Notes on Differential Equations

Wyatt Reeves

March 2021

Abstract

I'm compiling these notes for a Directed Reading Program I'm currently doing. These are for personal use so buyer beware!

1 Banach Spaces

One of the basic ways that we'll prove the existence of solutions to differential equations is to construct a Cauchy sequence of "approximate solutions" which come closer and closer to solving the specified differential equation. For this sequence to converge, our underlying space of functions will need to be complete: the theory of Banach spaces is the natural setting to study complete function spaces.

todo: rename base field F for consistency with later sections.

Definition 1.1. A **Normed Vector Space** is a pair $(V, ||\cdot||)$, where V is a vector space over field K which is either $\mathbf{R}$ or $\mathbf{C}$ and $||\cdot|| : V \to \mathbf{R}_{\geq 0}$ is a function satisfying the following three conditions:

1. $||v|| = 0$ if and only if $v = 0$.
2. $||\lambda v|| = |\lambda| \, ||v||$ for all $\lambda \in K$.
3. $||v + w|| \leq ||v|| + ||w||$.

In what follows, we will often abuse notation and use V to refer to both the NVS $(V, ||\cdot||)$ and its underlying vector space. We can naturally give a NVS $(V, ||\cdot||)$ the structure of a metric space by decreeing that the distance between any two vectors $v, w \in V$ is $||v - w||$ (it is left as an exercise to check that this distance function satisfies the appropriate axioms for a metric space).

todo: at some point I should probably mention/prove that the norm is continuous

Definition 1.2. A **Banach space** is a NVS V which is complete with respect to the metric associated to its norm.

Example 1.3. The vector space $\mathbf{R}^n$ equipped with the norm $||v|| = \sqrt{v_1^2 + \ldots + v_n^2}$ a Banach space (this should be familiar from undergraduate analysis). Surprisingly, it is the case that every *finite dimensional* Banach space is in some sense "basically the same" as this example. There are, however, many different *infinite dimensional* Banach spaces. We will meet many of them in these notes.

Definition 1.4. For a compact topological space K and a Banach space V, let $C(K, V)$ denote the vector space of continuous functions $K \to V$.

Theorem 1.5. *$C(K, V)$ becomes a Banach space when equipped with the norm $||\cdot||_\infty$, where*

$$||f||_\infty = \sup_{x \in K} ||f(x)||_V$$

Proof. For any $f \in C(K, V)$, because K is compact and f is continuous, we know that $||f(K)||_V$ is also compact and therefore a bounded subset of $\mathbf{R}$. This shows that $||f||_\infty$ is some finite number, so $||\cdot||_\infty$ is a well-defined function from $C(K, V)$ to $\mathbf{R}_{\geq 0}$. A straightforward application of the basic properties of sup shows that $(C(K, V), ||\cdot||_\infty)$ is a NVS (this is left as an exercise).

To see that $C(K, V)$ is complete with respect to $||\cdot||_\infty$, suppose that $\{f_n\}$ is a Cauchy sequence with respect to $||\cdot||_\infty$. In other words, $||f_i - f_j||_\infty$ goes to zero as i and j go to ∞. Then for every point $x \in K$, the sequence $\{f_n(x)\}$ is a Cauchy sequence in V, since

$$\begin{aligned} ||f_i(x) - f_j(x)||_V &\leq \sup_{x \in K} ||f_i(x) - f_j(x)||_V \\ &= ||f_i - f_j||_\infty \end{aligned}$$

so $||f_i(x) - f_j(x)||_V$ also goes to zero as i, j go to ∞. Since V is complete, this implies that $f_n(x)$ converges to some value y_x as n goes to ∞. Define $f : K \to V$ to be the function such that $f(x) = y_x$ for all $x \in K$. To conclude, we want to show that f_n converges to f and that f is continuous.

To show that f_n converges to f, fix $\varepsilon > 0$ and let N be large enough that $||f_i - f_j||_\infty < \varepsilon$ for all $i, j \geq N$. Then for any $k \geq N$ and any $x \in K$ we know that

$$\begin{aligned} ||f(x) - f_k(x)||_V &= \lim_{i \to \infty} ||f_i(x) - f_k(x)||_V \\ &\leq \lim_{i \to \infty} \sup_{x \in K} ||f_i(x) - f_k(x)||_V \\ &= \lim_{i \to \infty} ||f_i - f_k||_\infty \\ &\leq \varepsilon \end{aligned}$$

so

$$||f - f_k||_\infty \leq \sup_{x \in K} ||f(x) - f_k(x)||_V \leq \varepsilon.$$

This shows that $||f - f_k||_\infty$ goes to 0 as k goes to ∞ as desired.

To see that f is continuous, fix $x_0 \in K$ and $\varepsilon > 0$. By our previous argument there is some k such that $||f - f_k||_\infty < \varepsilon/3$. Since f_k is continuous, there is some open neighborhood U of x_0 such that $||f(x) - f(x_0)||_V < \varepsilon/3$ for any $x \in U$. We therefore know that

$$\begin{aligned} ||f(x) - f(x_0)||_V &\leq ||f(x) - f_k(x)||_V + ||f_k(x) - f_k(x_0)||_V + ||f_k(x_0) - f(x_0)||_V \\ &< ||f - f_k||_\infty + \varepsilon/3 + ||f_k - f||_\infty \\ &< \varepsilon. \end{aligned}$$

Since we can find such a U for any $\varepsilon > 0$, this implies that f is continuous at x_0, and since this is true for every x_0 in K, this implies that f is continuous. We have therefore shown that every Cauchy sequence f_n in $C(K, V)$ converges to an element f of $C(K, V)$, completing our proof that $C(K, V)$ is a Banach space. □

We will now prove the first of several important theorems that will allow us to prove the existence of solutions to differential equations in the context of Banach spaces.

Definition 1.6. If (X, d) is a metric space and $0 \leq \alpha < 1$, an $\alpha-$**contraction mapping** is a map $T : X \to X$ such that $d(Tx, Ty) \leq \alpha d(x, y)$ for every $x, y \in X$. A **contraction mapping** is a map that is an $\alpha-$contraction mapping for some $0 \leq \alpha < 1$.

Note that every contraction mapping is automatically continuous (this is an exercise).

Theorem 1.7 (Banach Fixed Point Theorem)**.** *Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction mapping. Then T has a unique fixed point.*

Proof. First, to see that the fixed point is unique, suppose that x and y are both fixed points of T, so $Tx = x$ and $Ty = y$. Then

$$d(x, y) = d(Tx, Ty) \leq \alpha d(x, y).$$

Since $\alpha < 1$, this implies that $d(x, y)$ must be 0, so $x = y$. Now to show the existence of a fixed point we must use the fact that X is complete. Pick any point $x \in X$ and consider the sequence $x_n = T^n x$. We will show that this sequence is Cauchy and it converges to the fixed point of T. Let δ denote $d(x, Tx)$. We want to bound $d(T^i x, T^j x)$ as i and j grow large. If $i = j$ then $d(T^i x, T^i x) = 0$, so suppose without loss of generality that $i > j$. Then

$$\begin{aligned} d(T^i x, T^j x) &\leq d(T^i x, T^{i-1} x) + d(T^{i-1} x, T^{i-2} x) + \ldots + d(T^{j+1} x, T^j x) \\ &\leq \alpha^{i-1}\delta + \alpha^{i-2}\delta + \ldots \alpha^j \delta \\ &= \alpha^j \delta \sum_{k=0}^{i-j-1} \alpha^k \\ &\leq \frac{\alpha^j \delta}{1 - \alpha}. \end{aligned}$$

For any $\varepsilon > 0$, since $\alpha < 1$ there is some large N such that for all $j \geq N$ we know that

$$\alpha^j < (1-\alpha)\varepsilon/\delta$$

so for all $i, j \geq N$ we see that $d(T^i x, T^j x) < \varepsilon$. Since this is true for all $\varepsilon > 0$ we see that the sequence $x_n = T^n x$ is Cauchy. Because we have assumed that (X, d) is a complete metric space, this implies that x_n converges to some value y. To see that y is a fixed point of T, note that since T is continuous,

$$\begin{aligned} Ty &= \lim_{n\to\infty} Tx_n \\ &= \lim_{n\to\infty} T^{n+1}x \\ &= \lim_{n\to\infty} x_{n+1} \\ &= y. \end{aligned}$$

We have therefore proven the existence and uniqueness of a fixed point of T in X, as desired. □

Exercises

Exercise 1.8. Given a NVS $(V, ||\cdot||)$, prove that the function $d(v, w) = ||v - w||$ satisfies the axioms of a distance function on a metric space.

Exercise 1.9. Prove that $(C(K), ||\cdot||_\infty)$ is a NVS for any compact topological space K.

Exercise 1.10. Prove that every contraction mapping is continuous.

Exercise 1.11. Prove that the mapping $Tx = x/2 + 1/x$ is a contraction mapping on $[1, \infty)$. Show that $\sqrt{2}$ is a fixed point of T. Use the Banach fixed point theorem to describe an iterative algorithm for computing $\sqrt{2}$.

Exercise 1.12. Show that if T is an α-contraction mapping on a complete metric space (X, d) and x is any point in X, the distance between x and the unique fixed point y of T is bounded above by

$$\frac{\delta}{1-\alpha},$$

where $\delta = d(x, Tx)$.

2 The Picard-Lindelöf Theorem

(todo) Give example of $x^2 \partial_x f + f = 0$

Definition 2.1. An **Initial Value Problem** $\mathcal{P}$ is the data (U, ξ, p, t_0), where U is an open subset of $\mathbf{R}^n$, $\xi : U \to \mathbf{R}^n$ is a continuous vector field on U, p is a point in U, and $t_0 \in \mathbf{R}$.

Definition 2.2. A **solution** of the initial value problem $\mathcal{P} = (U, \xi, p, t_0)$ is the data of an open interval J containing t_0 and a differentiable function $\gamma : J \to U$ such that $\gamma(t_0) = p$ and for all $t \in J$ it is the case that $\gamma'(t) = \xi(\gamma(t))$.

Definition 2.3. Given an initial value problem $\mathcal{P} = (U, \xi, p, t_0)$ and a closed interval I containing t_0, the **Picard-Lindelöf Integral Operator** $T_{\mathcal{P},I} : C(I, U) \to C(I, \mathbf{R}^n)$ is the map sending $\gamma \in C(I, U)$ to $T_{\mathcal{P},I}\gamma$, where

$$(T_{\mathcal{P},I}\gamma)(s) = p + \int_{t_0}^{s} \xi(\gamma(t))\ dt.$$

Lemma 2.4. *Let $\mathcal{P} = (U, \xi, p, t_0)$ be an initial value problem. Let J be an open interval containing t_0 and let $\gamma : J \to U$ be a continuous function. The following are equivalent:*

1. *γ is a solution to the initial value problem $\mathcal{P}$.*
2. *For every closed subinterval $I \subset J$ containing t_0, $\gamma|_I$ is a fixed point of the Picard Lindelöf integral operator $T_{\mathcal{P},I}$.*

Lemma 2.5. *Let $\mathcal{P} = (U, \xi, p, t_0)$ be an initial value problem. Let I be a closed interval containing t_0 and let $\gamma : I \to U$ be a continuous function. The following are equivalent:*

1. *γ is a fixed point of the Picard-Lindelöf integral operator $T_{\mathcal{P},I}$.*
2. *For every open subinterval $J \subset I$ containing t_0, $\gamma|_J$ is a solution to the initial value problem $\mathcal{P}$.*

Proof. Lemma 2.4 and Lemma 2.5 follow from the fundamental theorem of calculus. □

Theorem 2.6 (Picard-Lindelöf)**.** *Let $\mathcal{P} = (U, \xi, p, t_0)$ be a initial value problem, and suppose ξ is locally Lipschitz on U. Then there is some $\varepsilon > 0$ such that there exists a unique solution γ to the initial value problem $\mathcal{P}$ on $(t_0 - \varepsilon, t_0 + \varepsilon)$.*

Proof. Because ξ is locally Lipschitz, there is a constant L and a closed ball $B = \overline{B_R(p)}$ of radius $R > 0$ around p such that for all $x, y \in B$ we know

$$||\xi(x) - \xi(y)|| \leq L\, ||x - y||\,.$$

Also, since ξ is continuous and B is compact, we know that ξ is bounded on B: there is some M such that $||\xi(x)|| \leq M$ for all $x \in B$. Let a be any positive real number such that $a \leq R/M$ and $a < 1/L$ and let $I = [t_0 - a, t_0 + a]$. We wish to show that there is a unique function $\gamma : I \to U$ which is a fixed point of the Picard-Lindelöf integral operator $T_{\mathcal{P},I}$.

Using Lemma 2.4 and Lemma 2.5 this will be able to translate this result into an existence and uniqueness result for solutions to the initial value problem $\mathcal{P}$. Our strategy for showing that $T_{\mathcal{P},I}$ has a unique fixed point will be to use the Banach fixed point theorem.

Let $\gamma_0 : I \to \mathbf{R}^n$ be the constant function at p, so $\gamma_0(t) = p$ for all t. Let X be the closed ball of radius R around γ_0; by the definition of the norm on $C(I, \mathbf{R}^n)$ the functions in X are exactly the functions whose image lies in B. Also, because X is a closed subspace of a Banach space, X is itself complete. First, we claim that $T_{\mathcal{P},I}$ actually maps X to itself. To see this, note that

$$\begin{aligned} ||T_{\mathcal{P},I} f - \gamma_0||_\infty &= \sup_{t \in I} \left|\left| \int_{t_0}^t \xi(f(t)) \; dt \right|\right| \\ &\leq \sup_{t \in I} \int_{t_0}^t ||\xi(f(t))|| \; dt \\ &\leq \sup_{t \in I} \int_{t_0}^t M \; dt \\ &\leq Ma \\ &\leq R \end{aligned}$$

where the third-to-last inequality follows because $f \in X$ so we know that $f(t) \in B$ for all $t \in I$ so $||\xi(f(t))|| \leq M$ by our definition of M. Now we wish to show that $T_{\mathcal{P},I}$ is a contraction mapping. To see this, let f and g be in X. Then

$$\begin{aligned} ||T_{\mathcal{P},I} f - T_{\mathcal{P},I} g||_\infty &= \sup_{t \in I} \left|\left| \int_{t_0}^t \xi(f(t)) - \xi(g(t)) \; dt \right|\right| \\ &\leq \sup_{t \in I} \int_{t_0}^t ||\xi(f(t)) - \xi(g(t))|| \; dt \\ &\leq \sup_{t \in I} \int_{t_0}^t L \, ||f(t) - g(t)|| \; dt \\ &\leq \sup_{t \in I} \int_{t_0}^t L \, ||f - g||_\infty \; dt \\ &\leq aL \, ||f - g||_\infty \end{aligned}$$

which shows that $T_{\mathcal{P},I}$ is a contraction mapping since $aL < 1$ by our choice of a. The Banach fixed point theorem, Theorem 1.7, then tells us that $T_{\mathcal{P},I} : X \to X$ has a unique fixed point γ.

Now for any $\varepsilon < a$, letting $J = (t_0 - \varepsilon, t_0 + \varepsilon)$, we know that $\gamma|_J$ is a solution to the initial value problem $\mathcal{P}$ by Lemma 2.5. To show that $\gamma|_J$ is the unique solution on J, suppose that σ is any solution to the initial value problem $\mathcal{P}$ on J. Note that for any $0 < a' < \varepsilon < a$, if we let $I' = [t_0 + a', t_0 - a']$, then our above argument shows that $T_{\mathcal{P},I'}$ has a unique fixed point. By Lemma 2.4, $\sigma|_{I'}$ is a fixed point of $T_{\mathcal{P},I'}$, but we also know that $\gamma|_{I'}$ is a fixed point of $T_{\mathcal{P},I'}$ since γ is a fixed point of $T_{\mathcal{P},I}$ and $I' \subseteq I$. The uniqueness of the fixed point of $T_{\mathcal{P},I'}$ then implies that $\gamma|_{I'} = \sigma|_{I'}$. Since this is true for every $a' < \varepsilon$ we see that $\gamma|_J = \sigma$. □

We now wish to show that the solutions to our initial value problem vary continuously as the initial condition varies.

todo: this lemma actually needs to be used in the proof of Picard-Lindelöf

Lemma 2.7. *Let $\mathcal{P}_p = (U, \xi, p, t_0)$ be an initial value problem. Let M be the maximum of $||\xi||$ on the closed ball of radius R around p. For $a < M/R$, let $\delta = R - Ma$. Then if $||q - p|| < \delta$ and γ_q is a solution to the initial value problem $\mathcal{P}_q = (U, \xi, q, t_0)$ on $J = (t_0 - a, t_0 + a)$, it is the case that $\gamma_q(t) \in \overline{B_R(p)}$ for all $t \in J$.*

Proof. Suppose there is some $t \in J$ such that $||\gamma_q(t) - p|| > R$. Without loss of generality suppose that $t > 0$. Then because γ_q is continuous there is some smallest t' such that $||\gamma_q(t') - p|| = R$ but for any $t_0 \leq s < t'$ it's the case that $||\gamma_q(s) - p|| < R$. Then

$$\begin{aligned}
\big|\big|\gamma_q(t') - p\big|\big| &\leq \big|\big|\gamma_q(t') - q\big|\big| + ||q - p|| \\
&< \big|\big|\gamma_q(t') - \gamma_q(t_0)\big|\big| + \delta \\
&= \left|\left|\int_{t_0}^{t'} \gamma'(t)dt\right|\right| + \delta \\
&= \left|\left|\int_{t_0}^{t'} \xi(\gamma(t))dt\right|\right| + \delta \\
&\leq \int_{t_0}^{t'} Mdt + \delta \\
&= (t' - t_0)M + \delta \\
&< aM + \delta \\
&< R
\end{aligned}$$

where the crucial step here is that since $\gamma(t)$ is in $\overline{B_R(p)}$ for all $t_0 \leq t \leq t'$ we know that $||\xi(\gamma(t))|| \leq M$. We therefore obtain a contradiction, since by our assumption $R < R$. We then conclude that $||\gamma_q(t) - p|| \leq R$ for all $t \in J$. □

Proposition 2.8. *Let $\mathcal{P}_p = (U, \xi, p, t_0)$ be an initial value problem where ξ is locally Lipschitz continuous. There is some $a > 0$ such that for all $\varepsilon > 0$ there is a $\delta > 0$ such that for all q with $||q - p|| < \delta$, if γ_p and γ_q are the solutions to the initial value problems $\mathcal{P}_p$ and $\mathcal{P}_q = (U, \xi, q, t_0)$, respectively, then $||\gamma_p - \gamma_q||_\infty < \varepsilon$ on $[t_0 - a, t_0 + a]$.*

Proof. Let M, R, and L be as in the proof of Theorem 2.6. Choose $a > 0$ such that $a < M/R$ and $a < 1/L$. Then for any $\varepsilon > 0$ choose $\delta > 0$ such that $\delta < R - Ma$ and $\delta < (1 - aL)\varepsilon$. □

todo: Prove continuous dependence of solution on initial conditions.

2.1 Exercises

Problem 1 Show that the initial value problem

$$f' = \frac{2}{3} f^{1/3}$$

and $f(0) = 0$ does not have a unique solution. Show that the function $x \mapsto \frac{2}{3} x^{1/3}$ is not locally Lipschitz at $x = 0$.

Problem 2 Apply the Picard-Lindelöf operator "by hand" to the initial value problem $f' = f$ and $f(0) = 1$ to compute the Taylor series for e^t.

3 L^p Spaces

todo: include motivation for PDEs.

todo: include proof that these are NVS.

Theorem 3.1. *$L^p(\Omega)$ is a Banach space for any measure space Ω and any $1 \leq p < \infty$.*

Proof. Let f_n be a Cauchy sequence in $L^p(\Omega)$. Using the Cauchy property (todo: expand?), choose a subsequence f_{n_k} such that

$$||f_{n_{k+1}} - f_{n_k}||_p \leq 2^{-k}.$$

Now let g_i be the nonnegative function on Ω such that

$$g_i = |f_{n_1}| + \sum_{k=1}^{i-1} |f_{n_{k+1}} - f_{n_k}|$$

and let $g = \lim_{i \to \infty} g_i$. Note that

$$\begin{aligned} ||g_i||_p &\leq ||f_{n_1}||_p + \sum_{k=1}^{i-1} ||f_{n_{k+1}} - f_{n_k}||_p \\ &= ||f_{n_1}||_p + \sum_{k=1}^{i-1} 2^{-k}. \end{aligned}$$

Moreover, g_i converges to g monotonically from below, so since $x \mapsto x^p$ is a monotone continuous function on $\mathbf{R}_{\geq 0}$ it's also the case that g_i^p converges to g^p monotonically from below. As a result, the Monotone Convergence Theorem shows that

$$\begin{aligned}\int_\Omega g^p &= \lim_{i\to\infty}\int_\Omega g_i^p \\ &= \lim_{i\to\infty} ||g_i||_p^p \\ &\leq \lim_{i\to\infty}\left(||f_{n_1}||_p + \sum_{k=1}^{i-1} 2^{-k}\right)^p \\ &= \left(1 + ||f_{n_1}||_p\right)^p \\ &< \infty.\end{aligned}$$

Since g^p is in $L^1(\Omega)$, it must be finite almost everywhere, so it must also be the case that g is finite almost everywhere. Because of this, for almost every $x \in \Omega$ we know that

$$\sum_{i=1}^{\infty} f_{n_{k+1}}(x) - f_{n_k}(x)$$

converges absolutely, since

$$\sum_{i=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| = g(x) - |f_{n_1}(x)| < \infty.$$

As a result, we can define a function f on Ω almost everywhere as

$$f = f_{n_1} + \sum_{i=0}^{\infty}\left(f_{n_{k+1}} - f_{n_k}\right).$$

Since $|f| \leq g$, we know $|f|^p \leq g^p$, so since $g \in L^1(\Omega)$ we know that $|f|^p \in L^1(\Omega)$, which implies that $f \in L^p(\Omega)$. We now want to show that our original sequence converges to f. Note that for any k,

$$\begin{aligned}||f - f_{n_k}||_p &= \left|\left|\sum_{i=k}^{\infty} f_{n_{i+1}} - f_{n_i}\right|\right|_p \\ &\leq \sum_{i=k}^{\infty} 2^{-i} \\ &= 2^{1-k}.\end{aligned}$$

so $||f - f_{n_k}||_p \to 0$ as $k \to \infty$. Since our original sequence f_n was Cauchy and since a subsequence f_{n_k} converges to f, we can conclude that f_n also converges to f (todo: expand?). □

4 Hilbert Spaces

Definition 4.1. An **Inner Product Space** is a pair $(V, \langle -, - \rangle)$ where V is a vector space over a field F which is either $\mathbf{R}$ or $\mathbf{C}$ and the inner product $\langle -, - \rangle : V \times V \to F$ is a function satisfying the following three conditions:

1. $\langle v, v \rangle \in \mathbf{R}_{\geq 0}$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.
2. $\langle v, w + \lambda u \rangle = \langle v, w \rangle + \lambda \langle v, u \rangle$ for all $\lambda \in K$.
3. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

Here $\overline{z}$ is the complex conjugate of z. Note that when $F = \mathbf{R}$, the third axiom of an inner product space just says that $\langle v, w \rangle = \langle w, v \rangle$, so in this situation the inner product is symmetric. To fix notation, given an inner product space $(V, \langle -, - \rangle)$ we will write $||v|| = \sqrt{\langle v, v \rangle}$. The notation is suggestive: we will see in Proposition 4.3 that this function is indeed a norm.

todo: give examples

Proposition 4.2 (Cauchy-Schwartz Inequality)**.**

$$|\langle v, w \rangle| \leq ||v|| \, ||w||$$

Proof. For any $\lambda \in K$, note that

$$\begin{aligned}
0 &\leq \langle v + \lambda w, v + \lambda w \rangle \\
&= \langle v, v \rangle + \langle \lambda w, v \rangle + \langle v, \lambda w \rangle + \langle \lambda w, \lambda w \rangle \\
&= ||v||^2 + \lambda \langle v, w \rangle + \overline{\lambda \langle v, w \rangle} + |\lambda|^2 \, ||w||^2 \\
&= ||v||^2 + 2 \operatorname{Re}(\lambda \langle v, w \rangle) + |\lambda|^2 \, ||w||^2 .
\end{aligned}$$

The smaller the term on the right is, the more interesting of a bound this gives us. The expression we want to minimize is a quadratic equation in λ, and the standard formula for the minimum gives the value

$$\lambda = \frac{-\overline{\langle v, w \rangle}}{||w||^2} = \frac{-\langle w, v \rangle}{||w||^2}.$$

For this value of λ, we see that

$$\begin{aligned}
0 &\leq ||v||^2 - 2\frac{|\langle v, w \rangle|^2}{||w||^2} + \frac{|\langle v, w \rangle|^2}{||w||^4} \, ||w||^2 \\
&= ||v||^2 - \frac{|\langle v, w \rangle|^2}{||w||^2}
\end{aligned}$$

so

$$|\langle v,w\rangle|^2 \le ||v||^2\,||w||^2$$

so

$$|\langle v,w\rangle| \le ||v||\,||w||$$

since both sides are positive. □

todo: relate this to Hölder's inequality

Proposition 4.3. *Given an inner product space* $(V, \langle -,-\rangle)$*, if we define* $||v|| = \sqrt{\langle v,v\rangle}$ *as above, then* $(V, ||\cdot||)$ *is a normed vector space.*

Proof. First note that $||v|| = 0$ if and only if $\langle v,v\rangle = 0$ if and only if $v = 0$. This verifies the first axiom of a normed vector space. Also note that

$$\begin{aligned} ||\lambda v|| &= \sqrt{\langle \lambda v, \lambda v\rangle} \\ &= \sqrt{|\lambda|^2 \langle v,v\rangle} \\ &= |\lambda|\,||v|| \end{aligned}$$

This verifies the second axiom. The last thing to show is the triangle inequality $||v+w|| \le ||v|| + ||w||$. It suffices to show the inequality after squaring both sides:

$$\begin{aligned} ||v+w||^2 &= \langle v+w, v+w\rangle \\ &= \langle v,v\rangle + 2\operatorname{Re}(\langle v,w\rangle) + \langle v,w\rangle \\ &\le ||v||^2 + 2\,|\langle v,w\rangle| + ||w||^2 \\ &\le ||v||^2 + 2\,||v||\,||w|| + ||w||^2 \\ &= (||v|| + ||w||)^2 \end{aligned}$$

where the first inequality here is simply the fact that $\operatorname{Re} z \le |z|$ for any $z \in \mathbf{C}$ and the second inequality is the Cauchy-Schwartz inequality, Proposition 4.2. □

Definition 4.4. A **Hilbert Space** is an inner product space $(V, \langle -,-\rangle)$ such that the associated normed vector space is a Banach space.

todo: give examples

Proposition 4.5 (Parallelogram Law)**.** *If* $(V, \langle -,-\rangle)$ *is an inner product space, then*

$$||x+y||^2 + ||x-y||^2 = 2\,||x||^2 + 2\,||y||^2$$

Proof. Since V is an inner product space, we can rewrite the norm in terms of the inner product:

$$\begin{aligned}||x+y||^2 + ||x-y||^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= ||x||^2 + 2\,\mathrm{Re}(\langle x,y\rangle) + ||y||^2 + ||x||^2 - 2\,\mathrm{Re}(\langle x,y\rangle) + ||y||^2 \\ &= 2\,||x||^2 + 2\,||y||^2\end{aligned}$$

□

Remark 4.6. Note that even though the statement of Proposition 4.5 can be phrased purely in terms of the norm on V, it is false for a general normed vector space. For example, take $\mathbf{R}^2$ with the ℓ^1 norm.

Proposition 4.7. *If K is a closed and convex subset of a Hilbert space H, then for any point $x \in H$ there exists a unique point $y \in K$ that attains the minimum distance between x and a point in K.*

Proof. Like in Theorem 1.7, the uniqueness part of the statement does not require the assumption of completeness on the underlying vector space; it holds for an arbitrary inner product space as follows. Let D be the distance between K and x. In other words,

$$D = \inf_{y\in K} d(x,y).$$

Suppose that y_1 and y_2 both realize this infimum, so $d(y_1,x) = d(y_2,x) = D$. Then by the parallelogram law Proposition 4.5 applied to $(y_1 - x)$ and $(y_2 - x)$,

$$\begin{aligned}||y_1 - y_2||^2 &= 2\,||y_1 - x||^2 + 2\,||y_2 - x||^2 - ||y_1 + y_2 - 2x||^2 \\ &= 4D^2 - 4\left|\left|\frac{y_1+y_2}{2} - x\right|\right|^2 \\ &\le 4D^2 - 4D^2 \\ &= 0\end{aligned}$$

Here we've used the fact that K is convex, so $(y_1 + y_2)/2$ is in K, so by definition the definition of D its distance to x is at least D. Since $||y_1 - y_2||$ is always in $\mathbf{R}_{\geq 0}$ we see that $||y_1 - y_2|| = 0$, so $y_1 = y_2$. This proves uniqueness.

To prove the existence of a distance-minimizer we use a similar trick: let $y_n \in K$ be some point such that $d(y_n, x) \leq D + 1/n$. Since the distance function is continuous on H, if we can show that y_n converges to some $y \in H$, we'll be done: because $y_n \in K$ and K is closed we'll know that $y \in K$. Moreover

$$d(y,x) = \lim_{n\to\infty} d(y_n, x) = D$$

so y will be our desired distance-minimizer. To show that y_n converges to some $y \in H$ we'll use the parallelogram law Proposition 4.5 to argue that the sequence y_n is Cauchy and then conclude by the completeness of H.

Explicitly, applying Proposition 4.5 to $(y_n - x)$ and $(y_m - x)$ we get

$$\begin{aligned} ||y_n - y_m||^2 &= 2\,||y_n - x||^2 + 2\,||y_m - x||^2 - ||y_n + y_m - 2x||^2 \\ &= 2\left(D + \frac{1}{n}\right)^2 + 2\left(D + \frac{1}{m}\right)^2 - 4\left|\left|\frac{y_n + y_m}{2} - x\right|\right|^2 \\ &\leq 2D^2 + \frac{4D}{n} + \frac{2}{n^2} + 2D^2 + \frac{4D}{m} + \frac{2}{m^2} - 4D^2 \\ &= \frac{4D}{n} + \frac{2}{n^2} + \frac{4D}{m} + \frac{2}{m^2}. \end{aligned}$$

Since this goes to zero as n, m go to ∞ we see that the sequence y_n is Cauchy, as desired. □

Definition 4.8. If H is a Hilbert space and K is a closed convex subset of H, the **Projection Operator** $P_K : H \to K$ is the map that sends $x \in H$ to the point $y = P_K(x) \in K$ that minimizes the distance to x.

Remark 4.9. Note that the function P_K is well defined by Proposition 4.7.

The case when K is a closed linear subspace of H is particularly interesting.

Proposition 4.10. *If K is a closed linear subspace of H and x is some vector in H, the following are equivalent for all $y \in K$*

1. $y = P_K(x)$.
2. $\langle y - x, u\rangle = 0$ *for all* $u \in K$.

Proof. First we show that 1 implies 2, so suppose that $y = P_K(x)$. Given any $u \in K$, consider the function $f : \mathbf{R} \to \mathbf{R}_{\geq 0}$ such that

$$\begin{aligned} f(\lambda) &= ||y + \lambda u - x||^2 \\ &= ||y - x||^2 + 2\lambda\,\mathrm{Re}(\langle y - x, u\rangle) + \lambda^2\,||u||^2\,. \end{aligned}$$

Evidently, f is a quadratic function in λ. Furthermore, since $u \in K$ and since K is a linear subspace of H, we know that $y + \lambda u \in K$ for all λ. By our choice of y, this implies that $f(\lambda) \geq f(0)$ for all λ–so f attains its minimum at $\lambda = 0$. A quadratic function $f(\lambda) = a\lambda^2 + b\lambda + c$ attains its minimum at $\lambda = 0$ exactly when $b = 0$, so we conclude that $\mathrm{Re}(\langle y - x\rangle, u) = 0$. If $F = \mathbf{R}$, we're done. If $F = \mathbf{C}$ note that

$$\begin{aligned} \langle y - x, u\rangle &= \mathrm{Re}(\langle y - x, u\rangle) + i\,\mathrm{Im}(\langle y - x, u\rangle) \\ &= \mathrm{Re}(\langle y - x, u\rangle) + i\,\mathrm{Re}(\langle y - x, -iu\rangle) \\ &= 0 \end{aligned}$$

where the last equality follows because both u and iu are in K.

To show that 2 implies 1, suppose that $y \in K$ is such that $\langle y - x, u \rangle = 0$ for all $u \in K$. Then for any other $y' \in K$ we have

$$\begin{aligned} \left|\left|y' - x\right|\right|^2 &= \left\langle y' - x, y' - x \right\rangle \\ &= \left\langle (y' - y) + (y - x), (y' - y) + (y - x) \right\rangle \\ &= \left|\left|y' - y\right|\right|^2 + \left\langle y' - y, y - x \right\rangle + \left\langle y - x, y' - y \right\rangle + \left|\left|y - x\right|\right|^2 \\ &= \left|\left|y' - y\right|\right|^2 + \left|\left|y - x\right|\right|^2 \\ &\geq \left|\left|y - x\right|\right|^2. \end{aligned}$$

Here we've used the fact that $y' - y \in K$, so $\langle y - x, y' - y \rangle = \langle y' - y, y - x \rangle = 0$. This shows that y is the closest to x among all of the points of K, so $y = P_K(x)$ by definition. □

Corollary 4.11. *If K is a closed linear subspace of H then P_K is a continuous linear operator.*

Proof. To see that P_K is linear, note that by the implication 1 implies 2 of Proposition 4.10 we know that for any $u \in K$

$$\begin{aligned} \langle P_K(x) + \lambda P_K(y) - x - \lambda y, u \rangle &= \langle P_K(x) - x, u \rangle + \overline{\lambda} \langle P_K(y) - y, u \rangle \\ &= 0 \end{aligned}$$

so by the implication 2 implies 1 of Proposition 4.10 we know that

$$P_K(x + \lambda y) = P_K(x) + \lambda P_K(y).$$

This proves that P_K is linear. To show that P_K is continuous, note that

$$\begin{aligned} \left|\left|x\right|\right|^2 &= \langle x, x \rangle \\ &= \langle P_K(x) + (x - P_K(x)), P_K(x) + (x - P_K(x)) \rangle \\ &= \left|\left|P_K(x)\right|\right|^2 + 2\,\mathrm{Re}(\langle P_K(x), x - P_K(x) \rangle)\, \left|\left|P_K(x) - x\right|\right|^2 \\ &= \left|\left|P_K(x)\right|\right|^2 + \left|\left|P_K(x) - x\right|\right|^2 \end{aligned}$$

where we have used Proposition 4.10 to say that $\langle P_K(x), x - P_K(x) \rangle = 0$. This implies that $||P_K(x)|| \leq ||x||$, so P_K is continuous. □

todo: I have not talked about continuous linear operators yet.
todo: Riesz representation theorem.

5 Sobolev Spaces

todo: example of screening equation

todo: definition of weak derivative

todo: weak formulation of PDE

todo: Lax-Milgram theorem?