Chap1-6 Booster

Anson Li

Contents

1. Estimates

1.1. The Second Property

Let A be non-empty and bounded. The *infimum* of A is the real number $s_* = \inf A$ such that

Lower Bound Property $s_* \leq x$ for all $x \in A$, and
Second Property of the Infimum for all $\varepsilon > 0$ there exists $x_\varepsilon \in A$ with $x_\varepsilon \leq s_* + \varepsilon$.

It is easy to see that for a given number $c \in \mathbb{R}$, $c \leq s_*$ iff the lower bound property is satisfied, and $s_* \leq c$ iff the c satisfies the second property of the infimum.

Some applications: Brezis Corollary 2.11. Folland 1.10, 1.17.

Similarly, the *supremum* of A is the real number $s^* = \sup A$ such that

Upper Bound Property $x \leq s^*$ for all $x \in A$, and
Second Property of the Supremum for all $\varepsilon > 0$ there exists $x_\varepsilon \in A$ with $s^* + \varepsilon \leq x_\varepsilon$.

Definition 1.1

Let $A \subseteq \overline{\mathbb{R}}$ be non empty. A number $c \in \overline{\mathbb{R}}$ is said to satisfy the second property of the infimum if for every $\varepsilon > 0$, there exists $x_\varepsilon \in A$ such that $x_\varepsilon \leq c + \varepsilon$.

Proposition 1.2

Let $A \subseteq \overline{\mathbb{R}}$ non-empty, $c \in \overline{\mathbb{R}}$ satisfies the second property of the infimum, iff c is larger than $\inf A$.

Definition 1.3

If $c_1, c_2 \in \overline{\mathbb{R}}$, we say that c_1 is as small as c_2, if $c_1 \leq c_2$.

1.2. Accumulation Points

Let $A \subseteq \mathbb{R}$ be non-empty and bounded. Under what conditions can be assume $\inf(A) = \inf \operatorname{acc}(A)$, and $\sup(A) = \sup \operatorname{acc}(A)$? The equality clearly does not hold for all A, take $A = \{0\} \cup [1,2]$, then $\inf(A) = 0$ but $\operatorname{acc}(A) = [1, 2]$ so that $\inf \operatorname{acc}(A) = 1 \neq 0$.

In general however, we have the following;

$$\inf(A) \leq \inf \operatorname{acc}(A).$$

To prove this: let $y \in \operatorname{acc}(A)$, we know that y can be approximated (in the topology of $\mathbb{R}$) by elements of A, so let $y = \lim x_n$ for $(x_n) \subseteq A$. The infimum over A clearly bounds each x_n from below, so $\inf(A) \leq y$ for all y. Therefore $\inf(A) \leq \inf \operatorname{acc}(A)$.

Remark 1.4

Let $A \subseteq \mathbb{R}$ be non-empty and bounded. A number s_* is the *infimum* of A if for all $y \in A$, $s_* \leq y$, and for all $\varepsilon > 0$ there exists $y \in A$ where

$$s_* \leq y < s_* + \varepsilon.$$

We can replace the strict inequality by a regular inequality. Because the strict inequality *implies* the ordinary inequality, and if the condition with the ordinary inequality holds, we simply take $\varepsilon 2^{-1} > 0$ and obtain $y_{\varepsilon 2^{-1}}$ so that

$$s_* \leq y_{\varepsilon 2^{-1}} \leq s_* + \varepsilon 2^{-1} < s_* + \varepsilon.$$

The *second property of the infimum* is precisely, (which holds for $c \in \mathbb{R}$ iff $c \geq s_*$)

> There exists a sequence $(y_n) \subseteq A$ where $\limsup_{n\to\infty} y_n \leq s_*$.

1.3. Subsequential Limits

Let $(x_n) \subseteq \mathbb{R}$, there are two properties of the limit inferior worth mentioning:

$$x_n \to +\infty \quad \text{iff} \quad \liminf x_n = +\infty,$$

and x_n is bounded below iff $\liminf x_n \neq -\infty$.
Similarly, $x_n \to -\infty$ iff $\limsup x_n = -\infty$, and x_n is bounded above iff $\limsup x_n \neq +\infty$.

1.3.1 Additivity of Subsequential Limits

Let (x_n) and (y_n) be $\mathbb{R}$-valued sequences that are bounded below. Then the limit inferior is *super-additive*.

$$\liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n \leq \liminf_{n\to\infty}(x_n + y_n).$$

Suppose at least one of the x_n or $y_n \to +\infty$, without loss of generality: $\liminf_{n\to\infty} x_n = +\infty$, then the sum $(x_n + y_n)$ also converges to $+\infty$ as y_n is bounded below, and the right hand side is equal to $+\infty$.

If both of the limit inferiors are finite, we see that (upon adding two separate inequalities),

$$\inf_{n\geq m} x_n + \inf_{n\geq m} y_n \leq x_m + y_m.$$

For a fixed $k \geq 1$, we can take the infimum respect to $m \geq k$ on both sides:

$$\inf_{m\geq k}\left(\inf_{n\geq m} x_n + \inf_{n\geq m} y_n\right) \leq \inf_{m\geq k}(x_m + y_m)$$

and because $(\inf_{n\geq m} x_n)_m$ and $(\inf_{n\geq m} y_n)_m$ are increasing sequences, their termwise sum is also an increasing sequence. The left hand side simplifies to $\inf_{n\geq k} x_n + \inf_{n\geq k} y_n$ taking limits on both sides gives

$$\lim_{k\to\infty}\left(\inf_{n\geq k} x_n + \inf_{n\geq k} y_n\right) \leq \liminf_{n\to\infty}(x_n + y_n)$$

By continuity of addition, we can push the limit into each of the two terms on the left, and

$$\liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n \leq \liminf_{n\to\infty}(x_n + y_n).$$

Similarly, if (x_n) and (y_n) are real-valued sequences that are bounded above, then the limit superior is *subadditive*.

$$\limsup_{n\to\infty}(x_n + y_n) \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n.$$

We can prove this by taking a reflection, let $x_n' = -x_n$ and $y_n' = -y_n$, then:

$$\liminf_{n\to\infty}(x_n') + \liminf_{n\to\infty}(y_n') \leq \liminf_{n\to\infty}(x_n' + y_n'),$$

which is a reflection away from the desired inequality.

1.3.2 Multiplicativity of Subsequential Limits

Let $(x_n), (y_n) \subseteq \mathbb{R}$ be non-negative sequences. Then,

$$(\liminf x_n)(\liminf y_n) \leq \liminf(x_n y_n) \leq \limsup(x_n y_n) \leq (\limsup x_n)(\lim s$$

this can be proven by minimizing (resp. maximizing) the product termwise just like in the proof for additivity. For all $j \geq n$,

$$\left(\inf_{m\geq n} x_m\right)\left(\inf_{m\geq n} y_m\right) \leq (x_j y_j) \leq \left(\sup_{m\geq n} x_m\right)\left(\sup_{m\geq n} y_m\right).$$

Now we send $n \to \infty$,

1.3.3 Subsequential Limits of bdd. Sequences

Let $\{x_n\} \subseteq \mathbb{R}$ be bounded, and we write $E_n = \{x_m,\ m \geq n\}$, it is clear that $\limsup x_n = \inf_{n\geq 1} \sup E_n$, and $\liminf x_n = \sup_{n\geq 1} \inf E_n$. We see that

$$a < \limsup x_n < b \quad \text{iff} \quad \begin{cases} a < x_n & \text{frequently} \\ x_n < b & \text{eventually,} \end{cases}$$

and

$$c < \liminf x_n < d \quad \text{iff} \quad \begin{cases} x_n < d & \text{frequently} \\ c < x_n & \text{eventually.} \end{cases}$$

1.3.4 Suprema and Infima Calculations of bdd. Sets

Let $A = (a_1, \dots, a_n) \subseteq \mathbb{R}$ be a finite set, then

$$a \le \max(A) \le b \quad \text{iff} \quad \begin{cases} a \le x & \text{for some } x \in A \\ x \le b & \text{for all } x \in A, \end{cases}$$

and

$$c \le \min(A) \le d \quad \text{iff} \quad \begin{cases} x \le d & \text{for some } x \in A \\ c \le x & \text{for all } x \in A. \end{cases}$$

We extend this to the infinite case, but now one of our inequalities will have to be strict. Starting with the case with strict inequalities only, if $A \subseteq \mathbb{R}$ is non-empty and bounded,

$$a < \sup(A) < b \quad \text{iff} \quad \begin{cases} a < x & \text{for some } x \in A \\ x < b & \text{for all } x \in A, \end{cases}$$

similarly

$$c < \inf(A) < d \quad \text{iff} \quad \begin{cases} x < d & \text{for some } x \in A \\ c < x & \text{for all } x \in A. \end{cases}$$

The 'for all' bounds can be replaced with regular inequalities. This means

$$\sup(A) \le b \quad \text{iff} \quad x \le b \quad \forall x \in A, \text{ and}$$
$$c \le \inf(A) \quad \text{iff} \quad c \le x \quad \forall x \in A.$$

Finally, if the supremum (resp. infimum) of A is attained, then we can replace all of the estimates by regular inequalities. This holds if A is closed.

1.4. Boosted Triangle Inequalities

Let (X, d) be a metric space, the distance function $d : X \times X \to [0, +\infty)$ satisfies the *triangle inequality*.

$$d(x_1, x_2) \le d(x_1, z) + d(x_2, z) \quad \forall x_1, x_2, z \in X.$$

Lemma 1.5 Boosted Triangle Inequality

Let $x_1, x_2, z_1, z_2 \in X$, then

$$|d(x_1, z_1) - d(x_2, z_2)| \le d(x_1, x_2) + d(z_1, z_2).$$

Proof. Recall that $|a| = \max(a, -a)$, it suffices to show that $d(x_1, z_1) - d(x_2, z_2)$ is bounded above by the sum of distances because we can replace x_1 with x_2 and z_1 with z_2 to obtain the desired estimate. By adding and subtracting a factor of $d(x_1, z_2)$ and using the triangle-Lipschitz inequality twice, we obtain

$$d(x_1, z_1) - d(x_2, z_2) \pm d(x_1, z_2) \le \Big[d(x_1, z_1) - d(x_1, z_2) \Big] + \Big[d(x_1, z_2) - d(x_2, z_2) \Big].$$

The absolute values of two squared terms on the right hand side are controlled by $d(z_1, z_2)$ and $d(x_1, z_2)$ respectively, in symbols:

$$|d(x_1, z_1) - d(x_1, z_2)| \le d(z_1, z_2) \quad \text{and} \quad |d(x_1, z_2) - d(x_2, z_2)| \le d(x_1, x_2).$$

Combining the last two estimates finishes the proof. ■

Corollary 1.6 Continuity of Max, Min

Let $x_1, x_2, z_1, z_2 \in \mathbb{R}$, then

$$\Big| \max(|x_1|, |z_1|) - \max(|x_2|, |z_2|) \Big| \le |x_1 - x_2| + |z_1 - z_2|$$

the same estimate holds for max replaced with min. This means that we can estimate the difference of maxima (resp. minima) by comparing 'coordinate-wise'.

1.5. Maximum, Minimum Estimates

Remark 1.7

Let $\{f_j\}_1^n$ be a sequence of $\mathbb{R}$-valued functions, suppose also that $h(x)$ and $l(x)$ are $\mathbb{R}$-valued as well, define for all $j \le n$,

$$E_j^+ = \{z \in X,\ h(z) < f_j(z)\} \quad \text{and} \quad E_j^- = \{z \in X,\ f_j(z) < l(z)\}.$$

It is easy to see that

$$\cup_1^n E_j^+ = \{z \in X,\ h(z) < \max f_j(z)\}, \qquad \cap_1^n E_j^+ = \{z \in X,\ h(z) < \min f_j(z$$
$$\cup_1^n E_j^- = \{z \in X,\ \min f_j(z) < l(z)\}, \qquad \cap_1^n E_j^- = \{z \in X,\ \max f_j(z) < l(z$$

1.5.1 Positive and Negative Parts

Remark 1.8

For any $z \in \mathbb{R}$, write $z^+ = \max(z, 0)$, and $z^- = \max(-z, 0)$. It is clear that $|z| = z^+ + z^-$, and an easy geometric argument will show $-z^- \le z \le z^+ \le |z|$. We also have the representation $x = x^+ - x^-$ for all $x \in \mathbb{R}$.

Lemma 1.9

For any pair x, y of real numbers, $x \le y$ if and only if $x^+ \le y^+$ and $y^- \le x^-$.

Proof. If direction: To prove the first inequality, consider the two cases when $x \ge 0$ and $x < 0$. If $x \ge 0$, it forces $y \ge 0$, so $x^+ = x$ and $y^+ = y$, and the estimate follows from the original one, and if $x < 0$, this implies $x^+ = 0$, and because y^+ is non-negative, the estimate holds. For the second inequality, write $(-y) \le (-x)$, and $\max(-y, 0) \le \max(-x, 0)$ from the first part, and after expanding definitions, we obtain the desired estimate.

Only if direction: There are three cases, if $x > 0$ then $x^+ > 0$, and $y^+ > 0$ by the first estimate, (this is because $\max(a, 0) > 0$ iff there exists a strictly positive $b \in \{a, 0\}$.) so that $x^+ = x$, and $y^+ = y$, and $x \leq y$. Next, if $x = 0$, then $x^- = y^- = 0$, meaning y is forced to be non-negative, and hence $x = 0 \leq y$. Finally, if $x < 0$ then $x^- > 0$, and

$$x = -x^- \leq -y^- \leq y.$$

■

1.6. Binomial Estimates

This section is inspired from a particularl section in Rudin's Principles of Mathematical Analysis (see the pages of Chapter 3 on series), and Yosida's proof for The Weierstrass approximation theorem in 'Functional Analysis' – Yosida.

1.6.1 Binomial Theorem

Let $x, y \in \mathbb{R}$, the *binomial theorem* tells us $u(x, y) = (x + y)^n = \sum_0^n \binom{n}{p} x^p y^{n-p}$. We define $r_p(x, y) = \binom{n}{p} x^p y^{n-p}$ for $0 \leq p \leq n$. The notation $(n)_{\underline{k}} = n(n-1)\cdots(n-k+1) = \prod_0^k (n-j)$ is the *kth falling factorial of n*. If $F(t) = at^p$ is a monic polynomial, where $p \neq 0$,

$$F'(t) = pt^{-1}F(t) \quad \text{and} \quad F^{(k)}(t) = (p)_{\underline{k}} t^{-k} F(t).$$

Appealing to the chain rule, we get

$$\partial_x^k u(x, y) = (n)_{\underline{k}} (x+y)^{n-k} \quad \text{and} \quad \partial_x^k r_p(x, y) = (p)_{\underline{k}} t^{-k} r_p(x, y).$$

Eventually we want to fix $y = 1 - x$, so that $(x + y)^q = 1$ for all q, x, but we derive these formulas in full generality. Differentiating both sides of $u(x, y) = \sum_0^n r_p(x, y)$, and multiplying by x^k gives

$$(n)_{\underline{k}} (x + y)^{n-k} x^k = \sum_0^n (p)_{\underline{k}} r_p(x, y).$$

Set $y = (1 - x)$, so that $u(x) = x^p(1 - x)^{n-p}$ and $r_p(x) = \binom{n}{p} x^p (1 - x)^{n-p}$ and for $k = 1, 2$

$$nx = \sum_0^n p r_p(x) \quad \text{and} \quad n(n-1)x^2 = \sum_0^n p(p-1) r_p(x).$$

We see that a polynomial in terms of x turns into a sum of polynomials in p,

$$F(x) = \sum_0^d a_i x^i, \quad \text{then} \quad F(x) = \sum_0^n \sum_0^d \frac{(p)_{\underline{i}} a_i}{(n)_{\underline{i}}} r_p(x).$$

Conversely, a sum of polynomials in p can be converted into a polynomial $G(x)$

$$G(x) = \sum_0^n \sum_0^d b_j p^j r_p(x).$$

To compute $G(x)$, we require a conversion between $b_j p^j$ and $c_j (p)_{\underline{j}}$.

Lemma 1.10

$nx(1 - x) = \sum_0^n |p - nx|^2 r_p(x)$ for all $n \geq 0$, and x.

Proof. We write $|p - nx|^2 = p^2 + (nx)^2 - 2nxp$. Being mindful of the coefficients which are dependent of p and rewriting $|p - nx|^2$ in terms of falling powers of p,

$$p^2 + (nx)^2 - 2nxp = p(p-1) + (1 - 2nx) + (nx)^2 p^0.$$

So that $\sum_0^n |p - nx|^2 r_p(x) = n(n-1)x^2 + (1-2nx)p + (nx)^2 p^0$, as $\sum_0^n r_p(x) = u(x) = 1$ for all x. The expression simplifies

$$\begin{aligned}\textstyle\sum_0^n |p - nx|^2 r_p(x) &= n(n-1)x^2 + nx - (nx)^2 \\ &= -nx^2 + nx = nx(1 - x).\end{aligned}$$

■

Remark 1.11

To bound X in l^1, we can try bounding $Z = XY$ and $1/Y$ separately in l^p, l^q.

1.6.2 Sequences with Polynomial Growth

Lemma 1.12

Let $1 \leq x_n \lesssim n^p$ for some $p > 0$, then $\lim |x_n|^{1/n} = 1$.

Proof. Write $z_n = |x_n|^{1/n} - 1 \geq 0$, we want to show that $z_n \to 0$. We will prove that $|z_n| = \omega_n$.

nth power = sum of lower order powers Using the Binomial theorem, we can split the nth power into a sum of powers of lower order.

$$\sum_{k=0}^n \binom{n}{k} |z_n|^k = x_n \leq Cn^p$$

Pick $k > p$ large Let $k > p$ be sufficiently large such that, independently of $n \geq k$, and the binomial coefficient for a fixed k is of the form $(n - k)^k \lesssim_k \binom{n}{k} \lesssim_k n^k$. Therefore

$$|z_n| \lesssim \left(\frac{n^p}{\binom{n}{k}} \right)^{1/k} \lesssim_k \left(\frac{n^{p/k}}{(n-k)} \right) \lesssim_k n^{p/k-1} \to 0. \quad (1)$$

General Bounds for Binomial Expansion Equation (1) can be rewritten more carefully, if we consider for eventually large n (compared to k, which is held fixed)

$$\frac{(n-k)!}{k!} \leq \binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!},$$

the leftmost member can be supported below if we assume $n > 2k$, then

$$\frac{n^k}{2^k k!} \leq \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!},$$

and Equation (1) reads

$$|z_n| \le \left(\frac{n^p(2^k k!)}{n^k}\right)^{1/k} = 2(k!)^{1/k}\frac{n^{p/k}}{n} \lesssim_k n^{p/k-1}.$$

■

> **Remark 1.13**
> Generalize this to arbitrary exponents using the exponential function and Taylor expansion?

1.6.3 Applications of Binomial Estimates

- Let $x_n = n^{1/n}$, then $\lim x_n = 1$.

1.7. L^p differences

If f_j and f are in L^p for $p > 0$, we can estimate the pointwise difference:

$$|f_j - f|^p \le \big(|f_j| + |f|\big)^p,$$

each of the two individual terms are bounded above by $\max(|f_j|, |f|)$,

$$|f_j - f|^p \le \big(\max(|f_j|,|f|) + \max|f_j|,|f|)\big)^p \le 2^p\Big[\max(|f_j|,|f|)\Big]^p.$$

The function $t \mapsto |t|^p$ for $t \in [0, +\infty)$ is monotonically increasing for $p > 0$ (use first derivative test), and we can push the pth power into the arguments of the maximum,

$$|f_j - f|^p \le 2^p \max(|f_j|^p,\ |f|^p).$$

and because all norms on finite dimensional vector spaces are equivalent to the L^1 norm, one sees that the quantity $\max(|f_j|^p, |f|^p)$ is the l^∞, and can be bounded

$$\max(|f_j|^p, |f|^p) \le |f_j|^p + |f|^p.$$

This allows us to bound the integral of the difference by the sum of integrals

$$\int_X |f_j - f|^p \,\mathrm{d}x \le 2^p\left[\int_X |f_j|^p \,\mathrm{d}x + \int_X |f|^p \,\mathrm{d}x\right].$$

1.7.1 L^p uniformly bdd. sequences

Let f_j in L^p for $p > 0$ and $f_j \to f$ pointwise a.e, suppose also that $\limsup_{j\to\infty}\int_X |f_j|^p \,\mathrm{d}x \le C$, then

$$\int_X |f|^p \,\mathrm{d}x \le C \quad \text{or} \quad f \in L^p.$$

The proof is straight-forward using Fatou's Lemma:

$$\int_X \liminf_{j\to\infty} |f_j|^p \,\mathrm{d}x \le \liminf_{j\to\infty}\int_X |f_j|^p \,\mathrm{d}x \le C,$$

the result follows because $|f_j|^p \to |f|^p$ pointwise a.e.

1.7.2 Brezis-Lieb Lemma

We follow [BL83], see also [LL01].

1.8. Concavity of $\ln(x)$

1.9. Uniform Convergence and Modulus of Continuity

Let $(a_n) \subseteq \mathbb{R}$ be a sequence, suppose that $a_n \to 0$, then we write $a_n = \omega_n$. Next, if $a_{m,n}$ is a sequence in $\mathbb{N}^+$, we say that $a_{m,n} = \omega_m^n$ if for every n, $a_{m,n} = \omega_m$. It is easy to see if $a_{m,n} = \omega_{m,n}$,

$$\limsup_{m\to\infty} a_{m,n} = \omega_n \quad \text{and} \quad \limsup_{n\to\infty} a_{m,n} = \omega_m.$$

We say that $a_{m,n} = \omega_{m,n}$ if for every $\varepsilon > 0$ there exists $N \ge 1$ where $\min(m,n) \ge N$ implies $a_{m,n} \le \varepsilon$. If (x_n) is a Cauchy sequence, then $|x_n - x_m| = \omega_{m,n}$. It is clear that $a_{m,n} = \omega_{m,n}$ implies $a_{m,n} = \omega_n^m = \omega_m^n$, the converse is not true however. Finally, if $|f(x_0 + y) - f(x_0)| \to 0$ as $y \to 0$, we write $|f(x_0 + y) - f(x_0)| = \omega_{|y|}$.

We give a few examples. In complete metric spaces, sequences converge iff uniformly Cauchy. To see this, fix $x_n \to x$, we can control the distance $d(x_n, x_m) \le d(x, x_n) + d(x, x_m) = \omega_n + \omega_m = \omega_{n,m}$.

Let $f : X \to \mathbb{C}$ be a mapping from a metric, space. It is continuous it $x \in X$ iff

$$\sup_{d(x,z)\le\delta} |f(x) - f(z)| = \omega_\delta^x.$$

f is *uniformly continuous* iff

$$\sup_{d(x,z)\le\delta} |f(x) - f(z)| = \omega_\delta.$$

Let X be a topological space and E a Banach space. We use $B(X)$ to denote the space of bounded, E-valued functions from X.

$B(X)$ is complete. Fix a Cauchy sequence $(f_n) \subseteq B(X)$ in the uniform norm, we can borrow the completeness of E to obtain some function $f : X \to E$ where $f(x) = \lim_n f_n(x)$ for all x. To show that $\|f_n - f\|_u = \omega_n$, let $x \in X$ be arbitrary (the supremum will be taken over all such x later), then $|f_n(x) - f_m(x)| \le \|f_n - f_m\| = \omega_{m,n}$. Take the limit superior as $m \to \infty$, and $|f_n(x) - f(x)| \le \omega_n$ for all $x \in X$, and $\|f_n - f\|_u = \omega_n$.
$BC(X)$ is closed in $B(X)$, making $BC(X)$ a Banach space. Let $(f_n) \subseteq BC(X)$ converge uniformly to $f \in B(X)$, our goal is to show $f \in BC(X)$. Because (f_n) is Cauchy, we can use the same limit superior trick to eliminate one of

the variables. Fix $x_0 \in X$ and $x_1 \in B(r, x_0)$, we know that $\|f_n - f\|_u = \omega_n$, and

$$|f_n(x_1) - f(x_1)| = |f_n(x_0) - f(x_0)| = \omega_n.$$

For every $n \geq 1$, the function f_n is continuous at x_0, so $|f_n(x_1) - f_n(x_0)| = \omega_r^n$. Using the triangle inequality gives $|f(x_1) - f(x_0)| \leq 2\|f_n - f\|_u + |f_n(x_1) - f_n(x_0)|$; and

$$|f(x_1) - f(x_0)| \leq 2\omega_n + \omega_r^n.$$

The left hand side is independent of n, and is bounded by ω_r.

1.9.1 Absolute Convergence of Series

Lemma 1.14

Let $(f_n) \subseteq B(X)$ and $\sum_n f_n$ be absolutely convergent, then there exists $f \in B(X)$ where $s_n = \sum_{j\leq n} f_j$ converges to f uniformly.

Proof. For every $x \in X$, the series $\sum_{n\geq 1} |f_n(x)| \leq \sum_{n\geq 1} \|f_n\|_u$ is summable, and we define $f(x) = \lim_n s_n(x)$ pointwise. Because the series of norms is l^1, so that $\sum_{j>n} \|f_j\|_u = \omega_n$. The proof for uniform convergence consists of checking at every $x \in X$, keeping in mind that $f(x)$ is a limit of partial sums, **not a sum**, we write

$$|f(x) - s_n(x)| = |\lim_{m\to\infty} s_m(x) - s_n(x)|$$
$$\leq \lim_{m\to\infty} \sum_1^m \|f_j\|_u - \sum_1^n \|f_j\|_u,$$

which is bounded by $\sum_{j>n} \|f_j\|_u = \omega_n$. We see that $|f(x) - s_n(x)| = \omega_m$, and the left hand side is independent of x, therefore $\|f - s_n\|_u = \omega_n$. ∎

1.9.2 Uniform Convergence and Differentiation

Lemma 1.15

Let $f_n \in C^1(U, F)$ where $U \overset{\circ}{\subseteq} E$ and F are Banach spaces. If f_n' converges locally uniformly[a] to $g \in C(U, F)$, and $f_n \to f$ pointwise, then $f \in C^1(U, F)$ and $f' = g$.

[a] meaning $\forall x \in U$ there exists $U_x \overset{\circ}{\subseteq} U$, $f_n'|_{U_x} \to g|_{U_x}$ uniformly

Proof. We want to prove, for fixed x_0 and $|y| \leq r$ for r small,

$$|f(x_0 + y) - f(x_0) - g(x_0)(y)| \leq |y|\omega_r.$$

We can also assume $U_x = U$. The usual strategy is to approximate each of the pieces using separately, and using some limiting argument to remove the dependence on n. The uniform convergence of $f_n' \to g$ will be crucial.

Since $f_n \to f$ pointwise, we can write

$$|[f(x_0 + y) - f(x_0)] - [f_n(x_0 + y) - f_n(x_0)]| = \omega_n. \tag{2}$$

Let $l_y^{x_0} = \{x_0 + ty,\ t \in [0,\ 1]\}$ be the line segment between x_0 and $x_0 + y$, then the mean value theorem applied to f_n reads

$$|f_n(x_0 + y) - f_n(x_0) - Df_n(x_0)(y)| \leq |y| \sup_{z\in l_y^{x_0}} \|D_n f(z)\| = |y|\omega_r^n. \tag{3}$$

We can think of Df_n and g being interchangable 'under the limit' as $\|Df_n(x_0) - g(x_0)\|_u = \omega_n$. Combining this with Equations (2) and (3), gives us

$$|f(x_0 + y) - f(x_0) - g(x_0)(y)| \leq \omega_n + |y|\big(\omega_r^n + \omega_n\big) = |y|\omega_r,$$

which is what we set out to prove. ∎

1.10. Power Series

Remark 1.16 Power Series and Radius of Convergence

Given a complex Banach space E, lSee et $\{c_n\}_{n\geq 0} \subseteq \mathbb{C}$, we define the *power series* of c_n about $a \in E$ using the function $f(x) = \lim_n \sum_0^n c_j(x-a)^j$ for all $x \in E$ such that the series converges (absolutely or conditionally).

The nth *partial sum* of f is the E-valued function,

$$f_n(x) = \sum_0^n c_j(x-a)^j.$$

Define $C = \limsup_{n\to\infty} |c_n|^{1/n}$, then $R = C^{-1}$ is the *radius of convergence*. The open set $B(R, a) = \{x \in E,\ |x - a| < r\}$ is known as the *region of convergence*.

Lemma 1.17 Convergence of Power Series

Let $B = B(r, a)$ be as above, then for all $z \in B$, $f(z)$ converges absolutely, and for all $z \in \{z \in \mathbb{C},\ |z - a| > R\}$, $f(z)$ diverges. Moreover, $f \in C^\infty(B, E)$ and for all $k \geq 1$, $D^k f_n$ converges to $D^k f$ uniformly.

Proof. Let $z \in B$, rearranging we see that $\limsup |c_n|^{1/n} < |z|^{-1}$ and $|c_n|^{1/n} < |z|^{-1}$ eventually. (See section 1.3.4.) Taking the nth power termwise and then rearranging, we see that $|c_n||z|^n < 1$ eventually. Therefore $f(z)$ defines a absolutely convergent geometric series for all $z \in B$.

Recall if a series $(x_n) \subseteq \mathbb{C}$ is conditionally convergent iff its partial sums are Cauchy. Suppose $\sum x_n$ converges conditionally, we can write the nth term $x_n = s_n - s_{n-1}$ for $n \geq 2$, and $|x_n| = |s_n - s_{n-1}| \to 0$. If $|z| > R$, then $|c_n||z|^n > 1$ frequently, so that the series cannot be (conditionally) convergent.

Next, we show $f \in C^1(B,E)$ by differentiating f_n

$$f_n'(x)(\cdot) = \sum_0^n c_j \sum_1^j m\Big((x-a)^{i-1},\ \mathrm{id}_E(\cdot),\ (x-a)^{j-\underline{(j-i)}}\Big).$$

Which is bounded above by the following sum of sums (noting that we do not assume commutativity)

$$\|f_n'(x)\|_{L(E)} \le \sum_{j=0}^{n} |c_j||j||x-a|^{j-1}.$$

The results of Section 1.6.3 tell us the the partial sums formed by $\{f_n'\}$ have the same radius of convergence as $\{f_n\}$, because the power series defined by $\{d_j\}_{j\ge1}$ and $\{d_{j-1}\}_{j\ge1}$ have the same radius of convergence for any sequence $\{d_j\} \subseteq \mathbb{C}$.

The Lems. 1.14 and 1.15 and together with the fact that $\|f_n' - g\|_u = 0$ imply that $f \in C^1(B,E)$. Finally, to show $f \in C^\infty(B,E)$, we proceed by induction. The space of operators $L(E,E)$ is a Banach space of its own, so is $L(E,L(E,E)) = L^2(E,E)$, and $L^k(E,E)$ for $k \ge 1$. Suppose $D^{k-1}f_n$ converges uniformly to $D^{k-1}f$, we can rehearse the same argument, with

$$D^k f_n(x) = \sum_{j=0}^{n} c_j\big(j(j-1)\cdots(j-k+1)\big)(x-a)^{j-k}.$$

The formal series $x \mapsto \sum_j c_j d_{j,k}(x-a)^{j-k}$ has radius of convergence R, where

$$d_{j,k} = \big(j(j-1)\cdots(j-k+1)\big), \quad \text{and} \quad 1 \le d_{j,k} \le j^k \quad \text{for large } j.$$

Using results from section 1.6 we see that $\limsup_{j\to\infty} d_{j,k} = 1$ and $D^k f_n$ converges uniformly to some g_k such that $DD^{k-1}f = D^k f = g_k$, and therefore $f \in C^\infty(B,E)$. ■

1.11. Weierstrass Type Approximations

Building on the result in Lemma 1.10, we have the following lemma.

Lemma 1.18

Let $X = [0,\ 1]$, the space of complex-valued polynomials are uniformly dense in $C(X,\mathbb{C}) = C(X)$. More explicitly, for all $f \in C(X)$, and $\varepsilon > 0$, there exists an n-degree polynomial P_n such that $\|f(x) - P_n(x)\|_u \le \varepsilon$. This polynomial can be taken to be

$$P_n(x) = \sum_0^n f(p/n) r_p(x) \quad \text{and} \quad r_p(x) = \binom{n}{p} x^p (1-x)^{n=p}.$$

Proof. We use an anchor point argument.

Covering of X The uniform continuity of f tells us $\sup_{|x-y|\le\delta} |f(x) - f(y)| \le \omega_\delta = \omega_n$ for $n^{-1} \le \delta$. Covering the space X with cells of length n^{-1}, we write

$$X = \cup j = 0^{n-1}[jn^{-1},\ (j+1)n^{-1}].$$

Totally Bounded and Uniform Continuity Although not related to the proof, we notice that

$$\min_{0\le p\le n} |f(x) - f(p/n)| \le \sup_{|x-y|\le n^{-1}} |f(x) - f(y)| \le \omega_n.$$

Alternate covering for $X = [a,\ b]$ We set $X = \cup_0^n a_j + (a_{j+1} - a_j)[0,\ 1]$ and $a_0 = a$, $a_j = a + j(b-a)n^{-1}$. Pick $n \ge 1$ such that $(b-a)n^{-1} < \delta$. Also replace $f(p/n)$ with $f\big(a + p(b-a)n^{-1}\big)$.

Since $\sum_0^n r_p(x) \equiv 1$, for $x \in X$, we can treat $f(x)$ as a p-independent coefficient and estimate the difference of f and the pth sample $|f(x) - f(p/n)|$:

$$|f(x) - \textstyle\sum_0^n f(pn^{-1}) r_p(x)| = |\sum_0^n (f(x) - f(pn^{-1})) r_p(x)|.$$

To simplify the proof We can separate the sum into two pieces,

$$|\textstyle\sum_0^n (f(x) - f(\frac{p}{n})) r_p(x)| \le \left|\Big(\displaystyle\sum_{|x-\frac{p}{n}|\le\frac{1}{n}} + \sum_{|x-\frac{p}{n}|>\frac{1}{n}}\Big)(f(x) - f(\tfrac{p}{n})) r_p(x)\right|$$

Case 1 By uniform continuity, $\sum_{|x-\frac{p}{n}|\le\frac{1}{n}} |f(x) - f(p/n)||r_p(x)| \le \omega_n$, as $r_p(x) \ge 0$ for all $x \in X$.

Case 2 We can use Holder's inequality again, the difference is estimated by $\max_{0\le p\le n} |f(x) - f(p/n)| \le 2\|f\|_u$. So we only have to control the l^1 norm of

$$\sum_{|x-p/n|>1/n} |r_p(x)| = \sum_{|x-p/n|>1/n} |r_p(x)||x - \tfrac{p}{n}|^2 |x - p/n|^{-2}$$

$$\le \Big[\sum_{|x-p/n|>1/n} |r_p(x)||x - p/n|^2\Big]\Big[\sup_{\substack{|x-p/n|^2 > n^{-2},\\ 0\le p\le n}} |x - p/n|^{-2}\Big]$$

$$\le nx(1-x)n^{-2} = n^{-1}x(1-x).$$

Which is controlled by $\omega_n \sup|x(1-x)|$. And

$$|\sum_{|x-p/n|>1/n} f((x - f(p/n)) r_p(x)| \le 2\|f\|_u \omega_n \sup|x(1-x)| = \omega_n.$$

Combining Cases 1 and 2 reads

$$\textstyle\sum_0^n |f(x) - f(\frac{p}{n}) r_p(x)| \le \omega_n + \omega_n = \omega_n.$$

■

Note 1.19

Characteristics of Weierstrass-Type approximations:

convex sublevel sets satisfy uniform bound
partition of unity require topological argument and local finiteness

2. Approximations from a Generating Set

A lot of measure-theoretical constructs make use of the supremum or infimum of a functional defined on a smaller subset, and extending it to the entire space by sup/inf.

We construct an abstract approximation framework. Let Ω be an arbitrary set and $\Sigma \subseteq \Omega$ be a subset on which we define $J : \Sigma \to [0, +\infty]$.

Suppose further that, for every $x \in \Omega$, we associate to this x a non-empty subset that 'approximates' our final functional I, denoted by $\Sigma_x \subseteq \Sigma$. We then define the *upper and lower approximations* to J,

$$I^+(x) = \sup_{z \in \Sigma_x} J(x) \quad \text{and} \quad I^-(x) = \inf_{z \in \Sigma_z} J(z)$$

If we want our functional $I^\pm : \Omega \to [0, +\infty]$ to be *consistent* with the old functional J, we usually demand the following: (just to make life easy)

- For every $z \in \Sigma$, $z \in \Sigma_z$, and
- $J(z) = I^+(z) = I^-(z)$ for all $z \in \Sigma$, this makes $I^\pm$ an extension of J.

We can introduce an order relation on Ω, as *measured by* the functional $I = I^\pm$. Given $x, y \in \Omega$, we say that x is *less than* y as measured by I whenever $I(x) \leq I(y)$. This is written

$$x \lesssim_I y \quad \text{or} \quad x \leq y \text{ (meas. } I)$$

2.1. Proof Techniques

Technique 1 To show $x \leq y$ (meas. I), it is sufficient to show $\Sigma_x \subseteq \Sigma_y$ if $I = I^+$ (or $\Sigma_y \subseteq \Sigma_x$ whenever $I = I^-$).

Let $x \in \Omega$ be arbitrary, we wish to bound $I(x)$ by c (we agree that $c \geq 0$ means $c \in [0, +\infty]$ and 'bounding' means from above unless specified otherwise).

$$I^+(x) \leq c \quad \text{or} \quad I^-(x) \leq c.$$

Technique 2 On one hand, if $I = I^+$,

$$I^+(x) \leq c \quad \text{is implied by} \quad \Sigma_x \leq c \text{ (meas. } I),$$

or $\Sigma_x \subseteq \{I \leq c\}$. On the other hand,

$$I^-(x) \leq c \quad \text{is implied by} \quad \phi \geq c \text{ (meas. } I),\ \exists \phi \in \Sigma_x,$$

or equivalently: $\Sigma_x \cap \{I \geq c\} \neq \varnothing$.

Apply to Theorems 1.10 and 1.13 using this framework.

Technique 3 The most important technique by far is to control the terms leading up to the supremum/infimum.

Recall, if $A \subseteq \overline{\mathbb{R}}$ is a non-empty subset, and $s^* = \sup(A)$, $s_* = \inf(A)$, there exists increasing (resp. decreasing) sequences in A that converge to s^* (resp. s_*). From Rudin Theorem 3.19, let $(a_n), (b_n), (c_n)$ be sequences in $\overline{\mathbb{R}}$, where $b_n \to b \in \overline{\mathbb{R}}$, and for sufficiently large N,

$$a_n \leq b_n \leq c_n \quad \forall n \geq N,$$

then the limit of b_n can be controlled by the subsequential limits of (a_n) and (c_n), that is:

$$\limsup_{n \to \infty} a_n \leq b \leq \liminf_{n \to \infty} c_n.$$

Since $I^+(x)$ is defined by a supremum, it induces a sequence $(\phi_n) \subseteq \Sigma_x$ that increases to $I^+(x)$ (meas. I). To obtain a bound for $I^+(x)$, it therefore suffices to estimate $(J(\phi_n))_n$ using another sequence (c_n). We end up with the following condition.

Let $I = I^+$ and $x \in \Omega$, let $(\phi_n) \subseteq \Sigma_x$ increase to $I(x)$ (meas. I), suppose further that there exists a sequence $(\psi_n) \subseteq \Omega$, where

$$\phi_n \leq \psi_n \text{ (meas. } I) \quad \forall n \geq N$$

and that $\liminf_{n \to \infty} I(\psi_n) \leq c$. Then,

$$I(x) \leq \liminf_{n \to \infty} I(\psi_n) \leq c$$

This is essentially what Fatou's Lemma is saying.

2.2. Estimating Functionals

Let $\Sigma^{(1)}$ and $\Sigma^{(2)}$ be generating sets on Ω and $I^\pm_{(1)}$, $I^\pm_{(2)}$ correspondingly. Suppose we are given some $x \in X$ where for all $\phi \in \Sigma^{(1)}_x$, $\Sigma^{(2)}_x$ satisfies the *second property of the infimum* with respect to ϕ (meas. I), meaning

> There exists some sequence $(\psi_n) \subseteq \Sigma^{(2)}_x$, $\limsup_{n \to \infty} I(\psi_n) \leq I(\phi)$, or equivalently: for every ε there exists $\psi_\varepsilon \in \Sigma^{(2)}_x$ such that $\psi_\varepsilon \leq \phi + \varepsilon$ (meas. I).

If this is the case, then $I^-_{(2)}(x) \leq I^-_{(1)}(x)$.

A lot of the times we want to estimate the sizes of generated functionals. When is $I^-_{(1)} \leq I^+_{(2)}$?

Note 2.1 Approximation Framework: Overlapping Condition

Let $\Sigma^{(j)}$ be generating sets on Ω and $J_{(j)} : \Sigma^{(j)} \to \overline{\mathbb{R}}$ be functionals defined. Suppose that for every $x \in \Omega$, the two generating functionals satisfy an 'overlapping' condition:

$$J_{(1)}(\Sigma_x^{(1)}) \cap J_{(2)}(\Sigma_x^{(2)}) \neq \varnothing,$$

then

$$I_{(1)}^-(x) \le I_{(2)}^+(x) \quad \text{and} \quad I_{(2)}^-(x) \le I_{(1)}^+(x).$$

The above can be weakened to this 'cofinality' argument.

Note 2.2 Approximation Framework: Cofinality Argument

Suppose there exists $\phi_k \in \Sigma_{(k)}(x)$ $(k = 0,1)$such that $J_0(\phi_0) \le J_1(\phi_1)$, then

$$I_0^-(x) \le I_{(1)}^+(x)$$

Proposition 2.3

Let X be a set and $\{E_\alpha\}_{\alpha\in A}$ and $\{F_\beta\}_{\beta\in B}$ be two families of subsets of X, denote $E = \cup E_\alpha$ and $F = \cap F_\beta$, then

$$\inf_{z\in E} I(z) = \inf_{\alpha\in A} \inf_{z\in E_\alpha} I(z)$$

$$\inf_{z\in F} I(z) \ge \sup_{\beta\in B} \inf_{z\in F_\beta} I(z).$$

Proof. The whole proof is really tedious but I still want to write it out in full. Let $x, y \in E, F$, then for each α, β,
Every $x \in E$ belongs to some E_α, so

$$I(x) \ge \inf_{z\in E_\alpha} I(z) \le \inf_{\alpha\in A} \inf_{z\in E_\alpha} I(z).$$

This holds for every x, so $\inf_{x\in E} I(x) \ge \inf_\alpha \inf_{E_\alpha} I(z)$.

Every $y \in F$ belongs to every F_β, so

$$I(y) \ge \inf_{z\in F_\beta} I(z) \quad \forall \beta \in B,$$

so that $\inf_{y\in F} I(y) \ge \sup_{\beta\in B} \inf_{z\in F_\beta} I(z)$. Now fix $\alpha, \beta \in A, B$, since

$$E_\alpha \subseteq E \quad \text{and} \quad F \subseteq F_\beta,$$

we see that $\inf_{x\in E} I(x) \le \inf_{z\in E_\alpha} I(z)$. ■

We outline conditions for which the second equality holds.

2.3. Estimating Sub-sigma Algebras, Probability

Let $(X, \mathcal{M}, \mu)$ be a probability space, and if $X = A_1 \oplus A_2$, and $(A_j,\ \mathcal{M}_j,\ \mu_j)$ $j = 1,2$ where $\mathcal{M}_j = \{E \cap A_j,\ E \in \mathcal{M}\}$, and $\mu_j(E) = \mu(E \cap A_j)$. The inclusion $\iota_j : A_j \to X$ is always measurable, and $\mathcal{M}_j$ is the σ-algebra on A_j generated by ι_j.

The second scenario is when we keep the set X fixed, but we shrink the σ-algebra on X. Let $\mathcal{M}' \subseteq \mathcal{M}$ be another σ-algebra over X. If $\varphi = \mathrm{id}_X$, then it is $(\mathcal{M},\ \mathcal{M}')$ measurable, because $\varphi^{-1}(E') \in \mathcal{M}$ for all $E' \in \mathcal{M}'$.

This gives us a mapping that pushes measurable function on a smaller σ-algebra, to a larger one. Let $\varphi^* : L^+(X, \mathcal{M}') \to L^+(X, \mathcal{M}, \mu)$. We can also induce a measure on $(X,\ \mathcal{M}')$. Let $\mu' : \mathcal{M}' \to [0,\ 1]$ be the restriction of μ, i.e: $\mu' = \mu|_{\mathcal{M}'}$.

If $f \in L^1(X,\ \mathcal{M}',\ \mu')$, then $f \circ \varphi \in L^1(X,\ \mathcal{M},\ \mu)$.

Note 2.4

Convergence in measure looks like the same as 'vanish at infinity' for locally compact hausdorff spaces.

Note 2.5

can we check if something integrates to zero by the dual.

3. Measure Theory

Monotone Classes Some algebra structure (max, sup) containing the simple functions must generate L^+. This is used in Radon-Nikodym and Fubini's Theorem.

Proposition 3.1 Equivariant Measures

Let $(X,\ \mathcal{M},\ \mu)$, and $(Y,\ \mathcal{N},\ \eta)$ be measure spaces, suppose $\varphi : Y \to X$ is a measurable mapping that intertwines the two measures, i.e: $\mu(E) = \eta(\varphi^{-1}(E))$ for all $E \in \mathcal{M}$. Then for all $f \in L^+(X)$, $g \in L^1(X)$, $g \circ \varphi \in L^1(Y)$,

$$\int_X f(x)dx = \int_Y f \circ \varphi(y)dy$$

and

$$\int_X g(x)dx = \int_Y g \circ \varphi(y)dy.$$

3.1. h-intervals

We define the following subsets of $\mathbb{P}(\mathbb{R})$,

$$\mathcal{H}^- = \{(d,\ +\infty),\ -\infty < d\}$$
$$\mathcal{H}^+ = \{(-\infty,\ c],\ c < +\infty\}$$
$$\mathcal{H}_b = \{(a,b],\ -\infty < a < b < +\infty\}.$$

In what follows, the plus sign will always refer to the disjoint union of sets and not a coset. Let $\mathcal{H} = \mathcal{H}^- + \mathcal{H}^+\mathcal{H}_b + \{\varnothing\}$.

3.1.1 Outline for 1.15

Properties of F The following limits converge in $[0, +\infty]$.

$$\mu_0((a,b]) = F(b) - F(a) = \inf_{a^+>a}\inf_{b^+>b}(F(b^+) - F(a^+))$$
$$\mu_0((-\infty, c]) = \sup_{q\in\mathbb{Q}}(F(c) - F(q)) = \sup_{q\in\mathbb{Q}}\inf_{c^+>c}(F(c^+) - F(q))$$
$$\mu_0((d, +\infty)) = \sup_{q\in\mathbb{Q}}(F(q) - F(d)) = \sup_{q\in\mathbb{Q}}\inf_{d^+>d}(F(q) - F(d^+))$$

Abstracting Union Behaviour Let X be a space and $\Sigma, \overline{\Sigma} \subseteq \mathbb{P}(X)$.

An element $F \subseteq X$ can be written as the union of members in Σ iff

1. There exists $\{E_\alpha\} \subseteq \Sigma$, where $F \subseteq \bigcup E_\alpha$,
2. For this Σ-cover of F, for all $\alpha \in A$, $F \cap E_\alpha \in \Sigma$.

Applying this to $F \in \Sigma$ we get the following criterion.

> Every element in $\overline{\Sigma}$ is the union of elements in Σ, iff every $F \in \overline{\Sigma}$ corresponds to $\{E_\alpha\} \subseteq \Sigma$, $F \subseteq \bigcup E_\alpha$, and $F \cap E_\alpha \in \Sigma$ for all α.

We have the following.

> Let X be a set and $\Sigma, \overline{\Sigma} \subseteq \mathbb{P}(X)$. Suppose that for every pair $E \in \Sigma$, and $F \in \overline{\Sigma}$, $E \cap F \notin \Sigma$. Then no elements of Σ can be involved in the union of an element in $\overline{\Sigma}$.

Take $h^- = (c, +\infty)$ and $A \subseteq \mathbb{R}$ be bounded above, their intersection never an element in $\mathcal{H}^-$, and from this we conclude.

> For every $A = FDU(\mathcal{H}) = FDU(\mathcal{H}^-) + FDU(\mathcal{H}^+) + FDU(\mathcal{H}_b) + \varnothing$, if A is bounded below (resp. above), then the d.u. over $\mathcal{H}^+$ (resp. $\mathcal{H}^-$) is empty.

Abstracting Disjoint Union Behaviour Suppose for all $E_1, E_2 \in \Sigma$, that $E_1 \cap E_2 \neq \varnothing$, then every FDU consisting of elements in Σ contains at most one element of Σ.

We can apply this immediately to the unbounded rays $\mathcal{H}^\mp$, as:

$$(c, +\infty) \cap (c', +\infty) = (\max(c, c'), +\infty) \quad \text{and} \quad (-\infty, d] \cap (-\infty, d'] = (-\infty, \min(d, d')]$$

are both non-empty. Using the fact that the finite union of bounded above (resp. below) subsets is again bounded above (resp. below), we state the following.

> If $A = FDU(\mathcal{H})$, the d.u. contains *at most one* from each $\mathcal{H}^-$, $\mathcal{H}^+$. If A is unbounded above (resp. below) then the $\mathcal{H}^-$ d.u. (resp. $\mathcal{H}^+$) consists of *exactly one* element.

Ordering $\mathcal{H}_b$ intervals Let $E_1 = (a_1, b_1]$, and $E_2 = (a_2, b_2]$ be elements in $\mathcal{H}_b$, so that $a_1 < b_1$ and $a_2 < b_2$. It is easy to see that the two intervals overlap iff $\max(a_1, a_2) < \min(b_1, b_2)$, iff exactly one of the following is true

1. $\operatorname{argmax}(a_i) = \operatorname{argmin}(b_i)$, and $\max(a_i) < \min(b_i)$
2. $\operatorname{argmax}(a_i) \neq \operatorname{argmin}(b_i)$.

The second condition does not require $\max(a_i) < \min(b_i)$ to hold, without loss of generality suppose $i = 1$ is the argmax and argmin of the pairs of endpoints. Then $a_2 \leq a_1 < b_1 \leq b_2$. This is the same as saying $E_1 \subseteq E_2$. In either case, the intersection is the interval

$$E_1 \cap E_2 = (\max(a_1, a_2), \min(b_1, b_2)].$$

We shall give an consequence of the overlapping criterion of the second type.

> **Lemma 3.2 Properties of d.u. in hb**
>
> Let E_1 and E_2 be as above, and suppose $E_1 \cap E_2 = \varnothing$, then
>
> 1. $a_1 \neq a_2$, $b_1 \neq b_2$,
> 2. $a_1 < a_2$ iff $b_1 < b_2$.
>
> If $\{E_j\}_1^n$ is a disjoint sequence in $\mathcal{H}_b$, then the left (resp. right) endpoints are unique, a strict well-ordering exists for the left and right endpoints and are equal.
>
> If $E = (a, b] = FDU((a_\tau, b_\tau]) \in \mathcal{H}_b$, when the intervals are ordered by their endpoints, $a_1 = a$, $b_n = b$, and $a_{j+1} = b_j$ for $j = 1, \dots, n-1$.

Proof. Let $\overline{\varepsilon} > 0$ be so small that $a_1 + \varepsilon, b_1 - \varepsilon \in E_1$, $a_2 + \varepsilon, b_2 - \varepsilon \in E_2$ for all $\varepsilon \in (0, \overline{\varepsilon})$. If $a_1 = a_2$, then $a_1 + \varepsilon = a_2 + \varepsilon \in E_1 \cap E_2$, similarly for $b_1 = b_2$. For the next claim, suppose $a_1 < a_2$ and $b_2 \leq b_1$, then $\operatorname{argmax}(a_i) = \operatorname{argmin}(b_i) = 2$, so that $E_2 \subseteq E_1$ and the two intervals cannot be disjoint. Repeat the same argument for 'only if' direction.

Next, let $E = (a, b]$ be the d.u. of elements in $\mathcal{H}_b$ as in the lemma. Suppose $a_{j+1} \neq b_j$, we have two cases to consider. If, $b_j < a_{j+1}$, the points in $(b_j, a_{j+1}]$ are not in any interval (because if so this would give us a contradiction that the left end-points are well-ordered). And if $a_{j+1} < b_j$, there exists a point x in the intersection: $a_j < a_{j+1} < x \leq b_j < b_{j+1}$, so the intervals are not disjoint.

Finally, consider $a = \inf E = \inf_{1\leq j\leq n} \inf((a_j, b_j])$, $b = \sup E = \sup_{1\leq j\leq n} \sup((a_j, b_j])$ and the fact that the endpoints are ordered. ■

Let $A = FDU(\mathcal{H})$, we define the positive and negative unbounded components of A to be the intervals $A^{++} =$

$(c, +\infty)$, $A^{--} = (-\infty, d]$, where $c = \inf\{c' \in \mathbb{R}, (c', +\infty) \subseteq A\}$ and $d = \sup\{d' \in \mathbb{R}, (-\infty, d'] \subseteq A\}$.

Properties of d.u. of $\mathcal{H}$

> **Remark 3.3**
> F is increasing and right-continuous, so for all $x \in \mathbb{R}$, $F(x) = \inf_{x^+ \geq x} F(x^+) = \inf_{x^+ > x} F(x^+)$.

3.2. Convergence Theorems

MCT is proven in the following `.pdf` document. We state Fatou's Lemma for convenience: if $(f_n) \subseteq L^+$, then

$$\int_X \liminf_{n\to\infty} f_n(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \int_X f_n(x)\,\mathrm{d}x.$$

To discuss integration on L^1, we need a slight modification of Fatou's Lemma: if $(f_n) \subseteq L^1 \cap L^+$, then the same conclusion holds. The proof is straightforward, because the integral on $L^1 \cap L^+$ is the same (gives the same number) as the integral on L^+.

Indeed, if $h \in L^1 \cap L^+$, the integral on L^1 is the number $\int_X h^+(x)\,\mathrm{d}x - \int_X h^-(x)\,\mathrm{d}x$ (with the imaginary part suppressed), which is equal to $\int_X h^+(x)\,\mathrm{d}x$, the same number as the integral of h on L^+.

> **Proposition 3.4**
>
> **Ex. 2.18** Fatou's Lemma still holds if f_n measurable, $f_n \geq -g$ for some $g \in L^+ \cap L^1$. What is the analogue of Fatou's lemma for non-positive functions?
>
> **Ex. 2.19** Suppose $\{f_n\} \subseteq L^1(\mu)$ and $f_n \to f$ uniformly.
> a) If $\mu(X) < +\infty$, then $f \in L^1(\mu)$ and $\int f_n \to \inf f$.
> b) If $\mu(X) = +\infty$, the conclusions of a) can fail.
>
> **Ex. 2.20 (Generalized DCT)** Let $(f_n), (g_n), g \in L^1$, $f_n \to f$, $g_n \to g$ a.e, $\int_X |g_n|\,\mathrm{d}x \to \int_X |g|\,\mathrm{d}x$, and
> $$|f_n| \leq |g_n| \quad \text{pw a.e.},$$
> then, $\int_X f_n(x)\,\mathrm{d}x \to \int_X f(x)\,\mathrm{d}x$.
>
> **Ex. 2.21** Suppose $f_n, f \in L^1$ and $f_n \to f$ a.e. Then $\int |f_n - f|\,d\mu \to 0$ iff $\int |f_n|\,d\mu \to \int |f|\,d\mu$. (Use Exercise 20.)

Proof. Ex. 2.20) Unless the measure is complete, f is not necessarily measurable, but we can always identify f with a measurable function that agrees with f a.e. To show that $f \in L^1$, by taking limits on both sides

$$|f| \leq |g| \quad \text{pw a.e.} \quad \text{means} \quad \int_X |f|\,\mathrm{d}x \leq \int_X |g|\,\mathrm{d}x.$$

Assuming that f_n is real valued for the moment, our goal is to show that

$$\pm \int_X^{L^1} f(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \left[\pm \int_X^{L^1} f_n(x)\,\mathrm{d}x\right].$$

It is clear that $|g_n| \pm f_n \geq 0$ pw a.e., so that we can use Fatou's Lemma for $L^1 \cap L^+$ functions, denoting the integral on L^1 with a superscript:

$$\int_X^{L^1} \liminf_{n\to\infty} \left[|g_n(x)| \pm f_n(x)\right]\mathrm{d}x \leq \liminf_{n\to\infty} \int_X^{L^1} |g_n(x)| \pm f_n(x)\,\mathrm{d}x$$

The integrand on the left hand side converges pointwise a.e to $|g_n(x)| \pm f(x)$. Note: If the integrals involved are L^+ integrals, we cannot subtract $|g_n(x)|$ on both sides, as $\pm f_n(x)$ may not be non-negative; however we have fixed it by considering an L^1 version of Fatou's Lemma.

$$\pm \int_X^{L^1} f(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \left[\int_X^{L^1} |g_n(x)|\,\mathrm{d}x - \int_X^{L^1} |g(x)|\,\mathrm{d}x \pm \int_X^{L^1} f_n(x)\,\mathrm{d}x\right]$$

the difference of integrals within the limit is controlled by the error

$$\left|\int_X^{L^1} |g_n(x)|\,\mathrm{d}x - \int_X^{L^1} |g(x)|\,\mathrm{d}x\right| \to 0,$$

and this proves the statement for real f_n. If f_n is complex valued, we recall that:

- A sequence of complex numbers $(z_n) \subseteq \mathbb{C}$ converges to $z \in \mathbb{C}$ iff $\operatorname{Re} z_n \to \operatorname{Re} z$ and $\operatorname{Im} z_n \to \operatorname{Im} z$.
- The real (resp. complex) part of the integral of an L^1 function f is the L^1 integral of the real (resp. complex part) of f.

The hypothesis of the real DCT is satisfied for $\operatorname{Re} f$ (resp. $\operatorname{Im} f$), so that $\operatorname{Re} \int_X^{L^1} f_n(x)\,\mathrm{d}x \to \operatorname{Re} \int_X^{L^1} f(x)\,\mathrm{d}x$ (resp. imaginary part), and this proves the claim for complex f_n.

Ex. 2.21) Let $f_n \to f$ a.e., and $\|*\| f_{n_1} \to \|*\| f_1$. Then, for all $n \geq 1$, $|f_n - f| \leq |f_n| + |f|$. The right hand side of the previous estimate $\to 2|f|$ a.e., with $\int |f_n| + |f| \to 2\int |f|$ so that $\int |f_n - f| \to 0$ by Ex. 2.20. Conversely, if $\int |f_n - f| \to 0$, for almost everywhere x we have $||f_n(x)| - |f(x)|| \leq |f_n(x) - f(x)| \in L^+$. Using the L^1 inequality:

$$\left|\int |f_n| d\mu - \int |f| d\mu\right| \leq \int ||f_n| - |f|| d\mu$$

which is bounded above by $\liminf \int |f_n - f| = 0$. ∎

3.3. Convergence in Measure

> **Definition 3.5**
> Let f_n, f be complex-valued measurable functions on X.
>
> **convergence in meas.** for all $\varepsilon > 0$, $\mu(E(|f_n - f|, \varepsilon)) \to 0$, where $E(\phi, \varepsilon) = \{x \in X, |\phi(x)| \geq \varepsilon\}$
> **cauchy in meas.** for all $\varepsilon > 0$, $\mu(E(|f_n - f_m|, \varepsilon)) \to 0$ as

$m, n \to \infty$.

Proposition 3.6

Folland 2.30a If $f_n \to f$ in measure, there exists some subsequence $f_{n_j} \to f$ a.e.

Folland 2.30b If $\{f_n\}_1^\infty$ is Cauchy in measure, there exists a measurable f, $f_n \to f$ in measure; and a subsequence $f_{n_k} \to f$ a.e. If also $f_n \to g$ in measure, then $f = g$ a.e.[a]

Folland 2.29 If $f_n, f \in L^p$, $p \in [1, +\infty)$, $\|*\|f_n - f_p \to 0$, then $f_n \to f$ in measure.

Folland 2.31 If $f_n, f \in L^1$, $\|*\|f_n - f_1 \to 0$, then there is a subsequence $f_{n_j} \to f$ a.e.

[a] this f is the 'best possible' representative of the cauchy in measure sequence

Proof. 2.30a) Let $f_n \to f$ in measure, $\{c_j\}, \{d_j\} \in l^1 \cap l^{++}$. Pick an increasing sequence of numbers $\{n_j\}$ such that

$$\sup_{m \geq n_j} \mu(E(|f_m - f|,\ d_j)) \leq c_j, \quad E(g,\ \varepsilon) = \{x \in X,\ |g(x)| \geq \varepsilon\}.$$

Set $A_j = E(|f_{n_j} - f|,\ d_j)$, so that $\{\mu(A_j)\} \in l^1$ and $E = (\limsup A_j)^c$ has μ-null complement. For any $x \in E$, $x \notin A_j$ eventually, meaning there exists $j_x \in \mathbb{N}^+$ such that for all $j \geq j_x$,

$$\sup_{j \geq j_x} |f_{n_j}(x) - f(x)| \leq d_{j_x}.$$

So that $f_{n_j} \to f$ a.e.

2.30b) See book.

2.29) Modify the proof in the book (I don't know if this is true yet for $p > 1$)

2.31) Use 2.29 with $p = 1$, and 2.30a. ■

Proposition 3.7

Ex. 2.33 Let $f_n, f \in L^+$, and $f_n \to f$ in measure, then $\int f d\mu \leq \liminf \int f_n d\mu$.

Ex. 2.34 Let $f_n, g \in L^1$, and $|f_n| \leq g$, if $f_n \to f$ in measure, then $f \in L^1$ and $\int f_n d\mu \to \int f d\mu$.

Ex. 2.34 Boosted? Let $f_n, g_n, g \in L^1$, $g_n \to g$ a.e., $f_n \to f$ in measure, $|f_n| \leq g_n$ a.e., and $\int g_n d\mu \to \int g d\mu$. Then, $f \in L^1$, and $\int f_n d\mu \to \int f d\mu$.[a]

[a] I dont know if this is true.

Proof. Ex. 2.33) If $\liminf \int f_n d\mu = +\infty$, there is nothing to prove. Otherwise, there exists a subsequence $f_k \to f$ in measure, such that $\liminf \int f_k d\mu = \liminf \int f_n d\mu$. By Proposition 3.6, there exists a subsequence g_j such that $g_j \to f$ a.e., and by Fatou's Lemma

$$\int f d\mu \leq \liminf \int g_j d\mu = \liminf \int f_n d\mu.$$

Ex. 2.34) f is in L^1, because we can pick some subsequence $f_{n_j} \to f$ a.e., and $\int |f| d\mu \leq \liminf \int |f_{n_j}| d\mu \leq \liminf \int g_{n_j} d\mu \leq \int g d\mu$. Since $\int |g_n| \to \int |g|$, $\|*\|g_n - g_1 \to 0$ and $g_n \to g$ in measure by Folland 2.30a, so that

$$\int g + f d\mu \leq \liminf \int g_n + f_n d\mu$$

Suppose f_n is real valued, because $0 \leq f_n + g \to f + g$ in measure,

$$\int f + g d\mu \leq \left(\liminf \int f_n d\mu\right) + \int g d\mu.$$

Subtracting $\int g d\mu$ from both sides, and repeating the same argument with $g - f_n$,

$$\int f_n d\mu \to \int f d\mu.$$

This proves the case for real-valued f_n. If f_n is complex valued, and $f_n \to f$ in measure, then $\operatorname{Re} f_n \to \operatorname{Re} f$, $\operatorname{Im} f_n \to \operatorname{Im} f$ in measure. The previous argument on the real and imaginary parts proves the claim. ■

3.4. Signed Measures

Let $(X,\ \mathcal{M})$ be measurable space, a signed measure on $(X,\ \mathcal{M})$ is a set function $\nu : \mathcal{M} \to \overline{\mathbb{R}}$ that satisfies the three properties laid out in the book. There is a lot to be said about how properties 2 and 3 interact with each other.

3.4.1 Rearrangements

If $x \in \overline{\mathbb{R}}$, set $x^+ = \max(x,\ 0)$ and $x^- = \max(-x,\ 0)$. Fix $\{x_n\} \subseteq \mathbb{R}$, the sum $s = \sum_1^\infty x_n$ is *independent of rearrangements* if and only if at most one of $s^\pm = \sum_1^\infty x_n^\pm$ is unbounded. This is a consequence of Riemann's rearrangement theorem. Without loss of generality, if the 'positive' part is unbounded, and the negative part is bounded, then every rearrangement (grouping) must converge to $+\infty$. If both parts are bounded, then the series converges absolutely. This, by definition forces the equation $\nu(\cup E_j) = \sum \nu(E_j)$ to be independent of how you choose to order the union and sum. We also see that, if $\nu(A) = \pm\infty$, and $B \supseteq A$, measurable, then $\nu(B) = \nu(A)$. This means the measure $\nu(X)$ allows us to detect which part of ν (the positive or the negative) is unbounded.

3.4.2 Kernel of a Signed Measure

Definition 3.8

Let ν be a signed measure, for any $E \in \mathcal{M}$, we write $\eta_E(A) = \nu(E \cap A)$ with the same domain as ν.

$$\mathcal{M}_\nu^\pm = \{E \in \mathcal{M},\ \pm\eta_E \text{ is a positive measure on } \mathcal{M}\}$$

The elements of $\mathcal{M}_\nu^\pm$ are called the positive (resp. negative) subsets of ν. The kernel of ν is the collection

$$\text{Ker}(\nu) = \{E \in \mathcal{M},\ \eta_E \text{ is the zero measure on } \mathcal{M}\}.$$

Verify that the two definitions are the same as in the book.

Proposition 3.9

Let ν be a signed measure on $(X, \mathcal{M})$, then

- If E is in $\mathcal{M}_\nu^\pm$, so is any measurable subset of E.
- If $\{E_j\}_1^\infty \subseteq \mathcal{M}_\nu^\pm$, then $\cup E_j \in \mathcal{M}_\nu^\pm$.

Proof. Folland Proposition 3.2. ∎

The kernel of a signed measure ν is equal to $\mathcal{M}_\nu^+ \cap \mathcal{M}_\nu^-$. Then the two statements in Proposition 3.9 also hold for Ker(ν). This discussion will prove to be rather important so we restate:

Remark 3.10

Let ν be a signed measure, if $E \in \text{Ker}(\nu)$, so is any measurable subset of E. If $\{E_j\}_1^\infty \subseteq \text{Ker}(\nu)$, then $\cup E_j \in \text{Ker}(\nu)$.

We turn to mutual singularity and absolute continuity. The former is refers to the ability to separate one signed measure from another and should be interpreted as some a 'Hausdorff' condition. Absolute continuity on the other hand, has to do with 'where' the measure is supported. We state our two of our new and improved definitions at once.

Definition 3.11

Let ν, μ be signed measures,

- $\nu \perp \mu$ iff there exists $A \in \text{Ker}(\nu)$ such that A^c is in Ker(μ),
- $\nu \ll \mu$ iff $\text{Ker}(\mu) \subseteq \text{Ker}(\nu)$.

Verify that the two definitions are the same as the book.

It is useful to remember that if ν is a positive measure, then $\nu(E) = 0$ iff $E \in \text{Ker}(\nu)$.

Proposition 3.12 Exercise 3.9 Boosted

Let $\{\nu_j\}_1^\infty$ be a sequence of positive measures, and $\nu = \sum_1^\infty \nu_j$, then $\bigcap_1^\infty \text{Ker}(\nu_j) = \text{Ker}(\nu)$. Moreover,

- $\nu \ll \mu$ iff $\nu_j \ll \mu$ for all j,
- $\nu \perp \mu$ iff $\nu_j \perp \mu$ for all j.

Proof. We first prove $\bigcap_1^\infty \text{Ker}(\nu_j) = \text{Ker}(\nu)$. Fix $E \in \text{Ker}(\nu)$, because $\nu_j \ll \nu$ for $j \ge 1$, $E \in \text{Ker}(\nu_j)$. Conversely, if E is an element in the intersection, then $\nu_j(E) = 0$ for all j. By taking the supremum over finite subcollections of indices, we see that $\nu(E) = 0$ and $E \in \text{Ker}(\nu)$.

The second claim is an easy corollary of the first, because the subset condition is passed onto the individual members in the intersection. For the third claim, one direction is immediate: if $\nu \perp \mu$, then we find some $A \in \text{Ker}(\mu)$ such that $A^c \in \bigcap_1^\infty \text{Ker}(\nu_j)$. The reverse direction is a bit tricky, and involves constructing a measurable subset on which ν vanishes.

For every $j \ge 1$, let $B_j \in \text{Ker}(\nu_j)$ such that $B_j^c \in \text{Ker}(\mu)$. By Remark 3.10, their intersection $B = \bigcap B_j$ must be an element of Ker(ν); and by using Remark 3.10 again, we see that its complement, $B^c = \bigcup B_j^c$ is in Ker(μ) therefore $\nu \perp \mu$. ∎

The following is almost trivial but goes to show how useful this new perspective is.

Proposition 3.13 Exercise 3.2, 3.8 Boosted

Let ν be a signed measure, then $\text{Ker}(\nu) = \text{Ker}(|\nu|)$. If μ is another signed measure, then

- $\nu \ll \mu$ iff $|\nu| \ll |\mu|$ iff $\nu^\pm \ll \mu$,
- $\nu \perp \mu$ iff $|\nu| \perp |\mu|$ iff $\nu^\pm \perp \mu$,

If ν, μ_1, μ_2 are finite signed measure s, $\mu = \mu_1 + \mu_2$. Then $\nu \perp \mu_j$ $(j = 1, 2)$ implies $\nu \perp \mu$. The converse need not be true (take $\mu_2 = -\mu_1$); but is true if $\mu_1 \perp \mu_2$.

Proof. The first claim is immediate because of the remark preceding Proposition 3.12. Since $\nu \ll \mu$ and $\nu \perp \mu$ are characterized using the kernels of the two signed measures, the first equivalence is clear, and the second equivalence follows from Proposition 3.12. ∎

3.5. Complex Measures

Definition 3.14

$\mu : \mathcal{M} \to \mathbb{C}$ is a complex measure if $\mu(\varnothing) = 0$ and $\mu(\cup E_j) = \sum \mu(E_j)$ for disjoint $\{E_j\}$, where the sum converges absolutely.

3.5.1 Decomposition of Complex Measures

Let $\varphi_j : \mathbb{C} \to [0, +\infty)$ with $\varphi_1(a+bi) = a^+$, $\varphi_2(a+bi) = a^-$, $\varphi_3 = b^+$, $\varphi_4 = b^-$ where a, b are real numbers and

$x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$ are the positive and negative parts of a real number x.

For $j = 1, \dots, 4$, $\varphi_j(0) = 0$, so if $f : X \to \mathbb{C}$ is any measurable (resp. continuous) function, and $S_f = \{x \in X,\ f(x) \neq 0\}$, because $f(x) = 0$ implies $\varphi_j \circ f(x) = 0$, so that

$$S_{\varphi_j \circ f} \subseteq S_f.$$

Moreover, $\varphi_j \circ f$ is measurable (resp. continuous), and if f is continuous and has compact support, so does $\varphi_j \circ f$ for all f.

Also, if $\psi : [0, +\infty)^4 \to \mathbb{C}$ is the recombination mapping, meaning $\psi(a_1, a_2, b_1, b_2) = (a_1 - a_2) + i(b_1 - b_2)$, then $\psi(a_1, a_2, b_1, b_2) = 0$ implies $(a_1 - a_2) = 0$ and $(b_1 - b_2) = 0$.

Let $f_j : X \to \mathbb{R}^+$ $(j = 1, \dots, 4)$, $\psi(f_1, \dots, f_4) = F$. If $x \in \cap_1^4 f_j^{-1}(0)$, it is clear that $F(x) = 0$. Now let us introduce a 'disjoint support' condition. If $S_j = \operatorname{supp} f_j$, suppose

$$S_1 \cap S_2 = \varnothing \quad \text{and} \quad S_3 \cap S_4 = \varnothing,$$

then $F(x) = 0 \iff x \in \cap_1^4 f_j^{-1}(0)$. Indeed, $F(x) = 0 \implies f_1(x) = f_2(x)$, and $f_3(x) = f_4(x)$. It is impossible for the expressions to be non-zero, because the disjoint support condition means the diagonals of the images are $(0, 0)$.

> Is every complex measure the (complex) span of positive measures with $\mu(X) = 1$?

4. Point Set Topology

Let $(X, \mathcal{T})$ be a topological space, we call the elements of $\mathcal{T}$ open and write $U \overset{\circ}{\subseteq} X$ to mean $U \in \mathcal{T}$. If $E \subseteq X$, its subspace topology $\mathcal{T}_E$ is given by the intersection with open subsets of X, and $V \overset{\circ}{\subseteq} E$ refers to $V \in \mathcal{T}_E$.

A *neighbourhood of* $x \in X$ is a subset $E \subseteq X$ such that $x \in U \subseteq E$ for some $U \in \mathcal{T}$; and the set of neighbourhoods of x is $\mathcal{N}(x)$. Replacing x with $A \subseteq X$, $E \in \mathcal{N}(A)$ (meaning E is a neighbourhood of A) if there exists $U \overset{\circ}{\subseteq} X$ with $A \subseteq U \subseteq E$.

> Prove $\mathcal{N}(A) = \cap_{x \in A} \mathcal{N}(x)$, and $\mathcal{N}(\{x\}) = \mathcal{N}(x)$.

Often it will be useful to consider open neighbourhoods of x and A. We define $\mathcal{N}(x)^o = \mathcal{N}(x) \cap \mathcal{T}$, and $\mathcal{N}(A)^o = \mathcal{N}(A) \cap \mathcal{T}$.

4.1. Adherent Points

An extremely important concept in topology is the adherent point. An *adherent point* of $A \subseteq X$ is $x \in X$ where every neighbourhood of x intersects A. This means that x is 'stuck' to A and one is unable to separate x from A by taking smaller and smaller neighbourhoods about x. We immediately see that the set of adherent points of A is closed, and is equal to $\overline{A}$.

The boundary of A, denoted by ∂A consists of the points in x that adhere to both A and $X \setminus A$. One thinks of ∂A as being 'stuck' between A and $X \setminus A$.

4.2. Separation Axioms

We now contrast the previous section with the notions of separation. Suppose $E \in \mathcal{N}(x)$ satisfies $A \cap E = \varnothing$, we find an open $U \overset{\circ}{\subseteq} X$ where $x \in U \subseteq A^c$. Viewing $(U, \mathcal{T}_U)$ as a topological space, it contains none of the points in A, and motivates

Definition 4.1
$E \in \mathcal{N}(x)$ separates x from A whenever $A \cap E = \varnothing$.

For any open U, $x \in U$ and $y \notin U$, $\mathcal{N}(x)^o$ separates x from $\{y\}$, equivalently: points that *inside* U are separated from those that are *outside*.

Definition 4.2
The space X is T_1 if $\mathcal{N}(x)^o$ separates x from $\{y\}$ for all $x, y \in X$.

We can generalize Definition 4.1, by saying that $E \in \mathcal{N}(A)$ separates A from $B \subseteq X$ whenever $E \cap B = \varnothing$. We say that x, A is disjoint if $\{x\}, A$ is. The axioms T_2, T_3, T_4 are symmetric, and we will describe them at once.

Definition 4.3
T_2 X is Hausdorff if every disjoint pair x, y can be separated by disjoint open neighbourhoods of x, y.
T_3 X is regular if singletons are closed, and every disjoint, closed pair $\{x\}, A$ can be separated by disjoint open neighbourhoods of $\{x\}, A$.
T_4 X is normal if singletons are closed, and every disjoint, closed pair A, B can be separated by disjoint open neighbourhoods of A, B.

> Draw a picture describing the separation axioms T_i for $i = 0, 1, 2, 3, 4$.

Remark 4.4
The requirement that singletons are closed means $T_4 \implies T_3 \implies T_2 \implies T_1$. The topology of X is T_3 (resp. T_4) iff it is Hausdorff but one (resp. both) of the singletons can be replaced with closed sets.

> Prove X is regular iff for all $x \in V \overset{\circ}{\subseteq} X$, there exists $U \in \mathcal{N}(x)^o$ such that $x \in U \subseteq \overline{U} \subseteq V$. State and prove an equivalent condition for T_4.

The separation axioms describe how close X is to being a metric space. If X is a metric space, for all $x \in V \overset{\circ}{\subseteq} X$, we can easily find $r > 0$ such that $x \in B(r,x) \subseteq \overline{B(r,x)} \subseteq V$, so every metric space is regular (X is actually normal, see Exercise 4.3).

We conclude this section with a weaker notion of adherence: A accumulates at x if no neighbourhood of x separates x from $A \setminus \{x\}$.

> Prove x adheres to B iff no neighbourhood of x separates x from B.
>
> Prove $x \in \operatorname{acc} A$ iff x adheres to $A \setminus \{x\}$.

4.2.1 LSC Topology

4.3. Directed Sets

The family $\mathcal{N}(x)^o$ (more generally, $\mathcal{N}(x)$) becomes a directed set under reverse inclusion. A neighbourhood base at x is a cofinal subset of $\mathcal{N}(x)^o$; and it represents a smaller (possibly better-behaved) subcollection of $\mathcal{N}(x)^o$.

> Compare the definitions for directed sets, cofinal subset (see Exercise 4.30a) with the definitions and conditions laid out on p. 114 - 115.

In particular, if $\mathcal{M}(x)$ is a neighbourhood base at x, then it suffices to check that certain topological properties hold for $\mathcal{M}(x)$ to deduce that it holds for $\mathcal{N}(x)$.

> Prove $x \in \overline{A}$ (resp. $\operatorname{acc} A$ and resp. ∂A) iff no neighbourhood $U \in \mathcal{M}(x)$ separates x from A (resp. from $A \setminus \{x\}$ and resp. from A or from $X \setminus A$)

Remark 4.5

X is Hausdorff iff every pair x, y can be separated by disjoint sets in $\mathcal{M}(x), \mathcal{M}(y)$. A similar statement holds for T_3. In addition to the characterization above, X is T_2 if all of its two point subsets are disconnected (see Exercise 4.10) in the relative topology.

If $\mathcal{M}(x)$ is a neighbourhood base at x for all x, the union $\mathcal{M} = \cup_x \mathcal{M}(x)$ is a base of X; and X is said to be second countable if $\mathcal{M}$ can be taken to be countable. Finally, $E \subseteq X$ is a neighbourhood of x iff $U \subseteq E$ for some $U \in \mathcal{M}(x)$, and because E is open iff E it is a neighbourhood of every $x \in E$, we have an important characterization of open sets

> E is open iff it can be written as the union of sets in $\mathcal{M}$.[1]

[1] In some sense, knowing $\mathcal{M}$ implies knowing the entirety of $\mathcal{T}$.

4.3.1 Nets and first countable spaces

We will state and prove an extremely useful characterization of the adherent points of A which generalizes Propositions 4.6 and 4.18.

Proposition 4.6

Let $\mathcal{M}(x)$ be a neighbourhood base at x, then x adheres to B iff no neighbourhood $U \in \mathcal{M}(x)$ separates x from B iff some net $\langle z_\alpha \rangle_{\alpha \in \mathcal{M}(x)} \subseteq B$ converges to x. Specifically, if x admits a countable neighbourhood base, then x adheres to B iff some sequence in B converges to x.

Proof. We have proven the first equivalence. Suppose x adheres to B, we associate every neighbourhood $V \in \mathcal{M}(x)$ to an element $z_V \in V \cap B$. To show that $z_V \to x$, fix any open neighbourhood $U \in \mathcal{N}(x)^o$, because $\mathcal{M}(x)$ is cofinal, there exists $V \in \mathcal{M}(x)$ such that $V \subseteq U$, and for all $V' \gtrsim V$, $z_{V'} \in V' \subseteq V \subseteq U$. So that $z_V \to x$. Now suppose $V \in \mathcal{M}(x)$ separates x from B, then no net $\langle z_\alpha \rangle_{\alpha \in A} \subseteq B$ can intersect this neighbourhood of x.

By the first paragraph of page 116, if $\mathcal{M}(x)$ is countable, there exists a neighbourhood base $\{U_j\}_1^\infty$ at x which is well-ordered, countable $\mathcal{M}(x) = \{U_j^x\}_1^\infty$, such that $U_j \supseteq U_{j+1}$. Picking $z_j \in U_j \cap B$ finishes the proof. ∎

The previous proposition motivates the following description for topological spaces in which the notion of is characterized by sequences.

Definition 4.7

X is first countable if $\mathcal{M}(x)$ can be taken to be countable and decreasing for all x.

4.4. Dense Sets

We say that A is *dense* in X whenever $\overline{A} = X$, this means A has the property that for every point $x \in X$ and neighbourhood E, $E \cap A \neq \varnothing$. Furthermore, A is dense in X if and only if every non-empty open set intersects A.

More generally, A is *dense in* $U \overset{\circ}{\subseteq} X$ if $\overline{U \cap A} = \overline{U}$, Exercise 4.13 states that $A \subseteq X$ is dense (in X) iff it is dense in every open subset. Naturally, A is *nowhere dense in* $U \overset{\circ}{\subseteq} X$ if A is not dense in all $V \overset{\circ}{\subseteq} U$. See Theorem 5.9, it is extremely important in functional analysis.

Definition 4.8

$\mathcal{T}_0 = \{U \overset{\circ}{\subseteq} X,\ U \neq \varnothing\}$ the family of all non-empty open sets.

Definition 4.9

A subset $A \subseteq X$ is nowhere dense if $U \cap A = \varnothing$ for all $U \in \mathfrak{T}_0$.

Proposition 4.10

Show that A is nowhere dense iff $\overline{A}^o = \varnothing$.

Proof. Suppose A is nowhere dense, then $\overline{U \cap A} = \varnothing$ for all $U \in \mathfrak{T}_0$, and $\overline{U \cap A}^o = \varnothing$. ■

4.5. Continuity

The reason why the proof of Tietze's Extension Theorem works is because the norms of $\{g_j\}$ form a summable sequence. This is a technique which essentially boils down to an induction along with a limit process as long as we can approximate small elements in one space by sufficiently small elements in another space. See Lemma 5.5

4.6. Metric Spaces

Definition 4.11

A metric space is a pair $(X,\ d)$ where X is a set and a metric function $d : X \times X \to [0,\ +\infty)$

positive definiteness $d(x, y) = 0$ iff $x = y$,
triangle inequality For all $x_1, x_2, z \in X$, $d(x_1,\ x_2) \le d(x_1,\ z) + d(x_2,\ z)$.[a]
diameter $\operatorname{diam} E = \sup_{x,y\in E} d(x, y)$ for $E \subseteq X$
cauchy sequence $\{x_n\}$, where $\operatorname{diam} E_n \in c_0$, $E_n = \{x_m,\ m \ge n\}$.
complete metric space every cauchy sequence admits a limit to which it converges

[a] see Lemma 1.5

Proposition 4.12

distance to sets Let $A \subseteq X$, and $d_A(x) = \inf_{z\in A} d(x,\ z)$, then d_A is Lipschitz continuous, and $d_A(x) = 0$ iff $x \in \overline{A}$.
first countable, normal X is a first countable normal space.
shrinking diameter If $x \in U \overset{\circ}{\subseteq} X$, there exists $\varepsilon_{x,U} > 0$ such that for all $\varepsilon' \in (0,\ \varepsilon_{x,U}]$,

$$x \in B(\varepsilon',\ x) \subseteq \overline{B}(\varepsilon',\ x) \subseteq U.$$

Meaning the ball about x can be taken to be arbitrarily small. If X is a complete metric space, then every closed set A is again a complete metric space with the metric $d|_A : A \times A \to [0,\ +\infty)$.

Proof. Let $x, y \in X$, for every ε, by the second property of $d_A(x)$

$$d(y,\ z_\varepsilon) \le d(x,\ z_\varepsilon) + d(y,\ x) \le d_A(x) + d(y,\ x) + \varepsilon.$$

So that $d_A(y) \le d(y,\ x) + d_A(x)$, reversing the roles of x, y gives

$$|d_A(x) - d_A(y)| \le d(x,\ y).$$

It is clear that if $x \in \overline{A}$, then for all $\varepsilon > 0$, $B(\varepsilon,\ x) \cap A \ne \varnothing$, so $\inf_{z\in A} d(x,\ z) = 0$. Conversely, if given $\varepsilon > 0$, by the second property of $d_A(x)$,

$$d(x,\ z_{\varepsilon 2^{-1}}) \le d_A(x) + \varepsilon 2^{-1} = \varepsilon 2^{-1} < \varepsilon.$$

So that $B(\varepsilon,\ x) \cap A \ne \varnothing$ and $x \in \overline{A}$. ■

4.6.1 Baire Category Theorem

Definition 4.13

first category / meager X is the countable union of nowhere dense sets.
second category X is not the countable union of nowhere dense sets.
residual / comeager complement of a meager subset of X.

Proposition 4.14 Baire Category Theorem (Complete Metric Spaces)

If X is a complete metric space, and $\{U_n\}_1^\infty$ open and dense in X, then $\bigcap_1^\infty$ is dense in X. Moreover, X is of the second category.

Proof. If X is a T_3 space, for every $x \in U \overset{\circ}{\subseteq} X$, we can find $V_x \overset{\circ}{\subseteq} X$,

$$x \in V_x \subseteq \overline{V}_x \subseteq U.$$

If X is a complete metric space, $\operatorname{diam} V_x$ can be taken to be arbitrarily small and $\overline{V}_x$ is again a complete metric space in the subspace topology. Let $\{c_n\} \in l^1 \cap l^{++}$, we will construct a Cauchy sequence by induction. Given a non-empty $W \overset{\circ}{\subseteq} X$, set

$$x_1 \in X_1 \subseteq \overline{X}_1 \subseteq W \cap U_1$$

where $\operatorname{diam} X_1 \le c_1$.

Is it possible to use Folland Exercise 4.13? Because $\overline{V}_x$ is another complete metric space, and $U_j \cap V_x$ is dense in V_x? I am not sure.
Notice that $W \cap \overline{X}_1$ is non-empty, open ■

4.7. Proper Mappings

Let X be a set and $\kappa \subseteq \mathcal{P}(X)$ be a family of subsets of X such that $\varnothing \in \kappa$, and κ is closed under arbitrary intersections and finite unions.

Let X be a topological space and κ be the set of all compact subsets of X. If $f: X \to Y$ is a continuous mapping into a σ-compact Hausdorff space, and $\{U_n\}_1^\infty$ is a σ-compact exhaustion of Y, then f is a proper iff $f^{-1}(\overline{U}_{n_k}) \in \kappa$ for some subsequence n_k.

Lemma 4.15

Let X, Y be topological spaces and $f \in C(X, \mathbb{R})$, $g \in C(Y, \mathbb{R})$ proper. If at least one of X or Y is compact, or both f, g are bounded below, then $h = f + g$ is continuous proper.

Proof. If the first alternative holds, fix a compact $K \in \mathbb{R}$, since addition is completely continuous, $\widetilde{K} = K - f(X)$ is also compact. g is proper map, so $X \times g^{-1}(\widetilde{K})$ is compact. It suffices to show $h^{-1}(K) \subseteq X \times g^{-1}(\widetilde{K})$. Suppose $(x, y) \in h^{-1}(K)$, then $x \in X$ is clear. Because $f(x) + g(y) \in K$, it means $g(y) \in K - f(x) \subseteq \widetilde{K}$, or $y \in g^{-1}(\widetilde{K})$. By continuity, $h^{-1}(K)$ is a closed subset of a compact set in $X \times Y$, therefore it is compact and h is proper.

If the second alternative holds, it suffices to show that $h^{-1}([-b,\ +b])$ is compact for all $b > 0$. Because f, g are bounded below, there exists $a_1, a_2 > 0$ such that $-a_1 \leq \inf_x f(x)$ and $-a_2 \leq \inf_y g(y)$. For any $(x, y) \in h^{-1}([-b,\ +b])$ we have

$$f(x) \leq b - g(y) \leq b + a_2.$$

Similarly $g(y) \leq b - f(x) \leq b + a_1$, hence $h^{-1}([-b,\ +b]) \subseteq f^{-1}([-a_1,\ b + a_2]) \times g^{-1}([-a_2,\ b + a_1])$ and h is proper. ■

5. Functional Analysis

5.1. Operator Norm

$$S_\rho = \{x \in X,\ |x| = \rho\}$$
$$\overline{B}_\rho = \{x \in X,\ |x| \leq \rho\}$$
$$B_\rho = \{x \in X,\ |x| < \rho\}.$$

One should view these sets and their geometric structures as being inherited from the structure of $[0,\ +\infty)$. For any non-zero element $x \in X$, we can normalize $\frac{x}{\|x\|} = 1$, it is also clear that

$$S_\rho = \{\rho x,\ |x| = 1\} = \{\frac{\rho x}{\|x\|},\ x \neq 0\}.$$

5.2. Seminorms

Definition 5.1

A *seminorm* on a vector space X is a function $p: X \to [0, +\infty)$ satisfying:

$$p(x+y) \leq p(x) + p(y) \quad \text{and} \quad p(\alpha x) = |\alpha| p(x) \quad \forall \alpha \in \mathbb{C}.$$

The *zero-set* of a semi-norm p is always a subspace. To see this, define $p^0 = \{x \in X,\ p(x) = 0\}$. Fix $\alpha, \beta \in \mathbb{C}$ and $x_0, x_1 \in p^0$. By the triangle inequality and absolute homogeneity:

$$p(\alpha x_0 + \beta x_1) \leq |\alpha| p(x_0) + |\beta| p(x_1) = 0.$$

We begin this section with a discussion about the geometric properties of convex subsets of $\mathbb{R}$. Let $c > 0$, and $M_c = [0,\ c)$ be the open convex interval in $\mathbb{R}^+ = [0,\ +\infty)$. If $x,\ y$ in M_c, $t \in (0,\ 1)$, their convex combination $z_t = (t)x + (1-t)y$ clearly lies in c.

$$z_t \leq (t)c + (1-t)c < c.$$

Let $\alpha \in [0,\ 1]$, for any $x \in M_c$, αx is also in M_c. So that every point in M_c contains a straight line segment to the origin. Next, for every $x \geq 0$ there exists some large α (by the Archimedean property) such that

$$0 \leq \alpha^{-1} x < c.$$

We see that the infimum below is finite for every $x \geq 0$.

$$\inf\{\alpha,\ \alpha^{-1} x \in M_c\} = x/c.$$

It is useful here, to remember that convex subsets of $\mathbb{R}$ are precisely the intervals. Now conversely, let $C \overset{\circ}{\subseteq} \mathbb{R}^+$ be an open (relative to $\mathbb{R}^+$) convex subset containing the origin. First, literally by definition it is convex. Second, because $[0,\ \varepsilon) \subseteq C$ for sufficiently small ε, every point $x \geq 0$ can be dilated so that $\alpha^{-1} x \in C$. We define, as in the book

$$p: \mathbb{R}^+ \to \mathbb{R}^+ \quad p(x) = \inf\{\alpha,\ \alpha^{-1} x \in C\}.$$

In the one-dimensional case, it will turn out that p completely captures the geometric structure of C. The mental picture should be the following, for a fixed $x > 0$, as one increases α, the number $\alpha^{-1} x$ moves to the left. The quantity $p(x)$ smallest α such that it enters the interval.

Proposition 5.2

For $r, x \geq 0$, $p(x) < 1$ iff $x \in C$, $p(rx) = r p(x)$, $p(x+y) \leq p(x) + p(y)$.

Proof. We will only prove the triangle inequality. We encourage the reader to try out this simplified case, keeping the proofs as simple as possible because there is actual geometric insight in the theorems themselves. There is a general framework of proving estimates of the form:

$$q(x+y) \leq q(x) + q(y) \quad q(z) = \inf \Sigma_z.$$

One starts from the right hand side of the estimate, and approaches the two infima from above $\varepsilon_j > 0$ $(j = 1,2)$ and obtain ϕ_j such that $\phi_1 \le q(x) + \varepsilon_1$, similarly for y. If the two quantities ϕ_1, ϕ_2 produces $\phi_3 \le \phi_1 + \phi_2$, for some $\phi_3 \in \Sigma_{x+y}$, then the estimate will be proven.[2]

If $x/[p(x)+\varepsilon]$, $y/[p(y)+\varepsilon]$ are in C, how can we produce $(x+y)/\text{constant} \in C$? The first attempt would be to use the fact that for every point in C contains a straight line segment to the origin, so we let $r = \max(p(x),\ p(y)) + \varepsilon$, then

$$xr^{-1},\ yr^{-1} \in C \quad \text{hence} \quad (x+y)/(2r) \in C,$$

because $(x+y)/2r$ is the midpoint of a line segment. This proves $p(x+y) \le 2\max(p(x),\ p(y))$. Which is not what we are looking for unfortunately. Let us try again, the first step of the proof is obviously tight. Given the information we have so far about x, y and the subset C, we should not be able to deduce a better estimate. The culprit is in the second step, where we have made the choice of $r = \max(p(x),\ p(y)) + \varepsilon$. If $p(x) \ll p(y)$, the maximum of the two will be a terrible choice. We need an algebraic result.

> Given two positive numbers, the reciprocal of their sum lies on the line segment joining the two reciprocals. In short,
>
> $$\forall a, b > 0 \exists t_0 \in (0,\ 1),\ (a+b)^{-1} = t_0 a^{-1} + (1-t_0) b^{-1}.$$

Which is *exactly* what we want. [3] This means, if $r = p(x) + p(y) + 2\varepsilon$, then $\frac{x+y}{r} \in C$ and the claim is proven. ■

The functional constructed above separates the C from its complement by considering the sublevel set $\{p < 1\}$. For the general case where C is an open, convex subset of a Banach space containing the origin. The main ideas are the same, p is positively homogeneous, subadditive, and $C = \{p < 1\}$. However, we lose the property of distinguishing individual points, because $\dim E \ge 2$. The functional p as constructed should be thought of as the 'radial' component of a certain kind of 'polar' decomposition of a Banach space — one which the radial component measures how much you should inflate C to contain a point.

Furthermore, if we allow ourselves to place stronger assumptions on C, meaning C is absorbent, balanced and convex, then every seminorm p is of this form. It follows from the previous discussion that for all $c > 0$, the set $M = \{x \le_p c\}$

- is *convex*,
- is *balanced* — for all $x \in M$ and $|\alpha| \le 1$, $\alpha x \in M$, and
- is *absorbing* (Archimedean) — for all $x \in X$ there exists $\overline{\alpha} > 0$ such that $\alpha^{-1} x \in M$ for $\alpha \in (0,\ \overline{\alpha})$. Moreover,

$$p(x) = \inf\left\{\alpha c,\ \alpha^{-1} x \in M, \alpha > 0\right\}.$$

[2] this is used in Ex 5.12, 6.2, 6.3. Chapter 7 many times, often with sup instead of inf

[3] This algebraic lemma above underlies much of Chapter 6 and Harmonic Analysis.

5.3. Factor Spaces

Closest distance Let X be a $\mathbb{R}$ or $\mathbb{C}$ nvs and S a subset of X. Define $q_S : X \to [0, +\infty)$ to be the map $q_S(x) = \inf_{z\in S}\|x-z\|$, which computes the closest distance one can get to the set S. We claim that $x =_q 0$ iff $x \in \overline{S}$. The ($\Longleftarrow$) direction is clear, so fix $\varepsilon > 0$ and $z_\varepsilon \in S$ where $\|x-_\varepsilon\| \le \varepsilon$ so that $x \in \overline{S}$.
Seminorm to Norm on Factor Space Let X be a $\mathbb{R}$ or $\mathbb{C}$ vector space and $p : X \to [0, +\infty)$ a seminorm on X, set $M = \{x =_p 0\}$. Then $p' : X/M \to [0, +\infty)$ is a norm on X/M with $p'(x+M) = p(x)$ for all $x+M \in X/M$.

First, we show that p' is well defined. Fix $x+M \in X/M$ and for all $y \in M$, then

$$p(x) - p(y) \le p(x) = p'(x+M) \le p(x) + p(y).$$

Set $x_0 + M$ and $x_1 + M \in X/M$, and the triangle inequality follows from

$$p'((x_0+x_1)+M) = p(x_0+x_1) \le p'(x_0+M) + p'(x_1+M).$$

Similarly, if α is a scalar, then the homogenity of p' follows from that of p: $p'(\alpha x + M) = p(\alpha x) = |\alpha| p(x)$.
Norm to Seminorm on Factor Space Let M be a subspace of nvs X, then $q_M(x) = \inf_{z\in M}\|x - z\|$ is a seminorm on X/M. If M is a *closed* subspace of X, then q_M is a norm.

For all $x + \mathcal{M} \in X/\mathcal{M}$, we can 'define' its norm by a descending sequence of elements in $\mathcal{M}$, so that $(z_n) \subseteq \mathcal{M}$ and $\|x + z_n\| \searrow \|x + \mathcal{M}\|$. We do Exercise 12 first. Let $\|\cdot\|$ be a norm on X, fix $x + \mathcal{M}$ and $y + \mathcal{M}$ in $X/\mathcal{M}$, suppose also $x - y \in \mathcal{M}$, we want to show that the equation defines the same number. If $\|x + z_n\| \searrow \|x + \mathcal{M}\|$, then

$$\|x + z_n - (x-y)\| = \|y + z_n\| \ge \inf_{z\in\mathcal{M}}\|y+z\| = \|y\|_{X/\mathcal{M}}$$

Taking the limit inferior on both sides, we see that $\liminf\|x\|_{X/\mathcal{M}} \ge \|y\|_{X/\mathcal{M}}$. Now we relabel $(z_n) \subseteq \mathcal{M}$ and note that $y - x \in \mathcal{M}$, if $\|y + z_n\| \searrow \|y + \mathcal{M}\|$, then

$$\|y + z_n - (y-x)\| = \|x + z_n\| \ge \inf_{z\in\mathcal{M}}\|x+z\| = \|x\|_{X/\mathcal{M}}.$$

So that the quotient norm as defined above is independent of the representative being used. Note that the quotient norm is non-negative, as for all $x \in X, z \in \mathcal{M}$, $\|x + z\| \ge 0$, so the infimum over all such z is again non-negative.

Verification of the triangle inequality. Fix $x, y \in X$ and $(z_n), (z_n') \subseteq \mathcal{M}$ with $\|x + z_n\| \searrow \|x + \mathcal{M}\|$ and $\|y + z_n'\| \searrow \|y+\mathcal{M}\|$, for a fixed $n \ge 1$, we can apply the triangle inequality and take limits as follows:

$$\begin{aligned}\|(x+y)+\mathcal{M}\| &\le \|x+y+(z_n+z_n')\| \\ &\le \|x+z_n\| + \|y+z_n'\| \to \|x+\mathcal{M}\| + \|y+\mathcal{M}\|.\end{aligned}$$

Verification of Homogeneity. Fix $x \in X$ and $\alpha \in \mathbb{C}$, we can use the same sequential technique and pass the algebraic properties (as long as they are continuous under the limit) to each of the terms in the defining sequence.

$$\|\alpha(x+z_n)\| = |\alpha|\|x+z_n\| \searrow |\alpha|\|x+\mathcal{M}\|.$$

The limits converge and multiplication is continuous, hence we don't need split this proof into two estimates.

Verification of positive semi-definiteness. It is clear that $q_M \ge 0$, and for all $x \in \mathcal{M}$, $q_M(x) \le \|x - x\| = 0$, so q_M is a seminorm.

Verification of positive definiteness if M is closed. Suppose that $x \in X$ is a point where $\|x + \mathcal{M}\|_{X/\mathcal{M}} = 0$, so we there exists a sequence $(z_n) \subseteq \mathcal{M}$ with

$$x\|x+z_n\| \searrow 0, \quad \text{iff} \quad \|x - (-z_n)\| \to 0$$

and $x \in \mathcal{M}$ because $\mathcal{M}$ is closed.

Note 5.3

Just to make this clear, $\mathcal{M}$ is closed is necessary for the following claim

$$\|x+\mathcal{M}\|_{X/\mathcal{M}} = 0 \implies x \in \mathcal{M}.$$

Proposition 5.4

Ex 5.12b Let X be a Banach space, and $M \ne \{0\}$ a closed subspace of X, and X/M the factor space. For every $\varepsilon > 0$, there exists $x+M \in X/M$, $\|*\|x+M = 1-\varepsilon$, and $z \in X$, $\|*\|z = 1$, such that $z + M = x + M$.

Proof. Intuitively, an element $x + M \in X/M$ is approximated by sequences of $(x_n + m_n)$ where $x_n \in X$ and $m_n \in M$, where m_n should be thought of a small perturbation away from the equivalence class $x + M$. These sequences need not converge of course. For any fixed $n \ge 1$,

Reduce it to the case where: for every $x + M \in X/M$ with $\|*\|x + M = 1 - \varepsilon$, we can find $z \in X$, with $\|*\|z = 1$, such that $z + M = x + M$. This is different than what is asked in the question, because ε depends on $x + M$. We are allowed to do this because $X/M \ne \{0\}$, so for every $\varepsilon > 0$, we can always find some element $x+M$ whose quotient norm is $1-\varepsilon$.

The quotient norm of $x + M$ is strictly smaller than 1, this means we can fit a small perturbation of size at most δ, where $\delta \in (0, \varepsilon 2^{-1})$. There exists $y \in X$, $m \in M$:

$$z_\delta = y + m \quad \text{and} \quad \|*\|z_\delta \in [1-\varepsilon, 1-\varepsilon+\delta]$$

If $M \ne \{0\}$, we can assume that $m \ne 0$. Indeed, pick $a \in M$ with $\|*\|a \ne 0$, then

$$\|*\|z_\delta + a \le \|*\|z_\delta + \|*\|a \le (1-\varepsilon+\delta) + \|*\|a.$$

Because $\|*\|a$ is at our disposal, and can be made arbitrarily small, we can replace m with a, if $m = 0$. We will exaggerate the M-perturbation by using the line segment

$$|t|\|*\|m - \|y\| \le \|*\|y + tm.$$

So that $\|*\|y + tm \to +\infty$ for large t, and by continuity we obtain $t_0 > 1$ where $z = y + t_0 m$ has norm 1. Finally, if $M = \{0\}$ then $X/M \approx X$. ■

5.4. Successive Approximations

Lemma 5.5 Method of Successive Approximations

Let X, Y be Banach spaces, and $T \in L(Y,\ X)$. A subset $\Sigma \subseteq Y$ is said to satisfy the *reverse Lipschitz criterion* for X if there exists $C > 0$, $\gamma \in (0,\ 1)$, for any $\varepsilon > 0$ and $\|x\|_X \le \varepsilon$,

$$\|\phi\|_Y \le C\varepsilon, \quad \|x - T\phi\|_X \le \gamma\varepsilon \quad \text{for some} \quad \phi \in \Sigma.$$

Then, T is a surjection from $\{y = \sum_1^\infty \phi_j,\ \{\|*\|\phi_j\} \in l^1,\ \phi_j \in \Sigma\}$. That is, every $x \in X$ can be written $x = T(\psi)$; this element $\psi = \sum_1^\infty \phi_j$ can be taken to satisfy $\|\psi\| \le C\|x\|(1-\gamma)^{-1}$. Moreover, if Σ' satisfies the reverse Lipschitz criterion for a fixed $\varepsilon' > 0$, then $\Sigma = \{c\phi,\ c > 0,\ \phi \in \Sigma\}$ satisfies the criterion for all $\varepsilon > 0$.

Proof. Given $x \in X$, we can find $\phi_1 \in \Sigma$ with $\|\phi_1\| \le C\|x\|$ and $\|x_1\| \le \|x\|\gamma$; where $x_1 = x - T\phi_1$. Suppose $x, x_1, \dots, x_n$ have been chosen such that $\|\phi_j\| \le C\|x\|\gamma^{j-1}$, $\|x_j\| \le \|x\|\gamma^j$ for $j = 1, \dots, n$. We find a small ϕ_{n+1} that improves our estimate

$$\|*\|x - T(\textstyle\sum_1^n \phi_j) - T(\phi_{n+1}) \le \|x_n\|\gamma \le \|x\|\gamma^{n+1},$$
$$\|\phi_{n+1}\| \le C\|x_n\| \le C\|x\|\gamma^{n+1}.$$

Let $\psi = \sum_1^\infty \phi_j$, so that $\lim_n \|*\|\psi - \sum_1^n \phi_j = 0$. By continuity and linearity of T

$$\|*\|x - T\psi \le \liminf_n \|*\|x - T(\textstyle\sum_1^n \phi_j) \le \liminf_n \|x\|\gamma^n = 0.$$

To prove the last claim, if $x \ne 0$, $x' = (\varepsilon' r^{-1})x$ has norm less than ε' for sufficiently large $r > 0$. This induces $\phi' \in \Sigma'$ where $\|\phi'\| \le C\varepsilon'$, and $\|x' - \phi'\| \le \gamma\varepsilon'$. Multiplying by r across the three equations and by linearity of T

$$\forall \|x\| \le (\varepsilon' r), \quad \exists \|*\|\phi \le C(\varepsilon' r), \quad \|x - T\phi\| \le \gamma(\varepsilon' r)$$

for $\phi = r\phi' \in \Sigma$. ■

Proposition 5.6 Open Mapping Theorems

Let X, Y be Banach spaces, then every continuous linear surjection is an open map.

Manifold Every C^p submersion ($p \ge 1$) is an open map.

Invertible Operators The set of invertible operators is strongly open.
Graves The set of continuous linear surjections is strongly open.

5.5. Geometric Arguments

Remark 5.7

How do dilations affect the sublevel sets of seminorms? What about translations? We can center the sublevel at around a point x_0 to simulate some kind of ball.

This will be useful when discussing the weak topologies generated by functionals.

Lemma 5.8

Let X be a $\mathbb{R}$ Banach space and $p : X \to [0, \infty)$. For $c > 0$ we define

$$\overline{M}(c,\ x_0) = \{x \in X,\ x - x_0 \le_p c\} \quad \text{and} \quad M(c,\ x_0) = \{x \in X,\ x - x_0 <_p c\}.$$

If p is absolutely homogeneous, meaning $p(\alpha x) = |\alpha| p(x)$ for all $x \in X$, then for all $\alpha \ne 0$,

$$\alpha\overline{M}_p(c,\ x_0) = \overline{M}(|\alpha|c,\ \alpha x_0),$$

and

$$\alpha M(c,\ x_0) = M(|\alpha|c,\ \alpha x_0).$$

Suppose further that p is a seminorm (satisfies the triangle inequality in addition to homogeneity), then

5.5.1 Convexity

Remark 5.9

The *convex hull* of a subset $M \subseteq X$ is $\text{conv}(M) = \{\sum_1^n t_i x^i,\ x^i \in M\}$. Convex Set — Wikipedia.

Remark 5.10

Let $(t_i)_1^n \subseteq [0,\ 1]$, $\sum t_i = 1$ and $(x_i)_1^n \subseteq X$, then $|\sum_1^n t_i x_i| \le \max|x_i|$.

Remark 5.11 Diameter is preserved

Let $M \subseteq X$ and $\text{diam}(M) = \mu < +\infty$. We want to show $\text{diam}\,\text{conv}(M) = \text{diam}(M)$.

Since M is contained in its convex hull, $\text{diam}(M) \le \text{diam}\,\text{conv}(M)$. Let $(t_i)_1^n$ and $(s_j)_1^m$ be numbers in $[0,\ 1]$

that sum to 1, then

$$\left|\sum_1^n t_i x_i - \sum_1^m s_j y_j\right| \le \sum_{i,j} s_i t_j |x_i - y_j| \le \mu \sum_{i,j} s_i t_j.$$

The sum is equal to 1 by Fubini's theorem. More generally, if $x_{\underline{n}}$ are arbitrary elements in a normed vector space X, then $\|\sum t_i x_{\underline{n}}\| \le \max\|x_{\underline{n}}\|$.

Remark 5.12

A subset M is convex iff $tx+(1-t)y \in M$ for all $x, y \in M$. We prove this implies $\sum t_i x_i \in M$ for any convex combination $(t_i)_1^n \subseteq [0,\ 1]$. We can use induction. Suppose for $\sum_1^n t_i x_i \in M$ for arbitrary n-convex combinations of M, and fix $n+1$ elements $(t_i x_i)_1^{n+1}$

5.6. Closed Subspaces

Definition 5.13 Split Subspaces

Let E be a Banach space, a closed subspace $M \subseteq E$ is said to *split* whenever there exists another closed subspace $N \subseteq E$ with $M \cap N = 0$, and $M + N = E$.[a]

[a] We remark that linear complements are not unique (easy to see in $\mathbb{R}^2$, if x is any vector that spans a 1-dimensional subspace, then $\text{span}(y)$ is a linear complement of x provided that $y \notin \text{span}(x)$.

Proposition 5.14

If E_1 is a closed subspace of E that splits, where $E_1 \oplus E_2 = E$, and E_2 is closed, then $E_1 \times E_2 \cong E$. [a]

[a] the congruent symbol $\cong$ refers to a toplinear isomorphism.

Proof. Let $A(x, y) = x + y$ be the addition map, we equip [4] $E_1 \times E_2$ with $\|(x, y)\| = \|x\| + \|y\|$. Follows from the continuity of the norm that $\|A(x, y)\| = \|x + y\| \le \|(x, y)\|$ is a toplinear isomorphism by the open mapping theorem. ■

Remark 5.15 Split subspaces allow for projections

An immediate corollary of the last theorem is that if E_1 and E_2 splits in E, there exists continuous linear maps $P_j : E \to E_j$.

Proposition 5.16 Folland 5.2 Splitting

Ex 5.18a Let $\mathcal{M}$ be a closed subspace, and $x \in X \setminus \mathcal{M}$, then $\mathcal{M} + \mathbb{C}$ is again a closed subspace.
Ex 5.18b Every finite dimensional subspace is closed.

Ex 5.20 Every finite dimensional subspace is closed and splits.

Proof. 5.18a) If $\mathcal{M}$ is a closed subspace of X, and $x \in X \setminus \mathcal{M}$, by Theorem 5.8a, there exists some $f \in X^*$, such that $f(x) = \delta$, and $f(\mathcal{M}) = 0$. Let $\{z_n\} \subseteq \mathcal{M} + \mathbb{C}x$ converge to some $z \in X$, we prove that $z \in \mathcal{M} + \mathbb{C}x$. By the continuity of f,

$$f(z_n)\delta^{-1}x \to f(z)\delta^{-1}x.$$

Let $y_n = z_n - f(z_n)\delta^{-1}x$, and $y = z - f(z)\delta^{-1}x$. It follows that $y_n \to y \in \mathcal{M}$, because it is the linear combination of two convergent sequences. Therefore $z \in \mathcal{M} + \mathbb{C}x$.

5.18b) Let $\mathcal{M}$ be finite dimensional with basis $(a_i)_1^k$, we use induction to show that $\mathcal{M}$ is closed. This is trivial, as the base case $\varnothing$ is closed; and hence $\mathcal{M}$ is closed.

5.20) We now show $\mathcal{M}$ splits, for each basis vector, we can construct the projection map $\varphi_i : \mathcal{M} \to \mathbb{C}$ onto the i-th coordinate.[5] Let φ_i also represent its continuous extension to all of X. If $z \in X$, we can decompose it into two parts:

$$z = \left(\sum_1^k \varphi_i(z)a_i\right) + \left(z - \sum_1^k \varphi_i(z)a_i\right).$$

The second member is, almost by definition[6], in the kernel of all φ_j,

$$\begin{aligned}\varphi_j\left(z - \sum_1^k \varphi_i(z)a_i\right) &= \varphi_j(z) - \sum_{i=1}^k \varphi_j(\varphi_i(z)a_i)\\ &= \varphi_j(z) - \sum_{i=1}^k \varphi_i(z)\delta_i^j = 0.\end{aligned}$$

We set $N = \cap \text{Ker}(\varphi_i)$, which is a closed linear subspace of X and the proof is complete. ■

Proposition 5.17 Left and Right Inverses

Let T be a linear operator from E to F.

Left Inverse If T is an injection, T admits a left inverse iff the range of T is closed, and splits in F.
Right Inverse If T is a surjection, T admits a right inverse iff the kernel of T splits in E.

Proof. If T is an injection where $T(E)$ is closed and splits in F with a subspace $V \subseteq F$. Let $P : F \to T(E)$ be the projection. Since $P \circ T$ is a toplinear isomorphism, we denote its inverse by $W : T(E) \to E$.

We construct the left inverse $S : F \to E$, that projects $y \in F$ onto $T(E)$ then applies W (that is, $S(y) = W(P(p))$). If $x \in E$,

[4] the following is equivalent to $\max(\|x\|, \|y\|)$, because $\max(\|x\|, \|y\|) \le 2(\|x\| + \|y\|) \le 2\max(\|x\|, \|y\|)$.

[5] the motivation for this is that linear functionals 'live on' one dimension only, and we set $\varphi_i(a_j) = \delta_i^j$.

[6] where φ_j conveniently picks out the jth term among all i.

then $PTx = Tx$ because P is a projection (see [Rom07, pp.74]) and hence $STx = x$.

Conversely if T admits a left inverse $S : F \to E$, we show that its range is closed and splits.

Range of T is closed Let $\{Tx_n\} \subseteq T(E)$ be a convergent sequence in F. Because we have a left inverse in S, we can send the sequence back into E. If $Tx_n \to z$, then $x_n \to Sz \in E$, and hence $Tx_n \to TSz \in T(E)$.
$\operatorname{Ker}(S)$ is the linear complement of $T(E)$ For every $z \in F$, we decompose

$$z = TSz + (z - TSz),$$

where $S(z - TSz) = Sz - STSz = Sz - Sz = 0$.

If T is a surjection where $\operatorname{Ker}(T)$ splits, let $P : E \to \operatorname{Ker}(T)$ be the projection map. We construct a right inverse as follows. We associate any $y \in F$ to the element $x - P(x)$, where $Tx = y$. [7] The mapping $Sy = x - P(x)$ is a right inverse of T, as

$$TSy = Tx - TP(x) = Tx = y.$$

Moreover, S is linear, as $T \circ \sum_{\mathbb{R}}^{\wedge}(x_i) = \sum_{\mathbb{R}}^{\wedge} y_i$ whenever $y_i = Tx_i$, and the linear combination factors out from $\mathrm{id}_E \times P$:

$$S\left(\sum_{\mathbb{R}}^{\wedge} y_i\right) = \left(\sum_{\mathbb{R}}^{\wedge} x_i - P(\sum_{\mathbb{R}}^{\wedge} x_i)\right) = \sum_{\mathbb{R}}^{\wedge}(x_i - P(x_i)).$$

Conversely, if T admits a right inverse $S : F \to E$, we claim that

Range of S is closed Suppose $S(y_n)$ is a convergent sequence in E, recalling that T is a surjection we can assume $STx_n = S(y_n) \to x \in E$ for some sequence $\{x_n\} \subseteq E$ and element $x \in E$. By continuity of T, we can send these points into F

$$TSTx_n = Tx_n \to Tx \in F.$$

And hence $STx_n \to STx \in S(F)$.
Range of S is linear complement of $\operatorname{Ker} T$. Every $z \in E$ admits a decomposition $z = STz + (z - STz)$, where member in parentheses is in $\operatorname{Ker}(T)$, as $Tz - TSTz = Tz - Tz = 0$. Suppose $z \in S(F) \cap \operatorname{Ker} T$, then $z = Sy$ for some $y \in F$ and $Tz = 0$; and hence

$$0 = Tz = TSy = y \quad \text{implies} \quad 0 = Sy = z.$$

■

6. L^p Spaces

Proposition 6.1 Minkowski's Inequality for Integrals
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. If $f : (X \times Y) \to [0, +\infty]$ is a measurable mapping, then for every $r \geq 1$

$$\left(\int_X \left[\int_Y f(x,y)\mathrm{d}y\right]^r \mathrm{d}x\right)^{1/r} \leq \int_Y \left(\int_X |f(x,y)|^r \mathrm{d}x\right)^{1/r} \mathrm{d}y.$$

We say that the norm of the integral is bounded above by the integral of the norms.

Proof. Folland Proposition 6.19a ■

6.1. Retracts

Definition 6.2
Let Y be a topological space and $E \subseteq X$. A retract of E is a continuous map $r : Y \to E$ such that $r|_E = \mathrm{id}_E$.

Definition 6.3
If Y is a normed vector space, for any $\theta > 0$, the mapping $\varphi_\theta : Y \to M_\theta = \{y \in Y,\ |y| \leq \theta\}$ is called the θ-retract of Y.

$$\varphi_\theta|_{M_\theta} = \mathrm{id}_{M_\theta} \quad \text{and} \quad \varphi_\theta(y) = (\theta|y|^{-1})y \quad \forall y \notin M_\rho.$$

Definition 6.4
Let $f : X \to \mathbb{C}$ be measurable, $\theta > 0$, and $\varphi_\theta : \mathbb{C} \to M_\rho = \{z \in \mathbb{C},\ |z| \leq \rho\}$ be the ρ-retract. The maps

$$f_\theta = \varphi_\theta \circ f \quad \text{and} \quad f^\theta = f - f_\theta$$

are called the θ-retract and θ-tail of f.

Example 6.5
Let $f(t) = t$ defined on $\mathbb{R}$, its 1-retract is the map $g(t) = \max(-1,\ \min(1,\ t))$, while its 1-tail is $h(t) = \max(t - 1,\ 0) + \min(t + 1,\ 0)$

6.2. Distribution Functions

Definition 6.6
Let $f : X \to \mathbb{C}$ be measurable, its distribution function is a mapping $\lambda_f : (0, +\infty) \to [0, +\infty]$,

$$\lambda_f(\alpha) = \mu\big(\{x \in X,\ |f(x)| > \alpha\}\big).$$

Example 6.7
If $f : X \to \mathbb{C}$, $f(x) = c$, then $\lambda_f(\alpha) = \mu(X)\chi_{(0,\ |c|)}(\alpha)$.

[7] two things: 1) such an x always exists, because T is a surjection. 2) this is independent of x chosen. If $Tx = y = Tz$, then x and z differ by some element in $\operatorname{Ker}(T)$, which is precisely the part we chose to 'trim' from the correspondence.

Proposition 6.8

Let f be measurable,

- For all $p \in (0, +\infty)$, $\|*\|f_p^p = p\int_0^\infty(\alpha^{p-1})\lambda_f(\alpha)\mathrm{d}\alpha$
- If $f \in L^\infty$, $\operatorname{supp}(\lambda_f) = (0, \|*\|f_\infty]$.
- For $\theta > 0$,
  $\lambda_{f_\theta}(\alpha) = \chi_{(0,\theta)}(\alpha)\lambda_f(\alpha) \quad \text{and} \quad \lambda_{f^\theta}(\alpha) = \lambda_f(\alpha+\theta)$.

Proof. See Folland 6.23 to 6.25. ■

7. Minimax Principles

Definition 7.1 Palais-Smale Condition

Let E be a real Banach space, and $I \in C^1(E, \mathbb{R})$. We say that I satisfies the Palais-Smale condition if every bounded (meas. I) sequence $\{u_n\} \subseteq E$ with vanishing gradient $I'u_n \to 0$ admits a strongly convergent subsequence $\{u_m\}$.

Remark 7.2

The induced strongly convergent subsequence converges to a critical point of I.

Proposition 7.3 Mountain Pass Theorem

Let E be a real Banach space and $I \in C^1(E, \mathbb{R})$ satisfies the PS condition. Suppose $I(0) = 0$ and

- there exists $\rho, \alpha > 0$ such that $I|_{\partial B_\rho} \geq \alpha$,
- there is $e \in E \setminus \overline{B}_\rho$ such that $I(e) \leq 0$.

Then I possesses a critical value $c \geq \alpha$, where

$$c = \inf_{g \in \Gamma_0^e} \max_{u \in g([0,1])} I(u),$$

where $\Gamma_0^e = \{g \in C([0,1], E),\ g(0) = 0,\ g(1) = e)\}$ represents the set of continuous curves from 0 to e.

Two ideas in Proposition 7.3. First, we obtain an estimate for c (and c is finite). In this case we show $c \in [\alpha, +\infty)$. Second, using a deformation argument, if c is not a critical value. The time-one map of a modified pg field gives us

Proposition 7.4

Let E be a real Banach space and I satisfy the PS condition. If $c \in \mathbb{R}$ is not a critical value of I, for every $\overline{\varepsilon} > 0$ there exists $\varepsilon \in (0, \overline{\varepsilon})$ and $\eta \in C([0,1] \times E, E)$ such that

- The time-one map of η is supported within $I^{-1}([c-\overline{\varepsilon},\ c+\overline{\varepsilon}])$.
  $$\eta(1, u) = u \quad \forall u \notin I^{-1}([c-\overline{\varepsilon},\ c+\overline{\varepsilon}]).$$
- The time-one map of η sends $A_{c+\varepsilon}$ to $A_{c-\varepsilon}$.
  $$\eta(1, A_{c+\varepsilon}) \subseteq A_{c-\varepsilon}.$$

Proof of Proposition 7.3. The first claim is easy to verify. To show that $c < +\infty$, we only have to show $\Gamma_0^e \neq \varnothing$, because every continuous curve in attain their maximum (meas. I). Take the straight line from 0 to e. By an intermediate value theorem argument, every curve $g \in \Gamma_0^e$ must intersect ∂B_ρ, so that $g(t_0) \geq \alpha$ for some t_0. Hence $\max_{t\in[0,1]} g(t) \geq \alpha$ for every curve $g \in \Gamma_0^e$, and $c \geq \alpha$.

For the second part of the proof, if c is not a critical value, set $\overline{\varepsilon} = \alpha 2^{-1}$, we want a tightly supported time-one map $\eta(1,\cdot) = \eta(\cdot)$ with $\operatorname{supp}(\eta) \subseteq I^{-1}([c-\alpha 2^{-1},\ c+\alpha 2^{-1}])$.

More importantly, if $g \in \Gamma_0^e$, then $\eta(g)$ is also in Γ_0^e. This is because η is continuous, and the endpoints $0, e$ are non-positive,

$$I(\{0, e\}) \cap [c-\overline{\varepsilon},\ c+\overline{\varepsilon}] = \varnothing.$$

Now, approach c from above and obtain $g \in \Gamma_0^e$, $\max_t I(g(t)) \leq c+\varepsilon$, then $g([0,1]) \subseteq A_{c+\varepsilon}$ and $\max_t(\eta \circ g)(t) \leq c-\varepsilon$. This contradicts the definition of c because $\eta \circ g \in \Gamma_0^e$. ■

Proposition 7.5

Let E be a real Banach space and $I \in C^1(E, \mathbb{R})$ satisfying PS. If I is bounded from below, then $c = \inf_{z\in E} I(z)$ is a critical value of I.

Proof. Because I is bounded below, c is finite. To illustrate why the proof above works, we let $S = \{\{x\},\ x \in E\}$ be the collection of singletons in E. It is clear that $c = \inf_{K\in S}\max_{u\in K} I(u)$, and for $\overline{\varepsilon} > 0$, the deformation map $\eta : E \to E$ is supported around $I^{-1}([c-\overline{\varepsilon},\ c+\overline{\varepsilon}])$ and $\eta(A_{c+\varepsilon}) \subseteq A_{c-\varepsilon}$ for some $\varepsilon \in (0, \overline{\varepsilon})$.

Find some $z \in E$ such that $I(z) \leq c+\varepsilon$, then $\max \eta(\{z\} \leq c-\varepsilon$, contradicting the claim that c is the global minimum. ■

Proposition 7.6

Let E, I, ρ, α, e be as in Proposition 7.3, set $W = \{B \overset{\circ}{\subseteq} E,\ 0 \in B,\ e \notin \overline{B}\}$ as the open neighbourhoods about 0 whose closure separates 0 from e. Then,

$$b = \sup_{B\in W} \inf_{u\in\partial B} I(u)$$

is a critical value of I with $\alpha \leq b \leq c$.

Proof. First we show $\alpha \leq b \leq c$. We note that the infimum is taken over points on the boundary of B. The open ball

$B_\rho \in W$, and because $I|_{\partial B_\rho} \geq \alpha$, we see that $\alpha \leq b$.

To bound b from above, Lemma 7.7 shows that $g([0, 1]) \cap \partial B \neq \varnothing$ for any $g \in \Gamma_0^e$ and $B \in W$. Let w be a point of intersection as described above. Since w is on the boundary of B, $\inf_{z\in\partial B} I(z) \leq I(w)$, and since w is a point on the curve g, $I(w) \leq \sup_{t\in[0,1]} I(g(t))$. See **??** To show that b is a critical value, we can use an alternate $\hat{\eta}$ that moves along the positive gradient instead of the negative gradient of I. The image $\hat{\eta}(B)$ is open in E because $\hat{\eta}$ is a homeomorphism. ∎

Lemma 7.7
Let $g \in \Gamma_0^e = \{g \in C([0, 1], E), g(0) = 0, g(1) = e\}$, and $B \in W = \{B \overset{\circ}{\subseteq} E, 0 \in B, e \notin \overline{B}\}$. Then $X = g([0, 1])$ intersects the boundary of B.

Proof. Both ideas use the fact that given any open set U, we can split the space into the disjoint sum

$$E = U + \partial U + \overline{U}^c.$$

The first idea is rather complicated and uses the continuity of g to send boundary points back to boundary points in $[0, 1]$. Because $e \notin \overline{B}$, we can draw some $\delta > 0$ about the point e, and by continuity of the curve, there exists $\delta' > 0$ such that for all $t \in [1-\delta', 1]$, $g(t) \in B(\delta, e) \subseteq \overline{B}^c = B^{co}$. To this, let $\Sigma = \{t \in [0, 1], g(t') \notin \overline{B}\ \forall t' \in [t, 1]\}$ be the set of points such that once you pass $t \in \Sigma$, the curve no longer goes back into $\overline{B}$. This set is non-empty, bounded below because $t \notin \Sigma$ for sufficiently small t. Let $t_0 = \inf\Sigma > 0$, we can exhaust all of the possibilities.

If $t_0 \in B$, by continuity: $g(t_0+\delta') \in B$. If $t_0 \notin \overline{B}$, by continuity again $g(t_0-\delta') \notin \overline{B}$ and we conclude $g(t_0) \in \partial B$. The second idea is much more straight forward, since $X = g([0, 1])$ is connected, it cannot be written as the disjoint union of two non-empty open subsets $B \cap X$ and $B^{co} \cap X$, and we conclude $\partial B \cap X \neq \varnothing$. ∎

7.1. Rabinowitz 4: The Saddle Point Theorem

In this section, we introduce the Brouwer and its infinite-dimensional analogue, the Leray-Schauder degree. To begin, $\mathcal{O} = \{\text{open and bounded subsets of } \mathbb{R}^n\}$. Fix

Definition 7.8 Brouwer Degree

7.2. Rabinowitz 5: Some Generalizations of the Mountain Pass Theorem

7.3. Construction of pg field

The following lemma is needed in Proposition B.1 in [Rab86].

Lemma 7.9
For $r, s \in [1, +\infty)$, for every $a_1, a_2 > 0$, there exists $a_3 > 0$ where $(a_1 + a_2|z|^{r/s})^s \leq a_3(1+|z|^r)$.

Proof. We can prove the existence of such an a_3 by considering the behaviour of the quotient $g(t) = (a_1 + a_2|t|^{r/s})^s/(1+|t|^r)$ for $t \to 0$ and $t \to +\infty$. By inspection $g(t)$ is continuous at 0 with $g(0) = a_1^s$. For $t > 0$, $g(t) = (a_1|t|^{-r/s}+a_2)^s/(|t|^{-r}+1)$ which tends to a_2^s as $t \to \infty$. Since $\sup_{t\in[0,1]\cup[n,+\infty)} |g(t)|$ is finite for large n, we conclude that $\sup_t \geq 0|g(t)| = a_3 < +\infty$. ∎

8. Differential Equations

9. Radon Measures

In this section, X will be a LCH space.

Definition 9.1
Let $U \overset{\circ}{\subseteq} X$ and K compact. If $f \in C_c(X)$, $0 \leq f \leq 1$ and $\operatorname{supp}(f) \subseteq U \overset{\circ}{\subseteq} X$, we write $f \lesssim U$.

$$\Gamma_1(K) = \{f \in C_c(X),\ f \geq \chi_K\}$$
$$\Gamma_2(U) = \{f \in C_c(X),\ f \lesssim U\}.$$

If $E \subseteq X$, we write $\kappa(E) = \{K \subseteq E,\ K \text{ is compact}\}$ and $\kappa = \kappa(X)$.

The set of functions $\Gamma_1(K)$ and $\Gamma_2(U)$ satisfy the usual monotonicity and their geometric properties as well, which we will summarize. If $K \subseteq K'$, then $\Gamma_1(K') \subseteq \Gamma_1(K)$, and if $U \subseteq U'$, then $\Gamma_2(U) \subseteq \Gamma(U')$. If $f_1 \in \Gamma_1(K)$, $f_2 \in \Gamma_2(U)$,

$$\{f \in C_c(X),\ f \geq f_1\} \subseteq \Gamma_1(K)$$
$$\{f \in C_c(X),\ 0 \leq f \leq f_2\} \subseteq \Gamma_2(U).$$

Γ_1 is closed under addition and dilations by $c \geq 1$, while Γ_2 is a convex cone and is closed under addition whenever the supports are disjoint, that is

If $f_1, f_2 \in \Gamma_2(U)$, and $\operatorname{supp}(f_1) \cap \operatorname{supp}(f_2) = \varnothing$, then $f_1 + f_2 \in \Gamma_2(U)$.

We call the second family of functions, the subordinates of U, and will act as a family of lower approximants to $\mu(U)$ under the positive linear functional I_μ.

We note that if K is compact and $K \subseteq U \overset{\circ}{\subseteq} X$, and $f \in \Gamma_1(K) \cap \Gamma_2(U)$, then $f = 1$ on K, $0 \leq f \leq 1$, and $\operatorname{supp}(f) \subseteq U$; which is precisely the statement of Urysohn's Lemma.

Lemma 9.2 Urysohn (LCH)

If K compact, $K \subseteq U \overset{\circ}{\subseteq} X$, then there exists $f \in \Gamma_1(K) \cap \Gamma_2(U)$.

Proof. See Folland 4.32. ■

The following is a restatement of Urysohn's Lemma, which means that the family of compact subsets of $U \overset{\circ}{\subseteq} X$ satisfy a certain 'cofinality' condition.

Corollary 9.3

If $\{K_\alpha\} \subseteq \kappa(U)$, where U is open in X, to each K_α there exists a compact superset K'_α and $f_\alpha \in \Gamma_1(K_{\alpha'}) \cap \Gamma_2(U)$.

Definition 9.4

$\mathbb{R}^+ = [0, +\infty)$, $C_c(X)^+ = C_c(X, \mathbb{R}^+)$.

Definition 9.5

A linear functional $I : C_c(X) \to \mathbb{C}$ (not necessarily continuous) is called a positive linear functional if I maps $C_c(X)^+$ into $\mathbb{R}^+$.

Proposition 7.1 in Folland says that given a positive linear functional, for any compact K, the restriction of I onto the subspace of functions in $C_c(X)$ with $\operatorname{supp}(f) \subseteq K$ is a continuous linear functional. This is some kind of 'local' continuity, if we would like to equip $C_c(X)$ with a topology. The proof of Folland 7.1 exploits the 'order' structure for positive linear functionals, and a certain 'subtraction' trick borrowed from the proof of the dominated convergence theorem.

Note 9.6

Real-valued quantities, if bounded below can be 'transported' through I if some positivity condition is satisfied.

We can abstract this further. Let X be an ordered vector space and I a linear functional on X that satisfies some positivity condition. See Krein Extension Theorem for more. Of course this order structure has to respect scaling by positive numbers.

Proposition 9.7 Folland 7.1

Let I be a positive linear functional, then for all compact K, there exists $C_K > 0$ such that $|I(f)| \le C_K \|*\| f_u$ for all $f \in C_c(X)$ with $\operatorname{supp}(f) \subseteq K$.

Proof. Suppose f is real valued, because every complex-valued function can be decomposed $f = f_1 + i f_2$ for $f_j \in C_c(X, \mathbb{R})$. The order structure of $C_c(X, \mathbb{R})$ and $\mathbb{R}$ can be exploited as follows, using a 'subtraction 'trick.

$$(\|*\| f_u \phi - f) \in C_c(X, \mathbb{R}_+).$$

So that $\|*\| f I(\phi) \ge I(f)$ by positivity and homogeneity. Replacing f with $-f$ reads

$$|I(f)| \le I(\phi)\|*\| f_u.$$

Set $C_K = 2I(\phi)$. ■

Definition 9.8

Let μ be a Borel measure on X, we say that $E \in \mathcal{M}$ is

inner-regular $\mu(E) = \sup\{\mu(K),\ K \subseteq E,\ K \text{ is compact}\}$
outer-regular $\mu(E) = \inf\{\mu(U),\ E \subseteq U \overset{\circ}{\subseteq} X\}$.

A Radon measure on X is a Borel measure that is finite on compact sets, inner-regular on open sets, and outer-regular on all Borel sets.

Proposition 9.9

Let μ be a Borel measure, that is finite on compact sets. Then $C_c(X) \subseteq L^1(\mu)$.

Proof. Given $f \in C_c(X)$, we have the infinity bound $\|f\|_u$, along with the 1-bound χ_K where $K = \operatorname{supp}(f)$. So $\int |f| d\mu \le \|f\|_u \mu(K)$. ■

Definition 9.10

If μ is a Radon measure on X, we define $I_\mu : C_c(X) \to \mathbb{C}$ by

$$I_\mu(f) = \int f \, d\mu.$$

Proposition 9.11

If μ is a Radon measure, then I_μ is a positive linear functional. For every open set U,

$$\mu(U) = \sup\{I_\mu(f),\ f \in \Gamma(U)\} \tag{4}$$

Moreoever, the correspondence $\mu \mapsto I_\mu$ completely characterizes μ: if ν is another Radon measure and $I_\mu = I_\nu$, then $\mu = \nu$.

Proof. It is clear that I_μ is a positive linear functional. To prove Equation (4), to every open set U we assign two functionals:

plf $I_\mu : \Gamma_2(U) \to [0, +\infty)$,
compact subsets $\mu : \kappa(U) \to [0, +\infty)$.

Because μ is a Radon measure, it is inner-regular on open subsets, so that $\mu(U) = \sup_{K \in \kappa(U)} \mu(K)$. If $\{K_n\} \subseteq \kappa(U)$ is a sequence that increases to $\mu(U)$ (meas. μ), we can use Corollary 9.3 to obtain $f_n \in \Gamma_1(K_n) \cap \Gamma_2(U)$,

$$\mu(K_n) \le I_\mu(f_n) \le \mu(U).$$

So that our sequence of functions $f_n \nearrow \mu(U)$ (meas. I_μ), and Eq. (4) is proven. If ν is another Radon measure such that $I_\mu = I_\nu$, then the previous discussion shows that they agree on open sets, and by outer-regularity: $\mu = \nu$. ■

9.1. Riesz Representation Theorem

We turn to the Riesz Representation theorem for positive linear functionals. Later we will introduce other versions which will generalize this to non-positive linear functionals.

If I is any positive linear functional, and K compact, the infimum over all such $C_K > 0$ in Proposition 9.7 can be thought of some kind of 'measure' on the compact set K. On the other hand, if $U \overset{\circ}{\subseteq} X$,

$$\sup_{K\in\kappa(U)} \sup\left\{|I(\phi)|/\|*\|\phi_u,\ \phi \in C_c(X),\ \mathrm{supp}(\phi) \subseteq K\right\}$$

which by homogeneity and linearity of I, is the same as

$$\sup_{\phi\in\Gamma_2(U)} I(\phi).$$

Proposition 9.12 Riesz (Positive Linear Functionals)

Every positive linear functional I on $C_c(X)$ gives rise to a unique Radon measure μ such that $I = I_\mu$.

sup. over subordinates $\mu(U) = \sup\{I(\phi),\ \phi \in \Gamma_2(U)\}$, for $U \overset{\circ}{\subseteq} X$.
inf. over superordinates $\mu(K) = \inf\{I(\psi),\ \psi \in \Gamma_1(K)\}$ for $K \in \kappa$.

Note 9.13

The regularity conditions of I_μ and the regularity conditions on a Radon measure, feels like the meeting of a real-variable limit.

$$\inf_{U \overset{\circ}{\subseteq} X,\ E\subseteq U} \sup_{\phi\in\gamma_1(U)} I(\phi) = \sup_{K\subseteq E,\ K \text{ compact}} \inf_{\phi\in\Gamma_2(K)} I(\phi)$$

Note 9.14

More abstract approximation framework. Let Ω be a set and $\Sigma_{(k)} \subseteq \Omega'$ for $k = 0, 1$ with

$$J_{(k)} : \Sigma_{(k)} \to [0,\ +\infty].$$

Suppose further to every $x \in \Omega$, we are given a non-empty $\Sigma_{(k)}(x) \subseteq \Sigma_{(k)}$.

$$I^{\pm}_{(k)} : \Omega \to [0,\ +\infty]$$

$$I^{+}_{(k)}(x) = \sup_{\phi\in\Sigma_{(k)}(x)} J_{(k)}(\phi)$$
$$I^{-}_{(k)}(x) = \inf_{\phi\in\Sigma_{(k)}(x)} J_{(k)}(\phi)$$

10. Fourier Analysis

10.1. Weak Topologies

Recall that if $\{Y_\alpha\}$ is a family of topological spaces, and X is any set, the weak topology induced by a family of functions $\mathcal{F} = \{f_\alpha : X \to Y_\alpha\}$ is the topology $\mathcal{T}$ consisting precisely of $\{\varnothing,\ X\}$ and unions of finite intersections of

$$\mathcal{E}(\mathcal{F}) = \{f_\alpha^{-1}(U_\alpha),\ U_\alpha \overset{\circ}{\subseteq} Y_\alpha,\ \alpha \in A\}.$$

It is the weakest topology on X that makes every single f_α continuous. We will prove this.

Proposition 10.1

Let X, Y_α, $\mathcal{F}$, and let X have the weak topology $\mathcal{T}$ as described above. With this topology, every single f_α is continuous. And if X is equipped with another topology (say $\mathcal{T}'$) such that $f_\alpha \in C(\mathcal{T}',\ Y_\alpha)$ then $\mathcal{T} \subseteq \mathcal{T}'$.

Proof. Fix f_α, and $U_\alpha \overset{\circ}{\subseteq} Y_\alpha$, then $f_\alpha^{-1}(U_\alpha) \in \mathcal{T}$. So that f_α must be continuous. Moreover, if $\mathcal{T}'$ is another topology on X such that $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{T}'$, then it must contain $\{\varnothing,\ X\}$, along with the unions of finite intersections of members of $\mathcal{E}(\mathcal{F})$. ■

A

Ex. 4.17 If X is a set, $\mathcal{F}$ is a collection of $\mathbb{R}$-valued functions on X, $\mathcal{T}(\mathcal{F})$ the weak topology on X by this family of functions. Then $\mathcal{T}(\mathcal{F})$ is Hausdorff if the family $\mathcal{F}$ separates points: meaning for every $x \neq y$, there exists $f \in \mathcal{F}$, $f(x) \neq f(y)$.

Ex. 4.32 A topological space is Hausdorff iff every net in X converges at most to one point.

Ex 4.33 Let $\langle x_\alpha\rangle_{\alpha\in A}$ be a net in a topological space and let $E_\alpha = \{x_\beta,\ \beta \gtrsim \alpha\}$ denote the α-tail of the net, which is cofinal. Then x is a cluster point of the net iff $x \in \bigcap_{\alpha\in A} \overline{E}_\alpha$.

Ex 4.34 If X has the weak topology generated by $\mathcal{F} = \{f : X \to Y\}$ where Y is a topological space, then a net $\langle x_\alpha\rangle_{\alpha\in A} \subseteq X$ converges to x iff for every $f \in \mathcal{F}$, the $\langle f(x_\alpha)\rangle_{\alpha\in A} \to f(x)$.

10.1.1 Frechet Spaces

Definition 10.2

TVS a vector space with a topology such that the addition and scalar multiplication maps $a : X \times X \to X$, $m : \mathbb{C} \times X \to X$ are continuous.
cauchy net $\langle x_i \rangle_{i \in I}$, is cauchy whenever the product-directed net $\langle x_i - x_j \rangle_{(i,j) \in I \times I}$ converges to 0.
complete TVS every Cauchy net converges.
Frechet space a Cauchy-complete, Hausdorff topological vector space whose topology is defined by a countable family of seminorms.

Proposition 10.3

Ex. 4.17 If X is a set, $\mathcal{F}$ is a collection of $\mathbb{R}$-valued functions on X, $\mathcal{T}(\mathcal{F})$ the weak topology on X by this family of functions. Then $\mathcal{T}(\mathcal{F})$ is Hausdorff if the family $\mathcal{F}$ separates points: meaning for every $x \neq y$, there exists $f \in \mathcal{F}$, $f(x) \neq f(y)$.
Ex. 4.32 A topological space is Hausdorff iff every net in X converges at most to one point.
Ex 4.33 Let $\langle x_\alpha \rangle_{\alpha \in A}$ be a net in a topological space and let $E_\alpha = \{x_\beta,\ \beta \gtrsim \alpha\}$ denote the α-tail of the net, which is cofinal. Then x is a cluster point of the net iff $x \in \bigcap_{\alpha \in A} \overline{E}_\alpha$.
Ex 4.34 If X has the weak topology generated by $\mathcal{F} = \{f : X \to Y\}$ where Y is a topological space, then a net $\langle x_\alpha \rangle_{\alpha \in A} \subseteq X$ converges to x iff for every $f \in \mathcal{F}$, the $\langle f(x_\alpha) \rangle_{\alpha \in A} \to f(x)$.
Ex. 5.44 If X is a first countable TVS, every Cauchy sequence converges iff every Cauchy net converges.

10.2. Schwartz Space

Definition 10.4

Let $E \subseteq \mathbb{R}^n$ and $U \mathring{\subseteq} \mathbb{R}^n$. $C^\infty(U) = \{f : U \to \mathbb{C},\ \text{smooth}\}$, and $C_c^\infty(E) = \{f \in C_c^\infty(\mathbb{R}^n, \mathbb{C}),\ \operatorname{supp}(f) \subseteq E\}$. By default: $C^\infty = C^\infty(\mathbb{R}^n)$, and $C_c^\infty = C_c^\infty(\mathbb{R}^n)$.

$$\mathcal{S} = \{f \in C^\infty,\ \|*\| f_{(N,\alpha)} < +\infty\},$$

where $\|*\| f_{(N,\alpha)} = \sup_{x \in \mathbb{R}^n} (1 + |x|)^N |\partial^\alpha f(x)|$.

Similar to how we view weak L^p as the space of measurable functions whose distribution function decays sufficiently quickly against a polynomial, if $f \in \mathcal{S}$ and α is any multi-index

$$|\partial^\alpha f(x)| \leq \|*\| f_{(N,\alpha)} (1 + |x|)^{-N}$$

The polynomial factor $A^{-N}(x) = (1 + |x|)^{-N}$ is continuous and hence bounded about the origin. If p is a usual exponent, so that the norms of L^p are given by integration, $|x|^{-(n+\varepsilon)} \chi_{B^c} \in L^1(\mathbb{R}^n)$ for $B = \{x \in \mathbb{R}^n,\ |x| < \rho\}$ for some $c > 0$:

$$|\partial^\alpha f(x)|^p \leq \|*\| f_{(N,\alpha)}^p (1 + |x|)^{-(Np)}$$

It suffices to pick $Np > n$, so that

$$\|*\| \partial^\alpha f_p \leq \|*\| f_{(N,\alpha)} \|*\| (1 + |x|)^{-N}{}_p.$$

It is also clear that $\partial^\alpha f \in L^\infty$.

Proposition 10.5

Folland 8.2
Folland 8.3

10.3. Lp and Bounded Convergence

Definition 10.6

10.4. Translations and Convolutions

Definition 10.7

If f, g measurable functions on $\mathbb{R}^n$, for all $y \in \mathbb{R}^n$, $(\tau_y f)(x) = f(x - y)$. We also define the convolution between f, g as the function in x, whenever it converges for $a.e\ x$:

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y) g(y) dy.$$