Chapter 3 Manifolds, Vector Fields, and Differential Equations

John Chiasson
Boise State University

Contents

1

Vector Fields and Differential Equations

Recall the northern hemisphere patch $\mathbf{z}(x_1, x_2) : \mathcal{D}_1 \to \mathcal{U}_1$ for $\mathbf{S}^2$

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1 - (x_1^2 + x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D}_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \le x_1^2 + x_2^2 < 1\right\}$ and $\mathcal{U}_1 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 | z_3 > 0\right\}$. Consider the system of differential equations defined on the northern hemisphere coordinates given by

$$\begin{aligned} \frac{dx_1}{dt} &= -x_2 \\ \frac{dx_2}{dt} &= x_1 \end{aligned}$$

with $(x_1(0), x_2(0)) = (1/2, 0)$. The solution is

$$\begin{aligned} x_1(t) &= \frac{1}{2}\cos(t) \\ x_2(t) &= \frac{1}{2}\sin(t) \end{aligned}$$

and is shown in Figure 1.1.

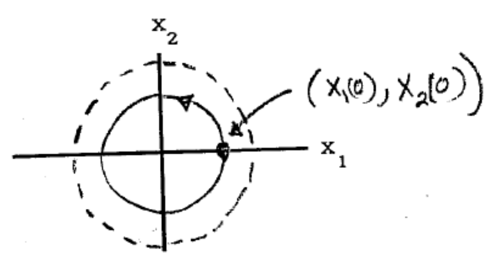

FIGURE 1.1. The northern hemisphere curve $(x_1(t), x_2(t)) = (\frac{1}{2}\cos(t), \frac{1}{2}\sin(t))$.

This curve in the northern hemisphere coordinates results in the curve on $\mathbf{S}^2$ given by

$$\mathbf{z}(x_1(t), x_2(t)) = \begin{bmatrix} \frac{1}{2}\cos(t) \\ \frac{1}{2}\sin(t) \\ \sqrt{\frac{3}{2}} \end{bmatrix} \in \mathbf{S}^2.$$

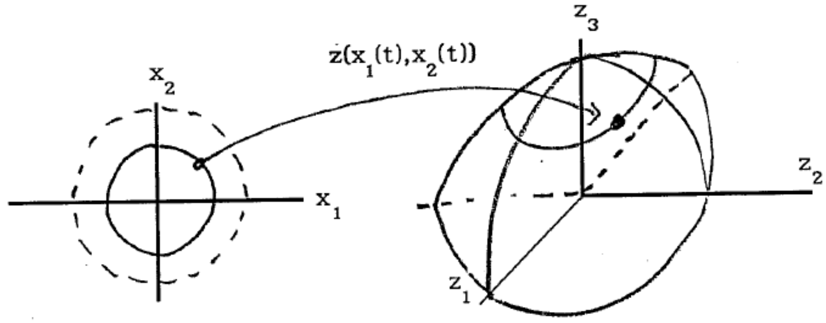

FIGURE 1.2. The curve $\begin{bmatrix} \frac{1}{2}\cos(t) \\ \frac{1}{2}\sin(t) \\ \sqrt{\frac{3}{2}} \end{bmatrix} \in \mathbf{S}^2$.

What is the differential equation describing this curve in the eastern hemisphere coordinate chart? The eastern hemisphere patch $\mathbf{z}^*(x_1^*, \hat{x_2}) : \mathcal{D}_2 \to \mathcal{U}_2$ for $\mathbf{S}^2$ is given by

$$\mathbf{z}^*(x_1^*, x_2^*) = \begin{bmatrix} x_2^* \\ \sqrt{1 - ((x_1^*)^2 + (x_2^*)^2)} \\ x_1^* \end{bmatrix} \in \mathbf{S}^2.$$

Here $\mathcal{D}_2 = \left\{(x_1^*, x_2^*) \in \mathbb{R}^2 |\ 0 \leq (x_1^*)^2 + (x_2^*)^2 < 1\right\}$ and $\mathcal{U}_2 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 | z_2 > 0\right\}$. For $-\pi/2 < t < \pi/2$ this curve in $\mathcal{U}_1 \cap \mathcal{U}_2 \subset \mathbf{S}^2$, that is, it is in both the northern and eastern coordinate charts. The change of coordinates from the northern hemisphere patch to the eastern hemisphere patch is

$$\begin{aligned} x_1^* &= \sqrt{1 - (x_1^2 + x_2^2)} \\ x_2^* &= x_1 \end{aligned}$$

with inverse

$$\begin{aligned} x_1 &= x_2^* \\ x_2 &= \sqrt{1 - ((x_1^*)^2 + (x_2^*)^2)}. \end{aligned}$$

Then

$$\begin{aligned} \frac{dx_1^*}{dt} &= \frac{1}{2}\frac{-2x_1}{\sqrt{1-(x_1^2+x_2^2)}}\frac{dx_1}{dt} + \frac{1}{2}\frac{-2x_2}{\sqrt{1-(x_1^2+x_2^2)}}\frac{dx_2}{dt} = \frac{-x_1}{\sqrt{1-(x_1^2+x_2^2)}}(-x_2) + \frac{-x_2}{\sqrt{1-(x_1^2+x_2^2)}}x_1 = 0 \\ \frac{dx_2^*}{dt} &= \frac{dx_1}{dt} = -x_2 = -\sqrt{1-((x_1^*)^2+(x_2^*)^2)} \end{aligned}$$

or

$$\begin{aligned} \frac{dx_1^*}{dt} &= 0 \\ \frac{dx_2^*}{dt} &= -\sqrt{1-((x_1^*)^2+(x_2^*)^2)} \end{aligned}$$

with initial conditions $(x_1^*(0), x_2^*(0)) = \left(\sqrt{1-(x_1^2(0)+x_2^2(0))}, x_1(0)\right) = (\sqrt{3}/2, 1/2)$.

Note that the differential equations in the x^* (eastern hemisphere) coordinates are nonlinear while in the x (northern hemisphere) coordinates they are linear. However, they describe the *same* curve on the manifold $\mathbf{S}^2$.

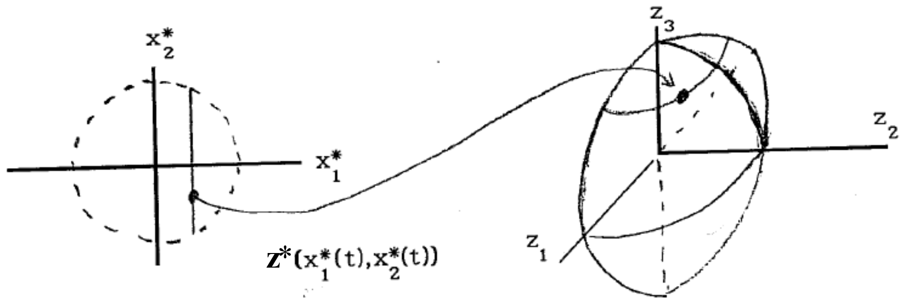

FIGURE 1.3. The curve $\begin{bmatrix} \frac{1}{2}\cos(t) \\ \frac{1}{2}\sin(t) \\ \sqrt{\frac{3}{2}} \end{bmatrix} \in \mathbf{S}^2$.

In contrast, suppose we started on the eastern hemisphere coordinate chart with the differential equations

$$\begin{aligned} \frac{dx_1^*}{dt} &= -x_2^* \\ \frac{dx_2^*}{dt} &= x_1^* \end{aligned}$$

and $(x_1^*(0), x_2^*(0)) = (1/2, 0)$. The solution is

$$\begin{aligned} x_1^*(t) &= \frac{1}{2}\cos(t) \\ x_2^*(t) &= \frac{1}{2}\sin(t) \end{aligned}$$

resulting in the curve

$$\mathbf{z}^*(x_1^*(t), x_2^*(t)) = \begin{bmatrix} x_2^*(t) \\ \sqrt{1 - ((x_1^*(t))^2 + (x_2^*(t))^2)} \\ x_1^*(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sin(t) \\ \sqrt{\frac{3}{2}} \\ \frac{1}{2}\cos(t) \end{bmatrix} \in \mathbf{S}^2.$$

As shown in Figure 1.4 below this is a *different* curve on the manifold from that just considered above.

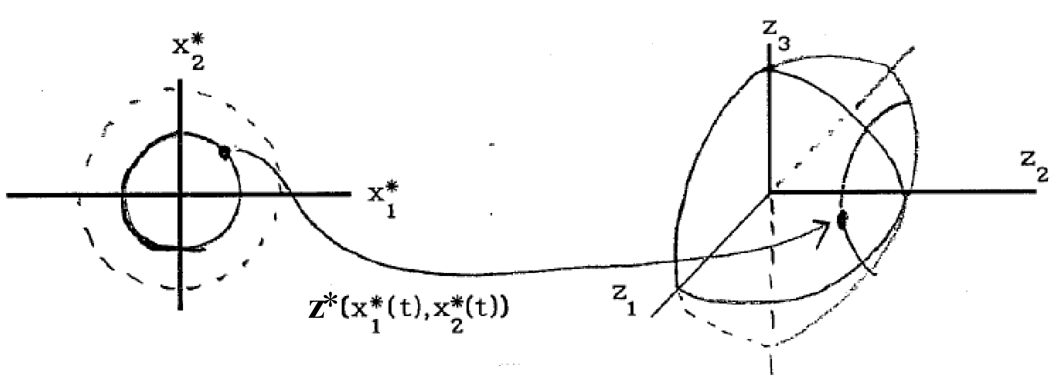

FIGURE 1.4. The curve $\begin{bmatrix} \frac{1}{2}\sin(t) \\ \sqrt{\frac{3}{2}} \\ \frac{1}{2}\cos(t) \end{bmatrix} \in \mathbf{S}^2$.

1.1 Vector Fields and Tangent Vectors

We again consider the differential equation

$$\begin{aligned}\frac{dx_1}{dt} &= -x_2\\ \frac{dx_2}{dt} &= x_1\end{aligned}$$

in the northern hemisphere coordinate system. With the initial conditions $(x_1(0), x_2(0)) = (x_{01}, x_{02})$ the solution is

$$\begin{aligned}x_1(t) &= x_{01}\cos(t) - x_{02}\sin(t)\\ x_2(t) &= x_{01}\sin(t) + x_{02}\cos(t).\end{aligned}$$

The corresponding curve $\mathbf{z}(x_1(t), x_2(t)) : \mathcal{D}_1 \to \mathcal{U}_1$ on $\mathbf{S}^2$ is given by

$$\mathbf{z}(x_1(t), x_2(t)) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \sqrt{1-(x_1^2(t)+x_2^2(t))} \end{bmatrix} = \begin{bmatrix} x_{01}\cos(t) - x_{02}\sin(t) \\ x_{01}\sin(t) + x_{02}\cos(t) \\ \sqrt{1-(x_{01}^2+x_{02}^2)} \end{bmatrix} \in \mathbf{S}^2.$$

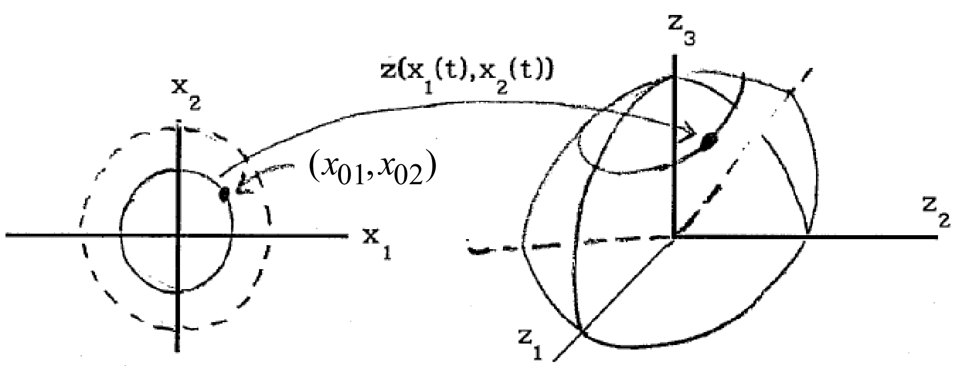

FIGURE 1.5.

The tangent to the curve at time t is

$$\begin{aligned}\frac{d}{dt}\mathbf{z}(x_1(t), x_2(t)) &= \frac{\partial \mathbf{z}(x_1,x_2)}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial \mathbf{z}(x_1,x_2)}{\partial x_2}\frac{dx_2}{dt}\\ &= \underbrace{\begin{bmatrix} 1 \\ 0 \\ \dfrac{-x_1}{\sqrt{1-(x_1^2+x_2^2)}} \end{bmatrix}}_{\mathbf{z}_{x_1}}(-x_2) + \underbrace{\begin{bmatrix} 0 \\ 1 \\ \dfrac{-x_2}{\sqrt{1-(x_1^2+x_2^2)}} \end{bmatrix}}_{\mathbf{z}_{x_2}}(x_1).\end{aligned}$$

The pair of coordinates $(x_1, x_2) \in \mathcal{D}_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 \middle|\ 0 \le x_1^2 + x_2^2 < 1\right\}$ corresponds to the point

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1-(x_1^2+x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

with the *unique* tangent vector at that point given by

$$(-x_2)\mathbf{z}_{x_1} + (x_1)\mathbf{z}_{x_2} \in \mathbf{T}_p(\mathbf{S}^2).$$

The component of this tangent vector are $-x_2$ and x_1. In other words, for each point $p \in \mathcal{U}_1$ of the northern hemisphere there is a unique tangent vector whose *components* are specified by the differential equation

$$\begin{aligned} \frac{dx_1}{dt} &= -x_2 \\ \frac{dx_2}{dt} &= \quad x_1. \end{aligned}$$

This set of vectors is called a vector field. Specifically, a *vector field* on an open subset $\mathcal{U}$ of a manifold $\mathcal{M}$ is an assignment to each point of $\mathcal{U}$ a unique tangent from the tangent space $\mathbf{T}_p(\mathcal{M})$. If the components of the vector field are differentiable functions of the coordinates then it is a *smooth* vector field. However, in all that follows the term *vector field* will be taken to mean a *smooth* vector field.

Consider a general differential equation on northern hemisphere coordinates given by

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2). \end{aligned}$$

This specifies the unique tangent vector $\in \mathbf{T}_p(\mathbf{S}^2)$ given by

$$f_1(x_1, x_2)\mathbf{z}_{x_1} + f_2(x_1, x_2)\mathbf{z}_{x_2}$$

at each point $p \in \mathcal{U}_1$ with local coordinates $\varphi(p) = (x_1, x_2)$. For any solution $(x_1(t), x_2(t))$ of the above differential equation we have

$$\frac{d}{dt}\mathbf{z}(x_1(t), x_2(t)) = \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_2}\frac{dx_2}{dt} = f_1(x_1, x_2)\mathbf{z}_{x_1} + f_2(x_1, x_2)\mathbf{z}_{x_2}.$$

Change of Basis Vectors for the Tangent Space $\mathbf{T}_p(\mathbf{S}^2)$

We look at how the basis vectors of the tangent space $\mathbf{T}_p(\mathbf{S}^2)$ transform going from one coordinate system to another. We will rediscover how the components of a tangent vector change going between coordinate systems. To do this consider the change of coordinates between the spherical and northern hemisphere patches on $\mathbf{T}_p(\mathbf{S}^2)$.

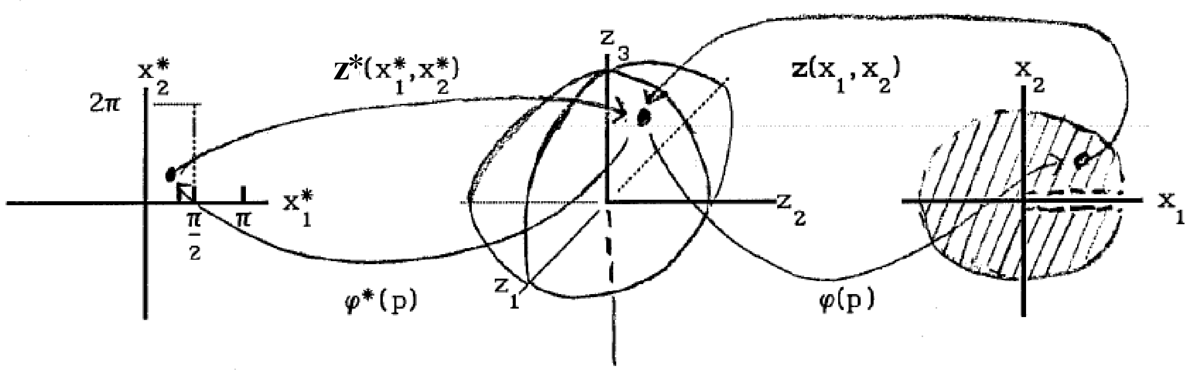

FIGURE 1.6. Spherical and northen hemisphere patches on $\mathbf{T}_p(\mathbf{S}^2)$

The spherical coordinate map $\mathbf{z}^*(x_1^*, x_2^*) : \mathcal{D}_1 \to \mathcal{U}_1$ for $\mathbf{S}^2$ is

$$\mathbf{z}^*(x_1^*, x_2^*) = \begin{bmatrix} \sin(x_1^*)\cos(x_2^*) \\ \sin(x_1^*)\sin(x_2^*) \\ \cos(x_1^*) \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D}_1 = \left\{(x_1^*, x_2^*) \in \mathbb{R}^2 |\ 0 < x_1^* < \pi, 0 < x_2^* < 2\pi\right\}$ and $\mathcal{U}_1 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 |\ \text{If } z_1 > 0 \text{ then } z_2 \neq 0\right\}$.

The northern hemisphere coordinate map $\mathbf{z}(x_1, x_2) : \mathcal{D}_2 \to \mathcal{U}_2$ for $\mathbf{S}^2$ is

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1-(x_1^2+x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D}_2 = \left\{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \leq x_1^2 + x_2^2 < 1\right\}$ and $\mathcal{U}_2 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 |\ z_3 > 0\right\}$.

The coordinate transformation from spherical to northern hemisphere coordinates $\varphi \circ \varphi^{*-1}$ is

$$\begin{aligned} x_1 &= \sin(x_1^*)\cos(x_2^*) \\ x_2 &= \sin(x_1^*)\sin(x_2^*) \end{aligned}$$

with inverse $\varphi^* \circ \varphi^{-1}$ given by

$$\begin{aligned} x_1^* &= \sqrt{x_1^2 + x_2^2} \\ x_2^* &= \tan^{-1}(x_1, x_2). \end{aligned}$$

Consider the calculation

$$\mathbf{z}(x_1, x_2)\Big|_{\substack{x_1=\sin(x_1^*)\cos(x_2^*) \\ x_2=\sin(x_1^*)\sin(x_2^*)}} = \left[\begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1-(x_1^2+x_2^2)} \end{bmatrix}\right]\Bigg|_{\substack{x_1=\sin(x_1^*)\cos(x_2^*) \\ x_2=\sin(x_1^*)\sin(x_2^*)}} = \begin{bmatrix} \sin(x_1^*)\cos(x_2^*) \\ \sin(x_1^*)\sin(x_2^*) \\ \cos(x_1^*) \end{bmatrix} = \mathbf{z}^*(x_1^*, x_2^*).$$

The basis vectors $\mathbf{z}^*_{x_1^*}$ and $\mathbf{z}^*_{x_2^*}$ in terms of the basis vectors $\mathbf{z}_{x_1}$ and $\mathbf{z}_{x_2}$ are given by

$$\begin{aligned} \mathbf{z}^*_{x_1^*} &= \frac{\partial \mathbf{z}^*(x_1^*, x_2^*)}{\partial x_1^*} = \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_1}\frac{\partial x_1}{\partial x_1^*} + \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_2}\frac{\partial x_2}{\partial x_1^*} = \mathbf{z}_{x_1}\frac{\partial x_1}{\partial x_1^*} + \mathbf{z}_{x_2}\frac{\partial x_2}{\partial x_1^*} \\ \mathbf{z}^*_{x_2^*} &= \frac{\partial \mathbf{z}^*(x_1^*, x_2^*)}{\partial x_2^*} = \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_1}\frac{\partial x_1}{\partial x_2^*} + \frac{\partial \mathbf{z}(x_1, x_2)}{\partial x_2}\frac{\partial x_2}{\partial x_2^*} = \mathbf{z}_{x_1}\frac{\partial x_1}{\partial x_2^*} + \mathbf{z}_{x_2}\frac{\partial x_2}{\partial x_2^*}. \end{aligned}$$

In matrix form we can write

$$\begin{bmatrix} \mathbf{z}^*_{x_1^*} \\ \mathbf{z}^*_{x_2^*} \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial x_1}{\partial x_1^*} & \dfrac{\partial x_2}{\partial x_1^*} \\ \dfrac{\partial x_1}{\partial x_2^*} & \dfrac{\partial x_2}{\partial x_2^*} \end{bmatrix}}_{\text{Change of Basis Matrix}} \begin{bmatrix} \mathbf{z}_{x_1} \\ \mathbf{z}_{x_2} \end{bmatrix} = \begin{bmatrix} \cos(x_1^*)\cos(x_2^*) & \cos(x_1^*)\sin(x_2^*) \\ -\sin(x_1^*)\sin(x_2^*) & \sin(x_1^*)\cos(x_2^*) \end{bmatrix} \begin{bmatrix} \mathbf{z}_{x_1} \\ \mathbf{z}_{x_2} \end{bmatrix}.$$

Let $a_1^*\mathbf{z}^*_{x_1^*} + a_2^*\mathbf{z}^*_{x_2^*}$ be a tangent vector from the tangent space $\mathbf{T}_p(\mathbf{S}^2)$ represented in the spherical coordinate patch. What is the representation of this vector in the northern hemisphere coordinate patch? The relationships between basis vectors $\left\{\mathbf{z}^*_{x_1^*}, \mathbf{z}^*_{x_2^*}\right\}$ and the basis vectors $\{\mathbf{z}_{x_1}, \mathbf{z}_{x_2}\}$ may showns as

$$\begin{array}{cc} \mathbf{z}^*_{x_1^*} & \mathbf{z}^*_{x_2^*} \\ \| & \| \\ \dfrac{\partial x_1}{\partial x_1^*}\mathbf{z}_{x_1} & \dfrac{\partial x_1}{\partial x_2^*}\mathbf{z}_{x_1} \\ + & + \\ \dfrac{\partial x_2}{\partial x_1^*}\mathbf{z}_{x_2} & \dfrac{\partial x_2}{\partial x_2^*}\mathbf{z}_{x_2} \end{array}$$

With $a_1^*\mathbf{z}_{x_1^*} + a_2^*\mathbf{z}_{x_2^*} = a_1\mathbf{z}_{x_1} + a_2\mathbf{z}_{x_2}$ and, using the above relationships between basis vectors, we have the diagram

$$\begin{array}{ccccc}
 & & a_1^*\mathbf{z}_{x_1^*} & + & a_2^*\mathbf{z}_{x_2^*} \\
 & & & \| & \\
a_1\mathbf{z}_{x_1} & = & a_1^*\dfrac{\partial x_1}{\partial x_1^*}\mathbf{z}_{x_1} & + & a_2^*\dfrac{\partial x_1}{\partial x_2^*}\mathbf{z}_{x_1} \\
 & & + & & + \\
a_2\mathbf{z}_{x_2} & = & a_1^*\dfrac{\partial x_2}{\partial x_1^*}\mathbf{z}_{x_2} & + & a_2^*\dfrac{\partial x_2}{\partial x_2^*}\mathbf{z}_{x_2} \\
 & & & \| & \\
 & & a_1\mathbf{z}_{x_1} & + & a_2\mathbf{z}_{x_2}
\end{array}$$

From the diagram we have

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial x_1}{\partial x_1^*} & \dfrac{\partial x_1}{\partial x_2^*} \\ \dfrac{\partial x_2}{\partial x_1^*} & \dfrac{\partial x_2}{\partial x_2^*} \end{bmatrix}}_{\text{change of components matrix}} \begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} \cos(x_1^*)\cos(x_2^*) & -\sin(x_1^*)\sin(x_2^*) \\ \cos(x_1^*)\sin(x_2^*) & \sin(x_1^*)\cos(x_2^*) \end{bmatrix} \begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix}.$$

We derived these change of components matrix previously using a curve $(x_1^*(t), x_2^*(t))$ given in the spherical coordinates which has coordinates

$$(x_1(t), x_2(t)) = \Big(\sin(x_1^*(t)\cos(x_2^*(t)),\ \sin(x_1^*(t))\sin(x_2^*(t))\Big).$$

There we differentiated both sides to obtain

$$\begin{bmatrix} \dfrac{dx_1}{dt} \\ \dfrac{dx_2}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial x_1}{\partial x_1^*} & \dfrac{\partial x_1}{\partial x_2^*} \\ \dfrac{\partial x_2}{\partial x_1^*} & \dfrac{\partial x_2}{\partial x_2^*} \end{bmatrix}}_{\text{change of components matrix}} \begin{bmatrix} \dfrac{dx_1^*}{dt} \\ \dfrac{dx_2^*}{dt} \end{bmatrix} = \begin{bmatrix} \cos(x_1^*)\cos(x_2^*) & -\sin(x_1^*)\sin(x_2^*) \\ \cos(x_1^*)\sin(x_2^*) & \sin(x_1^*)\cos(x_2^*) \end{bmatrix} \begin{bmatrix} \dfrac{dx_1^*}{dt} \\ \dfrac{dx_2^*}{dt} \end{bmatrix}.$$

To summarize, we write $a_1^*\mathbf{z}_{x_1^*} + a_2^*\mathbf{z}_{x_2^*} = a_1\mathbf{z}_{x_1} + a_2\mathbf{z}_{x_2}$ to mean the same vector represented in two different coordinate systems. Their components and basis vectors must be related as shown above for this to hold.

Exercise 1 *Northern Hemisphere to Spherical Coordinates*

Show that the spherical coordinates of a tangent vector (a_1^*, a_2^*) are related to its northern hemisphere components (a_1, a_2) by

$$\begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial x_1^*}{\partial x_1} & \dfrac{\partial x_1^*}{\partial x_2} \\ \dfrac{\partial x_2^*}{\partial x_1} & \dfrac{\partial x_2^*}{\partial x_2} \end{bmatrix}}_{\text{Change of components matrix}} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Show that the basis vectors $(\mathbf{z}^*_{x_1^*}, \mathbf{z}^*_{x_2^*})$ in spherical coordinates are related to the basis vectors $(\mathbf{z}_{x_1}, \mathbf{z}_{x_2})$ in the northern hemisphere by

$$\begin{bmatrix} \mathbf{z}_{x_1} \\ \mathbf{z}_{x_2} \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial x_1^*}{\partial x_1} & \dfrac{\partial x_2^*}{\partial x_1} \\ \dfrac{\partial x_1^*}{\partial x_2} & \dfrac{\partial x_2^*}{\partial x_2} \end{bmatrix}}_{\text{Change of basis matrix}} \begin{bmatrix} \mathbf{z}^*_{x_1^*} \\ \mathbf{z}^*_{x_2^*} \end{bmatrix}.$$

Example 1 *Series Connected DC Motor*
Recall the equations describing the series connected DC motor given by

$$\begin{aligned}\frac{di}{dt} &= -\frac{R}{L}i - \frac{K_bL_f}{L}i\omega + \frac{V_S}{L}\\ \frac{d\omega}{dt} &= \frac{K_TL_f}{J}i^2 - \frac{\tau_L}{J}\\ \frac{d\theta}{dt} &= \omega.\end{aligned}$$

Set $z_1 = \theta, z_2 = \omega, z_3 = i$, and $u = V_S/L$ to obtain

$$\frac{d}{dt}\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix} = \underbrace{\begin{bmatrix} z_2\\ c_1z_3^2 - d\\ -c_2z_3 - c_3z_3z_2\end{bmatrix}}_{f(z)} + \underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix}}_{g(z)}u$$

with $c_1 \triangleq K_TL_f/J, c_2 \triangleq R/L$, $c_3 \triangleq K_bL_f/L$, and $d \triangleq \tau_L/J$. Let

$$\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix} = \begin{bmatrix}\theta\\ \omega\\ i\end{bmatrix}$$

be considered as points in the manifold $\mathbf{E}^3$ and consider two different coordinate systems for it. The first coordinate system is the Cartesian coordinate map $\mathbf{z}(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbf{E}^3$ is

$$\mathbf{z}(x_1,x_2,x_3) = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix} \in \mathbf{E}^3.$$

In Example **??** of Chapter **??** it was shown that the change of coordinates given by

$$\begin{aligned}x_1^* &= x_1 = z_1\\ x_2^* &= x_2 = z_2\\ x_3^* &= c_1x_3^2 - d = c_1z_3^2 - d\end{aligned}$$

was the feedback linearizing transformation. So we take the inverse of this as the second coordinate system $\varphi^{*-1}(x^*) = \mathbf{z}^*(x^*) : \mathcal{D} \to \mathcal{U} \subset \mathbf{E}^3$ given by

$$\varphi^{*-1}(x^*) = \mathbf{z}^*(x^*) = \begin{bmatrix}x_1^*\\ x_2^*\\ \sqrt{(x_3^*+d)/c_1}\end{bmatrix} = \begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix} \in \mathbf{E}^3$$

where $\mathcal{D} = \{x^* \in \mathbb{R}^3 |\ x_3^* + d > 0\}$ and $\mathcal{U} = \{\mathbf{z} \in \mathbf{E}^3 |\ z_3 > 0\}$. Then $\varphi^* = \mathbf{z}^{*-1} : \mathcal{U} \to \mathcal{D}$ is written as

$$\varphi^*\left(\begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix}\right) = (z_1, z_2, c_1z_3^2 - d) = (x_1^*, x_2^*, x_3^*) \in \mathbb{R}^3.$$

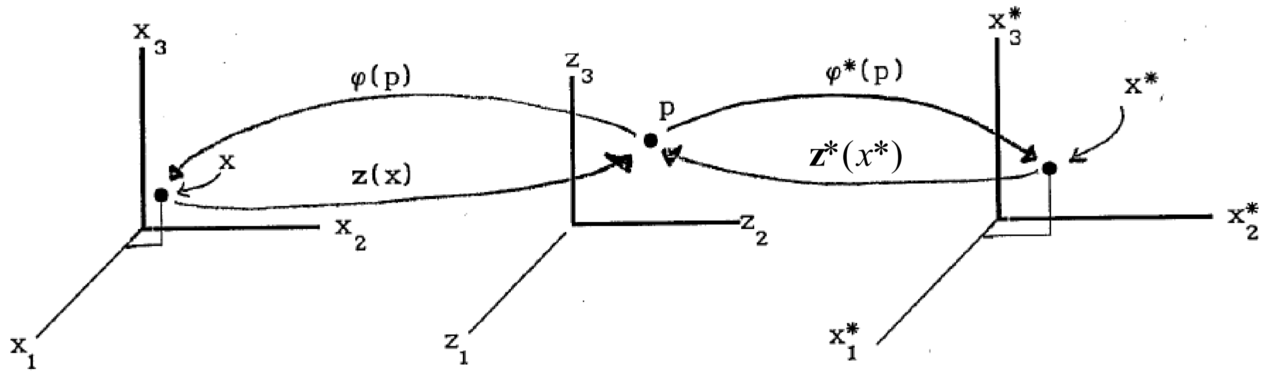

FIGURE 1.7. Coordinate systems for the series connected DC motor.

For each $u \in \mathbb{R}$ the tangent vector specified by the series connected DC motor is

$$\frac{dz}{dt} = f(z) + g(z)u \in \mathbf{T}_p(\mathbf{E}^3)$$

where $p = \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}^T$. The representation of the tangent vector $f(z)$ in Cartesian coordinates is

$$f_1(x)\mathbf{z}_{x_1} + f_2(x)\mathbf{z}_{x_2} + (f_3(x) + u)\mathbf{z}_{x_3}$$

with components $f_1(x) = x_2, f_2(x) = c_1x_3^2 - d, f_3(x) = -c_2x_3 - c_3x_3x_2$ and basis vectors

$$\mathbf{z}_{x_1} = \frac{\partial}{\partial x_1}\mathbf{z}(x_1, x_2, x_3) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{z}_{x_2} = \frac{\partial}{\partial x_2}\mathbf{z}(x_1, x_2, x_3) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{z}_{x_3} = \frac{\partial}{\partial x_3}\mathbf{z}(x_1, x_2, x_3) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

To determine the representation of this tangent vector in the x^* coordinates first note that the basis vectors in this coordinate system are

$$\mathbf{z}^*_{x_1^*} = \frac{\partial}{\partial x_1^*}\mathbf{z}^*(x_1^*, x_2^*, x_3^*) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{z}^*_{x_2^*} = \frac{\partial}{\partial x_2^*}\mathbf{z}^*(x_1^*, x_2^* x_3^*) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{z}^*_{x_3} = \frac{\partial}{\partial x_3^*}\mathbf{z}^*(x_1^*, x_2^*, x_3^*) = \begin{bmatrix} 0 \\ 0 \\ \dfrac{1}{2c_1\sqrt{(x_3^*+d)/c_1}} \end{bmatrix}.$$

The change of coordinates $T = \varphi^* \circ \varphi^{-1} : \{(x_1, x_2, x_3) \subset \mathbb{R}^3 | \ x_3 > 0\} \to \mathcal{D}$ is

$$(x_1^*, x_2^*, x_3^*) = \varphi^* \circ \varphi^{-1}(x_1, x_2, x_3) = \varphi^*(\varphi^{-1}(x_1, x_2, x_3)) = (x_1, x_2, c_1x_3^2 - d)$$

or

$$\begin{aligned} x_1^* &= T_1(x_1) = x_1 \\ x_2^* &= T_2(x) = x_2 \\ x_3^* &= T_3(x) = c_1x_3^2 - d. \end{aligned}$$

Then

$$
\begin{aligned}
\begin{bmatrix} f_1^*(x^*) \\ f_2^*(x^*) \\ f_3^*(x^*) \end{bmatrix}
&= \begin{bmatrix} \frac{\partial x_1^*}{\partial x_1} & \frac{\partial x_1^*}{\partial x_2} & \frac{\partial x_1^*}{\partial x_3} \\ \frac{\partial x_2^*}{\partial x_1} & \frac{\partial x_2^*}{\partial x_2} & \frac{\partial x_2^*}{\partial x_3} \\ \frac{\partial x_2^*}{\partial x_3} & \frac{\partial x_2^*}{\partial x_3} & \frac{\partial x_3^*}{\partial x_3} \end{bmatrix}
\begin{bmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2c_1x_3 \end{bmatrix}
\begin{bmatrix} x_2 \\ c_1x_3^2 - d \\ -c_2x_3 - c_3x_3x_2 \end{bmatrix} \Bigg|_{x=\varphi(\varphi^{*-1}(x^*))=\begin{bmatrix} x_1^* \\ x_2^* \\ \sqrt{(x_3^*+d)/c_1} \end{bmatrix}} \\
&= \begin{bmatrix} x_2^* \\ x_3^* \\ -2c_2(x_3^*+d) - 2c_3x_2^*(x_3^*+d) \end{bmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\begin{bmatrix} g_1^*(x^*) \\ g_2^*(x^*) \\ g_3^*(x^*) \end{bmatrix}
&= \begin{bmatrix} \frac{\partial x_1^*}{\partial x_1} & \frac{\partial x_1^*}{\partial x_2} & \frac{\partial x_1^*}{\partial x_3} \\ \frac{\partial x_2^*}{\partial x_1} & \frac{\partial x_2^*}{\partial x_2} & \frac{\partial x_2^*}{\partial x_3} \\ \frac{\partial x_2^*}{\partial x_3} & \frac{\partial x_2^*}{\partial x_3} & \frac{\partial x_3^*}{\partial x_3} \end{bmatrix}
\begin{bmatrix} g_1(x) \\ g_2(x) \\ g_3(x) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2c_1x_3 \end{bmatrix}
\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Bigg|_{x=\varphi(\varphi^{*-1}(x^*))=\begin{bmatrix} x_1^* \\ x_2^* \\ \sqrt{(x_3^*+d)/c_1} \end{bmatrix}} \\
&= \begin{bmatrix} 0 \\ 0 \\ 2c_1\sqrt{(x_3^*+d)/c_1} \end{bmatrix}.
\end{aligned}
$$

Then

$$
\begin{aligned}
f_1(x^*)\mathbf{z}_{x_1^*} + f_2(x^*)\mathbf{z}_{x_2^*} + f_3(x^*)\mathbf{z}_{x_3^*}
&= x_2^* \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_3^* \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \Big(-2c_2(x_3^*+d) - 2c_3x_2^*(x_3^*+d)\Big) \begin{bmatrix} 0 \\ 0 \\ \dfrac{1}{2c_1\sqrt{(x_3^*+d)/c_1}} \end{bmatrix} \\
&= \begin{bmatrix} x_2^* \\ x_3^* \\ (-c_2 - c_3x_2^*)\sqrt{(x_3^*+d)/c_1} \end{bmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
g_1(x^*)\mathbf{z}_{x_1^*} + g_2(x^*)\mathbf{z}_{x_2^*} + g_3(x^*)\mathbf{z}_{x_3^*}
&= 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 2c_1\sqrt{(x_3^*+d)/c_1} \begin{bmatrix} 0 \\ 0 \\ \dfrac{1}{2c_1\sqrt{(x_3^*+d)/c_1}} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.
\end{aligned}
$$

Exercise 2 Show that

$$\begin{aligned}
&f_1(x^*)\mathbf{z}_{x_1^*} + f_2(x^*)\mathbf{z}_{x_2^*} + f_3(x^*)\mathbf{z}_{x_3^*} + \Big(g_1(x^*)\mathbf{z}_{x_1^*} + g_2(x^*)\mathbf{z}_{x_2^*} + g_3(x^*)\mathbf{z}_{x_3^*}\Big)u \\
&\qquad = \left[\begin{array}{c} x_2^* \\ x_3^* \\ (-c_2 - c_3 x_2^*)\sqrt{(x_3^* + d)/c_1} + u \end{array}\right]_{x^*=\varphi^*\varphi^{-1}(x))} \\
&\qquad = \left[\begin{array}{c} x_2 \\ c_1 x_3^2 - d \\ -c_2 x_3 - c_3 x_3 x_2 + u \end{array}\right] \\
&\qquad = f_1(x)\mathbf{z}_{x_1} + f_2(x)\mathbf{z}_{x_2} + (f_3(x) + u)\mathbf{z}_{x_3} + \Big(g_1(x)\mathbf{z}_{x_1} + g_2(x)\mathbf{z}_{x_2} + g_3(x)\mathbf{z}_{x_3}\Big)u
\end{aligned}$$

1.2 How Mathematicians View Tangent Vectors

We now introduce the modern (abstract) approach to differential geometry. Once again consider tangent vectors on $\mathbf{S}^2$ using the northern hemisphere coordinate system. The northern hemisphere chart $\mathbf{z}(x_1, x_2) : \mathcal{D}_1 \to \mathcal{U}_1 \subset \mathbf{S}^2$ is

$$\mathbf{z}(x_1, x_2) = \left[\begin{array}{c} x_1 \\ x_2 \\ \sqrt{1 - (x_1^2 + x_2^2)} \end{array}\right] \in \mathbf{S}^2$$

where $\mathcal{D}_1 = \big\{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \le x_1^2 + x_2^2 < 1\big\}$ and $\mathcal{U}_1 \triangleq \big\{\mathbf{z} \in \mathbf{S}^2 |\ z_3 > 0\big\}$. With $p = \mathbf{z}(x_1, x_2)$, a set of basis vectors for the tangent space $\mathbf{T}_p(\mathbf{S}^2)$ is

$$\mathbf{z}_{x_1} = \frac{\partial}{\partial x_1}\mathbf{z}(x_1, x_2) = \left[\begin{array}{c} 1 \\ 0 \\ \dfrac{-x_1}{\sqrt{1 - (x_1^2 + x_2^2)}} \end{array}\right],\ \mathbf{z}_{x_2} = \frac{\partial}{\partial x_2}\mathbf{z}(x_1, x_2) = \left[\begin{array}{c} 0 \\ 1 \\ \dfrac{-x_2}{\sqrt{1 - (x_1^2 + x_2^2)}} \end{array}\right]$$

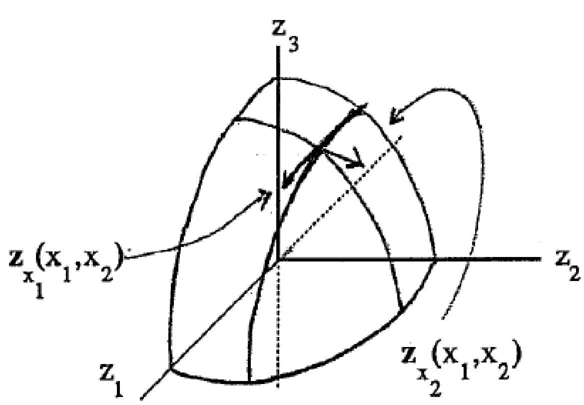

FIGURE 1.8. Northern hemisphere coordinate chart for $\mathbf{S}^2$ with the tangent vectors $\vec{z}_{x_1}$ and $\vec{z}_{x_2}$.

As previously mentioned, any tangent vector to the manifold for a point $p \in \mathcal{U}_1$ is a linear combination of $\mathbf{z}_{x_1}$ and $\mathbf{z}_{x_2}$. However, there is a problem with this description of a tangent space. To explain, imagine

there are two-dimensional (2-D) people living on the surface $\mathbf{S}^2$. By 2-D people is meant that they can move anywhere on the surface $\mathbf{S}^2$, but not off of it. In fact, they are not even aware of that the direction perpendicular to $\mathbf{S}^2$ exists. In this case the tangent vectors $\mathbf{z}_{x_1}$ and $\mathbf{z}_{x_2}$ don't make sense as they stick out off the surface as illustrated in Figure 1.8. The 2-D people cannot draw $\mathbf{z}_{x_1}$ and $\mathbf{z}_{x_2}$ as they cannot go off the surface. We need to reconcile this problem so that tangent vectors are part of the manifold $\mathbf{S}^2$ in such a way to that the 2-D people have knowledge of them. This is done by redefining tangent vectors in such a way that the 2-D people can use them and that they are, in some sense, equivalent to the old definition.
Remark Instead of the manifold $\mathbf{S}^2$ consider the manifold $\mathbf{E}^2$. In this case the Cartesian coordinate system $\mathbf{z}(x_1, x_2) : \mathbb{R}^2 \to \mathbf{E}^2$ is

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with tangent vectors at each point of the manifold given by

$$\mathbf{z}_{x_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{z}_{x_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

which lie in the manifold.

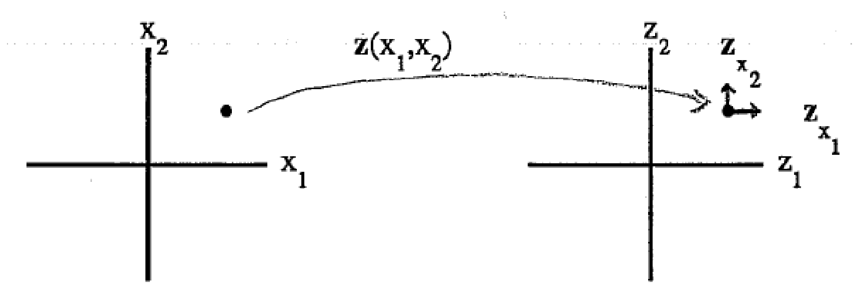

FIGURE 1.9. Euclidean space $\mathbf{E}^2$.

In this case 2-D people living on $\mathbf{E}^2$ can use $\{\mathbf{z}_{x_1}, \mathbf{z}_{x_2}\}$.

Digression Gradients

Before going forward with the modern definition of tangent vectors we digress to discuss gradients. Let $\mathcal{C}^k(\mathbb{R}^3)$ be the set of k-times differentiable functions on $\mathbb{R}^3$ and $\mathcal{C}^\infty(\mathbb{R}^3)$ be all infinitely differentiable functions on $\mathbb{R}^3$. Let $\mathcal{C}^k(p)$ and $\mathcal{C}^\infty(p)$ denote, respectively, all k-times differentiable and infinitely differentiable functions in a *neighborhood*[1] of a point p. For example, with $p = (x_{01}, x_{02}, x_{03}) \in \mathbb{R}^3$ let $h(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R}$ be in $\mathcal{C}^1(p)$ so that $\partial h(x)/\partial x_1, \partial h(x)/\partial x_2, \partial h(x)/\partial x_2$ all exist and are continuous in a neighborhood of p. The *gradient* is the operator $d : \mathcal{C}^1(p) \to \mathbb{R}^{1\times 3}$ given by

$$dh = \begin{bmatrix} \dfrac{\partial h(x)}{\partial x_1} & \dfrac{\partial h(x)}{\partial x_2} & \dfrac{\partial h(x)}{\partial x_3} \end{bmatrix}.$$

[1] Just take neighborhood of p to mean an open set that contains p.

Let $(x_1(t), x_2(t), x_3(t))$ be a smooth (differentiable) curve in $\mathbb{R}^3$ with $(x_1(0), x_2(0), x_3(0)) = (x_{01}, x_{02}, x_{03}) = p$. The scalar function $h \circ x(t) = h(x(t))$ has derivative

$$\begin{aligned}\left.\frac{dh}{dt}\right|_{t=0} &= \left.\frac{\partial h(x)}{\partial x_1}\right|_p \left.\frac{dx_1}{dt}\right|_{t=0} + \left.\frac{\partial h(x)}{\partial x_2}\right|_p \left.\frac{dx_2}{dt}\right|_{t=0} + \left.\frac{\partial h(x)}{\partial x_3}\right|_p \left.\frac{dx_3}{dt}\right|_{t=0} = \left.\begin{bmatrix} \frac{\partial h(x)}{\partial x_1} & \frac{\partial h(x)}{\partial x_2} & \frac{\partial h(x)}{\partial x_3} \end{bmatrix}\right|_p \left.\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix}\right|_{t=0} \\ &= \left.\left\langle dh, \frac{dx}{dt} \right\rangle\right|_{t=0}\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the dual product. $\left.\frac{dh}{dt}\right|_{t=0}$ is the rate of change of h at p in the direction $\left.\frac{dx}{dt}\right|_{t=0} = \left.(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt})\right|_{t=0}$. If $\left.\frac{dx}{dt}\right|_{t=0}$ points in the same direction as dh then the rate of change is at a maximum. Now consider the mapping (function) that takes $h \in \mathcal{C}^\infty(p)$ to the real number $\left.\left\langle dh, \frac{dx}{dt} \right\rangle\right|_{t=0}$. Specifically, define $X_p(h) : \mathcal{C}^\infty(p) \to \mathbb{R}$ by

$$X_p(h) \triangleq \left.\left\langle dh, \frac{dx}{dt} \right\rangle\right|_{t=0}.$$

For example, let $h(x_1, x_2, x_3) = x_1$ so

$$X_p(h) \triangleq \left.\left\langle dh, \frac{dx}{dt} \right\rangle\right|_{t=0} = \left\langle \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \left.\frac{dx}{dt}\right|_{t=0} \right\rangle = \left.\frac{dx_1}{dt}\right|_{t=0}.$$

Similarly, $h(x_1, x_2, x_3) = x_2$ gives $X_p(h) = \left.\frac{dx_2}{dt}\right|_{t=0}$ and $h(x_1, x_2, x_3) = x_3$ gives $X_p(h) = \left.\frac{dx_3}{dt}\right|_{t=0}$.

This mapping $X_p(h) : \mathcal{C}^\infty(p) \to \mathbb{R}$ is completely determined by the three numbers $\frac{dx_1}{dt}, \frac{dx_2}{dt}$, and $\frac{dx_3}{dt}$ (assumed to evaluated at $t = 0$). In other words, suppose there are three numbers a, b, c such that for all $h \in \mathcal{C}^\infty(p)$ we have

$$\begin{bmatrix} \frac{\partial h(x)}{\partial x_1} & \frac{\partial h(x)}{\partial x_2} & \frac{\partial h(x)}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} = \begin{bmatrix} \frac{\partial h(x)}{\partial x_1} & \frac{\partial h(x)}{\partial x_2} & \frac{\partial h(x)}{\partial x_3} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Then $a = \frac{dx_1}{dt}, b = \frac{dx_2}{dt}$, and $c = \frac{dx_3}{dt}$. To show this simply let $h(x)$ equal successively $h(x) = x_1, h(x) = x_2$, and $h(x) = x_3$.

A New Interpretation of the Tangent Vector on $\mathbf{S}^2$

We now go to the manifold $\mathbf{S}^2$ with the northern hemisphere coordinate chart and see how this new notion of a tangent vector can be used by 2-D people on $\mathbf{S}^2$. A 2-D person on $\mathbf{S}^2$ is only aware of functions defined on $\mathbf{S}^2$. The northern hemisphere patch $\mathbf{z}(x_1, x_2) : \mathcal{D}_1 \to \mathcal{U}_1 \subset \mathbf{S}^2$ is

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1 - (x_1^2 + x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D}_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \le x_1^2 + x_2^2 < 1\right\}$ and $\mathcal{U}_1 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 |\ z_3 > 0\right\}$.

Let $\mathbf{z}_0 = \mathbf{z}(x_{01}, x_{02}) = \left[\begin{array}{ccc} x_{01} & x_{02} & \sqrt{1-(x_{01}^2+x_{02}^2)} \end{array}\right]^T$ so it has local coordinates (x_{01}, x_{02}). The coordinate mapping $\varphi\left(\left[\begin{array}{ccc} z_1 & z_2 & z_3 \end{array}\right]^T\right) = (z_1, z_2)$ is only defined for those $\left[\begin{array}{ccc} z_1 & z_2 & z_3 \end{array}\right]^T$ with $z_1^2+z_2^2+z_3^2 = 1$, that is, $\left[\begin{array}{ccc} z_1 & z_2 & z_3 \end{array}\right]^T$ must be on the manifold $\mathbf{S}^2$. Let $h(p) = h(z_1, z_2, z_3)$ be a function defined in a neighborhood of $p = \mathbf{z}_0$. By this is meant $h(z)$ is defined only for $\mathbf{z} = \left[\begin{array}{ccc} z_1 & z_2 & z_3 \end{array}\right]^T$ with $z_1^2 + z_2^2 + z_3^2 = 1$ which are close to $\mathbf{z}_0 = \mathbf{z}(x_{01}, x_{02})$.

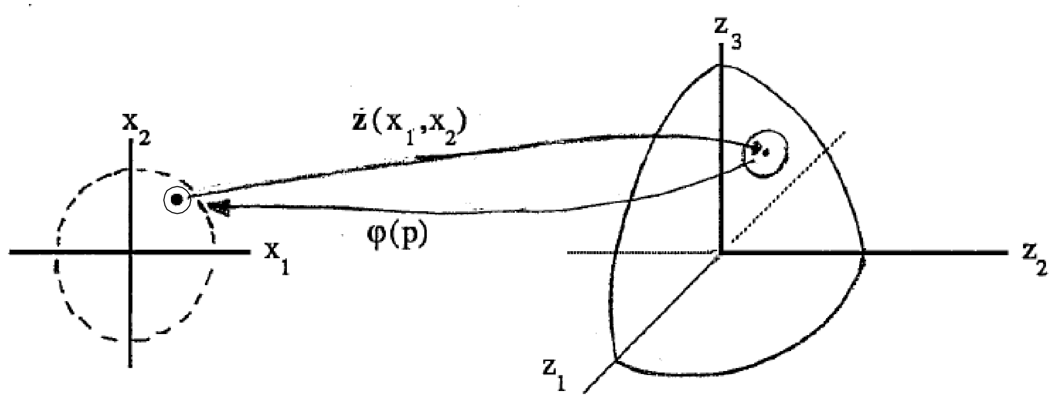

FIGURE 1.10. Northern hemisphere coordinate chart for $\mathbf{S}^2$. Left: Neighborhood of $x_0 = (x_{01}, x_{02})$. Right: Neighborhod of $\mathbf{z}_0 = \mathbf{z}(x_{01}, x_{02})$.

Now to the tangent vector. Let $(x_1(t), x_2(t))$ be a curve in $\mathcal{D}_1$ in a neighborhood of $x_0 = (x_{01}, x_{02})$ with $(x_1(0), x_2(0)) = (x_{01}, x_{02})$. Then $\mathbf{z}(x_1(t), x_2(t))$ is a curve on $\mathbf{S}^2$ going through $p = \left[\begin{array}{ccc} x_{01} & x_{02} & \sqrt{1-(x_{01}^2+x_{02}^2)} \end{array}\right]^T$ at $t = 0$. Define

$$h(t) \triangleq h(z_1(x_1(t), x_2(t)), z_2(x_1(t), x_2(t)), z_3(x_1(t), x_2(t)))$$

which is an ordinary function of time. The rate of change of h at p determined by the curve $\mathbf{z}(x(t), x(t))$ is

$$\begin{aligned} \frac{dh}{dt}_{|t=0} = \frac{d}{dt} h \circ \mathbf{z}(x_1(t), x_2(t))\Big|_{t=0} &= \frac{d}{dt} h(\mathbf{z}(x_1(t), x_2(t)))\Big|_{t=0} \\ &= \frac{\partial h(\mathbf{z}(x_1, x_2))}{\partial x_1}\Big|_{(x_{01}, x_{02})} \frac{dx_1}{dt}_{|t=0} + \frac{\partial h(\mathbf{z}(x_1, x_2))}{\partial x_2}\Big|_{(x_{01}, x_{02})} \frac{dx_2}{dt}_{|t=0}. \end{aligned}$$

Note that we do *not* write

$$\begin{aligned} \frac{dh}{dt}_{|t=0} &= \left[\begin{array}{ccc} \frac{\partial h(z)}{\partial z_1} & \frac{\partial h(z)}{\partial z_2} & \frac{\partial h(z)}{\partial z_3} \end{array}\right] \left[\begin{array}{cc} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} \\ \frac{\partial z_3}{\partial x_1} & \frac{\partial z_3}{\partial x_2} \end{array}\right]_{(x_{01}, x_{02})} \left[\begin{array}{c} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{array}\right]_{t=0} \\ &= dh \left[\begin{array}{cc} \mathbf{z}_{x_1} & \mathbf{z}_{x2} \end{array}\right] \left[\begin{array}{c} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{array}\right]_{|t=0} \\ &= \left\langle dh, \frac{d\mathbf{z}(t)}{dt} \right\rangle \text{ where } \frac{d\mathbf{z}(t)}{dt} = \mathbf{z}_{x_1}\frac{dx_1}{dt} + \mathbf{z}_{x_2}\frac{dx_2}{dt}. \end{aligned}$$

This is because the partial derivatives $\dfrac{\partial h(z)}{\partial z_1}, \dfrac{\partial h(z)}{\partial z_2}, \dfrac{\partial h(z)}{\partial z_3}$ as well as "tangent vectors" $\mathbf{z}_{x_1}, \mathbf{z}_{x_2}$ do not make sense! To explain, $\dfrac{\partial h(z)}{\partial z_2}$ is calculated as the limit

$$\frac{\partial h(z)}{\partial z_2} = \lim_{\Delta z \to 0} \frac{h(z_{01}, z_{02} + \Delta z, z_{03}) - h(z_{01}, z_{02}, z_{03})}{\Delta z}.$$

As illustrated in Figure 1.11 the point $[z_{01}, z_{02} + \Delta z, z_{03}]^T$ is not on $\mathbf{S}^2$ so $h(z_{01}, z_{02} + \Delta z, z_{03})$ is not defined (does not exist) for $\Delta z \neq 0$.

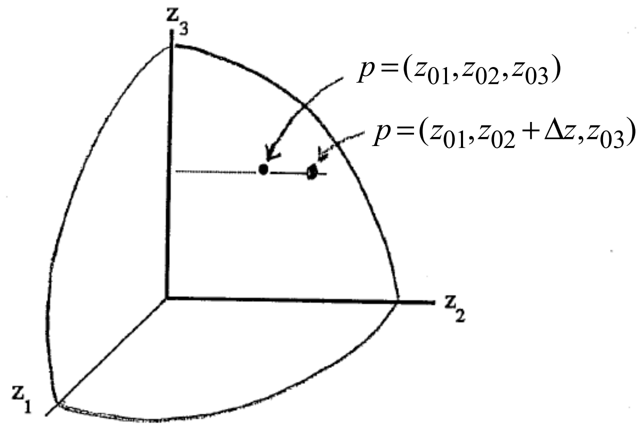

FIGURE 1.11. Calculation of $\dfrac{\partial h(z)}{\partial z_2} = \lim_{\Delta z \to 0} \frac{h(z_{01}, z_{02} + \Delta z, z_{03}) - h(z_{01}, z_{02}, z_{03})}{\Delta z}$.

In contrast,

$$\frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_1} = \frac{\partial h(\mathbf{z}(x_1, x_2))}{\partial x_1} = \frac{\partial}{\partial x_1} h(z_1(x_1, x_2), z_2(x_1, x_2), z_3(x_1, x_2))$$

is well defined as

$$\begin{aligned} &\frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_1} \\ &= \lim_{\Delta x_1 \to 0} \frac{h(z_1(x_1 + \Delta x_1, x_2), z_2(x_1 + \Delta x_1, x_2), z_3(x_1 + \Delta x_1, x_2)) - h(z_1(x_1, x_2), z_2(x_1, x_2), z_3(x_1, x_2))}{\Delta x_1}. \end{aligned}$$

In summary a 2-D person $\mathbf{S}^2$ is not aware of vectors like $\mathbf{z}_{x_1}$ $\mathbf{z}_{x2}$ as they stick off the manifold and only has knowledge of functions defined on $\mathbf{S}^2$.

Definition 1 *Differentiable Functions on* $\mathbf{S}^2$
Let $p = \mathbf{z}(x_{01}, x_{02}) \in \mathbf{S}^2$ for some coordinate system $\mathbf{z}(x_1, x_2)$. A function h on $\mathbf{S}^2$ is differentiable at $p \in \mathbf{S}^2$ if for every such coordinate system

$$h \circ \mathbf{z}(x_1, x_2)) = h(\mathbf{z}(x_1, x_2)))$$

is a differentiable function x_1 and x_2.

Definition 2 *Tangent Vectors on* $\mathbf{S}^2$

Let $\mathcal{F}(\mathbf{S}^2)$ denote the differentiable functions on $\mathbf{S}^2$, $\mathbf{z}(x_1, x_2) : \mathcal{D} \to \mathcal{U} \subset \mathbf{S}^2$ a coordinate system for $\mathbf{S}^2$, and $(x_1(t), x_2(t))$ be a curve in $\mathcal{D}$ with $(x_1(0), x_2(0)) = (x_{01}, x_{02})$. The tangent vector at $p = \mathbf{z}(x_{01}, x_{02})$ determined by the curve $\mathbf{z}(x_1(t), x_2(t))$ is the mapping $\mathbf{z}_p : \mathcal{F}(\mathbf{S}^2) \to \mathbb{R}$ given by

$$\begin{aligned}
\mathbf{z}_p : h \to \mathbf{z}_p(h) &\triangleq \frac{d}{dt}(h \circ \mathbf{z}(x_1(t), x_2(t)))\Big|_{t=0} \\
&= \frac{d}{dt} h(\mathbf{z}(x_1(t), x_2(t)))\Big|_{t=0} \\
&= \frac{\partial h \circ \mathbf{z}(x_1, x_2)}{\partial x_1}\Big|_{(x_{01}, x_{02})} \frac{dx_1}{dt}_{|t=0} + \frac{\partial h \circ \mathbf{z}(x_1, x_2)}{\partial x_2}\Big|_{(x_{01}, x_{02})} \frac{dx_2}{dt}_{|t=0}.
\end{aligned}$$

The two mappings

$$\frac{\partial}{\partial x_1} : h \in \mathcal{F}(\mathbf{S}^2) \to \frac{\partial h \circ \mathbf{z}(x_1, x_2)}{\partial x_1} \in \mathbb{R}$$

$$\frac{\partial}{\partial x_2} : h \in \mathcal{F}(\mathbf{S}^2) \to \frac{\partial h \circ \mathbf{z}(x_1, x_2)}{\partial x_2} \in \mathbb{R}$$

are *basis vectors* for the tangent space $\mathbf{T}_p(\mathbf{S}^2)$.

$\frac{dx_1}{dt}_{|t=0}$ and $\frac{dx_2}{dt}_{|t=0}$ are the *components* of the tangent vector.

Recall $p = \mathbf{z}(x_1, x_2)$ and $\varphi(p) = (x_1, x_2)$ are inverses of each other. The northern hemisphere patch $\varphi^{-1}(x_1, x_2) = \mathbf{z}(x_1, x_2) : \mathcal{D}_1 \to \mathcal{U}_1 \subset \mathbf{S}^2$ is

$$\varphi^{-1}(x_1, x_2) = \mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1 - (x_1^2 + x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D}_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \le x_1^2 + x_2^2 < 1\right\}$ and $\mathcal{U}_1 \triangleq \left\{\mathbf{z} \in \mathbf{S}^2 |\ z_3 > 0\right\}$. We have $\varphi : \mathcal{U}_1 \subset \mathbf{S}^2 \to \mathcal{D}_1$ given by

$$\varphi(p) = \left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}\right) = (z_1, z_2).$$

Mathematicians use the notation $\varphi^{-1}(x_1, x_2)$ instead of $\mathbf{z}(x_1, x_2)$! In this case we write the basis vectors as

$$\frac{\partial}{\partial x_1} : h \in \mathcal{F}(\mathbf{S}^2) \to \frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_1} \in \mathbb{R}$$

$$\frac{\partial}{\partial x_2} : h \in \mathcal{F}(\mathbf{S}^2) \to \frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_2} \in \mathbb{R}.$$

A tangent vector $\mathbf{z}_p$ at a point $p = \varphi^{-1}(x_1, x_2) \in \mathbf{S}^2$ is the mapping $\mathbf{z}_p : \mathcal{F}(\mathbf{S}^2) \to \mathbb{R}$

$$\mathbf{z}_p : h \to \mathbf{z}_p(h) \triangleq a_1 \frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_1} + a_2 \frac{\partial h \circ \varphi^{-1}(x_1, x_2)}{\partial x_2}$$

where a_1 and a_2 are the components of $\mathbf{z}_p$.

A New Interpretation of the Tangent Vector on $\mathbf{E}^3$

We now look at this new formulation of the tangent vector on the manifold $\mathbf{E}^3$. To so, we consider the spherical coordinate system on $\mathbf{E}^3$. With

$$\begin{aligned}\mathcal{D} &= \left\{(x_1,x_2,x_3)\in\mathbb{R}^3 |\ x_1>0, 0<x_2<\pi, 0<x_3<2\pi\right\}\\ \mathcal{U} &\triangleq \left\{\mathbf{z}\in\mathbf{E}^3 |\ \text{If } z_1>0 \text{ the } z_2\neq 0\right\},\end{aligned}$$

recall the spherical coordinate system $\mathbf{z}(x_1,x_2,x_3):\mathcal{D}\to\mathcal{U}$ given by

$$\varphi^{-1}(x_1,x_2,x_3)=\mathbf{z}(x_1,x_2,x_3)=\begin{bmatrix} x_1\sin(x_2)\cos(x_3)\\ x_1\sin(x_2)\sin(x_3)\\ x_1\cos(x_2)\end{bmatrix}\in\mathbf{E}^3.$$

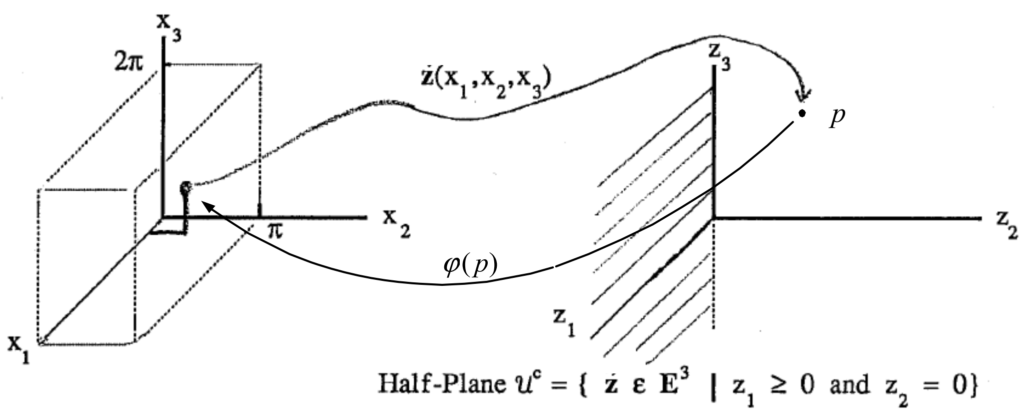

FIGURE 1.12. Spherical coordinates on $\mathbf{E}^3$.

Let $\mathcal{F}(\mathcal{U})$ denote the differentiable functions on $\mathcal{U}\subset\mathbf{E}^3$. Let $h(z_1,z_2,z_3)\in\mathcal{F}(\mathcal{U})$ with $p\in\mathcal{U}$. Let $(x_1(t),x_2(t),x_3(t))$ be a curve in $\mathcal{D}$ resulting in the curve $\varphi^{-1}(x_1(t),x_2(t),x_3(t))=\mathbf{z}(x_1(t),x_2(t),x_3(t))$ in $\mathcal{U}$ with $\varphi^{-1}(x_1(0),x_2(0),x_3(0))=p$. The tangent to the curve is the mapping $\mathbf{z}_p:\mathcal{F}(\mathcal{U})\to\mathbb{R}$ that takes $h\in\mathcal{F}(\mathcal{U})$ to the number

$$\begin{aligned}\mathbf{z}_p: h\to\mathbf{z}_p(h) &= \frac{d}{dt}h\circ\varphi^{-1}(x_1(t),x_2(t),x_3(t))\Big|_{t=0}\\ &= \frac{\partial h\circ\varphi^{-1}(x_1,x_2,x_3)}{\partial x_1}\frac{dx_1}{dt}_{|t=0}+\frac{\partial h\circ\varphi^{-1}(x_1,x_2,x_3)}{\partial x_2}\frac{dx_2}{dt}_{|t=0}+\frac{\partial h\circ\varphi^{-1}(x_1,x_2,x_3)}{\partial x_3}\frac{dx_3}{dt}_{|t=0}.\end{aligned}$$

The components of this tangent vector (mapping) are $\dfrac{dx_1}{dt}_{|t=0}$, $\dfrac{dx_2}{dt}_{|t=0}$, and $\dfrac{dx_3}{dt}_{|t=0}$.

A basis for the tangent space at p are the three mappings

$$\begin{aligned}\frac{\partial}{\partial x_1} &: h\to\frac{\partial}{\partial x_1}(h\circ\varphi^{-1})\\ \frac{\partial}{\partial x_2} &: h\to\frac{\partial}{\partial x_2}(h\circ\varphi^{-1})\\ \frac{\partial}{\partial x_3} &: h\to\frac{\partial}{\partial x_3}(h\circ\varphi^{-1}).\end{aligned}$$

For this particular manifold $\mathbf{E}^3$ we can expand the above tangent vector to obtain

$$\begin{aligned}
\mathbf{z}_p(h) &= \frac{d}{dt}(h\circ\varphi^{-1})(x_1(t),x_2(t),x_3(t))\Big|_{t=0} \\
&= \begin{bmatrix} \dfrac{\partial h}{\partial z_1} & \dfrac{\partial h}{\partial z_2} & \dfrac{\partial h}{\partial z_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_1}{\partial x_2} & \dfrac{\partial z_1}{\partial x_3} \\ \dfrac{\partial z_2}{\partial x_1} & \dfrac{\partial z_2}{\partial x_2} & \dfrac{\partial z_2}{\partial x_3} \\ \dfrac{\partial z_3}{\partial x_1} & \dfrac{\partial z_3}{\partial x_2} & \dfrac{\partial z_3}{\partial x_3} \end{bmatrix} \begin{bmatrix} \dfrac{dx_1}{dt} \\ \dfrac{dx_2}{dt} \\ \dfrac{dx_3}{dt} \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\partial h}{\partial z_1} & \dfrac{\partial h}{\partial z_2} & \dfrac{\partial h}{\partial z_3} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{z}_{x_1} & \mathbf{z}_{x_2} & \mathbf{z}_{x_3} \end{bmatrix} \begin{bmatrix} \dfrac{dx_1}{dt} \\ \dfrac{dx_2}{dt} \\ \dfrac{dx_3}{dt} \end{bmatrix}}_{\dfrac{d\mathbf{z}}{dt}}
\end{aligned}$$

With $\dfrac{d\mathbf{z}}{dt} = \mathbf{z}_{x_1}\dfrac{dx_1}{dt} + \mathbf{z}_{x_2}\dfrac{dx_2}{dt} + \mathbf{z}_{x_3}\dfrac{dx_3}{dt}$ (linear combination of $\mathbf{z}_{x_1}, \mathbf{z}_{x_2}$, and $\mathbf{z}_{x_3}$) we can write

$$\frac{d}{dt}(h\circ\varphi^{-1}) = \frac{d}{dt}(h\circ\mathbf{z}) = \left\langle dh, \frac{d\mathbf{z}}{dt}\right\rangle$$

and consider the tangent vector to be the mapping

$$\mathbf{z}_p : h \to \left\langle dh, \frac{d\mathbf{z}}{dt}\right\rangle$$

instead of $\dfrac{d\mathbf{z}}{dt} = \mathbf{z}_{x_1}\dfrac{dx_1}{dt} + \mathbf{z}_{x_2}\dfrac{dx_2}{dt} + \mathbf{z}_{x_3}\dfrac{dx_3}{dt}$. We make the correspondence

$$\begin{aligned}
\frac{\partial}{\partial x_1} &\leftrightarrow \mathbf{z}_{x_1} = \frac{\partial \mathbf{z}}{\partial x_1} \\
\frac{\partial}{\partial x_2} &\leftrightarrow \mathbf{z}_{x_2} = \frac{\partial \mathbf{z}}{\partial x_2} \\
\frac{\partial}{\partial x_3} &\leftrightarrow \mathbf{z}_{x_3} = \frac{\partial \mathbf{z}}{\partial x_3}.
\end{aligned}$$

That is, for any function $h(z_1,z_2,z_3)$ we have

$$\begin{aligned}
\frac{\partial(h\circ\varphi^{-1})}{\partial x_1} &= \left\langle dh, \frac{\partial \mathbf{z}}{\partial x_1}\right\rangle \\
\frac{\partial(h\circ\varphi^{-1})}{\partial x_2} &= \left\langle dh, \frac{\partial \mathbf{z}}{\partial x_2}\right\rangle \\
\frac{\partial(h\circ\varphi^{-1})}{\partial x_3} &= \left\langle dh, \frac{\partial \mathbf{z}}{\partial x_3}\right\rangle.
\end{aligned}$$

For any tangent vector we have the correspondence

$$\mathbf{z}_p \triangleq a_1\frac{\partial}{\partial x_1} + a_2\frac{\partial}{\partial x_2} + a_3\frac{\partial}{\partial x_3} \leftrightarrow a_1\mathbf{z}_{x_1} + a_2\mathbf{z}_{x_2} + a_3\mathbf{z}_{x_3}.$$

We reiterate: $\mathbf{z}_p \triangleq a_1\dfrac{\partial}{\partial x_1} + a_2\dfrac{\partial}{\partial x_2} + a_3\dfrac{\partial}{\partial x_3}$ is the mapping that takes $h \in \mathcal{F}(\mathcal{U})^2 \to \mathbb{R}$ given by

$$\mathbf{z}_p(h) \triangleq a_1\frac{\partial(h\circ\varphi^{-1})}{\partial x_1} + a_2\frac{\partial(h\circ\varphi^{-1})}{\partial x_2} + a_3\frac{\partial(h\circ\varphi^{-1})}{\partial x_3}.$$

Because this is the manifold $\mathbf{E}^3$ we can write $\mathbf{z}_p(h)$ as

$$\begin{aligned}
\mathbf{z}_p(h) &\triangleq a_1\frac{\partial(h\circ\varphi^{-1})}{\partial x_1} + a_2\frac{\partial(h\circ\varphi^{-1})}{\partial x_2} + a_3\frac{\partial(h\circ\varphi^{-1})}{\partial x_3} \\
&= \begin{bmatrix} \dfrac{\partial h}{\partial z_1} & \dfrac{\partial h}{\partial z_2} & \dfrac{\partial h}{\partial z_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial \mathbf{z}}{\partial x_1} & \dfrac{\partial \mathbf{z}}{\partial x_2} & \dfrac{\partial \mathbf{z}}{\partial x_3} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \\
&= \left\langle dh, \frac{d\mathbf{z}}{dt} \right\rangle
\end{aligned}$$

where $\dfrac{d\mathbf{z}}{dt} = a_1\mathbf{z}_{x_1} + a_2\mathbf{z}_{x_2} + a_3\mathbf{z}_{x_3}$.

1.3 Lie Derivatives and Tangent Vectors

We now make the connection between Lie derivatives and tangent vectors. In fact, we show they are one and the same! To explain, once again consider the northern hemisphere coordinate chart $\varphi^{-1}(x_1, x_2) = \mathbf{z}(x_1, x_2) : \mathcal{D} \to \mathcal{U}$ given by

$$\varphi^{-1}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1-(x_1^2+x_2^2)} \end{bmatrix} \in \mathbf{S}^2$$

where $\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2 |\ 0 \le x_1^2 + x_2^2 < 1\}$ and $\mathcal{U} \triangleq \{\mathbf{z} \in \mathbf{S}^2 |\ z_3 > 0\}$. Also $\varphi : \mathcal{U} \subset S^2 \to \mathcal{D}$

$$\varphi\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}\right) = (z_1, z_2).$$

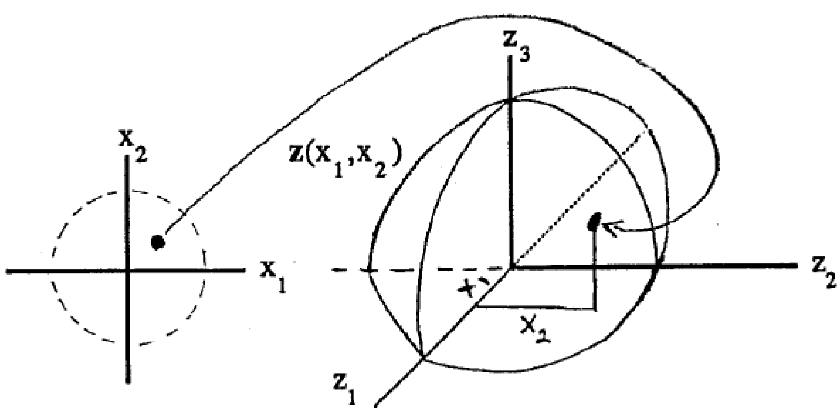

FIGURE 1.13. Northern hemisphere coordinate chart.

[2] $\mathcal{F}(\mathcal{U})$ are the differentiable functions on $\mathcal{U} \subset \mathbf{E}^3$.

Let h be a function defined on $\mathbf{S}^2$ so that

$$(h\circ\varphi^{-1})(x_1,x_2) = h\Big(x_1, x_2, \sqrt{1-(x_1^2+x_2^2)}\Big)$$

is defined on the coordinate patch $\mathcal{D}$. For any curve $(x_1(t), x_2(t))$ in $\mathcal{D}$, the curve $\varphi^{-1}(x_1(t), x_2(t))$ is on $\mathbf{S}^2$ and we have

$$\frac{d}{dt}(h\circ\varphi^{-1})(x_1(t), x_2(t)) = \frac{\partial(h\circ\varphi^{-1})}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial(h\circ\varphi^{-1})}{\partial x_2}\frac{dx_2}{dt}.$$

Recall that the tangent vector is now defined to be the map

$$\mathbf{z}_p : h \to \frac{\partial(h\circ\varphi^{-1})}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial(h\circ\varphi^{-1})}{\partial x_2}\frac{dx_2}{dt}.$$

More generally, with $f_1(x_1,x_2)$ and $f_2(x_1,x_2)$ scalar functions defined on $\mathcal{D}$, the tangent vector at the point $p = \varphi^{-1}(x_1,x_2) \in \mathbf{S}^2$ is given by the mapping

$$f_1(x_1,x_2)\frac{\partial}{\partial x_1} + f_2(x_1,x_2)\frac{\partial}{\partial x_2} : h \to f_1(x_1,x_2)\frac{\partial(h\circ\varphi^{-1})}{\partial x_1} + f_2(x_1,x_2)\frac{\partial(h\circ\varphi^{-1})}{\partial x_2}.$$

Further, with $\mathfrak{h}(x_1,x_2) \triangleq (h\circ\varphi^{-1})(x_1,x_2)$ and $f(x_1,x_2) = (f_1(x_1,x_2), f_2(x_1,x_2))$ we have the Lie derivative of $\mathfrak{h}(x_1,x_2)$ with respect to $f(x_1,x_2)$ given by

$$\mathcal{L}_f\mathfrak{h} = \frac{\partial\mathfrak{h}}{\partial x_1}f_1(x_1,x_2) + \frac{\partial\mathfrak{h}}{\partial x_2}f_2(x_1,x_2).$$

That is the Lie derivative operator $\mathcal{L}_f : \mathfrak{h} \to \mathcal{L}_f\mathfrak{h}$ is the tangent vector mapping.

Remark 1 Writing the tangent vector as $f_1(x_1,x_2)\dfrac{\partial}{\partial x_1} + f_2(x_1,x_2)\dfrac{\partial}{\partial x_2}$ tell us the local coordinates are (x_1,x_2) with a corresponding coordinate map φ^{-1} taking these coordinates to the point $\varphi^{-1}(x_1,x_2)$ on the manifold $\mathbf{S}^2$. Further, given a function h on the manifold the function $\mathfrak{h}(x_1,x_2) \triangleq (h\circ\varphi^{-1})(x_1,x_2)$ is known as φ^{-1} is known.

Remark 2 The mapping

$$\mathbf{z}_p : h \to \mathbf{z}_p(h) = f_1(x_1,x_2)\frac{\partial(h\circ\varphi^{-1})}{\partial x_1} + f_2(x_1,x_2)\frac{\partial(h\circ\varphi^{-1})}{\partial x_2}$$

is our new definition of tangent vector. Recall that we would prefer to look at the mapping as given by

$$\begin{aligned}
\mathbf{z}_p : h \to \mathbf{z}_p(h) &= \begin{bmatrix} \dfrac{\partial h(z)}{\partial z_1} & \dfrac{\partial h(z)}{\partial z_2} & \dfrac{\partial h(z)}{\partial z_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_1}{\partial x_2} \\ \dfrac{\partial z_2}{\partial x_1} & \dfrac{\partial z_2}{\partial x_2} \\ \dfrac{\partial z_3}{\partial x_1} & \dfrac{\partial z_3}{\partial x_2} \end{bmatrix} \begin{bmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{bmatrix} \\
&= dh \begin{bmatrix} \mathbf{z}_{x_1} & \mathbf{z}_{x_2} \end{bmatrix} \begin{bmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{bmatrix} \\
&= \langle dh, f_1(x_1,x_2)\mathbf{z}_{x_1} + f_2(x_1,x_2)\mathbf{z}_{x_2} \rangle
\end{aligned}$$

with $f_1(x_1,x_2)\mathbf{z}_{x_1} + f_2(x_1,x_2)\mathbf{z}_{x2}$ the tangent vector. However, this doesn't make sense because h is only defined on $\mathbf{S}^2$ ($\partial h/\partial z_1$, etc. are not defined) and $\mathbf{z}_{x_1}, \mathbf{z}_{x2}$ stick off the manifold.

Let's do a similar example in $\mathbf{E}^3$. Recall the spherical coordinate system $\varphi^{-1}(x_1, x_2, x_3)|\mathcal{D} \to \mathcal{U}$ given by

$$\varphi^{-1}(x_1, x_2, x_3) = \mathbf{z}(x_1, x_2, x_3) = \begin{bmatrix} x_1 \sin(x_2)\cos(x_3) \\ x_1 \sin(x_2)\sin(x_3) \\ x_1 \cos(x_2) \end{bmatrix} \in \mathbf{E}^3$$

where

$$\begin{aligned} \mathcal{D} &= \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 |\ x_1 > 0, 0 < x_2 < \pi, 0 < x_3 < 2\pi\right\} \\ \mathcal{U} &\triangleq \left\{\mathbf{z} \in \mathbf{E}^3 |\ \text{If } z_1 > 0 \text{ the } z_2 \neq 0\right\}. \end{aligned}$$

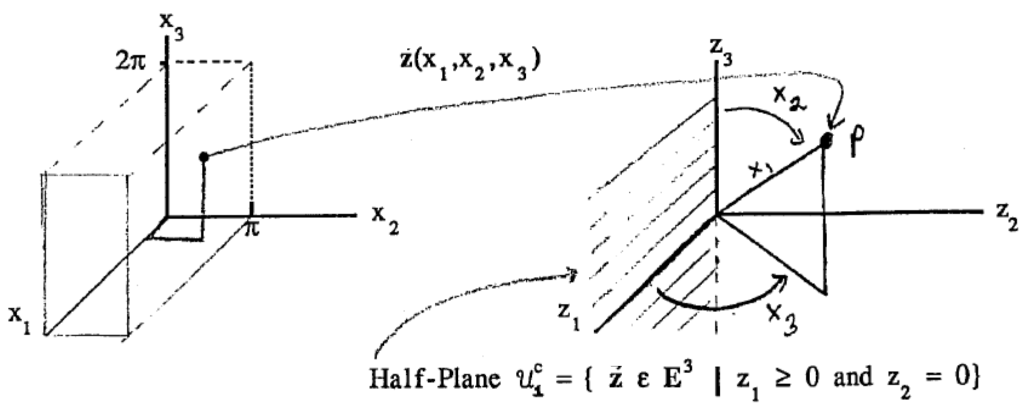

FIGURE 1.14. Spherical coordinates in $\mathbf{E}^3$.

Let $h(z_1, z_2, z_3)$ be a function on $\mathbf{E}^3$ so that

$$\mathfrak{h}(x_1, x_2, x_3) \triangleq (h \circ \varphi^{-1})(x_1, x_2, x_3) = h\left(\begin{bmatrix} x_1 \sin(x_2)\cos(x_3) \\ x_1 \sin(x_2)\sin(x_3) \\ x_1 \cos(x_2) \end{bmatrix}\right)$$

is the representation of h in spherical coordinates. For any curve $(x_1(t), x_2(t), x_3(t))$ in $\mathcal{D}$ at the point $p = \varphi^{-1}(x_1(t), x_2(t), x_3(t)) \in \mathcal{U} \subset \mathbf{E}^3$ we have the tangent vector mapping

$$\mathbf{z}_p : h \to \frac{d(h \circ \varphi^{-1})}{dt} = \frac{\partial(h \circ \varphi^{-1})}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial(h \circ \varphi^{-1})}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial(h \circ \varphi^{-1})}{\partial x_3}\frac{dx_3}{dt}.$$

That is, $\mathbf{z}_p$ is the map that takes a function h to the number $\dfrac{d(h \circ \varphi^{-1})}{dt}$ and is our new definition of tangent vector. The components of this tangent vector are $\dfrac{dx_1}{dt}, \dfrac{dx_2}{dt}$, and $\dfrac{dx_3}{dt}$. A set of basis vectors for the tangent space at $p = \varphi^{-1}(x_1(t), x_2(t), x_3(t))$ is

$$\begin{aligned} \frac{\partial}{\partial x_1} &: \quad h \to \frac{\partial(h \circ \varphi^{-1})}{\partial x_1} \\ \frac{\partial}{\partial x_2} &: \quad h \to \frac{\partial(h \circ \varphi^{-1})}{\partial x_2} \\ \frac{\partial}{\partial x_3} &: \quad h \to \frac{\partial(h \circ \varphi^{-1})}{\partial x_3}. \end{aligned}$$

Given $f(x) = (f_1(x), f_2(x), f_3(x))$ with $x \in \mathcal{D} \subset \mathbb{R}^3$ and $h(z_1, z_2, z_3)$ at function on $\mathbf{E}^3$ we have the tangent vector

$$f_1(x)\frac{\partial}{\partial x_1} + f_2(x)\frac{\partial}{\partial x_2} + f_3(x)\frac{\partial}{\partial x_3} : h \to f_1(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_1} + f_2(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_2} + f_3(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_3}$$

at the point $p = \varphi^{-1}(x_1, x_2, x_3)$. This is the same as the Lie derivative of $\mathfrak{h} \triangleq h \circ \varphi^{-1}$ with respect to f written as

$$\mathcal{L}_f : \mathfrak{h} \to \mathcal{L}_f \mathfrak{h} = \frac{\partial \mathfrak{h}}{\partial x_1} f_1(x) + \frac{\partial \mathfrak{h}}{\partial x_2} f_2(x) + \frac{\partial \mathfrak{h}}{\partial x_3} f_3(x)$$

Remark 3 Writing the tangent vector as $f_1(x)\frac{\partial}{\partial x_1} + f_2(x)\frac{\partial}{\partial x_2} + f_3(x)\frac{\partial}{\partial x_3}$ tell us the local coordinates are (x_1, x_2, x_3) with a corresponding coordinate map φ^{-1} taking these coordinates to the point $\varphi^{-1}(x_1, x_2, x_3)$ on the manifold $\mathbf{S}^2$. Further, given a function h on the manifold the function $\mathfrak{h}(x_1, x_2, x_3) \triangleq (h \circ \varphi^{-1})(x_1, x_2, x_3)$ is known as φ^{-1} is known.

Remark 4 For this particular manifold $\mathbf{E}^3$ we *can* rewrite

$$\mathbf{z}_p : h \to \mathbf{z}_p(h) = f_1(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_1} + f_2(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_2} + f_3(x)\frac{\partial(h \circ \varphi^{-1})}{\partial x_3}$$

as

$$\begin{aligned}
\mathbf{z}_p(h) &= \begin{bmatrix} \frac{\partial h(z)}{\partial z_1} & \frac{\partial h(z)}{\partial z_2} & \frac{\partial h(z)}{\partial z_3} \end{bmatrix} \begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \frac{\partial z_1}{\partial x_3} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} & \frac{\partial z_2}{\partial x_3} \\ \frac{\partial z_3}{\partial x_1} & \frac{\partial z_3}{\partial x_2} & \frac{\partial z_3}{\partial x_3} \end{bmatrix} \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \\ f_3(x_1, x_2, x_3) \end{bmatrix} \\
&= dh \begin{bmatrix} \mathbf{z}_{x_1} & \mathbf{z}_{x_2} & \mathbf{z}_{x_3} \end{bmatrix} \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \\ f_3(x_1, x_2, x_3) \end{bmatrix} \\
&= \langle dh, f_1(x_1, x_2, x_3)\mathbf{z}_{x_1} + f_2(x_1, x_2, x_3)\mathbf{z}_{x_2} + f_2(x_1, x_2, x_3)\mathbf{z}_{x_3} \rangle
\end{aligned}$$

and think of $f_1(x_1, x_2, x_3)\mathbf{z}_{x_1} + f_2(x_1, x_2, x_3)\mathbf{z}_{x_2} + f_2(x_1, x_2, x_3)\mathbf{z}_{x_3}$ as the tangent vector.

Transformation Law for Gradients

Consider two different coordinate systems for $\mathbf{z}(x) = \varphi^{-1}(x)$ and $\mathbf{z}(\bar{x}) = \bar{\varphi}^{-1}(\bar{x})$ for $\mathbf{E}^3$ as indicated in Figure 1.15. We can write the change of coordinates from x to $\bar{x}$ and vice versa as

$$\begin{aligned}
x &= (\varphi \circ \bar{\varphi}^{-1})(\bar{x}) = \varphi(\bar{\varphi}^{-1}(\bar{x})) \\
\bar{x} &= (\bar{\varphi} \circ \varphi^{-1})(x) = \bar{\varphi}(\varphi^{-1}(x)).
\end{aligned}$$

We also write this change of coordinates as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1(\bar{x}_1, \bar{x}_2, \bar{x}_3) \\ x_2(\bar{x}_1, \bar{x}_2, \bar{x}_3) \\ x_3(\bar{x}_1, \bar{x}_2, \bar{x}_3) \end{pmatrix} \tag{1.1}$$

and

$$\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} \bar{x}_1(x_1, x_2, x_3) \\ \bar{x}_2(x_1, x_2, x_3) \\ \bar{x}_3(x_1, x_2, x_3) \end{pmatrix} \tag{1.2}$$

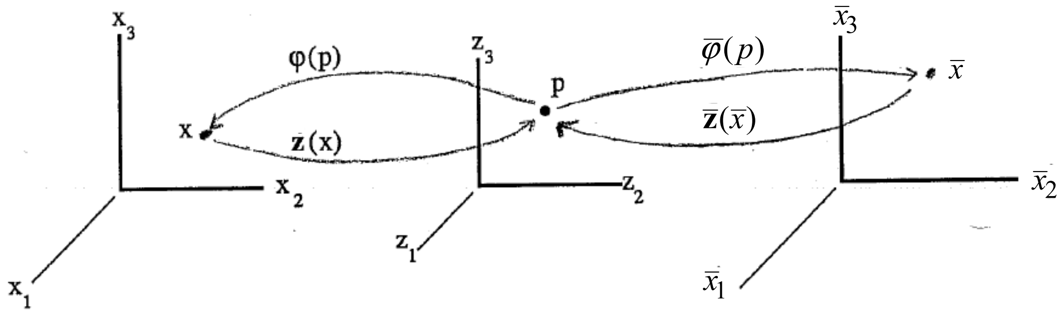

FIGURE 1.15. Two coordinate maps φ and $\bar{\varphi}$ for $\mathbf{E}^2$.

With h a function on $\mathbf{E}^3$, h in the x coordinates is $\mathfrak{h} = h \circ \varphi^{-1}(x)$ with gradient

$$d\mathfrak{h} = \left[\ \frac{\partial}{\partial x_1}(h \circ \varphi^{-1}) \quad \frac{\partial}{\partial x_2}(h \circ \varphi^{-1}) \quad \frac{\partial}{\partial x_3}(h \circ \varphi^{-1})\ \right]. \tag{1.3}$$

In the $\bar{x}$ coordinates h is represented by $\bar{\mathfrak{h}} = h \circ \bar{\varphi}^{-1}(\bar{x})$ with gradient

$$d\bar{\mathfrak{h}} = \left[\ \frac{\partial}{\partial \bar{x}_1}\left(h \circ \bar{\varphi}^{-1}\right) \quad \frac{\partial}{\partial \bar{x}_2}\left(h \circ \bar{\varphi}^{-1}\right) \quad \frac{\partial}{\partial \bar{x}_3}\left(h \circ \bar{\varphi}^{-1}\right)\ \right]. \tag{1.4}$$

How are these two gradients related? Well we have

$$\begin{aligned} \mathfrak{h}(x) &= h \circ \varphi^{-1}(x) \\ \bar{\mathfrak{h}}(\bar{x}) &= h \circ \bar{\varphi}^{-1}(\bar{x}) \end{aligned}$$

so that

$$\bar{\mathfrak{h}}(\bar{x})\big|_{\bar{x}=(\bar{\varphi}\circ\varphi^{-1})(x)} = h \circ \bar{\varphi}^{-1}(\bar{x})\big|_{\bar{x}=(\bar{\varphi}\circ\varphi^{-1})(x)} = h(\bar{\varphi}^{-1}(\bar{\varphi}(\varphi^{-1}(x))) = h(\varphi^{-1}(x)) = \mathfrak{h}(x).$$

By the chain rule we can write

$$\frac{\partial}{\partial x_1}\mathfrak{h}(x) = \frac{\partial}{\partial x_1}\left(\bar{\mathfrak{h}}(\bar{x})\big|_{\bar{x}=(\bar{\varphi}\circ\varphi^{-1})(x)}\right) = \sum_{j=1}^{3} \frac{\partial \bar{\mathfrak{h}}(\bar{x})}{\partial \bar{x}_j}\frac{\partial \bar{x}_j}{\partial x_1}$$

or

$$\begin{aligned} d\mathfrak{h} &= \left[\ \frac{\partial \bar{\mathfrak{h}}(\bar{x})}{\partial \bar{x}_1} \quad \frac{\partial \bar{\mathfrak{h}}(\bar{x})}{\partial \bar{x}_2} \quad \frac{\partial \bar{\mathfrak{h}}(\bar{x})}{\partial \bar{x}_3}\ \right] \begin{bmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \frac{\partial \bar{x}_1}{\partial x_2} & \frac{\partial \bar{x}_1}{\partial x_3} \\ \frac{\partial \bar{x}_2}{\partial x_1} & \frac{\partial \bar{x}_2}{\partial x_2} & \frac{\partial \bar{x}_2}{\partial x_3} \\ \frac{\partial \bar{x}_3}{\partial x_1} & \frac{\partial \bar{x}_3}{\partial x_2} & \frac{\partial \bar{x}_3}{\partial x_3} \end{bmatrix} && (1.5) \\ &= d\bar{\mathfrak{h}}\frac{\partial \bar{x}}{\partial x}. && (1.6) \end{aligned}$$

Let $(\bar{x}_1(t), \bar{x}_2(t), \bar{x}_3(t))$ be a curve in the $\bar{x}$ coordinates which corresponds to the curve

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} x_1(\bar{x}_1(t), \bar{x}_2(t), \bar{x}_3(t)) \\ x_2(\bar{x}_1(t), \bar{x}_2(t), \bar{x}_3(t)) \\ x_3(\bar{x}_1(t), \bar{x}_2(t), \bar{x}_3(t)) \end{pmatrix}. \tag{1.7}$$

Then

$$\begin{pmatrix} \dfrac{dx_1}{dt} \\ \dfrac{dx_2}{dt} \\ \dfrac{dx_3}{dt} \end{pmatrix} = \begin{bmatrix} \dfrac{\partial x_1}{\partial \bar{x}_1} & \dfrac{\partial x_1}{\partial \bar{x}_2} & \dfrac{\partial x_1}{\partial \bar{x}_3} \\ \dfrac{\partial x_2}{\partial \bar{x}_1} & \dfrac{\partial x_2}{\partial \bar{x}_2} & \dfrac{\partial x_2}{\partial \bar{x}_3} \\ \dfrac{\partial x_3}{\partial \bar{x}_1} & \dfrac{\partial x_3}{\partial \bar{x}_2} & \dfrac{\partial x_3}{\partial \bar{x}_3} \end{bmatrix} \begin{pmatrix} \dfrac{d\bar{x}_1}{dt} \\ \dfrac{d\bar{x}_2}{dt} \\ \dfrac{d\bar{x}_3}{dt} \end{pmatrix} \tag{1.8}$$

showing how the components of the tangent vector transform going from the $\bar{x}$ coordinates to the x coordinates. More generally we have

$$\begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} = \begin{bmatrix} \dfrac{\partial x_1}{\partial \bar{x}_1} & \dfrac{\partial x_1}{\partial \bar{x}_2} & \dfrac{\partial x_1}{\partial \bar{x}_3} \\ \dfrac{\partial x_2}{\partial \bar{x}_1} & \dfrac{\partial x_2}{\partial \bar{x}_2} & \dfrac{\partial x_2}{\partial \bar{x}_3} \\ \dfrac{\partial x_3}{\partial \bar{x}_1} & \dfrac{\partial x_3}{\partial \bar{x}_2} & \dfrac{\partial x_3}{\partial \bar{x}_3} \end{bmatrix} \begin{pmatrix} \bar{f}_1(\bar{x}) \\ \bar{f}_2(\bar{x}) \\ \bar{f}_3(\bar{x}) \end{pmatrix} \tag{1.9}$$

or compactly

$$f = \frac{\partial x}{\partial \bar{x}} \bar{f}. \tag{1.10}$$

Equation (1.5) is the transformation of the gradient in the $\bar{x}$ coordinates to the gradient in the x coordinates while Equation (1.10) is the transformation of the components of the tangent vector in the $\bar{x}$ coordinates to the components in the x coordinates. Note that these matrix transformations are inverses of each other. A quantity that transforms like the components of a tangent vector is called a *contravariant* vector while a quantity that transforms like a gradient is called a *covector*.

Invariance of the Lie Derivative Under Coordinate Transformations

Let $f(x) = (f_1(x), f_2(x), f_3(x))$ be the components of a tangent vector in the x coordinate system and $\bar{f}(\bar{x}) = \big(\bar{f}_1(\bar{x}), \bar{f}_2(\bar{x}), \bar{f}_3(\bar{x})\big)$ be the components of this same tangent vector in the $\bar{x}$ coordinates. With h a function defined on $\mathbf{E}^3$ it is represented in the x coordinates as $h \circ \varphi^{-1}$ and in the $\bar{x}$ coordinates as $h \circ \bar{\varphi}^{-1}$. Then

$$\begin{aligned} \mathcal{L}_f(h \circ \varphi^{-1}) &= \begin{bmatrix} \dfrac{\partial(h \circ \varphi^{-1})}{\partial x_1} & \dfrac{\partial(h \circ \varphi^{-1})}{\partial x_2} & \dfrac{\partial(h \circ \varphi^{-1})}{\partial x_3} \end{bmatrix} \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} \\ &= \begin{bmatrix} \dfrac{\partial(h \circ \bar{\varphi}^{-1})}{\partial \bar{x}_1} & \dfrac{\partial(h \circ \bar{\varphi}^{-1})}{\partial \bar{x}_2} & \dfrac{\partial(h \circ \bar{\varphi}^{-1})}{\partial \bar{x}_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial \bar{x}_1}{\partial x_1} & \dfrac{\partial \bar{x}_1}{\partial x_2} & \dfrac{\partial \bar{x}_1}{\partial x_3} \\ \dfrac{\partial \bar{x}_2}{\partial x_1} & \dfrac{\partial \bar{x}_2}{\partial x_2} & \dfrac{\partial \bar{x}_2}{\partial x_3} \\ \dfrac{\partial \bar{x}_3}{\partial x_1} & \dfrac{\partial \bar{x}_3}{\partial x_2} & \dfrac{\partial \bar{x}_3}{\partial x_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial x_1}{\partial \bar{x}_1} & \dfrac{\partial x_1}{\partial \bar{x}_2} & \dfrac{\partial x_1}{\partial \bar{x}_3} \\ \dfrac{\partial x_2}{\partial \bar{x}_1} & \dfrac{\partial x_2}{\partial \bar{x}_2} & \dfrac{\partial x_2}{\partial \bar{x}_3} \\ \dfrac{\partial x_3}{\partial \bar{x}_1} & \dfrac{\partial x_3}{\partial \bar{x}_2} & \dfrac{\partial x_3}{\partial \bar{x}_3} \end{bmatrix} \begin{pmatrix} \bar{f}_1(\bar{x}) \\ \bar{f}_2(\bar{x}) \\ \bar{f}_3(\bar{x}) \end{pmatrix} \\ &= \mathcal{L}_{\bar{f}}(h \circ \bar{\varphi}^{-1}) \end{aligned} \tag{1.11}$$

showing that the value of Lie derivative (tangent vector) is the same in all coordinate systems.

1.4 Submanifolds and the Implicit Function Theorem

Let's work in the manifold $\mathbf{E}^3$ with the Cartesian coordinate system $\varphi^{-1} : \mathbb{R}^3 \to \mathbf{E}^3$ given by

$$\varphi^{-1}(x_1, x_2, x_3) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbf{E}^3. \tag{1.12}$$

With c a constant, consider the implicit equation $F(z_1, z_2, z_3) = c$ defined on $\mathbf{E}^3$. For example, let $F(z_1, z_2, z_3) = z_1^2 + z_2^2 + z_3^2$ and consider the points $z \in \mathbf{E}^3$ satisfying $F(z_1, z_2, z_3) = c$ with $c > 0$. This implicit relationship allows us to solve for one of the variables in terms of the other two variables. Specifically, for all $z \in \mathbf{E}^3$ with $z_1^2 + z_2^2 \leq \sqrt{c}$, we have

$$\begin{aligned} z_3 &= \sqrt{c - (z_1^2 + z_2^2)} \\ z_3 &= -\sqrt{c - (z_1^2 + z_2^2)}. \end{aligned}$$

We use the function F to define a new manifold (submanifold) $\mathcal{M} \subset \mathbf{E}^3$ given by

$$\mathcal{M} \triangleq \left\{ \mathbf{z} \in \mathbf{E}^3 | \ F(z_1, z_2, z_3) - c = z_1^2 + z_2^2 + z_3^2 - c = 0 \right\} \tag{1.13}$$

which is just the surface of a sphere of radius $\sqrt{c}$. To make it into a manifold we need to have coordinate charts for it. We can define a coordinate chart by $\mathbf{z}(x_1, x_2) : \mathcal{D}_1 \to \mathcal{U}_1 \subset \mathbf{E}^3$ by

$$\mathbf{z}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{c - (x_1^2 + x_2^2)} \end{bmatrix} \in \mathbf{E}^3$$

where $\mathcal{D}_1 = \left\{ (x_1, x_2) \in \mathbb{R}^2 | \ 0 \leq x_1^2 + x_2^2 < 1 \right\}$ and $\mathcal{U}_1 \triangleq \{ \mathbf{z} \in \mathcal{M} | \ z_3 > 0 \}$.

Similarly, a coordinate chart for points of $\mathcal{M}$ with $z_3 < 0$ is given by $\mathbf{z}(\bar{x}_1, \bar{x}_2) : \mathcal{D}_2 \to \mathcal{U}_2 \subset \mathbf{E}^3$

$$\bar{\mathbf{z}}(\bar{x}_1, \bar{x}_2) = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ -\sqrt{c - (\bar{x}_1^2 + \bar{x}_2^2)} \end{bmatrix} \in \mathbf{E}^3 \tag{1.14}$$

where $\mathcal{D}_2 = \left\{ (\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2 | \ 0 \leq \bar{x}_1^2 + \bar{x}_2^2 < 1 \right\}$ and $\mathcal{U} \triangleq \{ \mathbf{z} \in \mathcal{M} | \ z_3 < 0 \}$. To get coordinate charts for points of $\mathcal{M}$ where $z_3 = 0$ one must solve for z_1 or z_2 in terms of the remaining variables.

We now generalize this example by considering $\mathbf{E}^3$ to be made up of spherically shaped submanifolds. To do this consider the coordinate system $\varphi^* : \mathcal{D}^* \to \mathcal{U}^*$ on $\mathbf{E}^3$ given by

$$\varphi^*\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \right) = \left(z_1, z_2, z_1^2 + z_2^2 + z_3^2 \right) = (x_1^*, x_2^*, x_3^*) \tag{1.15}$$

where $\mathcal{D}^* = \left\{ (x_1^*, x_2^*, x_3^*) \in \mathbb{R}^3 | \ 0 < (x_1^*)^2 + (x_2^*)^2 < x_3^* \right\}$ and $\mathcal{U}^* \triangleq \{ \mathbf{z} \in \mathbf{E}^3 | \ z_3 > 0 \}$. The inverse is

$$\varphi^{*-1}(x_1^*, x_2^*, x_3^*) = \begin{bmatrix} x_1^* \\ x_2^* \\ \sqrt{x_3^* - ((x_1^*)^2 + (x_2^*)^2)} \end{bmatrix}. \tag{1.16}$$

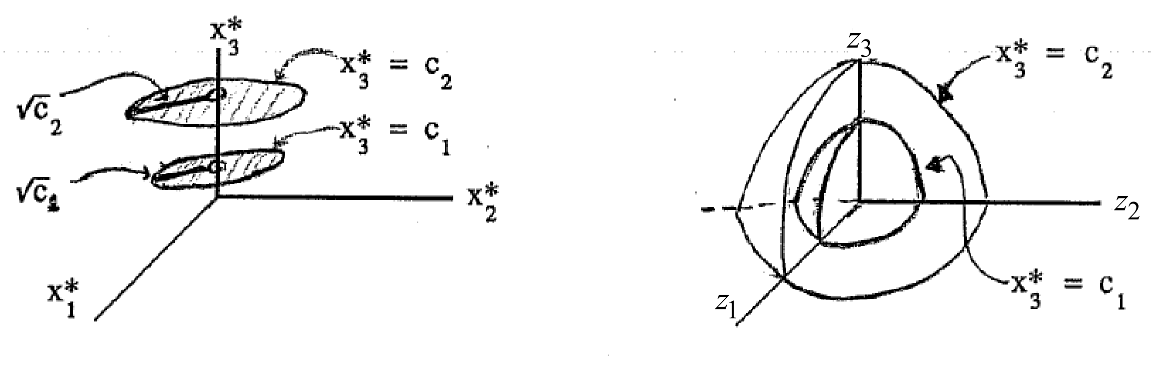

FIGURE 1.16. x and x^* coordinate systems.

With (x_1, x_2, x_3) the Cartesian coordinates for $\mathbf{E}^3$ the change of coordinates from x^* to x is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1^* \\ x_2^* \\ \sqrt{x_3^* - \left((x_1^*)^2 + (x_2^*)^2\right)} \end{pmatrix} \tag{1.17}$$

with inverse

$$\begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_1^2 + x_2^2 + x_3^2 \end{pmatrix}. \tag{1.18}$$

It turns out to be very common to define submanifolds of $\mathbf{E}^3$ implicitly as just described. The main tool needed to use this approach is the implicit theorem.

Theorem 1 *Implicit Function Theorem*

Let $F(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R}$ be a scalar valued continuously differentiable function. Let $\mathcal{M}$ be the subset of $\mathbb{R}^3$ defined by

$$\mathcal{M} \triangleq \left\{x \in \mathbb{R}^3 \middle| \, F(x_1, x_2, x_3) - c = 0\right\}.$$

Suppose

(1) $\mathcal{M}$ is nonempty.

(2) $\left.\dfrac{\partial F(x_1, x_2, x_3)}{\partial x_3}\right|_{x_0} \neq 0$ for $x_0 = (x_{01}, x_{02}, x_{03}) \in \mathcal{M}$.

Then there exists a neighborhood $\mathcal{D}$ of (x_{01}, x_{02}) and a scalar function $s(x_1, x_2) : \mathcal{D} \to \mathcal{M}$ such that

$$F(x_1, x_2, s(x_1, x_2)) - c = 0.$$

Proof. Define the transformation

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ F(x_1, x_2, x_3) - c \end{bmatrix}. \tag{1.19}$$

with

$$\begin{bmatrix} x_{01}' \\ x_{02}' \\ x_{03}' \end{bmatrix} \triangleq \begin{bmatrix} x_{01} \\ x_{02} \\ F(x_{01}, x_{02}, x_{03}) - c \end{bmatrix} = \begin{bmatrix} x_{01} \\ x_{02} \\ 0 \end{bmatrix}. \tag{1.20}$$

The Jacobian of this transformation is

$$J(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \dfrac{\partial F(x_1, x_2, x_3)}{\partial x_1} & \dfrac{\partial F(x_1, x_2, x_3)}{\partial x_2} & \dfrac{\partial F(x_1, x_2, x_3)}{\partial x_3} \end{bmatrix} \tag{1.21}$$

which is nonsingular at x_0 as $\det J(x_0) = \left.\dfrac{\partial F(x_1, x_2, x_3)}{\partial x_3}\right|_{x_0} \neq 0$. By the inverse function theorem this transformation has a unique inverse about the point $x_0' = \begin{bmatrix} x_{01} & x_{02} & 0 \end{bmatrix}^T$ of the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \triangleq \begin{bmatrix} x_1' \\ x_2' \\ h(x_1', x_2', x_3') \end{bmatrix}. \tag{1.22}$$

Substitute (1.22) into (1.19) to obtain

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} \triangleq \begin{bmatrix} x_1' \\ x_2' \\ F(x_1', x_2', h(x_1', x_2', x_3')) - c \end{bmatrix}. \tag{1.23}$$

This is true for all (x_1', x_2', x_3') in some neighborhood $\mathcal{U}'$ which contains the point $(x_{01}, x_{02}, 0)$. This is illustrated in Figure 1.17.

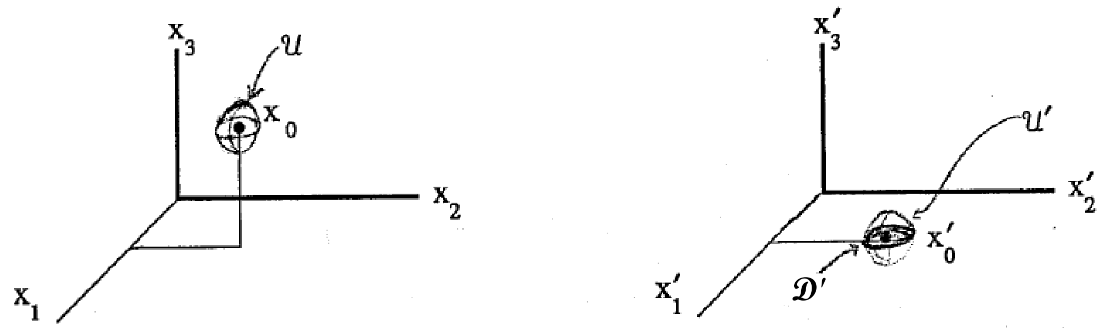

FIGURE 1.17. Transformation to the x' coordinate system.

Consequently Equation (1.23) holds for all $(x_1', x_2', 0)$ in a neighborhood $(x_{01}, x_{02}, 0)$. Specifically, let $\mathcal{D}' \triangleq \{(x_1', x_2') |\ (x_1', x_2', 0) \in \mathcal{U}'\}$ and set $x_3' = 0$ in (1.23) to obtain

$$0 = F(x_1', x_2', h(x_1', x_2', 0)) - c$$

for all $(x_1', x_2') \in \mathcal{D}'$ of (x_{01}, x_{02}). Define $s(x_1, x_2) \triangleq h(x_1, x_2, 0)$ and, as $x_1 = x_1', x_2 = x_2'$ we have

$$F(x_1, x_2, s(x_1, x_2)) - c \equiv 0.$$

■

Remark An immediate result of this theorem is that

$$\varphi^{-1}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \\ s(x_1, x_2) \end{bmatrix} \in \mathcal{M}$$

is a coordinate map for $\mathcal{M}$ around the point $p = \varphi^{-1}(x_{01}, x_{02}) = \begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \end{bmatrix}$.

Gradients, Tangent Vectors, and Manifolds

With the manifold $\mathbf{E}^3$ we can consider the Cartesian coordinates $(x_1, x_2, x_3) \in \mathbb{R}^3$ and the point $\begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}^T$ to be the same thing with $z_1 = x_1, z_2 = x_2, z_3 = x_3$. With this understanding, we now again consider a submanifold of $\mathbf{E}^3$ defined implicitly by a function $F(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R}$, that is, $\mathcal{M} \triangleq \{x \in \mathbb{R}^3 |\ F(x_1, x_2, x_3) - c = 0\}$ for some constant c. If $\left.\dfrac{\partial F(x_1, x_2, x_3)}{\partial x_3}\right|_{x_0} \neq 0$ for some point x_0 satisfying $F(x_0) - c = 0$ the implicit function theorem ensures we can find a surface

$$S(u_1, u_2) = \begin{bmatrix} s_1(u_1, u_2) \\ s_2(u_1, u_2) \\ s_3(u_1, u_2) \end{bmatrix} \tag{1.24}$$

with

$$S(u_{01}, u_{02}) = x_0 \tag{1.25}$$

and

$$F(s_1(u_1, u_2), s_2(u_1, u_2), s_3(u_1, u_2)) - c \equiv 0. \tag{1.26}$$

(In the proof of the implicit function theorem we took $u_1 = x_1, u_2 = x_2$ and $s_1(u_1, u_2) = u_1, s_2(u_1, u_2) = u_2$.)

Using the chain rule on Equation (1.26) we have

$$\begin{aligned}
\frac{\partial}{\partial u_1} F(s_1(u_1, u_2), s_2(u_1, u_2), s_3(u_1, u_2)) &= \frac{\partial F}{\partial x_1}\frac{\partial s_1}{\partial u_1} + \frac{\partial F}{\partial x_2}\frac{\partial s_2}{\partial u_1} + \frac{\partial F}{\partial x_3}\frac{\partial s_3}{\partial u_1} \equiv 0 \\
\frac{\partial}{\partial u_2} F(s_1(u_1, u_2), s_2(u_1, u_2), s_3(u_1, u_2)) &= \frac{\partial F}{\partial x_1}\frac{\partial s_1}{\partial u_2} + \frac{\partial F}{\partial x_2}\frac{\partial s_2}{\partial u_2} + \frac{\partial F}{\partial x_3}\frac{\partial s_3}{\partial u_2} \equiv 0.
\end{aligned}$$

These can be written in a geometric fashion via the dual product

$$\left\langle \frac{\partial F}{\partial x}, \frac{\partial S}{\partial u_1} \right\rangle \equiv 0 \tag{1.27}$$

$$\left\langle \frac{\partial F}{\partial x}, \frac{\partial S}{\partial u_2} \right\rangle \equiv 0 \tag{1.28}$$

where $\dfrac{\partial S}{\partial u_1} \in \mathbb{R}^3, \dfrac{\partial S}{\partial u_2} \in \mathbb{R}^3$ are contravariant (column) vectors and $\dfrac{\partial F}{\partial x} \in \mathbb{R}^{1\times 3}$ is a (row) covector.

The tangent space $\mathbf{T}_p(\mathcal{M})$ to the surface (submanifold) at the point $p = [s_1(u_1, u_2), s_2(u_1, u_2), s_3(u_1, u_2)]^T$ is spanned by $\dfrac{\partial S}{\partial u_1}$ and $\dfrac{\partial S}{\partial u_2}$. The above relationships show that the gradient $\left.\dfrac{\partial F}{\partial x}\right|_{p\in\mathcal{M}}$ is orthogonal (normal/perpendicular) to the tangent space $\mathbf{T}_p(\mathcal{M})$. In other words, if a manifold is defined implicitly by a function $F(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R}$ as $\mathcal{M} \triangleq \left\{x \in \mathbb{R}^3 | \ F(x_1, x_2, x_3) - c = 0\right\}$, then the gradient $\left.\dfrac{\partial F}{\partial x}\right|_{p\in\mathcal{M}}$ is perpendicular to $\mathbf{T}_p(\mathcal{M})$.

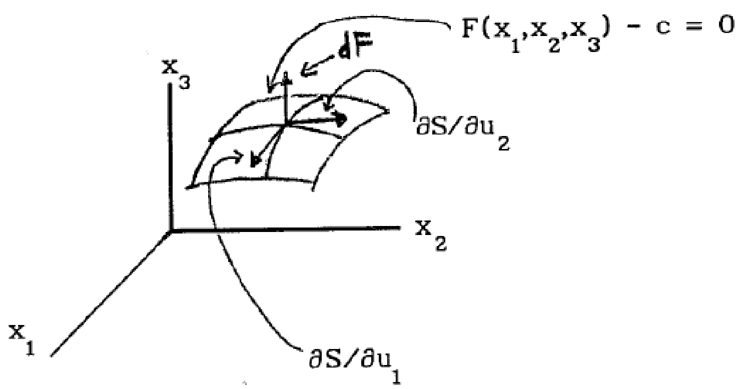

FIGURE 1.18. $\dfrac{\partial S}{\partial u_1}, \dfrac{\partial S}{\partial u_2}$ are tangent to the submanifold and $\dfrac{\partial F}{\partial x}$ is normal to the manifold.

Remark 5 The above discussion is really no surprise from a basic calculus course. Again let $F(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R}$ be a continuously differentiable function. Next consider a curve $(x_1(t), x_2(t), x_3(t))$ with $(x_1(0), x_2(0), x_3(0)) = (x_{01}, x_{02}, x_{03})$ and let $f(t) \triangleq F(x_1(t), x_2(t), x_3(t))$ so that we have

$$\frac{d}{dt} f(t) = \frac{d}{dt} F(x_1(t), x_2(t), x_3(t)) = \frac{\partial F}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial F}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial F}{\partial x_3}\frac{dx_3}{dt}$$

and, at $t = 0$,

$$f'(0) = \left\langle \left.\frac{\partial F}{\partial x}\right|_{x_0}, \dot{x}(0) \right\rangle.$$

The gradient $\left.\dfrac{\partial F}{\partial x}\right|_{x_0}$ points in the direction of the *maximum* rate of change of the function $F(x)$ at x_0. If the curve lies on the level curve defined by $F(x_1, x_2, x_3) - c = 0$ then $F(x_1(t), x_2(t), x_3(t)) \equiv c$ and it value does not change on $x(t)$. Thus $f'(0) = 0$ so $\dot{x}(0) = (\dot{x}_1(0), \dot{x}_2(0), \dot{x}_3(0))$ must be orthogonal to $\left.\dfrac{\partial F}{\partial x}\right|_{x_0}$. Further, as $(x_1(t), x_2(t), x_3(t)) \in \mathcal{M} = \{x \in \mathbb{R}^3 |\ F(x_1, x_2, x_3) - c = 0\}$ for all t, it must be that $\dot{x}(t) = (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t))$ is tangent to $\mathcal{M}$ for all t.

We now state the implicit function theorem for the general case.

Theorem 2 *Implicit Function Theorem - General Case*
Let $x \in \mathbb{R}^{n+m}$ and consider the m differentiable functions $F_1(x), F_2(x), ..., F_m(x)$. We look for solutions to the m equations

$$\begin{aligned} F_1(x_1, ..., x_n, x_{n+1}, ..., x_{n+m}) &= 0 \\ F_2(x_1, ..., x_n, x_{n+1}, ..., x_{n+m}) &= 0 \\ \vdots \quad &= \vdots \\ F_m(x_1, ..., x_n, x_{n+1}, ..., x_{n+m}) &= 0. \end{aligned} \tag{1.29}$$

Suppose at $x_0 = (x_{01}, ..., x_{0n}, x_{0,n+1}, ..., x_{0,n+m})$ we have $F_1(x_0) = 0, F_2(x_0) = 0, ..., F_m(x_0) = 0$ and

$$\det \begin{bmatrix} \dfrac{\partial F_1}{\partial x_{n+1}} & \cdots & \dfrac{\partial F_1}{\partial x_{n+m}} \\ \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_{n+1}} & \cdots & \dfrac{\partial F_m}{\partial x_{n+m}} \end{bmatrix}_{|x_0} \neq 0. \tag{1.30}$$

Then there exists m functions $s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), ..., s_m(x_1, ..., x_n)$ such that

$$\begin{aligned} F_1(x_1, ..., x_n, s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), ..., s_m(x_1, ..., x_n)) &= 0 \\ F_2(x_1, ..., x_n, s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), ..., s_m(x_1, ..., x_n)) &= 0 \\ \vdots \quad &= \vdots \\ F_m(x_1, ..., x_n, s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), ..., s_m(x_1, ..., x_n)) &= 0 \end{aligned} \tag{1.31}$$

for all $(x_1, ..., x_n)$ in a neighborhood of $(x_{01}, ..., x_{0n})$ with

$$\begin{aligned} x_{0,n+1} &= s_1(x_{01}, ..., x_{0n}) \\ x_{0,n+2} &= s_2(x_{01}, ..., x_{0n}) \\ \vdots \quad &= \quad \vdots \\ x_{0,n+m} &= s_m(x_{01}, ..., x_{0n}) \end{aligned} \tag{1.32}$$

Proof. Exercise - Just use the inverse function theorem. ∎

Remark 6 The implicit function theorem shows tells that if $x_0 \in \mathbb{R}^{n+m}$ satisfies

$$\mathcal{M} \triangleq \left\{ x \in \mathbb{R}^{n+m} |\ F_1(x) = 0, ..., F_m(x) = 0 \right\}$$

with

$$\det \begin{bmatrix} \dfrac{\partial F_1}{\partial x_{n+1}} & \cdots & \dfrac{\partial F_1}{\partial x_{n+m}} \\ \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_{n+1}} & \cdots & \dfrac{\partial F_m}{\partial x_{n+m}} \end{bmatrix}_{x_0} \neq 0$$

then there is a neighborhood of points of x_0 in $\mathcal{M}$ for which we have a coordinate chart given by

$$\varphi^{-1}(x_1, ..., x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ s_1(x_1, ..., x_n) \\ \vdots \\ s_m(x_1, ..., x_n) \end{bmatrix} \in \mathcal{M}.$$

We say $\mathcal{M}$ is n dimensional as it has n independent coordinates.

1.5 Feedback Linearizing Transformations and Integral Manifolds

Let's go back and take a look at the idea of feedback linearization and its relationship to tangent vectors and manifolds.

Consider the control system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \\ f_3(x_1, x_2, x_3) \end{bmatrix} + \begin{bmatrix} g_1(x_1, x_2, x_3) \\ g_2(x_1, x_2, x_3) \\ g_3(x_1, x_2, x_3) \end{bmatrix} u \tag{1.33}$$

or more compactly as

$$\frac{dx}{dt} = f(x) + g(x)u.$$

We want to find conditions for which a transformation of the form

$$\begin{aligned} x_1^* &= T_1(x_1, x_2, x_3) \\ x_2^* &= T_2(x_1, x_2, x_3) \\ x_3^* &= T_3(x_1, x_2, x_3) \end{aligned} \tag{1.34}$$

exists such that in the new coordinates the control system is given by

$$\frac{d}{dt}\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} x_2^* \\ x_3^* \\ f_3^*(x_1^*, x_2^*, x_3^*) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ g_3^*(x_1^*, x_2^*, x_3^*) \end{bmatrix} u. \tag{1.35}$$

We have

$$\begin{aligned} \frac{dx_1^*}{dt} &= \mathcal{L}_f(T_1) + u\mathcal{L}_g(T_1) \\ \frac{dx_2^*}{dt} &= \mathcal{L}_f(T_2) + u\mathcal{L}_g(T_2) \\ \frac{dx_3^*}{dt} &= \mathcal{L}_f(T_3) + u\mathcal{L}_g(T_3) \end{aligned} \tag{1.36}$$

In order for (1.36) to have the form (1.35) we define must define x_2^* as

$$x_2^* \triangleq \mathcal{L}_f(T_1)$$

with T_1 satisfying

$$\mathcal{L}_g(T_1) = 0$$

and we must define x_3^* as

$$x_3^* \triangleq \mathcal{L}_f(T_2) = \mathcal{L}_f^2(T_1)$$

with T_1 also satisfying

$$\mathcal{L}_g(T_2) = \mathcal{L}_g(\mathcal{L}_f(T_1)) = 0.$$

That is, we must find $T_1(x)$ such that

$$\begin{aligned} \mathcal{L}_g(T_1) &= 0 \qquad (1.37) \\ \mathcal{L}_g(\mathcal{L}_f(T_1)) &= 0. \qquad (1.38) \end{aligned}$$

These are second-order nonlinear partial differential equations in the unknown T_1. We can simplify these equations a bit by using the identity $\mathcal{L}_{[f,g]}(h) = \mathcal{L}_g\mathcal{L}_f(h) - \mathcal{L}_f\mathcal{L}_g(h)$ shown in Chapter 1. The conditions (1.37) and (1.38) become

$$\begin{aligned} \mathcal{L}_g(T_1) &= 0 \\ \mathcal{L}_{[f,g]}(T_1) &= \mathcal{L}_g\mathcal{L}_f(T_1) - \mathcal{L}_f\mathcal{L}_g(T_1) = 0 \end{aligned}$$

or

$$dT_1 \begin{bmatrix} g & ad_f g \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

where it is recalled that $ad_f g \triangleq [f, g]$. This has a nice geometric interpretation because it says that the gradient dT_1 must be perpendicular to $g, ad_f g$. The coordinate function $T_1(x)$ is a mapping from $\mathbb{R}^3 \to \mathbb{R}$. If we consider the submanifold $\mathcal{M} \triangleq \left\{x \in \mathbb{R}^3 | \ T_1(x_1, x_2, x_3) = c_1\right\}$ then we know that dT_1 is orthogonal to this two dimensional manifold. So the key problem to finding a feedback linearizing function is to find a function $T_1 : \mathbb{R}^3 \to \mathbb{R}$ such that for any c_1 the tangent space to $\mathcal{M} \triangleq \left\{x \in \mathbb{R}^3 | \ T_1(x_1, x_2, x_3) = c_1\right\}$ is spanned by $\{g, ad_f g\}$. The implicit function theorem tells us if certain conditions hold there is an explicit representation of the surface given by

$$S(u_1, u_2) = \begin{bmatrix} s_1(u_1, u_2) \\ s_2(u_1, u_2) \\ s_3(u_1, u_2) \end{bmatrix}$$

mapping a subset of $\mathbb{R}^2$ to $\mathcal{M}$. The tangent vectors to M are $\dfrac{\partial S}{\partial u_1}$ and $\dfrac{\partial S}{\partial u_2}$ and the feedback linearizing conditions tells us that

$$\{g, ad_f g\} \in span\left\{\frac{\partial S}{\partial u_1}, \frac{\partial S}{\partial u_2}\right\} = \left\{r_1\frac{\partial S}{\partial u_1} + r_2\frac{\partial S}{\partial u_2} \mid \text{with } r_1, r_2 \in \mathbb{R}\right\}.$$

We show how this all works out in the next chapter.

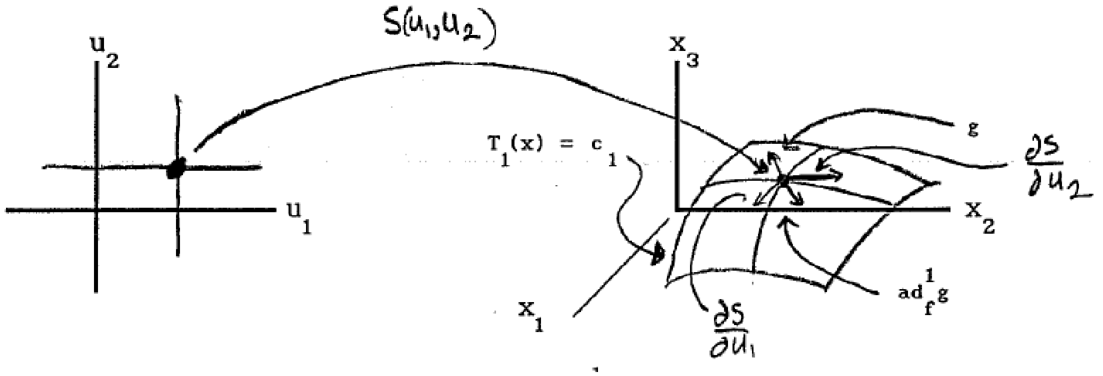

FIGURE 1.19. The vectors $\dfrac{\partial S}{\partial u_1}$ and $\dfrac{\partial S}{\partial u_2}$ span the same plane as the vectors g and $ad_f g$.

1.6 Problems

Problem 1 *Implicit Function Theorem*
Let $F_1(x_1, x_2, x_3, x_4, x_5)$ and $F_2(x_1, x_2, x_3, x_4, x_5)$ be continuously differentiable functions from an open set $\mathcal{U} \subset \mathbb{R}^5 \to \mathbb{R}$. Further suppose $x_0 = (x_{01}, x_{02}, x_{03}, x_{04}, x_{05}) \in \mathcal{U}$ satisfying $F_1(x_0) = F_2(x_0) = 0$ and

$$\left.\frac{\partial(F_1, F_2)}{\partial x_4 \partial x_5}\right|_{x_0} \triangleq \det \left.\left[\begin{array}{cc} \dfrac{\partial F_1}{\partial x_4} & \dfrac{\partial F_1}{\partial x_5} \\ \dfrac{\partial F_2}{\partial x_4} & \dfrac{\partial F_2}{\partial x_5} \end{array}\right]\right|_{x_0} \neq 0.$$

(a) Using the inverse function theorem show there exists functions $s_1(x_1, x_2, x_3)$ and $s_2(x_1, x_2, x_3)$ defined on a neighborhood $\mathcal{D} \subset \mathbb{R}^3$ containing (x_{01}, x_{02}, x_{03}) such that

$$\begin{aligned} x_{04} &= s_1(x_{01}, x_{02}, x_{03}) \\ x_{05} &= s_2(x_{01}, x_{02}, x_{03}) \end{aligned}$$

and for all $(x_1, x_2, x_3) \in \mathcal{D}$

$$\begin{aligned} F_1(x_1, x_2, x_3, s_1(x_1, x_2, x_3), s_2(x_1, x_2, x_3)) &\equiv 0 \\ F_2(x_1, x_2, x_3, s_1(x_1, x_2, x_3), s_2(x_1, x_2, x_3)) &\equiv 0. \end{aligned}$$

(b) Use part (a) to construct a coordinate chart for the manifold defined by

$$\mathcal{M} \triangleq \left\{x \in \mathbb{R}^5 |\ F_1(x_1, x_2, x_3, x_4, x_5) = 0,\ F_2(x_1, x_2, x_3, x_4, x_5) = 0\right\}$$

which contains x_0.

(c) Consider the system of equations

$$\begin{aligned} F_1(x) &= x_1^2 - x_2 x_4 = 0 \\ F_2(x) &= x_1 x_2 + x_4 x_5 = 0. \end{aligned}$$

Show that $x_0 = (-1, 1, 0, 1, 1)$ satisfies this system of equations. Show that $\left.\dfrac{\partial(F_1, F_2)}{\partial x_4 \partial x_5}\right|_{x_0} \neq 0$. Explicitly find $x_4 = s_1(x_1, x_2, x_3), x_5 = s_2(x_1, x_2, x_3)$ which are valid in a neighborhood of $(-1, 1, 0)$. Use s_1, s_2 to define a coordinate chart for $\mathcal{M}$.

1.7 References