Script 19

Differentiation in $\mathbb{R}^n$

8/4: **Definition 19.1.** A **linear transformation** $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is a function such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$,

(a) $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$;

(b) $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$.

That is, φ is a linear transformation if it respects the two operations in Definition 18.2.

Lemma 19.2. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\varphi(\mathbf{0}) = \mathbf{0}$.*

Proof. Suppose for the sake of contradiction that $\varphi(\mathbf{0}) \neq \mathbf{0}$. Then

$$\begin{aligned} \mathbf{0} &= \varphi(\mathbf{x}) - \varphi(\mathbf{x}) \\ &= \varphi(\mathbf{x} - \mathbf{x}) && \text{Definition 19.1} \\ &= \varphi(\mathbf{0}) \\ &\neq \mathbf{0} \end{aligned}$$

a contradiction. $\square$

Exercise 19.3. We denote $\mathbf{x} \in \mathbb{R}^2$ by $\mathbf{x} = (x, y)$. Determine whether the following functions are linear transformations:

(a) $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $\varphi(x, y) = x + y$.

Answer. φ is a linear transformation. $\square$

Proof. To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and for any $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^2$, and let λ be an arbitrary element of $\mathbb{R}$. Then

$$\begin{aligned} \varphi(\mathbf{x} + \mathbf{y}) &= (x_1 + y_1) + (x_2 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2) \\ &= \varphi(\mathbf{x}) + \varphi(\mathbf{y}) \end{aligned}$$

and

$$\begin{aligned} \varphi(\lambda \mathbf{x}) &= \lambda x_1 + \lambda x_2 \\ &= \lambda(x_1 + x_2) \\ &= \lambda \varphi(\mathbf{x}) \end{aligned}$$

as desired. $\square$

(b) $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, $\varphi(x, y) = (x, y + 1)$.

Answer. φ is not a linear transformation. □

Proof. By the definition of φ, $\varphi(\mathbf{0}) = (0, 1) \neq \mathbf{0}$. Thus, by the contrapositive of Lemma 19.2, φ is not a linear transformation, as desired. □

(c) $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(x, y) = (3x - y, x + 2y, 0)$.

Answer. φ is a linear transformation. □

Proof. To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and for any $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda\mathbf{x}) = \lambda\varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^2$, and let λ be an arbitrary element of $\mathbb{R}$. Then

$$\begin{aligned}
\varphi(\mathbf{x} + \mathbf{y}) &= (3[x_1 + y_1] - [x_2 + y_2], [x_1 + y_1] + 2[x_2 + y_2], 0) \\
&= ([3x_1 - x_2] + [3y_1 - y_2], [x_1 + 2x_2] + [y_1 + 2y_1], 0) \\
&= (3x_1 - x_2, x_1 + 2x_2, 0) + (3y_1 - y_2, y_1 + 2y_2, 0) \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
\varphi(\lambda\mathbf{x}) &= (3[\lambda x_1] - [\lambda x_2], [\lambda x_1] + 2[\lambda x_2], 0) \\
&= (\lambda[3x_1 - x_2], \lambda[x_1 + 2x_2], 0) \\
&= \lambda(3x_1 - x_2, x_1 + 2x_2, 0) \\
&= \lambda\varphi(\mathbf{x})
\end{aligned}$$

as desired. □

(d) $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(x, y) = (x^2, x + y, x + y^3)$.

Answer. φ is not a linear transformation. □

Proof. Consider $(1, 1) \in \mathbb{R}^2$ and let $2 \in \mathbb{R}$. Then

$$\begin{aligned}
\varphi(2(1, 1)) &= (4, 4, 10) \\
&\neq (2, 4, 4) \\
&= 2(1, 2, 2) \\
&= 2\varphi(1, 1)
\end{aligned}$$

as desired. □

Exercise 19.4.

(a) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a linear transformation. What does the graph of φ look like?

Answer. A line through the origin with finite slope. □

(b) Let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be a linear transformation. What does the graph of φ look like?

Answer. A plane through the origin with finite slope in both directions. □

Exercise 19.5.

(a) Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ and $\psi : \mathbb{R}^m \to \mathbb{R}^\ell$ be linear transformations. Prove that $\psi \circ \varphi$ is also a linear transformation.

Proof. To prove that $\psi \circ \varphi : \mathbb{R}_n \to \mathbb{R}^\ell$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $(\psi \circ \varphi)(\mathbf{x}+\mathbf{y}) = (\psi \circ \varphi)(\mathbf{x}) + (\psi \circ \varphi)(\mathbf{y})$ and $(\psi \circ \varphi)(\lambda\mathbf{x}) = \lambda(\psi \circ \varphi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then since φ and ψ are linear transformations themselves, we have that

$$
\begin{aligned}
(\psi \circ \varphi)(\mathbf{x}+\mathbf{y}) &= \psi(\varphi(\mathbf{x}+\mathbf{y})) \\
&= \psi(\varphi(\mathbf{x}) + \varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= \psi(\varphi(\mathbf{x})) + \psi(\varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= (\psi \circ \varphi)(\mathbf{x}) + (\psi \circ \varphi)(\mathbf{y})
\end{aligned}
$$

and

$$
\begin{aligned}
(\psi \circ \varphi)(\lambda\mathbf{x}) &= \psi(\varphi(\lambda\mathbf{x})) \\
&= \psi(\lambda\varphi(\mathbf{x})) && \text{Definition 19.1} \\
&= \lambda\psi(\varphi(\mathbf{x})) && \text{Definition 19.1} \\
&= \lambda(\psi \circ \varphi)(\mathbf{x})
\end{aligned}
$$

as desired. □

(b) Let $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations and let $\lambda \in \mathbb{R}$. Prove that $\varphi + \psi$ and $\lambda\varphi$ are linear transformations.

Proof. To prove that $\varphi + \psi$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $(\varphi+\psi)(\mathbf{x}+\mathbf{y}) = (\varphi+\psi)(\mathbf{x}) + (\varphi+\psi)(\mathbf{y})$ and $(\varphi+\psi)(\lambda\mathbf{x}) = \lambda(\varphi+\psi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then since φ and ψ are linear transformations themselves, we have that

$$
\begin{aligned}
(\varphi+\psi)(\mathbf{x}+\mathbf{y}) &= \varphi(\mathbf{x}+\mathbf{y}) + \psi(\mathbf{x}+\mathbf{y}) \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y}) + \psi(\mathbf{x}) + \psi(\mathbf{y}) && \text{Definition 19.1} \\
&= \varphi(\mathbf{x}) + \psi(\mathbf{x}) + \varphi(\mathbf{y}) + \psi(\mathbf{y}) \\
&= (\varphi+\psi)(\mathbf{x}) + (\varphi+\psi)(\mathbf{y})
\end{aligned}
$$

and

$$
\begin{aligned}
(\varphi+\psi)(\lambda\mathbf{x}) &= \varphi(\lambda\mathbf{x}) + \psi(\lambda\mathbf{x}) \\
&= \lambda\varphi(\mathbf{x}) + \lambda\psi(\mathbf{x}) && \text{Definition 19.1} \\
&= \lambda(\varphi(\mathbf{x}) + \psi(\mathbf{x})) \\
&= \lambda(\varphi+\psi)(\mathbf{x})
\end{aligned}
$$

as desired.

To prove that $\lambda\varphi$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\gamma \in \mathbb{R}$, $(\lambda\varphi)(\mathbf{x}+\mathbf{y}) = (\lambda\varphi)(\mathbf{x}) + (\lambda\varphi)(\mathbf{y})$ and $(\lambda\varphi)(\gamma\mathbf{x}) = \gamma(\lambda\varphi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let γ be an arbitrary element of $\mathbb{R}$. Then since φ is a linear transformation itself, we have that

$$
\begin{aligned}
(\lambda\varphi)(\mathbf{x}+\mathbf{y}) &= \lambda\varphi(\mathbf{x}+\mathbf{y}) \\
&= \lambda(\varphi(\mathbf{x}) + \varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= \lambda\varphi(\mathbf{x}) + \lambda\varphi(\mathbf{y}) \\
&= (\lambda\varphi)(\mathbf{x}) + (\lambda\varphi)(\mathbf{y})
\end{aligned}
$$

and

$$\begin{aligned}(\lambda\varphi)(\gamma\mathbf{x}) &= \lambda\varphi(\gamma\mathbf{x}) \\ &= \lambda\gamma\varphi(\mathbf{x}) && \text{Definition 19.1} \\ &= \gamma\lambda\varphi(\mathbf{x}) \\ &= \gamma(\lambda\varphi)(\mathbf{x})\end{aligned}$$

as desired. □

(c) Let $\pi_I : \mathbb{R}^m \to \mathbb{R}^k$ be the projection function from Definition 18.34. Prove that π_I is a linear transformation.

Proof. To prove that π_I is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $\pi_I(\mathbf{x}+\mathbf{y}) = \pi_I(\mathbf{x}) + \pi_I(\mathbf{y})$ and $\pi_I(\lambda\mathbf{x}) = \lambda\pi_I(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then we have that

$$\begin{aligned}\pi_I(\mathbf{x}+\mathbf{y}) &= (x_{i_1} + y_{i_k}, \dots, x_{i_k} + y_{i_k}) \\ &= (x_{i_1}, \dots, x_{i_k}) + (y_{i_1}, \dots, y_{i_k}) \\ &= \pi_I(\mathbf{x}) + \pi_I(\mathbf{y})\end{aligned}$$

and

$$\begin{aligned}\pi_I(\lambda\mathbf{x}) &= (\lambda x_{i_1}, \dots, \lambda x_{i_k}) \\ &= \lambda(x_{i_1}, \dots, x_{i_k}) \\ &= \lambda\pi_I(\mathbf{x})\end{aligned}$$

as desired. □

Definition 19.6. The $\boldsymbol{j^{\text{th}}}$ **standard basis vector** in $\mathbb{R}^n$ is the vector $\mathbf{e}_j$ defined by

$$(\mathbf{e}_j)_k = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

For example, the standard basis vectors for $\mathbb{R}^3$ are $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, and $\mathbf{e}_3 = (0,0,1)$. Notice that if $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, then $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$.

Definition 19.7. For any linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, we denote by $[\varphi]_{ij}$ the i^{th} component of the vector $\varphi(\mathbf{e}_j)$; i.e., $[\varphi]_{ij} = \varphi_i(\mathbf{e}_j)$.

Exercise 19.8.

(a) Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let $\mathbf{x} \in \mathbb{R}^n$. Find a formula for $\varphi(\mathbf{x})$ in terms of $[\varphi]_{ij}$, the components of $\mathbf{x}$, and the standard basis vectors in $\mathbb{R}^m$.

Proof. Since $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e_n}$ by Definition 19.6 and since φ is linear, we have that

$$\begin{aligned}\varphi(\mathbf{x}) &= \varphi(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e_n}) \\ &= \varphi(x_1\mathbf{e}_1) + \cdots + \varphi(x_n\mathbf{e}_n) \\ &= x_1\varphi(\mathbf{e}_1) + \cdots + x_n\varphi(\mathbf{e}_n) \\ &= x_1(\varphi_1(\mathbf{e}_1)\mathbf{e_1} + \cdots + \varphi_m(\mathbf{e}_1)\mathbf{e_m}) + \cdots + x_n(\varphi_1(\mathbf{e}_n)\mathbf{e_1} + \cdots + \varphi_m(\mathbf{e}_n)\mathbf{e_m}) \\ &= x_1([\varphi]_{11}\mathbf{e_1} + \cdots + [\varphi]_{m1}\mathbf{e_m}) + \cdots + x_n([\varphi]_{1n}\mathbf{e_1} + \cdots + [\varphi]_{mn}\mathbf{e_m}) \\ &= x_1\sum_{i=1}^m [\varphi]_{i1}\mathbf{e}_i + \cdots + x_n\sum_{i=1}^m [\varphi]_{in}\mathbf{e}_i\end{aligned}$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} [\varphi]_{ij} \mathbf{e}_i$$
$$= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j [\varphi]_{ij} \mathbf{e}_i$$

□

(b) For $1 \leq i \leq m$ and $1 \leq j \leq n$, let $A_{ij} \in \mathbb{R}$. Prove that there is a unique linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ such that $[\varphi]_{ij} = A_{ij}$ for all i, j.

Proof. Let φ be defined by

$$\varphi(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i$$

for all $\mathbf{x} \in \mathbb{R}^n$. Thus, by Definition 19.7, $[\varphi]_{ij} = A_{ij}$ for all i, j.

To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$. Then

$$\begin{aligned}
\varphi(\mathbf{x} + \mathbf{y}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j + y_j) A_{ij} \mathbf{e}_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j A_{ij} \mathbf{e}_i + y_j A_{ij} \mathbf{e}_i) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i + \sum_{i=1}^{m} \sum_{j=1}^{n} y_j A_{ij} \mathbf{e}_i \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
\varphi(\lambda \mathbf{x}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (\lambda x_j) A_{ij} \mathbf{e}_i \\
&= \lambda \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i \\
&= \lambda \varphi(\mathbf{x})
\end{aligned}$$

as desired.

Let $\psi : \mathbb{R}^n \to \mathbb{R}^m$ be any linear transformation satisfying $[\psi]_{ij} = A_{ij}$ for all i, j. To prove that $\varphi = \psi$, it will suffice to show that $\varphi(\mathbf{x}) = \psi(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Let $\mathbf{x}$ be an arbitrary element of $\mathbb{R}^n$. Then

$$\begin{aligned}
\varphi(\mathbf{x}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j [\psi]_{ij} \mathbf{e}_i && \text{Exercise 19.8a} \\
&= \psi(\mathbf{x})
\end{aligned}$$

as desired. □

Definition 19.9. We define an $m \times n$ matrix M to be an array of scalars

$$M = \{a_{ij}\} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

So a_{ij} denotes the scalar in row i, column j of the matrix. For every linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, there is a corresponding $m \times n$ matrix $\{[\varphi]_{ij}\}$. We denote $\{[\varphi]_{ij}\}$ by $[\varphi]$. Also, by Exercise 19.8, given a matrix of scalars, there is a unique linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ that corresponds to it.

Exercise 19.10.

(a) Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $\varphi(x, y, z) = (3x + 2y - z, 4x - 5y + 2z)$. Write down the matrix $[\varphi]$.

Answer. The matrix is

$$M = \begin{bmatrix} 3 & 2 & -1 \\ 4 & -5 & 2 \end{bmatrix}$$

□

(b) What is the linear transformation that corresponds to the following matrix?

$$\begin{bmatrix} -2 & 3 \\ 4 & 6 \\ 1 & 0 \end{bmatrix}$$

Answer. The linear transformation is $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\varphi(x, y) = (-2x + 3y, 4x + 6y, x)$$

□

Theorem 19.11. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is a constant $M_\varphi \in \mathbb{R}$ such that for all $\mathbf{x} \in \mathbb{R}^n$, we have $\|\varphi(\mathbf{x})\| \leq M_\varphi \|\mathbf{x}\|$.*

Lemma. *Let $a_1, \ldots, a_n \in \mathbb{R}$. Then*

$$\left(\sum_{i=1}^n a_i\right)^2 \leq n \sum_{i=1}^n a_i^2$$

Proof. We have that

$$\begin{aligned} \left(\sum_{i=1}^n a_i\right)^2 &= (1a_1 + \cdots + 1a_n)^2 \\ &\leq \left(\sqrt{1^2 + \cdots + 1^2} \cdot \sqrt{a_1^2 + \cdots + a_n^2}\right)^2 && \text{Lemma 18.9b} \\ &= \sqrt{n}^2 \sqrt{\sum_{i=1}^n a_i^2}^2 \\ &= n \sum_{i=1}^n a_i^2 \end{aligned}$$

as desired. □

Proof of Theorem 19.11. Let

$$M = \max_{i,j} |[\varphi]_{ij}| \qquad\qquad M_\varphi = M\sqrt{nm}$$

Then

$$\begin{aligned}
\|\varphi(\mathbf{x})\| &= \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_j[\varphi]_{ij}\right)^2} \\
&\leq \sqrt{\sum_{i=1}^{m} n \sum_{j=1}^{n}(x_j[\varphi]_{ij})^2} && \text{Lemma} \\
&= \sqrt{n}\cdot\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} x_j^2[\varphi]_{ij}^2} \\
&= \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n}\left(x_j^2\sum_{i=1}^{m}[\varphi]_{ij}^2\right)} \\
&\leq \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n}\left(x_j^2\sum_{i=1}^{m} M^2\right)} \\
&= \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n} mM^2x_j^2} \\
&= M\sqrt{nm}\cdot\sqrt{\sum_{j=1}^{n} x_j^2} \\
&= M_\varphi\|\mathbf{x}\| && \text{Definition 18.6}
\end{aligned}$$

as desired. □

8/7: **Corollary 19.12.** *Any linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous.*

Proof. To prove that φ is uniformly continuous, Definition 18.42 tells us that it will suffice to show that for every $\epsilon > 0$, there exists $\delta > 0$ such that if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\|\mathbf{x} - \mathbf{y}\| < \delta$, then $\|\varphi(\mathbf{x}) - \varphi(\mathbf{y})\| < \epsilon$. Let $\epsilon > 0$ be arbitrary. Since φ is a linear transformation, Theorem 19.11 asserts that there exists $M_\varphi \in \mathbb{R}$ such that $\|\varphi(\mathbf{x})\| \leq M_\varphi\|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$. With this result, choose $\delta = \frac{\epsilon}{M_\varphi}$. Now let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$ satisfying $\|\mathbf{x} - \mathbf{y}\| < \delta$. Then

$$\begin{aligned}
\|\varphi(\mathbf{x}) - \varphi(\mathbf{y})\| &= \|\varphi(\mathbf{x} - \mathbf{y})\| && \text{Definition 19.1} \\
&\leq M_\varphi\|\mathbf{x} - \mathbf{y}\| \\
&< M_\varphi \cdot \frac{\epsilon}{M_\varphi} \\
&= \epsilon
\end{aligned}$$

as desired. □

Lemma 19.13. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. If $\lim_{\mathbf{h}\to\mathbf{0}} \|\varphi(\mathbf{h})\|/\|\mathbf{h}\| = 0$, then φ is the zero transformation, i.e., $\varphi(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$.*

Proof. Suppose for the sake of contradiction that $\varphi(\mathbf{x}) \neq \mathbf{0}$ for some $\mathbf{x} \in \mathbb{R}^n$. Since $\varphi(\mathbf{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_j[\varphi]_{ij}\mathbf{e}_i \neq 0$ by Exercise 19.8, there exists at least one nonzero $[\varphi]_{ab}$. Consequently, since $\lim_{\mathbf{h}\to\mathbf{0}} \|\varphi(\mathbf{h})\|/\|\mathbf{h}\| = 0$, Definition 18.29 tells us that there exists $\delta > 0$ such that if $\mathbf{h} \in \mathbb{R}^n$ and $0 < \|\mathbf{h} - \mathbf{0}\| < \delta$, then

$|\|\varphi(\mathbf{h})\|/\|\mathbf{h}\| - 0| = \|\varphi(\mathbf{h})\|/\|\mathbf{h}\| < |[\varphi]_{ab}|$. Let $\mathbf{h} = (0, \dots, 0, h_b, 0, \dots, 0)$ where $0 < h_b < \delta$. It follows that $0 < \|\mathbf{h} - \mathbf{0}\| < \delta$. Therefore, since

$$\begin{aligned} \|\varphi(\mathbf{h})\| &= \left\| \sum_{i=1}^{m} \sum_{j=1}^{n} h_j [\varphi]_{ij} \mathbf{e}_i \right\| \\ &= \left\| \sum_{i=1}^{m} h_b [\varphi]_{ib} \mathbf{e}_i \right\| \\ &\geq \|h_b[\varphi]_{ab} \mathbf{e}_a\| \\ &= |h_b[\varphi]_{ab}| \end{aligned} \qquad \|\mathbf{h}\| = |h_b|$$

we have that

$$\begin{aligned} |[\varphi]_{ab}| &= \frac{|h_b[\varphi]_{ab}|}{|h_b|} \\ &\leq \frac{\|\varphi(\mathbf{h})\|}{\|\mathbf{h}\|} \\ &< |[\varphi]_{ab}| \end{aligned}$$

a contradiction. □

8/11: **Definition 19.14.** A function $f : A \to \mathbb{R}^m$ is differentiable at a point $\mathbf{a} \in A$ if there exists a linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - \varphi(\mathbf{h})\|}{\|\mathbf{h}\|} = 0$$

When such a linear transformation φ exists, it is called the **total derivative** (of f at $\mathbf{a}$) and is denoted by $Df(\mathbf{a})$.

Remark 19.15. For every $\mathbf{a} \in A$ at which f is differentiable, $Df(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation defined on all of $\mathbb{R}^n$. In particular, $Df(\mathbf{a})(\mathbf{x})$ is the derivative of f at $\mathbf{a} \in A$, evaluated at $\mathbf{x} \in \mathbb{R}^n$.

Proposition 19.16. *The derivative at $\mathbf{a} \in A$ of a function $f : A \to \mathbb{R}^m$ is unique. That is, if φ and ψ are two linear transformations that satisfy the limit of Definition 19.14, then $\varphi = \psi$. So, $Df(\mathbf{a})$ is well-defined.*

Proof. Let φ, ψ be linear transformations, each of which satisfies the limit of Definition 19.14. To prove that $\varphi = \psi$, it will suffice to show that $\varphi(\mathbf{x}) = \psi(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. To do so, Definition 19.1 tells us that it will suffice to verify that $(\varphi - \psi)(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$. To do this, Lemma 19.13 tells us that it will suffice to confirm that $\lim_{\mathbf{h} \to \mathbf{0}} \|(\varphi - \psi)(\mathbf{h})\|/\|\mathbf{h}\| = 0$. Let's begin.

To confirm that $\lim_{\mathbf{h} \to \mathbf{0}} \|(\varphi - \psi)(\mathbf{h})\|/\|\mathbf{h}\| = 0$, Definition 18.29 tells us that it will suffice to demonstrate that for every $\epsilon > 0$, there exists $\delta > 0$ such that if $\mathbf{h} \in A$ and $0 < \|\mathbf{h}\| < \delta$, then $\|(\varphi - \psi)(\mathbf{h})\|/\|\mathbf{h}\| < \epsilon$. Let $\epsilon > 0$ be arbitrary. Since φ and ψ both satisfy the limit of Definition 19.14, Definition 18.29 asserts that there exists a $\delta_1 > 0$ such that if $\mathbf{h} \in A$ and $0 < \|\mathbf{h}\| < \delta_1$, then

$$\frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - \varphi(\mathbf{h})\|}{\|\mathbf{h}\|} < \frac{\epsilon}{2}$$

and there exists a $\delta_2 > 0$ such that if $\mathbf{h} \in A$ and $0 < \|\mathbf{h}\| < \delta_2$, then

$$\frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - \psi(\mathbf{h})\|}{\|\mathbf{h}\|} < \frac{\epsilon}{2}$$

Choose $\delta = \min(\delta_1, \delta_2)$. Let $\mathbf{h}$ be an arbitrary element of A satisfying $0 < \|\mathbf{h}\| < \delta$. Then

$$\begin{aligned}
\frac{\|(\varphi - \psi)(\mathbf{h})\|}{\|\mathbf{h}\|} &= \frac{\|\varphi(\mathbf{h}) - \psi(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Definition 19.1} \\
&= \frac{\|\varphi(\mathbf{h}) - (f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}))\| + \|(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) - \psi(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Corollary 18.11} \\
&= \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \varphi(\mathbf{h})\|}{\|\mathbf{h}\|} + \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \psi(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Theorem 18.10} \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

as desired. ☐

Exercise 19.17.

(a) Let $f : A \to \mathbb{R}^m$ be a constant function, and $\mathbf{a} \in A$. Then $Df(\mathbf{a}) = \mathbf{0}$. Note that here $\mathbf{0}$ represents the zero transformation.

Proof. Let $f(\mathbf{x}) = \mathbf{y}$ for all $\mathbf{x} \in A$. To prove that $Df(\mathbf{a}) = \mathbf{0}$, Lemma 19.13 tells us that it will suffice to show that $\lim_{\mathbf{h}\to\mathbf{0}} \|Df(\mathbf{a})(\mathbf{h})\|/\|\mathbf{h}\| = 0$. But since $Df(\mathbf{a})$ exists by hypothesis, Definition 19.14 implies that

$$\begin{aligned}
0 &= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{y} - \mathbf{y} - Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|-Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Theorem 18.10}
\end{aligned}$$

as desired. ☐

(b) Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and $\mathbf{a} \in \mathbb{R}^n$. Then $Df(\mathbf{a}) = f$.

Proof. To prove that $Df(\mathbf{a}) = f$, Definition 19.1 and Lemma 19.13 tell us that it will suffice to show that $\lim_{\mathbf{h}\to\mathbf{0}} \|(f - Df(\mathbf{a}))(\mathbf{h})\|/\|\mathbf{h}\| = 0$. But since $Df(\mathbf{a})$ exists by hypothesis, Definition 19.14 implies that

$$\begin{aligned}
0 &= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{a}+\mathbf{h}-\mathbf{a}) - Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Definition 19.1} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{h}) - Df(\mathbf{a})(\mathbf{h})\|}{\|\mathbf{h}\|} \\
&= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|(f - Df(\mathbf{a}))(\mathbf{h})\|}{\|\mathbf{h}\|} && \text{Theorem 18.10}
\end{aligned}$$

as desired. ☐

Exercise 19.18. Let $f : \mathbb{R} \to \mathbb{R}$ and $a \in \mathbb{R}$. Show that f is differentiable at a in the sense of Definition 19.14 if and only if f is differentiable at a in the sense of Definition 12.1. Show that if f is differentiable at a (in either sense), then $Df(a)(x) = f'(a)x$. Are these two uses of the word "differentiable" consistent?

Proof. Suppose first that f is differentiable at a in the sense of Definition 19.14, with total derivative $Df(a)$. To prove that f is differentiable at a in the sense of Definition 12.1, it will suffice to show that $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = Df(a)(1)$. To do so, Definition 11.1 tells us that it will suffice to verify that for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $h \in \mathbb{R}$ and $0 < |h| < \delta$, then $|\frac{f(a+h)-f(a)}{h} - Df(a)(1)| < \epsilon$. Let $\epsilon > 0$ be arbitrary. Since $\lim_{h\to 0} \frac{|f(a+h)-f(a)-Df(a)(h)|}{|h|} = 0$ by Definition 19.14 and Remark 18.8, Definition 11.1 asserts that there exists a δ such that if $h \in \mathbb{R}$ and $0 < |h| < \delta$, then $|\frac{f(a+h)-f(a)-Df(a)(h)}{h}| < \epsilon$. Choose this δ to be our δ. Let h be an arbitrary element of $\mathbb{R}$ satisfying $0 < |h| < \delta$. Then

$$\begin{aligned}
\left|\frac{f(a+h)-f(a)}{h} - Df(a)(1)\right| &= \left|\frac{f(a+h)-f(a)-h\cdot Df(a)(1)}{h}\right| \\
&= \left|\frac{f(a+h)-f(a)-Df(a)(h)}{h}\right| && \text{Definition 19.1} \\
&< \epsilon
\end{aligned}$$

as desired.

By a symmetric argument, we can suppose that f is differentiable at a in the sense of Definition 12.1 and subsequently prove that $\lim_{h\to 0} \frac{\|f(a+h)-f(a)-f'(a)\cdot h\|}{\|h\|} = 0$.

By the proof of the forward direction and Definition 12.1, $f'(a) = Df(a)(1)$. It follows by Definition 19.1 that

$$\begin{aligned}
f'(a)x &= xDf(a)(1) \\
&= Df(a)(x)
\end{aligned}$$

as desired. ∎