1. Lecture 1

☠ Warning: ☠ These notes haven't been checked nearly as thoroughly as they should have been, and as such are likely to be full of errors. I don't remember if I fixed every mistake that was found during the lecture. These imperfect notes are only being provided so that you can recall and keep track of what was discussed in the lecture. Watch out for mistakes, and if something doesn't convince you and you find it hard to figure out, look up or ask me.

Let $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. Let $U \subset V$ be an open set, and consider a function

$$f : U \to W.$$

We say that f is differentiable at $x \in U$ if, informally, $f(x+h)$ can be approximated as $f(x) + L(h)$, for a linear map L. More precisely:

Definition 1.1. We say that f is differentiable at $x \in U$ if there exists a linear map $L : V \to W$ such that:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|_2} = 0, \tag{1}$$

where $\|h\|_2 = (h_1^2 + \cdots + h_n^2)^{1/2}$ is the length of $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$. When such an L exists, we will denote L by $Df|_x$, and call it the derivative of f at x. However, note that this definition of $Df|_x$ needs some checking to ensure that L is well-defined, i.e., there are not two different possibilities for L. This will be done in Lemma 1.6 below.

Notation 1.2. More generally, if $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$, then we set $\|h\|_p = (|h_1|^p + \cdots + |h_n|^p)^{1/p}$ for each $p \geq 1$, and also $\|h\|_\infty = \max(|h_1|, \ldots, |h_n|)$.

Exercise 1.3. Show that for $p \geq 1$ and $h \in \mathbb{R}^n$:

$$\|h\|_\infty \leq \|h\|_p \leq n^{1/p} \|h\|_\infty.$$

Definition 1.4. A normed vector space over $\mathbb{R}$ is a pair $(V, \|\cdot\|)$ consisting of a vector space V over $\mathbb{R}$ together with a norm $\|\cdot\| : V \to \mathbb{R}$, by which we mean a function $\|\cdot\| : V \to \mathbb{R}$ (where we will write $\|x\|$ instead of $\|\cdot\|(x)$) such that:

(i) $\|x\| \geq 0$ for all $x \in V$, and $\|x\| = 0$ if and only if $x = 0$;
(ii) For all $c \in \mathbb{R}$ and $x \in V$, $\|cx\| = |c|\|x\|$; and
(iii) For all $x, y \in \mathbb{R}$, $\|x + y\| \leq \|x\| + \|y\|$ (i.e., subadditivity).

It is easy to see from (i) and (iii) above that the norm $\|\cdot\|$ defines a metric on V, by $d(x, y) = \|x - y\|$, and hence also a topology on V.

Note that Definition 1.1 can be generalized as follows to normed vector spaces:

Definition 1.5. Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two finite dimensional normed vector spaces over $\mathbb{R}$, let $U \subset V$ be an open set, and $f : U \to W$ a function. We say that f is differentiable at $x \in U$ (with respect to the norms $\|\cdot\|_V$ on V and $\|\cdot\|_W$ on W) if there exists a linear transformation $L \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ such that:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|_V} = 0 \tag{2}$$

in W, where the limit is made sense of using the metric on W defined by $\|\cdot\|_W$. When f is differentiable at x, we will denote L by $Df|_x$, and call it the derivative of f at x. By Lemma 1.6 below, $Df|_x$ is well-defined.

Lemma 1.6. *The derivative $Df|_x$ in Definition 1.1 is well-defined, i.e., if $L_1, L_2 \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ satisfy:*

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - L_1 h}{\|h\|_V} = 0 = \lim_{h \to 0} \frac{f(x+h) - f(x) - L_2 h}{\|h\|_V},$$

then $L_1 = L_2$.

Proof. Subtracting, we get:

$$\lim_{h\to 0} \frac{(L_1 - L_2)(h)}{\|h\|_V} = 0.$$

Let us fix $0 \neq v \in V$, and show that $(L_1 - L_2)(v) = 0$. Indeed, letting $h = tv$ and taking the limit as $t \to 0+$, we get:

$$0 = \lim_{t\to 0+} \frac{(L_1 - L_2)(tv)}{\|tv\|_V} = \lim_{t\to 0+} \frac{t \cdot (L_1 - L_2)(v)}{t\|v\|_V} = \frac{(L_1 - L_2)(v)}{\|v\|_V} \lim_{t\to 0+} \frac{t}{t} = \frac{(L_1 - L_2)(v)}{\|v\|_V},$$

forcing $(L_1 - L_2)(v) = 0$. □

One problem with the above definition is that it depends, a priori, on the choice of the norms $\|\cdot\|_V$ and $\|\cdot\|_W$. However, we will only focus on finite dimensional vector spaces, and show that the choice of the norm is immaterial.

Definition 1.7. Let V be a vector space over $\mathbb{R}$, and let $\|\cdot\|, \|\cdot\|'$ be two norms on V. We say that $\|\cdot\|$ and $\|\cdot\|'$ are equivalent if they are bounded above by positive multiples of each other, i.e., if there exist constants $c_1, c_2 > 0$ such that for all $v \in V$:

$$c_1\|v\| \leq \|v\|' \leq c_2\|v\|.$$

Example 1.8. By Exercise 1.3, for any n, the norms $\|\cdot\|_p$ on $\mathbb{R}^n$, $p \geq 1$, are all equivalent.

To show that the notion of differentiability or derivative defined in Definition 1.5 is independent of the choice of norms involved in the context of finite dimensional vector spaces, we will show:

(i) The notion of differentiability or derivative in Definition 1.5 does not change if we replace the norms on V and W by equivalent norms; and
(ii) If V is a finite dimensional vector space over $\mathbb{R}$, then all norms on V are equivalent.

Exercise 1.9. (i) Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed linear spaces, and give V and W the topologies defined by $\|\cdot\|_V$ and $\|\cdot\|_W$, respectively. Let $T : V \to W$ be an $\mathbb{R}$-linear map. Then the following are equivalent:
- (a) T is continuous on V.
- (b) T is continuous at $0 \in V$.
- (c) There exists $M \in \mathbb{R}$ such that for all $v \in V$, $\|Tv\|_W \leq M\|v\|_V$.

Hint: (a) ⇒ (b) is easy. To get (b) ⇒ (c), take "$M = \epsilon/\delta$". (c) ⇒ (a) is easy: take "$\delta = \epsilon/M$".
Aside: When these equivalent conditions are satisfied, one can consider the $\|Tv\|_W \leq M\|v\|_V$ for all $v \in V$; this infimum is sometimes denoted by $\|T\|$ (note that $\|T\|$ depends on the choice of $\|\cdot\|_V$ and $\|\cdot\|_W$, but we are suppressing these dependences from notation). It is easy to see that sending $T \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ to $\|T\|$ is a norm on $\mathrm{Hom}_{\mathbb{R}}(V, W)$. In other words, we have obtained a norm on $\mathrm{Hom}_{\mathbb{R}}(V, W)$ starting from our norms $\|\cdot\|_V$ on V and $\|\cdot\|_W$ on W; this norm is called the operator norm on $\mathrm{Hom}_{\mathbb{R}}(V, W)$ determined by the norms $\|\cdot\|_V$ on V and $\|\cdot\|_W$ on W.

(ii) Two norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space V are equivalent if and only if they define the same topology on V, i.e., if and only if the identity map $V \to V$ is a homeomorphism, where the source V (i.e., the first copy of V) is given the topology defined by $\|\cdot\|$ and the target V (i.e., the second copy of V) is given the topology defined by $\|\cdot\|'$.
Hint: This is more or less immediate from (i).

(iii) Suppose $\|\cdot\|_V, \|\cdot\|'_V$ are two equivalent norms on a vector space V over $\mathbb{R}$, and $\|\cdot\|_W$ and $\|\cdot\|'_W$ are two equivalent norms on a vector space W over $\mathbb{R}$. Let $U \subset V$ be a set that is open with respect to $\|\cdot\|_V$, and hence by ((ii)) above also open with respect to $\|\cdot\|'_V$. Let $f : U \to V$ be a function, and $x \in V$. Then f is differentiable at x with respect to $\|\cdot\|_V$ and $\|\cdot\|_W$ if and only it is differentiable at x with respect to $\|\cdot\|'_V$ and $\|\cdot\|'_W$. Moreover, when this condition is satisfied, the derivative $Df|_x \in \mathrm{Hom}_{\mathbb{R}}(V, W)$, as defined using $\|\cdot\|_V$ and $\|\cdot\|_W$, equals $Df|_x \in \mathrm{Hom}_{\mathbb{R}}(V, W)$, as defined using $\|\cdot\|'_V$ and $\|\cdot\|'_W$.

Proposition 1.10. *Let V be a finite dimensional vector space over $\mathbb{R}$. Then any two norms $\|\cdot\|_V$ and $\|\cdot\|'_V$ on V are equivalent.*

Proof. Choosing a basis for V (by which we mean an ordered basis), we get an $\mathbb{R}$-linear isomorphism $T : \mathbb{R}^n \to V$, for some n. Using this isomorphism, we can transfer $\|\cdot\|_V$ and $\|\cdot\|'_V$ to $\mathbb{R}^n$, by $x \mapsto \|Tx\|_V$ and $x \mapsto \|Tx\|_{V'}$, respectively. It is enough to show that the two norms on $\mathbb{R}^n$ we get this way are equivalent.

In other words, it is enough to show that any two norms on $\mathbb{R}^n$ are equivalent. In other words, we may assume $V = \mathbb{R}^n$. It is enough then to show that any norm $\|\cdot\|$ on $V = \mathbb{R}^n$ is equivalent to the norm $\|\cdot\|_\infty$, where $\|\cdot\|_\infty$ is as in Notation 1.2, i.e., $\|\cdot\|_\infty(a_1,\dots,a_n) = \max(|a_1|,\dots,|a_n|)$.

One inequality is easy: if $e_1,\dots,e_n$ is the standard basis for $\mathbb{R}^n$ and c_2 is the maximum of the $\|e_i\|$, then

$$(3)\quad \|(a_1,\dots,a_n)\| = \|a_1e_1 + \cdots + a_ne_n\| \le |a_1|\|e_1\| + \cdots + |a_n|\|e_n\| \le nc_2 \max_{1\le i\le n} |a_i| = nc_2\|(a_1,\dots,a_n)\|_\infty$$

(thanks to one of the audience members for correcting me at this point).

We give $\mathbb{R}^n$ the usual topology, defined by the norm $\|\cdot\|_2$ or equivalently by $\|\cdot\|_\infty$. Now it is easy to see from (3) that $\|\cdot\|$ is continuous on $\mathbb{R}^n$ (e.g., $(\mathbb{R}^n, \|\cdot\|_\infty) \to (\mathbb{R}^n, \|\cdot\|)$ is continuous by (3) and Exercise 1.9(i), now compose with $\|\cdot\|$, which is continuous as a map on $(\mathbb{R}^n, \|\cdot\|)$).

Note that $S := \{v \in \mathbb{R}^n \mid \|v\|_\infty = 1\}$ is a compact subset not containing 0. Therefore, $\|\cdot\|$ is a nonzero continuous function on this compact set S, and hence takes a minimum value on S, say c_1^{-1} for some $c_1 \in \mathbb{R}$. Thus, whenever $\|v\|_\infty = 1$, we have $\|v\| \ge c_1^{-1}$, so that $c_1\|v\|_\infty \le \|v\|$ for all $v \in \mathbb{R}^n$. Combining with (3), the equivalence of $\|\cdot\|_\infty$ and $\|\cdot\|$ follows, and hence so does the proposition. □

Note that both the inequalities in the proof of the above proposition used the finite dimensionality of V; the latter inequality used the compactness of a unit sphere, which is a characteristic of finite dimensional normed linear spaces. Thanks to the above proposition, we can now canonically define a topology on any finite dimensional vector space over $\mathbb{R}$:

Notation 1.11. Let V be a finite dimensional vector space over $\mathbb{R}$. Then we will henceforth use on V the topology defined by any choice of a norm on V — the choice of the norm does not matter, i.e., different norms will give the same topology — by Proposition 1.10 and the fact that equivalent norms induce the same topology (Exercise 1.9(ii)).

Remark 1.12. Here is one way to discuss the above topology on a finite dimensional vector space $V/\mathbb{R}$: choose *any* linear isomorphism $V \to \mathbb{R}^n$ using a basis of V, and transfer the usual topology on $\mathbb{R}^n$ to V via this isomorphism. Since the usual topology on $\mathbb{R}^n$ is defined by the norm $\|\cdot\|_2$, this topology on V is is defined by the pull-back of $\|\cdot\|_2$ to V via $V \to \mathbb{R}^n$, i.e.: $V \to \mathbb{R}^n \xrightarrow{\|\cdot\|_2} \mathbb{R}$. Therefore, this topology coincides with that discussed in Notation 1.11, independently of the choice of the isomorphism $V \to \mathbb{R}^n$.

Exercise 1.13. If V and W are finite dimensional vector spaces topologized as above, it is easy to see that any linear map $V \to W$ is automatically continuous.
Hint: One can write the map as a surjection followed by an injection; the surjection can be thought of as a projection on choosing suitable bases, while the injection can be thought of as an embedding $\mathbb{R}^n \hookrightarrow \mathbb{R}^m$ of the form $(x_1,\dots,x_m) \mapsto (x_1,\dots,x_m,0,\dots,0)$.

Now, restricting to the finite dimensional case, which is what we will focus on, we can rephrase Definition 1.5 as:

Definition 1.14. Let V, W be two finite dimensional vector spaces over $\mathbb{R}$, let $U \subset V$ be an open set (Notation 1.11), and $f : U \to W$ a function. We say that f is differentiable at $x \in U$ if there exists a linear transformation $L \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ such that:

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|_V} = 0 \tag{4}$$

in W, where $\|\cdot\|_V$ is any choice of a norm on V, and where the limit is made sense of using the metric on W defined by any choice of a norm on $\|\cdot\|_W$. When f is differentiable at x, we will denote L by $Df|_x$, and call it the derivative of f at x. By Lemma 1.6, $Df|_x$ is well-defined.

The choices of the norms in the above definition do not matter, since we are using the fact that replacing norms on V and W by equivalent norms does not change the notion of differentiability or the derivative, and combining it with Proposition 1.10.

Notation 1.15. Henceforth, we will often consider situations as in Definition 1.14: V, W will be finite dimensional vector spaces over $\mathbb{R}$, $U \subset V$ will be an open set, and $f : U \to W$ a function. We might just write $V \supset U \xrightarrow{f} W$ to denote such a situation. In such a situation, we will say that f is differentiable on U, or just differentiable if U is clear from the context, if it is differentiable at x for all $x \in U$. If f is differentiable on U, i.e., if $Df|_x$ exists for all $x \in U$, we will denote by Df the map

$$U \to \mathrm{Hom}_{\mathbb{R}}(V, W)$$

given by $x \mapsto Df|_x$. Note that we write the value of Df at the point x as $Df|_x$, and not as $Df(x)$. Thus, while f takes values in W, Df takes values in $\mathrm{Hom}_{\mathbb{R}}(V, W)$.

Notation 1.16. Let $V \supset U \xrightarrow{f} W$ be as above. If f is differentiable, we have $Df : U \to \mathrm{Hom}_{\mathbb{R}}(V, W)$. Since $\mathrm{Hom}_{\mathbb{R}}(V, W)$ is a vector space, we can now ask if $Df : U \to \mathrm{Hom}_{\mathbb{R}}(V, W)$ is differentiable. If it is, we can repeat this procedure and define the second derivative $D^2 f = D(Df) : U \to \mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, W))$. Thus, we can define what it means for f to be n times differentiable; if it is, we will have:

$$D^n f : U \to \mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, \ldots \mathrm{Hom}_{\mathbb{R}}(V, W) \ldots),$$

where $\mathrm{Hom}_{\mathbb{R}}(V, -)$ is applied n times.

Notation 1.17.
- (i) Let V_1, V_2, W be vector spaces. A map $V_1 \times V_2 \to W$ is said to be bilinear if it is linear in each variable: $B(cv_1 + c'v_1', v_2) = cB(v_1, v_2) + c'B(v_1', v_2)$, and $B(v_1, cv_2 + c'v_2') = cB(v_1, v_2) + c'B(v_1, v_2')$.
- (ii) Let $V_1, \ldots, V_n, W$ be vector spaces. A map $V_1 \times \cdots \times V_n \to W$ is said to be multilinear if it is linear in each variable.
- (iii) Note that the set $V_1 \times \cdots \times V_n$ is also the set underlying the vector space $V_1 \oplus \cdots \oplus V_n$. Note that we write a multilinear map as $V_1 \times \cdots \times V_n \to W$, and not as $V_1 \oplus \cdots \oplus V_n \to W$, because multilinear maps are not vector space maps (i.e., linear maps). However, when we talk of the derivative $Df|_x$ of a function $f : V_1 \times \cdots \times V_n \to W$, we will write such a map as $Df|_x : V_1 \oplus \cdots \oplus V_n \to W$, because this time this map is linear.

Remark 1.18. Let $V \supset U \xrightarrow{f} W$ as above, and assume that f is n times differentiable on U, so we have

$$D^n f : U \to \mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, \ldots \mathrm{Hom}_{\mathbb{R}}(V, W) \ldots)),$$

where $\mathrm{Hom}_{\mathbb{R}}(V, -)$ is applied n times. How to think of this? We will now give a slightly different way of describing the kind of object that $D^n f$ is, whose significance will hopefully become more clear in the next lecture.

It is easy to check that we have a (with the right structures, linear) bijection

$$\mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, W)) \to \mathrm{Maps}_{\mathrm{bilinear}}(V \times V \to W),$$

mapping

$$T \in \mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, W)) \quad \textit{to} \quad (v_1, v_2) \mapsto (T(v_1))(v_2).$$

It is not hard to check that a similar consideration applies to $\mathrm{Hom}_{\mathbb{R}}(V, \mathrm{Hom}_{\mathbb{R}}(V, \dots \mathrm{Hom}_{\mathbb{R}}(V, W) \dots))$, and thus, we can think of $D^n f$ as a map

$$U \to \mathrm{Maps}_{\text{multilinear}}(\underbrace{V \times V \times \cdots \times V}_{n \text{ times}} \to W).$$

Example 1.19. (i) Consider the case where $V = W = \mathbb{R}$, let $\mathbb{R} \supset U \xrightarrow{f} \mathbb{R}$, and let $x \in U$. Suppose f is differentiable at x in the usual one-variable sense, with derivative $f'(x)$. Then in the equality:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - f'(x)h}{h} = 0$$

the same limit with the h in the denominator replaced by $|h|$ is also 0, so that $Df|_x \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}, \mathbb{R})$ exists and equals multiplication by $f'(x)$. Conversely, if $Df|_x \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}, \mathbb{R})$ exists, so that $Df|_x$ equals multiplication by a for some a, it is easy to see from the definitions that f is differentiable in the usual one-variable sense, and that $f'(x) = a$.

(ii) Consider the case where $V = W = \mathbb{C}$, let $\mathbb{C} \supset U \xrightarrow{f} \mathbb{C}$, and let $z_0 \in U$. Suppose f is holomorphic (i.e., analytic) at z_0 in the usual sense, with derivative $f'(z_0)$. Recall that this means:

$$\lim_{h \to 0} \frac{f(z_0+h) - f(z_0)}{h}$$

exists, where the limit "$h \to 0$" is made sense of in $\mathbb{C}$ and we note that the denominator has h rather than $|h|$, and that this limit equals $f'(z_0)$. From this it is easy to see that:

$$\lim_{h \to 0} \frac{f(z_0+h) - f(z_0) - f'(z_0)h}{|h|} = 0.$$

Note that the absolute value $|\cdot|$ is a norm on $\mathbb{C}$. Thus, it is clear that f is differentiable in the sense of Definition 1.14, and that $Df|_{z_0} \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$ equals multiplication by $f'(z_0)$ (and hence belongs to the subspace $\mathrm{Hom}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})$ of $\mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$).
But this time, the converse is not true: if f is differentiable at z_0, i.e., $Df|_{z_0} \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$ exists, it is not the case that f is holomorphic at z_0: for f to be holomorphic at z_0, we need not only that f is differentiable at z_0, but additionally that $Df|_{z_0} \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$ is given by multiplication by a complex number. Note that these are special elements of $\mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$, those whose bases with respect to $\{1, i\}$ is of the form:

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

The form of the above matrix suggests the Cauchy-Riemann equations, but we cannot discuss that unless we relate the derivative, as discussed above, to partial derivatives.

Lemma 1.20. *(i) Let $V \supset U \xrightarrow{f} W$, such that f is constant. Then f is differentiable on U, and $Df = 0$ as a function $U \to \mathrm{Hom}_{\mathbb{R}}(V, W)$.*

(ii) Suppose $f : V \to W$ is of the form $f(v) = c + Tv$, where $c \in W$ is a constant and $T \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ is a linear transformation. Then for all $x \in V$, f is differentiable at x, and $Df|_x = T$.

(iii) Suppose $f = B : V_1 \times V_2 \to W$ is a bilinear map. Then, for all $x = (x_1, x_2) \in V_1 \times V_2$ (which is the set underlying $V_1 \oplus V_2$), f is differentiable at x, and $Df|_x : V_1 \oplus V_2 \to W$ is given by:

$$Df|_x(v_1, v_2) = B(x_1, v_2) + B(v_1, x_2).$$

(iv) Suppose $f = M : V_1 \times \cdots \times V_n \to W$ is a multilinear map. Then for all $x = (x_1, \dots, x_n) \in V_1 \times \cdots \times V_n$, f is differentiable at x, and $Df|_x : V_1 \oplus \cdots \oplus V_n \to W$ is given by:

$$Df|_x(v_1, \dots, v_n) = \sum_{i=1}^{n} B(x_1, \dots, x_{i-1}, v_i, x_{i+1}, \dots, x_n).$$

Proof. (i) is easy. (ii) follows from:

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-Th}{\|h\|}=\lim_{h\to 0}\frac{c+T(x+h)-(c+Tx)-Th}{\|h\|}=\lim_{h\to 0}0=0,$$

where $\|\cdot\|$ is any norm on V.

Now let us prove (iii). We $\|\cdot\|$ stand for some chosen norms on V_1, V_2 and W, also for the norm on $V_1 \oplus V_2$ given by $\|(v_1, v_2)\| = \|v_1\| + \|v_2\|$, and again also stand for any norm on W. The following property of the bilinear map will help us: there exists $M > 0$ such that for all $v_1 \in V_1, v_2 \in V_2$:

$$\|B(v_1,v_2)\| \leq M\|v_1\|\|v_2\|. \tag{5}$$

We will only sketch the proof of this relation — we leave it as an easy exercise to fill in the details. First, one proves that the map $B : V_1 \times V_2 \to W$ is continuous: the idea is that if we identify $V_1 \cong \mathbb{R}^{n_1}, V_2 \cong \mathbb{R}^{n_2}$ and $W \cong \mathbb{R}^m$ using choices of bases, then a bilinear map $V_1 \times V_2 \to W$ identifies with a map $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^m$ which is coordinate-wise given by m homogeneous polynomials of degree 2 in $n_1 + n_2$ variables. Second, one proves that the 'unit balls' with respect to $\|\cdot\|$ on V_1 and V_2 are compact (since all norms are equivalent, this unit ball is easily seen to be closed and bounded), and one takes M to be the supremum of the $\|B(v_1, v_2)\|$ as v_1, v_2 range over the unit balls of V_1 and V_2, respectively. Thus, given general $v_1 \in V_1$ and $v_2 \in V_2$, we can write $v_1 = \|v_1\|v_1', v_2 = \|v_2\|v_2'$ with $\|v_1'\| = \|v_2'\| = 1$, and we have:

$$\|B(v_1,v_2)\| = \|B(\|v_1\|v_1', \|v_2\|v_2')\| = \|v_1\|\|v_2\|B(v_1',v_2') \leq M\|v_1\|\|v_2\},$$

since v_1', v_2' belong to the unit balls of V_1 and V_2, respectively, with respect to the chosen norms $\|\cdot\|$. This proves (5).

Recall that the derivative $Df|_x$ of $f : V_1 \times V_2 \to W$ at x is be written as $Df|_x : V_1 \oplus V_2 \to W$. In this proof, given $h \in V_1 \oplus V_2$, we will write it as (h_1, h_2) with $h_1 \in V_1, h_2 \in V_2$. We have:

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-(B(x_1,h_2)+B(h_1,x_2))}{\|h\|}=\lim_{h\to 0}\frac{B(x_1+h_1,x_2+h_2)-B(x_1,x_2)-B(x_1,h_2)-B(h_1,x_2)}{\|h\|},$$

which equals $\lim_{h\to 0}\frac{B(h_1,h_2)}{\|h\|}$.

This limit equals 0 since, by (5), we have:

$$\frac{\|B(h_1,h_2)\|}{\|h\|} \leq \frac{M\|h_1\|\|h_2\|}{\|h\|} \leq \frac{M\|h\|^2}{\|h\|} = M\|h\|,$$

which goes to 0 as $h \to 0$ in $V_1 \oplus V_2$.

The proof of (iv) is somewhat similar, we will leave it as an exercise. □

Lemma 1.21. *Let $V \supset U \xrightarrow{f} W$ and $x \in U$ be as above. Suppose $\lambda : W \to W'$ is a linear map. If f is differentiable at x, then so is $\lambda \circ f : U \to W'$, and we have $D(\lambda \circ f)|_x = (\lambda \circ Df|_x)$, where the right-hand side is the composite of $Df|_x : V \to W$ with $\lambda : W \to W'$ to get a map $V \to W'$.*

Proof. Easy. Use Exercise (1.13) to handle the limits. □

Lemma 1.22. *Let $V \supset U \xrightarrow{f} W$ be as above, with V, W finite dimensional vector spaces over $\mathbb{R}$ (and $U \subset V$ open). Assume that $W = W_1 \oplus \cdots \oplus W_n$, and write $f(v) = (f_1(v), \ldots, f_n(v))$, with $f_i(v) \in W_i$ for each i. Then f is differentiable at x if and only if each f_i is differentiable at x, and when this is so,*

$$Df|_x \in \mathrm{Hom}_{\mathbb{R}}(V,W) = \mathrm{Hom}_{\mathbb{R}}(V, W_1\oplus\cdots\oplus W_n) = \mathrm{Hom}_{\mathbb{R}}(V,W_1)\oplus\cdots\oplus\mathrm{Hom}_{\mathbb{R}}(V,W_n)$$

is given by the element

$$(Df_1|_x, \ldots, Df_n|_x) \in \mathrm{Hom}_{\mathbb{R}}(V,W_1)\oplus\cdots\oplus\mathrm{Hom}_{\mathbb{R}}(V,W_n).$$

Proof. We will prove one way implication, and the leave the other as an exercise. Suppose f is differentiable at x and $Df|_x = L \in \mathrm{Hom}_{\mathbb{R}}(V, W) = \mathrm{Hom}_{\mathbb{R}}(V, W_1) \oplus \cdots \oplus \mathrm{Hom}_{\mathbb{R}}(V, W_n)$. Note that L can then be thought of as $(\mathrm{pr}_1 \circ L, \ldots, \mathrm{pr}_n \circ L) \in \mathrm{Hom}_{\mathbb{R}}(V, W_1) \oplus \cdots \oplus \mathrm{Hom}_{\mathbb{R}}(V, W_n)$, where $\mathrm{pr}_i : W \to W_i$ is the projection with respect to the decomposition $W = W_1 \oplus \cdots \oplus W_n$.

Therefore, if f is differentiable at x, then the assertion that $f_1, \ldots, f_r$ are differentiable at x and that $Df|_x = (Df_1|_x, \ldots, Df_n|_x)$ follows from applying Lemma 1.21 with λ replaced by $\mathrm{pr}_1, \ldots, \mathrm{pr}_n$. Conversely, if each f_i is differentiable at x with derivative $L_i \in \mathrm{Hom}(V, W_i)$, so that:

$$\lim_{h \to 0} \frac{f_i(x+h) - f_i(x) - L_i h}{\|h\|} = 0,$$

then forming $L = \bigoplus_i L_i \in \mathrm{Hom}_{\mathbb{R}}(V, \bigoplus_i W_i) = \mathrm{Hom}_{\mathbb{R}}(V, W)$, we get:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|} = \lim_{h \to 0} \sum_{i=1}^{n} \frac{f_i(x+h) - f_i(x) - L_i h}{\|h\|} = 0,$$

as desired: what the first equality above uses is that the limit in W can be tested/computed by projecting onto each W_i. □

The above lemma shows that to study the derivatives of multivariable functions, we can often assume that the *codomain* W is one-dimensional: to check that $\mathbb{R}^n \supset U \overset{f=(f_1,\ldots,f_m)}{\to} \mathbb{R}^m$ is differentiable at some point, it comes down to checking that each $f_i : U \to \mathbb{R}$ is. But we usually cannot reduce to the case where V is one-dimensional.

Corollary 1.23. *Let $V \supset U \overset{f_1, f_2}{\to} W$ as above, and $x \in W$. Let $c_1, c_2 \in \mathbb{R}$. If f_1, f_2 are differentiable at x, then so is $c_1 f_1 + c_2 f_2$, and $D(c_1 f_1 + c_2 f_2)|_x = c_1 Df_1|_x + c_2 Df_2|_x$.*

Proof. Either check directly, or write $c_1 f_1 + c_2 f_2$ as the composite map:

$$U \overset{(f_1, f_2)}{\to} W \oplus W \overset{(w_1, w_2) \mapsto c_1 w_1 + c_2 w_2}{\to} W,$$

where the map on the left is differentiable at x by Lemma 1.22, with derivative at x given by $(Df_1|_x, Df_2|_x)$ using the notation of that lemma, so that the composite is also differentiable at x by Lemma 1.21, with derivative at x given by $c_1 Df_1|_x + c_2 Df_2|_x$. □