a finite dimensional space, all norms on \( X \) are equivalent. Therefore, from (6.6.13) and the assumption that the Gateaux derivatives are continuous,
\[
||h'(t)|| = \left\| \left( Dv_k f\left(x + \sum_{j=1}^{k-1} a_jv_j + ta_kv_k \right) - Dv_k f(x + ta_kv_k) \right) a_k \right\| \leq \varepsilon |a_k| \leq \varepsilon ||v||
\]
provided \( ||v|| \) is sufficiently small. Since \( \varepsilon \) is arbitrary, it follows from Lemma (6.6.1) the expression in (6.6.13) is \( o(||v||) \) because this expression equals a finite sum of terms of the form \( h(1) - h(0) \) where \( ||h'(t)|| \leq \varepsilon ||v|| \) whenever \( ||v|| \) is small enough. Thus
\[
f(x + v) - f(x) = \sum_{k=1}^{n} [f(x + a_kv_k) - f(x)] + o(||v||)
\]
= \sum_{k=1}^{n} D_{v_k}f(x) a_k + \sum_{k=1}^{n} [f(x + a_k v_k) - f(x) - D_{v_k}f(x) a_k] + o(v).

Consider the $k^{th}$ term in the second sum.

\[ f(x + a_k v_k) - f(x) - D_{v_k}f(x) a_k = a_k \left( \frac{f(x + a_k v_k) - f(x)}{a_k} - D_{v_k}f(x) \right) \]

where the expression in the parentheses converges to 0 as $a_k \to 0$. Thus whenever $||v||$ is sufficiently small,$^\star$

\[ ||f(x + a_k v_k) - f(x) - D_{v_k}f(x) a_k|| \leq \varepsilon |a_k| \leq \varepsilon ||v|| \]

which shows the second sum is also $o(v)$. Therefore,

\[ f(x + v) - f(x) = \sum_{k=1}^{n} D_{v_k}f(x) a_k + o(v). \]

Defining

\[ Df(x)v = \sum_{k=1}^{n} D_{v_k}f(x) a_k \]

where $v = \sum_k a_k v_k$, it follows $Df(x) \in \mathcal{L}(X,Y)$ and is given by the above formula. It remains to verify $x \to Df(x)$ is continuous.

\[ \| (Df(x) - Df(y)) v \| \]
\[ \leq \sum_{k=1}^{n} \| (D_{v_k}f(x) - D_{v_k}f(y)) a_k \| \]
\[ \leq \max \{ |a_k|, k = 1, \cdots, n \} \sum_{k=1}^{n} \| D_{v_k}f(x) - D_{v_k}f(y) \| \]
\[ = ||v|| \sum_{k=1}^{n} \| D_{v_k}f(x) - D_{v_k}f(y) \| \]

(Note that $||v|| = \max \{ |a_k|, k = 1, \cdots, n \}$ where $v = \sum_k a_k v_k$) and so

\[ \| Df(x) - Df(y) \| \leq \sum_{k=1}^{n} \| D_{v_k}f(x) - D_{v_k}f(y) \| \]

which proves the continuity of $Df$ because of the assumption the Gateaux derivatives are continuous.$^\blacksquare$

In particular, if $D_{v_k}f(x)$ exist and are continuous functions of $x$, this shows that $f$ is Gateaux differentiable and in fact the Gateaux derivatives are continuous. The following gives the corresponding result for functions defined on infinite dimensional spaces.
Theorem 6.6.4 Suppose \( f : U \to Y \) where \( U \) is an open set in \( X \), a normed linear space. Suppose that \( f \) is Gateaux differentiable on \( U \) and that the Gateaux derivative is continuous on an open set containing \( x \). Then \( f \) is Frechet differentiable at \( x \).

Proof: Denote by \( G(x) \in L(X,Y) \) the Gateaux derivative. Thus

\[
G(x)v = \lim_{\lambda \to 0} \frac{f(x + \lambda v) - f(x)}{\lambda}
\]

It is desired to show that \( G(x) = Df(x) \). Since \( G \) is continuous, one can obtain

\[
f(x + v) - f(x) = \int_0^1 G(x + tv)v dt
\]

where this is the ordinary Riemann integral.

\[
\left\| f(x + v) - f(x) - G(x)v \right\| \leq \int_0^1 \|G(x + tv)v - G(x)v\| dt \|v\|
\]

which is small provided \( \|v\| \) is sufficiently small. Thus \( G(x) = Df(x) \) as hoped. 

Recall the following.

Lemma 6.6.5 Let \( \|x\| = \sup_{\|y^*\| \leq 1} |\langle y^*, x \rangle| \).

Proof: Let \( f(kx) = k\|x\| \). Then

\[
\sup_{\|kx\| \leq 1} |\langle f, x \rangle| = \sup_{\|k\| \leq 1/\|x\|} |k|\|x\| = 1
\]

Then by Hahn Banach theorem, there is \( y^* \in X' \) which extends \( f \) and \( \|y^*\| \leq 1 \). Then

\[
\|x\| \geq \sup_{\|z^*\| \leq 1} |\langle z^*, x \rangle| \geq |\langle y^*, x \rangle| = \|x\|
\]

One does not need continuity of \( G \) near \( x \). It suffices to have continuity at \( x \). Let \( y^* \in Y' \). Then by the mean value theorem,

\[
\langle y^*, f(x + v) \rangle - \langle y^*, f(x) \rangle = \langle y^*, G(x + tv)v \rangle, \ t \in [0,1]
\]

Then

\[
\frac{1}{\|v\|} \|f(x + v) - f(x) - G(x)v\| = \frac{1}{\|v\|} \sup_{\|y^*\| \leq 1} |\langle y^*, f(x + v) - f(x) - G(x)v \rangle|
\]

\[
= \frac{1}{\|v\|} \sup_{\|y^*\| \leq 1} |\langle y^*, G(x + tv)v - G(x)v \rangle| \leq \sup_{|t| \leq 1} \|G(x + tv) - G(x)\|_{L(X,Y)}
\]

which converges to 0 as \( \|v\| \to 0 \) thanks to continuity of \( G \) at \( x \). This proves the following.
Theorem 6.6.6 Suppose \( f : U \to Y \) where \( U \) is an open set in \( X \), a normed linear space. Suppose that \( f \) is Gateaux differentiable on \( U \) and that the Gateaux derivative is continuous at \( x \). Then \( f \) is Frechet differentiable at \( x \) and \( Df(x)v = D_xf(x) \).

Example 6.6.7 Let \( X = C^2_0(\Omega) \) where \( \Omega \) is a bounded open set in \( \mathbb{R}^n \) consisting of those functions which are twice continuously differentiable and vanish near \( \partial \Omega \). The norm will be
\[ \|u\|_X = \|u\|_{\infty} + \max\{\|u_i\|_{\infty}, i\} + \max\{\|u_{ij}\|_{\infty}, i,j\} \]
Then let \( f : X \to \mathbb{R} \) be defined by
\[ f(u) = \frac{1}{2} \int_{\Omega} \nabla u \cdot \nabla vdx \]
Show \( f \) is differentiable at \( u \in X \).

Consider the Gateaux differentiability.
\[ \lim_{t \to 0} \frac{f(u + tv) - f(u)}{t} = \lim_{t \to 0} \frac{t \int_{\Omega} \nabla u \cdot \nabla vdx}{t} + \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla v \]
so it converges to
\[ \int_{\Omega} \nabla u \cdot \nabla vdx = - \int_{\Omega} \Delta uvdx \]
the last step comes from the divergence theorem. Clearly \( v \to - \int_{\Omega} \Delta uvdx \) is linear and \( \mathbb{R} \) valued.
\[ \left| - \int_{\Omega} \Delta uvdx \right| \leq \|v\|_X \int_{\Omega} |\Delta u| \, dx \leq \|v\|_X \, m(\Omega) \|u\|_X \]
Thus this appears to be in \( L(X, \mathbb{R}) \). This also shows that,
\[ \sup_{\|v\| \leq 1} |D_vf(u) - D_vf(\tilde{u})| \leq m(\Omega) \|u - \tilde{u}\|_X \]
and so \( u \to D_v(f)(u) \) is continuous as a map from \( X \) to \( L(X, \mathbb{R}) \) so it seems that this is a differentiable function and
\[ Df(u)(v) = - \int_{\Omega} \Delta uvdx \]

Definition 6.6.8 Let \( f : U \to Y \) where \( U \) is an open set in \( X \). Then \( f \) is called \( C^1(U) \) if it Gateaux differentiable and the Gateaux derivative is continuous on \( U \).

As shown, this implies \( f \) is differentiable and the Gateaux derivative is the Frechet derivative. It is good to keep in mind the following simple example or variations of it.

Example 6.6.9 Define
\[ f(x) = \begin{cases} x^2 \sin \left( \frac{1}{x} \right) & x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} \]
This function has the property that it is differentiable everywhere but is not \( C^1(\mathbb{R}) \). In fact the derivative fails to be continuous at 0.
6.7 Higher Order Derivatives

If \( f : U \subseteq X \to Y \) for \( U \) an open set, then

\[
x \to Df(x)
\]

is a mapping from \( U \) to \( \mathcal{L}(X,Y) \), a normed vector space. Therefore, it makes perfect sense to ask whether this function is also differentiable.

**Definition 6.7.1** The following is the definition of the second derivative.

\[
D^2f(x) \equiv D(Df(x)).
\]

Thus,

\[
Df(x + v) - Df(x) = D^2f(x)v + o(v).
\]

This implies

\[
D^2f(x) \in \mathcal{L}(X, \mathcal{L}(X,Y)), \quad D^2f(x)(u)(v) \in Y,
\]

and the map

\[
(u, v) \to D^2f(x)(u)(v)
\]

is a bilinear map having values in \( Y \). In other words, the two functions,

\[
u \to D^2f(x)(u)(v), \quad v \to D^2f(x)(u)(v)
\]

are both linear.

The same pattern applies to taking higher order derivatives. Thus,

\[
D^3f(x) \equiv D(D^2f(x))
\]

and \( D^3f(x) \) may be considered as a trilinear map having values in \( Y \). In general \( D^k f(x) \) may be considered a \( k \) linear map. This means the function

\[
(u_1, \cdots, u_k) \to D^k f(x)(u_1) \cdots (u_k)
\]

has the property

\[
u_j \to D^k f(x)(u_1) \cdots (u_j) \cdots (u_k)
\]

is linear.

Also, instead of writing

\[
D^2f(x)(u)(v), \quad \text{or} \quad D^3f(x)(u)(v)(w)
\]

the following notation is often used.

\[
D^2f(x)(u, v) \quad \text{or} \quad D^3f(x)(u, v, w)
\]

with similar conventions for higher derivatives than 3. Another convention which is often used is the notation

\[
D^k f(x) v^k
\]
instead of 

\[ D^k f(x)(v, \cdots, v). \]

Note that for every \( k \), \( D^k f \) maps \( U \) to a normed vector space. As mentioned above, \( Df(x) \) has values in \( \mathcal{L}(X,Y) \), \( D^2 f(x) \) has values in \( \mathcal{L}(X, \mathcal{L}(X,Y)) \), etc. Thus it makes sense to consider whether \( D^k f \) is continuous. This is described in the following definition.

**Definition 6.7.2** Let \( U \) be an open subset of \( X \), a normed vector space, and let \( f : U \to Y \). Then \( f \) is \( C^k(U) \) if \( f \) and its first \( k \) derivatives are all continuous. Also, \( D^k f(x) \) when it exists can be considered a \( Y \)-valued multi-linear function. Sometimes these are called tensors in case \( f \) has scalar values.

### 6.8 The Derivative And The Cartesian Product

There are theorems which can be used to get differentiability of a function based on existence and continuity of the partial derivatives. A generalization of this was given above. Here a function defined on a product space is considered. It is very much like what was presented above and could be obtained as a special case but to reinforce the ideas, I will do it from scratch because certain aspects of it are important in the statement of the implicit function theorem.

The following is an important abstract generalization of the concept of partial derivative presented above. Instead of taking the derivative with respect to one variable, it is taken with respect to several but not with respect to others. This vague notion is made precise in the following definition. First here is a lemma.

**Lemma 6.8.1** Suppose \( U \) is an open set in \( X \times Y \). Then the set, \( U_y \) defined by

\[ U_y \equiv \{ x \in X : (x,y) \in U \} \]

is an open set in \( X \). Here \( X \times Y \) is a finite dimensional vector space in which the vector space operations are defined componentwise. Thus for \( a, b \in \mathbb{F} \),

\[ a(x_1, y_1) + b(x_2, y_2) = (ax_1 + bx_2, ay_1 + by_2) \]

and the norm can be taken to be

\[ \|(x,y)\| \equiv \max (\|x\|, \|y\|) \]

**Proof:** In finite dimensions it doesn’t matter how this norm is defined because all are equivalent. It obviously satisfies most axioms of a norm. The only one which is not obvious is the triangle inequality. I will show this now.

\[
\|(x,y) + (x_1,y_1)\| \equiv \|(x + x_1, y + y_1)\| \equiv \max (\|x + x_1\|, \|y + y_1\|) \\
\leq \max (\|x\| + \|x_1\|, \|y\| + \|y_1\|) \\
\leq \max (\|x\|, \|y\|) + \max (\|x_1\|, \|y_1\|) \\
\equiv \|(x,y)\| + \|(x_1,y_1)\|
\]
Let \( x \in U_y \). Then \((x, y) \in U\) and so there exists \( r > 0 \) such that

\[
B((x, y), r) \subseteq U.
\]

This says that if \((u, v) \in X \times Y\) such that \(||(u, v) - (x, y)|| < r\), then \((u, v) \in U\).

Thus if

\[
||(u, y) - (x, y)|| = ||u - x|| < r,
\]

then \((u, y) \in U\). This has just said that \( B(x, r) \), the ball taken in \( X \) is contained in \( U_y \).

Or course one could also consider

\[
U_x \equiv \{ y : (x, y) \in U \}
\]

in the same way and conclude this set is open in \( Y \). Also, the generalization to many factors yields the same conclusion. In this case, for \( x \in \prod_{i=1}^n X_i \), let

\[
||x|| \equiv \max_i (||x_i||_{X_i} : x = (x_1, \cdots, x_n))
\]

Then a similar argument to the above shows this is a norm on \( \prod_{i=1}^n X_i \). Consider the triangle inequality.

\[
||(x_1, \cdots, x_n) + (y_1, \cdots, y_n)|| = \max_i (||x_i + y_i||_{X_i}) \leq \max_i (||x_i||_{X_i} + ||y_i||_{X_i})
\]

\[
\leq \max_i (||x_i||_{X_i}) + \max_i (||y_i||_{X_i})
\]

**Corollary 6.8.2** Let \( U \subseteq \prod_{i=1}^n X_i \) be an open set and let

\[
U_{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)} \equiv \{ x \in X_i : (x_1, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n) \in U \}
\]

Then \( U_{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)} \) is an open set in \( X_i \).

**Proof:** Let \( z \in U_{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)} \). Then \((x_1, \cdots, x_{i-1}, z, x_{i+1}, \cdots, x_n) \equiv x \in U\) by definition. Therefore, since \( U \) is open, there exists \( r > 0 \) such that \( B(x, r) \subseteq U \). It follows that for \( B(z, r) \) denoting the ball in \( X_i \) it follows that \( B(z, r)_{X_i} \subseteq U_{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)} \) because to say that \( ||z - w||_{X_i} < r \) is to say that

\[
||(x_1, \cdots, x_{i-1}, z, x_{i+1}, \cdots, x_n) - (x_1, \cdots, x_{i-1}, w, x_{i+1}, \cdots, x_n)|| < r
\]

and so \( w \in U_{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)} \).

Next is a generalization of the partial derivative.

**Definition 6.8.3** Let \( g : U \subseteq \prod_{i=1}^n X_i \rightarrow Y \), where \( U \) is an open set. Then the map

\[
z \rightarrow g(x_1, \cdots, x_{i-1}, z, x_{i+1}, \cdots, x_n)
\]

is a function from the open set in \( X_i \),

\[
\{ z : x = (x_1, \cdots, x_{i-1}, z, x_{i+1}, \cdots, x_n) \in U \}
\]
to $Y$. When this map is differentiable, its derivative is denoted by $D_i g(x)$. To aid in the notation, for $v \in X_i$, let $\theta_i v \in \prod_{i=1}^n X_i$ be the vector $(0, \cdots, v, \cdots, 0)$ where the $v$ is in the $i^{th}$ slot and for $v \in \prod_{i=1}^n X_i$, let $v_i$ denote the entry in the $i^{th}$ slot of $v$. Thus, by saying

$$z \to g(x_1, \cdots, x_{i-1}, z, x_{i+1}, \cdots, x_n)$$

is differentiable is meant that for $v \in X_i$ sufficiently small,

$$g(x + \theta_i v) - g(x) = D_i g(x) v + o(v).$$

Note $D_i g(x) \in \mathcal{L}(X_i, Y)$.

**Definition 6.8.4** Let $U \subseteq X$ be an open set. Then $f : U \to Y$ is $C^1(U)$ if $f$ is differentiable and the mapping

$$x \to Df(x),$$

is continuous as a function from $U$ to $\mathcal{L}(X, Y)$.

With this definition of partial derivatives, here is the major theorem. Note the resemblance with the matrix of the derivative of a function having values in $\mathbb{R}^m$ in terms of the partial derivatives.

**Theorem 6.8.5** Let $g, U, \prod_{i=1}^n X_i$, be given as in Definition 6.8.3. Then $g$ is $C^1(U)$ if and only if $D_i g$ exists and is continuous on $U$ for each $i$. In this case, $g$ is differentiable and

$$Dg(x)(v) = \sum_k D_k g(x) v_k$$

where $v = (v_1, \cdots, v_n)$.

**Proof:** Suppose then that $D_i g$ exists and is continuous for each $i$. Note that

$$\sum_{j=1}^k \theta_j v_j = (v_1, \cdots, v_k, 0, \cdots, 0).$$

Thus $\sum_{j=1}^n \theta_j v_j = v$ and define $\sum_{j=1}^0 \theta_j v_j \equiv 0$. Therefore,

$$g(x + v) - g(x) = \sum_{k=1}^n \left[ g\left(x + \sum_{j=1}^k \theta_j v_j\right) - g\left(x + \sum_{j=1}^{k-1} \theta_j v_j\right) \right] \quad (6.8.16)$$

Consider the terms in this sum.

$$g\left(x + \sum_{j=1}^k \theta_j v_j\right) - g\left(x + \sum_{j=1}^{k-1} \theta_j v_j\right) = g(x + \theta_k v_k) - g(x) +$$

$$\sum_{j=1}^{k-1} g\left(x + \sum_{j=1}^j \theta_j v_j\right) \theta_{k-j} v_{k-j} \quad (6.8.17)$$
\[
\left( g\left( x + \sum_{j=1}^{k} \theta_j v_j \right) - g\left( x + \theta_k v_k \right) \right) - \left( g\left( x + \sum_{j=1}^{k-1} \theta_j v_j \right) - g(x) \right)
\] (6.8.18)

and the expression in (6.8.18) is of the form \( h(v_k) - h(0) \) where for small \( w \in X_k \),

\[
h(w) \equiv g\left( x + \sum_{j=1}^{k-1} \theta_j v_j + \theta_k w \right) - g(x + \theta_k w).
\]

Therefore,

\[
Dh(w) = D_k g\left( x + \sum_{j=1}^{k-1} \theta_j v_j + \theta_k w \right) - D_k g\left( x + \theta_k w \right)
\]

and by continuity, \(||Dh(w)|| < \varepsilon\) provided \(||v||\) is small enough. Therefore, by Theorem 6.6.2, the mean value inequality, whenever \(||v||\) is small enough,

\[
||h(v_k) - h(0)|| \leq \varepsilon ||v||
\]

which shows that since \( \varepsilon \) is arbitrary, the expression in (6.8.18) is \( o(v) \). Now in (6.8.16)

\[
g(x + \theta_k v_k) - g(x) = D_k g(x) v_k + o(v_k) = D_k g(x) v_k + o(v).
\]

Therefore, referring to (6.8.10),

\[
g(x + v) - g(x) = \sum_{k=1}^{n} D_k g(x) v_k + o(v)
\]

which shows \( Dg(x) \) exists and equals the formula given in (6.8.17). Also \( x \rightarrow Dg(x) \) is continuous since each of the \( D_k g(x) \) are.

Next suppose \( g \) is \( C^1 \). I need to verify that \( D_k g(x) \) exists and is continuous. Let \( v \in X_k \) sufficiently small. Then

\[
g(x + \theta_k v) - g(x) = Dg(x) \theta_k v + o(\theta_k v)
\]

since \(||\theta_k v|| = ||v||\). Then \( D_k g(x) \) exists and equals

\[
Dg(x) \circ \theta_k
\]

Now \( x \rightarrow Dg(x) \) is continuous. It is clear that \( \theta_k : X_k \rightarrow \prod_{i=1}^{n} X_i \) is also continuous because \( \theta_k v \) places \( v \) in the \( k^{th} \) position and \( 0 \) in every other position. ■

Note that the above argument also works at a single point \( x \). That is, continuity at \( x \) of the partials implies \( Dg(x) \) exists and is continuous at \( x \).
6.9 Mixed Partial Derivatives

Let $U$ be an open set in $\prod_{i=1}^{n} X_i$ where the norm is the one described above and let $f : U \to Y$ be a function for which the higher order partial derivatives of the sort described above exist. As in the case of functions defined on open sets of $\mathbb{R}^n$ one can ask whether the mixed partials are equal.

Results of this sort were known to Euler in around 1734. The theorem was proved by Clairaut some time later. It turns out that the mixed partial derivatives, if continuous will end up being equal. It will also work in the more general situation just described.

Theorem 6.9.1 Let $U$ be an open subset of $\prod_{i=1}^{n} X_i$ where each $X_i$ is a normed linear space and $\|x\| = \max_i \|x_i\|$. Let $f : U \to Y$ have mixed partial derivatives $D_i D_j f$ and $D_j D_i f$. Then if these are continuous at $x \in U$, it follows they will be equal in the sense that $D_j D_i f (x) (u, v) = D_i D_j f (x) (v, u)$.

**Proof:** It suffices to assume that there are only two spaces and $U$ is an open subset of $X_1 \times X_2$ because one simply specializes to two of the variables in the general case. We denote the variable for $X_1$ as $x$ and the one from $X_2$ as $y$. Also, to simplify this, first assume $f$ has values in $\mathbb{R}$. Thus it will be denoted as $f$ rather than $f$. Since $U$ is open, there exists $r > 0$ such that $B ((x, y), r) \subseteq U$. Now let $t, s$ be small real numbers and consider

$$\Delta (s, t) = \frac{1}{st} \left( f(x + tu, y + sv) - f(x + tu, y) - (f(x, y + sv) - f(x, y)) \right)$$

Then $h'(t) = D_1 f (x + tu, y + sv) (u) - D_1 f (x + tu, y) (u)$. By the mean value theorem,

$$\Delta (s, t) = \frac{1}{s} h'(\theta t) = \frac{1}{s} D_1 f (x + \theta tu, y + sv) (u) - D_1 f (x + \theta tu, y) (u)$$

where $\theta \in (0, 1)$. Now use the mean value theorem again to obtain

$$\Delta (s, t) = D_2 D_1 f (x + \theta tu, y + \alpha sv) (u) (v), \quad \alpha \in (0, 1)$$

Similarly doing things in the other order writing

$$\Delta (s, t) = \frac{1}{st} \left( ((f(x + tu, y + sv) - f(x, y + sv)) - (f(x + tu, y) - f(x, y)) \right)$$

and taking the derivative first with respect to $s$ and next with respect to $t$, one can obtain

$$\Delta (s, t) = D_1 D_2 f \left( x + \hat{t} u, y + \hat{\alpha} s v \right) (v) (u)$$

where $\hat{t}, \hat{\alpha}$ are also in $(0, 1)$. Then letting $(s, t) \to (0, 0)$ and using continuity of the mixed partial derivatives, one obtains that

$$D_2 D_1 f (x, y) (u) (v) = D_1 D_2 f (x, y) (v) (u)$$
Letting \( v = u \) yields the desired result.

The general case follows right away by applying this result to \( \langle y^*, f \rangle \). Thus one obtains

\[
\langle y^*, D_2D_1f(x, y)(u) \rangle = \langle y^*, D_1D_2f(x, y)(v)(u) \rangle
\]

for every \( y^* \in Y' \). Hence, since \( Y' \) separates the points, it follows that the mixed partials are equal. \( \blacksquare \)

It is necessary to assume the mixed partial derivatives are continuous in order to assert they are equal. The following is a well known example \( [2] \).

**Example 6.9.2** Let

\[
f(x, y) = \begin{cases} 
    \frac{xy(x^2 - y^2)}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0) \\
    0 & \text{if } (x, y) = (0, 0)
\end{cases}
\]

From the definition of partial derivatives it follows immediately that \( f_x(0, 0) = f_y(0, 0) = 0 \). Using the standard rules of differentiation, for \( (x, y) \neq (0, 0) \),

\[
f_x = y \frac{x^4 - y^4 + 4x^2y^2}{(x^2 + y^2)^2}, \quad f_y = x \frac{x^4 - y^4 - 4x^2y^2}{(x^2 + y^2)^2}
\]

Now

\[
f_{xy}(0, 0) = \lim_{y \to 0} \frac{f_x(0, y) - f_x(0, 0)}{y} = \lim_{y \to 0} \frac{-y^4}{(y^2)^2} = -1
\]

while

\[
f_{yx}(0, 0) = \lim_{x \to 0} \frac{f_y(x, 0) - f_y(0, 0)}{x} = \lim_{x \to 0} \frac{x^4}{(x^2)^2} = 1
\]

showing that although the mixed partial derivatives do exist at \( (0, 0) \), they are not equal there.

Incidentally, the graph of this function appears very innocent. Its fundamental sickness is not apparent. It is like one of those whitened sepulchers mentioned in the Bible.
6.10 Implicit Function Theorem

Recall the following notation. \( \mathcal{L}(X,Y) \) is the space of bounded linear mappings from \( X \) to \( Y \) where here \((X, \|\cdot\|_X)\) and \((Y, \|\cdot\|_Y)\) are normed linear spaces. Recall that this means that for each \( L \in \mathcal{L}(X,Y) \)

\[
\|L\| \equiv \sup_{\|x\| \leq 1} \|Lx\| < \infty
\]

As shown earlier, this makes \( \mathcal{L}(X,Y) \) into a normed linear space. In case \( X \) is finite dimensional, \( \mathcal{L}(X,Y) \) is the same as the collection of linear maps from \( X \) to \( Y \). In what follows \( X,Y \) will be Banach spaces, complete normed linear spaces. Thus these are complete normed linear space and \( \mathcal{L}(X,Y) \) is the space of bounded linear maps. I will also cease trying to write the vectors in bold face partly to emphasize that these are not in \( \mathbb{R}^n \).

**Definition 6.10.1** Let \((X, \|\cdot\|_X)\) and \((Y, \|\cdot\|_Y)\) be two normed linear spaces. Then \( \mathcal{L}(X,Y) \) denotes the set of linear maps from \( X \) to \( Y \) which also satisfy the following condition. For \( L \in \mathcal{L}(X,Y) \),

\[
\lim_{\|x\| \leq 1} \|Lx\|_Y \equiv \|L\| < \infty
\]

Recall that this operator norm is less than infinity is always the case where \( X \) is finite dimensional. However, if you wish to consider infinite dimensional situations, you assume the operator norm is finite as a qualification for being in \( \mathcal{L}(X,Y) \). Then here is an important theorem.

**Theorem 6.10.2** If \( Y \) is a Banach space, then \( \mathcal{L}(X,Y) \) is also a Banach space.

**Proof:** Let \( \{L_n\} \) be a Cauchy sequence in \( \mathcal{L}(X,Y) \) and let \( x \in X \).

\[
\|L_n x - L_m x\| \leq \|x\| \|L_n - L_m\|.
\]

Thus \( \{L_n x\} \) is a Cauchy sequence. Let

\[
Lx = \lim_{n \to \infty} L_n x.
\]

Then, clearly, \( L \) is linear because if \( x_1, x_2 \) are in \( X \), and \( a, b \) are scalars, then

\[
L(ax_1 + bx_2) = \lim_{n \to \infty} L_n (ax_1 + bx_2) = \lim_{n \to \infty} (aL_n x_1 + bL_n x_2) = aLx_1 + bLx_2.
\]

Also \( L \) is bounded. To see this, note that \( \{|L_n|\} \) is a Cauchy sequence of real numbers because \( ||L_n|| - ||L_m|| \leq ||L_n - L_m|| \). Hence there exists \( K > \sup\{|L_n| : n \in \mathbb{N}\} \). Thus, if \( x \in X \),

\[
\|Lx\| = \lim_{n \to \infty} \|L_n x\| \leq K \|x\|.
\]

The following theorem is really nice. The series in this theorem is called the Neuman series.
Lemma 6.10.3 Let \((X, \|\cdot\|)\) is a Banach space, and if \(A \in \mathcal{L}(X,X)\) and \(\|A\| = r < 1\), then
\[
(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \in \mathcal{L}(X,X)
\]
where the series converges in the Banach space \(\mathcal{L}(X,X)\). If \(\mathcal{O}\) consists of the invertible maps in \(\mathcal{L}(X,X)\), then \(\mathcal{O}\) is open and if \(\mathcal{I}\) is the mapping which takes \(A\) to \(A^{-1}\), then \(\mathcal{I}\) is continuous.

**Proof:** First of all, why does the series make sense?
\[
\left\| \sum_{k=p}^{q} A^k \right\| \leq \sum_{k=p}^{q} \|A^k\| \leq \sum_{k=p}^{q} \|A\|^k \leq \sum_{k=p}^{\infty} r^k = \frac{r^p}{1 - r}
\]
and so the partial sums are Cauchy in \(\mathcal{L}(X,X)\). Therefore, the series converges to something in \(\mathcal{L}(X,X)\) by completeness of this normed linear space. Now why is it the inverse?
\[
\sum_{k=0}^{\infty} A^k (I - A) = \lim_{n \to \infty} \sum_{k=0}^{n} A^k (I - A)
= \lim_{n \to \infty} \left( \sum_{k=0}^{n} A^k - \sum_{k=1}^{n+1} A^k \right)
= \lim_{n \to \infty} (I - A^{n+1}) = I
\]
because \(\|A^{n+1}\| \leq \|A\|^{n+1} \leq r^{n+1}\). Similarly,
\[
(I - A) \sum_{k=0}^{\infty} A^k = \lim_{n \to \infty} (I - A^{n+1}) = I
\]
and so this shows that this series is indeed the desired inverse.

Next suppose \(A \in \mathcal{O}\) so \(A^{-1} \in \mathcal{L}(X,X)\). Then suppose \(\|A - B\| < \frac{r}{1 + \|A^{-1}\|}, r < 1\). Does it follow that \(B\) is also invertible?
\[
B = A - (A - B) = A \left[ I - A^{-1} (A - B) \right]
\]
Then \(\|A^{-1} (A - B)\| \leq \|A^{-1}\| \|A - B\| < r\) and so \(\left[I - A^{-1} (A - B)\right]^{-1}\) exists. Hence
\[
B^{-1} = \left[I - A^{-1} (A - B)\right]^{-1} A^{-1}
\]
Thus \(\mathcal{O}\) is open as claimed. As to continuity, let \(A, B\) be as just described. Then using the Neuman series,
\[
\|\mathcal{I}A - \mathcal{I}B\| = \left\| A^{-1} - \left[I - A^{-1} (A - B)\right]^{-1} A^{-1}\right\|
\]
\[ A^{-1} - \sum_{k=0}^{\infty} (A^{-1} (A - B))^k A^{-1} = \sum_{k=1}^{\infty} (A^{-1} (A - B))^k A^{-1} \]
\[ \leq \sum_{k=1}^{\infty} \|A^{-1}\|^{k+1} \|A - B\|^k = \|A - B\| \|A^{-1}\|^2 \sum_{k=0}^{\infty} \|A^{-1}\|^k \left( \frac{r}{1 + \|A^{-1}\|} \right)^k \]
\[ \leq \|B - A\| \|A^{-1}\|^2 \frac{1}{1 - r}. \]

Thus \( \mathcal{J} \) is continuous at \( A \in \mathcal{O} \).

**Lemma 6.10.4** Let \( \mathcal{O} \equiv \{ A \in \mathcal{L}(X, Y) : A^{-1} \in \mathcal{L}(Y, X) \} \)
and let \( \mathcal{J} : \mathcal{O} \to \mathcal{L}(Y, X), \mathcal{J} A \equiv A^{-1} \).

Then \( \mathcal{O} \) is open and \( \mathcal{J} \) is in \( C^m (\mathcal{O}) \) for all \( m = 1, 2, \ldots \). Also
\[ D\mathcal{J} (A) (B) = -\mathcal{J} (A) (B) \mathcal{J} (A). \quad (6.10.19) \]

In particular, \( \mathcal{J} \) is continuous.

**Proof:** Let \( A \in \mathcal{O} \) and let \( B \in \mathcal{L}(X, Y) \) with
\[ \|B\| \leq \frac{1}{2} \|A^{-1}\|^{-1}. \]
Then
\[ \|A^{-1}B\| \leq \|A^{-1}\| \|B\| \leq \frac{1}{2} \]
and so by Lemma 6.10.3,
\[ (I + A^{-1}B)^{-1} \in \mathcal{L}(X, X). \]

It follows that
\[ (A + B)^{-1} = (A (I + A^{-1}B))^{-1} = (I + A^{-1}B)^{-1} A^{-1} \in \mathcal{L}(Y, X). \]

Thus \( \mathcal{O} \) is an open set.

Thus
\[ (A + B)^{-1} = (I + A^{-1}B)^{-1} A^{-1} = \sum_{n=0}^{\infty} (-1)^n (A^{-1}B)^n A^{-1} = [I - A^{-1}B + o(B)] A^{-1} \]
which shows that $O$ is open and, also,

$$\mathcal{J}(A + B) - \mathcal{J}(A) = \sum_{n=0}^{\infty} (-1)^n (A^{-1}B)^n A^{-1} - A^{-1}$$

$$= -A^{-1}BA^{-1} + o(B)$$

$$= -\mathcal{J}(A)(B)\mathcal{J}(A) + o(B)$$

which demonstrates 160. It follows from this that we can continue taking derivatives of $\mathcal{J}$. For $||B_1||$ small,

$$- [D\mathcal{J}(A + B_1)(B) - D\mathcal{J}(A)(B)] =$$

$$\mathcal{J}(A + B_1)(B)\mathcal{J}(A + B_1) - \mathcal{J}(A)(B)\mathcal{J}(A)$$

$$= \mathcal{J}(A + B_1)(B)\mathcal{J}(A + B_1) - \mathcal{J}(A)(B)\mathcal{J}(A) +$$

$$\mathcal{J}(A)(B)\mathcal{J}(A + B_1) - \mathcal{J}(A)(B)\mathcal{J}(A) +$$

$$= [\mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)](B)\mathcal{J}(A + B_1) +$$

$$\mathcal{J}(A)(B)[\mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)]$$

$$= [\mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)](B)[\mathcal{J}(A) - \mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)] +$$

$$\mathcal{J}(A)(B)[\mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)]$$

$$= \mathcal{J}(A)(B_1)\mathcal{J}(A)(B)\mathcal{J}(A) + \mathcal{J}(A)(B)\mathcal{J}(A)(B_1)\mathcal{J}(A) + o(B_1)$$

and so

$$D^2\mathcal{J}(A)(B_1)(B) = \mathcal{J}(A)(B_1)\mathcal{J}(A)(B)\mathcal{J}(A) + \mathcal{J}(A)(B)\mathcal{J}(A)(B_1)\mathcal{J}(A)$$

which shows $\mathcal{J}$ is $C^2(O)$. Clearly we can continue in this way which shows $\mathcal{J}$ is in $C^m(O)$ for all $m = 1, 2, \ldots$.

Here are the two fundamental results presented earlier which will make it easy to prove the implicit function theorem. First is the fundamental mean value inequality. It was Theorem 5.10.4.

**Theorem 6.10.5** Suppose $U$ is an open subset of $X$ and $f : U \to Y$ has the property that $Df(x)$ exists for all $x$ in $U$ and that, $x + t(y - x) \in U$ for all $t \in [0, 1]$. (The line segment joining the two points lies in $U$.) Suppose also that for all points on this line segment,

$$||Df(x + t(y - x))|| \leq M.$$ 

Then

$$||f(y) - f(x)|| \leq M|y - x|.$$
Next recall the following theorem about fixed points of a contraction map. It was Corollary 1.8.3.

**Corollary 6.10.6** Let $B$ be a closed subset of the complete metric space $(X,d)$ and let $f : B \to X$ be a contraction map

$$d(f(x), f(\hat{x})) \leq rd(x, \hat{x}), \quad r < 1.$$ 

Also suppose there exists $x_0 \in B$ such that the sequence of iterates $\{f^n(x_0)\}_{n=1}^{\infty}$ remains in $B$. Then $f$ has a unique fixed point in $B$ which is the limit of the sequence of iterates. This is a point $x \in B$ such that $f(x) = x$. In the case that $B = B(x_0, \delta)$, the sequence of iterates satisfies the inequality

$$d(f^n(x_0), x_0) \leq \frac{d(x_0, f(x_0))}{1 - r}$$

and so it will remain in $B$ if

$$\frac{d(x_0, f(x_0))}{1 - r} < \delta.$$ 

Recall that for $X,Y$ normed vector spaces, the norm on $X \times Y$ is of the form $||\cdot|| = \max(||\cdot||_X, ||\cdot||_Y)$.

**Theorem 6.10.7** (implicit function theorem) Let $X,Y,Z$ be Banach spaces and suppose $U$ is an open set in $X \times Y$. Let $f : U \to Z$ be in $C^1(U)$ and suppose

$$f(x_0, y_0) = 0, \quad D_1f(x_0, y_0)^{-1} \in \mathcal{L}(Z,X). \quad (6.10.20)$$

Then there exist positive constants, $\delta, \eta$, such that for every $y \in B(y_0, \eta)$ there exists a unique $x(y) \in B(x_0, \delta)$ such that

$$f(x(y), y) = 0. \quad (6.10.21)$$

Furthermore, the mapping, $y \to x(y)$ is in $C^1(B(y_0, \eta))$.

**Proof:** Let $T(x, y) \equiv x - D_1f(x_0, y_0)^{-1} f(x, y)$. Therefore,

$$D_1T(x, y) = I - D_1f(x_0, y_0)^{-1} D_1f(x, y). \quad (6.10.22)$$

by continuity of the derivative which implies continuity of $D_1T$, it follows there exists $\delta > 0$ such that if $\|x-x_0\| < \delta$ and $\|y-y_0\| < \delta$, then

$$\|D_1T(x, y)\| < \frac{1}{2}, \quad D_1f(x, y)^{-1} \text{ exists} \quad (6.10.23)$$

The second claim follows from Lemma 6.10.4. By the mean value inequality, Theorem 6.10.5, whenever $x,x' \in B(x_0, \delta)$ and $y \in B(y_0, \delta)$,

$$\|T(x, y) - T(x', y)\| \leq \frac{1}{2} \|x - x'\|. \quad (6.10.24)$$
Also, it can be assumed δ is small enough that for some M and all such (x,y),
\[ \left\| D_1 f (x_0, y_0)^{-1} \right\| \left\| D_2 f (x, y) \right\| < M \]  
(6.10.25)

Next, consider only y such that \( \|y - y_0\| < \eta \) where \( \eta \) is so small that
\[ \|T(x_0, y) - x_0\| < \frac{\delta}{3} \]

Then for such y, consider the mapping \( T_y(x) = T(x, y) \). Thus by Corollary 6.10.4, for each \( n \in \mathbb{N} \),
\[ \delta > \frac{2}{3} \delta \geq \frac{|T_y(x_0) - x_0|}{1 - (1/2)} \geq |T_y^n(x_0) - x_0| \]

Then by 6.10.24, the sequence of iterations of this map \( T_y \) converges to a unique fixed point \( x(y) \) in the ball \( B(x_0, \delta) \). Thus, from the definition of \( T \), \( f(x(y), y) = 0 \). This is the implicitly defined function.

Next we show that this function is Lipschitz continuous. For \( y, \hat{y} \) in \( B(y_0, \eta) \),
\[ \|T(x, y) - T(x, \hat{y})\| = \|D_1 f (x_0, y_0)^{-1} f(x, y) - D_1 f (x_0, y_0)^{-1} f(x, \hat{y})\| \leq M \|y - \hat{y}\| \]
thanks to the above estimate 6.11.24 and the mean value inequality, Theorem 5.11.5. Note how convexity of \( B(y_0, \eta) \) which says that the line segment joining \( y, \hat{y} \) is contained in \( B(y_0, \eta) \) is important to use this theorem. Then from this,
\[ \|x(y) - x(\hat{y})\| = \|T(x(y), y) - T(x(\hat{y}), \hat{y})\| \leq \|T(x(y), y) - T(x(\hat{y}), \hat{y})\| \]
\[ + \|T(x(y), \hat{y}) - T(x(\hat{y}), \hat{y})\| \]
\[ \leq M \|y - \hat{y}\| + \frac{1}{2} \|x(y) - x(\hat{y})\| \]
Hence,
\[ \|x(y) - x(\hat{y})\| \leq 2M \|y - \hat{y}\| \]  
(6.10.26)

Finally consider the claim that this implicitly defined function is \( C^1 \).
\[ 0 = f(x + u, y + u) - f(x, y) \]
\[ = D_1 f(x, y)(x + u - x) + D_2 f(x, y)u \]
\[ + o(x + u - x, u) \]  
(6.10.27)

Consider the last term. \( o(x + u - x, u) / \|u\| \) equals
\[
\left\{ \begin{array}{ll}
\frac{o(x + u - x, u)}{\|x + u - x, u\|_{X \times Y}} & \text{if } \|x + u - x, u\|_{X \times Y} \neq 0 \\
0 & \text{if } \|x + u - x, u\|_{X \times Y} = 0
\end{array} \right.
\]
Now the Lipschitz condition just established shows that
\[ \max \left( \|x(y + u) - x(y)\|, \|u\| \right) \]
is bounded for nonzero $u$ sufficiently small that $y,y+u \in B(y_0,\eta)$. Therefore,
\[
\lim_{u \to 0} \frac{o(x(y+u)-x(y),u)}{\|u\|} = 0
\]
Then Theorem 6.10.9 shows that
\[
0 = D_1 f(x(y),y)(x(y+u)-x(y)) + D_2 f(x(y),y)u + o(u)
\]
Therefore, solving for $x(y+u)-x(y)$, it follows that
\[
x(y+u)-x(y) = -D_1 f(x(y),y)^{-1} D_2 f(x(y),y)u + D_1 f(x(y),y)^{-1} o(u)
\]
and now, the continuity of the partial derivatives $D_1 f, D_2 f$, continuity of the map $A \to A^{-1}$, along with the continuity of $y \to x(y)$ shows that $y \to x(y)$ is $C^1$ with derivative equal to $-D_1 f(x(y),y)^{-1} D_2 f(x(y),y)$. \(\blacksquare\)

It is easy to give a version of this theorem in which the function $f$ also depends on a parameter $\lambda \in \Lambda$, a metric space.

**Corollary 6.10.8** Let $X,Y,Z$ be Banach spaces and suppose $U$ is an open set in $X \times Y$. Let $f : U \times \Lambda \to Z$ satisfy $f(\cdot,\cdot,\lambda)$ is in $C^1(U)$ and suppose for each $\lambda$,
\[
f(x_0,y_0,\lambda) = 0, \ D_1 f(x_0,y_0,\lambda)^{-1} \in \mathcal{L}(Z,X).
\]  
(6.10.28)

Also suppose $(x,y) \to D_1 f(x,y,\lambda)$ is continuous uniformly in $\lambda$ and $D_2(x,y,\lambda)$ is uniformly bounded in $\lambda$ for $(x,y)$ sufficiently close to $(x_0,y_0)$. Then there exist positive constants, $\delta,\eta$, such that for every $y \in B(y_0,\eta)$ there exists a unique $x(y,\lambda) \in B(x_0,\delta)$ such that
\[
f(x(y,\lambda),y,\lambda) = 0.
\]  
(6.10.29)

Furthermore, the mapping, $y \to x(y,\lambda)$ is in $C^1(B(y_0,\eta))$ and $\lambda \to x(y,\lambda)$ is continuous.

**Proof:** It is just a repeat of the above proof except you use the uniform contraction principle, Corollary 6.10.8 to get the fixed point. \(\blacksquare\)

The next theorem is a very important special case of the implicit function theorem known as the inverse function theorem. Actually one can also obtain the implicit function theorem from the inverse function theorem. It is done this way in [55], [18] and in [4].

**Theorem 6.10.9** (inverse function theorem) Let $x_0 \in U$, an open set in $X$, and let $f : U \to Y$ where $X,Y$ are finite dimensional normed vector spaces. Suppose
\[
f \text{ is } C^1(U), \text{ and } Df(x_0)^{-1} \in \mathcal{L}(Y,X).
\]  
(6.10.30)

Then there exist open sets $W$, and $V$ such that
\[
x_0 \in W \subseteq U,
\]  
(6.10.31)

\[
f : W \to V \text{ is one to one and onto},
\]  
(6.10.32)

\[
f^{-1} \text{ is } C^1,
\]  
(6.10.33)
CHAPTER 6. THE DERIVATIVE

Proof: Apply the implicit function theorem to the function

\[ F(x, y) \equiv f(x) - y \]

where \( y_0 \equiv f(x_0) \). Thus the function \( y \to x(y) \) defined in that theorem is \( f^{-1} \).

Now let

\[ W \equiv B(x_0, \delta) \cap f^{-1}(B(y_0, \eta)) \]

and

\[ V \equiv B(y_0, \eta). \]

6.11 More Derivatives

When you consider a \( C^k \) function \( f \) defined on an open set \( U \), you obtain the following

\[ Df(x) \in \mathcal{L}(X, Y), D^2 f(x) \in \mathcal{L}(X, \mathcal{L}(X, Y)), D^3 f(x) \in \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, Y))) \]

and so forth. Thus they can each be considered as a linear transformation with values in some vector space. When you consider the vector spaces, you see that these can also be considered as multilinear functions on \( X \) with values in \( Y \). Now consider the product of two linear transformations \( A(y) B(y) w \), where everything is given to make sense and here \( w \) is an appropriate vector. Then if each of these linear transformations can be differentiated, you would do the following simple computation.

\[
(A(y + u) B(y + u) - A(y) B(y))(w)
\]

\[
= (A(y + u) B(y + u) - A(y) B(y + u) + A(y) B(y + u) - A(y) B(y))(w)
\]

\[
= ((DA(y) u + o(u)) B(y + u) + A(y) (DB(y) u + o(u)))(w)
\]

\[
= (DA(y) (u) B(y + u) + A(y) DB(y)(u) + o(u))(w)
\]

Then

\[
(u, w) \to (DA(y)(u) B(y) + A(y) DB(y)(u))(w)
\]

is clearly linear and

\[
(u, w) \to (DA(y)(u) B(y) + A(y) DB(y)(u))(w)
\]

is bilinear and continuous as a function of \( y \). By this we mean that for a fixed choice of \( (u, w) \) the resulting \( Y \) valued function just described is continuous. Now if each of \( A, B, DA, DB \) can be differentiated, you could replace \( y \) with \( y + \hat{u} \) and do a similar computation to obtain as many differentiations as desired, the \( k^{th} \) differentiation yielding a \( k \) linear function. You can do this as long as \( A \) and \( B \) have derivatives.

Now in the case of the implicit function theorem, you have

\[
Dx(y) = -D_1 f(x(y), y)^{-1} D_2 f(x(y), y). \quad (6.11.34)
\]
By Lemma \ref{lemma6104} and the implicit function theorem and the chain rule, this is the situation just discussed. Thus $D^2x(y)$ can be obtained. Then the formula for it will only involve $Dx$ which is known to be continuous. Thus one can continue in this way finding derivatives till $f$ fails to have them. The inverse map never creates difficulties because it is differentiable of order $m$ for any $m$ thanks to Lemma \ref{lemma6104}. Thus one can conclude the following corollary.

**Corollary 6.11.1** In the implicit and inverse function theorems, you can replace $C^1$ with $C^k$ in the statements of the theorems for any $k \in \mathbb{N}$.

### 6.12 Lyapunov Schmidt Procedure

You have $f : X \times \Lambda \to Y$ where here $X, \Lambda$ are Banach spaces. Suppose $(0, 0) \in X \times \Lambda$ and $f (0, 0) = 0$. Then if $D_1 f (0, 0)^{-1}$ is in $\mathcal{L} (Y, X)$, the implicit function theorem says that there exists $x (\lambda)$ a $C^p$ function such that locally $f (x (\lambda), \lambda) = 0$. So what if $D_1 f (0, 0)$ fails to be one to one? Sometimes this case is also considered. It may be that $D_1 f (0, 0)$ is one to one on some subspace and other nice things happen. In particular, suppose the following.

Letting $X_2 \equiv \ker D_1 f (0, 0)$ assume

$$X = X_1 \oplus X_2, \quad \dim (X_2) < \infty$$

where $X_1$ is a closed subspace. Thus $D_1 f (0, 0)$ is one to one on $X_1$. We let

$$Y_1 = D_1 f (0, 0) (X_1)$$

and suppose that $Y = Y_1 \oplus Y_2$ where $\dim (Y_2) < \infty$, $Y_1$ also a closed subspace.

$$X_1 \xrightarrow{D_1 f (0, 0)} Y_1 = D_1 f (0, 0) (X_1), \quad Y_1 \text{ closed} \quad < \infty$$

$$Y = Y_1 \oplus Y_2, \quad \dim (Y_2) < \infty$$

By the open mapping theorem, $D_1 f (0, 0)^{-1}$ is also continuous.

Let $Q$ be a continuous projection onto $Y_1$ which is assumed to exist\footnote{In Hilbert space, the existence of this projection map is obvious and it is assumed that it exists here.} so that $(I - Q)$ is a projection onto $Y_2$. Then the equation $f (x (\lambda), \lambda) = 0$ can be written as the pair

$$Qf (x, \lambda) = 0$$

$$(I - Q) f (x, \lambda) = 0$$

Consider the top. For $x = x_1 + x_2$ where $x_i \in X_i$, this is

$$Qf (x_1 + x_2, \lambda) = 0$$
CHAPTER 6. THE DERIVATIVE

Then if \( g(x_1, x_2, \lambda) = Qf(x_1 + x_2, \lambda) \), one has \( g : X_1 \times X_2 \times \Lambda \to Y_1 \)

\[
D_1g(x_1, x_2, \lambda) h = D_1Qf(x_1 + x_2, \lambda) h, \ h \in X_1.
\]

Thus \( D_1g(0, 0, 0)^{-1} \) is continuous by the open mapping theorem \( (D_1f(0, 0) \) is one to one on \( X_1 \)), and by the implicit function theorem, there is a solution to

\[
Qf(x_1 + x_2, \lambda) = 0
\]

for \( x_1 = x_1(x_2, \lambda) \). (Note how it is important that \( X_1 \) and \( Y_1 \) be Banach spaces.) Then the other equation yields

\[
(I - Q)f(x_1(x_2, \lambda) + x_2, \lambda) = 0
\]

and so for fixed \( \lambda \), this is a finite set of equations of a variable in a finite dimensional space.

This depends on being able to write \( X = X_1 \oplus X_2 \) where \( X_1 \) is closed, \( X_2 = \ker D_1 f(0, 0) \), a similar situation for \( Y = Y_1 \oplus Y_2 \). So when does this happen? Are there conditions on \( D_1 f(0, 0) \) which will cause it to occur?

There are such conditions. For example, \( D_1 f(0, 0) \) could be a Fredholm operator defined in Definition \[B.4.7]. The following are some easy examples in which all that nonsense about things being finite dimensional and part of a direct sum does not need to be considered.

**Example 6.12.1** Say \( X = \mathbb{R}^2 \) and \( \Lambda = \mathbb{R} \). Let \( f(x, y, \lambda) = x + xy + y^2 + \lambda \). Then

\[
D_1 f(0, 0, 0) = (1, 0)
\]

this \( 1 \times 2 \) matrix mapping \( \mathbb{R}^2 \) to \( \mathbb{R} \). Thus \( X_2 = (0, \alpha)^T : \alpha \in \mathbb{R} \) and \( X_1 = (\alpha, 0)^T : \alpha \in \mathbb{R} \). In this case, \( Y_1 = \mathbb{R} \) and so \( Q = I \). Thus the above reduces to the single equation

\[
f((\alpha, 0) + (0, \beta), \lambda) = 0
\]

and so since \( D_1 f(0, 0, 0) \) is one to one, \( x_1 = (\alpha, 0) = x_1((0, \beta), \lambda) \). Of course this is completely obvious because if you consider \( f \) in the natural way as a function of three variables, then the implicit function theorem immediately gives \( x = (y, \lambda) \) which is essentially the same result. We just write \((\alpha, 0)\) in place of \( \alpha \). The first independent variable is a function of the other two.

**Example 6.12.2** Here is another easy example. \( f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2 \)

\[
f(x, y, \lambda) = \begin{pmatrix} x + xy + y^2 + \sin(\lambda) \\ x + y^2 - x^2 + \lambda \end{pmatrix}
\]

Then

\[
D_1f(x, y, \lambda) = \begin{pmatrix} 1 + y & x + 2y \\ 1 - 2x & 2y \end{pmatrix}
\]
So
\[ D_1f((0,0),0) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \]

Then
\[ X_2 = \ker D_1f((0,0),0) = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\} \]

and
\[ X_1 = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\} \]

and clearly \( D_1f((0,0),0) \) is indeed one to one on \( X_1 \).

\[ D_1f(0,0)(X_1) = \left\{ \begin{pmatrix} y \\ y \end{pmatrix} : y \in \mathbb{R} \right\} = Y_1. \]

In this case, let
\[ Q(\alpha \beta) = \begin{pmatrix} \alpha + \beta \\ \alpha - \beta \\ \alpha + \beta \\ \alpha - \beta \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \]

so \( I - Q = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix} \). Thus the equations are

\[ Qf(x,\lambda) = 0 \]
\[ (I - Q)f(x,\lambda) = 0 \]

This reduces to

\[ \begin{pmatrix} -\frac{1}{2}x^2 + \frac{1}{2}xy + x + y^2 + \frac{1}{2}\lambda + \frac{1}{2}\sin\lambda \\ -\frac{1}{2}x^2 + \frac{1}{2}xy + x + y^2 + \frac{1}{2}\lambda + \frac{1}{2}\sin\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]

\[ \begin{pmatrix} \frac{1}{2}x^2 + \frac{1}{2}y - \frac{1}{2}\lambda + \frac{1}{2}\sin\lambda \\ -\frac{1}{2}x^2 - \frac{1}{2}y + \frac{1}{2}\lambda - \frac{1}{2}\sin\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]

Note how in both the top and the bottom, there is only one equation and one can solve for \( x \) in terms of \( y,\lambda \) near \((0,0,0)\) which is what the above general argument shows. Of course you can see this directly using the implicit function theorem. Then can you solve for \( y = y(\lambda) \)? This would involve trying to solve for \( y \) as a function of \( \lambda \) in the following where \( x(y,\lambda) \) comes from the first equations.

\[ \frac{1}{2}x^2(y,\lambda) + \frac{1}{2}y(x,\lambda) - \frac{1}{2}\lambda + \frac{1}{2}\sin\lambda = 0 \]

If you can do this, then you would have found \((x,y)\) as a function of \( \lambda \) for small \( \lambda \).

In this example, in the top equation, at \((0,0,0), x_y = 0. \) Also \( x_\lambda = 0 \) so \( x(y,\lambda) \approx -\lambda \) other than higher order terms for small \( y,\lambda \). Then in the bottom equation, for all variables very small, you would have \( \lambda^2 + y(-\lambda) - \lambda + \sin\lambda = 0 \), \( y(\lambda) = -\lambda + \frac{\sin\lambda}{\lambda} + \lambda \) at least approximately. Thus it seems there is a nonzero solution to the equation \( f(x,y,\lambda) = 0 \) which is valid for small \( \lambda, x, y \), this in addition to the zero solution. Note that for small nonzero \( \lambda, -1 + \frac{\sin\lambda}{\lambda} + \lambda \neq 0 \). It equals approximately \( \lambda - \frac{\lambda^3}{6} \) for small \( \lambda \) from the power series for \( \sin \).

In the next example, the same procedure gives a solution to a problem \( f((x,y),\lambda) = 0 \) such that for small \( \lambda, (x,y) \) is a function of \( \lambda \) which is nonzero and \( f((0,0),\lambda) = 0 \). Thus for small \( \lambda \), there are two solutions to the nonlinear system of equations.
Example 6.12.3 Let
\[ f((x, y), \lambda) = \begin{pmatrix} x + xy + y^2 + x \sin(\lambda) \\ x + y^2 - x^2 + x\lambda \end{pmatrix} \]

In this case \( f((0, 0), \lambda) = 0 \) even though \( \lambda \) might not be 0. The Lyapunov Schmidt procedure will be used to show that there are nonzero solutions \( x(\lambda), y(\lambda) \) such that \( f((x(\lambda), y(\lambda)), \lambda) = 0 \).

At origin,
\[ D_1f((0, 0), 0) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \]
Thus \( X_1 = \text{span}(e_1) \) and \( X_2 = \text{span}(e_2) \). Then \( Y_1 = \text{span}(e_1 + e_2) \) and \( Y_2 = \text{span}(e_1 - e_2) \). Also \( D_1f((0, 0), 0) \) is one to one on \( X_1 \) and its range is \( Y_1 \). Then let
\[ Q \left( \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right) = \begin{pmatrix} \frac{\alpha + \beta}{2} \\ \frac{\alpha - \beta}{2} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \]
\[ (I - Q) = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \]

Then \( Qf = 0 \) is yields the equation
\[ x + \frac{1}{2}x\lambda + \frac{1}{2}x\sin(\lambda) + \frac{1}{2}xy - \frac{1}{2}x^2 + y^2 = 0 \]
Also \((I - Q)f = 0\) yields the equation
\[ \frac{1}{2}x\sin(\lambda) - \frac{1}{2}x\lambda + \frac{1}{2}xy + \frac{1}{2}x^2 = 0 \]

Now consider \( x_y \) and \( x_\lambda \) at \((0, 0)\) from the first equation. Both of these are easily seen to be 0. Now consider \( x_{yy} \). After some computations, this is seen to be \( x_{yy} = -2 \). Similarly, \( x_{y\lambda}(0, 0) = 0, x_{\lambda\lambda}(0, 0) = 0 \) also. Thus up to terms of degree 3,
\[ x(y, \lambda) = -y^2 = \frac{1}{2}(-2)y^2 \]
Place this in the bottom equation.
\[ \frac{1}{2}y^2\lambda - \frac{1}{2}y^2\sin(\lambda) - \frac{1}{2}y^3 + \frac{1}{2}y^4 = 0 \]

Now the idea is to find \( y = y(\lambda) \), hopefully nonzero. Divide by \( y^2 \) and multiply by 2.
\[ y^2 - y + \lambda - \sin(\lambda) = 0 \]

Then for small \( \lambda \) this is approximately equal to
\[ y^2 - y + \frac{\lambda^3}{6} = 0 \]
Then a solution for $y$ for small $\lambda$ is

$$y = 1 + \sqrt{1 - \frac{2}{3} \lambda^3}$$

Of course there is another solution as well, when you replace the + with a minus sign. This is the one we want because when $\lambda = 0$ it reduces to $y = 0$. This shows that there exist solutions to the equations $f((x, y), \lambda) = 0$ which for small $\lambda$ are approximately

$$(x(\lambda), y(\lambda)) = \left(-y^2, \frac{1 - \sqrt{1 - \frac{2}{3} \lambda^3}}{2}\right)$$

In terms of $\lambda$ very small,

$$(x(\lambda), y(\lambda)) = \left(\frac{1}{6} \lambda^3 + \frac{1}{6} \sqrt{3} \sqrt{3 - 2 \lambda^3} - \frac{1}{2}, \frac{1 - \sqrt{1 - \frac{2}{3} \lambda^3}}{2}\right)$$

Using a power series in $\lambda$ to approximate these functions, this reduces to

$$(x(\lambda), y(\lambda)) = \left(-\frac{1}{36} \lambda^6 - \frac{1}{6} \lambda^3 + \frac{1}{36} \lambda^6 + \frac{1}{108} \lambda^9\right)$$

where higher order terms are neglected. Thus there exist other solutions than the zero solution even though $\lambda$ may be nonzero. Note that in this example, $f((0, 0), \lambda) = 0$.

### 6.13 Analytic Functions

In calculus, there was a difference between functions of a real variable and functions of a complex variable. In the latter case the existence of a single derivative implied the existence of all derivatives and in fact the Taylor series converged to the function. It is reasonable to ask if a similar phenomenon occurs in the case of complex Banach spaces versus real Banach spaces. This section presents a quick introduction to this topic based on the assumption that the reader has had some exposure to complex analysis. Some of the details involving questions of convergence and term by term differentiation are left to the reader. Also if $h$ maps an open subset of $\mathbb{C}$ to a complex Banach space $X$, and has a first derivative, then the usual Cauchy integral formula,

$$h(z) = \frac{1}{2\pi i} \int_{C} \frac{h(w)}{w-z} dw,$$

holds if $C$ is a circle contained, together with its interior, in the open set on which $h$ has a derivative. The integral can be defined as the ordinary Riemann integral using Riemann sums or it can be defined in terms of a Bochner integral. These details are routine and are left to the reader. There are several equivalent definitions of
an analytic function defined on a complex Banach space. The following is the one we will use since it resembles the familiar definition encountered in undergraduate complex variable courses.

**Definition 6.13.1** Let $X$ and $Y$ be complex Banach spaces and let $U \subseteq X$ be an open set. We say $f : U \to Y$ is analytic and bounded on $U$ if

$$z \to f(x + zh)$$ is analytic for $x \in U, h \in X$ and $|z|$ small enough

exists for all $x \in U$ and also $||f(x)|| \leq M < \infty$ for all $x \in U$. Here $z \in \mathbb{C}$ and $x, h \in X$.

Let $h \in X^l$ and consider all $z \in \mathbb{C}^l$ with $||z||_{\mathbb{C}^l} \equiv \max (|z_m|, m = 1, \cdots, l)$ sufficiently small. Let $C_1$ be a sufficiently small circle centered at 0. Then consider

$$z_m \to f(x + \sum_{m=1}^{l} z_m h_m)$$

which is analytic on and inside $C_1$. Thus using the Cauchy integral formula,

$$f\left(x + z_1 h_1 + \sum_{m=2}^{l} z_m h_m\right)$$

$$= \frac{1}{2\pi i} \int_{C_1} \frac{f(x + w_1 h_1 + \sum_{m=2}^{l} w_m h_m)}{(w_1 - z_1)} dw_1$$

$$= \left(\frac{1}{2\pi i}\right)^2 \int_{C_1} \frac{1}{w_1 - z_1},$$

$$\int_{C_1} f\left(x + w_1 h_1 + w_2 h_2 + \sum_{m=3}^{l} w_m h_m\right) dw_2 dw_1 =$$

$$\left(\frac{1}{2\pi i}\right)^l \int_{C_1} \cdots \int_{C_1} f(x + w_1 h_1 + w_2 h_2 + \cdots + w_l h_l) dw_1 \cdots dw_l.$$

Consider the case when $l = 2$.

$$\left(\frac{1}{2\pi i}\right)^2 \int_{C_1} \int_{C_1} f(x + w_1 h_1 + w_2 h_2) dw_2 dw_1 =$$

$$\left(\frac{1}{2\pi i}\right)^2 \int_{C_1} \int_{C_1} f(x + w_1 h_1 + w_2 h_2)\cdot$$

$$\sum_{k_2=0}^{\infty} \frac{z_2}{w_2^{k_2+1}} \sum_{k_1=0}^{\infty} \frac{z_1}{w_1^{k_1+1}} dw_2 dw_1 =$$
6.13. **ANALYTIC FUNCTIONS**

\[
\left(\frac{1}{2\pi i}\right)^2 \sum_{k_2=0}^{\infty} \sum_{k_1=0}^{\infty} \left(\int_{C_1} \int_{C_1} \frac{f(x + w_1 h_1 + w_2 h_2)}{w_2^{k_2+1} w_1^{k_1+1}} dw_2 dw_1\right) z_2^{k_2} z_1^{k_1}.
\]

Similarly, for arbitrary \( l \), and letting \( C \) be any circle centered at 0 with radius smaller than \( \frac{5}{7} \),

\[
f \left( x + \sum_{m=1}^{l} z_m h_m \right) = \sum_{k_1=0}^{\infty} \sum_{k_l=0}^{\infty} a_{k_1 \cdots k_l} (x, h_1, \cdots, h_l) z_1^{k_1} \cdots z_l^{k_l} \tag{6.13.35}
\]

where

\[
a_{k_1 \cdots k_l} (x, h_1, \cdots, h_l) = \left(\frac{1}{2\pi i}\right)^l \int_C \cdots \int_C \frac{f(x + \sum_{m=1}^{l} w_m h_m)}{\prod_{m=1}^{l} w_m^{k_m+1}} dw_1 \cdots dw_l. \tag{6.13.36}
\]

**Lemma 6.13.2** Let \( l \geq 1 \) and let \( t_m \in C \). Then if \( h \in X^l \), then whenever \( |z| \) is small enough, \( 6.13.35 \) holds. Also the coefficients satisfy

\[
a_{k_1 \cdots k_l} (x, t_l h_l, \cdots, t_1 h_1) = \left(\prod_{m=1}^{l} t_m^{k_m}\right) a_{k_1 \cdots k_l} (x, h_1, \cdots, h_l) \tag{6.13.37}
\]

and

\[
\|a_{k_1 \cdots k_l} (x, h_1, \cdots, h_l)\| \leq C \prod_{m=1}^{l} \|h_m\| \tag{6.13.38}
\]

for some constant \( C \).

**Proof:** Let \( C \) be small enough that the circles \( t_m C \) for all \( m = 1, \cdots, l \) and \( C \) have radius less than \( \frac{5}{7} \). First assume \( t_m \neq 0 \) for all \( m \). Then

\[
a_{k_1 \cdots k_l} (x, t_l h_l, \cdots, t_1 h_1)
\]

\[
= \left(\frac{1}{2\pi i}\right)^l \int_{C_l} \cdots \int_{C_1} \frac{f(x + \sum_{m=1}^{l} w_m t_m h_m)}{\prod_{m=1}^{l} (w_m t_m)^{k_m+1}} \prod_{m=1}^{l} t_m^{k_m+1} dw_1 \cdots dw_l
\]

Here we just multiplied and divided by \( \prod_{m=1}^{l} t_m^{k_m+1} \).

\[
= \left(\frac{1}{2\pi i}\right)^l \int_{C_l} \cdots \int_{C_1} \frac{f(x + \sum_{m=1}^{l} u_m h_m)}{\prod_{m=1}^{l} (u_m)^{k_m+1}} du_1 \cdots du_l \prod_{m=1}^{l} t_m^{k_m}
\]
CHAPTER 6. THE DERIVATIVE

\[= a_{k_1 \cdots k_l} (x, h_1, \cdots, h_1) \prod_{m=1}^{l} t_{m}^{k_m}.\]

Formally, \(w_i \in C\) and so \(t_i w_i \equiv u_i \in t_i C\). Then \(t_i dw_i = du_i\) and so \(dw_i = (1/t_i) du_i\).

This is why \(\prod_{m=1}^{l} t_{m}^{k_m+1}\) gets changed to \(\prod_{m=1}^{l} t_{m}^{k_m}\).

If \(t_m = 0\) for any \(m\), the result of both sides in the above equals zero due to the fact that \(\int_C w_{m+1} \prod_{m=1}^{l} h_{m}^{k_m} = 0\) whenever \(k_m \geq 1\).

To verify 6.13.38, use 6.13.37 to conclude

\[
\|a_{k_1 \cdots k_l}(x, h_1)\| \leq \left\| a_{k_1 \cdots k_l}(x, h_1) \right\| \prod_{m=1}^{l} \|h_m\|^{k_m}
\]

and \(\left\| a_{k_1 \cdots k_l}(x, h_1) \right\|\) is bounded by

\[
M \int_C \cdots \int_C \frac{1}{\prod_{m=1}^{l} |w_m|^{k_m+1}} \prod_{m=1}^{l} |w_m| \prod_{m=1}^{l} |w_m| \equiv C. \n\]

Lemma 6.13.3 Suppose

\[g(x + z h) = g(x) + \sum_{m=1}^{\infty} b_m(x, h) z^m\]

for all \(z\) small enough. Then

\[b_1(x, h_1 + h_2) = b_1(x, h_1) + b_1(x, h_2).\]

Proof: Recall that

\[f(x + \sum_{m=1}^{l} z_m h_m) = \sum_{k_l=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} a_{k_1 \cdots k_l}(x, h_1, \cdots, h_1) z_1^{k_1} \cdots z_l^{k_l}\]

and so one can write the following where \(g_{nm}\) is defined in the following expression.

\[g(x + z_1 h_1 + z_2 h_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} g_{mn}(x, h_1, h_2) z_1^m z_2^n.\]

Thus,

\[g(x + z_1 h_1) = \sum_{m=0}^{\infty} g_{m0}(x, h_1, h_2) z_1^m = g(x) + \sum_{m=1}^{\infty} b_m(x, h) z_1^m,\]
6.13. ANALYTIC FUNCTIONS

\[
g(x + z_2 h_2) = \sum_{n=0}^{\infty} g_{0n} (x, h_1, h_2) z_2^n = g(x) + \sum_{n=1}^{\infty} b_n (x, h) z_2^n
\]

which implies

\[
g_{n0} (x, h_1, h_2) = b_n (x, h_1), \; g_{0n} (x, h_1, h_2) = b_n (x, h_2).
\]

Now let \( z_1 = z_2 = z \). Then

\[
g(x + z (h_1 + h_2)) = g(x) + \sum_{n=0}^{\infty} b_n (x, h_1 + h_2) z^n
\]

\[= g(x) + z (g_{10} (x, h_1, h_2) + g_{01} (x, h_1, h_2)) + \text{higher order terms in } z.\]

Therefore,

\[
b_1 (x, h_1 + h_2) = g_{10} (x, h_1, h_2) + g_{01} (x, h_1, h_2)
\]

\[= b_1 (x, h_1) + b_1 (x, h_2). \]

**Lemma 6.13.4** Suppose \( a(x, h_1, \ldots, h_1) \) is multilinear, \((h_i \rightarrow a(x, h_1, \ldots, h_1) \) is linear),

\[
||a(x, h_1, \ldots, h_1)|| \leq C \prod_{m=1}^{l} ||h_m||,
\]

and

\[
D^{l-1} f(x) + h_1 (h_{l-1}) \cdots (h_1) - D^{l-1} f(x) (h_{l-1}) \cdots (h_1)
\]

\[-a(x, h_1, \ldots, h_1) = o(||h||).\]

Then \( D^{l} f(x) \) exists and

\[
D^{l} f(x) (h_1) (h_{l-1}) \cdots (h_1) = a(x, h_1, \ldots, h_1).
\]

**Proof:** If \( l = 1 \), the conclusion is obvious and is nothing more than the definition of the derivative.

\[
f(x + h) - f(x) - a(x, h) = o(||h||)
\]

and so from the definition of the derivative, \( a(x, h) = Df(x) h \).

Next let \( n = 2 \). By assumption,

\[
Df(x + h_1) (h_1) = Df(x) (h_1) - a(x, h_1, h_1) = o(||h||).
\]

Let \( L(x) \) be defined by

\[
L(x) (h_1) \equiv a(x, h, h_1).
\]

Then \( L(x) \in L(U, \mathcal{L}(X,Y)) \) because

\[
||L(x)|| \equiv \sup_{||h|| \leq 1} ||L(x) (h)|| \equiv \sup_{||h|| \leq 1} \sup_{||h_1|| \leq 1} ||L(x) (h_1) (h_1)|| \leq C.
\]
Also

\[ \|Df(x + h) - Df(x) - L(x)h\| \]

\[ \equiv \sup_{\|h_1\| \leq 1} \|Df(x + h_1)(h_1) - Df(x)(h_1) - L(x)(h_1)\| \]

\[ = \sup_{\|h_1\| \leq 1} \|Df(x)(h_1) - Df(x)(h_1) - a(x, h, h_1)\| = o(\|h\|) \]

and so \( L(x) = D^2f(x) \). Continuing in this way, we verify the conclusion of the lemma.

**Lemma 6.13.5** If \( f \) is analytic on \( U \), then \( f \in C^\infty(U) \). Also

**Proof:** By Lemma 6.13.3 applied to \( g = f \) and Lemma 6.13.2, \( Df(x) \) exists and

\[ Df(x)(h) = a_1(x, h). \]

These lemmas implied that \( h \to a_1(x, h) \) was linear. Suppose \( D^{l-1}f(x) \) exists for \( l \geq 2 \).

\[ f \left( x + \sum_{m=1}^{l} z_m h_m \right) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_1=0}^{\infty} a_{n_1 \cdots n_1}(x, h_1, \cdots, h_1) z_1^{n_1} \cdots z_l^{n_l}. \]

Differentiate with respect to \( z_1, \cdots, z_{l-1} \) to obtain

\[ D^{l-1}f \left( x + \sum_{m=1}^{l-1} z_m h_m + z_l h_l \right) (h_{l-1}) \cdots (h_1) = \]

\[ \sum_{n_1=0}^{\infty} \cdots \sum_{n_1=1}^{\infty} a_{n_1 \cdots n_1}(x, h_1, \cdots, h_1) \left( \prod_{m=1}^{l-1} n_m \right) z_1^{n_1-1} \cdots z_{l-1}^{n_{l-1}-1} z_l^{n_l}. \]

Take \( z_i = 0 \) for \( i = 1, \cdots, l-1 \). Then

\[ D^{l-1}f(x + z_l h_l)(h_{l-1}) \cdots (h_1) = \sum_{n_1=0}^{\infty} a_{n_1 \cdots 1}(x, h_1, \cdots, h_1) z_1^{n_1}. \quad (6.13.39) \]

Now we apply Lemma 6.13.3 to the function

\[ z_l \to D^{l-1}f(x + z_l h_l)(h_{l-1}) \cdots (h_1) \]

and conclude

\[ h_l \to a_{1 \cdots 1}(x, h_1, \cdots, h_1) \]

is linear. This involved taking \( n_l = 1 \) to get \( a_{1 \cdots 1}(x, h_1, \cdots, h_1) \). Thus from (6.13.39),

\[ D^{l-1}f(x + z_l h_l)(h_{l-1}) \cdots (h_1) - D^{l-1}f(x)(h_{l-1}) \cdots (h_1) \]

\[ = a_{1 \cdots 1}(x, h_1, \cdots, h_1) z_l + o(z_l h_l). \quad (6.13.40) \]
6.13. ANALYTIC FUNCTIONS

From this equation, it follows that

\[ a_{1\ldots l} \left( x, h_l \cdots h_i + \tilde{h}_i \cdots h_1 \right) z_l - a_{1\ldots l} \left( x, h_l \cdots h_i \right) z_l = a(z_l h_l) \]

because for each \( z_l \), the left side of (6.13.40) is linear in \( h_i \) for each \( i \leq l - 1 \). Dividing both sides of the above by \( z_l \) and then letting \( z_l \to 0 \), we see that \( a_{n_l1\ldots 1} \) is linear in each of the \( h_i \). Denoting \( z_l h_l \) by \( h_l \),

\[ D_l^{l-1} f(x + h_l) (h_{l-1}) \cdots (h_1) - D_l^{l-1} f(x) (h_{l-1}) \cdots (h_1) = a_{1\ldots l} \left( x, h_l, \cdots, h_1 \right) + o(\|h_l\|) \]

and so by Lemma 6.13.4, \( D_l f(x) \) exists and

\[ D_l f(x) (h_l) \cdots (h_1) = a_{1\ldots l} \left( x, h_l, \cdots, h_1 \right). \]

With these lemmas, the main result can be established. This is the generalization of the well known result for analytic functions.

**Theorem 6.13.6** Let \( X \) and \( Y \) be two complex Banach spaces and let \( U \) be an open set in \( X \). Then \( f : U \to Y \) is analytic on \( U \) if and only if \( Df(x) \) exists for each \( x \in U \) and in this case, \( f \in \mathcal{C}^\infty(U) \), and if \( h \in X \), then whenever \( z \) is small enough,

\[ f(x + zh) = f(x) + \sum_{n=1}^{\infty} \frac{D^n f(x) h^n z^n}{n!}. \]

**Proof:** We know

\[ f(x + zh) = f(x) + \sum_{n=1}^{\infty} a_n(x, h) z^n. \]

Differentiating, we obtain

\[ D^k f(x + zh) h^k = k! a_k(x, h) + \sum_{n=k+1}^{\infty} n (n - 1) \cdots (n - k + 1) z^{n-k}. \]

Letting \( z = 0 \) this shows

\[ D^k f(x) h^k = k! a_k(x, h) \]

and this proves half the theorem.

Conversely, if \( Df(x) \) exists on \( U \), it is clear that \( f \) is analytic on some ball, \( B(x, r) \subseteq U \), \( z \to f(y + zh) \) is analytic for \( y \in B(x, r) \) and small enough \( z \). Therefore the formula involving the series follows.
6.14 Ordinary Differential Equations

In this section we give an application to ordinary differential equations. To begin with, here are two Banach spaces which will be of use. Let \( Z \) be a complex Banach space and let \( X \) be the space of functions mapping \( B(0,1) \equiv D_1 \) to \( Z \) such that the functions are continuous on \( D_1 \) and analytic on \( B_1 \equiv B(0,1) \), the derivative is the restriction to \( B_1 \) of a continuous function defined on \( D_1 \), and the function equals 0 at 0.

\[
X \equiv \{ \phi \in C(D_1, X) : \phi(0) = 0 \}
\]

The norm on \( X \) will be

\[
||\phi||_X \equiv ||\phi||_\infty + ||\phi'||_\infty
\]

where

\[
||\phi||_\infty \equiv \sup \{||\phi(t)||_Z : t \in B_1\}.
\]

(Note that for a function continuous on \( D_1 \) it does not matter in the above definition of \( ||\cdot||_\infty \) whether we use \( B_1 \) or \( D_1 \) in the definition.) We define \( Y \) to be the space of continuous functions which are defined on \( D_1 \) having values in \( Z \) which are also analytic on \( B_1 \). The norm on \( Y \) is defined as

\[
||\phi||_\infty \equiv ||\phi||_Y.
\]

Note that \( B_1 \) is in \( \mathbb{C} \).

**Lemma 6.14.1** The spaces \( X \) and \( Y \) with the given norms are Banach spaces and if \( L : X \to Y \) is defined as \( L\phi(t) = \phi'(t) \) for all \( t \in B_1 \), then \( L \) is one to one, onto and continuous.

**Proof:** It is clear that \( X \) and \( Y \) are both normed linear spaces. It remains to show they are Banach spaces. Suppose \( \{\phi_n\} \) is a Cauchy sequence in \( X \). Then \( \phi_n \to \phi \) uniformly and \( \phi'_n \to \psi \) uniformly where \( \psi \) and \( \phi \) are continuous on \( D_1 \). We need to verify that \( \psi = \phi' \) on \( B_1 \). Letting \( C_1 \) be the unit circle, the Cauchy integral formula implies for \( t \in B_1 \),

\[
\phi(t) = \lim_{n \to \infty} \phi_n(t) = \lim_{n \to \infty} \frac{1}{2\pi i} \int_{C_1} \frac{\phi_n(w)}{w-t} \, dw = \frac{1}{2\pi i} \int_{C_1} \frac{\phi(w)}{w-t} \, dw
\]

which shows \( \psi'(t) \) exists on \( B_1 \). Also for \( t \in B_1 \),

\[
\psi(t) = \lim_{n \to \infty} \phi'_n(t) = \lim_{n \to \infty} \frac{1}{2\pi i} \int_{C_1} \frac{\phi_n(w)}{(w-t)^2} \, dw = \frac{1}{2\pi i} \int_{C_1} \frac{\phi(w)}{(w-t)^2} \, dw = \phi'(t).
\]

This shows \( X \) is a Banach space. A similar argument using the Cauchy integral theorem shows \( Y \) is a Banach space also. It is obvious that \( L \) is continuous. It remains to show \( L \) is one to one and onto.
Let \( \phi \in Y \). We need to show \( \phi = L\psi \) for some \( \psi \in X \). Let

\[
\psi (t) \equiv \int_{\Gamma} \phi (w) \, dw
\]

where \( \Gamma \) is any piecewise smooth curve from 0 to \( t \). By the Cauchy integral theorem, this definition is well defined and it is clear that \( \psi (0) = 0 \), \( \psi' (t) = \phi (t) \), and \( \psi \) is continuous on \( D_1 \). This shows \( L \) is onto.

It only remains to show \( L \) is one to one. Suppose \( L\phi = 0 \). Since \( \phi (0) = 0 \),

\[
\phi (t) = \int_0^t \phi' (ts) \, tds = 0
\]

if \( t \neq 0 \). But \( \phi (0) \) is given to equal zero. Thus \( L \) is one to one as claimed.

**Theorem 6.14.2** Let \( \Lambda \) and \( Z \) be complex Banach spaces and let \( W \) be an open subset of \( \mathbb{C} \times Z \times \Lambda \) containing \((0, y_0, \lambda)\). Also let \( f : W \to Z \) be analytic. Then there exists a unique \( y = y (y_0, \lambda) \) solving

\[
y' = f (t, y, \lambda), \quad y (0) = y_0
\]

valid for \( t \in D_\alpha \equiv B (0, |\alpha|) \) where \( \alpha = \alpha (y_0, \lambda) \). Furthermore, the map

\[
(t, y_0, \lambda) \to y (y_0, \lambda) (t)
\]

is analytic.

**Proof:** Let \( \alpha s = t \) and define \( \phi (s) \equiv y (t) - y_0 \). Then \( y \) is a solution to the system for \( t \in D_\alpha \) if and only if \( \phi \) is a solution for \( s \in D_1 \equiv B (0, 1) \) to the equations

\[
\phi' (s) = \alpha f (\alpha s, \phi (s) + y_0, \lambda), \quad \phi (0) = 0.
\]

Let \( X, Y, \) and \( L \) be given above and define

\[
\widetilde{W} \equiv \{(\alpha, \tilde{y}_0, \mu, \phi) \in \mathbb{C} \times Z \times \Lambda \times X : (s, \tilde{y}_0 + \phi (s), \mu) \in W \}
\]

for \( s \in D_1, (s, \tilde{y}_0 + \phi (s), \mu) \in W \).

For a given \( (\alpha, \tilde{y}_0, \mu, \phi) \in \widetilde{W} \),

\[
\{(s, \tilde{y}_0 + \phi (s), \mu) : s \in D_1\}
\]

is a compact subset of \( W \). This is because you have \( s \to (\alpha, \tilde{y}_0 + \phi (s), \mu) \) is the continuous image of a compact set which is assumed to be in \( W \). Consequently, the distance from this set to \( W \) is positive and so if \( (\beta, y_0, \lambda, \psi) \) is sufficiently close to \( (\alpha, \tilde{y}_0, \mu, \phi) \) in \( \mathbb{C} \times Z \times \Lambda \times X \) it follows \( (\beta, y_0, \lambda, \psi) \) is also in \( \widetilde{W} \). This shows \( \widetilde{W} \) is an open subset of \( \mathbb{C} \times Z \times \Lambda \times X \).

Now define \( F : \widetilde{W} \to Y \) (Recall that \( Y \) was a space of functions.) by

\[
F (\alpha, \tilde{y}_0, \mu, \phi) (s) \equiv L\phi (s) - \alpha f (\alpha s, \phi (s) + \tilde{y}_0, \mu).
\]
Then
\[ F(0, y_0, \lambda, 0) = L \phi = 0, \]
and \( F \) is analytic in \( \tilde{W} \). Also
\[ D_4 F(0, y_0, \lambda, 0) \psi = L \psi = \psi' \]
and so \( D_4 F(0, y_0, \lambda, 0) \in \mathcal{L}(X, Y) \), is one to one, onto and continuous by Lemma 6.14.1.

By the open mapping theorem, its inverse is also continuous. Therefore, the conditions of the implicit function theorem are satisfied and so there exists \( r > 0 \) such that if
\[ |\alpha| + \|\mu - \lambda\| + \|\tilde{y}_0 - y_0\| < r, \]
then there exists a unique \( \phi \in X \) such that
\[ F(\alpha, \tilde{y}_0, \mu, \phi) = 0, \]
and \( \phi \) is an analytic function of \((\alpha, \tilde{y}_0, \mu)\). Fixing \( 0 < \alpha < r \), it follows
\[ (\tilde{y}_0, \mu) \rightarrow y(\tilde{y}_0, \mu) \]
is analytic on an open subset of \( Z \times \Lambda \). Also \( t \rightarrow y(\tilde{y}_0, \mu)(t) \) is an analytic function because of the definition of \( y \) in terms of \( \phi, \phi(s) \equiv y(t) - y_0 \). It follows that for \( t \in B(0, |\alpha|) \),
\[ (\tilde{y}_0, \mu) \rightarrow y(\tilde{y}_0, \mu)(t) \text{ and } t \rightarrow y(\tilde{y}_0, \mu)(t) \]
are both analytic. ■

6.15 Tensor Products

**Definition 6.15.1** Denote by \( B(X \times Y, K) \) the bilinear forms having values in \( K \) a vector space. Define \( \otimes : X \times Y \rightarrow \mathcal{L}(B(X \times Y, K), K) \)
\[ x \otimes y(A) \equiv A(x, y) \]
where \( A \) is a bilinear form. Often \( K \) is a field, but it could be any vector space with no change.

**Remark 6.15.2** \( \otimes \) is bilinear. That is
\[
(ax_1 + bx_2) \otimes y = a(x_1 \otimes y) + b(x_2 \otimes y) \\
x \otimes (ay_1 + by_2) = a(x \otimes y_1) + b(x \otimes y_2)
\]
This follows right away from the definition. Note that \( x \otimes y \in \mathcal{L}(B(X \times Y, K), K) \) because it is linear on \( B(X \times Y, K) \).
\[ x \otimes y(aA + bB) \equiv aA(x, y) + bB(x, y) = a(x \otimes y)(A) + b(x \otimes y)(B) \]
6.15. TENSOR PRODUCTS

**Definition 6.15.3** Define \( X \otimes Y \) as the span of the \( x \otimes y \) in \( \mathcal{L} ( B ( X \times Y, K ), K ) \).

**Lemma 6.15.4** Let \( V \leq X \). That is, \( V \neq 0 \) is a subspace of \( X \). Let \( B_V \) be a Hamel basis for \( V \). Then there exists a Hamel basis for \( X, B_X \) which includes \( B_V \).

**Proof**: Consider \( ( Y, B_Y ) \) where \( Y \supseteq V \) and \( B_Y \supseteq B_V \). Call it \( \mathcal{F} \). Let \( ( Y, B_Y ) \leq ( \hat{Y}, B_{\hat{Y}} ) \) if and only if \( Y \leq \hat{Y} \) and \( B_Y \subseteq B_{\hat{Y}} \). Then let \( \mathcal{C} \) denote a maximal chain. Let \( ( \hat{X}, B_{\hat{X}} ) \) consist of the union of all \( Y \) in the chain along with the union of all \( B_Y \) in the chain. Then in fact this is something in \( \mathcal{F} \) thanks to this being a chain. Thus \( \hat{X} = \text{span} ( B_{\hat{X}} ) \). Then it must be the case that \( \hat{X} = X \). If not, there exists \( y \notin \text{span} ( B_{\hat{X}} ) \). But then \( ( \text{span} ( B_{\hat{X}}, y ), \{ B_{\hat{X}} \}, \{ y \} ) \) could be added to \( \mathcal{C} \) and obtain a strictly longer chain. ■

Next, 
\[
x \otimes ( ay + bz ) ( A ) \equiv A ( x, ay + bz ) = aA ( x, y ) + bA ( x, z ) \\
= ax \otimes y ( A ) + bx \otimes z ( A ) \equiv ( ax \otimes y + bx \otimes z ) ( A )
\]
and so \( \otimes \) distributes across addition from the left and you can factor out scalars. Similarly, it will work the same way from the right.

**Lemma 6.15.5** Let \( B_X \) be a Hamel basis for \( X \). Then define \( L : X \rightarrow Y \) as follows. For each \( x \in B_X \), let \( L x \equiv y_x \). Then for arbitrary \( z \), define \( Lz \) as follows. For \( z = \sum_{x \in B_X} c_x x \) where this denotes a finite sum consisting of the unique linear combination of basis vectors which equals \( x \),
\[
Lz \equiv \sum_{x \in B_X} c_x y_x
\]
Then \( L \) is linear.

**Proof**: There is only one finite linear combination equal to \( z \) thanks to linear independence. Thus \( L \) is well defined. Why is it linear? Let \( z = \sum_{x \in B_X} c_x x, w = \sum_{x \in B_X} c_w x \). Then \( L ( ax + bw ) = \)
\[
L \left( a \sum_{x \in B_X} c_x x + b \sum_{x \in B_X} c_w x \right) \equiv a \sum_{x \in B_X} c_x y_x + b \sum_{x \in B_X} c_w y_x = aLz + bLw
\]
Thus \( L \) is linear. ■

**Proposition 6.15.6** Let \( X, Y \) be vector spaces and let \( E \) and \( F \) be linearly independent subsets of \( X, Y \) respectively. Then \( \{ e \otimes f : e \in E, f \in F \} \) is linearly independent in \( X \otimes Y \).

**Proof**: Let \( \kappa \in K \). Say \( \sum_{i=1}^{n} \lambda_i e_i \otimes f_i = 0 \) in \( X \otimes Y \). This means it sends everything in \( B ( X \times Y; K ) \) to 0. Let \( \psi \in F^*, \psi ( f_k ) = 1 \) and \( \psi ( f_i ) = 0 \) for \( i \neq k \). Let \( \phi \in E^* \) be defined similarly. What you do is extend \( \{ e_i \} \) to a Hamel basis and
then define $\phi$ to equal 1 at $e_k$ and $\phi$ sends every other thing in the Hamel basis to 0. Then you look at $\phi(x) \psi(y) \kappa \equiv A(x,y)$. Then you have $0 = \sum_{i=1}^{n} \lambda_i e_i \otimes f_i(A) \equiv \lambda_k \phi(e_k) \psi(f_k) \kappa = \lambda_k \kappa$. Since $\kappa$ is arbitrary it must be that $\lambda_k = 0$. Thus these are linearly independent.

**Proposition 6.15.7** Suppose $u = \sum_{i=1}^{n} x_i \otimes y_i$ is in $X \otimes Y$ and is a shortest representation. Then $\{x_i\}, \{y_i\}$ are linearly independent. All such shortest representations have the same length. If $\sum_i x_i \otimes y_i = 0$ and $\{y_i\}$ are independent, then $x_i = 0$ for each $i$. In particular, $x \otimes y = 0$ iff $x = 0$ or $y = 0$.

**Proof:** Suppose the first part. If $\{y_i\}$ are not linearly independent, then one is a linear combination of the others. Say $y_n = \sum_{j=1}^{n-1} a_j y_j$. Then

$$\sum_{i=1}^{n} x_i \otimes y_i = \sum_{i=1}^{n-1} x_i \otimes y_i + x_n \otimes \sum_{j=1}^{n-1} a_j y_j = \sum_{i=1}^{n-1} x_i \otimes y_i + \sum_{i=1}^{n-1} a_i x_n \otimes y_i = \sum_{i=1}^{n-1} x_i \otimes y_i + a_i x_n \otimes y_i = \sum_{i=1}^{n-1} (x_i + a_i x_n) \otimes y_i$$

and so $n$ was not smallest after all. Similarly $\{x_i\}$ must be linearly independent. Now suppose that

$$\sum_{i=1}^{n} x_i \otimes y_i = \sum_{j=1}^{m} u_j \otimes v_j$$

and both are of minimal length. Why is $n = m$? We know that $\{x_i\}, \{y_i\}, \{u_j\}, \{v_j\}$ are independent. Let $\psi_k(y_k) = 1$, $\psi$ sends all other vectors to 0. Then do both sides to $A(x,y) \equiv \phi(x) \psi_k(y) \kappa$ where $\kappa \in K$ is given, $\phi \in X^*$ arbitrary.

$$\sum_{i=1}^{n} x_i \otimes y_i(A) = \sum_{i=1}^{n} \phi(x_i) \psi_k(y_i) \kappa = \phi(x_k) \kappa = \sum_{j=1}^{m} \phi(u_j) \psi(v_j) \kappa$$

Hence

$$\phi \left( x_k - \sum_{j=1}^{m} u_j \psi(v_j) \right) \kappa = 0$$

and this holds for any $\phi \in X^*$ and for any $\kappa$. Hence $x_k \in \text{span}(\{u_j\})$. Thus $m \geq n$ since this can be done for each $k$. $\{x_1, \ldots, x_n\} \subseteq \text{span}(u_1, \ldots, u_m)$ and the
left side is independent and is contained in the span of the right side. Similarly, \(\{u_1, \ldots, u_m\} \subseteq \text{span} \,(x_1, \ldots, x_n)\) and so \(m \leq n\). Thus \(m = n\).

Next suppose that \(\sum_{i=1}^n x_i \otimes y_i = 0\) and the \(\{y_i\}\) are linearly independent. Then let \(\psi_k(y_k) = 1\) and \(\psi_k\) sends the other vectors to 0. Then do the sum to \(A(x, y) = \phi(x) \psi(y) \kappa\) for \(\kappa \in K\). This yields \(\phi(x_k) \kappa = 0\) for every \(\phi\). Hence \(x_k = 0\). This is so for each \(k\). Similarly, if \(\sum_{i=1}^n x_i \otimes y_i = 0\) and \(\{x_i\}\) independent, then each \(y_i = 0\).

The next theorem is very interesting. It is concerned with bilinear forms having values in \(V\) a vector space \(\psi: X \times Y \to V\). Roughly, it says there is a unique linear map from \(X \otimes Y\) which delivers the given bilinear form.

**Theorem 6.15.8** Suppose \(\psi\) is a bilinear map in \(B(X \times Y; V)\) where \(V\) is a vector space. Let \(\otimes: X \times Y \to X \otimes Y\) be given by \(\otimes(x, y) \equiv x \otimes y\). Then there exists a unique linear map \(\hat{\psi} \in \mathcal{L}(X \otimes Y; V)\) such that \(\hat{\psi} \otimes = \psi\). In other words, the following diagram commutes.

\[
\begin{array}{ccc}
X \times Y & \xrightarrow{\otimes} & X \otimes Y \\
\downarrow & & \downarrow \\
V & \xrightarrow{\psi} & V \\
\end{array}
\]

That is, \(\hat{\psi}(x \otimes y) = \psi(x, y)\).

**Proof:** Let \(E\) be a Hamel basis for \(X\) and let \(F\) be one for \(Y\). Then by definition, \(\{e \otimes f : e \in E, f \in F\}\) spans \(X \otimes Y\). To see this, suppose you have \(x \otimes y\). Then \(x = \sum_i a_i e_i, y = \sum_j b_j f_j\) and so

\[
x \otimes y = \sum_i a_i e_i \otimes \sum_j b_j f_j = \sum_{i,j} a_i b_j e_i \otimes f_j
\]

Also, it was shown above in Proposition 6.15.6 that this set is linearly independent. Therefore, it is a Hamel basis and you can define \(\hat{\psi}\) as follows: \(\hat{\psi}(e \otimes f) = \psi(e, f)\) and extend \(\hat{\psi}\) linearly to get the desired linear transformation \(\hat{\psi}\). It is unique because its value on a Hamel basis is completely determined by the values of \(\psi\).

Thus \(\hat{\psi}\) gives the right thing on \(E \times F\). It gives the right thing on \(X \times Y\) also. To see this, suppose you have \(x \otimes y = \sum_{i,j} a_i b_j e_i \otimes f_j\) as above. Then

\[
\hat{\psi}(x \otimes y) = \sum_{i,j} a_i b_j \psi(e_i \otimes f_j)
\]

\[
= \sum_{i,j} a_i b_j \psi(e_i, f_j)
\]

\[
= \psi \left( \sum_i a_i e_i, \sum_j b_j f_j \right) = \psi(x, y)
\]
You could define by analogy $X_1 \otimes X_2 \otimes \cdots \otimes X_p$ and all the above results would have analogs in this situation. Thus $x_1 \otimes x_2 \otimes \cdots \otimes x_p$ acts on a $p$ linear form $A$ as follows.

$$(x_1 \otimes x_2 \otimes \cdots \otimes x_p) A \equiv A(x_1, x_2, \cdots, x_p)$$

Then the above theorem has an obvious corollary.

**Corollary 6.15.9** Suppose $\psi$ is a $p$ linear map in $B(X_1 \times \cdots \times X_p; V)$ where $V$ is a vector space. Let $\otimes^p : X_1 \times \cdots \times X_p \to X_1 \otimes X_2 \otimes \cdots \otimes X_p$ be given by $\otimes^p (x_1, x_2, \cdots, x_p) \equiv x_1 \otimes x_2 \otimes \cdots \otimes x_p$. Then there exists a unique linear map $\hat{\psi} \in \mathcal{L}(X_1 \otimes X_2 \otimes \cdots \otimes X_p; V)$ such that $\hat{\psi} \circ \otimes^p = \psi$. In other words, the following diagram commutes.

$$
\begin{array}{ccc}
X_1 \times \cdots \times X_p & \xrightarrow{\otimes^p} & X_1 \otimes X_2 \otimes \cdots \otimes X_p \\
& \downarrow \psi & \downarrow \hat{\psi} \\
& V & \\
\end{array}
$$

### 6.15.1 The Norm In Tensor Product Space

Let $X, Y$ be Banach spaces. Then we have the following definition.

**Definition 6.15.10** We define for $u \in X \otimes Y$

$$
\pi (u) \equiv \inf \left\{ \sum_i \|x_i\| \|y_i\| : u = \sum_i x_i \otimes y_i \right\}
$$

In this context, it is assumed that the elements of $X \otimes Y$ act on continuous bilinear forms. That is $A$ is bilinear and

$$
|A(x, y)| \leq \|A\| \|x\| \|y\|
$$

When we write $B(X \times Y; V)$ we mean the continuous ones. Here $V$ is a normed vector space and the underlying field for all is $F$.

**Proposition 6.15.11** $\pi (u)$ is well defined and is a norm. Also, $\pi (x \otimes y) = \|x\| \|y\|$.

**Proof:** It is obviously well defined. Let $\lambda \in F$ where $F$ is the field of interest. Assume $\lambda \neq 0$ since otherwise there is nothing to show. Then if $u = \sum_i x_i \otimes y_i$.

$$
\pi (\lambda u) \leq \sum_i \|\lambda x_i\| \|y_i\| = |\lambda| \sum_i \|x_i\| \|y_i\|
$$

Then taking inf, it follows that

$$
\pi (\lambda u) \leq |\lambda| \pi (u)
$$
Then from what was just shown for arbitrary nonzero \( \lambda \),
\[
\pi (u) = \pi (\lambda^{-1} \lambda u) \leq \frac{1}{|\lambda|} \pi (\lambda u)
\]
and so \(|\lambda| \pi (u) \leq \pi (\lambda u)\). Hence \( \pi (\lambda u) = |\lambda| \pi (u) \).

Let \( u = \sum_i x_i \otimes y_i, v = \sum_j \hat{x}_j \otimes \hat{y}_j \) such that
\[
\pi (u) + \varepsilon > \sum_i \|x_i\| \|y_i\|, \quad \pi (v) + \varepsilon > \sum_j \|\hat{x}_j\| \|\hat{y}_j\|.
\]
Then
\[
\pi (u + v) \leq \sum_i \|x_i\| \|y_i\| + \sum_j \|\hat{x}_j\| \|\hat{y}_j\| \leq \pi (u) + \pi (v) + 2\varepsilon.
\]
Since \( \varepsilon \) is arbitrary, \( \pi (u + v) \leq \pi (u) + \pi (v) \).

Now suppose \( \pi (u) = 0 \). Does it follow that \( u(A) = 0 \) for all \( A \in B (X \times Y, V) \)?
Let \( A \in B (X \times Y; V) \) and let \( u = \sum_i x_i \otimes y_i \) such that
\[
\varepsilon / \|A\| = \pi (u) + \varepsilon / \|A\| > \sum_i \|x_i\| \|y_i\|
\]
Then
\[
u(A) \equiv \sum_i A (x_i, y_i) \leq \sum_i \|A\| \|x_i\| \|y_i\| < (\varepsilon / \|A\|) \|A\| = \varepsilon
\]
and since \( \varepsilon \) is arbitrary, this requires \( u(A) = 0 \). Since \( A \) is arbitrary, this requires \( u = 0 \).

Next is the very interesting equality that \( \pi (x \otimes y) = \|x\| \|y\| \). It is obvious that
\[
\pi (x \otimes y) \leq \|x\| \|y\|
\]
because one way to write \( x \otimes y \) is \( x \otimes y \). Let \( \phi (x) = \|x\|, \psi (y) = \|y\| \) where \( \|\phi\|, \|\psi\| = 1 \). Here \( \phi, \psi \in X', Y' \) respectively. You get them from the Hahn Banach theorem. Then consider the continuous bilinear form \( A (\hat{x}, \hat{y}) \equiv \phi (\hat{x}) \psi (\hat{y}) \). Say \( x \otimes y = \sum_i x_i \otimes y_i \). There is a linear map \( \psi \in \mathcal{L}(X \otimes Y, V) \) such that \( \psi (\hat{x} \otimes \hat{y}) = A (\hat{x}, \hat{y}) \). You just specify this on all things of the form \( e \otimes f \) where \( e \in E \) a Hamel basis for \( X \) and \( f \in F \), a Hamel basis for \( Y \). Then it must hold for the linear span of these things which would yield the desired result. Hence, in particular,
\[
\|x\| \|y\| = \phi (x) \psi (y) = |A(x, y)| = |\psi (x \otimes y)| = \left| \sum \psi (x_i \otimes y_i) \right|
\]
\[
= \left| \sum_i A(x_i, y_i) \right| = \left| \sum_i \phi(x_i) \psi(y_i) \right| \leq \sum_i \|x_i\| \|y_i\|
\]
It follows that on taking inf of both sides over all such representations of \( x \otimes y \) that
\[
\|x\| \|y\| \leq \pi (x \otimes y). \quad \blacksquare
\]

There is no difference if you replace \( X \otimes Y \) with \( X_1 \otimes X_2 \otimes \cdots \otimes X_p \). One modifies the definition as follows.
Definition 6.15.12 We define for \( u \in X_1 \otimes X_2 \otimes \cdots \otimes X_p \)

\[
\pi (u) \equiv \inf \left\{ \sum_{i} \prod_{j=1}^{p} \| x^j_i \| : u = \sum_{i} x^1_i \otimes x^2_i \otimes \cdots \otimes x^p_i \right\}
\]

In this context, it is assumed that the elements of \( X_1 \otimes X_2 \otimes \cdots \otimes X_p \) act on continuous \( p \) linear forms. That is \( A \) is \( p \) linear and \( |A(x_1, \cdots, x_p)| \leq \|A\| \prod_i \|x_i\| \)

Corollary 6.15.13 \( \pi (u) \) is well defined and is a norm. Also,

\[
\pi \left( x^1 \otimes x^2 \otimes \cdots \otimes x^p \right) = \prod_k \| x^k \|.
\]

Recall Corollary

\[
X_1 \times \cdots \times X_p \\
\otimes^p \downarrow \psi \\
X_1 \otimes X_2 \otimes \cdots \otimes X_p \rightarrow V
\]

Is it the case that \( \hat{\psi} \) is continuous? Letting \( a \in X_1 \otimes X_2 \otimes \cdots \otimes X_p \), does it follow that \( \| \hat{\psi} (a) \|_V \leq \|a\|_{X_1 \otimes X_2 \otimes \cdots \otimes X_p} \)?

Let \( a = \sum_{i} x^1_i \otimes x^2_i \otimes \cdots \otimes x^p_i \). Then

\[
\left\| \hat{\psi} (a) \right\|_V = \left\| \hat{\psi} \left( \sum_{i} x^1_i \otimes x^2_i \otimes \cdots \otimes x^p_i \right) \right\|_V = \left\| \sum_{i} \hat{\psi} (x^1_i \otimes x^2_i \otimes \cdots \otimes x^p_i) \right\|_V \\
= \left\| \sum_{i} A (x^1_i, \cdots, x^p_i) \right\|_V \leq \|A\| \sum_{i} \prod_{m=1}^{p} \|x^m_i\|
\]

Then taking the inf over all such representations of \( a \), it follows that \( \left\| \hat{\psi} (a) \right\| \leq \|A\| \|a\|_{X_1 \otimes X_2 \otimes \cdots \otimes X_p} \). This proves the following theorem.

Theorem 6.15.14 Suppose \( A \) is a continuous \( p \) linear map in \( B (X_1 \times \cdots \times X_p; V) \) where \( V \) is a vector space. Let \( \otimes^p : X_1 \times \cdots \times X_p \rightarrow X_1 \otimes X_2 \otimes \cdots \otimes X_p \) be given by \( \otimes^p (x_1, x_2, \cdots, x_p) = x_1 \otimes x_2 \otimes \cdots \otimes x_p \). Then there exists a unique linear map \( \hat{\psi} \in \mathcal{L} (X_1 \otimes X_2 \otimes \cdots \otimes X_p; V) \) such that \( \hat{\psi} \circ \otimes^p = \psi \). In other words, the following diagram commutes.

\[
X_1 \times \cdots \times X_p \\
\otimes^p \downarrow \psi \\
X_1 \otimes X_2 \otimes \cdots \otimes X_p \rightarrow V
\]

Also \( \hat{\psi} \) is a continuous linear mapping.
6.15. TENSOR PRODUCTS

6.15.2 The Taylor Formula And Tensors

Let $V$ be a Banach space and let $P$ be a function defined on an open subset of $V$ which has values in a Banach space $W$. Recall the nature of the derivatives. $DP(x) \in \mathcal{L}(V,W)$. Thus $v \to DP(x)(v)$ is a differentiable map from $V$ to $W$. Then from the definition,

$$\hat{v} \to D^2P(x)(\hat{v})$$

is again a differentiable function with values in $W$. $D^2P(x) \in \mathcal{L}(V,\mathcal{L}(V,W))$. We can write it in the form

$$D^2P(x)(v,\hat{v})$$

Of course it is linear in both variables from the definition. Similarly, we denote $P^j(x)$ as a $j$ linear form which is the $j$th derivative. In fact it is a symmetric $j$ linear form as long as it is continuous. To see this, consider the case where $j = 3$ and $P$ has values in $\mathbb{R}$. Then let

$$P(x + tu + sv + rw) = h(t, s, r)$$

By equality of mixed partial derivatives, $h_{tsr}(0, 0, 0) = h_{str}(0, 0, 0)$ etc. Thus

$$h_{tsr}(t, s, r) = P^3(x + tu + sv + rw)(u, v, w)$$

and so

$$P^3(x)(u, v, w) = P^3(x)(v, u, w)$$

e tc. If it has values in $W$ you just replace $P$ with $\phi(P)$ where $\phi \in W'$ and use this result to conclude that $\phi(P^3(u, v, w)) = \phi(P^3(v, u, w))$ etc. Then since the dual space separates points, the desired result is obtained. Similarly, $D^3P(x)(v, u, w) = D^3P(x)(u, w, v)$ etc.

Recall the following proof of Taylor’s formula:

$$h(t) = P(u + tv)$$

$$h(t) = h(0) + h'(0)t + \cdots + \frac{1}{k!}h^k(0)t^k + \frac{1}{(k+1)!}h^{(k+1)}(s)t^{k+1}, |s| < |t|$$

Now

$$h'(t) = DP(u + tv)v,$$

$$h''(t) = D^2P(u + tv)(v)v,$$

$$h'''(t) = D^3P(u + tv)(v)(v)$$

e tc. Thus letting $t = 1,$

$$P(u + v) = P(u) + DP(u)v + \frac{1}{2!}D^2P(u)(v,v) + \cdots + \frac{1}{k!}D^kP(u)(v,\cdots,v) + o(|v|^k)$$

Now you can use the representation theorem Theorem 6.15.14 to write this in the following form.

$$P(u + v) = P(u) + P^1(u)v + \frac{1}{2!}P^2(u)v^\otimes 2 + \cdots + \frac{1}{k!}P^k(u)v^\otimes k + o(|v|^k)$$
where \( f^j(u) \in L(V^\otimes j, W) \). I think that the reason this is somewhat attractive is that the \( P^k \) are linear operators. Recall \( P^k(u) v^\otimes k = D P^k(u)(v, \ldots, v) \).

In the following exercises \( \uparrow \) indicates you should consider the previous problem.

### 6.16 Exercises

1. Suppose \( L \in L(X, Y) \) where \( X \) and \( Y \) are two finite dimensional normed vector spaces and suppose \( L \) is one to one. Show there exists \( r > 0 \) such that for all \( x \in X \),
   \[
   \|Lx\| \geq r \|x\|.
   \]
   **Hint:** Show that \( \|x\| \equiv \|Lx\| \) is a norm. Now suppose \( L \in L(X, Y) \) is one to one and onto for \( X, Y \) Banach spaces. Explain why the same result holds. **Hint:** Recall open mapping theorem.

2. Suppose \( U \subseteq X \) is an open subset of \( X \) a Banach space and that \( f : U \to Y \) is differentiable at \( x_0 \in U \) such that \( Df(x_0) \) is one to one and onto from \( X \) to \( Y \). \( (Df(x_0)^{-1} \in L(Y, X)) \) Then show that \( f(x) \neq f(x_0) \) for all \( x \) sufficiently near but not equal to \( x_0 \). In this case, you only know the derivative exists at \( x_0 \).

3. Suppose \( X, Y \) are Banach spaces and \( B \) is an open ball in \( X \), a Banach space, and \( f : B \to Y \) is differentiable. Suppose also there exists \( L \in L(X, Y) \) such that
   \[
   ||Df(x) - L|| < k
   \]
   for all \( x \in B \). Show that if \( x_1, x_2 \in B \),
   \[
   ||f(x_1) - f(x_2) - L(x_1 - x_2)|| \leq k \|x_1 - x_2\|.
   \]
   **Hint:** Consider \( Tx = f(x) - Lx \) and argue \( \|DT(x)\| < k \).

4. \( \uparrow \) Let \( U \) be an open subset of \( X \), \( f : U \to Y \) where \( X, Y \) are finite dimensional normed linear spaces and suppose \( f \in C^1(U) \) and \( Df(x_0) \) is one to one. Then show \( f \) is one to one near \( x_0 \). **Hint:** Show using the assumption that \( f \) is \( C^1 \) that there exists \( \delta > 0 \) such that if
   \[
   x_1, x_2 \in B(x_0, \delta),
   \]
   then
   \[
   |f(x_1) - f(x_2) - Df(x_0)(x_1 - x_2)| \leq \frac{k}{2} |x_1 - x_2| \tag{6.16.42}
   \]
   then use Problem \( \Box \). In case \( X, Y \) are Banach spaces, assume \( Df(x_0) \) is one to one and onto and continuous.

5. Let \( f(z) = e^z \) for \( z \) complex. Then \( Df(z) = e^z \neq 0 \). However, \( f \) is not one to one. Does this contradict the above problem?
6. Suppose $M \in \mathcal{L}(X,Y)$ where $X$ and $Y$ are finite dimensional linear spaces and suppose $M$ is onto. Show there exists $L \in \mathcal{L}(Y,X)$ such that

$$LMx = Px$$

where $P \in \mathcal{L}(X,X)$, and $P^2 = P$. Also show $L$ is one to one and onto and $M$ is one to one on $PX$. **Hint:** Let $\{y_1 \cdots y_n\}$ be a basis of $Y$ and let $Mx_i = y_i$. Then define

$$Ly = \sum_{i=1}^{n} \alpha_i x_i \text{ where } y = \sum_{i=1}^{n} \alpha_i y_i.$$ 

Show $\{x_1, \cdots, x_n\}$ is a linearly independent set and show you can obtain $\{x_1, \cdots, x_n, \cdots, x_m\}$, a basis for $X$ in which $Mx_j = 0$ for $j > n$. Then let

$$Px = \sum_{i=1}^{n} \alpha_i x_i$$

where

$$x = \sum_{i=1}^{m} \alpha_i x_i.$$ 

7. \uparrow Let $f : U \subseteq X \to Y$, $f$ is $C^1$, and $Df(x)$ is onto for each $x \in U$. Then show $f$ maps open subsets of $U$ onto open sets in $Y$. **Hint:** Let $P = LDf(x)$ as in Problem 3. Argue $L$ maps open sets from $Y$ to open sets of $X_1 \equiv PX$ and $L^{-1}$ maps open sets from $X_1$ to open sets of $Y$. Then $Lf(x + v) = Lf(x) + LDf(x)v + o(v)$. Now for $z \in X_1$, let $h(z) = Lf(x + z) - Lf(x)$. Then $h$ is $C^1$ on some small open subset of $X_1$ containing 0 and $Dh(0) = LDf(x)$ which is seen to be one to one and onto and in $\mathcal{L}(X_1, X_1)$. Therefore, if $r$ is small enough, $h(B(0,r))$ equals an open set in $X_1$, $V$. This is by the inverse function theorem. Hence $Lf(x + B(0,r)) - f(x) = V$ and so $f(x + B(0,r)) - f(x) = L^{-1}(V)$, an open set in $Y$.

8. There is a version of the above for Banach spaces. Let $f : U \subseteq X \to Y$, $f$ is $C^1$, and $Df(x)$ is onto for each $x \in U$. Suppose that whenever $A \in \mathcal{L}(X,Y)$ is onto, there exists a closed subspace $X_1$ such that $X = X_1 \oplus \ker(A)$. Then show $f$ maps open subsets of $U$ onto open sets in $Y$. **Hint:** Let $x \in U$. For $z \in X_1 \oplus \ker(Df(x))$ and small, consider $h(z) \equiv f(x + z) - f(x), h(0) = 0$. Then

$$h(z) - h(0) = Dh(0)z + o(z) = f(x + z) - f(x) = Df(x)z + o(z)$$

and so $Dh(0)z = Df(x)z$. Now $Df(x)$ is given to map $X$ onto $Y$. Thus $Dh(0)$ must also map $X_1$ onto $Y$. But $Df(x)$ must be one to one on $X_1$. Explain why. Now do a similar argument to the above using the inverse function theorem.
9. Suppose $U \subseteq \mathbb{R}^2$ is an open set and $f : U \to \mathbb{R}^3$ is $C^1$. Suppose $Df(s_0, t_0)$ has rank two and

$$f(s_0, t_0) = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}.$$ 

Show that for $(s, t)$ near $(s_0, t_0)$, the points $f(s, t)$ may be realized in one of the following forms.

\[
\{(x, y, \phi(x, y)) : (x, y) \text{ near } (x_0, y_0)\},
\]

or

\[
\{(x, \phi(x, z), z) : (x, z) \text{ near } (x_0, z_0)\}.
\]

This shows that parametrically defined surfaces can be obtained locally in a particularly simple form.

10. Let $f : U \to Y$, $Df(x)$ exists for all $x \in U$, $B(x_0, \delta) \subseteq U$, and there exists $L \in \mathcal{L}(X, Y)$, such that $L^{-1} \in \mathcal{L}(Y, X)$, and for all $x \in B(x_0, \delta)$

$$||Df(x) - L|| < \frac{r}{||L^{-1}||}, \quad r < 1.$$ 

Show that there exists $\varepsilon > 0$ and an open subset of $B(f(x_0), \varepsilon)$, $V$, such that $f : V \to B(f(x_0), \varepsilon)$ is one to one and onto. Also $Df^{-1}(y)$ exists for each $y \in B(f(x_0), \varepsilon)$ and is given by the formula

$$Df^{-1}(y) = \left[Df(f^{-1}(y))\right]^{-1}.$$ 

**Hint:** Let

$$T_y(x) \equiv T(x, y) \equiv x - L^{-1}(f(x) - y)$$

for $|y - f(x_0)| < \frac{(1-r)\delta}{2||L^{-1}||}$, consider $\{T^{\alpha}_y(x_0)\}$. This is a version of the inverse function theorem for $f$ only differentiable, not $C^1$.

11. Denote by $C([0, T], X)$ the space of functions which are continuous having values in $X$ and define a norm on this linear space as follows.

$$||f||_\lambda \equiv \max \{|f(t)| e^{\lambda t} : t \in [0, T]\}.$$ 

Show for each $\lambda \in \mathbb{R}$, this is a norm and that $C([0, T], X)$ is a complete normed linear space with this norm.

12. Let $X$ be a Banach space and let $f : [0, T] \times X \to X$ be continuous and suppose $f$ satisfies a Lipschitz condition,

$$|f(t, x) - f(t, y)| \leq K|x - y|$$
and let \( x_0 \in X \). Show there exists a unique solution to the Cauchy problem,

\[
x' = f(t, x), \quad x(0) = x_0,
\]

for \( t \in [0, T] \). **Hint:** Consider the map

\[
G : C([0, T] ; X) \rightarrow C([0, T] ; X)
\]

defined by

\[
Gx(t) \equiv x_0 + \int_0^t f(s, x(s)) \, ds,
\]

where the integral is defined componentwise. Show \( G \) is a contraction map for \( ||·||_{\lambda} \) given in Problem 11 for a suitable choice of \( \lambda \) and that therefore, it has a unique fixed point in \( C([0, T] ; X) \). Next argue, using the fundamental theorem of calculus, that this fixed point is the unique solution to the Cauchy problem.

13. ↑Use Theorem \( \lambda \lambda \) to give another proof of the above theorem. **Hint:** Use the same mapping and show that a large power is a contraction map.

14. Suppose you know that \( u(t) \leq a + \int_0^t k(s) u(s) \, ds \) where \( k(s) \geq 0 \) and \( k, ku \) are in \( L^1([0, T]) \). Show that then \( u(t) \leq a \exp \left( \int_0^t k(s) \, ds \right) \). This is a version of Gronwall’s inequality. **Hint:** Let \( W(t) = \int_0^t k(s) u(s) \, ds \). Then explain why \( W'(t) - k(t) W(t) \leq ak(t) \). Now use the usual technique of an integrating factor you saw in beginning differential equations.

15. Suppose in the above problem, you have \( a \) replaced with \( a(t) \) where \( t \to a(t) \) is increasing. Show that

\[
u(t) \leq a(t) e^{\int_0^t k(s) \, ds} \]

16. ↑Use the above Gronwall’s inequality to establish a result of continuous dependence on the initial condition and \( f \) in the ordinary differential equation of Problem 12.

17. The existence of partial derivatives does not imply continuity as was shown in an example. However, much more can be said than this. Consider

\[
f(x, y) = \begin{cases} (x^2 - y^4)^2 & \text{if } (x, y) \neq (0, 0), \\ 1 & \text{if } (x, y) = (0, 0). \end{cases}
\]

Show the directional derivative of \( f \) at \((0, 0)\) exists and equals 0 for every direction. The directional derivative in the direction \((v_1, v_2)\) is defined as

\[
\lim_{t \to 0} \frac{f(x + tv_1, y + tv_2) - f(x, y)}{t}.
\]

Now consider the curve \( x^2 = y^4 \) and the curve \( y = 0 \) to verify the function fails to be continuous at \((0, 0)\).
18. Let 
\[
f(x, y) = \begin{cases} 
 \frac{x^2 y^n}{x^2 + y^n} & \text{if } (x, y) \neq (0, 0), \\
0 & \text{if } (x, y) = (0, 0).
\end{cases}
\]
Show that this function is not continuous at (0, 0) but that it has all directional derivatives at (0, 0) and they all equal 0.

19. Let \( X_1 \) be a normed linear space having norm \( \| \cdot \|_i \). Then we can make \( \prod_{i=1}^n X_i \) into a normed linear space by defining a norm on \( x \in \prod_{i=1}^n X_i \) by
\[
\| x \| \equiv \max \{ \| x_i \|_i : i = 1, \ldots, n \}.
\]
Show this is a norm on \( \prod_{i=1}^n X_i \) as claimed.

20. Suppose \( f : U \subseteq X \times Y \to Z \) and \( D_2 f (x_0, y_0)^{-1} \in \mathcal{L} (X, Y) \) exists and \( f \) is \( C^1 \) so the conditions of the implicit function theorem are satisfied. Also suppose that all these are complex Banach spaces. Show that then the implicitly defined function \( y = y(x) \) is analytic. Thus it has infinitely many derivatives and can be given as a power series as described above.

21. Let \( H \) be a Hilbert space and let \( A \in \mathcal{L} (H, H) \). Suppose \( \ker (A) \) is finite dimensional. Explain why it is automatically a closed subspace of \( H \). Show that \( \ker (A) \oplus (\ker (A))^\perp = H \). Then show that \( A \) is one to one on \( (\ker (A))^\perp \).
In Lyapunov Schmidt method, you would let \( X_1 = (\ker (A))^\perp \) assuming \( A \) is \( D_1 f (0, 0) \).
Chapter 7

Fixed Point Theorems And More

7.1 The Schauder Fixed Point Theorem

First we give a proof of the Schauder fixed point theorem which is an infinite dimensional generalization of the Brouwer fixed point theorem. This is a theorem which lives in Banach space. There is also a version of this theorem valid in locally convex topological vector spaces where the theorem is sometimes called Tychonoff’s theorem. It was shown in an exercise that the closed unit ball fails to have the fixed point property in an infinite dimensional Hilbert space. That is, there is a continuous map from this closed ball to itself which has no fixed point.

Thus you need something other than continuity if you want to get a fixed point. This also shows that there is a retraction of $B$ onto $\partial B$ in any infinite dimensional separable Hilbert space. You get it the usual way. Take the line from $x$ to $f(x)$ and the retraction will be the function which gives the point on $\partial B$ which is obtained by extending this line till it hits the boundary of $B$. Thus for Hilbert spaces, those which have $\partial B$ a retraction of $B$ are exactly those which are infinite dimensional. The above reference claims this retraction property holds for any infinite dimensional normed linear space. I think it is fairly clear to see from the above example that this is not a surprising assertion. Recall that one of the proofs of the Brouwer fixed point theorem used the non existence of such a retraction, obtained using integration theory, to prove the theorem.

We let $K$ be a closed convex subset of $X$ a Banach space and let

\[ f \text{ be continuous, } f : K \to K, \text{ and } \overline{f(K)} \text{ is compact.} \]

**Lemma 7.1.1** *For each $r > 0$ there exists a finite set of points*

\[ \{y_1, \cdots, y_n\} \subseteq \overline{f(K)} \]
and continuous functions $\psi_i$ defined on $\overline{f(K)}$ such that for $x \in \overline{f(K)}$,

$$
\sum_{i=1}^{n} \psi_i(x) = 1, \tag{7.1.1}
$$

$$
\psi_i(x) = 0 \text{ if } x \notin B(y_i, r), \quad \psi_i(x) > 0 \text{ if } x \in B(y_i, r).
$$

If

$$
f_r(x) \equiv \sum_{i=1}^{n} y_i \psi_i(f(x)), \tag{7.1.2}
$$

then whenever $x \in K$,

$$
\|f(x) - f_r(x)\| \leq r.
$$

**Proof:** Using the compactness of $\overline{f(K)}$, there exists

$$
\{y_1, \cdots, y_n\} \subseteq \overline{f(K)} \subseteq K
$$

such that

$$
\{B(y_i, r)\}_{i=1}^{n}
$$

covers $\overline{f(K)}$. Let

$$
\phi_i(y) \equiv (r - \|y - y_i\|)^+.
$$

Thus $\phi_i(y) > 0$ if $y \in B(y_i, r)$ and $\phi_i(y) = 0$ if $y \notin B(y_i, r)$. For $x \in \overline{f(K)}$, let

$$
\psi_i(x) \equiv \phi_i(x) \left( \sum_{j=1}^{n} \phi_j(x) \right)^{-1}.
$$

Then (7.1.1) is satisfied. Indeed the denominator is not zero because $x$ is in one of the $B(y_i, r)$. Thus it is obvious that the sum of these equals 1 on $K$. Now let $f_r$ be given by (7.1.2) for $x \in K$. For such $x$,

$$
f(x) - f_r(x) = \sum_{i=1}^{n} (f(x) - y_i) \psi_i(f(x))
$$

Thus

$$
f(x) - f_r(x) = \sum_{\{i: f(x) \in B(y_i, r)\}} (f(x) - y_i) \psi_i(f(x))
+ \sum_{\{i: f(x) \notin B(y_i, r)\}} (f(x) - y_i) \psi_i(f(x))
= \sum_{\{i: f(x) - y_i \in B(0, r)\}} (f(x) - y_i) \psi_i(f(x))
+ \sum_{\{i: f(x) \notin B(y_i, r)\}} 0 \psi_i(f(x)) \in B(0, r)
$$
7.1. THE SCHAUDEL FIXED POINT THEOREM

because $0 \in B(0,r)$, $B(0,r)$ is convex, and $B(0,r)$. It is just a convex combination of things in $B(0,r)$. ■

Note that we could have had the $y_i$ in $f(K)$ in addition to being in $\overline{f(K)}$. This would make it possible to eliminate the assumption that $K$ is closed later on. All you really need is that $K$ is convex.

We think of $f_r$ as an approximation to $f$. In fact it is uniformly within $r$ of $f$ on $K$. The next lemma shows that this $f_r$ has a fixed point. This is the main result and comes from the Brouwer fixed point theorem in $\mathbb{R}^n$. It is an approximate fixed point.

Lemma 7.1.2 For each $r > 0$, there exists $x_r \in \text{convex hull of } f(K) \subseteq K$ such that $f_r(x_r) = x_r$, $\|f_r(x) - f(x)\| < r$ for all $x$.

Proof: If $f_r(x_r) = x_r$ and $x_r = \sum_{i=1}^{n} a_i y_i$ for $\sum_{i=1}^{n} a_i = 1$ and the $y_i$ described in the above lemma, we need

$$f_r(x_r) = \sum_{i=1}^{n} y_i \psi_i(f(x_r)) = \sum_{j=1}^{n} y_j \psi_j \left( f \left( \sum_{i=1}^{n} a_i y_i \right) \right) = \sum_{j=1}^{n} a_j y_j = x_r.$$ 

Also, if this is satisfied, then we have the desired fixed point.

This will be satisfied if for each $j = 1, \ldots, n$,

$$a_j = \psi_j \left( f \left( \sum_{i=1}^{n} a_i y_i \right) \right); \quad (7.1.3)$$

so, let

$$\Sigma_{n-1} = \left\{ a \in \mathbb{R}^n : \sum_{i=1}^{n} a_i = 1, \; a_i \geq 0 \right\}$$

and let $h : \Sigma_{n-1} \to \Sigma_{n-1}$ be given by

$$h(a)_j = \psi_j \left( f \left( \sum_{i=1}^{n} a_i y_i \right) \right).$$

Since $h$ is a continuous function of $a$, the Brouwer fixed point theorem applies and there exists a fixed point for $h$ which is a solution to $\Sigma_{n-1}$. ■

The following is the Schauder fixed point theorem.

Theorem 7.1.3 Let $K$ be a closed and convex subset of $X$, a normed linear space. Let $f : K \to K$ be continuous and suppose $\overline{f(K)}$ is compact. Then $f$ has a fixed point.
CHAPTER 7. FIXED POINT THEOREMS AND MORE

Proof: Recall that \( f(x_r) - f_r(x_r) \in B(0,r) \) and \( f_r(x_r) = x_r \) with \( x_r \in \text{convex hull of } f(K) \subseteq K \).

There is a subsequence, still denoted with subscript \( r \) such that \( f(x_r) \to x \in \overline{f(K)} \). Note that the fact that \( K \) is convex is what makes \( f \) defined at \( x_r \). \( x_r \) is in the convex hull of \( \overline{f(K)} \subseteq K \). This is where we use \( K \) convex. Then since \( f_r \) is uniformly close to \( f \), it follows that \( f_r(x_r) = x_r \to x \) also. Therefore,

\[
f(x) = \lim_{r \to 0} f(x_r) = \lim_{r \to 0} f_r(x_r) = \lim_{r \to 0} x_r = x. \]

We usually have in mind the mapping defined on a Banach space. However, the completeness was never used. Thus the result holds in a normed linear space.

There is a nice corollary of this major theorem which is called the Schaefer fixed point theorem or the Leray Schauder alternative principle [14].

Theorem 7.1.4 Let \( f : X \to X \) be a compact map. Then either

1. There is a fixed point for \( tf \) for all \( t \in [0,1] \) or

2. For every \( r > 0 \), there exists a solution to \( x = tf(x) \) for \( t \in (0,1) \) such that \( \|x\| > r \).

Proof: Suppose there is \( t_0 \in [0,1] \) such that \( t_0f \) has no fixed point. Then \( t_0 \neq 0, t_0f \) obviously has a fixed point if \( t_0 = 0 \). Thus \( t_0 \in (0,1] \). Then let \( r_M \) be the radial retraction onto \( B(0,M) \). By Schauder’s theorem there exists \( x \in B(0,M) \) such that \( t_0r_Mf(x) = x \). Then if \( \|f(x)\| \leq M \), \( r_M \) has no effect and so \( t_0f(x) = x \) which is assumed not to take place. Hence \( \|f(x)\| > M \) and so \( \|r_Mf(x)\| = M \) so \( \|x\| = t_0M \). Also \( t_0r_Mf(x) = t_0M \frac{f(x)}{\|f(x)\|} = x \) and so \( x = tf(x), \), \( t = t_0 \frac{M}{\|f(x)\|} < 1 \). Since \( M \) is arbitrary, it follows that the solutions to \( x = tf(x) \) for \( t \in (0,1) \) are unbounded. It was just shown that there is a solution to \( x = tf(x), t < 1 \) such that \( \|x\| = t_0M \) where \( M \) is arbitrary. Thus the second of the two alternatives holds.

As an example of the usefulness of the Schauder fixed point theorem, consider the following application to the theory of ordinary differential equations. In the context of this theorem, \( X = C([0,T] ; \mathbb{R}^n) \), a Banach space with norm given by

\[
\|x\| \equiv \max \{|x(t)| : t \in [0,T]\}.
\]

Theorem 7.1.5 Let \( f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n \) be continuous and suppose there exists \( L > 0 \) such that for all \( \lambda \in (0,1) \), if

\[
\dot{x'} = \lambda f(t,x), \quad x(0) = x_0 \quad (7.1.4)
\]

for all \( t \in [0,T] \), then \( \|x\| < L \). Then there exists a solution to

\[
\dot{x'} = f(t,x), \quad x(0) = x_0 \quad (7.1.5)
\]

for \( t \in [0,T] \).
7.1. THE SCHAUDER FIXED POINT THEOREM

**Proof:** Let

\[ N_x(t) \equiv \int_0^t f(s, x(s)) \, ds. \]

Thus a solution to the initial value problem exists if there exists a solution to

\[ x_0 + N(x) = x. \]

Let

\[ m \equiv \max \{ |f(t, x)| : (t, x) \in [0, T] \times \overline{B}(0, L) \}, \quad M \equiv |x_0| + mT \]

and let

\[ K \equiv \{ x \in C([0, T]; \mathbb{R}^n) : x(0) = x_0 \text{ and } |x| \leq M \}. \]

Now define

\[ A x \equiv \begin{cases} x_0 + N x & \text{if } |N x| \leq M - |x_0|, \\ x_0 + \frac{(M - |x_0|) N x}{||N x||} & \text{if } |N x| > M - |x_0|. \end{cases} \]

Then \( A \) is continuous and maps \( X \) to \( K \) because

\[ ||A x|| \leq |x_0| + ||N x|| \leq M \text{ if } ||N x|| \leq M - |x_0| \]

and otherwise,

\[ ||A x|| \leq |x_0| + \left( \frac{(M - |x_0|) ||N x||}{||N x||} \right) \leq |x_0| + M - |x_0| = M. \]

Also \( A(K) \) is equicontinuous because

\[ A x(t) - A x(t_1) = \int_{t_1}^t f(s, x(s)) \, ds \]

and the integrand is bounded because to be in \( K, ||x|| \) is bounded. Thus \( \overline{A(K)} \) is a compact set in \( X \) by the Ascoli Arzela theorem. By the Schauder fixed point theorem, \( A \) has a fixed point, \( x \in K \). I claim that \( ||N x|| \leq M - |x_0| \).

If \( ||N(x)|| > M - |x_0| \), then \( (M - |x_0|) / ||N x|| = \lambda < 1 \) and

\[ x_0 + \lambda N(x) = x \]

and so \( f \) holds. Therefore, by the assumed estimate on the solutions to \( f \), it follows that

\[ ||x|| < L \]

and so \( ||N x|| \leq mT = M - |x_0| \), a contradiction. Therefore, it must be the case that

\[ ||N(x)|| \leq M - |x_0| \]

which implies that

\[ x_0 + N(x) = x. \]

Since this is equivalent to \( f \), this proves the theorem. \( \blacksquare \)

Here is a neater proof which uses the Leray Schauder alternative, also called the Schaefer fixed point theorem presented above.
Theorem 7.1.6 Let \( f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \) be continuous and suppose there exists \( L > 0 \) such that for all \( \lambda \in (0, 1) \), if
\[
  x' = \lambda f(t, x), \quad x(0) = x_0
\]
for all \( t \in [0, T] \), then \( ||x|| < L \). Then there exists a solution to
\[
  x' = f(t, x), \quad x(0) = x_0
\]
for \( t \in [0, T] \).

Proof: Let \( F : X \to X \) where \( X \) described above.
\[
  F_y(t) = \int_0^t f(s, y(s) + x_0) \, ds
\]
Let \( B \) be a bounded set in \( X \). Then \( |f(s, y(s) + x_0)| \) is bounded for \( s \in [0, T] \) if \( y \in B \). Say \( |f(s, y(s) + x_0)| \leq C_B \). Hence \( F(B) \) is bounded in \( X \). Also, for \( y \in B, s < t \),
\[
  |F_y(t) - F_y(s)| \leq \left| \int_s^t f(s, y(s) + x_0) \, ds \right| \leq C_B |t - s|
\]
and so \( F(B) \) is pre-compact by the Ascoli Arzela theorem. By the Schaefer fixed point theorem, there are two alternatives. Either there are unbounded solutions \( y \) to
\[
  \lambda F(y) = y
\]
for various \( \lambda \in (0, 1) \) or for all \( \lambda \in [0, 1] \), there is a fixed point for \( \lambda F \). In the first case, there would be unbounded \( y_\lambda \) solving
\[
  y_\lambda(t) = \lambda \int_0^t f(s, y_\lambda(s) + x_0) \, ds
\]
Then let \( x_\lambda(s) \equiv y_\lambda(s) + x_0 \) and you get \( ||x_\lambda|| \) also unbounded for various \( \lambda \in (0, 1) \). The above implies
\[
  x_\lambda(t) - x_0 = \lambda \int_0^t f(s, x_\lambda(s)) \, ds
\]
so \( x'_\lambda = \lambda f(t, x_\lambda) \), \( x_\lambda(0) = x_0 \) and these would be unbounded for \( \lambda \in (0, 1) \) contrary to the assumption that there exists an estimate for these valid for all \( \lambda \in (0, 1) \). Hence the first alternative must hold and hence there is \( y \in X \) such that
\[
  Fy = y
\]
Then letting \( x(s) \equiv y(s) + x_0 \), it follows that
\[
  x(t) - x_0 = \int_0^t f(s, x(s)) \, ds
\]
and so \( x \) is a solution to the differential equation on \( [0, T] \). \( \blacksquare \)

Note that existence for solutions to \( 7.1.4 \) is not assumed, only estimates of possible solutions. These estimates are called \textit{a-priori} estimates. Also note this is a global existence theorem, not a local one for a solution defined on only a small interval.
7.2. THE KAKUTANI FIXED POINT THEOREM

7.2 The Kakutani Fixed Point Theorem

To begin with, if $A : X \rightarrow \mathcal{P}(Y)$ is a set-valued map, define the graph of $A$ by

$$G(A) \equiv \{(x, y) : y \in Ax\}.$$ 

Consider a map $A$ which maps $\mathbb{C}^n$ to $\mathcal{P}(\mathbb{C}^n)$ which satisfies

$$Ax \text{ is compact and convex.} \quad (7.2.8)$$

and also the condition that if $O$ is open and $O \supseteq Ax$, then there exists $\delta > 0$ such that if $y \in B(x, \delta)$, then $Ay \subseteq O$. \quad (7.2.9)

This last condition is sometimes referred to as upper semicontinuity. In words, $A$ is upper semicontinuous and has values which are compact and convex. This is equivalent to saying that if $Ax \in O$ and $x_n \rightarrow x$, then for large enough $n$, it follows that $Ax_n \subseteq O$.

**Lemma 7.2.1** Let $A$ satisfy (7.2.9). Then $AK$ is a subset of a compact set whenever $K$ is compact. Also the graph of $A$ is closed if $Ax$ is closed.

**Proof:** Let $x \in K$. Then $Ax$ is compact and contained in some open set whose closure is compact, $U_x$. By assumption that there exists an open set $V_x$ containing $x$ such that if $y \in V_x$, then $Ay \subseteq U_x$. Let $V_{x_1}, \ldots, V_{x_m}$ cover $K$. Then $AK \subseteq \bigcup_{k=1}^m \overline{U}_{x_k}$, a compact set.

To see the graph of $A$ is closed when $Ax$ is closed, let $x_k \rightarrow x, y_k \rightarrow y$ where $y_k \in Ax_k$. Then letting $O = Ax + B(0, r)$ it follows from that $y_k \in Ax_k \subseteq O$ for all $k$ large enough. Therefore, $y \in Ax + B(0, 2r)$ and since $r > 0$ is arbitrary and $Ax$ is closed it follows $y \in Ax$.

Also, there is a general consideration relative to upper semicontinuous functions.

**Lemma 7.2.2** If $f$ is upper semicontinuous on some set $K$ and $g$ is continuous and defined on $f(K)$, then $g \circ f$ is also upper semicontinuous.

**Proof:** Let $x_n \rightarrow x$ in $K$. Let $U \supseteq g \circ f(x)$. Is $g \circ f(x_n) \in U$ for all $n$ large enough? We have $f(x) \in g^{-1}(U)$, an open set. Therefore, if $n$ is large enough, $f(x_n) \in g^{-1}(U)$. It follows that for large enough $n$, $g \circ f(x_n) \in U$ and so $g \circ f$ is upper semicontinuous on $K$.

The next theorem is an application of the Brouwer fixed point theorem. First define an $n$ simplex, denoted by $[x_0, \ldots, x_n]$, to be the convex hull of the $n + 1$ points, $\{x_0, \ldots, x_n\}$ where $\{x_i - x_0\}_{i=1}^n$ are independent. Thus

$$[x_0, \ldots, x_n] \equiv \left\{ \sum_{i=1}^n t_i x_i : \sum_{i=1}^n t_i = 1, \ t_i \geq 0 \right\}.$$
If \( n \leq 2 \), the simplex is a triangle, line segment, or point. If \( n \leq 3 \), it is a tetrahedron, triangle, line segment or point. A collection of simplices is a tiling of \( \mathbb{R}^n \) if \( \mathbb{R}^n \) is contained in their union and if \( S_1, S_2 \) are two simplices in the tiling, with

\[
S_j = [x_0^j, \ldots, x_n^j],
\]

then

\[
S_1 \cap S_2 = [x_{k_0}, \ldots, x_{k_r}]
\]

where

\[
\{x_{k_0}, \ldots, x_{k_r}\} \subseteq \{x_0^1, \ldots, x_n^1\} \cap \{x_0^2, \ldots, x_n^2\}
\]

or else the two simplices do not intersect. The collection of simplices is said to be locally finite if, for every point, there exists a ball containing that point which also intersects only finitely many of the simplices in the collection. It is left to the reader to verify that for each \( \varepsilon > 0 \), there exists a locally finite tiling of \( \mathbb{R}^n \) which is composed of simplices which have diameters less than \( \varepsilon \). The local finiteness ensures that for each \( \varepsilon \) the vertices have no limit point. To see how to do this, consider the case of \( \mathbb{R}^2 \). Tile the plane with identical small squares and then form the triangles indicated in the following picture. It is clear something similar can be done in any dimension. Making the squares identical ensures that the little triangles are locally finite.

In general, you could consider \([0, 1]^n\). The point at the center is \((1/2, \ldots, 1/2)\). Then there are \(2n\) faces. Form the \(2n\) pyramids having this point along with the \(2^{n-1}\) vertices of the face. Then use induction on each of these faces to form smaller dimensional simplices tiling that face. Corresponding to each of these \(2n\) pyramids, it is the union of the simplices whose vertices consist of the center point along with those of these new simplices tiling the chosen face. In general, you can write any \(n\) dimensional cube as the translate of a scaled \([0, 1]^n\). Thus one can express each of identical cubes as a tiling of \(m(n)\) simplices of the appropriate size and thereby obtain a tiling of \(\mathbb{R}^n\) with simplices. A ball will intersect only finitely many of the cubes and hence finitely many of the simplices. To get their diameters small as desired, just use \([0, r]^n\) instead of \([0, 1]^n\).

Thus one can give a function any value desired on these vertices and extend appropriately to the rest of the simplex and obtain a continuous function.

The Kakutani fixed point theorem is a generalization of the Brouwer fixed point theorem from continuous single valued maps to upper semicontinuous maps which have closed convex values.
Theorem 7.2.3 Let $K$ be a compact convex subset of $\mathbb{R}^n$ and let $A : K \to \mathcal{P}(K)$ such that $Ax$ is a closed convex subset of $K$ and $A$ is upper semicontinuous. Then there exists $x$ such that $x \in Ax$. This is the “fixed point”.

Proof: Let there be a locally finite tiling of $\mathbb{R}^n$ consisting of simplices having diameter no more than $\varepsilon$. Let $P \mathbf{x}$ be the point in $K$ which is closest to $\mathbf{x}$. For each vertex $x_k$, pick $A_\varepsilon x_k \in AP x_k$ and define $A_\varepsilon$ on all of $\mathbb{R}^n$ by the following rule. If $x = \sum_{i=0}^{n} t_i x_i, t_i \in [0,1], \sum_i t_i = 1$, then

$$A_\varepsilon x \equiv \sum_{k=0}^{n} t_k A_\varepsilon x_k.$$ 

Now by construction $A_\varepsilon x_k \in AP x_k \in K$ and so $A_\varepsilon$ is a continuous map defined on $\mathbb{R}^n$ with values in $K$ thanks to the local finiteness of the collection of simplices. By the Brouwer fixed point theorem $A_\varepsilon$ has a fixed point $x_\varepsilon$ in $K$, $A_\varepsilon x_\varepsilon = x_\varepsilon$.

$x_\varepsilon = \sum_{k=0}^{n} t_\varepsilon^k A_\varepsilon x_\varepsilon^k, A_\varepsilon x_\varepsilon^k \in AP x_\varepsilon^k \subseteq K$

where a simplex containing $x_\varepsilon$ is

$$[x_0^\varepsilon, \ldots, x_n^\varepsilon], x_\varepsilon = \sum_{k=0}^{n} t_\varepsilon^k x_\varepsilon^k$$

Also, $x_\varepsilon \in K$ and is closer than $\varepsilon$ to each $x_k^\varepsilon$ so each $x_k^\varepsilon$ is within $\varepsilon$ of $K$. It follows that for each $k$, $|P x_\varepsilon^k - x_\varepsilon^k| < \varepsilon$ and so

$$\lim_{\varepsilon \to 0} |P x_\varepsilon^k - x_\varepsilon^k| = 0$$

By compactness of $K$, there exists a subsequence, still denoted with the subscript of $\varepsilon$ such that for each $k$, the following convergences hold as $\varepsilon \to 0$

$$t_\varepsilon^k \to t_k, A_\varepsilon x_\varepsilon^k \to y_k, P x_\varepsilon^k \to z_k, x_\varepsilon^k \to z_k$$

Any pair of the $x_\varepsilon^k$ are within $\varepsilon$ of each other. Hence, any pair of the $P x_\varepsilon^k$ are within $\varepsilon$ of each other because $P$ reduces distances. Therefore, in fact, $z_k$ does not depend on $k$.

$$\lim_{\varepsilon \to 0} P x_\varepsilon^k = \lim_{\varepsilon \to 0} x_\varepsilon^k = z, \lim_{\varepsilon \to 0} x_\varepsilon = \lim_{\varepsilon \to 0} \sum_{k=0}^{n} t_\varepsilon^k x_\varepsilon^k = \sum_{k=0}^{n} t_k z = z$$

By upper semicontinuity of $A$, for all $\varepsilon$ small enough,

$$AP x_\varepsilon^k \subseteq A z + B(0, r)$$
In particular, since \( A_\varepsilon x_k^\ast \in APx_k^\ast \),
\[
A_\varepsilon x_k^\ast \in Az + B(0, r)
\]
for \( \varepsilon \) small enough.

Since \( r \) is arbitrary and \( Az \) is closed, it follows
\[
y_k \in Az.
\]

It follows that since \( K \) is closed,
\[
x_\varepsilon \to z = \sum_{k=0}^{n} t_k y_k, \ t_k \geq 0, \sum_{k=0}^{n} t_k = 1
\]

Now by convexity of \( Az \) and the fact just shown that \( y_k \in Az \),
\[
z = \sum_{k=0}^{n} t_k y_k \in Az
\]
and so \( z \in Az \). This is the fixed point.

The variational principle of Ekeland is the following theorem \[29\]. You start with an approximate minimizer \( x_0 \). It says there is \( y_\lambda \) fairly close to \( x_0 \) such that if you subtract a “cone” from the value of \( \phi \) at \( y_\lambda \), then the resulting function is less than \( \phi(x) \) for all \( x \neq y_\lambda \). This cone is like a supporting plane for a convex function but pertains to functions which are certainly not convex.

7.3 A Variational Principle of Ekeland

**Definition 7.3.1** A function \( \phi : X \to (-\infty, \infty] \) is called proper if it is not constantly equal to \( \infty \). Here \( X \) is assumed to be a complete metric space. The function \( \phi \) is lower semicontinuous if
\[
x_n \to x \text{ implies } \phi(x) \leq \lim \inf_{n \to \infty} \phi(x_n)
\]
It is bounded below if there is some constant \( C \) such that \( C \leq \phi(x) \) for all \( x \).

The variational principle of Ekeland is the following theorem \[29\]. You start with an approximate minimizer \( x_0 \). It says there is \( y_\lambda \) fairly close to \( x_0 \) such that if you subtract a “cone” from the value of \( \phi \) at \( y_\lambda \), then the resulting function is less than \( \phi(x) \) for all \( x \neq y_\lambda \). This cone is like a supporting plane for a convex function but pertains to functions which are certainly not convex.
7.3. A VARIATIONAL PRINCIPLE OF EKELAND

**Theorem 7.3.2** Let $X$ be a complete metric space and let $\phi : X \to (-\infty, \infty]$ be proper, lower semicontinuous and bounded below. Let $x_0$ be such that

$$\phi(x_0) \leq \inf_{x \in X} \phi(x) + \varepsilon$$

Then for every $\lambda > 0$ there exists a $y_\lambda$ such that

1. $\phi(y_\lambda) \leq \phi(x_0)$
2. $d(y_\lambda, x_0) \leq \lambda$
3. $\phi(y_\lambda) - \frac{\varepsilon}{\lambda} d(x, y_\lambda) < \phi(x)$ for all $x \neq y_\lambda$

To motivate the proof, see the following picture which illustrates the first two steps. The $S_i$ will be sets in $X$ but are denoted symbolically by labeling them in $X \times (-\infty, \infty]$. 

Then the end result of this iteration would be a picture like the following.

Thus you would have $\phi(y_\lambda) - \frac{\varepsilon}{\lambda} d(y_\lambda, x) \leq \phi(x)$ for all $x$ which is seen to be what is wanted.

**Proof:** Let $x_1 = x_0$ and define

$$S_1 \equiv \{ z \in X : \phi(z) \leq \phi(x_1) - \frac{\varepsilon}{\lambda} d(z, x_1) \}$$

Then $S_1$ contains $x_1$ so it is nonempty. It is also clear that $S_1$ is a closed set. This follows from the lower semicontinuity of $\phi$. Let $x_2$ be a point of $S_1$, possibly different than $x_1$ and let

$$S_2 \equiv \{ z \in X : \phi(z) \leq \phi(x_2) - \frac{\varepsilon}{\lambda} d(z, x_2) \}$$
Continue in this way. Now let there be a sequence of points \( \{ x_k \} \) such that \( x_k \in S_{k-1} \) and define \( S_k \) by
\[
S_k \equiv \left\{ z \in X : \phi(z) \leq \phi(x_k) - \frac{\varepsilon}{\lambda} d(z, x_k) \right\}
\]
where \( x_k \) is some point of \( S_{k-1} \). Then \( x_k \) is a point of \( S_k \). Will this yield a nested sequence of nonempty closed sets? Yes, it appears that it would because if \( z \in S_k \) then
\[
\phi(z) \leq \phi(x_k) - \frac{\varepsilon}{\lambda} d(z, x_k) \leq \left( \phi(x_{k-1}) - \frac{\varepsilon}{\lambda} d(x_{k-1}, x_k) \right) - \frac{\varepsilon}{\lambda} d(z, x_k)
\]
showing that \( z \) has what it takes to be in \( S_{k-1} \). Thus we would obtain a sequence of nested, nonempty, closed sets according to this scheme.

Now here is how to choose the \( x_k \in S_{k-1} \). Let
\[
\phi(x_k) < \inf_{x \in S_{k-1}} \phi(x) + \frac{1}{2^k}
\]
Then for \( z \in S_{n+1} \subseteq S_n \),
\[
\phi(z) \leq \phi(x_{n+1}) - \frac{\varepsilon}{\lambda} d(z, x_{n+1})
\]
and so
\[
\frac{\varepsilon}{\lambda} d(z, x_{n+1}) \leq \phi(x_{n+1}) - \phi(z) \leq \inf_{x \in S_n} \phi(x) + \frac{1}{2^{n+1}} - \phi(z)
\]
\[
\leq \phi(z) + \frac{1}{2^{n+1}} - \phi(z) = \frac{1}{2^{n+1}}
\]
Thus every \( z \in S_{n+1} \) is within \( \frac{1}{2^{n+1}} \) of the single point \( x_{n+1} \) and so the diameter of \( S_n \) converges to 0 as \( n \to \infty \). By completeness of \( X \), there exists a unique \( y_\lambda \in \cap_n S_n \). Then it follows in particular that for \( x_0 = x_1 \) as above,
\[
\phi(y_\lambda) \leq \phi(x_0) - \frac{\varepsilon}{\lambda} d(y_\lambda, x_0) \leq \phi(x_0)
\]
which verifies the first of the above conclusions.

As to the second, \( \phi(x_0) \leq \inf_{x \in X} \phi(x) + \varepsilon \) and so, for any \( x \),
\[
\phi(y_\lambda) \leq \phi(x_0) - \frac{\varepsilon}{\lambda} d(y_\lambda, x_0) \leq \phi(x) + \varepsilon - \frac{\varepsilon}{\lambda} d(y_\lambda, x_0)
\]
this being true for \( x = y_\lambda \). Hence \( \frac{\varepsilon}{\lambda} d(y_\lambda, x_0) \leq \varepsilon \) and so \( d(y_\lambda, x_0) \leq \lambda \). Indeed,
\[
\phi(y_\lambda) \leq \phi(y_\lambda) + \varepsilon - \frac{\varepsilon}{\lambda} d(y_\lambda, x_0)
\]
\[
\frac{\varepsilon}{\lambda} d(y_\lambda, x_0) \leq \varepsilon, \text{ so } d(y_\lambda, x_0) \leq \lambda
\]
Finally consider the third condition. If it does not hold, then there exists $z \neq y_\lambda$ such that 
\[ \phi(y_\lambda) \geq \phi(z) + \frac{\varepsilon}{\lambda} d(z, y_\lambda) \]
so that 
\[ \phi(z) \leq \phi(y_\lambda) - \frac{\varepsilon}{\lambda} d(z, y_\lambda). \]
But then, by the definition of $y_\lambda$ as being in all the $S_n$, 
\[ \phi(y_\lambda) \leq \phi(x_n) - \frac{\varepsilon}{\lambda} d(x_n, y_\lambda) \]
and so 
\[ \phi(z) \leq \phi(x_n) - \frac{\varepsilon}{\lambda} (d(x_n, y_\lambda) + d(z, y_\lambda)) \]
\[ \leq \phi(x_n) - \frac{\varepsilon}{\lambda} d(x_n, z) \]
Since $n$ is arbitrary, this shows that $z \in \bigcap_n S_n$ but there is only one element of this intersection and it is $y_\lambda$ so $z$ must equal $y_\lambda$, a contradiction. ■

Note how if you make $\lambda$ very small, you could pick $\varepsilon$ very small such that the cone looks pretty flat.

### 7.3.1 Cariste Fixed Point Theorem

As mentioned in [29], the above result can be used to prove a fixed point theorem called the Cariste fixed point theorem.

**Theorem 7.3.3** Let $\phi$ be lower semicontinuous, proper, and bounded below on a complete metric space $X$ and let $F : X \to P(X)$ be set valued such that $F(x) \neq \emptyset$ for all $x$. Also suppose that for each $x \in X$, there exists $y \in F(x)$ such that 
\[ \phi(y) + d(x, y) \leq \phi(x) \]
Then there exists $x_0$ such that $x_0 \in F(x_0)$.

**Proof:** In the above Ekeland variational principle, let $\varepsilon = 1 = \lambda$. Then there exists $x_0$ such that for all $y \neq x_0$
\[ \phi(x_0) - d(y, x_0) < \phi(y), \quad \text{so} \quad \phi(x_0) < \phi(y) + d(y, x_0) \quad (7.3.10) \]
for all $y \neq x_0$. 

![Diagram](attachment:image.png)
Suppose $x_0 \notin F(x_0)$. From the assumption, there is $y \in F(x_0)$ (so $y \neq x_0$) such that
\[
\phi(y) + d(x_0, y) \leq \phi(x_0)
\]
Since $y \neq x_0$, it follows from \[\text{Equation 7.3.10}\]
\[
\phi(y) + d(x_0, y) \leq \phi(x_0) < \phi(y) + d(y, x_0)
\]
a contradiction. Hence $x_0 \in F(x_0)$ after all. ■

It is a funny theorem. It is easy to prove, but you look at it and wonder what it says. If $F$ is single valued, you would need to have a function $\phi$ such that for each $x$,
\[
\phi(F(x)) \leq \phi(x) - d(x, y)
\]
and if you have such a $\phi$ then you can assert there is a fixed point for $F$. Suppose $F$ is single valued and $d(Fx, Fy) \leq rd(x, y), 0 < r < 1$. Of course $F$ has a fixed point using easier techniques. However, this also follows from this result. Let
\[
\phi(x) = \frac{1}{1-r}d(x, F(x))
\]
Then is it true that for each $x$, there exists $y \in F(x)$ such that the inequality holds for all $x$? Is
\[
\frac{1}{1-r}d(F(x), F(F(x))) \leq \frac{1}{1-r}d(x, F(x)) - d(x, F(x))
\]
Yes, this is certainly so because the right side reduces to $\frac{1}{1-r}d(x, F(x))$. Thus this fixed point theorem implies the usual Banach fixed point theorem.

The Ekeland variational principle says that when $\phi$ is lower semicontinuous proper and bounded below, there exists $y$ such that
\[
\phi(y) - d(x, y) < \phi(x) \text{ for all } x \neq y
\]
In fact this can be proved from the Cariste fixed point theorem. Suppose the EVP does not hold. This would mean that for all $y$ there exists $x \neq y$ such that
\[
\phi(y) - d(x, y) \geq \phi(x)
\]
Thus, for all $x$ there exists $y \neq x$ such that
\[
\phi(x) - d(x, y) \geq \phi(y)
\]
The inequality is preserved if $x = y$. Then let
\[
F(x) \equiv \{ y \neq x : \phi(x) - d(x, y) \geq \phi(y) \} \neq \emptyset
\]
by assumption. This is the hypothesis for the Cariste fixed point theorem. Hence there exists $x_0 \in F(x_0) = \{ y \neq x_0 : \phi(x_0) - d(x_0, y) \geq \phi(y) \}$ but this cannot happen because you can’t have $x_0 \neq x_0$. Thus the Ekeland variational principle must hold after all.
7.3. A VARIATIONAL PRINCIPLE OF EKELAND

7.3.2 A Density Result

There are several applications of the Ekeland variational principle. For more of them, see [29]. One of these is to show that there is a point where $\phi'$ is small assuming $\phi$ is bounded below, lower semicontinuous, and Gateaux differentiable. Here

$$\langle \phi' (x), v \rangle = \lim_{h \to 0} \frac{\phi(x+hv) - \phi(x)}{h}, \quad \phi' (x) \in X'$$

It is sort of an approximate critical point at a point which causes $\phi$ to be near the infimum.

Theorem 7.3.4 Let $X$ be a Banach space and $\phi : X \to \mathbb{R}$ be Gateaux differentiable, bounded from below, and lower semicontinuous. Then for every $\varepsilon > 0$ there exists $x \in X$ such that

$$\phi (x) \leq \inf_{x \in X} \phi (x) + \varepsilon$$

Proof: From the Ekeland variational principle with $\lambda = 1$, there exists $x_\varepsilon$ such that

$$\phi (x_\varepsilon) \leq \inf_{x \in X} \phi (x) + \varepsilon$$

and for all $x$,

$$\phi (x_\varepsilon) < \phi (x) + \varepsilon \|x - x_\varepsilon\|$$

Then letting $x = x_\varepsilon + hv$ where $\|v\| = 1$,

$$\phi (x_\varepsilon + hv) - \phi (x_\varepsilon) > -\varepsilon |h|$$

Let $h < 0$. Then divide by it

$$\frac{\phi (x_\varepsilon + hv) - \phi (x_\varepsilon)}{h} < \varepsilon$$

Passing to a limit as $h \to 0$ yields

$$\langle \phi' (x_\varepsilon), v \rangle \leq \varepsilon$$

Now $v$ was arbitrary with norm 1 and so

$$\sup_{\|v\|=1} \langle \phi' (x_\varepsilon), v \rangle = \|\phi' (x_\varepsilon)\| \leq \varepsilon$$

There is another very interesting application of the Ekeland variational principle [29].

Theorem 7.3.5 Let $X$ be a Banach space and $\phi : X \to \mathbb{R}$ be Gateaux differentiable, bounded from below, and lower semicontinuous. Also suppose there exists $a, c > 0$ such that

$$a \|x\| - c \leq \phi (x) \text{ for all } x \in X$$
Then \( \{ \phi'(x) : x \in X \} \) is dense in the ball of \( X' \) centered at 0 with radius \( a \). Here \( \phi'(x) \in X' \) and is determined by

\[
\langle \phi'(x), v \rangle = \lim_{h \to 0} \frac{\phi(x + hv) - \phi(x)}{h}
\]

**Proof:** Let \( x^* \in X', \|x^*\| \leq a \). Let

\[
\psi(x) = \phi(x) - \langle x^*, x \rangle
\]

This is lower semicontinuous. It is also bounded from below because

\[
\psi(x) \geq \phi(x) - a \|x\| \geq (a \|x\| - c) - a \|x\| = -c
\]

It is also clearly Gateaux differentiable and lower semicontinuous because the piece added in is actually continuous. It is clear that the Gateaux derivative is just \( \phi'(x) - x^* \). By Theorem 7.3.4, there exists \( x \in X \) such that

\[
\phi'(x) - x^* \leq \epsilon
\]

Thus this theorem says that if \( \phi(x) \geq a \|x\| - c \) where \( \phi \) has the nice properties of the theorem it follows that \( \phi'(x) \) is dense in \( B(0, a) \) in the dual space \( X' \). It follows that if for every \( a \), there exists \( c \) such that

\[
\phi(x) \geq a \|x\| - c \text{ for all } x \in X
\]

then \( \{ \phi'(x) : x \in X \} \) is dense in \( X' \). This proves the following lemma.

**Lemma 7.3.6** Let \( X \) be a Banach space and \( \phi : X \to \mathbb{R} \) be Gateaux differentiable, bounded from below, and lower semicontinuous. Suppose for all \( a > 0 \) there exists \( c > 0 \) such that

\[
\phi(x) \geq a \|x\| - c \text{ for all } x
\]

Then \( \{ \phi'(x) : x \in X \} \) is dense in \( X' \).

If the above holds, then

\[
\frac{\phi(x)}{\|x\|} \geq a - \frac{c}{\|x\|}
\]

and so, since \( a \) is arbitrary, it must be the case that

\[
\lim_{\|x\| \to \infty} \frac{\phi(x)}{\|x\|} = \infty. \quad (7.3.11)
\]

In fact, this is sufficient. If not, there would exist \( a > 0 \) such that \( \phi(x_n) < a \|x_n\| - n \). Let \(-L \) be a lower bound for \( \phi(x) \). Then

\[
-L + n \leq a \|x_n\|
\]

and so \( \|x_n\| \to \infty \). Now it follows that

\[
a \geq \frac{\phi(x_n)}{\|x_n\|} + \frac{n}{\|x_n\|} \geq \frac{\phi(x_n)}{\|x_n\|} \quad (7.3.12)
\]

which is a contradiction to (7.3.11). This proves the following interesting density theorem.
Theorem 7.3.7 Let $X$ be a Banach space and $\phi : X \to \mathbb{R}$ be Gateaux differentiable, bounded from below, and lower semicontinuous. Also suppose the coercivity condition

$$\lim_{\|x\| \to \infty} \frac{\phi(x)}{\|x\|} = \infty$$

Then $\{\phi'(x) : x \in X\}$ is dense in $X'$. Here $\phi'(x) \in X'$ and is determined by

$$\langle \phi'(x), v \rangle = \lim_{h \to 0} \frac{\phi(x+hv) - \phi(x)}{h}$$

7.4 Critical Points

7.4.1 Mountain Pass Theorem In Hilbert Space

This is about a value at which there is a critical point. That is, there will be a $c$ such that $I(x) = c$ and $I'(x) = 0$. The Ekeland variational principle showed density of values of $I'(x)$. This is about finding a point where $I'(x) = 0$.

This is from Evans \[25\]. It is an interesting theorem. See also \[29\] for more general versions. It has to do with differentiable functions defined on a Hilbert space $H$. Thus $I : H \to \mathbb{R}$ will be differentiable. Then the following is the Palais Smale condition.

Definition 7.4.1 A functional $I$ satisfies the Palais Smale conditions means that if $\{I(u_k)\}$ is a bounded sequence and $I'(u_k) \to 0$, then $\{u_k\}$ is precompact. That is, it has a subsequence which converges.

It will be assumed that $I$ is $C^1(H; \mathbb{R})$ and also that $I'$ is Lipschitz on bounded sets. By $I'(u)$ is meant the element of $H$ such that

$$I(u + v) = I(u) + \langle I'(u), v \rangle_H + o(v)$$

Such exists because of the Riesz representation theorem. Note that, from the assumption that $I'$ is Lipschitz continuous, it follows that $I'$ is bounded on every bounded set.

First is a deformation theorem. The notation $[I(u) \in S]$ means $\{u : I(u) \in S\}$. Here is a picture which illustrates the main conclusion of the following theorem. The idea is that you modify the functional on some set making it smaller and leaving it unchanged off that set.
Theorem 7.4.2 Let $I$ be $C^1$, $I$ is non constant, satisfy the Palais Smale condition, and $I'$ is Lipschitz continuous on bounded sets. Also suppose that $c \in \mathbb{R}$ is such that either $[I(u) \in [c - \delta, c + \delta]] = \emptyset$ for some $\delta > 0$ or $[I(u) \in [c - \delta, c + \delta]] \neq \emptyset$ for all $\delta > 0$ and IF $I(u) = c$, then $I'(u) \neq 0$. Then for each sufficiently small $\varepsilon > 0$, there is a constant $\delta \in (0, \varepsilon)$ and a function $\eta : [0, 1] \times H \to H$ such that

1. $\eta(0, u) = u$
2. $\eta(1, u) = u$ on $[I(u) \notin (c - \varepsilon, c + \varepsilon)]$
3. $I(\eta(t, u)) \leq I(u)$
4. $\eta(1, [I(u) \leq c + \delta]) \subseteq [I(u) \leq c - \delta]$

The main part of this conclusion is the statement about $u \to \eta(1, u)$ contained in parts 2. and 4. The other two parts are there to facilitate these two although they are certainly interesting for their own sake.

Proof: Suppose $[I(u) \in [c - \delta, c + \delta]] = \emptyset$ for some $\delta > 0$. Then $[I(u) \leq c + \delta/2] \subseteq [I(u) \leq c - \delta/2]$ and you could take $\varepsilon = \delta$ and let $\eta(t, u) = u$. Therefore, assume $[I(u) \in [c - \delta, c + \delta]] \neq \emptyset$ for all $\delta > 0$. Since $I$ is nonconstant, $\varepsilon > 0$ can be chosen small enough that $[I(u) \notin (c - \varepsilon, c + \varepsilon)] \neq \emptyset$. Always let $\varepsilon$ be this small.

Claim 1: For all small enough $\varepsilon > 0$, if $u \in [I(u) \in [c - \varepsilon, c + \varepsilon]]$, $I'(u) \neq 0$ and in fact, for such $\varepsilon$, there exists $\sigma(\varepsilon) > 0$ such that $\sigma(\varepsilon) < \varepsilon$, $\|I'(u)\| > \sigma(\varepsilon)$ for all $u \in [I(u) \in [c - \varepsilon, c + \varepsilon]]$.

Proof of Claim 1: If claim is not so, then there is $\{u_k\}, \varepsilon_k, \sigma_k \to 0, \|I'(u_k)\| < \sigma_k$, and $I(u_k) \in [c - \varepsilon_k, c + \varepsilon_k]$ but $\|I'(u_k)\| \leq \sigma_k$. However, from the Palais Smale condition, there is a subsequence, still denoted as $u_k$ which converges to some $u$. Now $I(u_k) \in [c - \varepsilon_k, c + \varepsilon_k]$ and so $I(u) = c$ while $I'(u) = 0$ contrary to the hypothesis. This proves Claim 1. From now on, $\varepsilon$ will be sufficiently small.

Now define for $\delta < \varepsilon$ (The description of small $\delta$ will be described later.)

$$A \equiv [I(u) \notin (c - \varepsilon, c + \varepsilon)]$$
$$B \equiv [I(u) \in [c - \delta, c + \delta]]$$
Thus $A$ and $B$ are disjoint closed sets. Recall that it is assumed that $B \neq \emptyset$ since otherwise, there is nothing to prove. Also it is assumed throughout that $\varepsilon > 0$ is such that $A \neq \emptyset$ thanks to $I$ not being constant. Thus these are nonempty sets and we do not have to fuss with worrying about meaning when one is empty.

**Claim 2:** For any $u$, $\text{dist}(u, A) + \text{dist}(u, B) > 0$.

This is so because if not, then both would be zero and this requires that $u \in A \cap B$ since these sets are closed. But $A \cap B = \emptyset$.

Now define a function $$g(u) \equiv \frac{\text{dist}(u, A)}{\text{dist}(u, A) + \text{dist}(u, B)}$$

It is a continuous function of $u$ which has values in $[0, 1]$. Consider the ordinary differential initial value problem

$$\eta'(t, u) + g(u) h(\|I'(\eta(t, u))\|) I'(\eta(t, u)) = 0 \quad (7.4.13)$$

$$\eta(0, u) = u \quad (7.4.14)$$

where $r \to h(r)$ is a decreasing function which has values in $(0, 1]$ and equals 1 for $r \in [0, 1]$ and equals $1/r$ for $r > 1$. Here $u$ is given and the $\eta'$ is the time derivative is with respect to $t$. Thus, by assumption, the function $\eta \to g(u) h(\|I'(\eta)\|) I'(\eta)$

is Lipschitz continuous on bounded sets and so there exists a solution to the above initial value problem valid for all $t \in [0, 1]$. To see this, you can let $P$ be the projection map onto the closed ball of radius $M > \|u\|$ and the system

$$\eta'(t, u) + g(u) h(\|I'(P\eta(t, u))\|) I'(P\eta(t, u)) = 0$$

$$\eta(0, u) = u$$

Then by Lipschitz continuity, there is a global solution for all $t \geq 0$. Hence there is a local solution to $7.4.13, 7.4.14$. Note that

$$\|g(u) h(\|I'(P\eta(t, u))\|) I'(P\eta(t, u))\| = g(u) h(\|I'(P\eta(t, u))\|) \|I'(P\eta(t, u))\| \leq 1$$

Taking inner products with $\eta(t, u)$, and integrating $\int_0^t$ for this local solution,

$$\frac{1}{2} \|\eta(t, u)\|^2 - \frac{1}{2} \|u\|^2 + \int_0^t g(u) h(\|I'(P\eta(s, u))\|) \langle I'(P\eta(s, u)), \eta(s, u) \rangle ds = 0$$

$$\frac{1}{2} \|\eta(t, u)\|^2 \leq \frac{1}{2} \|u\|^2 + \int_0^t g(u) h(\|I'(P\eta(t, u))\|) \|I'(P\eta(t, u))\| \|\eta(s, u)\| ds$$

It follows that for $t \leq 1$,

$$\|\eta(t, u)\|^2 \leq \|u\|^2 + 2 \int_0^t \|\eta(s, u)\| ds$$
shown above, \( \eta \) is decreasing.

Thus we pick \( M > e \left( \| u \|^2 + 1 \right) \) and then we obtain that for \( t \in [0, 1] \), the projection map does not change anything. Hence there exists a solution to \( \eta \) on \( [0, 1] \) as desired.

Then for this solution, \( \eta (0, u) = u \) because of the above initial condition. If \( u \in \{ I (u) \notin [c - \varepsilon, c + \varepsilon] \} \), then \( u \in A \) and so \( g (u) = 0 \) so \( \eta (t, u) = u \) for all \( t \in [0, 1] \). This gives the first two conditions. Consider the third.

\[
\frac{d}{dt} (I (\eta (t, u))) = (I' (\eta), \eta') = - (I' (\eta), g (u) h (\| I' (\eta) \|) I' (\eta)) = - g (u) h (\| I' (\eta) \|) \| I' (\eta) \|^2
\]

and so this implies the third condition since it says that the function \( t \to I (\eta (t, u)) \) is decreasing.

It remains to consider the last condition. This involves choosing \( \delta \) still smaller if necessary. It is desired to verify that

\[
\eta (1, [I (u) \leq c + \delta]) \subseteq [I (u) \leq c - \delta]
\]

Suppose it is not so. Then there exists \( u \in [I (u) \leq c + \delta] \) but \( I (\eta (1, u)) > c - \delta \).

\[
c - \delta < I (\eta (1, u)) = I (u) - \int_0^1 (I' (\eta), g (u) h (\| I' (\eta) \|) I' (\eta)) dt
\]

\[
= I (u) - g (u) \int_0^1 h (\| I' (\eta) \|) \| I' (\eta (t, u)) \|^2 dt < c + \delta - g (u) \int_0^1 h (\| I' (\eta) \|) \| I' (\eta (t, u)) \|^2 dt
\]

Then

\[
c - 2\delta + g (u) \int_0^1 h (\| I' (\eta (t, u)) \|) \| I' (\eta (t, u)) \|^2 dt < c
\]

If \( I (u) \leq c - \delta \), there is nothing to show because in this case \( I (\eta (1, u)) \leq I (u) \leq c - \delta \). Hence we can assume that \( I (u) \geq c - \delta \) and also that \( I (u) \leq c + \delta \). Thus \( u \in B \) and so \( g (u) = 1 \). Thus

\[
c - 2\delta + \int_0^1 h (\| I' (\eta (t, u)) \|) \| I' (\eta (t, u)) \|^2 dt < c
\]

Also, it is being assumed that \( I (\eta (1, u)) > c - \delta \) and so by the third conclusion shown above, \( \eta (t, u) \in B \) for \( t \in [0, 1] \). We also know that for such values of
7.4. CRITICAL POINTS

\( \eta (t, u), \|I'(\eta(t,u))\| \geq \sigma(\varepsilon) \) from Claim 1. If \( \|I'(\eta(t,u))\| > 1 \), the integrand equals \( \|I'(\eta(t,u))\| \geq \sigma(\varepsilon) \). If \( \|I'(\eta(t,u))\| \leq 1 \), the integrand is \( \|I'(\eta(t,u))\|^2 \leq \sigma(\varepsilon)^2 \). Thus

\[
c - 2\delta + \int_0^1 \min \left( \sigma(\varepsilon), \sigma(\varepsilon)^2 \right) dt < c
\]

and the only restriction on \( \delta \) was that it should be smaller than \( \varepsilon \). Although it was not mentioned above, \( \delta \) was chosen so small that \( -2\delta + \min \left( \sigma(\varepsilon), \sigma(\varepsilon)^2 \right) > 0 \). Hence this yields a contradiction. Thus the last conclusion is verified. ■

Imagine a valley surrounded by a ring of mountains. On the other side of this ring of mountains, there is another low place. Then there must be some path from the valley to the exterior low place which goes through a point where the gradient equals 0, the gradient being the gradient of a function \( f \) which gives the altitude of the land. This is the idea of the mountain pass theorem. The critical point where \( \nabla f = 0 \) is the mountain pass.

**Theorem 7.4.3** Let \( H \) be a Hilbert space and let \( I : H \to \mathbb{R} \) be a \( C^1 \) functional having \( I' \) Lipschitz continuous and such that \( I \) satisfies the Palais Smale condition. Suppose \( I(0) = 0 \) and \( I(u) \geq a > 0 \) for all \( \|u\| = r \). Suppose also that there exists \( v, \|v\| = r \) such that \( I(v) \leq 0 \). Then define

\[ \Gamma \equiv \{ g \in C ([0, 1]; H) : g (0) = 0, g (1) = v \} \]

Let

\[ c \equiv \inf_{g \in \Gamma} \max_{0 \leq t \leq 1} I (g(t)) \]

Then \( c \) is a critical value of \( I \) meaning that there exists \( u \) such that \( I(u) = c \) and \( I'(u) = 0 \). In particular, there is \( u \neq 0 \) such that \( I'(u) = 0 \).

**Proof:** First note that \( c \geq a > 0 \). Suppose \( c \) is not a critical value. Then by the deformation theorem, for \( \varepsilon > 0, \varepsilon \) sufficiently small, there is \( \eta : H \to H \) and a \( \delta < \varepsilon \) small enough that

\[ \eta([I(u) \leq c+\delta]) \subseteq [I(u) \leq c-\delta] \]

and \( \eta \) leaves unchanged \( [I(u) \notin (c-\varepsilon, c+\varepsilon)] \). Then there is \( g \in \Gamma \) such that

\[ \max_{t \in [0,1]} I(g(t)) < c + \delta \]

Then in particular, \( I(g(t)) < c + \delta \) for every \( t \). Hence you look at \( \eta \circ g \). We know that \( g(0), g(1) \) are both in the set \( [I(u) \notin (c-\varepsilon, c+\varepsilon)] \) because they are both 0 and so \( \eta \) leaves these unchanged. Hence \( \eta \circ g \in \Gamma \) and

\[ I(\eta \circ g(t)) \leq c - \delta \]

for all \( t \in [0,1] \). Thus

\[ c = \inf_{g \in \Gamma} \max_{0 \leq t \leq 1} I(g(t)) \leq \max_{t \in [0,1]} I(\eta \circ g(t)) \leq c - \delta \]
which is clearly a contradiction. ■

The Palais-Smale conditions are pretty restrictive. For example, let \( I(x) = \cos x \). Thus \( I: \mathbb{R} \to \mathbb{R} \). Then let \( u_k = k\pi \). Clearly \( I(u_k) \) is bounded and \( \lim_{k \to \infty} I(u_k) = 0 \) but \( \{u_k\} \) is not precompact. However, here is a simple case which does satisfy the Palais-Smale conditions.

Example 7.4.4 Let \( I: \mathbb{R}^d \to \mathbb{R} \) satisfy \( \lim_{|x| \to \infty} I(x) = \infty \). Then \( I \) satisfies the Palais-Smale conditions.

The growth condition implies that if \( I(x_k) \) is bounded, then so is \( \{x_k\} \) and so this sequence is precompact. Nothing needs to be said about \( I'(x_k) \).

7.4.2 A Locally Lipschitz Selection, Pseudogradients

When you have a functional \( \phi \) defined on a Banach space \( X \), \( \phi'(x) \) is in \( X' \) and it isn’t obvious how you can understand it in terms of an element in \( X \) like what is done with Hilbert space using the Riesz representation theorem. However, there is something called a pseudogradient which is defined next.

Definition 7.4.5 Let \( \phi: X \to \mathbb{R} \) be \( C^1 \). Then \( v \) is a pseudogradient for \( \phi \) at \( x \) if the following hold.

1. \( \|v\|_X \leq 2 \|\phi'(x)\|_{X'} \).
2. \( \|\phi'(x)\|_{X'}^2 \leq \langle \phi'(x), v \rangle \).

A pseudogradient field \( V \) is a locally Lipschitz selection of \( G(x) \) where \( G(x) \) is defined to be the set of pseudogradients of \( \phi \) at \( x \). Thus \( V(x) \in G(x) \) and \( V(x) \) is a pseudogradient for \( \phi \) at each \( x \) a regular point of \( \phi \).

Note how this generalizes the case of Hilbert space. In the Hilbert space case, you have \( \phi'(x) \) which technically is in \( H' \) and you have the gradient, written here as \( \nabla \phi \) which is in \( H \) such that

\[
(\nabla \phi(x), v)_H \equiv \langle \phi'(x), v \rangle_{H', H}
\]

the existence of \( \nabla \phi(x) \) coming from the Riesz representation theorem which also gives that \( \nabla \phi(x) = R^{-1} \phi'(x) \) and so \( \|\nabla \phi(x)\|_H = \|\phi'(x)\|_{H'} \), so the above two conditions hold for the gradient field except for one thing. Why is \( x \to \nabla \phi(x) \) locally Lipschitz. We don’t know this, but with a pseudogradient field, we do. Also, the pseudogradient field is only required at regular points of \( \phi \) where \( \phi'(x) \neq 0 \). If you had strict inequalities holding in the above definition, then they would continue to hold for \( \hat{x} \) near \( x \). Thus if you had

\[
\|v\|_X < 2 \|\phi'(x)\|_{X'}, \quad \|\phi'(x)\|_{X'}^2 < \langle \phi'(x), v \rangle
\]

and \( G(x) \) were the set of such \( v \), then there would be an open set \( U \) containing \( x \) such that \( \bigcap_{\hat{x} \in U} G(\hat{x}) \neq \emptyset \). In fact, the intersection would contain \( v \).

This very nice lemma is from Gasinski L. and Papageorgiou N. [29]. It is a lovely application of Stone’s theorem and partitions of unity for a metric space.
Lemma 7.4.6 Let $Y$ be a metric space and let $X$ be a normed linear space. (We will want to add in $X$.) Let $\Gamma : Y \rightarrow P(X)$ such that $\Gamma(y)$ is a nonempty convex set. Suppose that for each $y \in Y$, there exists an open set $U$ containing $y$ such that

$$\emptyset \neq \cap_{\hat{y} \in U} \Gamma(\hat{y})$$

Then there exists a locally Lipschitz map $\gamma : Y \rightarrow X$ such that $\gamma(y) \in \Gamma(y)$ for all $y$.

Proof: Let $\mathcal{U}$ denote the collection of all open sets $U$ such that the nonempty intersection described above holds. Let $\mathcal{V}$ be a locally finite open refinement which also covers. Thus for any $V \in \mathcal{V}$

$$\emptyset \neq \cap_{\hat{y} \in V} \Gamma(\hat{y})$$

because it is a smaller intersection. Let $\{\phi_V\}_{V \in \mathcal{V}}$ be a partition of unity subordinate to the open covering $\mathcal{V}$. In fact, we can have $\phi_V$ locally Lipschitz. This follows from the above construction of the partition of unity in Theorem 4.1.1. Pick $x_V \in \cap_{\hat{y} \in V} \Gamma(\hat{y})$. Then consider

$$\gamma(y) = \sum_{V \in \mathcal{V}} x_V \phi_V(y)$$

It is clearly locally Lipschitz because near any point $y$, it is a finite sum of Lipschitz functions. Pick $y \in Y$. Then it is in some $V \in \mathcal{V}$. In fact, it is finitely many, $V_1, \cdots, V_n$ and for other $V \in \mathcal{V}$, $\phi_V(y) = 0$. Therefore,

$$\gamma(y) = \sum_{i=1}^n x_{V_i} \phi_{V_i}(y)$$

which is a convex combination of the $x_{V_i}$. Now $x_{V_i} \in \cap_{\hat{y} \in V_i} \Gamma(\hat{y}) \subseteq \Gamma(y)$, this for each $i$. Hence this is a convex combination of points in a nonempty convex set $\Gamma(y)$. Thus $\gamma(y) \in \Gamma(y)$. ■

The following lemma says that if $\phi$ is $C^1$ on $X$, then it has a pseudogradient field on $\{x : \phi'(x) \neq 0\}$, the set of regular points.

Lemma 7.4.7 Let $\phi$ be a $C^1$ function defined on $X$ a Banach space. Then there exists a pseudogradient field for $\phi$ on the set of regular points. ($V(x) \in G(x)$ and $x \rightarrow V(x)$ is locally Lipschitz on the set of regular points.)

Proof: First consider whether $G(x)$, the set of pseudogradients of $\phi$ at $x$ is nonempty for $\phi'(x) \neq 0$. From the definition of the operator norm, there exists $u$ such that $\|u\|_X = 1$ and $\langle \phi'(x), u \rangle \geq \delta \|\phi'(x)\|_X$, where $\delta \in (0,1)$. Then let $v = ru \|\phi'(x)\|_X$, where $r \in (1,2)$.

$$\langle \phi'(x), v \rangle = \langle \phi'(x), ru \|\phi'(x)\| \rangle = r \langle \phi'(x), u \rangle \|\phi'(x)\| \geq r\delta \|\phi'(x)\|^2$$
Then choose \( r, \delta \) such that \( r\delta > 1 \) and \( r < 2 \). Then if these were chosen this way in the above reasoning, it follows that
\[
\|v\| < 2 \|\phi'(x)\| \quad \text{and} \quad \langle \phi'(x), v \rangle > \|\phi'(x)\|^2.
\]
That \( \phi'(x) \neq 0 \) is needed to insure that the above strict inequalities hold.

Thus, letting \( Y \) be the metric space consisting of the regular points of \( \phi \), the continuity of \( \phi' \) implies that the above inequalities persist for all \( y \) close enough to \( x \). Thus there is an open set \( U \) containing \( x \) such that \( v \) satisfies the above inequalities for \( x \) replaced with arbitrary \( y \in U \). Thus
\[
v \in \cap_{y \in U} G(y)
\]
Since it is clear that each \( G(y) \) is convex, Lemma 7.4.6 implies the existence of a locally Lipschitz selection from \( G \). That is \( x \to V(x) \) is locally Lipschitz and \( V(x) \in G(x) \) for all regular \( x \).

It will be important to consider \( y' = f(y) \) where \( f \) is locally Lipschitz and \( y \) is just in a Banach space. This is more complicated than in Hilbert space because of the lack of a convenient projection map.

**Theorem 7.4.8** Let \( f : U \to X \) be locally Lipschitz where \( X \) is a Banach space and \( U \) is an open set. Then there exists a unique local solution to the IVP
\[
y' = f(y), \quad y(0) = y_0 \in U
\]

**Proof:** Let \( B \) be a closed ball of radius \( R \) centered at \( y_0 \) such that \( f \) has Lipschitz constant \( K \) on \( B \). Then
\[
y_1(t) = y_0 + \int_0^t f(y_0) \, ds
\]
and if \( y_n(t) \) has been obtained,
\[
y_{n+1}(t) = y_0 + \int_0^t f(y_n(s)) \, ds \quad (7.4.15)
\]
Now \( t < T \) where \( T \) is so small that \( \|f(y_0)\| t e^{KT} < R \).

**Claim:** \( \|y_n(t) - y_{n-1}(t)\| \leq \|f(y_0)\| t^n K^{n-1} \frac{1}{(n-1)!} \).

**Proof of claim:** First
\[
\|y_1(t) - y_0\| \leq \int_0^t \|f(y_0)\| \, ds \leq \|f(y_0)\| t
\]
Now suppose it is so for \( n \). Then
\[
\|y_{n+1}(t) - y_n(t)\| \leq \int_0^t \|f(y_n(s)) - f(y_{n-1}(s))\| \, ds
\]
By induction, \( y_n(s), y_{n-1}(s) \) are still in \( B \). This is because

\[
\|y_n(t) - y_0\| \leq \sum_{k=1}^{n} \|y_k(t) - y_{k-1}(t)\|
\]

\[
\leq \sum_{k=1}^{n} \|f(y_0)\| \frac{1}{(k-1)!} t^k K^{k-1}
\]

\[
\leq \|f(y_0)\| t e^{Kt} < R \tag{7.4.16}
\]

showing that \( y_n(t) \) stays in \( B \). Then since all values of the iterates remain in \( B \), induction gives

\[
\|y_{n+1}(t) - y_n(t)\| \leq \int_0^t K \|y_n(s) - y_{n-1}(s)\| ds
\]

\[
\leq K \int_0^t \|f(y_0)\| \frac{1}{(n-1)!} s^n K^{n-1} ds = K^n \frac{1}{(n-1)!} \|f(y_0)\| \int_0^t s^n ds
\]

\[
= K^n \frac{1}{n!} \|f(y_0)\| t^{n+1}
\]

which proves the claim. Since the inequality of the claim shows that \( \|y_n - y_{n-1}\| \) is summable, it follows that \( \{y_n\} \) is a Cauchy sequence in \( C([0,T], X) \). It satisfies \( \|y_n - y_0\| < R \) and so \( y_n \) converges uniformly to some \( y \in C([0,T], X) \). Hence one can pass to a limit in \( 7.4.15 \) and obtain

\[
y(t) = y_0 + \int_0^t f(y(s)) \, ds
\]

for \( t \in [0,T] \). Also \( \|y(t) - y_0\| \leq R \) and on \( B(y_0, R) \), \( f \) is Lipschitz continuous so Gronwall’s inequality gives uniqueness of solutions which remain in \( B \).

Here is an alternate proof which other than the ugly lemma, seems more elegant to me.

**Lemma 7.4.9** Define

\[
\gamma(x) = \begin{cases} 
  x & \text{if } \|x - y_0\| \leq R \\
  y_0 + \frac{x - y_0}{\|x - y_0\|} R & \text{if } \|x - y_0\| > R
\end{cases}
\]

Then \( \|\gamma(x) - \gamma(y)\| \leq 3 \|x - y\| \) for all \( x, y \in X \). Thus

\[
\|\gamma(x) - y_0\| \leq R.
\]

**Proof:** In case both of \( x, y \) are in \( B = \overline{B(y_0, R)} \), there is nothing to show. Suppose then that \( \|y - y_0\| \leq R \) but \( \|x - y_0\| > R \). Then, assuming \( y - y_0 \neq 0 \),

\[
\|\gamma(x) - \gamma(y)\| = \left\| y_0 + \frac{x - y_0}{\|x - y_0\|} R - y \right\| = \left\| \frac{x - y_0}{\|x - y_0\|} R - (y - y_0) \right\|
\]
Then consider for \( y \neq 0 \) the case where both \( x, y \) are in \( X \setminus \mathcal{B} \). In this case, you get

\[
\| x - y \| - \| y - y_0 \| \leq \| y - x \|
\]

Now

\[
B = (R - \| y - y_0 \|) < \| x - y_0 \| - \| y - y_0 \| \leq \| y - x \|
\]

\[
A \leq \left\| \frac{x - y_0}{\| x - y_0 \|} R - \frac{(y - y_0)}{\| y - y_0 \|} R \right\| \leq \frac{R}{\| x - y_0 \| \| y - y_0 \|} \left( \| (x - y_0) \| y - y_0 \| - (y - y_0) \| x - y_0 \| \right)
\]

\[
\leq \frac{R}{\| x - y_0 \| \| y - y_0 \|} (\| y - y_0 \| \| x - y \| + \| y - y_0 \| \| y - x \|)
\]

\[
\leq \frac{R}{\| x - y_0 \|} (\| x - y \| + \| y - x \|) < 2 \| y - x \|
\]

In case \( y = y_0 \), you have

\[
\| \gamma(x) - \gamma(y) \| = \left\| \frac{x - y_0}{\| x - y_0 \|} R \right\| = \left\| \frac{x - y}{\| x - y_0 \|} R \right\| < \| x - y \|
\]

The only other case is where both \( x, y \) are in \( X \setminus \mathcal{B} \). In this case, you get

\[
\| \gamma(x) - \gamma(y) \| = \left\| y_0 + \frac{x - y_0}{\| x - y_0 \|} R - \left( y_0 + \frac{y - y_0}{\| y - y_0 \|} R \right) \right\|
\]

\[
= \left\| \frac{x - y_0}{\| x - y_0 \|} R - \frac{y - y_0}{\| y - y_0 \|} R \right\| \leq 2 \| x - y \|
\]

by the same reasoning used above to estimate \( A \). ■

Alternate Proof of Theorem 7.4.9: Let \( B \) be a closed ball of radius \( R \) centered at \( y_0 \) such that \( f \) has Lipschitz constant \( K \) on \( B \). Let \( \gamma \) be as in Lemma 7.4.8. Consider \( g(x) \equiv f(\gamma(x)) \). Then

\[
\| g(x) - g(y) \| = \| f(\gamma(x)) - f(\gamma(y)) \| \leq K \| \gamma(x) - \gamma(y) \| \leq 3K \| x - y \|.
\]

Now consider for \( y \in C([0, T], X) \)

\[
F_y(t) \equiv y_0 + \int_0^t g(y(s)) \, ds
\]

Then

\[
\| F_y(t) - F_z(t) \| \leq \int_0^t K \| y(s) - z(s) \| \, ds
\]
Thus, iterating this inequality, it follows that a large enough power of $F$ is a contraction map. Therefore, there is a unique fixed point. Now letting $y$ be this fixed point,

$$
\|y(t) - y_0\| \leq \int_0^t 3K \|y(s) - y_0\| \, ds + \|f(y_0)\| \, T
$$

It follows from Gronwall’s inequality that

$$
\|y(t) - y_0\| \leq \|f(y_0)\| \, T e^{3KT}
$$

Choosing $T$ small enough, it follows that $\|y(t) - y_0\| < R$ on $[0, T]$ and so $\gamma$ has no effect. Thus this yields a local solution to the initial value problem.

In the case that $U = X$, the above argument shows that there exists a solution on some $[0, T)$ where $T$ is maximal.

$$
y(t) = y_0 + \int_0^t f(y(s)) \, ds, \quad t < T
$$

Suppose $T < \infty$. Suppose $\int_0^T \|f(y(s))\| \, ds < \infty$. Then you can consider $y_0 + \int_0^T f(y(s)) \, ds$ as an initial condition for the equation and obtain a unique solution $z$ valid on $[T, T + \delta]$. Then one could consider $\check{y}(t) = y(t)$ for $t < T$ and for $t \geq T$, $\check{y}(t) = z(t)$. Then for $t \in [T, T + \delta]$,

$$
\check{y}(t) = z(t) = y_0 + \int_0^T f(y(s)) \, ds + \int_T^t f(\check{y}(s)) \, ds
$$

and so in fact, for all $t \in [0, T + \delta]$,

$$
\check{y}(t) = y_0 + \int_0^t f(\check{y}(s)) \, ds
$$

contrary to the maximality of $T$. Hence it cannot be the case that $T < \infty$. Thus it must be the case that $\int_0^T \|f(y(s))\| \, ds = \infty$ if the solution is not global.

From the above observation, here is a corollary.

**Corollary 7.4.10** Let $f : X \to X$ be locally Lipschitz where $X$ is a Banach space. Then there exists a unique local solution to the IVP

$$
y' = f(y), \quad y(0) = y_0
$$

If $f$ is bounded, then in fact the solutions exists on $[0, T]$ for any $T > 0$.

**Proof:** Say $\|f(x)\| \leq M$ for all $M$. Then letting $[0, \hat{T})$ be the maximal interval, it must be the case that $\int_0^{\hat{T}} \|f(y(t))\| \, dt = \infty$, but this does not happen if $f$ is bounded.
Note that this conclusion holds just as well if $f$ has linear growth, $\|f(u)\| \leq a + b\|u\|$ for $a, b \geq 0$. One just uses an application of Gronwall’s inequality to verify a similar conclusion.

One can also give a simple modification of these theorems as follows.

**Corollary 7.4.11** Suppose $f : X \to X$ is continuous and $f$ is locally Lipschitz on $U$, an open subset of $X$, a Banach space. Suppose also that $f(x) = 0$ for all $x \notin U$ and that $\|f(x)\| < M$ for all $x \in X$. Then there exists a solution to the IVP

$$y' = f(y), \quad y(0) = y_0$$

for $t \in [0, T]$ for any $T > 0$.

**Proof:** Let $T$ be given. If $y_0 \notin U$, there is nothing to show. The solution is $y(t) \equiv y_0$. Suppose then that $y_0 \in U$. Then by Theorem 7.4.8, there exists a unique solution to the initial value problem on an interval $[0, \hat{T})$ of maximal length. If $\hat{T} = T$, then as $t_n \to T$, $\{y(t_n)\}$ must converge. This is because for $t_m < t_n$,

$$\|y(t_n) - y(t_m)\| \leq M |t_n - t_m|$$

showing that this is a Cauchy sequence. Since all such sequences lead to a Cauchy sequence, it must be the case that $\lim_{t \to T} y(t)$ exists. Thus it equals

$$y_0 + \int_0^T f(y(t)) \, dt$$

We let $y(T)$ equal the above and it follows from Gronwall’s inequality that there is a unique solution to the IVP on $[0, T]$ so the claim is true in this case.

Otherwise, if $\hat{T} < T$, then one can define

$$y(\hat{T}) \equiv y_0 + \int_0^{\hat{T}} f(y(s)) \, ds$$

If $y(\hat{T}) \in U$, then by the assumption that $f$ is bounded, one could consider a new initial condition and extend the solution further violating the maximality of the length of $[0, \hat{T})$. Therefore, it must be the case that $y(\hat{T}) \in U^C$. Then the solution is

$$\dot{y}(t) = \begin{cases} y(t), & t < \hat{T} \\ y(\hat{T}), & t > \hat{T} \end{cases}$$

because $f\left( y(\hat{T}) \right) = 0$ by assumption. □

One could also change the above argument for Corollary 7.4.11 to include the case that $f$ has linear growth.
7.4.3 Mountain Pass Theorem In Banach Space

In this section, is a more general version of the mountain pass theorem. It is generalized in two ways. First, the space is not a Hilbert space and second, the derivative of the functional is not assumed to be Lipschitz. Instead of using $I'$ one uses the pseudogradient in an appropriate differential equation. This is a significant generalization because there is no convenient projection map from $X'$ to $X$ like there is in Hilbert space. This is why the use of the pseudogradient is so interesting. For many more considerations of this sort of thing, see [29]. First is a deformation theorem. Here $I$ will be defined on a Banach space $X$ and $I'(x) \in X'$. First recall the Palais Smale conditions.

**Definition 7.4.12** A functional $I$ satisfies the Palais Smale conditions means that if $\{I(u_k)\}$ is a bounded sequence and $I'(u_k) \to 0$, then $\{u_k\}$ is precompact. That is, it has a subsequence which converges.

Here is a picture which illustrates the main conclusion of the following theorem. The idea is that you modify the functional on some set making it smaller and leaving it unchanged off that set.

\[
\begin{align*}
I(u) & \leq c - \delta & \text{if } I(u) \leq c + \delta.
\end{align*}
\]

**Theorem 7.4.13** Let $I$ be $C^1$, $I$ is non constant, satisfy the Palais Smale condition, and $I'$ is bounded on bounded sets. Also suppose that $c \in \mathbb{R}$ is such that either $I^{-1}([c-\delta,c+\delta]) = \emptyset$ for some $\delta > 0$ or $I^{-1}([c-\delta,c+\delta]) \neq \emptyset$ for all $\delta > 0$ and **IF** $I(u) = c$, then $I'(u) \neq 0$. Then for each sufficiently small $\varepsilon > 0$, there is a constant $\delta \in (0,\varepsilon)$ and a function $\eta : [0,1] \times X \to X$ such that

1. $\eta(0,u) = u$
2. $\eta(1,u) = u$ on $I^{-1}(X \setminus (c-\varepsilon,c+\varepsilon))$
3. $I(\eta(t,u)) \leq I(u)$
4. $\eta(1,I^{-1}(-\infty,c+\delta]) \subseteq I^{-1}(-\infty,c-\delta]$, so $I(\eta(1,u)) \leq c - \delta$ if $I(u) \leq c + \delta$. 


The main part of this conclusion is the statement about \( u \to \eta(1, u) \) contained in parts 2. and 4. The other two parts are there to facilitate these two although they are certainly interesting for their own sake.

**Proof:** Suppose \( I^{-1}\left(\left[c - \delta, c + \delta\right]\right) = \emptyset \) for some \( \delta > 0 \). Then

\[
I^{-1}\left(\left(-\infty, c + \frac{\delta}{2}\right]\right) \subseteq I^{-1}\left(\left(-\infty, c - \frac{\delta}{2}\right]\right)
\]

and you could take \( \varepsilon = \delta \) and let \( \eta(t, u) = u \). The conclusion holds with \( \delta = \delta/2 \).

Therefore, assume \( I^{-1}\left(\left[c - \delta, c + \delta\right]\right) \neq \emptyset \) for all \( \delta > 0 \). Since \( I \) is nonconstant, \( \varepsilon > 0 \) can be chosen small enough that

\[
I^{-1}(X \setminus (c - \varepsilon, c + \varepsilon)) \neq \emptyset.
\]

Always let \( \varepsilon \) be this small. Note that \( I \) nonconstant is part of the assumptions.

**Claim 1:** For all small enough \( \varepsilon > 0 \), if \( u \in I^{-1}(\left[c - \varepsilon, c + \varepsilon\right]) \), then \( I'(u) \neq 0 \) and in fact, for such \( \varepsilon \), there exists \( \sigma(\varepsilon) > 0 \), such that \( \sigma(\varepsilon) < \min(\varepsilon, 1), \|I'(u)\| > \sigma(\varepsilon) \) for all \( u \in I^{-1}(\left[c - \varepsilon, c + \varepsilon\right]) \).

**Proof of Claim 1:** If the claim is not so, then there is \( \{u_k\}, \varepsilon_k, \sigma_k \to 0, \|I'(u_k)\|\chi < \sigma_k \), and \( I(u_k) \in [c - \varepsilon_k, c + \varepsilon_k] \) but \( \|I'(u_k)\|\chi \leq \sigma_k \). However, from the Palais Smale condition, there is a subsequence, still denoted as \( u_k \) which converges to some \( u \). Now \( I(u_k) \in [c - \varepsilon_k, c + \varepsilon_k] \) and so \( I(u) = c \) while \( I'(u) = 0 \) contrary to the hypothesis. This proves Claim 1. From now on, \( \varepsilon \) will be sufficiently small that this holds.

Now define for \( \delta < \varepsilon \) (The precise description of small \( \delta \) will be described later. However, it will be \( \delta < \sigma(\varepsilon)/2 \), but this exact description is only used at the end.)

\[
A \equiv I^{-1}(X \setminus (c - \varepsilon, c + \varepsilon)) \\
B \equiv I^{-1}(\left[c - \delta, c + \delta\right])
\]

Thus \( A \) and \( B \) are disjoint closed sets. Recall that it is assumed that \( B \neq \emptyset \) since otherwise, there is nothing to prove. Also it is assumed throughout that \( \varepsilon > 0 \) is such that \( A \neq \emptyset \) thanks to \( I \) not being constant. Thus these are nonempty sets and we do not have to fuss with worrying about meaning when one is empty.

**Claim 2:** For any \( u \), \( \text{dist}(u, A) + \text{dist}(u, B) > 0 \).

This is so because if not, then both summands would be zero and this requires that \( u \in A \cap B \) since these sets are closed. But \( A \cap B = \emptyset \).

Now define a function

\[
g(u) \equiv \frac{\text{dist}(u, A)}{\text{dist}(u, A) + \text{dist}(u, B)}
\]

It is a continuous function of \( u \) which has values in \([0, 1]\). It is 1 on \( B \) and 0 on \( A \). Also define \( V(x) \) as a pseudogradient field for \( I \) on the regular points of \( I \). At points where \( I'(x) = 0 \), let \( V(x) = 0 \). Recall what this means:

\[
\|I'(x)\|_{\chi'}^2 \leq \langle I'(x), V(x) \rangle, \quad \|V(x)\|_{\chi} \leq 2 \|I'(x)\|_{\chi}.
\]
7.4. CRITICAL POINTS

and also \( V \) is locally Lipschitz on the regular points of \( I \). Thus \( x \rightarrow V(x) \) is continuous on \( X \), thanks to continuity of \( I' \), satisfies the above inequalities, and is locally Lipschitz on \( U = \{ x : I'(x) \neq 0 \} \). It exists because of Lemma \[\text{[4.11]}\].

Consider the ordinary differential initial value problem

\[
\eta'(t,u) + g(u) h(\|V(\eta(t,u))\|) V(\eta(t,u)) = 0 \quad (7.4.18)
\]

\[
\eta(0,u) = u \quad (7.4.19)
\]

where \( r \rightarrow h(r) \) is a decreasing function which has values in \((0,1]\) and equals 1 for \( r \in [0,1] \) and equals \( 1/r \) for \( r > 1 \).

Here \( u \) is given and the \( \eta' \) is the time derivative with respect to \( t \). By Corollary \[\text{[4.11]}\] there exists a solution to this for \( t \in [0,1] \).

Then for this solution, \( \eta(0,u) = u \) because of the above initial condition. If \( u \in \mathcal{I}^{-1}(X \setminus \{c - \varepsilon, c + \varepsilon\}) \), then \( u \in \mathcal{A} \) and so \( g(u) = 0 \) so \( \eta(t,u) = u \) for all \( t \in [0,1] \). This gives the first two conditions. Consider the third.

\[
\frac{d}{dt}(I(\eta(t,u))) = \langle I'(\eta), \eta' \rangle = -\langle I'(\eta), g(u) h(\|V(\eta(t,u))\|) V(\eta(t,u)) \rangle
\]

\[
= -g(u) h(\|V(\eta(t,u))\|) \langle I'(\eta), V(\eta(t,u)) \rangle
\]

\[
\leq -g(u) h(\|V(\eta(t,u))\|) \|I'(\eta)\|^2 \leq 0
\]

this last inequality from the inequalities of \[\text{[4.11]}\] and so this implies the third condition since it says that the function \( t \rightarrow I(\eta(t,u)) \) is decreasing.

It remains to consider the last condition. This involves an appropriate choice of small \( \delta \). It was chosen small and now it will be seen how small. It is desired to verify that

\[
\eta(1, \mathcal{I}^{-1}((\infty,c + \delta))) \subseteq \mathcal{I}^{-1}((\infty,c - \delta])
\]

Suppose it is not so. Then there exists \( u \) such that \( I(u) \in (c - \delta, c + \delta] \) but \( I(\eta(1,u)) > c - \delta \). We can assume that \( I(u) \in (c - \delta, c + \delta] \) because if \( I(u) \leq c - \delta \), then so is \( I(\eta(1,u)) \) from what was just shown. Hence \( g(u) = 1 \). Then using the fact that \( g(u) = 1 \),

\[
c - \delta < I(\eta(1,u)) = I(u) + \int_0^1 \frac{d}{dt}(I(\eta)) \, dt
\]

\[
= I(u) - \int_0^1 \langle I'(\eta), h(\|V(\eta)\|) V(\eta) \rangle \, dt
\]

\[
= I(u) + \int_0^1 -h(\|V(\eta)\|) \langle I'(\eta), V(\eta) \rangle \, dt
\]
\[ \leq \int_0^1 \left(-h(V(\eta))\|I'(\eta)\|^2\right) dt \]

Then
\[ c - \delta + \int_0^1 h(V(\eta))\|I'(\eta)\|^2 dt < I(u) \leq c + \delta \]

Thus
\[ c - 2\delta + \int_0^1 h(V(\eta))\|I'(\eta)\|^2 dt < c \]

Also, it is being assumed that \( I(\eta(1,u)) > c - \delta \) and so by the third conclusion shown above, \( \eta(t,u) \in B \) for \( t \in [0,1] \). We also know that for such values of \( \eta(t,u),\|I'(\eta(t,u))\| \geq \sigma(\varepsilon) \) from Claim 1. Now
\[ \|I'(x)\|^2_{X'} \leq \langle I'(x),V(x) \rangle \leq \|I'(x)\| \|V(x)\| \]

and so
\[ \|V(\eta(t,u))\|_X \geq \|I'(\eta(t,u))\|_{X'} \geq \sigma(\varepsilon). \tag{7.4.20} \]

Thus the above inequality yields
\[ c - 2\delta + \int_0^1 h(V(\eta))\sigma(\varepsilon)^2 dt < c \]

Now what is the value of \( h(V(\eta)) \)? From
\[ h(V(\eta(t,u))) \leq h(I'(\eta(t,u))) \leq h(\sigma(\varepsilon)) \leq \frac{1}{\sigma(\varepsilon)} \]

In fact, \( \sigma(\varepsilon) < 1 \) so \( h(\sigma(\varepsilon)) = 1 \) so the above estimate, while correct, is sloppy. Hence
\[ c - 2\delta + \int_0^1 \frac{1}{\sigma(\varepsilon)} \sigma(\varepsilon)^2 dt < c \]

So far it was only assumed \( \delta < \varepsilon \). As indicated above, \( \delta \) was chosen small enough that \( -2\delta + \sigma(\varepsilon) > 0 \). Hence this yields a contradiction. Thus the last conclusion is verified. \( \blacksquare \)

Imagine a valley surrounded by a ring of mountains. On the other side of this ring of mountains, there is another low place. Then there must be some path from the valley to the exterior low place which goes through a point where the gradient equals 0, the gradient being the gradient of a function \( f \) which gives the altitude of the land. This is the idea of the mountain pass theorem. The critical point where \( \nabla f = 0 \) is the mountain pass.

**Theorem 7.4.14** Let \( X \) be a Banach space and let \( I : X \to \mathbb{R} \) be a \( C^1 \) functional having \( I' \) bounded on bounded sets and such that \( I \) satisfies the Palais Smale condition. Suppose \( I(0) = 0 \) and \( I(u) \geq a > 0 \) for all \( \|u\| = r \). Suppose also that there exists \( v,\|v\| > r \) such that \( I(v) \leq 0 \). Then define
\[ \Gamma \equiv \{ g \in C([0,1] ; X) : g(0) = 0, g(1) = v \} \]
Let
\[ c = \inf_{g \in \Gamma} \max_{0 \leq t \leq 1} I(g(t)) \]

Then \( c \) is a critical value of \( I \) meaning that there exists \( u \) such that \( I(u) = c \) and \( I'(u) = 0 \). In particular, there is \( u \neq 0 \) such that \( I'(u) = 0 \).

**Proof:** First note that \( c \geq a > 0 \). Suppose \( c \) is not a critical value. Then either \( I^{-1}((c-\delta, c+\delta)) = \emptyset \) for some \( \delta > 0 \) in which case the conclusion of the deformation theorem, (Theorem 7.4.13) holds, or for all \( \delta > 0 \), \( I^{-1}((c-\delta, c+\delta)) \neq \emptyset \) and if \( I(u) = c \), then \( I'(u) \neq 0 \) in which case the deformation theorem also holds. Then by this theorem, for \( \varepsilon > 0, \varepsilon \) sufficiently small, \( \varepsilon < c \), there is \( \eta : X \to X \) and a \( \delta < \varepsilon \) small enough that
\[ \eta \left( I^{-1}((\infty, c+\delta]) \right) \subseteq I^{-1}((\infty, c-\delta]) \]
and \( \eta \) leaves unchanged \( I^{-1}(X \setminus (c-\varepsilon, c+\varepsilon)) \). Then there is \( g \in \Gamma \) such that
\[ \max_{t \in [0,1]} I(g(t)) < c + \delta \]

Then in particular, \( I(g(t)) < c + \delta \) for every \( t \). Hence you look at \( \eta \circ g \). We know that \( g(0), g(1) \) are both in the set \( I(u) \notin (c-\varepsilon, c+\varepsilon] \) because they are both 0 or less than 0 and so \( \eta \) leaves these unchanged. Hence \( \eta \circ g \in \Gamma \) and
\[ I(\eta \circ g(t)) \leq c - \delta \]
for all \( t \in [0,1] \). Thus
\[ c = \inf_{g \in \Gamma} \max_{0 \leq t \leq 1} I(g(t)) \leq \max_{t \in [0,1]} I(\eta \circ g(t)) \leq c - \delta \]
which is clearly a contradiction.

### 7.5 The min max Theorem

Here is the min max theorem. The proof given follows Brezis which is where I found it. A function \( f \) is convex if
\[ f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)y \]

It is concave if the inequality is turned around. It can be shown that in finite dimensions, convex functions are automatically continuous, similar for concave functions. Recall the following definition of upper and lower semicontinuous functions defined on a metric space and having values in \([\infty, \infty] \).

**Definition 7.5.1** A function is upper semicontinuous if whenever \( x_n \to x \), it follows that \( f(x) \geq \limsup_{n \to \infty} f(x_n) \) and it is lower semicontinuous if \( f(x) \leq \liminf_{n \to \infty} f(x_n) \).
Lemma 7.5.2 If $F$ is a set of functions which are upper semicontinuous, then
$g(x) \equiv \inf \{f(x) : f \in F\}$ is also upper semicontinuous. Similarly, if $F$ is a set of functions which are lower semicontinuous, then if $g(x) \equiv \sup \{f(x) : f \in F\}$ it follows that $g$ is lower semicontinuous.

Proof: Let $f \in F$ where these functions are upper semicontinuous. Then if $x_n \to x$, and $g(x) \equiv \inf \{f(x) : f \in F\}$,
$$f(x) \geq \limsup_{n \to \infty} f(x_n) \geq \limsup_{n \to \infty} g(x_n)$$
Since this is true for each $f \in F$, then it follows that you can take the infimum and obtain $g(x) \geq \limsup_{n \to \infty} g(x_n)$. Similarly, lower semicontinuity is preserved on taking sup.

Note that in a metric space, the above definitions upper and lower semicontinuity in terms of sequences are equivalent to the definitions that
$$f(x) \geq \lim_{r \to 0} \sup \{f(y) : y \in B(x,r)\}$$
$$f(x) \leq \lim_{r \to 0} \inf \{f(y) : y \in B(x,r)\}$$
respectively.

Here is a technical lemma which will make the proof shorter. It seems fairly interesting also.

Lemma 7.5.3 Suppose $H : A \times B \to \mathbb{R}$ is strictly convex in the first argument and concave in the second argument where $A, B$ are compact convex nonempty subsets of Banach spaces $E, F$ respectively and $x \to H(x,y)$ is lower semicontinuous while $y \to H(x,y)$ is upper semicontinuous. Let
$$H(g(y),y) \equiv \min_{x \in A} H(x,y)$$
Then $g(y)$ is uniquely defined and also for $t \in [0,1]$,
$$\lim_{t \to 0} g(y + t(z - y)) = g(y).$$

Proof: First suppose both $z, w$ yield the definition of $g(y)$. Then
$$H\left(\frac{z + w}{2}, y\right) < \frac{1}{2} H(z, y) + \frac{1}{2} H(w, y)$$
which contradicts the definition of $g(y)$. As to the existence of $g(y)$ this is nothing more than the theorem that a lower semicontinuous function defined on a compact set achieves its minimum.

Now consider the last claim about "hemicontinuity". For all $x \in A$, it follows from the definition of $g$ that
$$H(g(y + t(z - y)), y + t(z - y)) \leq H(x, y + t(z - y))$$
7.5. THE MIN MAX THEOREM

By concavity of $H$ in the second argument,

\begin{equation}
(1-t)H(g(y+t(z-y)),y)+tH(g(y+t(z-y)),z) \tag{7.5.21}
\end{equation}

\begin{equation}
\leq H(x,y+t(z-y)) \tag{7.5.22}
\end{equation}

Now let $t_n \to 0$. Does $g(y+t_n(z-y)) \to g(y)$? Suppose not. By compactness, $g(y+t_n(z-y))$ is in a compact set and so there is a further subsequence, still denoted by $t_n$ such that

$$g(y+t_n(z-y)) \to \hat{x} \in A$$

Then passing to a limit in (7.5.22), one obtains, using the upper semicontinuity in one and lower semicontinuity in the other the following inequality.

$$H(\hat{x},y) \leq \liminf_{n \to \infty} (1-t_n)H(g(y+t_n(z-y)),y) + \liminf_{n \to \infty} \inf_{n \to \infty} t_nH(g(y+t_n(z-y)),z)$$

$$\leq \liminf_{n \to \infty} \left( (1-t_n)H(g(y+t_n(z-y)),y) + t_nH(g(y+t_n(z-y)),z) \right)$$

$$\leq \limsup_{n \to \infty} H(x,y+t_n(z-y)) \leq H(x,y)$$

This shows that $\hat{x} = g(y)$ because this holds for every $x$. Since $t_n \to 0$ was arbitrary, this shows that in fact

$$\lim_{t \to 0^+} g(y+t(z-y)) = g(y) \blacksquare$$

Now with this preparation, here is the min-max theorem. A norm is called strictly convex if whenever $x \neq y$, $\|\frac{x+y}{2}\| < \|\frac{x}{2}\| + \|\frac{y}{2}\|$. 

**Theorem 7.5.4** Let $E, F$ be Banach spaces with $E$ having a strictly convex norm. Also suppose that $A \subseteq E, B \subseteq F$ are compact and convex sets and that $H : A \times B \to \mathbb{R}$ is such that

$x \to H(x,y)$ is convex

$y \to H(x,y)$ is concave

Thus $H$ is continuous in each variable in the case of finite dimensional spaces. Here assume that $x \to H(x,y)$ is lower semicontinuous and $y \to H(x,y)$ is upper semicontinuous. Then

$$\min_{x \in A} \max_{y \in B} H(x,y) = \max_{y \in B} \min_{x \in A} H(x,y)$$

This condition is equivalent to the existence of $(x_0,y_0) \in A \times B$ such that

$$H(x_0,y) \leq H(x_0,y_0) \leq H(x,y_0) \text{ for all } x, y \tag{7.5.23}$$
Proof: One part of the main equality is obvious.

\[
\max_{y \in B} H(x, y) \geq H(x, y) \geq \min_{x \in A} H(x, y)
\]

and so for each \(x\),

\[
\max_{y \in B} H(x, y) \geq \max_{y \in B} \min_{x \in A} H(x, y)
\]

and so

\[
\min_{x \in A} \max_{y \in B} H(x, y) \geq \max_{y \in B} \min_{x \in A} H(x, y)
\]

Next consider the other direction.

Define \(H_\varepsilon(x, y) \equiv H(x, y) + \varepsilon \|x\|\) where \(\varepsilon > 0\). Then \(H_\varepsilon\) is strictly convex in the first variable. This results from the observation that

\[
\frac{x + y}{2} < \left( \frac{\|x\| + \|y\|}{2} \right)
\]

Then by Lemma 7.5.3 there exists a unique \(x \equiv g(y)\) such that

\[
H_\varepsilon(g(y), y) = \min_{x \in A} H_\varepsilon(x, y)
\]

and also, whenever \(y, z \in A\),

\[
\lim_{t \to 0^+} g(y + t(z - y)) = g(y).
\]

Thus \(H_\varepsilon(g(y), y) = \min_{x \in A} H_\varepsilon(x, y)\). But also this shows that \(y \to H_\varepsilon(g(y), y)\) is the minimum of functions which are upper semicontinuous and so this function is also upper semicontinuous. Hence there exists \(y^*\) such that

\[
\max_{y \in B} H_\varepsilon(g(y), y) = H_\varepsilon(g(y^*), y^*) = \max_{y \in B} \min_{x \in A} H_\varepsilon(x, y)
\]

Thus from concavity in the second argument and what was just defined, for \(t \in (0, 1)\),

\[
H_\varepsilon(g(y^*), y^*) \geq H_\varepsilon(g((1-t)y^* + ty), (1-t)y^* + ty)
\]

\[
\geq (1-t) H_\varepsilon(g((1-t)y^* + ty), y^*) + tH_\varepsilon(g((1-t)y^* + ty), y)
\]

\[
\geq (1-t) H_\varepsilon(g(y^*), y^*) + tH_\varepsilon(g((1-t)y^* + ty), y)
\]

This is because \(\min_{x \in A} H_\varepsilon(x, y^*) = H_\varepsilon(g(y^*), y^*)\) so \(H_\varepsilon(g((1-t)y^* + ty), y^*) \geq H_\varepsilon(g(y^*), y^*)\). Then subtracting the first term on the right, one gets

\[
tH_\varepsilon(g(y^*), y^*) \geq tH_\varepsilon(g((1-t)y^* + ty), y)
\]

and cancelling the \(t\),

\[
H_\varepsilon(g(y^*), y^*) \geq H_\varepsilon(g((1-t)y^* + ty), y)
\]
Now apply Lemma 7.5.3 and let \( t \to 0^+ \). This along with lower semicontinuity yields

\[
H_\varepsilon(g(y^*), y^*) \geq \liminf_{t \to 0^+} H_\varepsilon(g((1-t)y^* + ty), y) = H_\varepsilon(g(y^*), y)
\]  
(7.5.27)

Hence for every \( x, y \)

\[
H_\varepsilon(x, y^*) \geq H_\varepsilon(g(y^*), y^*) \geq H_\varepsilon(g(y^*), y)
\]

Thus

\[
\min_x H_\varepsilon(x, y^*) \geq H_\varepsilon(g(y^*), y^*) \geq \max_y H_\varepsilon(g(y^*), y)
\]

and so

\[
\max_{y \in B} \min_{x \in A} H_\varepsilon(x, y) \geq \min_x H_\varepsilon(x, y^*) \geq \max_y H_\varepsilon(g(y^*), y) \geq \min_x \max_y H_\varepsilon(x, y)
\]

Thus, letting \( C \equiv \max \{ ||x|| : x \in A \} \)

\[
C + \max_{y \in B} \min_{x \in A} H(x, y) \geq \min_{x \in A} \max_{y \in B} H(x, y)
\]

Since \( \varepsilon \) is arbitrary, it follows that

\[
\max_{y \in B} \min_{x \in A} H(x, y) \geq \min_{x \in A} \max_{y \in B} H(x, y)
\]

This proves the first part because it was shown above in (7.5.24) that

\[
\min_{x \in A} \max_{y \in B} H(x, y) \geq \max_{y \in B} \min_{x \in A} H(x, y)
\]

Now consider (7.5.24) about the existence of a “saddle point” given the equality of \( \min \max \) and \( \max \min \). Let

\[
\alpha = \max_{y \in B} \min_{x \in A} H(x, y) = \min_{x \in A} \max_{y \in B} H(x, y)
\]

Then from

\[
y \to \min_{x \in A} H(x, y) \text{ and } x \to \max_{y \in B} H(x, y)
\]

being upper semicontinuous and lower semicontinuous respectively, there exist \( y_0 \) and \( x_0 \) such that

\[
\alpha = \min_{x \in A} H(x, y_0) = \max_{y \in B} \min_{x \in A} H(x, y) = \min_{x \in A} \max_{y \in B} H(x, y)
\]

Then

\[
\alpha = \max_{y \in B} H(x_0, y) \geq H(x_0, y_0)
\]

\[
\alpha = \min_{x \in A} H(x, y_0) \leq H(x_0, y_0)
\]
so in fact $\alpha = H(x_0, y_0)$ and from the above equalities,

$$H(x_0, y_0) = \alpha = \min_{x \in A} H(x, y_0) \leq H(x, y_0)$$

$$H(x_0, y_0) = \alpha = \max_{y \in B} H(x_0, y) \geq H(x_0, y)$$

and so

$$H(x_0, y) \leq H(x_0, y_0) \leq H(x, y_0)$$

Thus if the min-max condition holds, then there exists a saddle point, namely $(x_0, y_0)$.

Finally suppose there is a saddle point $(x_0, y_0)$ where

$$H(x_0, y) \leq H(x_0, y_0) \leq H(x, y_0)$$

Then

$$\min_{x \in A} \max_{y \in B} H(x, y) \leq \max_{y \in B} \min_{x \in A} H(x, y)$$

However, as noted above, it is always the case that

$$\max_{y \in B} \min_{x \in A} H(x, y) \leq \min_{x \in A} \max_{y \in B} H(x, y) \quad \blacksquare$$

Of course all of this works with no change if you have $E, F$ reflexive Banach spaces and the sets $A, B$ are just closed and bounded and convex. Then you just use the fact that the functional is weakly lower semicontinuous in the first variable and weakly upper semicontinuous in the second. Recall that lower semicontinuous and convex implies weakly lower semicontinuity. Then just use weak convergence instead of strong convergence in the above argument. Recall that closed bounded and convex sets with the weak topology can be considered metric spaces. I think the above is most interesting in finite dimensions. Of course in this case, you can simply assume the norm is the standard Euclidean norm and there is then no need to assume one of the norms is strictly convex. It comes automatically. Just use an equivalent norm which is strictly convex.

### 7.6 Exercises

1. Let $f: X \to [-\infty, \infty]$ where $X$ is a Banach space. This is said to be lower semicontinuous if whenever $x_n \to x$, it follows that $f(x) \leq \liminf_{n \to \infty} f(x_n)$. Show that this is the same as saying that the epigraph of $f$ is closed. Here we can make $X \times [-\infty, \infty]$ into a metric space in a natural way by using the product topology where the distance on $[-\infty, \infty]$ will be $d(\sigma, \alpha) \equiv |\arctan(\sigma) - \arctan(\alpha)|$. Here $\text{epi}(f) \equiv \{(x, \alpha) : \alpha \geq f(x)\}$. The function is upper semicontinuous if $\limsup_{n \to \infty} f(x_n) \leq f(x)$. What is a condition for $f$ to be upper semicontinuous? Do you need a Banach space to do this? Would it be sufficient to let $X$ be a metric space?
7.6. EXERCISES

2. Explain why the supremum of lower semicontinuous functions is lower semicontinuous and the infimum of upper semicontinuous functions is upper semicontinuous.

3. Let \( K \) be a nonempty closed and convex subset of \( \mathbb{R}^n \). Recall \( K \) is convex means that if \( x, y \in K \), then for all \( t \in [0, 1] \), \( tx + (1 - t)y \in K \). Show that if \( x \in \mathbb{R}^n \) there exists a unique \( z \in K \) such that
\[
|x - z| = \min \{ |x - y| : y \in K \}.
\]
This \( z \) will be denoted as \( P_\mathcal{X}x \). **Hint:** First note you do not know \( K \) is compact. Establish the parallelogram identity if you have not already done so,
\[
|u - v|^2 + |u + v|^2 = 2|u|^2 + 2|v|^2.
\]
Then let \( \{z_k\} \) be a minimizing sequence,
\[
\lim_{k \to \infty} |z_k - x|^2 = \inf \{ |x - y| : y \in K \} \equiv \lambda.
\]
Now using convexity, explain why
\[
\frac{|z_k - z_m|^2}{2} + \left| x - \frac{z_k + z_m}{2} \right|^2 = 2 \left| x - \frac{z_k}{2} \right|^2 + 2 \left| x - \frac{z_m}{2} \right|^2
\]
and then use this to argue \( \{z_k\} \) is a Cauchy sequence. Then if \( z_i \) works for \( i = 1, 2 \), consider \( (z_1 + z_2)/2 \) to get a contradiction.

4. In Problem 3 show that \( P_\mathcal{X}x \) satisfies and is in fact characterized as the solution to the following variational inequality:
\[
(x - P_\mathcal{X}x, y - P_\mathcal{X}x) \leq 0
\]
for all \( y \in K \). Then show that \( |P_\mathcal{X}x_1 - P_\mathcal{X}x_2| \leq |x_1 - x_2| \). **Hint:** For the first part note that if \( y \in K \), the function \( t \to |x - (P_\mathcal{X}x + t(y - P_\mathcal{X}x))|^2 \) achieves its minimum on \([0,1]\) at \( t = 0 \). For the second part,
\[
(x_1 - P_\mathcal{X}x_1) \cdot (P_\mathcal{X}x_2 - P_\mathcal{X}x_1) \leq 0, \ (x_2 - P_\mathcal{X}x_2) \cdot (P_\mathcal{X}x_1 - P_\mathcal{X}x_2) \leq 0.
\]
Explain why
\[
(x_2 - P_\mathcal{X}x_2 - (x_1 - P_\mathcal{X}x_1)) \cdot (P_\mathcal{X}x_2 - P_\mathcal{X}x_1) \geq 0
\]
and then use a some manipulations and the Cauchy Schwarz inequality to get the desired inequality. Thus \( P \) is called a retraction onto \( K \).

5. Browder’s lemma says: Let \( K \) be a convex closed and bounded set in \( \mathbb{R}^n \) and let \( A : K \to \mathbb{R}^n \) be continuous and \( f \in \mathbb{R}^n \). Then there exists \( x \in K \) such that for all \( y \in K \),
\[
(f - Ax, y - x) \leq 0
\]
CHAPTER 7. FIXED POINT THEOREMS AND MORE

show this is true. Hint: Consider $x \rightarrow P(f - Ax + x)$ where $P$ is the projection onto $K$. If there is a fixed point of this mapping, then $P(f - Ax + x) = x$. Now consider the variational inequality satisfied. This little lemma is the basis for a whole lot of nonlinear analysis involving nonlinear operators of various kinds.

6. Generalize the above problem as follows. Let $K$ be a convex closed and bounded set in $\mathbb{R}^n$ and let $A : K \rightarrow \mathcal{P}(\mathbb{R}^n)$ be upper semi-continuous having closed bounded convex values and $f \in \mathbb{R}^n$. Then there exists $x \in K$ and $z \in Ax$ such that for all $y \in K$,

$$\langle f - z, y - x \rangle \leq 0$$

show this is true. Also show that if $K$ is a closed convex and bounded set in $E$ a finite dimensional normed linear space and $A : K \rightarrow \mathcal{P}(E')$ is upper semicontinuous having closed bounded convex values and $f \in E'$, then there exists $x \in K$ and $z \in Ax$ such that for all $y \in K$,

$$\langle f - z, y - x \rangle \leq 0.$$

Hint: Use the construction for the proof of the Kakutani fixed point theorem and the above Browder’s lemma.

7. This problem establishes a remarkable result about existence for a system of inequalities based on the min max theorem. Let $E$ be a finite dimensional Banach space and let $K$ be a convex and compact subset of $E$. A set valued map $A : D(A) \subseteq K \rightarrow E'$ is called monotone if whenever $v_i \in Au$, it follows that $\langle v_1 - v_2, u_1 - u_2 \rangle \geq 0$. The graph, denoted as $\mathcal{G}(A)$ consists of all pairs $[u, v]$ such that $v \in Au$. This is a monotone subset of $E \times E'$. Let $z \in E'$ be fixed. Show that for $[u_i, v_i] \in \mathcal{G}(A)$, for $i = 1, 2, \cdots, n$ there exists a solution $x \in K$ to the system of inequalities

$$\langle z + v_i, u_i - x \rangle \geq 0, \quad i = 1, 2, \cdots, n$$

Hint: Let $P_n$ be all $\bar{\lambda} = (\lambda_1, \cdots, \lambda_n)$ such that each $\lambda_k \geq 0$ and $\sum_{k=1}^{n} \lambda_k = 1$. Let $H : P_n \times P_n \rightarrow \mathbb{R}$ be given by

$$H(\bar{\mu}, \bar{\lambda}) \equiv \sum_{i=1}^{n} \mu_i \left( z + v_i, \sum_{j=1}^{n} \lambda_j u_j - u_i \right)$$

(7.6.28)

Show that it is both convex and concave in both arguments. Then apply the min max theorem. Then argue that $H(\bar{\lambda}, \bar{\lambda}) \leq 0$ from monotonicity considerations. Letting $(\bar{\mu}_0, \bar{\lambda}_0)$ be the saddle point, you will have

$$H(\bar{\mu}, \bar{\lambda}_0) \leq H(\bar{\mu}_0, \bar{\lambda}_0) \leq H(\bar{\mu}_0, \bar{\lambda})$$

$$H(\bar{\mu}, \bar{\lambda}_0) \leq H(\bar{\mu}_0, \bar{\lambda}_0) \leq H(\bar{\mu}_0, \bar{\mu}_0) \leq 0$$

$$H(\bar{\mu}, \bar{\lambda}_0) \leq 0$$
Now choose $\tilde{\mu}$ judiciously while allowing $\tilde{\lambda}_0$ to be used to define $x$ which satisfies all the inequalities.

8. ↑ It gets even better. Let $K_{u,v} \equiv \{ x \in K : \langle z + v, u - x \rangle \geq 0 \}$. Show that $K_{u,v}$ is compact and that the sets $K_{u,v}$ have the finite intersection property. Therefore, there exists $x \in \cap_{[u,v] \in G(A)} K_{u,v}$. Explain why $\langle z + v, u - x \rangle \geq 0$ for all $[u,v] \in G(A)$. What would the inequalities be if $-A$ were monotone?

9. Problem 6 gave a solution to the inequality $\langle f - z, y - x \rangle \leq 0, z \in Ax$ under the condition that $A$ is upper semicontinuous. What are the differences between the result in the above problem and the result of Problem 6. You could replace $A$ with $-A$ in the earlier problem. If you did, would you get the result of the above problem?

10. Are there convenient examples of monotone set valued maps? Yes, there are. Let $X$ be a Banach space and let $\phi : X \to (-\infty, \infty]$ be convex, lower semicontinuous, and proper. See Problem 4 for a discussion of lower semicontinuous. Proper means that $\phi (x) < \infty$ for some $x$. Convex means the usual thing. $\phi (tx + (1-t)y) \leq t\phi (x) + (1-t)\phi (y)$ where $t \in [0,1]$. Then $x^* \in \partial \phi (x)$ means that

$$\langle x^*, z - x \rangle \leq \phi (z) - \phi (x), \text{ for all } z \in X$$

Show that if $x^* \in \partial \phi (x)$, then $\phi (x) < \infty$. The set of points $x$ where $\phi (x) < \infty$ is called the domain of $\phi$ denoted as $D(\phi)$. Also show that if $[x, x^*], [\bar{x}, \bar{x}^*]$ are two points of the graph of $\partial \phi$, then $\langle \bar{x}^* - x^*, \bar{x} - x \rangle \geq 0$ so that $\partial \phi$ is an example of a monotone graph. You might wonder whether this graph is nonempty. See the next problem for a partial answer to this question. Of course the above problem pertains to finite dimensional spaces so you could just take any $\phi : \mathbb{R}^n \to \mathbb{R}$ which is convex and differentiable. You can see that in this case the subgradient coincides with the derivative.

11. This problem gives a simple condition for the subgradient of a convex function to be onto. Let $X$ be a reflexive Banach space and suppose $\phi : X \to (-\infty, \infty]$ is convex, proper, lower semicontinuous, and for all $y^* \in X'$,

$$\lim_{\|x\| \to \infty} \phi (x) - \langle y^*, x \rangle = \infty$$

this last condition being called “coercive”. Show that under these conditions, you can conclude that $\partial \phi$ is not just nonempty for some $x$ but that in fact every $y^* \in X'$ is contained in some $\partial \phi (x)$. Thus $\partial \phi$ is actually onto. Hint: Consider the function $x \to \phi (x) - \langle y^*, x \rangle$. Argue that it is lower semicontinuous. Let

$$\lambda \equiv \inf \{ \phi (x) - \langle y^*, x \rangle : x \in X \}$$

Let $\{x_n\}$ be a minimizing sequence. Argue that from the coercivity condition, $\|x_n\|$ must be bounded. Now use the Eberlein Smulian theorem, Problem 47 on Page 234 or Problem 47 on Page 234 to verify that there is a weakly convergent subsequence $x_n \to x$ weakly. In finite dimensions, you just use the
Heine Borel theorem. You know the epigraph of \( \phi \) intersected with \( X \times \mathbb{R} \) is a convex and closed subset of \( X \times \mathbb{R} \). Explain why this is so. This will require a separation theorem in infinite dimensional space. In finite dimensional space, there isn’t much to show here. Next explain why \( \phi \) must be weakly lower semicontinuous. If you can’t do this part, just use the theorem that a function which is convex and lower semicontinuous is also weakly lower semicontinuous or specialize to finite dimensions and use advanced calculus. That is, if \( x_n \to x \) weakly, then \( \phi(x) \leq \lim\inf_{n \to \infty} \phi(x_n) \). Conclude that \( \lambda > -\infty \) and equals \( \phi(x) - \langle y^*, x \rangle \) which is no larger than \( \phi(z) - \langle y^*, z \rangle \). Now conclude that \( y^* \in \partial \phi(x) \).

12. Let \( \phi : \mathbb{R}^n \to \mathbb{R} \) be convex, proper lower semicontinuous, and bounded below. Show that the graph of \( \partial \phi \) is nonempty. **Hint:** Just consider \( \psi(x) = |x|^2 + \phi(x) \) and observe that this is coercive. Then argue using convexity that \( \partial \psi(x) = \partial \phi(x) + 2x \). (You don’t need to assume that \( \phi \) is bounded below but it is convenient to assume this.)

13. Suppose \( f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \) is continuous and an estimate of the following form holds.

\[
(f(t, x), x) \leq A + B |x|^2
\]

Show that there exists a solution to the initial value problem

\[
x' = f(t, x), \quad x(0) = x_0
\]

for \( t \in [0, T] \).

14. In the above problem, suppose that \( -f + \alpha I \) is monotone for large enough \( \alpha \) in addition to the estimate of that problem. Show that then there is only one solution to the problem. In fact, show that the solution depends continuously on the initial data.

15. It was shown that if \( f : X \to X \) is locally Lipschitz where \( X \) is a Banach space. Then there exists a unique local solution to the IVP

\[
y' = f(y), \quad y(0) = y_0
\]

If \( f \) is bounded, then in fact the solutions exists on \([0, T]\) for any \( T > 0 \). Show that it suffices to assume that \( ||f(y)|| \leq a + b \|y\| \).

16. Suppose \( f(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n \) is continuous and also that \( |f(t, x)| \leq M \) for all \((t, x)\). Show that there exists a solution to the initial value problem

\[
x' = f(t, x), \quad x(0) = x_0 \in \mathbb{R}^n
\]

for \( t \in [0, T] \). **Hint:** You might consider \( T : C([0, T], \mathbb{R}^n) \to C([0, T], \mathbb{R}^n) \) given by

\[
Fx(t) \equiv x_0 + \int_0^t f(s, x(s)) \, ds
\]

Argue that \( F \) has a fixed point using the Schauder fixed point theorem.
17. Remove the assumption that $|f(t, x)| \leq M$ at the expense of obtaining only a local solution.
   
   **Hint**: You can consider the closed set in $\mathbb{R}^n \ B = \overline{B(x_0, R)}$ where $R$ is some positive number. Let $P$ be the projection onto $B$.

18. In the Schauder fixed point theorem, eliminate the assumption that $K$ is closed. **Hint**: You can argue that the $\{y_i\}$ in the approximation can be in $f(K)$. 
Chapter 8

Degree Theory, An Introduction

This chapter is on the Brouwer degree, a very useful concept with numerous and important applications. The degree can be used to prove some difficult theorems in topology such as the Brouwer fixed point theorem, the Jordan separation theorem, and the invariance of domain theorem. It also is used in bifurcation theory and many other areas in which it is an essential tool. The degree will be developed for $\mathbb{R}^n$ first. When this is understood, it is not too difficult to extend to versions of the degree which hold in Banach space. There is more on degree theory in the book by Deimling [16] and much of the presentation here follows this reference. Another more recent book which is really good is [20]. This is a whole book on degree theory.

To give you an idea what the degree is about, consider a real valued $C^1$ function defined on an interval, $I$, and let $y \in f(I)$ be such that $f'(x) \neq 0$ for all $x \in f^{-1}(y)$. In this case the degree is the sum of the signs of $f'(x)$ for $x \in f^{-1}(y)$, written as $d(f,I,y)$.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{degree_example.png}
\caption{Example of degree calculation.}
\end{figure}

In the above picture, $d(f,I,y)$ is 0 because there are two places where the sign is 1 and two where it is $-1$.

The amazing thing about this is the number you obtain in this simple manner
is a specialization of something which is defined for continuous functions and which has nothing to do with differentiability.

There are many ways to obtain the Brouwer degree. The method I will use here is due to Heinz [32] and appeared in 1959. It involves first studying the degree for functions in $C^\infty$ and establishing all its most important topological properties with the aid of an integral. Then when this is done, it is extended to general continuous functions.

When you have the topological degree, you can get all sorts of amazing theorems like the invariance of domain theorem and others. The first section contains fundamental theorems from measure theory.

8.1 Sard’s Lemma

The following lemma is a wonderful application of the Vitali covering theorem.

**Lemma 8.1.1** Let $h$ be differentiable on $U$, an open set. If $T \subseteq U$ and $m_p(T) = 0$, then $m_p(h(T)) = 0$.

**Proof:** For $k \in \mathbb{N}$

$$T_k \equiv \{ x \in T : \| Dh(x) \| < k \}$$

and let $\varepsilon > 0$ be given. Since $T_k$ is a subset of a set of measure zero, it is measurable, but we don’t need to pay much attention to this fact. Now by outer regularity, there exists an open set $\tilde{V}$, containing $T_k$ which is contained in $\tilde{U}$ such that $m_p(\tilde{V}) < \varepsilon$.

Let $x \in T_k$. Then by differentiability,

$$h(x + v) = h(x) + Dh(x)v + o(v)$$

and so there exist arbitrarily small $r_x < 1$ such that $B(x, 5r_x) \subseteq V$ and whenever $\|v\| \leq 5r_x, \|o(v)\| < \frac{1}{k} \|v\|$. Thus

$$h(B(x, 5r_x)) \subseteq Dh(x)(B(0, 5r_x)) + h(x) + B(0, r_x) \subseteq B(0, k5r_x) + B(0, r_x) + h(x) \subseteq B(h(x), (5k + 1)r_x) \subseteq B(h(x), 6kr_x)$$

From the Vitali covering theorem, there exists a countable disjoint sequence of these balls, $\{ B(x_i, r_i) \}_{i=1}^\infty$ such that $\{ B(x_i, 5r_i) \}_{i=1}^\infty = \{ \tilde{B}_i \}_{i=1}^\infty$ covers $T_k$. Then letting $\overline{m_p}$ denote the outer measure determined by $m_p$,

$$\overline{m_p}(h(T_k)) \leq \overline{m_p}(h(\bigcup_{i=1}^\infty \tilde{B}_i)) \leq \overline{m_p}(\bigcup_{i=1}^\infty h(\tilde{B}_i))$$

$$\leq \sum_{i=1}^\infty \overline{m_p}(h(\tilde{B}_i)) \leq \sum_{i=1}^\infty m_p(B(h(x_i), 6kr_{x_i}))$$

$$= \sum_{i=1}^\infty m_p(B(x_i, 6kr_{x_i})) = (6k)^p \sum_{i=1}^\infty m_p(B(x_i, r_{x_i}))$$

$$\leq (6k)^p m_p(V) \leq (6k)^p \varepsilon$$
Since $\varepsilon > 0$ is arbitrary, this shows $m_p(h(T_k)) = \overline{m}_p(h(T_k)) = 0$. Now

$$m_p(h(T)) = \lim_{k \to \infty} m_p(h(T_k)) = 0. \blacksquare$$

The following is Sard’s lemma. In the proof, it does not matter which norm you use in defining balls but it may be easiest to consider the norm

$$||x|| \equiv \max \{ |x_i|, i = 1, \cdots, p \}$$

It is a very simple proof because it uses the Vitali covering theorem to remove technical considerations.

**Lemma 8.1.2 (Sard)** Let $U$ be an open set in $\mathbb{R}^p$ and let $h: U \to \mathbb{R}^p$ be differentiable. Let

$$Z \equiv \{ x \in U : \det D_h(x) = 0 \}.$$ 

Then $m_p(h(Z)) = 0$.

**Proof:** For convenience, assume the balls in the following argument come from $||\cdot||_{\infty}$. First note that $Z$ is a Borel set because $h$ is continuous and so the component functions of the Jacobian matrix are each Borel measurable. Hence the determinant is also Borel measurable.

Suppose that $U$ is a bounded open set. Let $\varepsilon > 0$ be given. Also let $V \supseteq Z$ with $V \subseteq U$ open, and

$$m_p(Z) + \varepsilon > m_p(V).$$

Now let $x \in Z$. Then since $h$ is differentiable at $x$, there exists $\delta_x > 0$ such that if $r < \delta_x$, then $B(x, r) \subseteq V$ and also $o(v) < \eta \|v\|$ for $\|v\| < r$. Thus

$$h(x + B(0, r)) = h(B(x, r)) \subseteq h(x) + D_h(x) (B(0, r)) + B(0, r\eta), \ \eta < 1.$$

Regard $D_h(x)$ as an $n \times n$ matrix, the matrix of the linear transformation $D_h(x)$ with respect to the usual coordinates. Since $x \in Z$, it follows that there exists an invertible matrix $A$ such that $AD_h(x)$ is in row reduced echelon form with a row of zeros on the bottom. Therefore,

$$m_p(A(h(B(x, r)))) \leq m_p(AD_h(x) (B(0, r)) + AB(0, r\eta)) \tag{8.1.1}$$

The diameter of $AD_h(x) (B(0, r))$ is no larger than $\|A\| \|D_h(x)\| 2r$ and it lies in $\mathbb{R}^{p-1} \times \{0\}$. The diameter of $AB(0, r\eta)$ is no more than $\|A\| (2r\eta)$. Therefore, the measure of the right side in (8.1.1) is no more than

$$[\|A\| \|D_h(x)\| 2r + \|A\| (2\eta)) r]^{p-1} (r\eta) \leq C (\|A\|, \|D_h(x)\|) (2r)^p \eta$$

That is,

$$m_p(A(h(B(x, r)))) \leq C (\|A\|, \|D_h(x)\|) (2r)^p \eta$$
Hence from the change of variables formula for linear maps,

\[ m_p(h(B(x,r))) \leq \eta \frac{C(||A||,||Dh(x)||)}{|\det(A)|} m_p(B(x,r)) \]

Then letting \( \delta_x \) be still smaller if necessary, corresponding to sufficiently small \( \eta \),

\[ m_p(h(B(x,r))) \leq \varepsilon m_p(B(x,r)) \]

The balls of this form constitute a Vitali cover of \( Z \). Hence, by the Vitali covering theorem, Corollary A.0.8, there exists \( \{ B_i \}_{i=1}^{\infty}, B_i = B_i(x_i,r_i), \) a collection of disjoint balls, each of which is contained in \( V \), such that \( m_p(h(B_i)) \leq \varepsilon m_p(B_i) \) and \( m_p(Z \cup \cup B_i) = 0 \). Hence from Lemma 8.1.1,

\[ m_p(h(Z) \cup \cup h(B_i)) \leq m_p(h(Z \cup \cup B_i)) = 0 \]

Therefore,

\[ m_p(h(Z)) \leq \sum_i m_p(h(B_i)) \leq \varepsilon \sum_i m_p(B_i) \leq \varepsilon (m_p(V)) \leq \varepsilon (m_p(Z) + \varepsilon) . \]

Since \( \varepsilon \) is arbitrary, this shows \( m_p(h(Z)) = 0 \). What if \( U \) is not bounded? Then consider \( Z_n = Z \cap B(0,n) \). From what was just shown, \( h(Z_n) \) has measure 0 and so it follows that \( h(Z) \) also does, being the countable union of sets of measure zero.

**8.2 Preliminary Results**

In this chapter \( \Omega \) will refer to a bounded open set.

**Definition 8.2.1** For \( \Omega \) a bounded open set, denote by \( C(\overline{\Omega}) \) the set of functions which are restrictions of functions in \( C^c_c(\mathbb{R}^n) \) to \( \overline{\Omega} \) and by \( C^m(\overline{\Omega}) \), \( m \leq \infty \) the space of restrictions of functions in \( C^m_c(\mathbb{R}^n) \) to \( \overline{\Omega} \). If \( f \in C(\overline{\Omega}) \) the symbol \( f \) will also be used to denote a function defined on \( \mathbb{R}^n \) equalling \( f \) on \( \overline{\Omega} \) when convenient. This saves the trouble of having to extend to all of \( \mathbb{R}^n \) using something like Theorem 4.2.3. The subscript \( c \) indicates that the functions have compact support. The norm in \( C(\overline{\Omega}) \) is defined as follows.

\[ ||f||_{\infty} = \sup \{ |f(x)| : x \in \overline{\Omega} \} . \]

If the functions take values in \( \mathbb{R}^n \) write \( C^m(\overline{\Omega};\mathbb{R}^n) \) or \( C(\overline{\Omega};\mathbb{R}^n) \) for these functions if there is no differentiability assumed. The norm on \( C(\overline{\Omega};\mathbb{R}^n) \) is defined in the same way as above,

\[ ||f||_{\infty} = \sup \{ |f(x)| : x \in \overline{\Omega} \} . \]

Of course if \( m = \infty \), the notation means that there are infinitely many derivatives. Also, \( C(\Omega;\mathbb{R}^n) \) consists of functions which are continuous on \( \Omega \) that have values in \( \mathbb{R}^n \) and \( C^m(\Omega;\mathbb{R}^n) \) denotes the functions which have \( m \) continuous derivatives defined on \( \Omega \).
Theorem 8.2.2 Let $\Omega$ be a bounded open set in $\mathbb{R}^n$ and let $f \in C_c(\mathbb{R}^n)$. Then there exists $g \in C^\infty_c(\mathbb{R}^n)$ with $\|g - f\|_{\infty, \mathbb{R}^n} < \varepsilon$.

Proof: Form $g = f * \psi_n$ for a mollifier $\psi_n$. This will approximate $f$ uniformly on $\Omega$, and will be in $C^\infty_c(\mathbb{R}^n)$. ■

Using the Weierstrass approximation theorem, you could also get $g$ to equal a polynomial for all $x \in \Omega$.

Applying this result to the components of a vector valued function yields the following corollary.

Corollary 8.2.3 If $f \in C(\overline{\Omega}; \mathbb{R}^n)$ for $\Omega$ a bounded subset of $\mathbb{R}^n$, then for all $\varepsilon > 0$, there exists $g \in C^\infty(\overline{\Omega}; \mathbb{R}^n)$ such that $\|g - f\|_\infty < \varepsilon$.

Lemma 4.3.1 on Page 93 will also play an important role in the definition of the Brouwer degree. Earlier it made possible an easy proof of the Brouwer fixed point theorem. Later in this chapter, it is used to show the definition of the degree is well defined. For convenience, here it is stated again.

Lemma 8.2.4 Let $g : U \rightarrow \mathbb{R}^n$ be $C^2$ where $U$ is an open subset of $\mathbb{R}^n$. Then

$$\sum_{j=1}^{n} \text{cof}((Dg)_{ij,j}) = 0,$$

where here $(Dg)_{ij} \equiv g_{i,j} \equiv \frac{\partial g_i}{\partial x_j}$. Also, $\text{cof}((Dg)_{ij}) = \frac{\partial \text{det}(Dg)}{\partial g_{ij,j}}$.

Another simple result which will be used whenever convenient is the following lemma.

Lemma 8.2.5 Let $K$ be a compact set and $C$ a closed set in a complete normed vector space such that $K \cap C = \emptyset$. Then

$$\text{dist}(K, C) > 0.$$ 

Proof: Let

$$d \equiv \inf \{||k - c|| : k \in K, c \in C\}$$

Let $\{k_n\}, \{c_n\}$ be such that

$$d + \frac{1}{n} > ||k_n - c_n||.$$ 

Since $K$ is compact, there is a subsequence still denoted by $\{k_n\}$ such that $k_n \rightarrow k \in K$. Then also

$$||c_n - c_m|| \leq ||c_n - k_n|| + ||k_n - k_m|| + ||c_m - k_m||.$$
If \( d = 0 \), then as \( m, n \to \infty \) it follows \( ||c_n - c_m|| \to 0 \) and so \( \{c_n\} \) is a Cauchy sequence which must converge to some \( c \in C \). But then \( ||c - k|| = \lim_{n \to \infty} ||c_n - k_n|| = 0 \) and so \( c = k \in C \cap K \), a contradiction to these sets being disjoint.

In particular the distance between a point and a closed set is always positive if the point is not in the closed set. Of course this is obvious even without the above lemma. The above lemmas will be used now to prove technical lemmas which are the basis for everything.

**Definition 8.2.6** Let \( g \in C^\infty (\overline{\Omega}; \mathbb{R}^n) \) where \( \Omega \) is a bounded open set. Also let \( \phi \varepsilon \) be a mollifier.

\[ \phi \varepsilon \in C^\infty_c (B(0, \varepsilon)), \phi \varepsilon \geq 0, \int \phi \varepsilon dx = 1. \]

First, here is a technical lemma which is the reason it all works out. It is a result on homotopy invariance for functions which are \( C^\infty \).

**Lemma 8.2.7** If \( h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \) is in \( C^\infty_c (\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n) \), and \( 0 \notin h (\partial \Omega \times [\alpha, \beta]) \) then for \( 0 < \varepsilon < \text{dist} (0, h (\partial \Omega \times [\alpha, \beta])) \),

\[ t \to \int_\Omega \phi \varepsilon (h (x, t)) \det D_1 h (x, t) dx \]

is constant for \( t \in (a, b) \), an open set which contains \([\alpha, \beta] \).

**Proof:** By continuity, we can get such an open interval, \((a, b)\) such that it contains \([\alpha, \beta]\) and \(0 \notin h (\partial \Omega \times [a, b])\). Let \( \varepsilon > 0 \) be such that for all \( t \in [a, b] \),

\[ B(0, \varepsilon) \cap h (\partial \Omega \times [a, b]) = \emptyset \quad (8.2.2) \]

Define for \( t \in (a, b) \),

\[ H(t) \equiv \int_\Omega \phi \varepsilon (h (x, t)) \det D_1 h (x, t) dx \]

Then if \( t \in (a, b) \),

\[ H'(t) = \int_\Omega \sum_\alpha \phi \varepsilon, \alpha (h (x, t)) h_{\alpha, t} (x, t) \det D_1 h (x, t) dx \]

\[ + \int_\Omega \phi \varepsilon (h (x, t)) \sum_\alpha j \det D_1 (h (x, t)), \alpha j h_{\alpha, j} dx \]

\[ \equiv A + B. \]

In this formula, the function \( \det \) is considered as a function of the \( n^2 \) entries in the \( n \times n \) matrix and the \( , \alpha j \) represents the derivative with respect to the \( \alpha j \)th entry \( h_{\alpha, j} \). Now as in the proof of Lemma 4.3.1 on Page 93,

\[ \det D_1 (h (x, t)), \alpha j = (\text{cof} D_1 (h (x, t))), \alpha j \]
and so
\[ B = \int_{\Omega} \sum_{\alpha} \sum_{j} \phi_{\varepsilon}(h(x,t)) (\text{cof } D_1(h(x,t)))_{\alpha j} h_{\alpha,j} \, dx. \]

By hypothesis
\[ x \to \phi_{\varepsilon}(h(x,t)) (\text{cof } D_1(h(x,t)))_{\alpha j} \]
is in \( C^\infty_c(\Omega) \) because if \( x \in \partial \Omega \), it follows that for all \( t \in [a,b] \)
\[ h(x,t) \notin B(0,\varepsilon) \]
and so \( \phi_{\varepsilon}(h(x,t)) = 0 \). Thus it equals 0 on \( \partial \Omega \).

Therefore, integrate by parts and write
\[ B = -\int_{\Omega} \sum_{\alpha} \sum_{j} \frac{\partial}{\partial x_j} (\phi_{\varepsilon}(h(x,t))) (\text{cof } D_1(h(x,t)))_{\alpha j} h_{\alpha,t} \, dx + \]
\[ -\int_{\Omega} \sum_{\alpha} \sum_{j} \phi_{\varepsilon}(h(x,t)) (\text{cof } D(h(x,t)))_{\alpha j,j} h_{\alpha,t} \, dx \]
The second term equals zero by Lemma 8.2.4. Simplifying the first term yields
\[ B = -\int_{\Omega} \sum_{\alpha} \sum_{j} \sum_{\beta} \phi_{\varepsilon,\beta}(h(x,t)) h_{\beta,j} h_{\alpha,t} (\text{cof } D_1(h(x,t)))_{\alpha j} \, dx \]
\[ = -\int_{\Omega} \sum_{\alpha} \sum_{j} \sum_{\beta} \phi_{\varepsilon,\beta}(h(x,t)) h_{\alpha,t} h_{\beta,j} (\text{cof } D_1(h(x,t)))_{\alpha j} \, dx \]
Now the sum on \( j \) is the dot product of the \( \beta^{th} \) row with the \( \alpha^{th} \) row of the cofactor matrix which equals zero unless \( \beta = \alpha \) because it would be a cofactor expansion of a matrix with two equal rows. When \( \beta = \alpha \), the sum on \( j \) reduces to \( \det (D_1(h(x,t))) \). Thus \( B \) reduces to
\[ -\int_{\Omega} \sum_{\alpha} \phi_{\varepsilon,\alpha}(h(x,t)) h_{\alpha,j} \sum_{j} h_{\alpha,j}(x,t) (\text{cof } D_1(h(x,t)))_{\alpha j} \, dx \]
\[ = -\int_{\Omega} \sum_{\alpha} \phi_{\varepsilon,\alpha}(h(x,t)) h_{\alpha,t} \det (D_1(h(x,t))) \, dx \]
Now \( A \) equals
\[ \int_{\Omega} \sum_{\alpha} \phi_{\varepsilon,\alpha}(h(x,t)) h_{\alpha,t}(x,t) \det D_1 h(x,t) \, dx \]
which is the same thing with opposite sign. Hence these sum to 0. Therefore, \( H'(t) = 0 \) and so \( H \) is a constant on \( (a,b) \supseteq [\alpha,\beta] \).

The following is a situation in which one only has continuity in \( t \). Of course the difficulty is that it is not yet clear whether the constant depends on \( \varepsilon \).
Corollary 8.2.8 If \( h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \) is in \( C_c(\mathbb{R}^n \times \mathbb{R}; \mathbb{R}^n) \), and for each \( t \in [0,1] \), \( h(\cdot, t) \in C_c^\infty(\mathbb{R}, \mathbb{R}^n) \) and \( 0 \not\in h(\partial\Omega \times [0,1]) \) then for \( \varepsilon \) small enough,

\[
t \to \int_{\Omega} \phi_\varepsilon(h(x,t)) \det D_1 h(x,t) \, dx
\]

is constant for \( t \in [0,1] \).

Proof: Let \( 0 < 6\delta < \text{dist}(0, h(\partial\Omega \times [0,1])) \). Now let \( 0 = t_0 < t_1 < \cdots < t_m = 1 \). Also let \( \hat{h}(x, t_k) = h(x, t_k) \) and for \( t \in (t_{k-1}, t_k) \),

\[
\hat{h}(x, t) \equiv h(x, t_{k-1}) + \frac{t - t_{k-1}}{t_k - t_{k-1}} (h(x, t_k) - h(x, t_{k-1}))
\]

Thus for fixed \( x \), this gives a piecewise linear function approximating \( t \to h(x, t) \). Let these \( t_k \) be close enough together that on each \( [t_{k-1}, t_k] \),

\[
\max_{t \in [t_{k-1}, t_k]} \left\| \hat{h}(\cdot, t) - h(\cdot, t) \right\|_\infty < \delta
\]

It follows that on \( [t_{k-1}, t_k] \), for \( x \in \partial\Omega \),

\[
\min_{t \in [t_{k-1}, t_k]} \left| \hat{h}(x, t) - 0 \right| > \delta
\]

This function \((x, t) \to \hat{h}(x, t)\) can be considered the restriction of a function in \( C_c^\infty(\mathbb{R}^n \times \mathbb{R}; \mathbb{R}^n) \) and \( 0 \not\in \hat{h}(\partial\Omega, t) \) for all \( t \in [t_{k-1}, t_k] \). By Lemma 8.2.7, for all \( \varepsilon \) small enough, say \( \varepsilon < \delta \),

\[
\int_{\Omega} \phi_\varepsilon(\hat{h}(x,t)) \det (D_1 \hat{h}(x,t)) \, dx
\]

is constant on this interval \( [t_{k-1}, t_k] \).

This has shown that for \( k, j \in \{0, 2, \cdots, m\} \),

\[
\int_{\Omega} \phi_\varepsilon(h(x,t_k)) \det (D_1 h(x,t_k)) \, dx = \int_{\Omega} \phi_\varepsilon(h(x,t_j)) \det (D_1 h(x,t_j)) \, dx
\]

You could include \( 1/2 \) in the partition. Then this would show that whatever the partition including this point,

\[
\int_{\Omega} \phi_\varepsilon(h(x,t_k)) \det (D_1 h(x,t_k)) \, dx = \int_{\Omega} \phi_\varepsilon(h(x,1/2)) \det (D_1 h(x,1/2)) \, dx
\]

Thus this integral is constant for \( t \in [0,1] \) as claimed. \( \blacksquare \)

This corollary is applied later to the situation where \( h(x,t) \) is of the form \( h(x, t) - y(t) \) where \( t \to y(t) \) is continuous and \( y(t) \not\in h(\partial\Omega, t) \).
8.3 Definitions And Elementary Properties

First is what is meant by two functions being homotopic.

Definition 8.3.1 \( U_y \equiv \{ f \in C(\overline{\Omega}; \mathbb{R}^n) : y \notin f(\partial \Omega) \} \). (Recall that \( \partial \Omega = \overline{\Omega} \setminus \Omega \))

For two functions, \( f, g \in U_y \),

\( f \sim g \) if there exists a continuous function,

\[ h : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n \]

such that \( h(x,1) = g(x) \) and \( h(x,0) = f(x) \) for \( x \in \overline{\Omega} \) and \( x \to h(x,t) \in U_y \) for all \( t \in [0,1] \) \( (y \notin h(\partial \Omega, t)) \). This function \( h \), is called a homotopy and \( f \) and \( g \) are homotopic.

Definition 8.3.2 For \( W \) an open set in \( \mathbb{R}^n \) and \( g \in C^1(W; \mathbb{R}^n) \) \( y \) is called a regular value of \( g \) if whenever \( x \in g^{-1}(y) \), \( \det(Dg(x)) \neq 0 \). Note that if \( g^{-1}(y) = \emptyset \), it follows that \( y \) is a regular value from this definition. Denote by \( S_y \) the set of singular values of \( g \), those \( y \) such that \( \det(Dg(x)) = 0 \) for some \( x \in g^{-1}(y) \).

Lemma 8.3.3 The relation \( \sim \) is an equivalence relation and, denoting by \( [f] \) the equivalence class determined by \( f \), it follows that \( [f] \) is an open subset of

\[ U_y \equiv \{ f \in C(\overline{\Omega}; \mathbb{R}^n) : y \notin f(\partial \Omega) \} . \]

Furthermore, \( U_y \) is an open set in \( C(\overline{\Omega}; \mathbb{R}^n) \). Let \( g \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \) and \( y \notin g(\partial \Omega) \) so \( \text{dist}(y, g(\partial \Omega)) > 5\delta \) for some positive \( \delta \). Then there exists \( g_1 \) such that \( \| g - g_1 \|_\infty < \delta \), \( y \) is a regular value of \( g_1 \), and

\[ \int_\Omega \phi_e ((1-t) g(x) + tg_1(x) - y) \det(D((1-t) g + tg_1))(x) \, dx \]

is constant for \( t \in [0,1] \).

**Proof:** In showing that \( \sim \) is an equivalence relation, it is easy to verify that \( f \sim f \) and that if \( f \sim g \), then \( g \sim f \). To verify the transitive property for an equivalence relation, suppose \( f \sim g \) and \( g \sim k \), with the homotopy for \( f \) and \( g \), the function, \( h_1 \) and the homotopy for \( g \) and \( k \), the function \( h_2 \). Thus \( h_1(x,0) = f(x) \), \( h_1(x,1) = g(x) \) and \( h_2(x,0) = g(x) \), \( h_2(x,1) = k(x) \). Then define a homotopy of \( f \) and \( k \) as follows.

\[
h(x,t) = \begin{cases} h_1(x,2t) \text{ if } t \in [0,\frac{1}{2}] \\ h_2(x,2t-1) \text{ if } t \in [\frac{1}{2},1] \\ \end{cases} .
\]

It is obvious that \( U_y \) is an open subset of \( C(\overline{\Omega}; \mathbb{R}^n) \). If \( g \in U_y \), then \( y \notin g(\partial \Omega) \) a compact set. Hence if \( f \) is close enough to \( g \), the same is true of \( f \).
Next consider the claim that \( \mathbf{f} \) is also an open set. If \( \mathbf{f} \in \mathcal{U}_y \), there exists \( \delta > 0 \) such that \( B(y, 2\delta) \cap \mathbf{f} (\partial\Omega) = \emptyset \). Let \( \mathbf{f}_1 \in C(\overline{\Omega}; \mathbb{R}^n) \) with \( ||\mathbf{f}_1 - \mathbf{f}||_\infty < \delta \). Then if \( t \in [0,1] \), and \( x \in \partial\Omega \)

\[
|f(x) + t(f_1(x) - f(x)) - y| \geq |f(x) - y| - t||f_1 - f||_\infty > 2\delta - t\delta > 0.
\]

Therefore, \( B(f, \delta) \subseteq f \) because if \( f_1 \in B(f, \delta) \), this shows that, letting \( h(x, t) \equiv f(x) + t(f_1(x) - f(x)) \), \( f_1 \sim f \).

Consider the last claim. There exists \( y_1 \in B(y, \delta) \) such that \( y_1 \) is a regular value of \( g \) thanks to Sard’s theorem. Let \( g_1(x) \equiv g(x) + y - y_1 \) for all \( x \in \overline{\Omega} \). Then \( ||g - g_1||_\infty < \delta \) and so it follows that for \( t \in [0,1] \), and \( x \in \partial\Omega \),

\[
|(1-t)g(x) + tg_1(x)) - y| = |g(x) - y + t(y - y_1)| \geq 5\delta - \delta > 0
\]

Therefore, from Lemma 8.2.3, the integral is constant. Also \( y \) is a regular value of \( g_1 \) because \( g_1(x) = y \) if and only if \( g(x) - y_1 \) and \( y_1 \) is a regular value for \( g \).

### 8.3.1 The Degree For \( C^\infty (\overline{\Omega}; \mathbb{R}^n) \)

Here I will give a definition of the degree which works for all functions in \( C^\infty (\overline{\Omega}; \mathbb{R}^n) \). These are the restrictions to \( \overline{\Omega} \) of functions in \( C^\infty_c (\mathbb{R}^n; \mathbb{R}^n) \).

**Definition 8.3.4** For \( g \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \), \( y \notin g(\partial\Omega) \)

\[
d(g,\Omega,y) = \lim_{\varepsilon \to 0} \int_\Omega \phi_\varepsilon (g(x) - y) \det Dg(x) \, dx
\]

The next lemma has to do with the existence of the limit in the definition of the degree and the fact that it is always an integer.

**Lemma 8.3.5** The above definition is well defined. In particular the limit exists. In fact

\[
\int_\Omega \phi_\varepsilon (g(x) - y) \det Dg(x) \, dx
\]

does not depend on \( \varepsilon \) whenever \( \varepsilon \) is small enough. If \( y \) is a regular value for \( g \) then for all \( \varepsilon \) small enough,

\[
\int_\Omega \phi_\varepsilon (g(x) - y) \det Dg(x) \, dx = \sum \{ \text{sgn} (\det Dg(x)) : x \in g^{-1}(y) \}
\]

(8.3.3)

If \( g, f \in \mathcal{U}_y \subseteq C^\infty (\overline{\Omega}; \mathbb{R}^n) \), and

\[
y \notin (tf + (1-t)g) (\partial\Omega), \quad \text{all } t \in [0,1]
\]

(8.3.4)
then
\[ d(f, \Omega, y) = d(g, \Omega, y) \]

If \( \text{dist} (y, g(\partial \Omega)) > 5\delta \) and \( y_1 \in B(y, \delta) \), then \( d(g, \Omega, y) = d(g, \Omega, y_1) \). Also, the appropriate integrals are equal. See the following picture for an illustration of this last claim. The degree has integer values.

**Proof:** If \( y \) is not a value of \( g \) then there is not much to show. For small enough \( \varepsilon \), you will get 0 in the integral.

The case where \( y \) is a regular value

First consider the case where \( y \) is a regular value of \( g \). I will show that in this case, the integral expression is eventually constant for small \( \varepsilon > 0 \) and equals the right side of 8.3.3. I claim the right side of this equation is actually a finite sum. This follows from the inverse function theorem because \( g^{-1}(y) \) is a closed, hence compact subset of \( \Omega \) due to the assumption that \( y \notin g(\partial \Omega) \). If \( g^{-1}(y) \) had infinitely many points in it, there would exist a sequence of distinct points \( \{x_k\} \subseteq g^{-1}(y) \). Since \( \Omega \) is bounded, some subsequence \( \{x_{k_l}\} \) would converge to a limit point \( x_{\infty} \). By continuity of \( g \), it follows \( x_{\infty} \in g^{-1}(y) \) also and so \( x_{\infty} \in \Omega \). Therefore, since \( y \) is a regular value, there is an open set, \( U_{x_{\infty}} \), containing \( x_{\infty} \) such that \( g \) is one to one on this open set contradicting the assertion that \( \lim_{l \to \infty} x_{k_l} = x_{\infty} \). Therefore, this set is finite and so the sum is well defined.

Thus the right side of 8.3.3 is finite when \( y \) is a regular value. Next I need to show the left side of this equation is eventually constant and equals the right side. By what was just shown, there are finitely many points, \( \{x_i\}_{i=1}^m = g^{-1}(y) \). By the inverse function theorem, there exist disjoint open sets \( U_i \) with \( x_i \in U_i \), such that \( g \) is one to one on \( U_i \) with \( \det (Dg(x)) \) having constant sign on \( U_i \) and \( g(U_i) \) is an open set containing \( y \). Then let \( \varepsilon \) be small enough that \( B(y, \varepsilon) \subseteq \bigcap_{i=1}^m g(U_i) \) and let \( V_i = g^{-1}(B(y, \varepsilon)) \cap U_i \).
Therefore, for any $\varepsilon$ this small,

$$\int_{\Omega} \phi_{\varepsilon}(g(x) - y) \det Dg(x) \, dx = \sum_{i=1}^{m} \int_{V_i} \phi_{\varepsilon}(g(x) - y) \det Dg(x) \, dx$$

The reason for this is as follows. The integrand on the left is nonzero only if $g(x) - y \in B(0, \varepsilon)$ which occurs only if $g(x) \in B(y, \varepsilon)$ which is the same as $x \in g^{-1}(B(y, \varepsilon))$. Therefore, the integrand is nonzero only if $x$ is contained in exactly one of the disjoint sets, $V_i$. Now using the change of variables theorem,

$$(z = g(x) - y, g^{-1}(y + z) = x)$$

$$= \sum_{i=1}^{m} \int_{g(V_i) - y} \phi_{\varepsilon}(z) \det Dg(g^{-1}(y + z)) \left| \det Dg^{-1}(y + z) \right| dz$$

By the chain rule, $I = Dg(g^{-1}(y + z)) Dg^{-1}(y + z)$ and so

$$\det Dg(g^{-1}(y + z)) \left| \det Dg^{-1}(y + z) \right|$$

$$= \text{sgn} \left( \det Dg(g^{-1}(y + z)) \right) \left| \det Dg(g^{-1}(y + z)) \right| \left| \det Dg^{-1}(y + z) \right|$$

$$= \text{sgn} \left( \det Dg(x) \right) = \text{sgn} \left( \det Dg(x_i) \right).$$

Therefore, this reduces to

$$\sum_{i=1}^{m} \text{sgn} \left( \det Dg(x_i) \right) \int_{g(V_i) - y} \phi_{\varepsilon}(z) \, dz = \sum_{i=1}^{m} \text{sgn} \left( \det Dg(x_i) \right) \int_{B(0, \varepsilon)} \phi_{\varepsilon}(z) \, dz = \sum_{i=1}^{m} \text{sgn} \left( \det Dg(x_i) \right).$$

In case $g^{-1}(y) = \emptyset$, there exists $\varepsilon > 0$ such that $g(\Omega) \cap B(y, \varepsilon) = \emptyset$ and so for $\varepsilon$ this small,

$$\int_{\Omega} \phi_{\varepsilon}(g(x) - y) \det Dg(x) \, dx = 0.$$
8.3. DEFINITIONS AND ELEMENTARY PROPERTIES

Showing the integral is constant for small $\varepsilon$

Let $y \notin g(\partial \Omega)$. By Lemma 8.3.3 $y$ is a regular value of $g_1$ where

$$\int_{\Omega} \phi_\varepsilon (g(x) + t(g_1(x) - g(x)) - y) \det (D(g + t(g_1 - g))(x)) \, dx$$

is constant for $t \in [0, 1]$. Letting $\varepsilon$ be small enough, and $t = 0$ and then 1,

$$\int_{\Omega} \phi_\varepsilon (g(x) - y) \det (Dg(x)) \, dx = \int_{\Omega} \phi_\varepsilon (g_1(x) - y) \det (Dg_1(x)) \, dx = d(g_1, \Omega, y)$$

Therefore, $\int_{\Omega} \phi_\varepsilon (g(x) - y) \det (Dg(x)) \, dx$ does not change for sufficiently small $\varepsilon$. Thus the limit exists and equals an integer. This shows the degree is an integer.

The next claim follows right away from the above. Suppose $0 \notin (tf + (1 - t)g)(\partial \Omega) - y$ for all $t \in [0, 1]$. Then choosing $\varepsilon$ small enough, it follows $d(f, \Omega, y) = d(g, \Omega, y)$ because the two integrals defining the degree for small $\varepsilon$ are equal, this by Lemma 8.3.3.

It also follows from the above argument and Lemma 8.3.3 that if dist $(y, g(\partial \Omega)) > 5\delta$ and if $y_1 \in B(y, \delta)$, then $d(g, \Omega, y_1) = d(g, \Omega, y)$ because the corresponding integrals are equal for all $\varepsilon$ small enough. $lacksquare$

8.3.2 Definition Of The Degree For Continuous Functions

With the above results, it is now possible to extend the definition of the degree to continuous functions which have no differentiability. It is desired to preserve the homotopy invariance. This requires the following definition.

**Definition 8.3.6** Let $y \in \mathbb{R}^n \setminus f(\partial \Omega)$ where $f \in C(\Omega; \mathbb{R}^n)$. Then

$$d(f, \Omega, y) \equiv d(g, \Omega, y)$$

where $y \notin g(\partial \Omega)$, $g \in C^\infty(\Omega; \mathbb{R}^n)$ and $f \sim g$.

**Theorem 8.3.7** The definition of the degree given in Definition 8.3.6 is well defined, equals an integer, and satisfies the following properties. In what follows, $I(x) = x$.

1. $d(I, \Omega, y) = 1$ if $y \in \Omega$.
2. If $\Omega_i \subseteq \Omega, \Omega_i$ open, and $\Omega_1 \cap \Omega_2 = \emptyset$ and if $y \notin f(\Omega \setminus (\Omega_1 \cup \Omega_2))$, then $d(f, \Omega_1, y) + d(f, \Omega_2, y) = d(f, \Omega, y)$.
3. For $y \in \mathbb{R}^n \setminus f(\partial \Omega)$, if $d(f, \Omega, y) \neq 0$ then $f^{-1}(y) \cap \Omega \neq \emptyset$.
4. If $t \rightarrow y(t)$ is continuous $h : \Omega \times [0, 1] \rightarrow \mathbb{R}^n$ is continuous and if $y(t) \notin h(\partial \Omega, t)$ for all $t$, then $t \rightarrow d(h(\cdot, t), \Omega, y(t))$ is constant.
5. \( d(f, \Omega, \cdot) \) is continuous and constant on every connected component of \( \mathbb{R}^n \setminus f(\partial \Omega) \).

6. \( d(g, \Omega, y) = d(f, \Omega, y) \) if \( g|_{\partial \Omega} = f|_{\partial \Omega} \).

**Proof:** First it is necessary to show the definition is well defined. There are two parts to this. First I need to show there exists \( g \) with the desired properties and then I need to show that it doesn’t matter which \( g \) I happen to pick. The first part is easy. Let \( \delta \) be small enough that

\[
B(y, \delta) \cap f(\partial \Omega) = \emptyset.
\]

Then by Lemma 8.3.3 there exists \( g \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \) such that \( ||g - f||_\infty < \delta \). It follows that for \( t \in [0,1] \),

\[
y \notin (tg + (1 - t)f)(\partial \Omega)
\]

and so \( g \sim f \). The reason is that if \( x \in \partial \Omega \),

\[
|tg(x) + (1 - t)f(x) - y| \geq |f(x) - y| - t|g(x) - f(x)| > \delta - \delta = 0
\]

Now consider the second part. Suppose \( g \sim f \) and \( g_1 \sim f \). Then by Lemma 8.3.3 again

\[g \sim g_1\]

Thus there is a function \( h : \overline{\Omega} \times [0,1] \rightarrow \mathbb{R}^n \) such that \( h(x, 0) = g(x) \) and \( h(x, 1) = g_1(x) \). The difficulty is that it is only known that this function is continuous. It is not known that \( h(\cdot, t) \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \). Let \( \psi_\varepsilon \) be a mollifier. Thus it is infinitely differentiable, has support in \( B(0, \varepsilon) \) and \( \int_{\mathbb{R}^n} \psi_\varepsilon(x) \, dx = 1 \). Then define

\[
h_\varepsilon(x, t) \equiv h(\cdot, t) * \psi_\varepsilon(x) \equiv \int_{\mathbb{R}^n} h(x - u, t) \psi_\varepsilon(u) \, du.
\]

Then as \( \varepsilon \to 0 \), the convergence is uniform on \( \overline{\Omega} \times [0,1] \). Now there exists \( \delta > 0 \) such that \( B(y, 5\delta) \cap h(\partial \Omega \times [0,1]) = \emptyset \) for all \( t \). Hence, by uniform convergence, for \( \varepsilon \) small enough, \( B(y, 6\delta) \cap h_\varepsilon(\partial \Omega, t) = \emptyset \) for all \( t \) and in fact, \( \max_{t \in [0,1]} \|h_\varepsilon(\cdot, t) - h(\cdot, t)\|_\infty < \delta \). Then by Lemma 8.2.7, it follows that the following integral is constant for \( t \in [0,1] \).

\[
\int_\Omega \phi_\varepsilon(h_\varepsilon(x, t) - y) \det(D_1 h_\varepsilon(\cdot, t)(x)) \, dx
\]

Thus, letting \( \varepsilon \) be smaller if necessary, \( d(h_\varepsilon(\cdot, 0), \Omega, y) = d(h_\varepsilon(\cdot, 1), \Omega, y) \). Since \( \|h_\varepsilon(\cdot, 1) - g_1\|_\infty < \delta \), it follows that \( y \notin t h_\varepsilon(x, 1) + (1 - t) g_1(x) \) for all \( x \in \partial \Omega \). Then by Lemma 8.3.3, formula 8.3.4,

\[d(h_\varepsilon(\cdot, 1), \Omega, y) = d(g_1, \Omega, y)\]

similarly,

\[d(h_\varepsilon(\cdot, 0), \Omega, y) = d(g, \Omega, y)\]
Therefore,
\[ d(g,\Omega, y) = d(h_\varepsilon (\cdot,0) \Omega, y) = d(h_\varepsilon (\cdot,1) \Omega, y) = d(g_1,\Omega, y) \]
which implies \( d(g_1,\Omega, y) = d(g,\Omega, y) \). Thus the definition is well defined.

Now consider the properties. The first, \( I \) is obvious since \( y \) is a regular value of \( I \).

Consider \( 2 \) about \( y / \in f(\Omega \setminus (\Omega_1 \cup \Omega_2)) \).

The assumption implies
\[ y / \in f(\partial \Omega) \cup f(\partial \Omega_1) \cup f(\partial \Omega_2) \]
Recall that \( y / \in f(\Omega \setminus (\Omega_1 \cup \Omega_2)) \). Say
\[ \text{dist} (y, f(\Omega_1 \cup \Omega_2)) > 5\delta, \delta > 0 \]
Then let \( g \in C^\infty (\Omega; \mathbb{R}^n) \) such that \( ||f - g||_\infty < \delta \). It follows that \( f \sim g \) and so by definition,
\[ d(f,\Omega, y) = d(g,\Omega, y), \ d(f,\Omega_1, y) = d(g,\Omega_1, y), \ d(f,\Omega_2, y) = d(g,\Omega_2, y) \]
since \( f(\Omega \setminus (\Omega_1 \cup \Omega_2)) \) includes \( f(\partial \Omega) \cup f(\partial \Omega_1) \cup f(\partial \Omega_2) \). One can use for a homotopy \( tf + (1 - t) g \) in every case.

Hence
\[ d(f,\Omega, y) = \int_{\Omega} \phi_\varepsilon (g(x) - y) Dg(x) \, dx \]
for all \( \varepsilon \) sufficiently small. However, \( \phi_\varepsilon (g(x) - y) = 0 \) if \( x / \in \Omega_1 \cup \Omega_2 \) whenever \( \varepsilon \) is sufficiently small, \( (\varepsilon < \delta \text{ will do}) \). Hence the above integral is
\[ \int_{\Omega_1} \phi_\varepsilon (g(x) - y) Dg(x) \, dx + \int_{\Omega_2} \phi_\varepsilon (g(x) - y) Dg(x) \, dx \]
for all \( \varepsilon \) small enough, and hence
\[ d(f,\Omega, y) = d(g,\Omega, y) = d(g,\Omega_1, y) + d(g,\Omega_2, y) \]
\[ = d(f,\Omega_1, y) + d(f,\Omega_1, y) \].

Property \( II \) is very important because it can be used to deduce the existence of solutions to a nonlinear equation. Suppose \( f^{-1}(y) \cap \Omega = \emptyset \). I will show this
requires \( d(f, \Omega, y) = 0 \). It is assumed \( y \notin f(\partial \Omega) \) and so if \( f^{-1}(y) \cap \Omega = \emptyset \), then 
\( y \notin f(\overline{\Omega}) \). Choosing \( g \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \) such that \( ||f - g||_\infty \) is sufficiently small, it can be assumed 
\( y \notin g(\overline{\Omega}), \ y \notin ((1 - t)f + t g)(\partial \Omega) \) for all \( t \in [0, 1] \).

Then it follows from the definition of the degree 
\[
d(f, \Omega, y) = d(g, \Omega, y) \equiv \lim_{\varepsilon \to 0} \int_{\Omega} \phi_\varepsilon (g(x) - y) Dg(x) \, dx = 0
\]
because eventually \( \varepsilon \) is smaller than the distance from \( y \) to \( g(\overline{\Omega}) \) and so 
\[
\phi_\varepsilon (g(x) - y) = 0
\]
for all \( x \in \Omega \).

Consider \( \mathcal{H} \). There is a \( \delta > 0 \) such that \( B(y(t), 6\delta) \cap h(\partial \Omega, t) \) for all \( t \). If this were not so, there would be a sequence \( \{t_k\} \subseteq [0,1]; \{x_k\} \subseteq \partial \Omega \) such that \( \text{dist}(y(t_k), h(x_k, t_k)) \leq 1/k \). Now taking a convergent subsequence \( t_k \to t \in [0,1], \ x_k \to x \in \partial \Omega \), it follows from continuity that \( y(t) = h(x, t) \) which is a contradiction. Let \( \psi_\varepsilon \) be a mollifier and let \( h_\varepsilon (x, t) \equiv h(\cdot, t) \ast \psi_\varepsilon (x) \). Then by the uniform convergence, whenever \( \varepsilon \) is sufficiently small, \( B(y(t), 5\delta) \cap h_\varepsilon (\partial \Omega, t) = \emptyset \) because for all \( t \), 
\[
\|h(\cdot, t) - h_\varepsilon (\cdot, t)\|_\infty < \delta.
\]
Therefore, \( h(\cdot, t) \sim h_\varepsilon (\cdot, t) \) for all \( t \) since \( y(t) \notin (1 - \lambda) h(\partial \Omega, t) + \lambda h_\varepsilon (\partial \Omega, t), \lambda \in [0, 1] \). To see this, let \( x \in \partial \Omega \)
\[
| (1 - \lambda) h(x, t) + \lambda h_\varepsilon (x, t) - y(t) |
\geq
| h(x, t) - y(t) | - \lambda | h(x, t) - h_\varepsilon (x, t) |
\geq
6\delta - \lambda\delta \geq 5\delta > 0
\]
Then from the definition of the degree above, which was shown above to be well defined, it follows that for all \( t \), 
\[
d(h(\cdot, t), \Omega, y(t)) = d(h_\varepsilon (\cdot, t), \Omega, y(t))
\]
and the expression on the right is constant in \( t \) thanks to Corollary 8.2.8.

Property \( \mathcal{H} \) about being constant on connected components is done by showing 
\( y \to d(f, \Omega, y) \) is continuous. Then, since it is integer valued, it must be constant on every connected component of \( f(\partial \Omega)^C \). Suppose \( \text{dist}(y, f(\partial \Omega)) = 5\delta > 0 \) and suppose \( \hat{y} \in B(y, \delta) \). Let \( y(t) \equiv ty + (1 - t)\hat{y}, t \in [0, 1] \). Then \( y(t) \notin f(\partial \Omega) \) and so by \( \mathcal{H} \)
\[
d(f, \Omega, y(t))
\]
is a constant. Letting \( t = 0 \) and then letting \( t = 1 \), it follows that 
\[
d(f, \Omega, y) = d(f, \Omega, \hat{y})
\]
showing that in fact \( y \to d(f, \Omega, y) \) is locally constant. Hence it is
8.4. BORSUK’S THEOREM

continuous and so it is also constant on every connected component of \( f(\partial \Omega)^C \) by Corollary 1.11.11.

Consider property 6 about the degree in which \( f = g \) on \( \partial \Omega \). This one is easy because for
\[
 y \in \mathbb{R}^n \setminus f(\partial \Omega) = \mathbb{R}^n \setminus g(\partial \Omega),
\]
and \( x \in \partial \Omega, \)
\[
t (f(x) + (1 - t)g(x)) - y = f(x) - y \neq 0
\]
for all \( t \in [0, 1] \) and so by 4, \( d(f, \Omega, y) = d(g, \Omega, y) \).

From the above, there is an easy corollary which gives related properties of the degree.

Corollary 8.3.8 The following additional properties of the degree are also valid.

1. If \( y \notin f(\Omega \setminus \Omega_1) \) and \( \Omega_1 \) is an open subset of \( \Omega \), then \( d(f, \Omega, y) = d(f, \Omega_1, y) \).

2. \( d(\cdot, \Omega, y) \) is defined and constant on
\[
\{ g \in C(\Omega; \mathbb{R}^n) : ||g - f||_\infty < r \}
\]
where \( r = \text{dist}(y, f(\partial \Omega)) \).

3. If \( \text{dist}(y, f(\partial \Omega)) \geq \delta \) and \( |z - y| < \delta \), then \( d(f, \Omega, y) = d(f, \Omega, z) \).

Proof: Consider 6. This really follows from 4 of previous theorem. You can take \( \Omega_2 = \emptyset \). I leave the details to you. To be more careful, you can modify the proof of 4 of the previous theorem slightly. Consider 6. To verify, let \( h(x, t) = tg(x)+ (1 - t)f(x) \). Then note that \( y \notin h(\partial \Omega, t) \) and use Property 6 of the previous theorem. Finally, consider 6. Let \( y(t) \equiv (1 - t)y + tz \). Then for \( x \in \partial \Omega \)
\[
|(1 - t)y + tz - f(x)| = |y - f(x) + t(z - y)|
\]
\[
\geq \delta - t|z - y| > \delta - \delta = 0
\]
Then by 6 of the previous theorem, \( d(f, \Omega, (1 - t)y + tz) \) is constant. When \( t = 0 \) you get \( d(f, \Omega, y) \) and when \( t = 1 \) you get \( d(f, \Omega, z) \).  

8.4 Borsuk’s Theorem

In this section is an important theorem which can be used to verify that \( d(f, \Omega, y) \neq 0 \). This is significant because when this is known, it follows from Theorem 8.3.7 that \( f^{-1}(y) \neq \emptyset \). In other words there exists \( x \in \Omega \) such that \( f(x) = y \).

Definition 8.4.1 A bounded open set, \( \Omega \) is symmetric if \( -\Omega = \Omega \). A continuous function, \( f : \Omega \to \mathbb{R}^n \) is odd if \( f(-x) = -f(x) \).
Suppose \( \Omega \) is symmetric and \( g \in C^\infty \left( \overline{\Omega}; \mathbb{R}^n \right) \) is an odd map for which \( 0 \) is a regular value. Then the chain rule implies \( Dg(-x) = Dg(x) \) and so \( d(g, \Omega, 0) \) must equal an odd integer because if \( x \in g^{-1}(0) \), it follows that \( -x \in g^{-1}(0) \) also and since \( Dg(-x) = Dg(x) \), it follows the overall contribution to the degree from \( x \) and \(-x\) must be an even integer. Also \( 0 \in g^{-1}(0) \) and so the degree equals an even integer added to \( \text{sgn} \left( \det Dg(0) \right) \), an odd integer, either \(-1\) or \(1\). It seems reasonable to expect that something like this would hold for an arbitrary continuous odd function defined on symmetric \( \Omega \). In fact this is the case and this is next. The following lemma is the key result used. This approach is due to Gromes [31]. See also Deimling [11] which is where I found this argument.

The idea is to start with a smooth odd map and approximate it with a smooth odd map which also has \( 0 \) a regular value.

**Lemma 8.4.2** Let \( g \in C^\infty \left( \overline{\Omega}; \mathbb{R}^n \right) \) be an odd map. Then for every \( \varepsilon > 0 \), there exists \( h \in C^\infty \left( \overline{\Omega}; \mathbb{R}^n \right) \) such that \( h \) is also an odd map, \( \|h - g\|_\infty < \varepsilon \), and \( 0 \) is a regular value of \( h \). Here \( \Omega \) is a symmetric bounded open set. In addition, \( d(g, \Omega, 0) \) is an odd integer.

**Proof:** In this argument \( \eta > 0 \) will be a small positive number and \( C \) will be a constant which depends only on the diameter of \( \Omega \). Let \( h_0(x) = g(x) + \eta x \) where \( \eta \) is chosen such that \( \det Dh_0(0) \neq 0 \). Note that \( h_0 \) is odd, close to \( g \) and so \( 0 \) is a value of \( h_0 \) thanks to \( h_0(0) = 0 \). The idea is to modify \( h_0 \) such that for all \( x \in h_0^{-1}(0), \det Dh_0(x) \neq 0 \).

Let \( \Omega_i \equiv \{ x \in \Omega : x_i \neq 0 \} \). In other words, leave out the plane \( x_i = 0 \) from \( \Omega \) in order to obtain \( \Omega_i \). A succession of modifications is about to take place on \( \Omega_1, \Omega_1 \cup \Omega_2, \) etc. Finally a function will be obtained on \( \cup_{i=1}^n \Omega_j \) which is everything except \( 0 \).

Define \( h_1(x) \equiv h_0(x) - y^1 x_1^3 \) where \( |y^1| < \eta \) and \( y^1 = (y^1, \cdots, y^n) \) is a regular value of the function, \( x \rightarrow \frac{h_0(x)}{x_1^3} \) for \( x \in \Omega_1 \). The existence of \( y^1 \) follows from Sard’s lemma because this function is in \( C^\infty \left( \Omega_1; \mathbb{R}^n \right) \). Thus \( h_1(x) = 0 \) if and only if \( y^1 = \frac{h_0(x)}{x_1^3} \). Since \( y^1 \) is a regular value, it follows that for such \( x \) satisfying \( h_1(x) = 0 \),

\[
\det \left( \frac{h_{0i,j}(x) x_1^3 - \frac{\partial}{\partial x_i} (x_1^3) h_{0i}(x)}{x_1^0} \right) = \\
\det \left( \frac{h_{0i,j}(x) x_1^3 - \frac{\partial}{\partial x_i} (x_1^3) y^1 x_1^3}{x_1^0} \right) \neq 0
\]

implying that

\[
\det \left( h_{0i,j}(x) - \frac{\partial}{\partial x_i} (x_1^3) y^1 \right) = \det(Dh_1(x)) \neq 0.
\]

This shows \( 0 \) is a regular value of \( h_1 \) on the set \( \Omega_1 \) and it is clear \( h_1 \) is an odd map in \( C^\infty \left( \overline{\Omega}; \mathbb{R}^n \right) \) and \( \|h_1 - g\|_\infty \leq C\eta \) where \( C \) depends only on the diameter of \( \Omega \).
Now suppose for some $k$ such that $1 \leq k < n$ there exists an odd mapping $h_k$ in $C^\infty (\overline{\Omega}; \mathbb{R}^n)$ such that $0$ is a regular value of $h_k$ on $\cup_{i=1}^k \Omega_i$ and $\|h_k - g\|_\infty \leq C \eta$. Sard’s theorem implies there exists $y^{k+1}$ a regular value of the function $x \rightarrow h_k(x)/x_{k+1}^3$ defined on $\Omega_{k+1}$ such that $\|y^{k+1}\| < \eta$ and let $h_{k+1} (x) \equiv h_k(x) - y^{k+1}x_{k+1}$. As before, $h_{k+1} (x) = 0$ if and only if $h_k(x)/x_{k+1}^3 = y^{k+1}$, a regular value of $x \rightarrow h_k(x)/x_{k+1}^3$. Consider such $x$ for which $h_{k+1}(x) = 0$. First suppose $x \in \Omega_{k+1}$. Then

$$\det \left( \frac{h_{k,j} \left( x \right) x_{k+1}^3 - \frac{\partial}{\partial x_j} \left( x_{k+1}^3 \right) y_{k+1}^{k+1}}{x_{k+1}^3} \right) \neq 0$$

which implies that whenever $h_{k+1}(x) = 0$ and $x \in \Omega_{k+1}$,

$$\det \left( h_{k,j} (x) - \frac{\partial}{\partial x_j} (x_{k+1}^3) y_{k+1}^{k+1} \right) = \det (Dh_{k+1}(x)) \neq 0. \quad (8.4.5)$$

However, if $x \in \cup_{i=1}^k \Omega_i$ but $x \notin \Omega_{k+1}$, then $x_{k+1} = 0$ since $\Omega_{k+1}$ consists of those $x$ such that $x_{k+1} \neq 0$, and so the left side of (8.4.5) reduces to $\det (h_{k,j}(x))$ which is not zero because $0$ is assumed a regular value of $h_k$ on $\cup_{i=1}^k \Omega_i$. Therefore, $0$ is a regular value for $h_{k+1}$ on $\cup_{i=1}^k \Omega_i$. (For $x \in \cup_{i=1}^k \Omega_i$, either $x \in \Omega_{k+1}$ or $x \notin \Omega_{k+1}$. If $x \in \Omega_{k+1}$, $0$ is a regular value by the construction above. In the other case, $0$ is a regular value by the induction hypothesis.) Also $h_{k+1}$ is odd and in $C^\infty (\overline{\Omega}; \mathbb{R}^n)$, and $\|h_{k+1} - g\|_\infty \leq C \eta$.

Let $h \equiv h_n$. Then $0$ is a regular value of $h$ for $x \in \cup_{j=1}^n \Omega_j$. The point of the $\Omega$ which is not in $\cup_{j=1}^n \Omega_j$ is $0$. If $x = 0$, then from the construction, $Dh(0) = Dh_0(0)$ and so $0$ is a regular value of $h$ for $x \in \Omega$. By choosing $\eta$ small enough, it follows $\|h - g\|_\infty < \varepsilon$.

For the last part, let $3 \delta = \text{dist} (g(\partial \Omega), 0)$ and let $h$ be as described above with $\|h - g\|_\infty < \delta$. Then $0 \notin (th + (1-t)g)(\partial \Omega)$ and so by the homotopy invariance of the degree, $t \rightarrow d (th + (1-t)g, \Omega, 0)$ is constant for $t \in [0, 1]$. Therefore,

$$d (g, \Omega, 0) = d (h, \Omega, 0)$$

So what is $d (h, \Omega, 0)$? Since $0$ is a regular value and $h$ is odd,

$$h^{-1}(0) = \{x_1, \ldots, x_r, -x_1, \ldots, -x_r, 0\}.$$

So consider $Dh(x)$ and $Dh(-x)$.

$$Dh(-x) + o(u) = h(-x + u) - h(-x) = -h(x + (u)) + h(x) = -(Dh(x)(u)) + o(-u) = Dh(x)(u) + o(u)$$
Hence $Dh(x) = Dh(-x)$ and so the determinants of these two are the same. It follows that
\[
d(h, \Omega, 0) = \sum_{i=1}^{r} \text{sgn} \left( \det \left( Dh(x_i) \right) \right) + \sum_{i=1}^{r} \text{sgn} \left( \det \left( Dh(-x_i) \right) \right) + \text{sgn} \left( \det \left( Dh(0) \right) \right)
\]
\[= 2m \pm 1 \text{ some integer } m\]
an odd integer.

**Theorem 8.4.3** (Borsuk) Let $f \in C(\overline{\Omega}; \mathbb{R}^n)$ be odd and let $\Omega$ be symmetric with $0 \notin f(\partial \Omega)$. Then $d(f, \Omega, 0)$ equals an odd integer.

**Proof:** Let $\psi_n$ be a mollifier which is symmetric, $\psi(-x) = \psi(x)$. Also recall that $f$ is the restriction to $\Omega$ of a continuous function, still denoted as $f$ which is defined on all of $\mathbb{R}^n$. Let $g$ be the odd part of this function. That is,
\[
g(x) \equiv \frac{1}{2}(f(x) - f(-x))
\]

Since $f$ is odd, $g = f$ on $\Omega$. Then
\[
g_n(-x) \equiv g \ast \psi_n(-x) = \int_{\mathbb{R}^n} g(-x - y) \psi_n(y) \, dy
\]
\[= -\int_{\mathbb{R}^n} g(x + y) \psi_n(y) \, dy = -\int_{\mathbb{R}^n} g(x - (y)) \psi_n(-y) \, dy = -g_n(x)
\]
Thus $g_n$ is odd and is infinitely differentiable. Let $3\delta = \text{dist}(f(\partial \Omega), 0)$ and let $n$ be large enough that $\|g_n - f\|_{\infty} < \delta$. Then $0 \notin (tg_n + (1-t)f)(\partial \Omega)$ for $t \in [0,1]$ and so by homotopy invariance,
\[
d(f, \Omega, 0) = d(g, \Omega, 0) = d(g_n, \Omega, 0)
\]
and by Lemma this is an odd integer.

### 8.5 Applications

With these theorems it is possible to give easy proofs of some very important and difficult theorems.

**Definition 8.5.1** If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ where $U$ is an open set. Then $f$ is locally one to one if for every $x \in U$, there exists $\delta > 0$ such that $f$ is one to one on $B(x, \delta)$.

As a first application, consider the invariance of domain theorem. This result says that a one to one continuous map takes open sets to open sets. It is an amazing result which is essential to understand if you wish to study manifolds. In fact, the following theorem only requires $f$ to be locally one to one. First here is a lemma which has the main idea.
Lemma 8.5.2 Let $g : \overline{B(0,r)} \rightarrow \mathbb{R}^n$ be one to one and continuous where here $B(0,r)$ is the ball centered at 0 of radius $r$ in $\mathbb{R}^n$. Then there exists $\delta > 0$ such that
\[ g(0) + B(0,\delta) \subseteq g(\overline{B(0,r)}). \]
The symbol on the left means: $\{g(0) + x : x \in B(0,\delta)\}$.

Proof: For $t \in [0,1]$, let
\[ h(x,t) \equiv g \left( \frac{x}{1+t} \right) - g \left( \frac{-tx}{1+t} \right). \]
Then for $x \in \partial B(0,r)$, $h(x,t) \neq 0$ because if this were so, the fact $g$ is one to one implies
\[ x = 0 \]
and this requires $x = 0$ which is not the case since $\|x\| = r$. Then $d(h(\cdot,t),B(0,r),0)$ is constant. Hence it is an odd integer for all $t$ thanks to Borsuk’s theorem, because $h(\cdot,1)$ is odd. Now let $B(0,\delta)$ be such that $B(0,\delta) \cap h(\partial \Omega,0) = \emptyset$. Then $d(h(\cdot,0),B(0,r),0) = d(h(\cdot,0),B(0,r),z)$ for $z \in B(0,\delta)$ because the degree is constant on connected components of $\mathbb{R}^n \setminus h(\partial \Omega,0)$. Hence $z = h(x,0) = g(x) - g(0)$ for some $x \in B(0,r)$. Thus
\[ g(B(0,r)) \supseteq g(0) + B(0,\delta). \]

Now with this lemma, it is easy to prove the very important invariance of domain theorem.

A function $f$ is locally one to one on an open set $\Omega$ if for every $x_0 \in \Omega$, there exists $B(x_0,r) \subseteq \Omega$ such that $f$ is one to one on $B(x_0,r)$.

Theorem 8.5.3 (invariance of domain) Let $\Omega$ be any open subset of $\mathbb{R}^n$ and let $f : \Omega \rightarrow \mathbb{R}^n$ be continuous and locally one to one. Then $f$ maps open subsets of $\Omega$ to open sets in $\mathbb{R}^n$.

Proof: Let $\overline{B(x_0,r)} \subseteq \Omega$ where $f$ is one to one on $\overline{B(x_0,r)}$. Let $g$ be defined on $B(0,r)$ given by
\[ g(x) \equiv f(x + x_0). \]
Then $g$ satisfies the conditions of Lemma 8.5.2, being one to one and continuous. It follows from that lemma there exists $\delta > 0$ such that
\[
\begin{align*}
f(\Omega) & \supseteq f(B(x_0,r)) = f(x_0 + B(0,r)) \\
& = g(B(0,r)) \supseteq g(0) + B(0,\delta) \\
& = f(x_0) + B(0,\delta) = B(f(x_0),\delta)
\end{align*}
\]
This shows that for any $x_0 \in \Omega$, $f(x_0)$ is an interior point of $f(\Omega)$ which shows $f(\Omega)$ is open.

With the above, one gets easily the following amazing result. It is something which is clear for linear maps but this is a statement about continuous maps.
Corollary 8.5.4 If \( n > m \) there does not exist a continuous one to one map from \( \mathbb{R}^n \) to \( \mathbb{R}^m \).

Proof: Suppose not and let \( f \) be such a continuous map, \( f(x) \equiv (f_1(x), \ldots, f_m(x))^T \).

Then let \( g(x) \equiv (f_1(x), \ldots, f_m(x), 0, \ldots, 0)^T \) where there are \( n - m \) zeros added in. Then \( g \) is a one to one continuous map from \( \mathbb{R}^n \) to \( \mathbb{R}^n \) and so \( g(\mathbb{R}^n) \) would have to be open from the invariance of domain theorem and this is not the case.

Corollary 8.5.5 If \( f \) is locally one to one and continuous, \( f: \mathbb{R}^n \to \mathbb{R}^n \), and \( \lim_{|x| \to \infty} |f(x)| = \infty \), then \( f \) maps \( \mathbb{R}^n \) onto \( \mathbb{R}^n \).

Proof: By the invariance of domain theorem, \( f(\mathbb{R}^n) \) is an open set. It is also true that \( f(\mathbb{R}^n) \) is a closed set. Here is why. If \( f(x_k) \to y \), the growth condition ensures that \( \{x_k\} \) is a bounded sequence. Taking a subsequence which converges to \( x \in \mathbb{R}^n \) and using the continuity of \( f \), it follows \( f(x) = y \). Thus \( f(\mathbb{R}^n) \) is both open and closed which implies \( f \) must be an onto map since otherwise, \( \mathbb{R}^n \) would not be connected.

The next theorem is the famous Brouwer fixed point theorem.

Theorem 8.5.6 (Brouwer fixed point) Let \( B = \overline{B(0, r)} \subseteq \mathbb{R}^n \) and let \( f: B \to B \) be continuous. Then there exists a point \( x \in B \), such that \( f(x) = x \).

Proof: Assume there is no fixed point. Consider \( h(x, t) \equiv tf(x) - x \) for \( t \in [0, 1] \). Then for \( |x| = r \),

\[
0 \notin tf(x) - x, t \in [0, 1]
\]

is constant. But when \( t = 0 \), this is \( d(-I, B, 0) = (-1)^n \neq 0 \). Hence \( d(f - I, B, 0) \neq 0 \) so there exists \( x \) such that \( f(x) - x = 0 \).

It is easy to generalize this to an arbitrary closed bounded convex set \( K \) in \( \mathbb{R}^n \) as follows. You use the fact that \( K \) is a retract, Theorem 4.2.5. Thus there exists a mapping \( P \) which is continuous and maps all of \( \mathbb{R}^n \) to \( K \) and fixed points of \( K \). In particular, \( P \) maps \( \overline{B(0, r)} \) to \( K \). Thus there exists \( x \in \overline{B(0, r)} \) such that \( f(P(x)) = x \). But then \( x \in K \) and so \( P(x) = x \). Thus \( f(x) = x \).

You can also use standard stuff from Hilbert space to get this. Let \( K \) be a closed bounded convex set and let \( f: K \to K \) be continuous. Let \( P \) be the projection map onto \( K \). Then \( P \) is continuous because \( |Px - Py| \leq |x - y| \). Recall why this
is. From the material on Hilbert space, \((x-Px, y-Px) \leq 0\) for all \(y \in K\). Indeed, this characterizes \(Px\). Therefore,

\[(x-Px, Py-Px) \leq 0, \quad (y-Py, Px-Py) \leq 0 \quad \text{so} \quad (y-Py, Py-Px) \geq 0\]

Hence, subtracting the first from the last,

\[(y-Py-(x-Px), Py-Px) \geq 0\]

consequently,

\[|x-y||Py-Px| \geq (y-x, Py-Px) \geq |Py-Px|^2\]

and so \(|Py-Px| \leq |y-x|\) as claimed.

Now let \(r\) be so large that \(K \subseteq B(0, r)\). Then consider \(f \circ P\). This map takes \(B(0, r) \rightarrow B(0, r)\). In fact it maps \(B(0, r)\) to \(K\). Therefore, being the composition of continuous functions, it is continuous and so has a fixed point in \(B(0, r)\) denoted as \(x\). Hence \(f(P(x)) = x\). Now, since \(f\) maps into \(K\), it follows that \(x \in K\). Hence \(Px = x\) and so \(f(x) = x\). This has proved the following general Brouwer fixed point theorem.

**Theorem 8.5.7** Let \(f : K \rightarrow K\) be continuous where \(K\) is compact and convex and nonempty, \(K \subseteq \mathbb{R}^n\). Then \(f\) has a fixed point.

**Definition 8.5.8** \(f\) is a retract of \(B(0, r)\) onto \(\partial B(0, r)\) if \(f\) is continuous,

\[f\left(B(0, r)\right) \subseteq \partial B(0, r)\]

and \(f(x) = x\) for all \(x \in \partial B(0, r)\).

**Theorem 8.5.9** There does not exist a retraction of \(B(0, r)\) onto its boundary, \(\partial B(0, r)\).

**Proof:** Suppose \(f\) were such a retraction. Then for all \(x \in \partial B(0, r)\), \(f(x) = x\) and so from the properties of the degree, the one which says if two functions agree on \(\partial \Omega\), then they have the same degree,

\[1 = d(I, B(0, r), 0) = d(f, B(0, r), 0)\]

which is clearly impossible because \(f^{-1}(0) = \emptyset\) which implies \(d(f, B(0, r), 0) = 0\).

You should now use this theorem to give another proof of the Brouwer fixed point theorem.

The proofs of the next two theorems make use of the Tietze extension theorem, Theorem [7.1.1].

**Theorem 8.5.10** Let \(\Omega\) be a symmetric open set in \(\mathbb{R}^n\) such that \(\emptyset \in \Omega\) and let \(f : \partial \Omega \rightarrow V\) be continuous where \(V\) is an \(m\) dimensional subspace of \(\mathbb{R}^n\), \(m < n\). Then \(f(-x) = f(x)\) for some \(x \in \partial \Omega\).
Proof: Suppose not. Using the Tietze extension theorem or Theorem 4.2.3, extend \( f \) to all of \( \mathbb{R}^n \), \( f(\overline{\Omega}) \subseteq V \). (Here the extended function is also denoted by \( f \).) Let \( g(x) = f(x) - f(-x) \). Then \( 0 \notin g(\partial \Omega) \) and so for some \( r > 0 \), \( B(0,r) \subseteq \mathbb{R}^n \setminus g(\partial \Omega) \). For \( z \in B(0,r) \),

\[
d(g, \Omega, z) = d(g, \Omega, 0) \neq 0
\]

because \( B(0,r) \) is contained in a component of \( \mathbb{R}^n \setminus g(\partial \Omega) \) and Borsuk’s theorem implies that \( d(g, \Omega, 0) \neq 0 \) since \( g \) is odd. Hence

\[
V \supseteq g(\Omega) \supseteq B(0,r)
\]

and this is a contradiction because \( V \) is \( m \) dimensional. \( \blacksquare \)

This theorem is called the Borsuk Ulam theorem. Note that it implies there exist two points on opposite sides of the surface of the earth which have the same atmospheric pressure and temperature, assuming the earth is symmetric and that pressure and temperature are continuous functions. The next theorem is an amusing result which is like combing hair.

Theorem 8.5.11 Let \( n \) be odd and let \( \Omega \) be an open bounded set in \( \mathbb{R}^n \) with \( 0 \in \Omega \). Suppose \( f : \partial \Omega \to \mathbb{R}^n \setminus \{0\} \) is continuous. Then for some \( x \in \partial \Omega \) and \( \lambda \neq 0 \), \( f(x) = \lambda x \).

Proof: Using the Tietze extension theorem or Theorem 4.2.3, extend \( f \) to all of \( \mathbb{R}^n \). Also denote the extended function by \( f \). Suppose for all \( x \in \partial \Omega \), \( f(x) \neq \lambda x \) for all \( \lambda \in \mathbb{R} \). Then

\[
0 \notin tf(x) + (1 - t)x, \ (x,t) \in \partial \Omega \times [0,1]
\]

Thus there exists a homotopy of \( f \) and \( I \) and a homotopy of \( f \) and \( -I \). Then by the homotopy invariance of degree,

\[
d(f, \Omega, 0) = d(I, \Omega, 0), \ d(f, \Omega, 0) = d(-I, \Omega, 0).
\]

But this is impossible because \( d(I, \Omega, 0) = 1 \) but \( d(-I, \Omega, 0) = (-1)^n = -1 \). \( \blacksquare \)

8.6 The Product Formula

This section is on the product formula for the degree which is used to prove the Jordan separation theorem. To begin with here is a lemma which is similar to an earlier result except here there are \( r \) points.

Lemma 8.6.1 Let \( y_1, \ldots, y_r \) be points not in \( f(\partial \Omega) \) and let \( \delta > 0 \). Then there exists \( \bar{f} \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \) such that \( \|\bar{f} - f\|_\infty < \delta \) and \( y_i \) is a regular value for \( \bar{f} \) for each \( i \).
8.6. THE PRODUCT FORMULA

Proof: Let \( f_0 \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \), \( ||f_0 - f||_\infty < \frac{\delta}{2} \). Let \( \bar{y}_1 \) be a regular value for \( f_0 \) and \( |\bar{y}_1 - y_1| < \frac{\delta}{3r} \). Let \( f_1(x) \equiv f_0(x) + y_1 - \bar{y}_1 \). Thus \( y_1 \) is a regular value of \( f_1 \) because \( Df_1(x) = Df_0(x) \) and if \( f_1(x) = y_1 \), this is the same as having \( f_0(x) = \bar{y}_1 \) where \( \bar{y}_1 \) is a regular value of \( f_0 \). Then also

\[
||f - f_1||_\infty \leq ||f - f_0||_\infty + ||f_0 - f_1||_\infty \\
= ||f - f_0||_\infty + |\bar{y}_1 - y_1| \\
< \frac{\delta}{3r} + \frac{\delta}{2}.
\]

Suppose now there exists \( f_k \in C^\infty(\overline{\Omega}; \mathbb{R}^n) \) with each of the \( y_i \) for \( i = 1, \ldots, k \) a regular value of \( f_k \) and

\[
||f - f_k||_\infty < \frac{\delta}{2} + \frac{k}{r} \left( \frac{\delta}{3} \right).
\]

Then letting \( S_k \) denote the singular values of \( f_k \), Sard’s theorem implies there exists \( \bar{y}_{k+1} \) such that

\[
|\bar{y}_{k+1} - y_{k+1}| < \frac{\delta}{3r}
\]

and

\[
\bar{y}_{k+1} \notin S_k \cup \bigcup_{i=1}^k (S_k + y_{k+1} - y_i).
\]

(8.6.6) Let

\[
f_{k+1}(x) \equiv f_k(x) + y_{k+1} - \bar{y}_{k+1}.
\]

(8.6.7) If \( f_{k+1}(x) = y_i \) for some \( i \leq k \), then

\[
f_k(x) + y_{k+1} - y_i = \bar{y}_{k+1}
\]

and so \( f_k(x) \) is a regular value for \( f_k \) since by \( \text{(8.6.6)} \) \( \bar{y}_{k+1} \notin S_k \). Therefore, for \( i \leq k \), \( y_i \) is a regular value of \( f_{k+1} \) since by \( \text{(8.6.7)} \), \( Df_{k+1} = Df_k \). Now suppose \( f_{k+1}(x) = y_{k+1} \). Then

\[
y_{k+1} = f_k(x) + y_{k+1} - \bar{y}_{k+1}
\]

so \( f_k(x) = \bar{y}_{k+1} \) implying that \( f_k(x) = \bar{y}_{k+1} \notin S_k \). Hence \( \text{det } Df_{k+1}(x) = \text{det } Df_k(x) \neq 0 \). Thus \( y_{k+1} \) is also a regular value of \( f_{k+1} \). Also,

\[
||f_{k+1} - f|| \leq ||f_{k+1} - f_k|| + ||f_k - f|| \\
\leq \frac{\delta}{3r} + \frac{\delta}{2} + \frac{k}{r} \left( \frac{\delta}{3} \right) = \frac{\delta}{2} + \frac{k+1}{r} \left( \frac{\delta}{3} \right).
\]

Let \( \bar{f} \equiv f \). Then

\[
||f - f||_\infty < \frac{\delta}{2} + \left( \frac{\delta}{3} \right) < \delta
\]

and each of the \( y_i \) is a regular value of \( \bar{f} \). ■
Definition 8.6.2 Let the connected components of \( \mathbb{R}^n \setminus f(\partial \Omega) \) be denoted by \( K_i \). From the properties of the degree listed in Theorem 8.3.7, \( d(f, \Omega, \cdot) \) is constant on each of these components. Denote by \( d(f, \Omega, K_i) \) the constant value on the component, \( K_i \).

The product formula considers the situation depicted in the following diagram in which \( y \notin g(f(\partial \Omega)) \) and the \( K_i \) are the connected components of \( \mathbb{R}^n \setminus f(\partial \Omega) \).

The following diagram may be helpful in remembering what it says.

\[
\begin{array}{ccc}
\Omega & \xrightarrow{f} & f(\Omega) \\
R^n \setminus f(\partial \Omega) = \bigcup_i K_i & \xrightarrow{g} & R^n \setminus y
\end{array}
\]

Lemma 8.6.3 Let \( f \in C(\overline{\Omega}; \mathbb{R}^n) \), \( g \in C^\infty(\mathbb{R}^n, \mathbb{R}^n) \), and \( y \notin g(f(\partial \Omega)) \). Suppose also that \( y \) is a regular value of \( g \). Then the following product formula holds where \( K_i \) are the bounded components of \( \mathbb{R}^n \setminus f(\partial \Omega) \).

\[
d(g \circ f, \Omega, y) = \sum_{i=1}^\infty d(f, \Omega, K_i) d(g, K_i, y).
\]

All but finitely many terms in the sum are zero.

Proof: First note that if \( K_i \) is unbounded, \( d(f, \Omega, K_i) = 0 \) because there exists a point, \( z \in K_i \) such that \( f^{-1}(z) = \emptyset \) due to the fact that \( f(\overline{\Omega}) \) is compact and is consequently bounded. Thus it makes no difference in the above formula whether the \( K_i \) are arbitrary components or only bounded components. Let \( \{x_j\} \) denote the points of \( g^{-1}(y) \) which are contained in \( K_i \), the \( i^{th} \) bounded component of \( \mathbb{R}^n \setminus f(\partial \Omega) \). Then \( m_i < \infty \) because if not, there would exist a limit point \( x \) for this sequence. Then \( g(x) = y \) and so \( x \notin f(\partial \Omega) \). Thus \( \det(Dg(x)) \neq 0 \) and so by the inverse function theorem, \( g \) would be one to one on an open ball containing \( x \) which contradicts having \( x \) a limit point.

Note also that \( g^{-1}(y) \cap f(\overline{\Omega}) \) is a compact set covered by the components of \( \mathbb{R}^n \setminus f(\partial \Omega) \) because by assumption, \( g^{-1}(y) \cap f(\partial \Omega) = \emptyset \). It follows \( g^{-1}(y) \cap f(\overline{\Omega}) \) is covered by finitely many of these components. It is not in \( f(\partial \Omega) \).
The only terms in the above sum which are nonzero are those corresponding to \( K_i \) having nonempty intersection with \( g^{-1}(y) \cap f(\Omega) \). The other components contribute 0 to the above sum because if \( K_i \cap g^{-1}(y) = \emptyset \), it follows from Theorem 8.3.7 that \( d(g, K_i, y) = 0 \). If \( K_i \) does not intersect \( f(\Omega) \), then \( d(f, \Omega, K_i) = 0 \). Therefore, the above sum is actually a finite sum since \( g^{-1}(y) \cap f(\Omega) \), being a compact set, is covered by finitely many of the \( K_i \). Thus there are no convergence problems.

Let \( d(f, \Omega, K_i) = d(f, \Omega, u_i^j) \) where the \( \{u_i^j\}_{j=1}^{m_i} \) are the points in \( g^{-1}(y) \cap K_i \). By Lemma S.B.1 there exists \( \tilde{f} \) such that \( \|\tilde{f} - f\|_\infty \) is very small and each of the \( u_i^j \) are regular values for \( \tilde{f} \). If \( \|\tilde{f} - f\|_\infty \) is small enough, then \( (f + t(\tilde{f} - f))(\partial \Omega) \) does not contain any of the \( u_i^j \). This is so because by the definition of \( u_i^j \) they are in some \( K_i \) and these are connected components of \( R^n \setminus f(\partial \Omega) \). Thus

\[
\sum_i \sum_j d(f, \Omega, K_i) d(f, \Omega, u_i^j) = \sum_i \sum_j d(g \circ f, \Omega, y) d(g \circ \tilde{f}, \Omega, x_i^j)
\]

This is so because by the definition of \( u_i^j \) they are in some \( K_i \) and these are connected components of \( R^n \setminus f(\partial \Omega) \). Thus

\[
\sum_i \sum_j d(g \circ f, \Omega, y) d(g \circ \tilde{f}, \Omega, x_i^j) = \sum_i d(g, K_i, y) d(f, \Omega, K_i)
\]

With this lemma, the following is the product formula.
Theorem 8.6.4 (product formula) Let \( \{K_i\}_{i=1}^\infty \) be the bounded components of \( \mathbb{R}^n \setminus f(\partial \Omega) \) for \( f \in C(\overline{\Omega};\mathbb{R}^n) \), let \( g \in C(\mathbb{R}^n,\mathbb{R}^n) \), and suppose that \( y \notin g(f(\partial \Omega)) \). Then

\[
d(g \circ f, \Omega, y) = \sum_{i=1}^{\infty} d(g, K_i, y) d(f, \Omega, K_i).
\]

(8.6.9)

All but finitely many terms in the sum are zero.

Proof: Let \( \sup \{|\tilde{g}(z) - g(z)| : z \in f(\overline{\Omega})\} \) be sufficiently small that

\[
y \notin (g \circ f + t(\tilde{g} \circ f - g \circ f))(\partial \Omega), \ t \in [0, 1]
\]

\( \tilde{g} \) being \( C^\infty(\mathbb{R}^n,\mathbb{R}^n) \) with \( y \) a regular value of \( \tilde{g} \). It follows that

\[
d(g \circ f, \Omega, y) = d(\tilde{g} \circ f, \Omega, y).
\]

(8.6.10)

Now also, the \( K_i \) are the open components of \( \mathbb{R}^n \setminus f(\partial \Omega) \) and so \( \partial K_i \subseteq f(\partial \Omega) \) (if \( x \in \partial K_i \), then if \( x \notin f(\partial \Omega) \), it would be in a ball contained in one of the \( K_j \) and so could not be in \( \partial K_i \) and so if \( z \in \partial K_i \), then \( g(z) \in g(f(\partial \Omega)) \). Consequently, for \( t \in [0, 1] \),

\[
y \notin (g + t(\tilde{g} - g))(\partial K_i)
\]

(\( y \) is not in the larger set \( (g \circ f + t(\tilde{g} \circ f - g \circ f))(\partial \Omega) \)) which shows that, by homotopy invariance,

\[
d(g, K_i, y) = d(\tilde{g}, K_i, y).
\]

(8.6.11)

Therefore, by Lemma 8.6.3,

\[
d(g \circ f, \Omega, y) = d(\tilde{g} \circ f, \Omega, y) = \sum_{i=1}^{\infty} d(\tilde{g}, K_i, y) d(f, \Omega, K_i)
\]

\[
= \sum_{i=1}^{\infty} d(g, K_i, y) d(f, \Omega, K_i)
\]

and the sum has only finitely many non-zero terms.

Note there are no convergence problems because these sums are actually finite sums because, as in the previous lemma, \( g^{-1}(y) \cap f(\overline{\Omega}) \) is a compact set covered by the components of \( \mathbb{R}^n \setminus f(\partial \Omega) \) and so it is covered by finitely many of these components. For the other components, \( d(f, \Omega, K_i) = 0 \) or else \( d(g, K_i, y) = 0 \).

The following theorem is the Jordan separation theorem, a major result. A homeomorphism is a function which is one to one onto and continuous having continuous inverse. Before the theorem, here is a helpful lemma.

Lemma 8.6.5 Let \( \Omega \) be a bounded open set in \( \mathbb{R}^n \), \( f \in C(\overline{\Omega};\mathbb{R}^n) \), and suppose \( \{\Omega_i\}_{i=1}^{\infty} \) are disjoint open sets contained in \( \Omega \) such that

\[
y \notin f(\overline{\Omega} \setminus \cup_{j=1}^{\infty} \Omega_j)
\]
8.6. THE PRODUCT FORMULA

Then
\[ d(f, \Omega, y) = \sum_{j=1}^{\infty} d(f, \Omega_j, y) \]
where the sum has all but finitely many terms equal to 0.

**Proof:** By assumption, the compact set \( f^{-1}(y) = \{ x \in \Omega : f(x) = y \} \) has empty intersection with
\[ \Omega \setminus \bigcup_{j=1}^{\infty} \Omega_j \]
and so this compact set is covered by finitely many of the \( \Omega_j \), say \( \{ \Omega_1, \ldots, \Omega_n \} \) and
\[ y \notin f(\bigcup_{j=n}^{\infty} \Omega_j). \]
By Theorem 8.3.4 and letting \( O = \bigcup_{j=n}^{\infty} \Omega_j \),
\[ d(f, \Omega, y) = \sum_{j=1}^{n-1} d(f, \Omega_j, y) + d(f, O, y) = \sum_{j=1}^{\infty} d(f, \Omega_j, y) \]
because \( d(f, O, y) = 0 \) as is \( d(f, \Omega_j, y) \) for every \( j \geq n \). \( \blacksquare \)

**Lemma 8.6.6** Define \( \partial U \) to be those points \( x \) with the property that for every \( r > 0 \), \( B(x, r) \) contains points of \( U \) and points of \( U^C \). Then for \( U \) an open set,
\[ \partial U = \overline{U} \setminus U \]
Let \( C \) be a closed subset of \( \mathbb{R}^n \) and let \( K \) denote the set of components of \( \mathbb{R}^n \setminus C \). Then if \( K \) is one of these components, it is open and
\[ \partial K \subseteq C \]

**Proof:** Let \( x \in \overline{U} \setminus U \). If \( B(x, r) \) contains no points of \( U \), then \( x \notin \overline{U} \). If \( B(x, r) \) contains no points of \( U^C \), then \( x \in U \) and so \( x \notin \overline{U} \setminus U \). Therefore, \( U \setminus U \subseteq \partial U \).
Now let \( x \in \partial U \). If \( x \in U \), then since \( U \) is open there is a ball containing \( x \) which is contained in \( U \) contrary to \( x \in \partial U \). Therefore, \( x \notin U \). If \( x \) is not a limit point of \( U \), then some ball containing \( x \) contains no points of \( U \) contrary to \( x \in \partial U \). Therefore, \( x \in \overline{U} \setminus U \) which shows the two sets are equal.

Why is \( K \) open for \( K \) a component of \( \mathbb{R}^n \setminus C \)? This is obvious because in \( \mathbb{R}^n \) an open ball is connected. Thus if \( k \in K \), letting \( B(k, r) \subseteq C \), it follows \( K \cup B(k, r) \) is connected and contained in \( C \). Thus \( K \cup B(k, r) \) is connected, contained in \( C \), and therefore is contained in \( K \) because \( K \) is maximal with respect to being connected and contained in \( C \).

Now for \( K \) a component of \( \mathbb{R}^n \setminus C \), why is \( \partial K \subseteq \partial C? \) Let \( x \in \partial K \). If \( x \notin C \), then \( x \in K_1 \), some component of \( \mathbb{R}^n \setminus C \). If \( K_1 \neq K \) then \( x \) cannot be a limit point of \( K \) and so it cannot be in \( \partial K \). Therefore, \( K = K_1 \) but this also is a contradiction because if \( x \in \partial K \) then \( x \notin K \). \( \blacksquare \)

I will give a shorter version of the proof and a longer version. First is the shorter version which leaves out a few details which may or may not be clear. Sometimes, it seems to me that when you put in too many details, you lose the forest by stumbling around hitting trees. It may still have too many details.
Theorem 8.6.7 (Jordan separation theorem) Let \( f \) be a homeomorphism of \( C \) and \( f(C) \) where \( C \) is a compact set in \( \mathbb{R}^n \). Then \( \mathbb{R}^n \setminus C \) and \( \mathbb{R}^n \setminus f(C) \) have the same number of connected components.

**Proof:** Denote by \( K \) the bounded components of \( \mathbb{R}^n \setminus C \) and denote by \( L \), the bounded components of \( \mathbb{R}^n \setminus f(C) \). Also, using the Tietze extension theorem, there exists \( \bar{f} \) an extension of \( f \) to all of \( \mathbb{R}^n \) which maps into a bounded set and let \( \bar{f}^{-1} \) be an extension of \( f^{-1} \) to all of \( \mathbb{R}^n \) which also maps into a bounded set. Pick \( K \in K \) and take \( y \in K \). Then \( \partial K \subseteq C \) and so

\[
y \notin \bar{f}^{-1} \left( \bar{f}(\partial K) \right)
\]

Since \( \bar{f}^{-1} \circ \bar{f} \) equals the identity \( I \) on \( \partial K \), it follows from the properties of the degree that

\[
1 = d(I, K, y) = d \left( \bar{f}^{-1} \circ \bar{f}, K, y \right).
\]

Recall that if two functions agree on the boundary, then they have the same degree. Let \( H \) denote the set of bounded components of \( \mathbb{R}^n \setminus f \left( \partial K \right) \). These will be as large as those in \( L \) and if a set in \( L \) intersects one of these larger \( H \in H \) then \( H \) contains the component in \( L \). By the product formula,

\[
1 = d \left( \bar{f}^{-1} \circ \bar{f}, K, y \right) = \sum_{H \in H} d \left( \bar{f}, K, H \right) d \left( \bar{f}^{-1}, H, y \right),
\]

the sum being a finite sum from the product formula. That is, there are finitely many \( H \) involved in the sum, the other terms being zero.

What about those sets of \( H \) which contain no set of \( L \)? These sets also have empty intersection with all sets of \( L \). Therefore, for \( H \) one of these, \( H \subseteq f(C) \). Therefore,

\[
d \left( \bar{f}^{-1}, H, y \right) = d \left( f^{-1}, H, y \right) = 0
\]

because \( y \in K \) a component of \( \mathbb{R}^n \setminus C \), but for \( u \in H \subseteq f(C) \), \( f^{-1}(u) \in C \) so \( f^{-1}(u) \neq y \) implying that \( d \left( f^{-1}, H, y \right) = 0 \). Thus in (8.6.12) all such terms are zero.

Then letting \( \mathcal{H}_1 \) be those sets of \( \mathcal{H} \) which contain (intersect) some sets of \( L \), the above sum reduces to

\[
\sum_{H \in \mathcal{H}_1} d \left( \bar{f}, K, H \right) d \left( \bar{f}^{-1}, H, y \right) = \sum_{H \in \mathcal{H}_1} d \left( \bar{f}, K, H \right) \sum_{L \in \mathcal{L}_H} d \left( \bar{f}^{-1}, L, y \right)
\]

where \( \mathcal{L}_H \) are those sets of \( \mathcal{L} \) contained in \( H \). If \( \mathcal{L}_H = \emptyset \), the above shows that the second sum is 0 with the convention that \( \sum_{\emptyset} = 0 \). Now \( d \left( \bar{f}, K, H \right) = d \left( \bar{f}, K, L \right) \) where \( L \in \mathcal{L}_H \). Therefore,

\[
\sum_{H \in \mathcal{H}_1} \sum_{L \in \mathcal{L}_H} d \left( \bar{f}, K, H \right) d \left( \bar{f}^{-1}, L, y \right) = \sum_{H \in \mathcal{H}_1} \sum_{L \in \mathcal{L}_H} d \left( \bar{f}, K, L \right) d \left( \bar{f}^{-1}, L, y \right)
\]
8.6. THE PRODUCT FORMULA

As noted above, there are finitely many \( H \in H \) which are involved. \( \mathbb{R}^n \setminus f(C) \subseteq \mathbb{R}^n \setminus f(\partial K) \) and so every \( L \) must be contained in some \( H \in H_1 \). It follows that the above reduces to

\[
\sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, y \right)
\]

Thus from (8.6.12)

\[
1 = \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, y \right) = \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, K \right) \tag{8.6.13}
\]

Let \( |\mathcal{K}| \) denote the number of components in \( \mathcal{K} \) and similarly, \( |\mathcal{L}| \) denotes the number of components in \( \mathcal{L} \). Thus

\[
|\mathcal{K}| = \sum_{K \in \mathcal{K}} 1 = \sum_{K \in \mathcal{K}} \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, K \right)
\]

Similarly,

\[
|\mathcal{L}| = \sum_{L \in \mathcal{L}} 1 = \sum_{L \in \mathcal{L}} \sum_{K \in \mathcal{K}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, K \right)
\]

If \( |\mathcal{K}| < \infty \), then

\[
\sum_{K \in \mathcal{K}} \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, K \right) < \infty.
\]

The summation which equals 1 is a finite sum and so is the outside sum. Hence we can switch the order of summation and get

\[
|\mathcal{K}| = \sum_{L \in \mathcal{L}} \sum_{K \in \mathcal{K}} d(\bar{f}, K, L) d\left( \bar{f}^{-1}, L, K \right) = |\mathcal{L}|
\]

A similar argument applies if \( |\mathcal{L}| < \infty \). Thus if one of these numbers is finite, so is the other and they are equal. It follows that \( |\mathcal{L}| = |\mathcal{K}| \). \( \blacksquare \)

Now is the same proof with more details included.

**Theorem 8.6.8 (Jordan separation theorem)** Let \( f \) be a homeomorphism of \( C \) and \( f(C) \) where \( C \) is a compact set in \( \mathbb{R}^n \). Then \( \mathbb{R}^n \setminus C \) and \( \mathbb{R}^n \setminus f(C) \) have the same number of connected components.

**Proof:** Denote by \( \mathcal{K} \) the bounded components of \( \mathbb{R}^n \setminus C \) and denote by \( \mathcal{L} \), the bounded components of \( \mathbb{R}^n \setminus f(C) \). Also, using the Tietze extension theorem, there exists \( \bar{f} \) an extension of \( f \) to all of \( \mathbb{R}^n \) which maps into a bounded set and let \( \bar{f}^{-1} \) be an extension of \( f^{-1} \) to all of \( \mathbb{R}^n \) which also maps into a bounded set. Pick \( K \in \mathcal{K} \) and take \( y \in K \). Then

\[
y \notin \bar{f}^{-1}(f(\partial K))
\]

because by Lemma 8.6.4, \( \partial K \subseteq C \) and on \( C, \bar{f} = f \). Thus the right side is of the form

\[
\bar{f}^{-1}\left( \frac{\subseteq f(C)}{f(\partial K)} \right) = f^{-1}(f(\partial K)) \subseteq C
\]
and \( y \notin C \). Since \( \overline{f^{-1} \circ f} \) equals the identity \( I \) on \( \partial K \), it follows from the properties of the degree that

\[
1 = d(I, K, y) = d \left( \overline{f^{-1} \circ f}, K, y \right).
\]

Recall that if two functions agree on the boundary, then they have the same degree. Let \( \mathcal{H} \) denote the set of bounded components of \( \mathbb{R}^n \setminus f(\partial K) \). (These will be as large as those in \( \mathcal{L} \)) By the product formula,

\[
1 = d \left( \overline{f^{-1} \circ f}, K, y \right) = \sum_{H \in \mathcal{H}} d(f, K, H) d \left( \overline{f^{-1}}, H, y \right), \tag{8.6.14}
\]

the sum being a finite sum from the product formula. It might help to consult the following diagram.

Now letting \( x \in L \in \mathcal{L} \), if \( S \) is a connected set containing \( x \) and contained in \( \mathbb{R}^n \setminus f(C) \), then it follows \( S \) is contained in \( \mathbb{R}^n \setminus f(\partial K) \) because \( \partial K \subseteq C \). Therefore, every set of \( \mathcal{L} \) is contained in some set of \( \mathcal{H} \). Furthermore, if any \( L \in \mathcal{L} \) has nonempty intersection with \( H \in \mathcal{H} \) then it must be contained in \( H \). This is because

\[
L = (L \cap H) \cup (L \cap \partial H) \cup \left( L \cap \overline{f(C)} \right).
\]

Now by Lemma \ref{lem:8.6.6},

\[
L \cap \partial H \subseteq L \cap f(\partial K) \subseteq L \cap f(C) = \emptyset.
\]

Since \( L \) is connected, \( L \cap \overline{f(C)} = \emptyset \). Letting \( \mathcal{L}_H \) denote those sets of \( \mathcal{L} \) which are contained in \( H \) equivalently having nonempty intersection with \( H \), if \( p \in H \setminus \cup \mathcal{L}_H = H \setminus \cup \mathcal{L} \), then \( p \in H \cap f(C) \) and so

\[
H = (\cup \mathcal{L}_H) \cup (H \cap f(C)) \tag{8.6.15}
\]

**Claim 1:**

\[
\overline{f^{-1}} \setminus \cup \mathcal{L}_H \subseteq f(C).
\]

**Proof of the claim:** Suppose \( p \in \overline{f^{-1}} \setminus \cup \mathcal{L}_H \) but \( p \notin f(C) \). Then \( p \notin L \) \( \in \mathcal{L} \). It must be the case that \( L \) has nonempty intersection with \( H \) since otherwise \( p \) could not be in \( \overline{f^{-1}} \). However, as shown above, this requires \( L \subseteq H \) and now by \ref{lem:8.6.10} and \( p \notin \cup \mathcal{L}_H \), it follows \( p \notin f(C) \) after all. This proves the claim.

**Claim 2:** \( y \notin \overline{f^{-1}}(\overline{H} \setminus \cup \mathcal{L}_H) \). Recall \( y \in K \in \mathcal{K} \) the bounded components of \( \mathbb{R}^n \setminus C \).
Proof of the claim: If not, then \( f^{-1}(z) = y \) where \( z \in \overline{H} \cup \mathcal{L}_H \subseteq f(C) \) and so \( z = f(w) \) for some \( w \in C \) and so \( y = f^{-1}(f(w)) = w \in C \) contrary to \( y \in K \), a component of \( \mathbb{R}^n \setminus C \).

Now every set of \( \mathcal{L} \) is contained in some set of \( \mathcal{H} \). What about those sets of \( \mathcal{H} \) which contain no set of \( \mathcal{L} \) so that \( \mathcal{L}_H = \emptyset \)? From 8.6.15 it follows \( \mathcal{H} \subseteq f(C) \).

Therefore, \( d(\overline{f^{-1}}, H, y) = d(f^{-1}, H, y) = 0 \) because \( y \in K \), a component of \( \mathbb{R}^n \setminus C \). Therefore, letting \( \mathcal{H}_1 \) denote those sets of \( \mathcal{H} \) which contain some set of \( \mathcal{L} \), 8.6.14 is of the form

\[
1 = \sum_{H \in \mathcal{H}_1} d(\overline{f}, K, H) d(\overline{f^{-1}}, H, y). 
\]

and it is still a finite sum because the terms in the sum are 0 for all but finitely many \( H \in \mathcal{H}_1 \). I want to expand \( d(\overline{f^{-1}}, H, y) \) as a sum of the form

\[
\sum_{L \in \mathcal{L}_H} d(\overline{f^{-1}}, L, y)
\]

using Lemma 8.6.3. Therefore, I must verify

\[
y \notin \overline{f^{-1}}(\overline{H} \cup \mathcal{L}_H)
\]

but this is just Claim 2. By Lemma 8.6.7 I can write the above sum in place of \( d(\overline{f^{-1}}, H, y) \). Therefore,

\[
1 = \sum_{H \in \mathcal{H}_1} d(\overline{f}, K, H) d(\overline{f^{-1}}, H, y) = \sum_{H \in \mathcal{H}_1} d(\overline{f}, K, H) \sum_{L \in \mathcal{L}_H} d(\overline{f^{-1}}, L, y)
\]

where there are only finitely many \( H \) which give a nonzero term and for each of these, there are only finitely many \( L \) in \( \mathcal{L}_H \) which yield \( d(\overline{f^{-1}}, L, y) \neq 0 \). Now the above equals

\[
= \sum_{H \in \mathcal{H}_1} \sum_{L \in \mathcal{L}_H} d(\overline{f}, K, H) d(\overline{f^{-1}}, L, y). \tag{8.6.16}
\]

By definition,

\[
d(\overline{f}, K, H) = d(\overline{f}, K, x)
\]

where \( x \) is any point of \( H \). In particular \( d(\overline{f}, K, H) = d(\overline{f}, K, L) \) for any \( L \in \mathcal{L}_H \). Therefore, the above reduces to

\[
= \sum_{L \in \mathcal{L}} d(\overline{f}, K, L) d(\overline{f^{-1}}, L, y). \tag{8.6.17}
\]
Here is why. There are finitely many \( H \in \mathcal{H} \) for which the term in the double sum of 8.6.16 is not zero, say \( H_1, \ldots, H_m \). Then the above sum in 8.6.17 equals

\[
\sum_{k=1}^{m} \sum_{L \in \mathcal{L}_{H_k}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, y) + \sum_{L \in \mathcal{L} \setminus \bigcup_{k=1}^{m} \mathcal{L}_{H_k}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, y)
\]

The second sum equals 0 because those \( L \) are contained in some \( H \in \mathcal{H} \) for which

\[
0 = \sum_{L \in \mathcal{L}_{H}} d(\bar{f}, K, H) d(\bar{f}^{-1}, H, y) = d(\bar{f}, K, H) \sum_{L \in \mathcal{L}_{H}} d(\bar{f}^{-1}, L, y)
\]

\[
= \sum_{L \in \mathcal{L}_{H}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, y).
\]

Therefore, the sum in 8.6.17 reduces to

\[
\sum_{k=1}^{m} \sum_{L \in \mathcal{L}_{H_k}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, y)
\]

which is the same as the sum in 8.6.16. Therefore, 8.6.17 does follow. Then the sum in 8.6.17 reduces to

\[
\sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K)
\]

and all but finitely many terms in the sum are 0.

By the same argument,

\[
1 = \sum_{K \in \mathcal{K}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K)
\]

and all but finitely many terms in the sum are 0. Letting \(|\mathcal{K}|\) denote the number of elements in \( \mathcal{K} \), similar for \( \mathcal{L} \),

\[
|\mathcal{K}| = \sum_{K \in \mathcal{K}} 1 = \sum_{K \in \mathcal{K}} \left( \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K) \right)
\]

\[
|\mathcal{L}| = \sum_{L \in \mathcal{L}} 1 = \sum_{L \in \mathcal{L}} \left( \sum_{K \in \mathcal{K}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K) \right)
\]

Suppose \(|\mathcal{K}| < \infty \). Then you can switch the order of summation in the double sum for \(|\mathcal{K}| \) and so

\[
|\mathcal{K}| = \sum_{K \in \mathcal{K}} \left( \sum_{L \in \mathcal{L}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K) \right)
\]

\[
= \sum_{L \in \mathcal{L}} \left( \sum_{K \in \mathcal{K}} d(\bar{f}, K, L) d(\bar{f}^{-1}, L, K) \right) = |\mathcal{L}|
\]
Let $d\Omega$ and $\Omega$ be an open connected bounded set in $\mathbb{R}^n, n \geq 1$ such that $\mathbb{R}^n \setminus \partial \Omega$ consists of two, three if $n = 1$, connected components. Let $f \in C(\Omega; \mathbb{R}^n)$ be continuous and one to one. Then $f(\Omega)$ is the bounded component of $\mathbb{R}^n \setminus f(\partial \Omega)$ and for $y \in f(\Omega)$, $d(f, \Omega, y)$ either equals 1 or $-1$.

**Proof:** First suppose $n \geq 2$. By the Jordan separation theorem, $\mathbb{R}^n \setminus f(\partial \Omega)$ consists of two components, a bounded component $B$ and an unbounded component $U$. Using the Tietze extension theorem, there exists $g$ defined on $\mathbb{R}^n$ such that $g = f^{-1}$ on $f(\Omega)$. Thus on $\partial \Omega$, $g \circ f = I$. It follows from this and the product formula that

$$1 = d(I, \Omega, g(y)) = d(g \circ f, \Omega, g(y)) = d(g, B, g(y)) d(f, \Omega, B) + d(f, \Omega, U) d(g, U, g(y)) = d(g, B, g(y)) d(f, \Omega, B)$$

The reduction happens because $d(f, \Omega, U) = 0$ as explained above. Since $U$ is unbounded, there are points in $U$ which cannot be in the compact set $f(\Omega)$. For such, the degree is 0 but the degree is constant on $U$, one of the components of $f(\partial \Omega)$. Therefore, $d(f, \Omega, B) \neq 0$ and so for every $z \in B$, it follows $z \in f(\Omega)$. Thus $B \subseteq f(\Omega)$. On the other hand, $f(\Omega)$ cannot have points in both $U$ and $B$ because it is a connected set. Therefore $f(\Omega) \subseteq B$ and this shows $B = f(\Omega)$. Thus $d(f, \Omega, B) = d(f, \Omega, y)$ for each $y \in B$ and the above formula shows this equals either 1 or $-1$ because the degree is an integer. In the case where $n = 1$, the argument is similar but here you have 3 components in $\mathbb{R}^1 \setminus f(\partial \Omega)$ so there are more terms in the above sum although two of them give 0.

### 8.7 A Function With Values In Smaller Dimensions

Recall that we have the degree defined $d(f, \Omega, y)$ for continuous functions on $\overline{\Omega}$ and $y \notin f(\partial \Omega)$. It had properties as follows.

1. $d(I, \Omega, y) = 1$ if $y \in \Omega$.

2. If $\Omega_i \subseteq \Omega, \Omega_1 \cap \Omega_2 = \emptyset$ and if $y \notin f(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2))$, then $d(f, \Omega_1, y) + d(f, \Omega_2, y) = d(f, \Omega, y)$. 

3. If \( y \notin f(\Omega \setminus \Omega_1) \) and \( \Omega_1 \) is an open subset of \( \Omega \), then
\[
d(f, \Omega, y) = d(f, \Omega_1, y).
\]

4. For \( y \in \mathbb{R}^n \setminus f(\partial \Omega) \), if \( d(f, \Omega, y) \neq 0 \) then \( f^{-1}(y) \cap \Omega \neq \emptyset \).

5. If \( t \to y(t) \) is continuous \( h: \overline{\Omega} \times [0,1] \to \mathbb{R}^n \) is continuous and if \( \partial \Omega \neq \emptyset \), then \( t \to d(h(\cdot, t), \Omega, y(t)) \) is constant.

6. \( d(\cdot, \Omega, y) \) is defined and constant on
\[
\{ g \in C(\overline{\Omega}; \mathbb{R}^n) : \| g - f \|_\infty < r \}
\]
where \( r = \text{dist}(y, f(\partial \Omega)) \).

7. \( d(f, \Omega, \cdot) \) is constant on every connected component of \( \mathbb{R}^n \setminus f(\partial \Omega) \).

8. \( d(g, \Omega, y) = d(f, \Omega, y) \) if \( g_{|\partial \Omega} = f_{|\partial \Omega} \).

**Theorem 8.7.1** Let \( \Omega \) be a bounded open set in \( \mathbb{R}^n \) and let \( f \in C(\overline{\Omega}; \mathbb{R}^m) \) where \( \mathbb{R}^m = \{ x \in \mathbb{R}^n : x_k = 0 \text{ for } k > m \} \). Thus \( x \) concludes with a column of \( n-m \) zeros. Let \( y \in \mathbb{R}^m \setminus (I-f)(\partial \Omega) \). Then \( d(I-f, \Omega, y) = d(I-f)_{|\Omega \cap \Omega'} \Omega \cap \mathbb{R}^m, y \) \).

**Proof:** To save space, let \( g = I-f \). Then there is no loss of generality in assuming at the outset that \( y \) is a regular value for \( g \). Indeed, everything above was reduced to this case. Then for \( x \in g^{-1}(y) \) and letting \( x_m \) be the first \( m \) variables for \( x \),
\[
Dg(x) = \begin{pmatrix}
D_{x_m} g(x) \\
0 \\
I_{n-m}
\end{pmatrix}
\]
Then it follows that
\[
0 \neq \det(Dg(x)) = \det(Dx_m^* g(x))
\]
\[
= \det(Dx_m g(x)) = d(I-f)_{|\Omega \cap \Omega'} \Omega \cap \mathbb{R}^m, y \)
\]
This last is just the determinant of the derivative of the function which results from restricting \( g \) to the first \( m \) variables. Now \( y \in \mathbb{R}^m \) and \( f \) also is given to have values in \( \mathbb{R}^m \), so if \( g(x) = y \), then you have \( x - f(x) = y \) which requires \( x \in \mathbb{R}^n \) also. Therefore, \( g^{-1}(y) \) consists of points in \( \mathbb{R}^n \) only. Thus, \( y \) is also a regular value of the function which results from restricting \( g \) to \( \mathbb{R}^m \cap \Omega \).
\[
d(I-f, \Omega, y) = d(g, \Omega, y)
\]
\[
= \sum_{x \in g^{-1}(y)} \text{sign}(\det(Dg(x)))
= \sum_{x \in g^{-1}(y)} \text{sign}(\det(Dx_m g(x))) = d(I-f)_{|\Omega \cap \Omega'} \Omega \cap \mathbb{R}^m, y \)
\]
\]
Recall that for \( g \in C^\infty (\overline{\Omega}; \mathbb{R}^n) \),
\[
d (g, \Omega, y) \equiv \lim_{\varepsilon \to 0} \int_{\Omega} \phi_\varepsilon (g (x) - y) \det Dg (x) \, dx
\]

In fact, it can be shown that the degree is unique based on its Properties, 1,2,5 above. It involves reducing to linear maps and then some complicated arguments involving linear algebra. It is done in [10]. Here we will be a little less ambitious. The following lemma will be useful when extending the degree to finite dimensional normed linear spaces and from there to Banach spaces. It is motivated by the following diagram.

\[
\begin{array}{ccc}
\theta^{-1} (y) & \overset{\theta^{-1}}{\leftarrow} & y \\
\uparrow \theta^{-1} \circ g \circ \theta & & \uparrow g \\
\theta^{-1} (\Omega) & \overset{\theta}{\to} & \Omega
\end{array}
\]

**Lemma 8.7.2** Let \( y \notin g (\partial \Omega) \) and let \( \theta \) be an isomorphism of \( \mathbb{R}^n \). That is, \( \theta \) is one to one onto and linear. Then
\[
d (\theta^{-1} \circ g \circ \theta, \theta^{-1} (\Omega), \theta^{-1} y) = d (g, \Omega, y)
\]

**Proof:** It suffices to consider \( g \in C^\infty (\overline{\Omega}; \mathbb{R}^n) \) for which \( y \) is a regular value because you can get such a \( \hat{g} \) with \( \| \hat{g} - g \|_\infty < \delta \) where
\[
B (y, 2\delta) \cap g (\partial \Omega) = \emptyset.
\]

Thus \( B (y, \delta) \cap \hat{g} (\partial \Omega) = \emptyset \) and so \( d (g, \Omega, y) = d (\hat{g}, \Omega, y) \). One can assume similarly that \( \| \hat{g} - g \|_\infty \) is sufficiently small that
\[
d (\theta^{-1} g \circ \theta, \theta^{-1} (\Omega), \theta^{-1} y) = d (\theta^{-1} \hat{g} \circ \theta, \theta^{-1} (\Omega), \theta^{-1} y)
\]
because both \( \theta \) and \( \theta^{-1} \) are continuous. Thus it suffices to consider at the outset \( g \in C^\infty (\overline{\Omega}; \mathbb{R}^n) \). Then from the definition of degree for \( C^\infty \) maps,
\[
d (\theta^{-1} \circ g \circ \theta, \theta^{-1} (\Omega), \theta^{-1} y)
\]
\[
= \lim_{\varepsilon \to 0} \int_{\theta^{-1} \Omega} \phi_\varepsilon ( (\theta^{-1} \circ g \circ \theta) (z) - \theta^{-1} y) \det D (\theta^{-1} \circ g \circ \theta) (z) \, dz
\]

Now \( D (\theta^{-1} \circ g \circ \theta) (z) = \theta^{-1} D (g \circ \theta) (z) = \theta^{-1} Dg (\theta (z)) \theta z \). Changing the variables \( x = \theta z, z = \theta^{-1} x \), this last integral equals
\[
\int_{\Omega} \phi_\varepsilon ( (\theta^{-1} g \circ \theta) (\theta^{-1} (x)) - \theta^{-1} y) \det Dg (x) \, d\theta |\det \theta| |\det \theta^{-1}|^2 \, dx
\]
\[
= \int_{\Omega} \phi_\varepsilon (\theta^{-1} g (x) - \theta^{-1} y) |\det \theta^{-1}| \det Dg (x) \, dx
\]

Recall that \( \phi_\varepsilon \) is a mollifier which is nonzero only in \( B (0, \varepsilon) \). Now
\[
g^{-1} (y) = \{ x_1, \cdots, x_m \} = (\theta^{-1} \hat{g})^{-1} (\theta^{-1} y)
\]
and so $g(x_i) = y$ and $\theta^{-1}g(x_i) = \theta^{-1}y$. By the inverse function theorem, there exist disjoint open sets $U_i$ with $x_i \in U_i$, such that $\theta^{-1}g$ is one to one on $U_i$ with $\det(D(\theta^{-1}g)(x)) = \det(\theta^{-1}) \det Dg(x)$ having constant sign on $U_i$ and $\theta^{-1}g(U_i)$ is an open set containing $\theta^{-1}y$. Then let $\varepsilon$ be small enough that $B(\theta^{-1}y, \varepsilon) \subseteq \cap_{i=1}^m \theta^{-1}g(U_i)$ and let $V_i = (\theta^{-1}g)^{-1}(B(\theta^{-1}y, \varepsilon)) \cap U_i$. Thus for small $\varepsilon$, the $V_i$ are disjoint open sets in $\Omega$ and

$$\int_{\Omega} \phi_{\varepsilon}(\theta^{-1}g(x) - \theta^{-1}y) \det Dg(x) \det \theta^{-1} \, dx = \sum_{i=1}^m \int_{V_i} \phi_{\varepsilon} \det Dg(x) \det \theta^{-1} \, dx$$

Now just let $z = g(x) - y$ and change the variables.

$$= \sum_{i=1}^m \det \theta^{-1} \int_{g(V_i)-y} \phi_{\varepsilon}(\theta^{-1}z) \det Dg(g^{-1}(y+z)) \det Dg^{-1}(y+z) \, dz$$

By the chain rule, $I = Dg(g^{-1}(y+z)) Dg^{-1}(y+z)$ and so

$$\det Dg(g^{-1}(y+z)) \det Dg^{-1}(y+z) = \sgn(\det Dg(g^{-1}(y+z))) \cdot$$

$$\det Dg(g^{-1}(y+z)) \det Dg^{-1}(y+z) = \sgn(\det Dg^{-1}(y+z))$$

and so it all reduces to

$$\sum_{i=1}^m \sgn(\det Dg(x_i)) \int_{g(V_i)-y} \phi_{\varepsilon}(\theta^{-1}z) \, dz = \sum_{i=1}^m \sgn(\det Dg(x_i)) \int_{\theta B(0,\varepsilon)} |\det \theta^{-1}| \phi_{\varepsilon} \theta^{-1}z \, dz$$

$$= \sum_{i=1}^m \sgn(\det Dg(x_i)) \int_{B(0,\varepsilon)} \phi_{\varepsilon}(w) |\det \theta^{-1}| \det \theta \, dw$$

$$= \sum_{i=1}^m \sgn(\det Dg(x_i)) = d(g, \Omega, y). \blacksquare$$

What about functions which have values in finite dimensional vector spaces?
Theorem 8.7.3 Let $\Omega$ be an open bounded set in $V$ a real normed $n$ dimensional vector space. Then there exists a topological degree $d(f, \Omega, y)$ for $f \in C(\overline{\Omega}, V)$, $y \notin f(\partial \Omega)$ which satisfies all the properties of the degree for functions having values in $\mathbb{R}^n$ described above,

1. $d(f, \Omega, y) = 1$ if $y \in \Omega$.
2. If $\Omega_1 \subseteq \Omega, \Omega_2$ open, and $\Omega_1 \cap \Omega_2 = \emptyset$ and if $y \notin f(\overline{\Omega \setminus (\Omega_1 \cup \Omega_2)})$, then $d(f, \Omega_1, y) + d(f, \Omega_2, y) = d(f, \Omega, y)$.
3. If $y \notin f(\overline{\Omega \setminus \Omega_1})$ and $\Omega_1$ is an open subset of $\Omega$, then
   $$d(f, \Omega, y) = d(f, \Omega_1, y).$$
4. For $y \in \mathbb{R}^n \setminus f(\partial \Omega)$, if $d(f, \Omega, y) \neq 0$ then $f^{-1}(y) \cap \Omega \neq \emptyset$.
5. If $t \rightarrow y(t)$ is continuous $h : \overline{\Omega} \times [0, 1] \rightarrow \mathbb{R}^n$ is continuous and if $y(t) \notin h(\partial \Omega, t)$ for all $t$, then $t \rightarrow d(h(\cdot, t), \Omega, y(t))$ is constant.
6. $d(\cdot, \Omega, y)$ is defined and constant on
   $$\{g \in C(\overline{\Omega}; \mathbb{R}^n) : \|g - f\|_\infty < r\}$$
   where $r = \text{dist}(y, f(\partial \Omega))$.
7. $d(f, \Omega, \cdot)$ is constant on every connected component of $\mathbb{R}^n \setminus f(\partial \Omega)$.
8. $d(g, \Omega, y) = d(f, \Omega, y)$ if $g|_{\partial \Omega} = f|_{\partial \Omega}$.

Proof: There is an isomorphism $\theta : \mathbb{R}^n \rightarrow V$ which also preserves all topological properties. This follows from the properties of finite dimensional vector spaces. In fact, every algebraic isomorphism is automatically a homeomorphism preserving all topological properties. Then it is pretty easy to see what the degree should be.

$$d(f, \Omega, y) \equiv d(\theta^{-1} \circ f \circ \theta, \theta^{-1}(\Omega), \theta^{-1}y)$$

Then by standard material on finite dimensional vector spaces, the norm on $V$ is equivalent to the norm defined by $|v| \equiv |\theta^{-1}v|_{\mathbb{R}^n}$. Hence all of those properties hold. By Lemma 8.7.2 this definition does not depend on the particular isomorphism used. If $\theta$ is another one, then one would need to verify that

$$d(\theta^{-1} \circ f \circ \theta, \theta^{-1}(\Omega), \theta^{-1}y) = d(\hat{\theta}^{-1} \circ f \circ \hat{\theta}, \hat{\theta}^{-1}(\Omega), \hat{\theta}^{-1}y)$$

However, you could use that lemma to conclude that

$$d(\hat{\theta}^{-1} \circ f \circ \hat{\theta}, \hat{\theta}^{-1}(\Omega), \hat{\theta}^{-1}y) = d(\alpha^{-1} \circ \hat{\theta}^{-1} \circ f \circ \alpha, \alpha^{-1}\hat{\theta}^{-1}(\Omega), \alpha^{-1}\hat{\theta}^{-1}y)$$

where $\alpha$ is such that $\hat{\theta} \circ \alpha = \theta$. Then this verifies the appropriate equation. □

Next one considers what happens when the function $f - I$ has values in a smaller dimensional subspace.
**Theorem 8.7.4** Let $\Omega$ be a bounded open set in $V$ an $n$ dimensional normed linear space and let $f \in C(\Omega; V_m)$ where $V_m$ is an $m$ dimensional subspace. Let $y \in V_m \setminus (I - f)(\partial \Omega)$. Then $d(I - f, \Omega, y) = d((I - f)|_{\Omega \cap V_m}, \Omega \cap V_m, y)$.

**Proof:** Letting $\{v_1, \ldots, v_m\}$ be a basis for $V_m$, let a basis for $V$ be

$\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$

Let $\theta$ be the isomorphism which satisfies $\theta e_i = v_i$ where the $e_i$ denotes the standard basis vectors for $\mathbb{R}^n$. Then from the above,

$$d(I - f, \Omega, y) = d((I - f)|_{\theta^{-1}(\Omega) \cap V_m}, \theta^{-1}(\Omega) \cap \mathbb{R}^m, \theta^{-1}y) \equiv d((I - f)|_{\Omega \cap V_m}, \Omega \cap V_m, y). \blacksquare$$

### 8.8 The Leray Schauder Degree

This is a very important generalization to Banach spaces. It turns out you can define the degree of $I - F$ where $F$ is a compact mapping. To recall what one of these is, here is the definition.

**Definition 8.8.1** Let $\Omega$ be a bounded open set in $X$ a Banach space and let $F : \Omega \to X$ be continuous. Then $F$ is called compact if $F(B)$ is precompact whenever $B$ is bounded. That is, if $\{x_n\}$ is a bounded sequence, then there is a subsequence $\{x_{n_k}\}$ such that $\{F(x_{n_k})\}$ converges.

**Theorem 8.8.2** Let $F : \Omega \to X$ as above be compact. Then for each $\varepsilon > 0$, there exists $F_\varepsilon : \partial \Omega \to X$ such that $F$ has values in a finite dimensional subspace of $X$ and $\sup_{x \in \Omega} \|F_\varepsilon(x) - F(x)\| < \varepsilon$. In addition to this, $(I - F)^{-1}(\text{compact set}) = \text{compact set}$. (This is called “proper”.)

**Proof:** It is known that $\overline{F(\Omega)}$ is compact. Therefore, there is an $\varepsilon$ net for $F(\Omega)$, $\{Fx_k\}_{k=1}^n$ satisfying

$$\overline{F(\Omega)} \subseteq \cup_k B(Fx_k, \varepsilon)$$

Now let

$$\phi_k(Fx) \equiv (\varepsilon - \|Fx - Fx_k\|)^+$$

Thus this is equal to 0 if $\|Fx_k - Fx\| \geq \varepsilon$ and is positive if $\|Fx_k - Fx\| < \varepsilon$. Then consider

$$F_\varepsilon(x) \equiv \sum_{k=1}^n F(x_k) \frac{\phi_k(Fx)}{\sum_i \phi_i(Fx)}$$
8.8. THE LERAY SCHAUDER DEGREE

It clearly has values in span \(\{F_{x_k}\}_{k=1}^\infty\). How close is it to \(F(x)\)? Say \(Fx \in B(Fx_k, \varepsilon)\). Then for such \(x\), \(\|F(x) - F(x_k)\| < \varepsilon\) by definition. Hence

\[
\|F(x) - F_\varepsilon(x)\| = \sum_{k: \|F(x) - Fx_k\| < \varepsilon} \|F(x_k) - F(x)\| \frac{\phi_k(Fx)}{\sum_i \phi_i(Fx)} < \varepsilon \sum_k \phi_k(Fx) = \varepsilon
\]

Of course \(x\) is arbitrary and so

\[
\sup_{x \in \Omega} \|F_\varepsilon(x) - F(x)\| < \varepsilon.
\]

Next consider the second claim. Let \(K\) be compact. Consider \(\{x_k\} \subseteq (I - F)^{-1}(K)\). It is necessary to show that it has a convergent subsequence. Then \(\{(I - F)(x_k)\}\) is a sequence in \(K\) and so it has a convergent subsequence still denoted with subscript \(k\) such that \((I - F)(x_k) \to y\). The \(x_k\) are in a bounded set \(\Omega\) and so, from compactness of \(F\), there is a further subsequence, still denoted with subscript \(k\) such that \(Fx_k \to z\). It follows that \(x_k \to y - z\) and hence every sequence in \((I - F)^{-1}(K)\) has a convergent subsequence.

**Corollary 8.8.3** Let \(F : \overline{\Omega} \to X\) as above be compact. Then for each \(\varepsilon > 0\), there exists \(F_\varepsilon : \overline{\Omega} \to X\) such that \(F_\varepsilon\) has values in a finite dimensional subspace of \(X\) and \(\sup_{x \in \Omega} \|F_\varepsilon(x) - F(x)\| < \varepsilon\). In addition to this, \((I - F)^{-1}\) (compact set) = compact set. (This is called “proper”.) If \(\Omega\) is symmetric and \(F\) is odd \((F(-x) = -F(x))\), then one can also assume \(F_\varepsilon\) is also odd.

**Proof:** Suppose \(\Omega\) is symmetric in that \(x \in \Omega\) iff \(-x \in \Omega\). Suppose also that \(F\) is odd. Thus \(F(\Omega)\) is also symmetric. Thus \(\overline{F(\Omega)}\) is compact and symmetric. If \(y \in F(\Omega)\), then \(y = Fx\) and so \(-y = -F(x) = F(-x) \in F(\Omega)\). Choose the \(\varepsilon\) net to be symmetric. That is, you have \((Fx)_k\) in the net if and only if \(- (Fx)_k\) is in the net. Just add them in if needed. Therefore, there is an \(\varepsilon\) net for \(F(\Omega)\), \((Fx)_k\) satisfying

\[
\overline{F(\Omega)} \subseteq \bigcup_k B(Fx_k, \varepsilon), \{Fx_k\} \text{ is symmetric.}
\]

Number these so that

\[
Fx_{-k} = -Fx_k = F(-x_k), |k| \leq m_\varepsilon.
\]

Now let

\[
\phi_k(Fx) \equiv (\varepsilon - \|Fx - Fx_k\|)^+.
\]
\[ \phi_{-k}(Fx) = (\varepsilon - ||Fx - Fx_k||)^+ \]
\[ = (\varepsilon - ||Fx + Fx_k||)^+ \]
\[ = (\varepsilon - ||Fx - (-Fx_k)||)^+ \]
\[ = \phi_k(-Fx) \]

that is, \( \phi_{-k} \) is centered at \(-Fx_k\) while \( \phi_k \) is centered at \( Fx_k \), each function equal to 0 off \( B(Fx_k, \varepsilon)\) and is positive on \( B(Fx_k, \varepsilon) \). Then consider

\[ F_x(x) = \sum_{k=-m_x}^{m_x} F(x_k) \frac{\phi_k(Fx)}{\sum_i \phi_i(Fx)} \]

\[ F_x(-x) = -\sum_{k=-m_x}^{m_x} F(-x_k) \frac{\phi_k(F(-x))}{\sum_i \phi_i(F(-x))} = -\sum_{k=-m_x}^{m_x} F(x_{-k}) \frac{\phi_k(F(x))}{\sum_i \phi_i(F(x))} \]

The rest of the argument is the same.

Now let \( F: \Omega \to X \) be compact and consider \( I - F \). Is \( (I - F)(\partial\Omega) \) closed? Suppose \( (I - F)x_k \to y \). Then \( K \equiv y \cup \{(I - F)x_k\}_{k=1}^\infty \) is a compact set because if you have any open cover, one of the open sets contains \( y \) and hence it contains all \( (I - F)x_k \) except for finitely many which can then be covered by finitely many open sets in the open cover. Hence, since \( (I - F) \) is proper, \( (I - F)^{-1}(K) \) is compact. It follows that there is a subsequence, still called \( x_k \) such that \( x_k \to x \in (I - F)^{-1}(K) \). Then by continuity of \( F \),

\[ (I - F)(x_k) \to (I - F)(x) \]
\[ (I - F)(x_k) \to y \]

It follows \( y = (I - F)x \) and so in fact \( (I - F)(\partial\Omega) \) is closed.

**Lemma 8.8.4** If \( F: \Omega \to X \) is compact and \( \Omega \) is a bounded open set in \( X \), then \( (I - F)(\partial\Omega) \) is closed.

**Justification for definition of Leray Schauder Degree**

Now let \( y \notin (I - F)(\partial\Omega) \), a closed set. Hence dist \((y, (I - F)(\partial\Omega)) > 4\delta > 0\). Now let \( F_k \) be a sequence of approximations to \( F \) which have values in an increasing sequence of finite dimensional subsets \( V_k \) each of which contains \( y \). Thus \( \lim_{k \to \infty} \sup_{x \in \Omega} ||F(x) - F_k(x)|| = 0 \). Consider

\[ d(I - F_k|_{V_k}, \Omega \cap V_k, y) \]
Each of these is a well defined integer according to Theorem 8.7.3. For all $k$ large enough,
\[ \sup_{x \in \mathbb{R}} \| (I - F_k)(x) - (I - F_k)(x) \| < \delta \]
Hence, for all such $k$,
\[ B(y, 3\delta) \cap (I - F_k)(\partial \Omega) = \emptyset, \text{ that is } \dist(y, (I - F_k)(\partial \Omega)) > 3\delta \quad (8.8.18) \]
Note that this implies
\[ \dist (y, (I - F_k)(\partial (\Omega \cap V))) > 3\delta \]
for any subspace $V$. If $k < l$ are two such indices, then consider
\[ d(I - F_k|V_k, \Omega \cap V_k, y), d(I - F_l|V_l, \Omega \cap V_l, y) \]
Are they equal? Let $V = V_k + V_l$. Then by Theorem 8.7.4,
\[ d(I - F_l|V_l, \Omega \cap V_l, y) = d(I - F_k|V_k, \Omega \cap V_k, y) \]
So what about $d(I - F_l|V, \Omega \cap V, y), d(I - F_k|V, \Omega \cap V, y)$? Are these equal?
\[ \sup_{x \in \mathbb{R}^n} \| F_l(x) - F_k(x) \| \leq \sup_{x \in \mathbb{R}^n} \| F_l(x) - F(x) \| + \sup_{x \in \mathbb{R}^n} \| F(x) - F_k(x) \| < 2\delta \]
This implies for
\[ h(x, t) = t (I - F_l)(x) + (1 - t) (I - F_k)(x), \]
and $x \in \overline{\Omega \cap V}, y \notin h(\partial (\Omega \cap V), t)$ for all $t \in [0, 1]$. To see this, let $x \in \partial \Omega$
\[ \| t (I - F_l)(x) + (1 - t) (I - F_k)(x) - y \| = \| t (I - F_k)(x) + t (F_k x - F_l x) + (1 - t) (I - F_k)(x) - y \| = \| (I - F_k)(x) + t (F_k x - F_l x) \| \geq 3\delta - t2\delta \geq \delta \]
Hence
\[ d(I - F_l|V, \Omega \cap V, y) = d(I - F_k|V, \Omega \cap V, y) \]
and so
\[ \lim_{k \to \infty} d(I - F_k|V_k, \Omega \cap V_k, y) \]
exists. A similar argument shows that this limit is independent of the sequence $\{F_k\}$ of approximating functions having values in a finite dimensional space. Thus we have the following definition of the Leray Schauder degree.

8.8. THE LERAY SCHAUDER DEGREE

277
Definition 8.8.5 Let \( X \) be a Banach space and let \( F : X \to X \) be compact. That is, \( F(\Omega) \) is precompact whenever \( \Omega \) is bounded. Let \( \Omega \) be a bounded open set in \( X \) and let \( y \notin (I - F)(\partial \Omega) \). Let \( F_k \) be a sequence of operators which have values in finite dimensional spaces \( V_k \) such that \( V_k \subseteq V_{k+1} \cdots, y \in V_k \), and \( \lim_{k \to \infty} \sup_{x \in \Omega} \| F(x) - F_k(x) \| = 0 \). Then
\[
D(I - F, \Omega, y) = \lim_{k \to \infty} d(I - F_k|_{V_k}, \Omega \cap V_k, y)
\]
In fact, the sequence on the right is eventually constant. So
\[
D(I - F, \Omega, y) = d(I - F_k|_{V_k}, \Omega \cap V_k, y)
\]
for all \( k \) sufficiently large.

The main properties of the Leray Schauder degree follow from the corresponding properties of Brouwer degree.

Theorem 8.8.6 Let \( D \) be the Leray Schauder degree just defined and let \( \Omega \) be a bounded open set \( y \notin (I - F)(\partial \Omega) \) where \( F \) is always a compact mapping. Then the following properties hold:

1. \( D(I, \Omega, y) = 1 \)

2. If \( \Omega_i \subseteq \Omega \) where \( \Omega_i \) is open, \( \Omega_1 \cap \Omega_2 = \emptyset \), and \( y \notin \overline{\Omega_i \cup \Omega_j} \) then
\[
D(I - F, \Omega, y) = D(I - F, \Omega_1, y) + D(I - F, \Omega_2, y)
\]

3. If \( t \to y(t) \) is continuous \( h : \overline{\Omega} \times [0, 1] \to X \) is continuous, \( (x, t) \to h(x, t) \) is compact, (It takes bounded subsets of \( \overline{\Omega} \times [0, 1] \) to precompact sets in \( X \)) and if \( y(t) \notin (I - h)(\partial \Omega, t) \) for all \( t \), then \( t \to D((I - h)(\cdot, t), \Omega, y(t)) \) is constant.

Proof: The mapping \( x \to 0 \) is clearly compact. Then an approximating sequence is \( F_k, F_k x = 0 \) for all \( k \). Then
\[
D(I, \Omega, y) = \lim_{k \to \infty} d(I|_{V_k}, \Omega \cap V_k, y) = 1
\]

For the second part, let \( k \) be large enough that for \( U = \Omega, \Omega_1, \Omega_2 \),
\[
D(I - F, U, y) = d(I - F_k|_{V_k}, U \cap V_k, y)
\]
where \( F_k \) is the sequence of approximating functions having finite dimensional range. Then the result follows from the Brouwer degree. In fact,
\[
D(I - F, \Omega, y) = d(I - F_k|_{V_k}, \Omega \cap V_k, y)
= d(I - F_k|_{V_k}, \Omega_1 \cap V_k, y) + d(I - F_k|_{V_k}, \Omega_2 \cap V_k, y)
= D(I - F, \Omega_1, y) + D(I - F, \Omega_2, y)
\]
this does the second claim of the theorem. Now consider the third one about homotopy invariance.

Claim: If dist \((y, (I - F) \partial \Omega) \geq 6\delta\), and if \(\|y - z\| < \delta\), then

\[
D (I - F, \Omega, y) = D (I - F, \Omega, z)
\]

Proof of claim: Let \(F_k\) be the approximations and include both \(y, z\) in all the finite dimensional subspaces \(V_k\). Then for \(k\) large enough, \(\sup_{x \in \overline{T}} \|F(x) - F_k(x)\| < \delta\) and also,

\[
D (I - F, \Omega, y) = d ((I - F_k)|_{V_k}, \Omega \cap V_k, y)
\]

\[
D (I - F, \Omega, z) = d ((I - F_k)|_{V_k}, \Omega \cap V_k, z)
\]

Now for \(x \in \partial (\Omega \cap V_k)\),

\[
\|(I - F_k)(x) - y\| \geq \|(I - F)(x) + (F(x) - F_k(x)) - y\|
\]

\[
\geq \|(I - F)(x) - y\| - \|F(x) - F_k(x)\| > 6\delta - \delta = 5\delta
\]

Hence dist \((y, (I - F_k) \partial \Omega) \geq 5\delta\) while \(\|y - z\| < \delta\). Hence

\[
d ((I - F_k)|_{V_k}, \Omega \cap V_k, y) = d ((I - F_k)|_{V_k}, \Omega \cap V_k, z)
\]

by Theorem 5.3.3.

From compactness of \(h\), there is an \(\varepsilon\) net for \(h (\overline{\Omega} \times [0, 1]), \{h (x_k, t_k)\}\) such that

\[
h (\overline{\Omega} \times [0, T]) \subseteq \bigcup_{k=1}^{n} B (h (x_k, t_k), \varepsilon).
\]

Say the \(t_k\) are ordered. Then, as before,

\[
\phi_k (x) \equiv (\varepsilon - \|h (x, t) - h (x, t_k)\|)^+
\]

\[
h_k (x, t) \equiv \sum_{k=1}^{n} h (x_k, t_k) \phi_k (h (x, t)) \sum_{i} \phi_i (h (x, t))
\]

Then this is clearly continuous and has values in span \(\{h (x_k, t_k)\}_{k=1}^{n}\). How well does it approximate? Say \(h (x, t) \in h (\overline{\Omega} \times [0, T])\). Then it is in some

\[
B (h (x_k, t_k), \varepsilon),
\]

maybe several. Thus letting \(K (x, t)\) be those indices \(k\) such that

\[
h (x, t) \in B (h (x_k, t_k), \varepsilon)
\]

\[
\|h_e (x, t) - h (x, t)\| \leq \sum_{k \in K (x, t)} \|h (x_k, t_k) - h (x, t)\| \frac{\phi_k (h (x, t))}{\sum_i \phi_i (h (x, t))}
\]

\[
\leq \varepsilon \sum_{k=1}^{n} \frac{\phi_k (h (x, t))}{\sum_i \phi_i (h (x, t))} = \varepsilon
\]
Now here is a claim.

**Claim:** There exists \( \delta > 0 \) such that for all \( t \in [0,1] \),

\[
\text{dist} (y(t), (I - h) (\partial \Omega, t)) > 6\delta
\]

**Proof of claim:** If not, there is \( (x_n, t_n) \in \partial \Omega \times [0,1] \) such that

\[
\| y(t_n) - (I - h)(x_n, t_n) \| < 1/n
\]

Then \( h(x_n, t_n) \) is in a compact set because of compactness of \( h \). Also, the \( y(t_n) \) are in a compact set because \( y \) is continuous and \( y([0,T]) \) must therefore be compact. It follows that \( (x_n, t_n) \) must be in a compact subset of \( \partial \Omega \times [0,1] \). It follows there is a subsequence, still denoted as \( (x_n, t_n) \) which converges to \( (x, t) \) in \( \partial \Omega \times [0,1] \), then by continuity, \( \| y(t) - (I - h)(x, t) \| = 0 \) contrary to assumption. This proves the claim.

As with \( h \) there exists a sequence \( \{ y_k(t) \} \) such that \( y_k(t) \rightharpoonup y(t) \) uniformly in \( t \in [0,1] \) but \( y_k \) has values in a finite dimensional subspace of \( X, Y_k \). Choose \( k_0 \) large enough that for all \( t \in [0,1] \), \( \| y(t) - y_{k_0}(t) \| < \delta \). Thus by the first claim,

\[
D(h(\cdot, t), \Omega, y(t)) = D(h(\cdot, t), \Omega, y_{k_0}(t))
\]

for all \( t \). Also,

\[
\text{dist} (y_0(t), (I - h) (\partial \Omega, t)) > 5\delta
\]

From the above, let \( h_k \to h \) uniformly on \( \Omega \times [0,1] \) but \( h_k \) has values in a finite dimensional subspace \( V_k \). Let all the \( V_k \) contain the values of \( y_{k_0} \) and so, for all \( k \) large enough,

\[
\sup_{\Omega \times [0,T]} \| h(t,x) - h_k(t,x) \| < \delta
\]

so for such \( k \),

\[
\text{dist} (y_0(t), (I - h_k) (\partial (\Omega \cap V_k), t))
\geq \text{dist} (y_0(t), (I - h_k) (\partial \Omega, t)) > 4\delta
\]

Then

\[
D(h(\cdot, t), \Omega, y(t)) = D(h(\cdot, t), \Omega, y_{k_0}(t))
= \lim_{k \to \infty} d(h_k(\cdot, t) |_{\Omega \cap V_k}, (\Omega \cap V_k, y_{k_0}(t))
\]

and \( d(h_k(\cdot, t) |_{\Omega \cap V_k}, (\Omega \cap V_k, y_{k_0}(t)) \) is constant in \( t \) for all large enough \( k \). Thus

\[
D(h(\cdot, t), \Omega, y(t)) = \lim_{k \to \infty} a_k, \ a_k \text{ independent of } t. \blacksquare
\]

One of the nice results which follows right away from this is the Schauder fixed point theorem.

**Theorem 8.8.7** Let \( B = \overline{B(0,r)} \) and let \( F : B \to B \) be compact. Then \( F \) has a fixed point.
8.8. THE LERAY-SCHAUDER DEGREE

Proof: Suppose it does not. Then consider $D(I - tF, B(0, r), 0)$. If $t = 1$, $0 \notin (I - tF)(\partial B)$ since otherwise, there would be a fixed point. If $t < 1$ there is no point of $\partial B$ which $I - tF$ sends to 0 because if so,

$$x - tFx = 0, \|x\| = 1, \|Fx\| \leq 1.$$  

Therefore, by homotopy invariance, $t \to D(I - tF, B(0, r), 0)$ is constant for $t \in [0, 1]$. It must equal $D(I - F, B(0, r), 0) = D(I, B(0, r), 0) = 1$.

Therefore, there exists $x \in B(0, r)$ such that $(I - F)(x) = 0$ so $F$ which means $F$ has a fixed point after all. □

One can get an improved version of this easily.

Theorem 8.8.8 Let $K$ be a closed bounded convex subset of a Banach space $X$ and suppose $F : K \to K$ is compact. Then $F$ has a fixed point.

Proof: By Theorem 4.2.5, $K$ is a retract. Thus there is a continuous function $R : X \to K$ which leaves points of $K$ unchanged. Then you consider $F \circ R$. It is still a compact mapping obviously. Let $B(0, r)$ be so large that it contains $K$. Then from the above theorem, it has a fixed point in $B(0, r)$ denoted as $x$. Then $F(R(x)) = x$. But $F(R(x)) \in K$ and so $x \in K$. Hence $Rx = x$ and so $Fx = x$. □

There is an easy modification of the above which is often useful. If $F : X \to F(X)$ where $F(X)$ is bounded and in a compact set, and $F$ is a compact map, then you could consider $F : \text{conv}(F(X)) \to F(X) \subseteq \text{conv}(F(X))$ where here $\text{conv}(F(X))$ is a closed bounded convex subset of $X$. Then by the Schauder theorem, there is a fixed point for $F$.

Here is an easy application of this theorem to ordinary differential equations.

Theorem 8.8.9 Let $g : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous. Let $F : C([0, T]; \mathbb{R}^n) \to C([0, T]; \mathbb{R}^n)$ be given by

$$F(y)(t) = y_0 + \int_0^t g(s, y(s)) \, ds$$

Suppose that whenever

$$y(s) = F(y)(s), \text{ for } s \leq t,$$

it follows that $\max_{s \in [0, t]} |y(s)| < M$, $|y_0| < M$. Then there exists a solution to the integral equation

$$y(t) = y_0 + \int_0^t g(s, y(s)) \, ds$$

for $t \in [0, T]$. 
There is a fixed point for $F$ for every $x$. Let \[ M \] be the radial projection in $\mathbb{R}^n$ onto $B(0, M)$. Then $F \circ r_M$ is compact because $|g(s, r_M y)|$ is bounded. It also maps to a compact subset of $C([0, T]; \mathbb{R}^n)$ thanks to the Arzela Ascoli theorem. Then by the Schauder fixed point theorem, there exists a solution $y = F \circ r_M$ to

\[ y(t) = y_0 + \int_0^t g(s, r_M y(s)) \, ds \]

Then for $s \in [0, \hat{T}]$ where $\hat{T}$ is the largest such that $\|y(s)\| \leq M$ for $s \in [0, \hat{T}]$. Thus on $[0, \hat{T}], r_M$ has no effect. If $\hat{T} < T$, then by the estimate, $\|y(\hat{T})\| < M$. Hence $\hat{T}$ is not really the last. Thus $\hat{T} = T$. ■

Here is the Schauder alternative theorem, also called the Schaefer fixed point theorem $\Omega$.

**Theorem 8.8.10** Let $f : X \to X$ be a compact map. Then either

1. There is a fixed point for $tf$ for all $t \in [0, 1]$ or
2. For every $r > 0$, there exists a solution to $x = tf(x)$ for $t \in (0, 1)$ such that $\|x\| > r$.

**Proof:** Suppose there is $t_0 \in [0, 1]$ such that $t_0f$ has no fixed point. Then $t_0 \neq 0, t_0f$ obviously has a fixed point if $t_0 = 0$. Thus $t_0 \in (0, 1]$. Then let $r_M$ be the radial retraction onto $B(0, M)$. By Schauder’s theorem there exists $x \in B(0, M)$ such that $t_0 r_M f(x) = x$. Then if $\|f(x)\| \leq M$, $r_M$ has no effect and so $t_0 f(x) = x$ which is assumed not to take place. Hence $\|f(x)\| > M$ and so $\|r_M f(x)\| = M$ so $\|x\| = t_0 M$. Also $t_0 r_M f(x) = t_0 M \frac{f(x)}{\|f(x)\|} = x$ and so $x = tf(x), \hat{t} = t_0 \frac{M}{\|f(x)\|} < 1$. Since $M$ is arbitrary, it follows that the solutions to $x = tf(x)$ for $t \in (0, 1)$ are unbounded. It was just shown that there is a solution to $x = tf(x), \hat{t} < 1$ such that $\|x\| = t_0 M$ where $M$ is arbitrary. Thus the second of the two alternatives holds. ■

There is a lot more on degree theory in $\Omega$. Here is a very interesting theorem from this reference which pertains specifically to infinite dimensional spaces.

**Theorem 8.8.11** Let $X$ be an infinite dimensional Banach space and let $0 \notin \partial \Omega$ where $\Omega$ is an open bounded subset of $X$. Let $F : \Omega \to X$ be compact. Suppose that for all $x \in \partial \Omega$, $F x \neq \lambda x$ for $\lambda \in [0, 1]$ and that $0 \notin F(\partial \Omega)$. Then $D (I - F, \Omega, 0) = 0$.

**Proof:** Recall that $D (I - F, \Omega, 0) = \lim_{k \to \infty} d (I - F_k, \Omega \cap V_k, 0)$ where $F_k$ has values in a finite dimensional subspace $V_k$,

\[ \sup_{x \in \Omega} \|F_k(x) - F(x)\| < 1/k \]

Since the dimension of $X$ is infinite, it can always be assumed that span $(-F_k(\Omega))$ is a proper subspace of $V_k$ and this will be assumed. This is where it is significant that the dimension of $X$ is infinite. Also recall that in the limit, eventually
d (I - F_k, \Omega \cap V_k, 0) is a constant. Then the fact that \( Fx \neq \lambda x \) for all \( x \in \partial \Omega \) will persist for \( F_k \) for all \( k \) large enough.

If not, then there exists \( x_k \in \partial \Omega \), \( F_k x_k = \lambda_k x_k \) for some \( x_k \in \partial \Omega \) and \( \lambda_k \in [0, 1] \). Then there is a subsequence \( \lambda_k \to \lambda_0 \in [0, 1] \). Then

\[
Fx_k - \lambda_k x_k \to 0
\]

because it is uniformly close to \( F_k x_k - \lambda_k x_k \). Now by assumption \( 0 \notin F(\partial \Omega) \). If \( \lambda_0 = 0 \), then you would have \( Fx_k \to 0 \) which does not happen because \( 0 \) is at a positive distance from \( F(\partial \Omega) \). Hence for all \( k \) large enough,

\[
Fx_k \neq \lambda x
\]

for all \( \lambda \in [0, 1] \). Pick \( k \) sufficiently large that in the limit for the Leray Schauder degree \( d (I - F_k, \Omega \cap V_k, 0) \) remains constant. Then for \( \lambda \in [0, 1] \),

\[
d (\lambda I - F_k, \Omega \cap V_k, 0) = d (-F_k, \Omega \cap V_k, 0) = d (-F_k, \Omega \cap V_k, p)
\]

for all \( p \notin \text{span} (-F_k (\Omega)) \) which is also close enough to 0. Hence, since the degree of this last equals 0 for such \( p \), it follows that

\[
d (I - F_k, \Omega \cap V_k, 0) = 0
\]

Hence \( D (I - F, \Omega, 0) = 0 \) as claimed. □

This theorem implies a very strange fixed point theorem. It is strange because it only applies to infinite dimensions.

**Corollary 8.8.12** Let \( X \) be an infinite dimensional Banach space. Let \( 0 \in \Omega_0 \subseteq \Omega \) be two open sets. Let \( F : \Omega \to X \) be a compact mapping which satisfies

1. \( \|Fx\| \leq \|x\| \) for \( x \in \partial \Omega_0 \)
2. \( \|Fx\| \geq \|x\| \) for \( x \in \partial \Omega \)

Then \( F \) has a fixed point in \( \overline{\Omega \setminus \Omega_0} \).

**Proof:** First note that \( \overline{\Omega \setminus \Omega_0} \) is like an annulus with both edges included. Suppose \( F \) does not have a fixed point in \( \overline{\Omega \setminus \Omega_0} \). What if \( t = 1 \) and \( x \in \partial \Omega_0 \)? Could \( 0 = (I - F) (x) \)? If so, the fixed point is obtained so assume this is not so. Then for \( t < 1 \) and \( x \in \partial \Omega \), if you have \( x = t^{-1} Fx \), this would mean that \( \|Fx\| = t \|x\| \) and \( \|Fx\| < \|x\| \) which is assumed not to happen. Hence for \( t \in (0, 1] \) we can assume that \( 0 \notin (I - t^{-1} F) (\partial \Omega) \). If \( (I - F) (x) = 0 \) for \( x \notin \partial \Omega_0 \) then the fixed point has been found. For \( t \in [0, 1] \), you can’t have \( (I - t F) (x) = 0 \) for \( x \in \partial \Omega_0 \) because then you would have \( \|x\| = t \|Fx\| \) and so \( \|Fx\| > \|x\| \) which is assumed not to happen. Therefore, we can assume that for \( x \in \partial \Omega_0 \), \( (I - t F) (x) \neq 0 \). Therefore, \( D (I - F, \Omega_0, 0) = D (I, \Omega_0, 0) = 1 \) by homotopy invariance. Also from properties of the degree,

\[
D (I - F, \Omega_0, 0) + D (I - F, \Omega \setminus \Omega_0, 0) = D (I - F, \Omega, 0)
\]
Recall that this is true if \(0 \notin (I - F)\left(\overline{\Omega} - \Omega_0\right)\) which is assumed to take place when we assume there is no fixed point. It is desired to use the above theorem so we need to consider \(\overline{F}\left(\partial\Omega\right)\) and whether 0 is in this set. Condition 2 implies 0 is not in this set. Then Theorem 8.8.11 implies that \(D(I - F, \Omega, 0) = 0\) and so \(D\left(I - F, \Omega \setminus \overline{\Omega_0}, 0\right) = -1\). Hence there is a fixed point in \(\Omega \setminus \Omega_0\) after all contrary to the assumption that there was no such thing. \(\blacksquare\)

This only works in infinite dimensions. Consider an annulus in \(\mathbb{R}^2\) and let \(F\) be a rotation through an angle of 30 degrees. It clearly has no fixed point but the above conditions are satisfied. This seems very interesting, something which happens in infinite dimensions but not in finite dimensions.

**8.9 Exercises**

1. Show the Brouwer fixed point theorem is equivalent to the nonexistence of a continuous retraction onto the boundary of \(B(0, r)\).

2. Using the Jordan separation theorem, prove the invariance of domain theorem \(n \geq 2\). Thus an open ball goes to some open. **Hint:** You might consider \(B(x, r)\) and show \(f\) maps the inside to one of two components of \(\mathbb{R}^n \setminus f(\partial B(x, r))\). etc.

3. Give a version of Proposition 8.6.9 which is valid for the case where \(n = 1\).

4. It was shown that if \(f\) is locally one to one and continuous, \(f : \mathbb{R}^n \to \mathbb{R}^n\), and

\[
\lim_{|x| \to \infty} |f(x)| = \infty,
\]

then \(f\) maps \(\mathbb{R}^n\) onto \(\mathbb{R}^n\). Suppose you have \(f : \mathbb{R}^m \to \mathbb{R}^n\) where \(f\) is one to one and \(\lim_{|x| \to \infty} |f(x)| = \infty\). Show that \(f\) cannot be onto.

5. Can there exist a one to one onto continuous map, \(f\) which takes the unit interval \([0, 1]\) to the unit disk \(B(0, 1)\)? **Hint:** Think in terms of invariance of domain.

6. Let \(m < n\) and let \(B_m(0, r)\) be the ball in \(\mathbb{R}^m\) and \(B_n(0, r)\) be the ball in \(\mathbb{R}^n\). Show that there is no one to one continuous map from \(B_m(0, r)\) to \(B_n(0, r)\). **Hint:** It is like the above problem.

7. Consider the unit disk,

\[
\{(x, y) : x^2 + y^2 \leq 1\} \equiv D
\]

and the annulus

\[
\left\{(x, y) : \frac{1}{2} \leq x^2 + y^2 \leq 1\right\} \equiv A
\]

Is it possible there exists a one to one onto continuous map \(f\) such that \(f(D) = A\)? Thus \(D\) has no holes and \(A\) is really like \(D\) but with one hole punched
out. Can you generalize to different numbers of holes? \textbf{Hint:} Consider the invariance of domain theorem. The interior of $D$ would need to be mapped to the interior of $A$. Where do the points of the boundary of $A$ come from? Consider Theorem [???].

8. Suppose $C$ is a compact set in $\mathbb{R}^n$ which has empty interior and $f : C \to \Gamma \subseteq \mathbb{R}^n$ is one to one onto and continuous with continuous inverse. Could $\Gamma$ have nonempty interior? Show also that if $f$ is one to one and onto $\Gamma$ then if it is continuous, so is $f^{-1}$.

9. Let $K$ be a nonempty closed and convex subset of $\mathbb{R}^n$. Recall $K$ is convex means that if $x, y \in K$, then for all $t \in [0, 1]$, $tx + (1-t)y \in K$. Show that if $x \in \mathbb{R}^n$ there exists a unique $z \in K$ such that

$$|x - z| = \min \{|x - y| : y \in K\}.$$  

This $z$ will be denoted as $P_x$. \textbf{Hint:} First note you do not know $K$ is compact. Establish the parallelogram identity if you have not already done so,

$$|u - v|^2 + |u + v|^2 = 2|u|^2 + 2|v|^2.$$

Then let $\{z_k\}$ be a minimizing sequence,

$$\lim_{k \to \infty} |z_k - x|^2 = \inf \{|x - y| : y \in K\} \equiv \lambda.$$

Now using convexity, explain why

$$\left| \frac{z_k - z_m}{2} \right|^2 + \left| x - \frac{z_k + z_m}{2} \right|^2 = 2 \left| \frac{x - z_k}{2} \right|^2 + 2 \left| \frac{x - z_m}{2} \right|^2$$

and then use this to argue $\{z_k\}$ is a Cauchy sequence. Then if $z_i$ works for $i = 1, 2$, consider $(z_1 + z_2)/2$ to get a contradiction.

10. In Problem [???] show that $P_x$ satisfies the following variational inequality.

$$(x-Px) \cdot (y-Px) \leq 0$$

for all $y \in K$. Then show that $|P_x_1 - P_x_2| \leq |x_1 - x_2|$. \textbf{Hint:} For the first part note that if $y \in K$, the function $t \to |x - (P_x + t(y-Px))|^2$ achieves its minimum on $[0, 1]$ at $t = 0$. For the second part,

$$(x_1-Px_1) \cdot (P_x_2-Px_1) \leq 0, (x_2-Px_2) \cdot (P_x_1-Px_2) \leq 0.$$  

Explain why

$$(x_2-Px_2 - (x_1-Px_1)) \cdot (P_x_2-Px_1) \geq 0$$

and then use some manipulations and the Cauchy Schwarz inequality to get the desired inequality. Thus $P$ is called a retraction onto $K$.  

---
11. Establish the Brouwer fixed point theorem for any convex compact set in \( \mathbb{R}^n \).
   **Hint:** If \( K \) is a compact and convex set, let \( R \) be large enough that the closed ball, \( D(0, R) \supseteq K \). Let \( P \) be the projection onto \( K \) as in Problem 10 above. If \( f \) is a continuous map from \( K \) to \( K \), consider \( f \circ P \). You want to show \( f \) has a fixed point in \( K \).

12. Suppose \( D \) is a set which is homeomorphic to \( \overline{B(0,1)} \). This means there exists a continuous one to one map, \( h \) such that \( h\left(B(0,1)\right) = D \) such that \( h^{-1} \) is also one to one. Show that if \( f \) is a continuous function which maps \( D \) to \( D \) then \( f \) has a fixed point. Now show that it suffices to say that \( h \) is one to one and continuous. In this case the continuity of \( h^{-1} \) is automatic. Sets which have the property that continuous functions taking the set to itself have at least one fixed point are said to have the fixed point property. Work Problem 7 using this notion of fixed point property. What about a solid ball and a donut? Could these be homeomorphic?

13. Suppose \( \Omega \) is any open bounded subset of \( \mathbb{R}^n \) which contains \( 0 \) and that \( f : \overline{\Omega} \rightarrow \mathbb{R}^n \) is continuous with the property that \( f(x) \cdot x \geq 0 \) for all \( x \in \partial \Omega \). Show that then there exists \( x \in \overline{\Omega} \) such that \( f(x) = 0 \). Give a similar result in the case where the above inequality is replaced with \( \leq \).
   **Hint:** You might consider the function \( h(t, x) \equiv tf(x) + (1-t)x \).

14. Suppose \( \Omega \) is an open set in \( \mathbb{R}^n \) containing \( 0 \) and suppose that \( f : \overline{\Omega} \rightarrow \mathbb{R}^n \) is continuous and \( |f(x)| \leq |x| \) for all \( x \in \partial \Omega \). Show \( f \) has a fixed point in \( \overline{\Omega} \).
   **Hint:** Consider \( h(t, x) \equiv t(x - f(x)) + (1-t)x \) for \( t \in [0,1] \). If \( t = 1 \) and some \( x \in \partial \Omega \) is sent to \( 0 \), then you are done. Suppose therefore, that no fixed point exists on \( \partial \Omega \). Consider \( t < 1 \) and use the given inequality.

15. Let \( \Omega \) be an open bounded subset of \( \mathbb{R}^n \) and let \( f, g : \overline{\Omega} \rightarrow \mathbb{R}^n \) both be continuous such that

\[
|f(x)| - |g(x)| > 0
\]

for all \( x \in \partial \Omega \). Show that then

\[
d(f - g, \Omega, 0) = d(f, \Omega, 0)
\]

Show that if there exists \( x \in f^{-1}(0) \), then there exists \( x \in (f - g)^{-1}(0) \).
   **Hint:** You might consider \( h(t, x) \equiv (1-t)f(x) + t(f(x) - g(x)) \) and argue \( 0 \notin h(t, \partial \Omega) \) for \( t \in [0,1] \).

16. Suppose \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is continuous and satisfies

\[
|f(x) - f(y)| \geq \alpha |x - y|, \quad \alpha > 0,
\]
Show that \( f \) must map \( \mathbb{R}^n \) onto \( \mathbb{R}^n \). \textbf{Hint:} First show \( f \) is one to one. Then use invariance of domain. Next show, using the inequality, that the points not in \( f(\mathbb{R}^n) \) must form an open set because if \( y \) is such a point, then there can be no sequence \( \{ f(x_n) \} \) converging to it. Finally recall that \( \mathbb{R}^n \) is connected.

It is obvious that \( f \) is one to one. This follows from the inequality. If \( U \) are the points not in the image of \( f \), then \( U \) must be open because if not, then for \( y \) one of these points, there would be a sequence \( f(x_n) \to y \). Then by the inequality, \( \{ x_n \} \) is a Cauchy sequence and so it converges to \( x \). Thus \( f(x) = \lim_{n \to \infty} f(x_n) = y \). Now by invariance of domain, \( f(\mathbb{R}^n) \) is open. However, \( \mathbb{R}^n \) is connected and so in fact, \( U \) is empty.

17. Let \( f : \mathbb{C} \to \mathbb{C} \) where \( \mathbb{C} \) is the field of complex numbers. Thus \( f \) has a real and imaginary part. Letting \( z = x + iy \),

\[
f(z) = u(x, y) + iv(x, y)
\]

Recall that the norm in \( \mathbb{C} \) is given by \( |x + iy| = \sqrt{x^2 + y^2} \) and this is the usual norm in \( \mathbb{R}^2 \) for the ordered pair \((x, y)\). Thus complex valued functions defined on \( \mathbb{C} \) can be considered as \( \mathbb{R}^2 \) valued functions defined on some subset of \( \mathbb{R}^2 \). Such a complex function is said to be analytic if the usual definition holds. That is

\[
f'(z) = \lim_{h \to 0} \frac{f(z + h) - f(z)}{h}.
\]

In other words,

\[
f(z + h) = f(z) + f'(z)h + o(h)
\]

at a point \( z \) where the derivative exists. Let \( f(z) = z^n \) where \( n \) is a positive integer. Thus \( z^n = p(x, y) + iq(x, y) \) for \( p, q \) suitable polynomials in \( x \) and \( y \). Show this function is analytic. Next show that for an analytic function and \( u \) and \( v \) the real and imaginary parts, the Cauchy Riemann equations hold.

\[
u_x = v_y, \quad u_y = -v_x.
\]

In terms of mappings show \( \text{SMIR} \) has the form

\[
\begin{pmatrix}
u(x + h_1, y + h_2) \\
v(x + h_1, y + h_2)
\end{pmatrix}
= \begin{pmatrix}u(x, y) \\
v(x, y)
\end{pmatrix} + \begin{pmatrix}u_x(x, y) & u_y(x, y) \\
v_x(x, y) & v_y(x, y)
\end{pmatrix} \begin{pmatrix}h_1 \\
h_2
\end{pmatrix} + o(h)
\]

\[
= \begin{pmatrix}u(x, y) \\
v(x, y)
\end{pmatrix} + \begin{pmatrix}u_x(x, y) & -v_x(x, y) \\
v_x(x, y) & u_y(x, y)
\end{pmatrix} \begin{pmatrix}h_1 \\
h_2
\end{pmatrix} + o(h)
\]

where \( h = (h_1, h_2)^T \) and \( h \) is given by \( h_1 + ih_2 \). Thus the determinant of the above matrix is always nonnegative. Letting \( B_r \) denote the ball \( B((0, 0), r) = B((0, 0), r) \) show

\[
d(f, B_r, 0) = n.
\]
where \( f(z) = z^n \). In terms of mappings on \( \mathbb{R}^2 \),
\[
f(x, y) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}.
\]

Thus show
\[
d(f, B_r, 0) = n.
\]

**Hint:** You might consider
\[
g(z) \equiv \prod_{j=1}^{\infty} (z - a_j)
\]
where the \( a_j \) are small real distinct numbers and argue that both this function and \( f \) are analytic but that \( 0 \) is a regular value for \( g \) although it is not so for \( f \). However, for each \( a_j \) small but distinct \( d(f, B_r, 0) = d(g, B_r, 0) \).

18. Using Problem 17, prove the fundamental theorem of algebra as follows. Let \( p(z) \) be a nonconstant polynomial of degree \( n 
\[
p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots
\]
Show that for large enough \( r, |p(z)| > |p(z) - a_n z^n| \) for all \( z \in \partial B(0, r) \).

Now from Problem 17 you can conclude \( d(p, B_r, 0) = d(f, B_r, 0) = n \) where \( f(z) = a_n z^n \).

19. The proof of Sard’s lemma made use of the hard Vitali covering theorem. Here is another way to do something similar. Let \( U \) be a bounded open set and let \( f : U \to \mathbb{R}^n \) be in \( C^1(U) \). Let \( S \) denote the set of \( x \in U \) such that \( Df(x) \) has rank less than \( n \). Thus it is a closed set. Let \( U_m = \{ x \in U : \| Df(x) \| \leq m \} \), a closed set. It suffices to show that for \( S_m = U_m \cap S, f(S_m) \) has measure zero because \( f(S) = \bigcup_m f(S_m) \) these sets increasing in \( m \). By definition of differentiability,
\[
\lim_{k \to \infty} \sup_{\|v\| \leq 1/k} \frac{\| f(x + v) - f(x) - Df(x)v \|}{\|v\|} = 0
\]
for each \( x \in U \). Explain why the above function of \( x \) is measurable. Now by Eggoroff’s theorem, there is measurable set \( A \) of measure less than \( m^n \frac{\varepsilon}{m^n |\text{meas}A|} \) such that off \( A \), the convergence is uniform. Let \( C_k \) be a countable union of non overlapping half open rectangles one of which is of the form \( \prod_{i=1}^{\infty} (a_i, b_i] \) such that each has diameter less than \( 2^{-k} \). Consider the half open rectangles which have nonempty intersection with \( S_m \setminus A, I_k \). Then repeat the argument given in the first section of this chapter. Show that for \( k \) large enough, the rank condition and uniform convergence above implies that \( m_n (\cup \{ f(I) : I \in I_k \}) \) is less than \( \varepsilon \). Now show that \( f(A) \) is contained in a set of measure no more than \( m^n 10^n \frac{\varepsilon}{m^n |\text{meas}A|} = 2\varepsilon \). Thus \( f(S_m) \) has measure no more than \( 3\varepsilon \). Since \( \varepsilon \) is arbitrary, this establishes the desired conclusion.
20. Let $X$ be a Banach space and let $\Omega$ be a symmetric and bounded open set. Let $F : \Omega \to X$ be odd and compact $0 \notin (I - F)(\partial \Omega)$. Show using Corollary 8.8.3 that $D(I - F, \Omega, 0)$ is an odd integer.

21. Let $F$ be compact. Suppose $I - F$ is one to one on $B(0, r)$. Then using similar reasoning to the finite dimensional case, show that there is a $\delta > 0$ such that

$$(I - F)(0) + B(0, \delta) \subseteq (I - F)(B(0, r))$$

22. Let $F$ be compact. Suppose $I - F$ is locally one to one on an open set $\Omega$. Show that $(I - F)$ maps open sets to open sets. This is a version of invariance of domain.

23. Suppose $(I - F)$ is locally one to one and $F$ is compact. Suppose also that $\lim_{\|x\| \to \infty} \|F(x)\| / \|x\| = \infty$. Show that $(I - F)$ is onto.

24. As a variation of the above problem, suppose $F : X \to X$ is compact and

$$\lim_{\|x\| \to \infty} \frac{\|F(x)\|}{\|x\|} = 0$$

Then $I - F$ is onto. Note that $I - F$ is not one to one.

25. Suppose $F$ is compact and $\|(I - F)x - (I - F)y\| \geq \alpha \|x - y\|$. Show that then $(I - F)$ is onto.

26. The Jordan curve theorem is: Let $C$ denote the unit circle,

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$ 

Suppose $\gamma : C \to \Gamma \subseteq \mathbb{R}^2$ is one to one onto and continuous. Then $\mathbb{R}^2 \setminus \Gamma$ consists of two components, a bounded component (called the inside) $U_1$ and an unbounded component (called the outside), $U_2$. Also the boundary of each of these two components of $\mathbb{R}^2 \setminus \Gamma$ is $\Gamma$ and $\Gamma$ has empty interior. Using the Jordan separation theorem, prove this important result.

27. This problem is from [20]. Recall Theorem 8.8.11. It allowed you to say that $D(I - F, \Omega, 0) = 0$ provided $0 \notin F(\partial \Omega)$ and $\lambda x \neq Fx$ for all $x \in \partial \Omega$, $\lambda \in [0, 1]$. This was for $F$ compact and defined on an infinite dimensional space $X$. Suppose now that $F$ is compact and $F : \Omega \to X$ where $0 \in \Omega$ an open set in $X$. Suppose also that $F(0) = 0$ and that

$$\lim_{x \to 0} \sup_{r > 0} \inf \left\{ \frac{\|F(x)\|}{\|x\|} : \|x\| \leq r \right\} = \infty$$

Show that there is a sequence $\alpha_n \to 0$ each $\alpha_n \neq 0$, and for some $x_n \neq 0$, $x_n - \alpha_n F(x_n) = 0$. Note that when $\alpha = 0$, there is only one solution to $(I - \alpha F)(x) = 0$, but this says that there are many small $\alpha_n \neq 0$ for which there is a nonzero solution to $(I - \alpha_n F)(x) = 0$. That is there exist arbitrarily
small $\alpha_n$ such that $(I - \alpha_n) F (x_n) = 0$. This says that 0 is a bifurcation point for $I - \alpha F$. **Hint:** Let $\alpha_n \downarrow 0$ and pick $r_n$ such that for all $\|x\| = r_n$,

$$\|\alpha_n F(x)\| > \|x\|$$

Thus $0 \notin \overline{\alpha_n F(\partial B(0,r_n))}$ and also $\alpha_n F(x) \neq \lambda x$ for all $x \in \partial B(0,r_n)$. Use the theorem to conclude that

$$D (I - \alpha_n F, B(0,r_n), 0) = 0$$

and then consider the homotopy $I - \alpha_n t F$. If it sends no point of $\partial B(0, r_n)$ to 0 then you would have

$$D (I - \alpha_n t F, B(0,r_n), 0) = D (I, B(0,r_n), 0) = 1$$

Neat result, but where are the examples?
Chapter 9

Nonlinear Operators In Banach Space

In this chapter, a discussion of various kinds of nonlinear operators. Some standard references on these operators are [17], [18], [8], [9], [6], [41], [54], [10] and references listed there. The most important examples of these operators seem to be due to Brezis in the 1960's and these things have been generalized and used by many others since this time. I am following many of these, but the stuff about maximal monotone operators is mainly from Barbu [6]. I am trying to include all the necessary basic results such as fixed point theorems which are needed to prove the main theorems and also to rewrite in a manner understandable to me.

A word about notation: When \( f \in X' \) and \( x \in X \), it is often the case that \( f(x) \) is written as \( \langle f, x \rangle \) and this convention will be used here.

It seems like the main issue is the following. When does \( \langle f_n, x_n \rangle \) converge to \( \langle f, x \rangle \) given that \( f_n \) and \( x_n \) both converge weakly to \( f \) and \( x \) respectively? There is no problem in finite dimensions because in finite dimensions, there is only one meaning for convergence. However, in infinite dimensions, there certainly is a problem as can be instantly realized by consideration of the Riemann Lebesgue lemma, for example. You know that \( \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx \to 0 \) so \( \sin(nx) \) converges weakly to 0 but \( \int_{-\pi}^{\pi} \sin^2(nx) \, dx \) certainly does not converge to 0.

The idea behind all of these considerations is that \( f_n \) is to come from some nonlinear operator which has properties which will allow one to successfully pass to a limit. When the operator is linear, there usually is no problem because the graph is a subspace and so if it is closed, it will also be weakly closed. Thus, if \( x_n \to x \) weakly and \( Lx_n \to f \) weakly, then \( f = Lx \). However, nothing like this happens with nonlinear operators. Consideration of when this happens is the purpose of this catalogue of nonlinear operators, and also to generalize to set valued operators. First is a section on single valued nonlinear operators and then the case of set valued nonlinear operators is discussed.
9.1 Some Nonlinear Single Valued Operators

Here is an assortment of nonlinear operators which are useful in applications to nonlinear partial differential equations. Generalizations of the notion of a pseudomonotone map will be presented later to include the case of set valued pseudomonotone maps. This is on the single valued version of some of these and these ideas originate with Brezis in the 1960’s. A good description is given in Lions [41].

**Definition 9.1.1** For $V$ a real Banach space, $A : V \rightarrow V'$ is a pseudomonotone map if whenever

$$ u_n \rightharpoonup u $$

and

$$ \limsup_{n \to \infty} \langle Au_n, u_n - u \rangle \leq 0 \quad \text{(9.1.2)} $$

it follows that for all $v \in V$,

$$ \liminf_{n \to \infty} \langle Au_n, u_n - v \rangle \geq \langle Au, u - v \rangle. \quad \text{(9.1.3)} $$

The half arrows denote weak convergence.

If $V$ is finite dimensional, then pseudomonotone maps are continuous. Also the property of being pseudomonotone is preserved when restriction is made to finite dimensional spaces. The notation is explained in the following diagram.

$$ W' \xrightarrow{i^*} V' $$

$$ W \xrightarrow{i} V $$

The map $i$ is just the inclusion map. $iw \equiv w$ and $i^*$ is the usual adjoint map. $\langle i^* f, w \rangle_{W', W} \equiv \langle f, iw \rangle_{V', V} = \langle f, w \rangle_{V' V}$. Thus $i^*Ai(w) \in W'$ and it is defined by

$$ \langle i^*Ai(w), z \rangle_{W', W} \equiv \langle Aw, z \rangle_{V', V} $$

in other words, you restrict $A$ to $W$ and only consider what the resulting functional does to things in $W$.

**Proposition 9.1.2** Let $V$ be finite dimensional and let $A : V \rightarrow V'$ be pseudomonotone and bounded (meaning $A$ maps bounded sets to bounded sets). Then $A$ is continuous. Also, if $A : V \rightarrow V'$ is pseudomonotone and bounded, and if $W \subseteq V$ is a finite dimensional subspace, then $i^*Ai$ is pseudomonotone as a map from $W$ to $W'$.

**Proof:** Say $u_n \rightarrow u$. Does it follow that $Au_n \rightarrow Au$? If not, then there is a subsequence such that $Au_n \rightarrow \xi \neq Au$ thanks to $\{Au_n\}$ being bounded. Then the lim sup condition holds obviously. In fact the limit of $\langle Au_n, u_n - u \rangle$ exists and equals 0. Hence for all $v$,

$$ \liminf_{n \to \infty} \langle Au_n, u_n - v \rangle \geq \langle Au, u - v \rangle $$
Therefore, 
\[ \langle \xi, u - v \rangle \geq \langle Au, u - v \rangle \]
for all \( v \) and so in fact \( \xi = Au \) after all. Thus \( A \) must be continuous.

As to the second part of this proposition, if you have \( w_n \rightharpoonup w \) in \( W \), then in fact convergence takes place strongly because weak and strong convergence are the same in finite dimensions. Hence the same argument given above holds to show that \( i^*Ai \) is continuous.

**Definition 9.1.3** \( A : V \to V' \) is monotone if for all \( v, u \in V \),
\[ \langle Au - Av, u - v \rangle \geq 0, \]
and \( A \) is Hemicontinuous if for all \( v, u \in V \),
\[ \lim_{t \to 0^+} \langle A (u + t (v - u)) , u - v \rangle = \langle Au - Av, u - v \rangle. \]

**Theorem 9.1.4** Let \( V \) be a Banach space and let \( A : V \to V' \) be monotone and hemicontinuous. Then \( A \) is pseudomonotone.

**Proof:** Let \( A \) be monotone and Hemicontinuous. First here is a claim.

**Claim:** If 9.1.1 and 9.1.2 hold, then \( \lim_{n \to \infty} \langle Au_n, u_n - u \rangle = 0. \)

**Proof of the claim:** Since \( A \) is monotone,
\[ \langle Au_n - Au, u_n - u \rangle \geq 0 \]
so
\[ \langle Au_n, u_n - u \rangle \geq \langle Au, u_n - u \rangle. \]

Therefore,
\[ 0 = \lim_{n \to \infty} \inf \langle Au, u_n - u \rangle \leq \lim_{n \to \infty} \inf \langle Au_n, u_n - u \rangle \leq \lim_{n \to \infty} \sup \langle Au_n, u_n - u \rangle \leq 0. \]

Now using that \( A \) is monotone again, then letting \( t > 0 \),
\[ \langle Au_n - A(u + t(v - u)), u_n - u + t(u - v) \rangle \geq 0 \]
and so
\[ \langle Au_n, u_n - u + t(u - v) \rangle \geq \langle A(u + t(v - u)), u_n - u + t(u - v) \rangle. \]

Taking the lim inf on both sides and using the claim and \( t > 0 \),
\[ t \lim_{n \to \infty} \inf \langle Au_n, u - v \rangle \geq t \langle A(u + t(v - u)), (u - v) \rangle. \]

Next divide by \( t \) and use the Hemicontinuity of \( A \) to conclude that
\[ \lim_{n \to \infty} \inf \langle Au_n, u - v \rangle \geq \langle Au, u - v \rangle. \]
From the claim,
\[
\liminf_{n \to \infty} \langle Au_n, u - v \rangle = \liminf_{n \to \infty} \left( \langle Au_n, u_n - v \rangle + \langle Au_n, u - u_n \rangle \right)
\]
\[
= \liminf_{n \to \infty} \langle Au_n, u_n - v \rangle \geq \langle Au, u - v \rangle.
\]

Monotonicity is very important in the above proof. The next example shows that even if the operator is linear and bounded, it is not necessarily pseudomonotone.

**Example 9.1.5** Let \( H \) be any Hilbert space (complete inner product space, more on these later) and let \( A : H \to H' \) be given by
\[
\langle Ax, y \rangle \equiv (-x, y)_H.
\]
Then \( A \) fails to be pseudomonotone.

**Proof:** Let \( \{x_n\}_{n=1}^\infty \) be an orthonormal set of vectors in \( H \). Then Parsevall’s inequality implies
\[
\|x\|^2 \geq \sum_{n=1}^\infty |(x_n, x)|^2
\]
and so for any \( x \in H \), \( \lim_{n \to \infty} (x_n, x) = 0 \). Thus \( x_n \to 0 \equiv x \). Also
\[
\limsup_{n \to \infty} \langle Ax_n, x_n - x \rangle = \limsup_{n \to \infty} \langle Ax_n, x_n - 0 \rangle = \limsup_{n \to \infty} \left(-\|x_n\|^2\right) = -1 \leq 0.
\]
If \( A \) were pseudomonotone, we would need to be able to conclude that for all \( y \in H \),
\[
\liminf_{n \to \infty} \langle Ax_n, x_n - y \rangle \geq \langle Ax, x - y \rangle = 0.
\]
However,
\[
\liminf_{n \to \infty} \langle Ax_n, x_n - 0 \rangle = -1 < 0 = \langle A0, 0 - 0 \rangle.
\]
The following proposition is useful.

**Proposition 9.1.6** Suppose \( A : V \to V' \) is pseudomonotone and bounded where \( V \) is separable. Then it must be demicontinuous. This means that if \( u_n \to u \), then \( Au_n \rightharpoonup Au \). In case that \( V \) is reflexive, you don’t need the assumption that \( V \) is separable.

**Proof:** Since \( u_n \to u \) is strong convergence and since \( Au_n \) is bounded, it follows
\[
\limsup_{n \to \infty} \langle Au_n, u_n - u \rangle = \lim_{n \to \infty} \langle Au_n, u_n - u \rangle = 0.
\]
Suppose this is not so that $A u_n$ converges weakly to $A u$. Since $A$ is bounded, there exists a subsequence, still denoted by $n$ such that $A u_n \to \xi$ weak $\ast$. I need to verify $\xi = A u$. From the above, it follows that for all $v \in V$

$$\langle A u, u - v \rangle \leq \lim \inf_{n \to \infty} \langle A u_n, u_n - v \rangle = \lim \inf_{n \to \infty} \langle A u_n, u - v \rangle = \langle \xi, u - v \rangle$$

Hence $\xi = A u$. ■

There is another type of operator which is more general than pseudomonotone.

**Definition 9.1.7** Let $A : V \to V'$ be an operator. Then $A$ is called type M if whenever $u_n \to u$ and $A u_n \to \xi$, and

$$\lim \sup_{n \to \infty} \langle A u_n, u_n \rangle \leq \langle \xi, u \rangle$$

it follows that $A u = \xi$.

**Proposition 9.1.8** If $A$ is pseudomonotone, then $A$ is type M.

**Proof:** Suppose $A$ is pseudomonotone and $u_n \to u$ and $A u_n \to \xi$, and

$$\lim \sup_{n \to \infty} \langle A u_n, u_n \rangle \leq \langle \xi, u \rangle$$

Then

$$\lim \sup_{n \to \infty} \langle A u_n, u_n - u \rangle = \lim \sup_{n \to \infty} \langle A u_n, u_n \rangle - \langle \xi, u \rangle \leq 0$$

Hence

$$\lim \inf_{n \to \infty} \langle A u_n, u_n - v \rangle \geq \langle A u, u - v \rangle$$

for all $v \in V$. Consequently, for all $v \in V$,

$$\langle A u, u - v \rangle \leq \lim \inf_{n \to \infty} \langle A u_n, u_n - v \rangle = \lim \inf_{n \to \infty} (\langle A u_n, u - v \rangle + \langle A u_n, u_n - u \rangle) = \langle \xi, u - v \rangle + \lim \inf_{n \to \infty} \langle A u_n, u_n - u \rangle \leq \langle \xi, u - v \rangle$$

and so $A u = \xi$. ■

An interesting result is the following which states that a monotone linear function added to a type M is also type M.

**Proposition 9.1.9** Suppose $A : V \to V'$ is type M and suppose $L : V \to V'$ is monotone, bounded and linear. Then $L + A$ is type M. Let $V$ be separable or reflexive so that the weak convergences in the following argument are valid.
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

Proof: Suppose $u_n \rightarrow u$ and $Au_n + Lu_n \rightarrow \xi$ and also that
\[
\limsup_{n \rightarrow \infty} \langle Au_n + Lu_n, u_n \rangle \leq \langle \xi, u \rangle.
\]
Does it follow that $\xi = Au + Lu$? Suppose not. There exists a further subsequence, still called $n$ such that $Lu_n \rightarrow Lu$. This follows because $L$ is linear and bounded. Then from monotonicity,
\[
\langle Lu_n, u_n \rangle \geq \langle Lu_n, u \rangle + \langle L(u), u_n - u \rangle.
\]
Hence with this further subsequence, the lim sup is no larger and so
\[
\limsup_{n \rightarrow \infty} \langle Au_n, u_n \rangle + \lim_{n \rightarrow \infty} ((\langle Lu_n, u \rangle + \langle L(u), u_n - u \rangle) \leq \langle \xi, u \rangle.
\]
and so
\[
\limsup_{n \rightarrow \infty} \langle Au_n, u_n \rangle \leq \langle \xi - Lu, u \rangle.
\]
It follows since $A$ is type $M$ that $Au = \xi - Lu$, which contradicts the assumption that $\xi \neq Au + Lu$.

There is also the following useful generalization of the above proposition.

Corollary 9.1.10 Suppose $A : V \rightarrow V'$ is type $M$ and suppose $L : W \rightarrow W'$ is monotone, bounded and linear where $V \subseteq W$ and $V$ is dense in $W$ so that $W' \subseteq V'$. Then for $u_0 \in W$ define $M(u) \equiv L(u - u_0)$. Then $M + A$ is type $M$. Let $V$ be separable or reflexive so that the weak convergences in the following argument are valid.

Proof: Suppose $u_n \rightarrow u$ and $Au_n + Mu_n \rightarrow \xi$ and also that
\[
\limsup_{n \rightarrow \infty} \langle Au_n + Mu_n, u_n \rangle \leq \langle \xi, u \rangle.
\]
Does it follow that $\xi = Au + Mu$? Suppose not. By assumption, $u_n \rightarrow u$ and so,
\[
u_n - u_0 \rightarrow u - u_0 \text{ weak convergence in } W
\]
since $L$ is bounded, there is a further subsequence, still called $n$ such that
\[
Mu_n = L(u_n - u_0) \rightarrow L(u - u_0) = Mu.
\]
Since $M$ is monotone,
\[
\langle Mu_n - Mu, u_n - u \rangle \geq 0.
\]
Thus
\[
\langle Mu_n, u_n \rangle - \langle Mu_n, u \rangle - \langle Mu, u_n \rangle + \langle Mu, u \rangle \geq 0
\]
and so
\[
\langle Mu_n, u_n \rangle \geq \langle Mu_n, u \rangle + \langle Mu, u_n - u \rangle.
\]
Hence with this further subsequence, the lim sup is no larger and so
\[ \langle \xi, u \rangle \geq \limsup_{n \to \infty} \langle Au_n + Mu_n, u_n \rangle \]
\[ \geq \limsup_{n \to \infty} (\langle Au_n, u_n \rangle + \langle Mu_n, u \rangle + \langle Mu_n, u_n - u \rangle) \]
\[ = \limsup_{n \to \infty} \langle Au_n, u_n \rangle + \lim_{n \to \infty}(\langle Mu_n, u \rangle + \langle Mu_n, u_n - u \rangle) \leq \langle \xi, u \rangle \]
and so
\[ \limsup_{n \to \infty} \langle Au_n, u_n \rangle \leq \langle \xi - Mu, u \rangle \]
It follows since \( A \) is type \( M \) that \( Au = \xi - Mu \), which contradicts the assumption that \( \xi \neq Au + Mu \).

The following is Browder’s lemma. It is a very interesting application of the Brouwer fixed point theorem.

**Lemma 9.1.11** *(Browder)* Let \( K \) be a convex closed and bounded set in \( \mathbb{R}^n \) and let \( A : K \to \mathbb{R}^n \) be continuous and \( f \in \mathbb{R}^n \). Then there exists \( x \in K \) such that for all \( y \in K \),
\[ (f - Ax, y - x)_{\mathbb{R}^n} \leq 0 \]
If \( K \) is convex, closed, bounded subset of \( V \) a finite dimensional vector space, then the same conclusion holds. If \( f \in V' \), there exists \( x \in K \) such that for all \( y \in K \),
\[ \langle f - Ax, y - x \rangle_{V', V} \leq 0 \]

**Proof:** Let \( P_K \) denote the projection onto \( K \). Thus \( P_K \) is Lipschitz continuous,
\[ x \to P_K (f - Ax + x) \]
is a continuous map from \( K \) to \( K \). By the Brouwer fixed point theorem, it has a fixed point \( x \in K \). Therefore, for all \( y \in K \),
\[ (f - Ax + x - x, y - x) = (f - Ax, y - x) \leq 0 \]
As to the second claim. Consider the following diagram.
\[
\begin{array}{ccc}
\mathbb{R}^n & \xrightarrow{\theta^*} & V' \\
\mathbb{R}^n & \xrightarrow{\theta} & V \\
\end{array}
\]
where
\[ \theta(x) = \sum_{i=1}^{n} x_i v_i \]
Thus \( \theta \) and \( \theta^* \) are both continuous linear and one to one and onto. Hence there is \( x \in \theta^{-1}K \) a closed convex and bounded subset of \( \mathbb{R}^n \) such that \( x = \theta^{-1}u, u \in K \), and
\[ (\theta^* f - \theta^* A \theta (\theta^{-1} u), \theta^{-1} y - \theta^{-1} u)_{\mathbb{R}^n} = \langle f - Au, y - u \rangle_{V', V} \leq 0 \]
for all \( y \in K \).

From this lemma, there is an interesting theorem on surjectivity.
Proposition 9.1.12 Let $A : V \to V'$ be continuous and coercive where $V$ is a finite dimensional normed linear space and
\[
\lim_{\|v\| \to \infty} \frac{\langle A(v + v_0), v \rangle}{\|v\|_{V'}} = \infty
\]
for some $v_0$. Then for all $f \in V'$, there exists $v \in V$ such that $Av = f$.

Proof: Define the closed convex sets $B_n \equiv B(v_0, n)$. By Browder's lemma, there exists $x_n$ such that
\[
\langle f - Av_n, y - v_n \rangle \leq 0
\]
for all $y \in B_n$. Then taking $y = v_0$,
\[
\langle Av_n, v_n - v_0 \rangle \leq \langle f, v_n - v_0 \rangle
\]
letting $w_n = v_n - v_0$,
\[
\langle A(w_n + v_0), w_n \rangle \leq \langle f, w_n \rangle
\]
and so
\[
\frac{\langle A(w_n + v_0), w_n \rangle}{\|w_n\|} \leq \|f\|
\]
which implies that the $\|w_n\|$ and hence the $\|v_n\|$ are bounded. It follows that for large $n$, $v_n$ is an interior point of $B_n$. Therefore,
\[
\langle f - Av_n, z \rangle_{V', V} \leq 0
\]
for all $z$ in some open ball centered at $v_0$. Hence $f - Av_n = 0$.

Lemma 9.1.13 Let $A : V \to V'$ be type $M$ and bounded and suppose $V$ is reflexive or $V$ is separable. Then $A$ is demicontinuous.

Proof: Suppose $u_n \to u$ and $Au_n$ fails to converge weakly to $Au$. Then there is a further subsequence, still denoted as $u_n$ such that $Au_n \rightharpoonup \zeta \neq Au$. Then thanks to the strong convergence, you have
\[
\limsup_{n \to \infty} \langle Au_n, u_n \rangle = \langle \zeta, u \rangle
\]
which implies $\zeta = Au$ after all.

With these lemmas and the above proposition, there is a very interesting surjectivity result.

Theorem 9.1.14 Let $A : V \to V'$ be type $M$, bounded, and coercive
\[
\lim_{\|u\| \to \infty} \frac{\langle A(u + u_0), u \rangle}{\|u\|} = \infty,
\]
for some $u_0$, where $V$ is a separable reflexive Banach space. Then $A$ is surjective.
9.1. SOME NONLINEAR SINGLE VALUED OPERATORS

Proof: Since $V$ is separable, there exists an increasing sequence of finite dimensional subspaces $\{V_n\}$ such that $\bigcup_n V_n = V$ and each $V_n$ contains $u_0$. Say $\text{span} (v_1, \cdots , v_n) = V_n$. Then consider the following diagram.

\[ V_n' \xleftarrow{i^*} V' \xrightarrow{i} V \]

The map $i$ is the inclusion map. Consider the map $i^*Ai$. By Lemma 9.1.13 this map is continuous.

\[ \frac{\langle i^*Ai (v + u_0) , v \rangle_{V_n'}}{\|v\|} = \frac{\langle A (v + u_0) , v \rangle_{V',V}}{\|v\|} \]

Hence $i^*Ai$ is coercive. Let $f \in V'$. Then from Proposition 9.1.12, there exists $x_n$ such that

\[ i^*Aiv_n = i^*f \]

In other words,

\[ \langle Av_n, y \rangle_{V',V} = \langle f, y \rangle_{V',V} \quad (9.1.5) \]

for all $y \in V_n$. Letting $y \equiv v_n - u_0 \equiv w_n$,

\[ \langle A(w_n + u_0) , w_n \rangle = \langle f, w_n \rangle \]

Then from the coercivity condition 9.1.4, the $w_n$ are bounded independent of $n$. Hence this is also true of the $v_n$. Since $V$ is reflexive, there is a subsequence, still called $\{v_n\}$ which converges weakly to $v \in V$. Since $A$ is bounded, it can also be assumed that $Av_n \rightharpoonup \zeta \in V'$. Then

\[ \limsup_{n \to \infty} \langle Av_n, v_n \rangle = \limsup_{n \to \infty} \langle f, v_n \rangle = \langle f, v \rangle \]

Also, passing to the limit in 9.1.5,

\[ \langle \zeta, y \rangle = \langle f, y \rangle \]

for any $y \in V_n$, this for any $n$. Since the union of these $V_n$ is dense, it follows that the above equation holds for all $y \in V$. Therefore, $f = \zeta$ and so

\[ \limsup_{n \to \infty} \langle Av_n, v_n \rangle = \limsup_{n \to \infty} \langle f, v_n \rangle = \langle f, v \rangle = \langle \zeta, v \rangle \]

Since $A$ is type $M$, $Av = \zeta = f$. 

You can generalize pseudomonotone slightly without any trouble.

Definition 9.1.15 Let $V$ be a Banach space and let $K$ be a closed convex nonempty subset of $V$. Then $A : K \to V'$ is pseudomonotone if similar conditions hold as above. That is, if

\[ u_n \rightharpoonup u \quad (9.1.6) \]
and
\[ \limsup_{n \to \infty} \langle Au_n, u_n - u \rangle \leq 0 \] (9.1.7)
it follows that for all \( v \in K \),
\[ \liminf_{n \to \infty} \langle Au_n, u_n - v \rangle \geq \langle Au, u - v \rangle. \] (9.1.8)

Then it is easy to give a nice result on variational inequalities.

**Proposition 9.1.16** Let \( K \) be a closed convex nonempty subset of \( V \) a separable reflexive Banach space. Let \( A : K \to V' \) be pseudomonotone and bounded. Also assume that either \( K \) is bounded or there is a coercivity condition
\[ \lim_{\|u\| \to \infty} \frac{\langle Au, u - u_0 \rangle}{\|u\|} = \infty, \quad u_0 \in K \]
then for \( f \in V' \), there exists \( u \in K \) such that for all \( v \in K \),
\[ \langle Au, u - v \rangle \leq \langle f, u - v \rangle \]

**Proof:** Let \( V_n \) be finite dimensional spaces whose union is dense in \( V, \ldots, V_n \subseteq V_{n+1} \cdot \cdot \cdot \), each containing \( u_0, n > \|u_0\| \). By a repeat of the proof of Proposition 9.1.2, \( i^*Ai \) will be continuous on \( K \). Therefore, by Browder's lemma, there exists \( u_n \in K_n \equiv K \cap B(0, n) \cap V_n \) such that for all \( v \in K_n \),
\[ \langle i^*f - i^*Ai u_n, v - u_n \rangle_{V_n', V_n} = \langle f - Au_n, v - u_n \rangle_{V', V} \leq 0 \]
Now assume we don’t know that \( K \) is bounded. In case it is bounded, the argument simplifies. In the harder case, the coercivity condition implies that the \( u_n \) are bounded in \( V \). This follows from letting \( v = u_0 \) in the above inequality. Thus
\[ \langle f, u_n - u_0 \rangle \geq \langle Au_n, u_n - u_0 \rangle \]
Hence
\[ \frac{\langle Au_n, u_n - u_0 \rangle}{\|u_n\|} \leq \frac{\|f\| \|u_n - u_0\|}{\|u_n\|} \]
The right side is bounded and so it follows that the left side is also bounded. Therefore, \( \|u_n\| \) must be bounded. Taking a subsequence and using the assumption that \( V \) is reflexive, we can obtain
\[ u_n \to u \text{ weakly in } V \]
By the fact that convex closed sets are weakly closed also, it follows that \( u \in K \). Also, given \( M \), eventually all \( \|u_n\| \) and \( \|u\| \) are less than \( M \). Now from the inequality,
\[ \langle Au_n, u_n - v \rangle \leq \langle f, u_n - v \rangle \]
Thus
\[ \langle Au_n, u_n - u \rangle + \langle Au_n, u - v \rangle \leq \langle f, u_n - u \rangle + \langle f, u - v \rangle \]
9.2. DUALITY MAPS

Then taking \( \limsup_{n \to \infty} \) one gets

\[
\limsup_{n \to \infty} \langle Au_n, u_n - u \rangle + \langle \xi, u - v \rangle \leq \langle f, u - v \rangle
\]

This holds for \( v \in K_m \) where \( m \) is arbitrary. Hence one could let \( v_m \to u \). Thus eventually \( \|v_m\| < M \) and so for large \( m, v_m \in K_m \). Then it follows that

\[
\limsup_{n \to \infty} \langle Au_n, u_n - u \rangle \leq 0.
\]

Consequently, by the assumption that \( A \) is pseudomonotone on \( K \), for every \( v \in K \),

\[
\langle Au, u - v \rangle \leq \liminf_{n \to \infty} \langle Au_n, u_n - v \rangle
\]

for all \( v \in K \). Then from the inequality obtained from Browder’s lemma,

\[
\langle Au_n, u_n - v \rangle \leq \langle f, u_n - v \rangle
\]

and so * implies on taking \( \liminf \) that for all \( v \in K \),

\[
\langle Au, u - v \rangle \leq \langle f, u - v \rangle
\]

9.2 Duality Maps

The duality map is an attempt to duplicate some of the features of the Riesz map in Hilbert space which is discussed in the chapter on Hilbert space.

**Definition 9.2.1** A Banach space is said to be strictly convex if whenever \( \|x\| = \|y\| \) and \( x \neq y \), then

\[
\left\| \frac{x + y}{2} \right\| < \|x\|.
\]

\( F : X \to X' \) is said to be a duality map if it satisfies the following: a.) \( ||F(x)|| = ||x||^{p-1} \). b.) \( F(x)(x) = ||x||^p \), where \( p > 1 \).

Duality maps exist. Here is why. Let

\[
F(x) = \left\{ x^* : ||x^*|| \leq ||x||^{p-1} \text{ and } x^*(x) = ||x||^p \right\}
\]

Then \( F(x) \) is not empty because you can let \( f(\alpha x) = \alpha ||x||^p \). Then \( f \) is linear and defined on a subspace of \( X \). Also

\[
\sup_{||\alpha x|| \leq 1} |f(\alpha x)| = \sup_{||\alpha x|| \leq 1} |\alpha||x||^p \leq ||x||^{p-1}
\]

Also from the definition,

\[
f(x) = ||x||^p
\]
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

and so, letting $x^*$ be a Hahn Banach extension, it follows $x^* \in F(x)$.

Also, $F(x)$ is closed and convex. It is clearly closed because if $x^*_n \to x^*$, the condition on the norm clearly holds and also the other one does too. It is convex because

$$\|x^* \lambda + (1 - \lambda) y^*\| \leq \lambda \|x^*\| + (1 - \lambda) \|y^*\| \leq \lambda \|x\|^{p-1} + (1 - \lambda) \|x\|^{p-1}$$

If the conditions hold for $x^*$, then we can show that in fact $\|x^*\| = \|x\|^{p-1}$.

This is because

$$\|x^*\| \geq \left(\frac{x}{\|x\|}\right) = \left(\frac{1}{\|x\|}\right) \|x^* (x)\| = \|x\|^{p-1}.$$

Now how many things are in $F(x)$ assuming the norm on $X'$ is strictly convex? Suppose $x_1^*$ and $x_2^*$ are two things in $F(x)$. Then by convexity, so is $(x_1^* + x_2^*)/2$.

Hence by strict convexity, if the two are different, then

$$\left\|\frac{x_1^* + x_2^*}{2}\right\| = \|x\|^{p-1} \leq \frac{1}{2} \|x_1^*\| + \frac{1}{2} \|x_2^*\| = \|x\|^{p-1}$$

which is a contradiction. Therefore, $F$ is an actual mapping.

What are some of its properties? First is one which is similar to the Cauchy-Schwarz inequality. Since $p - 1 = p/p'$,

$$\sup_{\|y\| \leq 1} |\langle Fx, y \rangle| = \|x\|^{p/p'}$$

and so for arbitrary $y \neq 0$,

$$|\langle Fx, y \rangle| = \|y\| \left|\left\langle Fx, \frac{y}{\|y\|} \right\rangle\right| \leq \|y\| \|x\|^{p/p'}$$

Next we can show that $F$ is monotone.

$$\langle Fx - Fy, x - y \rangle = \langle Fx, x \rangle - \langle Fx, y \rangle - \langle Fy, x \rangle + \langle Fy, y \rangle$$

$$\geq \|x\|^p + \|y\|^p - \|y\| \|x\|^{p/p'} - \|y\|^{p/p'} \|x\|$$

$$\geq \|x\|^p + \|y\|^p - \left(\frac{\|y\|^p}{p} + \frac{\|x\|^p}{p'}\right) - \left(\frac{\|y\|^p}{p'} + \frac{\|x\|^p}{p}\right) = 0$$

Next it can be shown that $F$ is hemicontinuous. By the construction, $F'(x + ty)$ is bounded as $t \to 0$. Let $y$ be a subsequence such that

$$F(x + ty) \to \xi$$

Then we ask: Does $\xi$ do what it needs to do in order to be $F(x)$? The answer is yes. First of all $\|F(x + ty)\| = \|x + ty\|^{p-1} \to \|x\|^{p-1}$. The set

$$\left\{ x^* : \|x^*\| \leq \|x\|^{p-1} + \varepsilon \right\}$$
is closed and convex and so it is weak ∗ closed as well. For all small enough \( t \), it follows \( F(x + ty) \) is in this set. Therefore, the weak limit is also in this set and it follows \( ||ξ|| ≤ ||x||^{p-1} + ε. \) Since \( ε \) is arbitrary, it follows \( ||ξ|| ≤ ||x||^{p-1}. \) Is \( ξ(x) = ||x||^p? \) We have
\[
||x||^p = \lim_{t \to 0} ||x + ty||^p = \lim_{t \to 0} ⟨F(x + ty), x + ty⟩ = \lim_{t \to 0} ⟨F(x + ty), x⟩ = ⟨ξ, x⟩
\]
and so, \( ξ \) does what it needs to do to be \( F(x) \). This would be clear if \( ||ξ|| = ||x||^{p-1}. \) However, \( ||ξ, x|| = ||x||^p \) and so \( ||ξ|| ≥ \left( ξ, \frac{x}{||x||} \right) = ||x||^{p-1}. \) This shows \( ξ \) does everything it needs to do to equal \( F(x) \) and so it is \( F(x). \)

Since this conclusion follows for any convergent sequence, it follows that \( F(x + ty) \) converges to \( F(x) \) weakly as \( t \to 0. \) This is what it means to be hemicontinuous. This proves the following theorem. One can show also that \( F \) is demicontinuous which means strongly convergent sequences go to weakly convergent sequences. Here is a proof for the case where \( p = 2. \) You can clearly do the same thing for arbitrary \( p. \)

**Lemma 9.2.2** Let \( F \) be a duality map for \( p = 2 \) where \( X, X' \) are reflexive and have strictly convex norms. (If \( X \) is reflexive, there is always an equivalent strictly convex norm.\(^3\)) Then \( F \) is demicontinuous.

**Proof:** Say \( x_n \to x. \) Then does it follow that \( Fx_n \rightharpoonup Fx? \) Suppose not. Then there is a subsequence, still denoted as \( x_n \) such that \( x_n \to x \) but \( Fx_n \rightharpoonup y \neq Fx \) where here \( \rightharpoonup \) denotes weak convergence. This follows from the Eberlein Smulian theorem, Problem \( 8.4. \) on Page \( 66. \) Then
\[
⟨y, x⟩ = \lim_{n \to ∞} ⟨Fx_n, x_n⟩ = \lim_{n \to ∞} ||x_n||^2 = ||x||^2
\]
Also, there exists \( z, ||z|| = 1 \) and \( ⟨y, z⟩ ≥ ||y|| − ε. \) Then
\[
||y|| − ε ≤ ⟨y, z⟩ = \lim_{n \to ∞} ⟨Fx_n, z⟩ ≤ \lim_{n \to ∞} ||Fx_n|| = \lim_{n \to ∞} ||x_n|| = ||x||
\]
and since \( ε \) is arbitrary, \( ||y|| ≤ ||x||. \) It follows from the above construction of \( Fx, \) that \( y = Fx \) after all, a contradiction.

**Theorem 9.2.3** Let \( X \) be a reflexive Banach space with \( X' \) having strictly convex norm.\(^4\) Then for \( p > 1, \) there exists a mapping \( F : X \to X' \) which is bounded, monotone, hemicontinuous, coercive in the sense that \( \lim_{||x|| \to ∞} ⟨Fx, x⟩ / ||x|| = ∞, \) which also satisfies the inequalities
\[
||⟨Fx, y⟩|| ≤ ||⟨Fx, x⟩||^{1/p'} ||⟨Fy, y⟩||^{1/p}
\]
\(^1\)It is known that if the space is reflexive, then there is an equivalent norm which is strictly convex. However, in most examples, this strict convexity is obvious.
Note that these conclusions about duality maps show that they map onto the dual space.

The duality map was onto and it was monotone. This was shown above. Consider the form of a duality map for the \( L^p \) spaces. Let \( F : L^p \to (L^p)' \) be the one which satisfies
\[
\|Ff\| = \|f\|^{p-1}, \quad \langle Ff, f \rangle = \|f\|^p
\]
Then in this case,
\[
Ff = |f|^{p-2} \overline{f}
\]
This is because it does what it needs to do.

\[
\|Ff\|_{L^{p'}} = \left( \int_\Omega \left( |f|^{p-1} \right)^{p'} \, d\mu \right)^{1/p'} = \left( \int_\Omega \left( |f|^{p/p'} \right)^{p'} \, d\mu \right)^{1/p'} = \left( \int_\Omega |f|^p \, d\mu \right)^{1-(1/p)} = \left( \int_\Omega |f|^p \, d\mu \right)^{p-1} = \|f\|^{p-1}_{L^p}
\]
while it is obvious that
\[
\langle Ff, f \rangle = \int_\Omega |f|^p \, d\mu = \|f\|_{L^p(\Omega)}^p.
\]

Now here is an interesting inequality which I will only consider in the case where the quantities are real valued.

**Lemma 9.2.4** Let \( p \geq 2 \). Then for \( a, b \) real numbers,
\[
\left( |a|^{p-2} a - |b|^{p-2} b \right) (a - b) \geq C |a - b|^p
\]
for some constant \( C \) independent of \( a, b \).

**Proof:** There is nothing to show if \( a = b \). Without loss of generality, assume \( a > b \). Also assume \( p > 2 \). There is nothing to show if \( p = 2 \). I want to show that there exists a constant \( C \) such that for \( a > b \),
\[
\frac{|a|^{p-2} a - |b|^{p-2} b}{|a - b|^{p-1}} \geq C
\]
(9.2.9)
First assume also that \( b \geq 0 \). Now it is clear that as \( a \to \infty \), the quotient above converges to 1. Take the derivative of this quotient. This yields
\[
(p - 1) |a - b|^{p-2} \frac{|a|^{p-2} |a - b| - |a|^{p-2} a - |b|^{p-2} b}{|a - b|^{2p-2}}
\]
Now remember \( a > b \). Then the above reduces to
\[
(p - 1) |a - b|^{p-2} b |b|^{p-2} - |a|^{p-2}
\]
Since \( b \geq 0 \), this is negative and so 1 would be a lower bound. Now suppose \( b < 0 \). Then the above derivative is negative for \( b < a \leq -b \) and then it is positive for \( a > -b \). It equals 0 when \( a = -b \). Therefore the quotient in (9.2.9) achieves its minimum value when \( a = -b \). This value is

\[
\frac{|-b|^{p-2} (-b) - |b|^{p-2} b}{|b|^{p-1}} = \frac{|b|^{p-2} - 2b}{|b|^{p-1}} = \frac{|b|^{p-2}}{|b|^{p-2}} = \frac{1}{2^{p-2}}.
\]

Therefore, the conclusion holds whenever \( p \geq 2 \). That is

\[
\langle F - Fv, u - v \rangle \geq 0
\]

and it equals 0 if and only if \( u - v \).

**Proof:** First why is it monotone? By definition of \( F \), \( \langle F(u), u \rangle = \|u\|^2 \) and \( \|F(u)\| = \|u\| \). Then

\[
|\langle Fu, v \rangle| = \left| \left\langle Fu, \frac{v}{\|v\|} \right\rangle \right| \|v\| \leq \|Fu\| \|v\| = \|u\| \|v\|
\]

Hence

\[
\langle Fu - Fv, u - v \rangle = \|u\|^2 + \|v\|^2 - \langle Fu, v \rangle - \langle Fv, u \rangle \geq \|u\|^2 + \|v\|^2 - 2 \|u\| \|v\| \geq 0
\]

Now suppose \( \|x\| = \|y\| = 1 \) but \( x \neq y \). Then

\[
\left\langle Fx, \frac{x + y}{2} \right\rangle \leq \left\| \frac{x + y}{2} \right\| < \frac{\|x\| + \|y\|}{2} = 1
\]

It follows that

\[
\frac{1}{2} \langle Fx, x \rangle + \frac{1}{2} \langle Fx, y \rangle = \frac{1}{2} + \frac{1}{2} \langle Fx, y \rangle < 1
\]

and so

\[
\langle Fx, y \rangle < 1
\]

For arbitrary \( x, y, x/\|x\| \neq y/\|y\| \)

\[
\langle Fx, y \rangle = \|x\| \|y\| \left\langle F \left( \frac{x}{\|x\|} \right), \left( \frac{y}{\|y\|} \right) \right\rangle
\]
It is easy to check that $F(\alpha x) = \alpha F(x)$. Therefore,

$$|\langle Fx, y \rangle| = \|x\| \|y\| \left( F\left( \frac{x}{\|x\|} \right), \left( \frac{y}{\|y\|} \right) \right) < \|x\| \|y\|$$

Now say that $x \neq y$ and consider

$$\langle Fx - Fy, x - y \rangle$$

First suppose $x = \alpha y$. Then the above is

$$\langle F(\alpha y) - Fy, (\alpha - 1)y \rangle = (\alpha - 1) \left( \langle F(\alpha y), y \rangle - \|y\|^2 \right)$$

$$= (\alpha - 1) \left( \langle \alpha F(y), y \rangle - \|y\|^2 \right)$$

$$= (\alpha - 1)^2 \|y\|^2 > 0$$

The other case is that $x/\|x\| \neq y/\|y\|$ and in this case,

$$\langle Fx - Fy, x - y \rangle = \|x\|^2 + \|y\|^2 - \langle Fx, y \rangle - \langle Fy, x \rangle$$

$$> \|x\|^2 + \|y\|^2 - 2 \|x\| \|y\| \geq 0$$

Thus $F$ is strictly monotone as claimed.

As mentioned, this will hold for any $p > 1$. Here is a proof in the case that the Banach space is real which is the usual case of interest. First here is a simple observation.

**Observation 9.2.6** Let $p > 1$. Then $x \rightarrow |x|^{p-2} x$ is strictly monotone. Here $x \in \mathbb{R}$.

To verify this observation,

$$\frac{d}{dx} \left( (x^2)^{\frac{p-2}{2}} x \right) = \frac{1}{x^2} (p - 1) (x^3)^{\frac{1}{2p}} > 0$$

**Theorem 9.2.7** Let $X$ be a real reflexive Banach space and $X, X'$ have strictly convex norms as discussed above. Let $F$ be the duality map for $p > 1$. Then $F$ is strictly monotone. This means

$$\langleFu - Fv, u - v\rangle \geq 0$$

and it equals 0 if and only if $u = v$.

**Proof:** First why is it monotone? By definition of $F$, $\langle F(u), u \rangle = \|u\|^p$ and $\|F(u)\| = \|u\|^{p-1}$. Then

$$|\langle Fu, v \rangle| = \left| \langle Fu, \frac{v}{\|v\|} \right| \|v\| \leq \|Fu\| \|v\| = \|u\|^{p-1} \|v\|$$
Hence

\[ (F_u - F_v, u - v) = \|u\|^p + \|v\|^p - (F_{u, v}) - (F_{v, u}) \]
\[
\geq \|u\|^p + \|v\|^p - \|u\|^{p-1} \|v\| - \|u\| \|v\|^{p-1}
\]
\[
\geq \|u\|^p + \|v\|^p - \left( \frac{\|u\|^p}{p'} + \frac{\|v\|^p}{p} \right) - \left( \frac{\|u\|^p}{p} + \frac{\|v\|^p}{p'} \right) = 0
\]

Now suppose \( \|x\| = \|y\| = 1 \) but \( x \neq y \). Then

\[
\langle F(x, \frac{x + y}{2}) \rangle \leq \|x\|^{p-1} \left\| \frac{x + y}{2} \right\| < \frac{\|x\| + \|y\|}{2} = 1
\]

It follows that

\[
\frac{1}{2} \langle F(x, x) + \frac{1}{2} (F(x, y) = \frac{1}{2} + \frac{1}{2} (F(x, y) < 1
\]

and so

\[ \langle F(x, y) < 1 \]

It is easy to check that for nonzero \( \alpha, F(\alpha x) = |\alpha|^{p-2} \alpha F(x) \). This is because

\[
\left\| |\alpha|^{p-2} \alpha F(x) \right\| = |\alpha|^{p-1} \|x\|^{p-1} = \|\alpha x\|^{p-1}
\]
\[
\langle |\alpha|^{p-2} \alpha F(x, \alpha x) \rangle = |\alpha|^{p} \|x\|^p = \|\alpha x\|^p
\]

and so, since \( |\alpha|^{p-2} \alpha F(x) \) acts like \( F(\alpha x) \), it is \( F(\alpha x) \). It follows that for arbitrary \( x, y \), such that \( x/\|x\| \neq y/\|y\| \)

\[
\langle F(x, y) = \|x\|^{p-1} \|y\| \left\langle F\left( \frac{x}{\|x\|} \right), \left( \frac{y}{\|y\|} \right) \right\rangle
\]

Therefore,

\[
\langle F(x, y) = \|x\|^{p-1} \|y\| \left\langle F\left( \frac{x}{\|x\|} \right), \left( \frac{y}{\|y\|} \right) \right\rangle < \|x\|^{p-1} \|y\|
\]

Now say that \( x \neq y \) and consider

\[
\langle F(x - Fy, x - y) \rangle
\]

First suppose \( x = \alpha y \). This is the case where \( x \) is a multiple of \( y \). Then the above is

\[
(F(\alpha y) - Fy, (\alpha - 1) y) = (\alpha - 1) (F(\alpha y), y) - \|y\|^p
\]
\[
= (\alpha - 1) \left( |\alpha|^{p-2} \alpha \|y\|^p - \|y\|^p \right) = (\alpha - 1) \left( |\alpha|^{p-2} \alpha - 1 \right) \|y\|^p > 0
\]

by the above observation that \( x \to |x|^{p-2} x \) is strictly monotone. Similarly, \( \langle F(x - Fy, x - y) \rangle > 0 \) if \( y = \alpha x \) for \( \alpha \neq 1 \).
Thus the desired result holds in the case that one vector is a multiple of the other. The other case is that neither vector is a multiple of the other. Thus, in particular, 
\[ \frac{x}{\|x\|} \neq \frac{y}{\|y\|}, \]
and in this case, it follows from
\[ \langle Fx - Fy, x - y \rangle = \|x\|^p + \|y\|^p - \langle Fx, y \rangle - \langle Fy, x \rangle \]
\[ > \|x\|^p + \|y\|^p - \|x\|^{p-1}\|y\| - \|y\|^{p-1}\|x\| \]
\[ \geq \|x\|^p + \|y\|^p - \left( \frac{\|x\|^p}{p'} + \frac{\|y\|^p}{p} \right) - \left( \frac{\|y\|^p}{p'} + \frac{\|x\|^p}{p} \right) = 0 \]
Thus \( F \) is strictly monotone as claimed.

Another useful observation about duality maps for \( p = 2 \) is that
\[ \|y^\ast\|_{V'} = \langle F_{F^{-1}y^\ast}, y^\ast \rangle = \langle F_{F^{-1}y^\ast}, y^\ast \rangle = \frac{\|y^\ast\|_{V'}^2}{\|y^\ast\|^2} \]
also from similar reasoning,
\[ \langle y^\ast, F_{F^{-1}y^\ast} \rangle = \langle F F^{-1}y^\ast, F^{-1}y^\ast \rangle = \|F^{-1}y^\ast\|^2 = \|y^\ast\|_{V'}^2 \]

### 9.3 Penalization And Projection Operators

In this section, \( X \) will be a reflexive Banach space such that \( X, X' \) has a strictly convex norm. Let \( K \) be a closed convex set in \( X \). Then the following lemma is obtained.

**Lemma 9.3.1** Let \( K \) be closed and convex nonempty subset of \( X \) a reflexive Banach space which has strictly convex norm. Then there exists a projection map \( P \) such that \( Px \in K \) and for all \( y \in K \),
\[ \|y - x\| \geq \|x - Px\| \]

**Proof:** Let \( \{y_n\} \) be a minimizing sequence for \( y \rightarrow \|y - x\| \) for \( y \in K \). Thus
\[ d \equiv \inf \{\|y - x\| : y \in K\} = \lim_{n \rightarrow \infty} \|y_n - x\| \]
Then obviously \( \{y_n\} \) is bounded. Hence there is a subsequence, still denoted by \( n \) such that \( y_n \rightarrow w \in K \). Then
\[ \|w - x\| \leq \lim \inf_{n \rightarrow \infty} \|y_n - x\| = d \]
How many closest points to \( x \) are there? Suppose \( w_1 \) is another one. Then
\[ \frac{w_1 + w}{2} - x = \frac{w_1 - x + w - x}{2} < \frac{w_1 - x}{2} + \frac{w - x}{2} = d \]
contradicting the assumption that both \( w, w_1 \) are closest points to \( x \). Therefore, \( Px \) consists of a single point.
Denote by $F$ the duality map such that $\langle Fx, x \rangle = \|x\|^2$. This is described earlier but there is also a very nice treatment which is somewhat different in [3]. Everything can be generalized and is in [41] but here I will only consider this case. First here is a useful result.

**Proposition 9.3.2** Let $F$ be the duality map just described. Let $\phi(x) \equiv \|x\|^2$. Then $F(x) = \partial \phi(x)$.

**Proof:** This follows from

$$\langle Fx, y - x \rangle \leq \langle Fx, y \rangle - \langle Fx, x \rangle \leq \langle Fx, x \rangle^{1/2} \langle Fy, y \rangle^{1/2} - \langle Fx, x \rangle \leq \langle Fy, y \rangle^{1/2} - \langle Fx, x \rangle^{1/2} \leq \|y\|^2 - \|x\|^2.$$  

Next is a really nice result about the characterization of $Px$ in terms of $F$.

**Proposition 9.3.3** Let $K$ be a nonempty closed convex set in $X$ a reflexive Banach space in which both $X, X'$ have strictly convex norms. Then $w \in K$ is equal to $Px$ if and only if

$$\langle F(x - w), y - w \rangle \leq 0$$

for every $y \in K$.

**Proof:** First suppose the condition. Then for $y \in K$, it follows from the above proposition about the subgradient,

$$\frac{1}{2} \|x - y\|^2 - \frac{1}{2} \|x - w\|^2 \geq \langle F(x - w), w - y \rangle \geq 0$$

and so since this holds for all $y$ it follows that

$$\|x - y\| \geq \|x - w\|$$

for all $y$ which says that $w = Px$.

Next, using the subgradient idea again, for $\theta \in [0, 1]$, suppose $w = Px$ then for $y \in K$ arbitrary,

$$0 \geq \frac{1}{2} \|x - w\|^2 - \frac{1}{2} \|x - (w + \theta (y - w))\|^2 \geq \langle F(x - (w + \theta (y - w))), \theta (y - x) \rangle$$

Now divide by $\theta$ and let $\theta \downarrow 0$ and use the hemicontinuity of $F$ given above. Then

$$0 \geq \langle F(x - w), y - x \rangle \ ■$$

**Definition 9.3.4** An operator of penalization is an operator $f : X \rightarrow X'$ such that $f = 0$ on $K$, $f$ is monotone and nonzero off $K$ as well as demicontinuous. (Strong convergence goes to weak convergence.) Actually, in applications, it is usually easy to give an ad hoc description of an appropriate penalization operator.
Proposition 9.3.5 Let \( K \) be a closed convex nonempty subset of \( X \) a reflexive Banach space such that \( X, X' \) have strictly convex norms. Then
\[
f(x) \equiv F(x - Px)
\]
is an operator of penalization. Here \( P \) is the projection onto \( K \).

Proof: First, observe that \( f(x) \) is 0 on \( K \) and nonzero off \( K \). Why is it monotone?
\[
\langle F(x - Px) - F(x_1 - Px_1), x - x_1 \rangle = \langle F(x - Px) - F(x_1 - Px_1), x - Px - (x_1 - Px_1) \rangle
\]
The first term is \( \geq 0 \) because \( F \) is monotone. As to the second, it equals
\[
\langle F(x - Px), Px - Px_1 \rangle + \langle F(x_1 - Px_1), Px_1 - Px \rangle
\]
and both of these are \( \geq 0 \) because of Proposition 9.3.3 which characterizes the projection map.

Now why is this hemicontinuous? Let \( x_n \to x \). Then \( Px_n \) is clearly bounded. Taking a subsequence, it can be assumed that \( Px_n \to \xi \) weakly. Is \( \xi = Px? \)
\[
\|x - Px\| \leq \|x - x_n\| + \|x_n - Px_n\| \quad \|x_n - Px_n\| \leq \|x_n - Px\| \leq \|x - x_n\| + \|x - Px\|
\]
It follows that
\[
\|x - Px\| - \|x_n - Px_n\| \leq \|x - x_n\| \quad \|x_n - Px_n\| - \|x - Px\| \leq \|x - x_n\|
\]
Hence \( \|x_n - Px_n\| \to \|x - Px\| \). However, from convexity and strong lower semicontinuity implying weak lower semicontinuity,
\[
\|x - \xi\| \leq \liminf_{n \to \infty} \|x_n - Px_n\| = \|x - Px\|
\]
and so \( \xi = Px \) because there is only one value in \( Px \). This has shown that, thanks to uniqueness of \( Px \), \( x_n \to x \) implies \( Px_n \to Px \) weakly.

Next we show that \( f \) is demicontinuous. Suppose \( x_n \to x \). Then from what was just shown, \( Px_n \to Px \) weakly. Thus \( x_n - Px_n \to x - Px \) weakly. Then
\[
\limsup_{n \to \infty} \langle F(x_n - Px_n), x_n - Px_n - (x - Px) \rangle
\]
from Proposition 9.3.3 which characterizes the projection map. It follows that, since \( F \) is monotone hemicontinuous and bounded, it is also pseudomonotone and so for all \( v \)
\[
\liminf_{n \to \infty} \langle F(x_n - Px_n), (x_n - Px_n) - v \rangle
\]
Now \( F(x_n - Px_n) \) is bounded. If it converges to \( \xi \), then
\[
\lim_{n \to \infty} \inf \langle F(x_n - Px_n), (x_n - Px_n) - (x - Px) \rangle 
\leq \lim_{n \to \infty} \sup \left[ \langle F(x_n - Px_n), (x_n - Px_n) - (x - Px) \rangle + \langle F(x_n - Px_n), (x - Px) - v \rangle \right] 
\leq \langle \xi, (x - Px) - v \rangle
\]
It follows that
\[
\langle \xi, (x - Px) - v \rangle \geq \lim_{n \to \infty} \inf \langle F(x_n - Px_n), (x_n - Px_n) - (x - Px) \rangle 
\geq \langle F(x - Px), (x - Px) - v \rangle
\]
Since \( v \) is arbitrary, it follows that \( \xi = F(x - Px) \). Hence \( F(x_n - Px_n) \to F(x - Px) \) weakly. Thus this is demicontinuous.

### 9.4 Set-Valued Maps, Pseudomonotone Operators

In the abstract theory of partial differential equations and variational inequalities, it is important to consider set-valued maps from a Banach space to the power set of its dual. In this section we give an introduction to this theory by proving a general result on surjectivity for a class of such operators. I am writing some of these theorems for the context of complex spaces, but we are mainly concerned with real spaces. However, it indicates what you would do for the complex case. Basically you just consider real parts. There isn’t any real mathematical significance in this generalization.

**Lemma 9.4.1** Suppose \( A : \mathbb{C}^n \to \mathcal{P}(\mathbb{C}^n) \) satisfies \( Ax \) is compact and convex, and \( A \) is upper semicontinuous, \( 7.2.9 \) and \( K \) is a nonempty compact convex set in \( \mathbb{C}^n \). Then if \( y \in \mathbb{C}^n \) there exists \([x, w] \in G(A)\) such that \( x \in K \) and
\[
\text{Re}(y - w, z - x) \leq 0
\]
for all \( z \in K \).

**Proof:** Tile \( \mathbb{C}^n \) with \( 2n \) simplices such that the collection is locally finite and each simplex has diameter less than \( \varepsilon < 1 \). This collection of simplices is determined by a countable collection of vertices. For each vertex \( x \), pick \( A_x x \in Ax \) and define \( A_x \) on all of \( \mathbb{C}^n \) by the following rule. If
\[
x \in [x_0, \ldots, x_{2n}],
\]
so \( x = \sum_{i=0}^{2n} t_i x_i \), then
\[
A_x x \equiv \sum_{k=0}^{2n} t_k A_x x_k.
\]
Thus $A_e$ is a continuous map defined on $\mathbb{C}^n$ thanks to the local finiteness of the collection of simplices. Let $P_K$ denote the projection on the convex set $K$. By the Brouwer fixed point theorem, there exists a fixed point, $x_e \in K$ such that

$$P_K (y - A_e x_e + x_e) = x_e.$$  

By Corollary C.1.9 this requires

$$\text{Re} (y - A_e x_e, z - x_e) \leq 0$$

for all $z \in K$.

Suppose $x_e \in [x_0, \ldots, x_{2n}]$ so $x_e = \sum_{k=0}^{2n} t_k x_k$. Then since $x_e$ is contained in $K$, a compact set, and the diameter of each simplex is less than 1, it follows that $A_e x_k^e$ is contained in $A(K + B(0,1))$, which is contained in a compact set thanks to Lemma 7.2.1. The reason is that $A$ is assumed to take bounded sets to bounded sets and $K + B(0,1)$ is a bounded set.

From the Heine Borel theorem, there exists a sequence $\varepsilon \rightarrow 0$ such that

$$t_k^e \rightarrow t_k, x_e \rightarrow x \in K, A_e x_k^e \rightarrow y_k$$

for $k = 0, \ldots, 2n$. Since the diameter of the simplex containing $x_e$ converges to 0, it follows

$$x_k^e \rightarrow x, A_e x_k^e \rightarrow y_k.$$  

By upper semicontinuity, it follows that for all $r > 0$, $A x_k^e \subseteq A x + B(0,r)$ for all $\varepsilon$ small enough. Since $A_e x_k^e \in A x_k^e$, and $A x$ is closed, this implies $y_k \in A x$. Since $A x$ is convex,

$$\sum_{k=1}^{2n} t_k y_k \in A x.$$  

Hence for all $z \in K$,

$$\left( y - \sum_{k=1}^{2n} t_k y_k, z - x \right) = \lim_{\varepsilon \rightarrow 0} \left( y - \sum_{k=1}^{2n} t_k A_e x_k^e, z - x_e \right)$$

$$= \lim_{\varepsilon \rightarrow 0} (y - A_e x_e, z - x_e) \leq 0.$$  

Let $w = \sum_{k=1}^{2n} t_k y_k$. 

If you replaced $A$ with $A \circ P_K$ where $A$ is only defined on $K$, then if you carried out the above argument you would get the same result because $x \in K$. Thus there is an obvious generalization. You only need to have $A$ defined on $K$.

**Lemma 9.4.2** Suppose in addition to 7.2.8 and 7.2.9, (compact convex valued and upper semicontinuous) $A$ is coercive,

$$\lim_{|x| \rightarrow \infty} \inf \left\{ \frac{\text{Re} (y, x)}{|x|} : y \in A x \right\} = \infty.$$  

Then $A$ is onto.
9.4. SET-VALUED MAPS, PSEUDOMONOTONE OPERATORS

Proof: Let \( y \in \mathbb{C}^n \) and let \( K_r \equiv B(0, r) \). By Lemma 9.4.1 there exists \( x_r \in K_r \) and \( w_r \in Ax_r \) such that

\[
\text{Re} \left( y \right) \left( z - x_r \right) \leq 0 \tag{9.4.11}
\]

for all \( z \in K_r \). Letting \( z = 0 \),

\[
\text{Re} \left( w_r \right) \leq \text{Re} \left( y, x_r \right).
\]

Therefore,

\[
\inf \left\{ \frac{\text{Re} \left( w \right)}{|x_r|} \mid w \in Ax_r \right\} \leq |y|.
\]

It follows from the assumption of coercivity that \( |x_r| \) is bounded independent of \( r \). Therefore, picking \( r \) strictly larger than this bound, \( 9.4.11 \) implies

\[
\text{Re} \left( y - w_r, v \right) \leq 0
\]

for all \( v \) in some open ball containing \( 0 \). Therefore, for all \( v \) in this ball

\[
\text{Re} \left( y - w_r, v \right) = 0
\]

and hence this holds for all \( v \in \mathbb{C}^n \) and so \( y = w_r \in Ax_r \). ■

Lemma 9.4.3 Let \( F \) be a finite dimensional Banach space of dimension \( n \), and let \( T \) be a mapping from \( F \) to \( P \left( F' \right) \) such that \( 7.2.8 \) and \( 7.2.9 \) both hold for \( F' \) in place of \( C^n \). Then if \( T \) is also coercive,

\[
\lim_{||u|| \to \infty} \inf \left\{ \frac{\text{Re} \left( y^* \left( u \right) \right)}{||u||} \mid y^* \in Tu \right\} = \infty, \tag{9.4.12}
\]

it follows \( T \) is onto.

Proof: Let \( || \cdot || \) be an equivalent norm for \( F \) such that there is an isometry of \( \mathbb{C}^n \) and \( F, \theta \). Now define \( A : \mathbb{C}^n \to P \left( \mathbb{C}^n \right) \) by \( Ax \equiv \theta^* T \theta x \).

\[
\begin{array}{ccc}
P \left( F' \right) & \theta^* & \mathbb{C}^n \\
T & \uparrow & \uparrow A \\
F & \theta & \mathbb{C}^n
\end{array}
\]

Thus \( y \in Ax \) means that there exists \( z^* \in T \theta x \) such that

\[
\left( w, y \right)_{C^n} = z^* \left( \theta w \right)
\]

for all \( w \in \mathbb{C}^n \). Then \( A \) satisfies the conditions of Lemma 9.4.1 and so \( A \) is onto. Consequently \( T \) is also onto. ■

With these lemmas, it is possible to prove a very useful result about a class of mappings which map a reflexive Banach space to the power set of its dual space. For more theorems about these mappings and their applications, see [10]. In the discussion below, we will use the symbol, \( \rightarrow \), to denote weak convergence.
Definition 9.4.4 Let $V$ be a Reflexive Banach space. We say $T : V \rightarrow \mathcal{P}(V')$ is pseudomonotone if the following conditions hold.

\[ Tu \text{ is closed, nonempty, convex.} \tag{9.4.13} \]

If $F$ is a finite dimensional subspace of $V$, then if $u \in F$ and $W \supseteq Tu$ for $W$ a weakly open set in $V'$, then there exists $\delta > 0$ such that

\[ v \in B(u, \delta) \cap F \text{ implies } Tv \subseteq W. \tag{9.4.14} \]

If $u_k \rightharpoonup u$ and if $u_k^* \in Tu_k$ is such that

\[ \limsup_{k \to \infty} u_k^*(u_k - u) \leq 0, \]

then for all $v \in V$, there exists $u^*(v) \in Tu$ such that

\[ \liminf_{k \to \infty} u_k^*(u_k - v) \geq u^*(v)(u - v). \tag{9.4.15} \]

We say $T$ is coercive if

\[ \lim_{||v|| \to \infty} \inf \left\{ \frac{z^*(v)}{||v||} : z^* \in Tu \right\} = \infty. \tag{9.4.16} \]

In the case that $T$ takes bounded sets to bounded sets so it is a bounded set valued operator, it turns out you don’t have to consider the second of the above conditions about the upper semicontinuity. It follows from the other conditions. It is convenient to use the notation

\[ \langle u^*, v \rangle \equiv u^*(v), u^* \in V', v \in V. \]

and this will be used interchangably with the earlier notation from now on.

Lemma 9.4.5 Let $T : X \rightarrow \mathcal{P}(X')$ satisfy conditions $9.4.13$ and $9.4.15$ above and suppose $T$ is bounded. Then if $x_n \rightharpoonup x$ in $X$, and if $U$ is a weakly open set containing $Tx$, then $Tx_n \subseteq U$ for all $n$ large enough. If fact the limit condition $9.4.13$ can be weakened to the following more general condition: If $u_k \rightharpoonup u$, then there exists a subsequence still denoted as $\{u_k\}$, such that if $u_k^* \in Tu_k$ satisfies

\[ \limsup_{k \to \infty} u_k^*(u_k - u) \leq 0, \]

then for all $v \in V$, there exists $u^*(v) \in Tu$ such that

\[ \liminf_{k \to \infty} u_k^*(u_k - v) \geq u^*(v)(u - v). \tag{9.4.17} \]

In other words, a convergent sequence has a subsequence for which the pseudomonotone limit condition holds.
9.4. SET-VALUED MAPS, PSEUDOMONOTONE OPERATORS

Proof: If this is not true, there exists \(x_n \to x\), a weakly open set \(U\), containing \(Tx\) and \(z_n \in Tx_n\), but \(z_n \notin U\). Taking a subsequence if necessary, we obtain a sequence which satisfies the limit condition that if the \(\limsup\) is bounded above by 0, then the condition on the \(\liminf\) holds for that sequence, and \(z_n \to z \notin U\) in addition to this. Then

\[
\limsup_{n \to \infty} \langle z_n, x_n - x \rangle = 0
\]

so if \(y \in X\) there exists \(z(y) \in Tx\) such that

\[
\langle z, x - y \rangle = \liminf_{n \to \infty} \langle z_n, x_n - y \rangle \geq \langle z(y), x - y \rangle.
\]

Letting \(w = x - y\), this shows, since \(y \in X\) is arbitrary, that the following inequality holds for every \(w \in X\).

\[
\langle z, w \rangle \geq \langle z(x - w), w \rangle, \quad z(x - w) \in Tx.
\]

In particular, we may replace \(w\) with \(-w\) and obtain

\[
\langle z, -w \rangle \geq \langle z(x + w), -w \rangle,
\]

which implies

\[
\langle z(x - w), w \rangle \leq \langle z, w \rangle \leq \langle z(x + w), w \rangle.
\]

Therefore, there exists

\[
\lambda \in \mathbb{R}
\]

such that

\[
\langle z, w \rangle = \langle z(y), w \rangle.
\]

But this is a contradiction to \(z \notin Ax\) because if \(z \notin Ax\) there exists \(w \in X\) such that for all \(z_1 \in Ax\),

\[
\langle z, w \rangle > \langle z_1, w \rangle.
\]

Therefore, \(z \in Ax\) which contradicts the assumption that \(z_n\) and consequently \(z\) are not contained in \(U\). ■

This more general limit condition is sometimes useful if not essential to use. Here is a definition of something slightly different than the above.

**Definition 9.4.6** Say \(T : V \to \mathcal{P}(V')\) is modified bounded pseudomonotone if the following conditions hold.

\[
Tu \text{ is closed, nonempty, convex.} \tag{9.4.18}
\]

\(T\) is bounded meaning it takes bounded sets to bounded sets.

If \(u_k \to u\) then there exists a subsequence, still denoted as \(\{u_k\}\) such that if \(u_k^* \in Tu_k\) and

\[
\limsup_{k \to \infty} u_k^*(u_k - u) \leq 0,
\]

then for all \(v \in V\), there exists \(u^*(v) \in Tu\) such that

\[
\liminf_{k \to \infty} u_k^*(u_k - v) \geq u^*(v)(u - v). \tag{9.4.19}
\]
One of the nice properties of pseudomonotone maps is that when you add two of them, you get another one. I will give a proof in the case that the two pseudomonotone maps are both bounded. It is probably true in general, but as just noted, it is less trouble to verify if you don’t have to worry about as many conditions. I will also assume the spaces are all real so it will not be necessary to constantly write the real part.

**Theorem 9.4.7** Suppose $A, B : X \to P(X')$ are both pseudomonotone and bounded. Then so is their sum. If $A$ is modified bounded pseudomonotone and $B$ is bounded pseudomonotone, then $A + B$ is modified bounded pseudomonotone.

**Proof:** It is clear that $Ax + Bx$ is closed and convex because this is true of both of the sets in the sum. It is also bounded because both terms in the sum are bounded. It only remains to verify the limit condition. Suppose then that $u_n \to u$ weakly

Will the limit condition hold for $A + B$ when applied to this further subsequence? Suppose $z_n \in Ax_n, w_n \in Bx_n$ and

$$\lim \sup_{n \to \infty} \langle z_n + w_n, u_n - u \rangle \leq 0$$

Will the lim inf condition hold? From the above,

$$\lim \sup_{n \to \infty} \langle z_n + w_n, u_n - u \rangle \leq \lim \sup_{n \to \infty} \langle z_n, u_n - u \rangle + \lim \sup_{n \to \infty} \langle w_n, u_n - u \rangle \quad (9.4.20)$$

and so, if the second term $\leq 0$, you would have, since $B$ is modified bounded pseudomonotone,

$$\lim \inf_{n \to \infty} \langle w_n, u_n - u \rangle \geq \langle w(u), u - u \rangle = 0$$

Hence you would have $\lim \inf_{n \to \infty} \langle w_n, u_n - u \rangle \geq 0 \geq \lim \sup_{n \to \infty} \langle w_n, u_n - u \rangle$ and so $\lim_{n \to \infty} \langle w_n, u_n - u \rangle = 0$. Hence

$$\lim \sup_{n \to \infty} \langle z_n + w_n, u_n - u \rangle = \lim \sup_{n \to \infty} \langle z_n, u_n - u \rangle \leq 0$$

Then using that $A$ is pseudomonotone, $\lim_{n \to \infty} \langle z_n, u_n - u \rangle = 0$ also. Then from this it is routine to establish the pseudomonotone limit condition for the sum $A + B$. In fact, you would have for any $v$,

$$\lim \inf_{n \to \infty} \langle z_n, u_n - v \rangle \geq \langle z(u), u - v \rangle$$

$$\lim \inf_{n \to \infty} \langle w_n, u_n - v \rangle \geq \langle w(u), u - v \rangle$$

Then you would get

$$\lim \inf_{n \to \infty} \langle z_n + w_n, u_n - v \rangle = \lim \inf_{n \to \infty} (\langle z_n, u_n - v \rangle + \langle w_n, u_n - v \rangle) \geq \lim \inf_{n \to \infty} (\langle z_n, u_n - v \rangle) + \lim \inf_{n \to \infty} (\langle w_n, u_n - v \rangle) \geq \langle z(u), u - v \rangle + \langle w(u), u - v \rangle$$
and \( z(u) + w(u) \in (A + B)(u) \). Thus the limit condition will hold if either \( \limsup_{n \to \infty} \langle z_n, u_n - u \rangle \) or \( \limsup_{n \to \infty} \langle w_n, u_n - u \rangle \) is \( \leq 0 \). Therefore, if the limit condition fails, you must have both of these strictly positive. Take a subsequence, still denoted with subscript \( n \), such that

\[
\limsup_{n \to \infty} \langle z_n, u_n - u \rangle = \lim_{n \to \infty} \langle z_n, u_n - u \rangle = \delta > 0.
\]

Then, using this subsequence (\( \limsup \) gets smaller when you go to a subsequence.), and the assumption,

\[
0 \geq \limsup_{n \to \infty} \langle z_n + w_n, u_n - u \rangle = \delta + \limsup_{n \to \infty} \langle w_n, u_n - u \rangle \geq \delta
\]

which is a contradiction. Thus the case where both are strictly positive cannot occur. The last claim follows from first taking a subsequence for which the pseudomonotone limit condition holds for \( A \) and then repeating the argument.

It is not entirely clear whether the sum of modified bounded pseudomonotone operators is modified bounded pseudomonotone. This is because when you go to a subsequence, the \( \limsup \) gets smaller and so it is not entirely clear whether the subsequence for \( A \) will continue to yield the limit condition if a further subsequence is taken.

The following is mostly in [46].

**Theorem 9.4.8** Let \( V \) be a reflexive Banach space and let \( T : V \to \mathcal{P}(V') \) be pseudomonotone, bounded, and coercive. Then \( T \) is onto. More generally, the same holds if \( T \) is modified bounded pseudomonotone.

**Proof:** The proof is for modified bounded pseudomonotone since this is more general. Let \( F \) be the set of finite dimensional subspaces of \( V \) and let \( F \in \mathcal{F} \). Then define \( T_F \) as

\[
T_F \equiv i_F^* Ti_F
\]

where here \( i_F \) is the identity map from \( F \) to \( V \). Then \( T_F \) satisfies the conditions of Lemma 9.4.5 thanks to Lemma 9.4.6 and so \( T_F \) is onto \( \mathcal{P}(F') \). Let \( w^* \in V' \). Then since \( T_F \) is onto, there exists \( u_F \in F \) such that

\[
i_F^* w^* \in i_F^* Ti_F u_F.
\]

Thus for each finite dimensional subspace \( F \), there exists \( u_F \in F \) such that for all \( v \in F \),

\[
\langle w^*, v \rangle = \langle u_F^*, v \rangle, \quad u_F^* \in Tu_F.
\]

(9.4.21)

Replacing \( v \) with \( u_F \), in (9.4.21),

\[
\frac{\langle u_F^*, u_F \rangle}{||u_F||} = \frac{\langle w^*, u_F \rangle}{||u_F||} \leq ||w^*||.
\]

Therefore, the assumption that \( T \) is coercive implies \( \{u_F : F \in \mathcal{F} \} \) is bounded in \( V \). Now define

\[
W_F \equiv \cup \{u_{F'} : F' \supseteq F \}.
\]
Then $W_F$ is bounded and if $\overline{W_F} = \text{weak closure of } W_F$, then

$$\{ \overline{W_F} : F \in \mathcal{F} \}$$

is a collection of nonempty weakly compact (since $V$ is reflexive and the $u_F$ were just shown bounded) sets having the finite intersection property because $W_F \neq \emptyset$ for each $F$. (If $F_i, i = 1, \cdots, n$ are finite dimensional subspaces, let $F$ be a finite dimensional subspace which contains all of these. Then $W_F \neq \emptyset$ and $W_F \subseteq \cap_{i=1}^n \overline{W_{F_i}}$.) Thus there exists

$$u \in \cap \{ \overline{W_F} : F \in \mathcal{F} \}.$$

I will show $w^* \in Tu$. If $w^* \notin Tu$, a closed convex set, there exists $v \in V$ such that

$$\langle w^*, u - v \rangle < \langle u^*, u - v \rangle$$

(9.4.22) for all $u^* \in Tu$. This follows from the separation theorems. (These theorems imply there exists $z \in V$ such that

$$\langle w^*, z \rangle < \langle u^*, z \rangle$$

for all $u^* \in Tu$. Define $u - v \equiv z$.)

Now let $F \supseteq \{u, v\}$. Since $u \in \overline{W_F}$, a weakly sequentially compact set, there exists a sequence, $\{u_k\}$, such that

$$u_k \rightharpoonup u, \ u_k \in W_F.$$

and in addition, for this subsequence, the pseudomonotone limit condition holds.

Then since $F \supseteq \{u, v\}$, there exists $u_k^* \in Tu_k$ such that

$$\langle u_k^*, u_k - u \rangle = \langle w^*, u_k - u \rangle.$$

Therefore,

$$\limsup_{k \to \infty} \langle u_k^*, u_k - u \rangle = \limsup_{k \to \infty} \langle w^*, u_k - u \rangle = 0.$$

It follows by the assumption that $T$ is modified bounded pseudomonotone and the pseudomonotone limit condition holds for this sequence; the following holds for the $v$ defined above in 9.4.22.

$$\liminf_{k \to \infty} \langle u_k^*, u_k - v \rangle \geq \langle u^*(v), u - v \rangle, \ u^*(v) \in Tu.$$

But since $v \in F$, $\langle u_k^*, u_k - v \rangle = \langle w^*, u_k - v \rangle$ and so

$$\liminf_{k \to \infty} \langle u_k^*, u_k - v \rangle = \liminf_{k \to \infty} \langle w^*, u_k - v \rangle = \langle w^*, u - v \rangle,$$

so from 9.4.22 $\langle w^*, u - v \rangle < \langle u^*, u - v \rangle$ for all $u^* \in Tu$,

$$\langle w^*, u - v \rangle = \liminf_{k \to \infty} \langle u_k^*, u_k - v \rangle \geq \langle u^*(v), u - v \rangle > \langle w^*, u - v \rangle,$$

a contradiction. Thus, $w^* \in Tu$. ■
9.5 Generalized Gradients

This is an interesting theorem, but one might wonder if there are easy to verify examples of such possibly set valued mappings. In what follows consider only real spaces because the essential ideas are included in this case which is also the case of most use in applications. Of course, you might with some justification, make the claim that the following is not really very easy to verify any more than the original definition.

**Definition 9.5.1** Let \( V \) be a real reflexive Banach space and let \( f : V \to \mathbb{R} \) be a locally Lipschitz function, meaning that \( f \) is Lipschitz near every point of \( V \) although \( f \) need not be Lipschitz on all of \( V \). Under these conditions,

\[
f^0 (x, y) = \lim_{\mu \to 0+} \sup_{h \to 0} \frac{f(x + h + \mu y) - f(x + h)}{\mu}
\]

(9.5.23)

and \( \partial f(x) \subseteq X' \) is defined by

\[
\partial f(x) \equiv \{ x^* \in X' : x^* (y) \leq f^0 (x, y) \text{ for all } y \in X \}.
\]

(9.5.24)

The set just described is called the generalized gradient. In \( 9.5.23 \) we mean the following by the right hand side.

\[
\lim_{(r, \delta) \to (0,0)} \sup \left\{ \frac{f(x + h + \mu y) - f(x + h)}{\mu} : \mu \in (0, r), h \in B(0, \delta) \right\}
\]

I will show, following [46], that these generalized gradients of locally Lipschitz functions are sometimes pseudomonotone. First here is a lemma.

**Lemma 9.5.2** Let \( f \) be as described in the above definition. Then \( \partial f(x) \) is a closed, bounded, convex, and non empty subset of \( V' \). Furthermore, for \( x^* \in \partial f(x) \),

\[
\|x^*\| \leq \text{Lip}_x (f).
\]

(9.5.25)

**Proof:** It is left as an exercise to verify the assertions that \( \partial f(x) \) is closed, and convex. It follows directly from the definition. To verify this set is bounded, let \( \text{Lip}_x (f) \) denote a Lipschitz constant valid near \( x \in V \) and let \( x^* \in \partial f(x) \). Then choosing \( y \) with \( \|y\| = 1 \) and \( x^* (y) \geq \frac{1}{2} \|x^*\| \),

\[
\frac{1}{2} \|x^*\| = x^* (y) \leq f^0 (x, y).
\]

(9.5.26)

Also, for small \( \mu \) and \( h \),

\[
\left| \frac{f(x + h + \mu y) - f(x + h)}{\mu} \right| \leq \text{Lip}_x (f) \|y\| = \text{Lip}_x (f).
\]

Therefore, \( f^0 (x, y) \leq \text{Lip}_x (f) \) and so \( 9.5.20 \) shows \( \|x^*\| \leq 2 \text{Lip}_x (f) \).
The interesting part of this Lemma is that $\partial f (x) \neq \emptyset$. To verify this first note that the definition of $f^0$ implies that $y \to f^0 (x, y)$ is a gauge function. Now fix $y \in \mathbb{V}$ and define on $\mathbb{R}$ a linear map $x_0^*$ by

$$ x_0^* (\alpha y) \equiv \alpha f^0 (x, y). $$

Then if $\alpha \geq 0$,

$$ x_0^* (\alpha y) = \alpha f^0 (x, y) = f^0 (x, \alpha y). $$

If $\alpha < 0$,

$$ x_0^* (\alpha y) = \alpha f^0 (x, y) = \lim_{\mu \to 0^+, h \to 0} \frac{(-\alpha) f (x + h) - (-\alpha) f (x + h + \mu y)}{\mu}, $$

$$ (-\alpha) \lim_{\mu \to 0^+, h \to 0} \frac{f (x + h - \mu y) - f (x + h)}{\mu} \leq (-\alpha) f^0 (x, -y) = f^0 (x, \alpha y). $$

Therefore, $x_0^* (\alpha y) \leq f^0 (x, \alpha y)$ for all $\alpha$. By the Hahn Banach theorem there is an extension of $x_0^*$ to all of $\mathbb{V}$, $x^*$ which satisfies,

$$ x^* (y) \leq f^0 (x, y) $$

for all $y$. It remains to verify $x^*$ is continuous. This follows easily from

$$ |x^* (y)| = \max (x^* (-y), x^* (y)) \leq \max (f^0 (x, y), f^0 (x, -y)) \leq \text{Lip}_x (f) || y ||, $$

which verifies \textit{9.5.25} and proves the lemma.

This lemma has verified the first condition needed in the definition of pseudomonotone. The next lemma verifies that these generalized subgradients satisfy the second of the conditions needed in the definition. In fact somewhat more than is needed in the definition is shown.

\textbf{Lemma 9.5.3} Let $U$ be weakly open in $\mathbb{V}'$ and suppose $\partial f (x) \subseteq U$. Then $\partial f (z) \subseteq U$ whenever $z$ is close enough to $x$.

\textbf{Proof:} Suppose to the contrary there exists $z_n \to x$ but $z_n^* \in \partial f (z_n) \setminus U$. From the first lemma, we may assert that $|| z_n^* || \leq 2\text{Lip}_x (f)$ for all $n$ large enough. Therefore, there is a subsequence, still denoted by $n$ such that $z_n^*$ converges weakly to $z^* \notin U$.

\textbf{Claim:} $f^0 (x, y) \geq \limsup_{n \to \infty} f^0 (x_n, y)$.

\textbf{Proof of the claim:} There exists $\delta > 0$ such that if $\mu, || h || < \delta$, then

$$ \epsilon + f^0 (x, y) \geq \frac{f (x + h + \mu y) - f (x + h)}{\mu}. $$


Thus, for $\|h\| < \delta$,

$$\varepsilon + f^0(x, y) \geq \frac{f(x_n + (x - x_n) + h + \mu y) - f(x_n + (x - x_n) + h)}{\mu}.$$ 

Now let $\|h'\| < \frac{\delta}{2}$ and let $n$ be so large that $\|x - x_n\| < \frac{\delta}{2}$. Suppose $\|h'\| < \frac{\delta}{2}$. Then choosing $h \equiv h' - (x - x_n)$, it follows the above inequality holds because $\|h\| < \delta$. Therefore, if $\|h'\| < \frac{\delta}{2}$, and $n$ is sufficiently large,

$$\varepsilon + f^0(x, y) \geq \frac{f(x_n + h' + \mu y) - f(x_n + h')}{\mu}.$$ 

Consequently, for all $n$ large enough,

$$\varepsilon + f^0(x, y) \geq f^0(x_n, y)$$

which proves the claim.

Now with the claim,

$$z^*(y) = \limsup_{n \to \infty} z_n^*(y) \leq \limsup_{n \to \infty} f^0(x_n, y) \leq f^0(x, y)$$

so $z^* \in \partial f(x)$ contradicting the assumption that $z^* \notin U$. Thus, it is necessary to assume more on $f^0$ in order to obtain the third axiom defining pseudomonotone. The following theorem describes the situation.

**Theorem 9.5.4** Let $f : V \to V'$ be locally Lipschitz and suppose it satisfies the condition that whenever $x_n$ converges weakly to $x$

and

$$\limsup_{n \to \infty} f^0(x_n, x - x_n) \geq 0$$

it follows that

$$\limsup_{n \to \infty} f^0(x_n, z - x_n) \leq f^0(x, z - x)$$

for all $z \in V$. Then $\partial f$ is pseudomonotone.

**Proof:** We will verify both are satisfied thanks to Lemmas 9.4.1 and 9.4.2. It remains to verify 9.4.5. To do so, I will adopt the convention that $x^* \in \partial f(x)$. Suppose

$$\limsup_{n \to \infty} x^*_n (x_n - x) \leq 0.$$ 

This implies $\liminf_{n \to \infty} x^*_n (x_n - x) \geq 0$. Thus,

$$0 \leq \liminf_{n \to \infty} x^*_n (x_n - x) \leq \liminf_{n \to \infty} f^0(x_n, x - x_n) \leq \limsup_{n \to \infty} f^0(x_n, x - x_n),$$

where

$$\limsup_{n \to \infty} f^0(x_n, x - x_n).$$
which implies, by the above assumption that for all $z$,
\[
\limsup x_n^* (z - x_n) \leq \limsup f^0 (x_n, z - x_n) \leq f^0 (x, z - x). \tag{9.5.28}
\]
In particular, this holds for $z = x$ and this implies $\limsup x_n^* (x - x_n) \leq 0$ which along with (9.5.27) yields
\[
\lim_{n \to \infty} x_n^* (x_n - x) = 0 \tag{9.5.29}
\]
Now let $z$ be arbitrary. There exists a subsequence, $n_k$, depending on $z$ such that
\[
\lim_{k \to \infty} x_{n_k}^* (x_{n_k} - z) = \liminf x_{n_k}^* (x_{n_k} - z).
\]
Now from Lemma 9.5.2 and its proof, the $||x_n^*||$ are all bounded by $\text{Lip}_x (f)$ whenever $n$ is large enough. Therefore, there is a further subsequence, still denoted by $n_k$ such that $x_{n_k}^*$ converges weakly to $x^* (z)$.

We need to verify that $x^* (z) \in \partial f (x)$. To do so, let $y$ be arbitrary. Then from the definition,
\[
x_n^* (y - x_n) \leq f^0 (x_n, y - x_n). \tag{9.5.30}
\]
From (9.5.29), we can take the lim sup of both sides and obtain, using (9.5.28)
\[
x^* (z) (y - x) \leq \limsup f^0 (x_n, y - x_n) \leq f^0 (x, y - x).
\]
Since $y$ is arbitrary, this shows $x^* (z) \in \partial f (x)$ and proves the theorem.

## 9.6 Maximal Monotone Operators

Here it is assumed that the spaces are all real spaces to simplify the presentation.

**Definition 9.6.1** Let $A : D (A) \subseteq X \to \mathcal{P} (X)$ be a set valued map. It is said to be monotone if whenever $y_i \in Ax_i$,
\[
\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0
\]
Denote by $\mathcal{G} (A)$ the graph of $A$ consisting of all pairs $(x, y)$ where $y \in Ax$. Such a monotone operator is said to be maximal monotone if
\[
F + A
\]
is onto where $F$ is the duality map with $p = 2$.

Actually, it is more usual to say that the graph is maximal monotone if the graph is monotone and there is no monotone graph which properly contains the given graph. However, the two conditions are equivalent and I am more used to using the version in the above definition.

There is a fundamental result about these which is given next.
Theorem 9.6.2 Let \( X, X' \) be reflexive and have strictly convex norms. Let \( A \) be a monotone set valued map as just described. Then if \( \lambda F + A \) is onto for some \( \lambda > 0 \), then whenever
\[
\langle y - z, x - u \rangle \geq 0 \text{ for all } [x, y] \in \mathcal{G}(A)
\]
it follows that \( z \in Au \text{ and } u \in D(A) \). That is, the graph is maximal.

Proof: Suppose that for all \([x, y] \in \mathcal{G}(A)\),
\[
\langle y - z, x - t \rangle \geq 0
\]
Does it follow that \( z \in At \)? By assumption, \( z + \lambda F(t) = \lambda F\hat{x} + \hat{\xi}, \hat{\xi} \in A\hat{x} \). Then replacing \( y \) with \( \hat{\xi} \) and \( x \) with \( \hat{x} \),
\[
\langle \hat{\xi} - (\lambda F\hat{x} + \hat{\xi} - \lambda Ft), \hat{x} - t \rangle \geq 0
\]
and so
\[
\lambda \langle Ft - F\hat{x}, t - \hat{x} \rangle \leq 0
\]
which implies from Theorem 9.2.5 that \( t = \hat{x} \) and so the graph of \( A \) is indeed maximal monotone.
\[
z + \lambda F(t) = \lambda F\hat{x} + \hat{\xi} \Rightarrow z = \hat{\xi} \in A\hat{x} = At
\]

Note that this would have worked with no change if the duality map had been for arbitrary \( p > 1 \).

9.6.1 Maximal Monotone Operators

In fact, these two conditions are equivalent. This is shown in [4]. We give a proof of this here.

Next is the theorem about the graph being maximal being equivalent to the operator being maximal monotone. It is a very convenient result to have. The proof is a modified version of one in Barbu [3]. I was unable to completely follow his argument. It is based on the following lemma also in Barbu. This is a little like the Browder lemma but is based on the min max theorem above. It is also a very interesting argument.

Lemma 9.6.3 Let \( E \) be a finite dimensional Banach space and let \( K \) be a convex and compact subset of \( E \). Let \( \mathcal{G}(A) \) be a monotone subset of \( E \times E' \) such that \( D(A) \subseteq K \) and \( B \) is a single valued monotone and continuous operator from \( E \) to \( E' \). Then there exists \( x \in K \) such that
\[
\langle Bx + v, u - x \rangle_{E', E} \geq 0 \text{ for all } [u, v] \in \mathcal{G}(A).
\]
If \( B \) is coercive
\[
\lim_{\|x\| \to \infty} \frac{\langle Bx, x \rangle}{\|x\|} = \infty,
\]
and \( 0 \in D(A) \), then one can assume only that \( K \) is convex and closed.
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

Proof: Let $T : E \to K$ be the multivalued operator defined by

$$Ty \equiv \left\{ x \in K : \langle By + v, u - x \rangle_{E', E} \geq 0 \text{ for all } [u, v] \in \mathcal{G}(A) \right\}$$

Here $y \in E$ and it is desired to show that $Ty \neq \emptyset$ for all $y \in K$. For $[u, v] \in \mathcal{G}(A)$, let

$$K_{u,v} = \left\{ x \in K : \langle By + v, u - x \rangle_{E', E} \geq 0 \right\}$$

Then $K_{u,v}$ is a closed, hence compact subset of $K$. The thing to do is to show that $\bigcap_{u,v} K_{u,v} \equiv Ty \neq \emptyset$ whenever $y \in K$. Then one argues that $T$ is set valued, has convex compact values and is upper semicontinuous. Then one applies the Kakutani fixed point theorem to get $x \in Tx$.

Since these sets $K_{u,v}$ are compact, it suffices to show that they satisfy the finite intersection property. Thus for $\{[u_i, v_i]\}_{i=1}^n$ a finite set of elements of $\mathcal{G}(A)$, it is necessary to show that there exists a solution $x$ to the inequalities

$$\langle u_i - x, By + v_i \rangle \geq 0, \quad i = 1, 2, \ldots, n$$

and then it follows from finite intersection property that there exists $x \in \bigcap_{[u,v] \in \mathcal{G}(A)} K_{u,v}$ which is what was desired. Let $P_n$ be all $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n)$ such that each $\lambda_k \geq 0$ and $\sum_{k=1}^n \lambda_k = 1$. Let $H : P_n \times P_n \to \mathbb{R}$ be given by

$$H(\vec{\mu}, \vec{\lambda}) \equiv \sum_{i=1}^n \mu_i \left( \langle By + v_i, \sum_{j=1}^n \lambda_j u_j - u_i \rangle \right) \quad (9.6.31)$$

Then this is both convex and concave in both $\vec{\lambda}, \vec{\mu}$ and so by Theorem [7.5.4], there exists $\vec{\mu}_0, \vec{\lambda}_0$ both in $P_n$ such that for all $\vec{\mu}, \vec{\lambda}$,

$$H(\vec{\mu}, \vec{\lambda}_0) \leq H(\vec{\mu}_0, \vec{\lambda}_0) \leq H(\vec{\mu}_0, \vec{\lambda}) \quad (9.6.32)$$

However, plugging in $\vec{\mu} = \vec{\lambda}$ in (9.6.31),

$$H(\vec{\lambda}, \vec{\lambda}) = \sum_{i=1}^n \lambda_i \left( \langle By + v_i, \sum_{j=1}^n \lambda_j u_j - u_i \rangle \right) = \sum_{i=1}^n \left( \langle By + v_i, \sum_{j=1}^n \lambda_i \lambda_j u_j - \lambda_i u_i \rangle \right)$$

$$= \sum_{i=1}^n \left( \langle By + v_i, \sum_{j=1}^n (\lambda_i \lambda_j u_j - \lambda_i \lambda_j u_i) \rangle \right) = \left( \langle By_i, \sum_{i=1}^n \sum_{j=1}^n (\lambda_i \lambda_j u_j - \lambda_i \lambda_j u_i) \rangle \right) + \sum_{i=1}^n \left( \langle v_i, \sum_{j=1}^n (\lambda_i \lambda_j u_j - \lambda_i \lambda_j u_i) \rangle \right)$$
9.6. MAXIMAL MONOTONE OPERATORS

The first term obviously equals 0. Consider the second. This term equals
\[ \sum_i \sum_j \lambda_i \lambda_j \langle v_i, (u_j - u_i) \rangle \]
The terms equal 0 when \( j = i \) or they come in pairs
\[
\lambda_i \lambda_j \langle v_i, (u_j - u_i) \rangle + \lambda_i \lambda_j \langle (v_i, (u_j - u_i)) \rangle
\]
\[
= \lambda_i \lambda_j \langle (v_i, (u_j - u_i)) \rangle - \langle v_j, (u_j - u_i) \rangle \leq 0
\]
by monotonicity of \( A \). Hence \( H(\vec{\lambda}, \vec{\lambda}) \leq 0 \). Then from \[ \ref{9.6.32} \], for all \( \vec{\mu} \)
\[
H(\vec{\mu}, \vec{\lambda}_0) \leq H(\vec{\mu}_0, \vec{\lambda}_0) \leq H(\vec{\mu}_0, \vec{\mu}_0) \leq 0
\]
It follows that
\[
\sum_{i=1}^m \mu_i \left( By + v_i, \sum_{j=1}^n \lambda^0_i u_j - u_i \right) \leq 0
\]
\[
\sum_{i=1}^m \mu_i \left( By + v_i, u_i - \sum_{j=1}^n \lambda^0_j u_j \right) \geq 0
\]
where \( \vec{\lambda}_0 \equiv (\lambda^0_1, \cdots, \lambda^0_n) \). This is true for any choice of \( \vec{\mu} \). In particular, you could let \( \vec{\mu} \) equal 1 in the \( i^{th} \) position and 0 elsewhere and conclude that for all \( i = 1, \cdots, n, \)
\[
\left( By + v_i, u_i - \sum_{j=1}^n \lambda^0_j u_j \right) \geq 0
\]
so you let \( x = \sum_{j=1}^n \lambda^0_j u_j \) and this shows that \( Ty \neq \emptyset \) because the sets \( K_{u,v} \) have the finite intersection property.

Thus \( T : K \rightarrow \mathcal{P}(K) \) and for each \( y \in K, Ty \neq \emptyset \). In fact this is true for any \( y \) but we are only considering \( y \in K \). Now \( Ty \) is clearly a closed subset of \( K \). It is also clearly convex. Is it upper semicontinuous? Let \( y_k \to y \) and consider \( Ty_k + B(0, r) \). Is \( Ty_k \in Ty + B(0, r) \) for all \( k \) large enough? If not, then there is a subsequence, denoted as \( z_k \in Ty_k \) which is outside this open set \( Ty + B(0, r) \). Then taking a further subsequence, still denoted as \( z_k \), it follows that \( z_k \to z \notin Ty + B(0, r) \). Now
\[
\langle By_k + v, u - z_k \rangle \geq 0 \text{ all } [u, v] \in \mathcal{G}(A)
\]
Therefore, from continuity of \( B, \)
\[
\langle By + v, u - z \rangle \geq 0 \text{ all } [u, v] \in \mathcal{G}(A)
\]
which means \( z \in Ty \) contrary to the assumption that \( T \) is not upper semicontinuous. Since \( T \) is upper semicontinuous and maps to compact convex sets, it follows from Theorem \[ \ref{7.2.3} \] that \( T \) has a fixed point \( x \in Tx \). Hence there exists a solution \( x \) to
\[
\langle Bx + v, u - x \rangle \geq 0 \text{ all } [u, v] \in \mathcal{G}(A)
\]
Next suppose that \( K \) is only closed and convex but \( B \) is coercive and \( 0 \in D(A) \). Then let \( K_n = B(0,n) \cap K \) and let \( A_n \) be the restriction of \( A \) to \( B(0,n) \). It follows that there exists \( x_n \in K_n \) such that for all \( [u,v] \in \mathcal{G}(A_n) \),

\[
\langle Bx_n + v, u - x_n \rangle \geq 0
\]

Then since \( 0 \in D(A) \), one can pick \( v_0 \in A_0 \) and obtain

\[
\langle Bx_n + v_0, u - x_n \rangle \geq 0, \quad \langle v_0, -x_n \rangle \geq \langle Bx_n, x_n \rangle
\]

from which it follows from coercivity of \( B \) that the \( x_n \) are bounded independent of \( n \). Say \( \|x_n\| < C \). Then for all \( n \) large enough \( \|u\| < n \) and so

\[
\langle Bx_n + v, u - x_n \rangle \geq 0
\]

Then letting \( n \to \infty \) and using the continuity of \( B \),

\[
\langle Bx + v, u - x \rangle \geq 0
\]

Since \( [u,v] \) was arbitrary, this proves the lemma.

**Observation 9.6.4** If you have a monotone set valued function, then its graph can always be considered a subset of the graph of a maximal monotone graph. If \( A \) is monotone, then let \( F \) be \( \mathcal{G}(B) \) such that \( \mathcal{G}(B) \supseteq \mathcal{G}(A) \) and \( B \) is monotone. Partially order by set inclusion. Then let \( C \) be a maximal chain. Let \( \mathcal{G}(\hat{A}) = \cup C \).

If \( [x_i, y_i] \in \mathcal{G}(\hat{A}) \), then both are in some \( B \in C \). Hence \( (y_1 - y_2, x_1 - x_2) \geq 0 \) so monotone and must be maximal monotone because if \( (z - v, x - u) \geq 0 \) for all \( [u,v] \in \mathcal{G}(\hat{A}) \) and \( [x, z] \notin \hat{A} \), then you could include this ordered pair and contradict maximality of the chain \( C \).

Next is an interesting theorem which comes from this lemma. It is an infinite dimensional version of the above lemma.

**Theorem 9.6.5** Let \( X \) be a reflexive Banach space and let \( K \) be a closed convex subset of \( X \). Let \( A, B \) be monotone such that

1. \( D(A) \subseteq K, 0 \in D(A) \).
2. \( B \) is single valued, hemicontinuous, bounded and coercive mapping \( X \) to \( X' \).

Then there exists \( x \in K \) such that

\[
\langle Bx + v, u - x \rangle \geq 0 \quad \text{for all} \ [u,v] \in \mathcal{G}(A)
\]

Before giving the proof, here is an easy lemma.
Lemma 9.6.6 Let $E$ be finite dimensional and let $B : E \to E'$ be monotone and hemicontinuous. Then $B$ is continuous.

Proof: The space can be considered a finite dimensional Hilbert space $(\mathbb{R}^n)$ and so weak and strong convergence are exactly the same. First it is desired to show that $B$ is bounded. Suppose it is not. Then there exists $\|x_k\|_E = 1$ but $\|Bx_k\|_{E'} \to \infty$. Since finite dimensional, there is a subsequence still denoted as $x_k$ such that $x_k \to x, \|x\|_E = 1$.

$$\langle Bx_k - Bx, x_k - x \rangle \geq 0$$

Hence

$$\left\langle Bx_k - Bx, \frac{\|Bx_k\|_{E'}}{\|Bx_k\|_{E'}}, x_k - x \right\rangle \geq 0$$

Then taking another subsequence, written with index $k$, it can be assumed that

$$Bx_k/\|Bx_k\| \to y^* \in E', \|y^*\|_{E'} = 1$$

Hence,

$$\langle y^*, x_k - x \rangle \geq 0$$

for all $x \in E$, but this requires that $y^* = 0$, a contradiction. Thus $B$ is monotone, hemicontinuous, and bounded. It follows from Theorem 9.1.4 which says that monotone and hemicontinuous operators are pseudomonotone and Proposition 9.1.6 which says that bounded pseudomonotone operators are demicontinuous that $B$ is demicontinuous, hence continuous because, as just noted above, weak and strong convergence are the same for finite dimensional spaces. In case $B$ is bounded, then this follows from Proposition 9.4.4 above. It is pseudomonotone and bounded hence demicontinuous and weak and strong convergence is the same in finite dimensions.

Proof of Theorem 9.6.5: Let $\{X_n\}$ be an increasing sequence of finite dimensional subspaces. Let $\hat{A}$ be maximal monotone on $\cup_n X_n$ and extending $A$. By this is meant that the graph of $\hat{A}$ contains the graph of $A$ restricted to $\cup_n X_n$, $\hat{A}$ is monotone and there is no other larger graph with these properties. See the above observation. Let $j_n : X_n \to X$ be the inclusion map and $j^*_n : X' \to X'_n$ be the dual map. Then $j^*_n \hat{A}j_n \equiv A_n$ and $j^*_n B j_n \equiv B_n$ have monotone graphs from $X_n$ to $P(X'_n)$ with $B_n$ being continuous and single valued. This follows from the hemicontinuity and the above lemma which states that on finite dimensional spaces, hemicontinuity and monotonicity imply continuity. Then

$$[u, v] \in \mathcal{G}(A_n)$$

means

$$u \in D(A) \cap X_n \text{ and } v \in j_n^* \hat{A} j_n (u) = j_n^* \hat{A} (u) \text{ since } u \in X_n$$

Then from Lemma 9.4.3, there exists $x_n \in X_n$ such that

$$\langle B_n x_n + v_n, u_n - x_n \rangle_{X', X} \geq 0 \text{ all } [u_n, v_n] \in \mathcal{G}(A_n)$$
That is, there exists \( x_n \in K \cap X_n \) such that for all \( u \in D(\hat{A}) \cap X_n, [u,v] \in \mathcal{G}(\hat{A}) \)
\[
\langle Bx_n + v, u - x_n \rangle_{X',X} \geq 0
\]
(9.6.33)

Then
\[
\langle v, u - x_n \rangle \geq \langle Bx_n, x_n - u \rangle
\]
(9.6.34)

From the assumption that \( 0 \in D(\hat{A}) \), one can let \( u = 0 \) and then pick \( v_0 \in \hat{A}0 \).
Then the above reduces to
\[
\langle v_0, -x_n \rangle \geq \langle Bx_n, x_n \rangle
\]

By coercivity of \( B \), these \( x_n \) are all bounded and so by the Eberlien Smulian theorem, there is a subsequence \( \{x_n\} \) which satisfies
\[
x_n \to x \text{ weakly in } X
\]
\[
Bx_n \to y \text{ weakly in } X'
\]

Then from (9.6.34)
\[
\langle v, u - x_n \rangle + \langle Bx_n, u \rangle \geq \langle Bx_n, x_n \rangle
\]

Then it follows that
\[
\langle v, u - x_n \rangle + \langle Bx_n, u \rangle - \langle Bx_n, x \rangle \geq \langle Bx_n, x_n - x \rangle
\]

It follows that
\[
\limsup_{n \to \infty} \langle Bx_n, x_n - x \rangle \leq \langle v, u - x \rangle + \langle y, u \rangle - \langle y, x \rangle
\]
\[
= \langle v + y, u - x \rangle
\]

Claim: \( \limsup_{n \to \infty} \langle Bx_n, x_n - x \rangle \leq 0 \).

Proof of claim: This is so if \( \langle v + y, u - x \rangle \leq 0 \) for some \( [u,v] \in \mathcal{G}(\hat{A}) \). If \( \langle v + y, u - x \rangle \) is greater than 0 for all \( [u,v] \), then since \( \hat{A} \) is maximal, it would follow that \( -y \in Ax \). Now consider (9.6.33)
\[
\langle v, u - x \rangle \geq \limsup_{n \to \infty} \langle Bx_n, x_n \rangle - \langle y, u \rangle
\]

Since \( x \in D(\hat{A}) \), you could put in \( u = x \) in the above and obtain
\[
0 \geq \limsup_{n \to \infty} \langle Bx_n, x_n \rangle - \langle y, x \rangle = \limsup_{n \to \infty} \langle Bx_n, x_n - x \rangle
\]
which shows the claim is true.

Since \( B \) is monotone and hemicontinuous, it satisfies the pseudomonotone condition, Theorem (9.1.4). Hence for any \( z \),
\[
\langle y, x - z \rangle \geq \limsup_{n \to \infty} \langle Bx_n, x_n - x \rangle + \limsup_{n \to \infty} \langle Bx_n, x - z \rangle
\]
\[ \geq \limsup_{n \to \infty} (\langle Bx_n, x_n - x \rangle + \langle Bx_n, x - z \rangle) \]
\[ \geq \liminf_{n \to \infty} (\langle Bx_n, x_n - z \rangle) \geq \langle Bx, x - z \rangle \]

Since \( z \) is arbitrary, this shows that \( y = Bx \). It follows from 9.6.33 that for any \([u, v] \in G(\hat{A})\),
\[ \langle Bx_n + v, u - x_n \rangle = \langle Bx_n + v, u - x \rangle + \langle Bx_n + v, x - x_n \rangle \geq 0 \]
\[ \langle Bx_n + v, u - x \rangle \geq \langle Bx_n, x_n - x \rangle \geq \langle Bx, x_n - x \rangle \]

Now take a limit of both sides and use the fact that \( y = Bx \) to obtain
\[ \langle Bx + v, u - x \rangle \geq 0 \]
for all \([u, v] \in G(\hat{A})\). Here \( \hat{A} \) extends \( A \) on \( \cup_n X_n \). Why does it follow from this that there exists an \( x \) such that the inequality holds for all \([u, v] \in G(A)\)?

Let \( V \) be a finite dimensional subspace.
\[ K_V \equiv \left\{ x \in K : \langle Bx + v, u - x \rangle \geq 0 \text{ for all } [u, v] \in G(A), u \in V \right\} \]

Then from the above argument, \( K_V \neq \emptyset \). You just choose your subspaces \( X_n \) to all include \( V \). Also, from coercivity of \( B \) and the above argument, these \( K_V \) are all bounded and weakly closed. Hence they are weakly compact. Then if you have finitely many of them, you can let your subspaces include each \( V \) and conclude that these \( K_V \) have finite intersection property and so there exists \( x \in \cap V K_V \) which gives the desired \( x \). \( \blacksquare \)

Note that there is only one place where \( 0 \in D(A) \) was used and it was to get the estimate. In the argument,
\[ \langle v, u - x_n \rangle \geq \langle Bx_n, x_n - u \rangle \]
and it was convenient to be able to take \( u = 0 \). However, you could also assume other things on \( B \) such as that it satisfies an estimate of the form
\[ \|Bx\| \leq C\|x\| + C \]
and if you did this, you could also obtain the necessary estimate as follows.
\[ \langle v, u - x_n \rangle \geq \langle Bx_n, x_n - u \rangle \]
\[ \langle v, u - x_n \rangle + \langle Bx_n, u \rangle \geq \langle Bx_n, x_n \rangle \]
\[ \|v\| (\|u\| + \|x_n\|) + (C \|x_n\| + C \|u\|) \geq \langle Bx_n, x_n \rangle \]

and then pick some \([u, v]\). Thus the following corollary comes right away. This would have worked just as well if you had an estimate of the form
\[ \|Bx\| \leq C \|x\|^{p-1} + C, \ p > 1 \]
Corollary 9.6.7 Let $X$ be a reflexive Banach space and let $K$ be a closed convex subset of $X$. Let $A, B$ be monotone such that

1. $D(A) \subseteq K$

2. $B$ is single valued, hemicontinuous, bounded and coercive mapping $X$ to $X'$ which satisfies the estimate

$$\|Bx\| \leq C\|x\| + C$$

or more generally

$$\|Bx\| \leq C\|x\|^{p-1} + C, \quad p > 1$$

Then there exists $x \in K$ such that

$$\langle Bx + v, u - x \rangle_{X',X} \geq 0 \text{ for all } [u, v] \in \mathcal{G}(A)$$

Now here is the equivalence between maximal monotone graph and having $F + A$ be onto. It was already shown that if $\lambda F + A$ is onto, then the graph of $A$ is maximal monotone in the sense that there is no monotone operator whose graph properly contains the graph of $A$. This was Theorem 9.6.2 above which is stated here as a reminder of what it said.

Theorem 9.6.8 Let $X, X'$ be reflexive and have strictly convex norms. Let $A$ be a monotone set valued map as just described. Then if

$$\lambda F + A \text{ onto},$$

for some $\lambda > 0$, then whenever

$$\langle y - z, x - u \rangle \geq 0 \text{ for all } [x, y] \in \mathcal{G}(A)$$

it follows that $z \in Au$ and $u \in D(A)$. That is, the graph is maximal.

Theorem 9.6.9 Let $X$ be a strictly convex reflexive Banach space. Suppose the graph of $A : X \to P(X)$ is maximal monotone in the sense that it is monotone and no monotone graph can properly contain the graph of $A$. Then for all $\lambda > 0, \lambda F + A$ is onto. Conversely, if for some $\lambda > 0, \lambda F + A$ is onto, then the graph of $A$ is maximal with respect to being monotone.

Proof: In Theorem 9.6.8, let $Bx \equiv \lambda F(x) - y_0$. Then from the properties of the duality map, Theorem 9.2.3 above, it follows that $B$ satisfies the necessary conditions to use the result of Corollary 9.6.7 with $K = X$. This $B$ is monotone hemicontinuous, and coercive. Thus there exists $x$ such that for all $[u, v] \in \mathcal{G}(A)$,

$$\langle \lambda F(x) - y_0 + v, u - x \rangle_{X',X} \geq 0$$

$$\langle v - (y_0 - \lambda F(x)), u - x \rangle_{X',X} \geq 0$$

By maximality of the graph, it follows that $x \in D(A)$ and

$$y_0 - \lambda F(x) \in A(x), \quad y_0 = \lambda F(x) + A(x)$$
9.6. MAXIMAL MONOTONE OPERATORS

so $\lambda F + A$ is onto as claimed. The converse was proved in Theorem 9.6.9.

Note that this theorem holds if $F$ is a duality map for $p > 1$. That is, $\langle Fx, x \rangle = \|x\|^p, \|Fx\| = \|x\|^{p-1}$.

Suppose $A : X \rightarrow \mathcal{P}(X)$ is maximal monotone. Then let $z \in X$ and define a new mapping $\hat{A}$ as follows.

$$D(\hat{A}) \equiv \{x : x - x_0 \in D(A)\}, \hat{A}(x) \equiv A(x - x_0)$$

**Proposition 9.6.10** Let $A, \hat{A}$ be as just defined. Then $\hat{A}$ is also maximal monotone.

**Proof:** From Theorem 9.6.9 it suffices to show that graph of $\hat{A}$ is monotone and is maximal. Suppose then that $x_1, x_2 \in A(x_i - x_0)$ and so the above is $\geq 0$. Next suppose for all $[x, x^*] \in G(\hat{A})$,

$$\langle x^* - z^*, x - z \rangle \geq 0$$

Does it follow that $[z, z^*] \in G(\hat{A})$? The above says that

$$\langle x^* - z^*, x - x_0 - (z - x_0) \rangle \geq 0$$

whenever $x - x_0 \in D(A)$ and $x^* \in A(x - x_0)$. Hence, since $A$ is given to be maximal monotone, $z - x_0 \in D(A)$ and $z^* \in A(z - x_0)$ which says that $z^* \in \hat{A}(z)$. Thus $\hat{A}$ is maximal monotone by the Theorem 9.6.9.

9.6.2 Surjectivity Theorems

As an interesting example of this theorem, here is another result in Barbu [1]. It is interesting because it is not assumed $B$ is bounded.

**Theorem 9.6.11** Let $B : X \rightarrow X'$ be monotone hemicontinuous. Then $B$ is maximal monotone. If $B$ is coercive, then $B$ is also onto. Here $X$ is a strictly convex reflexive Banach space.

**Proof:** Suppose $B$ is not maximal monotone. Then there exists $(x_0, x_0^*) \in X \times X'$ such that for all $x$,

$$\langle Bx - x_0^*, x - x_0 \rangle \geq 0$$

and yet $x_0^* \neq Bx_0$. This is going to be a contradiction. Let $u \in X$ and consider $x_t \equiv tx_0 + (1 - t) u, t \in (0, 1)$. Then consider

$$\langle Bx_t - x_0^*, x_t - x_0 \rangle$$
However, \( x_t - x_0 = tx_0 + (1 - t) u - x_0 = (1 - t) (u - x_0) \) and so, for each \( t \in (0, 1) \),
\[
0 \leq \langle Bx_t - x_0^*, x_t - x_0 \rangle = (1 - t) \langle Bx_t - x_0^*, u - x_0 \rangle
\]
Divide by \((1 - t)\) and then let \( t \uparrow 1 \). This yields the following by hemicontinuity.
\[
\langle Bx_0 - x_0^*, u - x_0 \rangle \geq 0
\]
which holds for all \( u \). Hence \( Bx_0 = x_0^* \) after all. Thus \( B \) is indeed maximal monotone.

Next suppose \( B \) is coercive. Let \( F \) be the duality map (or the duality map for arbitrary \( p > 1 \)). Then from Theorem 9.6.9 there exists a solution \( x_\lambda \) to
\[
\lambda Fx_\lambda + Bx_\lambda = x_0^* \in X'
\]
(9.6.35)
Then the \( x_\lambda \) are bounded because, doing both sides to \( x_\lambda \),
\[
\lambda \| x_\lambda \|^2 + \langle Bx_\lambda, x_\lambda \rangle = \langle x_0^*, x_\lambda \rangle
\]
and so
\[
\frac{\langle Bx_\lambda, x_\lambda \rangle}{\| x_\lambda \|} \leq \| x_0^* \|
\]
Thus the coercivity of \( B \) implies that the \( x_\lambda \) are bounded. There exists a subsequence such that
\[
x_\lambda \to x \text{ weakly.}
\]
Then from the equation \( \| \lambda Fx_\lambda \| = \lambda \| x_\lambda \| \) and so,
\[
Bx_\lambda \to x_0^* \text{ strongly.}
\]
Since \( B \) is monotone and hemicontinuous, it satisfies the pseudomonotone condition, Theorem 9.6.4. The above strong convergence implies
\[
\lim_{\lambda \to 0} \langle Bx_\lambda, x_\lambda - x \rangle = 0
\]
Hence for all \( y \),
\[
\lim_{\lambda \to 0} \inf \langle Bx_\lambda, x_\lambda - y \rangle = \lim_{\lambda \to 0} \inf \langle Bx_\lambda, x - y \rangle = \langle x_0^*, x - y \rangle \geq \langle Bx, x - y \rangle
\]
Since \( y \) is arbitrary, this shows that \( x_0^* = Bx \) and so \( B \) is onto as claimed. \( \blacksquare \)

Again, note that it really didn’t matter about the particular duality map used, although the usual one was featured in the argument.

There are some more things which can be said about maximal monotone operators. To include some of these, here is a very interesting lemma found in [I].

**Lemma 9.6.12** Let \( X \) be a Banach space and suppose that
\[
x_n \to 0, \quad \| x_n^* \| \to \infty
\]
Then denoting by \( D_r \), the closed disk centered at \( 0 \) with radius \( r \). It follows that for every \( D_r \), there exists \( y_0 \in D_r \) and a subsequence with index \( n_k \) such that
\[
\langle x_{n_k}^*, x_{n_k} - y_0 \rangle \to -\infty
\]
Proof: Suppose this is not true. Then there exists $D_r$ which has the property that for all $u \in D_r$,
\[
(x^*_n, x_n - u) \geq C_u
\]
for all $n$. Now let
\[
E_k \equiv \{ y \in D_r : (x^*_n, x_n - y) \geq -k \text{ for all } n \}
\]
Then this is a closed set, being the intersection of closed sets. Also, by assumption, the union of these $E_k$ equals $D_r$ which is a complete metric space. Hence one of these $E_k$ must have nonempty interior by the Bair category theorem, say for $k_0$. Say $B(y, \varepsilon) \subseteq D_r$. Then for all $\|u - y\| < \varepsilon$,
\[
(x^*_n, x_n - u) \geq -k_0 \text{ for all } n
\]
Of course $-y \in D_r$ also, and so there is $C$ such that
\[
(x^*_n, x_n + y) \geq C \text{ for all } n
\]
Then
\[
(x^*_n, 2x_n + y - u) \geq C - k_0 \text{ for all } n
\]
whenever $\|y - u\| < \varepsilon$. Now recall that $x_n \to 0$. Consider only $u$ such that $\|y - u\| < \varepsilon/2$. Therefore, for all $n$ large enough, the expression $2x_n + y - u$ for such $u$ contains a small ball centered at the origin, say $D_\delta$. (The set of all $y - u$ for $u$ closer to $y$ than $\varepsilon/2$ is the ball $B(0, \varepsilon/2)$ and then the $2x_n$ does not move it by much provided $n$ is large enough.) Therefore,
\[
(x^*_n, v) \geq C - k_0
\]
for all $\|v\| \leq \delta$. This contradicts the assumption that $\|x^*_n\| \to \infty$. ■

**Corollary 9.6.13** Let $X$ be a Banach space and suppose that
\[
x_n \to x, \quad \|x^*_n\| \to \infty
\]
Then denoting by $D_r$ the closed disk centered at $x$ with radius $r$. It follows that for every $D_r$, there exists $y_0 \in D_r$ and a subsequence with index $n_k$ such that
\[
(x^*_n, x_{n_k} - y_0) \to -\infty
\]

**Proof:** It follows that $x_{n_k} - x \to 0$. Therefore, from Lemma 9.6.12, for every $r > 0$, there exists $y_0 \in B(0, r)$ and a subsequence $x_{n_k}$ such that
\[
(x^*_n, x_{n_k} - y_0) \to -\infty
\]
Thus
\[
(x^*_n, x_{n_k} - (x + y_0)) \to -\infty
\]
Just let $y_0 = x + \hat{y}_0$. Then $y_0 \in D_r$ and satisfies the desired conditions. ■
A set valued mapping $A : D(A) \to \mathcal{P}(X)$ is locally bounded at $x \in \overline{D(A)}$ if whenever $x_n \to x$, $x_n \in D(A)$ it follows that

$$\limsup_{n \to \infty} \{\|x_n^*\| : x_n^* \in Ax_n\} < \infty.$$ 

**Lemma 9.6.15** A set valued operator $A$ is locally bounded at $x \in \overline{D(A)}$ if and only if there exists $r > 0$ such that $A$ is bounded on $B(x, r) \cap D(A)$.

**Proof:** Say the limit condition holds. Then if no such $r$ exists, it follows that $A$ is unbounded on every $B(x, r) \cap D(A)$. Hence, you can let $r_n \to 0$ and pick $x_n \in B(x, r_n) \cap D(A)$ with $x_n^* \in Ax_n$ such that $\|x_n^*\| > n$, violating the limit condition. Hence some $r$ exists such that $A$ is bounded on $B(x, r) \cap D(A)$. Conversely, suppose $A$ is bounded on $B(x, r) \cap D(A)$ by $M$. Then if $x_n \to x$, it follows that for all $n$ large enough, $x_n \in B(x, r)$ and so if $x_n^* \in Ax_n$, $\|x_n^*\| \leq M$. Hence $\limsup_{n \to \infty} \{\|x_n^*\| : x_n^* \in Ax_n\} \leq M < \infty$ which verifies the limit condition.

With this definition, here is a very interesting result.

**Theorem 9.6.16** Let $A : D(A) \to X'$ be maximal monotone. Then if $x$ is an interior point of $D(A)$, it follows that $A$ is locally bounded at $x$.

**Proof:** You could use Corollary 9.6.13. If $x$ is an interior point of $D(A)$, and $A$ is not locally bounded, then there exists $x_n \to x$ and $x_n^* \in Ax_n$ such that $\|x_n^*\| \to \infty$. Then by Corollary 9.6.13, there exists $y_0$ close to $x$, in $D(A)$ and a subsequence $x_{n_k}$ such that

$$\langle x_{n_k}^*, x_{n_k} - y_0 \rangle \to -\infty$$

Letting $y_0^* \in Ay_0$,

$$\langle x_{n_k}^* - y_0^*, x_{n_k} - y_0 \rangle \geq 0$$

and so

$$\langle x_{n_k}^*, x_{n_k} - y_0 \rangle \geq \langle y_0^*, x_{n_k} - y_0 \rangle$$

and the right side is bounded below because it converges to $\langle y_0^*, x - y_0 \rangle$ and this is a contradiction.

Does the same proof work if $x$ is a limit point of $D(A)$? No. Suppose $x$ is a limit point of $D(A)$. If $A$ is not locally bounded, then there exists $x_n \to x, x_n \in D(A)$ and $x_n^* \in Ax_n$ and $\|x_n^*\| \to \infty$. Then there is $y_0$ close to $x$ such that $\langle x_{n_k}^*, x_{n_k} - y_0 \rangle \to -\infty$ but now everything crashes in flames because it is not known that $y_0 \in D(A)$.

It follows from the above theorem that if $A$ is defined on all of $X$ and is maximal monotone, then it is locally bounded everywhere. Recall the definition of a pseudomonotone operator.

**Definition 9.6.17** A set valued operator $B$ is quasi-bounded if whenever $x \in D(B)$ and $x^* \in Bx$ are such that

$$\|\langle x^*, x \rangle\|, \|x\| \leq M,$$

it follows that $\|x^*\| \leq K_M$. Bounded would mean that if $\|x\| \leq M$, then $\|x^*\| \leq K_M$. Here you only know this if there is another condition.
By Proposition 9.6.18 an example of a quasi-bounded operator is a maximal monotone operator $G$ for which $0 \in \text{int} \,(D\,(G))$. Then there is a useful result which gives examples of quasi-bounded operators.

**Proposition 9.6.18** Let $A : D\,(A) \subseteq X \to \mathcal{P}(X')$ be maximal monotone and suppose $0 \in \text{int} \,(D\,(A))$. Then $A$ is quasi-bounded.

**Proof:** From local boundedness, Theorem 9.6.16, there exists $\delta, C > 0$ such that
\[
\sup \{ \|x^*\| : x^* \in A(x) \text{ for } \|x\| \leq \delta \} < C
\]
Now suppose that $\|x\|, |\langle x^*, x \rangle| \leq M$. Then letting $\|y\| \leq \delta, y^* \in Ay,$
\[
0 \leq \langle x^* - y^*, x - y \rangle = \langle x^*, x \rangle - \langle x^*, y \rangle - \langle y^*, x \rangle + \langle y^*, y \rangle
\]
and so for $\|y\| \leq \delta$,
\[
\langle x^*, y \rangle \leq \langle x^*, x \rangle - \langle y^*, x \rangle + \langle y^*, y \rangle \leq M + MC + C\delta
\]
Hence, $\|x^*\| \leq M + MC + C\delta \equiv K_M$. ■

This is actually quite a restrictive requirement and leaves out a lot which would be interesting.

**Definition 9.6.19** Let $V$ be a Reflexive Banach space. We say $T : V \to \mathcal{P}(V')$ is pseudomonotone if the following conditions hold.

\[
Tu \text{ is closed, nonempty, convex.}
\]

(9.6.36)

If $F$ is a finite dimensional subspace of $V$, then if $u \in F$ and $W \supseteq Tu$ for $W$ a weakly open set in $V'$, then there exists $\delta > 0$ such that
\[
v \in B(u, \delta) \cap F \text{ implies } Tv \subseteq W.
\]

(9.6.37)

If $u_k \rightharpoonup u$ and if $u_k^* \in Tu_k$ is such that
\[
\limsup_{k \to \infty} u_k^*(u_k - u) \leq 0,
\]
then for all $v \in V$, there exists $u^*(v) \in Tu$ such that
\[
\liminf_{k \to \infty} u_k^*(u_k - v) \geq u^*(v) (u - v).
\]

(9.6.38)

Then here is an interesting result.

**Theorem 9.6.20** Suppose $A : X \to \mathcal{P}(X')$ is maximal monotone. That is, $D\,(A) = X$. Then $A$ is pseudomonotone.
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

**Proof:** Consider the first condition. Say $x_1^* \in Ax$. Let $u^* \in Au$. For $\lambda \in [0, 1]$, 

$$\langle \lambda x_1^* + (1 - \lambda) x_2^* - u^*, x - u \rangle = \lambda \langle x_1^* - u^*, x - u \rangle + (1 - \lambda) \langle x_2^* - u^*, x - u \rangle \geq 0$$

and so, since $[u, u^*]$ is arbitrary, it follows that $\lambda x_1^* + (1 - \lambda) x_2^* \in Ax$. Thus $Ax$ is convex. Is it closed? Say $x_n^* \in Ax$ and $x_n^* \to x^*$. Is it the case that $x^* \in D(A)$? Let $[u, u^*] \in G(A)$ be arbitrary. Then 

$$\langle x^* - u^*, x - u \rangle = \lim_{n \to \infty} \langle x_n^* - u^*, x_n - u \rangle \geq 0$$

and so $Ax$ is also closed.

Consider the second condition. It is to show that if $x_n \to x$ in $V$ a finite dimensional subspace and if $U$ is a weakly open set containing 0, then eventually $Ax_n \subseteq Ax + U$. Suppose then that this is not the case. Then there exists $x_n^*$ outside of $Ax + U$ but in $Ax_n$. Since $A$ is locally bounded at $x$, it follows that the $\|x_n^*\|$ are bounded. Thus there is a subsequence, still denoted as $x_n$ and $x_n^*$ such that $x_n^* \to x^*$ weakly and $x^* /\in Ax + U$. Now let $[u, u^*] \in G(A)$.

$$\langle x^* - u^*, x - u \rangle = \lim_{n \to \infty} \langle x_n^* - u^*, x_n - u \rangle \geq 0$$

and since $[u, u^*]$ is arbitrary, it follows that $x^* \in Ax$ and so is inside $Ax + U$. Thus the second condition holds also.

Consider the third. Say $x_k \to x$ weakly and letting $x_k^* \in Ax_k$, suppose

$$\lim \sup_{k \to \infty} \langle x_k^*, x_k - x \rangle \leq 0,$$

Is it the case that there exists $x^* (y) \in Ax$ such that

$$\lim \inf_{k \to \infty} \langle x_k^*, x_k - y \rangle \geq \langle x^* (y), x - y \rangle?$$

The proof goes just like it did earlier in the case of single valued pseudomonotone operators. It is just a little more complicated. First, let $x^* \in Ax$.

$$\langle x_k^* - x^*, x_k - x \rangle \geq 0$$

and so

$$\lim \inf_{k \to \infty} \langle x_k^*, x_k - x \rangle \geq \lim \inf_{k \to \infty} \langle x^*, x_k - x \rangle = 0 \geq \lim \sup_{k \to \infty} \langle x_k^*, x_k - x \rangle$$

Thus

$$\lim_{k \to \infty} \langle x_k^*, x_k - x \rangle = 0.$$

Now let $x_t^* \in A(x + t (y - x))$, $t \in (0, 1)$, where here $y$ is arbitrary. Then

$$\langle x_n^* - x_t^*, x_n - x + t (x - y) \rangle \geq 0$$
Hence
\[
\liminf_{n \to \infty} \langle x^*_n, x_n - x + t(x - y) \rangle \geq \liminf_{n \to \infty} \langle x^*_t, x_n - x + t(x - y) \rangle
\]
and so from the above limit,
\[
t \liminf_{n \to \infty} \langle x^*_n, x - y \rangle \geq \liminf_{n \to \infty} \langle x^*_t, x_n - y \rangle \geq \langle x^*_t, x - y \rangle
\]
Cancel the t.
\[
\liminf_{n \to \infty} \langle x^*_n, x - y \rangle = \liminf_{n \to \infty} \langle x^*_n, x_n - y \rangle \geq \langle x^*_t, x - y \rangle
\]
Now you have a fixed y and \( x^*_t \in A(x + t(y - x)) \). The subspace determined by \( x, y \) is finite dimensional. Also it was shown above that A is locally bounded at \( x \) and so there is a subsequence, still denoted as \( x^*_t \) such that \( x^*_t \to x^*(y) \) weakly. Now from the upper semicontinuity on finite dimensional spaces shown above, for every \( S \) a finite subset of \( X \) and \( \varepsilon > 0 \), it follows that for all \( t \) small enough,
\[
x^*_t \in Ax + B_S(0, \varepsilon)
\]
Thus \( x^*(y) \in Ax \). Hence, there exists \( x^*(y) \in Ax \) such that
\[
\liminf_{n \to \infty} \langle x^*_n, x_n - y \rangle \geq \langle x^*(y), x - y \rangle \]
Suppose \( T \) is a bounded pseudomonotone operator and \( S \) is a maximal monotone operator, both defined on a strictly convex reflexive Banach space. What of their sum? Is \( (T + S)(x) \) convex and closed? Say \( t_n \in Tx \) and \( s_n \in Sx \) is it the case that \( t_1 + (1 - \theta) (s_2 + t_2) \in (T + S)(x) \) whenever \( \theta \in [0, 1] \)? Of course this is so. Thus \( T + S \) has convex values. Does it have closed values? Suppose \( \{s_n + t_n\} \) converges to \( z \in X', s_n \in Sx, t_n \in Tx \). Is \( z \in (T + S)(x) \)? Taking a subsequence, and using the assumption that \( T \) is bounded, it can be assumed that \( t_n \to t \in Tx \) weakly. Therefore, \( s_n \) must also converge weakly and so it converges to some \( s = z - t \in Sx \). Convex and closed implies weakly closed. Thus \( T + S \) has closed convex values. Is it upper semicontinuous on finite dimensional subspaces? Suppose \( x_n \to x \) in a finite dimensional subspace \( F \). Does it follow that
\[
(S + T)x_n \subseteq (S + T)x + B(0, r)
\]
for all \( n \) sufficiently large? It is known that \( Sx_n \subseteq Sx + B(0, r/2) \) and \( Tx_n \subseteq Tx + B(0, r/2) \) whenever \( n \) is sufficiently large and so it follows that
\[
(S + T)x_n \subseteq (S + T)x + B(0, r/2) + B(0, r/2) \subseteq (S + T)x + B(0, r)
\]
whenever \( n \) is large enough.

What of the pseudomonotone condition? Suppose
\[
\limsup_{n \to \infty} \langle u^*_n + v^*_n, x_n - x \rangle \leq 0
\]
where \( u^*_n \in S x_n \) and \( v^*_n \in T x_n \) where \( x_n \to x \) weakly. Is it the case that for every \( y \), there exists \( u^* \in S x \) and \( v^* \in T x \) such that
\[
\lim \inf_{n \to \infty} \langle u^*_n + v^*_n, x_n - y \rangle \geq \langle u^* + v^*, x - y \rangle?
\]
By monotonicity,
\[
0 \geq \lim \sup_{n \to \infty} \langle u^*_n + v^*_n, x_n - x \rangle \geq \lim \sup_{n \to \infty} \langle u^*_n, x_n - x \rangle
\]
Hence
\[
\lim \sup_{n \to \infty} \langle v^*_n, x_n - x \rangle \leq 0
\]
which implies
\[
\lim_{n \to \infty} \langle v^*_n, x_n - x \rangle = 0
\]
showing that
\[
\lim_{n \to \infty} \langle v^*_n, x_n - x \rangle \geq 0 \geq \lim_{n \to \infty} \langle v^*_n, x_n - x \rangle
\]
It follows that if \( y \) is given, there exists \( v^* \in T (x) \) such that
\[
\lim_{n \to \infty} \langle v^*_n, x_n - y \rangle \geq \langle v^*, x - y \rangle
\]
Now let \( u^*_t \in S (x + t (y - x)) \) for \( t > 0 \). Thus
\[
\langle u^*_n - u^*_t, x_n - x + t (x - y) \rangle \geq 0
\]
\[
\langle u^*_n, x_n - x + t (x - y) \rangle \geq \langle u^*_t, x_n - x + t (x - y) \rangle
\]
Then using the above and the convergence in 1.1.33,
\[
\lim_{n \to \infty} \langle u^*_n + v^*_n, x_n - y \rangle \geq \lim_{n \to \infty} \langle u^*_n, x_n - y \rangle
\]
\[
= \langle u^*_t, x - y \rangle + \langle v^*, x - y \rangle
\]
Now as before where it was shown that maximal monotone and defined on \( X \) implied pseudomonotone, and the theorem which says that maximal monotone operators are locally bounded on the interior of their domains, it follows that there exists a sequence, still denoted as \( u^*_t \) which converges to something called \( u^* \). Then as before, the subspace spanned by \( x, y \) is finite dimensional and so from upper semicontinuity, for all \( t \) small enough,
\[
u^*_t \in S (x) + B (0, r)
\]
Note that weak convergence is the same as strong on finite dimensional spaces. Since this is true for all \( r \) and \( S (x) \) is closed, it follows that \( u^* \in S (x) \). Thus, passing to a limit as \( t \to 0 \) one gets \( u^* \in S (x) \), \( v^* \in T (x) \), and
\[
\lim_{n \to \infty} \langle u^*_n + v^*_n, x_n - y \rangle \geq \langle u^* + v^*, x - y \rangle
\]
This proves the following generalization of Theorem 9.6.20.
Theorem 9.6.21 Let $T, S : X \to \mathcal{P}(X')$ where $X$ is a strictly convex reflexive Banach space and suppose $T$ is bounded and pseudomonotone while $S$ is maximal monotone. Then $T + S$ is pseudomonotone.

Also, there is an interesting result which is based on the obvious observation that if $A$ is maximal monotone, then so is $\hat{A}(x) \equiv A(x_0 + x)$.

Lemma 9.6.22 Let $A$ be maximal monotone. Then for each $\lambda > 0$,
\[
x \to \lambda F(x - x_0) + Ax
\]
is onto.

Proof: Let $\hat{A}(x) \equiv A(x_0 + x)$ so as earlier, $\hat{A}$ is maximal monotone. Then let $y^* \in X'$. Then there exists $y$ such that $\hat{A}(y) + \lambda F(y) \ni y^*$. Now define $x \equiv y + x_0$. Then
\[
\hat{A}(y) + \lambda F(y) \ni y^*, \quad \hat{A}(x - x_0) + \lambda F(x - x_0) \ni y^*, \quad A(x) + \lambda F(x - x_0) \ni y^* \quad \Box
\]

Definition 9.6.23 Let $A : D(A) \to \mathcal{P}(X')$ be maximal monotone. Let $A^{-1} : A(D(A)) \to \mathcal{P}(X')$ be defined as follows.
\[
x \in A^{-1}x^* \text{ if and only if } x^* \in Ax
\]

Observation 9.6.24 $A^{-1}$ is also maximal monotone. This is easily seen as follows. $[x, y] \in \mathcal{G}(A)$ if and only if $[y, x] \in \mathcal{G}(A^{-1})$.

Earlier, it was shown that if $B$ is monotone and hemicontinuous and coercive, then it was onto. It was not necessary to assume that $B$ is bounded. The same thing holds for $A$ maximal monotone. This will follow from the next result. Recall that a maximal monotone operator is locally bounded at every interior point of its domain which was shown above. Also it appears to not be possible to show that a maximal monotone operator is locally bounded at a limit point of $D(A)$. The following result is in [3] although he claims a better result than what I am proving here in which it is only necessary to verify $A^{-1}$ is locally bounded at every point of $A(D(A))$. However, I was unable to follow the argument and so I am proving another theorem with the same argument he uses. It looks like a typo to me but I often have trouble following hard theorems so I am not sure. Anyway, the following is the best I can do. I think it is still a very interesting result.

Theorem 9.6.25 Suppose $A^{-1}$ is locally bounded at every point of $\overline{A(D(A))}$. Then in fact $A(D(A)) = X'$ and in fact $\overline{A(D(A))} = A(D(A))$.

Proof: This is done by showing that $A(D(A))$ is both open and closed. Since it is nonempty, it must be all of $X'$ because $X'$ is connected. First it is shown that $A(D(A))$ is closed. Suppose $y_n \in Ax_n$ and $y_n \to y$. Does it follow that $y \in A(D(A))$? Since $y$ is a limit point of $A(D(A))$, it follows that $A^{-1}$ is locally bounded at $y$. Thus there is a subsequence still denoted by $y_n$ such that $y_n \to y$ and
for \( x_n \in A^{-1}y_n \) or in other words, \( y_n \in Ax_n \), it follows that \( x_n \) is bounded. Hence there exists a subsequence, still denoted with the subscript \( n \) such that \( x_n \to x \) weakly and \( y_n \to y \) strongly. Hence if \([u,v] \in \mathcal{G}(A)\),

\[
\langle y - v, x - u \rangle = \lim_{n \to \infty} \langle y_n - v, x_n - x \rangle \geq 0
\]

Since \([u,v] \) is arbitrary and \( A \) is maximal monotone, it follows that \( y \in Ax \) or in other words, \( x \in A^{-1}y \) and \( y \in A(D(A)) \). Thus \( A(D(A)) \) is closed.

Next consider why \( A(D(A)) \) is open. Let \( y_0 \in A(D(A)) \). Then there exists \( D_r \equiv \overline{B}(y_0, r) \) centered at \( y_0 \) such that \( A^{-1} \) is bounded on \( D_r \). Since \( A \) is maximal monotone, for each \( y \in X' \) there is a solution \( x_\varepsilon \) to the inclusion

\[
y \in \varepsilon F(x_\varepsilon - x_0) + Ax_\varepsilon, \; y_\varepsilon \equiv y - \varepsilon F(x_\varepsilon - x_0) \in Ax_\varepsilon
\]

Consider only \( y \in B(y_0, \frac{r}{2}) \).

\[
\langle (y - \varepsilon F(x_\varepsilon - x_0)) - y_0, x_\varepsilon - x_0 \rangle \geq 0
\]

Then using \( \langle Fz, z \rangle = \|z\|^2 \),

\[
\|y - y_0\| \|x_\varepsilon - x_0\| \geq \langle y - y_0, x_\varepsilon - x_0 \rangle \geq \varepsilon \|x_\varepsilon - x_0\|^2
\]

and so \( \varepsilon \|x_\varepsilon - x_0\| = \varepsilon \|F(x_\varepsilon - x_0)\| \leq \|y - y_0\| < r/2 \). Thus \( y_\varepsilon \) stays in \( B(y_0, r) \).

This is because \( y \) is closer to \( y_0 \) than \( r/2 \) while \( y_\varepsilon \) is within \( r/2 \) of \( y \). It follows that the \( x_\varepsilon \) are bounded and so \( x_\varepsilon - x_0 \) is bounded and so \( \varepsilon F(x_\varepsilon - x_0) \to 0 \). Thus \( y_\varepsilon \to y \) strongly. Since the \( x_\varepsilon \) are bounded, there exists a further subsequence, still denoted as \( x_\varepsilon \) such that \( x_\varepsilon \to x \), some point of \( X \). Then if \([u,v] \in \mathcal{G}(A)\),

\[
\langle y_\varepsilon - v, x_\varepsilon - u \rangle \geq 0
\]

and letting \( \varepsilon \to 0 \) using the strong convergence of \( y_\varepsilon \) one obtains

\[
\langle y - v, x - u \rangle \geq 0
\]

which shows that \( y \in Ax \). Thus \( B(y_0, \frac{r}{2}) \subseteq A(D(A)) \equiv D(A^{-1}) \) and so \( A(D(A)) \) is open.

The proof featured the usual duality map.

Note that as part of the proof \( A(D(A)) \) was shown to be closed so although it was assumed at the outset that \( A^{-1} \) was locally bounded on \( A(D(A)) \), this is the same as saying that \( A^{-1} \) is locally bounded on \( A(D(A)) \).

**Corollary 9.6.26** Suppose \( A : D(A) \to \mathcal{P}(X') \) is maximal monotone and coercive. Then \( A \) is onto.

**Proof:** From Theorem 9.6.25 it suffices to show that \( A^{-1} \) is locally bounded at \( y^* \in A(D(A)) \). The case of an interior point follows from Theorem 9.6.10. Assume
then that \( y^* \) is a limit point of \( A(D(A)) \). Of course this includes the case of interior points. Then there exists \( y^*_n \rightarrow y^* \) where \( y^*_n \in Ax_n \). Then

\[
\frac{\langle y^*_n, x_n \rangle}{\|x_n\|} \leq \|y^*_n\|
\]

and the right side is bounded. Hence by coercivity, so is \( \|x_n\| \). Therefore, there is a further subsequence, still denoted as \( x_n \) such that \( x_n \rightarrow x \) weakly while \( y^*_n \rightarrow y^* \) strongly. Then letting \( [u,v^*] \in G(A) \),

\[
\langle y^* - v^*, x - u \rangle = \lim_{n \rightarrow \infty} \langle y^*_n - v^*, x_n - u \rangle \geq 0
\]

Hence \( y^* \in Ax \) and \( y^* \in A(D(A)) \). Thus \( A^{-1} \) is locally bounded on \( A(D(A)) \) and so \( A \) is onto from the above theorem.

### 9.6.3 Approximation Theorems

This section continues following Barbu [9]. Always it is assumed that the situation is of a real reflexive Banach space \( X \) having strictly convex norm and its dual \( X' \). As observed earlier, there exists a solution \( x_{\lambda} \) to the inclusion

\[
0 \in F(x_{\lambda} - x) + \lambda A(x_{\lambda})
\]

To see this, you consider \( \hat{A}(y) \equiv A(x + y) \). Then \( \hat{A} \) is also maximal monotone and so there exists a solution to

\[
0 \in F(\hat{x}) + \lambda \hat{A}(\hat{x}) = F(\hat{x}) + \lambda A(x + \hat{x})
\]

Now let \( x_{\lambda} = x + \hat{x} \) so \( \hat{x} = x_{\lambda} - x \). Hence

\[
0 \in F(x_{\lambda} - x) + \lambda A x_{\lambda}
\]

Here you could have \( F \) the duality map for any given \( p > 1 \).

The symbol \( \limsup_{n,n \rightarrow \infty} a_{mn} \) means \( \limsup_{N \rightarrow \infty} (\sup_{m \geq N, n \geq N} a_{mn}) \). Then here is a simple observation.

**Lemma 9.6.27** Suppose \( \limsup_{n,n \rightarrow \infty} a_{mn} \leq 0 \). Then \( \limsup_{m \rightarrow \infty} (\limsup_{n \rightarrow \infty} a_{mn}) \leq 0 \).

**Proof:** There exists \( N \) such that if both \( m, n \geq N, a_{mn} \leq \varepsilon \). Then

\[
\limsup_{n \rightarrow \infty} a_{mn} = \lim_{n \rightarrow \infty} \sup_{n > N} a_{mn} \leq \varepsilon
\]

Thus also

\[
\limsup_{m \rightarrow \infty} \left( \limsup_{n \rightarrow \infty} a_{mn} \right) = \lim_{m \rightarrow \infty, m \geq N} \left( \limsup_{n \rightarrow \infty} a_{mn} \right) \leq \varepsilon.
\]

The argument will be based on the following lemma.
Lemma 9.6.28  Let $A : D (A) \to \mathcal{P} (X')$ be maximal monotone and let $v_n \in Au_n$ and $u_n \to u$, $v_n \to v$ weakly.

Also suppose that

$$\lim_{m,n \to \infty} \langle v_n - v_m,u_n - u_m \rangle \leq 0$$

or

$$\lim_{n \to \infty} \langle v_n - v,u_n - u \rangle \leq 0$$

Then $[u,v] \in G (A)$ and $\langle v_n,u_n \rangle \to \langle v,u \rangle$.

**Proof:** By monotonicity,

$$\lim_{m,n \to \infty} \langle v_n - v_m,u_n - u_m \rangle = 0$$

Suppose then that $\langle v_n,u_n \rangle$ fails to converge to $\langle v,u \rangle$. Then there is a subsequence, still denoted with subscript $n$ such that $\langle v_n,u_n \rangle \to \mu \neq \langle v,u \rangle$. Let $\varepsilon > 0$. Then there exists $M$ such that if $n,m > M$, then

$$|\langle v_n,u_n \rangle - \mu| < \varepsilon, |\langle v_n - v_m,u_n - u_m \rangle| < \varepsilon$$

Then if $m,n > M$,

$$|\langle v_n - v_m,u_n - u_m \rangle| = |\langle v_n,u_n \rangle + \langle v_m,u_m \rangle - \langle v_n,u_m \rangle - \langle v_m,u_n \rangle| < \varepsilon$$

Hence it is also true that

$$|\langle v_n,u_n \rangle + \langle v_m,u_m \rangle - \langle v_n,u_m \rangle - \langle v_m,u_n \rangle| \leq |2\mu - (\langle v_n,u_m \rangle + \langle v_m,u_n \rangle)| < 3\varepsilon$$

Now take a limit first with respect to $n$ and then with respect to $m$ to obtain

$$|2\mu - (\langle v,u \rangle + \langle v,u \rangle)| < 3\varepsilon$$

Since $\varepsilon$ is arbitrary, $\mu = \langle v,u \rangle$ after all. Hence the claim that $\langle v_n,u_m \rangle \to \langle v,u \rangle$ is verified. Next suppose $[x,y] \in \mathcal{G} (A)$ and consider

$$\langle v - y,u - x \rangle = \langle v,u \rangle - \langle v,x \rangle - \langle y,u \rangle + \langle y,x \rangle$$

$$= \lim_{n \to \infty} (\langle v_n,u_n \rangle - \langle v_n,x \rangle - \langle y,u_n \rangle + \langle y,x \rangle)$$

$$= \lim_{n \to \infty} \langle v_n - y,u_n - x \rangle \geq 0$$

and since $[x,y]$ is arbitrary, it follows that $v \in Au$.

Next suppose $\limsup_{n \to \infty} \langle v_n - v,u_n - u \rangle \leq 0$. It is not known that $[u,v] \in \mathcal{G} (A)$.

$$\limsup_{n \to \infty} [\langle v_n,u_n \rangle - \langle v,u \rangle] \leq 0$$

$$\limsup_{n \to \infty} \langle v_n,u_n \rangle - \langle v,u \rangle \leq 0$$
Thus $\limsup_{n \to \infty} \langle v_n, u_n \rangle \leq \langle v, u \rangle$. Now let $[x, y] \in \mathcal{G}(A)$

$$
\langle v - y, u - x \rangle = \langle v, u \rangle - \langle v, x \rangle - \langle y, u \rangle + \langle y, x \rangle
$$

$$
\geq \limsup_{n \to \infty} [\langle v_n, u_n \rangle - \langle v_n, x \rangle - \langle y, u_n \rangle + \langle y, x \rangle]
$$

$$
\geq \liminf_{n \to \infty} [\langle v_n - y, u_n - x \rangle] \geq 0
$$

Hence $[u, v] \in \mathcal{G}(A)$. Now

$$
\limsup_{n \to \infty} \langle v_n - v, u_n - u \rangle \leq 0 \leq \liminf_{n \to \infty} \langle v_n - v, u_n - u \rangle
$$

the second coming from monotonicity and the fact that $v \in Au$. Therefore,

$$
\lim_{n \to \infty} \langle v_n - v, u_n - u \rangle = 0
$$

which shows that $\lim_{n \to \infty} \langle v_n, u_n \rangle = \langle v, u \rangle$.

**Definition 9.6.29** Let $x_\lambda$ just defined

$$
0 \in F(x_\lambda - x) + \lambda A x_\lambda
$$

be denoted by $J_\lambda x$ and define also

$$
A_\lambda(x) = -\lambda^{-(p-1)} F(x_\lambda - x) = -\lambda^{-(p-1)} F(J_\lambda x - x).
$$

This is for $F$ a duality map with $p > 1$. Thus for the usual duality map, you would have

$$
A_\lambda(x) = -\lambda^{-1} F(J_\lambda x - x)
$$

Recall how this $x_\lambda$ is defined. In general,

$$
0 \in F(J_\lambda x - x) + \lambda^{p-1} A x_\lambda
$$

Thus, from the definition,

$$
A_\lambda(x) \in A(J_\lambda x)
$$

Formally, and to help remember what is going on, you are looking at a generalization of

$$
A_\lambda x = \frac{A}{1 + \lambda A} x = \frac{1}{\lambda} \left( x - (I + \lambda A)^{-1} x \right)
$$

This is in the case where $F = I$ to keep things simpler. You have $0 = x_\lambda - x + \lambda A x_\lambda$ and so formally $x_\lambda = (I + \lambda A)^{-1} x$. Thus you are looking at $\frac{1}{\lambda} \left( x - x_\lambda \right) = \frac{1}{\lambda} \left( x - (I + \lambda A)^{-1} x \right) = A_\lambda x$. In fact, this is exactly what you do when you are in a single Hilbert space. This is just a generalization to mappings between Banach spaces and their duals.

Then there are some things which can be said about these operators. It is presented for the general duality map for $p > 1$. 

---

9.6. MAXIMAL MONOTONE OPERATORS
Theorem 9.6.30  The following hold. Here $X$ is a reflexive Banach space with strictly convex norm. $A : D(A) \to P(X')$ is maximal monotone. Then

1. $J_{\lambda}$ and $A_{\lambda}$ are bounded single valued operators defined on $X$. Bounded means they take bounded sets to bounded sets. Also $A_{\lambda}$ is a monotone operator.

2. $A_{\lambda}, J_{\lambda}$ are demicontinuous. That is, strongly convergent sequences are mapped to weakly convergent sequences.

3. For every $x \in D(A)$, $\|A_{\lambda}(x)\| \leq \|Ax\| \equiv \inf \{\|y^*\| : y^* \in Ax\}$. For every $x \in \text{conv}(D(A))$, it follows that $\lim_{\lambda \to 0} J_{\lambda}(x) = x$. The new symbol means the closure of the convex hull. It is the closure of the set of all convex combinations of points of $D(A)$.

Proof:  1.) It is clear that these are single valued operators. What about the assertion that they are bounded? Let $y^* \in Ax_{\lambda}$ such that the inclusion defining $x_{\lambda}$ becomes an equality. Thus

\[ F(x_{\lambda} - x) + \lambda^{p-1} y^* = 0 \]

Then let $x_0 \in D(A)$ be given.

\[ \langle F(x_{\lambda} - x), x_{\lambda} - x \rangle + \lambda^{p-1} \langle y^*, x_{\lambda} - x_0 \rangle + \lambda^{p-1} \langle y^*, x_0 - x \rangle = 0 \]

Then by monotoncity of $A$,

\[ \|x_{\lambda} - x\|^p + \lambda^{p-1} \langle y^*, x_{\lambda} - x_0 \rangle + \lambda^{p-1} \langle y^*, x_0 - x \rangle \leq 0 \]

It follows that

\[ \|x_{\lambda} - x\|^p \leq \lambda^{p-1} \|y^*\| \|x_{\lambda} - x_0\| + \lambda^{p-1} \|y^*\| \|x_0 - x\| \]

Hence if $x$ is in a bounded set, it follows the resulting $x_{\lambda} = J_{\lambda}x$ remain in a bounded set. Now from the definition of $A_{\lambda}$, it follows that this is also a bounded operator. Why is $A_{\lambda}$ monotone?

\[
0 \leq \langle A_{\lambda}x - A_{\lambda}y, x - y \rangle = \langle A_{\lambda}x - A_{\lambda}y, x - J_{\lambda}x - (y - J_{\lambda}y) \rangle + \langle A_{\lambda}x - A_{\lambda}y, J_{\lambda}x - J_{\lambda}y \rangle \\
= \left\langle \lambda^{-(p-1)} F(J_{\lambda}x - x) - \lambda^{-(p-1)} F(J_{\lambda}y - y), J_{\lambda}x - x - (J_{\lambda}y - y) \right\rangle + \langle A_{\lambda}x - A_{\lambda}y, J_{\lambda}x - J_{\lambda}y \rangle
\]

and both terms are nonnegative, the first because $F$ is monotone so indeed $A_{\lambda}$ is monotone.

2.) What of the demicontinuity of $A_{\lambda}$? This one is really tricky. Suppose $x_n \to x$. Does it follow that $A_{\lambda}x_n \to A_{\lambda}x$ weakly? The proof will be based on a pair of equations. These are

\[
\lim_{m,n \to \infty} \langle F(J_{\lambda}x_n - x_n) - F(J_{\lambda}x_m - x_m), J_{\lambda}x_n - x_n - (J_{\lambda}x_m - x_m) \rangle = 0
\]
and
\[ \lim_{m,n \to \infty} \langle A_\lambda(x_n) - A_\lambda(x_m), J_\lambda x_n - J_\lambda x_m \rangle = 0 \]

When these have been established, Lemma 9.6.28 is used to get the desired result for a subsequence. It will be shown that every sequence has a subsequence which gives the right sort of weak convergence and from this the desired weak convergence of \( A_\lambda x_n \) to \( A_\lambda x \) follows.

\[ 0 \in F(J_\lambda x_n - x_n) + \lambda^{p-1}A(J_\lambda x_n) \]
\[ 0 \in F(J_\lambda x - x) + \lambda^{p-1}A(J_\lambda x) \]

\[-\lambda^{-(p-1)}F(J_\lambda x - x) \equiv A_\lambda(x) \in A(J_\lambda x) \]
\[-\lambda^{-(p-1)}F(J_\lambda x_n - x_n) \equiv A_\lambda(x_n) \in A(J_\lambda x_n) \]

Note also that for a given \( x \) there is only one solution \( J_\lambda x \) to \( 0 \in F(J_\lambda x - x) + \lambda^{p-1}A(J_\lambda x) \). By monotonicity of \( F \),

\[ 0 \leq \langle F(J_\lambda x_n - x_n) - F(J_\lambda x_m - x_m), x_m - x_n + J_\lambda x_n - J_\lambda x_m \rangle \]

Then from the above,

\[ \langle F(J_\lambda x_n - x_n) - F(J_\lambda x_m - x_m), x_n - x_m \rangle \]
\[ \leq \langle F(J_\lambda x_n - x_n) - F(J_\lambda x_m - x_m), J_\lambda x_n - J_\lambda x_m \rangle \]

Now from the boundedness of these operators, the left side of the above inequality converges to 0 as \( n, m \to \infty \). Thus

\[ \lim_{m,n \to \infty} \inf \langle F(J_\lambda x_n - x_n) - F(J_\lambda x_m - x_m), J_\lambda x_n - J_\lambda x_m \rangle \geq 0 \] (9.6.40)

\[ \lim_{m,n \to \infty} \inf \langle -\lambda^{p-1}A_\lambda(x_n) - (-\lambda^{p-1}A_\lambda(x_m)), J_\lambda x_n - J_\lambda x_m \rangle \geq 0 \]

\[ \lim_{m,n \to \infty} \inf \left( \lambda^{p-1}A_\lambda(x_m) - \lambda^{p-1}A_\lambda(x_n), J_\lambda x_n - J_\lambda x_m \right) \geq 0 \]

The expression on the left in the above is non positive. Multiplying by \(-1\),

\[ 0 \geq \lim_{m,n \to \infty} \sup \langle A_\lambda(x_n) - A_\lambda(x_m), J_\lambda x_n - J_\lambda x_m \rangle \]
\[ \geq \lim_{m,n \to \infty} \inf \langle A_\lambda(x_n) - A_\lambda(x_m), J_\lambda x_n - J_\lambda x_m \rangle \geq 0 \] (9.6.41)

Thus, in fact, the expression in (9.6.30) converges to 0. By boundedness considerations and the strong convergence given,

\[ \lim_{m,n \to \infty} \langle F(J_\lambda x_n - x_n) - F(J_\lambda x_m - x_m), J_\lambda x_n - x_n - (J_\lambda x_m - x_m) \rangle = 0 \]
(9.6.42)
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

From boundedness again, there is a subsequence still denoted with the subscript \( n \) such that
\[
J_{\lambda}x_n - x_n \to a - x, \quad F (J_{\lambda}x_n - x_n) \to b \quad \text{both weakly.}
\]
Since \( F \) is maximal monotone, (Theorem [46.24]) it follows from Lemma [46.28] that \([a-x,b] \in G (F) \) and so in fact \( F (a-x) = b \). Thus this has just shown that \( F (J_{\lambda}x_n - x_n) \to F (a-x) \). Next consider [46.24]. We have \( J_{\lambda}x_n \to a \) weakly and \( A_{\lambda} (x_n) = -\lambda^{-p-1}F (J_{\lambda}x_n - x_n) \to -\lambda^{-p-1}b \) weakly. Then from Lemma [46.28] again, \([a,-\lambda^{-p-1}b] \in G (A) \) so \(-\lambda^{-p-1}b \in A (a) \) so \( b \in -\lambda^{-p-1}A (a) \). But it was just shown that \( b = F (a-x) \) and so
\[
F (a-x) \in -\lambda^{-p-1}A (a) \quad \text{so} \quad 0 \in F (a-x) + \lambda^{-p-1}A (a) \quad \text{so} \quad a = J_{\lambda}x.
\]
As noted at the beginning, there is only one solution to this inclusion for a given \( x \) and it is \( a = J_{\lambda}x \). This has shown that in terms of weak convergence,
\[
A_{\lambda} (x_n) \to -\lambda^{-p-1}b = -\lambda^{-p-1}F (a-x) = -\lambda^{-p-1}F (J_{\lambda}x - x) \equiv A_{\lambda} (x)
\]
This has shown that \( A_{\lambda} \) is demicontinuous. Also it has shown that \( J_{\lambda} \) is also demicontinuous. (This result is a lot nicer in Hilbert space.)

3.) Why is \( \| A_{\lambda} (x) \| \leq |Ax| \) whenever \( x \in D (A) \)?
\[
A_{\lambda} (x) = -\lambda^{-p-1}F (J_{\lambda}x - x)
\]
where \( 0 \in F (J_{\lambda}x - x) + \lambda^{-p-1}A (J_{\lambda}x) \). Therefore, \( A_{\lambda} (x) \in A (J_{\lambda}x) \). Then letting \([u,v] \in G (A) \),
\[
0 \leq \langle v - A_{\lambda} (x), u - J_{\lambda}x \rangle
\]
In particular, if \( y \in Ax \)
\[
0 \leq \langle y - A_{\lambda} (x), x - J_{\lambda}x \rangle = \left\langle y + \lambda^{-p-1}F (J_{\lambda}x - x), x - J_{\lambda}x \right\rangle
\]
Hence
\[
\lambda^{-p-1} \| J_{\lambda}x - x \|^p \leq \| y \| \| J_{\lambda}x - x \|
\]
and so
\[
\lambda^{-p-1} \| J_{\lambda}x - x \|^{p-1} = \lambda^{-p-1} \| F (J_{\lambda}x - x) \| = \| A_{\lambda} (x) \| \leq \| y \|
\]
and since \( y \in Ax \) is arbitrary, \( \| A_{\lambda} (x) \| \leq |Ax| = \inf \{ \| y \| : y \in Ax \} \).

Next consider the claim that for all \( x \in \text{conv} (D (A)) \), it follows that
\[
\lim_{\lambda \to 0} J_{\lambda} (x) = x.
\]
Let \([u,v] \in G (A) \) and \( x \) is arbitrary.
\[
0 \leq \langle v - A_{\lambda} (x), u - J_{\lambda}x \rangle = \left\langle v + \lambda^{-p-1}F (J_{\lambda}x - x), u - J_{\lambda}x \right\rangle
\]
Let holds for any \( x \) must be the case that \( \text{map for } p > F \) the above results on approximation. It will include the general case of prescribed above, is maximal monotone with domain \( A \).

Now if \( x \) is in a convex reflexive Banach space. Then \( \text{Corollary 9.6.31} \) and so in fact, \( \lim \sup \lambda \).

Thus
\[
\|J_\lambda x - x\|^p \leq \lambda^{p-1} \langle v, u - x \rangle + \langle F(J_\lambda x - x), u - x \rangle + \lambda^{p-1} \langle v, x - J_\lambda x \rangle \quad (9.6.43)
\]
for \( x \) arbitrary and \( u \) anything in \( D(A) \). It follows that \( 9.6.33 \) holds for any \( u \in \text{conv } (D(A)) \). Say \( u = x_n \in \text{conv } (D(A)) \) where \( x_n \to x \). Then
\[
\|J_\lambda x - x\|^p \leq \lambda^{p-1} \langle v, x_n - x \rangle + \langle F(J_\lambda x - x), x_n - x \rangle + \lambda^{p-1} \langle v, x - J_\lambda x \rangle
\]
\[
\leq \lambda^{p-1} \|v\| \|x_n - x\| + \|J_\lambda x - x\|^{p-1} \|x_n - x\| + \lambda^{p-1} \|v\| \|J_\lambda x - x\|
\]
You have something like this:
\[
y_\lambda = \|J_\lambda x - x\|, a_n = \|x_n - x\|,
\]
\[
y_\lambda^p \leq \lambda^{p-1} \|v\| a_n + \lambda^{p-1} a_n + \lambda^{p-1} \|v\| y_\lambda, \quad y_\lambda \geq 0
\]
where \( p > 1 \) and \( a_n \to 0 \). Then
\[
\lim \sup_{\lambda \to 0} y_\lambda^p \leq \lim \sup_{\lambda \to 0} y_\lambda^{p-1} a_n
\]
and so,
\[
\lim \sup_{\lambda \to 0} y_\lambda \leq a_n
\]
Hence
\[
\lim \sup_{\lambda \to 0} \|J_\lambda x - x\| \leq \|x_n - x\|
\]
Since \( x_n \) is arbitrary, it follows that for every \( \varepsilon > 0 \),
\[
\lim \sup_{\lambda \to 0} \|J_\lambda x - x\| \leq \varepsilon
\]
and so in fact, \( \lim \sup_{\lambda \to \infty} \|J_\lambda x - x\| = 0. \)

Now here is an interesting corollary.

**Corollary 9.6.31** Let \( A \) be maximal monotone. \( A : X \to X' \) where \( X \) is a strictly convex reflexive Banach space. Then \( \overline{D(A)} \) is convex.

**Proof:** It is known that \( J_\lambda : X \to D(A) \) for any \( \lambda \). Also, if \( x \in \text{conv } (D(A)) \), then it was shown that \( J_\lambda x \to x \). Clearly
\[
\text{conv } (D(A)) \supseteq D(A)
\]
Now if \( x \) is in the set on the left, \( J_\lambda x \to x \) and so in fact, since \( J_\lambda x \in D(A) \), it must be the case that \( x \in D(A) \). Thus the two sets are the same and so in fact, \( D(A) \) is closed and convex. \( \blacksquare \)

Note that this implies that \( A(D(A)) \) is also convex. This is because \( A^{-1} \) described above, is maximal monotone with domain \( A(D(A)) \).

Next is a useful generalization of some of the earlier material used to establish the above results on approximation. It will include the general case of \( F \) a duality map for \( p > 1 \).
Proposition 9.6.32 Suppose $A : X \rightarrow P(X')$ where $X$ is a reflexive Banach space with strictly convex norm. Suppose also that $A$ is maximal monotone. Then if $\lambda_n \rightarrow 0$ and if $x_n \rightarrow x$ weakly, $A_{\lambda_n} x_n \rightarrow x^*$ weakly, and
\[
\lim_{n,m \rightarrow \infty} \sup \langle A_{\lambda_n} x_n - A_{\lambda_m} x_m, x_n - x_m \rangle \leq 0
\]
Then
\[
\lim_{n,m \rightarrow \infty} \langle A_{\lambda_n} x_n - A_{\lambda_m} x_m, x_n - x_m \rangle = 0,
\]
$[x, x^*] \in G(A)$, and $\langle A_{\lambda_n} x_n, x_n \rangle \rightarrow \langle x^*, x \rangle$.

Proof: Let $\alpha = \limsup_{n \rightarrow \infty} \langle A_{\lambda_n} x_n, x_n \rangle$. It is finite because the expression is bounded independent of $n$. Then
\[
\lim_{m \rightarrow \infty} \left( \limsup_{n \rightarrow \infty} \left( \langle A_{\lambda_n} x_n, x_n \rangle + \langle A_{\lambda_m} x_m, x_m \rangle \right) - \left[ \langle A_{\lambda_n} x_n, x_m \rangle + \langle A_{\lambda_m} x_m, x_n \rangle \right] \right) \leq 0
\]
Thus
\[
\lim_{m \rightarrow \infty} \left( \alpha + \langle A_{\lambda_m} x_m, x_m \rangle - \left[ \langle x^*, x_m \rangle + \langle A_{\lambda_m} x_m, x \rangle \right] \right) \leq 0
\]
and so
\[
2\alpha - 2 \langle x^*, x \rangle \leq 0
\]
The next simple observation is that
\[
\|A_{\lambda_n} x_n\| = \left\| \lambda_n^{-(p-1)} F(J_{\lambda_n} x_n - x_n) \right\| \leq C
\]
due to the weak convergence. Hence $\lambda_n^{-(p-1)} \|J_{\lambda_n} x_n - x_n\|^{p-1} \leq C$ and so
\[
\|J_{\lambda_n} x_n - x_n\| \leq \lambda_n C^{1/(p-1)}. \tag{9.6.44}
\]
Thus if $[u, u^*] \in G(A)$,
\[
\liminf_{n \rightarrow \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle = \liminf_{n \rightarrow \infty} \langle A_{\lambda_n} x_n - u^*, J_{\lambda_n} x_n - u \rangle \geq 0
\]
because $A_{\lambda} x \in AJ_{\lambda} x$. However, the left side satisfies
\[
0 \leq \liminf_{n \rightarrow \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle \leq \limsup_{n \rightarrow \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle = \limsup_{n \rightarrow \infty} \left[ \langle A_{\lambda_n} x_n, x_n \rangle - \langle A_{\lambda_n} x_n, u \rangle - \langle u^*, x_n \rangle + \langle u^*, u \rangle \right]
= \alpha - \langle x^*, u \rangle - \langle u^*, x \rangle + \langle u^*, u \rangle \leq \langle x^*, x \rangle - \langle x^*, u \rangle - \langle u^*, x \rangle + \langle u^*, u \rangle
= \langle x^* - u^*, x - u \rangle
\]
and this shows that $[x, x^*] \in G(A)$ since $[u, u^*]$ was arbitrary.
Next let \([u, u^*] \in \mathcal{G}(A)\). Then thanks to (9.6.4),

\[
0 \leq \lim_{n \to \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle = \lim_{n \to \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle
\]

\[
\leq \limsup_{n \to \infty} \langle A_{\lambda_n} x_n - u^*, x_n - u \rangle
\]

\[
= \limsup_{n \to \infty} \langle A_{\lambda_n} x_n, x_n \rangle - \langle A_{\lambda_n} x_n, u \rangle - \langle u^*, x_n \rangle + \langle u^*, u \rangle
\]

\[
\leq \langle x^*, x \rangle - \langle x^*, u \rangle - \langle u^*, x \rangle + \langle u^*, u \rangle = \langle x^* - u^*, x - u \rangle
\]

In particular, you could let \([u, u^*] = [x, x^*]\) and conclude that

\[
\lim_{n \to \infty} \langle A_{\lambda_n} x_n - x^*, x_n - x \rangle = \lim_{n \to \infty} \langle A_{\lambda_n} x_n, x_n \rangle - \langle A_{\lambda_n} x_n, x \rangle + \langle x^*, x \rangle - \langle x^*, x_n \rangle
\]

\[
= \lim_{n \to \infty} \langle A_{\lambda_n} x_n, x_n \rangle - \langle x^*, x \rangle + \langle x^*, x \rangle - \langle x^*, x \rangle = 0
\]

which shows that \(\lim_{n \to \infty} \langle A_{\lambda_n} x_n, x_n \rangle = \langle x^*, x \rangle\). Then it follows from this that

\[
\lim_{n,m \to \infty} \langle A_{\lambda_n} x_n - A_{\lambda_m} x_m, x_n - x_m \rangle = 0 \quad \blacksquare
\]

For the rest of this, the usual duality map for \(p = 2\) will be used. It may be that one could change this, but I don’t have a need to do it right now so from now on, \(F\) will be the usual thing.

### 9.6.4 Sum Of Maximal Monotone Operators

To begin with, here is a nice lemma.

#### Lemma 9.6.33

Let \(0 \in D(A)\) and let \(A\) be maximal monotone and let \(B : X \to X'\) be monotone hemicontinuous, bounded, and coercive. Then \(B + A\) is also maximal monotone. Also \(B + A\) is onto.

**Proof:** By Theorem (9.6.3), there exists \(x \in \overline{D(A)}\) such that for all \([u, u^*] \in \mathcal{G}(A)\),

\[
\langle Bx + Fx - y^* + u^*, u - x \rangle \geq 0
\]

Hence for all \([u, u^*]\),

\[
\langle u^* - (y^* - (Bx + Fx)), u - x \rangle \geq 0
\]

It follows that

\[
y^* - (Bx + Fx) \in Ax
\]

and so \(y^* \in Bx + Ax + Fx\) showing that \(B + A\) is maximal monotone because it added to \(F\) is onto. As to the last claim, just don’t add in \(F\) in the argument. Thus for all \([u, u^*]\),

\[
\langle Bx - y^* + u^*, u - x \rangle \geq 0
\]

Then the rest is as before. You find that \(y^* - Bx \in Ax. \quad \blacksquare\)
Suppose instead of \(9.6.33\) let
\[
\text{Hence, by Lemma } F
\]
Then
\[
y^* = \tilde{A}x + \tilde{B}x \equiv A(x_0 + x) + B(x_0 + x)
\]

**Lemma 9.6.35** Let 0 be on the interior of \(D(A)\) and also in \(D(B)\). Also let 0 \(\in B(0)\) and 0 \(\in A(0)\). Then if \(A, B\) are maximal monotone, so is \(A + B\).

**Proof:** Note that, since 0 \(\in A(0)\), if \(x^* \in Ax\), then \(\langle x^*, x \rangle \geq 0\). Also note that \(|B(0)| \leq |B(0)| = 0\) and so also \(\langle B(x), x \rangle \geq 0\). It is necessary to show that \(F + A + B\) is onto. However, \(B(\lambda)\) is monotone hemicontinuous, bounded and coercive. Hence, by Lemma 9.6.34, \(B(\lambda) + A\) is maximal monotone. If \(x^* \in X'\) is given, there exists a solution to
\[
x^* \in Fx_\lambda + B_\lambda x_\lambda + Ax_\lambda
\]
Do both sides to \(x_\lambda\) and let \(x^*_\lambda \in Ax_\lambda\) be such that equality holds in the above.
\[
x^* = Fx_\lambda + B_\lambda x_\lambda + x^*_\lambda \tag{9.6.45}
\]
Then
\[
\langle x^*, x_\lambda \rangle = \|x_\lambda\|^2 + \langle x^*_\lambda, x_\lambda \rangle
\]
It follows that
\[
\|x_\lambda\| \leq \|x^*\|, \quad \langle x^*_\lambda, x_\lambda \rangle \leq \langle x^*, x_\lambda \rangle \leq \|x^*\| \|x_\lambda\| \leq \|x^*\|^2 \tag{9.6.46}
\]
Next, 0 is on the interior of \(D(A)\) and so from Theorem 9.6.11, there exists \(\rho > 0\) such that if \(y^* \in Ax\) for \(\|x\| \leq \rho\), then \(\|y^*\| < M\) and in fact, all such \(x\) are in \(D(A)\). Now let
\[
y_\lambda = \frac{1}{2\|x^*_\lambda\|} F^{-1}(x^*_\lambda) \text{ so } \|y_\lambda\| < \rho
\]
Thus \(y_\lambda \in D(A)\) and if \(y^*_\lambda \in Ay_\lambda\), then \(\|y^*_\lambda\| < M\). Then for such bounded \(y^*_\lambda\),
\[
0 \leq \langle y^*_\lambda - x^*_\lambda, y_\lambda - x_\lambda \rangle = \langle y^*_\lambda, y_\lambda \rangle - \langle x^*_\lambda, x_\lambda \rangle
\]
Then
\[
\frac{1}{2} \|x^*_\lambda\|^2 = \left\langle x^*_\lambda, \frac{1}{2\|x^*_\lambda\|} F^{-1}(x^*_\lambda) \right\rangle = \langle x^*_\lambda, y_\lambda \rangle \leq \langle y^*_\lambda, y_\lambda \rangle - \langle y^*_\lambda, x_\lambda \rangle + \langle x^*_\lambda, x_\lambda \rangle
\]
\[
\leq M\rho + M \|x_\lambda\| + \langle x^*_\lambda, x_\lambda \rangle
\]
9.6. MAXIMAL MONOTONE OPERATORS

From (9.6.46),

$$\|x_\lambda\| \leq 2 \left( M \rho + M \|x^*\| + \|x^*\|^2 \right)$$

Thus from (9.6.46), $x_\lambda, x^*_\lambda, Fx_\lambda$ are all bounded. Hence it follows from (9.6.45) that $B_\lambda x_\lambda$ is also bounded. Therefore, there is a sequence, $\lambda_n \to 0$ such that

$$x_{n\lambda} \to z \text{ weakly}$$
$$x_\lambda \to w^* \text{ weakly}$$
$$Fx_\lambda \to u^* \text{ weakly}$$
$$B_{\lambda_n} x_{\lambda_n} \to b^* \text{ weakly}$$

Using (9.6.45), it follows that

$$\langle Fx_{\lambda_n} + x^*_\lambda + B_{\lambda_n} x_{\lambda_n} - (Fx_{\lambda_m} + x^*_\lambda + B_{\lambda_m} x_{\lambda_m}), x_{\lambda_n} - x_{\lambda_m} \rangle = 0$$

Thus

$$\langle Fx_{\lambda_n} + x^*_\lambda, x_{\lambda_n} - x_{\lambda_m} \rangle + \langle B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \rangle = 0$$

Now $F + A$ is surely monotone and so

$$\lim_{m,n \to \infty} \sup \langle B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \rangle \leq 0$$

By Proposition 9.6.32, $b^* \in Bz$ and

$$\lim_{m,n \to \infty} \langle B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \rangle = 0$$

Then returning to (9.6.47),

$$\lim_{m,n \to \infty} \langle Fx_{\lambda_n} + x^*_\lambda - (Fx_{\lambda_m} + x^*_\lambda), x_{\lambda_n} - x_{\lambda_m} \rangle \leq 0$$

Now from Lemma 9.6.33, $F + A$ is maximal monotone. Hence Proposition 9.6.32 applies again and it follows that $u^* + w^* \in Fz + Az$. Then passing to the limit as $n \to \infty$ in

$$x^* = Fx_{\lambda_n} + B_{\lambda_n} x_{\lambda_n} + x^*_\lambda$$

it follows that

$$x^* = u^* + b^* + w^* = Fz + Az + Bz$$

and this shows that $A + B$ is maximal monotone because $x^*$ was arbitrary. ■

You don’t need to assume all that stuff about $0 \in A(0), 0 \in B(0), 0$ on interior of $D(A)$ and so forth.

**Theorem 9.6.36** Suppose $A, B$ are maximal monotone and the interior of $D(A)$ has nonempty intersection with $D(B)$. Then $A + B$ is maximal monotone.
9.6.34

**Proof:** Let \( x_0 \) be on the interior of \( D(A) \) and also in \( D(B) \). Let \( \hat{A}(x) = A(x_0 + x) - x_0^* \in A(x_0) \). Thus \( 0 \in D(\hat{A}) \) and \( 0 \in \hat{A}(0) \). Do the same thing for \( B \) to get \( \hat{B} \) defined similarly. Are these still maximal monotone? Suppose for all \([u, u^*] \in \mathcal{G}(\hat{A})\)

\[
\langle y^* - u^*, y - u \rangle \geq 0
\]

Does it follow that \( y^* \in \hat{A}y \)? It is given that \( u^* \in A(x_0 + u) \). The above implies for all \([u, u^*] \in \mathcal{G}(\hat{A})\)

\[
\langle y^* + x_0^* - (u^* + x_0^*), (y + x_0) - (u + x_0) \rangle \geq 0
\]

and since \( u + x_0 \) is a generic element of \( D(A) \) for \( u \in D(\hat{A}) \), the above implies \( y^* + x_0^* \in A(y + x_0) \) and so \( y \in A(y + x_0) - x_0^* \equiv \hat{A}(y) \). Hence the graph is maximal. Similar for \( \hat{B} \). Thus the lemma can be applied to \( \hat{A}, \hat{B} \) to conclude that the sum of these is maximal monotone. Now a repeat of the above reasoning which shows that \( A \) is maximal monotone shows that the fact that \( \hat{A} + \hat{B} \) is maximal monotone implies that \( A + B \) is also. You just shift with \( -x_0 \) instead of \( x_0 \). It amounts to nothing more than the observation that maximal graphs don’t lose their maximality by shifting their ranges and domains. ■

Suppose \( B, A \) are maximal monotone. Does there always exist a solution \( x \) to

\[
x^* \in Fx + B_\lambda x + Ax \quad (9.6.48)
\]

Consider the monotone hemicontinuous and bounded operator \( F + B_\lambda \) defined by

\[
(\hat{F} + \hat{B}_\lambda)(x) = (\hat{F} + \hat{B}_\lambda)(x + x_0)
\]

also coercive for some \( x_0 \in D(A) \)? If so, the existence of the desired solution to the above inclusion follows from Corollary [9.6.33]. Then for all \( \|x\| \) large enough that \( \|x + x_0\| > \|x_0\| \),

\[
\frac{\langle F(x + x_0) + B_\lambda(x + x_0), x \rangle}{\|x\|} = \frac{\langle F(x + x_0), x \rangle}{\|x\|} + \frac{\langle B_\lambda(x + x_0) - B_\lambda(x_0), x \rangle}{\|x\|} + \frac{\langle B_\lambda(x_0), x \rangle}{\|x\|}
\]

\[
\geq \frac{1}{2} \frac{\|F(x + x_0), x\|}{\|x + x_0\|} - \|B_\lambda(x_0)\|
\]

\[
\geq \frac{1}{2} \frac{\langle F(x + x_0), x \rangle}{\|x + x_0\|} - \frac{1}{2} \frac{\|F(x + x_0), x_0\|}{\|x + x_0\|} - \|B_\lambda(x_0)\|
\]

\[
\geq \frac{1}{2} \frac{\langle F(x + x_0), x \rangle}{\|x + x_0\|} - \frac{1}{2} \frac{\|F(x + x_0), x_0\|}{\|x_0\|} - \|B_\lambda(x_0)\|
\]

\[
\geq \frac{1}{2} \frac{\langle F(x + x_0), x_0 \rangle}{\|x + x_0\|} - \frac{1}{2} \|x + x_0\| - \|B_\lambda(x_0)\|
\]
9.6. MAXIMAL MONOTONE OPERATORS

\[ \frac{1}{2} \| x + x_0 \|^2 - \frac{1}{2} \| x + x_0 \| - \| B_\lambda (x_0) \| \]

which shows that

\[ \lim_{\| x \| \to \infty} \frac{\langle F(x + x_0) + B_\lambda (x + x_0), x \rangle}{\| x \|} = \infty \]

and so by Corollary 9.6.34, there exists a solution to 9.6.48. This shows half of the following interesting theorem which is another version of the above major result.

**Theorem 9.6.37** Suppose \( A, B \) are maximal monotone operators. Then for each \( x^* \in X' \), there exists a solution \( x_\lambda \) to

\[ x^* \in Fx_\lambda + B_\lambda x_\lambda + Ax_\lambda, \; \lambda > 0 \quad (9.6.49) \]

If for \( \lambda \in (0, \delta) \), \( \{B_\lambda x_\lambda\} \) is bounded, then there exists a solution \( x \) to

\[ x^* \in Fx + Bx + Ax \]

**Proof:** The existence of a solution to the inclusion 9.6.49 comes from the above discussion. The last claim follows from almost a repeat of the last part of the proof of the above theorem. Since \( \{B_\lambda x_\lambda\} \) is given to be bounded for \( \lambda \in (0, \delta) \), there is a sequence, \( \lambda_n \to 0 \) such that

\[ x_{\lambda_n} \to z \text{ weakly} \]
\[ x^*_\lambda \to w^* \text{ weakly} \]
\[ Fx_\lambda \to u^* \text{ weakly} \]
\[ B_\lambda x_\lambda \to b^* \text{ weakly} \]

Using 9.6.34, it follows that

\[ \langle Fx_{\lambda_n} + x^*_{\lambda_n} + B_{\lambda_n} x_{\lambda_n} - (Fx_{\lambda_m} + x^*_{\lambda_m} + B_{\lambda_m} x_{\lambda_m}), x_{\lambda_n} - x_{\lambda_m} \rangle = 0 \]

Thus

\[ \langle Fx_{\lambda_n} + x^*_{\lambda_n} - (Fx_{\lambda_m} + x^*_{\lambda_m}), x_{\lambda_n} - x_{\lambda_m} \rangle \]
\[ + \langle B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \rangle = 0 \quad (9.6.50) \]

Now \( F + A \) is surely monotone and so

\[ \lim_{m,n \to \infty} \sup \{ B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \} \leq 0 \]

By Proposition 9.6.32, \( b^* \in Bz \) and

\[ \lim_{m,n \to \infty} \langle B_{\lambda_n} x_{\lambda_n} - B_{\lambda_m} x_{\lambda_m}, x_{\lambda_n} - x_{\lambda_m} \rangle = 0 \]

Then returning to 9.6.31,

\[ \lim_{m,n \to \infty} \sup \langle Fx_{\lambda_n} + x^*_{\lambda_n} - (Fx_{\lambda_m} + x^*_{\lambda_m}), x_{\lambda_n} - x_{\lambda_m} \rangle \leq 0 \]
Now from Corollary 9.6.31, \( F + A \) is maximal monotone (In fact, \( F + A \) is onto). Hence Proposition 9.6.32 applies again and it follows that \( u^* + w^* \in Fz + Az \). Then passing to the limit as \( n \to \infty \) in
\[
x^* = Fx_{\lambda_n} + B_{\lambda_n}x_{\lambda_n} + x_{\lambda_n}^*
\]
it follows that
\[
x^* = u^* + b^* + w^* = Fz + Az + Bz \quad \blacksquare
\]

### 9.6.5 Convex Functions, An Example

As before, \( X \) will be a Banach space in what follows. Sometimes it will be a reflexive Banach space and in this case, it will be assumed that the norm is strictly convex.

**Definition 9.6.38** Let \( \phi : X \to (-\infty, \infty] \). Then \( \phi \) is convex if whenever \( t \in [0,1] \), \( x, y \in X \),
\[
\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)
\]
The epigraph of \( \phi \) is defined by
\[
\text{epi}(\phi) \equiv \{(x,y) : y \geq \phi(x)\}
\]
When \( \text{epi}(\phi) \) is closed in \( X \times (-\infty, \infty] \), we say that \( \phi \) is lower semicontinuous, l.s.c. The function is called proper if \( \phi(x) < \infty \) for some \( x \). The collection of all such \( x \) is called \( D(\phi) \), the domain of \( \phi \).

This definition of lower semicontinuity is equivalent to the usual definition.

**Lemma 9.6.39** The above definition of lower semicontinuity is equivalent to the assertion that whenever \( x_n \to x \), it follows that \( \phi(x) \leq \liminf_{n \to \infty} \phi(x_n) \). In case that \( \phi \) is convex, lower semicontinuity is equivalent to weak lower semicontinuity. That is \( \text{epi}(\phi) \) is closed if and only if \( \text{epi}(\phi) \) is weakly closed. In this case, the limit condition: If \( x_n \to x \) weakly, then \( \phi(x) \leq \liminf_{n \to \infty} \phi(x_n) \) is valid.

**Proof:** Suppose the limit condition holds. Why is \( \text{epi}(\phi) \) closed? Why is \( X \times (-\infty, \infty] \setminus \text{epi}(\phi) \equiv \text{epi}(\phi)^C \) open? Let \( (x, \alpha) \in \text{epi}(\phi)^C \). Then \( \alpha < \phi(x) \). Consider \( B(x, r) \times (\alpha - \frac{\delta}{2}, \alpha + \frac{\delta}{2}) \). If every such open set contains a point of \( \text{epi}(\phi) \), then there exists \( x_n \to x \), \( y_n < \alpha + \frac{\delta}{2} \), \( y_n \geq \phi(x_n) \). Hence, from the limit condition,
\[
\phi(x) \leq \liminf_{n \to \infty} \phi(x_n) \leq \liminf_{n \to \infty} y_n \leq \alpha + \frac{\delta}{2} < \alpha + \frac{\delta}{2} < \phi(x)
\]
a contradiction. It follows that there exists \( r > 0 \) such that \( B(x, r) \times (\alpha - \frac{\delta}{2}, \alpha + \frac{\delta}{2}) \cap \text{epi}(\phi) = \emptyset \). Since \( \text{epi}(\phi)^C \) is open, it follows that \( \text{epi}(\phi) \) is closed.

Next suppose \( \text{epi}(\phi) \) is closed. Why does the limit condition hold? Suppose \( x_n \to x \). Then \( (x_n, \phi(x_n)) \in \text{epi}(\phi) \). There is a subsequence such that
\[
\alpha \equiv \liminf_{n \to \infty} \phi(x_n) = \lim_{k \to \infty} \phi(x_{n_k})
\]
and so \((x_{n_k}, \phi(x_{n_k})) \to (x, \alpha)\). Since \(\text{epi}(\phi)\) is closed, this means \((x, \alpha) \in \text{epi}(\phi)\). Hence
\[
\alpha \equiv \lim_{n \to \infty} \inf \phi(x_n) \geq \phi(x).
\]

Consider the last claim. In this case, \(\text{epi}(\phi)\) is convex. If it is closed, then it is weakly closed thanks to separation theorems: If \((x, \alpha) \in \text{epi}(\phi)^C\), then \(\alpha < \infty\) and so there exists \((x^*, \beta) \in (X \times \mathbb{R})^t\) and \(l\) such that for all \((t, \gamma) \in \text{epi}(\phi)\),
\[
x^*(t) + \beta\gamma > l > x^*(x) + \alpha\beta
\]
Then \(B_{(x^*, \beta)}((x, \alpha), \delta)\) is a weakly open set containing \((x, \alpha)\). For \(\delta\) small enough, it does not intersect \(\text{epi}(\phi)\) since if not so, there would exist \((t_n, \gamma_n) \in \text{epi}(\phi) \cap B_{(x^*, \beta)}((x, \alpha), \frac{1}{n})\) and so
\[
x^*(t_n) + \beta\gamma_n \to x^*(x) + \alpha\beta
\]
contrary to the above inequality. Thus \(\text{epi}(\phi)\) is weakly closed. Also, if \(\text{epi}(\phi)\) is weakly closed, then it is obviously strongly closed.

What of the limit condition using weak convergence instead of strong convergence? Say \(x_n \to x\) weakly. Does it follow that if \(\text{epi}(\phi)\) is weakly closed that \(\phi(x) \leq \lim \inf_{n \to \infty} \phi(x_n)\)? It is just as above. There is a subsequence such that
\[
\alpha \equiv \lim_{n \to \infty} \inf \phi(x_n) = \lim_{k \to \infty} \phi(x_{n_k})
\]
and so \((x_{n_k}, \phi(x_{n_k})) \to (x, \alpha)\) weakly. Since \(\text{epi}(\phi)\) is weakly closed, this means \((x, \alpha) \in \text{epi}(\phi)\). Hence
\[
\alpha \equiv \lim_{n \to \infty} \inf \phi(x_n) \geq \phi(x). \blacksquare
\]

There is also another convenient characterization of what it means for a function to be lower semicontinuous.

**Lemma 9.6.40** Let \(\phi : X \to (-\infty, \infty]\). Then \(\phi\) is lower semicontinuous if and only if \(\phi^{-1}((a, \infty])\) is open for any \(a \in \mathbb{R}\).

**Proof:** Suppose first that \(\text{epi}(\phi)\) is closed. Consider \(x \in \phi^{-1}((a, \infty])\). Thus \(\phi(x) > a\). Thus \((x, a) \in \text{epi}(\phi)^C\) because \(a < \phi(x)\). Since \(\text{epi}(\phi)\) is closed, there exists \(r, \varepsilon > 0\) such that
\[
B(x, r) \times (a - \varepsilon, a + \varepsilon) \subseteq \text{epi}(\phi)^C
\]
Hence if \(y \in B(x, r)\), it follows that \(\phi(y) \geq a + \varepsilon\) since otherwise there would be a point of \(\text{epi}(\phi)^C\) in this open set \(B(x, r) \times (a - \varepsilon, a + \varepsilon)\). Hence \(B(x, r) \subseteq \phi^{-1}((a, \infty])\).

Conversely, suppose \(\phi^{-1}((a, \infty])\) is open for any \(a\) and let \((x, b) \in \text{epi}(\phi)^C\). Then \(\phi(x) > b\). Thus there exists \(B(x, r)\) such that for \(y \in B(x, r)\), it follows that \(\phi(y) > b\). That is, \(y \in \phi^{-1}((b, \infty])\). So consider \(B(x, r) \times (-\infty, b)\). If \((y, a) \in\)
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

$B(x, r) \times (-\infty, b)$, then since $\phi(y) > b, \alpha < \phi(y)$ and so there is no point of intersection between epi($\phi$) and this open set $B(x, r) \times (-\infty, b)$.

Of course one can define upper semicontinuous the same way that $\phi^{-1}(-\infty, a)$ is open. Thus a function is continuous if and only if it is both upper and lower semicontinuous.

In case $X$ is reflexive, the limit condition implies that epi($\phi$) is weakly closed. Suppose $(x, \alpha)$ is a weak limit point of epi($\phi$). Then by the Eberlein Smulian theorem, Problem 14 on Page 66, there is a subsequence of points of $X$, $(x_n, \alpha_n)$ which converges weakly to $(x, \alpha)$. Thus if the limit condition holds,

$$\phi(x) \leq \liminf_{n \to \infty} \phi(x_n) \leq \liminf_{n \to \infty} \alpha_n = \alpha$$

and so $(x, \alpha) \in$ epi($\phi$). If $X$ is not reflexive, this isn’t all that clear because it is not clear that a limit point is the limit of a sequence. However, one could consider a limit condition involving nets and get a similar result.

**Definition 9.6.41** Let $\phi : X \to (-\infty, \infty]$ be convex lower semicontinuous, and proper. Then

$$\partial\phi(x) \equiv \{x^* : \phi(y) - \phi(x) \geq \langle x^*, y - x \rangle \text{ for all } y\}$$

The domain of $\partial\phi$, denoted as $D(\partial\phi)$ is just the set of all $x$ for which $\partial\phi(x) \neq \emptyset$. Note that $D(\partial\phi) \subseteq D(\phi)$ since if $x \notin D(\phi)$, the defining inequality could not hold for all $y$ because the left side would be $-\infty$ for some $y$.

**Theorem 9.6.42** For $X$ a real Banach space, let $\phi(x) \equiv \frac{1}{2}||x||^2$. Then $F(x) = \partial\phi(x)$. Here $F$ was the set valued map satisfying $x^* \in Fx$ means

$$||x^*|| = ||Fx||, \langle Fx, x \rangle = ||x||^2$$

**Proof:** Let $x^* \in F(x)$. Then

$$\langle x^*, y - x \rangle = \langle x^*, y \rangle - \langle x^*, x \rangle \leq ||x|| ||y|| - ||x||^2 \leq \frac{1}{2} ||y||^2 - \frac{1}{2} ||x||^2.$$ (9.6.51)

This shows $F(x) \subseteq \partial\phi(x)$.

Now let $x^* \in \partial\phi(x)$. Then for all $t \in \mathbb{R}$,

$$\langle x^*, ty + x \rangle = \langle x^*, (ty + x) - x \rangle \leq \frac{1}{2} \left(||x + ty||^2 - ||x||^2\right).$$

Now if $t > 0$, divide both sides by $t$. This yields

$$\langle x^*, y \rangle \leq \frac{1}{2t} \left(||x|| + t ||y||\right) - \frac{1}{2t} ||x||^2 = \frac{1}{2t} \left(2t ||x|| ||y|| + t^2 ||y||^2\right)$$
Letting $t \to 0$,
\[
\langle x^*, y \rangle \leq ||x|| \cdot ||y||.
\] (9.6.52)

Next suppose $t = -s$, where $s > 0$ in [9.6.51]. Then, since when you divide by a negative, you reverse the inequality, for $s > 0$
\[
\langle x^*, y \rangle \geq \frac{1}{2s} \left[ ||x||^2 - 2||x - sy|| \cdot ||sy|| + ||sy||^2 - ||x - sy|| \right] .
\] (9.6.53)

Taking a limit as $s \to 0$ yields
\[
\langle x^*, y \rangle \geq -||x|| \cdot ||y||.
\] (9.6.55)

It follows from [9.6.55] and [9.6.52] that
\[
||\langle x^*, y \rangle|| \leq ||x|| \cdot ||y||
\]
and that, therefore, $||x^*|| \leq ||x||$ and $||\langle x^*, x \rangle|| \leq ||x||^2$. Now return to [9.6.54] and let $y = x$. Then
\[
\langle x^*, x \rangle \geq \frac{1}{2s} \left[ -2||x - sx|| \cdot ||sx|| + ||sx||^2 \right] \\
= -||x||^2 \left( 1 - s \right) + s ||x||^2
\]
Letting $s \to 1$,
\[
\langle x^*, x \rangle \geq ||x||^2 .
\]

Since it was already shown that $||\langle x^*, x \rangle|| \leq ||x||^2$, this shows $\langle x^*, x \rangle = ||x||^2$ and also $||x^*|| \leq ||x||$. Thus
\[
||x^*|| \geq \langle x^*, \frac{x}{||x||} \rangle = ||x||
\]
so in fact $x^* \in F(x)$. □

The next result gives conditions under which the subgradient is onto. This means that if $y^* \in X'$, then there exists $x \in X$ such that $y^* \in \partial \phi (x)$.

**Theorem 9.6.43** Suppose $X$ is a reflexive Banach space and suppose $\phi : X \to (-\infty, \infty)$ is convex, proper, l.s.c., and for all $y^* \in X'$, $x \to \phi (x) - \langle y^*, x \rangle$ is coercive. Then $\partial \phi$ is onto.

**Proof:** The function $x \to \phi (x) - y^* (x) \equiv \psi (x)$ is convex, proper, l.s.c., and coercive. Let
\[
\lambda \equiv \inf \{ \phi (x) - \langle y^*, x \rangle : x \in X \}
\]
and let \( \{x_n\} \) be a minimizing sequence satisfying
\[
\lambda = \lim_{n \to \infty} \phi(x_n) - \langle y^*, x_n \rangle
\]
By coercivity,
\[
\lim_{||x|| \to \infty} \phi(x) - \langle y^*, x \rangle = \infty
\]
and so this minimizing sequence is bounded. By the Eberlein Smulian theorem, Problem 14 on Page 66, there is a weakly convergent subsequence \( x_{nk} \to x \). By Lemma 9.6.39,
\[
\lambda = \phi(x) - \langle y^*, x \rangle \leq \liminf_{k \to \infty} \phi(x_{nk}) - \langle y^*, x_{nk} \rangle = \lambda
\]
so there exists \( x \) which minimizes \( x \to \phi(x) - \langle y^*, x \rangle \equiv \psi(x) \). Therefore, \( 0 \in \partial \psi(x) \) because
\[
\psi(y) - \psi(x) \geq 0 = \langle 0, y - x \rangle
\]
Thus, \( 0 \in \partial \psi(x) = \partial \phi(x) - y^* \). 

Now let \( \phi \) be a convex proper lower semicontinuous function defined on \( X \) where \( X \) is a reflexive Banach space with strictly convex norm. Consider \( \partial \phi \). Is it maximal monotone? Is it the case that \( F + \partial \phi \) is onto? First of all, is \( \partial \phi \) monotone? Let \( x^* \in \partial \phi(x), y^* \in \partial \phi(y) \). Then
\[
\phi(y) - \phi(x) \geq \langle x^*, y - x \rangle
\]
\[
\phi(x) - \phi(y) \geq \langle y^*, x - y \rangle
\]
Hence adding these yields
\[
\langle y^* - x^*, x - y \rangle \leq 0, \quad \langle y^* - x^*, y - x \rangle \geq 0.
\]
Yes, \( \partial \phi \) is certainly monotone. Is it maximal monotone?

**Theorem 9.6.44** Let \( \phi \) be convex, proper, and lower semicontinuous on \( X \) where \( X \) is a reflexive Banach space having strictly convex norm. Then \( \partial \phi \) is maximal monotone.

**Proof:** It is necessary to show that \( F + \partial \phi \) is onto. To do this, let
\[
\psi(x) \equiv \frac{1}{2} ||x||^2 + \phi(x) - \langle y^*, x \rangle
\]
where \( y^* \) is a given element of \( X' \) and the idea is to show that \( y^* \in F(x) + \partial \phi(x) \) for some \( x \). Then by separation theorems, \( \phi(x) \geq b + \langle z^*, x \rangle \) for some \( b, z^* \). Hence it is clear that \( \psi \) is convex, lower semicontinuous and coercive in the sense that
\[
\lim_{||x|| \to \infty} \psi(x) = \infty
\]
It follows that any minimizing sequence for \( \psi \) is bounded. Hence by the weak lower semicontinuity, this function has a minimum at \( x_0 \) say. Thus
\[
\frac{1}{2} \|x_0\|^2 + \phi(x_0) - \langle y^*, x_0 \rangle \leq \frac{1}{2} \|x\|^2 + \phi(x) - \langle y^*, x \rangle
\]
for all \( x \). Then
\[
\frac{1}{2} \|x_0\|^2 - \frac{1}{2} \|x\|^2 + \langle y^*, x - x_0 \rangle \leq \phi(x) - \phi(x_0)
\]
Now from Theorem 9.6.45,
\[
\langle F(x), x_0 - x \rangle \leq \frac{1}{2} \|x_0\|^2 - \frac{1}{2} \|x\|^2
\]
and so, the above reduces to
\[
\langle F(x), x_0 - x \rangle + \langle y^*, x - x_0 \rangle \leq \phi(x) - \phi(x_0)
\]
Next let \( x = x_0 + t(z - x_0), t \in (0, 1), \) where \( z \) is arbitrary. Then
\[-t \langle F(x_0 + t(z - x_0)), z - x_0 \rangle + t \langle y^*, z - x_0 \rangle \leq \phi(x_0 + t(z - x_0)) - \phi(x_0)\]
and so, by convexity,
\[-t \langle F(x_0 + t(z - x_0)), z - x_0 \rangle + t \langle y^*, z - x_0 \rangle \leq (1 - t) \phi(x_0) + t \phi(z) - \phi(x_0)\]
\[t \langle y^*, z - x_0 \rangle \leq t (\phi(z) - \phi(x_0)) + t \langle F(x_0 + t(z - x_0)), z - x_0 \rangle\]
Now cancel the \( t \) on both sides to obtain
\[
\langle y^*, z - x_0 \rangle \leq \phi(z) - \phi(x_0) + \langle F(x_0 + t(z - x_0)), z - x_0 \rangle
\]
By the fact that \( F \) is hemicontinuous, actually demicontinuous, one can let \( t \downarrow 0 \) and obtain
\[
\langle y^*, z - x_0 \rangle \leq \phi(z) - \phi(x_0) + \langle F(x_0), z - x_0 \rangle
\]
This says that \( y^* - F(x_0) \in \partial \phi(x_0) \) from the definition of what \( \partial \phi(x_0) \) means. ■

There is a much harder approach to this theorem which is based on a theorem about when the subgradient of a sum equals the sum of the subgradients. This major theorem is given next. Much of the above is in [2] but I don’t remember where I found the following proof.

**Theorem 9.6.45** Let \( \phi_1 \) and \( \phi_2 \) be convex, l.s.c. and proper having values in \((-\infty, \infty]\). Then
\[
\partial (\lambda \phi_1)(x) = \lambda \partial \phi_1(x), \quad \partial (\phi_1 + \phi_2)(x) \supseteq \partial \phi_1(x) + \partial \phi_2(x)
\] (9.6.56)
if \( \lambda > 0 \). If there exists \( \bar{x} \in \text{dom} \phi_1 \cap \text{dom} \phi_2 \) and \( \phi_1 \) is continuous at \( \bar{x} \) then for all \( x \in X \),
\[
\partial (\phi_1 + \phi_2)(x) = \partial \phi_1(x) + \partial \phi_2(x).
\] (9.6.57)
\textbf{Proof:} It is obvious so we only need to show \textbf{5.6.37}. Suppose \( \varphi \) is as described. It is clear \textbf{5.6.36} holds whenever \( x \notin \text{dom} (\phi_1) \cap \text{dom} (\phi_2) \) since then \( \partial (\phi_1 + \phi_2) = \emptyset \). Therefore, assume

\[ x \in \text{dom} (\phi_1) \cap \text{dom} (\phi_2) \]

in what follows. Let \( x^* \in \partial (\phi_1 + \phi_2) (x) \). Is \( x^* \) is the sum of an element of \( \partial \phi_1 (x) \) and \( \partial \phi_2 (x) \)? Does there exist \( x_1^* \) and \( x_2^* \) such that for every \( y \),

\[ x^* (y - x) = x_1^* (y - x) + x_2^* (y - x) \leq \phi_1 (y) - \phi_1 (x) + \phi_2 (y) - \phi_2 (x) ? \]

If so, then

\[ \phi_1 (y) - \phi_1 (x) - x^* (y - x) \geq \phi_2 (x) - \phi_2 (y) . \]

Define

\[ C_1 = \{(y, a) \in X \times \mathbb{R} : \phi_1 (y) - \phi_1 (x) - x^* (y - x) \leq a \} , \]

\[ C_2 = \{(y, a) \in X \times \mathbb{R} : a \leq \phi_2 (x) - \phi_2 (y) \} . \]

I will show \( \text{int} (C_1) \cap C_2 = \emptyset \) and then by Problem 14 on Page 366 there exists an element of \( X' \) which does something interesting.

Both \( C_1 \) and \( C_2 \) are convex and nonempty. Say \( y_1, y_2 \in C_1 \) and \( t \in [0, 1] \). Then

\[ \phi_1 (((ty_1) + (1 - t) y_2) - \phi_1 (x) - x^* (((ty_1) + (1 - t) y_2) - x) \]

\[ \leq t \phi (y_1) + (1 - t) \phi (y_2) - (t \phi_1 (x) + (1 - t) \phi (x)) \]

\[ - (t x^* (y_1 - x) + (1 - t) x^* (y_2 - x)) \]

\[ \leq ta + (1 - t) a = a \]

so \( C_1 \) is indeed convex. The case of \( C_2 \) is similar.

\( C_1 \) is nonempty because it contains \( \varphi, \phi_1 (\varphi) - \phi_1 (x) - x^* (\varphi - x) \) since

\[ \phi_1 (\varphi) - \phi_1 (x) - x^* (\varphi - x) \leq \phi_1 (\varphi) - \phi_1 (x) - x^* (\varphi - x) \]

\( C_2 \) is also nonempty because it contains \( \varphi, \phi_2 (x) - \phi_2 (\varphi) \) since

\[ \phi_2 (x) - \phi_2 (\varphi) \leq \phi_2 (x) - \phi_2 (\varphi) \]

In addition to this,

\[ (\varphi, \phi_1 (\varphi) - x^* (\varphi - x) - \phi_1 (x) + 1) \in \text{int} (C_1) \]

due to the assumed continuity of \( \phi_1 \) at \( \varphi \) and so \( \text{int} (C_1) \neq \emptyset \). If \( (y, a) \in \text{int} (C_1) \) then

\[ \phi_1 (y) - x^* (y - x) - \phi_1 (x) \leq a - \varepsilon \]
whenever $\varepsilon$ is small enough. Therefore, if $(y, a)$ is also in $C_2$, the assumption that $x^* \in \partial (\phi_1 + \phi_2) (x)$ implies
\[ a - \varepsilon \geq \phi_1 (y) - x^* (y - x) - \phi_1 (x) \geq \phi_2 (x) - \phi_2 (y) \geq a, \]
a contradiction. Therefore $\text{int} (C_1) \cap C_2 = \emptyset$ and by Problem 12 on Page 66, there exists $(w^*, \beta) \in X' \times \mathbb{R}$ with
\[ (w^*, \beta) \neq (0, 0), \] (9.6.58)
and
\[ w^* (y) + \beta a \geq w^* (y_1) + \beta a_1, \] (9.6.59)
whenever $(y, a) \in C_1$ and $(y_1, a_1) \in C_2$.

Claim: $\beta > 0$.

Proof of claim: If $\beta < 0$ let
\[ a = \phi_1 (\overline{x}) - x^* (\overline{x} - x) - \phi_1 (x) + 1, \]
\[ a_1 = \phi_2 (x) - \phi_2 (\overline{x}), \text{ and } y = y_1 = \overline{x}. \]
Then from (9.6.58)
\[ \beta (\phi_1 (\overline{x}) - x^* (\overline{x} - x) - \phi_1 (x) + 1) \geq \beta (\phi_2 (x) - \phi_2 (\overline{x})). \]
Dividing by $\beta$ yields
\[ \phi_1 (\overline{x}) - x^* (\overline{x} - x) - \phi_1 (x) + 1 \leq \phi_2 (x) - \phi_2 (\overline{x}) \]
and so
\[ \phi_1 (\overline{x}) + \phi_2 (\overline{x}) - (\phi_1 (x) + \phi_2 (x)) + 1 \leq x^* (\overline{x} - x) \]
\[ \leq \phi_1 (\overline{x}) + \phi_2 (\overline{x}) - (\phi_1 (x) + \phi_2 (x)), \]
a contradiction. Therefore, $\beta \geq 0$.

Now suppose $\beta = 0$. Letting
\[ a = \phi_1 (\overline{x}) - x^* (\overline{x} - x) - \phi_1 (x) + 1, \]
\[ (\overline{x}, a) \in \text{int} (C_1), \]
and so there exists an open set $U$ containing 0 and $\eta > 0$ such that
\[ \overline{x} + U \times (a - \eta, a + \eta) \subseteq C_1. \]
Therefore, (9.6.58) applied to $(\overline{x} + z, a) \in C_1$ and $(\overline{x}, \phi_2 (x) - \phi_2 (\overline{x})) \in C_2$ for $z \in U$ yields
\[ w^* (\overline{x} + z) \geq w^* (\overline{x}) \]
for all $z \in U$. Hence $w^* (z) = 0$ on $U$ which implies $w^* = 0$, contradicting (9.6.58).
This proves the claim.
Now with the claim, it follows $\beta > 0$ and so, letting $z^* = w^*/\beta$, \[(9.6.60)\]
\[z^*(y) + a \geq z^*(y_1) + a_1\]
whenever $(y, a) \in C_1$ and $(y_1, a_1) \in C_2$. In particular,
\[(9.6.61)\]
\[(y, \phi_1(y) - \phi_1(x) - x^*(y - x)) \in C_1\]
because
\[
\phi_1(y) - \phi_1(x) - x^*(y - x) \leq \phi_1(y) - x^*(y - x) - \phi_1(x)
\]
and
\[(9.6.62)\]
\[(y_1, \phi_2(x) - \phi_2(y_1)) \in C_2.\]
by similar reasoning so letting $y = x$,
\[
\begin{align*}
z^*(x) + \left(\phi_1(x) - x^*(x - x) - \phi_1(x)\right) &\geq z^*(y_1) + \phi_2(x) - \phi_2(y) + \\
&\geq 0.
\end{align*}
\]
Therefore,
\[
z^*(y_1 - x) \leq \phi_2(y_1) - \phi_2(x)
\]
for all $y_1$ and so $z^* \in \partial \phi_2(x)$. Now let $y_1 = x$ in \[(9.6.61)\] and using \[(9.6.60)\] and \[(9.6.62)\], it follows
\[
z^*(y) + \phi_1(y) - x^*(y - x) - \phi_1(x) \geq z^*(x)
\]
\[
\phi_1(y) - \phi_1(x) \geq x^*(y - x) - z^*(y - x)
\]
and so $x^* - z^* \in \partial \phi_1(x)$ so $x^* = z^* + (x^* - z^*) \in \partial \phi_2(x) + \partial \phi_1(x). \blacksquare$

\textbf{Corollary 9.6.46} Let $\phi : X \to (-\infty, \infty]$ be convex, proper, and lower semicontinuous. Here $X$ is a Banach space. Then $\partial \phi$ is maximal monotone.

\textbf{Proof:} Let $\psi(x) = \frac{1}{2} \|x\|^2$. There exists $x^*$ and some number $b$ such that $\phi(x) \geq b + \langle x^*, x \rangle$. Therefore, $\psi + \phi$ is convex, lower semicontinuous, and bounded. It follows $\partial (\psi + \phi)$ is onto by Theorem \ref{thm:9.6.42}. However, $\psi$ is continuous everywhere, in particular at every point of the domain of $\phi$. Therefore, $\partial \psi + \partial \phi = \partial (\phi + \psi)$ and by Theorem \ref{thm:9.6.42} this shows that $F + \partial \phi$ is onto. \blacksquare

It seems to me that the above are the most important results about convex proper lower semicontinuous functions. However, there are many other very interesting properties known.

\textbf{Proposition 9.6.47} Let $\phi : X \to (-\infty, \infty]$ be convex proper and lower semicontinuous. Then $D(\partial \phi)$ is dense in $D(\phi)$ and so $D(\partial \phi) = \overline{D(\phi)}$.

\textbf{Proof:} Let $x_\lambda$ be the solution to $0 \in F(x_\lambda - x) + \lambda \partial \phi(x_\lambda)$. Here $x \in D(\phi)$. Say $u^*_\lambda \in \partial \phi(x_\lambda)$ such that the inclusion becomes an equality. Then
\[
0 = \langle F(x_\lambda - x) + \lambda u^*_\lambda, x_\lambda - x \rangle = \|x_\lambda - x\|^2 - \lambda \langle u^*_\lambda, x_\lambda - x \rangle
\]
\[
\geq \|x_\lambda - x\|^2 - \lambda (\phi(x) - \phi(x_\lambda))
\]
Hence, letting $z^*, b$ be such that $\phi(y) \geq b + \langle z^*, y - x \rangle$,
\[
\lambda(\phi(x) - [b + (z^*, x - x)]) \geq \lambda(\phi(x) - \phi(x)) \geq \|x - x\|^2
\]
\[
\lambda \psi(x) - \lambda b \geq \|x - x\|^2 - \lambda \|z\| \|x - x\|
\]
\[
\geq \|x - x\|^2 - \lambda \left(\frac{\|z^*\|^2}{2} + \frac{\|x - x\|^2}{2}\right)
\]
Thus
\[
\lambda \psi(x) - \lambda b + \lambda \|z^*\|^2 = \left(1 - \frac{\lambda}{2}\right) \|x - x\|^2
\]
It follows that $\lambda x \to x$. This shows that $D(\psi) \subseteq D(\partial \psi)$ and so $D(\psi) \subseteq D(\partial \psi)$.

There is a really amazing theorem, Moreau’s theorem. It is in [1], [4] and [6]. It involves approximating a convex function with one which is differentiable, at least in the case where you have a Hilbert space. In the general case considered in this chapter, the function is continuous.

**Theorem 9.6.48** Let $\phi$ be a convex lower semicontinuous proper function defined on $X$. Define $A \equiv \partial \phi, A_\lambda = (\partial \phi)_\lambda$
\[
\phi_\lambda(x) \equiv \min_{y \in X} \left(\frac{1}{2\lambda} \|x - y\|^2 + \phi(y)\right)
\]

Then the function is well defined, convex, Gateaux differentiable,
\[
D_z \phi_\lambda(x) \equiv \lim_{t \downarrow 0} \frac{\phi_\lambda(x + tz) - \phi_\lambda(x)}{t} = \langle A_\lambda x, z \rangle
\]
so the Gateaux derivative is just $A_\lambda x$ and for all $x \in X,$
\[
\lim_{\lambda \to 0} \phi_\lambda(x) = \phi(x),
\]
In addition,
\[
\phi_\lambda(x) = \frac{1}{2\lambda} \|x - J_\lambda x\|^2 + \phi(J_\lambda(x)) \quad (9.6.63)
\]
where $J_\lambda x$ is as before, the solution to
\[
0 \in F(J_\lambda x - x) + \lambda \partial \phi(J_\lambda x)
\]

**Proof:** First of all, why does the minimum take place? By the convexity, closed epigraph, and assumption that $\phi$ is proper, separation theorems apply and one can say that there exists $z^*$ such that for all $y \in H$,
\[
\frac{1}{2\lambda} \|x - y\|^2 + \phi(y) \geq \frac{1}{2\lambda} \|x - y\|^2 + (z^*, y) + c \quad (9.6.64)
\]
It follows easily that a minimizing sequence is bounded and so from lower semicontinuity which implies weak lower semicontinuity due to convexity, there exists $y_x$ such that

$$\min_{y \in H} \left( \frac{1}{2\lambda} \|x - y\|^2 + \phi(y) \right) = \left( \frac{1}{2\lambda} \|x - y_x\|^2 + \phi(y_x) \right)$$

Why is $\phi_\lambda$ convex? For $\theta \in [0, 1]$,

$$\phi_\lambda (\theta x + (1 - \theta) z) \equiv \frac{1}{2\lambda} \|\theta x + (1 - \theta) z - y_{(\theta x + (1 - \theta) z)}\|^2 + \phi(y_{(\theta x + (1 - \theta) z)})$$

$$\leq \frac{1}{2\lambda} \|\theta x + (1 - \theta) z - (\theta y_x + (1 - \theta) y_z)\|^2 + \phi(\theta y_x + (1 - \theta) y_z)$$

$$\leq \frac{\theta}{2\lambda} \|x - y_x\|^2 + \frac{1 - \theta}{2\lambda} \|y_z - y_x\|^2 + \theta \phi(y_x) + (1 - \theta) \phi(y_z)$$

$$= \theta \phi_\lambda (x) + (1 - \theta) \phi_\lambda (z)$$

So is there a formula for $y_x$? Since it involves minimization of the functional, it follows that

$$0 \in -\frac{1}{\lambda} F(x - y_x) + \partial \phi(y_x) = \frac{1}{\lambda} F(y_x - x) + \partial \phi(y_x)$$

Recall that if $\psi(x) = \frac{1}{2} \|x\|^2$, then $\partial \psi(x) = F(x)$. Thus

$$y_x = J_\lambda x$$

because this was how $J_\lambda x$ was defined. Therefore,

$$\phi_\lambda (x) = \frac{1}{2\lambda} \|x - J_\lambda x\|^2 + \phi(J_\lambda (x)) = \frac{\lambda}{2} \|A_\lambda x\|^2 + \phi(J_\lambda x), \ A = \partial \phi$$

It follows from this equation that

$$\phi(J_\lambda x) \leq \phi_\lambda (x) \leq \phi (x), \quad (9.6.65)$$

the second inequality following from taking $y = x$ in the definition of $\phi_\lambda$.

Next consider the claim about $\phi_\lambda (x) \uparrow \phi (x)$. First suppose that $x \in D (\phi)$.

Then from Proposition 9.6.47, $x \in D (\partial \phi)$ and so from the material on approximations, Theorem 9.6.30, it follows that $J_\lambda x \to x$. Hence from 9.6.65 and lower semicontinuity of $\phi$,

$$\phi(x) \leq \lim \inf_{\lambda \to 0} \phi(J_\lambda x) \leq \lim \inf_{\lambda \to 0} \phi_\lambda (x) \leq \lim \sup_{\lambda \to 0} \phi_\lambda (x) \leq \phi (x)$$

showing that in this case, $\lim_{\lambda \to 0} \phi_\lambda (x) = \phi (x)$. Next suppose $x \notin D (\phi)$ so that $\phi (x) = \infty$. Why does $\phi_\lambda (x) \to \infty$? Suppose not. Then from the description of $\phi_\lambda$
9.6. MAXIMAL MONOTONE OPERATORS

Given above and using the fact that the epigraph is closed and convex, there would exist a subsequence, still denoted as \( \lambda \) such that

\[
C \geq \phi_\lambda (x) = \frac{1}{2\lambda} \|x - J_\lambda x\|^2 + \phi (J_\lambda (x)) \geq \frac{1}{2\lambda} \|x - J_\lambda x\|^2 + \langle z^*, x - J_\lambda x \rangle + b
\]

Then multiplying by \( \lambda \), it follows that for a suitable constant \( M \),

\[
\|x - J_\lambda x\|^2 \leq M\lambda + \lambda M \|x - J_\lambda x\|
\]

And so a use of the quadratic formula implies

\[
\|x - J_\lambda x\| \leq \frac{M}{2} \left( 1 + \sqrt{5} \right) \lambda
\]

Hence \( J_\lambda x \to x \) and so from \( \text{Lem.} \) it follows from lower semicontinuity again that

\[
\infty = \phi (x) = \lim_{\lambda \to 0} \inf \phi (J_\lambda x) \leq \lim_{\lambda \to 0} \inf \phi_\lambda (x) \leq \lim_{\lambda \to 0} \sup \phi_\lambda (x) \leq \phi (x)
\]

And so again, \( \lim_{\lambda \to 0} \phi_\lambda (x) = \infty \). Also note that if \( \lambda > \mu \), then

\[
\min_{y \in \lambda} \left( \frac{1}{2\lambda} \|x - y\|^2 + \phi (y) \right) \leq \min_{y \in \lambda} \left( \frac{1}{2\mu} \|x - y\|^2 + \phi (y) \right)
\]

Because for a given \( y \), \( \frac{1}{2\lambda} \|x - y\|^2 + \phi (y) \leq \frac{1}{2\mu} \|x - y\|^2 + \phi (y) \). Thus \( \phi_\lambda (x) \uparrow \phi (x) \).

Next consider the claim about the Gateaux differentiability. Using the description \( \text{Lem.} \)

\[
\phi_\lambda (y) - \phi_\lambda (x) = \frac{1}{2\lambda} \|y - J_\lambda y\|^2 + \phi (J_\lambda (y)) - \left( \frac{1}{2\lambda} \|x - J_\lambda x\|^2 + \phi (J_\lambda (x)) \right)
\]

(9.6.66)

Using the fact that if \( \psi (x) = \|x\|^2 \), then \( \partial \psi (x) = Fx \), and that \( A_\lambda x \in \partial \phi (J_\lambda x) \),

\[
\begin{align*}
& \geq \lambda^{-1} \langle F (x - J_\lambda x), (y - J_\lambda y) - (x - J_\lambda x) \rangle + \langle A_\lambda x, J_\lambda (y) - J_\lambda (x) \rangle \\
= & \langle A_\lambda x, (y - J_\lambda y) - (x - J_\lambda x) \rangle + \langle A_\lambda x, J_\lambda (y) - J_\lambda (x) \rangle = \langle A_\lambda x, y - x \rangle
\end{align*}
\]

Hence

\[
(\phi_\lambda (y) - \phi_\lambda (x)) - \langle A_\lambda x, y - x \rangle \geq 0
\]

Also from \( \text{Lem.} \)

\[
\frac{1}{2\lambda} \|y - J_\lambda y\|^2 - \frac{1}{2\lambda} \|x - J_\lambda x\|^2 = - \left( \frac{1}{2\lambda} \|x - J_\lambda x\|^2 - \frac{1}{2\lambda} \|y - J_\lambda y\|^2 \right)
\]

\[
\leq - \frac{1}{\lambda} \langle F (y - J_\lambda y), (x - J_\lambda x) - (y - J_\lambda y) \rangle = \langle A_\lambda y, (y - J_\lambda y) - (x - J_\lambda x) \rangle
\]

Similarly, from \( \text{Lem.} \),

\[
\phi (J_\lambda (y)) - \phi (J_\lambda (x)) = - (\phi (J_\lambda (x)) - \phi (J_\lambda (y)))
\]
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

\[ \leq - \langle A_\lambda (y) J_\lambda (x) - J_\lambda (y) \rangle = \langle A_\lambda (y), J_\lambda (y) - J_\lambda (x) \rangle \]

It follows that

\[ \langle A_\lambda (y), J_\lambda (y) - J_\lambda (x) \rangle \]
\[ \geq (\phi_\lambda (y) - \phi_\lambda (x)) \geq \langle A_\lambda x, y - x \rangle \]

and so

\[ \langle A_\lambda (y), y - x \rangle \]
\[ \geq (\phi_\lambda (y) - \phi_\lambda (x)) \geq \langle A_\lambda x, y - x \rangle \]

Therefore,

\[ \langle A_\lambda (y) - A_\lambda (x), y - x \rangle \]
\[ \geq (\phi_\lambda (y) - \phi_\lambda (x)) - \langle A_\lambda x, y - x \rangle \geq 0 \]

Next let \( y = x + tz \) for \( t > 0 \). Then

\[ t \langle A_\lambda (x + tz) - A_\lambda (x), z \rangle \geq (\phi_\lambda (x + tz) - \phi_\lambda (x)) - t \langle A_\lambda x, z \rangle \geq 0 \]

Using the demicontinuity of \( A_\lambda \), you can divide by \( t \) and pass to a limit to obtain

\[ \lim_{t \downarrow 0} \frac{\phi_\lambda (x + tz) - \phi_\lambda (x)}{t} = \langle A_\lambda x, z \rangle \]

A much better theorem is available in case \( X = X' = H \) a Hilbert space. In this case \( \phi_\lambda \) is also Frechet differentiable. See Theorem 10.4.22 which is presented later. Everything is much nicer in the Hilbert space setting because \( F \) is just replaced with the identity and the approximations are defined more easily.

\[ 0 \in J_\lambda x - x + \lambda AJ_\lambda x, \]
\[ x \in J_\lambda x + \lambda AJ_\lambda x = (I + \lambda A) J_\lambda x \]
\[ J_\lambda x = (I + \lambda A)^{-1} x \]

Then one can show that \( J_\lambda \) is Lipschitz continuous and many other nice things happen.

Next is an interesting result about when the sum of a maximal monotone operator and a subgradient is also maximal monotone. A version of this is well known in the case of a single Hilbert space. In the case of a single Hilbert space, this result can be used to produce very regular solutions to evolution equations for functions which have values in the Hilbert space. You would get this by letting \( X = X' \) equal to a Hilbert space and your maximal monotone operator \( A \) would be defined on \( L^2 (0,T; H) = X \) a space of Hilbert space valued functions which are square integrable. Then you could take \( Lu = u' \) with domain equal to those functions in \( X \) which are equal to 0 at the left end of the interval for example. This is done more generally later. In this case the duality map is just the identity. The next theorem includes the case of two different spaces. I am not sure whether this is a useful result at this time, in terms of evolution equations. However, it is good to have conditions which show that the sum of two maximal monotone operators is maximal monotone.
9.6. MAXIMAL MONOTONE OPERATORS

Theorem 9.6.49 Let $X$ be a reflexive Banach space with strictly convex norm and let $\Phi$ be non negative, convex, proper, and lower semicontinuous. Suppose also that $A : D(A) \to P(X')$ is a maximal monotone operator and there exists

$$\xi \in D(A) \cap D(\Phi).$$

(9.6.67)

Suppose also that

$$\Phi(J_{\lambda}x) \leq \Phi(x) + C\lambda$$

(9.6.68)

Then $A + \partial \Phi$ is maximal monotone.

Proof: Recall that

$$A_{\lambda}x = -\lambda^{-1}F(J_{\lambda}x - x),$$

where $0 \in F(J_{\lambda}x - x) + \lambda \partial A(J_{\lambda}x)$

Let $y^* \in X'$. From Theorem (9.6.67) there exists $x_{\lambda} \in H$ such that

$$y^* \in Fx_{\lambda} + A_{\lambda}x_{\lambda} + \partial \Phi(x_{\lambda}).$$

It is desired to show that $A_{\lambda}x_{\lambda}$ is bounded. From the above,

$$y^* - Fx_{\lambda} - A_{\lambda}x_{\lambda} \in \partial \Phi(x_{\lambda})$$

(9.6.69)

and so

$$\langle y^* - Fx_{\lambda} - A_{\lambda}x_{\lambda}, J_{\lambda}x_{\lambda} - x_{\lambda} \rangle \leq \Phi(J_{\lambda}x_{\lambda}) - \Phi(x_{\lambda}) \leq C\lambda$$

(9.6.70)

which implies

$$\langle y^* - Fx_{\lambda} - A_{\lambda}x_{\lambda}, (-\lambda)F^{-1}(A_{\lambda}x) \rangle \leq \Phi(J_{\lambda}x_{\lambda}) - \Phi(x_{\lambda}) \leq C\lambda$$

and so

$$\langle y^* - Fx_{\lambda} - A_{\lambda}x_{\lambda}, -F^{-1}(A_{\lambda}x) \rangle \leq C$$

Hence

$$\langle y^* - Fx_{\lambda}, -F^{-1}(A_{\lambda}x_{\lambda}) \rangle + \|A_{\lambda}x_{\lambda}\|^2 \leq C$$

(9.6.71)

I claim $\{\|x_{\lambda}\|\}$ are bounded independent of $\lambda$.

By (9.6.68) and monotonicity of $A_{\lambda}$,

$$\Phi(\xi) - \Phi(x_{\lambda}) \geq \langle y^* - F_{\lambda}x_{\lambda} - A_{\lambda}x_{\lambda}, \xi - x_{\lambda} \rangle$$

$$\geq \langle y^* - F_{\lambda}x_{\lambda}, \xi - x_{\lambda} \rangle - \langle A_{\lambda}x_{\lambda}, \xi - x_{\lambda} \rangle$$

$$\geq \langle y^* - F_{\lambda}x_{\lambda}, \xi - x_{\lambda} \rangle - \langle A_{\lambda}x_{\lambda}, \xi - x_{\lambda} \rangle$$

$$= \langle y^*, \xi \rangle - \langle y^*, x_{\lambda} \rangle - \langle F_{\lambda}x_{\lambda}, \xi \rangle + \|x_{\lambda}\|^2 - \|\xi - x_{\lambda}\|^2 - \|A_{\lambda}\| - \|A_{\lambda}\|$$

$$\geq -\|y^*\| \|\xi\| - \|y^*\| \|x_{\lambda}\| - \|x_{\lambda}\|^2 - \|\xi\|^2 - \|A_{\lambda}\| - \|A_{\lambda}\|$$

Therefore, there exist constants, $C_1$ and $C_2$, depending on $\xi$ and $y^*$ but not on $\lambda$ such that

$$\Phi(\xi) \geq \Phi(x_{\lambda}) + \|x_{\lambda}\|^2 - C_1 \|x_{\lambda}\|^2 - C_2.$$ 

Since $\Phi \geq 0$, the above shows that $\|x_{\lambda}\|$ is indeed bounded. Now from (9.6.67) it follows that $\{A_{\lambda}x_{\lambda}\}$ is bounded for small positive $\lambda$. By Theorem (9.6.68) there exists a solution $x$ to

$$y^* \in F_{\lambda}x + A_{\lambda}x + \partial \Phi(x)$$

and since $y^*$ is arbitrary, this shows that $A + \partial \Phi$ is maximal monotone.
9.7 Perturbation Theorems

In this section we give surjectivity of the sum of a pseudomonotone set valued map with a linear maximal monotone map and also with another maximal monotone operator added in. It generalizes the surjectivity results given earlier because one could have 0 for the maximal monotone linear operator. The theorems developed here lead to nice results on evolution equations because the linear maximal monotone operator can be something like a time derivative and $X$ can be some sort of an $L^p$ space for functions having values in a suitable Banach space. This is presented later in the material on Bochner integrals.

The notation $\langle z^*, u \rangle_{V', V}$ will mean $z^* (u)$ in this section. We will not worry about the order either. Thus

$$\langle u, z^* \rangle \equiv z^* (u) \equiv \langle z^*, u \rangle$$

This is just convenient in writing things down. Also, it is assumed that all Banach spaces are real to simplify the presentation. It is also usually assumed that the Banach spaces are reflexive. Thus we can regard

$$(V \times V')' = V' \times V$$

and $\langle \langle y^*, x \rangle, (u, v^*) \rangle \equiv \langle y^*, u \rangle + \langle x, v^* \rangle$. It is known [4] that for a reflexive Banach space, there is always an equivalent strictly convex norm. It is therefore, assumed that the norm for the reflexive Banach space is strictly convex.

**Definition 9.7.1** Let $L : D (L) \subseteq V \to V'$ be a linear map where we always assume $D (L)$ is dense in $V$. Then

$$D (L^*) \equiv \{ u \in V : |(Lv, u)| \leq C \| v \| \text{ for all } v \in D (L) \}$$

For such $u$, it follows that on a dense subset of $V$, namely $D (L), v \to \langle Lz, u \rangle$ is a continuous linear map. Hence there exists a unique element of $V'$, denoted as $L^* u$ such that for all $v \in D (L),$

$$\langle Lv, u \rangle_{V', V} = \langle L^* u, v \rangle_{V', V}$$

Thus

$$L : D (L) \subseteq V \to V'$$

$$L^* : D (L^*) \subseteq V \to V'$$

There is an interesting description of $L^*$ in terms of $L$ which will be quite useful.

**Proposition 9.7.2** Let $\tau : V \times V' \to V' \times V$ be given by $\tau (a, b) \equiv (-b, a)$. Also for $S \subseteq X$ a reflexive Banach space,

$$S^\perp \equiv \{ z^* \in X' : \langle z^*, s \rangle = 0 \text{ for all } s \in S \}$$

Also denote by $\mathcal{G} (L) \equiv \{(x, Lx) : x \in D (L)\}$. Then

$$\mathcal{G} (L^*) = (\tau \mathcal{G} (L))^\perp$$
Proof: Let \((x, L^*x) \in \mathcal{G}(L^*)\). This means that
\[ |\langle Ly, x \rangle| \leq C \|y\| \quad \text{for all } y \in D(L) \]
and \(\langle Ly, x \rangle = \langle L^*x, y \rangle\) for all \(y \in D(L)\). Let \((y, Ly) \in \mathcal{G}(L)\). Then \(\tau(y, Ly) = (-Ly, y)\). Then
\[ \langle (x, L^*x), (-Ly, y) \rangle = \langle x, -Ly \rangle + \langle L^*x, y \rangle = -\langle x, Ly \rangle + \langle x, L^*y \rangle = 0 \]
Thus \(\mathcal{G}(L^*) \subseteq (\tau \mathcal{G}(L))^\perp\). Next suppose \((x, y^*) \in (\tau \mathcal{G}(L))^\perp\). This means that if \((u, Lu) \in \mathcal{G}(L)\), then
\[ \langle (x, y^*), (-Lu, u) \rangle \equiv \langle x, -Lu \rangle + \langle y^*, u \rangle = 0 \]
and so for all \(u \in D(L)\),
\[ \langle y^*, u \rangle = \langle x, Lu \rangle \]
and so \(x \in D(L^*)\). Hence for all \(u \in D(L)\),
\[ \langle y^*, u \rangle = \langle x, Lu \rangle = \langle L^*x, u \rangle \]
Then, since \(D(L)\) is dense, it follows that \(y^* = L^*x\) and so \((x, y) \in \mathcal{G}(L^*)\). Thus these are the same. \(\blacksquare\)

Theorem 9.4.8 is a very nice surjectivity result for set valued pseudomonotone operators. We recall what it said here. Recall the meaning of coercive.

\[ \lim_{\|v\| \to \infty} \inf \left\{ \frac{\langle z^*, v \rangle}{\|v\|} : z^* \in Tv \right\} \]

In this section, we use the convenient notation \(\langle z^*, x \rangle_{\mathcal{V}^*, \mathcal{V}} \equiv z^*(x)\).

**Theorem 9.7.3** Let \(V\) be a reflexive Banach space and let \(T : V \to \mathcal{P}(V')\) be pseudomonotone, bounded and coercive. Then \(T\) is onto. More generally, this continues to hold if \(T\) is modified bounded pseudomonotone.

Recall the definition of pseudomonotone.

**Definition 9.7.4** For \(X\) a reflexive Banach space, we say \(A : X \to \mathcal{P}(X')\) is pseudomonotone if the following hold.

1. The set \(Au\) is nonempty, closed and convex for all \(u \in X\).

2. If \(F\) is a finite dimensional subspace of \(X\), \(u \in F\), and if \(U\) is a weakly open set in \(V'\) such that \(Au \subseteq U\), then there exists a \(\delta > 0\) such that if \(v \in B_\delta(u) \cap F\) then \(Av \subseteq U\). (Weakly upper semicontinuous on finite dimensional subspaces.)

3. If \(u_i \to u\) weakly in \(X\) and \(u_i^* \in Au_i\) is such that
\[ \limsup_{i \to \infty} \langle u_i^*, u_i - u \rangle \leq 0, \quad (9.7.72) \]
then, for each \(v \in X\), there exists \(u^* (v) \in Au\) such that
\[ \liminf_{i \to \infty} \langle u_i^*, u_i - v \rangle \geq \langle u^*(v), u - v \rangle. \quad (9.7.73) \]
Also recall the definition of modified bounded pseudomonotone. It is just the above except that the limit condition is replaced with the following condition: If $u_i \to u$ weakly in $X$ then there exists a subsequence, still denoted as $\{u_i\}$ such that if
\[
\limsup_{i \to \infty} \langle u_i^*, u_i - u \rangle \leq 0,
\]
then, for each $v \in X$, there exists $u^*(v) \in Au$ such that
\[
\liminf_{i \to \infty} \langle u_i^*, u_i - v \rangle \geq \langle u^*(v), u - v \rangle.
\]

Also recall that this more general limit condition along with the assumption and the assumption that $A$ is bounded is sufficient to obtain condition. This was Lemma 9.4.5 proved earlier and stated here for convenience.

**Lemma 9.7.5** Let $A : X \to \mathcal{P}(X')$ satisfy conditions 1 and 2 above and suppose $A$ is bounded. Then if whenever $x_n \to x$ in $X$, there is a subsequence $\{x_{nk}\}$ such that for this subsequence, the pseudomonotone limit condition
\[
\lim_{n \to \infty} \langle z_{nk}, x_{nk} - x \rangle \leq 0 \implies \liminf_{n \to \infty} \langle z_{nk}, x_{nk} - y \rangle \geq \langle z(y), x - y \rangle
\]
for $z_{nk} \in Ax_{nk}$ and $z(y)$ is some element of $Ax$. Then if $U$ is a weakly open set containing $Ax$, then $Ax_n \subseteq U$ for all $n$ large enough.

**Definition 9.7.6** Now let $L : D(L) \subseteq V \to V'$ such that $L$ is linear, monotone, $D(L)$ is dense in $V$, $L$ is closed, and $L^*$ is monotone. Let $A : V \to \mathcal{P}(V')$ be a bounded operator. Then $A$ is called $L$ pseudomonotone if $Av$ is closed and convex in $V'$ and for any sequence $\{u_n\} \subseteq D(L)$ such that $u_n \to u$ weakly in $V$ and $Lu_n \to Lu$ weakly in $V'$, and for $z_n^* \in Au_n$,
\[
\limsup_{n \to \infty} \langle z_n^*, u_n - u \rangle \leq 0
\]
then for every $v \in V$, there exists $z^*(v) \in Au$ such that
\[
\liminf_{n \to \infty} \langle z_n^*, u_n - v \rangle \geq \langle z^*(v), u - v \rangle
\]
It is called $L$ modified bounded pseudomonotone if the above limit condition holds for some subsequence whenever $u_n \to u$ weakly and $Lu_n \to Lu$ weakly.

**Lemma 9.7.7** Suppose $X$ is the Banach space
\[
X = D(L), \quad \parallel u \parallel_X \equiv \parallel u \parallel_V + \parallel Lu \parallel_V,
\]
where $L$ is as described in the above definition. Also assume that $A$ is bounded. Then if $A$ is $L$ pseudomonotone, it follows that $A$ is pseudomonotone as a map from $X$ to $\mathcal{P}(X')$. If $A$ is $L$ modified bounded pseudomonotone, then $A$ is modified bounded pseudomonotone as a map from $X$ to $\mathcal{P}(X')$. 
9.7. PERTURBATION THEOREMS

Proof: Is $A$ bounded? Of course because the norm of $X$ is stronger than the norm on $V$. Is $Au$ convex and closed? This also follows because $X \subseteq V$. It is clear that $Au$ is convex. If $\{z_n\} \subseteq Au$ and $z_n \to z$ in $X'$, then does it follow that $z \in Au$? Since $A$ is bounded, there is a further subsequence which converges weakly to $w$ in $V'$. However, $Au$ is convex and closed so it is weakly closed. Hence $w \in Au$ and also $w = z$. It only remains to verify the pseudomonotone limit condition. Suppose then that $u_n \to u$ weakly in $X$ and for $z_n^* \in Au_n$,

$$\limsup_{n \to \infty} \langle z_n^*, u_n - u \rangle \leq 0$$

Then it follows that $Lu_n \to Lu$ weakly in $V'$ and $u_n \to u$ weakly in $V$ so $u \in X$. Hence the assumption that $A$ is $L$ pseudomonotone implies that for every $v \in V$, and for every $v \in X$, there exists $z^*(v) \in Au \subseteq V' \subseteq X'$ such that

$$\liminf_{n \to \infty} \langle z_n^*, u_n - v \rangle \geq \langle z^*(v), u - v \rangle$$

The last claim goes the same way. You just have to take a subsequence. ■

Then we have the following major surjectivity result. In this theorem, we will assume for simplicity that all spaces are real spaces. Versions of this appear to be due to Brezis [3] and Lions [10]. Of course the theorem holds for complex spaces as well. You just need to use $\text{Re} \langle \, \rangle$ instead of $\langle \, \rangle$.

Theorem 9.7.8 Let $L : D(L) \subseteq V \to V'$ where $D(L)$ is dense, $L$ is monotone, $L$ is closed, and $L^*$ is monotone, $L$ a linear map. Let $A : V \to \mathcal{P}(V')$ be $L$ pseudomonotone, bounded, coercive. Then $L + A$ is onto. Here $V$ is a reflexive Banach space such that the norms for $V$ and $V'$ are strictly convex. In case that $A$ is strictly monotone ($\langle Au - Av, u - v \rangle > 0$ implies $u \neq v$) the solution $u$ to $f \in Lu + Au$ is unique. If, in addition to this, $\langle Au - Av, u - v \rangle \geq r \langle \|u - v\|_V \rangle$ where $U$ is some Banach space containing $V$, and $r$ is a positive strictly increasing function for which $\lim_{t \to 0^+} r(t) = 0$, then the map $f \to u$ where $f \in Lu + Au$ is continuous as a map from $V'$ to $U$. The conclusion holds if $A$ is only $L$ modified bounded pseudomonotone.

Proof: Let $F$ be the duality map for $p = 2$. Consider the Banach space $X$ given by

$$X = D(L), \|u\|_X \equiv \|u\|_V + \|Lu\|_{V'}$$

This is isometric with the graph of $L$ with the graph norm and so $X$ is reflexive. Now define a set valued map $G_{\varepsilon}$ on $X$ as follows. $z^* \in G_{\varepsilon}(u)$ means there exists $w^* \in Au$ such that

$$\langle z^*, v \rangle_{X', X} = \varepsilon \langle Lv, F^{-1}(Lu) \rangle_{V', V} + \langle Lu, v \rangle_{V', V} + \langle w^*, v \rangle_{V', V}$$

It follows from Lemma 15.25 that $G_{\varepsilon}$ is the sum of a set valued $L$ modified bounded pseudomonotone operator with an operator which is demicontinuous, bounded, and
monotone, hence pseudomonotone. Thus by Lemma 9.4.7 it is $L$ modified bounded pseudomonotone. Is it coercive?

\[
\lim_{\|u\|_X \to \infty} \inf \left\{ \frac{\langle z^*, u \rangle + \varepsilon \langle F^{-1}(Lu), F^{-1}(Lu) \rangle_{V', V} + \langle Lu, u \rangle_{V', V}}{\|u\|_X} : z^* \in Au \right\} = \infty
\]

It equals

\[
\lim_{\|u\|_X \to \infty} \inf \left\{ \frac{\langle z^*, u \rangle + \varepsilon \|F^{-1}(Lu)\|_{V'}^2}{\|u\|_X} : z^* \in Au \right\}
\]

and this is

\[
\geq \lim_{\|u\|_X \to \infty} \inf \left\{ \frac{\langle z^*, u \rangle + \varepsilon \|Lu\|_{V'}^2}{\|u\|_V + \|Lu\|_{V'}} : z^* \in Au \right\}
\]

because $L$ is monotone. Now let $M$ be an arbitrary positive number. By assumption, there exists $R$ such that if $\|u\|_V > R$, then

\[
\inf \left\{ \frac{\langle z^*, u \rangle}{\|u\|_V} : z^* \in Au \right\} > M
\]

and so for every $z^* \in Au$,

\[
\frac{\langle z^*, u \rangle}{\|u\|_V} > M, \quad \langle z^*, u \rangle > M \|u\|_V
\]

Thus if $\|u\|_V > R$,

\[
\inf \left\{ \frac{\langle z^*, u \rangle + \varepsilon \|Lu\|_{V'}^2}{\|u\|_V + \|Lu\|_{V'}} : z^* \in Au \right\} \geq \frac{M \|u\|_V + \varepsilon \|Lu\|_{V'}^2}{\|u\|_V + \|Lu\|_{V'}}
\]

I claim that if $\|u\|_X$ is large enough, the above is larger than $M/2$. If not, then there exists $\{u_n\}$ such that $\|u_n\|_X \to \infty$ but the right side is less than $M/2$. First say $\|Lu_n\|$ is bounded. then there is an obvious contradiction since the right hand side then converges to $M$. Thus it can be assumed that $\|Lu_n\|_{V'} \to \infty$. Hence, for all $n$ large enough, $\varepsilon \|Lu\|_{V'}^2 > M \|Lu_n\|_{V'}$. However, this implies the right side is larger than

\[
\frac{M \|u_n\|_V + M \|Lu_n\|_{V'}}{\|u_n\|_V + \|Lu_n\|_{V'}} = M > M/2
\]

This is a contradiction. Hence the right side is larger than $M/2$ for all $n$ large enough. It follows since $M$ is arbitrary, that

\[
\lim_{\|u\|_X \to \infty} \inf \left\{ \frac{\langle z^*, u \rangle + \varepsilon \|Lu\|_{V'}^2}{\|u\|_V + \|Lu\|_{V'}} : z^* \in Au \right\} = \infty
\]
It follows from Theorem \ref{thm:9.4.8} that if $f \in V'$, there exists $u_\varepsilon$ such that for all $v \in D(L) = X$,

$$
\varepsilon \langle Lv, F^{-1}(Lu_\varepsilon) \rangle_{V',V} + \langle Lu_\varepsilon, v \rangle_{V',V} + \langle w^*_\varepsilon, v \rangle_{V',V} = \langle f, v \rangle, \quad w^*_\varepsilon \in Au_\varepsilon \quad (9.7.76)
$$

First we get an estimate.

$$
\varepsilon \langle Lu_\varepsilon, F^{-1}(Lu_\varepsilon) \rangle_{V',V} + \langle Lu_\varepsilon, u_\varepsilon \rangle_{V',V} + \langle w^*_\varepsilon, u_\varepsilon \rangle_{V',V} = \langle f, u_\varepsilon \rangle
$$

Hence it follows from the coercivity of $A$ that $\|u_\varepsilon\|_V$ is bounded independent of $\varepsilon$. Thus the $w^*_\varepsilon$ are also bounded in $V'$ because it is assumed that $A$ is bounded. Now from the equation solved \ref{eq:9.7.76}, it follows that $F^{-1}(Lu_\varepsilon) \in D(L^*)$. Thus the first term is just $\varepsilon \langle L^*(F^{-1}(Lu_\varepsilon)), v \rangle_{V',V}$. It follows, since $D(L) = X$ is dense in $V$ that

$$
\varepsilon L^*(F^{-1}(Lu_\varepsilon)) + Lu_\varepsilon + w^*_\varepsilon = f \quad (9.7.77)
$$

Then act on $F^{-1}(Lu_\varepsilon)$ on both sides. From monotonicity of $L^*$, this yields $\|Lu_\varepsilon\|_V$, is bounded independent of $\varepsilon > 0$. Thus there is a subsequence still denoted with a subscript of $\varepsilon$ such that

$$
u_\varepsilon \rightharpoonup u \text{ in } V
$$

$$
Lu_\varepsilon \rightharpoonup Lu \text{ in } V'
$$

This because of the fact that the graph of $L$ is closed, hence weakly closed. Thus $u \in X$. Also

$$
w^*_\varepsilon \rightharpoonup w^* \text{ in } V'.
$$

It follows that we can pass to a limit in \ref{eq:9.7.77} and obtain

$$
Lu + w^* = f \quad (9.7.78)
$$

Now by assumption on $A$, it is $L$ modified bounded pseudomonotone and so there is a subsequence, still denoted as $u_\varepsilon$ such that the pseudomonotone limit condition holds. This will be what is referred to in what follows. Then

$$
\langle \varepsilon L^*(F^{-1}(Lu_\varepsilon)), u_\varepsilon - u \rangle + \langle Lu_\varepsilon, u_\varepsilon - u \rangle + \langle w^*_\varepsilon, u_\varepsilon - u \rangle = \langle f, u_\varepsilon - u \rangle
$$

and so,

$$
\varepsilon \langle F^{-1}(Lu_\varepsilon), Lu_\varepsilon - Lu \rangle + \langle Lu_\varepsilon, u_\varepsilon - u \rangle + \langle w^*_\varepsilon, u_\varepsilon - u \rangle = \langle f, u_\varepsilon - u \rangle
$$

using the monotonicity of $L$,

$$
\varepsilon \langle Lu_\varepsilon - Lu, F^{-1}(Lu_\varepsilon) - F^{-1}(Lu) \rangle + \varepsilon \langle Lu_\varepsilon - Lu, F^{-1}(Lu) \rangle + \langle Lu, u_\varepsilon - u \rangle + \langle w^*_\varepsilon, u_\varepsilon - u \rangle \leq \langle f, u_\varepsilon - u \rangle
$$
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

Now using monotonicity of $F^{-1}$,

$$\varepsilon \langle Lu_{\varepsilon} - Lu, F^{-1}(Lu) \rangle + \langle Lu, u_{\varepsilon} - u \rangle + \langle w_{\varepsilon}^*, u_{\varepsilon} - u \rangle \leq \langle f, u_{\varepsilon} - u \rangle$$

and so, passing to a limit as $\varepsilon \to 0$,

$$\limsup_{\varepsilon \to 0} \langle w_{\varepsilon}^*, u_{\varepsilon} - u \rangle \leq 0$$

It follows that for all $v \in X = D(L)$ there exists $w^*(v) \in Au$

$$\liminf_{\varepsilon \to 0} \langle w_{\varepsilon}^*, u_{\varepsilon} - v \rangle \geq \langle w^*(v), u - v \rangle$$

But the left side equals

$$\liminf_{\varepsilon \to 0} [\langle w_{\varepsilon}^*, u_{\varepsilon} - u \rangle + \langle w_{\varepsilon}^*, u - v \rangle] \leq \limsup_{\varepsilon \to 0} \langle w_{\varepsilon}^*, u_{\varepsilon} - u \rangle + \langle w^*, u - v \rangle \leq \langle w^*, u - v \rangle$$

and so

$$\langle w^*, u - v \rangle \geq \langle w^*(v), u - v \rangle$$

for all $v$.

Is $w^* \in Au$? Suppose not. Then $Au$ is a closed convex set and $w^*$ is not in it.

Hence, since $V$ is reflexive, there exists $z \in V$ such that whenever $y^* \in Au$, $\langle w^*, z \rangle < \langle y^*, z \rangle$. Now simply choose $v$ such that $u - v = z$ and it follows that

$$\langle w^*(v), u - v \rangle > \langle w^*, u - v \rangle \geq \langle w^*(v), u - v \rangle$$

which is clearly a contradiction. Hence $w^* \in Au$. Thus from Lemma 9.7.8, this has shown that $L + A$ is onto.

Consider the claim about uniqueness and continuous dependence. Say you have $f_i \in Lu_i + Au_i, i = 1, 2$. Let $z_i^* \in Au_i$ be such that equality holds in the two inclusions. Then

$$f_1 - f_2 = z_1^* - z_2^* + Lu_1 - Lu_2$$

It follows that

$$\langle f_1 - f_2, u_1 - u_2 \rangle = \langle z_1^* - z_2^* + Lu_1 - Lu_2, u_1 - u_2 \rangle \geq r(\|u_1 - u_2\|)$$

Thus if $f_1 = f_2$, then $u_1 = u_2$. If $f_n \to f$ in $V'$, then $r(\|u - u_n\|) \to 0$ where $u_n$ goes with $f_n$ and $u$ with $f$ as just described, and so $u_n \to u$ because the coercivity estimate given above shows that the $u_n$ and $u$ are all bounded. Thus the map just described is continuous.

The following lemma is interesting in terms of the hypotheses of the above theorem. [□]

**Lemma 9.7.9** Let $L : D(L) \to X'$ where $D(L)$ is dense and $L$ is a closed operator. Then $L$ is maximal monotone if and only if both $L, L^*$ are monotone.
Proof: Suppose both \( L, L^* \) are monotone. One must show that \( \lambda F + L \) is onto. However, \( F \) is monotone and hemicontinuous (actually demicontinuous) and coercive. Hence the fact that \( \lambda F + L \) is onto follows from Theorem 4.7.8. Next suppose \( L \) is maximal monotone. If \( L \) is maximal monotone, then for every \( \varepsilon > 0 \) there exists a solution \( u_\varepsilon \) such that \( \varepsilon Lu_\varepsilon + F(u_\varepsilon - u) = 0 \). Here \( u \in D(L^*) \). This is from Lemma 5.6.22. It is originally due to Browder [11]. Then

\[
\varepsilon \langle Lu_\varepsilon, u_\varepsilon \rangle + \langle F(u_\varepsilon - u), u_\varepsilon \rangle = 0
\]

and so \( \langle F(u_\varepsilon - u), u_\varepsilon \rangle \leq 0 \). Then

\[
\langle F(u_\varepsilon - u), u_\varepsilon - u \rangle \leq \langle F(u_\varepsilon - u), u \rangle
\]

so \( \|u_\varepsilon - u\|^2 \leq \|u_\varepsilon - u\| \|u\| \) and so

\[
\|u_\varepsilon - u\| \leq \|u\|
\]

Thus the \( u_\varepsilon \) are bounded.

Next let \( v \in D(L) \).

\[
\|u_\varepsilon - u\|^2 = \langle F(u_\varepsilon - u), u_\varepsilon - u \rangle = \langle F(u_\varepsilon - u), u_\varepsilon - u \rangle + \langle F(u_\varepsilon - u), v - u \rangle
\]

\[
\leq \varepsilon \langle Lv, v - u \rangle + \langle F(u_\varepsilon - u), v - u \rangle \leq \varepsilon \langle Lv, v - u \rangle + \langle F(u_\varepsilon - u), v - u \rangle
\]

Hence

\[
\limsup_{\varepsilon \to 0} \|u_\varepsilon - u\|^2 \leq \limsup_{\varepsilon \to 0} (\varepsilon \langle Lv, v - u \rangle + \langle F(u_\varepsilon - u), v - u \rangle)
\]

\[
\leq \limsup_{\varepsilon \to 0} \langle F(u_\varepsilon - u), v - u \rangle \leq \limsup_{\varepsilon \to 0} \|u_\varepsilon - u\| \|v - u\|
\]

and so \( u_\varepsilon \to u \) strongly. Also

\[
\langle F(u_\varepsilon - u), u_\varepsilon \rangle = -\varepsilon \langle Lu_\varepsilon, u_\varepsilon \rangle \leq 0
\]

Then

\[
\langle L^*u, u \rangle = \lim_{\varepsilon \to 0} \langle L^*u, u_\varepsilon \rangle = \lim_{\varepsilon \to 0} \langle Lu_\varepsilon, u \rangle = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \langle -F(u_\varepsilon - u), u \rangle
\]

\[
= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \langle F(u_\varepsilon - u), u_\varepsilon - u \rangle - \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \langle F(u_\varepsilon - u), u_\varepsilon \rangle
\]

Both of these last terms are nonnegative, the first obviously and the second from the above where it was shown that \( \langle F(u_\varepsilon - u), u_\varepsilon \rangle \leq 0 \). ■

In the hypotheses of Theorem 4.7.8 one could have simply said that \( L \) is closed, linear, densely defined and maximal monotone. One can also show that if \( L \) is maximal monotone, then it must be densely defined. This is done in [3].

One can go further in obtaining a perturbation theorem like the above. Let linear \( L \) be densely defined with \( L \) closed and \( L, L^* \) monotone. In short, \( L \) is densely defined and maximal monotone, \( L : X \to X' \). Let \( A \) be a set valued \( L \) pseudomonotone operator which is coercive and bounded. Also let \( B : D(B) \to P(X) \) be maximal monotone. It is of interest to consider whether \( L + A + B \) is onto \( X' \). In considering this, I will add further assumptions as needed. First note that \( \langle Lx, x \rangle = \langle Lx - L0, x - 0 \rangle \geq 0 \).
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

Definition 9.7.10 Define
\[ \limsup_{m,n \to \infty} a_{m,n} \equiv \lim_{k \to \infty} \sup \{ a_{m,n} : \min(m,n) \geq k \} \]

Then \( \limsup_{m,n \to \infty} a_{m,n} \geq \limsup_{n \to \infty} (\limsup_{m \to \infty} a_{m,n}) \). To see this, suppose \( a > \limsup_{m,n \to \infty} a_{m,n} \). Then there exist \( k \) such that whenever \( m,n > k \),
\[ a_{m,n} < a \]
It follows that for \( m \geq k \),
\[ \limsup_{n \to \infty} a_{m,n} \leq a \]

Hence
\[ \limsup_{m \to \infty} \left( \limsup_{n \to \infty} a_{m,n} \right) \leq a \]
Since \( a > \limsup_{m,n \to \infty} a_{m,n} \) is arbitrary, it follows that
\[ \limsup_{m \to \infty} \left( \limsup_{n \to \infty} a_{m,n} \right) \leq \limsup_{m,n \to \infty} a_{m,n} \]

Then the following lemma is useful. I found this result in a paper by Gasinski, Migorski and Ochal [28]. They begin with the following interesting lemma or something like it which is similar to some of the ideas used in the section on approximation of maximal monotone operators.

Lemma 9.7.11 Suppose \( A \) is a set valued operator, \( A : X \to \mathcal{P}(X) \) and \( u^*_n \in Au_n \). Suppose also that \( u_n \to u \) weakly and \( u^*_n \to u^* \) weakly. Suppose also that
\[ \limsup_{m,n \to \infty} \langle u^*_n - u^*_m, u_n - u_m \rangle \leq 0 \]

Then one can conclude that
\[ \limsup_{n \to \infty} \langle u^*_n, u_n - u \rangle \leq 0 \]

Proof: Let \( \alpha \equiv \limsup_{n \to \infty} \langle u^*_n, u_n \rangle \). It is a finite number because these sequences are bounded. Then using the weak convergence,
\[
\begin{align*}
0 & \geq \limsup_{m \to \infty} \left( \limsup_{n \to \infty} \langle u^*_n - u^*_m, u_n - u_m \rangle \right) \\
& = \limsup_{m \to \infty} \left( \limsup_{n \to \infty} \left( \langle u^*_n, u_n \rangle + \langle u^*_m, u_m \rangle - \langle u^*_n, u_m \rangle - \langle u^*_m, u_n \rangle \right) \right) \\
& = \limsup_{m \to \infty} \left( \alpha + \langle u^*_n, u_n \rangle - \langle u^*_m, u_m \rangle - \langle u^*_m, u \rangle \right) \\
& = \left( \alpha + \alpha - \langle u^*, u \rangle - \langle u^*, u \rangle \right) = 2\alpha - 2 \langle u^*, u \rangle \\
& \leq 2(\alpha - \langle u^*, u \rangle) = 2\alpha - 2 \langle u^*, u \rangle
\end{align*}
\]
Now
\[ \limsup_{n \to \infty} \langle u^*_n, u_n - u \rangle = \alpha - \langle u^*, u \rangle \leq 0. \]
9.7. PERTURBATION THEOREMS

To begin with, consider the approximate problem which is to determine whether \( L + A + B \) is onto. Here \( B_\lambda x = -\lambda^{-1}F(x_\lambda - x) \) where \( 0 \in F(x_\lambda - x) + \lambda Bx \). In the notation given above, \( B_\lambda x = -\lambda^{-1}F(J_\lambda x - x) \). Then by Theorem 9.6.30, \( B_\lambda \) is monotone, demicontinuous, and bounded. In addition, we assume \( 0 \in D(B) \).

Then

\[
\langle B_\lambda x, x \rangle \geq \langle B_\lambda 0, x \rangle \geq -|B(0)| \|x\| \tag{9.7.79}
\]

**Lemma 9.7.12** Let \( A \) be pseudomonotone, bounded and coercive and let \( 0 \in D(B) \). Then if \( y^* \in X' \), there exists a solution \( x_\lambda \) to

\[
y^* \in Lx_\lambda + Ax_\lambda + B_\lambda x_\lambda
\]

**Proof:** From the inequality 9.7.79, \( A + B_\lambda \) is coercive. It is also bounded and pseudomonotone. It is pseudomonotone from Theorem 9.6.21. Therefore, there exists a solution \( x_\lambda \) by Theorem 9.7.8. ■

Acting on \( x_\lambda \) and using the inequality 9.7.79, it follows that these solutions \( x_\lambda \) lie in a bounded set. The details follow. Letting \( z^*_\lambda \in Ax_\lambda \) be such that equality holds in the above inclusion,

\[
y^* = Lx_\lambda + z^*_\lambda + B_\lambda x_\lambda \tag{9.7.80}
\]

Thus, from coercivity, \( \|x_\lambda\| \) are bounded. Then since \( A \) is bounded, the \( z^*_\lambda \) are all bounded also independent of \( \lambda \). The top line shows also that

\[
\langle y^*, x_\lambda \rangle = \frac{\langle Lx_\lambda, x_\lambda \rangle + \langle z^*_\lambda, x_\lambda \rangle + \langle B_\lambda x_\lambda, x_\lambda \rangle}{\|x_\lambda\|} \geq \frac{\langle Lx_\lambda, x_\lambda \rangle + \langle z^*_\lambda, x_\lambda \rangle - |B(0)| \|x_\lambda\|}{\|x_\lambda\|} \geq \frac{\langle z^*_\lambda, x_\lambda \rangle}{\|x_\lambda\|} - |B(0)|
\]

Thus, from coercivity, \( \|x_\lambda\| \) are bounded. Then since \( A \) is bounded, the \( z^*_\lambda \) are all bounded also independent of \( \lambda \). The top line shows also that

\[
\langle y^*, x_\lambda \rangle = \langle Lx_\lambda, x_\lambda \rangle + \langle z^*_\lambda, x_\lambda \rangle + \langle B_\lambda x_\lambda, x_\lambda \rangle \geq \langle z^*_\lambda, x_\lambda \rangle + \langle B_\lambda x_\lambda, x_\lambda \rangle \geq \langle B_\lambda x_\lambda, x_\lambda \rangle - M \geq -|B(0)| \|x_\lambda\| - M \tag{9.7.81}
\]

where \( |\langle z^*_\lambda, x_\lambda \rangle| \leq \hat{M} \) for all \( \lambda \). Hence there is a constant \( M \) such that

\[
|\langle B_\lambda x_\lambda, x_\lambda \rangle| \leq M
\]

**Definition 9.7.13** A set valued operator \( B \) is quasi-bounded if whenever \( x \in D(B) \) and \( x^* \in Bx \) are such that

\[
|\langle x^*, x \rangle|, \ |x| \leq M,
\]

it follows that \( \|x^*\| \leq K_M \). Bounded would mean that if \( \|x\| \leq M \), then \( \|x^*\| \leq K_M \). Here you only know this if there is another condition.
Lemma 9.7.14 In the above situation, suppose the maximal monotone operator $B$ is quasi-bounded and $\langle B_\lambda x_\lambda, x_\lambda \rangle \leq M$. Then the $B_\lambda x_\lambda$ are bounded. Also

$$\|J_\lambda x_\lambda - x_\lambda\|^2 \leq M\lambda$$

**Proof:** Now $B_\lambda x_\lambda \in BJ_\lambda x_\lambda$

$$- |B(0)||x_\lambda| \leq \langle B_\lambda x_\lambda, x_\lambda \rangle = \langle B_\lambda x_\lambda, J_\lambda x_\lambda \rangle + \langle B_\lambda x_\lambda, x_\lambda - J_\lambda x_\lambda \rangle$$

$$= \langle B_\lambda x_\lambda, J_\lambda x_\lambda \rangle + \langle \lambda^{-1}F(J_\lambda x_\lambda - x_\lambda), J_\lambda x_\lambda - x_\lambda \rangle$$

$$= \langle B_\lambda x_\lambda, J_\lambda x_\lambda \rangle + \lambda^{-1}\|J_\lambda x_\lambda - x_\lambda\|^2 \leq M$$

This inequality shows that $J_\lambda x_\lambda - x_\lambda \to 0$ and so $J_\lambda x_\lambda$ is bounded as is $x_\lambda$ which was shown above. Also $B_\lambda x_\lambda \in BJ_\lambda x_\lambda$ and since $B$ is quasi-bounded, it follows that $B_\lambda x_\lambda$ is bounded.

Assume from now on that $B$ is quasi-bounded. Then the estimate 9.7.81 and this lemma shows that $B_\lambda x_\lambda$ is also bounded independent of $\lambda$. Thus, adjusting the constants, there exists an estimate of the form

$$\|x_\lambda\| + \|J_\lambda x_\lambda\| + \|B_\lambda x_\lambda\| + \|z_\lambda^*\| + \|Lx_\lambda\| \leq C, \quad \|x_\lambda - J_\lambda x_\lambda\| \leq \sqrt{\lambda}M \quad (9.7.82)$$

Let $\lambda = 1/n$. Also denote by $J_n$ the the operator $J_{1/n}$ to save notation. There exists a subsequence

$$x_n \to x \text{ weakly},$$

$$J_n x_n \to x \text{ weakly},$$

$$B_n x_n \to g^* \text{ weakly},$$

$$z_n^* \to z^* \text{ weakly},$$

$$Lx_n \to Lx \text{ weakly}$$

Now from the inclusion satisfied,

$$0 = \langle z_n^* - z_m^*, x_n - x_m \rangle + \langle B_n x_n - B_m x_m, x_n - x_m \rangle \quad (9.7.83)$$

Consider that last term. $B_n x_n \in BJ_n x_n$ similar for $B_m x_m$. Hence this term is of the form

$$\langle B_n x_n - B_m x_m, x_n - x_m \rangle = \langle B_n x_n - B_m x_m, J_n x_n - J_m x_m \rangle$$

$$\langle B_n x_n - B_m x_m, (x_n - J_n x_n) - (x_m - J_m x_m) \rangle$$

From the estimate 9.7.81,

$$\langle B_n x_n - B_m x_m, x_n - x_m \rangle \geq \langle B_n x_n - B_m x_m, (x_n - J_n x_n) - (x_m - J_m x_m) \rangle$$
and
\[
|\langle B_n x_n - B_m x_m, (x_n - J_n x_n) - (x_m - J_m x_m) \rangle| \leq 2C \left( \sqrt{\frac{1}{n}} + \sqrt{\frac{1}{m}} \right)
\]

Then from 9.7.83,
\[
0 \geq \langle z^*_n - z^*_m, x_n - x_m \rangle + e_{n,m}
\]
where \(e_{n,m} \to 0\) as \(n, m \to \infty\). Hence
\[
\limsup_{m,n\to\infty} \langle z^*_n - z^*_m, x_n - x_m \rangle \leq 0
\]
From Lemma 9.7.11,
\[
\limsup_{n\to\infty} \langle z^*_n, x_n - x \rangle \leq 0
\]
Hence, since \(A\) is pseudomonotone, for every \(y\), there exists \(z^* (y) \in Ax\) such that
\[
\liminf_{n\to\infty} \langle z^*_n, x_n - y \rangle \geq \langle z^* (y), x - y \rangle
\]
In particular, if \(x = y\), this shows that
\[
\liminf_{n\to\infty} \langle z^*_n, x_n - x \rangle \geq 0 \geq \limsup_{n\to\infty} \langle z^*_n, x_n - x \rangle
\]
showing that
\[
\lim_{n\to\infty} \langle z^*_n, x_n \rangle = \langle z^*, x \rangle
\]
Next, returning to the inclusion solved,
\[
0 = Lx_n + z^*_n + B_n x_n
\]
Act on \((x_n - x)\). Then from monotonicity of \(L\),
\[
0 \geq \langle Lx, x_n - x \rangle + \langle z^*_n, x_n - x \rangle + \langle B_n x_n, x_n - x \rangle
\]
Thus, taking \(\limsup\) of both sides,
\[
\limsup_{n\to\infty} \langle B_n x_n, x_n - x \rangle = \limsup_{n\to\infty} \langle B_n x_n, J_n x_n - x \rangle \leq 0
\]
Hence
\[
\limsup_{n\to\infty} \langle B_n x_n, J_n x_n \rangle \leq \langle g^*, x \rangle
\]
Letting \([a,b^*] \in G (B)\),
\[
\langle B_n x_n - b^*, J_n x_n - a \rangle = \langle B_n x_n, J_n x_n \rangle - \langle B_n x_n, a \rangle - \langle b^*, J_n x_n \rangle + \langle b^*, a \rangle
\]
Then taking \(\limsup\),
\[
0 \leq \limsup_{n\to\infty} \langle B_n x_n - b^*, J_n x_n - a \rangle \leq \langle g^*, x \rangle - \langle g^*, a \rangle - \langle b^*, x \rangle + \langle b^*, a \rangle = \langle g^* - b^*, x - a \rangle
\]
CHAPTER 9. NONLINEAR OPERATORS IN BANACH SPACE

It follows that \( g^* \in B(x) \) and \( x \in D(B) \).

Thus, passing to the limit in the equation \( 9.7.80 \) where, as explained \( \lambda = 1/n \), one obtains

\[
y^* = Lu + z^* + g^*
\]

where \( z^* \in Ax \) and \( g^* \in Bx \). This proves the following nice generalization of the above perturbation theorem.

**Theorem 9.7.15** Let \( B \) be maximal monotone from \( X \) to \( \mathcal{P}(X') \), \( 0 \in D(B) \), and \( B \) is quasi-bounded as explained above. Let \( A : X \to \mathcal{P}(X') \) be pseudomonotone, bounded, and coercive. Also let \( L \) be a densely defined linear operator such that both \( L \) and \( L^* \) are monotone. (That is, \( L \) is linear and maximal monotone.) Then \( L + A + B \) is onto \( X' \).

### 9.8 Exercises

1. Let \( X, V \) be reflexive Banach spaces with \( X \subseteq V \) and \( X \) dense in \( V \). Suppose that \( A : V \to V' \) is pseudomonotone. Show that \( A : X \to X' \) is also pseudomonotone.

2. A bilinear form \( (u, v) \to a(u, v), u, v \in V \) a separable reflexive Banach space is just a real valued map which is linear in both variables, hence the term bilinear. The Lax Milgram theorem says that if you have such a bilinear form which is coercive, meaning that

\[
a(u, u) \geq \delta \|u\|^2, \quad |a(u, v)| \leq C \|u\| \|v\|
\]

then for every \( f \in V' \), there exists a unique \( u \in V \) such that \( a(u, v) = \langle f, v \rangle \) for all \( v \in V \). Prove this. Also show continuous dependence on \( f \).

3. There was a theorem that if \( A \) is monotone and hemicontinuous, then it is pseudomonotone. Thus if it is also coercive on a separable reflexive Banach space \( V \) then it was surjective onto \( V' \). An operator is strictly monotone if it is monotone and also

\[
\langle Au - Av, u - v \rangle \geq 0
\]

and equals 0 if and only if \( u = v \). Suppose you have such an operator on a separable reflexive Banach space which is strictly monotone. Consider the mapping \( f \to u \) where \( Au = f \). Show that this mapping is demicontinuous from the strong topology of \( V' \) to the weak topology of \( V \). That is, if \( f_n \to f \) in \( V' \), then \( u_n \to u \) (weak convergence) in \( V \).

4. Use the above problem to prove the following simple result on measurability. Suppose \( f(\omega) \in V' \) and \( \omega \to f(\omega) \) is measurable with respect to a measurable space \( (\Omega, \mathcal{F}) \). Also suppose \( A \) is a strictly monotone operator for each \( \omega \). Let \( Au(\omega) = f(\omega) \). Show that \( \omega \to u(\omega) \) must be measurable into \( V \).
Chapter 10

Maximal Monotone Operators, Hilbert Space

10.1 Lipschitz Functions

**Definition 10.1.1** A function \( f : [a, b] \to \mathbb{R} \) is Lipschitz if there is a constant \( K \) such that for all \( x, y \),

\[
|f(x) - f(y)| \leq K |x - y|.
\]

More generally, \( f \) is Lipschitz on a subset of \( \mathbb{R}^n \) if for all \( x, y \) in this set,

\[
|f(x) - f(y)| \leq K |x - y|.
\]

**Lemma 10.1.2** Suppose \( f : [a, b] \to \mathbb{R} \) is Lipschitz continuous and increasing. Then \( f' \) exists a.e., is in \( L^1([a,b]) \), and

\[
f(x) = f(a) + \int_a^x f'(t) \, dt.
\]

If \( f : \mathbb{R} \to \mathbb{R} \) is Lipschitz, then it is in \( L^1_{loc}(\mathbb{R}) \).

**Proof:** The Dini derivates are defined as follows.

\[
D^+ f(x) \equiv \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h}, \quad D_- f(x) \equiv \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h}
\]

\[
D^- f(x) \equiv \limsup_{h \to 0^+} \frac{f(x) - f(x-h)}{h}, \quad D_+ f(x) \equiv \liminf_{h \to 0^+} \frac{f(x) - f(x-h)}{h}
\]

For convenience, just let \( f \) equal \( f(a) \) for \( x < a \) and equal \( f(b) \) for \( x > b \). Let \((a, b)\) be an open interval and let

\[
N_{ab} \equiv \{ x \in (a, b) : D^+ f(x) > q > p > D_- f(x) \}
\]

381
Let \( V \subseteq (a, b) \) be an open set containing \( N_{pq} \) such that \( m(V) < m(N_{pq}) + \varepsilon \). By assumption, if \( x \in N_{pq} \), there exist arbitrarily small \( h \) such that

\[
\frac{f(x + h) - f(x)}{h} < p,
\]

These intervals \([x, x + h]\) are then a Vitali covering of \( N_{pq} \). It follows from Corollary A.0.8 that there is a disjoint union of countably many, \( \{[x_i, x_i + h_i]\}_{i=1}^{\infty} \) which cover all of \( N_{pq} \) except for a set of measure zero. Thus also the open intervals \( \{(x_i, x_i + h_i)\}_{i=1}^{\infty} \) also cover all of \( N_{pq} \) except for a set of measure zero. Now for points \( x' \) of \( N_{pq} \) so covered, there are arbitrarily small \( h \) such that

\[
\frac{f(x' + h') - f(x')}{h'} > q
\]

and \([x', x' + h']\) is contained in one of these original open intervals \((x_i, x_i + h_i)\). By the Vitali covering theorem again, Corollary A.0.8, it follows that there exists a countable disjoint sequence \( \{[x'_{j}, x'_{j} + h'_{j}]\}_{j=1}^{\infty} \) which covers all of \( N_{pq} \) except for a set of measure zero, each of these \([x'_{j}, x'_{j} + h'_{j}]\) being contained in some \((x_i, x_i + h_i)\) . Then it follows that

\[
qm(N_{pq}) \leq q \sum_{j} h'_{j} \leq \sum_{j} f(x'_{j} + h'_{j}) - f(x'_{j}) \leq \sum_{i} f(x_i + h_i) - f(x_i)
\]

\[
\leq p \sum_{i} h_i \leq pm(V) \leq pm(N_{pq}) + \varepsilon
\]

Since \( \varepsilon > 0 \) is arbitrary, this shows that \( qm(N_{pq}) \leq pm(N_{pq}) \) and so \( m(N_{pq}) = 0 \). Now taking the union of all \( N_{pq} \) for \( p, q \in \mathbb{Q} \), it follows that for a.e. \( x, D_{+}f(x) = D_{+}f(x) \) and so the derivative from the right exists. Similar reasoning shows that the set of derivative zero is the right derivative from the left also exists. You just do the same argument with \( D_{-}f(x) \) and \( D_{-}f(x) \) to obtain the existence of a derivative from the left. Next you can use the same argument to verify that \( D_{-}f(x) = D_{+}f(x) \) off a set of measure zero. This is outlined next. Define a new \( N_{pq} \):

\[
N_{pq} = \{x \in (a, b) : D_{+}f(x) > q > p > D_{-}f(x)\}
\]

Let \( V \) be an open set containing \( N_{pq} \) such that \( m(V) < m(N_{pq}) + \varepsilon \). For each \( x \in N_{pq} \) there are arbitrarily small \( h \) such that

\[
\frac{f(x) - f(x - h)}{h} < p
\]

Then as before, there is a countable disjoint sequence of closed intervals contained in \( V, \{[x_i - h_i, x_i]\}_{i=1}^{\infty} \) such that their union includes all of \( N_{pq} \) except a set of measure zero. Thus this is also true of the open intervals \( \{(x_i - h_i, x_i)\}_{i=1}^{\infty} \). Then for the points of \( N_{pq} \) covered by these open intervals \( x' \), there are arbitrarily small \( h' \) such that

\[
\frac{f(x' + h') - f(x')}{h'} > q.
\]
and each \([x', x' + h']\) is contained in an interval \((x_i - h_i, x_i)\). Then by the Vitali covering theorem again, Corollary 4.1.3 there are countably many disjoint closed intervals \(\{[x_j', x_j' + h_j']\}_{j=1}^{\infty}\) whose union includes all of \(N_{pq}\) except for a set of measure zero such that each of these is contained in some \((x_i - h_i, x_i)\) described earlier. Then as before,

\[
qm(N_{pq}) \leq q \sum_j h_j' \leq \sum_j f(x_j' + h_j') - f(x_j') \leq \sum_i f(x_i) - f(x_i - h_i)
\]

\[
\leq p \sum_i h_i \leq pm(V) \leq p(m(N_{pq}) + \varepsilon)
\]

Then as before, this shows that \(qm(N_{pq}) \leq pm(N_{pq})\) and so \(m(N_{pq}) = 0\). Then taking the union of all such for \(p, q \in \mathbb{Q}\) yields \(D^+_f(x) = D^-f(x)\) for a.e. \(x\). Taking the union of all these sets of measure zero and considering points not in this union, it follows that \(f'(x)\) exists for a.e. \(x\). Thus \(f'(t) \geq 0\) and is a limit of measurable even continuous functions for a.e. \(x\) so \(f'\) is clearly measurable. The issue is whether \(f(y) - f(x) = \int x \mathcal{A}_{[x,y]}(t) f'(t) \, dt\). Up to now, the only thing used has been that \(f\) is increasing.

Let \(h > 0\).

\[
\int_a^x \frac{f(t) - f(t-h)}{h} \, dt = \frac{1}{h} \int_a^x f(t) \, dt - \frac{1}{h} \int_a^{x-h} f(t) \, dt
\]

\[
= \frac{1}{h} \int_a^x f(t) \, dt - \frac{1}{h} \int_a^x f(t) \, dt
\]

\[
= \frac{1}{h} \int_{x-h}^x f(t) \, dt - \frac{1}{h} \int_{x-h}^x f(t) \, dt
\]

Therefore, by continuity of \(f\) it follows from Fatou’s lemma that

\[
\int_a^x D_+ f(t) \, dt = \int_a^x f'(t) \, dt \leq \lim_{h \to 0^+} \int_a^x \frac{f(t) - f(t-h)}{h} \, dt = f(x) - f(a)
\]

and this shows that \(f'\) is in \(L^1\). This part only used the fact that \(f\) is increasing and continuous. That \(f\) is Lipschitz has not been used.

If it were known that there is a dominating function for \(t \to \frac{f(t) - f(t-h)}{h}\), then you could simply apply the dominated convergence theorem in the above inequality instead of Fatou’s lemma and get the desired result. But from Lipschitz continuity, you have

\[
\left| \frac{f(t) - f(t-h)}{h} \right| \leq K
\]

and so one can indeed apply the dominated convergence theorem and conclude that

\[
\int_a^x f'(t) \, dt = f(x) - f(a)
\]
The last claim follows right away from consideration of intervals since the restriction
of a Lipschitz function is Lipschitz. ■

With the above lemmas, the following is the main theorem about absolutely
continuous functions.

The following simple corollary is a case of Rademacher’s theorem.

**Corollary 10.1.3** Suppose \( f : [a, b] \to \mathbb{R} \) is Lipschitz continuous,\[
|f(x) - f(y)| \leq K|x - y|.
\]
Then \( f'(x) \) exists a.e. and
\[
f(x) = f(a) + \int_a^x f'(t) \, dt.
\]

**Proof:** If \( f \) were increasing, this would follow from the above lemma. Let \( g(x) = 2Kx - f(x) \). Then \( g \) is Lipschitz with a different Lipschitz constant and also if \( x < y \),
\[
g(y) - g(x) = 2Ky - f(y) - (2Kx - f(x)) \\
\geq 2K(y - x) - K|y - x| = k|y - x| \geq 0
\]
and so Lemma 10.1.2 applies to \( g \) and this shows that \( f'(t) \) exists for a.e. \( t \) and \( g'(x) = 2K - f'(x) \). Also
\[
2K(x - a) - (f(x) - f(a)) = g(x) - g(a) = 2Kx - f(x) - (2Ka - f(a)) = \int_a^x (2K - f'(t)) \\
= 2K(x - a) - \int_a^x f'(t) \, dt
\]
showing that \( f(x) - f(a) = \int_a^x f'(t) \, dt \). ■

### 10.2 Basic Theory

Here is provided a short introduction to some of the most important properties of
maximal monotone operators in Hilbert space. The following definition describes
them. It is more specialized than the earlier material on maximal monotone opera-
tors from a Banach space to its dual and therefore, better results can be obtained.
More on this can be read in [9] and [54].

**Definition 10.2.1** Let \( H \) be a real Hilbert space and let \( A : D(A) \to \mathcal{P}(H) \) have
the following properties.

1. For each \( y \in H \) there exists \( x \in D(A) \) such that \( y \in x + Ax \).
2. $A$ is monotone. That is, if $z \in Ax$ and $w \in Ay$ then

$$(z - w, x - y) \geq 0$$

Such an operator is called a maximal monotone operator.

It turns out that whenever $A$ is maximal monotone, so is $\lambda A$ for all $\lambda > 0$.

**Lemma 10.2.2** Suppose $A$ is maximal monotone. Then so is $\lambda A$. Also $J_{\lambda} \equiv (I + \lambda A)^{-1}$ makes sense for each $\lambda > 0$ and is Lipschitz continuous.

**Proof:** To begin with consider $(I + A)^{-1}$. Suppose

$$x_1, x_2 \in (I + A)^{-1}(y)$$

Then $y \in (I + A)x_i$ and so $y - x_i \in Ax_i$. By monotonicity

$$(y - x_1 - (y - x_2), x_1 - x_2) \geq 0$$

and so

$$0 \geq |x_1 - x_2|^2$$

which shows $J_1 \equiv (I + A)^{-1}$ makes sense. In fact this is Lipschitz with Lipschitz constant 1. Here is why. $x \in (I + A)J_1x$ and $y \in (I + A)J_1y$. Then

$$x - J_1x \in AJ_1x, \quad y - J_1y \in AJ_1y$$

and so by monotonicity

$$0 \leq (x - J_1x - (y - J_1y), J_1x - J_1y)$$

which yields

$$|J_1x - J_1y|^2 \leq (x - y, J_1x - J_1y) \leq |x - y||J_1x - J_1y|$$

which yields the result.

Next consider the claim that $\lambda A$ is maximal monotone. The monotone part is immediate. The only thing in question is whether $I + \lambda A$ is onto. Let $r \in (-1, 1)$ and pick $f \in H$. Consider solving the equation for $u$

$$(1 + r)u + Au \ni (1 + r)f \quad (10.2.1)$$

This is equivalent to finding $u$ such that

$$(I + A)u \ni (1 + r)f - ru$$

or in other words finding $u$ such that

$$u = J_1((1 + r)f - ru)$$
However, if 
\[ Tu \equiv J_1 ((1 + r) f - ru), \]
then since \(|r| < 1\), \(T\) is a contraction mapping and so there exists a unique solution to (10.2.1). Thus 
\[ u + \frac{1}{1 + r} Au \ni f \]
It follows for any \(|r| < 1, (1 + r)^{-1} A\) is maximal monotone. This takes care of all \(\lambda \in (\frac{1}{2}, \infty)\). Now do the same thing for \((2/3) A\) to get the result for all \(\lambda \in (\left(\frac{2}{3}\right) \left(\frac{1}{2}\right), \infty)\). Next consider the same argument to \((2/3)^2 A\) to get the desired result for all \(\lambda \in \left(\frac{2}{3} \left(\frac{1}{2}\right), \infty\right)\). Continuing this way shows \(\lambda A\) is maximal monotone for all \(\lambda > 0\). Also from the first part of the proof \((I + \lambda A)^{-1}\) is Lipschitz continuous with Lipschitz constant 1.

A maximal monotone operator can be approximated with a Lipschitz continuous operator which is also monotone and has certain salubrious properties. This operator is called the Yosida approximation and as in the case of linear operators it is obtained by formally considering

\[ A \frac{1}{1 + \lambda A} \]

If you do the division formally you get the definition for \(A_\lambda\),

\[ A_\lambda x \equiv \frac{1}{\lambda} x - \frac{1}{\lambda} J_\lambda x \tag{10.2.2} \]

where \(J_\lambda = (I + \lambda A)^{-1}\) as above. It is obvious that \(A_\lambda\) is Lipschitz continuous with Lipschitz constant no more than 2/\(\lambda\). Actually you can show 1/\(\lambda\) also works but this is not important here.

**Lemma 10.2.3** \(A_\lambda x \in AJ_\lambda x\) and \(|A_\lambda x| \leq |y|\) for all \(y \in Ax\) whenever \(x \in D(A)\). Also \(A_\lambda\) is monotone.

**Proof:** Consider the first claim. From the definition,

\[ A_\lambda x \equiv \frac{1}{\lambda} x - \frac{1}{\lambda} J_\lambda x \]

\[ \frac{1}{\lambda} x - \frac{1}{\lambda} J_\lambda x \in AJ_\lambda x? \]

\[ x - J_\lambda x \in \lambda AJ_\lambda x? \]

\[ x \in J_\lambda x + \lambda AJ_\lambda x? \]
10.2. BASIC THEORY

Is
\[ x \in (I + \lambda A) J_\lambda x? \]

Certainly so. This is how \( J_\lambda \) is defined.

Now consider the second claim. Let \( y \in Ax \) for some \( x \in D(A) \). Then by monotonicity and what was just shown
\[ 0 \leq (A_\lambda x - y, J_\lambda x - x) = -\lambda (A_\lambda x - y, A_\lambda x) \]
and so
\[ |A_\lambda x|^2 \leq (y, A_\lambda x) \leq |y| |A_\lambda x| \]

Finally, to show \( A_\lambda \) is monotone,
\[ (A_\lambda x - A_\lambda y, x - y) = \]
\[ \left( \frac{1}{\lambda} x - \frac{1}{\lambda} J_\lambda x - \left( \frac{1}{\lambda} y - \frac{1}{\lambda} J_\lambda y \right), x - y \right) \]
\[ = \frac{1}{\lambda} |x - y|^2 - \frac{1}{\lambda} (J_\lambda x - J_\lambda y, x - y) \]
\[ \geq \frac{1}{\lambda} |x - y|^2 - \frac{1}{\lambda} |x - y| |J_\lambda x - J_\lambda y| \]
\[ \geq \frac{1}{\lambda} |x - y|^2 - \frac{1}{\lambda} |x - y|^2 = 0 \]
and this proves the lemma.

**Proposition 10.2.4** Suppose \( D(A) \) is dense in \( H \). Then for all \( x \in H \),
\[ |J_\lambda x - x| \to 0 \]

**Proof:** From the above, if \( u \in D(A) \) and \( y \in Au \), then
\[ \left| \frac{1}{\lambda} u - \frac{1}{\lambda} J_\lambda u \right| \leq |y| \]
Hence \( J_\lambda u \to u \). Now for \( x \) arbitrary,
\[ |J_\lambda x - x| \leq |J_\lambda x - J_\lambda u| + |J_\lambda u - u| + |u - x| \]
\[ < 2\varepsilon + |J_\lambda u - u| \]
where the last term converges to 0 as \( \lambda \to 0 \). Since \( \varepsilon \) is arbitrary, this shows the proposition.

Thus in the case where \( D(A) \) is dense, if you have
\[ x \in \varepsilon Ax + x_\varepsilon \]
so that \( x_\varepsilon = J_\varepsilon x \), then \( |x - x_\varepsilon| \to 0 \).
The next lemma gives a way to determine whether a pair \([x, y]\) is in the graph of \(A\) defined as
\[
\{[x, y] : y \in Ax\} \equiv G(A)
\]
Here I am writing \([,\,]\) rather than \((,\,)\) to avoid confusion with the inner product. It is the conclusion of this lemma which accounts for the use of the term “maximal”. It essentially says there is no larger monotone graph which includes the one for \(A\).

**Lemma 10.2.5** Suppose \((y_1 - y, x_1 - x) \geq 0\) for all \([x, y] \in G(A)\) where \(A\) is maximal monotone. Then \(x_1 \in D(A)\) and \(y_1 \in Ax_1\). Also if \([x_k, y_k] \in G(A)\) and \(x_k \to x, y_k \rightharpoonup y\) where the half arrow denotes weak convergence, then \([x, y] \in G(A)\).

**Proof:** I want to show \(y_1 \in Ax_1\) or in other words I want to show
\[
x_1 + \lambda y_1 \in x_1 + \lambda Ax_1
\]
or in other words
\[
J_\lambda(x_1 + \lambda y_1) = x_1.
\]
This is the motivation for the following argument.

From Lemma 10.2.3 \(A_\lambda(x_1 + \lambda y_1) \in AJ_\lambda(x_1 + \lambda y_1)\) and so by the above assumption
\[
0 \leq (y_1 - A_{\lambda}(x_1 + \lambda y_1), x_1 - J_{\lambda}(x_1 + \lambda y_1))
\]
\[
= \left( y_1 - \left( \frac{1}{\lambda}(x_1 + \lambda y_1) - \frac{1}{\lambda}J_{\lambda}(x_1 + \lambda y_1) \right), x_1 - J_{\lambda}(x_1 + \lambda y_1) \right)
\]
\[
= \left( \left( -\frac{1}{\lambda}x_1 + \frac{1}{\lambda}J_{\lambda}(x_1 + \lambda y_1) \right), x_1 - J_{\lambda}(x_1 + \lambda y_1) \right)
\]
\[
= \frac{-1}{\lambda} (x_1 - J_{\lambda}(x_1 + \lambda y_1), x_1 - J_{\lambda}(x_1 + \lambda y_1))
\]
which requires
\[
x_1 = J_{\lambda}(x_1 + \lambda y_1)
\]
and this says \(x_1 \in D(A)\) because \(J_{\lambda}\) maps into \(D(A)\). Also it says
\[
x_1 + Ax_1 \ni x_1 + \lambda y_1
\]
and so \(y_1 \in Ax_1\).

This makes the last claim pretty easy. Suppose \(x_k \to x\) where \(x_k \in D(A)\) and that \(y_k \in Ax_k\) and \(y_k \rightharpoonup y\). I need to verify \(y = Ax\) and \(x \in D(A)\). Let \([u, v] \in G(A)\). Then
\[
(y - v, x - u) = \lim_{k \to \infty} (y_k - v, x_k - u) \geq 0
\]
and so, by the first part, \(x \in D(A)\) and \(y \in Ax\). Why does that limit hold? It is because
\[
|(y - v, x - u) - (y_k - v, x_k - u)|
\]
10.2. BASIC THEORY

\[ \leq |(y - v, x - u) - (y_k - v, x - u)| + |(y_k - v, x_k - x)| \]

The second term is no larger than

\[ |y_k - v| |x_k - x| \]

which converges to 0 since \( y_k \) is weakly convergent, hence bounded. The first term converges to 0 because of the assumption that \( y_k \) converges weakly to \( y \).

What about the sum of maximal monotone operators? This might not be maximal monotone but what you can say is the following.

**Proposition 10.2.6** Let \( A \) be maximal monotone and let \( B \) be Lipschitz and monotone. Then \( A + B \) is maximal monotone.

**Proof:** First suppose \( B \) has a Lipschitz constant less than 1. The monotonicity is obvious. I need to show that for any \( y \) there exists \( x \in D(A) \) such that

\[ y \in x + Bx + Ax \]

This happens if and only if

\[ y - Bx \in (I + A)x \]

if and only if \( x = (I + A)^{-1}(y - Bx) \). Let

\[ Tx \equiv (I + A)^{-1}(y - Bx) \]

Then \( T \) is clearly a contraction mapping because \( (I + A)^{-1} \) is Lipschits with Lipschitz constant 1. Therefore, there exists a unique fixed point and this shows \( A + B \) is maximal monotone. Now the same argument applied to \( A + B \) shows that \( A + 2B \) is maximal monotone. Continuing this way \( A + nB \) is maximal monotone. Now for arbitrary \( B \) let \( n \) be large enough that \( n^{-1}B \) has Lipschitz constant less than 1. Then as just explained, \( A + n (n^{-1}B) = A + B \) is maximal monotone. This proves the proposition.

The following is a useful result for determining conditions under which \( A + B \) is maximal monotone or more particularly whether a given \( y \) is in \((I + A + B)(H)\) where \( A,B \) are both maximal monotone.

**Theorem 10.2.7** Let \( A \) and \( B \) be maximal monotone, let

\[ y \in x_\lambda + B_\lambda x_\lambda + Ax_\lambda, \]

and suppose \( B_\lambda x_\lambda \) is bounded independent of \( \lambda \). Then there exists \( x \in D(A) \cap D(B) \) such that \( y = x + Ax + Bx \).

**Proof:** First of all, it follows from Proposition 10.2.6 that there exists a unique \( x_\lambda \). Note

\[ y - x_\lambda - B_\lambda x_\lambda \in Ax_\lambda \]

\[ y - x_\mu - B_\mu x_\mu \in Ax_\mu \]
and so by monotonicity of $A$,

$$(x_\mu - x_\lambda + B_\mu x_\mu - B_\lambda x_\lambda, x_\lambda - x_\mu) \geq 0$$

and so

$$|x_\lambda - x_\mu|^2 \leq (B_\mu x_\mu - B_\lambda x_\lambda, x_\lambda - x_\mu) = -(B_\lambda x_\lambda - B_\mu x_\mu, x_\lambda - x_\mu) \quad (10.2.3)$$

I want to write as many things as possible in terms of the $B_\lambda$ and $B_\mu$. Denote as $J_\lambda (B)$ the operator $(I + \lambda B)^{-1}$. Then

$$B_\lambda x_\lambda = \frac{1}{\lambda} (x_\lambda - J_\lambda (B) x_\lambda)$$

and so

$$x_\lambda = \lambda B_\lambda x_\lambda + J_\lambda (B) x_\lambda$$

Thus (10.2.3) becomes

$$|x_\lambda - x_\mu|^2 = - (B_\lambda x_\lambda - B_\mu x_\mu, \lambda B_\lambda x_\lambda + J_\lambda (B) x_\lambda - (\mu B_\mu x_\mu + J_\mu (B) x_\mu))$$

$$= - (B_\lambda x_\lambda - B_\mu x_\mu, \lambda B_\lambda x_\lambda - \mu B_\mu x_\mu) + (B_\mu x_\mu - B_\lambda x_\lambda, J_\lambda (B) x_\lambda - J_\mu (B) x_\mu)$$

$$= - (B_\lambda x_\lambda - B_\mu x_\mu, \lambda B_\lambda x_\lambda - \lambda B_\mu x_\mu) - (B_\lambda x_\lambda - B_\mu x_\mu, (\lambda - \mu) B_\mu x_\mu)$$

Now recall $B_\mu x \in BJ_\mu (B) x$. Then by monotonicity the first and last terms to the right of the equal sign in the above are negative. Therefore,

$$|x_\lambda - x_\mu|^2 \leq |(B_\lambda x_\lambda - B_\mu x_\mu, (\lambda - \mu) B_\mu x_\mu)| \leq C |\lambda - \mu|$$

where $C$ is some constant which comes from the assumption the $B_\lambda x_\lambda$ are bounded.

Therefore, letting $\lambda$ denote a sequence converging to 0 it follows

$$\lim_{\lambda \to 0} x_\lambda = x_1 \in H$$

for some $x$, the convergence being strong convergence. Also taking a further subsequence and using weak compactness it can be assumed

$$B_\lambda x_\lambda \rightharpoonup z_1$$

where this time the convergence is weak. Taking another subsequence, it can also be assumed

$$y - x_\lambda - B_\lambda x_\lambda \rightharpoonup z_2 \quad (10.2.4)$$
the convergence being weak convergence. Recall $B_\lambda x_\lambda \in BJ_\lambda (B) x_\lambda$ and also note that by assumption there is a constant $C$ independent of $\lambda$ such that

$$C \geq |B_\lambda x_\lambda| \geq \frac{1}{\lambda} (x_\lambda - J_\lambda (B) x)$$

which shows

$$J_\lambda (B) x_\lambda \to x_1$$

also. Now it follows from Lemma 10.2.5 that $x_1 \in D (B)$ and $z_1 \in Bx_1$. Recall

$$y - x_\lambda - B_\lambda x_\lambda \in Ax_\lambda$$

and so by the same lemma again,

$$x_1 \in D (A), z_2 \in Ax_1$$

By 10.2.4 it follows

$$y - x_1 - z_1 = z_2 \in Ax_1$$

Thus

$$y = x_1 + z_1 + z_2 \in x_1 + Bx_1 + Ax_1$$

and this proves the theorem.

### 10.3 Evolution Inclusions

One of the interesting things about maximal monotone operators is the concept of evolution inclusions. To facilitate this, here is a little lemma.

**Lemma 10.3.1** Let $f : [0, T] \to \mathbb{R}$ be continuous and suppose

$$D^+ f (t) \equiv \lim_{h \to 0^+} \sup f (t + h) - f (t) - g (t)$$

where $g$ is a continuous function. Then

$$f (t) - f (0) \leq \int_0^t g (s) \, ds.$$

**Proof:** Suppose this is not so. Then let

$$S \equiv \left\{ t \in [0, T] : f (t) - f (0) > \int_0^t g (s) \, ds \right\}$$

and it would follow that $S \neq \emptyset$. Let $a = \inf S$. Then there exists a decreasing sequence $h_n \to 0$ such that

$$f (a + h_n) - f (0) > \int_0^{a + h_n} g (s) \, ds$$

(10.3.5)
CHAPTER 10. MAXIMAL MONOTONE OPERATORS, HILBERT SPACE

First suppose \( a = 0 \). Then dividing by \( h_n \) and taking the limit,

\[
g(0) > D^+ f(0) \geq g(0),
\]

a contradiction. Therefore, assume \( a > 0 \). Then by continuity

\[
f(a) - f(0) \geq \int_0^a g(s) \, ds
\]

If strict inequality holds, then \( a \neq \inf S \). It follows

\[
f(a) - f(0) = \int_0^a g(s) \, ds
\]

and so from

\[
\frac{f(a + h_n) - f(a)}{h_n} > \frac{1}{h_n} \int_a^{a+h_n} g(s) \, ds.
\]

Then doing \( \limsup_{n \to \infty} \) to both sides,

\[
g(a) > D^+ f(a) \geq g(a)
\]

the same sort of contradiction obtained earlier. Thus \( S = \emptyset \) and this proves the lemma.

The following is the main result.

**Theorem 10.3.2** Let \( H \) be a Hilbert space and let \( A \) be a maximal monotone operator as described above. Let \( f : [0, T] \to H \) be continuous such that \( f' \in L^2(0, T; H) \). Then there exists a unique solution to the evolution inclusion

\[
y' + Ay \ni f, \quad y(0) = y_0 \in D(A)
\]

Here \( y' \) exists a.e., \( y(t) \in D(A) \) a.e., \( y \) is continuous.

**Proof:** Let \( y_\lambda \) be the solution to

\[
y_\lambda' + A_\lambda y_\lambda = f, \quad y_\lambda(0) = y_0
\]

I will base the entire proof on estimating the solutions to the corresponding integral equation

\[
y_\lambda(t) - y_0 + \int_0^t A_\lambda y_\lambda(s) \, ds = \int_0^t f(s) \, ds \quad \text{(10.3.6)}
\]

Let \( h, k \) be small positive numbers. Then

\[
y_\lambda(t + h) - y_\lambda(t) + \int_t^{t+h} A_\lambda y_\lambda(s) \, ds = \int_t^{t+h} f(s) \, ds \quad \text{(10.3.7)}
\]
Next consider the difference operator

\[ D_k g(t) = \frac{g(t+k) - g(t)}{k} \]

Do this \( D_k \) to both sides of 10.3.7 where \( k < h \). This gives

\[ D_k (y_\lambda (t+h) - y_\lambda (t)) + \frac{1}{k} \left( \int_{t+h}^{t+h+k} A_\lambda y_\lambda (s) \, ds - \int_{t}^{t+k} A_\lambda y_\lambda (s) \, ds \right) = \frac{1}{k} \left( \int_{t+h}^{t+h+k} f(s) \, ds - \int_{t}^{t+k} f(s) \, ds \right) \tag{10.3.8} \]

Now multiply both sides by \( y_\lambda (t+h+k) - y_\lambda (t+k) \). Consider the first term. To simplify the ideas consider instead

\[ (D_k g(t), g(t+k)) = \frac{1}{k} \left( |g(t+k)|^2 - (g(t), g(t+k)) \right) \geq \frac{1}{k} \left( |g(t+k)|^2 - |g(t)||g(t+h)| \right) \geq \frac{1}{k} \left( \frac{1}{2} |g(t+k)|^2 - \frac{1}{2} |g(t)|^2 \right) \tag{10.3.9} \]

Then applying this simple observation to 10.3.8

\[
\frac{1}{2} \frac{1}{k} \left( |y_\lambda (t+h+k) - y_\lambda (t+k)|^2 - |y_\lambda (t+h) - y_\lambda (t)|^2 \right) + \\
\left( \frac{1}{k} \left( \int_{t+h}^{t+h+k} A_\lambda y_\lambda (s) \, ds - \int_{t}^{t+k} A_\lambda y_\lambda (s) \, ds \right), y_\lambda (t+h+k) - y_\lambda (t+k) \right) + \\
\leq \left( \frac{1}{k} \left( \int_{t+h}^{t+h+k} f(s) \, ds - \int_{t}^{t+k} f(s) \, ds \right), y_\lambda (t+h+k) - y_\lambda (t+k) \right)
\]

Taking \( \limsup_{k \to 0} \) of both sides yields

\[
\frac{1}{2} D^+ \left( |y_\lambda (t+h) - y_\lambda (t)|^2 \right) + (A_\lambda y_\lambda (t+h) - A_\lambda y_\lambda (t), y_\lambda (t+h) - y_\lambda (t)) \leq (f(t+h) - f(t), y_\lambda (t+h) - y_\lambda (t))
\]

Now recall that \( A_\lambda \) is monotone. Therefore,

\[
D^+ \left( |y_\lambda (t+h) - y_\lambda (t)|^2 \right) \leq |f(t+h) - f(t)|^2 + |y_\lambda (t+h) - y_\lambda (t)|^2
\]

From Lemma 10.3.2 it follows that for all \( \varepsilon > 0 \),

\[
|y_\lambda (t+h) - y_\lambda (t)|^2 - |y_\lambda (h) - y_0|^2 \\
\leq \int_{t}^{t+h} |f(s) - f(t)|^2 \, ds + \int_{t}^{t+h} |y_\lambda (s) - y_\lambda (t)|^2 \, ds + \varepsilon t
\]
and so since \( \varepsilon \) is arbitrary, the term \( \varepsilon t \) can be eliminated. By Gronwall’s inequality,

\[
|y_\lambda (t + h) - y_\lambda (t)|^2 \leq e^t \left( |y_\lambda (h) - y_0|^2 + \int_0^t |f (s + h) - f (s)|^2 \, ds \right). \tag{10.3.10}
\]

The last integral equals

\[
\int_0^t \left[ \int_s^{s+h} |f' (r)| \, dr \right]^2 \, ds \leq \int_0^t h \left[ \int_s^{s+h} |f' (r)| \, dr \right]^2 \, ds
\]

\[
= h \left[ \int_0^h \int_0^t |f' (r)| \, ds \, dr + \int_h^t \int_0^r |f' (r)| \, ds \, dr + \int_t^{t+h} \int_r^{t+h} |f' (r)| \, ds \, dr \right]^2 \, ds
\]

and now it follows that for all \( t + h < T \),

\[
|y_\lambda (t + h) - y_\lambda (t)|^2 \leq e^T \left( \left| \frac{y_\lambda (h) - y_0}{h} \right|^2 + \| f' \|_{L^2(0,T; H)}^2 \right). \tag{10.3.11}
\]

Now return to \( 10.3.7 \)

\[
\left| \frac{y_\lambda (h) - y_0}{h} \right| \leq \frac{1}{h} \int_0^h A_\lambda y_\lambda (s) \, ds + \frac{1}{h} \int_0^h f (s) \, ds
\]

Then taking \( \limsup_{h \to 0} \) of both sides

\[
\limsup_{h \to 0} \frac{y_\lambda (h) - y_0}{h} \leq |A_\lambda y_0| + |f (0)|
\]

From Lemma \( 10.2.3 \), \( |A_\lambda y_0| \leq |a| \) for all \( a \in Ay_0 \). This is where \( y_0 \in D (A) \) is used. Thus from \( 10.3.11 \), there exists a constant \( C \) independent of \( t \) and \( h \) and \( \lambda \) such that

\[
\left| \frac{y_\lambda (t + h) - y_\lambda (t)}{h} \right|^2 \leq C
\]

From the estimate just obtained and \( 10.3.11 \), this implies

\[
\frac{y_\lambda (t + h) - y_\lambda (t)}{h} + \frac{1}{h} \int_t^{t+h} A_\lambda y_\lambda (s) \, ds = \frac{1}{h} \int_t^{t+h} f (s) \, ds \quad \tag{10.3.12}
\]

Now letting \( h \to 0 \), it follows that for all \( t \in [0,T) \), there exists a constant \( C \) independent of \( t, \lambda \) such that

\[
|A_\lambda y_\lambda (t)| \leq C \quad \tag{10.3.13}
\]

This is a very nice estimate. The next task is to show uniform convergence of the \( y_\lambda \) as \( \lambda \to 0 \). From \( 10.3.7 \)

\[
(D_h (y_\lambda (t) - y_\mu (t)), y_\lambda (t + h) - y_\mu (t + h)) +
\]
10.3. EVOLUTION INCLUSIONS

\[
\left( \frac{1}{h} \int_{t}^{t+h} (A_\lambda y_\lambda (s) - A_\mu y_\mu (s)) \, ds, y_\lambda (t + h) - y_\mu (t + h) \right) = 0
\]

Then from the argument in 10.3.9,

\[
\frac{1}{h} \int_{t}^{t+h} \left( |y_\lambda (t + h) - y_\mu (t + h)|^2 - |y_\lambda (t) - y_\mu (t)|^2 \right) + \left( \frac{1}{h} \int_{t}^{t+h} (A_\lambda y_\lambda (s) - A_\mu y_\mu (s)) \, ds, y_\lambda (t + h) - y_\mu (t + h) \right) \leq 0
\]

Now take \( \limsup_{h \to 0} \) to obtain

\[
\frac{1}{2} D^+ |y_\lambda (t) - y_\mu (t)|^2 + (A_\lambda y_\lambda (t) - A_\mu y_\mu (t), y_\lambda (t) - y_\mu (t)) \leq 0
\]

Using the definition of \( A_\lambda \) this equals

\[
\frac{1}{2} D^+ |y_\lambda (t) - y_\mu (t)|^2 + (A_\lambda y_\lambda (t) - A_\mu y_\mu (t), \lambda A_\lambda y_\lambda (t) + J_\lambda y_\lambda (t) - (\mu A_\mu y_\mu (t) + J_\mu y_\mu (t))) \leq 0
\]

Now this last term splits into the following sum

\[
(A_\lambda y_\lambda (t) - A_\mu y_\mu (t), \lambda A_\lambda y_\lambda (t) - \mu A_\mu y_\mu (t)) + (A_\lambda y_\lambda (t) - A_\mu y_\mu (t), J_\lambda y_\lambda (t) - J_\mu y_\mu (t))
\]

By Lemma 10.2.3 the second of these terms is nonnegative. Also from the estimate 10.3.13, the first term converges to 0 uniformly in \( t \) as \( \lambda, \mu \to 0 \). Then by Lemma 10.3.1 it follows that if \( \lambda \) is any sequence converging to 0, \( y_\lambda (t) \) is uniformly Cauchy. Let

\[
y(t) \equiv \lim_{\lambda \to 0} y_\lambda (t).
\]

Thus \( y \) is continuous because it is the uniform limit of continuous functions. Since \( A_\lambda y_\lambda (t) \) is uniformly bounded, it also follows

\[
y(t) = \lim_{\lambda \to 0} J_\lambda y_\lambda (t) \quad \text{uniformly in } t. \quad (10.3.14)
\]

Taking a further subsequence, you can assume

\[
A_\lambda y_\lambda \to z \quad \text{weak } * \quad \text{in } L^\infty (0, T; H). \quad (10.3.15)
\]

Thus \( z \in L^\infty (0, T; H) \). Recall \( A_\lambda y_\lambda \in AJ_\lambda y_\lambda \).

Now \( A \) can be considered a maximal monotone operator on \( L^2 (0, T; H) \) according to the rule

\[
Ay(t) \equiv A(y(t))
\]

where

\[
D(A) \equiv \{ f \in L^2 (0, T; H) : f(t) \in D(A) \; \text{a.e. } t \}
\]
CHAPTER 10. MAXIMAL MONOTONE OPERATORS, HILBERT SPACE

By Lemma 10.2.5 applied to $A$ considered as a maximal monotone operator on $L^2(0,T;H)$ and using 10.3.14 and 10.3.15, it follows $y(t) \in D(A)$ a.e. $t$ and $z(t) \in Ay(t)$ a.e. $t$. Then passing to the limit in 10.3.6 yields

$$y(t) - y_0 + \int_0^t z(s) \, ds = \int_0^t f(s) \, ds. \quad (10.3.16)$$

Then by fundamental theorem of calculus, $y'(t)$ exists a.e. $t$ and

$$y' + z = f, \quad y(0) = y_0$$

where $z(t) \in Ay(t)$ a.e.

It remains to verify uniqueness. Suppose $[y_1, z_1]$ is another pair which works. Then from 10.3.16,

$$y(t) - y_1(t) + \int_0^t (z(r) - z_1(r)) \, dr = 0$$
$$y(s) - y_1(s) + \int_0^s (z(r) - z_1(r)) \, dr = 0$$

Therefore for $s < t$,

$$y(t) - y_1(t) - (y(s) - y_1(s)) = \int_s^t (z(r) - z_1(r)) \, dr$$

and so

$$||y(t) - y_1(t)| - |y(s) - y_1(s)|| \leq K|s - t|$$

for some $K$ depending on $||z||_{L^\infty}, ||z_1||_{L^\infty}$. Since $y, y_1$ are bounded, it follows that $t \to |y(t) - y_1(t)|^2$ is also Lipschitz. Therefore by Corollary 10.3.16, it is the integral of its derivative which exists a.e. So what is this derivative? As before,

$$(D_h (y(t) - y_1(t)), y(t + h) - y_1(t + h))$$
$$+ \left( \frac{1}{h} \int_t^{t+h} (z(s) - z_1(s)) \, ds, y(t + h) - y_1(t + h) \right) = 0$$

and so

$$\frac{1}{h} \left( \frac{|y(t + h) - y_1(t + h)|^2}{2} - \frac{|y(t) - y_1(t)|^2}{2} \right)$$
$$+ \left( \frac{1}{h} \int_t^{t+h} (z(s) - z_1(s)) \, ds, y(t + h) - y_1(t + h) \right) \leq 0$$

Then taking $\lim_{h \to 0}$ it follows that for a.e. $t$ (Lebesgue points of $z - z_1$ intersected with the points where $|y - y_1|^2$ has a derivative)

$$\frac{1}{2} \frac{d}{dt} |y(t) - y_1(t)|^2 + (z(t) - z_1(t), y(t) - y_1(t)) \leq 0$$
Thus for a.e. $t$,

$$\frac{d}{dt} |y(t) - y_1(t)|^2 \leq 0$$

and so

$$|y(t) - y_1(t)|^2 - |y_0 - y_0|^2 = \int_0^t \frac{d}{dt} |y(s) - y_1(s)|^2 \, ds \leq 0.$$ 

\[\square\]

## 10.4 Subgradients

### 10.4.1 General Results

**Definition 10.4.1** Let $X$ be a real locally convex topological vector space. For $x \in X$, \(\delta \phi(x) \subseteq X'\), possibly \(\emptyset\). This subset of $X'$ is defined by $y^* \in \delta \phi(x)$ means for all $z \in X$,

$$y^* (z - x) \leq \phi(z) - \phi(x).$$

Also $x \in \delta \phi^* (y^*)$ means that for all $z^* \in X'$,

$$(z^* - y^*) (x) \leq \phi^* (z^*) - \phi^* (y^*).$$

We define \(\text{dom}(\delta \phi) \equiv \{ x : \delta \phi(x) \neq \emptyset \}\).

The subgradient is an attempt to generalize the derivative. For example, a function may have a subgradient but fail to be differentiable at some point. A good example is $f(x) = |x|$. At $x = 0$, this function fails to have a derivative but it does have a subgradient. In fact, $\delta f(0) = [-1, 1]$.

To begin with consider the question of existence of the subgradient of a convex function. There is a very simple criterion for existence. It is essentially that the subgradient is nonempty at every point of the interior of the domain of $\phi$. First recall Problems 6 and 7 on Page 65 which say the interior of a convex set is convex and if nonempty, then every point of the convex set can be obtained as the limit of a sequence of points of the interior.

**Theorem 10.4.2** Let $\phi : X \to (-\infty, \infty]$ be convex and suppose for some $u \in \text{dom}(\phi)$, $\phi$ is continuous. Then $\delta \phi(x) \neq \emptyset$ for all $x \in \text{int}(\text{dom}(\phi))$. Thus

$$\text{dom}(\delta \phi) \supseteq \text{int}(\text{dom}(\phi)).$$

**Proof:** Let $x_0 \in \text{int}(\text{dom}(\phi))$ and let

$$A \equiv \{(x_0, \phi(x_0))\}, B \equiv \text{epi}(\phi) \cap X \times \mathbb{R}.$$

Then $A$ and $B$ are both nonempty and convex. Recall epi($\phi$) can contain a point like $(x, \infty)$. Since $\phi$ is continuous at $u \in \text{dom}(\phi)$,

$$(u, \phi(u) + 1) \in \text{int}(\text{epi}\phi \cap X \times \mathbb{R}).$$
Thus \( \text{int}(B) \neq \emptyset \) and also \( \text{int}(B) \cap A = \emptyset \). By Problem 10 on Page 65, \( \text{int}(B) \) is convex and so by Problem 12 on Page 66 there exists \( x^* \in X' \) and \( \beta \in \mathbb{R} \) such that
\[
(x^*, \beta) \neq (0, 0) \quad (10.4.17)
\]
and for all \( (x, a) \in \text{int} B \),
\[
x^* (x) + \beta a > x^* (x_0) + \beta \phi (x_0). \quad (10.4.18)
\]

From Problem 1 on Page 62, whenever \( x \in \text{dom} (\phi) \),
\[
x^* (x) + \beta \phi (x) \geq x^* (x_0) + \beta \phi (x_0).
\]

If \( \beta = 0 \), this would mean \( x^*(x - x_0) \geq 0 \) for all \( x \in \text{dom} (\phi) \). Since \( x_0 \in \text{int} (\text{dom} (\phi)) \), this implies \( x^* = 0 \), contradicting 10.4.17. If \( \beta < 0 \), apply 10.4.18 to the case when \( a = \phi (x_0) + 1 \) and \( x = x_0 \) to obtain a contradiction. It follows \( \beta > 0 \) and so
\[
\phi (x) - \phi (x_0) \geq - \frac{x^*}{\beta} (x - x_0)
\]
which says \(-x^*/\beta \in \delta \phi (x_0)\). 

**Definition 10.4.3** Let \( \phi : X \to (-\infty, \infty] \) be some function, not necessarily convex but satisfying \( \phi (y) < \infty \) for some \( y \in X \). Define \( \phi^* : X' \to (-\infty, \infty] \) by
\[
\phi^* (x^*) \equiv \sup \{ x^* (y) - \phi (y) : y \in X \}.
\]

This function, \( \phi^* \), defined above, is called the conjugate function of \( \phi \) or the polar of \( \phi \). Note \( \phi^* (x^*) \neq -\infty \) because \( \phi (y) < \infty \) for some \( y \).

**Theorem 10.4.4** Let \( X \) be a real Banach space. Then \( \phi^* \) is convex and l.s.c.

**Proof:** Let \( \lambda \in [0, 1] \). Then
\[
\phi^* (\lambda x^* + (1 - \lambda) y^*) = \sup \{ (\lambda x^* + (1 - \lambda) y^*) (y) - \phi (y) : y \in X \}
\]

\[
\sup \{ \lambda (x^* (y) - \phi (y)) + (1 - \lambda) (y^* (y) - \phi (y)) : y \in X \}
\]

\[
\leq \lambda \phi^* (x^*) + (1 - \lambda) \phi^* (y^*).
\]

It remains to show the function is l.s.c. Consider \( f_y (x^*) \equiv x^* (y) - \phi (y) \). Then \( f_y \) is obviously convex. Also
\[
\text{epi} (\phi^*) = \cap_{y \in X} \text{epi} (f_y).
\]

Therefore, if we can show \( \text{epi} (f_y) \) is closed, we are done. If \( (x^*, a) \notin \text{epi} (f_y) \), then \( a < x^* (y) - \phi (y) \) and, by continuity, this condition will persist for all \( y^* \) near \( x^* \). Thus \( \text{epi} (f_y) \) is closed and this proves the theorem.

Note this theorem holds with no change in the proof if \( X \) is only a locally convex topological vector space and \( X' \) is given the weak * topology.
We define $\phi^{**}$ on $X$ by
\[
\phi^{**}(x) \equiv \sup \{ x^*(x) - \phi^*(x^*), x^* \in X' \}.
\]

**Theorem 10.4.6** $\phi^{**}(x) \leq \phi(x)$ for all $x$ and if $\phi$ is convex and l.s.c., $\phi^{**}(x) = \phi(x)$ for all $x \in X$.

**Proof:**
\[
\phi^{**}(x) \equiv \sup \{ x^*(x) - \sup \{ x^*(y) - \phi(y) : y \in X \} : x^* \in X' \}
\leq \sup \{ x^*(x) - (x^*(x) - \phi(x)) \} = \phi(x).
\]

Next suppose $\phi$ is convex and l.s.c. so its epigraph is closed. If $\phi^{**}(x_0) < \phi(x_0)$, then $(x_0, \phi^{**}(x_0))$ is not in the closed set $\text{epi}(\phi)$. Separation theorems imply there exists $(x'_0, \beta) \in (X \times \mathbb{R})'$ such that for all $y \in X$,
\[
x'_0(y) + \beta \phi(y) < c' < x'_0(x_0) + \beta \phi^{**}(x_0).
\]

Letting $y = x_0$, it follows $\beta < 0$. Then dividing by $\beta$ and renaming $x'_0$ with $x'_0/\beta$, it follows for all $y \in X$, an inequality of the following form holds.
\[
x'_0(y) + \phi(y) > c > x'_0(x_0) + \phi^{**}(x_0).
\]

It follows from the definition of $\phi^{**}$ and the fact that $\phi^{**}(x) = \phi(x)$ for all $x$ and this proves the theorem.

The following corollary is descriptive of the situation just discussed. It says that to find $\text{epi}(\phi^{**})$ it suffices to take the intersection of all closed convex sets which contain $\text{epi}(\phi)$.

**Corollary 10.4.7** $\text{epi}(\phi^{**})$ is the smallest closed convex set containing $\text{epi}(\phi)$.

**Proof:** $\text{epi}(\phi^{**}) \supseteq \text{epi}(\phi)$ from Theorem 10.4.4. Also $\text{epi}(\phi^{**})$ is closed by the proof of Theorem 10.4.5. Suppose $\text{epi}(\phi) \subseteq K \subseteq \text{epi}(\phi^{**})$ and $K$ is convex and closed. Let
\[
\psi(x) \equiv \min \{ a : (x, a) \in K \}.
\]

$\{a : (x, a) \in K\}$ is a closed subset of $(-\infty, \infty)$ so the minimum exists.) $\psi$ is also a convex function with $\text{epi}(\psi) = K$. To see $\psi$ is convex, let $\lambda \in [0, 1]$. Then, by the convexity of $K$,
\[
\lambda(x, \psi(x)) + (1 - \lambda)(y, \psi(y))
\]
It follows from the definition of \( \psi \) that

\[
\psi (\lambda x + (1 - \lambda) y) \leq \lambda \psi (x) + (1 - \lambda) \psi (y).
\]

Then

\[
\phi^{**} \leq \psi \leq \phi
\]

and so from the definitions,

\[
\phi^{**} \geq \psi^{*} \geq \phi^{*}
\]

which implies from the definitions and Theorem 10.4.6 that

\[
\phi^{**} = \phi^{**} \leq \psi^{**} = \psi \leq \phi^{**}.
\]

Therefore, \( \psi = \phi^{**} \) and \( \text{epi} (\phi^{**}) \) is the smallest closed convex set containing \( \text{epi} (\phi) \) as claimed.

There is an interesting symmetry which relates \( \delta \phi, \delta \phi^{*}, \phi, \) and \( \phi^{*} \).

**Theorem 10.4.8** Suppose \( \phi \) is convex, l.s.c. (lower semicontinuous or in other words having a closed epigraph), and proper. Then

\( y^{*} \in \delta \phi (x) \) if and only if \( x \in \delta \phi^{*} (y^{*}) \)

where this last expression means

\[
(z^{*} - y^{*}) (x) \leq \phi^{*} (z^{*}) - \phi^{*} (y^{*})
\]

for all \( z^{*} \) and in this case,

\[
y^{*} (x) = \phi^{*} (y^{*}) + \phi (x).
\]

**Proof:** If \( y^{*} \in \delta \phi (x) \) then \( y^{*} (z - x) \leq \phi (z) - \phi (x) \) and so

\[
y^{*} (z) - \phi (z) \leq y^{*} (x) - \phi (x)
\]

for all \( z \in X \). Therefore,

\[
\phi^{*} (y^{*}) \leq y^{*} (x) - \phi (x) \leq \phi^{*} (y^{*}).
\]

Hence

\[
y^{*} (x) = \phi^{*} (y^{*}) + \phi (x). \tag{10.4.20}
\]

Now if \( z^{*} \in X' \) is arbitrary, shows

\[
(z^{*} - y^{*}) (x) = z^{*} (x) - y^{*} (x) = z^{*} (x) - \phi (x) - \phi^{*} (y^{*}) \leq \phi^{*} (z^{*}) - \phi^{*} (y^{*})
\]

and this shows \( x \in \delta \phi^{*} (y^{*}) \).

Now suppose \( x \in \delta \phi^{*} (y^{*}) \). Then for \( z^{*} \in X' \),

\[
(z^{*} - y^{*}) (x) \leq \phi^{*} (z^{*}) - \phi^{*} (y^{*})
\]
so
\[ z^* (x) - \phi^* (z^*) \leq y^* (x) - \phi^* (y^*) \]
and so, taking sup over all \( z^* \), and using Theorem 10.4.6,
\[ \phi^{**} (x) = \phi (x) \leq y^* (x) - \phi^* (y^*) \leq \phi^{**} (x). \]
Thus
\[ y^* (x) = \phi^* (y^*) + \phi^{**} (x) = \phi^* (y^*) + \phi (x) \geq y^* (z) - \phi (z) + \phi (x) \]
for all \( z \in X \) and this implies for all \( z \in X \),
\[ \phi (z) - \phi (x) \geq y^* (z - x) \]
so \( y^* \in \delta \phi (x) \) and this proves the theorem.

Definition 10.4.9 If \( X \) is a Banach space define \( u \in H^1 (0,T;X) \) if there exists \( g \in L^2 (0,T;X) \) such that
\[ u (t) = u (0) + \int_0^t g (s) \, ds \]
When this occurs define \( u' (\cdot) \equiv g (\cdot) \).

The next Lemma is quite interesting for its own sake but it is also used in the next theorem.

Lemma 10.4.10 Suppose \( g \in L^2 (0,T;X) \). Then as \( h \to 0 \),
\[ \frac{1}{h} \int_{(\cdot)^\perp}^{(\cdot)^\perp + h} g (s) \, ds \chi_{[0,T-h]} (\cdot) \to g \]
in \( L^2 (0,T;X) \).

Proof: Let
\[ \tilde{g} (u) \equiv \begin{cases} g (u) & \text{if } u \in [0,T], \\ 0 & \text{if } u \notin [0,T] \end{cases}, \quad \phi_h (r) \equiv \frac{1}{h} \chi_{[-h,0]} (r). \]
Thus \( \tilde{g} \in L^2 (\mathbb{R};X) \) and
\[ \tilde{g} * \phi_h (t) \equiv \int_{\mathbb{R}} \tilde{g} (t-s) \phi_h (s) \, ds. \]
Then
\[ \| \tilde{g} * \phi_h - \tilde{g} \|_{L^2 (\mathbb{R};X)} \leq \left( \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \| \tilde{g} (t) - \tilde{g} (t-s) \|_X \phi_h (s) \, ds \right)^2 \, dt \right)^{1/2}. \]
which by Minkowski’s inequality for integrals is no larger than

\[
\leq \int_{\mathbb{R}} \phi_h(s) \left( \int_{\mathbb{R}} \| g(t) - \tilde{g}(t-s) \|^2_X dt \right)^{1/2} ds
\]

\[
= \frac{1}{h} \int_{-h}^{0} \left( \int_{\mathbb{R}} \| \tilde{g}(t) - \tilde{g}(t-s) \|^2_X dt \right)^{1/2} ds < \frac{1}{h} \int_{-h}^{0} \varepsilon ds = \varepsilon
\]

whenever \( h \) is small enough. This follows from continuity of translation in \( L^2(\mathbb{R}; X) \), a consequence of the regularity of the measure. Thus, \( \tilde{g} \ast \phi_h \rightarrow \tilde{g} \) in \( L^2(\mathbb{R}; X) \). Now

\[
\tilde{g} \ast \phi_h(t) - \frac{1}{h} \int_{t}^{t+h} g(s) ds \mathcal{X}_{[0,T-h]}(t) = \begin{cases} 
0 & \text{if } t \in [0, T-h] \\
\frac{1}{h} \int_{t}^{t+h} g(s) ds & \text{if } t \notin [0, T-h]
\end{cases}
\]

and therefore,

\[
\left\| \tilde{g} \ast \phi_h(\cdot) - \frac{1}{h} \int_{t}^{t+h} g(s) ds \mathcal{X}_{[0,T-h]}(\cdot) \right\|_{L^2(\mathbb{R}; X)} = 
\left( \int_{-h}^{0} \left\| \int_{t}^{t+h} g(u) du \right\|^2_X dt \right)^{1/2} + \left( \int_{T-h}^{T} \left\| \int_{t}^{t+h} g(u) du \right\|^2_X dt \right)^{1/2}
\leq \frac{1}{h} \left( \int_{-h}^{0} \left( \int_{-h}^{h} \| g(u) \|^2_X du \right)^2 dt \right)^{1/2} + \frac{1}{h} \left( \int_{T-h}^{T} \left( \int_{-h}^{h} \| g(u) \|^2_X du \right)^2 dt \right)^{1/2}
\]

which by Minkowski’s inequality for integrals is no larger than

\[
\leq \frac{1}{h} \int_{-h}^{h} \left( \int_{-h}^{h} |g(u)|^2_X du \right)^{1/2} + \frac{1}{h} \int_{T-h}^{T} \left( \int_{-h}^{h} |g(u)|^2_X du \right)^{1/2} du
\]

\[
\leq \frac{1}{\sqrt{h}} \int_{-h}^{h} \| g(u) \|_X du + \frac{1}{\sqrt{h}} \int_{T-h}^{T} \| g(u) \|_X du
\]

\[
\leq \frac{(2h)^{1/2}}{\sqrt{h}} \left( \int_{-h}^{h} |g(u)|^2_X du \right)^{1/2} + \frac{(2h)^{1/2}}{\sqrt{h}} \left( \int_{T-h}^{T} |g(u)|^2_X du \right)^{1/2},
\]
an expression which converges to zero as \( h \to 0 \) because of the dominated convergence theorem and the fact that \( \| g \|_X \) is in \( L^1(\mathbb{R}; X) \). Therefore,

\[
\frac{1}{h} \int_{(\cdot)}^{(\cdot)+h} g(s) \, ds \mathcal{X}_{[0,T-h]}(\cdot) \to \tilde{g}
\]

in \( L^2(\mathbb{R}; X) \) and consequently in \( L^2(0,T; X) \) as well. But \( \tilde{g} = g \) on \([0,T] \). ■

The following theorem is a form of the chain rule in which the derivative is replaced by the subgradient.

**Theorem 10.4.11** Suppose \( u \in H^1(0,T; X) \), \( z \in L^2(0,T; X') \), and \( z(t) \in \delta \phi(u(t)) \) a.e. \( t \in [0,T] \). Then the function, \( t \to \phi(u(t)) \) is in \( L^1(0,T) \) and its weak derivative equals \( z(u') \).

**Proof:** Modify \( u \) on a set of measure zero such that \( \delta \phi(u(t)) \neq \emptyset \) for all \( t \). Next modify \( z \) on a set of measure zero such that for \( \tilde{u} \) and \( \tilde{z} \) the modified functions, \( \tilde{z}(t) \in \delta \phi(\tilde{u}(t)) \) for all \( t \). First I claim \( t \to \phi(\tilde{u}(t)) \) is in \( L^1(0,T) \). Pick \( t_0 \in [0,T] \) and let

\[
\tilde{z}(t_0) \in \delta \phi(\tilde{u}(t_0)).
\]

Then for \( t \in [0,T] \),

\[
\tilde{z}(t_0) (\tilde{u}(t) - \tilde{u}(t_0)) + \phi(\tilde{u}(t_0)) \leq \phi(\tilde{u}(t)) \leq \tilde{z}(t) (\tilde{u}(t) - \tilde{u}(t_0)) + \phi(\tilde{u}(t_0))
\]

since

\[
\tilde{z}(t_0) (\tilde{u}(t) - \tilde{u}(t_0)) \leq \phi(\tilde{u}(t_0)) - \phi(\tilde{u}(t))
\]

and

\[
\tilde{z}(t_0) (\tilde{u}(t_0) - \tilde{u}(t)) \leq \phi(\tilde{u}(t_0)) - \phi(\tilde{u}(t)).
\]

Then \( t \to \phi(\tilde{u}(t)) \) is in \( L^1(0,T) \) since \( \tilde{z} \in L^2(0,T; X') \) and \( \tilde{u} \in L^2(0,T; X) \). Also, for \( t \in [0,T-h] \),

\[
\mathcal{X}_{[0,T-h]}(\tilde{z}) \left( \frac{\tilde{u}(t+h) - \tilde{u}(t)}{h} \right) \leq \mathcal{X}_{[0,T-h]}(\tilde{z}) \left( \frac{\tilde{u}(t+h) - \tilde{u}(t)}{h} \right)
\]

\[
\leq \mathcal{X}_{[0,T-h]}(\tilde{z}) (t+h) \left( \frac{\tilde{u}(t+h) - \tilde{u}(t)}{h} \right)
\]

Now \( \mathcal{X}_{[0,T-h]}(\cdot) \tilde{z}(\cdot+h) \to z(\cdot) \) in \( L^2(0,T; X') \) by continuity of translation. Also,

\[
\mathcal{X}_{[0,T-h]}(\cdot) \left( \frac{\tilde{u}(\cdot+h) - \tilde{u}(\cdot)}{h} \right) = \mathcal{X}_{[0,T-h]}(\cdot) \left( \frac{u(\cdot+h) - u(\cdot)}{h} \right)
\]

\[
= \mathcal{X}_{[0,T-h]}(\cdot) \frac{1}{h} \int_{(\cdot)}^{(\cdot)+h} u'(s) \, ds
\]

in \( L^2(0,T; X) \) and so by Lemma \( \text{lem} \)

\[
\mathcal{X}_{[0,T-h]}(\cdot) \left( \frac{\tilde{u}(\cdot+h) - \tilde{u}(\cdot)}{h} \right) \to z(u')
\]
in $L^1(0,T)$.

It follows from the definition of weak derivatives that in the sense of weak derivatives,

$$\frac{d}{dt}(\phi(u(\cdot))) = z(u') \in L^1(0,T).$$

Note that by the problem 4 on Page 132 this implies that for a.e. $t \in [0,T]$, $\phi(u(t))$ is equal to a continuous function, $\phi \circ u$, and that

$$(\phi \circ u)(t) - (\phi \circ u)(0) = \int_0^t z(u') ds.$$

There are other rules of calculus which have a generalization to subgradients. The following theorem is on such a generalization. It generalizes the theorem which states that the derivative of a sum equals the sum of the derivatives.

**Theorem 10.4.12** Let $\phi_1$ and $\phi_2$ be convex, l.s.c. and proper having values in $(-\infty, \infty)$. Then

$$\delta(\lambda \phi_i)(x) = \lambda \delta \phi_i(x), \delta(\phi_1 + \phi_2)(x) \supseteq \delta \phi_1(x) + \delta \phi_2(x) \quad (10.4.22)$$

if $\lambda > 0$. If there exists $\overline{x} \in \text{dom}(\phi_1) \cap \text{dom}(\phi_2)$ and $\phi_1$ is continuous at $\overline{x}$ then for all $x \in X$,

$$\delta(\phi_1 + \phi_2)(x) = \delta \phi_1(x) + \delta \phi_2(x). \quad (10.4.23)$$

**Proof:** (10.4.22) is obvious so we only need to show (10.4.23). Suppose $\overline{x}$ is as described. It is clear (10.4.22) holds whenever $x \notin \text{dom}(\phi_1) \cap \text{dom}(\phi_2)$ since then both sides equal $\emptyset$. Therefore, assume

$$x \in \text{dom}(\phi_1) \cap \text{dom}(\phi_2)$$

in what follows. Let $x^* \in \delta(\phi_1 + \phi_2)(x)$. Is $x^*$ the sum of an element of $\delta \phi_1(x)$ and $\delta \phi_2(x)$? Does there exist $x_1^*$ and $x_2^*$ such that for every $y$,

$$x^*(y - x) = x_1^*(y - x) + x_2^*(y - x) \leq \phi_1(y) - \phi_1(x) + \phi_2(y) - \phi_2(x)?$$

If so, then

$$\phi_1(y) - \phi_1(x) - x^*(y - x) \geq \phi_2(x) - \phi_2(y).$$

Define

$$C_1 \equiv \{(y, a) \in X \times \mathbb{R} : \phi_1(y) - \phi_1(x) - x^*(y - x) \leq a\},$$

$$C_2 \equiv \{(y, a) \in X \times \mathbb{R} : a \leq \phi_2(x) - \phi_2(y)\}.$$

I will show $\text{int}(C_1) \cap C_2 = \emptyset$ and then by separation theorems there exists an element of $X'$ which does something interesting.
10.4. SUBGRADIENTS

Both $C_1$ and $C_2$ are convex and nonempty. $C_1$ is nonempty because it contains $(\overline{x}, \phi_1(\overline{x}) - \phi_1(x) - x^*(\overline{x} - x))$ since

$$\phi_1(\overline{x}) - \phi_1(x) - x^*(\overline{x} - x) \leq \phi_1(\overline{x}) - \phi_1(x) - x^*(\overline{x} - x)$$

$C_2$ is also nonempty because it contains $(\overline{x}, \phi_2(x) - \phi_2(\overline{x}))$ since

$$\phi_2(x) - \phi_2(\overline{x}) \leq \phi_2(x) - \phi_2(\overline{x})$$

In addition to this,

$$(\overline{x}, \phi_1(\overline{x}) - x^*(\overline{x} - x) - \phi_1(x) + 1) \in \text{int}(C_1)$$

due to the assumed continuity of $\phi_1$ at $\overline{x}$ and so $\text{int}(C_1) \neq \emptyset$. If $(y, a) \in \text{int}(C_1)$ then

$$\phi_1(y) - x^*(y - x) - \phi_1(x) \leq a - \varepsilon$$

evertheless $\varepsilon$ is small enough. Therefore, if $(y, a)$ is also in $C_2$, the assumption that $x^* \in \delta(\phi_1 + \phi_2)(x)$ implies

$$a - \varepsilon \geq \phi_1(y) - x^*(y - x) - \phi_1(x) \geq \phi_2(x) - \phi_2(y) \geq a,$$

a contradiction. Therefore $(C_1) \cap C_2 = \emptyset$ and so by separation theorems, there exists $(w^*, \beta) \in X' \times \mathbb{R}$ with

$$(w^*, \beta) \neq (0, 0), \quad (10.4.24)$$

and

$$w^*(y) + \beta a \geq w^*(y_1) + \beta a_1, \quad (10.4.25)$$

whenever $(y, a) \in C_1$ and $(y_1, a_1) \in C_2$.

**Claim:** $\beta > 0$.

**Proof of claim:** If $\beta < 0$ let

$$a = \phi_1(\overline{x}) - x^*(\overline{x} - x) - \phi_1(x) + 1,$$

$$a_1 = \phi_2(x) - \phi_2(\overline{x}), \text{ and } y = y_1 = \overline{x}.$$

Then from

$$\beta(\phi_1(\overline{x}) - x^*(\overline{x} - x) - \phi_1(x) + 1) \geq \beta(\phi_2(x) - \phi_2(\overline{x})).$$

Dividing by $\beta$ yields

$$\phi_1(\overline{x}) - x^*(\overline{x} - x) - \phi_1(x) + 1 \leq \phi_2(x) - \phi_2(\overline{x})$$

and so

$$\phi_1(\overline{x}) + \phi_2(\overline{x}) - (\phi_1(x) + \phi_2(x)) + 1 \leq x^*(\overline{x} - x)$$

$$\leq \phi_1(\overline{x}) + \phi_2(\overline{x}) - (\phi_1(x) + \phi_2(x)),$$

a contradiction. Therefore, $\beta \geq 0$.
Now suppose \( \beta = 0 \). Letting
\[
a = \phi_1(x) - x^*(x - x) - \phi_1(x) + 1,
\]
and so there exists an open set \( U \) containing 0 and \( \eta > 0 \) such that
\[
(x, a) \in \operatorname{int}(C_1)
\]
and \( (\pi, z, a) \in C_1 \) for \( z \in U \). Therefore, \( 10.4.25 \) applied to \( (x, a) \in C_1 \) and \( (x, \phi(x) - \phi(x)) \in C_2 \) for \( \phi \) yields
\[
w^*(x) \geq w^*(x)
\]
for all \( x \in U \). Hence \( w^*(x) = 0 \) on \( U \) which implies \( w^* = 0 \), contradicting \( 10.4.24 \). This proves the claim.

Now with the claim, it follows \( \beta > 0 \) and so, letting
\[
z^* = w^*/\beta,
\]
then
\[
z^*(y) + a \geq z^*(y_1) + a_1
\]
whenever \( (y, a) \in C_1 \) and \( (y_1, a_1) \in C_2 \). In particular,
\[
(y, \phi_1(y) - \phi_1(x) - x^*(y - x)) \in C_1
\]
because
\[
\phi_1(y) - \phi_1(x) - x^*(y - x) \leq \phi_1(y) - x^*(y - x) - \phi_1(x)
\]
and
\[
(y_1, \phi_2(x) - \phi_2(y_1)) \in C_2.
\]
by similar reasoning so letting \( y = x \),
\[
z^*(x) + \left( \phi_1(x) - x^*(x - x) - \phi_1(x) \right) \geq z^*(y_1) + \phi_2(x) - \phi_2(y_1).
\]
Therefore,
\[
z^*(y_1 - x) \leq \phi_2(y_1) - \phi_2(x)
\]
for all \( y_1 \) and so \( z^* \in \delta \phi_2(x) \). Now let \( y_1 = x \) in \( 10.4.25 \) and using \( 10.4.26 \) and \( 10.4.27 \), it follows
\[
z^*(y) + \phi_1(y) - x^*(y - x) - \phi_1(x) \geq z^*(x)
\]
and
\[
\phi_1(x) - \phi_1(x) \geq x^*(y - x) - z^*(y - x)
\]
for all \( x \) and so \( x^* - z^* \in \delta \phi_1(x) \) so \( x^* = z^* + (x^* - z^*) \in \delta \phi_2(x) + \delta \phi_1(x) \) and this proves the theorem.
Next is a very important example known as the duality map from a Banach space to its dual space. Before doing this, consider a Hilbert space $H$. Define a map $R$ from $H$ to $H'$, called the Riesz map, by the rule

$$R(x)(y) \equiv (y,x).$$

By the Riesz representation theorem, this map is onto and one to one with the properties

$$R(x)(x) = ||x||^2, \text{ and } ||Rx||^2 = ||x||^2.$$

The duality map from a Banach space to its dual is an attempt to generalize this notion of Riesz map to an arbitrary Banach space.

**Definition 10.4.13** For $X$ a Banach space define $F : X \to \mathcal{P}(X')$ by

$$F(x) \equiv \left\{ x^* \in X' : x^*(x) = ||x||^2, ||x^*|| \leq ||x|| \right\}. \quad (10.4.29)$$

**Lemma 10.4.14** With $F(x)$ defined as above, it follows that

$$F(x) = \left\{ x^* \in X' : x^*(x) = ||x||^2, ||x^*|| = ||x|| \right\}$$

and $F(x)$ is a closed, nonempty, convex subset of $X'$.

**Proof:** If $x^*$ is in the set described in (10.4.29),

$$x^* \left( \frac{x}{||x||} \right) = ||x||$$

and so $||x^*|| \geq ||x||$. Therefore

$$x^* \in \left\{ x^* \in X' : x^*(x) = ||x||^2, ||x^*|| = ||x|| \right\}.$$

This shows this set and the set of (10.4.29) are equal. It is also clear the set of (10.4.29) is closed and convex. It only remains to show this set is nonempty.

Define $f : \mathbb{R}x \to \mathbb{R}$ by $f(\alpha x) = \alpha ||x||^2$. Then the norm of $f$ on $\mathbb{R}x$ is $||x||$ and $f(x) = ||x||^2$. By the Hahn Banach theorem, $f$ has an extension to all of $X x^*$, and this extension is in the set of (10.4.29), showing this set is nonempty as required.

The next theorem shows this duality map is the subgradient of $\frac{1}{2} ||x||^2$.

**Theorem 10.4.15** For $X$ a real Banach space, let $\phi(x) \equiv \frac{1}{2} ||x||^2$. Then $F(x) = \delta\phi(x)$.

**Proof:** Let $x^* \in F(x)$. Then

$$\langle x^*, y - x \rangle = \langle x^*, y \rangle - \langle x^*, x \rangle \leq ||x|| ||y|| - ||x||^2 \leq \frac{1}{2} ||y||^2 - \frac{1}{2} ||x||^2.$$
408 CHAPTER 10. MAXIMAL MONOTONE OPERATORS, HILBERT SPACE

This shows $F(x) \subseteq \delta \phi(x)$.

Now let $x^* \in \delta \phi(x)$. Then for all $t \in \mathbb{R}$,

$$
\langle x^*, ty \rangle = \langle x^*, (ty + x) - x \rangle \leq \frac{1}{2} (||x + ty||^2 - ||x||^2).
$$

(10.4.30)

Now if $t > 0$, divide both sides by $t$. This yields

$$
\langle x^*, y \rangle \leq \frac{1}{2t} \left( ||x||^2 + ||ty||^2 - ||x||^2 \right)
$$

(10.4.31)

Next suppose $t = -s$, where $s > 0$ in (10.4.31). Then, since when you divide by a negative, you reverse the inequality, for $s > 0$

$$
\langle x^*, y \rangle \geq \frac{1}{2s} \left[ ||x||^2 - ||x - sy||^2 \right] \geq \frac{1}{2s} \left[ (||x||^2 - 2||x - sy|| ||sy|| + ||sy||^2 - ||x - sy||^2) \right].
$$

(10.4.32)

$$
= \frac{1}{2s} \left[ -2||x - sy|| ||sy|| + ||sy||^2 \right].
$$

(10.4.33)

Taking a limit as $s \to 0$ yields

$$
\langle x^*, y \rangle \geq -||x|| ||y||.
$$

(10.4.34)

It follows from (10.4.32) and (10.4.33) that

$$
||\langle x^*, y \rangle|| \leq ||x|| ||y||
$$

and that, therefore, $||x^*|| \leq ||x||$ and $||\langle x^*, x \rangle|| \leq ||x||^2$. Now return to (10.4.32) and let $y = x$. Then

$$
\langle x^*, x \rangle \geq \frac{1}{2s} \left[ -2||x - sx|| ||sx|| + ||sx||^2 \right] = -||x||^2 (1 - s) + s||x||^2
$$

Letting $s \to 1$,

$$
\langle x^*, x \rangle \geq ||x||^2.
$$

Since it was already shown that $||\langle x^*, x \rangle|| \leq ||x||^2$, this shows $\langle x^*, x \rangle = ||x||^2$ and also $||x^*|| \leq ||x||$. Thus

$$
||x^*|| \geq \left\langle x^*, \frac{x}{||x||} \right\rangle = ||x||
$$

so in fact $x^* \in F(x)$. ■

The next result gives conditions under which the subgradient is onto. This means that if $y^* \in X^*$, then there exists $x \in X$ such that $y^* \in \delta \phi(x)$.
Theorem 10.4.16 Suppose $X$ is a reflexive Banach space and suppose $\phi : X \to (-\infty, \infty]$ is convex, proper, l.s.c., and for all $y^* \in X'$, $x \to \phi (x) - y^*(x)$ is coercive. Then $\partial \phi$ is onto.

Proof: The function $x \to \phi (x) - y^*(x) \equiv \psi (x)$ is convex, proper, l.s.c., and coercive. Let

$$\lambda = \inf \{ \phi (x) - y^* (x) : x \in X \}$$

and let $\{ x_n \}$ be a minimizing sequence satisfying

$$\lambda = \lim_{n \to \infty} \phi (x_n) - y^* (x_n)$$

By coercivity,

$$\lim_{||x|| \to \infty} \phi (x) - y^* (x) = \infty$$

and so this minimizing sequence is bounded. By the Eberlein Smulian theorem, Problem 14 on Page 66, there is a weakly convergent subsequence $x_{nk} \to x$. By Problem 11 on Page 66, $\phi$ is also weakly lower semicontinuous. Therefore,

$$\lambda = \phi (x) - y^* (x) \leq \liminf_{k \to \infty} \phi (x_{nk}) - y^* (x_{nk}) = \lambda$$

so there exists $x$ which minimizes $x \to \phi (x) - y^* (x) \equiv \psi (x)$. Therefore, $0 \in \partial \psi (x)$ because

$$\psi (y) - \psi (x) \geq 0 = 0 (y - x)$$

by Theorem 10.4.12, $0 \in \partial \psi (x) = \delta \phi (x) - y^*$ and this proves the theorem.

Corollary 10.4.17 Suppose $X$ is a reflexive Banach space and $\phi : X \to (-\infty, \infty]$ is convex, proper, and l.s.c. Then for each $y^* \in X'$ there exist $x \in X$, $x_1^* \in F (x)$, and $x_2^* \in \partial \phi (x)$ such that

$$y^* = x_1^* + x_2^*$$

Proof: Apply Theorem 10.4.16 to the convex function $\frac{1}{2} ||x||^2 + \phi (x)$ and use Theorems 10.4.14 and 10.4.15.

10.4.2 Hilbert Space

In this section the subgradients are of a slightly different form and defined on a subset of $H$, a real Hilbert space. In Hilbert space the duality map is just the Riesz map defined earlier by

$$Rx (y) \equiv (y, x).$$

Definition 10.4.18 $\dom (\partial \phi) \equiv \dom (\delta \phi)$ and for $x \in \dom (\partial \phi)$,

$$\partial \phi (x) \equiv R^{-1} \delta \phi (x).$$

Thus $y \in \partial \phi (x)$ if and only if for all $z \in H$,

$$Ry (z - x) = (y, z - x) \leq \phi (z) - \phi (x).$$
Recall the definition of a maximal monotone operator.

**Definition 10.4.19** A mapping $A : D(A) \subseteq H \to P(H)$ is called monotone if whenever $y_i \in Ax_i$,

$$(y_1 - y_2, x_1 - x_2) \geq 0.$$  

A monotone map is called maximal monotone if whenever $z \in H$, there exists $x \in D(A)$ and $y \in A(x)$ such that $z = y + x$. Put more simply, $I + A$ maps $D(A)$ onto $H$.

The following lemma states, among other things, that when $\phi$ is a convex, proper, l.s.c. function defined on a Hilbert space, $\partial \phi$ is maximal monotone.

**Lemma 10.4.20** If $\phi$ is a convex, proper, l.s.c. function defined on a Hilbert space, then $\partial \phi$ is maximal monotone and $(I + \partial \phi)^{-1}$ is a Lipschitz continuous map from $H$ to $\text{dom}(\partial \phi)$ having Lipschitz constant 1.

**Proof:** Let $y \in H$. Then $Ry \in H'$ and by Corollary 10.4.17, there exists $x \in \text{dom}(\delta \phi)$ such that $Rx + \delta \phi(x) \ni Ry$. Multiplying by $R^{-1}$ we see $y \in x + \delta \phi(x)$. This shows $I + \partial \phi$ is onto. If $y_i \in \partial \phi(x_i)$, then $Ry_i \in \delta \phi(x_i)$ and so by the definition of subgradients,

$$(y_1 - y_2, x_1 - x_2) = R(y_1 - y_2)(x_1 - x_2) = Ry_1(x_1 - x_2) - Ry_2(x_1 - x_2) \geq \phi(x_1) - \phi(x_2) = 0$$

showing $\partial \phi$ is monotone. Now suppose $x_i \in (I + \partial \phi)^{-1}(y)$. Then $y - x_i \in \partial \phi(x_i)$ and by monotonicity of $\partial \phi$,

$$- |x_1 - x_2|^2 = (y - x_1 - (y - x_2), x_1 - x_2) \geq 0$$

and so $x_1 = x_2$. Thus $(I + \partial \phi)^{-1}$ is well defined. If $x_i = (I + \partial \phi)^{-1}(y_i)$, then by the monotonicity of $\partial \phi$,

$$(y_1 - x_1 - (y_1 - x_2), x_1 - x_2) \geq 0$$

and so

$$|y_1 - y_2| |x_1 - x_2| \geq |x_1 - x_2|^2$$

which shows

$$|(I + \partial \phi)^{-1}(y_1) - (I + \partial \phi)^{-1}(y_2)| \leq |y_1 - y_2|.$$  

Here is another proof.

**Lemma 10.4.21** Let $\phi$ be convex, proper and lower semicontinuous on $X$ a reflexive Banach space having strictly convex norm, then for each $\alpha > 0$,

$$I + \alpha \partial \phi$$

is onto.
Proof: By separation theorems applied to the epigraph of $\phi$, and since $\phi$ is proper, there exists $w^*$ such that

$$(w^*, x) + b \leq \alpha \phi(x)$$

for all $x$. Pick $y \in H$. Then consider

$$\frac{1}{2} |y - x|^2 + \alpha \phi(x)$$

This functional of $x$ is bounded below by

$$\frac{1}{2} |y - x|^2 + (w^*, x) + b$$

Thus it is clearly coercive. Hence any minimizing sequence has a weakly convergent subsequence. It follows from lower semicontinuity that there exists $x_0$ which minimizes this functional. Hence, if $z \neq x_0$,

$$0 \leq \frac{1}{2} |y - z|^2 + \alpha \phi(z) - \left( \frac{1}{2} |y - x_0|^2 + \alpha \phi(x_0) \right)$$

Then writing $|y - z|^2 = |y - x_0|^2 + |z - x_0|^2 - 2(y - x_0, z - x_0)$,

$$= \frac{1}{2} |y - x_0|^2 + \frac{1}{2} |z - x_0|^2 - (y - x_0, z - x_0) + \alpha \phi(z) - \frac{1}{2} |y - x_0|^2 - \alpha \phi(x_0)$$

Thus, letting $z$ be replaced with $x_0 + t(z - x_0)$ for small positive $t$,

$$t(y - x_0, z - x_0) \leq \frac{t^2}{2} |z - x_0|^2 + \alpha \phi(x_0 + t(z - x_0)) - \alpha \phi(x_0)$$

Using convexity of $\phi$,

$$\leq \frac{t^2}{2} |z - x_0|^2 + t\alpha \phi(z) - t\alpha \phi(x_0)$$

Divide by $t$ and let $t \to 0$ to obtain that

$$(y - x_0, z - x_0) \leq \alpha \phi(z) - \alpha \phi(x_0)$$

and so

$$y - x_0 \in \partial (\alpha \phi(x_0))$$

Thus $y = x_0 + \alpha \partial \phi(x_0)$ because $\partial (\alpha \phi) = \alpha \partial \phi$. ■

There is a really amazing theorem, Moreau’s theorem. It is in [1], [2] and [5]. It involves approximating a convex function with one which is differentiable.
Theorem 10.4.22 Let \( \phi \) be a convex lower semicontinuous proper function defined on \( H \). Define

\[
\phi_\lambda (x) \equiv \min_{y \in H} \left( \frac{1}{2\lambda} |x - y|^2 + \phi (y) \right)
\]

Then the function is well defined, convex, Frechet differentiable, and for all \( x \in H \),

\[
\lim_{\lambda \to 0} \phi_\lambda (x) = \phi (x),
\]

\( \phi_\lambda (x) \) increasing as \( \lambda \) decreases. In addition,

\[
\phi_\lambda (x) = \frac{1}{2\lambda} |x - J_\lambda x|^2 + \phi (J_\lambda (x))
\]

where \( J_\lambda x \equiv (I + \lambda \partial \phi)^{-1} (x) \). The Frechet derivative at \( x \) equals \( A_\lambda x \) where

\[
A_\lambda = \frac{1}{\lambda} - \frac{1}{\lambda} (I + \lambda \partial \phi)^{-1} = \frac{1}{\lambda} - \frac{1}{\lambda} J_\lambda
\]

Also, there is an interesting relation between the domain of \( \phi \) and the domain of \( \partial \phi \)

\[
\text{dom} \phi \subseteq \text{dom} (\partial \phi) \subseteq \overline{\text{dom} (\partial \phi)}
\]

Proof: First of all, why does the minimum take place? By the convexity, closed epigraph, and assumption that \( \phi \) is proper, separation theorems apply and one can say that there exists \( z^* \) such that for all \( y \in H \),

\[
\frac{1}{2\lambda} |x - y|^2 + \phi (y) \geq \frac{1}{2\lambda} |x - y|^2 + (z^*, y) + c \quad (10.4.35)
\]

It follows easily that a minimizing sequence is bounded and so from lower semicontinuity which implies weak lower semicontinuity, there exists \( y_x \) such that

\[
\min_{y \in H} \left( \frac{1}{2\lambda} |x - y|^2 + \phi (y) \right) = \left( \frac{1}{2\lambda} |x - y_x|^2 + \phi (y_x) \right)
\]

Why is \( \phi_\lambda \) convex? For \( \theta \in [0, 1] \),

\[
\phi_\lambda (\theta x + (1 - \theta) z) = \frac{1}{2\lambda} \left| \theta x + (1 - \theta) z - y_{(\theta x + (1 - \theta) z)} \right|^2 + \phi (y_{(\theta x + (1 - \theta) z)})
\]

\[
\leq \frac{1}{2\lambda} \left| \theta x + (1 - \theta) z - (\theta y_x + (1 - \theta) y_z) \right|^2 + \phi (\theta y_x + (1 - \theta) y_z)
\]

\[
\leq \frac{\theta}{2\lambda} |x - y_x|^2 + \frac{1 - \theta}{2\lambda} |z - y_z|^2 + \theta \phi (y_x) + (1 - \theta) \phi (y_z)
\]

\[
= \theta \phi_\lambda (x) + (1 - \theta) \phi_\lambda (z)
\]

So is there a formula for \( y_x \)? Since it involves minimization of the functional, it follows as in Lemma 10.4.21 that

\[
\frac{1}{\lambda} (x - y_x) \in \partial \phi (y_x)
\]
Thus
\[ x \in y_x + \lambda \partial \phi(y_x) \]
and so
\[ y_x = J_\lambda x. \]
Thus
\[ \phi_\lambda(x) = \frac{1}{2\lambda} |x - J_\lambda x|^2 + \phi(J_\lambda x) = \frac{\lambda}{2} |A_\lambda x|^2 + \phi(J_\lambda x) \]
Note that \( J_\lambda x \in D(\partial \phi) \) and so it must also be in \( D(\phi) \). Now also
\[ A_\lambda x \equiv \frac{x}{\lambda} - \frac{1}{\lambda} J_\lambda x \in \partial \phi(J_\lambda x) \]
This is so if and only if
\[ x \in J_\lambda x + \lambda \partial \phi(J_\lambda x) = (I + \lambda \partial \phi)(J_\lambda x) = (I + \lambda \partial \phi)(I + \lambda \partial \phi)^{-1} x \]
which is clearly true by definition.

Next consider the claim about differentiability.
\[
\phi_\lambda(y) - \phi_\lambda(x) = \frac{\lambda}{2} |A_\lambda y|^2 + \phi(J_\lambda y) - \left( \frac{\lambda}{2} |A_\lambda x|^2 + \phi(J_\lambda x) \right)
\]
= \[ \frac{\lambda}{2} \left( |A_\lambda y|^2 - |A_\lambda x|^2 \right) + \phi(J_\lambda y) - \phi(J_\lambda x) \]
\[ \geq \frac{\lambda}{2} \left( |A_\lambda y|^2 - |A_\lambda x|^2 \right) + (A_\lambda x, J_\lambda y - J_\lambda x) \]
\[ \geq \frac{\lambda}{2} \left( |A_\lambda y|^2 - |A_\lambda x|^2 \right) + \lambda |A_\lambda x|^2 - \frac{\lambda}{2} |A_\lambda x|^2 - \frac{\lambda}{2} |A_\lambda y|^2 + (A_\lambda x, y - x) \]
\[ = (A_\lambda x, y - x) = (A_\lambda x - A_\lambda y, y - x) + (A_\lambda y, y - x) \] (10.4.36)

Then it follows that
\[ -(A_\lambda x - A_\lambda y, y - x) \geq \phi_\lambda(x) - \phi_\lambda(y) - (A_\lambda y, x - y) \]
However, \( A_\lambda \) is Lipschitz continuous with constant \( 1/\lambda \) and so
\[ \frac{1}{\lambda} |x - y|^2 \geq \phi_\lambda(x) - \phi_\lambda(y) - (A_\lambda y, x - y) \] (10.4.37)

Then switching \( x, y \) in the equation
\[ \frac{1}{\lambda} |x - y|^2 \geq \phi_\lambda(y) - \phi_\lambda(x) - (A_\lambda x, y - x) \] (10.4.38)
But also that term on the end in (10.4.37) equals \((A\lambda y, y - x)\) and so it is also the case that
\[
\frac{1}{\lambda} |x - y|^2 \geq \phi_\lambda (x) - \phi_\lambda (y) + (A\lambda x, y - x)
\]
\[
= - (\phi_\lambda (y) - \phi_\lambda (x) - (A\lambda x, y - x))
\]
(10.4.39)

From (10.4.38) and (10.4.39) it follows that
\[
\frac{1}{\lambda} |x - y|^2 \geq |\phi_\lambda (y) - \phi_\lambda (x) - (A\lambda x, y - x)|
\]
which shows that \(D\phi_\lambda(x) = A\lambda x\). This proves the differentiability part.

Next recall that for any maximal monotone operator \(A\), if you have \(x \in D(A)\),
\[
\lim_{\lambda \to 0} J_\lambda x = x
\]
Recall why this was so. If \(x \in D(A)\), then
\[
x - J_\lambda x \in \lambda A x
\]
and so, \(|x - J_\lambda x| \to 0\) as \(\lambda \to 0\). If \(x\) is only in \(
\bar{D}(A)\), it also works because for \(y \in D(A)\)
\[
|x - J_\lambda x| \leq |x - y| + |y - J_\lambda y| + |J_\lambda y - J_\lambda x|
\]
\[
\leq 2|x - y| + |y - J_\lambda y|
\]
If \(\varepsilon\) is given, simply pick \(|y - x| < \varepsilon/2\) and then
\[
|x - J_\lambda x| \leq \varepsilon + |y - J_\lambda y|
\]
and the last converges to 0. Therefore, \(J_\lambda x \to x\) on \(\bar{D}(A)\).

Returning to the proof of the theorem, if \(x \in \bar{D}(\partial \phi)\) then recall that
\[
\phi_\lambda (x) = \frac{1}{2\lambda} |x - J_\lambda x|^2 + \phi(J_\lambda x)
\]
and so,
\[
\liminf_{\lambda \to 0} \phi_\lambda (x) \geq \liminf_{\lambda \to 0} \phi(J_\lambda x) \geq \phi(x) \geq \limsup_{\lambda \to 0} \phi_\lambda (x)
\]
which shows the desired result in case \(x \in \bar{D}(\partial \phi)\). Now consider the case where \(x \notin \bar{D}(\partial \phi)\). In this case, there is a positive lower bound \(\delta\) to \(|x - J_\lambda x|\) because each \(J_\lambda x \in D(\partial \phi)\). Then from the definition and what was shown above,
\[
\phi_\lambda (x) = \frac{\lambda}{2} |A\lambda x|^2 + \phi(J_\lambda x) \geq \frac{\lambda}{2} |A\lambda x|^2 + (z^*, J_\lambda x) + c
\]
\[
\geq \frac{\lambda}{2} |A\lambda x|^2 + (z^*, J_\lambda x - x) + (z^*, x) + c
\]
10.5. A Perturbation Theorem

In this section is a simple perturbation theorem found in [9] and [54].

Recall that for $B$ a maximal monotone operator, $B_\lambda$, the Yosida approximation, is defined by

$$B_\lambda x \equiv \frac{1}{\lambda} (x - J_\lambda x), \quad J_\lambda x \equiv (I + \lambda B)^{-1} x.$$ 

This follows from Theorem 10.2.7 on Page 380.

**Theorem 10.5.1** Let $A$ and $B$ be maximal monotone operators and let $x_\lambda$ be the solution to

$$y \in x_\lambda + B_\lambda x_\lambda + Ax_\lambda.$$ 

Then $y \in x + Bx + Ax$ for some $x \in D(A) \cap D(B)$ if $B_\lambda x_\lambda$ is bounded independent of $\lambda$.

The following is the perturbation theorem of this section. See [9] and [54].

**Theorem 10.5.2** Let $H$ be a real Hilbert space and let $\Phi$ be non-negative, convex, proper, and lower semicontinuous. Suppose also that $A$ is a maximal monotone operator and there exists

$$\xi \in D(A) \cap D(\Phi).$$ 

Suppose also that for $J_\lambda x \equiv (I + \lambda A)^{-1} x$,

$$\Phi(J_\lambda x) \leq \Phi(x) + C_\lambda$$

Then $A + \partial \Phi$ is maximal monotone.

**Proof:** Letting $A_\lambda$ be the Yosida approximation of $A$,

$$A_\lambda x = \frac{1}{\lambda} (x - J_\lambda x),$$

Hence $\phi_{\lambda}(x) \to \infty$ and since $\phi(x) \geq \phi_{\lambda}(x)$ by construction, it follows that $\phi(x) = \infty$. The construction of $\phi_{\lambda}$ also shows that as $\lambda$ decreases, $\phi_{\lambda}(x)$ increases.

Note that the last part of the argument shows that if $x \not\in D(\partial \phi)$, then $x \not\in D(\phi)$.

Hence this shows that $D(\partial \phi) \subseteq D(\phi) \subseteq D(\partial \phi)$.
and letting \( y \in H \), it follows from the Hilbert space version of Proposition 10.2.6 that there exists \( x_\lambda \in H \) such that
\[
y \in x_\lambda + A_\lambda x_\lambda + \partial \Phi (x_\lambda).
\]
Consequently,
\[
y - x_\lambda - A_\lambda x_\lambda \in \partial \Phi (x_\lambda)
\]
and so
\[
(y - x_\lambda - A_\lambda x_\lambda, J_\lambda x_\lambda - x_\lambda) \leq \Phi (J_\lambda x_\lambda) - \Phi (x_\lambda) \leq C \lambda
\]
which implies
\[
- (y - x_\lambda - A_\lambda x_\lambda, A_\lambda x_\lambda) = |A_\lambda x_\lambda|^2 - |y - x_\lambda| |A_\lambda x_\lambda| \leq C.
\]
By 10.5.42 and monotonicity of \( A_\lambda \),
\[
\Phi (\xi) - \Phi (x_\lambda) \geq \left( \frac{y - x_\lambda - A_\lambda x_\lambda, \xi - x_\lambda}{\in \partial \Phi (x_\lambda)} \right)
\]
\[
(y - x_\lambda, \xi - x_\lambda) - (A_\lambda x_\lambda, \xi - x_\lambda)
\]
\[
\geq (y - x_\lambda, \xi - x_\lambda) - (A_\lambda \xi, \xi - x_\lambda)
\]
\[
\geq (y - \xi, \xi - x_\lambda) + |\xi - x_\lambda|^2 - (A_\lambda \xi, \xi - x_\lambda)
\]
\[
= |\xi - x_\lambda|^2 + (y - \xi - A_\lambda \xi, \xi - x_\lambda)
\]
\[
\geq |\xi - x_\lambda|^2 - C_{\xi y} |\xi - x_\lambda|
\]
where \( C_{\xi y} \) depends on \( \xi \) and \( y \) but is independent of \( \lambda \) because of the assumption that \( \xi \in D (A) \cap D (\Phi) \) and Lemma 10.2.5 which gives a bound on \( |A_\lambda \xi| \) in terms of \( |y| \) for \( y \in Ax \). Therefore, there exist constants, \( C_1 \) and \( C_2 \), depending on \( \xi \) and \( y \) but not on \( \lambda \) such that
\[
\Phi (\xi) \geq \Phi (x_\lambda) + |x_\lambda|^2 - C_1 |x_\lambda| - C_2.
\]
Since \( \Phi \geq 0 \), this shows that \( |x_\lambda| \) is bounded independent of \( \lambda \).
\[
2 \left( \Phi (\xi) + C_2 + \frac{C_1^2}{2} \right) \geq \Phi (x_\lambda) + |x_\lambda|^2.
\]
This shows \( |x_\lambda| \) is bounded independent of \( \lambda \). Therefore, by 10.5.43, \( |A_\lambda x_\lambda| \) is bounded independent of \( \lambda \). By Theorem 10.5.1, this shows there exists \( x \in D (\partial \Phi) \cap D (A) \) such that
\[
y \in Ax + \partial \Phi (x) + x
\]
and so \( A + \partial \Phi \) is maximal monotone since \( y \in H \) was arbitrary. ■
10.6 An Evolution Inclusion

In this section is a theorem on existence and uniqueness for the initial value problem

\[ x' + \partial\phi (x) \ni f, \ x(0) = x_0. \]

Suppose \( \phi \) is a mapping from \( H \) to \([0, \infty]\) which satisfy the following axioms.

\[ \phi \text{ is convex and lower semicontinuous, and proper,} \quad (10.6.45) \]

**Lemma 10.6.1** For \( x \in L^2 (0, T; H) \), \( t \to \phi (x) \) is measurable.

**Proof:** This follows because \( \phi \) is Borel measurable and so \( \phi \circ x \) is also measurable. \( \blacksquare \)

Now define the following function \( \Phi \), on the Hilbert space, \( L^2 (0, T; H) \).

\[ \Phi (x) \equiv \begin{cases} \int_0^T \phi (x(t)) \, dt & \text{if } x(t) \in D \text{ for a.e. } t \\ +\infty & \text{otherwise} \end{cases} \quad (10.6.46) \]

**Lemma 10.6.2** \( \Phi \) is convex, nonnegative, and lower semicontinuous on \( L^2 (0, T; H) \).

**Proof:** Since \( \phi \) is nonnegative and convex, it follows that \( \Phi \) is also nonnegative and convex. It remains to verify lower semicontinuity. Suppose, \( x_n \to x \) in \( L^2 (0, T; H) \) and let

\[ \lambda = \lim \inf_{n \to \infty} \Phi (x_n). \]

Is \( \lambda \geq \Phi (x) \)? Then is suffices to assume \( \lambda < \infty \). Suppose not. Then \( \lambda < \Phi (x) \). Taking a subsequence, we can have \( \lambda = \lim_{n \to \infty} \Phi (x_n) \) and we can take a further subsequence for which convergence of \( x_n \) to \( x \) is pointwise a.e. Then

\[ \lambda < \Phi (x) \equiv \int_0^T \phi (x(t)) \, dt \leq \int_0^T \lim \inf_{n \to \infty} \phi (x_n(t)) \, dt \leq \lim \inf_{n \to \infty} \int_0^T \phi (x_n(t)) \, dt = \lim \inf_{n \to \infty} \Phi (x_n) = \lambda \]

which is a contradiction. \( \blacksquare \)

Define

\[ D(L) \equiv \{ x \in L^2 (0, T; H) : \text{such that} \}

\[ x(t) = x_0 + \int_0^t x'(s) \, ds \text{ where } x' \in L^2 (0, T; H) \} \quad (10.6.47) \]

and for \( x \in D(L) \),

\[ Lx \equiv x'. \]

Then \( L \) is maximal monotone. To see this, consider the equation

\[ \lambda x' + x = z, \ x(0) = x_0 \]
It clearly has a solution so $\lambda L + I$ is onto. In fact, the solution is

$$x = e^{-\frac{x}{\lambda}} x_0 + \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \int_0^t e^{\frac{x}{\lambda} s} (s) \, ds$$

Also,

$$(Lx - Ly, x - y)_{L^2(0,T;H)} = \int_0^T ((x' - y'), (x - y) \, H \, dt$$

$$= \int_0^T \left( x'(t) - y'(t), \int_0^t x'(s) - y'(s) \, ds \right) \, dt$$

$$= \frac{1}{2} \int_0^T \frac{d}{dt} \left( \left| \int_0^t x'(s) - y'(s) \, ds \right|^2 \right) \, dt$$

$$= \left| \int_0^t x'(s) - y'(s) \, ds \right|^2_H \geq 0$$

Thus we have the following lemma.

**Lemma 10.6.3** $L$ is maximal monotone and if $z \in L^2(0,T;H)$, then $J_\lambda z$ is given by

$$J_\lambda [z](t) \equiv (I + \lambda L)^{-1} ([z])(t) = e^{-\frac{x}{\lambda}} x_0 + \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \int_0^t e^{\frac{x}{\lambda} s} z(s) \, ds. \quad (10.6.48)$$

The main theorem is the following.

**Theorem 10.6.4** Let $x_0 \in D \equiv D(\phi)$. Then $L + \partial \Phi$ is maximal monotone so there exists a unique solution to

$$Lx + x + \partial \Phi (x) \ni f \quad (10.6.49)$$

for every $f \in L^2(0,T;H)$. Thus there exists $x \in L^2(0,T;H)$ such that $x' \in L^2(0,T;H), x(0) = x_0 \in D(\phi)$, and

$$x' + x + \partial \Phi (x) \ni f, \quad x(0) = x_0$$

**Proof:** This is from Theorem 10.5.2. Since $x_0 \in D$, it follows that $\phi(x_0) < \infty$.

Let $z \in D(\Phi)$, the effective domain of $\Phi$. Then $\int_0^T \phi(z(t)) \, dt < \infty$, so by convexity of $\phi$ and 10.6.48,

$$\phi(J_\lambda z(t)) \leq e^{-\frac{t}{\lambda}} \phi(x_0) + \frac{1}{\lambda} e^{-\frac{t}{\lambda}} \int_0^t e^{\frac{s}{\lambda}} \phi(z(s)) \, ds. \quad (10.6.50)$$

Then

$$\Phi(J_\lambda z) =$$
10.6. AN EVOLUTION INCLUSION

\[
\int_0^T \phi(J_\lambda z(t)) \, dt \leq \phi(x_0) \lambda + \int_0^T \frac{1}{\lambda} \int_0^t e^{-(t-s)/\lambda} \phi(z(s)) \, ds \, dt
\]

\[
\leq \lambda \phi(x_0) + \frac{1}{\lambda} \int_0^T \phi(z(s)) \int_s^T e^{-(t-s)/\lambda} \, dt \, ds
\]

\[
\leq \lambda \phi(x_0) + \left( \int_0^T \phi(z(s)) \, ds \right) \frac{1}{\lambda} \int_0^\infty e^{-t/\lambda} \, dt
\]

\[
= \phi(x_0) \lambda + \int_0^T \phi(z(s)) \, ds
\]

\[
= \phi(x_0) \lambda + \Phi(z)
\]

The conditions of Theorem 10.5.2 are satisfied. This proves \( L + \partial \Phi \) is maximal monotone on \( L^2(0,T;H) \) and consequently there exists a unique solution to the differential inclusion of the theorem. ■

Then the main result is the following.

**Theorem 10.6.5** Let \( f \in L^2(0,T;H) \) and \( x_0 \in D \). Let \( \phi \) be as described above, a lower semicontinuous convex proper function defined on \( H \). Then there exists a unique solution \( x \in L^2(0,T;H), x' \in L^2(0,T;H) \), to

\[
x' + \partial \Phi(x) \ni f \text{ in } L^2(0,T;H), \ x(0) = x_0
\]

This satisfies the pointwise condition

\[
x'(t) + \partial \phi(x(t)) \ni f(t) \text{ for a.e. } t, \ x(0) = x_0
\]

**Proof:** From Theorem 10.5.2 there exists a unique solution to

\[
x' + \partial \Phi(x_v) + x_v \ni f + v \text{ in } L^2(0,T;H), \ x_v(0) = x_0
\]

whenever \( v \in L^2(0,T;H) \). Then a simple argument based on fundamental theorem of calculus implies that for a.e. \( t \),

\[
x'(t) + \partial \phi(x(t)) + x_v(t) \ni f(t) + v(t)
\]

Then for given \( v,u \) one can act on \( x_v(t) - x_u(t) \) and integrate. This yields

\[
\frac{1}{2} \| x_v(t) - x_u(t) \|^2_H + \int_0^t |x_v - x_u|^2 \, ds \leq \int_0^t |v(s) - u(s)|^2_H \, ds
\]

It follows that a sufficiently high power of the mapping \( u \to x_u \) is a contraction map on \( C([0,T];H) \) and so by Theorem 1.8.5 there exists a unique fixed point \( v \) in \( C([0,T];H) \). Thus \( x_v = v \) and so

\[
v' + \partial \Phi(v) \ni f \text{ in } L^2(0,T;H), \ v(0) = x_0
\]
The details on showing that a high enough power of the mapping \( v \to x_v \) is a contraction goes as follows. Let \( \psi(v) = x_v \). Then you have

\[
|\psi^n(v)(t) - \psi^n(u)(t)|^2 \leq 2 \int_0^t |\psi^{n-1}(v)(s) - \psi^{n-1}(u)(s)|^2 \, ds
\]

and so

\[
\leq 2 \int_0^t 2 \int_0^s |\psi^{n-2}(v)(r) - \psi^{n-2}(u)(r)|^2 \, dr \, ds
\]

Now you just iterate this to eventually find that for large \( n \),

\[
|\psi^n(v)(t) - \psi^n(u)(t)|^2 \leq 2^n \int_0^t \int_0^{s_1} \cdots \int_0^{s_m} ds_{m+1} \cdots ds_1 \|v - u\|_{C([0,T],H)}^2
\]

\[
\leq \frac{2^{n+1}}{(n-1)!} (1 + T)^{n+1} \|v - u\|_{C([0,T],H)}^2
\]

Thus

\[
\|\psi^n(v) - \psi^n(u)\| \leq \frac{2^{n+1}}{(n-1)!} (1 + T)^{n+1} \|v - u\|_{C([0,T],H)}^2 \leq \frac{1}{2} \|v - u\|_{C([0,T],H)}^2
\]

whenever \( n \) is sufficiently large. You can doubtless obtain a more precise inequality, but such effort only obfuscates the essential ideas.

### 10.7 A More Complicated Perturbation Theorem

In this section is a simple perturbation theorem which is a small generalization of one found in [9] and [54].

Recall that for \( B \) a maximal monotone operator, \( B_\lambda \), the Yosida approximation, is defined by

\[
B_\lambda x \equiv \frac{1}{\lambda} (x - J_\lambda x), \quad J_\lambda x \equiv (I + \lambda B)^{-1} x.
\]

This follows from Theorem 10.2.7 on Page 389.

**Theorem 10.7.1** Let \( A \) and \( B \) be maximal monotone operators and let \( x_\lambda \) be the solution to

\[
y \in x_\lambda + B_\lambda x_\lambda + Ax_\lambda.
\]

Then \( y \in x + Bx + Ax \) for some \( x \in D(A) \cap D(B) \) if \( B_\lambda x_\lambda \) is bounded independent of \( \lambda \).
10.7. A MORE COMPLICATED PERTURBATION THEOREM

**Theorem 10.7.2** Let $H$ be a real Hilbert space and let $\Phi$ be non negative, convex, proper, and lower semicontinuous. Suppose also that $A$ is a maximal monotone operator and there exists

$$\xi \in D(A) \cap D(\Phi).$$

(10.7.51)

Suppose also that for $J_\lambda x \equiv (I + \lambda A)^{-1} x$,

$$\Phi(J_\lambda x) \leq \Phi(x) + C(x) \lambda$$

(10.7.52)

where for some constants, $K_1, K_2$,

$$K_2 + K_1 \left( \Phi(x) + |x|^2 \right) \geq C(x).$$

(10.7.53)

Then $A + \partial \Phi$ is maximal monotone.

**Proof:** Letting $A_\lambda$ be the Yosida approximation of $A$,

$$A_\lambda x = \frac{1}{\lambda} (x - J_\lambda x),$$

and letting $y \in H$, it follows from the Hilbert space version of Proposition 10.2.6 there exists $x_\lambda \in H$ such that

$$y \in x_\lambda + A_\lambda x_\lambda + \partial \Phi(x_\lambda).$$

Consequently,

$$y - x_\lambda - A_\lambda x_\lambda \in \partial \Phi(x_\lambda)$$

(10.7.54)

and so

$$(y - x_\lambda - A_\lambda x_\lambda, J_\lambda x_\lambda - x_\lambda) \leq \Phi(J_\lambda x_\lambda) - \Phi(x_\lambda) \leq C(x) \lambda$$

(10.7.55)

which implies

$$- (y - x_\lambda - A_\lambda x_\lambda, A_\lambda x_\lambda) \leq C(x) \lambda.$$  

(10.7.56)

I claim $\{C(x_\lambda)\}$ and $\{|x_\lambda|\}$ are bounded independent of $\lambda$.

By (10.7.55) and monotonicity of $A_\lambda$,

$$\Phi(\xi) - \Phi(x_\lambda) \geq \left(\frac{\in \partial \Phi(x_\lambda)}{y - x_\lambda - A_\lambda x_\lambda, \xi - x_\lambda}\right)$$

$$\geq (y - x_\lambda, \xi - x_\lambda) - (A_\lambda x_\lambda, \xi - x_\lambda)$$

$$\geq (y - x_\lambda, \xi - x_\lambda) - (A_\lambda \xi, \xi - x_\lambda)$$

$$\geq (y - \xi, \xi - x_\lambda) + |\xi - x_\lambda|^2 - (A_\lambda \xi, \xi - x_\lambda)$$

$$\geq |\xi - x_\lambda|^2 - C_\xi |\xi - x_\lambda|$$
where \( C_{\xi y} \) depends on \( \xi \) and \( y \) but is independent of \( \lambda \) because of the assumption that \( \xi \in D(A) \cap D(\Phi) \) and Lemma 10.2.3 which gives a bound on \( |A_\lambda \xi| \) in terms of \( |y| \) for \( y \in Ax \). Therefore, there exist constants, \( C_1 \) and \( C_2 \), depending on \( \xi \) and \( y \) but not on \( \lambda \) such that

\[
\Phi(\xi) \geq \Phi(x_\lambda) + |x_\lambda|^2 - C_1 |x_\lambda| - C_2.
\]

Since \( \Phi \geq 0 \),

\[
2 \left( \Phi(\xi) + C_2 + \frac{C_1^2}{2} \right) \geq \Phi(x_\lambda) + |x_\lambda|^2.
\]

This shows \( |x_\lambda| \) is bounded independent of \( \lambda \). Therefore, by Lemma 10.7.53

\[
K_2 + 2K_1 \left( \Phi(\xi) + C_2 + \frac{C_1^2}{2} \right) \geq K_2 + K_1 \left( \Phi(x_\lambda) + |x_\lambda|^2 \right) \geq C(x_\lambda),
\]

showing that both \( |x_\lambda| \) and \( C(x_\lambda) \) are bounded independent of \( \lambda \). Therefore, from Lemma 10.7.53, it follows \( A_\lambda x_\lambda \) is bounded independent of \( \lambda \). By Theorem 10.7.1, this shows there exists \( x \in D(\partial \Phi) \cap D(A) \) such that

\[
y \in Ax + \partial \Phi(x) + x
\]

and so \( A + \partial \Phi \) is maximal monotone since \( y \in H \) was arbitrary. 

### 10.8 An Evolution Inclusion

In this section is a theorem on existence and uniqueness for the initial value problem

\[
x' + \partial_2 \phi(t, x) \ni f, \quad x(0) = x_0.
\]

Suppose \( \{\phi(t, \cdot)\}_{t \in [0, T]} \) is a family of functions mapping \( H \) to \([0, \infty]\) which satisfy the following axioms.

\[
\phi(t, \cdot) \text{ is convex and lower semicontinuous}, \quad (10.8.57)
\]

\[
D(\phi(t, \cdot)) = D, \quad \text{independent of } t \in [0, T], \quad (10.8.58)
\]

There exists a constant, \( K \), such that for all \( x \in D, \)

\[
|\phi(t, x) - \phi(s, x)| \leq K \left( \phi(r, x) + |x|^2 + 1 \right) |t - s| \quad (10.8.59)
\]

for all \( r \in [0, T] \).

**Lemma 10.8.1** Under the conditions, 10.8.57 - 10.8.59, \( \phi : H \times [0, T] \to [0, \infty] \) is lower semicontinuous.
10.8. AN EVOLUTION INCLUSION

Proof: Let \((x_n, t_n) \to (x, t)\) and let \(\lambda \equiv \liminf_{n \to \infty} \phi (t_n, x_n)\). Is
\[
\phi (t, x) \leq \lambda?
\]
It suffices to assume \(\lambda < \infty\) and by taking a subsequence, \(x_n \in D\) for all \(n\) and
\[
\phi (t_n, x_n) \to \lambda.
\]
Then
\[
\liminf_{n \to \infty} \phi (t_n, x_n) = \liminf_{n \to \infty} [\phi (t_n, x_n) - \phi (t, x_n) + \phi (t, x_n)].
\] (10.8.60)
Now
\[
\limsup_{n \to \infty} |\phi (t_n, x_n) - \phi (t, x_n)| \leq \limsup_{n \to \infty} K (\phi (t_n, x_n) + |x_n|^2 + 1) |t_n - t| = 0.
\]
Therefore, from \[10.8.60\]
\[
\lambda = \liminf_{n \to \infty} \phi (t_n, x_n) = \liminf_{n \to \infty} \phi (t, x_n) \geq \phi (t, x)
\]
because of the assumption that \(\phi (t, \cdot)\) is lower semicontinuous. ■

In all that follows \([x]\) is an element of \(L^2 (0, T; H)\). Thus \([x]\) is the equivalence class of measurable square integrable functions which equal \(x\) a.e. This seems a little fussy but since the existence results are based on surjectivity theorems and the Hilbert space they apply to is \(L^2 (0, T; H)\), it seems best to emphasize the equivalence classes of functions by using this notation, at least while proving theorems on existence and uniqueness.

Corollary 10.8.2 For \([x] \in L^2 (0, T; H), t \to \phi (t, x (t))\) is measurable.

Proof: This follows because, due to Lemma \[10.8.1\], \(\phi\) is Borel measurable and so \(\phi \circ x\) is also measurable.

Now define the following function, \(\Phi\), on the Hilbert space, \(L^2 (0, T; H)\).
\[
\Phi ([x]) \equiv \left\{ \begin{array}{ll} \int_0^T \phi (t, x (t)) \, dt & \text{if } x (t) \in D \text{ for all } t \text{ for some } x (\cdot) \in [x] \\ + \infty & \text{otherwise} \end{array} \right. \quad (10.8.61)
\]
Note that since the functions \(\phi (t, \cdot)\) are proper, the top condition is equivalent to the condition
\[
\int_0^T \phi (t, x (t)) \, dt \text{ if } x (t) \in D \text{ a.e. for all } x (\cdot) \in [x].
\]

Lemma 10.8.3 \(\Phi\) is convex, nonnegative, and lower semicontinuous on \(L^2 (0, T; H)\).
424 CHAPTER 10. MAXIMAL MONOTONE OPERATORS, HILBERT SPACE

**Proof:** Since each \( \phi (t, \cdot) \) is nonnegative and convex, it follows that \( \Phi \) is also nonnegative and convex. It remains to verify lower semicontinuity. Suppose, \( [x_n] \to [x] \) in \( L^2 (0, T; H) \) and let

\[
\lambda = \lim \inf_{n \to \infty} \Phi ([x_n]) .
\]

Is \( \lambda \geq \Phi ([x]) \)? It suffices to assume \( \lambda < \infty \), \( x_n (t) \in D \) for all \( t \), and \( x_n (t) \to x (t) \) a.e. say for \( t \notin N \) where \( N \) has measure zero. Let

\[
\bar{x} (t) = \begin{cases} 
  x (t) & \text{if } t \notin N \\
  x_1 (t) & \text{if } t \in N 
\end{cases}
\]

Then \( \bar{x} = [x] \) and \( \bar{x} (t) \in D \) for all \( t \). Then by pointwise convergence and Fatou’s lemma,

\[
\Phi ([x]) = \Phi ([\bar{x}]) = \int_0^T \phi (t, \bar{x} (t)) \, dt \leq \int_0^T \lim \inf_{n \to \infty} \phi (t, x_n (t)) \, dt
\]

\[
\leq \lim \inf_{n \to \infty} \int_0^T \phi (t, x_n (t)) \, dt = \lim \inf_{n \to \infty} \Phi ([x_n]) \equiv \lambda .
\]

Define

\[
D (L) = \{ [x] \in L^2 (0, T; H) : \text{for some } x \in [x] \text{ such that } x (t) = x_0 + \int_0^t x' (s) \, ds \text{ where } [x'] \in L^2 (0, T; H) \}
\]

and for \( [x] \in D (L) \),

\[
L [x] \equiv [x'] .
\]

Then \( L \) is maximal monotone. To see this, consider the equation

\[
\lambda x' + x = f, \ x (0) = x_0
\]

It clearly has a solution so \( \lambda L + I \) is onto. Also picking representatives,

\[
(Lx - Ly, x - y)_{L^2(0, T; H)} = \int_0^T ((x' - y'), x - y) \, dt
\]

\[
= \int_0^T \left( \int_0^t x' (s) - y' (s) \, ds \right) \, dt
\]

\[
= \frac{1}{2} \int_0^T \frac{d}{dt} \left( \left\| \int_0^t x' (s) - y' (s) \, ds \right\|^2 \right) \, dt
\]

\[
= \left\| \int_0^t x' (s) - y' (s) \, ds \right\|^2_H \geq 0
\]

The following lemma is easily obtained.
10.8. AN EVOLUTION INCLUSION

Lemma 10.8.4 Let $L$ is maximal monotone and if $[z] \in L^2(0,T;H)$, then the equivalence class, $[J_\lambda [z]]$ is determined by the function,

$$J_\lambda [z] (t) \equiv (I + \lambda L)^{-1} ([z]) (t) = e^{-\lambda t} x_0 + \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} z (s) \, ds. \quad (10.8.63)$$

The main theorem is the following.

Theorem 10.8.5 Let $x_0 \in D$. Then $L + \partial \Phi$ is maximal monotone so there exists a unique solution to

$$L [x] + [x] + \partial \Phi ([x]) \ni [f] \quad (10.8.64)$$

for every $[f] \in L^2 (0,T;H)$.

Proof: This is from Theorem 10.8.2. Since $x_0 \in D$, it follows from 10.8.64 that $\phi (t, x_0)$ is bounded.

Let $[z] \in D (\Phi)$, the effective domain of $\Phi$. Then there exists $z \in [z]$ such that $z \in D$ for all $t$, and $\int_0^T \phi (t,z(t)) \, dt < \infty$, so by convexity of $\phi (t, \cdot)$ and 10.8.63,

$$\phi (t, J_\lambda [z] (t)) \leq e^{-\lambda t} \phi (t,x_0) + \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} \phi (t,z(s)) \, ds. \quad (10.8.65)$$

Now the first term in 10.8.64 is bounded so consider the second. The integral in this term is of the form

$$\int_0^t e^{\lambda s} \phi (s,z(s)) \, ds + \int_0^t e^{\lambda s} (\phi (t,z(s)) - \phi (s,z(s))) \, ds. \quad (10.8.66)$$

Since $[z] \in D (\Phi)$, $\phi (s,z(s)) < \infty$ for all $s$ and also the first integral in 10.8.65 is finite. By 10.8.73, the second term in 10.8.66 is dominated by

$$C_\lambda \int_0^t K \left(1 + \phi (s,z(s)) + |z(s)|^2 \right) \, dt < \infty.$$

This shows $\phi (t, J_\lambda [z] (t)) < \infty$ for all $t$ and so $\Phi ([J_\lambda [z]])$ is given by the top line of 10.8.64. Therefore, by convexity of $\phi (t, \cdot)$ and Jensen’s inequality,

$$\Phi ([J_\lambda [z]]) = \int_0^T \phi \left(t, e^{-\lambda t} x_0 + \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} z(s) \, ds \right) \, dt$$

$$\leq \int_0^T \left(e^{-\lambda t} \phi (t,x_0) + \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} \phi (t,z(s)) \, ds \right) \, dt$$

$$= \int_0^T e^{-\lambda t} \phi (t,x_0) \, dt + \int_0^T \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} \phi (s,z(s)) \, ds \, dt$$

$$+ \int_0^T \frac{1}{\lambda} e^{-\lambda t} \int_0^t e^{\lambda s} (\phi (t,z(s)) - \phi (s,z(s))) \, ds \, dt. \quad (10.8.67)$$
By (10.7.67) the last term is dominated by
\[\int_{0}^{T} \int_{0}^{t} \frac{1}{\lambda} e^{-\frac{(t-s)}{\lambda}} K \left( 1 + \phi(s, z(s)) + |z(s)|^2 \right) |t-s| ds dt = \]
\[\int_{0}^{T} \int_{s}^{T} \frac{1}{\lambda} e^{-\frac{(t-s)}{\lambda}} t - s \ ds dt \left( 1 + \phi(s, z(s)) + |z(s)|^2 \right) ds \]
\[\leq C \lambda + C \lambda \Phi([z]) + ||z||^2 \]
(10.8.68)
for some constant, C. From (10.8.68), \(\phi(t, x_0)\) is bounded and so the first term in (10.8.67) is dominated by an expression of the form \(C \lambda\). Now consider the middle term of (10.8.67). Since \(\phi\) is nonnegative,
\[\int_{0}^{T} \frac{1}{\lambda} e^{-\frac{t}{\lambda}} \int_{0}^{t} e^{\frac{s}{\lambda}} \phi(s, z(s)) ds dt = \int_{0}^{T} \int_{s}^{T} \frac{1}{\lambda} e^{-\frac{(t-s)}{\lambda}} ds dt \phi(s, z(s)) ds \]
\[\leq \int_{0}^{T} \int_{0}^{\infty} e^{-u} \phi(s, z(s)) ds = \Phi([z]) \]
(10.8.69)
It follows
\[\Phi([J_\lambda [z]]) \leq \Phi([z]) + C \lambda + C \lambda \Phi([z]) + ||z||^2 \]
The conditions of Theorem (10.7.6) are satisfied with \(K_1 = K_2 = C\). This proves \(L + \partial \Phi\) is maximal monotone on \(L^2(0, T; H)\) and consequently there exists a unique solution to the differential inclusion of the theorem.

Of course it is desirable to be able to say that \([y] \in \partial \Phi([x])\) if and only if \(y(t) \in \partial \Phi(t, x(t))\) for some \(x \in [x]\). To obtain this, here are two more assumptions. For all \(x \in H\),
\[t \to J_1(t) x \text{ is measurable,} \]
(10.8.70)
where \(J_1(t) x\) is the solution, \(y\), to \(y(t) + \partial_2 \phi(t, y(t)) \ni x\), and there exists \([\xi] \in L^2(0, T; H)\) such that
\([J_1(\cdot) [\xi]] \in L^2(0, T; H)\).
(10.8.71)

**Lemma 10.8.6** If (10.8.70) and (10.8.71) hold, and if \([y] \in L^2(0, T; H)\), then \([y] \in \partial \Phi([x])\) if and only if there exists \(x \in [x]\) such that \(\partial_2 \phi(t, x(t)) \neq \emptyset\) for all \(t\) and \(y(t) \in \partial_2 \phi(t, x(t))\) a.e.

**Proof:** First suppose \(y(t) \in \partial_2 \phi(t, x(t))\) a.e. and \(\partial_2 \phi(t, x(t)) \neq \emptyset\) for all \(t\) where \(x \in [x]\). Then for all \([w] \in L^2(0, T; H)\),
\[(y, [w])_{L^2(0, T; H)} = \int_{0}^{T} (y(t), w(t))_H dt \]
\[\leq \int_{0}^{T} (\phi(t, x(t) + w(t)) dt - \int_{0}^{T} \phi(t, x(t)) dt \leq \Phi([x] + [w]) - \Phi([x]). \]
To prove the converse, define $A : D (\partial \Phi) \to \mathcal{P} (L^2 [0, T; H])$ as follows.

$$[y] \in A [x] \text{ if and only if for some } x \in [x],$$

$$\partial_2 \phi (t, x (t)) \neq \emptyset \text{ for all } t \text{ and } y (t) \in \partial_2 \phi (t, x (t)) \text{ a.e. } t.$$ 

It follows $A$ is monotone. I will show $A$ is maximal monotone. From the first part of the proof, the graph of $A$ is contained in the graph of $\partial \Phi$. Since $A$ is maximal, this will imply $A = \partial \Phi$ and prove the lemma.

It remains to show $A$ is maximal monotone. By 10.8.70, for each $x \in H, J_1 (t) x$ is measurable. Now from 10.8.71, and using the fact that $J_1 (t)$ is a contraction,

$$|J_1 (t) x - J_1 (t) \xi (t)| \leq |x - \xi (t)|$$

and so $[J_1 (\cdot) x]$ is in $L^2 (0, T; H)$. Now if

$$s (t) = \sum_{i=1}^{n} \lambda_{E_i} (t) x_i$$

is a simple function,

$$J_1 (t) s (t) = \sum_{i=1}^{n} \lambda_{E_i} (t) J_1 (t) x,$$

and $[J_1 (\cdot) s]$ is in $L^2 (0, T; H)$. If $[f] \in L^2 (0, T; H)$ is arbitrary, take a sequence of simple functions, $s_n$ converging to $f$ pointwise and $[s_n] \to [f]$ in $L^2 (0, T; H)$. Then

$$|J_1 (t) s_n (t) - J_1 (t) f (t)| \leq |s_n (t) - f (t)|$$

and it follows $J_1 (t) s_n (t)$ converges pointwise to $J_1 (t) f (t)$ showing that $t \to J_1 (t) f (t)$ is measurable. Now the equivalence class of functions equal to this one a.e. is in $L^2 (0, T; H)$ by Fatou’s lemma and the assumption that the simple functions, $s_n$ converge in $L^2 (0, T; H)$. This shows $A$ is maximal and proves the lemma.

Conditions 10.8.70 and 10.8.71 are just what is needed to obtain the conclusion of Lemma 10.8.6 but it may not be clear how to verify these conditions easily. The following lemma gives sufficient conditions which are easy to verify which imply 10.8.70 and 10.8.71.

**Lemma 10.8.7** Suppose there exists $[\xi] \in L^2 (0, T; H)$ such that

$$J_1 (t) \xi (t), \phi (t, J_1 (t) \xi (t))$$

are bounded independent of $t \in [0, T]$ and $t \to J_1 (t) \xi (t)$ is measurable. Then the conclusion of Lemma 10.8.6 holds.

**Proof:** Let $y (t) = J_1 (t) \xi (t)$. Thus

$$y (t) + \partial_2 \phi (t, y (t)) \ni \xi (t).$$
Now suppose \( x \in H \) and let
\[
x(s) + z(s) = x \tag{10.8.72}
\]
where \( z(s) \in \partial_2 \phi(s, x(s)) \), so \( x(s) = J_1(s)x \). Take the inner product of both sides with \( x(s) - y(s) \) to obtain
\[
(x(s), x(s) - y(s))_H + (z(s), x(s) - y(s))_H = (x, x(s) - y(s))_H
\]
and therefore,
\[
\frac{1}{2} |x(s)|^2_H - \frac{1}{2} |y(s)|^2_H \leq \phi(s, y(s)) - \phi(s, x(s))
\]
\[
+ |x|_H |x(s)|_H + |y|_H |y(s)|_H \leq \frac{1}{4} |x(s)|^2_H + c|x|_H^2 + \frac{1}{2} |y(s)|^2_H
\]
\[
+ \phi(s, y(s)) - \phi(s, x(s)).
\]
Consequently,
\[
\phi(s, x(s)) + \frac{1}{4} |x(s)|^2_H \leq |y(s)|^2_H + c|x|_H^2 + \phi(s, y(s)) < C \tag{10.8.73}
\]
a constant depending on \( x \). Replacing \( s \) with \( t \) in (10.8.74) and subtracting yields
\[
x(t) - x(s) + z(t) - z(s) = 0.
\]
Now taking the inner product of this with \( x(t) - x(s) \) it follows from (10.8.75).
\[
|x(s) - x(t)|^2_H = (z(s) - z(t), x(t) - x(s))_H
\]
\[
\leq \phi(s, x(t)) - \phi(s, x(s)) + \phi(t, x(s)) - \phi(t, x(t))
\]
\[
\leq \left( K \left[ \phi(t, x(t)) + |x(t)|_H^2 + 1 \right] + K \left[ \phi(s, x(s)) + |x(s)|_H^2 + 1 \right] \right) |t - s|
\]
which shows by (10.8.73) that \( \cdot \) is Lipschitz continuous and is therefore measurable which verifies (10.8.70). The assumptions of the lemma include (10.8.74). It follows the conclusion of Lemma 10.8.10 holds.

**Remark 10.8.8** Note that if \( \phi(t, \cdot) \) has a minimum at \( \xi(t) \) and if \( t \to \xi(t) \) and \( t \to \phi(t, \xi(t)) \) are bounded and measurable, then
\[
\xi(t) + 0 = \xi(t)
\]
and \( 0 \in \partial_2 \phi(t, \xi(t)) \). Therefore, in this case \( J_1(t)\xi(t) = \xi(t) \) and so the hypotheses of Lemma 10.8.10 hold.

**Corollary 10.8.9** Assume (10.8.74) - (10.8.77) and (10.8.70). Let \( x_0 \in D \) and let \( [f] \in L^2(0, T; H) \). Then there exists a unique function, \( x \), satisfying
\[
x \text{ and } [x'] \text{ are in } L^2(0, T; H)
\]
which is a solution to
\[
x' + \partial_2 \phi(t, x) \ni f \text{ a.e., } x(0) = x_0, \ x(t) = x_0 + \int_0^t x'(s) \, ds \tag{10.8.74}
\]
10.8. AN EVOLUTION INCLUSION

**Proof:** Let \([v] \in L^2 (0, T; H)\) and let \([x]\) be the unique solution to
\[
L [x] + [x] + \partial \Phi ([x]) \ni [f] + [v].
\] (10.8.75)

Letting \([x_i]\) be the solution corresponding to (10.8.75) in which \(v\) is replaced with \(v_i\), and \(x_i \in [x_i]\) is such that
\[
x_i (t) = x_0 + \int_0^t x_i' (s) \, ds, \quad i = 1, 2,
\]
from Lemma 10.8.6 and 10.8.7 that for each \(t \in [0, T]\),
\[
\frac{1}{2} |x_1 (t) - x_2 (t)|_H^2 + \frac{1}{2} \int_0^t |x_1 - x_2|_H^2 \, ds \leq \frac{1}{2} \int_0^t |v_1 (s) - v_2 (s)|_H^2 \, ds
\]
and so
\[
|x_1 (t) - x_2 (t)|_H^2 \leq \int_0^t |v_1 (s) - v_2 (s)|_H^2 \, ds.
\]

Now define a mapping, \(\Lambda : L^2 (0, T; H) \to L^2 (0, T; H)\) by \(\Lambda [v] = [x]\) where \([x]\) is the solution to (10.8.75). Then, if \([v_i]\) is in \(L^2 (0, T; H)\) and \([x_i]\) is the corresponding solution to (10.8.75),
\[
||\Lambda [v_1] - \Lambda [v_2]||_{L^2 (0, T; H)}^2 \equiv \int_0^t |x_1 (s) - x_2 (s)|_H^2 \, ds \leq \int_0^t \int_0^s |v_1 (r) - v_2 (r)|_H^2 \, dr \, ds.
\]

Iterating this inequality, by replacing \(\Lambda\) with \(\Lambda^k\), it follows that for all \(k\) large enough, \(\Lambda^k\) is a contraction map on \(L^2 (0, T; H)\). Thus there exists a unique fixed point for \(\Lambda, [x]\). Thus
\[
L [x] + [x] + \partial \Phi ([x]) \ni [f] + [x].
\]

Let \(x \in [x]\) be such that
\[
x (t) = x_0 + \int_0^t x' (s) \, ds.
\]
By Lemma 10.8.7
\[
x' + x + \partial_2 \phi (t, x) \ni f + x
\]
This function, \(x (\cdot)\) is the unique solution to (10.8.74) because if \(x_1\) is another solution, then \([x_1] = [x]\) and since both functions are continuous, they must coincide. \(\blacksquare\)
CHAPTER 10. MAXIMAL MONOTONE OPERATORS, HILBERT SPACE
Chapter 11

The Yankov von Neumann Aumann theorem

The Yankov von Neumann Aumann theorem deals with the projection of a product measurable set. It is a very difficult but interesting theorem. The material of this chapter is taken from [13], [14], [4], and [34]. We use the standard notation that for $S$ and $F$ σ algebras, $S \times F$ is the σ algebra generated by the measurable rectangles, the product measure σ algebra. The next result is fairly easy and the proof is left for the reader.

**Lemma 11.0.10** Let $(X, d)$ be a metric space. Then if $d_1(x, y) = \frac{d(x, y)}{1 + d(x, y)}$, it follows that $d_1$ is a metric on $X$ and the basis of open balls taken with respect to $d_1$ yields the same topology as the basis of open balls taken with respect to $d$.

**Theorem 11.0.11** Let $(X_i, d_i)$ denote a complete metric space and let $X \equiv \prod_{i=1}^{\infty} X_i$. Then $X$ is also a complete metric space with the metric

$$
\rho(x, y) \equiv \sum_{i=1}^{\infty} 2^{-i} \frac{d_i(x_i, y_i)}{1 + d_i(x_i, y_i)}.
$$

Also, if $X_i$ is separable for each $i$ then so is $X$.

**Proof:** It is clear from the above lemma that $\rho$ is a metric on $X$. We need to verify $X$ is complete with this metric. Let $\{x^n\}$ be a Cauchy sequence in $X$. Then it is clear from the definition that $\{x^n_i\}$ is a Cauchy sequence for each $i$ and converges to $x_i \in X_i$. Therefore, letting $\varepsilon > 0$ be given, we choose $N$ such that

$$
\sum_{k=N}^{\infty} 2^{-k} < \frac{\varepsilon}{2},
$$

we choose $M$ large enough that for $n > M$,

$$
2^{-i} \frac{d_i(x^n_i, x_i)}{1 + d_i(x^n_i, x_i)} < \frac{\varepsilon}{2(N + 1)}
$$

431
A complete separable metric space is called a Polish space. Let \( k \) for accomplished by finding to the metric, \( \rho \). \( \prod \) dense in \( D \) can see above theorem. Thus for \( n \) is onto and continuous. Here \( N \) Theorem 11.0.13 Definition 11.0.12 D \( \equiv \cup_{k=1}^{\infty} D_k \). Thus \( D_k \) is a countable subset of \( X_i \), \( D_i \equiv \{ r_k^i \}_{k=1}^{\infty} \). Then let

\[
D_k \equiv D_1 \times \cdots \times D_k \times \{ r_1^{k+1} \} \times \{ r_1^{k+2} \} \times \cdots
\]

Thus \( D_k \) is a countable subset of \( X \). Let \( D \equiv \cup_{k=1}^{\infty} D_k \). Then \( D \) is countable and we can see \( D \) is dense in \( X \) as follows. The projection of \( D_k \) onto the first \( k \) entries is dense in \( \prod_{i=1}^{k} X_i \) and for \( k \) large enough the remaining component's contribution to the metric, \( \rho \) is very small. Therefore, obtaining \( d \in D \) close to \( x \in X \) may be accomplished by finding \( d \in D \) such that \( d \) is close to \( x \) in the first \( k \) components for \( k \) large enough. Note that we do not use \( \prod_{k=1}^{\infty} D_k \)!

**Definition 11.0.12** A complete separable metric space is called a Polish space.

**Theorem 11.0.13** Let \( X \) be a Polish space. Then there exists \( f : \mathbb{N}^N \to X \) which is onto and continuous. Here \( \mathbb{N}^N \equiv \prod_{i=1}^{\infty} \mathbb{N} \) and a metric is given according to the above theorem. Thus for \( n, m \in \mathbb{N}^N \),

\[
\rho (n, m) \equiv \sum_{i=1}^{\infty} 2^{-i} \frac{|n_i - m_i|}{1 + |n_i - m_i|}.
\]

**Proof:** Since \( X \) is Polish, there exists a countable covering of \( X \) by closed sets having diameters no larger than \( 2^{-1}, \{ B(i) \}_{i=1}^{\infty} \). Each of these closed sets is also a Polish space and so there exists a countable covering of \( B(i) \) by a countable collection of closed sets, \( \{ B(i, j) \} \). \( \cup_{j=1}^{\infty} \) each having diameter no larger than \( 2^{-2} \) where \( B(i, j) \subset B(i) \neq \emptyset \) for all \( j \). Continue this way. Thus

\[
B(n_1, n_2, \cdots, n_m, i) = \cup_{j=1}^{\infty} B(n_1, n_2, \cdots, n_m, i)
\]

and each of \( B(n_1, n_2, \cdots, n_m, i) \) is a closed set contained in \( B(n_1, n_2, \cdots, n_m) \) whose diameter is at most half of the diameter of \( B(n_1, n_2, \cdots, n_m) \). Now we define our mapping from \( \mathbb{N}^N \) to \( X \). If \( n = \{ n_k \}_{k=1}^{\infty} \in \mathbb{N}^N \), we let \( f(n) \equiv \cap_{m=1}^{\infty} B(n_1, n_2, \cdots, n_m) \). Since the diameters of these sets converge to 0, there exists a unique point in this countable intersection and this is \( f(n) \).

We need to verify \( f \) is continuous. Let \( n \in \mathbb{N}^N \) be given and suppose \( m \) is very close to \( n \). The only way this can occur is for \( n_k \) to coincide with \( m_k \) for many \( k \). Therefore, both \( f(n) \) and \( f(m) \) must be contained in \( B(n_1, n_2, \cdots, n_m) \) for some fairly large \( m \). This implies, from the above construction that \( f(m) \) is as close to \( f(n) \) as \( 2^{-m} \), proving \( f \) is continuous. To see that \( f \) is onto, note that from the construction, if \( x \in X \), then \( x \in B(n_1, n_2, \cdots, n_m) \) for some choice of \( n_1, n_2, \cdots, n_m \) for each \( m \). Note nothing is said about \( f \) being one to one. It probably is not one to one.
Definition 11.0.14 We call a topological space, \( X \), a Suslin space if \( X \) is a Hausdorff space and there exists a Polish space, \( Z \) and a continuous function \( f \) which maps \( Z \) onto \( X \). These are also called analytic sets in some contexts but we will use the term Suslin space in referring to them.

Corollary 11.0.15 \( X \) is a Suslin space, if and only if there exists a continuous mapping from \( \mathbb{N}^\mathbb{N} \) onto \( X \).

Proof: We know there exists a Polish space \( Z \) and a continuous function, \( h : Z \to X \) which is onto. By the above theorem there exists a continuous map, \( g : \mathbb{N}^\mathbb{N} \to Z \) which is onto. Then \( h \circ g \) is a continuous map from \( \mathbb{N}^\mathbb{N} \) onto \( X \). The “if” part of this theorem is accomplished by noting that \( \mathbb{N}^\mathbb{N} \) is a Polish space.

Lemma 11.0.16 Let \( X \) be a Suslin space and suppose \( X_i \) is a subspace of \( X \) which is also a Suslin space. Then \( \bigcup_{i=1}^\infty X_i \) and \( \bigcap_{i=1}^\infty X_i \) are also Suslin spaces. Also every Borel set in \( X \) is a Suslin space.

Proof: Let \( f_i : Z_i \to X_i \) where \( Z_i \) is a Polish space and \( f_i \) is continuous and onto. Without loss of generality we may assume the spaces \( Z_i \) are disjoint because if not, we could replace \( Z_i \) with \( Z_i \times \{i\} \). Now we define a metric, \( \rho \), for \( Z = \bigcup_{i=1}^\infty Z_i \) as follows.

\[
\rho(x, y) \equiv 1 \text{ if } x \in Z_i, y \in Z_k, i \neq k \\
\rho(x, y) = \frac{d_i(x, y)}{1 + d_i(x, y)} \text{ if } x, y \in Z_i.
\]

Here \( d_i \) is the metric on \( Z_i \). It is easy to verify \( \rho \) is a metric and that \( (Z, \rho) \) is a Polish space. Now we define \( f : Z \to \bigcup_{i=1}^\infty X_i \) as follows. For \( x \in Z_i \), \( f(x) \equiv f_i(x) \). This is well defined because the \( Z_i \) are disjoint. If \( y \) is very close to \( x \) it must be that \( x \) and \( y \) are in the same \( Z_i \) otherwise this could not happen. Therefore, continuity of \( f \) follows from continuity of \( f_i \). This shows countable unions of Suslin subspaces of a Suslin space are Suslin spaces.

If \( H \subseteq X \) is a closed subset, then, letting \( f : Z \to X \) be onto and continuous, it follows \( f : f^{-1}(H) \to H \) is onto and continuous. Since \( f^{-1}(H) \) is closed, it follows \( f^{-1}(H) \) is a Polish space. Therefore, \( H \) is a Suslin space.

Now we show countable intersections of Suslin spaces are Suslin. It is clear that \( \theta : \prod_{i=1}^\infty Z_i \to \prod_{i=1}^\infty X_i \) given by \( \theta(z) \equiv x = \{x_i\} \) where \( x_i = f_i(z_i) \) is continuous and onto. Therefore, \( \prod_{i=1}^\infty X_i \) is a Suslin space. Now let \( P = \{y \in \prod_{i=1}^\infty f_i(Z_i) : y_i = y_j \text{ for all } i, j\} \). Then \( P \) is a closed subspace of a Suslin space and so it is Suslin. Then we define \( h : P \to \bigcap_{i=1}^\infty X_i \) by \( h(y) \equiv f_i(y_i) \). This shows \( \bigcap_{i=1}^\infty X_i \) is Suslin because \( h \) is continuous and onto. \( (h \circ \theta : \theta^{-1}(P) \to \bigcap_{i=1}^\infty X_i \) is continuous and \( \theta^{-1}(P) \) being a closed subset of a Polish space is Polish.)

Next let \( U \) be an open subset of \( X \). Then \( f^{-1}(U) \), being an open subset of a Polish space, can be obtained as an increasing limit of closed sets, \( K_n \). Therefore, \( U = \bigcup_{n=1}^\infty f(K_n) \). Each \( f(K_n) \) is a Suslin space because it is the continuous image of a Polish space, \( K_n \). Therefore, by the first part of the lemma, \( U \) is a Suslin space.

Now let

\[
\mathcal{F} = \left\{ E \subseteq X : \text{ both } E^C \text{ and } E \text{ are Suslin} \right\}.
\]
We see that $\mathcal{F}$ is closed with respect to taking complements. The first part of this lemma shows $\mathcal{F}$ is closed with respect to countable unions. Therefore, $\mathcal{F}$ is a $\sigma$ algebra and so, since it contains the open sets, must contain the Borel sets.

It turns out that Suslin spaces tend to be measurable sets. In order to develop this idea, we need a technical lemma.

**Lemma 11.0.17** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and denote by $\mu^*$ the outer measure generated by $\mu$. Thus

$$\mu^*(S) \equiv \inf \left\{ \mu(E) : E \supseteq S, \ E \in \mathcal{F} \right\}.$$  

Then $\mu^*$ is regular, meaning that for every $S$, there exists $E \in \mathcal{F}$ such that $E \supseteq S$ and $\mu(E) = \mu^*(S)$. If $S_n \uparrow S$, it follows that $\mu^*(S_n) \uparrow \mu^*(S)$. Also if $\mu(\Omega) < \infty$, then a set, $E$ is measurable if and only if

$$\mu^*(\Omega) \geq \mu^*(E) + \mu^*(\Omega \setminus E).$$

**Proof:** First we verify that $\mu^*$ is regular. If $\mu^*(S) = \infty$, let $E = \Omega$. Then $\mu^*(S) = \mu(E)$ and $E \supseteq S$. On the other hand, if $\mu^*(S) < \infty$, then we can obtain $E_n \in \mathcal{F}$ such that $\mu^*(S) + \frac{1}{n} \geq \mu(E_n)$ and $E_n \supseteq S$. Now let $F_n = \bigcap_{i=1}^n E_i$. Then $F_n \supseteq S$ and so $\mu^*(S) + \frac{1}{n} \geq \mu(F_n) \geq \mu^*(S)$. Therefore, letting $F = \bigcap_{k=1}^\infty F_k \in \mathcal{F}$ it follows $\mu(F) = \lim_{n \to \infty} \mu(F_n) = \mu^*(S)$.

Let $E_n \supseteq S_n$ be such that $E_n \in \mathcal{F}$ and $\mu(E_n) = \mu^*(S_n)$. Also let $E_\infty \supseteq S$ such that $\mu(E_\infty) = \mu^*(S)$ and $E_\infty \in \mathcal{F}$. Now consider $B_n \equiv \bigcup_{k=1}^n E_k$. We claim

$$\mu(B_n) = \mu(S_n).$$

(11.0.1)

Here is why:

$$\mu(E_1 \setminus E_2) = \mu(E_1) - \mu(E_1 \cap E_2) = \mu^*(S_1) - \mu^*(S_1) = 0.$$  

Therefore,

$$\mu(B_2) = \mu(E_1 \cup E_2) = \mu(E_1 \setminus E_2) + \mu(E_2) = \mu(E_2) = \mu^*(S_2).$$

Continuing in this way we see that $\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \setminus E)$. Let $F \supseteq S$, $F \in \mathcal{F}$, and $\mu(F) = \mu^*(S)$. Then since $\mu^*$ is subadditive,

$$\mu^*(\Omega \setminus F) \leq \mu^*(E \setminus F) + \mu^*(\Omega \cap E \cap F^c).$$

(11.0.2)
Since $F$ is measurable,

$$\mu^*(E) = \mu^*(E \cap F) + \mu^*(E \setminus F)$$

(11.0.3)

and

$$\mu^*(\Omega \setminus E) = \mu^*(F \setminus E) + \mu^*(\Omega \cap E^C \cap F^C)$$

(11.0.4)

and by the hypothesis,

$$\mu^*(\Omega) \geq \mu^*(E) + \mu^*(\Omega \setminus E).$$

(11.0.5)

Therefore,

$$\mu(\Omega) \geq \mu^*(E) + \mu^*(\Omega \setminus E)$$

$$\mu(\Omega) \geq \mu^*(E \cap F) + \mu^*(E \setminus F) + \mu^*(\Omega \setminus E)$$

$$\geq \mu^*(\Omega \setminus F) + \mu^*(F \setminus E) + \mu^*(E \cap F)$$

$$\geq \mu^*(\Omega \setminus F) + \mu^*(F) = \mu(\Omega)$$

showing that all the inequalities must be equal signs. Hence, referring to the top and fourth lines above,

$$\mu(\Omega) = \mu^*(E \cap F) + \mu^*(F \setminus E) + \mu^*(E \cap F).$$

Subtracting $\mu^*(\Omega \setminus F) = \mu(\Omega \setminus F)$ from both sides gives

$$\mu^*(S) = \mu(F) = \mu^*(F \setminus E) + \mu^*(E \cap F) \geq \mu^*(S \setminus E) + \mu^*(E \cap S).$$

The next theorem is a major result. It states that the Suslin subsets are measurable under appropriate conditions.

**Theorem 11.0.18** Let $\Omega$ be a metric space and let $(\Omega, \mathcal{F}, \mu)$ be a complete Borel measure space with $\mu(\Omega) < \infty$. Denote by $\mu^*$ the outer measure generated by $\mu$. Then if $A$ is a Suslin subset of $\Omega$, it follows that $A$ is $\mu^*$ measurable.

**Proof:** We need to verify that

$$\mu^*(\Omega) \geq \mu^*(A) + \mu^*(\Omega \setminus A).$$

We know from Corollary 11.0.13, there exists a continuous map, $f : \mathbb{N}^\mathbb{N} \rightarrow A$ which is onto. Let

$$E(k) \equiv \{ n \in \mathbb{N}^\mathbb{N} : n_1 \leq k \}.$$

Then $E(k) \uparrow \mathbb{N}^\mathbb{N}$ and so from Lemma 11.0.14 we know $\mu^*(f(E(k))) \uparrow \mu^*(A)$. Therefore, there exists $m_1$ such that

$$\mu^*(f(E(m_1))) > \mu^*(A) - \frac{\varepsilon}{2}.$$
CHAPTER 11. THE YANKOV VON NEUMANN AUßMANN THEOREM

Now $E(k)$ is clearly not compact but it is trying to be as far as the first component is concerned. Now we let

$$E(m_1, k) \equiv \{ n \in \mathbb{N}^2 : n_1 \leq m_1 \text{ and } n_2 \leq k \}.$$ 

Thus $E(m_1, k) \uparrow E(m_1)$ and so we can pick $m_2$ such that

$$\mu^*(f(E(m_1, m_2))) > \mu^*(f(E(m_1))) - \frac{\varepsilon}{2^2}.$$ 

We continue in this way obtaining a decreasing list of sets,

$$f(E(m_1, m_2, \ldots, m_k)),$$

such that

$$\mu^*(f(E(m_1, m_2, \ldots, m_k))) > \mu^*(f(E(m_1, m_2, \ldots, m(k-1)))) - \frac{\varepsilon}{2^k}.$$ 

Therefore,

$$\mu^*(f(E(m_1, m_2, \ldots, m_k))) - \mu^*(A) > \sum_{l=1}^{k} - \left(\frac{\varepsilon}{2^l}\right) > -\varepsilon.$$ 

Now define a closed set,

$$C \equiv \bigcap_{k=1}^{\infty} f(E(m_1, m_2, \ldots, m_k)).$$ 

The sets $f(E(m_1, m_2, \ldots, m(k-1), m_k))$ are decreasing as $k \to \infty$ and so

$$\mu^*(C) = \lim_{k \to \infty} \mu^*\left(f(E(m_1, m_2, \ldots, m(k-1), m_k))\right) \geq \mu^*(A) - \varepsilon.$$ 

We wish to verify that $C \subseteq A$. If we can do this we will be done because $C$, being a closed set, is measurable and so

$$\mu^*(\Omega) = \mu^*(C) + \mu^*(\Omega \setminus C) \geq \mu^*(A) - \varepsilon + \mu^*(\Omega \setminus A).$$ 

Since $\varepsilon$ is arbitrary, this will conclude the proof. Therefore, we only need to verify that $C \subseteq A$.

What we know is that each $f(E(m_1, m_2, \ldots, m(k-1), m_k))$ is contained in $A$. We do not know their closures are contained in $A$. We let $m \equiv \{m_i\}_{i=1}^{\infty}$ where the $m_i$ are defined above. Then letting

$$K \equiv \{ n \in \mathbb{N}^2 : n_i \leq m_i \text{ for all } i \},$$ 

we see that $K$ is a closed, hence complete subset of $\mathbb{N}^2$ which is also totally bounded due to the definition of the distance. Therefore, $K$ is compact and so $f(K)$ is also compact, hence closed due to the assumption that $\Omega$ is a Hausdorff space and we know that $f(K) \subseteq A$. We verify that $C = f(K)$. We know $f(K) \subseteq C$. 

Suppose therefore, \( p \in C \). From the definition of \( C \), we know there exists \( r^k \in E (m_1, m_2, \ldots, m_k) \) such that \( d(f(r^k), p) < \frac{1}{k} \). Denote by \( \mathbf{r}^k \) the element of \( \mathbb{N}^\mathbb{N} \) which consists of modifying \( r^k \) by taking all components after the \( k^{th} \) equal to one. Thus \( \mathbf{r}^k \in K \). Now \( \{r^k\} \) is in a compact set and so taking a subsequence we can have \( \mathbf{r}^k \to \mathbf{r} \in K \). But from the metric on \( \mathbb{N}^\mathbb{N} \), it follows that \( \rho(\mathbf{r}^k, \mathbf{r}) < \frac{1}{2k-2} \). Therefore, \( r^k \to r \) also and so \( f(r^k) \to f(r) = p \). Therefore, \( p \in f(K) \) and this proves the theorem.

Note we could have proved this under weaker assumptions. If we had assumed only that every point has a countable basis (first axiom of countability) and \( \Omega \) is Hausdorff, the same argument would work. We will need the following definition.

**Definition 11.0.19** Let \( F \) be a \( \sigma \) algebra of sets from \( \Omega \) and let \( \mu \) denote a finite measure defined on \( F \). We let \( F_\mu \) denote the completion of \( F \) with respect to \( \mu \). Thus we let \( \mu^* \) be the outer measure determined by \( \mu \) and \( F_\mu \) will be the \( \sigma \) algebra of \( \mu^* \) measurable subsets of \( \Omega \). We also define \( bF \) by

\[
\mathcal{F} = \cap \{ F_\mu : \mu \text{ is a finite measure defined on } F \}.
\]

Also, if \( X \) is a topological space, we will denote by \( B(X) \) the Borel sets of \( X \).

With this notation, we can give the following simple corollary of Theorem

**Corollary 11.0.20** Let \( \Omega \) be a compact metric space and let \( A \) be a Suslin subset of \( \Omega \). Then \( A \in B(\Omega) \).

**Proof**: Let \( \mu \) be a finite measure defined on \( B(\Omega) \). By Theorem 11.0.18 \( A \in B(\Omega)_\mu \). Since this is true for every finite measure, \( \mu \), it follows \( A \in B(\Omega) \) as claimed.

We give another technical lemma about the completion of measure spaces.

**Lemma 11.0.21** Let \( \mu \) be a finite measure on a \( \sigma \) algebra, \( \Sigma \). Then \( A \in \Sigma_\mu \) if and only if there exists \( A_1 \in \Sigma \) and \( N_1 \) such that \( A = A_1 \cup N_1 \) where there exists \( N \in \Sigma \) such that \( \mu(N) = 0 \) and \( N_1 \subseteq N \).

**Proof**: Suppose first \( A = A_1 \cup N_1 \) where these sets are as described. Let \( S \in \mathcal{P}(\Omega) \) and let \( \mu^* \) denote the outer measure determined by \( \mu \). Then since \( A_1 \in \Sigma \subseteq \Sigma_\mu \)

\[
\mu^*(S) \leq \mu^*(S \setminus A) + \mu^*(S \cap A) \\
\leq \mu^*(S \setminus A_1) + \mu^*(S \cap A_1) + \mu^*(N_1) \\
= \mu^*(S \setminus A_1) + \mu^*(S \cap A_1) = \mu^*(S)
\]

showing that \( A \in \Sigma_\mu \).
Now suppose \( A \in \Sigma_\mu \). Then there exists \( B_1 \supseteq A \) such that \( \mu^*(B_1) = \mu^*(A) \), and \( B_1 \in \Sigma \). Also there exists \( A_1^C \in \Sigma \) with \( A_1^C \supseteq A^C \) and \( \mu(A_1^C) = \mu^*(A^C) \). Then \( A_1 \subseteq A \subseteq B_1 \)

\[
A \subseteq A_1 \cup (B_1 \setminus A_1).
\]

Now

\[
\mu(A_1) + \mu^*(A^C) = \mu(A_1) + \mu(A_1^C) = \mu(\Omega)
\]

and so

\[
\mu(B_1 \setminus A_1) = \mu^*(B_1 \setminus A_1) = \mu^*(B_1 \setminus A) + \mu^*(A \setminus A_1) = \mu^*(B_1) - \mu^*(A) + \mu^*(A) - \mu^*(A_1) = \mu^*(A) - \mu^*(A^C) = 0
\]

because \( A \in \Sigma_\mu \) implying \( A = A_1 \cup (B_1 \setminus A_1) \cap A \) and \( N_1 \subseteq N = (B_1 \setminus A_1) \in \Sigma \) with \( \mu(N) = 0 \). ■

Next we need another definition.

**Definition 11.0.22** We say \((\Omega, \Sigma)\), where \( \Sigma \) is a \( \sigma \) algebra of subsets of \( \Omega \), is separable if there exists a sequence \( \{A_n\}_{n=1}^\infty \subseteq \Sigma \) such that \( \sigma(\{A_n\}) = \Sigma \) and if \( w \neq w' \), then there exists \( A \in \Sigma \) such that \( X_A(\omega) \neq X_A(\omega') \). This last condition is referred to by saying \( \{A_n\} \) separates the points of \( \Omega \). Given two measure spaces, \((\Omega, \Sigma)\) and \((\Omega', \Sigma')\), we say they are isomorphic if there exists a function, \( f : \Omega \to \Omega' \) which is one to one and \( f(E) \in \Sigma' \) whenever \( E \in \Sigma \) and \( f^{-1}(F) \in \Sigma \) whenever \( F \in \Sigma' \).

The interesting thing about separable measure spaces is that they are isomorphic to a very simple sort of measure space in which topology plays a significant role.

**Lemma 11.0.23** Let \((\Omega, \Sigma)\) be separable. Then there exists \( E \in \{0, 1\}^N \) such that \((\Omega, \Sigma)\) and \((E, B(\Omega))\) are isomorphic.

**Proof:** First we show \( \{A_n\} \) separates the points. We already know \( \Sigma \) separates the points. If this is not so, there exists \( \omega, \omega_1 \in \Omega \) such that for all \( n, X_{A_n}(\omega) = X_{A_n}(\omega_1) \). Then let

\[
\mathcal{F} = \{F \in \Sigma : X_F(\omega) = X_F(\omega_1)\}
\]

Thus \( A_n \in \mathcal{F} \) for all \( n \). It is also clear that \( \mathcal{F} \) is a \( \sigma \) algebra and so \( \mathcal{F} = \Sigma \) contradicting the assumption that \( \Sigma \) separates points. Now we define a function from \( \Omega \) to \( \{0, 1\}^N \) as follows.

\[
f(\omega) = \{X_{A_n}(\omega)\}_{n=1}^\infty
\]

We also let \( E \equiv f(\Omega) \). Since the \( \{A_n\} \) separate the points, we see that \( f \) is one to one. A subbasis for the topology of \( \{0, 1\}^N \) consists of sets of the form \( \prod_{i=1}^\infty H_i \)
where \( H_i = \{0, 1\} \) for all \( i \) except one, when \( i = j \) and \( H_j \) equals either \( \{0\} \) or \( \{1\} \). Therefore, \( f^{-1}(\text{subbasic open set}) \in \Sigma \) because if \( H_j \) is the exceptional set then this equals \( A_j \) if \( H_j = \{1\} \) and \( A_j^c \) if \( H_j = \{0\} \). Intersections of these subbasic sets with \( E \) gives a countable subbasis for \( E \) and so the inverse image of all sets in a countable subbasis for \( E \) are in \( \Sigma \), showing that \( f^{-1}(\text{open set}) \in \Sigma \). Now we consider \( f(A_n) \).

\[
f(A_n) \equiv \{\{\lambda_k\}_{k=1}^{\infty} : \lambda_n = 1\} \cap E,
\]

an open set in \( E \). Hence \( f(A_n) \in B(E) \). Now letting

\[
\mathcal{F} \equiv \{G \subseteq \Omega : f(G) \in B(E)\},
\]

we see that \( \mathcal{F} \) is a \( \sigma \) algebra which contains \( \{A_n\}_{n=1}^{\infty} \) and so \( \mathcal{F} \supseteq \sigma(\{A_n\}) = \Sigma \). Thus \( f(F) \in B(E) \) for all \( A \in \Sigma \). ■

**Lemma 11.0.24** Let \( \phi : (\Omega_1, \Sigma_1) \rightarrow (\Omega_2, \Sigma_2) \) where \( \phi^{-1}(U) \in \Sigma_1 \) for all \( U \in \Sigma_2 \). Then if \( F \in \Sigma_2 \), it follows \( \phi^{-1}(F) \in \Sigma_1 \).

**Proof:** Let \( \mu \) be a finite measure on \( \Sigma_1 \) and define a measure \( \phi(\mu) \) on \( \Sigma_2 \) by the rule

\[
\phi(\mu)(F) \equiv \mu(\phi^{-1}(F)).
\]

Now let \( A \in \Sigma_{2\phi(\mu)} \). Then by Lemma [11.0.11], \( A = A_1 \cup N_1 \) where there exists \( N \in \Sigma_2 \) with \( \phi(\mu)(N) = 0 \) and \( A_1 \in \Sigma_2 \). Therefore, from the definition of \( \phi(\mu) \), we have \( \mu(\phi^{-1}(N)) = 0 \) and therefore, \( \phi^{-1}(A) = \phi^{-1}(A_1) \cup \phi^{-1}(N_1) \) where \( \phi^{-1}(N_1) \subseteq \phi^{-1}(N) \in \Sigma_1 \) and \( \mu(\phi^{-1}(N)) = 0 \). Therefore, \( \phi^{-1}(A) \in \Sigma_{1\mu} \) and so if \( F \in \Sigma_2 \), then

\[
F \in \cap \{\Sigma_{2\nu} : \nu \text{ is a finite measure on } \Sigma_2\} \subseteq \cap \{\Sigma_{2\phi(\mu)} : \mu \text{ is a finite measure on } \Sigma_1\},
\]

and so \( \phi^{-1}(F) \in \Sigma_{1\mu} \). Since \( \mu \) is arbitrary, this shows \( \phi^{-1}(F) \in \Sigma_{1\mu} \).

The next lemma is a special case of the Yankov von Neumann Aumann projection theorem. It contains the main idea of the proof of the more general theorem.

**Lemma 11.0.25** Let \( (\Omega, \Sigma) \) be separable and let \( X \) be a Suslin space. Let \( G \in \Sigma \times B(X) \). \( (\Sigma \times B(X) \) is the \( \sigma \) algebra of product measurable sets, the smallest \( \sigma \) algebra containing the measurable rectangles.) Then

\[
\text{pro}_{\Omega}(G) \in \widehat{\Sigma}.
\]

**Proof:** Let \( f : (\Omega, \Sigma) \rightarrow (E, B(E)) \) be the isomorphism of Lemma [11.0.23]. We have the following claim.

**Claim:** \( f \times I_X \) maps \( \Sigma \times B(X) \) to \( B(E) \times B(X) \).

**Proof of the claim:** First of all, assume \( A \times B \) is a measurable rectangle where \( A \in \Sigma \) and \( B \in B(X) \). Then by the assumption that \( f \) is an isomorphism, \( f(A) \in B(E) \) and so

\[
f \times I_X (A \times B) \in B(E) \times B(X).
\]
Now let

\[ \mathcal{F} \equiv \{ P \in \Sigma \times B (X) : f \times I_X (P) \in B (E) \times B (X) \} . \]

Then we see that \( \mathcal{F} \) is a \( \sigma \) algebra and contains the elementary sets. (\( \mathcal{F} \) is closed with respect to complements because \( f \) is one to one.) Therefore, \( \mathcal{F} = \Sigma \times B (X) \) and this proves the claim.

Therefore, since \( G \in \Sigma \times B (X) \), we see

\[ f \times I_X (G) \in B (E) \times B (X) \subseteq B (E \times X) . \]

The set inclusion follows from the observation that if \( A \in B (E) \) and \( B \in B (X) \) then \( A \times B \) is in \( B (E \times X) \) and the collection of sets in \( B (E) \times B (X) \) which are in \( B (E \times X) \) is a \( \sigma \) algebra.

Therefore, there exists \( D \), a Borel set in \( E \times X \) such that \( f \times I_X (G) = D \cap (E \times X) \). Now from this it follows from Lemma \ref{lem:yvna11.0.19} that \( D \) is a Suslin space. Letting \( Y \) be \( \{0,1\}^\mathbb{N} \), it follows that \( \text{proj}_Y (D) \) is a Suslin space in \( Y \). By Corollary \ref{cor:yvna11.0.20}, we see that \( \text{proj}_Y (D) \subseteq B (Y) \). Now

\[ \text{proj}_Y (G) = \{ \omega \in \Omega : \text{there exists } x \in X \text{ with } (\omega, x) \in G \} \]

\[ = \{ \omega \in \Omega : \text{there exists } x \in X \text{ with } (f (\omega), x) \in f \times I_X (G) \} \]

\[ = f^{-1}(\{ y \in Y : \text{there exists } x \in X \text{ with } (y, x) \in D \}) \]

\[ = f^{-1}(\text{proj}_Y (D)) . \]

Now \( \text{proj}_Y (D) \subseteq B (Y) \) and so Lemma \ref{lem:yvna11.0.21} shows \( f^{-1}(\text{proj}_Y (D)) \subseteq \hat{\Sigma} \).

Now we are ready to prove the Yankov von Neumann Aumann projection theorem. First we must present another technical lemma.

**Lemma 11.0.26** Let \( X \) be a Hausdorff space and let \( G \in \Sigma \times B (X) \) where \( \Sigma \) is a \( \sigma \) algebra of sets of \( \Omega \). Then there exists \( \Sigma_0 \subseteq \Sigma \) a countably generated \( \sigma \) algebra such that \( G \in \Sigma_0 \times B (X) \).

**Proof:** First suppose \( G \) is a measurable rectangle, \( G = A \times B \) where \( A \in \Sigma \) and \( B \in B (X) \). Letting \( \Sigma_0 \) be the finite \( \sigma \) algebra, \( \{ \emptyset, A, A^c, \Omega \} \), we see that \( G \in \Sigma_0 \times B (X) \). Similarly, if \( G \) equals an elementary set, then the conclusion of the lemma holds for \( G \). Let

\[ \mathcal{F} \equiv \{ H \in \Sigma \times B (X) : H \in \Sigma_0 \times B (X) \} \]

for some countably generated \( \sigma \) algebra, \( \Sigma_0 \). We just saw that \( \mathcal{F} \) contains the elementary sets. If \( H \in \mathcal{F} \), then \( H^c \in \Sigma_0 \times B (X) \) for the same \( \Sigma_0 \) and so \( \mathcal{F} \) is closed with respect to complements. Now suppose \( H_n \in \mathcal{F} \). Then for each \( n \), there exists a countably generated \( \sigma \) algebra, \( \Sigma_0_n \) such that \( H_n \in \Sigma_0_n \times B (X) \). Then \( \bigcup_{n=1}^{\infty} H_n \in \sigma (\{ \Sigma_0_n \times B (X) \}) \). We will be done when we show

\[ \sigma (\{ \Sigma_0_n \times B (X) \}_{n=1}^{\infty}) \subseteq \sigma (\{ \Sigma_0_n \}_{n=1}^{\infty}) \times B (X) \]
because it is clear that \( \sigma (\{\Sigma_n\}_{n=1}^\infty) \) is countably generated. We see that
\[
\sigma (\{\Sigma_n \times B (X)\}_{n=1}^\infty)
\]
is generated by sets of the form \( A \times B \) where \( A \in \Sigma_n \) and \( B \in B (X) \). But each such set is also contained in \( \sigma (\{\Sigma_n\}_{n=1}^\infty \times B (X)) \) and so the desired inclusion is obtained. Therefore, \( \mathcal{F} \) is a \( \sigma \) algebra and so since \( \mathcal{F} \) was shown to contain the measurable rectangles, this verifies \( \mathcal{F} = \Sigma \times B (X) \) and this proves the lemma.

**Theorem 11.0.27** Let \((\Omega, \Sigma)\) be a measure space and let \( G \in \Sigma \times B (X) \) where \( X \) is a Suslin space. Then

\[
\text{proj}_\Omega (G) \in \hat{\Sigma}.
\]

**Proof:** By the previous lemma, \( G \in \Sigma_0 \times B (X) \) where \( \Sigma_0 \) is countably generated. If \((\Omega, \Sigma_0)\) were separable, we could then apply Lemma 11.0.25 and be done. Unfortunately, we don’t know \( \Sigma_0 \) separates the points of \( \Omega \). Therefore, we define an equivalence class on the points of \( \Omega \) as follows. We say \( \omega \sim \omega_1 \) if and only if \( \mathcal{X}_A (\omega) = \mathcal{X}_A (\omega_1) \) for all \( A \in \Sigma_0 \). Now the nice thing to notice about this equivalence relation is that if \( \omega \in A \in \Sigma_0 \), and if \( \omega \sim \omega_1 \), then \( 1 = \mathcal{X}_A (\omega) = \mathcal{X}_A (\omega_1) \) implying \( \omega_1 \in A \) also. Therefore, every set of \( \Sigma_0 \) is the union of equivalence classes. It follows that for \( A \in \Sigma_0 \), and \( \pi \) the map given by \( \pi \omega \equiv [\omega] \) where \([\omega]\) is the equivalence class determined by \( \omega \),

\[
\pi (A) \cap \pi (\Omega \setminus A) = \emptyset.
\]
Suppose now that \( H_n \in \Sigma_0 \times B (X) \). If \( ([\omega], x) \in \cap_{n=1}^\infty \pi \times I_X (H_n) \), then for each \( n \),

\[
([\omega], x) = (\pi w_n, x)
\]
for some \((\omega_n, x) \in H_n\). But this implies \( \omega \sim \omega_n \) and so from the above observation that the sets of \( \Sigma_0 \) are unions of equivalence classes, it follows that \( (\omega, x) \in H_n \). Therefore, \((\omega, x) \in \cap_{n=1}^\infty H_n \) and so \( ([\omega], x) = \pi \times I_X (\omega, x) \) where \((\omega, x) \in \cap_{n=1}^\infty H_n \). This shows that
\[
\pi \times I_X (\cap_{n=1}^\infty H_n) \supseteq \cap_{n=1}^\infty \pi \times I_X (H_n).
\]
In fact these two sets are equal because the other inclusion is obvious. We will denote by \( \Omega_1 \) the set of equivalence classes and \( \Sigma_1 \) will be the subsets, \( S_1 \), of \( \Omega_1 \) such that \( S_1 = \{[\omega] : \omega \in S \in \Sigma_0\} \). Then \((\Omega_1, \Sigma_1)\) is clearly a measure space which is separable. Let

\[
\mathcal{F} \equiv \{H \in \Sigma_0 \times B (X) : \pi \times I_X (H), \pi \times I_X (H^C) \in \Sigma_1 \times B (X)\}.
\]
We see that the measurable rectangles, \( A \times B \) where \( A \in \Sigma_0 \) and \( B \in B (X) \) are in \( \mathcal{F} \), that from the above observation on countable intersections, \( \mathcal{F} \) is closed with respect to countable unions and closed with respect to complements. Therefore, \( \mathcal{F} \) is a \( \sigma \) algebra and so \( \mathcal{F} = \Sigma_0 \times B (X) \). By Lemma 11.0.25 \((\Omega_1, \Sigma_1)\) is isomorphic to \((E, B (E))\) where \( E \) is a subspace of \( \{0, 1\}^\mathbb{N} \). Denoting the isomorphism by \( h \),
it follows as in Lemma 11.0.25 that \( h \times I_X \) maps \( \Sigma_1 \times B(X) \) to \( B(E) \times B(X) \). Therefore, we see \( f \equiv h \circ \pi \) is a mapping from \( \Omega \) to \( E \times B(X) \) which has the property that \( f \times I_X \) maps \( \Sigma_0 \times B(X) \) to \( B(E) \times B(X) \). Now from the proof of Lemma 11.0.25 starting with the claim, we see that \( G \in \Sigma_0 \). However, if \( \mu \) is a finite measure on \( \Sigma \), then \( \Sigma_0 \subseteq \Sigma \) and so \( \Sigma_0 \subseteq \Sigma \).

\[ 11.1 \text{ Multifunctions And Their Measurability} \]

\subsection*{11.1.1 The General Case}

Let \( X \) be a separable complete metric space and let \( (\Omega, \mathcal{C}, \mu) \) be a set, a \( \sigma \)-algebra of subsets of \( \Omega \), and a measure \( \mu \) such that \( \mu \) is a complete \( \sigma \)-finite measure space. Also let \( \Gamma : \Omega \rightarrow \mathcal{P}(F(X)) \), the closed subsets of \( X \).

**Definition 11.1.1** We define \( \Gamma^{-}(S) = \{ \omega : \Gamma(\omega) \cap S \neq \emptyset \} \)

We will consider a theory of measurability of set valued functions. The following theorem is the main result in the subject. In this theorem the third condition is what we will refer to as measurable.

**Theorem 11.1.2** The following are equivalent.

1. For all \( B \) a Borel set in \( X \), \( \Gamma^{-}(B) \in \mathcal{C} \).
2. For all \( F \) closed in \( X \), \( \Gamma^{-}(F) \in \mathcal{C} \).
3. For all \( U \) open in \( X \), \( \Gamma^{-}(U) \in \mathcal{C} \).
4. There exists a sequence, \( \{\sigma_n\} \) of measurable functions satisfying \( \sigma_n(\omega) \in \Gamma(\omega) \) such that for all \( \omega \in \Omega \),
   \[ \Gamma(\omega) = \{\sigma_n(\omega) : n \in \mathbb{N}\} \]
   These functions are called measurable selections.
5. For all \( x \in X \), \( \omega \rightarrow \text{dist}(x, \Gamma(\omega)) \) is a measurable real valued function.
6. \( \mathcal{G}(\Gamma) \equiv \{ (\omega, x) : x \in \Gamma(\omega) \} \subseteq \mathcal{C} \times B(X) \).

**Proof:** It is obvious that 1.) \( \Rightarrow \) 2.). To see that 2.) \( \Rightarrow \) 3.) note that \( \Gamma^{-}(\bigcup_{i=1}^{\infty} F_i) = \bigcup_{i=1}^{\infty} \Gamma^{-}(F_i) \). Since any open set in \( X \) can be obtained as a countable union of closed sets, this implies 2.) \( \Rightarrow \) 3.).

Now we verify that 3.) \( \Rightarrow \) 4.). For convenience, drop the assumption that \( \Gamma(\omega) \) is closed in this part of the argument. It will just be set valued and satisfy the measurability condition. A measurable selection will be obtained in \( \Gamma(\omega) \). Let \( \{x_n\}_{n=1}^{\infty} \) be a countable dense subset of \( X \). For \( \omega \in \Omega \), let \( \psi_1(\omega) = x_n \) where \( n \) is
that $\sigma$ such that $a$ measurable set. By what was just shown, there exists a measurable selection, $\psi$ such that the conclusion of 4.) holds. To do this we define for $\Gamma (\omega )$ closed and measurable, let

$$\Omega_n \equiv \{ \omega \in \Omega : \psi_1 (\omega ) = x_n \}.$$ 

Then $\Omega_n \in C$ and $\Omega_n \cap \Omega_m = \emptyset$ for $n \neq m$ and $\bigcup_{n=1}^{\infty} \Omega_n = \Omega$. Let

$$D_n \equiv \{ x_k : x_k \in B (x_n, 1) \}.$$ 

Now for each $n$, and $\omega \in \Omega_n$, let $\psi_2 (\omega ) = x_k$ where $k$ is the smallest index such that $x_k \in D_n$ and $B \left( x_k, \frac{1}{2} \right) \cap \Gamma (\omega ) \neq \emptyset$. Thus $\text{dist} (\psi_2 (\omega ), \Gamma (\omega )) < \frac{1}{\delta}$ and

$$d (\psi_2 (\omega ), \psi_1 (\omega )) < 1.$$ 

Continue this way obtaining $\psi_k$ a measurable function such that

$$\text{dist} (\psi_k (\omega ), \Gamma (\omega )) < \frac{1}{2^{k-1}}, \quad d (\psi_k (\omega ), \psi_{k+1} (\omega )) < \frac{1}{2^{k-2}}.$$ 

Then for each $\omega$, $\{ \psi_k (\omega ) \}$ is a Cauchy sequence converging to a point, $\sigma (\omega ) \in \overline{\Gamma (\omega )}$. This has shown that if $\Gamma$ is measurable, there exists a measurable selection, $\sigma (\omega ) \in \overline{\Gamma (\omega )}$. Of course, if $\Gamma (\omega )$ is closed, then $\sigma (\omega ) \in \Gamma (\omega )$. Note that this had nothing to do with the measure. It remains to show there exists a sequence of these measurable selections $\sigma_n$ such that the conclusion of 4.) holds. To do this we define for $\Gamma (\omega )$ closed and measurable,

$$\Gamma_{ni} (\omega ) \equiv \begin{cases} \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) & \text{if } \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) \neq \emptyset, \\ \Gamma (\omega ) & \text{otherwise.} \end{cases}$$ 

Thus

$$\Gamma (\omega ) \cap B \left( x_n, 2^{-(i+1)} \right) \subseteq \Gamma_{ni} (\omega ) \subseteq \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right).$$ 

First we show that $\Gamma_{ni}$ is measurable. Let $U$ be open. Then

$$\{ \omega : \Gamma_{ni} (\omega ) \cap U \neq \emptyset \} = \{ \omega : \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) \cap U \neq \emptyset \} \cup$$

$$\left[ \{ \omega : \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) = \emptyset \} \cap \{ \omega : \Gamma (\omega ) \cap U \neq \emptyset \} \right]$$

$$= \{ \omega : \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) \cap U \neq \emptyset \} \cup$$

$$\left[ \{ \omega : \Gamma (\omega ) \cap B \left( x_n, 2^{-i} \right) \neq \emptyset \} \cap \{ \omega : \Gamma (\omega ) \cap U \neq \emptyset \} \right],$$

a measurable set. By what was just shown, there exists $\sigma_{ni}$, a measurable function such that $\sigma_{ni} (\omega ) \in \overline{\Gamma_{ni} (\omega )} \subseteq \Gamma (\omega )$ for all $\omega \in \Omega$. If $x \in \Gamma (\omega )$, then

$$x \in B \left( x_n, 2^{-(i+2)} \right)$$
whenever \( x_n \) is close enough to \( x \). Thus both \( x, \sigma_{n(i+2)}(\omega) \) are in \( B(x_n, 2^{-(i+2)}) \) and so \( |\sigma_{n(i+1)}(\omega) - x| < 2^{-i} \). It follows that condition 4.) holds. Note that this had nothing to do with the measure.

Now we verify that 4.) \( \Rightarrow \) 3.). Suppose there exist measurable selections \( \sigma_n(\omega) \in \Gamma(\omega) \) satisfying condition 4.). Let \( U \) be open. Then

\[
\{ \omega : \Gamma(\omega) \cap U \neq \emptyset \} = \cup_{n=1}^{\infty} \sigma_n^{-1}(U) \in C.
\]

Now we verify that 4.) \( \Rightarrow \) 5.). Let \( F(\omega) \equiv \text{dist}(x, \Gamma(\omega)) \). Then letting \( U \) be an open set in \([0, \infty), F(\omega) \in U \) if and only if \( d(x, \sigma_n(\omega)) \in U \) for some \( \sigma_n(\omega) \). Let \( h_n(\omega) \equiv d(x, \sigma_n(\omega)) \). Then \( h_n \) is measurable and \( F^{-1}(U) = \cup_{n=1}^{\infty} h_n^{-1}(U) \in C \). This shows that for all \( x \in X, \omega \rightarrow \text{dist}(x, \Gamma(\omega)) \) is measurable and this proves 5.).

Now we verify that 5.) \( \Rightarrow \) 4.). We know \( \text{dist}(x, \Gamma(\cdot)) \) is measurable and we show \( \{ \omega : \Gamma(\omega) \cap U \neq \emptyset \} \subset C \) whenever \( U \) is open and then use 3.) \( \Rightarrow \) 4.). Since \( X \) is separable, there exists \( B(x_i, r_i) \) such that \( U = \cup_{i=1}^{\infty} B(x_i, r_i) \). Then

\[
\{ \omega : \Gamma(\omega) \cap U \neq \emptyset \} = \cup_{i=1}^{\infty} \{ \omega : \Gamma(\omega) \cap B(x_i, r_i) \neq \emptyset \}
\]

\[
= \cup_{i=1}^{\infty} \{ \omega : \text{dist}(x_i, \Gamma(\omega)) < r_i \} \in C.
\]

Therefore, 5.) \( \Rightarrow \) 4.) as claimed.

Now we must prove 5.) \( \Rightarrow \) 6.). We note that \( \omega \rightarrow \text{dist}(x, \Gamma(\omega)) \) is measurable and \( x \rightarrow \text{dist}(x, \Gamma(\omega)) \) is continuous. Also, the graph of \( \Gamma, \mathcal{G}(\Gamma) \) is given by

\[
\mathcal{G}(\Gamma) = \{(\omega, x) : \text{dist}(x, \Gamma(\omega)) = 0 \}.
\]

We wish to show that \( (\omega, x) \rightarrow \text{dist}(x, \Gamma(\omega)) \) is product measurable because then \( \mathcal{G}(\Gamma) \), being the inverse image of \( \{0\} \) will be product measurable. Let \( \{x_k\} \) be a countable dense set in \( X \) and let

\[
\phi_k(\omega, x) \equiv \text{dist}(x_n, \Gamma(\omega))
\]

where \( n \) is the first index such that \( x \in B(x_n, 2^{-k}) \). Then \( \phi_k(\omega, x) \rightarrow \text{dist}(x, \Gamma(\omega)) \) due to the continuity of \( x \rightarrow \text{dist}(x, \Gamma(\omega)) \) and so we must argue that \( \phi_k \) is product measurable. On

\[
E_n \equiv \Omega \times (B(x_n, 2^{-k}) \setminus \cup_{m<n} B(x_m, 2^{-k}))
\]

\[
\phi_k(\omega, x) = \text{dist}(x_n, \Gamma(\omega)) \). Thus, on this set, \( \phi_k \) equals a measurable function of \( \omega \) and does not depend on \( x \) on \( E_n \). It follows that there are measurable simple \( \mathcal{C} \) measurable functions, \( s_n(\omega) \) which increase pointwise to \( \text{dist}(x_n, \Gamma(\omega)) \) on \( E_n \). Thus \( s_n(\omega)X_{E_n}(x) \) increases to \( \phi_k(\omega, x) \) on \( E_n \) showing that \( \phi_k X_{E_n} \) is product measurable with respect to \( \mathcal{C} \times \sigma(\tau) \) since \( E_n \) is a measurable rectangle with respect to \( \mathcal{C} \) and \( \sigma(\tau) \). Therefore, \( \phi_k \) is product measurable and so \( (\omega, x) \rightarrow \text{dist}(x, \Gamma(\omega)) \) is also product measurable.

It remains to prove 6.) \( \Rightarrow \) 1.). This follows from Theorem 4.).
\[ = \text{proj}_\Omega (G(\Gamma) \cap (\Omega \cap B)) . \]

But from Theorem 11.0.27, \( \text{proj}_\Omega (G(\Gamma) \cap (\Omega \cap B)) \in \mathcal{C} \subset \mathcal{C}_\mu = \mathcal{C} . \) The last part results from \((\Omega, \mathcal{C}, \mu)\) being a complete measure space. Note that without this assumption we could not draw the conclusion desired. This required consideration of the measure.

For much more on multifunctions, you should see the book by Hu and Papageorgiou. The above proof follows the presentation in this book.

### 11.1.2 A Special Case Which Is Easier

The above is a pretty long and difficult argument to show that \( \Gamma^{-1}(U) \in \mathcal{C} \) for all \( U \) open is equivalent to \( \Gamma^{-1}(F) \) for all \( F \) closed. However, there is a special case for which this is much easier to show. Suppose \( \Gamma(\omega) \) is not just closed but also compact. Then as above, if \( \Gamma^{-1}(F) \in \mathcal{C} \) for all \( F \) closed, then \( \Gamma^{-1}(U) = \cup_n \Gamma^{-1}(F_n) \) where \( F_n \) is an increasing sequence of closed sets whose union is \( U \). This follows from the observation that

\[ \Gamma(\omega) \cap U = \cup_n \Gamma(\omega) \cap F_n \]

and so to say the set on the left is nonempty is to say that at least one of the sets on the right is nonempty. Thus if \( \Gamma^{-1}(F) \in \mathcal{C} \) for all \( F \) closed, then \( \Gamma^{-1}(U) \in \mathcal{C} \) for all \( U \) open. This requires no special considerations.

Now suppose \( \Gamma(\omega) \) is compact for every \( \omega \) and that \( \Gamma^{-1}(U) \in \mathcal{C} \) for every \( U \) open. Then let \( F \) be a closed set and let \( \{U_n\} \) be a decreasing sequence of open sets whose intersection equals \( F \) such that also, for all \( n \), \( U_n \supseteq U_{n+1} \). Then

\[ \Gamma(\omega) \cap F = \cap_n \Gamma(\omega) \cap U_n = \cap_n \Gamma(\omega) \cap U_n \]

Now because of compactness, the set on the left is nonempty if and only if each set on the right is also nonempty. Thus \( \Gamma^{-1}(F) = \cap_n \Gamma^{-1}(U_n) \in \mathcal{C} \). Thus in this special case, it is much easier to see that these two conditions for measurability are equivalent.
Appendix A

Covering Theorems

The Vitali covering theorem is a profound result about coverings of a set in \( \mathbb{R}^p \) with open balls. The balls can be defined in terms of any norm for \( \mathbb{R}^p \). For example, the norm could be

\[ ||x|| \equiv \max \{ |x_k| : k = 1, \ldots, p \} \]

or the usual norm

\[ |x| = \sqrt{\sum_{k=1}^{p} |x_k|^2} \]

or any other. The proof given here is from Basic Analysis [38]. Before beginning the proof, here is a useful lemma.

**Lemma A.0.3** In a normed linear space,

\[ B(x, r) = \{ y : ||y - x|| \leq r \} \]

**Proof:** It is clear that \( B(x, r) \subseteq \{ y : ||y - x|| \leq r \} \) because if \( y \in B(x, r) \), then there exists a sequence of points of \( B(x, r) \), \( \{ x_n \} \) such that \( ||x_n - y|| \rightarrow 0 \), \( ||x_n|| < r \). However, this requires that \( ||x_n|| \rightarrow ||y|| \) and so \( ||y|| \leq r \). Now let \( y \) be in the right side. It suffices to consider \( ||y - x|| = 1 \). Then you could consider for \( t \in (0, 1) \),

\[ z(t) = x + t(y - x) \]

Then

\[ ||z(t) - x|| = t||y - x|| = tr < r \]

and so \( z(t) \in B(x, r) \). But also,

\[ ||z(t) - y|| = (1 - t)||y - x|| = (1 - t)r \]

so \( \lim_{t \rightarrow 0} ||z(t) - y|| = 0 \) showing that \( y \in B(x, r) \).

Thus the usual way we think about the closure of a ball is completely correct in a normed linear space. Recall that this lemma is not always true in the context of a metric space. Recall the discrete metric for example in which the distance between different points is 1 and distance between a point and itself is 0. In what follows we will use the result of this lemma without comment.
Lemma A.0.4 Let $\mathcal{F}$ be a countable collection of open balls satisfying
\[ \infty > M \equiv \sup \{ r : B(p, r) \in \mathcal{F} \} > 0 \]
and let $k \in (0, \infty)$. Then there exists $\mathcal{G} \subseteq \mathcal{F}$ such that
\begin{align*}
\text{If } B(p, r) \in \mathcal{G} \text{ then } r &> k, \quad (1.0.1) \\
\text{If } B_1, B_2 \in \mathcal{G} \text{ then } \overline{B_1} \cap \overline{B_2} &\neq \emptyset, \quad (1.0.2) \\
\mathcal{G} \text{ is maximal with respect to } &1.0.1 \text{ and } 1.0.2. \quad (1.0.3)
\end{align*}
By this is meant that if $\mathcal{H}$ is a collection of balls satisfying $1.0.1$ and $1.0.2$, then $\mathcal{H}$ cannot properly contain $\mathcal{G}$.

Proof: If no ball of $\mathcal{F}$ has radius larger than $k$, let $\mathcal{G} = \emptyset$. Assume therefore, that some balls have radius larger than $k$. Let $\mathcal{F} \equiv \{ B_i \}_{i=1}^\infty$. Now let $B_{n_1}$ be the first ball in the list which has radius greater than $k$. If every ball having radius larger than $k$ has closure which intersects $\overline{B_{n_1}}$, then stop. The maximal set is $\{B_{n_1}\}$. Otherwise, let $B_{n_2}$ be the next ball having radius larger than $k$ for which $\overline{B_{n_2}} \cap \overline{B_{n_1}} = \emptyset$. Continue this way obtaining $\{B_{n_i}\}_{i=1}^\infty$, a finite or infinite sequence of balls having radius larger than $k$ whose closures are disjoint. Then let $\mathcal{G} \equiv \{ B_{n_i} \}$. To see $\mathcal{G}$ is maximal with respect to $1.0.1$ and $1.0.2$, suppose $B \in \mathcal{F}$, $B$ has radius larger than $k$, and $\mathcal{G} \cup \{B\}$ satisfies $1.0.1$ and $1.0.2$. Then at some point in the process, $B$ would have been chosen because it would be the ball of radius larger than $k$ which has the smallest index at some point in the construction. Therefore, $B \in \mathcal{G}$ and this shows $\mathcal{G}$ is maximal with respect to $1.0.1$ and $1.0.2.$ \hfill \blacksquare

Proposition A.0.5 Let $\mathcal{F}$ be a collection of open balls, and let
\[ A \equiv \cup \{ B : B \in \mathcal{F} \}. \]
Suppose
\[ \infty > M \equiv \sup \{ r : B(p, r) \in \mathcal{F} \} > 0. \]
Then there exists $\mathcal{G} \subseteq \mathcal{F}$ such that $\mathcal{G}$ consists of balls whose closures are disjoint and
\[ A \subseteq \cup \{ \overline{B} : B \in \mathcal{G} \} \]
where for $B = B(x, r)$ a ball, $\overline{B}$ denotes the open ball $B(x, 5r)$.

Proof: First of all, it follows from Problem 28 on Page 185 that there is a countable subset of $\mathcal{F}$ which also covers $A$. Thus it can be assumed that $\mathcal{F}$ is countable.

By Lemma A.0.1 there exists $\mathcal{G}_1 \subseteq \mathcal{F}$ which satisfies $1.0.1$, $1.0.3$, and $1.0.2$ with $k = \frac{2M}{3}$.

Suppose $\mathcal{G}_1, \ldots, \mathcal{G}_{m-1}$ have been chosen for $m \geq 2$. Let $\overline{\mathcal{G}_i}$ denote the collection of closures of the balls of $\mathcal{G}_i$. Then let $\mathcal{F}_m$ be those balls of $\mathcal{F}$, such that if $B$ is one of these balls, $\overline{B}$ has empty intersection with every closed ball of $\overline{\mathcal{G}_i}$ for each
Then using Lemma A.0.4, let $G_m$ be a maximal collection of balls from $F_m$ with the property that each ball has radius larger than $(\frac{2}{3})^m M$ and their closures are disjoint. Let $G \equiv \cup_{k=1}^{\infty} G_k$. Thus the closures of balls in $G$ are disjoint. Let $x \in B(p, r) \in F \setminus G$. Choose $m$ such that 

$$\left(\frac{2}{3}\right)^m M < r \leq \left(\frac{2}{3}\right)^{m-1} M$$

Then $B(p, r)$ must have nonempty intersection with the closure of some ball from $G_1 \cup \cdots \cup G_m$ because if it didn’t, then $G_m$ would fail to be maximal. Denote by $B(p_0, r_0)$ a ball in $G_1 \cup \cdots \cup G_m$ whose closure has nonempty intersection with $B(p, r)$. Thus

$$r_0, r > \left(\frac{2}{3}\right)^m M.$$ 

Consider the picture, in which $w \in B(p_0, r_0) \cap B(p, r)$.

Then for $x \in B(p, r)$, 

$$\|x - p_0\| \leq \|x - p\| + \|p - w\| + \|w - p_0\| \leq r_0$$

$$\leq r + r_0 \leq 2 \left(\frac{2}{3}\right)^{m-1} M + r_0 \leq 2 \left(\frac{3}{2}\right)^m M + r_0 \leq 4r_0$$

Thus $B(p, r)$ is contained in $B(p_0, 4r_0)$. It follows that the closures of the balls of $G$ are disjoint and the set $\{\hat{B} : B \in G\}$ covers $A$. ■

Here is the concept of a Vitali covering.

**Definition A.0.6** Let $S$ be a set and let $C$ be a covering of $S$ meaning that every point of $S$ is contained in a set of $C$. This covering is said to be a Vitali covering if for each $\varepsilon > 0$ and $x \in S$, there exists a set $B \in C$ containing $x$, the diameter of $B$ is less than $\varepsilon$, and there exists an upper bound to the set of diameters of sets of $C$.

Recall the outer measure determined by a measure. When the measure is $m_p$ that is $p$ dimensional Lebesgue measure, denote this outer measure as $\hat{m}_p$. Thus

$$\hat{m}_p(S) \equiv \inf \{m_p(F) : F \supseteq S, F \text{ measurable}\}$$

Recall that $\hat{m}_p = m_p$ on the $\sigma$ algebra of Lebesgue measurable sets. Also, if $E$ is any set, there exists $F \supseteq E$ such that $F$ is measurable and $m_p(F) = \hat{m}_p(E)$. Then the following is also called the Vitali covering theorem.
Theorem A.0.7 Let \( E \subseteq \mathbb{R}^p \) be a bounded set and let \( \mathcal{F} \) be a collection of open balls, of bounded radii such that \( \mathcal{F} \) covers \( E \) in the sense of Vitali. Then there exists a countable collection of balls from \( \mathcal{F} \) whose closures are disjoint, denoted by \( \{B_j\}_{j=1}^\infty \), such that \( \tilde{m}_p(E \setminus \cup_{j=1}^\infty \overline{B_j}) = \tilde{m}_p(E \setminus \cup_{j=1}^\infty B_j) = 0 \).

**Proof:** From the definition of Lebesgue measure,

\[
m_p(B(x, \alpha r)) = m_p(B(0, \alpha r)) = \alpha^p m_p(B(0, r)) = \alpha^p m_p(B(x, r)),
\]

This is especially clear if the norm is \( \|\cdot\|_{\infty} \) because in this case, the balls are just \( p \)-dimensional cubes centered at \( x \). It is also true for any other norm, which will be made clear from the change of variables formula. Let \( S(x, r) \equiv \{y : |y - x| = r\} \).

Then for each \( \varepsilon < r \),

\[
m_p(S(x, r)) \subseteq m_p(B(x, r + \varepsilon)) - m_p(B(x, r - \varepsilon)) = m_p(B(0, r + \varepsilon)) - m_p(B(0, r - \varepsilon)) = \left( \left( \frac{r + \varepsilon}{r} \right)^p - \left( \frac{r - \varepsilon}{r} \right)^p \right) (m_p(B(0, r))).
\]

Hence \( m_p(S(x, r)) = 0 \).

If \( m_p(E) = 0 \), there is nothing to prove, so assume the outer measure of this set is positive. Let \( F \supseteq E \) such that \( F \) is measurable and \( m_p(F) = \tilde{m}_p(E) \). By outer regularity of Lebesgue measure, there exists \( U \), an open set which satisfies

\[
m_p(F) > (1 - 10^{-p})m_p(U), \ U \supseteq F.
\]

Each point of \( F \) is contained in balls of \( \mathcal{F} \) of arbitrarily small radii and so there exists a covering of \( F \) with balls of \( \mathcal{F} \) whose closures are contained in \( U \). Therefore, by Proposition A.0.5, there exist balls, \( \{B_i\}_{i=1}^\infty \subseteq \mathcal{F} \) such that their closures are disjoint and

\[
F \subseteq \cup_{j=1}^\infty \overline{B_j}, \ \overline{B_j} \subseteq U.
\]

Therefore,

\[
m_p(F \setminus \cup_{j=1}^\infty \overline{B_j}) \leq m_p(U) - m_p(\cup_{j=1}^\infty \overline{B_j}) < (1 - 10^{-p})^{-1} m_p(F) - \sum_{j=1}^\infty m_p(\overline{B_j}) = (1 - 10^{-p})^{-1} m_p(F) - 5^{-p} \sum_{j=1}^\infty m_p(\overline{B_j}) \leq (1 - 10^{-p})^{-1} m_p(F) - 5^{-p} m_p(F) = m_p(F) \theta_p
\]
where
\[ \theta_p \equiv (1 - 10^{-p})^{-1} - 5^{-p} = \frac{5^p - (1 - 10^{-p})}{5^p (1 - 10^{-p})} = \frac{50^p - 10^p + 1}{50^p - 5^p} < 1. \]

Thus, there exists \( m_1 \) large enough that
\[ m_p (F \setminus \bigcup_{j=1}^{m_1} \overline{B_j}) < \theta_p m_p (F) \]

Now consider \( F \setminus \bigcup_{j=1}^{m_1} \overline{B_j} \) and apply the same reasoning to it that was done to \( F \).
Thus there exists \( m_2 > m_1 \) such that
\[ m_p (F \setminus \bigcup_{j=1}^{m_2} \overline{B_j}) < \theta_p m_p (F \setminus \bigcup_{j=1}^{m_1} \overline{B_j}) < \theta_p^2 m_p (F) \]

Continuing this way, there exists an increasing subsequence \( m_k \) such that
\[ m_p (F \setminus \bigcup_{j=1}^{m_k} \overline{B_j}) < \theta_p m_p (F \setminus \bigcup_{j=1}^{m_1} \overline{B_j}) \]

and since \( \theta_p < 1 \), and \( m_p (F) < \infty \), this implies \( m_p (F \setminus \bigcup_{j=1}^{\infty} \overline{B_j}) = 0 \). Now
\[
0 \leq \bar{m}_p (E \setminus \bigcup_{j=1}^{\infty} \overline{B_j}) \leq \bar{m}_p (E \setminus \bigcup_{j=1}^{\infty} B_j) \leq m_p (F \setminus \bigcup_{j=1}^{\infty} \overline{B_j}) = 0.
\]

You don’t need to assume that \( E \) is bounded in order to draw the above conclusion.

**Corollary A.0.8** Let \( E \subseteq \mathbb{R}^p \) be a set and let \( \mathcal{F} \) be a collection of balls which come from some norm, open or not, but having bounded radii such that \( \mathcal{F} \) covers \( E \) in the sense of Vitali. Then there exists a countable collection of balls from \( \mathcal{F} \) having disjoint closures, denoted by \( \{B_j\}_{j=1}^{\infty} \), such that \( \bar{m}_p (E \setminus \bigcup_{j=1}^{\infty} B_j) = 0 \).

**Proof:** Consider \( A_n = B(0, n) \setminus B(0, n-1), n = 1, 2, \cdots \). Let \( E_n = E \cap A_n \). Then \( \bigcup_{n=1}^{\infty} E_n \cup N = E \) where \( N \) is a set of measure zero. From Theorem A.1.7, there exist balls of \( \mathcal{F} \), having disjoint closures denoted by \( \{B^n_i\}_{i=1}^{\infty} \), such that \( \bar{m}_p (E_n \setminus \bigcup_{i=1}^{\infty} B^n_i) = 0 \) and each \( B^n_i \subseteq A_n \). Then \( \{B^n_i, (i, n) \in \mathbb{N} \times \mathbb{N}\} \) is a suitable collection of balls having disjoint closures. In fact,
\[
\bar{m}_p (E \setminus \bigcup_{n=1}^{\infty} B^n_i) = \bar{m}_p (\bigcup_n (E_n \setminus \bigcup_{i=1}^{\infty} B^n_i)) \leq \sum_n \bar{m}_p (E_n \setminus \bigcup_{i=1}^{\infty} B^n_i) = 0.
\]
Appendix B

Banach Spaces

B.1 Introduction

In this part of the book are the basic theorems from linear functional analysis. This is a very big topic. See Dunford and Schwartz for example. It goes on for three volumes. That which is presented here is what I personally have found most useful. It should give the necessary background for the nonlinear analysis presented later. Nonlinear analysis really does have a different flavor than standard linear functional analysis, but some background is needed, mainly things like the definition of a Banach space.

B.2 Theorems Based On Baire Category

B.2.1 Baire Category Theorem

Some examples of Banach spaces that have been discussed up to now are \( \mathbb{R}^n \), \( C^n \), and \( L^p (\Omega) \). Theorems about general Banach spaces are proved in this chapter. The main theorems to be presented here are the uniform boundedness theorem, the open mapping theorem, the closed graph theorem, and the Hahn Banach Theorem. The first three of these theorems come from the Baire category theorem which is about to be presented. They are topological in nature. The Hahn Banach theorem has nothing to do with topology. Banach spaces are all normed linear spaces and as such, they are all metric spaces because a normed linear space may be considered as a metric space with \( d(x, y) \equiv \|x - y\| \). You can check that this satisfies all the axioms of a metric. As usual, if every Cauchy sequence converges, the metric space is called complete.

Definition B.2.1 A complete normed linear space is called a Banach space.

The following remarkable result is called the Baire category theorem. To get an idea of its meaning, imagine you draw a line in the plane. The complement of this
line is an open set and is dense because every point, even those on the line, are limit points of this open set. Now draw another line. The complement of the two lines is still open and dense. Keep drawing lines and looking at the complements of the union of these lines. You always have an open set which is dense. Now what if there were countably many lines? The Baire category theorem implies the complement of the union of these lines is dense. In particular it is nonempty. Thus you cannot write the plane as a countable union of lines. This is a rather rough description of this very important theorem. The precise statement and proof follow.

**Theorem B.2.2** Let \((X, d)\) be a complete metric space and let \(\{U_n\}_{n=1}^{\infty}\) be a sequence of open subsets of \(X\) satisfying \(\bigcup_{n=1}^{\infty} U_n = X\) (\(U_n\) is dense). Then \(D \equiv \bigcap_{n=1}^{\infty} U_n\) is a dense subset of \(X\).

**Proof:** Let \(p \in X\) and let \(r_0 > 0\). I need to show \(D \cap B(p, r_0) \neq \emptyset\). Since \(U_1\) is dense, there exists \(p_1 \in U_1 \cap B(p, r_0)\), an open set. Let \(p_1 \in B(p_1, r_1) \subseteq B(p_1, r_1) \subseteq U_1 \cap B(p, r_0)\) and \(r_1 < 2^{-1}\). This is possible because \(U_1 \cap B(p, r_0)\) is an open set and so there exists \(r_1\) such that \(B(p_1, 2r_1) \subseteq U_1 \cap B(p, r_0)\). But

\[
B(p_1, r_1) \subseteq B(p_1, r_1) \subseteq B(p_1, 2r_1)
\]

because \(B(p_1, r_1) = \{x \in X : d(x, p) \leq r_1\}\). (Why?)

![Diagram](r_0 \quad p \quad p_1)

There exists \(p_2 \in U_2 \cap B(p_1, r_1)\) because \(U_2\) is dense. Let

\[
p_2 \in B(p_2, r_2) \subseteq B(p_2, r_2) \subseteq U_2 \cap B(p_1, r_1) \subseteq U_1 \cap U_2 \cap B(p, r_0).
\]

and let \(r_2 < 2^{-2}\). Continue in this way. Thus

\[
r_n < 2^{-n},
\]

\[
B(p_n, r_n) \subseteq U_1 \cap U_2 \cap ... \cap U_n \cap B(p, r_0),
\]

\[
B(p_n, r_n) \subseteq B(p_{n-1}, r_{n-1}).
\]

The sequence, \(\{p_n\}\) is a Cauchy sequence because all terms of \(\{p_k\}\) for \(k \geq n\) are contained in \(B(p_n, r_n)\), a set whose diameter is no larger than \(2^{-n}\). Since \(X\) is complete, there exists \(p_\infty\) such that

\[
\lim_{n \to \infty} p_n = p_\infty.
\]

Since all but finitely many terms of \(\{p_n\}\) are in \(\overline{B(p_m, r_m)}\), it follows that \(p_\infty \in \overline{B(p_m, r_m)}\) for each \(m\). Therefore, \(p_\infty \in \bigcap_{m=1}^{\infty} \overline{B(p_m, r_m)} \subseteq \bigcap_{i=1}^{\infty} U_i \cap B(p, r_0)\).
The following corollary is also called the Baire category theorem.

**Corollary B.2.3** Let $X$ be a complete metric space and suppose $X = \bigcup_{i=1}^{\infty} F_i$, where each $F_i$ is a closed set. Then for some $i$, interior $F_i \neq \emptyset$.

**Proof:** If all $F_i$ has empty interior, then $F_i^C$ would be a dense open set. Therefore, from Theorem B.2.2, it would follow that $\emptyset = (\bigcup_{i=1}^{\infty} F_i)^C = \bigcap_{i=1}^{\infty} F_i^C \neq \emptyset$.

The set $D$ of Theorem B.2.2 is called a $G_\delta$ set because it is the countable intersection of open sets. Thus $D$ is a dense $G_\delta$ set.

Recall that a norm satisfies:

a.) $||x|| \geq 0$, $||x|| = 0$ if and only if $x = 0$.

b.) $||x + y|| \leq ||x|| + ||y||$.

c.) $||cx|| = |c||x||$ if $c$ is a scalar and $x \in X$.

From the definition of continuity, it follows easily that a function is continuous if

$$\lim_{n \to \infty} x_n = x$$

implies

$$\lim_{n \to \infty} f(x_n) = f(x).$$

**Theorem B.2.4** Let $X$ and $Y$ be two normed linear spaces and let $L : X \to Y$ be linear ($L(ax + by) = aL(x) + bL(y)$ for $a, b$ scalars and $x, y \in X$). The following are equivalent

a.) $L$ is continuous at 0

b.) $L$ is continuous

c.) There exists $K > 0$ such that $||Lx||_Y \leq K ||x||_X$ for all $x \in X$ ($L$ is bounded).

**Proof:**

a.)$\Rightarrow$b.) Let $x_n \to x$. It is necessary to show that $Lx_n \to Lx$. But $(x_n - x) \to 0$ and so from continuity at 0, it follows

$$L(x_n - x) = Lx_n - Lx \to 0$$

so $Lx_n \to Lx$. This shows a.) implies b.).

b.)$\Rightarrow$c.) Since $L$ is continuous, $L$ is continuous at 0. Hence $||Lx||_Y < 1$ whenever $||x||_X \leq \delta$ for some $\delta$. Therefore, suppressing the subscript on the $|| ||$, we have

$$||L \left( \frac{\delta x}{||x||} \right) || \leq 1.$$

Hence

$$||Lx|| \leq \frac{1}{\delta} ||x||.$$

c.)$\Rightarrow$a.) follows from the inequality given in c.).
**Definition B.2.5** Let $L : X \rightarrow Y$ be linear and continuous where $X$ and $Y$ are normed linear spaces. Denote the set of all such continuous linear maps by $\mathcal{L}(X,Y)$ and define

$$||L|| = \sup\{||Lx|| : ||x|| \leq 1\}.$$  \hspace{1cm} (2.2.1)

This is called the operator norm.

Note that from Theorem B.2.4 $||L||$ is well defined because of part c.) of that Theorem.

The next lemma follows immediately from the definition of the norm and the assumption that $L$ is linear.

**Lemma B.2.6** With $||L||$ defined in 2.2.1, $\mathcal{L}(X,Y)$ is a normed linear space. Also $||Lx|| \leq ||L|| ||x||$.

**Proof:** Let $x \neq 0$ then $x/||x||$ has norm equal to 1 and so

$$\left|\left| L \left( \frac{x}{||x||} \right) \right|\right| \leq ||L||.$$  \hspace{1cm}

Therefore, multiplying both sides by $||x||$, $||Lx|| \leq ||L|| ||x||$. This is obviously a linear space. It remains to verify the operator norm really is a norm. First of all, if $||L|| = 0$, then $Lx = 0$ for all $||x|| \leq 1$. It follows that for any $x \neq 0$, $0 = L \left( \frac{x}{||x||} \right)$ and so $Lx = 0$. Therefore, $L = 0$. Also, if $c$ is a scalar,

$$||cL|| = \sup_{||x|| \leq 1} ||cL(x)|| = |c| \sup_{||x|| \leq 1} ||Lx|| = |c||L||.$$  \hspace{1cm}

It remains to verify the triangle inequality. Let $L, M \in \mathcal{L}(X,Y)$.

$$||L + M|| = \sup_{||x|| \leq 1} ||(L + M)(x)|| \leq \sup_{||x|| \leq 1} (||Lx|| + ||Mx||)$$  \hspace{1cm}

$$\leq \sup_{||x|| \leq 1} ||Lx|| + \sup_{||x|| \leq 1} ||Mx|| = ||L|| + ||M||.$$  \hspace{1cm}

This shows the operator norm is really a norm as hoped. \hspace{1cm}

For example, consider the space of linear transformations defined on $\mathbb{R}^n$ having values in $\mathbb{R}^m$. The fact the transformation is linear automatically imparts continuity to it. You should give a proof of this fact. Recall that every such linear transformation can be realized in terms of matrix multiplication.

Thus, in finite dimensions the algebraic condition that an operator is linear is sufficient to imply the topological condition that the operator is continuous. The situation is not so simple in infinite dimensional spaces such as $C(X;\mathbb{R}^n)$. This explains the imposition of the topological condition of continuity as a criterion for membership in $\mathcal{L}(X,Y)$ in addition to the algebraic condition of linearity.

**Theorem B.2.7** If $Y$ is a Banach space, then $\mathcal{L}(X,Y)$ is also a Banach space.
B.2. THEOREMS BASED ON BAIRE CATEGORY

Proof: Let \( \{L_n\} \) be a Cauchy sequence in \( \mathcal{L}(X,Y) \) and let \( x \in X \).

\[ ||L_n x - L_m x|| \leq ||x|| \cdot ||L_n - L_m||. \]

Thus \( \{L_n x\} \) is a Cauchy sequence. Let

\[ Lx = \lim_{n \to \infty} L_n x. \]

Then, clearly, \( L \) is linear because if \( x_1, x_2 \) are in \( X \), and \( a, b \) are scalars, then

\[ L(ax_1 + bx_2) = \lim_{n \to \infty} L_n(ax_1 + bx_2) = \lim_{n \to \infty} (aL_n x_1 + bL_n x_2) = aLx_1 + bLx_2. \]

Also \( L \) is continuous. To see this, note that \( \{||L_n||\} \) is a Cauchy sequence of real numbers because

\[ |||L_n|| - ||L_m|||| \leq ||L_n - L_m||. \]

Hence there exists \( K > \sup\{||L_n|| : n \in \mathbb{N}\} \). Thus, if \( x \in X \),

\[ ||Lx|| = \lim_{n \to \infty} ||L_n x|| \leq K ||x||. \]

\[ \Box \]

B.2.2 Uniform Boundedness Theorem

The next big result is sometimes called the Uniform Boundedness theorem, or the Banach-Steinhaus theorem. This is a very surprising theorem which implies that for a collection of bounded linear operators, if they are bounded pointwise, then they are also bounded uniformly. As an example of a situation in which pointwise bounded does not imply uniformly bounded, consider the functions \( f_\alpha(x) \equiv \chi_{(0,1)}(x)x^{-1} \) for \( \alpha \in (0,1) \). Clearly each function is bounded and the collection of functions is bounded at each point of \((0,1)\), but there is no bound for all these functions taken together. One problem is that \((0,1)\) is not a Banach space. Therefore, the functions cannot be linear.

Theorem B.2.8 Let \( X \) be a Banach space and let \( Y \) be a normed linear space. Let \( \{L_\alpha\}_{\alpha \in \Lambda} \) be a collection of elements of \( \mathcal{L}(X,Y) \). Then one of the following happens.

a.) \( \sup\{||L_\alpha|| : \alpha \in \Lambda\} < \infty \)

b.) There exists a dense \( G_\delta \) set, \( D \), such that for all \( x \in D \),

\[ \sup\{||L_\alpha x|| : \alpha \in \Lambda\} = \infty. \]

Proof: For each \( n \in \mathbb{N} \), define

\[ U_n = \{x \in X : \sup\{||L_\alpha x|| : \alpha \in \Lambda\} > n\}. \]

Then \( U_n \) is an open set because if \( x \in U_n \), then there exists \( \alpha \in \Lambda \) such that

\[ ||L_\alpha x|| > n. \]
But then, since \( L_\alpha \) is continuous, this situation persists for all \( y \) sufficiently close to \( x \), say for all \( y \in B(x, \delta) \). Then \( B(x, \delta) \subseteq U_n \) which shows \( U_n \) is open.

Case b.) is obtained from Theorem \[\text{B.2.2}\] if each \( U_n \) is dense.

The other case is that for some \( n, U_n \) is not dense. If this occurs, there exists \( x_0 \) and \( r > 0 \) such that for all \( x \in B(x_0, r) \), \( \|L_\alpha x\| \leq n \) for all \( \alpha \). Now if \( y \in B(0, r), x_0 + y \in B(x_0, r) \). Consequently, for all such \( y \), \( \|L_\alpha(x_0 + y)\| \leq n \). This implies that for all \( \alpha \in \Lambda \) and \( \|y\| < r \),

\[
\|L_\alpha y\| \leq n + \|L_\alpha(x_0)\| \leq 2n.
\]

Therefore, if \( \|y\| \leq 1, \|\frac{r}{2}y\| < r \) and so for all \( \alpha \),

\[
\|L_\alpha \left(\frac{r}{2}y\right)\| \leq 2n.
\]

Now multiplying by \( r/2 \) it follows that whenever \( \|y\| \leq 1, \|L_\alpha (y)\| \leq 4n/r \). Hence case a.) holds.

### B.2.3 Open Mapping Theorem

Another remarkable theorem which depends on the Baire category theorem is the open mapping theorem. Unlike Theorem \[\text{B.2.8}\] it requires both \( X \) and \( Y \) to be Banach spaces.

**Theorem B.2.9** Let \( X \) and \( Y \) be Banach spaces, let \( L \in \mathcal{L}(X,Y) \), and suppose \( L \) is onto. Then \( L \) maps open sets onto open sets.

To aid in the proof, here is a lemma.

**Lemma B.2.10** Let \( a \) and \( b \) be positive constants and suppose

\[
B(0, a) \subseteq \overline{L(B(0, b))}.
\]

Then

\[
\overline{L(B(0, b))} \subseteq L(B(0, 2b)).
\]

**Proof of Lemma B.2.10** Let \( y \in \overline{L(B(0, b))} \). There exists \( x_1 \in B(0, b) \) such that \( \|y - Lx_1\| < \frac{a}{2} \). Now this implies

\[
2y - 2Lx_1 \in B(0, a) \subseteq \overline{L(B(0, b))}.
\]

Thus \( 2y - 2Lx_1 \in \overline{L(B(0, b))} \) just like \( y \) was. Therefore, there exists \( x_2 \in B(0, b) \) such that \( \|2y - 2Lx_1 - Lx_2\| < a/2 \). Hence \( \|4y - 4Lx_1 - 2Lx_2\| < a \), and there exists \( x_3 \in B(0, b) \) such that \( \|4y - 4Lx_1 - 2Lx_2 - Lx_3\| < a/2 \). Continuing in this way, there exist \( x_1, x_2, x_3, x_4, \ldots \) in \( B(0, b) \) such that

\[
\|2^n y - \sum_{i=1}^{n} 2^{n-(i-1)}L(x_i)\| < a
\]
which implies
\[ \| y - \sum_{i=1}^{n} 2^{-(i-1)} L(x_i) \| = \| y - L \left( \sum_{i=1}^{n} 2^{-(i-1)} (x_i) \right) \| < 2^{-n} a \] (2.2.2)

Now consider the partial sums of the series, \( \sum_{i=m}^{n} 2^{-(i-1)} x_i \).
\[ \| \sum_{i=m}^{n} 2^{-(i-1)} x_i \| \leq b \sum_{i=m}^{\infty} 2^{-(i-1)} = b 2^{-m+2}. \]

Therefore, these partial sums form a Cauchy sequence and so since \( X \) is complete, there exists \( x = \sum_{i=1}^{\infty} 2^{-(i-1)} x_i \). Letting \( n \to \infty \) in (2.2.2) yields \( \| y - Lx \| = 0 \). Now
\[ \| x \| = \lim_{n \to \infty} \| \sum_{i=1}^{n} 2^{-(i-1)} x_i \| \]
\[ \leq \lim_{n \to \infty} \sum_{i=1}^{n} 2^{-(i-1)} \| x_i \| < \lim_{n \to \infty} \sum_{i=1}^{n} 2^{-(i-1)} b = 2b. \]

\textbf{Proof of Theorem 2.2.3:} \( Y = \bigcup_{n=1}^{\infty} L(B(0,n)) \). By Corollary 2.2.10, the set, \( \overline{L(B(0,n))} \) has nonempty interior for some \( n_0 \). Thus \( B(y,r) \subseteq \overline{L(B(0,n_0))} \) for some \( y \) and some \( r > 0 \). Since \( L \) is linear \( B(-y,r) \subseteq \overline{L(B(0,n_0))} \) also. Here is why. If \( z \in B(-y,r) \), then \(-z \in B(y,r)\) and so there exists \( x_n \in B(0,n_0) \) such that \( Lx_n \to -z \). Therefore, \( L(-x_n) \to z \) and \(-x_n \in B(0,n_0) \) also. Therefore \( z \in \overline{L(B(0,n_0))} \). Then it follows that
\[ B(0,r) \subseteq B(y,r) + B(-y,r) \]
\[ \equiv \{ y_1 + y_2 : y_1 \in B(y,r) \text{ and } y_2 \in B(-y,r) \} \]
\[ \subseteq \overline{L(B(0,2n_0))} \]

The reason for the last inclusion is that from the above, if \( y_1 \in B(y,r) \) and \( y_2 \in B(-y,r) \), there exists \( x_n, z_n \in B(0,n_0) \) such that
\[ Lx_n \to y_1, \quad Lz_n \to y_2. \]

Therefore,
\[ \| x_n + z_n \| \leq 2n_0 \]
and so \( (y_1 + y_2) \in \overline{L(B(0,2n_0))} \).

By Lemma 2.2.10, \( \overline{L(B(0,2n_0))} \subseteq L(B(0,4n_0)) \) which shows
\[ B(0,r) \subseteq L(B(0,4n_0)). \]

Letting \( a = r(4n_0)^{-1} \), it follows, since \( L \) is linear, that \( B(0,a) \subseteq L(B(0,1)) \). It follows since \( L \) is linear,
\[ L(B(0,r)) \supseteq B(0,ar). \] (2.2.3)
Now let $U$ be open in $X$ and let $x + B(0, r) = B(x, r) \subseteq U$. Using 2.2.3,

$$L(U) \supseteq L(x + B(0, r))$$

$$= Lx + L(B(0, r)) \supseteq Lx + B(0, ar) = B(Lx, ar).$$

Hence

$$Lx \in B(Lx, ar) \subseteq L(U).$$

which shows that every point, $Lx \in LU$, is an interior point of $LU$ and so $LU$ is open.

This theorem is surprising because it implies that if $|\cdot|$ and $||\cdot||$ are two norms with respect to which a vector space $X$ is a Banach space such that $|\cdot| \leq K ||\cdot||$, then there exists a constant $k$, such that $||\cdot|| \leq k |\cdot|$. This can be useful because sometimes it is not clear how to compute $k$ when all that is needed is its existence.

To see the open mapping theorem implies this, consider the identity map $I_x = x$.

Then $I : (X, ||\cdot||) \to (X, |\cdot|)$ is continuous and onto. Hence $I$ is an open map which implies $I^{-1}$ is continuous. Theorem B.2.4 gives the existence of the constant $k$.

### B.2.4 Closed Graph Theorem

**Definition B.2.11** Let $f : D \to E$. The set of all ordered pairs of the form $\{(x, f(x)) : x \in D\}$ is called the graph of $f$.

**Definition B.2.12** If $X$ and $Y$ are normed linear spaces, make $X \times Y$ into a normed linear space by using the norm $|| (x, y) || = \max (||x||, ||y||)$ along with component-wise addition and scalar multiplication. Thus $a(x, y) + b(z, w) \equiv (ax + bz, ay + bw)$.

There are other ways to give a norm for $X \times Y$. For example, you could define $||(x, y)|| = ||x|| + ||y||$

**Lemma B.2.13** The norm defined in Definition B.2.12 on $X \times Y$ along with the definition of addition and scalar multiplication given there make $X \times Y$ into a normed linear space.

**Proof:** The only axiom for a norm which is not obvious is the triangle inequality. Therefore, consider

$$||(x_1, y_1) + (x_2, y_2)|| = ||(x_1 + x_2, y_1 + y_2)||$$

$$= \max (||x_1 + x_2||, ||y_1 + y_2||)$$

$$\leq \max (||x_1|| + ||x_2||, ||y_1|| + ||y_2||)$$

$$\leq \max (||x_1||, ||y_1||) + \max (||x_2||, ||y_2||)$$

$$= ||(x_1, y_1)|| + ||(x_2, y_2)||.$$

It is obvious $X \times Y$ is a vector space from the above definition.
Lemma B.2.14 If $X$ and $Y$ are Banach spaces, then $X \times Y$ with the norm and vector space operations defined in Definition B.2.12 is also a Banach space.

Proof: The only thing left to check is that the space is complete. But this follows from the simple observation that $\{(x_n, y_n)\}$ is a Cauchy sequence in $X \times Y$ if and only if $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in $X$ and $Y$ respectively. Thus if $\{(x_n, y_n)\}$ is a Cauchy sequence in $X \times Y$, it follows there exist $x$ and $y$ such that $x_n \to x$ and $y_n \to y$. But then from the definition of the norm, $(x_n, y_n) \to (x, y)$.

Lemma B.2.15 Every closed subspace of a Banach space is a Banach space.

Proof: If $F \subseteq X$ where $X$ is a Banach space and $\{x_n\}$ is a Cauchy sequence in $F$, then since $X$ is complete, there exists a unique $x \in X$ such that $x_n \to x$. However this means $x \in F = F$ since $F$ is closed.

Definition B.2.16 Let $X$ and $Y$ be Banach spaces and let $D \subseteq X$ be a subspace. A linear map $L : D \to Y$ is said to be closed if its graph is a closed subspace of $X \times Y$. Equivalently, $L$ is closed if $x_n \to x$ and $Lx_n \to y$ implies $x \in D$ and $y = Lx$.

Note the distinction between closed and continuous. If the operator is closed the assertion that $y = Lx$ only follows if it is known that the sequence $\{Lx_n\}$ converges. In the case of a continuous operator, the convergence of $\{Lx_n\}$ follows from the assumption that $x_n \to x$. It is not always the case that a mapping which is closed is necessarily continuous. Consider the function $f (x) = \tan (x)$ if $x$ is not an odd multiple of $\frac{\pi}{2}$ and $f (x) \equiv 0$ at every odd multiple of $\frac{\pi}{2}$. Then the graph is closed and the function is defined on $\mathbb{R}$ but it clearly fails to be continuous. Of course this function is not linear. You could also consider the map,

$$
\frac{d}{dx} : \{y \in C^1 ([0, 1]) : y (0) = 0 \} \equiv D \to C ([0, 1])
$$

where the norm is the uniform norm on $C ([0, 1])$, $\|y\|_\infty$. If $y \in D$, then

$$
y (x) = \int_0^x y' (t) \, dt.
$$

Therefore, if $\frac{dy_n}{dx} \to f \in C ([0, 1])$ and if $y_n \to y$ in $C ([0, 1])$ it follows that

$$
y_n (x) = \int_0^x \frac{dy_n (t)}{dx} \, dt \\
y (x) = \int_0^x f (t) \, dt
$$

and so by the fundamental theorem of calculus $f (x) = y' (x)$ and so the mapping is closed. It is obviously not continuous because it takes $y (x)$ and $y (x) + \frac{1}{n} \sin (nx)$ to two functions which are far from each other even though these two functions are very close in $C ([0, 1])$. Furthermore, it is not defined on the whole space, $C ([0, 1])$.

The next theorem, the closed graph theorem, gives conditions under which closed implies continuous.
**Theorem B.2.17** Let $X$ and $Y$ be Banach spaces and suppose $L : X \to Y$ is closed and linear. Then $L$ is continuous.

**Proof:** Let $G$ be the graph of $L$. $G = \{(x, Lx) : x \in X\}$. By Lemma [5.4.22], it follows that $G$ is a Banach space. Define $P : G \to X$ by $P(x, Lx) = x$. $P$ maps the Banach space $G$ onto the Banach space $X$ and is continuous and linear. By the open mapping theorem, $P$ maps open sets onto open sets. Since $P$ is also one to one, this says that $P^{-1}$ is continuous. Thus $||P^{-1}x|| \leq K||x||$. Hence

$$||Lx|| \leq \max (||x||, ||Lx||) \leq K||x||$$

By Theorem [5.4.24] on Page 172, this shows $L$ is continuous and proves the theorem.

The following corollary is quite useful. It shows how to obtain a new norm on the domain of a closed operator such that the domain with this new norm becomes a Banach space.

**Corollary B.2.18** Let $L : D \subseteq X \to Y$ where $X, Y$ are a Banach spaces, and $L$ is a closed operator. Then define a new norm on $D$ by

$$||x||_D \equiv ||x||_X + ||Lx||_Y.$$  

Then $D$ with this new norm is a Banach space.

**Proof:** If $\{x_n\}$ is a Cauchy sequence in $D$ with this new norm, it follows both $\{x_n\}$ and $\{Lx_n\}$ are Cauchy sequences and therefore, they converge. Since $L$ is closed, $x_n \to x$ and $Lx_n \to Lx$ for some $x \in D$. Thus $||x_n - x||_D \to 0$.

**B.3 Quotient Spaces**

A useful idea is that of a quotient space. It is a way to create another Banach space from a given Banach space and a closed subspace. It generalizes similar concepts which are routine in linear algebra.

**Definition B.3.1** Let $X$ be a Banach space and let $V$ be a closed subspace of $X$. Then $X/V$ denotes the set of equivalence classes determined by the equivalence relation which says $x \sim y$ means $x - y \in V$. An individual equivalence class will be denoted by any of the following symbols. $x + V$, $[x]$, or $[x]_V$. Vector space operations are defined as follows:

$$(x + V) + y + V \equiv x + y + V$$

or in other symbols,

$$[x] + [y] \equiv [x + y]$$

and for $\alpha \in \mathbb{F}$,

$$\alpha [x] \equiv [\alpha x].$$

Also a norm is defined by

$$||[x]|| \equiv \inf \{||x + v|| : v \in V\}.$$
B.3. QUOTIENT SPACES

463

It is left as an exercise to verify the above algebraic operations are well defined. With the above definition, here is the major theorem about quotient spaces.

**Theorem B.3.2** Let $X$ be a Banach space and let $V$ be a closed subspace of $X$. Then with the above definitions of vector space operations, $X/V$ is a Banach space. In the case where $V = \ker (A)$ for $A \in \mathcal{L}(X,Y)$ for $Y$ another Banach space, define $\hat{A} : X/V \to A(X) \subseteq Y$ by $\hat{A}([x]) \equiv Ax$. Then $\hat{A}$ is continuous and $1 - 1$. In fact, $\|\hat{A}\| \leq \|A\|$

**Proof:** First of all, consider the claim that the given norm really is a norm. First note that $\|x + V\| \geq 0$ and $\|x + V\| = 0$ only if $x \in V$ because $V$ is closed. Therefore, $x + V = 0 + V$. Next,

$$\|\alpha [x]\| \equiv \|\alpha x\| \equiv \inf \{\|\alpha x + v\| : v \in V\} \leq \inf \{\|\alpha x + \alpha v\| : v \in V\} = |\alpha| \inf \{\|x + v\| : v \in V\} = |\alpha| \|x\|$$

Consider the triangle inequality,

$$\|x + y\| = \inf \{\|x + y + v\| : v \in V\} \leq \|x + v_1\| + \|y + v_2\|$$

for any choice of $v_1$ and $v_2$. Therefore, taking the infimum of both sides over $v_2$ yields

$$\|x + y\| \leq \|x + v_1\| + \|y\|$$

and then taking the infimum over all $v_1$ yields

$$\|x + y\| \leq \|x\| + \|y\|$$

Next consider the claim that $X/V$ is a Banach space. Letting $\{x_n\}$ be a Cauchy sequence in $X/V$, I will show a subsequence of this converges to a point in $X/V$. This is done by defining a suitable sequence in $X$ and then using completeness of $X$. By choosing a subsequence, it can be assumed that $\|x_n - x_{n+1}\| < 2^{-n}$. Let $z_1 \equiv x_1$. Then choose $v_2 \in V$ such that $\|x + v_2 - z_1\| < 2^{-1}$. Let $z_2 = x_2 + v_2$. Suppose $\{z_1, \ldots, z_n\}$ have been chosen, each having the property that $\|z_k - z_{k+1}\| < 2^{-k}$ then let $v_{n+1}$ be chosen such that $\|x_{n+1} + v_{n+1} - z_n\| < 2^{-n}$ and let $z_{n+1} \equiv x_{n+1} + v_{n+1}$. Thus $\{z_n\}$ is a Cauchy sequence in $X$ and so it converges to $x \in X$. Then

$$\|x - [x_n]\| \leq \|x - (x_{n+1} + v_n)\| = \|x - z_{n+1}\|$$

and so $\lim_{n \to \infty} [x_n] = [x]$.

Next consider the claim about $\hat{A}$. This is well defined and linear because if $[x] = [x_1]$, then $x - x_1 \in \ker (A)$ and so $Ax = Ax_1$. Thus $\hat{A}([x]) = A([x_1])$. It is linear because

$$\hat{A}(\alpha [x] + \beta [y]) = \hat{A}(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha \hat{A}([x]) + \beta \hat{A}([y])$$
Next consider the claim that \( \hat{A} \) is continuous. Letting \( v \in V \),

\[
\| \hat{A}(x) \| = \| Ax \| = \| A(x + v) \| \leq \| A \| \| x + v \|
\]

and so, taking the infimum over all \( v \in V \),

\[
\left\| \hat{A}(x) \right\| \leq \| A \| \| x \|
\]

and this shows \( \| \hat{A} \| \leq \| A \| \). ■

### B.4 Operators With Closed Range

When is \( T(X) \) a closed subset of \( Y \) for \( T \in \mathcal{L}(X,Y) \)? One way this happens is when \( T = I - C \) for \( C \) compact.

**Definition B.4.1** Let \( C \in \mathcal{L}(X,Y) \) where \( X, Y \) are two Banach spaces. Then \( C \) is called a compact operator if \( C \) (bounded set) = (precompact set). This means that every sequence has a subsequence which converges. Equivalently, it means the closure is compact.

**Lemma B.4.2** Suppose \( C \in \mathcal{L}(X,X) \) is compact. Then \( (I - C)(X) \) is closed.

**Proof:** Let \( (I - C)x_n \to y \). Let \( z_n \in \ker (I - C) \) such that

\[
\text{dist} (x_n, \ker (I - C)) \leq \| x_n - z_n \| \\
\leq \left( 1 + \frac{1}{n} \right) \text{dist} (x_n, \ker (I - C))
\]

**Case 1:** \( \| x_n - z_n \| \to \infty \).

In this case, you get \( (I - C)(x_n - z_n) \to y \) and so there is a subsequence such that \( C \left( \frac{x_n - z_n}{\| x_n - z_n \|} \right) \) converges. Also \( \frac{x_n - z_n}{\| x_n - z_n \|} \) converges to the same thing. Let it be called w. Thus

\[
\begin{aligned}
\frac{x_n - z_n}{\| x_n - z_n \|} &\to w, \quad C \left( \frac{x_n - z_n}{\| x_n - z_n \|} \right) \to Cw \\
C \left( \frac{x_n - z_n}{\| x_n - z_n \|} \right) &\to w \text{ so } Cw = w, \ w \in \ker (I - C)
\end{aligned}
\]

\[
\frac{x_n - z_n}{\| x_n - z_n \|} - w = \frac{1}{\| x_n - z_n \|} \left( x_n - z_n - w \| x_n - z_n \| \right)
\]

\[
\geq \frac{1}{\| x_n - z_n \|} \text{dist} (x_n, \ker (I - C))
\]

\[
\geq \left( (1 + \frac{1}{n}) \text{dist} (x_n, \ker (I - C)) \right) \text{dist} (x_n, \ker (I - C))
\]
B.4. OPERATORS WITH CLOSED RANGE

Now passing to a limit,

\[ 0 \geq \lim_{n \to \infty} \frac{1}{1 + 1/n} = 1 \]

so Case 1 cannot occur.

Case 2: A subsequence of \( \|x_n - z_n\| \) is bounded.

Let \( n \) denote the subscript for the subsequence. Then there is a further subsequence still denoted with \( n \) such that \( C(x_n - z_n) \) converges. Then also \( (x_n - z_n) \) converges because \( (I - C)(x_n) = (I - C)(x_n - z_n) \) is given to converge. Let \( (x_n - z_n) \to x \). Then

\[ y = \lim_{n \to \infty} (I - C)x_n = \lim_{n \to \infty} (I - C)(x_n - z_n) = (I - C)x \]

and so \( y \in (I - C)(X) \) showing that \( (I - C)(X) \) is closed. ■

Here is a useful lemma.

Lemma B.4.3 Suppose \( W \) and \( V \) are closed subspaces of a Banach space \( X \) and \( V \not\subset W \) (\( V \) is a proper subset of \( W \)) while \( (\lambda I - L)(W) \subset V, \lambda \neq 0 \). Then there exists \( w \in W \setminus V \) such that \( \|w\| = 1 \) and

\[ \text{dist}(Lw, LV) \geq 1/2 \]

Proof: Let \( w_0 \in W \setminus V \). Then let \( v \in V \) be such that \( \|\lambda w_0 - v\| \leq 2 \text{dist}(\lambda w_0, V) \). Then let

\[ w = \frac{\lambda w_0 - v}{\|\lambda w_0 - v\|} \]

It follows that \( \|w\| = 1 \) and is in \( W \setminus V \). Now let \( x \in V \). Then

\[
Lx - Lw = \lambda (x - w) + \underbrace{(L - \lambda I)(x - w)}_{\in V}
= \lambda x + (L - \lambda I)(x - w) - \lambda w
\]

\[
= \frac{1}{\|\lambda w_0 - v\|} (\lambda x \|\lambda w_0 - v\| + (L - \lambda I)(x - w) \|\lambda w_0 - v\| - \lambda \|\lambda w_0 - v\| w)
\]

\[
= \frac{1}{\|\lambda w_0 - v\|} (\lambda x \|\lambda w_0 - v\| + (L - \lambda I)(x - w) \|\lambda w_0 - v\| - \lambda (\lambda w_0 - v))
\]

\[
= \frac{1}{\|\lambda w_0 - v\|} \left( \lambda x \|\lambda w_0 - v\| + (L - \lambda I)(x - w) \|\lambda w_0 - v\| + \lambda v - \lambda w_0 \right)
\]

Thus

\[
\|Lx - Lw\| \geq \frac{1}{\|\lambda w_0 - v\|} \|\lambda x \|\lambda w_0 - v\| + (L - \lambda I)(x - w) \|\lambda w_0 - v\| + \lambda v - \lambda w_0 \|
\]

\[
\geq \frac{1}{\text{dist}(\lambda w_0, V)} \text{dist}(\lambda w_0, V) = \frac{1}{2} \quad \blacksquare
\]

Here is another fairly elementary lemma a little like the above.
Lemma B.4.4 Let $Y$ be an infinite dimensional Banach space. Then there exists a sequence $\{x_n\}$ in the unit sphere $S$, $\|x_n\| = 1$, such that $\|x_n - x_m\| \geq \frac{1}{2}$ whenever $n \neq m$.

Proof: Pick $x_1 \in S$. Now the span of $x_1$ is not everything and so there exists $u_2 \notin \text{span} \ (x_1)$. Let $u_2$ be a point of span $(x_1)$ such that $\|u_2 - w_2\| \leq 2 \text{dist} \ (u_2, \text{span} \ (x_1))$. Then $x_2 = \frac{u_2 - w_2}{\|u_2 - w_2\|}$. Then

$$\|x_1 - x_2\| = \left\| \frac{\|u_2 - w_2\|}{\|u_2 - w_2\|} x_1 - (u_2 - w_2) \right\| \geq \frac{\text{dist} \ (u_2, \text{span} \ (x_1))}{2 \text{dist} \ (u_2, \text{span} \ (x_1))} = \frac{1}{2}$$

Now repeat the argument with span $(x_1, x_2)$ in place of span $(x_1)$ and continue to get the desired sequence. $\blacksquare$

Lemma B.4.5 Let $L$ be a compact linear map. Then the eigenspace of $L$ is finite dimensional for each eigenvalue $\lambda \neq 0$.

Proof: Consider $(L - \lambda I)^{-1} (0) \cap S$ where $S$ is the unit sphere. The eigenspace is just $(L - \lambda I)^{-1} (0)$. Let $Y$ be this inverse image. If $Y$ is infinite dimensional, then the above Lemma [58] applies. There exists $\{x_n\} \subseteq (L - \lambda I)^{-1} (0) \cap S$ where $\|x_n - x_m\| \geq 1/2$ for all $n \neq m$. Then there is a subsequence, still denoted with subscript $n$ such that $\{Lx_n\}$ is a Cauchy sequence. Thus $Lx_n = \lambda x_n$ and so, since $\lambda \neq 0$, it follows that $\{x_n\}$ is also a Cauchy sequence and converges to some $x$. But this is impossible because of the construction of the $\{x_n\}$ which prevents there being any Cauchy sequence. Thus $Y$ must be finite dimensional. $\blacksquare$

This lemma is useful in proving the following major spectral theorem about the eigenvalues of a compact operator. I found this theorem in Deimling [11].

Theorem B.4.6 Let $L \in \mathcal{L} (X, X)$ with $L$ compact. Let $\Lambda$ be the eigenvalues of $L$. That is $\lambda \in \Lambda$ means there exists $x \neq 0$ such that $Lx = \lambda x$. It is assumed the field of scalars is $\mathbb{R}$ or $\mathbb{C}$. Let $R_\lambda \equiv L - \lambda I$. Then the following hold.

1. If $\mu \in \Lambda$ then $|\mu| \leq \|L\|$, $\Lambda$ is at most countable and has no limit points other than possibly 0.

2. $R_\lambda$ is a homeomorphism onto $X$ whenever $\lambda \notin \Lambda \cup \{0\}$.

3. For all $\lambda \in \Lambda \setminus \{0\}$, there exists a smallest $k = k (\lambda)$,

   (a) $R_\lambda^k X \oplus N (R_\lambda^k) = X$ where $N (R_\lambda^k)$ is the vectors $x$ such that $R_\lambda^k x = 0$. $R_\lambda^k X$ is closed, $\dim \ (N (R_\lambda^k)) < \infty$.

   (b) $R_\lambda^k X$ and $N (R_\lambda^k)$ are invariant under $L$ and $R_\lambda |_{R_\lambda^k X}$ is a homeomorphism onto $R_\lambda^k X$.

   (c) $N (R_\mu^k) \subseteq R_\lambda^k X$ for all $\lambda, \mu \in \Lambda \setminus \{0\}$ where $\lambda \neq \mu$. 

B.4. OPERATORS WITH CLOSED RANGE

On the other hand, \( \{R_k^*X\} \) are decreasing in \( k \). By similar reasoning using Lemma and the observation that \( R_\lambda (R_k^*X) \supseteq R_{k+1}^*X \) (in fact they are equal) it follows that the \( \{R_k^*X\} \) are also eventually constant, say for \( m \geq l \).

Now if you have \( y \in N (R_k^*) \cap R_k^*X \), then \( y = R_k^*w \) and also \( R_k^*y = 0 \). Hence \( R_k^*w = 0 \) and so, \( w \in N (R_{2k}^*) = N (R_k^*) \) which implies \( R_k^*w = 0 \) and so \( y = 0 \). It follows \( N (R_k^*) \cap R_k^*X = \{0\} \).

Now suppose \( l > k \). Then there exists \( y \in R_{l-1}^*X \setminus R_{l-1}^*X \) and so \( R_{l-1}y \in R_{l-1}^*X = R_{l-1}R_{l-1}^*X \). So \( R_{l-1}y = R_{l-1}z \) for some \( z \in R_{l-1}^*X \). Thus \( y - z = 0 \) because \( y \notin R_{l-1}^*X \) but \( z \) is. However, \( R_{l-1} (y - z) = 0 \) and so

\[
(y - z) \in N (R_{l-1}) \cap R_{l-1}^*X \subseteq N (R_k^*) \cap R_k^*X
\]

which cannot happen from the above which showed that \( N (R_k^*) \cap R_k^*X = \{0\} \). Thus \( l \leq k \).

Next suppose \( l < k \). Then you would have \( R_l^*X = R_k^*X \) and \( N (R_k^*) \supseteq N (R_l^*) \). Thus there exists \( y \in N (R_k^*) \) but not in \( N (R_l^*) \). Hence \( R_k^*y = 0 \) but \( R_l^*y \neq 0 \). However, \( R_{l+1}y \) is in \( R_{l+1}^*X \) from the definition of \( l \) and so there is \( u \) such that \( R_{l+1}y = R_{l+1}^*u \). Thus

\[
0 = R_{l+1}^*y = R_{l+1}^{k-l}u = R_{l+1}^{k-l}R_{l+1}^*y = R_{l+1}^{k-l}R_{l+1}^*R_l^*y = R_{l+1}^{k-l}y
\]

Now it follows that \( u \in N (R_{l+1}^{k-l}) = N (R_k^*) \). This is a contradiction because it says that \( R_k^*u = 0 \) but right above the displayed equation, we had \( R_k^*y \in R_k^*X \) and \( R_k^*y \neq 0 \). Thus, with the above paragraph, \( k = l \).

What about the claim that \( R_\lambda \) restricted to \( R_k^*X \) is a homeomorphism? It maps \( R_k^*X \) to \( R_{k+1}^*X \). Also, if \( R_\lambda (y) = 0 \) for \( y \in R_k^*X \), then \( R_k^*y = 0 \) also and so \( y \in R_k^*X \cap N (R_k^*) \). It was shown above that this implies \( y = 0 \). Thus \( R_\lambda \) appears to be one to one. By assumption, it is continuous. Also from Lemma

\[
R_k^*X \text{ is closed.}
\]
This follows from the observation that

\[ R^k_\lambda = (L - \lambda I)^k = \sum_{j=0}^{k} \binom{k}{j} L^j (-\lambda I)^{k-j} = (-\lambda)^k I + \sum_{j=1}^{k} \binom{k}{j} L^j (-\lambda I)^{k-j} \]  

(2.4.4)

which is a multiple of \( I - C \) where \( C \) is a compact map. Then by the open mapping theorem, it follows that \( R_\lambda \) is a homeomorphism onto \( R^{k+1}_\lambda X = R^k_\lambda X \).

What about \( R^k_\lambda X \oplus N (R^k_\lambda) = X \)? It only remains to verify that \( R^k_\lambda X + N (R^k_\lambda) = X \) because the only vector in the intersection was shown to be 0. Thus if you have \( x + y = 0 \) where \( x \) is in one of these and \( y \) in the other, then \( x = -y \) so each is in both and hence both are 0. Pick \( x \in X \). Then \( R^k_\lambda x \in R^k_\lambda (R^k_\lambda X) = R^k_\lambda X \). Therefore, \( R^k_\lambda x = R^k_\lambda (R^k_\lambda y) \) for some \( y \) and so \( R^k_\lambda (x - R^k_\lambda y) = 0 \). Hence

\[ x - R^k_\lambda y \in N (R^k_\lambda) \]

showing that \( x \in R^k_\lambda X + N (R^k_\lambda) \).

It is obvious that \( R^k_\lambda X \) and \( N (R^k_\lambda) \) are invariant under \( L \). If \( \lambda_0 \notin \Lambda \setminus \{0\} \), then \( L - \lambda_0 I \) is one to one and so the compactness of \( L \) and Lemma 4.4.4 implies that \( (L - \lambda_0 I) X \) is closed. Hence the open mapping theorem implies \( L - \lambda_0 I \) is a homeomorphism onto \( (L - \lambda_0 I) X \). Is this last all of \( X \)? There is nothing in the above argument which involved an essential assumption that \( \lambda \in \Lambda \). Hence, repeating this argument, you see that \( (L - \lambda_0 I) X \oplus N (L - \lambda_0 I) = X \), but \( N (L - \lambda_0 I) = 0 \). Hence \( (L - \lambda_0 I) X = X \) and so indeed \( (L - \lambda_0 I) \) is a homeomorphism.

For \( \mu \in \Lambda \), \( Lx = \mu x \) and so \( |\mu| ||x|| \leq ||L|| \|x\| \) so \( |\mu| \leq ||L|| \). Why is \( \Lambda \) at most countable and has only one possible limit point at 0? It was shown that \( R_\lambda \) is a homeomorphism when restricted to \( R^k_\lambda X \). It follows that for \( x \in R^k_\lambda X \), \( ||R_\lambda x|| > \delta \|x\| \) for some \( \delta > 0 \), this for every such \( x \in R^k_\lambda X \). Now consider \( \mu \) close to \( \lambda \) and consider \( R_\mu \). Then for \( x \in R^k_\lambda X \), \( ||R_\mu x|| = ||(R_\lambda + (\lambda - \mu)) x|| \geq \delta \|x\| - |\lambda - \mu| \|x\| > \frac{\delta}{2} \|x\| \) provided \( |\lambda - \mu| < \frac{\delta}{2} \). Thus for \( \mu \) close enough to \( \lambda \), \( R_\mu \) is one to one on \( R^k_\lambda X \). But also \( R_\mu \) is one to one on \( N (R^k_\lambda) \). Let see why this is so. Suppose \( (L - \mu I) x = 0 \) for \( x \in N (R^k_\lambda) \). Then

\[ 0 = (L - \mu I + (\mu - \lambda) I)^k x = (\mu - \lambda)^k x + \sum_{j=1}^{k} \binom{k}{j} (L - \mu I)^j (\mu - \lambda)^{k-j} x \]

and the second term involving the sum yields 0. Since \( R^k_\lambda X \oplus N (R^k_\lambda) = X \), this shows that \( (L - \mu I) \) is one to one for \( \mu \) near \( \lambda \). It follows that for \( \mu \) near \( \lambda \), \( \mu \notin \Lambda \). Thus the only possible limit point is 0. Note that there is no restriction on the size of \( \mu \) for \( (L - \mu I) \) to be one to one on \( N (R^k_\lambda) \).

Why is \( \dim (N (R^k_\lambda)) < \infty \) for each \( \lambda \neq 0 \). This follows from 4.4.4. \( R^k_\lambda \) is a multiple of \( I - C \) for \( C \) a compact operator. Hence this is finite dimensional by Lemma 4.4.4.
What about $N(R^k_\mu) \subseteq R^k_\lambda X$ for $\mu$ an eigenvalue different than $\lambda$? Say $R^k_\mu x = 0$. Then, does it follow that $x \in R^k_\lambda X$? From what was just shown

$$x = y + z, \ y \in R^k_\lambda X, \ z \in N(R^k_\lambda)$$

Then

$$0 = R^p_\mu x = R^p_\mu y + R^p_\mu z$$

Here $p = k(\mu)$. This is where it is important that $\mu \in \Lambda$. However, $N(R^k_\lambda)$ and $R^k_\lambda X$ are invariant under $R^p_\mu$ since it is clear that $R_\lambda$ and $R_\mu$ commute. Thus $R^p_\mu y = -R^p_\mu z$ and $R^p_\mu y \in R^k_\lambda X, -R^p_\mu z \in N(R^k_\lambda)$ and these are equal. Hence they are both 0. Now it was just shown that $R_\mu$ is one to one on $N(R^k_\lambda)$ and so $z = 0$. Hence $x = y \in R^k_\lambda X$.

Note that in the last step, we can’t conclude that $y = 0$ because we only know that $R_\mu$ is one to one on $R^k_\lambda X$ if $\mu$ is sufficiently close to $\lambda$. The above is about compact mappings from a single space to itself. However, there are also mappings which have closed range which map from one space to another. The Fredholm operators have this property that their image is closed. These are discussed next.

Suppose $T \in \mathcal{L}(X,Y)$. Then $TX$ is a subspace of $Y$ and so it has a Hamel basis $\mathcal{B}$. Extending $\mathcal{B}$ to a Hamel basis for $Y$ yields $\mathcal{C}$. Then $Y = \text{span}(\mathcal{B}) \oplus \text{span}(\mathcal{C} \setminus \mathcal{B})$. Thus $Y = TX \oplus E$. For more on this, see [29].

**Definition B.4.7** Let $T \in \mathcal{L}(X,Y)$. Then this is a Fredholm operator means

1. $\dim(\ker(T)) < \infty$
2. $\dim(E) < \infty$ where $Y = TX \oplus E$

**Proposition B.4.8** Let $T \in \mathcal{L}(X,Y)$. Then $TX$ is closed if and only if there exists $\delta > 0$ such that

$$||Tx|| \geq \delta \text{ dist}(x, \ker(T)).$$

**Proof:** First suppose $TX$ is closed. Let $\hat{T} : X/\ker(T) \to Y$ be defined as $\hat{T}([x]) \equiv Tx$. Then by Theorem 3.3.4, $\hat{T}$ is one to one and continuous and $X/\ker(T)$ is a Banach space, $||\hat{T}|| \leq ||T||$. Also $\hat{T}$ has the same range as $T$. Thus $TX$ is the same as $\hat{T}(X/\ker(T))$ and $\hat{T} \in \mathcal{L}(X/\ker(T),Y)$. By the open mapping theorem, $\hat{T}$ is continuous and has continuous inverse. Recall

$$||[x]|| \equiv \inf \{||x + z|| : z \in \ker T\} = \text{ dist}(x, \ker(T))$$

Then

$$\text{ dist}(x, \ker(T)) = ||[x]|| = ||\hat{T}^{-1}\hat{T}[x]|| \leq ||\hat{T}^{-1}|| ||\hat{T}[x]|| = ||\hat{T}^{-1}|| ||Tx||$$

and so,

$$||Tx|| \geq \delta \text{ dist}(x, \ker(T))$$
where $\delta = 1/\| T^{-1} \|$.

Next suppose the inequality holds. Why will $TX$ be closed? Say $\{Tx_n\}$ is a sequence in $TX$ converging to $y$. Then by the inequality,

$$\|Tx_n - Tx_m\| \geq \delta \text{dist} (x_n - x_m, \ker (T)) = \delta \| [x_n] - [x_m]\|_{X/\ker (T)}$$

showing that $\{[x_n]\}$ is a Cauchy sequence in $X/\ker (T)$. Therefore, since this is a Banach space, there exists $[x]$ such that $[x_n] \to [x]$ in $X/\ker (T)$ and so $\hat{T} ([x_n]) \to \hat{T} ([x])$ in $Y$. But this is the same as saying that $T(x_n) \to T(x)$. It follows that $y = Tx$ and so $TX$ is indeed closed.

\textbf{Theorem B.4.9} If $T$ is a Fredholm operator, then $TX$ is closed in $Y$.

\textbf{Proof:} Recall that $Y = TX \oplus E$ where $E$ is a closed subspace of $Y$. In fact, $E$ is finite dimensional, but it is only needed that $E$ is closed. Let $T_0 \in \mathcal{L}(X \times E, TX \oplus E)$ be given by

$$T_0 (x, e) \equiv Tx + e$$

Let the norm on $X \times E$ be

$$\|(x, e)\|_{X \times E} \equiv \max \{\|x\|_X, \|e\|_E\}$$

Thus $T_0 (x, e) = 0$ implies both $Tx = 0$ and $e = 0$. Thus $\ker (T_0) = \ker (T) \times \{0\}$.

Also, $T_0 (X \times E)$ is closed in $Y$ because in fact it is all of $Y$, $TX \oplus E$. By Proposition \textbf{B.4.8}, there exists $\delta > 0$ such that

$$\|T_0 (x, e)\|_Y \geq \delta \text{dist} ((x, e), \ker (T_0)) = \delta \text{dist} ((x, e), \ker (T) \times \{0\}) \geq \delta \text{dist} (x, \ker (T))$$

Then

$$\|Tx\|_Y \equiv \|T_0 (x, 0)\|_Y \geq \delta \text{dist} (x, \ker (T))$$

and by Proposition \textbf{B.4.8}, $TX$ is closed. \blacksquare

Actually, the above proves the following corollary.

\textbf{Corollary B.4.10} If $TX \oplus E$ is closed in $Y$ and $E$ is a closed subspace of $Y$, then $TX$ is closed. Here $T \in \mathcal{L}(X, Y)$.

Note that it appears that $\dim (\ker (T)) < \infty$ was not really needed.

Let $B$ be a Hamel basis for $TX$ and consider $A \equiv \{x : Tx \in B\}$. Then this is a linearly independent set of vectors in $X$. Suppose now that $\ker (T) = \text{span} (z_1, \cdots, z_n)$ where $\{z_1, \cdots, z_n\}$ is linearly independent so here the assumption that $\ker (T)$ has finite dimensions is being used. Then if $x \in X, Tx \in TX$ and so there are finitely many vectors $x_i \in A$ such that

$$Tx = \sum_i c_i Tx_i.$$
Hence
\[ T \left( x - \sum_i c_i x_i \right) = 0 \]
so
\[ x - \sum_i c_i x_i = \sum_{j=1}^n a_j z_j \]

Hence \( X = \text{span} (A) + \ker (T) \). In fact, \( \{ A, \{ z_1, \cdots, z_n \} \} \) is linearly independent as is easily seen and so this is a basis for \( X \). Hence
\[ X = \text{span} (A) \oplus \ker (T) \equiv X_1 \oplus \ker (T) \]

Is \( X_1 \) closed? Define \( S : TX \to X_1 \) as follows: \( Sy = x \in X_1 \) such that \( Tx = y \). Since \( T \) is one to one on \( X_1 \), there is only one such \( x \). Is \( S \) continuous? Yes, this is so by the open mapping theorem. It is just the inverse of a continuous one to one linear onto map. Now this reduces to the situation discussed above in Corollary B.4.10. You have \( S \in \mathcal{L}(TX, X_1) \) and \( S(TX) \oplus \ker (T) \) is all of \( X \) and so it is closed in \( X \). Therefore, \( S(TX) = X_1 \) is closed. This, along with the above proves the following.

**Theorem B.4.11** Let \( T \in \mathcal{L}(X, Y) \) be a Fredholm operator and suppose \( \ker (T) \) is finite dimensional and that \( Y = TX \oplus E \) where \( E \) is a finite dimensional subspace or more generally closed. Then \( TX \) is closed and also for \( X = X_1 \oplus \ker (T) \), it follows that \( X_1 \) is closed.

## B.5 Closed Subspaces

**Theorem B.5.1** Let \( X \) be a Banach space and let \( V = \text{span} (x_1, \cdots, x_n) \). Then \( V \) is a closed subspace of \( X \).

**Proof:** Without loss of generality, it can be assumed \( \{ x_1, \cdots, x_n \} \) is linearly independent. Otherwise, delete those vectors which are in the span of the others till a linearly independent set is obtained. Let
\[ x = \lim_{p \to \infty} \sum_{k=1}^n c_k^p x_k \in V. \quad (2.5.5) \]

First suppose \( c^p \equiv (c_1^p, \cdots, c_n^p) \) is not bounded in \( F^n \). Then \( d^p \equiv c^p / ||c^p||_{F^n} \) is a unit vector in \( F^n \) and so there exists a subsequence, still denoted by \( d^p \) which converges to \( d \) where \( ||d|| = 1 \). Then
\[ 0 = \lim_{p \to \infty} \frac{x}{||c^p||} = \lim_{p \to \infty} \sum_{k=1}^n d_k^p x_k = \sum_{k=1}^n d_k x_k \]
where \( \sum_k |d_k|^2 = 1 \) in contradiction to the linear independence of the \( \{x_1, \cdots, x_n\} \). Hence it must be the case that \( c^p \) is bounded in \( F^n \). Then taking a subsequence, still denoted as \( p \), it can be assumed \( c^p \to c \) and then in Appendix B it follows

\[
x = \sum_{k=1}^{n} c_k x_k \in \text{span} \{x_1, \cdots, x_n\}.
\]

Proposition B.5.2 Let \( E \) be a separable Banach space. Then there exists an increasing sequence of subspaces, \( \{F_n\} \) such that \( \dim (F_{n+1}) - \dim (F_n) \leq 1 \) and equals 1 for all \( n \) if the dimension of \( E \) is infinite. Also \( \bigcup_{n=1}^\infty F_n \) is dense in \( E \). In the case where \( E \) is infinite dimensional, \( F_n = \text{span} (e_1, \cdots, e_n) \) where for each \( n \)

\[
\text{dist} (e_{n+1}, F_n) \geq \frac{1}{2} \tag{2.5.6}
\]

and defining,

\[
G_k \equiv \text{span} \{e_j : j \neq k\}
\]

\[
\text{dist} (e_n, G_k) \geq \frac{1}{4}. \tag{2.5.7}
\]

**Proof:** Since \( E \) is separable, so is \( \partial B (0, 1) \), the boundary of the unit ball. Let \( \{w_k\}_{k=1}^\infty \) be a countable dense subset of \( \partial B (0, 1) \).

Let \( e_1 = w_1 \). Let \( F_1 = \text{Fspan} e_1 \). Suppose \( F_n \) has been obtained and equals \( \text{span} (e_1, \cdots, e_n) \) where \( \{e_1, \cdots, e_n\} \) is independent, \( ||e_k|| = 1 \), and

\[
\text{dist} (e_n, \text{span} (e_1, \cdots, e_{n-1})) \geq \frac{1}{2}.
\]

For each \( n \), \( F_n \) is closed by Theorem \( B.5.1 \).

If \( F_n \) contains \( \{w_k\}_{k=1}^\infty \), let \( F_m = F_n \) for all \( m > n \). Otherwise, pick \( w \in \{w_k\} \) to be the point of \( \{w_k\}_{k=1}^\infty \) having the smallest subscript which is not contained in \( F_n \). Then \( w \) is at a positive distance, \( \lambda \) from \( F_n \) because \( F_n \) is closed. Therefore, there exists \( y \in F_n \) such that \( \lambda \leq ||y - w|| \leq 2\lambda \). Let \( e_{n+1} = \frac{w - y}{||w - y||} \). It follows

\[
w = ||w - y|| e_{n+1} + y \in \text{span} (e_1, \cdots, e_{n+1}) = F_{n+1}
\]

Then if \( x \in \text{span} (e_1, \cdots, e_n) \),

\[
||e_{n+1} - x|| = \left|\left| \frac{w - y}{||w - y||} - x \right|\right|
\]  
= \left|\left| \frac{w - y}{||w - y||} - \frac{||y - w||}{||w - y||} x \right|\right|
\]  
\geq \frac{1}{2\lambda} ||w - y - ||w - y|| x||
\]  
\geq \frac{\lambda}{2\lambda} = \frac{1}{2}.
This has shown the existence of an increasing sequence of subspaces, \( \{ F_n \} \) as described above. It remains to show the union of these subspaces is dense. First note that the union of these subspaces must contain the \( \{ w_k \}_{k=1}^\infty \) because if \( w_m \) is missing, then it would contradict the construction at the \( m^{th} \) step. That one should have been chosen. However, \( \{ w_k \}_{k=1}^\infty \) is dense in \( \partial B(0,1) \). If \( x \in E \) and \( x \neq 0 \), then \( \frac{x}{||x||} \in \partial B(0,1) \) then there exists

\[
w_m \in \{ w_k \}_{k=1}^\infty \subseteq \cup_{n=1}^\infty F_n
\]
such that \( ||w_m - \frac{x}{||x||}|| < \frac{\varepsilon}{||x||} \). But then

\[
||x|| w_m - x|| < \varepsilon
\]
and so \( ||x|| w_m \) is a point of \( \cup_{n=1}^\infty F_n \) which is within \( \varepsilon \) of \( x \). This proves \( \cup_{n=1}^\infty F_n \) is dense as desired.

Let \( y \in G_k \). Thus for some \( n \),

\[
y = \sum_{j=1}^{k-1} c_j e_j + \sum_{j=k+1}^n c_j e_j
\]
and I need to show \( ||y - e_k|| \geq 1/4 \). Without loss of generality, \( c_n \neq 0 \) and \( n > k \). Suppose 2.5.7 does not hold for some such \( y \) so that

\[
\left\| e_k - \left( \sum_{j=1}^{k-1} c_j e_j + \sum_{j=k+1}^n c_j e_j \right) \right\| < \frac{1}{4}, \quad (2.5.8)
\]
Then from the construction,

\[
\frac{1}{4} > |c_n| \left\| e_k - \left( \sum_{j=1}^{k-1} \frac{c_j}{c_n} e_j + \sum_{j=k+1}^{n-1} \frac{c_j}{c_n} e_j + e_n \right) \right\|
\geq |c_n| \frac{1}{2}
\]
and so \( |c_n| < 1/2 \). Consider the left side of 2.5.7. By the construction

\[
\left\| c_n (e_k - e_n) + (1 - c_n) e_k - \left( \sum_{j=1}^{k-1} c_j e_j + \sum_{j=k+1}^{n-1} c_j e_j \right) \right\|
\geq |1 - c_n| - |c_n| \left\| (e_k - e_n) - \left( \sum_{j=1}^{k-1} \frac{c_j}{c_n} e_j + \sum_{j=k+1}^{n-1} \frac{c_j}{c_n} e_j \right) \right\|
\geq |1 - c_n| - |c_n| \frac{1}{2} \geq 1 - \frac{3}{2} |c_n| > 1 - \frac{3}{2} \frac{1}{2} = \frac{1}{4},
\]
a contradiction. This proves the desired estimate.
Appendix C

Hilbert Spaces

Hilbert spaces are just complete inner product spaces.

C.1 Basic Theory

Definition C.1.1 Let $X$ be a vector space. An inner product is a mapping from $X \times X$ to $\mathbb{C}$ if $X$ is complex and from $X \times X$ to $\mathbb{R}$ if $X$ is real, denoted by $(x,y)$ which satisfies the following.

\begin{align*}
(x,x) &\geq 0, \quad (x,x) = 0 \text{ if and only if } x = 0, \quad (3.1.1) \\
(x,y) &= \overline{(y,x)}. \quad (3.1.2)
\end{align*}

For $a, b \in \mathbb{C}$ and $x, y, z \in X$,

\[(ax + by, z) = a(x, z) + b(y, z). \quad (3.1.3)\]

Note that $(3.1.2)$ and $(3.1.3)$ imply \((x, ay + bz) = a(x, y) + b(x, z)\). Such a vector space is called an inner product space.

The Cauchy Schwarz inequality is fundamental for the study of inner product spaces.

Theorem C.1.2 (Cauchy Schwarz) In any inner product space

\[|(x, y)| \leq ||x|| \cdot ||y||.\]

Proof: Let $\omega \in \mathbb{C}, |\omega| = 1$, and $\overline{\omega}(x, y) = |(x, y)| = \text{Re}(x, y\omega)$. Let

\[F(t) = (x + ty\omega, x + t\omega y).\]

If $y = 0$ there is nothing to prove because

\[(x, 0) = (x, 0 + 0) = (x, 0) + (x, 0)\]
and so \((x, 0) = 0\). Thus, it can be assumed \(y \neq 0\). Then from the axioms of the inner product,

\[
F(t) = ||x||^2 + 2t \text{Re}(x, \omega y) + t^2||y||^2 \geq 0.
\]

This yields

\[
||x||^2 + 2t ||(x, y)|| + t^2||y||^2 \geq 0.
\]

Since this inequality holds for all \(t \in \mathbb{R}\), it follows from the quadratic formula that

\[
4|(x, y)|^2 - 4||x||^2||y||^2 \leq 0.
\]

This yields the conclusion and proves the theorem.

**Proposition C.1.3** For an inner product space, \(||x|| \equiv (x, x)^{1/2}\) does specify a norm.

**Proof:** All the axioms are obvious except the triangle inequality. To verify this,

\[
||x + y||^2 \equiv (x + y, x + y) \equiv ||x||^2 + ||y||^2 + 2 \text{Re}(x, y)
\]

\[
\leq ||x||^2 + ||y||^2 + 2 ||(x, y)||
\]

\[
\leq ||x||^2 + ||y||^2 + 2 ||x|| ||y|| = (||x|| + ||y||)^2.
\]

The following lemma is called the parallelogram identity.

**Lemma C.1.4** In an inner product space,

\[
||x + y||^2 + ||x - y||^2 = 2||x||^2 + 2||y||^2.
\]

The proof, a straightforward application of the inner product axioms, is left to the reader.

**Lemma C.1.5** For \(x \in H\), an inner product space,

\[
||x|| = \sup_{||y|| \leq 1} |(x, y)|
\]

**Proof:** By the Cauchy Schwarz inequality, if \(x \neq 0\),

\[
||x|| \geq \sup_{||y|| \leq 1} |(x, y)| \geq \left( x, \frac{x}{||x||} \right) = ||x||.
\]

It is obvious that (3.1.4) holds in the case that \(x = 0\).

**Definition C.1.6** A Hilbert space is an inner product space which is complete. Thus a Hilbert space is a Banach space in which the norm comes from an inner product as described above.

In Hilbert space, one can define a projection map onto closed convex nonempty sets.
Definition C.1.7 A set, $K$, is convex if whenever $\lambda \in [0,1]$ and $x,y \in K$, $\lambda x + (1-\lambda)y \in K$.

Theorem C.1.8 Let $K$ be a closed convex nonempty subset of a Hilbert space, $H$, and let $x \in H$. Then there exists a unique point $Px \in K$ such that $\|Px - x\| \leq \|y - x\|$ for all $y \in K$.

Proof: Consider uniqueness. Suppose that $z_1$ and $z_2$ are two elements of $K$ such that for $i = 1,2$,

$$\|z_i - x\| \leq \|y - x\|$$

for all $y \in K$. Also, note that since $K$ is convex, $\frac{z_1 + z_2}{2} \in K$.

Therefore, by the parallelogram identity,

$$\|z_1 - x\|^2 \leq \|\frac{z_1 + z_2}{2} - x\|^2 = \|\frac{z_1 - x}{2} + \frac{z_2 - x}{2}\|^2$$

$$= 2(\|\frac{z_1 - x}{2}\|^2 + \|\frac{z_2 - x}{2}\|^2) - \|\frac{z_1 - z_2}{2}\|^2$$

$$= \frac{1}{2}\|z_1 - x\|^2 + \frac{1}{2}\|z_2 - x\|^2 - \|\frac{z_1 - z_2}{2}\|^2$$

$$\leq \|z_1 - x\|^2 - \|\frac{z_1 - z_2}{2}\|^2,$$

where the last inequality holds because of letting $z_1 = z_2$ and $y = z_1$. Hence $z_1 = z_2$ and this shows uniqueness.

Now let $\lambda = \inf\{\|x - y\| : y \in K\}$ and let $y_n$ be a minimizing sequence. This means $\{y_n\} \subseteq K$ satisfies $\lim_{n \to \infty} \|x - y_n\| = \lambda$. Now the following follows from properties of the norm.

$$\|y_n - x + y_m - x\|^2 = 4(\|\frac{y_n + y_m}{2} - x\|^2)$$

Then by the parallelogram identity, and convexity of $K$, $\frac{y_n + y_m}{2} \in K$, and so

$$\| (y_n - x) - (y_m - x) \|^2 = 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4(\|\frac{y_n + y_m}{2} - x\|^2)$$

$$\leq 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4\lambda^2.$$

Since $\|x - y_n\| \to \lambda$, this shows $\{y_n - x\}$ is a Cauchy sequence. Thus also $\{y_n\}$ is a Cauchy sequence. Since $H$ is complete, $y_n \to y$ for some $y \in H$ which must be in $K$ because $K$ is closed. Therefore

$$\|x - y\| = \lim_{n \to \infty} \|x - y_n\| = \lambda.$$

Let $Px = y$. 

Corollary C.1.9 Let $K$ be a closed, convex, nonempty subset of a Hilbert space, $H$, and let $x \in H$. Then for $z \in K$, $z = Px$ if and only if
\[ \text{Re}(x - z, y - z) \leq 0 \quad (3.1.6) \]
for all $y \in K$.

Before proving this, consider what it says in the case where the Hilbert space is $\mathbb{R}^n$.

Condition 3.1.6 says the angle, $\theta$, shown in the diagram is always obtuse. Remember from calculus, the sign of $\langle x, y \rangle$ is the same as the sign of the cosine of the included angle between $x$ and $y$. Thus, in finite dimensions, the conclusion of this corollary says that $z = Px$ exactly when the angle of the indicated angle is obtuse. Surely the picture suggests this is reasonable.

The inequality 3.1.6 is an example of a variational inequality and this corollary characterizes the projection of $x$ onto $K$ as the solution of this variational inequality.

Proof of Corollary: Let $z \in K$ and let $y \in K$ also. Since $K$ is convex, it follows that if $t \in [0, 1]$,
\[ z + t(y - z) = (1 - t)z + ty \in K. \]

Furthermore, every point of $K$ can be written in this way. (Let $t = 1$ and $y \in K$.) Therefore, $z = Px$ if and only if for all $y \in K$ and $t \in [0, 1]$,
\[ ||x - (z + t(y - z))||^2 = ||(x - z) - t(y - z)||^2 \geq ||x - z||^2 \]
for all $t \in [0, 1]$ and $y \in K$ if and only if for all $t \in [0, 1]$ and $y \in K$
\[ ||x - z||^2 + t^2 ||y - z||^2 - 2t \text{Re}(x - z, y - z) \geq ||x - z||^2 \]
If and only if for all $t \in [0, 1]$,
\[ t^2 ||y - z||^2 - 2t \text{Re}(x - z, y - z) \geq 0. \quad (3.1.7) \]
Now this is equivalent to holding for all $t \in (0, 1)$. Therefore, dividing by $t \in (0, 1)$, \[ t ||y - z||^2 - 2 \text{Re}(x - z, y - z) \geq 0 \]
for all $t \in (0, 1)$ which is equivalent to 3.1.6. \[ \square \]

Corollary C.1.10 Let $K$ be a nonempty convex closed subset of a Hilbert space, $H$. Then the projection map, $P$ is continuous. In fact,
\[ |Px - Py| \leq |x - y|. \]
**Proof:** Let \( x, x' \in H \). Then by Corollary C.1.9,

\[
\Re (x' - Px', Px - Px') \leq 0, \quad \Re (x - Px, Px' - Px) \leq 0
\]

Hence

\[
0 \leq \Re (x - Px, Px - Px') - \Re (x' - Px', Px - Px') = \Re (x - x', Px -Px') - |Px - Px'|^2
\]

and so

\[
|Px - Px'|^2 \leq |x - x'| |Px - Px'|.
\]

The next corollary is a more general form for the Brouwer fixed point theorem.

**Corollary C.1.11** Let \( f : K \to K \) where \( K \) is a convex compact subset of \( \mathbb{R}^n \). Then \( f \) has a fixed point.

**Proof:** Let \( K \subseteq \overline{B(0,R)} \) and let \( P \) be the projection map onto \( K \). Then consider the map \( f \circ P \) which maps \( \overline{B(0,R)} \) to \( \overline{B(0,R)} \) and is continuous. By the Brouwer fixed point theorem for balls, this map has a fixed point. Thus there exists \( x \) such that

\[
f \circ P (x) = x
\]

Now the equation also requires \( x \in K \) and so \( P(x) = x \). Hence \( f(x) = x \).

**Definition C.1.12** Let \( H \) be a vector space and let \( U \) and \( V \) be subspaces. \( U \oplus V = H \) if every element of \( H \) can be written as a sum of an element of \( U \) and an element of \( V \) in a unique way.

The case where the closed convex set is a closed subspace is of special importance and in this case the above corollary implies the following.

**Corollary C.1.13** Let \( K \) be a closed subspace of a Hilbert space, \( H \), and let \( x \in H \). Then for \( z \in K \), \( z = Px \) if and only if

\[
(x - z, y) = 0
\]

for all \( y \in K \). Furthermore, \( H = K \oplus K^\perp \) where

\[
K^\perp = \{ x \in H : (x, k) = 0 \text{ for all } k \in K \}
\]

and

\[
||x||^2 = ||x - Px||^2 + ||Px||^2.
\]

**Proof:** Since \( K \) is a subspace, the condition \( C.1.10 \) implies \( \Re(x - z, y) \leq 0 \) for all \( y \in K \). Replacing \( y \) with \(-y\), it follows \( \Re(x - z, -y) \leq 0 \) which implies \( \Re(x - z, y) \geq 0 \) for all \( y \). Therefore, \( \Re(x - z, y) = 0 \) for all \( y \in K \). Now let
\(|\alpha| = 1\) and \(\alpha(x - z, y) = |(x - z, y)|\). Since \(K\) is a subspace, it follows \(\overline{\alpha}y \in K\) for all \(y \in K\). Therefore,

\[
0 = \text{Re}(x - z, \overline{\alpha}y) = (x - z, \overline{\alpha}y) = \alpha(x - z, y) = |(x - z, y)|.
\]

This shows that \(z = Px\), if and only if \(\text{II}\).

For \(x \in H\), \(x = x - Px + Px\) and from what was just shown, \(x - Px \in K^\perp\) and \(Px \in K\). This shows that \(K^\perp + K = H\). Is there only one way to write a given element of \(H\) as a sum of a vector in \(K\) with a vector in \(K^\perp\)? Suppose \(y + z = y_1 + z_1\) where \(z, z_1 \in K^\perp\) and \(y, y_1 \in K\). Then \((y - y_1) = (z_1 - z)\) and so from what was just shown, \((y - y_1, y - y_1) = (y - y_1, z_1 - z) = 0\) which shows \(y_1 = y\) and consequently \(z_1 = z\). Finally, letting \(z = Px\),

\[
|z|^2 = (x - z + z, x - z + z) = |x - z|^2 + (x - z, z) + (z, x - z) + |z|^2
\]

\[
= |x - z|^2 + |z|^2
\]

The following theorem is called the Riesz representation theorem for the dual of a Hilbert space. If \(z \in H\) then define an element \(f \in H'\) by the rule \((x, z) \equiv f(x)\). It follows from the Cauchy Schwarz inequality and the properties of the inner product that \(f \in H'\). The Riesz representation theorem says that all elements of \(H'\) are of this form.

**Theorem C.1.14** Let \(H\) be a Hilbert space and let \(f \in H'\). Then there exists a unique \(z \in H\) such that

\[
f(x) = (x, z)
\]

for all \(x \in H\).

**Proof:** Letting \(y, w \in H\) the assumption that \(f\) is linear implies

\[
f(yf(w) - f(y)w) = f(w)f(y) - f(y)f(w) = 0
\]

which shows that \(yf(w) - f(y)w \in f^{-1}(0)\), which is a closed subspace of \(H\) since \(f\) is continuous. If \(f^{-1}(0) = H\), then \(f\) is the zero map and \(z = 0\) is the unique element of \(H\) which satisfies \(\text{II}\). If \(f^{-1}(0) \neq H\), pick \(u \notin f^{-1}(0)\) and let \(w \equiv u - Pu \neq 0\). Thus Corollary \(\text{II}\) implies \((y, w) = 0\) for all \(y \in f^{-1}(0)\). In particular, let \(y = x f(w) - f(x)w\) where \(x \in H\) is arbitrary. Therefore,

\[
0 = (f(w)x - f(x)w, w) = f(w)(x, w) - f(x)||w||^2.
\]

Thus, solving for \(f(x)\) and using the properties of the inner product,

\[
f(x) = (x, \frac{f(w)w}{||w||^2})
\]

Let \(z = \frac{f(w)w}{||w||^2}\). This proves the existence of \(z\). If \(f(x) = (x, z_i)\) \(i = 1, 2, \) for all \(x \in H\), then for all \(x \in H\), then \((x, z_1 - z_2) = 0\) which implies, upon taking \(x = z_1 - z_2\) that \(z_1 = z_2\). \(\square\)

If \(R : H \to H'\) is defined by \(Rx(y) \equiv (y, x)\), the Riesz representation theorem above states this map is onto. This map is called the Riesz map. It is routine to show \(R\) is linear and \(|Rx| = |x|\).
C.2 Approximations In Hilbert Space

The Gram Schmidt process applies in any Hilbert space.

**Theorem C.2.1** Let \( \{x_1, \ldots, x_n\} \) be a basis for \( M \) a subspace of \( H \) a Hilbert space. Then there exists an orthonormal basis for \( M \), \( \{u_1, \ldots, u_n\} \) which has the property that for each \( k \leq n \), \( \text{span}(x_1, \ldots, x_k) = \text{span}(u_1, \ldots, u_k) \). Also if \( \{x_1, \ldots, x_n\} \subseteq H \), then \( \text{span}(x_1, \ldots, x_n) \) is a closed subspace.

**Proof:** Let \( \{x_1, \ldots, x_n\} \) be a basis for \( M \). Let \( u_1 \equiv x_1 / |x_1| \). Thus for \( k = 1 \), \( \text{span}(u_1) = \text{span}(x_1) \) and \( \{u_1\} \) is an orthonormal set. Now suppose for some \( k < n \), \( u_1, \ldots, u_k \) have been chosen such that \( (u_j, u_l) = \delta_{jl} \) and \( \text{span}(x_1, \ldots, x_k) = \text{span}(u_1, \ldots, u_k) \). Then define

\[
  u_{k+1} \equiv \frac{x_{k+1} - \sum_{j=1}^{k} (x_{k+1} \cdot u_j) u_j}{|x_{k+1} - \sum_{j=1}^{k} (x_{k+1} \cdot u_j) u_j|},
\]

where the denominator is not equal to zero because the \( x_j \) form a basis and so \( x_{k+1} \notin \text{span}(x_1, \ldots, x_k) = \text{span}(u_1, \ldots, u_k) \).

Thus by induction,

\[ u_{k+1} \in \text{span}(u_1, \ldots, u_k, x_{k+1}) = \text{span}(x_1, \ldots, x_k, x_{k+1}) \].

Also, \( x_{k+1} \in \text{span}(u_1, \ldots, u_k, u_{k+1}) \) which is seen easily by solving (3.2.11) for \( x_{k+1} \) and it follows

\[ \text{span}(x_1, \ldots, x_k, x_{k+1}) = \text{span}(u_1, \ldots, u_k, u_{k+1}) \].

If \( l \leq k \),

\[
  (u_{k+1} \cdot u_l) = C \left( (x_{k+1} \cdot u_l) - \sum_{j=1}^{k} (x_{k+1} \cdot u_j) (u_j \cdot u_l) \right) \\
  = C \left( (x_{k+1} \cdot u_l) - \sum_{j=1}^{k} (x_{k+1} \cdot u_j) \delta_{lj} \right) \\
  = C \left( (x_{k+1} \cdot u_l) - (x_{k+1} \cdot u_l) \right) = 0.
\]

The vectors, \( \{u_j\}_{j=1}^{n} \), generated in this way are therefore an orthonormal basis because each vector has unit length.

Consider the second claim about finite dimensional subspaces. Without loss of generality, assume \( \{x_1, \ldots, x_n\} \) is linearly independent. If it is not, delete vectors
until a linearly independent set is obtained. Then by the first part, span \((x_1, \cdots, x_n) = \text{span} (u_1, \cdots, u_n) \equiv M\) where the \(u_i\) are an orthonormal set of vectors. Suppose \(\{y_k\} \subseteq M\) and \(y_k \to y \in H\). Is \(y \in M\)? Let

\[ y_k = \sum_{j=1}^{n} c^k_j u_j \]

Then let \(c^k \equiv (c^k_1, \cdots, c^k_n)^T\). Then

\[
|c^k - c^l|^2 = \sum_{j=1}^{n} |c^k_j - c^l_j|^2 = \left( \sum_{j=1}^{n} (c^k_j - c^l_j) u_j, \sum_{j=1}^{n} (c^k_j - c^l_j) u_j \right) = ||y_k - y_l||^2
\]

which shows \(\{c^k\}\) is a Cauchy sequence in \(\mathbb{R}^n\) and so it converges to \(c \in \mathbb{R}^n\). Thus

\[ y = \lim_{k \to \infty} y_k = \lim_{k \to \infty} \sum_{j=1}^{n} c^k_j u_j = \sum_{j=1}^{n} c_j u_j \in M. \]

This completes the proof.

**Theorem C.2.2** Let \(M\) be the span of \(\{u_1, \cdots, u_n\}\) in a Hilbert space \(H\) and let \(y \in H\). Then \(P y\) is given by

\[ Py = \sum_{k=1}^{n} (y, u_k) u_k \]  \hspace{1cm} (3.2.12)

and the distance is given by

\[
\sqrt{||y||^2 - \sum_{k=1}^{n} (y, u_k)^2}. \]  \hspace{1cm} (3.2.13)

**Proof:**

\[
\left( y - \sum_{k=1}^{n} (y, u_k) u_k, u_p \right) = (y, u_p) - \sum_{k=1}^{n} (y, u_k) (u_k, u_p) = (y, u_p) - (y, u_p) = 0
\]

It follows that

\[
\left( y - \sum_{k=1}^{n} (y, u_k) u_k, u \right) = 0
\]

for all \(u \in M\) and so by Corollary \(\text{Cor. } C.1.13\) this verifies \(3.2.12\).
The square of the distance, $d$ is given by

$$
d^2 = \left( y - \sum_{k=1}^{n} (y, u_k) u_k, y - \sum_{k=1}^{n} (y, u_k) u_k \right) = |y|^2 - 2 \sum_{k=1}^{n} |(y, u_k)|^2 + \sum_{k=1}^{n} |(y, u_k)|^2
$$

and this shows (3.2.13).

What if the subspace is the span of vectors which are not orthonormal? There is a very interesting formula for the distance between a point of a Hilbert space and a finite dimensional subspace spanned by an arbitrary basis.

**Definition C.2.3** Let $\{x_1, \cdots, x_n\} \subseteq H$, a Hilbert space. Define

$$
G(x_1, \cdots, x_n) \equiv \begin{pmatrix}
(x_1, x_1) & \cdots & (x_1, x_n) \\
\vdots & \ddots & \vdots \\
(x_n, x_1) & \cdots & (x_n, x_n)
\end{pmatrix} \quad (3.2.14)
$$

Thus the $ij$th entry of this matrix is $(x_i, x_j)$. This is sometimes called the Gram matrix. Also define $G(x_1, \cdots, x_n)$ as the determinant of this matrix, also called the Gram determinant.

$$
G(x_1, \cdots, x_n) \equiv \left| \begin{array}{cccc}
(x_1, x_1) & \cdots & (x_1, x_n) \\
\vdots & \ddots & \vdots \\
(x_n, x_1) & \cdots & (x_n, x_n)
\end{array} \right| \quad (3.2.15)
$$

The theorem is the following.

**Theorem C.2.4** Let $M = \text{span} \{x_1, \cdots, x_n\} \subseteq H$, a Real Hilbert space where $\{x_1, \cdots, x_n\}$ is a basis and let $y \in H$. Then letting $d$ be the distance from $y$ to $M$,

$$
d^2 = \frac{G(x_1, \cdots, x_n, y)}{G(x_1, \cdots, x_n)} \quad (3.2.16)
$$

**Proof:** By Theorem C.2.1, $M$ is a closed subspace of $H$. Let $\sum_{k=1}^{n} \alpha_k x_k$ be the element of $M$ which is closest to $y$. Then by Corollary C.1.13,

$$
\left( y - \sum_{k=1}^{n} \alpha_k x_k, x_p \right) = 0
$$

for each $p = 1, 2, \cdots, n$. This yields the system of equations,

$$
(y, x_p) = \sum_{k=1}^{n} (x_p, x_k) \alpha_k, p = 1, 2, \cdots, n \quad (3.2.17)
$$
Also by Corollary C.1.13,
\[ ||y||^2 = \left( y - \sum_{k=1}^{n} \alpha_k x_k \right)^2 + \left( \sum_{k=1}^{n} \alpha_k x_k \right)^2 \]
and so, using 3.2.17,
\[ ||y||^2 = d^2 + \sum_{j} \left( \sum_{k} \alpha_k (x_k, x_j) \right) \alpha_j \]
\[ \equiv d^2 + y_x^T \alpha \quad (3.2.18) \]
in which
\[ y_x^T \equiv ((x, x_1), \ldots, (x, x_n)) \], \[ \alpha^T \equiv (\alpha_1, \ldots, \alpha_n) \].

Then 3.2.17 and 3.2.18 imply the following system
\[
\begin{pmatrix}
G(x_1, \ldots, x_n) & 0 \\
y_x^T & 1 \\
\end{pmatrix}
\begin{pmatrix}
\alpha \\
d^2 ||y||^2 \\
\end{pmatrix}
= 
\begin{pmatrix}
y_x \\
||y||^2 \\
\end{pmatrix}
\]

By Cramer’s rule,
\[
d^2 = \frac{\det \begin{pmatrix} G(x_1, \ldots, x_n) & y_x \\ y_x^T & ||y||^2 \end{pmatrix}}{\det \begin{pmatrix} G(x_1, \ldots, x_n) & 0 \\ y_x^T & 1 \end{pmatrix}}
= \frac{\det \begin{pmatrix} G(x_1, \ldots, x_n) & y_x \\ y_x^T & ||y||^2 \end{pmatrix}}{\det \begin{pmatrix} G(x_1, \ldots, x_n) \\ \det \begin{pmatrix} G(x_1, \ldots, x_n) 
\end{pmatrix} \\
\end{pmatrix}} = \frac{G(x_1, \ldots, x_n, y)}{G(x_1, \ldots, x_n)}
\]
and this proves the theorem.

C.3 Orthonormal Sets

The concept of an orthonormal set of vectors is a generalization of the notion of the standard basis vectors of \( \mathbb{R}^n \) or \( \mathbb{C}^n \).

**Definition C.3.1** Let \( H \) be a Hilbert space. \( S \subseteq H \) is called an orthonormal set if \( ||x|| = 1 \) for all \( x \in S \) and \( (x, y) = 0 \) if \( x, y \in S \) and \( x \neq y \). For any set, \( D \),
\[ D^\perp \equiv \{ x \in H : (x, d) = 0 \text{ for all } d \in D \} \]
If \( S \) is a set, \( \text{span}(S) \) is the set of all finite linear combinations of vectors from \( S \).
C.3. ORTHONORMAL SETS

You should verify that \( D^\perp \) is always a closed subspace of \( H \).

**Theorem C.3.2** In any separable Hilbert space, \( H \), there exists a countable orthonormal set, \( S = \{ x_i \} \) such that the span of these vectors is dense in \( H \). Furthermore, if \( \text{span} (S) \) is dense, then for \( x \in H \),

\[
x = \sum_{i=1}^{\infty} (x, x_i) x_i \equiv \lim_{n \to \infty} \sum_{i=1}^{n} (x, x_i) x_i.
\] (3.3.20)

**Proof:** Let \( \mathcal{F} \) denote the collection of all orthonormal subsets of \( H \). \( \mathcal{F} \) is nonempty because \( \{ x \} \in \mathcal{F} \) where \( \| x \| = 1 \). The set, \( \mathcal{F} \) is a partially ordered set with the order given by set inclusion. By the Hausdorff maximal theorem, there exists a maximal chain, \( \mathcal{C} \) in \( \mathcal{F} \). Then let \( S \equiv \cup \mathcal{C} \). It follows \( S \) must be a maximal orthonormal set of vectors. Why? It remains to verify that \( S \) is countable span (\( S \)) is dense, and the condition, (3.3.20) holds. To see \( S \) is countable note that if \( x, y \in S \), then

\[
|x - y|^2 = \| x \|^2 + \| y \|^2 - 2 \text{Re}(x, y) = \| x \|^2 + \| y \|^2 = 2.
\]

Therefore, the open sets, \( B (x, \frac{1}{2}) \) for \( x \in S \) are disjoint and cover \( S \). Since \( H \) is assumed to be separable, there exists a point from a countable dense set in each of these disjoint balls showing there can only be countably many of the balls and that consequently, \( S \) is countable as claimed.

It remains to verify (3.3.20) and that span (\( S \)) is dense. If span (\( S \)) is not dense, then span (\( S \)) is a closed proper subspace of \( H \) and letting \( y \notin \text{span} (S) \),

\[
z \equiv \frac{y - Py}{\| y - Py \|} \in \text{span} (S)^\perp.
\]

But then \( S \cup \{ z \} \) would be a larger orthonormal set of vectors contradicting the maximality of \( S \).

It remains to verify (3.3.20). Let \( S = \{ x_i \}_{i=1}^{\infty} \) and consider the problem of choosing the constants, \( c_k \) in such a way as to minimize the expression

\[
\left\| x - \sum_{k=1}^{n} c_k x_k \right\|^2 = \| x \|^2 + \sum_{k=1}^{n} |c_k|^2 - \sum_{k=1}^{n} c_k (x, x_k) - \sum_{k=1}^{n} c_k (x, x_k).
\]

This equals

\[
\| x \|^2 + \sum_{k=1}^{n} |c_k - (x, x_k)|^2 - \sum_{k=1}^{n} |(x, x_k)|^2
\]

and therefore, this minimum is achieved when \( c_k = (x, x_k) \) and equals

\[
\| x \|^2 - \sum_{k=1}^{n} |(x, x_k)|^2
\]
Now since \( \text{span}(S) \) is dense, there exists \( n \) large enough that for some choice of constants, \( c_k \),

\[
\left\| x - \sum_{k=1}^{n} c_k x_k \right\|^2 < \varepsilon.
\]

However, from what was just shown,

\[
\left\| x - \sum_{i=1}^{n} (x, x_i) x_i \right\|^2 \leq \left\| x - \sum_{k=1}^{n} c_k x_k \right\|^2 < \varepsilon
\]

showing that \( \lim_{n \to \infty} \sum_{i=1}^{n} (x, x_i) x_i = x \) as claimed. ■

The proof of this theorem contains the following corollary.

**Corollary C.3.3** Let \( S \) be any orthonormal set of vectors and let

\[
\{x_1, \ldots, x_n\} \subseteq S.
\]

Then if \( x \in H \)

\[
\left\| x - \sum_{k=1}^{n} c_k x_k \right\|^2 \geq \left\| x - \sum_{i=1}^{n} (x, x_i) x_i \right\|^2
\]

for all choices of constants, \( c_k \). In addition to this, Bessel’s inequality

\[
\|x\|^2 \geq \sum_{k=1}^{n} |(x, x_k)|^2.
\]

If \( S \) is countable and \( \text{span}(S) \) is dense, then letting \( \{x_i\}_{i=1}^{\infty} = S \), C.3.20 follows.
Appendix D

The Hausdorff Maximal Theorem

The purpose of this appendix is to prove the equivalence between the axiom of choice, the Hausdorff maximal theorem, and the well-ordering principle. The Hausdorff maximal theorem and the well-ordering principle are very useful but a little hard to believe; so, it may be surprising that they are equivalent to the axiom of choice. First it is shown that the axiom of choice implies the Hausdorff maximal theorem, a remarkable theorem about partially ordered sets.

A nonempty set is partially ordered if there exists a partial order, \( \prec \), satisfying 
\[ x \prec x \]
and
\[ \text{if } x \prec y \text{ and } y \prec z \text{ then } x \prec z. \]

An example of a partially ordered set is the set of all subsets of a given set and \( \subseteq \). Note that two elements in a partially ordered sets may not be related. In other words, just because \( x, y \) are in the partially ordered set, it does not follow that either \( x \prec y \) or \( y \prec x \). A subset of a partially ordered set, \( C \), is called a chain if \( x, y \in C \) implies that either \( x \prec y \) or \( y \prec x \). If either \( x \prec y \) or \( y \prec x \) then \( x \) and \( y \) are described as being comparable. A chain is also called a totally ordered set. \( C \) is a maximal chain if whenever \( \tilde{C} \) is a chain containing \( C \), it follows the two chains are equal. In other words \( C \) is a maximal chain if there is no strictly larger chain.

**Lemma D.0.4** Let \( F \) be a nonempty partially ordered set with partial order \( \prec \). Then assuming the axiom of choice, there exists a maximal chain in \( F \).

**Proof:** Let \( \mathcal{X} \) be the set of all chains from \( F \). For \( C \in \mathcal{X} \), let
\[ S_C = \{ x \in F \text{ such that } C \cup \{ x \} \text{ is a chain strictly larger than } C \}. \]
If \( S_C = \emptyset \) for any \( C \), then \( C \) is maximal. Thus, assume \( S_C \neq \emptyset \) for all \( C \in X \). Let \( f(C) \in S_C \). (This is where the axiom of choice is being used.) Let

\[
g(C) = C \cup \{f(C)\}.
\]

Thus \( g(C) \supseteq C \) and \( g(C) \setminus C = \{f(C)\} = \{\text{a single element of } F\} \). A subset \( T \) of \( X \) is called a tower if

\[
\emptyset \in T,
\]

\( C \in T \) implies \( g(C) \in T \),

and if \( S \subseteq T \) is totally ordered with respect to set inclusion, then

\[
\cup S \in T.
\]

Here \( S \) is a chain with respect to set inclusion whose elements are chains.

Note that \( X \) is a tower. Let \( T_0 \) be the intersection of all towers. Thus, \( T_0 \) is a tower, the smallest tower. Are any two sets in \( T_0 \) comparable in the sense of set inclusion so that \( T_0 \) is actually a chain? Let \( C_0 \) be a set of \( T_0 \) which is comparable to every set of \( T_0 \). Such sets exist, \( \emptyset \) being an example. Let

\[
B \equiv \{D \in T_0 : D \supseteq C_0 \text{ and } f(C_0) \notin D\}.
\]

The picture represents sets of \( B \). As illustrated in the picture, \( D \) is a set of \( B \) when \( D \) is larger than \( C_0 \) but fails to be comparable to \( g(C_0) \). Thus there would be more than one chain ascending from \( C_0 \) if \( B \neq \emptyset \), rather like a tree growing upward in more than one direction from a fork in the trunk. It will be shown this can’t take place for any such \( C_0 \) by showing \( B = \emptyset \).

This will be accomplished by showing \( \widetilde{T}_0 \equiv T_0 \setminus B \) is a tower. Since \( T_0 \) is the smallest tower, this will require that \( \widetilde{T}_0 = T_0 \) and so \( B = \emptyset \).

Claim: \( \widetilde{T}_0 \) is a tower and so \( B = \emptyset \).

Proof of the claim: It is clear that \( \emptyset \in \widetilde{T}_0 \) because for \( \emptyset \) to be contained in \( B \) it would be required to properly contain \( C_0 \) which is not possible. Suppose \( D \in \widetilde{T}_0 \).

The plan is to verify \( g(D) \in \widetilde{T}_0 \).

Case 1: \( f(D) \in C_0 \). If \( D \subseteq C_0 \), then since both \( D \) and \( \{f(D)\} \) are contained in \( C_0 \), it follows \( g(D) \subseteq C_0 \) and so \( g(D) \notin B \). On the other hand, if \( D \supseteq C_0 \), then since \( D \in \widetilde{T}_0 \), \( f(C_0) \in D \) and so \( g(D) \) also contains \( f(C_0) \) implying \( g(D) \notin B \). These are the only two cases to consider because \( C_0 \) is comparable to every set of \( T_0 \).
Case 2: \( f(D) \notin C_0 \). If \( D \subseteq C_0 \) it can’t be the case that \( f(D) \notin C_0 \) because if this were so, \( g(D) \) would not compare to \( C_0 \).

Hence if \( f(D) \notin C_0 \), then \( D \supseteq C_0 \). If \( D = C_0 \), then \( f(D) = f(C_0) \in g(D) \) so \( g(D) \notin B \). Therefore, assume \( D \supseteq C_0 \). Then, since \( D \) is in \( T_0 \), \( f(C_0) \in D \) and so \( f(C_0) \in g(D) \). Therefore, \( g(D) \in T_0 \).

Now suppose \( S \) is a totally ordered subset of \( T_0 \) with respect to set inclusion. Then if every element of \( S \) is contained in \( C_0 \), so is \( \cup S \) and so \( \cup S \in T_0 \). If, on the other hand, some chain from \( S \), \( C \), contains \( C_0 \) properly, then since \( C \notin B \), \( f(C_0) \in C \subseteq \cup S \) showing that \( \cup S \notin B \) also. This has proved \( T_0 \) is a tower and since \( T_0 \) is the smallest tower, it follows \( T_0 = T_0 \). This has shown roughly that no splitting into more than one ascending chain can occur at any \( C_0 \) which is comparable to every set of \( T_0 \). Next it is shown that every element of \( T_0 \) has the property that it is comparable to all other elements of \( T_0 \). This is done by showing that these elements which possess this property form a tower.

Define \( T_1 \) to be the set of all elements of \( T_0 \) which are comparable to every element of \( T_0 \). (Recall the elements of \( T_0 \) are chains from the original partial order.)

**Claim:** \( T_1 \) is a tower.

**Proof of the claim:** It is clear that \( \emptyset \in T_1 \) because \( \emptyset \) is a subset of every set. Suppose \( C_0 \in T_1 \). It is necessary to verify that \( g(C_0) \in T_1 \). Let \( D \in T_0 \) (Thus \( D \subseteq C_0 \) or else \( D \supseteq C_0 \)) and consider \( g(C_0) = C_0 \cup \{ f(C_0) \} \). If \( D \subseteq C_0 \), then \( D \subseteq g(C_0) \) so \( g(C_0) \) is comparable to \( D \). If \( D \supseteq C_0 \), then \( D \supseteq g(C_0) \) by what was just shown (\( B = \emptyset \)). Hence \( g(C_0) \) is comparable to \( D \). Since \( D \) was arbitrary, it follows \( g(C_0) \) is comparable to every set of \( T_0 \). Now suppose \( S \) is a chain of elements of \( T_1 \) and let \( D \) be an element of \( T_0 \). If every element in the chain, \( S \), is contained in \( D \), then \( \cup S \) also contains in \( D \). On the other hand, if some set, \( C \), from \( S \) contains \( D \) properly, then \( \cup S \) also contains \( D \). Thus \( \cup S \in T_1 \) since it is comparable to every \( D \in T_0 \).

This shows \( T_1 \) is a tower and proves therefore, that \( T_0 = T_1 \). Thus every set of \( T_0 \) compares with every other set of \( T_0 \) showing \( T_0 \) is a chain in addition to being a tower.

Now \( \cup T_0, g(\cup T_0) \in T_0 \). Hence, because \( g(\cup T_0) \) is an element of \( T_0 \), and \( T_0 \) is a chain of these, it follows \( g(\cup T_0) \subseteq \cup T_0 \). Thus

\[
\cup T_0 \supseteq g(\cup T_0) \supseteq \cup T_0,
\]

a contradiction. Hence there must exist a maximal chain after all. This proves the lemma.

If \( X \) is a nonempty set, \( \leq \) is an order on \( X \) if

\[
x \leq x,
\]

and
APPENDIX D. THE HAUSDORFF MAXIMAL THEOREM

and if \( x, y \in X \), then

\[
\text{either } x \leq y \text{ or } y \leq x
\]

and

\[
\text{if } x \leq y \text{ and } y \leq z \text{ then } x \leq z.
\]

\( \leq \) is a well order and say that \((X, \leq)\) is a well-ordered set if every nonempty subset of \( X \) has a smallest element. More precisely, if \( S \neq \emptyset \) and \( S \subseteq X \) then there exists an \( x \in S \) such that \( x \leq y \) for all \( y \in S \). A familiar example of a well-ordered set is the natural numbers.

**Lemma D.0.5**  The Hausdorff maximal principle implies every nonempty set can be well-ordered.

**Proof:**  Let \( X \) be a nonempty set and let \( a \in X \). Then \( \{a\} \) is a well-ordered subset of \( X \). Let

\[
F = \{ S \subseteq X : \text{ there exists a well order for } S \}.
\]

Thus \( F \neq \emptyset \). For \( S_1, S_2 \in F \), define \( S_1 \prec S_2 \) if \( S_1 \subseteq S_2 \) and there exists a well order for \( S_2 \), \( \leq_2 \) such that

\[
(S_2, \leq_2) \text{ is well-ordered}
\]

and if

\[
y \in S_2 \setminus S_1 \text{ then } x \leq_2 y \text{ for all } x \in S_1,
\]

and if \( \leq_1 \) is the well order of \( S_1 \) then the two orders are consistent on \( S_1 \). Then observe that \( \prec \) is a partial order on \( F \). By the Hausdorff maximal principle, let \( C \) be a maximal chain in \( F \) and let

\[
X_\infty = \cup C.
\]

Define an order, \( \leq_1 \), on \( X_\infty \) as follows. If \( x, y \) are elements of \( X_\infty \), pick \( S \in C \) such that \( x, y \) are both in \( S \). Then if \( \leq_S \) is the order on \( S \), let \( x \leq y \) if and only if \( x \leq_S y \). This definition is well defined because of the definition of the order, \( \prec \). Now let \( U \) be any nonempty subset of \( X_\infty \). Then \( S \cap U \neq \emptyset \) for some \( S \in C \). Because of the definition of \( \leq_1 \), if \( y \in S_2 \setminus S_1 \), \( S_1 \in C \), then \( x \leq y \) for all \( x \in S_1 \). Thus, if \( y \in X_\infty \setminus S \) then \( x \leq y \) for all \( x \in S \) and so the smallest element of \( S \cap U \) exists and is the smallest element in \( U \). Therefore \( X_\infty \) is well-ordered. Now suppose there exists \( z \in X \setminus X_\infty \). Define the following order, \( \leq_1 \), on \( X_\infty \cup \{z\} \).

\[
x \leq_1 y \text{ if and only if } x \leq y \text{ whenever } x, y \in X_\infty
\]

\[
x \leq_1 z \text{ whenever } x \in X_\infty.
\]

Then let

\[
\tilde{C} = \{ S \in C \text{ or } X_\infty \cup \{z\} \}.
\]

Then \( \tilde{C} \) is a strictly larger chain than \( C \) contradicting maximality of \( C \). Thus \( X \setminus X_\infty = \emptyset \) and this shows \( X \) is well-ordered by \( \leq \). This proves the lemma.

With these two lemmas the main result follows.
D.1. THE HAMEL BASIS

Theorem D.0.6 The following are equivalent.

The axiom of choice

The Hausdorff maximal principle

The well-ordering principle.

Proof: It only remains to prove that the well-ordering principle implies the axiom of choice. Let $I$ be a nonempty set and let $X_i$ be a nonempty set for each $i \in I$. Let $X = \cup \{X_i : i \in I\}$ and well order $X$. Let $f(i)$ be the smallest element of $X_i$. Then

$$f \in \prod_{i \in I} X_i.$$ 

D.1 The Hamel Basis

A Hamel basis is nothing more than the correct generalization of the notion of a basis for a finite dimensional vector space to vector spaces which are possibly not of finite dimension.

Definition D.1.1 Let $X$ be a vector space. A Hamel basis is a subset of $X$, $\Lambda$ such that every vector of $X$ can be written as a finite linear combination of vectors of $\Lambda$ and the vectors of $\Lambda$ are linearly independent in the sense that if $\{x_1, \cdots, x_n\} \subseteq \Lambda$ and

$$\sum_{k=1}^{n} c_k x_k = 0$$

then each $c_k = 0$.

The main result is the following theorem.

Theorem D.1.2 Let $X$ be a nonzero vector space. Then it has a Hamel basis.

Proof: Let $x_1 \in X$ and $x_1 \neq 0$. Let $\mathcal{F}$ denote the collection of subsets of $X$, $\Lambda$ containing $x_1$ with the property that the vectors of $\Lambda$ are linearly independent as described in Definition D.1.1 partially ordered by set inclusion. By the Hausdorff maximal theorem, there exists a maximal chain, $\mathcal{C}$ Let $\Lambda = \cup \mathcal{C}$. Since $\mathcal{C}$ is a chain, it follows that if $\{x_1, \cdots, x_n\} \subseteq \mathcal{C}$ then there exists a single $\Lambda' \in \mathcal{C}$ containing all these vectors. Therefore, if

$$\sum_{k=1}^{n} c_k x_k = 0$$

it follows each $c_k = 0$. Thus the vectors of $\Lambda$ are linearly independent. Is every vector of $X$ a finite linear combination of vectors of $\Lambda$?
APPENDIX D. THE HAUSDORFF MAXIMAL THEOREM

Suppose not. Then there exists \( z \) which is not equal to a finite linear combination of vectors of \( \Lambda \). Consider \( \Lambda \cup \{ z \} \). If

\[
 cz + \sum_{k=1}^{m} c_k x_k = 0
\]

where the \( x_k \) are vectors of \( \Lambda \), then if \( c \neq 0 \) this contradicts the condition that \( z \) is not a finite linear combination of vectors of \( \Lambda \). Therefore, \( c = 0 \) and now all the \( c_k \) must equal zero because it was just shown \( \Lambda \) is linearly independent. It follows \( \mathcal{C} \cup \{ \Lambda \cup \{ z \} \} \) is a strictly larger chain than \( \mathcal{C} \) and this is a contradiction. Therefore, \( \Lambda \) is a Hamel basis as claimed. This proves the theorem.

D.2 Exercises

1. Zorn’s lemma states that in a nonempty partially ordered set, if every chain has an upper bound, there exists a maximal element, \( x \) in the partially ordered set. \( x \) is maximal, means that if \( x \prec y \), it follows \( y = x \). Show Zorn’s lemma is equivalent to the Hausdorff maximal theorem.

2. Show that if \( Y, Y_1 \) are two Hamel bases of \( X \), then there exists a one to one and onto map from \( Y \) to \( Y_1 \). Thus any two Hamel bases are of the same size.

3. Using the Baire category theorem of the chapter on Banach spaces show that any Hamel basis of a Banach space is either finite or uncountable.

4. Consider the vector space of all polynomials defined on \([0,1]\). Does there exist a norm, \(|\cdot|\) defined on these polynomials such that with this norm, the vector space of polynomials becomes a Banach space (complete normed vector space)?
Bibliography


BIBLIOGRAPHY


[56] **Smart D.R.** *Fixed point theorems* Cambridge University Press, 1974.


Index

$C^k$, 141
$G_\delta$, 177
$L^p (\Omega; X)$, 143
$\langle \cdot, \cdot \rangle$, 52
$G$, 154
$L(X, Y)$, 19

adjoint map, 107
Alexander subbasis theorem, 111
algebra, 101
analytic function, 155
analytic sets, 155
application of Bair theorem, 157
apriori estimates, 119
Ascoli Arzela theorem, 106, 111

Banach Alaoglu theorem, 101
Banach space, 101
Banach Steinhaus theorem, 147
basis of a topology, 91
Bessel’s inequality, 145
bifurcation point, 241
Bochner integrable, 108
Borsuk Ulam theorem, 258
bounded continuous linear functions, 105
box topology, 81
Brouwer fixed point theorem, 91, 241, 193
Browder lemma, 141

Cantor diagonalization procedure, 114
Cariste fixed point theorem, 157
Cauchy Schwarz inequality, 145
chain rule, 141
closed graph theorem, 151
closed range, 161
closed set, 152
closure of a set, 152
compact, 111
compact map
finite dimensional approximation, 141
compact operator, 140
compact set, 111
conjugate function, 108, 109
connected, 57
connected components, 57
connected set
continuous image, 90
continuity
uniform, 85
continuous function, 155
contraction map, 241
fixed point, 119
fixed point theorem, 161
high power, 79
convex, 497
set, 119
convex hull, 497
convex sets, 79
countable basis, 152
demicontinuous, 108, 109
derivative
chain rule, 141
continuous, 141
continuous Gateaux derivatives, 152
Frechet, 152
Gateaux, 152, 154
generalized partial, 145
higher order, 150, 151
matrix, 152
partial, 152
second, 152
well defined, 139
derivatives, 139
differentiable, 139
continuous, 140
continuous partials, 153
derivatives, 139
differential equations, 176
global existence, 194
Dini derivatives, 281
directional derivative, 145
dominated convergence theorem, 115
dual space, 48
duality map
strictly monotone, 304

demicontinuous, 303
duality mapping, 301
duality maps
application of duality, 301
properties, 303

Eberlein Smulian theorem, 73
Ekeland variational principle, 42
epsilon net, 11
equality of mixed partial derivatives, 156
equicontinuous, 21
extending off closed set, 90
extension
mapping, 90
extreme value theorem, 13
finite intersection property, 15
fixed point property, 77
fixed point theorem
Cariste, 40
Kakutani, 104
Frechet derivative, 159
Fredholm operator
Banach space, 159

Gateaux derivative, 143

gauge function, 57
generalized gradient
pseudomonotone, 321
generalized gradients, 319
Gram determinant, 483
Gram matrix, 483
Hahn Banach theorem, 52
Hamel basis, 49
Hausdorff maximal principle, 111
Hausdorff maximal theorem, 111
Hausdorff metric, 57
Hausdorff space, 41
Heine Borel theorem, 1
hemiconcave, 49
hemiconvex, 49
higher order derivative
multilinear form, 161
higher order derivatives, 161
implicit function theorem, 161
inverse function theorem, 188
Hilbert space, 76
Holder space
compact embedding, 72
not separable, 1

Holder spaces, 77
homotopic, 249
homotopy, 249
implicit function theorem, 161
higher order derivatives, 165
increasing function
differentiability, 151
inner product space, 176
invariance of domain, 176
inverse function theorem, 163
higher order derivatives, 167
invertible maps, 158
different spaces, 159
isometric, 158
Jordan curve theorem, 257
Jordan Separation theorem, 262
Jordan separation theorem, 264
Kakutani fixed point theorem, 199
Leray Schauder alternative, 194
Leray Schauder degree, 277
properties, 278
limit
continuity, 138
infinite limits, 136
limit of a function, 135
INDEX

limit point, 32, 135
limits
combinations of functions, 136
limits and continuity, 138
Lindeloff property, 10, 34
Lipschitz
functions, 381
Lipschitz continuous, 24
Lipschitz function
integral of its derivative, 384
little o notation, 139
locally bounded, 6
locally compact, 34
locally compact, 32
locally finite, 8
Lyapunov Schmidt procedure, 165
maximal monotone
approximation, 344
onto, 339
pseudomonotone, 294
maximal monotone
Banach space, 322
equivalent conditions, 330
interior of domain, 334
maximal monotone operator, 322
mean value inequality, 144
mean value theorem, 144, 160
measurable
multifunctions, 442
measurable representative, 118
measurable selection, 443
measure space
isomorphic, 438
separable, 438
metric space, 7
closed sets, 8
balls, 8
compact, 11
convergence, 8
open sets, 8
product, 15
min max theorem, 225
Minkowski functional, 55
mixed partial derivatives, 155
mollifier, 80
monotone, 293
monotone operators, 322
Banach space, 322
monotone and hemicontinuous coercive, 326
monotone graph extended, 326
monotone operators, 322
Moreau's theorem, 363
mountain pass theorem
Banach space, 222
Hilbert space, 211
multi-index, 71
multifunctions, 442
measurability, 442
Neuman series, 157, 158
non equal mixed partials
example, 156
normal topological space, 34
not pseudomonotone, 294
one point compactification, 34
open cover, 34
open mapping theorem, 458
open sets, 31
operator norm, 456
operators
closed range, 464
orthonormal set, 484
pap97, 34
paracompact space, 83
partial derivatives, 142, 153
continuous, 153
partial order, 40
partially ordered set, 487
partition of unity, 87
metric space, 87, 89
penalization operators, 308
pointwise compact, 21
pointwise convergence, 49
pointwise limits of measurable functions, 103
polar, 445
Polish space, 48
polish space, 72
polynomial, 41
precompact, 46, 63
product formula, 241
product topology, 46, 63
projection in Hilbert space, 278
pseudogradient, 212
pseudomonotne
 bounded, 314
pseudomonotne, 292
  generalized perturbation, 258
  modified bounded, 257
  perturbation, 277
  sum, 277
pseudomonotone
 onto, 47
pseudomonotone operator, 314, 315, 335
set valued, 414
  measurability, 442
  locally bounded, 334
  simple function, 44
  singular values, 424
spectral theory
  Banach space, 90
  compact operator, 110
Stone’s theorem, 57
  strictly convex, 61
  strongly measurable, 424
subbasis, 36
subgradients
  sum, 359
sum of maximal monotone and pseudomonotone, 484

quasi-bounded, 424

Rademacher’s theorem, 483
Radon Nikodym property, 484
  refinement of a cover, 83
  regular topological space, 84
  regular values, 484
  relative topology, 47
  retract
    Banach space, 79
    closed and convex set, 79
  Riesz map, 181
  Riesz representation theorem, 125
    Hilbert space, 181
  Riesz representation theorem $L^p$
    $\sigma$ finite case, 128

Sard’s lemma, 441
Sard’s theorem, 441
Schaefer fixed point
  differential equations, 189
Schaefer fixed point theorem, 189
Schauder fixed point
  approximate fixed point, 198
Schauder fixed point theorem, 198
  degree theory, 235
  second derivative, 63
sections of open sets, 194
separable metric space, 48
  separated, 48
  sequentially compact, 48
  set valued functions
    measurability, 48
  set valued map
    locally bounded, 48
simple function, 63
singular values, 424

Tietze extension theorem, 60
  totally bounded set, 61
  totally ordered set, 451
  type M, 475
    demicontinuous, 498

uniform boundedness theorem, 487
uniform continuity, 14, 28
uniform continuity and compactness, 28
uniform contraction principle, 26
uniform convergence, 28
uniform convergence and continuity, 28
uniformly bounded, 19
uniformly continuous, 28
uniformly equicontinuous, 19
uniqueness of limits, 135
upper semicontinuity
  set valued map, 194
upper semicontinuous composition, 194

Suslin space, 433

Type M, 475
INDEX

variational inequality, 478
vector measures, 118
vector valued function
  limit theorems, 136
Vitali covering, 449
  closed balls, 451
Vitali covering theorem, 450, 451
weak* measurable, 106
weakly measurable, 99
Weierstrass
  Stone Weierstrass theorem, 72
well ordered sets, 490
Yankov von Neumann Aumann, 439, 440