1. Basic Definitions

A real or complex\[^1\] vectorspace $V$ with a real-valued function, the norm,

$$|| : V \rightarrow \mathbb{R}$$

with properties

- $|x + y| \leq |x| + |y|$ (triangle inequality)
- $|\alpha x| = |\alpha| \cdot |x|$ ($\alpha$ complex, $x \in V$)
- $|x| = 0 \Rightarrow x = 0$ (positivity)

is a normed complex vectorspace, or simply normed space. Because of the triangle inequality, the function

$$d(x, y) = |x - y|$$

is a metric. The symmetry comes from

$$d(y, x) = |y - x| = |(-1) \cdot (x - y)| = |-1| \cdot |x - y| = |x - y| = d(x, y)$$

[\[^1\] In fact, for many purposes, the scalars need not be $\mathbb{R}$ or $\mathbb{C}$, need not be locally compact, and need not even be commutative. The basic results hold for Banach spaces over non-discrete, complete, normed division rings. This allows scalars like the $p$-adic field $\mathbb{Q}_p$, or Hamiltonian quaternions $\mathbb{H}$, and so on.]
When $V$ is complete with respect to this metric, $V$ is a Banach space.

Hilbert spaces are Banach spaces, but many natural Banach spaces are not Hilbert spaces, and may fail to enjoy useful properties of Hilbert spaces. Riesz’ lemma below is sometimes a sufficient substitute.

Most norms on Banach spaces do not arise from inner products. Norms arising from inner products recover the inner product via the polarization identities

$$4\langle x, y \rangle = |x + y|^2 - |x - y|^2 \quad \text{(real vector space)}$$

$$4\langle x, y \rangle = |x + y|^2 - |x - y|^2 + i|x + iy|^2 - i|x - iy|^2 \quad \text{(complex vector space)}$$

Given a norm on a vector space, if the polarization expression gives an inner product, then the norm is produced by that inner product. However, checking whether the polarization expression is bilinear or hermitian, may be awkward or non-intuitive.

2. Riesz’ Lemma

The following essentially elementary inequality is sometimes an adequate substitute for corollaries of the Hilbert-space minimum principle and its corollaries. Once one sees the proof, it is not surprising, but,

**[2.1] Lemma:** (Riesz) For a non-dense subspace $X$ of a Banach space $Y$, given $r < 1$, there is $y \in Y$ with $|y| = 1$ and $\inf_{x \in X} |x - y| \geq r$.

**Proof:** Take $y_1$ not in the closure of $X$, and put $R = \inf_{x \in X} |x - y_1|$. Thus, $R > 0$. For $\varepsilon > 0$, let $x_1 \in X$ be such that $|x_1 - y_1| < R + \varepsilon$. Put $y = (y_1 - x_1)/|x_1 - y_1|$, so $|y| = 1$. And

$$\inf_{x \in X} |x - y| = \inf_{x \in X} |x + \frac{x_1}{|x_1 - y_1|} - \frac{y_1}{|x_1 - y_1|}| = \inf_{x \in X} |x - y_1| = \frac{R}{R + \varepsilon}$$

By choosing $\varepsilon > 0$ small, $R/(R + \varepsilon)$ can be made arbitrarily close to 1. ///

3. Counter-examples to unique norm-minimizing element

The (true) minimum principle for Hilbert spaces is that a closed, convex subset has a unique element of minimum norm. This has many important elementary corollaries special to Hilbert spaces, such as existence of orthogonal complements to subspaces, and often fails for Banach spaces.

An important historical example of failure of functionals to attain their infs on closed, convex subsets of Banach spaces is the falsity of the Dirichlet principle as originally naively proposed. [2]

The (true) minimizing principle in a Hilbert space $V$ is that, in a closed, convex, non-empty subset $E \subset V$, there is a unique element of least norm. As an example corollary, for non-dense subspace $W$ of a Hilbert

[2] The Dirichlet principle, invoked by Riemann but observed by Weierstraß to be false as stated, would assert that a solution of $\Delta u = f$ on an open set $\Omega \subset \mathbb{R}^n$, with boundary condition $u|_{\partial \Omega} = g$ on $\partial \Omega$, is a minimizer of the energy integral

$$E(u) = \int_\Omega \frac{1}{2} |\nabla u|^2 + |u|^2$$

on the Banach subspace of $C^2(\Omega)$ functions $u$ satisfying $u|_{\partial \Omega} = g$. However, the infimum need not be attained in that Banach space. Hilbert justified Dirichlet’s principle in certain circumstances. Beppo Levi (1906) observed that using energy integrals to form the norm (squared) of a pre-Hilbert space in $C^2(\Omega)$, and completing to a Hilbert space, does guarantee existence of a solution in that Hilbert space.
space $V$, there is $v \in V$ with $|v| = 1$ and $\inf_{w \in W} |v - w| = 1$, by taking $v$ to be a unit-length vector in the orthogonal complement to $W$. This minimization property typically fails in Banach spaces, as follows.

[3.1] Example: Many minimizing elements can exist: in the Banach space $L^1[a,b]$, in the closed, convex subset $E = \{ f : \int_a^b f = 1 \}$, there are infinitely-many norm-minimizing elements.

[3.2] Example: In the Banach space $Y = C^o[0,2]$, with closed convex subset

$$E = \{ f \in C^o[0,2] : \int_0^1 f(x) \, dx - \int_1^2 f(x) \, dx = 1 \}$$

there is no norm-minimizing element. To this end, let

$$s(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ -1 & \text{for } 1 \leq x \leq 2 \end{cases}$$

and

$$\lambda(f) = \int_0^2 f(x) \cdot s(x) \, dx$$

Certainly $C^o[0,2] \subset L^2[0,2]$, so by Cauchy-Schwarz-Bunyakowsky,

$$|\lambda(f)| = |\langle f, s \rangle| \leq |f|_{L^2[0,2]} \cdot |s|_{L^2[0,2]} = |f|_{L^2[0,2]} \cdot \sqrt{2}$$

with equality only for $f$ a scalar multiple of $s$. Also, certainly

$$|f|_{L^2[0,2]} \leq \left( \int_0^2 |f|_Y \right)^{\frac{1}{2}} = |f|_Y \cdot \sqrt{2} \quad \text{(for } f \in C^o[0,2])$$

Since $s$ is not continuous, non-zero $f \in Y$ is never a constant multiple of $s$, so Cauchy-Schwarz-Bunyakovsky gives a strict inequality

$$|\lambda(f)| < |f|_{L^2[0,2]} \cdot \sqrt{2} \leq |f|_Y \cdot 2 \quad \text{(for all } 0 \neq f \in Y)$$

Thus,

$$\frac{1}{2} < |f|_Y \quad \text{(for } f \in E)$$

Yet it is easy to arrange continuous functions $f$ with $\lambda(f) = 1$ and sup-norm $|f|_Y$ approaching $1/2$ from above, by approximating $\frac{1}{2}s(x)$ by continuous functions. For example, form a continuous, piecewise-linear function

$$g(x) = \begin{cases} \frac{1}{2} & \text{for } 0 \leq x \leq 1 - \varepsilon \\ \frac{1}{2} - \frac{x - (1 - \varepsilon)}{2\varepsilon} & \text{for } 1 - \varepsilon \leq x \leq 1 + \varepsilon \\ -\frac{1}{2} & \text{for } 1 + \varepsilon \leq x \leq 2 \end{cases}$$

The sup-norm of $g$ is obviously $\frac{1}{2}$, and $\lambda(g) = 1 - \frac{1}{2}\varepsilon$. Thus, functions $f = g/(1 - \frac{1}{2}\varepsilon)$ have $\lambda(f) = 1$ and sup norms approaching $\frac{1}{2}$ from above. This proves the claimed failure. ///
4. Normed spaces of linear maps

There is a natural norm on the set of continuous linear maps $T : X \to Y$ from one normed space $X$ to another normed space $Y$. Even when $X, Y$ are Hilbert spaces, the set of continuous linear maps $X \to Y$ is generally only a Banach space.

Let $\text{Hom}^o(X, Y)$ denote the collection of continuous linear maps from the normed vectorspace $X$ to the normed vectorspace $Y$. Use the same notation $|| \cdot ||$ for the norms on both $X$ and $Y$, since context will make clear which is meant.

A linear (not necessarily continuous) map $T : X \to Y$ from one normed space to another has uniform operator norm

$$|T| = |T|_{\text{uniform}} = \sup_{|x| \leq 1} |Tx|$$

where we allow the value $+\infty$. Such $T$ is called bounded if $|T| < +\infty$. There are several obvious variants of the expression for the uniform norm:

$$|T| = \sup_{|x| \leq 1} |Tx| = \sup_{|x| < 1} |Tx| = \sup_{|x| \neq 0} \frac{|Tx|}{|x|}$$

[4.1] Proposition: For a linear map $T : X \to Y$ from one normed space to another, the following conditions are equivalent:

- $T$ is continuous.
- $T$ is continuous at 0.
- $T$ is bounded.

Proof: First, show that continuity at a point $x_o$ implies continuity everywhere. For another point $x_1$, given $\varepsilon > 0$, take $\delta > 0$ so that $|x - x_o| < \delta$ implies $|Tx - Tx_o| < \varepsilon$. Then for $|x' - x_1| < \delta$

$$|(x' + x_o - x_1) - x_o| < \delta$$

By linearity of $T$,

$$|Tx' - Tx_1| = |T(x' + x_o - x_1) - Tx_o| < \varepsilon$$

which is the desired continuity at $x_1$.

Now suppose that $T$ is continuous at 0. For $\varepsilon > 0$ there is $\delta > 0$ so that $|x| < \delta$ implies $|Tx| < \varepsilon$. For $x \neq 0$,

$$\left| \frac{\delta}{2|x|} x \right| < \delta$$

so

$$\left| T \frac{\delta}{2|x|} \cdot x \right| < \varepsilon$$

Multiplying out and using the linearity, boundedness is obtained:

$$|Tx| < \frac{2\varepsilon}{\delta} \cdot |x|$$

[3] Another traditional notation for the collection of continuous linear maps from $X$ to $Y$ is $B(X, Y)$, where $B$ stands for bounded.
Finally, prove that boundedness implies continuity at 0. Suppose there is $C$ such that $|Tx| < C|x|$ for all $x$. Then, given $\varepsilon > 0$, for $|x| < \varepsilon/C$

$$|Tx| < C|x| < C \cdot \frac{\varepsilon}{C} = \varepsilon$$

which is continuity at 0. ///

The space $\operatorname{Hom}^o(X, Y)$ of continuous linear maps from a normed space $X$ to another normed space $Y$ has a natural structure of vectorspace by

$$(\alpha T)(x) = \alpha \cdot (Tx) \quad \text{and} \quad (S + T)x = Sx + Tx$$

for $\alpha \in \mathbb{C}$, $S, T \in \operatorname{Hom}^o(X, Y)$, and $x \in X$.

**[4.2] Proposition:** With the uniform operator norm, the space $\operatorname{Hom}^o(X, Y)$ of continuous linear operators from a normed space $X$ to a Banach space $Y$ is complete, whether or not $X$ itself is complete. Thus, $\operatorname{Hom}^o(X, Y)$ is a Banach space.

**Proof:** Let $\{T_i\}$ be a Cauchy sequence of continuous linear maps $T : X \to Y$. Try defining the limit operator $T$ in the natural fashion, by

$$Tx = \lim_i T_{x_i}$$

First, check that this limit exists. Given $\varepsilon > 0$, take $i_o$ large enough so that $|T_i - T_j| < \varepsilon$ for $i, j > i_o$. By the definition of the uniform operator norm,

$$|T_i x - T_j x| < |x|\varepsilon$$

Thus, the sequence of values $T_{x_i}$ is Cauchy in $Y$, so has a limit in $Y$. Call the limit $Tx$.

We need to prove that the map $x \to Tx$ is continuous and linear. The arguments are inevitable. Given $c \in \mathbb{C}$ and $x \in X$, for given $\varepsilon > 0$ choose index $i$ so that for $j > i$ both $|Tx - T_j x| < \varepsilon$ and $|Tcx - T_j cx| < \varepsilon$. Then

$$|Tcx - cTx| \leq |Tcx - T_j cx| + |cT_j x - cTx| = |Tcx - T_j cx| + |c| \cdot |T_j x - Tx| < (1 + |c|)\varepsilon$$

This is true for every $\varepsilon$, so $Tcx = cTx$. Similarly, given $x, x' \in X$, for $\varepsilon > 0$ choose an index $i$ so that for $j > i$ $|Tx - T_j x| < \varepsilon$ and $|Ty - T_j y| < \varepsilon$ and $|T(x + y) - T_j(x + y)| < \varepsilon$. Then

$$|T(x + y) - Tx - Ty| \leq |T(x + y) - T_j(x + y)| + |T_j x - Tx| + |T_j y - Ty| < 3\varepsilon$$

This holds for every $\varepsilon$, so $T(x + y) = Tx + Ty$.

For continuity, show that $T$ is bounded. Choose an index $i_o$ so that for $i, j \geq i_o$

$$|T_i - T_j| \leq 1$$

This is possible since the sequence of operators is Cauchy. For such $i, j$

$$|T_i x - T_j x| \leq |x|$$

for all $x$. Thus, for $i \geq i_o$

$$|T_i x| \leq |(T_i - T_{i_o}) x| + |T_{i_o} x| \leq |x|(1 + |T_{i_o}|)$$

Taking a limsup,

$$\limsup_i |T_i x| \leq |x|(1 + |T_{i_o}|)$$

This implies that $T$ is bounded, and so is continuous.
Finally, we should see that $T x = \lim_i T_i x$ is the operator-norm limit of the $T_i$. Given $\varepsilon > 0$, let $i_o$ be sufficiently large so that $|T_i x - T_j x| < \varepsilon$ for all $i, j \geq i_o$ and for all $|x| \leq 1$. Then $|T x - T x_i| \leq \varepsilon$ and

$$\sup_{|x| \leq 1} |T x - T_i x| \leq \sup_{|x| \leq 1} \varepsilon = \varepsilon$$

giving the desired outcome.

---

5. Dual spaces of normed spaces

This section considers an important special case of continuous linear maps between normed spaces, namely continuous linear maps from Banach spaces to scalars. All assertions are special cases of those for continuous linear maps to general Banach spaces, but deserve special attention.

For $X$ a normed vectorspace with norm $||$, a continuous linear map $\lambda : X \to \mathbb{C}$ is a (continuous linear) functional on $X$. Let

$$X^* = \text{Hom}^o(X, \mathbb{C})$$

denote the collection of all such (continuous) functionals.

As more generally, for any linear map $\lambda : X \to \mathbb{C}$ of a normed vectorspace to $\mathbb{C}$, the norm $|\lambda|$ is

$$|\lambda| = \sup_{|x| \leq 1} |\lambda x|$$

where $|\lambda x|$ is the absolute value of the value $\lambda x \in \mathbb{C}$. We allow the value $+\infty$. Such a linear map $\lambda$ is bounded if $|\lambda| < +\infty$.

As a special case of the corresponding general result:

**[5.1] Corollary:** For a $k$-linear map $\lambda : X \to k$ from a normed space $X$ to $k$, the following conditions are equivalent:

- The map $\lambda$ is continuous.
- The map $\lambda$ is continuous at one point.
- The map $\lambda$ is bounded.

**Proof:** These are special cases of the earlier proposition where the target was a general Banach space.

---

The dual space

$$X^* = \text{Hom}^o(X, \mathbb{C})$$

of $X$ is the collection of continuous linear functionals on $X$. This dual space has a natural structure of vectorspace by

$$(\alpha \lambda)(x) = \alpha \cdot (\lambda x) \quad \text{and} \quad (\lambda + \mu)x = \lambda x + \mu x$$

for $\alpha \in \mathbb{C}$, $\lambda, \mu \in X^*$, and $x \in X$. It is easy to check that the norm

$$|\lambda| = \sup_{|x| \leq 1} |\lambda x|$$

really is a norm on $X^*$, in that it meets the conditions

- **Positivity:** $|\lambda| \geq 0$ with equality only if $\lambda = 0$.
- **Homogeneity:** $|\alpha \lambda| = |\alpha| \cdot |\lambda|$ for $\alpha \in k$ and $\lambda \in X^*$. As a special case of the discussion of the uniform norm on linear maps, we have
[5.2] **Corollary:** The dual space $X^*$ of a normed space $X$, with the natural norm, is a Banach space. That is, with respect to the natural norm on continuous functionals, it is complete.

---

### 6. Banach-Steinhaus/uniform-boundedness theorem

This result is non-trivial in the sense that it uses the Baire category theorem.

#### [6.1] Theorem: (Banach-Steinhaus/uniform boundedness)

For a family of continuous linear maps $T_\alpha : X \to Y$ from a Banach space $X$ to a normed space $Y$, either there is a uniform bound $M < \infty$ so that $|T_\alpha| \leq M$ for all $\alpha$, or there is $x \in X$ such that

$$
\sup_\alpha \frac{|T_\alpha x|}{|x|} = +\infty
$$

In the latter case, in fact, there is a dense $G_\delta$ of such $x$.

**Proof:** Let $p(x) = \sup_\alpha |T_\alpha x|$. We allow the possibility that $p(x) = +\infty$. Being the sup of continuous functions, $p$ is lower semi-continuous: for each integer $n$, the set $U_n = \{x : p(x) > n\}$ is open.

On one hand, if every $U_n$ is dense in $X$, by Baire category the intersection is dense, so is non-empty. By definition, it is a dense $G_\delta$. On that set $p$ is $+\infty$.

On the other hand, if one of the $U_n$ is not dense, then there is a ball $B$ of radius $r > 0$ about a point $x_o$ which does not meet $U_n$. For $|x - x_o| < r$ and for all $\alpha$

$$
|T_\alpha (x - x_o)| \leq |T_\alpha x| + |T_\alpha x_o| \leq 2n
$$

As $x - x_o$ varies over the open ball of radius $r$ the vector $x' = (x - x_o)/r$ varies over the open ball of radius 1, and

$$
|T_\alpha x'| = \left| T_\alpha \left( \frac{x - x_o}{r} \right) \right| \leq 2n/r
$$

Thus, $|T_\alpha| \leq 2n/r$, which is the uniform boundedness.

---

### 7. Open mapping theorem

The open mapping theorem is non-trivial, since it invokes the Baire category theorem.

#### [7.1] Theorem: (open mapping)

For a continuous linear surjection $T : X \to Y$ of Banach spaces, there is $\delta > 0$ such that for all $y \in Y$ with $|y| < \delta$ there is $x \in X$ with $|x| \leq 1$ such that $Tx = y$. In particular, $T$ is an open map.

#### [7.2] Corollary: A bijective continuous linear map of Banach spaces is an isomorphism.

**Proof:** In the corollary the non-trivial point is that $T$ is open, which is the point of the theorem. The linearity of the inverse is easy.

For every $y \in Y$ there is $x \in X$ so that $Tx = y$. For some integer $n$ we have $n > |x|$, so $Y$ is the union of the sets $TB(n)$, with usual open balls

$$
B(n) = \{x \in X : |x| < n\}
$$

By Baire category, the closure of some one of the sets $TB(n)$ contains a non-empty open ball

$$
V = \{y \in Y : |y - y_o| < r\}
$$
for some $r > 0$ and $y_0 \in Y$. Since we are in a metric space, the conclusion is that every point of $V$ occurs as the limit of a Cauchy sequence consisting of elements from $TB(n)$.

Certainly

$$\{ y \in Y : |y| < r \} \subset \{ y_1 - y_2 : y_1, y_2 \in V \}$$

Thus, every point in the ball $B'_r$ of radius $r$ centered at 0 in $Y$ is the sum of two limits of Cauchy sequences from $TB(n)$. Thus, surely every point in $B'_r$ is the limit of a single Cauchy sequence from the image $TB(2n)$ of the open ball $B(2n)$ of twice the radius. That is, the closure of $TB(2n)$ contains the ball $B'(r)$.

Using the linearity of $T$, the closure of $TB(\rho)$ contains the ball $B'(\rho/2n)$ in $Y$.

Given $|y| < 1$, choose $x_1 \in B(2n/r)$ so that $|y - Tx_1| < \varepsilon$. Choose $x_2 \in B(\varepsilon \cdot 2n/\varepsilon)$ so that

$$|(y - Tx_1) - Tx_2| < \varepsilon/2$$

Choose $x_3 \in B(\varepsilon \cdot 2n/\varepsilon)$ so that

$$|(y - Tx_1 - Tx_2) - Tx_3| < \varepsilon/2^2$$

Choose $x_4 \in B(\varepsilon \cdot 2n/\varepsilon)$ so that

$$|(y - Tx_1 - Tx_2 - Tx_3) - Tx_4| < \varepsilon/2^3$$

and so on. The sequence

$$x_1, x_1 + x_2, x_1 + x_2 + x_3, \ldots$$

is Cauchy in $X$. Since $X$ is complete, the limit $x$ of this sequence exists in $X$, and $Tx = y$. We find that

$$x \in B\left(\frac{2n}{r}\right) + B\left(\frac{\varepsilon}{2} \cdot \frac{2n}{\varepsilon}\right) + B\left(\frac{\varepsilon}{2^2} \cdot \frac{2n}{\varepsilon}\right) + \ldots \subset B\left(\frac{(1 + 2\varepsilon)2n}{r}\right)$$

Thus,

$$TB((1 + \varepsilon)\frac{2n}{r}) \supset \{ y \in Y : |y| < 1 \}$$

This proves open-ness at 0. ///

8. Closed graph theorem

The closed graph theorem uses the open mapping theorem, so invokes Baire category, so is non-trivial.

It is straightforward to show[4] that a continuous map $f : X \to Y$ of Hausdorff topological spaces has closed graph

$$\Gamma_f = \{ (x, y) : f(x) = y \} \subset X \times Y$$

Similarly, a topological space $X$ is Hausdorff if and only if the diagonal $X^\Delta = \{ (x, x) : x \in X \}$ is closed in $X \times X$. [5]

[4] To show that a continuous map $f : X \to Y$ of topological spaces with $Y$ Hausdorff has closed graph $\Gamma_f$, show the complement is open. Take $(x, y) \notin \Gamma_f$. Let $V_1$ be a neighborhood of $f(x)$ and $V_2$ a neighborhood of $y$ such that $V_1 \cap V_2 = \emptyset$, using Hausdorff-ness. By continuity of $f$, for $x'$ in a suitable neighborhood $U$ of $x$, the image $f(x')$ is inside $V_1$. Thus, the neighborhood $U \times V_2$ of $(x, y)$ does not meet $\Gamma_f$.

[5] To show that closed-ness of the diagonal $X^\Delta$ in $X \times X$ implies $X$ is Hausdorff, let $x_1 \neq x_2$ be points in $X$. Then there is a neighborhood $U_1 \times U_2$ of $(x_1, x_2)$, with $U_i$ a neighborhood of $x_i$, not meeting the diagonal. That is, $(x, x') \in U_1 \times U_2$ implies $x \neq x'$. That is, $U_1 \cap U_2 = \emptyset$. 8
[8.1] **Theorem:** A linear map $T : V \to W$ of Banach spaces is continuous if it has closed graph

$$\Gamma = \Gamma_T = \{(v, w) : Tv = w\}$$

**Proof:** It is routine to check that $V \times W$ with norm $|v \times w| = |v| \cdot |w|$ is a Banach space. Since $\Gamma$ is a closed subspace of $V \times W$, it is a Banach space itself with the restriction of this norm.

The projection $\pi_V : V \times W \to V$ is a continuous linear map. The restriction $\pi_V|_{\Gamma}$ of $\pi_V$ to $\Gamma$ is still surjective, because it $T$ is an everywhere-defined function on $V$. By the open mapping theorem, $\pi_V|_{\Gamma}$ is open. Thus, the bijection $\pi_V|_{\Gamma}$ is a homeomorphism. Letting $\pi_W : V \times W \to W$ be the projection to $W$,

$$T = \pi_W \circ (\pi_V|_{\Gamma})^{-1} : V \to W$$

expresses $T$ as a composite of continuous functions. ///

[8.2] **Remark:** The proof introduced two readily verifiable, useful ideas: a product of Banach spaces is a Banach space, and a closed vector subspace of a Banach space is a Banach space.

---

### 9. Hahn-Banach Theorem

Hahn-Banach does not use completeness, much less Baire category. The salient feature is convexity, and the scalars must be $\mathbb{R}$ or $\mathbb{C}$. Indeed, the Hahn-Banach theorem seems to be a result about real vectorspaces. Note that a $\mathbb{C}$-vectorspace may immediately be considered as a $\mathbb{R}$-vectorspace simply by forgetting some of the structure.

For $Y$ a vector subspace of $X$, and for $S : Y \to Z$ a linear map to another vectorspace $Z$, a linear map $T : X \to Z$ is an extension of $S$ to $X$ when the restriction $T|_Y$ of $T$ to $Y$ is $S$.

[9.1] **Theorem:** (Hahn-Banach) Let $X$ be a normed vectorspace with scalars $\mathbb{R}$ or $\mathbb{C}$, $Y$ be a subspace, and $\lambda$ be a continuous linear functional on $Y$. Then there is an extension $\Lambda$ of $\lambda$ to $X$ such that

$$|\Lambda| = |\lambda|$$

[9.2] **Corollary:** Given $x \neq y$ in a normed space $X$, neither a scalar multiple of the other, there is a continuous linear functional $\lambda$ on $X$ so that $\lambda x = 1$ while $\lambda y = 0$. ///

[9.3] **Corollary:** Let $Y$ be a closed subspace of a normed space $X$, and $x_o \not\in Y$. Then there is a continuous linear functional $\lambda$ on $X$ which is 0 on $Y$, has $|\lambda| = 1$, and $\lambda(x_o) = |x_o|$. ///

**Proof:** We treat the case that the scalars are $\mathbb{R}$, and reduce the complex case to this.

The critical part is to extend a linear functional by just one dimension. That is, for given $x_o \not\in Y$ make an extension $\lambda'$ of $\lambda$ to $Y' = Y + \mathbb{R}x_o$. Every vector in $Y'$ has a unique expression as $y + cx_o$ with $c \in \mathbb{R}$, so define functionals by

$$\mu(y + cx_o) = \lambda y + c\ell \quad \text{(for arbitrary } \ell \in \mathbb{R})$$

The issue is to choose $\ell$ so that $|\mu| = |\lambda|$. Certainly $\lambda = 0$ is extendable by $\Lambda = 0$, so we consider the case that $|\lambda| \neq 0$. We can divide by $|\lambda|$ to suppose that $|\lambda| = 1$.

The condition $|\mu| = |\lambda|$ is a condition on $\ell$:

$$|\lambda y + c\ell| \leq |y + cx_o| \quad \text{(for every } y \in Y)$$
We have simplified to the situation that we know this does hold for \( c = 0 \). So for \( c \neq 0 \), divide through by \( |c| \) and replace \( y \in Y \) by \( cy \), so that the condition becomes

\[
|\lambda y + \ell| \leq |y + x_o| \quad \text{ (for every } y \in Y) \]

Replacing \( y \) by \( -y \), the condition on \( \ell \) is that

\[
|\ell - \lambda y| \leq |y - x_o| \quad \text{ (for every } y \in Y) \]

For a single \( y \in Y \), the condition on \( \ell \) is that

\[
\lambda y - |y - x_o| \leq \ell \leq \lambda y + |y - x_o| \]

To have a common solution \( \ell \), it is exactly necessary that every lower bound be less than every upper bound. To see that this is so, start from

\[
\lambda y_1 - \lambda y_2 = \lambda (y_1 - y_2) \leq |\lambda (y_1 - y_2)| \leq |y_1 - y_2| \leq |y_1 - x_o| + |y_2 - x_o| \]

by the triangle inequality. Subtracting \( |y_1 - x_o| \) from both sides and adding \( \lambda y_2 \) to both sides,

\[
\lambda y_1 - |y_1 - x_o| \leq \lambda y_2 + |y_2 - x_o| \]

as desired. That is, we have proven the existence of at least one extension from \( Y \) to \( Y' = Y + \mathbb{R} x_o \) with the same norm.

An equivalent of the Axiom of Choice will extend to the whole space while preserving the norm, as follows. Consider the set of pairs \((Z, \zeta)\) where \( Z \) is a subspace containing \( Y \) and \( \zeta \) is a continuous linear functional on \( Z \) extending \( \lambda \) and with \( |\zeta| \leq 1 \). Order these by

\[
(Z, \zeta) \leq (Z', \zeta') \quad \text{ when } Z \subset Z' \text{ and } \zeta' \text{ extends } \zeta. \]

For a totally ordered collection \((Z_\alpha, \zeta_\alpha)\) of such,

\[
Z' = \bigcup_\alpha Z_\alpha
\]

is a subspace of \( X \). In general, of course, the union of a family of subspaces would not be a subspace, but these are nested.

We obtain a continuous linear functional \( \zeta' \) on this union \( Z' \), extending \( \lambda \) and with \( |\zeta'| \leq 1 \), as follows. Any finite batch of elements already occur inside some \( Z_\alpha \). Given \( z \in Z' \), let \( \alpha \) be any index large enough so that \( z \in Z_\alpha \), and put

\[
\zeta'(z) = \zeta_\alpha(z)
\]

The family is totally ordered, so the choice of \( \alpha \) does not matter so long as it is sufficiently large. Certainly for \( c \in \mathbb{R} \)

\[
\zeta'(cz) = \zeta_\alpha(cz) = c \zeta_\alpha(z) = c \zeta'(z)
\]

For \( z_1 \) and \( z_2 \) and \( \alpha \) large enough so that both \( z_1 \) and \( z_2 \) are in \( Z_\alpha \),

\[
\zeta'(z_1 + z_2) = \zeta_\alpha(z_1 + z_2) = \zeta_\alpha(z_1) + \zeta_\alpha(z_2) = \zeta'(z_1) + \zeta'(z_2)
\]

proving linearity. Thus, there is a maximal pair \((Z', \zeta')\). The earlier argument shows that \( Z' \) must be all of \( X \), since otherwise we could construct a further extension, contradicting the maximality. This completes the proof for the case that the scalars are the real numbers.
To reduce the complex case to the real case, the main trick is that, for $\lambda_o$ a real-linear real-valued functional, the functional
\[ \lambda x = \lambda_o(x) - i\lambda(ix) \]
is complex-linear, and has the same norm as $\lambda_o$. In particular, when
\[ \lambda_o(x) = \text{Re}\lambda(x) = \frac{\lambda x + \overline{\lambda x}}{2} \]
is the real part of $\lambda$ we recover $\lambda$ itself by this formula.

Granting this, given $\lambda$ on a complex subspace, take its real part $\lambda_o$, a real-linear functional, and extend $\lambda_o$ to a real-linear functional $\Lambda_o$ with the same norm. Then the desired extension of $\lambda$ is
\[ \Lambda x = \Lambda_o(x) - i\Lambda(ix) \]
proving the theorem in the complex case.

Consider the construction
\[ \lambda x = \lambda_o(x) - i\lambda(ix) \]
Since $\lambda_o(x + y) = \lambda_o x + \lambda_o y$ it follows that $\lambda$ also has this additivity property. For $a, b$ real,
\[
\lambda((a + bi)x) = \lambda_o((a + bi)x) - i\lambda_o(i(a + bi)x) = \lambda_o(ax) + \lambda_o(ix) - i\lambda_o(iax) - i\lambda_o(-bx) \\
= a\lambda_o x + b\lambda_o(ix) - i\lambda_o(iy) + ib\lambda_o x = (a + bi)\lambda_o x - i(a + bi)\lambda_o(ix) = (a + bi)\lambda(x)
\]
This gives the linearity.

Regarding the norm: since $\lambda_o$ is real-valued, always
\[ |\lambda_o(x)| \leq \sqrt{\lambda_o(x)^2 + \lambda_o(ix)^2} = |\lambda x| \]
On the other hand, given $x$ there is a complex number $\mu$ of absolute value 1 so that $\mu\lambda(x) = |\lambda x|$. And
\[ \lambda_o(x) = \lambda(x) + \overline{\lambda(x)} \]
Then
\[ |\lambda(x)| = |\mu\lambda(x)| = |\lambda\mu x| = \lambda_o(\mu x) - i\lambda_o(i\mu x) \]
Since the left-hand side is real, and since $\lambda_o$ is real-valued, $\lambda_o(\mu x) = 0$. Thus,
\[ |\lambda(x)| = \lambda_o(\mu x) \]
Since $|\mu x| = |x|$, we have equality of norms of the functionals $\lambda_o$ and $\lambda$. This completes the justification of the reduction of the complex case to the real case.