Mathematics 595 (CAP/TRA) Fall 2005

3 Calculus in normed vector spaces

3.1 Topological and normed vector spaces

In this section, we introduce the notions of topological vector space and normed vector space, which tie together the algebraic theory of linear spaces with the analytic/topological theory of metric/topological spaces. This provides us with sufficient machinery to develop an abstract theory of (differential) calculus.

Definition 3.1.1. A *topological vector space* (TVS) is a vector space V over $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ equipped with a topology $\mathcal{T}$ so that addition and scalar multiplication are continuous. More precisely, the maps $A : V \times V \to V$, $A(v, w) = v + w$ and $M : \mathbb{F} \times V \to V$, $M(\alpha, v) = v$, are continuous. (Here we use the Tychonov product topology on $V \times V$.)

Normed vector spaces (NVS) are basic examples of TVS's. Recall that $(V, || \cdot ||)$ is a *normed vector space* if $|| \cdot ||$ is a norm on V, i.e., $|| \cdot || : V \to \mathbb{R}$, $||v|| = 0$ if and only if $v = 0$, $||\alpha v|| = |\alpha|\,||v||$ for all $\alpha \in \mathbb{F}$ and $v \in V$, and $||v + w|| \le ||v|| + ||w||$ for all $v, w \in V$. Every norm on a vector space generates a (translation invariant) metric by the formula $d(v, w) = ||v - w||$, which in turn induces a topology on $\mathcal{T}_d$.

Proposition 3.1.2. *Every normed vector space is a topological vector space.*

Proof. It's enough to verify that A and M are continuous according to the ϵ-δ definition of continuity in (V, d), since the topology on V comes from the metric d. Moreover, the Tychonov product topology on $V \times V$ is the topology generated by the usual (Euclidean) product metric $d \times d$, i.e.,

$$(d \times d)((v_1, w_1), (v_2, w_2)) = \sqrt{||v_1 - w_1||^2 + ||v_2 - w_2||^2}.$$

Similarly, the product topology on $\mathbb{F} \times V$ is the topology generated by the Euclidean product metric $d_{\mathbb{F}} \times d$, where $d_{\mathbb{F}}$ is the usual Euclidean metric in $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

First, we check that addition is continuous. Let $(v_0, w_0) \in V \times V$ and let $\epsilon > 0$. Set $\delta = \epsilon/\sqrt{2}$. If $(v, w) \in V \times W$ satisfies $(d \times d)((v, w), (v_0, w_0)) < \delta$, then

$$||v + w - v_0 - w_0|| \le ||v - v_0|| + ||w - w_0|| \le \sqrt{2}(d \times d)((v, w), (v_0, w_0)) < \sqrt{2}\,\delta = \epsilon.$$

Next, we check that scalar multiplication is continuous. Let $v_0 \in V$ and $\alpha_0 \in \mathbb{F}$ and let $\epsilon > 0$. Set

$$\delta = \min\left\{1, \frac{1}{1 + |\alpha_0| + ||v_0||}\epsilon\right\}.$$

If $v \in V$ and $\alpha \in \mathbb{F}$ satisfy $(d_{\mathbb{F}} \times d)((\alpha_0, v_0), (\alpha, v)) < \delta$, then

$$\begin{aligned} ||\alpha v - \alpha_0 v_0|| &\le ||\alpha(v - v_0)|| + ||(\alpha - \alpha_0) v_0|| \\ &= |\alpha|\,||v - v_0|| + |\alpha - \alpha_0|\,||v_0|| \\ &< (|\alpha_0| + \delta)\delta + ||v_0||\,\delta \\ &\le (1 + |\alpha_0| + ||v_0||)\delta \le \epsilon. \end{aligned}$$

□

There are certain basic normed vector spaces which we will refer to repeatedly. Some of these have already been introduced earlier in these notes. For the sake of completeness, we present here a list of old and new examples. All of these are infinite-dimensional (as vector spaces).

Examples 3.1.3. (1) The space $C([0,1])$ of continuous functions on the interval $[0,1]$, equipped with the maximum norm $||f||_\infty = \max_{0\le t\le 1}|f(t)|$. We can also consider the spaces $C^k([0,1])$, $k\in\mathbb{N}$, as normed vector spaces in two ways: (i) as subspaces of $C([0,1])$, equipped with the same norm, or (ii) with the norm

$$||f||_{k,\infty} = \sum_{i=1}^k \max_{0\le t\le 1}|f^{(i)}(t)|.$$

Note that $V = (C^k([0,1]), ||\cdot||_\infty)$ and $W = (C^k([0,1]), ||\cdot||_{k,\infty})$ are different as normed vector spaces; when $k=1$ the sequence (f_n), $f_n(t)=\sin(nt)$, is bounded in V but not in W.

(2) Recall that we defined the Lebesgue space L as (a quotient of) the completion of $C([0,1])$ with respect to the norm $||f||_1 = \int_0^1 |f(t)|\,dt$. We denote this space by $L^1([0,1])$, and define spaces $L^p([0,1])$, $1\le p<\infty$, in the same way, by completing $C([0,1])$ with respect to the norm $||f||_p = (\int_0^1 |f(t)|^p\,dt)^{1/p}$.

The space $L^\infty([0,1])$ must be defined in a different manner since it does not contain $C([0,1])$ densely. We let $L^\infty([0,1])$ denote the collection of essentially bounded functions in $L = L^1([0,1])$, where we say that an equivalence class $[f]\in L$ is *essentially bounded* if there exists a constant $M<\infty$ so that the set $\{t\in[0,1] : |f(t)|>M\}$ has zero measure.

Exercise 3.1.4. Prove that the notion of essential boundedness for an equivalence class $[f]$ in L is independent of the choice of function representing this equivalence class.

We equip $L^\infty([0,1])$ with the maximum norm

$$||f||_\infty = \inf\{M : \{t\in[0,1] : |f(t)|>M\} \text{ has zero measure}\}.$$

(3) The sequence space analogs of the L^p spaces of the previous example are the spaces ℓ^p, $1\le p\le\infty$. We let

$$\ell^p = \begin{cases} \{x=(x_1,x_2,\ldots) : \sum_n |x_n|^p < \infty\}, & 1\le p<\infty, \\ \{x=(x_1,x_2,\ldots) : \sup_n |x_n| < \infty\}, & p=\infty. \end{cases}$$

We equip ℓ^∞ with the maximum norm $||\cdot||_\infty$, and we equip ℓ^p with the p-norm

$$||x||_p = \left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p}.$$

(4) The subspace $c_0\subset\ell^\infty$ consists of all sequences which converge to zero:

$$c_0 = \{x=(x_1,x_2\ldots)\in\ell^\infty : \lim_n |x_n| = 0\}.$$

(5) The space ℓ^∞ can be generalized to base sets of arbitrary cardinality. Let Z be any set and let $\ell^\infty(Z)$ be the vector space of all functions $f: Z\to\mathbb{R}$ for which $||f||_\infty := \sup_{z\in Z}|f(z)| < \infty$. Then $(\ell^\infty(Z), ||\cdot||_\infty)$ is a normed vector space.

The following definition identifies the basic class of spaces which we will consider in this chapter.

Definition 3.1.5. A *Banach space* is a complete normed vector space.

Completeness here means completeness of the associated (translation-invariant) metric space.

The following proposition gives a useful test of whether a given NVS is Banach.

Proposition 3.1.6. *A normed vector space is a Banach space if and only if every absolutely convergent series of vectors is in fact convergent.*

Proof. More precisely, we must show that a normed vector space $(V, ||\cdot||)$ is complete if and only if, whenever (v_n) is a sequence in V satisfying

$$\sum_{n=1}^{\infty} ||v_n|| < \infty, \tag{3.1.7}$$

then the sequence of partial sums $(\sum_{k=1}^{n} v_k)$ converges in V.

Suppose first that $(V, ||\cdot||)$ is complete, and let (v_n) be a sequence in V satisfying (3.1.7). For each $\epsilon > 0$ there exists N so that $\sum_{n \geq N} ||v_n|| < \epsilon$, whence

$$\left|\left|\sum_{k=1}^{n} v_k - \sum_{k=1}^{m} v_k\right|\right| \leq \sum_{k=m+1}^{n} ||v_k|| < \epsilon$$

for all $n > m \geq N$. Thus the sequence of partial sums is Cauchy, and hence converges since V is complete.

To prove the converse, assume that every absolutely convergent sequence has a convergent sequence of partial sums, and let (w_n) be a Cauchy sequence in V. By passing to a subsequence if necessary, we may assume that $||w_n - w_{n-1}|| < 2^{-n}$ for each n. Define $v_1 = w_1$ and $v_n = w_n - w_{n-1}$ for each $n \geq 2$. Then $\sum_{n=1}^{\infty} ||v_n|| < \infty$ so $(\sum_{k=1}^{n} v_k)$ converges. Since $\sum_{k=1}^{n} v_k = w_n$, (w_n) converges. Thus V is complete. □

All of the examples described above in Example 3.1.3 are Banach spaces.

3.2 Continuous linear maps and the dual space

Definition 3.2.8. Let $(V, ||\cdot||)$ and $(W, ||\cdot||)$ be normed vector spaces. A linear map $T : V \to W$ is said to be *bounded* if there exists a constant $C > 0$ so that

$$||T(v)|| \le C||v|| \tag{3.2.9}$$

for all $v \in V$.

In other words, T is bounded if it is a Lipschitz map, when viewed as a map from V to W, where V and W are equipped with the translation invariant metrics coming from the norms.

In general metric spaces, the class of Lipschitz functions is significantly smaller than the class of all continuous functions. The following theorem shows that, for linear maps between normed vector spaces, these two classes agree.

Theorem 3.2.10. *A linear map $T : V \to W$ is continuous if and only if it is bounded.*

Proof. Every bounded map is Lipschitz and hence continuous. Suppose that $T : V \to W$ is continuous and linear. Then $T(0) = 0$; by continuity there exists $\delta > 0$ so that $||v|| < \delta$ implies $||T(v)|| < 1$. For $v \in V$ nonzero, let $w = \frac{\delta}{2||v||}v$. Then $||w|| < \delta$ so

$$||T(w)|| = \frac{1}{2}\delta\frac{||T(v)||}{||v||} < 1$$

and (3.2.9) holds with $C = 2/\delta$. □

For the rest of this chapter we change our notation from Chapter 1, and use $L(V, W)$ to denote the space of all bounded linear maps from V to W. (Recall that we earlier used this notation for the space of all linear maps. In the finite dimensional case, these two spaces agree.) Note that $L(V, W)$ is a vector space. Can we give it the structure of a normed vector space?

Definition 3.2.11. Let $T \in L(V, W)$. The *operator norm* of T, denoted $||T||$, is the infimum of all constants C so that (3.2.9) holds for all $v \in V$. Equivalently,

$$||T|| = \sup_{v \in V : v \ne 0} \frac{||T(v)||}{||v||} = \sup_{v \in V : ||v|| \le 1} ||T(v)||.$$

Exercise 3.2.12. Prove that $T \mapsto ||T||$ defines a norm on $L(V, W)$.

From now on, unless otherwise stated, $L(V, W)$ is equipped with this norm.

Theorem 3.2.13. *If W is a Banach space, then $L(V, W)$ is a Banach space.*

Proof. Let (T_n) be a Cauchy sequence in $L(V, W)$, and let $v \in V$. Given $\epsilon > 0$ there exists N so that $||T_n - T_m|| < \epsilon/(1 + ||v||)$ whenever $n, m \ge N$. By the definition of the norm in $L(V, W)$, $||T_n(v) - T_m(v)|| < \epsilon$ for all such n, m. Thus $(T_n(v))$ is a Cauchy sequence in W, and hence converges to an element of W. We denote this element by $T(v)$; this defines a function from V to W. Since each T_n is linear, so is T. To see that T is bounded, we fix $v \in V$, $v \neq 0$, and compute

$$\frac{||T(v)||}{||v||} = \lim_{n\to\infty} \frac{||T_n(v)||}{||v||} \leq \lim_{n\to\infty} ||T_n|| < \infty,$$

since (T_n) is Cauchy. Thus $T \in L(V,W)$. Finally, to see that $T_n \to T$ in $L(V,W)$, we fix $v \in V$, $||v|| \leq 1$, and compute

$$\begin{aligned} ||T(v) - T_n(v)|| &= ||\lim_m T_m(v) - T_n(v)|| \\ &\leq \sup_{m\geq n} ||T_m(v) - T_n(v)|| \\ &\leq \sup_{m\geq n} ||T_m - T_n|| < \epsilon \end{aligned}$$

if $n \geq N$, since (T_n) is Cauchy. Taking the supremum over all $v \in V$, $||v|| \leq 1$, yields

$$n \geq N \quad \Rightarrow \quad ||T - T_n|| < \epsilon,$$

i.e., $T_n \to T$ in $L(V,W)$. □

Suppose that V and W are Banach spaces, and that $T \in L(V,W)$ is an isomorphism, i.e., T is invertible. The Open Mapping Theorem (which we will not prove in this course) asserts that T is an open map: $T(U)$ is open in W whenever U is open in V. Thus T^{-1} is continuous, and hence bounded: $T^{-1} \in L(W,V)$. In other words, a bounded linear isomorphism between Banach spaces is necessarily bi-Lipschitz. If $||T|| = ||T^{-1}|| = 1$, i.e., T is an isometry, T is called an *isometric isomorphism*. Isometrically isomorphic Banach spaces are indistinguishable from both an algebraic and a metric point of view.

Exercise 3.2.14. Let W be a subspace of a normed vector space V. Define a map $||\cdot|| : V/W \to [0,\infty)$ by $||v+W|| = \inf\{||v+w|| : w \in W\}$. Prove that $||\cdot||$ is a norm on V/W if W is closed. In this case, prove that the canonical quotient map $Q : V \to V/W$, $Q(v) = v + W$ is bounded, with $||Q|| \leq 1$.

Exercise 3.2.15. Let $(V_i, ||\cdot||_i)$ be a sequence of normed vector spaces, and let $V = \prod_i V_i = \{v = (v_1, v_2, \ldots) : v_i \in V_i \text{ for all } i\}$ be the product space. Prove that

$$||v||_p := \left(\sum_i ||v_i||_i^p\right)^{1/p}, \qquad 1 \leq p < \infty$$

and

$$||v||_\infty := \sup_i ||v_i||_i$$

define norms on

$$\oplus_p V_i := \{v \in V : ||v||_p < \infty\}$$

and

$$\oplus_\infty V_i = \{v \in V : ||v||_\infty < \infty\}$$

respectively. If each $(V_i, ||\cdot||_i)$ is a Banach space, prove that $(\oplus_p V_i, ||\cdot||_p)$ is a Banach space for each $1 \leq p \leq \infty$. Prove that

$$\{v \in \oplus_\infty V_i : ||v_i|| \to 0\}$$

is a Banach subspace of $(\oplus_\infty V_i, ||\cdot||_\infty)$. Finally, prove that $P_i : \oplus_p V_i \to V_i$, $P_i(v) = v_i$, is a bounded linear map with $||P_i|| \leq 1$.

The case $W = \mathbb{R}$ deserves special mention.

Definition 3.2.16. Let V be a normed vector space. The space $L(V, \mathbb{R})$ is called the *dual space* to V, and is denoted V^*.

By Theorem 3.2.13, V^* is Banach whether or not V is Banach.

Exercise 3.2.17. Prove: if V is a finite-dimensional normed vector space, then V^* is equivalent with V.

Example 3.2.18. The dual of $L^p([0,1])$, $1 \leq p < \infty$, is isometrically isomorphic to $L^q([0,1])$, where q is the Hölder conjugate of p: $\frac{1}{p} + \frac{1}{q} = 1$. Indeed, each $g \in L^q([0,1])$ defines a linear map on $L^p([0,1])$ by the formula

$$f \mapsto \int_0^1 f(t)g(t)\,dt.$$

Let us call this element T_g. Hölder's inequality shows that

$$\left| \int_0^1 f(t)g(t)\,dt \right| \leq ||f||_p ||g||_q; \tag{3.2.19}$$

thus $T_g \in L^p([0,1])^*$ with $||T_g|| \leq ||g||_q$. On the other hand, choosing $f(t) = |g(t)|^{q-1} \operatorname{sign} g(t)$, where $\operatorname{sign} g(t)$ denotes the sign of $g(t)$, makes (3.2.19) an equality, which shows that $||T_g| \geq ||g||_q$. (This proof works if $p > 1$, $q < \infty$; we leave the final case as an exercise.) The converse statement (every bounded linear functional on L^p arises in the above manner from an element of L^q is typically proved in graduate courses in real analysis.

The dual of $L^\infty([0,1])$ is **not** $L^1([0,1])$ but is significantly larger. It has been identified as a certain space of measures. We will not discuss this identification at all in these notes.

Theorem 3.2.20. *(a) The dual of c_0 is isometrically isomorphic with ℓ^1.*

(b) The dual of ℓ^p, $1 \leq p < \infty$, is isometrically isomorphic with ℓ^q, $q = p/(p-1)$.

Proof. First, we prove (a). For each $w = (w_i) \in \ell^1$, the map $v \mapsto \sum_i v_i w_i$ defines a bounded linear map T_w on c_0, with $||T_w|| \leq ||w||_1$. Thus ℓ^1 embeds in $(c_0)^*$. Conversely, if $T \in (c_0)^*$ define $w_i = T(e^i)$ for each i, where e^i is the ith basis element in ℓ^∞. Let $w = (w_i)$ denote the resulting sequence. Apply the boundedness condition for T to the sequence of elements $\sum_{i=1}^n (\operatorname{sign} w_i) e_i$, where $\operatorname{sign} w = w/|w|$ if $w \neq 0$ and $\operatorname{sign} 0 = 0$. The result is

$$\sum_{i=1}^n |w_i| \leq ||T||;$$

letting $n \to \infty$ gives $||w||_1 \leq ||T||$. Thus $w \in \ell^1$. Note also that $T_w = T$. It follows that $w \mapsto T_w$ is an isometric isomorphism.

Case (b) is similar. We do the case $1 < p < \infty$ and leave the case $p = 1$ to the reader. For $w = (w_i) \in \ell^q$ define $v \mapsto \sum_i v_i w_i$ as before. This is a bounded linear map T_w on ℓ^p, with $||T_w|| \leq ||w||_q$. Thus $(\ell^p)^*$ embeds in ℓ^q. Conversely, if $T \in (\ell^p)^*$ define $w_i = T(e^i)$ for each i and $w = (w_i)$. Apply the boundedness condition for T to the sequence of elements $\sum_{i=1}^n (\operatorname{sign} w_i)|w_i|^{q-1} e_i$ (note that $q < \infty$ since $p > 1$) to conclude

$$\sum_{i=1}^n |w_i|^q \leq ||T|| \left(\sum_{i=1}^n |w_i|^{p(q-1)} \right)^{1/p} = ||T|| \left(\sum_{i=1}^n |w_i|^q \right)^{1/p} .$$

Dividing by $(\sum_{i=1}^n |w_i|^q)^{1/p}$ and passing to the limit as $n \to \infty$ shows that $||w||_q \leq ||T||$ and completes the proof. $\square$

3.3 Geometry of the dual space

The dual space V^* can be identified with a certain collection of subspaces of V. This provides for a geometric understanding of the dual space which complements its algebraic definition.

To begin, assume only that V is a vector space.

Definition 3.3.21. A *hyperplane* in V is a codimension one subspace W, i.e., $\dim(V/W) = 1$.

There is a one-to-one correspondence between hyperplanes and nonzero linear functionals on V: to each linear map $T : V \to \mathbb{R}$, $T \neq 0$, associate the hyperplane $W = \ker(T)$, and conversely, to each hyperplane W, choose an isomorphism $\Phi : V/W \to \mathbb{R}$ and associate to W the linear map $T = \Phi \circ Q$, where $Q : V \to V/W$ is the standard quotient map. Note that $\ker(T) = \ker(Q) = W$.

Suppose that S and T are two nonzero linear functionals with $\ker(S) = \ker(T)$. Choose $v_0 \in V$ with $S(v_0) = 1$, then $T(v_0) \neq 0$. For $v \in V$, we have $v - S(v)v_0 \in \ker(S) = \ker(T)$, i.e., $T = v_0\, S$. Putting all of this together, we conclude

Proposition 3.3.22. *A subspace of V is a hyperplane if and only if it is the kernel of a nonzero linear functional. Two linear functionals have the same kernel if and only if one is a nonzero multiple of the other.*

Now assume that $(V, ||\cdot||)$ is a normed vector space. Which hyperplanes correspond to *continuous* linear maps?

Proposition 3.3.23. *Every hyperplane W in a normed vector space V is either closed or dense. Under the above equivalence, closed hyperplanes correspond to continuous linear functionals.*

Proof. Consider $\overline{W}$, the closure of W. Since addition and scalar multiplication are continuous, $\overline{W}$ is also a subspace of V. Since $W \subset \overline{W}$ and $\dim(V/W) = 1$, we must have $\dim(V/\overline{W}) = 0$ or $= 1$. In the former case, $\overline{W} = V$ and W is dense. In the latter case, $\overline{W} = W$ and W is closed.

To prove the second part, note that $\ker(T) = T^{-1}(\{0\})$. Thus $\ker(T)$ is closed if T is continuous. Conversely, if $W = \ker(T)$ is closed, then $W = \ker(\Phi \circ Q)$, where $Q : V \to V/\ker(T)$ is the canonical quotient map and Φ is an isomorphism from $V/\ker(T)$ to $\mathbb{R}$. Since $V/\ker(T)$ is finite-dimensional, Φ is continuous. By Exercise 3.2.14, Q is continuous. The proof of Proposition 3.3.22 shows that $T = \alpha(\Phi \circ Q)$ for some $\alpha \in \mathbb{R}$; thus T is continuous. $\square$

Examples of closed hyperplanes are easy to find. Any codimension one hyperplane in $\mathbb{R}^n$ (i.e., a level set of a linear function $f : \mathbb{R}^n \to \mathbb{R}$) is such an example. Dense hyperplanes are rather more difficult to construct.

Example 3.3.24. Let $V = c_0$ and let e^m be the mth coordinate vector in V: $e^m_n = 1$ if $m = n$ and $e^m_n = 0$ otherwise. Let $v^0 = (v^0_n) \in c_0$ be the vector with $v^0_n = \frac{1}{n}$. The collection $\{v^0, e^1, e^2, \ldots\}$ is linearly independent (remember that linear dependences must involve only **finite** linear combinations). Let $B = \{v^0, e^1, e^2, \ldots\} \cup \{b^i : i \in I\}$ be a basis for c_0. Define $T : c_0 \to \mathbb{R}$ by setting

$$T(\alpha_0 v^0 + \sum_{m=1}^{\infty} \alpha_m e^m + \sum_{i \in I} \beta_i b^i) = \alpha_0.$$

Since $e^m \in \ker(T)$ for all m, $\ker(T)$ is dense in c_0. But clearly $T \neq 0$ so $\ker(T) \neq V$.

3.4 The Hahn–Banach theorem and applications

The Hahn–Banach theorem is a basic workhorse of functional analysis. It answers the question: when can a linear map defined on a subspace of a Banach space V and satisfying a suitable growth condition admit an extension to all of V which satisfies the same growth condition?

Definition 3.4.25. Let V be a real vector space. A map $p : V \to \mathbb{R}$ is called a *sublinear functional* if $p(v + w) \leq p(v) + p(w)$ for all $v, w \in V$ and $p(\alpha v) = \alpha p(v)$ for all $v \in V$ and all $\alpha \geq 0$.

For example, every norm on V is a sublinear functional.

Theorem 3.4.26 (Hahn–Banach). *Let W be a subspace of a vector space V, let $t : W \to \mathbb{R}$ be a linear functional, and let $p : V \to \mathbb{R}$ be a sublinear functional so that $|t| \leq p$ on W. Then there exists a linear functional $T : V \to \mathbb{R}$ so that $T|_W = t$ and $|T| \leq p$ on V.*

Before giving the proof, we indicate a few corollaries and applications.

Corollary 3.4.27. *Let W be a subspace of a normed vector space V and let $t \in W^*$. Then there exists $T \in V^*$ so that $T|_W = t$ and $||T|| = ||t||$.*

Corollary 3.4.27 follows immediately from the Hahn–Banach theorem by choosing $p(v) = ||t|| \, ||v||$.

Corollary 3.4.28. *Let W be a closed subspace of a normed vector space V and let $v_0 \in V \setminus W$. Then there exists $T \in V^*$ with $T|_W = 0$, $T(v_0) = \operatorname{dist}(v_0, W)$, and $||T|| = 1$.*

Exercise 3.4.29. Prove Corollary 3.4.28 by applying the Hahn–Banach theorem to the quotient space $\tilde{V} = V/W$, with subspace $\tilde{W}$ equal to the one-dimensional space spanned by the vector $\tilde{v}_0 = v_0 + W$ and linear functional $\tilde{t} : \tilde{W} \to \mathbb{R}$ given by $\tilde{t}(\alpha\tilde{v}_0) = \alpha||\tilde{v}_0||$, and with $p(\tilde{v}) = ||\tilde{v}||$. (See Exercise 3.2.14 for the definition of the norm on the quotient of a normed vector space by a closed subspace. Note that $||\tilde{v}_0|| = \operatorname{dist}(v_0, W)$.)

Corollary 3.4.30. *Let $(V, ||\cdot||)$ be a normed vector space and let W be a subspace of V. Then*

$$\overline{W} = \bigcap \{\ker(T) : T \in V^* \text{ and } W \subset \ker(T)\}. \tag{3.4.31}$$

In particular, W is dense in V if and only if the only element of V^* which annihilates W is the zero functional.

To prove Corollary 3.4.30, let X be the subspace on the right hand side of (3.4.31). Continuity of elements of V^* ensures that $\overline{W} \subset X$. If $v_0 \notin \overline{W}$, then $d = \operatorname{dist}(v_0, W) > 0$. By Corollary 3.4.28 there exists $T \in V^*$ so that $T(v_0) > 0$ and $T|_W = 0$. Thus $v_0 \notin X$. Hence $X \subset \overline{W}$ and the proof is complete.

We now begin the proof of the Hahn–Banach theorem. We begin by stating a certain special case of the theorem.

Lemma 3.4.32. *Suppose that the hypotheses of the Hahn–Banach theorem are satisfied and, in addition, that W is a hyperplane in V. Then the conclusion of the Hahn–Banach theorem is true.*

Assuming the validity of this lemma temporarily, let us give the proof of the Hahn–Banach theorem.

Proof of Theorem 3.4.26. The existence of the extension T is established by a nonconstructive argument, using Zorn's Lemma. Consider the family $\mathcal{F}$ of all pairs (S,X), where X is a subspace of V containing W and $S : X \to \mathbb{R}$ is a linear functional satisfying $S|_W = t$ and $|S| \le p$ on X. We equip $\mathcal{F}$ with the following natural partial ordering: $(S_1,X_1) \ll (S_2,X_2)$ if and only if $X_1 \subset X_2$ and $S_2|_{X_1} = S_1$. In order to apply Zorn's Lemma to this poset, we must show that every nonempty chain in $\mathcal{F}$ has an upper bound in $\mathcal{F}$.

Let $\mathcal{C} = (S_i,X_i)_{i\in I}$ be a nonempty chain in $\mathcal{F}$. Define $X = \bigcup_i X_i$. Clearly X is a subspace of V containing W. Define a map $S : X \to \mathbb{R}$ by setting $S(x) = S_i(x)$ if $x \in X_i$. By the definition of the partial ordering in the previous paragraph, S is well-defined. It is easy to check that S is linear. Then (S,X) is an upper bound for $\mathcal{C}$ in $\mathcal{F}$.

By Zorn's Lemma, there exists a maximal element (S,X) in $\mathcal{F}$, i.e., a linear functional $S : X \to \mathbb{R}$, $S|_W = t$, and $|S| \le p$. To complete the proof, it suffices to show that $X = V$. If not, choose $y \in V \setminus X$ and consider the subspace $Y := X \oplus \{y\} = \{x + \alpha y : x \in X, \alpha \in \mathbb{F}\}$. Since X is a hyperplane in Y, Lemma 3.4.32 can be applied to construct an extension of S to Y, contradicting the maximality of (S,X). The proof is finished, modulo the proof of Lemma 3.4.32. ☐

Proof of Lemma 3.4.32. Fix $v_0 \in V \setminus W$ so $V = W \oplus \{v_0\} = \{w + \alpha v_0 : w \in W, \alpha \in \mathbb{R}\}$. Assume that $t : W \to \mathbb{R}$ is as in the statement of the theorem. To begin the proof, let's try to understand what form an extension $T : V \to \mathbb{R}$ would have to take. Indeed, we would have

$$T(\alpha v_0 + w_1) = \alpha T(v_0) + t(w_1) \le p(\alpha v_0 + w_1)$$

for all $\alpha > 0$ and $w_1 \in W$, whence

$$T(v_0) \le -\alpha^{-1}(t(w_1) + p(\alpha v_0 + w_1)) = -t(w_1/\alpha) + p(v_0 + w_1/\alpha).$$

Since $w_1/\alpha \in W$, this implies that

$$T(v_0) \le -t(w_1) + p(v_0 + w_1) \qquad \forall w_1 \in W.$$

Similarly,

$$T(-\alpha v_0 + w_2) = -\alpha T(v_0) + t(w_2) \le p(-\alpha v_0 + w_2)$$

for all $\alpha > 0$ and $w_2 \in W$, whence

$$T(v_0) \ge t(w_2) - p(-v_0 + w_2) \qquad \forall w_2 \in W$$

by a similar argument.

From the above computations, we see that a necessary condition for the existence of such an extension is

$$t(w_2) - p(v_0 + w_2) \le -t(w_1) + p(-v_0 + w_1) \tag{3.4.33}$$

for all $w_1, w_2 \in W$. Indeed, this inequality is needed in order to ensure that $T(v_0)$ is well-defined. But it is also sufficient, since in this case we may choose $T(v_0)$ to be any value in the range

$$[\sup_{w_2 \in W} t(w_2) - p(v_0 + w_2), \inf_{w_1 \in W} -t(w_1) + p(-v_0 + w_1)]$$

and then extend to all of V by linearity.

To see why (3.4.33) holds true, we compute

$$\begin{aligned} t(w_1) + t(w_2) &= t(w_1 + w_2) \le p(w_1 + w_2) \\ &= p(w_1 - v_0 + v_0 + w_2) \le p(-v_0 + w_1) + p(v_0 + w_2). \end{aligned}$$

□

Double duals and reflexivity. The *double dual* of V is the space V^{**}. It is the dual of the dual space V^*, and is always a Banach space.

Proposition 3.4.34. *There is a canonical isometric embedding of V into V^{**}.*

Proof. For each $v \in V$, we consider the map $\Xi_v : V^* \to \mathbb{R}$ given by

$$\Xi_v(T) = T(v).$$

The map Ξ_v is linear, and

$$||\Xi_v|| = \inf\{C : |T(v)| \le C||T|| \,\forall T\} \le ||v||$$

for all $v \in V$, $v \neq 0$. On the other hand, Corollary 3.4.28 of the Hahn-Banach theorem, applied with $W = \{0\}$ and $v_0 = v$, guarantees the existence of $T^v \in V^*$ with $||T^v|| = 1$ and $T^v(v) = ||v||$. If $C < ||v||$, then $|T^v(v)| > C||T^v||$ so $C < ||\Xi_v||$. Thus $||\Xi_v|| \ge ||v||$ and so $v \mapsto \Xi_v$ is an isometric embedding. □

Definition 3.4.35. A Banach space V is *reflexive* if the map $v \mapsto \Xi_v$ of the previous proposition is onto (thus V^{**} is isometrically isomorphic with V).

Every finite-dimensional Banach space is reflexive. It follows from the identification in Example 3.2.18 and Exercise 3.2.20 that $L^p([0,1])$ and ℓ^p are reflexive if $1 < p < \infty$. On the other hand, c_0 is not reflexive, since $(c_0)^{**}$ is isometrically isomorphic with ℓ^∞. Similarly, ℓ^1 is not reflexive, since $(\ell^1)^{**}$ is isometrically isomorphic with $(\ell^\infty)^*$.

Remark 3.4.36. The definition of reflexivity requires that the particular map $v \mapsto \Xi_v$ given of Proposition 3.4.34 be onto. There are non-reflexive Banach spaces V for which there exists a surjective map from V to V^{**} (not the map $v \mapsto \Xi_v$).

An application: Banach limits. Let

$$c = \{v = (v_n) \in \ell^\infty : \lim_n v_n \text{ exists}\}.$$

Exercise 3.4.37. Prove that c is a closed subspace of ℓ^∞, hence a Banach space.

The "limit" operator l assigns to each element in c its limit:

$$l(v) = \lim_n v_n.$$

This is a bounded linear functional on c, with $||l|| = 1$. This obviously makes sense for elements of c. What could it mean for elements of the larger space ℓ^∞? It may seem unlikely that any sort of limit operator could have a meaning in this case; consider, for example, an element $v = (v_n) \in \ell^\infty$ for which the coordinates v_n enumerate the rationals in $(0, 1)$. On the other hand, we have the following

Theorem 3.4.38. *The linear functional l on c can be extended to a linear functional L on ℓ^∞ with the following properties:*

- *(a)* $||L|| = 1$,
- *(b)* *L is monotone:* $v \in \ell^\infty$, $v_n \geq 0 \Rightarrow L(v) \geq 0$, *and*
- *(c)* $L(v_1, v_2, \ldots) = L(v_2, v_3, \ldots)$ *for all* $(v_1, v_2, \ldots) \in \ell^\infty$.

We will give a proof for Theorem 3.4.38 using the Hahn–Banach theorem. For a different proof using ultrafilters, consult the link

`ConstructionOfBanachLimitUsingLimitAlongAnUltrafilter.html`

at `planetmath.org/encyclopedia`.

Proof. Let $\sigma : \ell^\infty \to \ell^\infty$ be the shift map:

$$\sigma(v_1, v_2, \ldots) = (v_2, v_3, \ldots).$$

Define $W = (\mathrm{Id} - \sigma)(\ell^\infty) = \{v - \sigma(v) : v \in \ell^\infty\}$. W is a subspace of ℓ^∞. Let $\mathbf{1} = (1, 1, 1, \ldots) \in \ell^\infty$.

We claim that $\mathrm{dist}(\mathbf{1}, W) = 1$. Clearly $\mathrm{dist}(\mathbf{1}, W) \leq 1$. To prove the other inequality, let $v \in \ell^\infty$. Notice that $(v - \sigma(v))_n = v_n - v_{n+1}$. If $(v - \sigma(v))_n \leq 0$ for any n, then $||\mathbf{1} - (v - \sigma(v))||_\infty \geq 1$. On the other hand, if $(v - \sigma(v))_n \geq 0$ for all n, then $v = (v_n)$ is a nonincreasing sequence. Since v is bounded, $\lim_n v_n$ exists, whence $\lim_n (v - \sigma(v))_n = 0$. Again, we find

$$||\mathbf{1} - (v - \sigma(v))||_\infty = \sup_n |1 - (v - \sigma(v))_n| \geq 1.$$

Thus $\mathrm{dist}(\mathbf{1}, W) \geq 1$.

By Corollary 3.4.28, there exists a linear functional $L : \ell^\infty \to \mathbb{R}$ satisfying $||L|| - 1$, $L(\mathbf{1}) = 1$, and $L|_W = 0$. Thus parts (a) and (c) are satisfied. To prove (b), suppose that $v \in \ell^\infty$ has $v_n \geq 0$ for all n but $L(v) < 0$. By scaling v if necessary, we may assume that $0 \le v_n \le 1$ for all n. Then $||\mathbf{1} - v||_\infty \le 1$ and $L(\mathbf{1} - v) = 1 - L(v) > 1$, contradicting (a).

Finally, we must show that $L|_c = l$. Since $L(\alpha\mathbf{1}) = \alpha$ for any α, it's enough to show that $L|_{c_0} = 0$. For $v \in c_0$, define $v^1 = \sigma(v)$ and $v^{m+1} = \sigma(v^m)$ for $m \ge 1$. Observe that for each m we have $v^m - v = (v^m - v^{m-1}) + \cdots + (v^1 - v) \in W$ and so $L(v) = L(v^m)$. Since $v \in c_0$, $\lim_m ||v^m||_\infty = 0$ so $|L(v)| = \lim_m |L(v^m)| = 0$. Thus $v \in \ker(L)$. ☐

A linear functional $L : \ell^\infty \to \mathbb{R}$ of the type constructed in 3.4.38 is called a *Banach limit operator*. Banach limits have a variety of uses in functional analysis which we will not discuss.

Remark 3.4.39. Here is an interesting observation: **every** Banach limit assigns value $\frac{1}{2}$ to the sequence $v = (1, 0, 1, 0, \ldots)$. This is despite the fact that this sequence is not convergent. To prove this, use the shift invariance and linearity of Banach limits to conclude that $2L(v) = L(\mathbf{1}) = 1$.

An element $v \in \ell^\infty$ with the property that $L(v)$ is independent of the choice of Banach limit L is called *almost convergent*. Lorentz (1948) gave the following characterization for almost convergent sequences: a sequence $v = (v_n)$ is almost convergent to a limit L if and only if

$$\lim_{p\to\infty} \frac{v_n + \cdots + v_{n+p-1}}{p} = L$$

uniformly in n, i.e., for all $\epsilon > 0$ there exists P so that for all $p \ge P$ and all n, we have

$$\left| \frac{v_n + \cdots + v_{n+p-1}}{p} - L \right| < \epsilon.$$

3.5 Differentiation and integration for maps between Banach spaces

Throughout this section, V and W denote Banach spaces, and Ω denotes an open set in V.

Definition 3.5.40. A map $f : \Omega \to W$ is said to be *differentiable* at a point $v_0 \in \Omega$ if there exists $T \in L(V, W)$ so that

$$\lim_{v \to v_0, v \neq v_0} \frac{||f(v) - f(v_0) - T(v - v_0)||}{||v - v_0||} = 0.$$

We write $Df(v_0) = T$ and call this the *derivative* of f at v_0.

The derivative $Df(v_0)$ is uniquely determined if it exists. Suppose that S and T are two elements of $L(V, W)$ satisfying

$$\lim_{v \to v_0, v \neq v_0} \frac{||f(v) - f(v_0) - S(v - v_0)||}{||v - v_0||} = \lim_{v \to v_0, v \neq v_0} \frac{||f(v) - f(v_0) - T(v - v_0)||}{||v - v_0||} = 0.$$

Then

$$\lim_{v \to v_0, v \neq v_0} \frac{||(S - T)(v - v_0)||}{||v - v_0||} = 0.$$

It easily follows that $||S - T|| = 0$, so $S = T$.

Warning! The derivative $Df(v_0)$ is an element of $L(V, W)$, i.e., a continuous linear map from V to W. This does **not** mean that $Df(v_0)$ depends continuously on $v_0 \in V$. (See the definition of $C^1(V, W)$ below.)

Example 3.5.41. Let $T \in L(V, W)$ and let $b \in V$. The derivative of the affine map $f : V \to W$, $f(v) = T(v) + b$ is $Df(v_0) = T$.

Theorem 3.5.42 (Chain Rule). *Let V, W, Z be Banach spaces, let $\Omega \subset V$ and $\Omega' \subset W$ be open sets, let $f : \Omega \to W$ and $g : \Omega' \to Z$ be continuous functions and let $v_0 \in \Omega$ and $w_0 = f(v_0) \in \Omega'$. Assume that f is differentiable at v_0 and that g is differentiable at w_0. Then $g \circ f$ is differentiable at v_0 and*

$$D(g \circ f)(v_0) = Dg(w_0)Df(v_0).$$

(Note that this is an element of $L(V, Z)$.)

Proof. By preliminary translations, we may assume that $v_0 = 0$, $w_0 = f(v_0) = 0$, and $g(w_0) = 0$. For the sake of simplicity of notation, write $T = Df(0) \in L(V, W)$ and $S = Dg(0) \in L(W, Z)$. Define the error functions

$$\epsilon_f(v) := f(v) - T(v)$$

and

$$\epsilon_g(w) := g(w) - S(w).$$

Note that $\epsilon_f(v)/||v|| \to 0$ as $v \to 0$ and $\epsilon_g(w)/||w|| \to 0$ as $w \to 0$ by the above definition of differentiability. For $v \in V$ define $w_1 := f(v) = T(v) + \epsilon_f(v)$. Notice that $\epsilon_g(w_1) = 0$ if $w_1 = 0$, and also that $w_1 \to 0$ as $v \to 0$. For $w_1 \neq 0$ we compute

$$\begin{aligned}(g \circ f)(v) - (S \circ T)(v) &= g(w_1) - S(T(v)) \\ &= S(w_1) - S(T(v)) + \epsilon_g(w_1) \\ &= S(\epsilon_f(v)) + \epsilon_g(w_1).\end{aligned}$$

Therefore

$$\begin{aligned}\frac{||(g \circ f)(v) - (S \circ T)(v)||}{||v||} &\leq \frac{||S(\epsilon_f(v))|| + ||\epsilon_g(w_1)||}{||v||} \\ &\leq ||S|| \cdot \frac{||\epsilon_f(v)||}{||v||} + \left(||T|| + \frac{||\epsilon_f(v)||}{||v||}\right) \frac{||\epsilon_g(w_1)||}{||w_1||}.\end{aligned}$$

By preceding remarks, this converges to zero as $v \to 0$. ☐

Definition 3.5.43. A map $f : \Omega \to W$ is said to be *continuously differentiable* at $v_0 \in \Omega$ if Df is continuous at v_0, and *continuously differentiable* if it is continuously differentiable at all points of Ω. The class of such maps is denoted $C^1(\Omega, W)$.

What about higher derivatives?

Definition 3.5.44. A differentiable map $f : \Omega \to W$ is *twice differentiable* at $v_0 \in \Omega$ if $Df : \Omega \to L(V, W)$ is differentiable at v_0. We denote this derivative as $D^2 f(v_0)$. If $D^2 f$ is continuous at v_0 we say that f is *twice continuously differentiable* at v_0, and we let $C^2(\Omega, W)$ be the space of all functions which are twice continuously differentiable at each point of Ω.

Note that $D^2 f(v_0) \in L(V, L(V, W))$. We pause for a quick algebraic lemma.

Lemma 3.5.45. *$L(V, L(V, W))$ is isometrically isomorphic with $L(V \times V, W)$, where $V \times V$ is equipped with the maximum (ℓ^∞) norm.*

Proof. Given $T \in L(V, L(V, W))$, define $\Phi(T) \in L(V \times V, W)$ to be

$$\Phi(T)(v_1, v_2) = T(v_1)(v_2).$$

This is clearly a continuous linear bijection. To see that it is an isometry, we compute the norm:

$$\begin{aligned}||T|| = \sup_{||v_1||=1} ||T(v_1)|| &= \sup_{||v_1||=1, ||v_2||=1} ||T(v_1)(v_2)|| \\ &= \sup_{||v_1||=1, ||v_2||=1} ||\Phi(T)(v_1, v_2)|| = ||\Phi(T)||.\end{aligned}$$

(The choice of norm on $V \times V$ is essential here.) ☐

Thus we can identify $D^2 f(v_0)$ as a continuous bilinear map from $V \times V$ to W. We will use the notations $D^2 f(v_0)(v_1)(v_2) = D^2 f(v_0)(v_1, v_2)$ interchangeably.

We delay the proof of the following fact until the next section.

Proposition 3.5.46. *$D^2f(v_0)$ is* **symmetric***: $D^2f(v_0)(v_1, v_2) = D^2f(v_0)(v_2, v_1)$ for all $v_1, v_2 \in V$.*

In a similar fashion, we define the kth order derivative $D^k f(v_0)$ for each $k \in \mathbb{N}$. It is a continuous multilinear map from V^k to W which is symmetric under the canonical action of the symmetric group on k letters on V^k. As always, we let $C^k(\Omega, W)$ denote the space of k times continuously differentiable maps from Ω to W. Finally,

$$C^\infty(\Omega, W) = \bigcap_{k \in \mathbb{N}} C^k(\Omega, W).$$

3.6 Taylor formulas for maps in $C^k(\Omega, W)$

In this section we consider the Taylor formula for maps in $C^k(\Omega, W)$, where Ω is an open set in a Banach space V and W is another Banach space. We'll come to this formula in three stages: first, in the special case when $W = \mathbb{R}$, next, in the special case when $V = \mathbb{R}$ and $\Omega \subset \mathbb{R}$ is an open interval, and finally, in the general case.

Theorem 3.6.47. *(a) Let $f \in C^{k+1}(\Omega, \mathbb{R})$, let $v_0 \in \Omega$, and let $B(v_0, r) \subset \Omega$. Then for all $v \in V$ with $||v|| < r$ we have*

$$f(v_0 + v) = \sum_{j=0}^{k} \frac{1}{j!} D^j f(v_0)(\underbrace{v, \ldots, v}_{j}) + \frac{1}{(k+1)!} D^{k+1} f(v_0 + \theta v)(\underbrace{v, \ldots, v}_{k+1})$$

for some $0 < \theta < 1$.

(b) Let $f \in C^k(\Omega, \mathbb{R})$, let $v_0 \in \Omega$, and let $B(v_0, r) \subset \Omega$. Then for all $v \in V$ with $||v|| < r$ we have

$$f(v_0 + v) = \sum_{j=0}^{k} \frac{1}{j!} D^j f(v_0)(\underbrace{v, \ldots, v}_{j}) + o(||v||^k),$$

where $o(||v||^k)$ denotes a term which tends to zero faster than $||v||^k$ as $v \to 0$.

Lemma 3.6.48. *Let $f : \Omega \to \mathbb{R}$ be n times differentiable, let $v_0 \in \Omega$, and let $v \in V$ be such that $[v_0, v_0 + v] = \{v_0 + tv : 0 \leq t \leq 1\} \subset \Omega$. Define $g : [0, 1] \to \mathbb{R}$ by $g(t) = f(v_0 + tv)$. Then for each $j = 1, \ldots, n$, the map g is j times differentiable at each point in $[0, 1]$, with*

$$g^{(j)}(t) = D^j f(v_0 + tv)(\underbrace{v, \ldots, v}_{j}), \quad 0 \leq t \leq 1.$$

Proof. By induction. The base case $j = 1$ is covered by the chain rule:

$$g'(t) = Df(v_0 + tv)(v).$$

We now assume that the result holds for some index $j < n$, i.e.,

$$g^{(j)}(t) = D^j f(v_0 + tv)(\underbrace{v, \ldots, v}_{j}) = \Xi_v^j(D^j f(v_0 + tv)),$$

where $\Xi_v^j \in (V^j)^{**}$ is the evaluation map $\Xi_v^j(T) = T(v, \ldots, v)$ (compare Proposition 3.4.34). Applying the chain rule again gives

$$g^{(j+1)}(t) = \Xi_v^j(D^{j+1} f(v_0 + tv)(v)) = D^{j+1} f(v_0 + tv)(\underbrace{v, \ldots, v}_{j+1}).$$

□

Proof of Theorem 3.6.47. To prove part (a), define $g(t) = f(v_0+tv)$ as in the lemma. Then g is $(k+1)$ times differentiable on $[0,1]$ with classical kth order Taylor series approximation

$$g(1) = \sum_{j=0}^{k} \frac{1}{j!} g^{(j)}(0) + \frac{1}{(k+1)!} g^{(k+1)}(\theta)$$

for some $0 < \theta < 1$. Using the formulas for $g^{(j)}(t)$ from the lemma finishes the proof of part (a).

Turning to part (b), we apply (a) with k replaced by $k-1$, obtaining

$$f(v_0+v) = \sum_{j=0}^{k-1} \frac{1}{j!} D^j f(v_0)(\underbrace{v, \ldots, v}_{j}) + \frac{1}{k!} D^k f(v_0+\theta v)(\underbrace{v, \ldots, v}_{k})$$

for some $0 < \theta < 1$. Thus

$$f(v_0+v) = \sum_{j=0}^{k} \frac{1}{j!} D^j f(v_0)(\underbrace{v, \ldots, v}_{j}) + \frac{1}{k!}(D^k f(v_0+\theta v) - D^k f(v_0))(\underbrace{v, \ldots, v}_{k}).$$

Since $f \in C^k(\Omega, \mathbb{R})$, $D^k f$ is continuous at v_0, so

$$\lim_{v \to 0} ||D^k f(v_0+\theta v) - D^k f(v_0)|| = 0$$

and

$$\begin{aligned} \left| \frac{1}{k!}(D^k f(v_0+\theta v) - D^k f(v_0))(\underbrace{v, \ldots, v}_{k}) \right| &\leq \frac{1}{k!} ||D^k f(v_0+\theta v) - D^k f(v_0)|| \cdot ||v||^k \\ &= o(||v||^k) \end{aligned}$$

as $v \to 0$. □

To consider Taylor type formulas with Banach space targets, we need to discuss the integration of Banach space valued functions. We use Riemann integrals.

Definition 3.6.49. Let $f : [a,b] \to W$ be a function taking values in a Banach space W. For each partition $\pi = \{a = t_0 < t_1 < \cdots < t_N = b\}$ and each vector $\zeta = (\zeta_1, \ldots, \zeta_N)$ with $\zeta_i \in [t_{i-1}, t_i]$, the associated *Riemann sum* is

$$R(f, \pi, \zeta) = \sum_{i=1}^{N} (t_i - t_{i-1}) f(\zeta_i),$$

where the indicated operations are scalar multiplication and addition in W. Then the *Riemann integral* of f over $[a,b]$ is the limit

$$\int_a^b f(t)\, dt = \lim_{\mathrm{mesh}(\pi) \to 0} R(f, \pi, \zeta)$$

if it exists independent of the choice of ζ, where $\mathrm{mesh}(\pi) = \max_i |t_i - t_{i-1}|$.

Proposition 3.6.50. *If $f : [a,b] \to W$ is continuous, then $F(t) := \int_a^t f(s)\,ds$ exists for each $t \in [a,b]$. Furthermore, $F \in C^1((a,b),W)$ with $DF(t)(1) = f(t)$ for all $a < t < b$.*

Corollary 3.6.51. *If $f : [a,b] \to W$ is continuously differentiable, then $\int_a^b Df(t)(1)\,dt = f(b) - f(a)$.*

Proof. $\boxed{\text{PROOF GOES HERE}}$ □

Now we can state the Taylor formula for maps $f \in C^{k+1}((a,b),W)$. For such a map, and for $0 \le j \le k+1$ and $a < t < b$, we write $f^{(j)}(t) = D^j f(t)(1)$.

Theorem 3.6.52. *For $f \in C^{k+1}((a,b),W)$ and $a < t_0 < b$, we have*

$$f(t_0 + t) = \sum_{j=0}^{k} \frac{1}{j!} f^{(j)}(t_0) t^j + R_k(t_0, t),$$

for t so small that $a < t_0 + t < b$, where

$$R_k(t_0, t) = \frac{1}{k!} \int_0^t f^{(k+1)}(t_0 + s)(t - s)^k\,ds.$$

Proof. Again, we induct on k. For $k = 0$ the statement reads

$$f(t_0 + t) = f(t_0) + \int_0^t f'(t_0 + s)\,ds = f(t_0) + \int_0^t Df(t_0 + s)(1)\,ds$$

which follows from Corollary 3.6.51. Assuming the statement for some $k-1$, we prove it for k. It suffices to show that

$$R_{k-1}(t_0, t) = \frac{1}{k!} f^{(k)}(t_0) t^k + R_k(t_0, t).$$

We prove this by integration by parts:

$$\begin{aligned} R_{k-1}(t_0, t) &= \frac{1}{(k-1)!} \int_0^t f^{(k)}(t_0 + s)(t - s)^{k-1}\,ds \\ &= \frac{1}{(k-1)!} \left(-f^{(k)}(t_0 + s) \frac{(t-s)^k}{k} \Big|_0^t + \int_0^t f^{(k+1)}(t_0 + s) \frac{(t-s)^k}{k}\,ds \right) \\ &= \frac{1}{k!} f^{(k)}(t_0) t^k + \frac{1}{k!} \int_0^t f^{(k+1)}(t_0 + s)(t - s)^k\,ds. \end{aligned}$$

□

Finally, we have a Taylor formula for maps in $C^{k+1}(\Omega, W)$ for $\Omega \subset V$ open and V, W arbitrary Banach spaces.

Theorem 3.6.53. *Let Ω be an open set in a Banach space V. For $f \in C^{k+1}(\Omega, W)$, $v_0 \in \Omega$, $B(v_0, r) \subset \Omega$, and $v \in V$ with $||v|| < r$, we have*

$$f(v_0 + v) = \sum_{j=0}^{k} \frac{1}{j!} D^j f(v_0)(\underbrace{v, \ldots, v}_{j}) + R_k(v_0, v),$$

where

$$||R_k(v_0, v)|| \leq \frac{1}{(k+1)!} \max_{0 \leq t \leq 1} ||D^{k+1} f(v_0 + tv)|| \, ||v||^{k+1}.$$

The relation between Theorems 3.6.53 and 3.6.52 is the same as the relationship between Theorem 3.6.47 and the classical one-variable Taylor formula (see the proof of Theorem 3.6.47).

Proof of Theorem 3.6.53. Define $g : [-1-\delta, 1+\delta] \to W$ by $g(t) = f(v_0 + tv)$, where δ is chosen so small that $(1+\delta)||v|| < r$. Then

$$g \in C^{k+1}((-1-\delta, 1+\delta), W)$$

and

$$g(1) = \sum_{j=0}^{k} \frac{1}{j!} g^{(j)}(0) + R_k(0, 1).$$

An easy computation gives

$$g^{(j)}(t) = D^k(v_0 + tv)(v, \ldots, v)$$

so

$$\begin{aligned} ||R_k(0,1)|| &= \frac{1}{k!} || \int_0^1 g^{(k+1)}(s)(1-s)^k \, ds || \\ &= \frac{1}{k!} || \int_0^1 D^{k+1} f(v_0 + sv)(v, \ldots, v)(1-s)^k \, ds || \\ &\leq \frac{1}{k!} \max_{0 \leq t \leq 1} ||D^{k+1} f(v_0 + tv)|| \, ||v||^{k+1} \int_0^1 (1-s)^k \, ds \\ &= \frac{1}{(k+1)!} \max_{0 \leq t \leq 1} ||D^{k+1} f(v_0 + tv)|| \, ||v||^{k+1}. \end{aligned}$$

□

3.7 Inverse and Implicit Function Theorems

in Banach spaces via FPT

3.8 Rectifiable curves in Banach spaces

recall: maps from $[a, b]$ to V, integration, FTC

curves, rectifiability, jost Lemma 11.5, arc length function, C1 implies rectifiable and length = integral of differential, jost Thm 11.16 (arc length param), line integrals

Remark: everything works in (complete) metric spaces also

3.9 Further embedding theorems

separable X embeds isometrically in $C([0, 1])$? (needs Banach-Alaoglu)

separable X embeds 12-bi-Lipschitz in c_0 (Aharoni)

Urysohn?