Chapter 3. Banach Spaces

Finite Dimensional Normed Linear Spaces

3.1 Example: Recall, from linear algebra, that when U and V are non-trivial finite dimensional inner product spaces over $\mathbb{R}$ and $F : U \to V$ is a linear map, the closed unit ball in U is compact (so that $\|Fx\|$ attains its maximum on the closed unit ball) and we have

$$\|F\| = \max\left\{\|Fx\| \,\middle|\, x \in U, \|x\| = 1\right\} = \|Fu\| = \sqrt{\lambda}$$

where λ is the largest eigenvalue of $F^*F : U \to U$ and u is a unit eigenvector for λ.

3.2 Theorem: *Let U be an n-dimensional normed linear space over $\mathbb{R}$. Let $\{u_1, \cdots, u_n\}$ be any basis for U and let $F : \mathbb{R}^n \to U$ be the associated vector space isomorphism given by $F(t) = \sum_{k=1}^{n} t_k u_k$. Then both F and F^{-1} are Lipschitz continuous.*

Proof: Let $M = \left(\sum_{k=1}^{n} \|u_k\|^2\right)^{1/2}$. For $t \in \mathbb{R}^n$ we have

$$\begin{aligned} \big\|F(t)\big\| &= \Big\|\sum_{k=1}^{n} t_k u_k\Big\| \leq \sum_{k=1}^{n} |t_k|\,\|u_k\| \quad \text{, by the Triangle Inequality,} \\ &\leq \Big(\sum_{k=1}^{n} {t_k}^2\Big)^{1/2}\Big(\sum_{k=1}^{n} \|u_k\|^2\Big)^{1/2} \quad \text{, by the Cauchy-Schwarz Inequality,} \\ &= M\|t\|. \end{aligned}$$

For all $s, t \in \mathbb{R}^n$, $\big\|F(s) - F(t)\big\| = \big\|F(s-t)\big\| \leq M\,\|s-t\|$, so F is Lipschitz continuous.

Note that the map $N : U \to \mathbb{R}$ given by $N(x) = \|x\|$ is (uniformly) continuous, indeed we can take $\delta = \epsilon$ in the definition of continuity. Since F and N are both continuous, so is the composite $G = N \circ F : \mathbb{R}^n \to \mathbb{R}$, which given by $G(t) = \big\|F(t)\big\|$. By the Extreme Value Theorem, the map G attains its minimum value on the unit sphere $\big\{t \in \mathbb{R}^n \big| \|t\| = 1\big\}$, which is compact. Let $m = \min_{\|t\|=1} G(t) = \min_{\|t\|=1} \big\|F(t)\big\|$. Note that $m > 0$ because when $t \neq 0$ we have $F(t) \neq 0$ (since F is a bijective linear map) and hence $\|F(t)\| \neq 0$. For $t \in \mathbb{R}^n$, if $\|t\| > 1$ then we have $\big\|\frac{t}{\|t\|}\big\| = 1$ so, by the choice of m,

$$\big\|F(t)\big\| = \|t\|\,\Big\|F\big(\tfrac{t}{\|t\|}\big)\Big\| \geq \|t\| \cdot m > m.$$

It follows that for all $t \in \mathbb{R}^n$, if $\big\|F(t)\big\| \leq m$ then $\|t\| \leq 1$. Since F is bijective, it follows that for $x \in U$, if $\|x\| \leq m$ then $\big\|F^{-1}(x)\big\| \leq 1$. Thus for all $x \in U$, if $x = 0$ then $\|F^{-1}(x)\| = 0 = \frac{\|x\|}{m}$ and if $x \neq 0$ then since $\big\|\frac{mx}{\|x\|}\big\| = m$ we have

$$\big\|F^{-1}(x)\big\| = \tfrac{\|x\|}{m}\,\big\|F^{-1}\big(\tfrac{mx}{\|x\|}\big)\big\| \leq \tfrac{\|x\|}{m}.$$

For all $x, y \in U$, we have $\big\|F^{-1}(x) - F^{-1}(y)\big\| = \big\|F^{-1}(x-y)\big\| \leq \frac{1}{m}\,\|x-y\|$, so F^{-1} is Lipschitz continuous.

3.3 Corollary: *When U and V are normed linear spaces with U finite-dimensional, every linear map $F : U \to V$ is Lipschitz continuous.*

3.4 Corollary: *When U is a finite-dimensional vector space, any two norms on U induce the same topology, and a sequence converges in one norm if and only if it converges in the other, and a sequence is Cauchy in one norm if and only if it is Cauchy in the other.*

3.5 Definition: *Let Y be a metric space and let $\emptyset \neq X \subseteq Y$. Recall that for $y \in Y$ we define the* **distance** *between y and X to be*

$$d(y,X) = \inf\left\{d(y,x)\big|x \in X\right\}.$$

Recall (or prove) that when X is compact, the minimum value of $d(y,x)$, $x \in X$ is attained and so we can choose $x \in X$ such that $d(y,x) = d(y,X)$.

3.6 Theorem: *Let W be a normed linear space and let $U \subseteq W$ be a finite-dimensional subspace. Then for every $w \in W$ there exists $u \in U$ such that $d(w,u) = d(w,U)$.*

Proof: Let $w \in W$. If $w \in U$ we can take $u = w$ to get $d(w,u) = 0 = d(w,U)$. Suppose that $w \notin U$. Let $d = d(w,U)$ and note that since U is closed we have $d > 0$ (since we can choose $r > 0$ so that $B(w,r) \cap U = \emptyset$ and then $d \geq r$). Let $K = \overline{B}(w, d+1) \cap U$. Note that $d(w,K) = d(w,U)$. Indeed, since $K \subseteq U$ we have $d(w,K) \geq d(w,U) = d$ and, on the other hand, for any $0 < \epsilon < 1$ we can choose $u \in U$ with $d \leq d(w,u) < d + \epsilon < d + 1$, and then we have $u \in K$ hence $d(w,K) \leq d(w,u) < d + \epsilon$. Since K is closed and bounded in U, and U is a finite dimensional vector space (so we have a bijective map $F : \mathbb{R}^n \to U$ with F and F^{-1} both Lipschitz continuous), it follows that K is compact. Since K is compact we can choose $u \in K$ such that $d(w,u) = d(w,K) = d(w,U)$.

3.7 Lemma: *Let W be a normed linear space and let $U \subsetneqq W$ be a proper closed subspace. For every $0 < r < 1$ there exists an element $w \in W \setminus U$ with $\|w\| = 1$ such that $d(w,U) \geq r$.*

Proof: Let $0 < r < 1$. Since $U \subsetneqq W$ we can choose $v \in W \setminus U$. Let $d = d(v,U)$ and note that since U is closed we have $d > 0$. Since $d = \inf\left\{\|u - v\| \,\big|\, u \in U\right\}$ we can choose $u \in U$ such that $d \leq \|v - u\| < \frac{d}{r}$. Let $w = \frac{v-u}{\|v-u\|}$. Then we have $\|w\| = 1$ and for all $x \in U$ we have

$$\|x - w\| = \left\|x - \tfrac{v-u}{\|v-u\|}\right\| = \tfrac{1}{\|v-u\|} \cdot \left\|\|v-u\|x + u - v\right\| \geq \tfrac{r}{d} \cdot d = r.$$

3.8 Theorem: *(Riesz's Theorem) Let U be a normed linear space. Then U is finite-dimensional if and only if the closed unit ball in U is compact.*

Proof: Suppose that U is finite-dimensional. Let $\mathcal{B} = \{u_1, u_2, \cdots, u_n\}$ be a basis for U and let $F : \mathbb{R}^n \to U$ be the isomorphism given by $F(t) = \sum_{k=1}^{n} t_k u_k$. By Theorem 3.2, F and F^{-1} are continuous. Since F is continuous, $F^{-1}\big(\overline{B}(0,1)\big)$ is closed, and since F^{-1} is continuous, $F^{-1}\big(\overline{B}(0,1)\big)$ is bounded (by Theorem 1.26). Since $F^{-1}\big(\overline{B}(0,1)\big)$ is closed and bounded in $\mathbb{R}^n$, it is compact. Since F is a homeomorphism and $F^{-1}\big(\overline{B}(0,1)\big)$ is compact, it follows that $\overline{B}(0,1)$ is also compact.

Suppose that U is infinite dimensional. Choose $u_1 \in U$ with $\|u_1\| = 1$ and let $U_1 = \text{Span}\{u_1\}$. Since U_1 is finite-dimensional, it is closed and it is a proper subspace of U so, by the above lemma, we can choose $u_2 \in U \setminus U_1$ with $\|u_2\| = 1$ such that $d(u_2, U_1) \geq \frac{1}{2}$, and note that this implies that $d(u_2, u_1) \geq \frac{1}{2}$. Let $U_2 = \text{Span}\{u_1, u_2\}$ and note that U_2 is a proper closed subspace of U. By the lemma we can choose $u_3 \in U \setminus U_2$ with $\|u_3\| = 1$ such that $d(u_3, U_2) \geq \frac{1}{2}$, and note that this implies that $d(u_3, u_1) \geq \frac{1}{2}$ and $d(u_3, u_2) \geq \frac{1}{2}$. Repeat this procedure to obtain a sequence $(u_n)_{n \geq 1}$ such that $\|u_n\| = 1$ for all $n \in \mathbf{Z}^+$ and $d(u_n, u_k) \geq \frac{1}{2}$ for $1 \leq k < n$. Then (u_n) is a sequence in the closed unit ball $\overline{B}(0,1)$ which has no convergent subsequence, and hence $\overline{B}(0,1)$ is not compact.

The Hahn-Banach Theorem

3.9 Definition: Let W be a vector space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$ and let $p:W\to\mathbb{R}$. We say that p is **subadditive** when $p(x+y)\leq p(x)+p(y)$ for all $x,y\in W$, we say that p is **homogeneous** when $p(tx)=|t|\,p(x)$ for all $x\in W$ and all $t\in\mathbb{C}$, and we say that p is **positively homogeneous** when $p(tx)=t\,p(x)$ for all $x\in W$ and all $t\in\mathbb{R}$ with $t\geq 0$. A **seminorm** on a vector space W is a subadditive homogeneous map $p:W\to\mathbb{R}$.

3.10 Theorem: *(The Hahn-Banach Theorem for Real Vector Spaces) Let W be a vector space over $\mathbb{R}$, let $U\subseteq W$ be a subspace, and let $p:W\to\mathbb{R}$ be subadditive and positively homogeneous. Then every linear map $f:U\to\mathbb{R}$ with $f(x)\leq p(x)$ for all $x\in U$ extends to a linear map $g:W\to\mathbb{R}$ with $g(x)\leq p(x)$ for all $x\in W$.*

Proof: We claim that when $w\in W\setminus U$ and $V=U+\text{Span}\,\{w\}=\big\{u+tw\,\big|\,u\in U,t\in\mathbb{R}\big\}$, every linear map $f:U\to\mathbb{R}$ with $f(x)\leq p(x)$ for all $x\in U$ extends to a linear map $g:V\to\mathbb{R}$ with $g(x)\leq p(x)$ for all $x\in V$. Note that such an extension g is determined by the value $g(w)\in\mathbb{R}$ and must be given by $g(u+tw)=f(u)+t\,g(w)$ for all $u\in U$ and $t\in\mathbb{R}$. We shall choose $r=g(w)\in\mathbb{R}$ so that the map $g(u+tw)=f(u)+tr$ satisfies the requirement that $g(v)\leq p(v)$ for all $v=u+tw\in V$. Note that for all $x,y\in U$ we have

$$f(x)-f(y)=f(x-y)\leq p(x-y)=p\big((x+w)+(-y-w)\big)\leq p(x+w)+p(-y-w)$$

(by subadditivity) and hence $-p(-y-w)-f(y)\leq p(x+w)-f(x)$. It follows that

$$\sup\big\{-p(-y-w)-f(y)\,\big|\,y\in U\big\}\leq\inf\big\{p(x+w)-f(x)\,\big|\,x\in U\big\}$$

so we can choose $r\in\mathbb{R}$ such that

$$-p(-y-w)-f(y)\leq r\leq p(x+w)-f(x)\ \text{ for all }\ x,y\in U.$$

We define $g:V\to\mathbb{R}$ by $g(u+tw)=f(u)+tr$ for all $u\in U$ and $t\in\mathbb{R}$. We must show that $g(u+tw)\leq p(u+tw)$ for all $u\in U$ and $t\in\mathbb{R}$. Let $u\in U$ and $t\in\mathbb{R}$. If $t=0$ then we have $g(u+tw)=g(u)=f(u)\leq p(u)=p(u+tw)$. If $t>0$ then since $r\leq p\big(\frac{u}{t}+w\big)-f\big(\frac{u}{t}\big)=\frac{1}{t}\big(p(u+tw)-f(u)\big)$ (by positive homogeneity) we have $tr\leq p(u+tw)-f(u)$ hence $g(u+tw)=f(u)+tr\leq p(u+tw)$. Finally, if $t<0$ then since $r\geq -p\big(-\frac{u}{t}-w\big)-f\big(\frac{u}{t}\big)=\frac{1}{t}\big(p(u+tw)-f(u)\big)$ (by positive homogeneity) we have $tr\leq p(u_tw)-f(u)$, hence $g(u+tw)=f(u)+tr\leq p(u+tw)$, as required. This completes the proof of our claim.

We can now complete the proof of Part (1) using Zorn's Lemma. Let S be the set of all linear extensions of f dominated by p, that is the set of all linear maps $g:V\to\mathbb{R}$, where V is a subspace of W containing U, such that $g(x)=f(x)$ for all $x\in U$ and $g(x)\leq p(x)$ for all $x\in V$. Define an order on S by stipulating that $g_1\leq g_2$ when g_2 is an extension of g_1 (or equivalently when the graph of g_2 contains the graph of g_1). Note that every chain $C=\big\{g_\alpha:V_\alpha\to\mathbb{R}\,\big|\,\alpha\in A\big\}$ in S has an upper bound, namely the map $g:V\to\mathbb{R}$, where $V=\bigcup\limits_{\alpha\in A}V_\alpha$, given by $g(x)=g_\alpha(x)$ for any $\alpha\in A$ for which $x\in V_\alpha$ (you should verify, as an exercise, that the map g is well-defined and linear with $g(x)=f(x)$ for all $x\in U$ and $g(x)\leq p(x)$ for all $x\in V$, and that g is an upper bound for C). By Zorn's Lemma, S has a maximal element $g:V\to\mathbb{R}$. By our previous claim, if we had $V\subsetneqq W$ we could choose $w\in W\setminus V$ and extend g to a linear map h defined on $V'=V+\text{Span}\,\{w\}$ with $h(x)\leq p(x)$ for all $x\in V'$, but this would contradict the maximality of g in S. Thus we must have $V=W$ and so the maximal element g in S is an extension of f to all of W.

3.11 Theorem: *(The Hahn-Banach Theorem for Real or Complex Vector Spaces) Let W be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, let $U \subseteq W$ be a subspace, and let $p : W \to \mathbb{R}$ be a seminorm. Then every linear map $f : U \to \mathbb{F}$ with $\big|f(x)\big| \leq p(x)$ for all $x \in U$ extends to a linear map $g : W \to \mathbb{F}$ with $\big|g(x)\big| \leq p(x)$ for all $x \in W$.*

Proof: In the case that $\mathbb{F} = \mathbb{R}$, this follows immediately from the Hahn-Banach Theorem for Real Vector Spaces, because we can extend $f : U \to \mathbb{R}$ to a linear map $g : W \to \mathbb{R}$ with $g(x) \leq p(x)$ for all $x \in W$, and then (since g is linear and p is a seminorm) we also have $-g(x) = g(-x) \leq p(-x) = p(x)$, so that $\big|g(x)\big| \leq p(x)$, for all $x \in W$.

Suppose that $\mathbb{F} = \mathbb{C}$. Let $f : U \to \mathbb{C}$ be $\mathbb{C}$-linear with $\big|f(x)\big| \leq p(x)$ for all $x \in W$. Write $f(x) = u(x) + i\,v(x)$ where $u, v : U \to \mathbb{R}$, and note that u and v are $\mathbb{R}$-linear and we have

$$u(ix) = \text{Re}\,\big(f(ix)\big) = \text{Re}\,\big(i\,f(x)\big) = \text{Re}\,\big(i\,(u(x) + i\,v(x))\big) = -v(x)$$

so that $v(x) = -\,u(ix)$ and $f(x) = u(x) + i\,v(x) = u(x) - i\,u(ix)$. Since $u(x) \leq \big|g(x)\big| \leq p(x)$ for all $x \in U$, using the Hahn-Banach Theorem for Real Vector Spaces, we can extend u to an $\mathbb{R}$-linear map $w : W \to \mathbb{R}$ with $w(x) \leq p(x)$ for all $x \in W$. Define $g : W \to \mathbb{C}$ by $g(x) = w(x) - i\,w(ix)$. Verify that g is $\mathbb{C}$-linear, and note that g extends f because $f(x) = u(x) - i\,u(ix)$ for all $x \in U$. It remains to show that $\big|g(x)\big| \leq p(x)$ for all $x \in W$. Let $x \in W$. Write $g(x) = r\,e^{i\,\theta}$ with $r > 0$ so that $\big|g(x)\big| = r = e^{-i\,\theta}\,g(x) = g\big(e^{-i\,\theta}\,x\big)$. Then we have

$$\big|g(x)\big| = \text{Re}\,\big(|g(x)|\big) = \text{Re}\,\big(g\big(e^{-i\,\theta}\,x\big)\big) = w\big(e^{-i\,\theta}\,x\big) \leq p\big(e^{-i\,\theta}\,x\big) = p(x),$$

as required.

3.12 Theorem: *(The Hahn-Banach Theorem for Bounded Linear Functionals) Let W be a normed linear space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $U \subseteq W$ be a subspace. Then every bounded linear map $f \in U^*$ extends to a bounded linear map $g \in W^*$ with $\|g\| = \|f\|$.*

Proof: Let $f \in U^*$, that is let $f : U \to \mathbb{F}$ be a bounded linear map. Define $p : W \to \mathbb{R}$ by $p(x) = \|f\|\,\|x\|$. Then $p(x+y) = \|f\|\,\|x+y\| \leq \|f\|\big(\|x\| + \|y\|\big) = p(x) + p(y)$, and $p(tx) = \|f\|\,\|tx\| = |t|\,\|f\|\,\|x\| = |t|\,p(x)$, so p is a seminorm. By the above theorem, we can extend f to a linear map $g : W \to \mathbb{F}$ with $\big|g(x)\big| \leq p(x) = \|f\|\,\|x\|$ for all $x \in W$. Since $\big|g(x)\big| \leq \|f\|\,\|x\|$ for all $x \in W$, we have $\|g\| \leq \|f\|$ (so in particular, g is a bounded linear map, that is $g \in W^*$). And since $g(x) = f(x)$ for all $x \in U$ we have $\|g\| = \sup\big\{|g(x)|\,\big|\,x \in W, \|x\| = 1\big\} \geq \sup\big\{|g(x)|\,\big|\,x \in U, \|x\| = 1\big\} = \|f\|$.

3.13 Corollary: *Let W be a normed linear space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $0 \neq w \in W$. Then there exists a bounded linear functional $g \in W^*$ with $g(w) = \|w\|$ and $\|g\| = 1$.*

Proof: Let $U = \text{Span}\,\{w\}$ and define $f : U \to \mathbb{F}$ by $f(tw) = t\,\|w\|$. Then $f \in U^*$ with $f(w) = \|w\|$ and $\|f\| = 1$. By the Hahn-Banach Theorem (for Bounded Linear Functionals), f extends to a bounded linear functional $g \in W^*$ with $\|g\| = \|f\| = 1$.

3.14 Corollary: *Let W be a normed linear space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, let $U \subsetneqq W$ be a proper closed subspace, and let $w \in W \setminus U$. Then there exists a bounded linear functional $g \in W^*$ with $\|g\| = 1$ such that $g(w) = d(w, U)$ and $g(u) = 0$ for all $u \in U$.*

Proof: Let $d = d(w, U)$ and note that $d > 0$ because U is closed and $w \notin U$. Let $V = U + \text{Span}\,\{w\} = \big\{u + tw \,\big|\, u \in U, t \in \mathbb{F}\big\}$. Define $f \in V^*$ by $f(u + tw) = td$. Note that $f(u) = 0$ for all $u \in U$ and $f(w) = d$. We claim that $\|f\| = 1$. Recall (or verify) that for all $t \in \mathbb{F}$ we have $d(tw, U) = |t|\, d(w, U)$. It follows that for all $u \in U$ and $t \in \mathbb{F}$ we have $\big|f(u+tw)\big| = |t|\, d = |t|\, d(w, U) = d(tw, U) \leq d(tw, -u) = \|u+tw\|$ and hence $\|f\| \leq 1$. On the other hand, for all $0 < r < 1$, since $d = d(w, U) = \inf\big\{d(w, x)\big| x \in U\big\}$, we can choose $u \in U$ so that $d \leq d(w, -u) < \frac{d}{r}$ and then we have $\big|f(u+w)\big| = d > r\, d(w, -u) = r\,\|u+w\|$ and hence $\|f\| > r$. Thus $\|f\| = 1$, as claimed. By the Hahn-Banach Theorem, we can extend $f \in V^*$ to a bounded linear map $g \in W^*$ with $\|g\| = \|f\| = 1$.

3.15 Corollary: *Let W be a normed linear space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. If W^* is separable then W is separable.*

Proof: Suppose that W^* is separable. Choose a sequence $(f_n)_{n \geq 1}$ in W^* such that the set $\big\{f_n \,\big|\, n \in \mathbf{Z}^+\big\}$ is dense in W^*. For each $n \in \mathbf{Z}^+$ choose $u_n \in W$ with $\|u_n\| = 1$ such that $\big|f_n(u_n)\big| > \frac{1}{2}\|f_n\|$. Let $U = \text{Span}\,\big\{u_n \,\big|\, n \in \mathbf{Z}^+\big\}$. Recall (or verify) that when $U \subseteq W$ is a subspace of a normed linear space W, the closure $\overline{U}$ of U in W is also a subspace. We claim that $\overline{U} = W$. Suppose, for a contradiction, that $\overline{U} \subsetneqq W$. Choose $w \in W \setminus \overline{U}$. By Corollary 3.14, we can choose $g \in W^*$ with $\|g\| = 1$ such that $g(w) = d(w, \overline{U})$ and $g(v) = 0$ for all $v \in \overline{U}$. In particular, note that $g(u_n) = 0$ for all $n \in \mathbf{Z}^+$. Since $\big\{f_n \,\big|\, n \in \mathbf{Z}^+\big\}$ is dense in W^* we can choose an index $n \in \mathbf{Z}^+$ such that $\|f_n - g\| < \frac{1}{3}$. Then we have

$$1 = \|g\| = \|g - f_n + f_n\| \leq \|g - f_n\| + \|f_n\| < \tfrac{1}{3} + \|f_n\|$$

hence $\|f_n\| > \frac{2}{3}$. Since $\|f_n\| > \frac{2}{3}$, $\big|f_n(u_n)\big| > \frac{1}{2}\|f_n\|$, $g(u_n) = 0$, $\|u_n\| = 1$ and $\|f_n - g\| < \frac{1}{3}$, we have

$$\tfrac{1}{3} < \tfrac{1}{2}\|f_n\| < \big|f_n(u_n)\big| = \big|f_n(u_n) - g(u_n)\big| = \big|(f_n - g)(u_n)\big| \leq \|f_n - g\| < \tfrac{1}{3}$$

which gives the desired contradiction. Thus $\overline{U} = W$ as claimed.

Finally, note that when $\mathbb{F} = \mathbb{R}$, the set $\text{Span}_{\mathbb{Q}}\big\{u_1, u_2, u_3, \cdots\big\}$ is countable and dense in U, hence also dense in $\overline{U} = W$, and when $\mathbb{F} = \mathbb{C}$, the set $\text{Span}_{\mathbb{Q}[i]}\big\{u_1, u_2, \cdots\big\}$ is countable and dense in U, hence also in $\overline{U} = W$.

3.16 Note: In Part 1 of Theorem 1.28 (the Riesz Representation Theorem for the ℓ_p Spaces), we saw that the map $F : \ell_1 \to {\ell_\infty}^*$ given by $F(b)(a) = \sum\limits_{k=1}^{\infty} a_k b_k$ is an injective norm-preserving linear map. Note that F cannot be surjective because if it was an isomorphism of normed linear spaces then, since ℓ_1 is separable ${\ell_\infty}^*$ would also be separable and hence, by the above corollary, ℓ_∞ would be separable (but it is not). Similarly, when $a, b \in \mathbb{R}$ with $a < b$, the injective norm preserving map $F : L_1[a, b] \to L_\infty[a, b]^*$ given by $F(g)(f) = \int_a^b fg$ (as seen in Theorem 1.31, the Riesz Representation Theorem for the L_p Spaces) cannot be surjective because $L_1[a, b]$ is separable but $L_\infty[a, b]$ is not.

The Hahn-Banach Separation Theorem

3.17 Definition: Let U be a real vector space and let $A \subseteq U$. A point $a \in A$ is called an **internal point** of A when for every $u \in U$ there exists $r > 0$ such that $a + tu \in A$ for all $t \in (-r, r)$. The set of internal points of A is called the **core** (or the **algebraic interior**, or the **radial kernel**) of A, and is denoted by Core(A). Note that when U is a normed linear space, the interior of A is contained in the core of A.

3.18 Definition: Let U be a real vector space. Let $A \subseteq U$ be convex with $0 \in \text{Core}(A)$. We define the **Minkowski functional** of A to be the map $p = p_A : U \to \mathbb{R}$ given by

$$p(x) = \inf \left\{ r > 0 \,\middle|\, \tfrac{1}{r}\, x \in A \right\}.$$

Note that the set $\left\{ r > 0 \,\middle|\, \tfrac{1}{r}\, x \in A \right\}$ is nonempty because $0 \in \text{Core}(A)$.

3.19 Theorem: *(The Minkowski Functional) Let U be a real vector space and let $A \subseteq U$ be convex with $0 \in \text{Core}(A)$. Then the Minkowski functional of A is positively homogeneous and subadditive.*

Proof: Let $p = p_A$ be the Minkowski functional of A. Then p is positively homogeneous because for $x \in U$ and $t > 0$ we have

$$p(x) = \inf \left\{ r>0 \,\middle|\, \tfrac{1}{r}\, tx \in A \right\} = \inf \left\{ ts \,\middle|\, s>0, \tfrac{1}{s}\, x \in A \right\} = t \cdot \inf \left\{ s>0 \,\middle|\, \tfrac{1}{s}\, x \in A \right\} = t\, p(x).$$

To show that p is subadditive, let $x, y \in U$ and let $\epsilon > 0$. Choose $s \in S = \left\{ r>0 \,\middle|\, \tfrac{1}{r}\, x \in A \right\}$ such that $p(x) \le s < p(x) + \frac{\epsilon}{2}$ and choose $t \in T = \left\{ r>0 \,\middle|\, \tfrac{1}{r}\, y \in A \right\}$ with $p(y) \le t < p(y) + \frac{\epsilon}{2}$. Since $\frac{1}{s}\, x \in A$ and $\frac{1}{t}\, y \in A$ and A is convex, we have

$$\tfrac{1}{s+t}(x + y) = \tfrac{s}{s+t} \cdot \tfrac{1}{s}\, x + \tfrac{t}{s+t} \cdot \tfrac{1}{t}\, y \in A$$

so that $s + t \in R = \left\{ r>0 \,\middle|\, \tfrac{1}{r}(x+y) \in A \right\}$. Thus $p(x+y) = \inf R \le s+t < p(x) + p(y) + \epsilon$. Since $p(x + y) < p(x) + p(y) + \epsilon$ for all $\epsilon > 0$, it follows that $p(x + y) \le p(x) + p(y)$.

3.20 Theorem: *(The Hahn-Banach Separation Theorem) Let U be a real vector space. Let A and B be disjoint nonempty convex subsets of U, with $\text{Core}(A) \ne \emptyset$. Then there exists a nonzero linear map $f : U \to \mathbb{R}$ such that $f(x) \le f(y)$ for every $x \in A$ and $y \in B$.*

Proof: Let $a \in \text{Core}(A)$, let $b \in B$, and let C be the convex set $C = A - B - a + b$. Since $A \cap B = \emptyset$ we have $0 \notin A - B$ so $b - a \notin C$. Since $a \in \text{Core}(A)$ we have $0 \in \text{Core}(A - a)$, and since $A - a \subseteq (A - a) - (B - b) = C$, we also have $0 \in \text{Core}(C)$. Let $p : U \to \mathbb{R}$ be the Minkowski functional of C, given by $p(x) = \inf \left\{ r > 0 \,\middle|\, \tfrac{1}{r}\, x \in C \right\}$. Since $0 \in C$ and $b - a \notin C$ and C is convex, we have $t(b - a) \notin C$ for all $t \ge 1$ and so $p(b - a) \ge 1$. On the other hand, we have $p(x) \le 1$ for all $x \in C$.

Let $f : \text{Span}\{b - a\} \to \mathbb{R}$ be the linear map given by $f\big(t(b - a)\big) = t\, p(b - a)$. When $t > 0$, since p is positively homogeneous we have $f\big(t(b-a)\big) = t\, p(b-a) = p\big(t(b-a)\big)$, and when $t \le 0$, since p is nonnegative we have $f\big(t(b - a)\big) = t\, p(b - a) \le 0 \le p\big(t(b - a)\big)$, and so $f(x) \le p(x)$ for all $x \in \text{Span}\{b - a\}$. By the Hahn Banach Theorem (for Real Vector Spaces), we can extend f to a linear map $f : U \to \mathbb{R}$ with $f(x) \le p(x)$ for all $x \in U$. For all $x \in A$ and $y \in B$, since $x - y - a + b \in C$ we have

$$1 \ge p(x-y+a-b) \ge f(x-y-a+b) = f(x)-f(y)+f(b-a) \ge f(x)-f(y)+1$$

so that $f(x) \le f(y)$.

3.21 Exercise: Let $U = \mathbb{R}^\infty$, let $A = \left\{ a = \sum\limits_{k=1}^{n} a_k e_k \,\middle|\, n \in \mathbf{Z}^+, a_n > 0 \right\}$, and let $B = \{0\}$. Show that A and B are disjoint nonempty convex subsets of U, but there is no nonzero linear map $f : U \to \mathbb{R}$ with $f(x) \le f(y)$ for all $x \in A$ and $y \in B$.

3.22 Definition: Let $a \le b$ and let $f : [a,b] \to \mathbb{R}$. For a partition $P = (x_0, x_1, \cdots, x_n)$ of $[a,b]$ (so we have $a = x_0 < x_1 < \cdots < x_n = b$), the **variation** of f for the partition P is

$$V(f,P) = \sum_{k=1}^{n} \left| f(x_k) - f(x_{k-1} \right|$$

and the **variation** of f on the interval $[a,b]$ is

$$V\big(f,[a,b]\big) = \sup\Big\{ V(f,P) \Big| P \text{ is a partition of } [a,b] \Big\}.$$

We say that f is of **bounded variation** on $[a,b]$ when $V\big(f,[a,b]\big) < \infty$, and we write

$$\mathcal{BV}[a,b] = \Big\{ f : [a,b] \to \mathbb{R} \Big| f \text{ is of bounded variation} \Big\}.$$

3.23 Theorem: *Let $a \le b$ and let $f : [a,b] \to \mathbb{R}$. Then f is of bounded variation on $[a,b]$ if and only if f is rectifiable (meaning that the graph of f has finite length).*

Proof: Recall that the length of the graph of f on $[a,b]$ is defined as follows. For a partition $P = (x_0, x_1, \cdots, x_n)$ of $[a,b]$, we define

$$L\big(f,P\big) = \sum_{k=1}^{n} \sqrt{\big(f(x_k) - f(x_{k-1})\big)^2 + (x_k - x_{k-1})^2}$$

and then the length of the graph of f on $[a,b]$ is given by

$$L(f,[a,b]\big) = \sup\Big\{ L(f,P) \Big| \ P \text{ is a partition of } [a,b] \Big\}.$$

For any partition $P = (x_0, \cdots, x_n)$ of $[a,b]$, since

$$\big| f(x_k) - f(x_{k-1}) \big| \le \sqrt{\big(f(x_k) - f(x_{k-1})\big)^2 + (x_k - x_{k-1})^2}$$

for all indices k, it follows that $V(f,P) \le L(f,P)$. Since $V(f,P) \le L(f,P)$ for all partitions P, it follows that $V(f,[a,b]) \le L(f,[a,b])$. On the other hand, for all partitions $P = (x_0, \cdots, x_n)$ of $[a,b]$ we have

$$\sqrt{\big(f(x_k) - f(x_{k-1})\big)^2 + (x_k - x_{k-1})^2} \le \big| f(x_k) - f(x_{k-1}) \big| + \big(x_k - x_{k-1}\big)$$

for all k, it follows that $L(f,P) \le V(f,P) + \sum_{k=1}^{n} (x_k - x_{k-1}) = V(f,P) + (b-a)$. Since $L(f,P) \le V(f,P) + (b-a)$ for all P, it follows that $L(f,[a,b]\big) \le V\big(f,[a,b]\big) + (b-a)$.

3.24 Definition: Let $g \in \mathcal{BV}[a,b]$. For a partition $P = (x_0, x_1, \cdots, x_n)$ of $[a,b]$, write $\|P\| = \max\big\{ x_k - x_{k-1} \big| 1 \le k \le n \big\}$. For $f \in \mathcal{C}[a,b] = \mathcal{C}\big([a,b],\mathbb{R}\big)$, we define the **Riemann-Stieltjes integral** of f on $[a,b]$ with respect to the weight function g to be

$$\int_a^b f\,dg = \int_a^b f(x)\,dg(x) = \lim_{\|P\| \to 0} \sum_{k=1}^{n} f(t_k)\big(g(x_k) - g(x_{k-1}\big).$$

This means that $\int_a^b f\,dg$ is the (unique) real number such that for every $\epsilon > 0$ there exists $\delta > 0$ such that for every partition $P = (x_0, \cdots, x_n)$ of $[a,b]$ with $\|P\| < \delta$, and for all $t_1, t_2, \cdots, t_n$ with each $t_k \in [x_{k-1}, x_k]$ we have

$$\left| \int_a^b f\,dg - \sum_{k=1}^{n} f(t_k)\big(g(x_k) - g(x_{k-1})\big) \right| < \epsilon.$$

3.25 Exercise: Verify, as an exercise, that when $g \in \mathcal{BV}[a,b]$ and $f \in \mathcal{C}[a,b]$, the Riemann-Stieltjes integral $\displaystyle\int_a^b f\,dg$ exists and is uniqe with $\displaystyle\left|\int_a^b f\,dg\right| \leq V\big(g,[a,b]\big)\cdot \|f\|_\infty$.

3.26 Theorem: *(The Riesz Representation Theorem) For every $L \in \big(\mathcal{C}[a,b]\big)^*$ there exists $g \in \mathcal{BV}[a,b]$ with $g(a) = 0$ and $V\big(g,[a,b]\big) = \|L\|$ such that for all $f \in \mathcal{C}[a,b]$ we have*

$$L(f) = \int_a^b f\,dg.$$

Proof: Let $L \in \big(\mathcal{C}[a,b]\big)^*$, that is let $L : \mathcal{C}[a,b] \to \mathbb{R}$ be a bounded linear functional. By the Heine-Borel Theorem, we can extend L to a bounded linear map $M : \mathcal{B}[a,b] \to \mathbb{R}$, that is to $M \in \big(\mathcal{B}[a,b]\big)^*$, with $\|M\| = \|L\|$. For $a < x \leq b$, let $s_x : [a,b] \to \mathbb{R}$ be the step function given by $s_x(t) = 1$ for $a \leq t \leq x$ and $s_x(t) = 0$ for $x < t \leq b$, and let $s_a(t) = 0$ for all $t \in [a,b]$. Define $g : [a,b] \to \mathbb{R}$ by $g(x) = M(s_x)$.

We claim that $g \in \mathcal{BV}[a,b]$ with $V\big(g,[a,b]\big) \leq \|M\| = \|L\|$. Let $P = (x_0, x_1, \cdots, x_n)$ be any partition of $[a,b]$. For $y \in \mathbb{R}$, let $\sigma(y)$ be the sign of y, given by $\sigma(y) = \frac{y}{|y|}$ when $y \neq 0$ with $\sigma(0) = 0$. For $1 \leq k \leq n$ let $\epsilon_k = \sigma\big(g(x_k) - g_(x_{k-1})\big)$ so that we have $\big|g(x_k) - g(x_{k-1})\big| = \epsilon_k\big(g(x_k) - g(x_{k-1})$. Then

$$\begin{aligned}\sum_{k=1}^n \big|g(x_k) - g(x_{k-1})\big| &= \sum_{k=1}^n \epsilon_k\big(g(x_k) - g(x_{k-1})\big) = \sum_{k=1}^n \epsilon_k\big(M(s_{x_k}) - M(x_{k-1})\big) \\ &= M\Big(\sum_{k=1}^n \epsilon_k(s_{x_k} - s_{x_{k-1}})\Big) \leq \|M\|\Big\|\sum_{k=1}^n \epsilon_k(s_{x_k} - s_{x_{k-1}})\Big\|_\infty \leq \|L\|\end{aligned}$$

since $\|M\| = \|L\|$ and $\Big\|\sum\limits_{k=1}^n \epsilon_k(s_{x_k} - s_{x_{k-1}})\Big\|_\infty \leq 1$ because the function $\sum\limits_{k=1}^n \epsilon_k(s_{x_k} - s_{x_{k-1}})$ only takes the values 0 and ± 1. Thus $g \in \mathcal{BV}[a,b]$ with $V\big(g,[a,b]\big) \leq \|L\|$, as claimed.

Note that if we can show that $M(f) = \int_a^b f\,dg$ for all $f \in \mathcal{C}[a,b]$ then, from Exercise 3.25, when $\|f\|_\infty \leq 1$ we have $V\big(g,[a,b]\big) \geq \big|\int_a^b f\,dg\big| = \big|M(f)\big|$, and it follows that $V\big(g,[a,b]\big) \geq \|M\| = \|L\|$. Thus it remains to show that $M(f) = \int_a^b f\,dg$ for all $f \in \mathcal{C}[a,b]$. Let $f \in \mathcal{C}[a,b]$. Let $n \in \mathbf{Z}^+$. Since f is uniformly continuous, we can choose $\delta > 0$ such that $\big|f(x) - f(y)\big| < \frac{1}{n}$ for all $x, y \in [a,b]$ with $|x - y| < \delta$. Choose a partition $P_n = (x_0, x_1, \cdots, x_\ell)$ of $[a,b]$ with $\|P\| < \delta$ and $\|P\| < \frac{1}{n}$. Let $I_1 = [x_0, x_1]$ and $I_k = (x_{k-1}, x_k]$ for $1 < k \leq \ell$. Let $f_n = \sum\limits_{k=1}^\ell f(x_k)\big(s_{x_k} - s_{x_{k-1}}\big)$, so f_n is the step function given by $f_n(t) = f(x_k)$ for $t \in I_k$. Note that for all $t \in I_k$ we have $|x_k - t| \leq |x_k - x_{k-1}| \leq \|P\| < \delta$ hence $\big|f(x_k) - f(t)\big| < \frac{1}{n}$, that is $\big|f_n(t) - f(t)\big| < \frac{1}{n}$, and it follows that $\|f_n - f\|_\infty \leq \frac{1}{n}$. Also note that

$$M(f_n) = \sum_{k=1}^\ell f(x_k)\big(M(s_{x_k}) - M(s_{x_{k-1}})\big) = \sum_{k=1}^\ell f(x_k)\big(g(x_k) - g(x_{k-1})\big)$$

which is one of the sums used to approximate the Riemjann-Steiltjes integral $\int_a^b f\,dg$. We do this construction for each $n \in \mathbf{Z}^+$ to obtain a sequence of partitions P_n with $\|P_n\| \to 0$, and a sequence of step functions f_n on these partitions with $\|f_n - f\|_\infty < \frac{1}{n}$ for all $n \in \mathbf{Z}^+$ so that $f_n \to f$ in $\mathcal{B}[a,b]$. Since $M : \mathcal{B}[a,b] \to \mathbb{R}$ is continuous, we have

$$M(f) = M\big(\lim_{n\to\infty} f_n\big) = \lim_{n\to\infty} M(f_n) = \lim_{\|P\|\to 0} \sum_{k=1}^\ell f(x_k)\big(g(x_k) - g(x_{k-1})\big) = \int_a^b f\,dg.$$

The Open Mapping Theorem and The Closed Graph Theorem

3.27 Theorem: *(The Open Mapping Theorem) Let U and V be Banach spaces. Let $F \in \mathcal{B}(U,V)$ be surjective. Then F is open (meaning that the set $FA = \{Fa \,|\, a \in A\}$ is open in V for every open set A in U).*

Proof: We claim that for all $R > 0$ there exists $r > 0$ such that $B(0,r) \subseteq \overline{FB(0,R)}$. Note that $U = \bigcup\limits_{k=1}^{\infty} B(0,k)$. Since F is onto, $V = F\big(\bigcup\limits_{n=1}^{\infty} B(0,n)\big) = \bigcup\limits_{n=1}^{\infty} FB(0,n)$. Since V is complete, the Baire Category Theorem implies that one of the sets $\overline{FB(0,n)}$ has non-empty interior. By scaling, $\overline{FB(0,r)}$ has non-empty interior for all $r > 0$ so, in particular, $\overline{FB(0,1)}$ has non-empty interior. Choose $c \in V$ and $r > 0$ such that $B(c,2r) \subseteq \overline{FB(0,1)}$. Since $FB(0,1)$ is dense in $\overline{FB(0,1)}$, we can choose $a \in B(0,1)$ and $b = Fa \in FB(0,1)$ with $\|b - c\| < r$ and then we have $B(b,r) \subseteq B(c,2r) \subseteq \overline{FB(0,1)}$.

Let $y \in B(0,r)$ and let $\epsilon > 0$. Since $b + y \in B(b,r) \subseteq \overline{FB(0,1)}$, we can choose $z \in FB(0,1)$ such that $\|Fz - b - y\| < \epsilon$. Since $z \in FB(0,1)$ we can choose $x \in B(0,1)$ such that $Fx = z$. Since $x \in B(0,1)$ and $a \in B(0,1)$ we have $x - a \in B(0,2)$, and we have $\|F(x-a) - y\| = \|Fx - Fa - y\| = \|z = b - y\| < \epsilon$. This proves that $y \in \overline{FB(0,2)}$ hence (since $y \in B(0,r)$ was arbitrary) $B(0,r) \subseteq \overline{FB(0,2)}$. By scaling, it follows that for all $R > 0$ there exists $r > 0$ such that $B(0,r) \subseteq \overline{FB(0,R)}$, as claimed.

We claim that for all $R > 0$ there exists $r > 0$ such that $B(0,r) \subseteq FB(0,R)$. By our previous claim, we can choose $r > 0$ such that $B(0,r) \subseteq \overline{FB(0,\frac{1}{2})}$. By scaling, it follows that $B\big(0,\frac{r}{2^k}\big) \subseteq \overline{FB\big(0,\frac{1}{2^{k+1}}\big)}$ for all $k \geq 0$. Let $y \in B(0,r) \subseteq \overline{FB(0,\frac{1}{2})}$. Choose $x_1 \in B\big(0,\frac{1}{2}\big)$ such that $\|y - Fx_1\| < \frac{r}{2}$. Since $y - Fx_1 \in B\big(0,\frac{r}{2}\big) \subseteq \overline{FB\big(0,\frac{1}{4}\big)}$, we can choose $x_2 \in B\big(0,\frac{1}{4}\big)$ such that $\|(y - Fx_1) - Fx_2\| < \frac{r}{4}$, that is $\|y - F(x_1 + x_2)\| < \frac{r}{4}$. Repeat this procedure obtain elements $x_k \in B\big(0,\frac{1}{2^k}\big)$ such that $\|y - F(u_n)\| < \frac{r}{2^n}$ where $u_n = \sum\limits_{k=1}^{n} x_k$. Note that for $\ell > n$ we have

$$\|u_\ell - u_n\| = \Big\| \sum_{k=n+1}^{\ell} x_k \Big\| \leq \sum_{k=n+1}^{\ell} \|x_k\| < \sum_{k=n+1}^{\ell} \tfrac{1}{2^k} < \sum_{k=n+1}^{\infty} \tfrac{1}{2^k} = \tfrac{1}{2^n},$$

and it follows that the sequence (u_n) is Cauchy in U, and hence it converges because U is complete. Let $u = \lim\limits_{n\to\infty} u_n = \sum\limits_{k=1}^{\infty} x_k$. Note that

$$\|u_n\| = \Big\| \sum_{k=1}^{n} x_k \Big\| \leq \sum_{k=1}^{n} \|x_k\| < \|x_1\| + \sum_{k=2}^{n} \tfrac{1}{2^k} < \|x_1\| + \sum_{k=2}^{\infty} \tfrac{1}{2^k} = \|x_1\| + \tfrac{1}{2}$$

and hence $\|u\| = \big\| \lim\limits_{n\to\infty} u_n \big\| = \lim\limits_{n\to\infty} \|u_n\| \leq \|x_1\| + \frac{1}{2} < \frac{1}{2} + \frac{1}{2} = 1$ so that we have $u \in B(0,1)$. Since $\|y - Fu_n\| < \frac{r}{2^n}$ it follows that $Fu_n \to y$ in V. Since F is bounded, hence continuous, we have $F(u) = F\big(\lim\limits_{n\to\infty} u_n\big) = \lim\limits_{n\to\infty} F(u_n) = y$. Thus $y \in FB(0,1)$ and hence (since y was arbitrary) $B(0,r) \subseteq FB(0,1)$. By scaling, it follows that for all $R > 0$ there exists $r > 0$ such that $B(0,r) \subseteq FB(0,R)$.

Finally, we show that F is open. Let $A \subseteq U$ be open. Let $v \in FA$ and choose $u \in A$ such that $Fu = v$. Since A is open, we can choose $R > 0$ so that $B(u,R) \subseteq A$. By our above claim, we can choose $r > 0$ such that $B(0,r) \subseteq FB(0,R)$. By linearity, we have

$$B(v,R) = v + B(0,r) \subseteq v + FB(0,R) = F\big(u + B(0,R)\big) = FB(u,R) \subseteq FA.$$

3.28 Definition: Let X be a metric space equipped with two metrics d_1 and d_2. We say the two metrics are **equivalent** when they induce the same topology on X, that is when every open ball $B_2(a,\epsilon)$ contains an open ball $B_1(a,\delta)$ and vice versa. Equivalently, the two metrics are equivalent when the identity map $I:(X,d_1)\to(X,d_2)$ is a homeomorphism. Similarly, when U is a vector space which equipped with two norms $\|\ \|_1$ and $\|\ \|_2$, we say the two norms are **equivalent** when they induce the same topology on U. By Theorem 1.26, the norms are equivalent when there exist $\ell,m\geq 0$ such that for all $x\in U$ we have $\|x\|_2\leq \ell\,\|x\|_1$ and $\|x\|_1\leq m\|x\|_2$.

3.29 Corollary: *Let U be a vector space equipped with two norms $\|\ \|_1$ and $\|\ \|_2$. Suppose that U is complete under both norms. If there exists $\ell\geq 0$ such that $\|x\|_2\leq \ell\,\|x\|_1$ for all $x\in U$ then the two norms are equivalent. Equivalently, if the identity map $I:(U,d_1)\to (U,d_2)$ is continuous then it is a homeomorphism.*

Proof: If the identity map $I:(U,d_1)\to(U,d_2)$ is continuous, then it is bounded (by Theorem 1.26) and surjective (obviously), so it is open (by the Open Mapping Theorem), and so its inverse $I:(U,d_2)\to(U,d_1)$ is continuous.

3.30 Definition: When X and Y are topological spaces, the **product topology** on $X\times Y$ is the topology given by taking the basic open sets to be sets of the form $A\times B$ where A is open in X and B is open in Y (this will be explained in Chapter 4). When X and Y are metric spaces, there are various ways that one can define a metric on $X\times Y$ so that the induced topology is the product topology. Let us define the **product metric** on $X\times Y$ by

$$d\big((x_1,y_1),(x_2,y_2)\big)=d(x_1,x_2)+d(y_1,y_2).$$

Not only does this metric induce the product topology on $X\times Y$, it also behaves as expected with sequences: if (x_n) is a sequence in X and (y_n) is a sequence in Y and $a\in X$ and $b\in Y$, then $(x_n,y_n)\to(a,b)$ in $X\times Y$ if and only if $x_n\to a$ in X and $y_n\to b$ in Y. Also, (x_n,y_n) is Cauchy in $X\times Y$ if and only if (x_n) is Cauchy in X and (y_n) is Cauchy in Y.

3.31 Definition: Let X and Y be metric spaces and let $f:A\subseteq X\to Y$ be a function. The **graph** of f is the set

$$\text{Graph}\,(f)=\Big\{(x,y)\,\Big|\,x\in A, y=f(x)\Big\}=\Big\{(a,f(a)\,\Big|\,a\in A\Big\}.$$

We say that f has a **closed graph** (or simply that f is **closed**), when the graph of f is closed in $X\times Y$ using the product topology. Note that f has a closed graph when for every sequence (x_n) in A,

$$\text{if } x_n\to a\in X \text{ and } f(x_n)\to b \text{ in } Y \text{ then } a\in A \text{ and } b=f(a).$$

3.32 Theorem: *(The Closed Graph Theorem) Let U and V be Banach spaces, and let $F:U\to V$ be a linear operator. If the graph of F is closed then F is continuous.*

Proof: Denote the norm on U by $\|\ \|_1$ and the norm on V by $\|\ \|_2$. Define a second norm $\|\ \|_3$ on U by $\|x\|_3=\|x\|_1+\|Fx\|_2$. Let (x_n) be a Cauchy sequence in $(U,\|\ \|_3)$. Then (x_n) is Cauchy in $(U,\|\ \|_1)$ and (Fx_n) is Cauchy in $(V,\|\ \|_2)$ (because $\|x_\ell-x_n\|_1\leq\|x_\ell-x_n\|_3$ and $\|x_\ell-x_n\|_2\leq\|x_\ell-x_n\|_3$). Since $(U,\|\ \|_1)$ and $(V,\|\ \|_2)$ are both complete, it follows that (x_n) converges in $(U,\|\ \|_1)$ and (Fx_n) converges in $(V,\|\ \|_2)$, say $x_n\to a$ in $(U,\|\ \|_1)$ and $Fx_n\to b$ in $(V,\|\ \|_2)$. Since F has a closed graph, we have $b=Fa$, and it follows that $x_n\to a$ in $(U,\|\ \|_3)$ because $\|x_n-a\|_3=\|x_n-a\|1+\|Fx_n-b\|_2$. Thus $(U,\|\ \|_3)$ is complete. Since U is complete under both $\|\ \|_1$ and $\|\ \|_3$ and we have $\|x\|_1\leq\|x\|_3$ for all $x\in U$, it follows from Corollary 3.29 that there exists $\ell\geq 0$ such that $\|x\|_3\leq\ell\|x\|_1$ for all $x\in U$. Thus F is bounded, hence continuous.