Pages containing errors in Modern Analysis

Kenneth Kuttler

Chapter 1
Set Theory and General Topology

Chapter 2
Compactness and Continuous Functions

Chapter 3

Banach Spaces

3.1 Baire category theorem

3.2 Uniform boundedness closed graph and open mapping theorems

3.3 Hahn Banach theorem

3.4 Exercises

1. first problem
2. second
3. third
4. fourth
5. fifth
6. sixth
7. seventh
8. number eight

$\sum_{k=-n}^{n} a_k e^{ikx}$. Show $S_n f(x) = \int_{-\pi}^{\pi} f(y) D_n(x-y)dy$ where

$$D_n(t) = \frac{\sin((n+\frac{1}{2})t)}{2\pi \sin(\frac{t}{2})}.$$

This is called the Dirichlet kernel.

9. ↑ Let $Y = \{f$ such that f is continuous, defined on $\mathbb{R}$, and 2π periodic$\}$. Define $||f||_Y = \sup\{|f(x)| : x \in [-\pi, \pi]\}$. Show that $(Y, ||\ ||_Y)$ is a Banach space. Let $x \in \mathbb{R}$ and define $L_n(f) = S_n f(x)$. Show $L_n \in Y'$ but $\lim_{n\to\infty} ||L_n|| = \infty$. **Hint:** Let $f(y)$ approximate $sign(D_n(x-y))$.
10. ↑ Show there exists a dense G_δ subset of Y such that for f in this set, $|S_n f(x)|$ is unbounded. Show there is a dense G_δ subset of Y having the property that $|S_n f(x)|$ is unbounded on a dense G_δ subset of $\mathbb{R}$. This shows Fourier series can fail to converge pointwise to continuous periodic functions in a fairly spectacular way.
11. Let X be a normed linear space and let M be a convex open set containing 0. Define

$$\rho(x) = \inf\{t > 0 : \frac{x}{t} \in M\}.$$

Show ρ is a gauge function defined on X. This particular example is called a Minkowski functional. Recall a set, M, is convex if $\lambda x + (1-\lambda)y \in M$ whenever $\lambda \in [0,1]$ and $x, y \in M$.

12. ↑ This problem explores the use of the Hahn Banach theorem in establishing separation theorems. Let M be an open convex set containing 0. Let $x \notin M$. Show there exists $x^* \in X'$ such that $\operatorname{Re} x^*(x) \geq 1 > \operatorname{Re} x^*(y)$ for all $y \in M$. **Hint:** If $y \in M, \rho(y) < 1$. Show this. If $x \notin M,\ \rho(x) \geq 1$. Try $f(\alpha x) = \alpha\rho(x)$ for $\alpha \in \mathbb{R}$. Then extend f to F, show F is continuous, then fix it so F is the real part of $x^* \in X'$.
13. A Banach space is said to be strictly convex if whenever $||x|| = ||y||$ and $x \neq y$, then

$$\left|\left|\frac{x+y}{2}\right|\right| < ||x||.$$

$F : X \to X'$ is said to be a duality map if it satisfies the following: a.) $||F(x)|| = ||x||$. b.) $F(x)(x) = ||x||^2$. Show that if X' is strictly convex, then such a duality map exists. **Hint:** Let $f(\alpha x) = \alpha ||x||^2$ and use Hahn Banach theorem, then strict convexity.

Chapter 4

Hilbert Spaces

Chapter 5
Calculus in Banach Space

Chapter 6

Locally Convex Topological Vector Space

6.1 Separation theorems

6.2 The weak and weak$*$ topologies

6.3 Set-valued maps

vertex, $\mathbf{x}$, pick $A_\epsilon \mathbf{x} \in A\mathbf{x}$ and define A_ϵ on all of $\mathbb{C}^n$ by the following rule. If

$$\mathbf{x} \in [\mathbf{x}_0, \cdots, \mathbf{x}_{2n}],$$

so $\mathbf{x} = \sum_{i=0}^{2n} t_i \mathbf{x}_i$, then

$$A_\epsilon \mathbf{x} \equiv \sum_{k=0}^{2n} t_k A_\epsilon \mathbf{x}_k.$$

Thus A_ϵ is a continuous map defined on $\mathbb{C}^n$ thanks to the local finiteness of the collection of simplicies. Let P_K denote the projection on the convex set K. By the Brouwer fixed point theorem, there exists a fixed point, $\mathbf{x}_\epsilon \in K$ such that

$$P_K (\mathbf{y} - A_\epsilon \mathbf{x}_\epsilon + \mathbf{x}_\epsilon) = \mathbf{x}_\epsilon.$$

By Corollary 4.8 this requires

$$\operatorname{Re} (\mathbf{y} - A_\epsilon \mathbf{x}_\epsilon, \mathbf{z} - \mathbf{x}_\epsilon) \leq 0$$

for all $\mathbf{z} \in K$.

Suppose $\mathbf{x}_\epsilon \in [\mathbf{x}_0^\epsilon, \cdots, \mathbf{x}_{2n}^\epsilon]$ so $\mathbf{x}_\epsilon = \sum_{k=0}^{2n} t_k^\epsilon \mathbf{x}_k^\epsilon$. Then since $\mathbf{x}_\epsilon$ is contained in K, a compact set, and the diameter of each simplex is less than 1, it follows that $A_\epsilon \mathbf{x}_k^\epsilon$ is contained in $A(\overline{K + B(\mathbf{0},1)})$, which is contained in a compact set thanks to Lemma 6.29. Taking a subsequence, we may obtain from the Heine Borel theorem that for some sequence, $\epsilon \to 0$

$$t_k^\epsilon \to t_k, \mathbf{x}_\epsilon \to \mathbf{x}, A_\epsilon \mathbf{x}_k^\epsilon \to \mathbf{y}_k$$

for $k = 0, \cdots, 2n$. Since the diameter of the simplex containing $\mathbf{x}_\epsilon$ converges to 0, it follows

$$\mathbf{x}_k^\epsilon \to \mathbf{x},\ A_\epsilon \mathbf{x}_k^\epsilon \to \mathbf{y}_k.$$

Since the graph of A is closed and $A_\epsilon \mathbf{x}_k^\epsilon \in A\mathbf{x}_k^\epsilon$, this implies $\mathbf{y}_k \in A\mathbf{x}$. Since $A\mathbf{x}$ is convex,

$$\sum_{k=1}^{2n} t_k \mathbf{y}_k \in A\mathbf{x}.$$

Hence for all $\mathbf{z} \in K$,

$$\operatorname{Re}\left(\mathbf{y} - \sum_{k=1}^{2n} t_k \mathbf{y}_k, \mathbf{z} - \mathbf{x}\right) = \lim_{\epsilon \to 0} \operatorname{Re}\left(\mathbf{y} - \sum_{k=1}^{2n} t_k^\epsilon A_\epsilon \mathbf{x}_k^\epsilon, \mathbf{z} - \mathbf{x}_\epsilon\right)$$
$$= \lim_{\epsilon \to 0} \operatorname{Re} (\mathbf{y} - A_\epsilon \mathbf{x}_\epsilon, \mathbf{z} - \mathbf{x}_\epsilon) \leq 0.$$

Let $\mathbf{w} = \sum_{k=1}^{2n} t_k \mathbf{y}_k$. This proves the lemma.

Chapter 7

Measures and Measurable Functions

7.1 σ Algebras

7.2 Monotone classes and algebras

Corollary 7.8 *Let* $(Z_1, \mathcal{R}_1, \mathcal{E}_1)$ *and* $(Z_2, \mathcal{R}_2, \mathcal{E}_2)$ *be as just described in Lemma 7.6. Then* $(Z_1 \times Z_2, \mathcal{R}, \mathcal{E})$ *also satisfies the conditions of Lemma 7.6 if* $\mathcal{R}$ *is defined as*

$$\mathcal{R} \equiv \{R_1 \times R_2 : R_i \in \mathcal{R}_i\}$$

and

$$\mathcal{E} \equiv \{ \textit{ finite disjoint unions of sets of } \mathcal{R}\}.$$

Consequently, $\mathcal{E}$ *is an algebra of sets.*

Proof: It is clear $\emptyset, Z_1 \times Z_2 \in \mathcal{R}$. Let $R_1^1 \times R_2^1$ and $R_1^2 \times R_2^2$ be two elements of $\mathcal{R}$.

$$R_1^1 \times R_2^1 \cap R_1^2 \times R_2^2 = R_1^1 \cap R_1^2 \times R_2^1 \cap R_2^2 \in \mathcal{R}$$

by assumption.

$$\begin{gathered} R_1^1 \times R_2^1 \setminus \left(R_1^2 \times R_2^2\right) = \\ R_1^1 \times \left(R_2^1 \setminus R_2^2\right) \cup \left(R_1^1 \setminus R_1^2\right) \times \left(R_2^1 \cap R_2^2\right) \\ = R_1^1 \times A_2 \cup A_1 \times R_2 \end{gathered}$$

where $A_2 \in \mathcal{E}_2$, $A_1 \in \mathcal{E}_1$, and $R_2 \in \mathcal{R}_2$. Since the two sets in the above expression on the right do not intersect, and each A_i is a finite union of disjoint elements of $\mathcal{R}_i$, it follows the above expression is in $\mathcal{E}$. This proves the corollary. The following example will be referred to frequently.

Example 7.9 *Consider for* $\mathcal{R}$*, sets of the form* $I = (a, b] \cap (-\infty, \infty)$ *where* $a \in [-\infty, \infty]$ *and* $b \in [-\infty, \infty]$*. Then, clearly,* $\emptyset, (-\infty, \infty) \in \mathcal{R}$ *and it is not hard to see that all conditions for Corollary 7.7 are satisfied. Applying this corollary repeatedly, we find that for*

$$\mathcal{R} \equiv \left\{ \prod_{i=1}^{n} I_i : I_i = (a_i, b_i] \cap (-\infty, \infty) \right\}$$

and $\mathcal{E}$ *is defined as finite disjoint unions of sets of* $\mathcal{R}$*,*

$$(\mathbb{R}^n, \mathcal{R}, \mathcal{E})$$

satisfies the conditions of Corollary 7.7 and in particular $\mathcal{E}$ *is an algebra of sets of* $\mathbb{R}^n$*. It is clear that the same would hold if* I *were of the form* $[a, b) \cap (-\infty, \infty)$*.*

Chapter 8
The Abstract Lebesgue Integral

Chapter 9

The Construction of Measures

9.1 Outer measures

But we know Formula (2) holds because A is measurable. Apply the Definition 9.1 to $S \cap T$ instead of S.

The next theorem is the main result on outer measures. It is a very general result which applies whenever one has an outer measure on the power set of any set. This theorem will be referred to as Caratheodory's procedure in the rest of the book.

Theorem 9.4 *The collection of μ measurable sets, $\mathcal{S}$, forms a σ algebra and*

$$\text{If } F_i \in \mathcal{S},\ F_i \cap F_j = \emptyset, \text{ then } \mu(\cup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} \mu(F_i). \tag{3}$$

If $\cdots F_n \subseteq F_{n+1} \subseteq \cdots$, then if $F = \cup_{n=1}^{\infty} F_n$ and $F_n \in \mathcal{S}$, it follows that

$$\mu(F) = \lim_{n\to\infty} \mu(F_n). \tag{4}$$

If $\cdots F_n \supseteq F_{n+1} \supseteq \cdots$, and if $F = \cap_{n=1}^{\infty} F_n$ for $F_n \in \mathcal{S}$ then if $\mu(F_1) < \infty$, we may conclude that

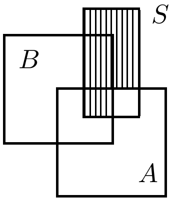
$$\mu(F) = \lim_{n\to\infty} \mu(F_n). \tag{5}$$

Also, $(\mathcal{S}, \mu)$ is complete. By this we mean that if $F \in \mathcal{S}$ and if $E \subseteq \Omega$ with $\mu(E \setminus F) + \mu(F \setminus E) = 0$, then $E \in \mathcal{S}$.

Proof: First note that $\emptyset$ and Ω are obviously in $\mathcal{S}$. Now suppose that $A, B \in \mathcal{S}$. We show $A \setminus B = A \cap B^C$ is in $\mathcal{S}$. Using the assumption that $B \in \mathcal{S}$ in the second equation below, in which $S \cap A$ plays the role of S in the definition for B being μ measurable,

$$\begin{aligned}\mu(S \cap (A \cap B^C)) + \mu(S \setminus (A \cap B^C)) &= \mu(S \cap A \cap B^C) + \mu(S \cap (A^C \cup B)) \\ &= \mu(S \cap (A^C \cup B)) + \mu(S \cap A) - \mu(S \cap A \cap B).\end{aligned} \tag{6}$$

The following picture of $S \cap (A^C \cup B)$ may be of use.



From the picture, and the measurablility of A, we see that Formula (6) is no larger than

$$\leq \mu(S \cap A \cap B) + \mu(S \setminus A) + \mu(S \cap A) - \mu(S \cap A \cap B) = \mu(S).$$

This has shown that if $A, B \in \mathcal{S}$, then $A \setminus B \in \mathcal{S}$. Since $\Omega \in \mathcal{S}$, this shows that $A \in \mathcal{S}$ if and only if $A^C \in \mathcal{S}$. Now if $A, B \in \mathcal{S}$, $A \cup B = (A^C \cap B^C)^C = (A^C \setminus B)^C \in \mathcal{S}$. By induction, if $A_1, \cdots, A_n \in \mathcal{S}$, then so is $\cup_{i=1}^{n} A_i$. If $A, B \in \mathcal{S}$, with $A \cap B = \emptyset$,

$$\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B).$$

By induction, if $A_i \cap A_j = \emptyset$ and $A_i \in \mathcal{S}$, $\mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$.

Now let $A = \cup_{i=1}^\infty A_i$ where $A_i \cap A_j = \emptyset$ for $i \neq j$.

$$\sum_{i=1}^{\infty} \mu(A_i) \geq \mu(A) \geq \mu(\cup_{i=1}^n A_i) = \sum_{i=1}^{n} \mu(A_i).$$

Since this holds for all n, we conclude, since μ is assumed to be an outer measure, that $\mu(A) = \sum_{i=1}^\infty \mu(A_i)$ which establishes Formula (3). Part (4) follows from part (3) just as in the proof of Theorem 7.12.

In order to establish (5), let the F_n be as given there. Then, since $(F_1 \setminus F_n)$ increases to $(F_1 \setminus F)$, we may use part (4) to conclude

$$\lim_{n\to\infty} (\mu(F_1) - \mu(F_n)) = \mu(F_1 \setminus F).$$

Now $\mu(F_1 \setminus F) + \mu(F) \geq \mu(F_1)$ and so $\mu(F_1 \setminus F) \geq \mu(F_1) - \mu(F)$. Hence

$$\lim_{n\to\infty} (\mu(F_1) - \mu(F_n)) \geq \mu(F_1) - \mu(F)$$

which implies that, since $F \subseteq F_n$ for all n,

$$\mu(F) \leq \lim_{n\to\infty} \mu(F_n) \leq \mu(F).$$

It remains to show $\mathcal{S}$ is closed under countable unions. We already know that if $A \in \mathcal{S}$, then $A^C \in \mathcal{S}$ and $\mathcal{S}$ is closed under finite unions. Let $A_i \in \mathcal{S}$, $A = \cup_{i=1}^\infty A_i$, $B_n = \cup_{i=1}^n A_i$. Then

$$\begin{aligned} \mu(S) &= \mu(S \cap B_n) + \mu(S \setminus B_n) \qquad (7) \\ &= (\mu\lfloor S)(B_n) + (\mu\lfloor S)(B_n^C). \end{aligned}$$

By Lemma 9.3 we know B_n is $(\mu\lfloor S)$ measurable and so is B_n^C. We want to show $\mu(S) \geq \mu(S \setminus A) + \mu(S \cap A)$. If $\mu(S) = \infty$, there is nothing to prove. Assume $\mu(S) < \infty$. Then we apply Parts (5) and (4) to Formula (7) and let $n \to \infty$. Thus

$$B_n \uparrow A,\ B_n^C \downarrow A^C$$

and this yields $\mu(S) = (\mu\lfloor S)(A) + (\mu\lfloor S)(A^C) = \mu(S \cap A) + \mu(S \setminus A)$.

Thus $A \in \mathcal{S}$ and this proves Parts (3), (4), and (5).

Chapter 10

Lebesgue Measure

10.1 Lebesgue Measure

10.2 Iterated integrals

10.3 Change of variables

10.4 Polar coordinates

Proof: Let E be an open set in X and let

$$\mathcal{S}_E \equiv \{F \text{ Borel in } Y \text{ such that } E \times F \text{ is Borel in } X \times Y\}.$$

Then $\mathcal{S}_E$ contains the open sets and is clearly closed with respect to countable unions. Let $F \in \mathcal{S}_E$. Then

$$E \times F^C \cup E \times F = E \times Y = \text{ a Borel set.}$$

Therefore, since $E \times F$ is Borel, it follows $E \times F^C$ is Borel. Therefore, $\mathcal{S}_E$ is a σ algebra. It follows $\mathcal{S}_E =$ Borel sets, and so, we have shown- open $\times$ Borel $=$ Borel. Now let F be a fixed Borel set in Y and define

$$\mathcal{S}_F \equiv \{E \text{ Borel in } X \text{ such that } E \times F \text{ is Borel in } X \times Y\}.$$

The same argument which was just used shows $\mathcal{S}_F$ is a σ algebra containing the open sets. Therefore, $\mathcal{S}_F =$ the Borel sets, and this proves the lemma since F was an arbitrary Borel set.

Now we define the unit sphere in $\mathbb{R}^n$, S^{n-1}, by

$$S^{n-1} \equiv \{\mathbf{w} \in \mathbb{R}^n : |\mathbf{w}| = 1\}.$$

Then S^{n-1} is a compact metric space using the usual metric on $\mathbb{R}^n$. We define a map

$$\theta : S^{n-1} \times (0, \infty) \to \mathbb{R}^n \setminus \{\mathbf{0}\}$$

by

$$\theta(\mathbf{w}, \rho) \equiv \rho \mathbf{w}.$$

It is clear that θ is one to one and onto with a continuous inverse. Therefore, if $\mathcal{B}_1$ is the set of Borel sets in $S^{n-1} \times (0, \infty)$, and $\mathcal{B}$ are the Borel sets in $\mathbb{R}^n \setminus \{\mathbf{0}\}$, it follows

$$\mathcal{B} = \{\theta(F) : F \in \mathcal{B}_1\}. \tag{8}$$

Observe also that the Borel sets of S^{n-1} satisfy the conditions of Lemma 7.6 with Z defined as $S^{n-1}$and the same is true of the sets $(a, b] \cap (0, \infty)$ where $0 \leq a, b \leq \infty$ if Z is defined as $(0, \infty)$. By Corollary 7.7, finite disjoint unions of sets of the form

$$\left\{E \times I : E \text{ is Borel in } S^{n-1}\right.$$
$$\left.\text{and } I = (a, b] \cap (0, \infty) \text{ where } 0 \leq a, b \leq \infty\right\}$$

form an algebra of sets, $\mathcal{A}$. It is also clear that $\sigma(\mathcal{A})$ contains the open sets and so $\sigma(\mathcal{A}) = \mathcal{B}_1$ because every set in $\mathcal{A}$ is in $\mathcal{B}_1$ thanks to Lemma 10.18. Let $A_r \equiv S^{n-1} \times (0, r]$ and let

$$\mathcal{M} \equiv \left\{F \in \mathcal{B}_1 : \int_{\mathbb{R}^n} \mathcal{X}_{\theta(F \cap A_r)} dm_n\right.$$

Chapter 11
Product Measure

11.1 Product Measure

Let $K_x \subseteq (P \cap X_n)$ and $K_y \subseteq (Q \cap Y_n)$ be such that

$$\mu(K_x) + \varepsilon > \mu(P \cap X_n)$$

and

$$\lambda(K_y) + \varepsilon > \lambda(Q \cap Y_n).$$

By Theorem 1.36 $K_x \times K_y$ is compact and from the definition of product measure,

$$\begin{gathered}(\mu \times \lambda)(K_x \times K_y) = \mu(K_x)\lambda(K_y) \\ \geq \mu(P \cap X_n)\lambda(Q \cap Y_n) - \varepsilon(\lambda(Q \cap Y_n) + \mu(P \cap X_n)) + \varepsilon^2.\end{gathered}$$

Since ε is arbitrary, this verifies that $(\mu \times \lambda)$ is inner regular on $S \cap R_n$ whenever S is an elementary set. Similarly, $(\mu \times \lambda)$ is outer regular on $S \cap R_n$ whenever S is an elementary set. Thus $\mathcal{G}_n$ contains the elementary sets.

Next we show that $\mathcal{G}_n$ is a monotone class. If $S_k \downarrow S$ and $S_k \in \mathcal{G}_n$, let K_k be a compact subset of $S_k \cap R_n$ with

$$(\mu \times \lambda)(K_k) + \varepsilon 2^{-k} > (\mu \times \lambda)(S_k \cap R_n).$$

Let $K = \cap_{k=1}^{\infty} K_k$. Then

$$S \cap R_n \setminus K \subseteq \cup_{k=1}^{\infty}(S_k \cap R_n \setminus K_k).$$

Therefore

$$\begin{aligned}(\mu \times \lambda)(S \cap R_n \setminus K) &\leq \sum_{k=1}^{\infty}(\mu \times \lambda)(S_k \cap R_n \setminus K_k) \\ &\leq \sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon.\end{aligned}$$

Now let $V_k \supseteq S_k \cap R_n$, V_k is open and

$$(\mu \times \lambda)(S_k \cap R_n) + \varepsilon > (\mu \times \lambda)(V_k).$$

Let k be large enough that

$$(\mu \times \lambda)(S_k \cap R_n) - \varepsilon < (\mu \times \lambda)(S \cap R_n).$$

Then $(\mu \times \lambda)(S \cap R_n) + 2\varepsilon > (\mu \times \lambda)(V_k)$. This shows $\mathcal{G}_n$ is closed with respect to intersections of decreasing sequences of its elements. The consideration of increasing sequences is similar. By the monotone class theorem, $\mathcal{G}_n = \mathcal{S} \times \mathcal{F}$.

Now let $S \in \mathcal{S} \times \mathcal{F}$ and let $l < (\mu \times \lambda)(S)$. Then $l < (\mu \times \lambda)(S \cap R_n)$ for some n. It follows from the first part of this proof that there exists a

Chapter 12

The L^p spaces

Chapter 13

Representation Theorems

Chapter 14

Fundamental Theorem of Calculus

Chapter 15
General Radon Measures

Chapter 16
Fourier Transforms

Chapter 17
Probability

Chapter 18
Weak Derivatives

Chapter 19

Hausdorff Measures

$$l(\mathbf{x}_1) - g(\mathbf{x}_1) \geq m(A_{P_i\mathbf{x}_1}) \geq 2|y_1|, \tag{3}$$

$$l(\mathbf{x}_2) - g(\mathbf{x}_2) \geq m(A_{P_i\mathbf{x}_2}) \geq 2|y_2|. \tag{4}$$

Claim: $|y_1 - y_2| \leq |l(\mathbf{x}_1) - g(\mathbf{x}_2)|$ or $|y_1 - y_2| \leq |l(\mathbf{x}_2) - g(\mathbf{x}_1)|$.

Proof of Claim: If not,

$$\begin{gathered} 2|y_1 - y_2| > |l(\mathbf{x}_1) - g(\mathbf{x}_2)| + |l(\mathbf{x}_2) - g(\mathbf{x}_1)| \\ \geq |l(\mathbf{x}_1) - g(\mathbf{x}_1) + l(\mathbf{x}_2) - g(\mathbf{x}_2)| \\ = l(\mathbf{x}_1) - g(\mathbf{x}_1) + l(\mathbf{x}_2) - g(\mathbf{x}_2). \\ \geq 2\,|y_1| + 2\,|y_2| \end{gathered}$$

by (3) and (4) contradicting the triangle inequality.

Now suppose $|y_1 - y_2| \leq |l(\mathbf{x}_1) - g(\mathbf{x}_2)|$. From the claim,

$$\begin{aligned} |\mathbf{x}_1 - \mathbf{x}_2| &= (|P_i\mathbf{x}_1 - P_i\mathbf{x}_2|^2 + |y_1 - y_2|^2)^{1/2} \\ &\leq (|P_i\mathbf{x}_1 - P_i\mathbf{x}_2|^2 + |l(\mathbf{x}_1) - g(\mathbf{x}_2)|^2)^{1/2} \\ &\leq (|P_i\mathbf{x}_1 - P_i\mathbf{x}_2|^2 + (|z_1 - z_2| + 2\varepsilon)^2)^{1/2} \\ &\leq diam(A) + O(\sqrt{\varepsilon}) \end{aligned}$$

where z_1 and z_2 are such that $P_i\mathbf{x}_1 + z_1\mathbf{e}_i \in A$, $P_i\mathbf{x}_2 + z_2\mathbf{e}_i \in A$, and

$$|z_1 - l(\mathbf{x}_1)| < \varepsilon \text{ and } |z_2 - g(\mathbf{x}_2)| < \varepsilon.$$

If $|y_1 - y_2| \leq |l(\mathbf{x}_2) - g(\mathbf{x}_1)|$, then we use the same argument but let

$$|z_1 - g(\mathbf{x}_1)| < \varepsilon \text{ and } |z_2 - l(\mathbf{x}_2)| < \varepsilon,$$

Since $\mathbf{x}_1, \mathbf{x}_2$ are arbitrary elements of $S(A, \mathbf{e}_i)$ and ε is arbitrary, this proves Formula 2.

The next lemma says that if A is already symmetric with respect to the jth direction, then this symmetry is not destroyed by taking $S\,(A, \mathbf{e}_i)$.

Lemma 19.5 *Suppose A is a Borel set in $\mathbb{R}^n$ such that $P_j\mathbf{x} + \mathbf{e}_j x_j \in A$ if and only if $P_j\mathbf{x} + (-x_j)\mathbf{e}_j \in A$. Then if $i \neq j$, $P_j\mathbf{x} + \mathbf{e}_j x_j \in S(A, \mathbf{e}_i)$ if and only if $P_j\mathbf{x} + (-x_j)\mathbf{e}_j \in S(A, \mathbf{e}_i)$.*

Chapter 20
The Area Formula

Chapter 21

The Coarea Formula

Chapter 22

Fourier Analysis in $\mathbb{R}^n$

Chapter 23
Integration for Vector Valued Functions

Chapter 24
Convex Functions

24.1 Continuity properties of convex functions

24.2 Separation properties

24.3 Conjugate functions

24.4 Subgradients

for all $z \in X$. Therefore,

$$\phi^*(y^*) \leq y^*(x) - \phi(x) \leq \phi^*(y^*).$$

Hence

$$y^*(x) = \phi^*(y^*) + \phi(x). \tag{7}$$

Now if $z^* \in X'$ is arbitrary, Formula 7 shows

$$(z^* - y^*)(x) = z^*(x) - \phi^*(y^*) - \phi(x) \leq \phi^*(z^*) - \phi^*(y^*)$$

and this shows $x \in \delta\phi^*(y^*)$.

Now suppose $x \in \delta\phi^*(y^*)$. Then for $z^* \in X'$,

$$(z^* - y^*)(x) \leq \phi^*(z^*) - \phi^*(y^*)$$

and so, taking $\sup$ over all z^*, and using Theorem 24.17,

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

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 24.22 *If X is a Banach space, $u \in H^1(0,T;X)$ if there exists $g \in L^2(0,T;X)$ such that*

$$u(t) = u(0) + \int_0^t g(s)\,ds$$

and we 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. We leave its proof as an interesting exercise for the reader.

Lemma 24.23 *Suppose $g \in L^2(0,T;X)$. Then*

$$\int_{(\cdot)}^{(\cdot)+h} g(s)\,ds\mathcal{X}_{[0,T-h]}(\cdot) \to g$$

in $L^2(0,T;X)$.

The following theorem is a form of the chain rule in which the derivative is replaced by the subgradient.

Theorem 24.24 *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, $t \rightarrow \phi(u(t))$ 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 $\widetilde{u}$ and $\widetilde{z}$ the modified functions, $\widetilde{z}(t) \in \delta\phi(\widetilde{u}(t))$ for all t. First we show $t \rightarrow \phi(\widetilde{u}(t))$ is in $L^1(0,T)$. Pick $t_0 \in [0,T]$ and let $\widetilde{z}(t_0) \in \delta\phi(\widetilde{u}(t_0))$. Then for $t \in [0,T]$,

$$\widetilde{z}(t_0)(\widetilde{u}(t) - \widetilde{u}(t_0)) + \phi(\widetilde{u}(t_0)) \leq \phi(\widetilde{u}(t)) \leq \widetilde{z}(t)(\widetilde{u}(t) - \widetilde{u}(t_0)) + \phi(\widetilde{u}(t_0))$$

since $\widetilde{z}(t)(\widetilde{u}(t_0) - \widetilde{u}(t)) \leq \phi(\widetilde{u}(t_0)) - \phi(\widetilde{u}(t))$. This inequality shows $t \rightarrow \phi(\widetilde{u}(t))$ is in $L^1(0,T)$ since $\widetilde{z} \in L^2(0,T;X')$ and $\widetilde{u} \in L^2(0,T;X)$. Also, for $t \in [0,T-h]$,

$$\mathcal{X}_{[0,T-h]}(t)\,\widetilde{z}(t)\left(\frac{\widetilde{u}(t+h) - \widetilde{u}(t)}{h}\right) \leq \mathcal{X}_{[0,T-h]}(t)\,\frac{\phi(\widetilde{u}(t+h)) - \phi(\widetilde{u}(t))}{h}$$

$$\leq \mathcal{X}_{[0,T-h]}(t)\,\widetilde{z}(t+h)\left(\frac{\widetilde{u}(t+h) - \widetilde{u}(t)}{h}\right)$$

Now $\mathcal{X}_{[0,T-h]}(\cdot)\,\widetilde{z}(\cdot + h) \rightarrow z(\cdot)$ in $L^2(0,T;X')$ by continuity of translation. Also,

$$\mathcal{X}_{[0,T-h]}(\cdot)\,\frac{\widetilde{u}(\cdot+h) - \widetilde{u}(\cdot)}{h} = \mathcal{X}_{[0,T-h]}(\cdot)\,\frac{u(\cdot+h) - u(\cdot)}{h}$$

$$= \mathcal{X}_{[0,T-h]}(\cdot)\,\frac{1}{h}\int_{(\cdot)}^{(\cdot)+h} u'(s)\,ds$$

in $L^2(0,T;X)$ and so by Lemma 24.23,

$$\mathcal{X}_{[0,T-h]}(\cdot)\,\frac{\phi(\widetilde{u}(\cdot+h)) - \phi(\widetilde{u}(\cdot))}{h} \rightarrow z(u')$$

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 Theorem 18.2, 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 24.25 *Let ϕ_1 and ϕ_2 be convex, $l.s.c.$ and proper. Then*

$$\delta(\lambda\phi_i)(x) = \lambda\delta\phi_i(x), \ \delta(\phi_1+\phi_2)(x) \supseteq \delta\phi_1(x) + \delta\phi_2(x) \tag{8}$$

if $\lambda > 0$. If there exists $\overline{x} \in dom(\phi_1) \cap dom(\phi_2)$ and ϕ_1 is continuous at $\overline{x}$ then for all $x \in X$,

$$\delta(\phi_1+\phi_2)(x) = \delta\phi_1(x) + \delta\phi_2(x). \tag{9}$$

Proof: Formula (8) is obvious so we only need to show Formula (9). Suppose $\overline{x}$ is as described. It is clear Formula (9) holds whenever $x \notin dom(\phi_1) \cap dom(\phi_2)$ since then both sides equal $\emptyset$. Therefore, we will assume $x \in dom(\phi_1) \cap dom(\phi_2)$ in what follows. Let $x^* \in \delta(\phi_1+\phi_2)(x)$. We need to show x^* is the sum of an element of $\delta\phi_1(x)$ and $\delta\phi_2(x)$. Define

$$C_1 \equiv \{(y,a) \in X \times \mathbb{R} : \phi_1(y) - x^*(y-x) - \phi_1(x) \leq a\},$$
$$C_2 \equiv \{(y,a) \in X \times \mathbb{R} : a \leq \phi_2(x) - \phi_2(y)\}.$$

Both C_1 and C_2 are convex and nonempty. In addition to this,

$$(\overline{x}, \phi_1(\overline{x}) - x^*(\overline{x}-x) - \phi_1(x) + 1) \in int(C_1)$$

due to the assumed continuity of ϕ_1 at $\overline{x}$. If $(y,a) \in int(C_1)$ then

$$\phi_1(y) - x^*(y-x) - \phi_1(x) \leq a - \epsilon$$

whenever ϵ is small enough. Therefore, if (y,a) is also in C_2, the assumption that $x^* \in \delta(\phi_1+\phi_2)(x)$ implies

$$a - \epsilon \geq \phi_1(y) - x^*(y-x) - \phi_1(x) \geq \phi_2(x) - \phi_2(y) \geq a,$$

a contradiction. Therefore $int(C_1) \cap C_2 = \emptyset$ and so by Corollary 6.11 and Lemma 24.10, there exists $(w^*, \beta) \in X' \times \mathbb{R}$ with

$$(w^*, \beta) \neq (0,0), \tag{10}$$

and

$$w^*(y) + \beta a \geq w^*(y_1) + \beta a_1, \tag{11}$$

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

$$\beta(\phi_1(\overline{x}) - x^*(\overline{x}-x) - \phi_1(x) + 1) \geq \beta(\phi_2(x) - \phi_2(\overline{x})).$$

Dividing by β yields

$$\phi_1(\overline{x}) - x^*(\overline{x}-x) - \phi_1(x) + 1 \leq \phi_2(x) - \phi_2(\overline{x})$$

and so

$$\begin{aligned}\phi_1(\overline{x}) + \phi_2(\overline{x}) - (\phi_1(x) + \phi_2(x)) + 1 &\leq x^*(\overline{x} - x) \\ &\leq \phi_1(\overline{x}) + \phi_2(\overline{x}) - (\phi_1(x) + \phi_2(x)),\end{aligned}$$

a contradiction. Therefore, $\beta \geq 0$.

Now suppose $\beta = 0$. Letting

$$a = \phi_1(\overline{x}) - x^*(\overline{x} - x) - \phi_1(x) + 1,$$

$$(\overline{x}, a) \in int(C_1),$$

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, Formula (11) 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 Formula (10). This proves the claim.

Now with the claim, it follows $\beta > 0$ and so, letting $z^* = w^*/\beta$, Formula (11) implies

$$z^*(y) + a \geq z^*(y_1) + a_1$$

whenever $(y, a) \in C_1$ and $(y_1, a_1) \in C_2$. In particular,

$$\begin{gathered}(y, \phi_1(y) - x^*(y - x) - \phi_1(x)) \in C_1 \text{ and} \\ (y_1, \phi_2(x) - \phi_2(y_1)) \in C_2. \qquad (12)\end{gathered}$$

So letting $y = x$,

$$z^*(x) + (\phi_1(x) - x^*(x - x) - \phi_1(x)) \geq z^*(y_1) + \phi_2(x) - \phi_2(y_1).$$

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 Formula (12). Then

$$z^*(y) + \phi_1(y) - x^*(y - x) - \phi_1(x) \geq z^*(x)$$

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.

24.5 Hilbert space

24.6 Exercises

1. For A a maximal monotone operator defined on a Hilbert space H, let
$$G(A) \equiv \{[x,y] : x \in D(A) \text{ and } y \in Ax\}.$$
Show that for $\lambda > 0$, λA is also maximal monotone. We define $J_\lambda(A)$, written as J_λ for short, by
$$J_\lambda(A)(x) \equiv (I + \lambda A)^{-1} x.$$
Show
$$|J_\lambda x - J_\lambda y| \leq |x - y|.$$
Hint: For $r \in (-1,1)$ and $f \in H$, show there exists a solution, u, to the equation,
$$(1+r)u + Au \ni (1+r)f,$$
as follows. Let
$$J_1 = (I + A)^{-1}$$
and show J_1 is Lipschitz continuous with Lipschitz constant 1. This equation has a solution if and only if
$$u = J_1((1+r)f - ru) = Tu.$$
Show T is a contraction map.
2. ↑ Define for A maximal monotone,
$$A_\lambda x \equiv \frac{1}{\lambda}x - \frac{1}{\lambda}J_\lambda x.$$
Show A_λ is Lipschitz continuous with Lipschitz constant no more than $\frac{2}{\lambda}$. Also verify that
$$A_\lambda x \in AJ_\lambda x,$$
and
$$|A_\lambda x| \leq |y|$$
for all $y \in Ax$ if $x \in D(A)$. This operator, A_λ, is called the Yosida approximation to A.
3. ↑ Suppose
$$(y_1 - y, x_1 - x) \geq 0$$
for all $[x,y] \in G(A)$ where A is maximal monotone. Show that this implies $x_1 \in D(A)$ and $y_1 \in Ax_1$. **Hint:** Try to show
$$J_\lambda(x_1 + \lambda y_1) = x_1$$
because then it will follow $x_1 \in D(A)$ and $y_1 \in Ax_1$. To verify this, use the assumption and Problem 2 to conclude
$$0 \leq (y_1 - A_\lambda(x_1 + \lambda y_1), x_1 - J_\lambda(x_1 + \lambda y_1)).$$

Then simplify to find

$$0 \leq -\frac{1}{\lambda}\left(x_1 - J_\lambda\left(x_1 + \lambda y_1\right), x_1 - J_\lambda\left(x_1 + \lambda y_1\right)\right).$$

The problem shows the graphs of these operators are maximal with respect to also being monotone and this is the reason for the name, maximal monotone.

4. ↑ Suppose $[x_k, y_k] \in G(A)$ and

$$x_k \to x, y_k \rightharpoonup y$$

where the half arrow denotes weak convergence. Show that then $[x, y] \in G(A)$.

5. ↑ Let A be maximal monotone and let B be Lipschitz and monotone. Then $A + B$ is maximal monotone. **Hint:** First suppose B has Lipschitz constant less than one. Then consider

$$Tx \equiv (I + A)^{-1}(y - Bx).$$

Show T is a contraction map and consequently has a fixed point x which satisfies

$$y \in x + Ax + Bx.$$

Next let $A + B$ play the role of A to conclude that $A + B + B$ is maximal monotone. Continuing in this way, show that any Lipschitz constant is all right.

6. ↑ 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 λ. Show there exists

$$x_1 \in D(A) \cap D(B)$$

such that

$$y \in x_1 + Bx_1 + Ax_1.$$

Hint: $y - x_\lambda - B_\lambda x_\lambda \in Ax_\lambda$ and so

$$\begin{aligned} |x_\lambda - x_\mu|^2 &\leq (B_\mu x_\mu - B_\lambda x_\lambda, x_\lambda - x_\mu) \\ &= -(B_\lambda x_\lambda - B_\mu x_\mu, x_\lambda - x_\mu) \\ &= -(B_\lambda x_\lambda - B_\mu x_\mu, J_\lambda(B)x_\lambda - J_\mu(B)x_\mu) - \\ &\quad (B_\lambda x_\lambda - B_\mu x_\mu, \lambda B_\lambda x_\lambda - \mu B_\mu x_\mu) \\ &\leq |(B_\lambda x_\lambda - B_\mu x_\mu, \lambda B_\lambda x_\lambda - \mu B_\mu x_\mu)|. \end{aligned}$$

Conclude $\{x_\lambda\}$ is Cauchy as $\lambda \to 0$. Select a subsequence

$$x_\lambda \to x_1, B_\lambda x_\lambda \rightharpoonup z_1, \text{ and } y - x_\lambda - B_\lambda x_\lambda \rightharpoonup z_2.$$

Use Problem 4 and the observation that $J_\lambda(B)x_\lambda - x_\lambda \to 0$ to conclude

$$z_1 \in Bx_1, z_2 \in Ax_1,$$

and

$$y = x + z_1 + z_2.$$

7. ↑ Let A be maximal monotone and let $B = \partial\phi$ where ϕ is proper, lower semicontinuous, and convex. Suppose
$$\phi\left(J_\lambda\left(A\right)x\right) \leq \phi\left(x\right) + C\lambda$$
and there exists $\xi \in D\left(A\right) \cap D\left(\phi\right)$. Then $A + \partial\phi$ is maximal monotone. **Hint:** Let $y \in H$ be arbitrary and let x_λ be given by
$$y \in x_\lambda + \partial\phi\left(x_\lambda\right) + A_\lambda x_\lambda$$
and show $A_\lambda x_\lambda$ is bounded. Using Problem 6 it will follow $A + \partial\phi$ is maximal monotone. To do this, note
$$\left(y - x_\lambda - A_\lambda x_\lambda, J_\lambda\left(x_\lambda\right) - x_\lambda\right) \leq C\lambda$$
because $y - x_\lambda - A_\lambda x_\lambda \in \partial\phi\left(x_\lambda\right)$. Thus,
$$-\left(y - x_\lambda - A_\lambda x_\lambda, A_\lambda x_\lambda\right) \leq C. \tag{7}$$
Also since $A_\lambda\xi$ is bounded independent of λ, (Problem 2), and A_λ is monotone,
$$\begin{aligned}\phi\left(\xi\right) - \phi\left(x_\lambda\right) &\geq \left(y - x_\lambda - A_\lambda x_\lambda, \xi - x_\lambda\right)\\ &\geq \left(y - x_\lambda, \xi - x_\lambda\right) - \left(A_\lambda x_\lambda, \xi - x_\lambda\right)\\ &\geq \left(y - x_\lambda, \xi - x_\lambda\right) - \left(A_\lambda\xi, \xi - x_\lambda\right) \geq \left|x_\lambda\right|^2 - C\left|x_\lambda\right|\end{aligned}$$
for some C independent of λ. Hence $C \geq \phi\left(x_\lambda\right) + \left|x_\lambda\right|^2$. By Theorem ?? we can find $\left|x_\lambda\right|$ is bounded and then Formula (7) shows $A_\lambda x_\lambda$ is bounded.

8. Let ϕ be a proper convex function defined on a normed linear space. Show ϕ is lower semicontinuous if and only if whenever $u_n \to u$, $\phi\left(u\right) \leq \liminf_{n\to\infty}\phi\left(u_n\right)$.

9. Let $L : D\left(L\right) \subseteq L^2\left(\Omega\right) \to L^2\left(\Omega; \mathbb{R}^n\right)$ be given by $Lu \equiv \nabla u$ where $D\left(L\right)$ is defined to be the space of functions in $L^2\left(\Omega\right)$ whose weak

Appendix A
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 we give a proof that the axiom of choice implies the Hausdorff maximal theorem, a remarkable theorem about partially ordered sets.

We say that 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 $\prec \equiv \subseteq$. Note that we can not conclude that any two elements in a partially ordered set are related. In other words, just because $x,\ y$ are in the partially ordered set, it does not follow that either $x \prec y$ or $y \prec x$. We call a subset of a partially ordered set, $\mathcal{C}$, a chain if $x,\ y \in \mathcal{C}$ implies that either $x \prec y$ or $y \prec x$. Sometimes this is called a totally ordered set. We say $\mathcal{C}$ is a maximal chain if whenever $\widetilde{\mathcal{C}}$ is a chain containing $\mathcal{C}$, it follows the two chains are equal. In other words $\mathcal{C}$ is a maximal chain if there is no strictly larger chain.

Lemma 24.1 *Let $\mathcal{F}$ be a nonempty partially ordered set with partial order $\prec$. Then assuming the axiom of choice, there exists a maximal chain in $\mathcal{F}$.*

Proof: Let $\mathcal{X}$ be the set of all chains from $\mathcal{F}$. For $\mathcal{C} \in \mathcal{X}$, let

$$S_{\mathcal{C}} = \{x \in \mathcal{F} \text{ such that } \mathcal{C} \cup \{x\} \text{ is a chain strictly larger than } \mathcal{C}\}.$$

If $S_{\mathcal{C}} = \emptyset$ for any $\mathcal{C}$, then $\mathcal{C}$ is maximal and we are done. Thus, assume $S_{\mathcal{C}} \neq \emptyset$ for all $\mathcal{C} \in \mathcal{X}$. Let $f(\mathcal{C}) \in S_{\mathcal{C}}$. (This is where the axiom of choice is being used.) Let

$$g(\mathcal{C}) = \mathcal{C} \cup \{f(\mathcal{C})\}.$$

Thus $g(\mathcal{C}) \supsetneq \mathcal{C}$ and $g(\mathcal{C}) \setminus \mathcal{C} = \{f(\mathcal{C})\} = \{\text{a single element of } \mathcal{F}\}$. We call a subset $\mathcal{T}$ of $\mathcal{X}$ a tower if

$$\emptyset \in \mathcal{T},$$
$$\mathcal{C} \in \mathcal{T} \text{ implies } g(\mathcal{C}) \in \mathcal{T},$$

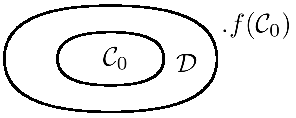
and if $\mathcal{S} \subseteq \mathcal{T}$ is totally ordered with respect to set inclusion, then

$$\cup \mathcal{S} \in \mathcal{T}.$$

Note that $\mathcal{X}$ is a tower. Let $\mathcal{T}_0$ be the intersection of all towers. Thus, $\mathcal{T}_0$ is a tower, the smallest tower. We wish to show that any two sets in $\mathcal{T}_0$ are comparable in the sense of set inclusion so that $\mathcal{T}_0$ is actually a chain. This will proceed in two steps. First let $\mathcal{C}_0$ be a set of $\mathcal{T}_0$ which is comparable to every set of $\mathcal{T}_0$. Such sets exist, $\emptyset$ being an example. Let

$$\mathcal{B} \equiv \{\mathcal{D} \in \mathcal{T}_0 : \mathcal{D} \supsetneq \mathcal{C}_0 \text{ and } f(\mathcal{C}_0) \notin \mathcal{D}\}.$$

The picture represents sets of $\mathcal{B}$.



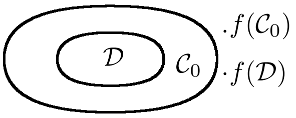
We wish to show $\mathcal{B} = \emptyset$. This will be accomplished by showing $\widetilde{\mathcal{T}_0} \equiv \mathcal{T}_0 \setminus \mathcal{B}$ is a tower. Since $\mathcal{T}_0$ is the smallest tower, this will require that $\widetilde{\mathcal{T}_0} = \mathcal{T}_0$ and so $\mathcal{B} = \emptyset$. Note that for $\mathcal{D} \in \mathcal{T}_0$, to say $\mathcal{D} \in \widetilde{\mathcal{T}_0}$ is the same as saying $\mathcal{D} \notin \mathcal{B}$.

Claim: $\widetilde{\mathcal{T}_0}$ is a tower and so $\mathcal{B} = \emptyset$.

Proof of the claim: It is clear that $\emptyset \in \widetilde{\mathcal{T}_0}$. Suppose $\mathcal{D} \in \widetilde{\mathcal{T}_0}$. We need to verify $g(\mathcal{D}) \in \widetilde{\mathcal{T}_0}$.

Case 1: $f(\mathcal{D}) \in \mathcal{C}_0$. If $\mathcal{D} \subseteq \mathcal{C}_0$, then if $g(\mathcal{D}) \nsubseteq \mathcal{C}_0$, it would follow $g(\mathcal{D}) \supsetneq \mathcal{C}_0$ due to the assumption that $\mathcal{C}_0$ is comparable. However, both $\{f(\mathcal{D})\}$ and $\mathcal{D}$ are contained in $\mathcal{C}_0$ and so $g(\mathcal{D}) \subseteq \mathcal{C}_0$ which implies $g(\mathcal{D}) \notin \mathcal{B}$. On the other hand, if $\mathcal{D} \supsetneq \mathcal{C}_0$, then since $\mathcal{D} \in \widetilde{\mathcal{T}_0}$, we know $f(\mathcal{C}_0) \in \mathcal{D}$ and so $g(\mathcal{D})$ also contains $f(\mathcal{C}_0)$ implying $g(\mathcal{D}) \notin \mathcal{B}$.

Case 2: $f(\mathcal{D}) \notin \mathcal{C}_0$. If $\mathcal{D} \subsetneq \mathcal{C}_0$ then we can't have $f(\mathcal{D}) \notin \mathcal{C}_0$ because if this were so, $g(\mathcal{D})$ would not compare to $\mathcal{C}_0$.



Hence if $f(\mathcal{D}) \notin \mathcal{C}_0$, then $\mathcal{D} \supseteq \mathcal{C}_0$. If $\mathcal{D} = \mathcal{C}_0$, then $g(\mathcal{D}) = g(\mathcal{C}_0)$ so $g(\mathcal{D}) \notin \mathcal{B}$. Therefore, assume $\mathcal{D} \supsetneq \mathcal{C}_0$. Then, since $\mathcal{D}$ is in $\widetilde{\mathcal{T}_0}$, $f(\mathcal{C}_0) \in \mathcal{D}$ and so $f(\mathcal{C}_0) \in g(\mathcal{D})$. Therefore, $g(\mathcal{D}) \in \widetilde{\mathcal{T}_0}$.

Now suppose $\mathcal{S}$ is a totally ordered subset of $\widetilde{\mathcal{T}_0}$. Then if every element of $\mathcal{S}$ is contained in $\mathcal{C}_0$, so is $\cup\mathcal{S}$ and so $\cup\mathcal{S} \in \widetilde{\mathcal{T}_0}$. If, on the other hand, some chain from $\mathcal{S}$, $\mathcal{C}$, contains $\mathcal{C}_0$ properly, then since $\mathcal{C} \notin \mathcal{B}$, $f(\mathcal{C}_0) \in \mathcal{C} \subseteq \cup\mathcal{S}$ showing that $\cup\mathcal{S} \notin \mathcal{B}$ also. This has proved $\widetilde{\mathcal{T}_0}$ is a tower and since $\mathcal{T}_0$ is the smallest tower, it follows $\widetilde{\mathcal{T}_0} = \mathcal{T}_0$.

Now we define $\mathcal{T}_1$ to be the set of all chains from $\mathcal{T}_0$ which are comparable to every chain from $\mathcal{T}_0$.

Claim: $\mathcal{T}_1$ is a tower.

Proof of the claim: It is clear that $\emptyset \in \mathcal{T}_1$. Suppose $\mathcal{C}_0 \in \mathcal{T}_1$. We need to verify that $g(\mathcal{C}_0) \in \mathcal{T}_1$. Let $\mathcal{D} \in \mathcal{T}_0$ and consider $g(\mathcal{C}_0) \equiv \mathcal{C}_0 \cup \{f(\mathcal{C}_0)\}$. If $\mathcal{D} \subseteq \mathcal{C}_0$, then $\mathcal{D} \subseteq g(\mathcal{C}_0)$. If $\mathcal{D} \supsetneq \mathcal{C}_0$, then $\mathcal{D} \supseteq g(\mathcal{C}_0)$ by what was just shown $\left(\widetilde{\mathcal{T}_0} = \mathcal{T}_0\right)$. Hence $g(\mathcal{C}_0)$ is comparable to every set of $\mathcal{T}_1$. Now suppose $\mathcal{S}$ is a chain of elements of $\mathcal{T}_1$ and let $\mathcal{D}$ be an element of $\mathcal{T}_0$. If every element of $\mathcal{S}$ is contained in $\mathcal{D}$, then $\cup\mathcal{S}$ is also contained in $\mathcal{D}$. On the other hand, if some set, $\mathcal{C}$, from $\mathcal{S}$ contains $\mathcal{D}$ properly, then $\cup\mathcal{S}$ also contains $\mathcal{D}$. Thus $\cup\mathcal{S} \in \mathcal{T}_1$.

This shows $\mathcal{T}_1$ is a tower and proves therefore, that $\mathcal{T}_0 = \mathcal{T}_1$. Thus every set of $\mathcal{T}_0$ compares with every other set of $\mathcal{T}_0$ showing $\mathcal{T}_0$ is a chain in addition to being a tower.

Now $\cup\mathcal{T}_0, g(\cup\mathcal{T}_0) \in \mathcal{T}_0$. Hence, because $g(\cup\mathcal{T}_0)$ is a chain in $\mathcal{T}_0$, and $\mathcal{T}_0$ is a chain, it follows $g(\cup\mathcal{T}_0) \subseteq \cup\mathcal{T}_0$. Thus

$$\cup\mathcal{T}_0 \supseteq g(\cup\mathcal{T}_0) \supsetneq \cup\mathcal{T}_0,$$

a contradiction. Hence there must exist a maximal chain after all. This proves the lemma.

If X is a nonempty set, we say $\leq$ is an order on X if

$$x \leq x,$$

and if $x,\ y \in X$, then

$$\text{either } x \leq y \text{ or } y \leq x$$

and

$$\text{if } x \leq y \text{ and } y \leq z \text{ then } x \leq z.$$

We say that $\leq$ 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 24.2 *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

$$\mathcal{F} = \{S \subseteq X : \text{ there exists a well order for } S\}.$$

Thus $\mathcal{F} \neq \emptyset$. We will say that for $S_1,\ S_2 \in \mathcal{F}$, $S_1 \prec S_2$ 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 we observe that $\prec$ is a partial order on $\mathcal{F}$. By the Hausdorff maximal principle, we let $\mathcal{C}$ be a maximal chain in $\mathcal{F}$ and let

$$X_\infty = \cup\mathcal{C}.$$

We also define an order, $\leq$, on X_∞ as follows. If $x,\ y$ are elements of X_∞, we pick $S \in \mathcal{C}$ such that $x,\ y$ are both in S. Then if $\leq_S$ is the order on S, we 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_∞. Then $S \cap U \neq \emptyset$ for some $S \in \mathcal{C}$. Because of the definition of $\leq$, if $y \in S_2 \setminus S_1,\ S_i \in \mathcal{C}$, 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_∞ 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

$$\widetilde{\mathcal{C}} = \{S \in \mathcal{C} \text{ or } X_\infty \cup \{z\}\}.$$

Then $\widetilde{\mathcal{C}}$ is a strictly larger chain than $\mathcal{C}$ contradicting maximality of $\mathcal{C}$. Thus $X \setminus X_\infty = \emptyset$ and this shows X is well-ordered by $\leq$. This proves the lemma.

With these two lemmas we can now state the main result.

Theorem 24.3 *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

$$f \in \prod_{i \in I} X_i.$$

A.1 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. When we say x is maximal, we mean that if $x \prec y$, it follows $y = x$. Show Zorn's lemma is equivalent to the Hausdorff maximal theorem.
2. Let X be a vector space. We say $Y \subseteq X$ is a Hamel basis if every element of X can be written in a unique way as a finite linear combination of elements in Y. Show that every vector space has a Hamel basis and 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.
3. ↑ Using the Baire category theorem of Chapter 3 show that any Hamel basis of a Banach space is either finite or uncountable.

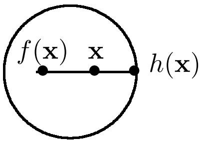
Appendix B
Stone’s Theorem and Partitions of Unity

Appendix C

Taylor Series and Analytic Functions

Appendix D
The Brouwer fixed point theorem

Proof: Suppose there is no fixed point for f. Then define $h(\mathbf{x})$ as shown in the following picture.



Then $h : B \to \partial B$ is a continuous retraction contradicting Lemma 24.63. This proves the theorem.

Corollary 24.65 *Let K be any compact convex subset of $\mathbb{R}^n$ and let $f : K \to K$ be continuous. Then f has a fixed point.*

Proof: Let $K \subseteq B(0,r)$ and define $g : \overline{B(0,r)} \to \overline{B(0,r)}$ by

$$g(\mathbf{x}) \equiv f \circ proj_K(\mathbf{x}).$$

Then g is continuous so $g(\mathbf{x}) = \mathbf{x}$ for some $\mathbf{x} \in B(0,r)$. Thus

$$f(proj_K(\mathbf{x})) = \mathbf{x}.$$

Since $f(K) \subseteq K$, it follows that $\mathbf{x} \in K$ and so $proj_K(\mathbf{x}) = \mathbf{x}$. Hence $f(\mathbf{x}) = \mathbf{x}$. This proves the corollary.