Chapter 4. Introduction to Derivatives

In this chapter, we give an informal introduction to differentiation of vector-valued functions of several variables. We state some definitions and theorems, and we provide some computational examples, but the proofs are postponed until the following chapter.

4.1 Definition: Let $U \subseteq \mathbb{R}^n$ be open in $\mathbb{R}^n$, let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, and let at $a \in U$, say $a = (a_1, \cdots, a_n)$. We define the k^{th} **partial derivative** of f at a to be

$$\frac{\partial f}{\partial x_k}(a) = {g_k}'(a_k) \text{ , where } g_k(t) = f(a_1, \cdots, a_{k-1}, t, a_{k+1}, \cdots a_n) \, ,$$

or equivalently,

$$\frac{\partial f}{\partial x_k}(a) = {h_k}'(0) \text{ , where } h_k(t) = f(a_1, \cdots, a_{k-1}, a_k + t, a_{k+1}, \cdots a_n) \, ,$$

provided that the derivatives exist. Note that g_k and h_k are functions of a single variable.

Sometimes $\frac{\partial f}{\partial x_k}$ is written as f_{x_k} or as f_k. When we write $u = f(x)$, we can also write $\frac{\partial f}{\partial x_k}$ as $\frac{\partial u}{\partial x_k}$, u_{x_k} or u_k. When $n = 3$ and we write x, y and z instead of x_1, x_2 and x_3, the partial derivatives $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ and $\frac{\partial f}{\partial x_3}$ are written as $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$, or as f_x, f_y and f_z. When $n = 1$ so there is only one variable $x = x_1$ we have $\frac{\partial f}{\partial x}(a) = \frac{df}{dx}(a) = f'(a)$.

4.2 Example: Let $f(x, y) = x^3 y + 2xy^2$. Find $\frac{\partial f}{\partial x}(1, 2)$ and $\frac{\partial f}{\partial y}(1, 2)$.

Solution: Let $g_1(t) = f(t, 2) = 2t^3 + 8t$. Then $g_1'(t) = 6t^2 + 8$ so $\frac{\partial f}{\partial x}(1, 2) = g'(1) = 14$. Let $g_2(t) = f(1, t) = t + 2t^2$. Then $g_2'(t) = 1 + 4t$ so $\frac{\partial f}{\partial y}(1, 2) = g_2'(2) = 9$.

4.3 Note: Rather than explicitly determining the functions $g_k(t)$ as we did in the above solution, we can calculate the partial derivative $\frac{\partial f}{\partial x_k}(a)$ by simply treating the variables x_i with $i \neq k$ as constants, and differentiating f as if it were a function of the single variable x_k.

4.4 Example: Let $f(x, y, z) = (x - z^2) \sin(x^2 y + z)$. Find $\frac{\partial f}{\partial x}(x, y, z)$ and $\frac{\partial f}{\partial x}(3, \frac{\pi}{2}, 0)$.

Solution: Treating y and z as constants, we obtain

$$\frac{\partial f}{\partial x}(x, y, z) = \sin(x^2 y + z) + (x - z^2) \cos(x^2 y + z)(2xy)$$

and so $\frac{\partial f}{\partial x}(3, \frac{\pi}{2}, 0) = \sin \frac{9\pi}{2} + 3 \cos \frac{9\pi}{2}(3\pi) = 1$.

4.5 Definition: Let $U \subseteq \mathbb{R}^n$ be open in $\mathbb{R}^n$, let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and let $a \in U$. Write $u = f(x) = \big(f_1(x), f_2(x), \cdots, f_m(x)\big)^T$ with $x = (x_1, x_2, \cdots, x_n)^T$. We define the **derivative matrix**, or the **Jacobian matrix**, of f at a to be the matrix

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

and we define the **linearization** of f at a to be the affine map $L : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$L(x) = f(a) + Df(a)(x - a)$$

provided that all the partial derivatives $\frac{\partial f_k}{\partial x_l}(a)$ exist.

4.6 Definition: Let U be open in $\mathbb{R}^n$ and let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. We say that f is $\mathcal{C}^1$ in U when all the partial derivatives $\frac{\partial f_k}{\partial f_l}$ exist and are continuous in U. The **second order partial derivatives** of f are the functions

$$\frac{\partial^2 f_j}{\partial x_k \partial x_l} = \frac{\partial\left(\frac{\partial f_j}{\partial x_l}\right)}{\partial x_k}.$$

We also write $\frac{\partial^2 f_j}{\partial {x_k}^2} = \frac{\partial^2 f_j}{\partial x_k \partial x_k}$. We say that f is $\mathcal{C}^2$ when all the partial derivatives $\frac{\partial^2 f_j}{\partial x_k \partial x_l}$ exist and are continuous in U. Higher order derivatives can be defined similarly, and we say f is $\mathcal{C}^k$ when all the k^{th} order derivatives $\frac{\partial^k f_j}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}$ exist and are continuous in U.

4.7 Definition: Let $a \in U$ where U is an open set in $\mathbb{R}$, and let $f : U \subseteq \mathbb{R} \to \mathbb{R}^m$, say $x = f(t) = \big(x_1(t), x_2(t), \cdots, x_m(t)\big)$. Then we write $f'(a) = Df(a)$ and we have

$$f'(a) = Df(a) = \begin{pmatrix} \frac{\partial x_1}{\partial t}(a) \\ \vdots \\ \frac{\partial x_m}{\partial t}(a) \end{pmatrix} = \begin{pmatrix} {x_1}'(a) \\ \vdots \\ {x_m}'(a) \end{pmatrix}.$$

The vector $f'(a)$ is called the **tangent vector** to the curve $x = f(t)$ at the point $f(a)$. In the case that t represents time and $f(t)$ represents the position of a moving point, $f'(a)$ is also called the **velocity** of the moving point at time $t = a$.

4.8 Definition: Let $a \in U$ where U is an open set in $\mathbb{R}^n$ and let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. We define the **gradient** of f at a to be the vector

$$\nabla f(a) = Df(a)^T = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

4.9 Definition: Let $U \subseteq \mathbb{R}^n$ be open in $\mathbb{R}^n$, let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, and let $a \in U$. We say that f is **differentiable** at a when there exists an affine map $L : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U \Big(|x-a| \leq \delta \implies \big|f(x) - L(x)\big| \leq \epsilon |x-a|\Big).$$

We say that f is differentiable in U when f is differentiable at every point $a \in U$.

4.10 Theorem: *Let $U \subseteq \mathbb{R}^n$ be open, let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and let $a \in U$. Then*

(1) If f is differentiable at a then the partial derivatives of f at a all exist, and the affine map L which appears in the definition of the derivative is the linearization of f at a.
(2) If f is differentiable in U then f is continuous in U.
(3) If f is $\mathcal{C}^1$ in U then f is differentiable in U.
(4) If f is $\mathcal{C}^2$ in U then $\frac{\partial^2 f_j}{\partial x_k \partial x_\ell} = \frac{\partial^2 f_j}{\partial x_\ell \partial x_k}$ for all j, k, ℓ.

Proof: The proof will be given in the next two chapters.

4.11 Note: Let $a \in U$ where U is open in $\mathbb{R}^n$ and let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at a. The definition of the derivative, together with Part (1) of the above theorem, imply that the function f is approximated by its linearization near $x = a$, that is when $x \cong a$ we have

$$f(x) \cong L(x) = f(a) + Df(a)(x - a)\,.$$

The geometric objects (curves and surfaces etc) Graph(f), Null(f), $f^{-1}(k)$ and Range(f) are all approximated by the affine spaces Graph(L), Null(L), $L^{-1}(k)$ and Range(L). Each of these affine spaces is called the (affine) **tangent space** of its corresponding geometric object: the space Graph(L) is called the (affine) tangent space of the set Graph(f) at the point $\big(a, f(a)\big)$; when $f(a) = 0$, the space Null(L) is called the (affine) tangent space of Null(f) at the point a, and more generally when $f(a) = k$, so that $a \in f^{-1}(k)$, the space $L^{-1}(k)$ is called the (affine) tangent space to $f^{-1}(k)$ at the point a; and the space Range(L) is called the (affine) tangent space of the set Range(f) at the point $f(a)$. When a tangent space is 1-dimensional we call it a **tangent line** and when a tangent space is 2-dimensional we call it a **tangent plane**.

4.12 Example: Find an explicit, an implicit and a parametric equation for the tangent line to the curve in $\mathbb{R}^2$ which is defined explicitly by the equation $y = f(x)$, implicitly by the equation $g(x, y) = k$, and parametrically by the equation $(x, y) = \alpha(t) = \big(x(t), y(t)\big)$.

Solution: The curve in $\mathbb{R}^2$ defined explicitly by $y = f(x)$ has a tangent line at the point $(a, f(a))$ which is given explicitly by $y = L(x)$, that is

$$y = f(a) + f'(a)(x - a).$$

When $g(a, b) = k$, the curve in $\mathbb{R}^2$ defined implicitly by the equation $g(x, y) = k$ has a tangent line at the point (a, b) which is given implicitly by the equation $L(x, y) = k$, that is by $f(a, b) + \big(\frac{\partial f}{\partial x}(a, b)\,,\ \frac{\partial f}{\partial y}(a, b)\big)(x - a, y - b)^T = k$, or equivalently by

$$\tfrac{\partial f}{\partial x}(a, b)(x - a) + \tfrac{\partial f}{\partial y}(a, b)(y - b) = 0.$$

The curve in $\mathbb{R}^2$ defined parametrically by $(x, y) = \alpha(t) = \big(x(t), y(t)\big)$ or, more accurately, by $(x, y)^T = \alpha(t) = \big(x(t), y(t)\big)^T$ has a tangent line at the point $\alpha(a) = \big(x(a), y(a)\big)^T$ which is given parametrically by $(x, y)^T = L(t) = \alpha(a) + \alpha'(a)(t - a)$, that is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(a) \\ y(a) \end{pmatrix} + \begin{pmatrix} x'(a) \\ y'(a) \end{pmatrix} (t - a).$$

4.13 Example: Find an explicit, an implicit, and a parametric equation for the tangent line to the curve in $\mathbb{R}^3$ which is defined explicitly by $(x, y) = f(z) = \big(x(z), y(z)\big)$, implicitly by $u(x, y, z) = k$ and $v(x, y, z) = l$, and parametrically by $(x, y, z) = \alpha(t) = \big(x(t), y(t), z(t)\big)$.

Solution: The curve in $\mathbb{R}^3$ given explicitly by $(x, y) = f(z) = \big(x(z), y(z)\big)$ or, more accurately, by $(x, y)^T = f(z) = \big(x(z), y(z)\big)^T$, has a tangent plane at the point $(x(c), y(c), c)$ which is given explicitly by $(x, y)^T = L(z) = f(c) + Df(c)(z - c)$, that is by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(c) \\ y(c) \end{pmatrix} + \begin{pmatrix} x'(c) \\ y'(c) \end{pmatrix} (z - c)$$

When $u(a, b, c) = k$ and $v(a, b, c) = \ell$ and we write $g(x, y, z) = \big(u(x, y, z), v(x, y, z)\big)^T$, the curve in $\mathbb{R}^3$ given implicitly by $g(x, y, z) = (k, \ell)^T$, has a tangent line at (a, b, c) given implicitly by $L(x,y,z)=(k,\ell)^T$, that is b $g(a,b,c)+Dg(a,b,c)\big(x–a,y–b,z–c\big)^T=(k,\ell)^T$, or equivalently by

$$\begin{pmatrix} \frac{\partial u}{\partial x}(a,b,c) & \frac{\partial u}{\partial y}(a,b,c) & \frac{\partial u}{\partial z}(a,b,c) \\ \frac{\partial v}{\partial x}(a,b,c) & \frac{\partial v}{\partial y}(a,b,c) & \frac{\partial v}{\partial z}(a,b,c) \end{pmatrix} \begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The curve in $\mathbb{R}^3$ given parametrically by $(x,y,z)^T=\alpha(a)=\big(x(a),y(a),z(a)\big)^T$ has a tangent line at $\alpha(a)$ which is given parametrically by $(x,y,z)^T=L(t)=\alpha(a)+\alpha'(a)(t-a)$, that is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(a) \\ y(a) \\ z(a) \end{pmatrix} + \begin{pmatrix} x'(a) \\ y'(a) \\ z'(a) \end{pmatrix}(t-a).$$

4.14 Example: Find an explicit, an implicit and a parametric equation for the tangent plane to the surface in $\mathbb{R}^3$ which is defined explicitly by $z=f(x,y)$, implicitly by $g(x,y,z)=k$, and parametrically by $(x,y,z)=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)$.

Solution: The surface in $\mathbb{R}^3$ given explicitly by $z=f(x,y)$ has a tangent plane at the point $\big(a,b,f(a,b)\big)$ given explicitly by $z=L(x,y)=f(a,b)+Df(a,b)\big(x–a,y–b\big)^T$, that is

$$z=f(a,b)+\frac{\partial f}{\partial x}(a,b)(x-a)+\frac{\partial f}{\partial y}(a,b)(y-b).$$

When $g(a,b,c)=k$, the surface in $\mathbb{R}^3$ given implicitly by $g(x,y,z)=k$ has tangent plane at (a,b,c) given implicitly by $L(x,y,z)=k$, that is $g(a,b,c)+Dg(a,b,c)\big(x–a,y–b,z–c\big)^T=k$ or equivalenty

$$\frac{\partial g}{\partial x}(a,b,c)(x-a)+\frac{\partial g}{\partial y}(a,b,c)(y-b)+\frac{\partial g}{\partial z}(a,b,c)(z-c)=0.$$

The surface in $\mathbb{R}^3$ defined parametrically by $(x,y,z)=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)$ or, more accurately, by $(x,y,z)^T=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)^T$ has a tangent plane at $\sigma(a,b)$ which is given parametrically by $(x,y,z)^T=L(s,t)=\sigma(a,b)+D\sigma(a,b)\big(s–a,t–b\big)^T$, that is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(a,b) \\ y(a,b) \\ z(a,b) \end{pmatrix} + \begin{pmatrix} \frac{\partial x}{\partial s}(a,b) & \frac{\partial x}{\partial t}(a,b) \\ \frac{\partial y}{\partial s}(a.b) & \frac{\partial y}{\partial t}(a,b) \\ \frac{\partial z}{\partial s}(a,b) & \frac{\partial z}{\partial t}(a,b) \end{pmatrix} \begin{pmatrix} s-a \\ t-b \end{pmatrix}.$$

4.15 Example: Find a parametric equation for the tangent line to the helix given by $(x,y,z)=(2\cos t,2\sin t,t)$ at the point where $t=\frac{\pi}{3}$, and find the point where this tangent line crosses the xz-plane.

Solution: Let $f(t)=(2\cos t,2\sin t,t)$ and note that $f'(t)=\big(-2\sin t,2\cos t,1\big)$. We have $f\big(\frac{\pi}{3}\big)=\big(1,\sqrt{3},\frac{\pi}{3}\big)$ and $f'\big(\frac{\pi}{3}\big)=(-\sqrt{3},1,1)$ and so the tangent line at the point $f\big(\frac{\pi}{3}\big)$ is given parametrically by $(x,y,z)=L(t)=\big(1,\sqrt{3},\frac{\pi}{3}\big)+\big(-\sqrt{3},1,1\big)\big(t-\frac{\pi}{3}\big)$. The point of intersection with the xz-plane occurs when $y=0$, that is when $\sqrt{3}+t-\frac{\pi}{3}=0$, so we take $t=\frac{\pi}{3}-\sqrt{3}$ to obtain $(x,y,z)=L\big(\frac{\pi}{3}-\sqrt{3}\big)=\big(1,\sqrt{3},\frac{\pi}{3}\big)-\sqrt{3}\big(-\sqrt{3},1,1\big)=\big(4,0,\frac{\pi}{3}-\sqrt{3}\big)$.

4.16 Example: Find an explicit equation for the tangent plane to the surface $z = \dfrac{e^{x^2+2xy}}{\sqrt{2+y}}$ at the point $(2,-1)$.

Solution: Let $f(x,y) = \frac{e^{x^2+2xy}}{\sqrt{2+y}}$. Then

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{e^{x^2+2y}(2x+2y)}{\sqrt{2+y}} \\ \frac{\partial f}{\partial y}(x,y) &= \frac{e^{x^2+2y}(2x)\sqrt{2+y} - e^{x^2+2xy}\frac{1}{2\sqrt{2+y}}}{2+y} \end{aligned}$$

so we have $f(2,-1) = 1$, and $\frac{\partial f}{\partial x}(2,-1) = 2$ and $\frac{\partial f}{\partial y}(2,-1) = \frac{7}{2}$. Thus the equation to the tangent plane is $z = 1 + 2(x-2) + \frac{7}{2}(y+1)$, or equivalently $4x + 7y - 2z = -1$.

4.17 Example: Find an implicit equation for the tangent line to the curve given by $2\sqrt{y+x^2} + \ln(y-x^2) = 6$ at the point $(2,5)$.

Solution: Let $g(x,y) = 2\sqrt{y+x^2} + \ln(y-x^2)$ and note that $g(2,5) = 2\sqrt{9} + \ln 1 = 6$. We have $\frac{\partial g}{\partial x}(x,y) = \frac{2x}{\sqrt{y+x^2}} - \frac{2x}{y-x^2}$ and $\frac{\partial g}{\partial y}(x,y) = \frac{1}{\sqrt{y+x^2}} + \frac{1}{y-x^2}$ so that $\frac{\partial g}{\partial x}(2,5) = \frac{4}{3} - \frac{4}{1} = -\frac{8}{3}$ and $\frac{\partial g}{\partial y}(2,5) = \frac{1}{3} + \frac{1}{1} = \frac{4}{3}$, so the tangent line at $(2,5)$ is given by $-\frac{8}{3}(x-2) + \frac{4}{3}(y-5) = 0$ or, equivalently, by $2(x-2) = (y-5)$ or by $y = 2x+1$.

4.18 Example: Find a parametric equation for the tangent line to the curve of intersection of the paraboloid $z = 2 - x^2 - y^2$ with the cone $y = \sqrt{x^2+z^2}$ at the point $p = (1,1,0)$.

Solution: Note that the paraboloid is given by $x^2 + y^2 + z = 2$ and the cone is given by $x^2 - y^2 + z^2 = 0$, with $y \geq 0$. Let $u(x,y,z) = x^2 + y^2 + z$ and $v(x,y,z) = x^2 - y^2 + z^2$ and let $g(x,y,z) = \big(u(x,y,z), v(x,y,z)\big)^T$ so that the curve of intersection is given implicitly by $g(x,y,z) = (2,0)^T$. Note that $g(1,1,0) = (2,0)^T$ and

$$\begin{aligned} Dg(x,y,z) &= \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x & 2y & 1 \\ 2x & -2y & 2z \end{pmatrix} \\ Dg(1,1,0) &= \begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix} \end{aligned}$$

The tangent line at $(1,1,0)$ is given implicitly by $Dg(1,1,0)\big(x-1, y-1, z\big)^T = (0,0)^T$ that is

$$\begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x-1 \\ y-1 \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is equivalent to the pair of equations $2(x-1) + 2(y-1) + z = 0$ and $2(x-1) - 2(y-1) = 0$. We remark that these are the equations of the tangent planes to the two given surfaces at $(1,1,0)$. The two equations are equivalent to $2x + 2y + z = 4$ and $x - y = 0$. We let $y = t$, then the second equation gives $x = y = t$, and the first equation gives $z = 4 - 2x - 2y = 4 - 4t$, so the line is given parametrically by $(x,y,z) = (0,0,4) + t(1,1,-4)$.

4.19 Exercise: Find an explicit equation for the tangent plane to the surface given by $(x,y,z) = \big(r\cos t\,,\ r\sin t\,,\ \frac{3}{1+r^2}\big)$ at the point where $(r,t) = \big(\sqrt{2}, \frac{\pi}{4}\big)$.

4.20 Theorem: *(The Chain Rule) Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$, let $g : V \subseteq \mathbb{R}^m \to \mathbb{R}^l$, and let $h(x) = g(f(x))$. If f is differentiable at a and g is differentiable at $f(a)$ then h is differentiable at a and $Dh(a) = Dg(f(a))Df(a)$.*

Proof: A proof will be given in the next chapter.

4.21 Exercise: Let $z = f(x,y) = 4x^2 - 8xy + 5y^2$, $(u,v) = g(z) = \big(\sqrt{z-1}\,,\, 5\ln z\big)$ and $h(x,y) = g\big(f(x,y)\big)$. Find $Dh(2,1)$.

4.22 Exercise: Let $(x,y) = f(r,\theta) = \big(r\cos\theta\,,\, r\sin\theta\big)$, let $z = g(x,y)$ and let $z = h(r,\theta) = g\big(f(r,\theta)\big)$. If $h(r,\theta) = r^2 e^{\sqrt{3}(\theta - \frac{\pi}{6})}$ then find $\nabla g(\sqrt{3},1)$.

4.23 Exercise: Let $(x,y,z) = f(s,t)$ and $(u,v) = g(x,y,z)$. Find a formula for $\frac{\partial u}{\partial t}$.

4.24 Definition: Let $a \in U$ where U is an open set in $\mathbb{R}^n$, let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at a, and let $v \in \mathbb{R}^n$. We define the **directional derivative of** f **at** a **with respect to** v, written as $D_v f(a)$, as follows: pick any differentiable curve $\alpha(t)$ with $\alpha(0) = a$ and $\alpha'(0) = v$ $\big($for example, we could pick $\alpha(t) = a + v\,t\big)$, and define $D_v f(a)$ to be the rate of change of the function f at $t = 0$ as we move along the curve α. To be precise, let $\beta(t) = f\big(\alpha(t)\big)$, note that $\beta'(t) = Df\big(\alpha(t)\big)\alpha'(t)$, and then define $D_v f(a)$ to be

$$\begin{aligned} D_v f(a) &= \beta'(0) \\ &= Df(\alpha(0))\,\alpha'(0) \\ &= Df(a)\,v \\ &= \nabla f(a) \cdot v\,. \end{aligned}$$

Notice that the formula for $D_v f(a)$ does not depend on the choice of the curve $\alpha(t)$. The (directional) **derivative of** f **in the direction of** v is defined to be $D_w f(a)$ where w is the unit vector in the direction of v, that is $w = \frac{v}{|v|}$.

4.25 Exercise: Let $f(x,y,z) = x\sin(y^2 - 2xz)$ and let $\alpha(t) = \big(\sqrt{t}\,,\, \frac{1}{2}t\,,\, e^{(t-4)/4}\big)$. Find the rate of change of f as we move along the curve $\alpha(t)$ when $t = 4$.

4.26 Theorem: *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at $a \in U$. Say $f(a) = b$. The gradient $\nabla f(a)$ is perpendicular to the level set $f(x) = b$, it is in the direction in which f increases most rapidly, and its length is the rate of increase of f in that direction.*

Proof: The proof will be given in the next chapter.

4.27 Note: Let $a \in U$ where U is an open set in $\mathbb{R}^n$, and let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable. The k^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$f_{x_k}(a) = \frac{\partial f}{\partial x_k}(a) = \Big(\frac{\partial f_1}{\partial x_k}(a), \cdots, \frac{\partial f_m}{\partial x_k}(a)\Big)^T \in \mathbb{R}^m,$$

which is the tangent vector to the curve $\beta_k(t) = f\big(\alpha_k(t)\big)$ at $t = 0$, where α_k is the curve through a in the direction of the standard basis vector e_k given by $\alpha_k(t) = a + te_k$.

The ℓ^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$\nabla f_\ell(a) = \Big(\frac{\partial f_\ell}{\partial x_1}(a), \cdots, \frac{\partial f_\ell}{\partial x_n}(a)\Big)^T$$

which is orthogonal to the level set $f_\ell(x) = f_\ell(a)$, and points in the direction in which f_ℓ increases most rapidly, and its length is the rate of increase of f_ℓ in that direction.