The Lamé and Metric Coefficients for Curvilinear Coordinates in $\mathbb{R}^3$

Victor I. Piercey
Univserity of Arizona

November 21, 2007

Abstract

This paper will explore how to perform differential calculus operations on vectors in general curvilinear coordinates. We will begin with the case that the coordinates are orthogonal. Our differential operations will depend on knowing the so-called Lamé coefficients associated with the coordinates. When we pass to the general case, the differential operations involve the metric coefficients and Christoffel symbols. Along the way we will encounter the classical origins of modern concepts such as covariant and contravariant tensors, Riemannian metrics, dual bases, and Christoffel symbols.

Contents

The Lamé and Metric Coefficients for Curvilinear Coordinates in $\mathbb{R}^3$

Victor I. Piercey
November 21, 2007

1 Introduction

On the inside of the cover of my undergraduate electromagnetism textbook [PS02], there were cryptic formulae for differential operators in cylindrical and spherical coordinates. For example, in spherical coordinates, we have the gradient, divergence, curl and Laplacian given as follows:

$$\begin{aligned}
\nabla f &= \hat{r}\frac{\partial f}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi} \\
\nabla \cdot \mathbf{A} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 A_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta A_\theta) + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial \phi} \\
\nabla \times \mathbf{A} &= \frac{\hat{r}}{r\sin\theta}\left[\frac{\partial}{\partial \theta}(\sin\theta A_\phi) - \frac{\partial A_\theta}{\partial \phi}\right] + \frac{\hat{\theta}}{r}\left[\frac{1}{\sin\theta}\frac{\partial A_r}{\partial \phi} - \frac{\partial}{\partial r}(r A_\phi)\right] + \frac{\hat{\phi}}{r}\left[\frac{\partial}{\partial r}(r A_\theta) - \frac{\partial A_r}{\partial \theta}\right] \\
\nabla^2 f &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}.
\end{aligned}$$

The derivation of these formulae can be quite tedious. Since electromagnetism frequently involves cylindrical symmetry as well, it was necessary to derive expressions for these operators in cylindrical coordinates. In the derivations, one notices that certain expressions come up over and over again. This makes one suspect that there is something more general behind the computations. This is indeed correct. Cylindrical and spherical coordinate systems in $\mathbb{R}^3$ are examples of orthogonal curvilinear coordinate systems in $\mathbb{R}^3$. Such coordinate systems come equipped with a set of functions, called the Lamé coefficients. If the curvilinear coordinates are not orthogonal, the more general metric coefficients are required. The expressions for the gradient, divergence, curl and Laplacian operators in curvilinear coordinates can all be expressed in terms of these coefficients. This allows us to do the computations once and only once for every orthogonal curvilinear coordinate system, or more generally any curvilinear coordinate system. We will explore these concepts in this paper, with the goal of recovering the expressions above with greater ease. We will start with the orthogonal case, and proceed to the general case.

In what follows, we make the following conventions. If $f(x, y, z) = c$ is a level surface, then f is continuously differentiable with nonvanishing gradient. The triple scalar product will be denoted by:

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}].$$

Observe that three vectors $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ are linearly independent and thus form a basis for $\mathbb{R}^3$ if and only if $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] \neq 0$ (as the triple scalar product can be computed as the determinant of a matrix with $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ as its first, second and third rows respectively). A basis $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ for $\mathbb{R}^3$ will be considered right-handed when $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] > 0$ and left-handed when $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] < 0$. In modern terminology, a right handed basis is one with positive orientation and a left-handed basis has negative orientation (with respect to the standard basis).

ACKNOWLEDGEMENTS: The material below comes from private lectures given by Steve Sy at Michigan State University. Any errors. of course, are my own.

2 Terminology

2.1 Orthogonal Coordinate Surfaces

Suppose we have three level surfaces given by

$$w_1 = f(x, y, z) = c_1$$

$$w_2 = g(x, y, z) = c_2$$

$$w_3 = h(x, y, z) = c_3.$$

Definition 2.1 *Three level surfaces, as above, are called orthogonal coordinate surfaces if they satisfy the following two conditions:*

$$\boldsymbol{\nabla} w_1 \cdot \boldsymbol{\nabla} w_2 = \boldsymbol{\nabla} w_1 \cdot \boldsymbol{\nabla} w_3 = \boldsymbol{\nabla} w_2 \cdot \boldsymbol{\nabla} w_3 = 0 \tag{2.1}$$

$$[\boldsymbol{\nabla} w_1, \boldsymbol{\nabla} w_2, \boldsymbol{\nabla} w_3] > 0 \tag{2.2}$$

What we consider are families of such level surfaces allowing the constants c_1, c_2 and c_3 to vary within the range of f, g and h respectively.

As an example, consider the following three level surfaces:

$$r = f(x, y, z) = \sqrt{x^2 + y^2 + z^2} = c_1$$

$$\theta = g(x, y, z) = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = c_2$$

$$\phi = h(x, y, z) = \tan^{-1}\left(\frac{y}{x}\right) = c_3$$

The first level surface gives us a family of spheres centered at the origin, the second level surface gives us a family of planes making an angle θ with the z-axis (with θ between 0 and π) and the third level set gives us a family of planes making an angle ϕ with the positive x-axis. We will ignore singularities, which essentially have measure zero. With this convention, up to a null set we can tessellate all of $\mathbb{R}^3$ with these level surfaces. These surfaces will be connected with spherical coordinates, as we shall soon see.

Before looking at the intersections of these surfaces, we introduce one last definition.

Definition 2.2 *The level surface $w_1 = f(x, y, z) = c_1$ is called the w_1-coordinate surface for c_1.*

We have analogous definitions for $w_2 = g(x, y, z) = c_2$ and $w_3 = h(x, y, z) = c_3$.

2.2 Orthogonal Coordinate Curves

Note that any two orthogonal coordinate surfaces intersect in a curve. These curves are called *orthogonal coordinate curves.* In particular, we fix the following definitions.

Definition 2.3 *The w_1-coordinate curve for (c_2, c_3) is the intersection of the w_2-coordinate surface for c_2 and the w_3-coordinate surface for c_3.*

Definition 2.4 *The w_2-coordinate curve for (c_1, c_3) is the intersection of the w_1-coordinate surface for c_1 and the w_3-coordinate surface for c_3.*

Definition 2.5 *The w_3-coordinate curve for (c_1, c_2) is the intersection of the w_1-coordinate surface for c_1 and the w_2-coordinate surface for c_2.*

Since these curves line in the orthogonal coordinate surfaces, they intersect orthogonally.

2.3 Orthogonal Curvilinear Coordinates

All three orthogonal coordinate surfaces and all three orthogonal coordinate curves intersect in a single point P. The point P is uniquely specified by the constants c_1, c_2 and c_3. Therefore, the point P is uniquely specified by the values of w_1, w_2 and w_3.

Definition 2.6 *If $w_1 = f(x, y, z)$, $w_2 = g(x, y, z)$ and $w_3 = h(x, y, z)$ are functions whose level sets are orthogonal coordinate surfaces, then for a point $P \in \mathbb{R}^3$ we say that the triple (w_1, w_2, w_3) that uniquely specifies P represents P in orthogonal curvilinear coordinates.*

In modern terminology, the three functions f, g and h are a (local) coordinate system for $\mathbb{R}^3$. Condition 2.2 guarantees that the coordinate system is right-handed. If one considers the example above, we recover our usual system of spherical coordinates.

Finally, note that at each point P, the orthogonal coordinate curves form a set of “orthogonal coordinate axes” emanating from P.

3 Tangent Vectors and Lamé Coefficients

3.1 Orthogonal Coordinate Curves II

We start with the vector field of position. This is a vector field defined on $\mathbb{R}^3$ in Cartesian coordinates.

Definition 3.1 *The vector field of position $\boldsymbol{R}$ in $\mathbb{R}^3$ is defined with respect to the standard basis $\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$ by*

$$\boldsymbol{R} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}.$$

We assume that we have orthogonal curvilinear coordinates given by three functions

$$w_1 = f(x, y, z)$$

$$w_2 = g(x, y, z)$$

$$w_3 = h(x, y, z)$$

that can be inverted at a point P. This allows us to solve this system for x, y and z to obtain

$$x = k_1(w_1, w_2, w_3)$$

$$y = k_2(w_1, w_2, w_3)$$

$$z = k_3(w_1, w_2, w_3)$$

where k_i is continuously differentiable for $i = 1, 2, 3$. It follows that we can rewrite the vector field of position in the given orthogonal curvilinear coordinates as

$$\boldsymbol{R} = k_1(w_1, w_2, w_3)\hat{\boldsymbol{\imath}} + k_2(w_1, w_2, w_3)\hat{\boldsymbol{\jmath}} + k_3(w_1, w_2, w_3)\hat{\boldsymbol{k}}.$$

Thus $\boldsymbol{R}$ is a vector-valued function given by

$$\boldsymbol{R} = \boldsymbol{K}(w_1, w_2, w_3).$$

Continuing with the example of spherical coordinates, we have

$$x = k_1(r, \theta, \phi) = r \sin\theta \cos\phi$$

$$y = k_2(r, \theta, \phi) = r \sin\theta \sin\phi$$

$$z = k_3(r, \theta, \phi) = r \cos\theta$$

and the vector field of position is given by

$$\boldsymbol{R} = r \sin\theta \cos\phi \hat{\boldsymbol{\imath}} + r \sin\theta \sin\phi \hat{\boldsymbol{\jmath}} + r \cos\theta \hat{\boldsymbol{k}}.$$

We can now use the vector field of position to obtain explicit formulae for the orthogonal coordinate curves. Specifically, if we write

$$\boldsymbol{R} = \boldsymbol{K}(w_1, w_2, w_3)$$

then the w_1-coordinate curve for (c_2, c_3) is given by

$$\boldsymbol{S}_1(w_1) = \boldsymbol{K}(w_1, c_2, c_3)$$

the w_2-coordinate curve for (c_1, c_3) is given by

$$\boldsymbol{S}_2(w_2) = \boldsymbol{K}(c_1, w_2, c_3)$$

and the w_3-coordinate curve for (c_1, c_2) is given by

$$\boldsymbol{S}_3(w_3) = \boldsymbol{K}(c_1, c_2, w_3).$$

3.2 Tangent Vectors

Consider the w_1-coordinate curve for (c_2, c_3) given by $\boldsymbol{S}_1$ above. Then the vector

$$\boldsymbol{S}_1'(w_1) = \frac{\mathrm{d}}{\mathrm{d}w_1}\left(\boldsymbol{K}(w_1, c_2, c_3)\right)$$

is tangent to the curve, but does not necessarily have unit length. This is because $\boldsymbol{S}_1$ is not necessarily parameterized by arc length. It follows that

$$\frac{\partial}{\partial w_1}\boldsymbol{K}(w_1, w_2, w_3) = \frac{\partial \boldsymbol{R}}{\partial w_1}$$

is a vector tangent to any w_1-coordinate curve, but may not necessarily have unit length. Similarly, $\dfrac{\partial \boldsymbol{R}}{\partial w_2}$ and $\dfrac{\partial \boldsymbol{R}}{\partial w_3}$ are vectors tangent to any w_2 and w_3 coordinate curves respectively, but do not necessarily have unit length.

We define the following in accordance with the discussion above:

$$\boldsymbol{h}_1 = \frac{\partial R}{\partial w_1}$$

$$\boldsymbol{h}_2 = \frac{\partial R}{\partial w_2}$$

and

$$\boldsymbol{h}_3 = \frac{\partial R}{\partial w_3}.$$

In our example with spherical coordinates:

$$\boldsymbol{h}_1 = \sin\theta\cos\phi\hat{\boldsymbol{\imath}} + \sin\theta\sin\phi\hat{\boldsymbol{\jmath}} + \cos\theta\hat{\boldsymbol{k}}$$

$$\boldsymbol{h}_2 = r\cos\theta\cos\phi\hat{\boldsymbol{\imath}} + r\cos\theta\sin\phi\hat{\boldsymbol{\jmath}} - r\sin\theta\hat{\boldsymbol{k}}$$
$$\boldsymbol{h}_3 = -r\sin\theta\sin\phi\hat{\boldsymbol{\imath}} + r\sin\theta\cos\phi\hat{\boldsymbol{\jmath}}.$$

3.3 Lamé Coefficients

Define the following:

$$h_1 = |\boldsymbol{h}_1| = \left|\frac{\partial \boldsymbol{R}}{\partial w_1}\right|$$

$$h_2 = |\boldsymbol{h}_2| = \left|\frac{\partial \boldsymbol{R}}{\partial w_2}\right|$$

and

$$h_3 = |\boldsymbol{h}_3| = \left|\frac{\partial \boldsymbol{R}}{\partial w_3}\right|.$$

Definition 3.2 *The quantities h_1, h_2 and h_3 above are called the Lamé coefficients for the orthogonal curvilinear coordinate system with coordinates (w_1, w_2, w_3).*

Note that since $\boldsymbol{h}_1$, $\boldsymbol{h}_2$ and $\boldsymbol{h}_3$ do not necessarily have unit or even constant length, the Lamé coefficients are generally functions on $\mathbb{R}^3$.

Let us compute the Lamé coefficients for spherical coordinates. Using the results previously obtained:

$$\begin{aligned}
h_1 &= |\boldsymbol{h}_1| \\
&= \sqrt{\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta} \\
&= \sqrt{\sin^2\theta(\cos^2\phi + \sin^2\phi) + \cos^2\theta} \\
&= \sqrt{\sin^2\theta + \cos^2\theta} \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
h_2 &= |\boldsymbol{h}_2| \\
&= \sqrt{r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta} \\
&= \sqrt{r^2[\cos^2\theta(\cos^2\phi + \sin^2\phi) + \sin^2\theta]} \\
&= \sqrt{r^2(\cos^2\theta + \sin^2\theta)} \\
&= \sqrt{r^2} \\
&= r
\end{aligned}$$

since we always assume $r \geq 0$. Finally:

$$\begin{aligned}
h_3 &= |\boldsymbol{h}_3| \\
&= \sqrt{r^2\sin^2\theta\sin^2\phi + r^2\sin^2\theta\cos^2\phi} \\
&= \sqrt{r^2\sin^2\theta(\sin^2\phi + \cos^2\phi)} \\
&= \sqrt{r^2\sin^2\theta} \\
&= r\sin\theta.
\end{aligned}$$

3.4 Unit Tangent Vectors

We can now obtain unit tangent vectors to the orthogonal coordinate curves by dividing the tangent vectors by the Lamé coefficients. Accordingly we define

$$\hat{\boldsymbol{h}}_1 = \frac{\boldsymbol{h}_1}{h_1}$$

$$\hat{\boldsymbol{h}}_2 = \frac{\boldsymbol{h}_2}{h_2}$$

and

$$\hat{\boldsymbol{h}}_3 = \frac{\boldsymbol{h}_3}{h_3}.$$

Then $\hat{\boldsymbol{h}}_1$, $\hat{\boldsymbol{h}}_2$ and $\hat{\boldsymbol{h}}_3$ are unit tangent vectors to the w_1, w_2 and w_3 coordinate curves, respectively. Note that by condition 2.2, the set $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ is a right-handed set of mutually orthogonal unit vectors. In other words, $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ at a point P form a right-handed orthonormal basis for a copy of $\mathbb{R}^3$ footed at P In modern terminology, these vectors form an orthonormal frame for the tangent bundle of $\mathbb{R}^3$.

Sometimes we use the notation $\hat{\boldsymbol{e}}_{w_i}$ for $\hat{\boldsymbol{h}}_i$. We will do so with the example of spherical coordinates. Dividing each $\boldsymbol{h}_i$ by the corresponding Lamé coefficient h_i we get the following for spherical coordinates:

$$\hat{\boldsymbol{e}}_r = \sin\theta\cos\phi\hat{\boldsymbol{\imath}} + \sin\theta\sin\phi\hat{\boldsymbol{\jmath}} + \cos\theta\hat{\boldsymbol{k}}$$
$$\hat{\boldsymbol{e}}_\theta = \cos\theta\cos\phi\hat{\boldsymbol{\imath}} + \cos\theta\sin\phi\hat{\boldsymbol{\jmath}} - \sin\theta\hat{\boldsymbol{k}}$$

and

$$\hat{\boldsymbol{e}}_\phi = -\sin\phi\hat{\boldsymbol{\imath}} + \cos\phi\hat{\boldsymbol{\jmath}}.$$

3.5 Orthogonal Coordinates Test

Suppose we have a general curvilinear coordinate system where we have defined

$$\boldsymbol{h}_1 = \frac{\partial \boldsymbol{R}}{\partial w_1}, \quad \boldsymbol{h}_2 = \frac{\partial \boldsymbol{R}}{\partial w_2}, \quad \boldsymbol{h}_3 = \frac{\partial \boldsymbol{R}}{\partial w_3}.$$

If the vectors $\{\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3\}$ are mutually orthogonal, then the coordinate curves are mutually orthogonal and hence the coordinate surfaces are mutually orthogonal. Then with the appropriate handedness, the coordinate system is orthogonal. If the handedness is not correct, we can interchange any pair of indices to obtain a right-handed system.

This gives us the orthogonal coordinates test. Given a set of equations

$$\begin{aligned} x &= k_1(w_1, w_2, w_3) \\ y &= k_2(w_1, w_2, w_3) \\ z &= k_3(w_1, w_2, w_3) \end{aligned}$$

that define a coordinate system, one can check to see if (w_1, w_2, w_3) are orthogonal by checking to see if

$$\frac{\partial \boldsymbol{R}}{\partial w_1}\cdot\frac{\partial \boldsymbol{R}}{\partial w_2} = \frac{\partial \boldsymbol{R}}{\partial w_1}\cdot\frac{\partial \boldsymbol{R}}{\partial w_3} = \frac{\partial \boldsymbol{R}}{\partial w_2}\cdot\frac{\partial \boldsymbol{R}}{\partial w_3} = 0.$$

One can check fairly easily now that spherical coordinates are orthogonal curvilinear coordinates.

As another example, consider paraboloidal coordinates in $\mathbb{R}^3$ defined by the equations

$$\begin{aligned} x &= uv\cos\phi \\ y &= uv\sin\phi \\ z &= \frac{1}{2}u^2 - \frac{1}{2}v^2 \end{aligned}$$

where $u, v \geq 0$ and $\phi \in [0, 2\pi)$. In such coordinates the vector field of position is

$$\boldsymbol{R} = (uv\cos\phi)\hat{\boldsymbol{\imath}} + (uv\sin\phi)\hat{\boldsymbol{\jmath}} + \left(\frac{1}{2}u^2 - \frac{1}{2}v^2\right)\hat{\boldsymbol{k}}.$$

We compute $\boldsymbol{h}_i$ for $i = 1, 2, 3$. To wit:

$$\begin{aligned}\boldsymbol{h}_1 &= \frac{\partial \boldsymbol{R}}{\partial u} \\ &= (v\cos\phi)\hat{\boldsymbol{\imath}} + (v\sin\phi)\hat{\boldsymbol{\jmath}} - v\hat{\boldsymbol{k}}\end{aligned}$$

$$\begin{aligned}\boldsymbol{h}_2 &= \frac{\partial \boldsymbol{R}}{\partial v} \\ &= (u\cos\phi)\hat{\boldsymbol{\imath}} + (u\sin\phi)\hat{\boldsymbol{\jmath}} + u\hat{\boldsymbol{k}}\end{aligned}$$

$$\begin{aligned}\boldsymbol{h}_3 &= \frac{\partial \boldsymbol{R}}{\partial \phi} \\ &= (-uv\sin\phi)\hat{\boldsymbol{\imath}} + (uv\cos\phi)\hat{\boldsymbol{\jmath}}.\end{aligned}$$

We now check orthogonality:

$$\begin{aligned}\boldsymbol{h}_1 \cdot \boldsymbol{h}_2 &= uv\cos^2\phi + uv\sin^2\phi - uv \\ &= uv - uv \\ &= 0\end{aligned}$$

$$\begin{aligned}\boldsymbol{h}_1 \cdot \boldsymbol{h}_3 &= -uv^2\cos\phi\sin\phi + uv^2\cos\phi\sin\phi \\ &= 0\end{aligned}$$

$$\begin{aligned}\boldsymbol{h}_2 \cdot \boldsymbol{h}_3 &= -u^2v\cos\phi\sin\phi + u^2v\cos\phi\sin\phi \\ &= 0.\end{aligned}$$

Therefore the paraboloidal coordinate system, up to handedness, is an orthogonal curvilinear coordinate system. Note that the Lamé coefficients are $h_1 = (\sqrt{2})v$, $h_2 = (\sqrt{2})u$, and $h_3 = uv$.

3.6 Integrating in Orthogonal Curvilinear Coordinates

For simplicity, we will just consider integrating a function $f : \mathbb{R}^3 \to \mathbb{R}$ over a compact region $\Omega \subset \mathbb{R}^3$. Suppose we have orthogonal curvilinear coordinates given by equations

$$\begin{aligned}x &= k_1(w_1, w_2, w_3) \\ y &= k_2(w_1, w_2, w_3) \\ z &= k_3(w_1, w_2, w_3).\end{aligned}$$

Then by the change of variables formula:

$$\int_\Omega f(x, y, z)\ \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_{K(\Omega)} f(w_1, w_2, w_3)|\det(J)|\ \mathrm{d}w_1\mathrm{d}w_2\mathrm{d}w_3$$

where J is the Jacobain of the coordinate change $K = (k_1, k_2, k_3)$.

Proposition 3.3 *The determinant* $|\det(J)| = h_1 h_2 h_3$*, where* h_1, h_2 *and* h_3 *are the Lamé coefficients for the coordinate system.*

PROOF: The rows of the matrix J are the vectors $\boldsymbol{h}_i$. It follows that the determinant of J is the triple scalar product $[\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3]$. Hence:

$$|\det(J)| = |[\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3]| = |h_1 h_2 h_3 [\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3]| = h_1 h_2 h_3$$

as claimed. ∎

As a consequence, it follows that we can integrate in orthogonal curvilinear coordinates according to the following:

$$\int_\Omega f(x, y, z)\ \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_{K(\Omega)} f(w_1, w_2, w_3) h_1 h_2 h_3\ \mathrm{d}w_1\mathrm{d}w_2\mathrm{d}w_3.$$

For example, to integrate in spherical coordinates we have

$$\int f(x, y, z)\ \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int f(r, \theta, \phi) r^2 \sin\theta\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$$

and in paraboloidal coordinates we have

$$\int f(x, y, z)\ \mathrm{d}x\mathrm{d}y\mathrm{d}z = 2\int f(u, v, \phi) u^2 v^2\ \mathrm{d}u\mathrm{d}v\mathrm{d}\phi.$$

4 Vectors in Orthogonal Curvilinear Coordinates

4.1 Basis

Since $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ is a right handed set of mutually orthogonal unit vectors, at each point in $P \in \mathbb{R}^3$ they may be used as a basis for a copy of $\mathbb{R}^3$ footed at P. In modern language, $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ is an orthonormal frame for the tangent bundle of $\mathbb{R}^3$.

A word of caution is in order. The vectors $\hat{\boldsymbol{h}}_i$ are not fixed, they vary over the points of $\mathbb{R}^3$. So a vector field given by

$$\boldsymbol{V} = \sum_i v^i \hat{\boldsymbol{h}}_i$$

that is constant in the frame $\hat{\boldsymbol{h}}_i$ is not, in general, a constant vector field. Consider as an example, spherical coordinates. Suppose we have a vector field given by

$$\boldsymbol{V} = 2\hat{\boldsymbol{e}}_r - 3\hat{\boldsymbol{e}}_\theta + 5\hat{\boldsymbol{e}}_\phi.$$

The vector field $\boldsymbol{V}$ is constant in spherical coordinates. In Cartesian coordinates $\boldsymbol{V}$ is given by

$$\begin{aligned}
\boldsymbol{V} &= (2\sin\theta\cos\phi - 3\cos\theta\cos\phi - 5\sin\phi)\hat{\boldsymbol{\imath}} + (2\sin\theta\sin\phi - 3\cos\theta\sin\phi + 5\cos\theta)\hat{\boldsymbol{\jmath}} + (3\cos\theta + 5\sin\theta)\hat{\boldsymbol{k}} \\
&= \frac{1}{\sqrt{(x^2+y^2)(x^2+y^2+z^2)}}\left[\left(2x\sqrt{x^2+y^2} - 3xz - 5y\sqrt{x^2+y^2+z^2}\right)\hat{\boldsymbol{\imath}}\right. \\
&+ \left.\left((2y+5z)\sqrt{x^2+y^2} - 3yz\right)\hat{\boldsymbol{\jmath}} + \left(3x\sqrt{x^2+y^2} + 5(x^2+y^2)\right)\hat{\boldsymbol{k}}\right]
\end{aligned}$$

which is certainly not constant!

4.2 Dot and Cross Products in Orthogonal Curvilinear Coordinates

The dot product and cross product of vectors in orthogonal curvilinear coordinates works the same way it does in Cartesian coordinates.

Since the vectors $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ form an orthogonal basis of unit vectors,

$$\hat{\boldsymbol{h}}_i \cdot \hat{\boldsymbol{h}}_j = \delta_{ij}$$

for every $i, j \in \{1, 2, 3\}$ where δ_{ij} is the Kronecker delta. Therefore for arbitrary vectors $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$\begin{aligned}
\boldsymbol{v} \cdot \boldsymbol{w} &= \left(\sum_i v_i \hat{\boldsymbol{h}}_i\right) \cdot \left(\sum_j w_j \hat{\boldsymbol{h}}_j\right) \\
&= \sum_i \sum_j v_i w_j \hat{\boldsymbol{h}}_i \cdot \hat{\boldsymbol{h}}_j \\
&= \sum_i \sum_j v_i w_j \delta_{ij} \\
&= \sum_i v_i w_i.
\end{aligned}$$

Since $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ is a right-handed system of orthogonal basis vectors, as in the Cartesian case:

$$\hat{\boldsymbol{h}}_i \times \hat{\boldsymbol{h}}_j = \sum_k \epsilon_{ijk} \hat{\boldsymbol{h}}_k$$

where ϵ_{ijk} is the Levi-Cevita tensor:

$$\epsilon_{ijk} = \begin{cases} 1 & : \text{ if } i, j, k \text{ is a cyclic permutation of 1,2,3} \\ -1 & : \text{ if } i, j, k \text{ is an acyclic permutation 1,2,3} \\ 0 & : \text{ otherwise} \end{cases}$$

Then for arbitrary vectors $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$\begin{aligned}
\boldsymbol{v} \times \boldsymbol{w} &= \left(\sum_i v_i \hat{\boldsymbol{h}}_i\right) \times \left(\sum_j w_j \hat{\boldsymbol{h}}_j\right) \\
&= \sum_i \sum_j v_i w_j \hat{\boldsymbol{h}}_i \times \hat{\boldsymbol{h}}_j \\
&= \sum_i \sum_j \sum_k \epsilon_{ijk} v_i w_j \hat{\boldsymbol{h}}_k
\end{aligned}$$

exactly as in the Cartesian case.

4.3 Relationship to the Cartesian Basis

In this section we will write the vectors $\hat{\boldsymbol{h}}_i$ in Cartesian coordinates and we will write the Cartesian basis vectors in the new basis $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$. In order to do the manipulations with the indices, the Cartesian basis will be written as $\{\hat{\boldsymbol{\delta}}_1, \hat{\boldsymbol{\delta}}_2, \hat{\boldsymbol{\delta}}_3\}$ with the obvious correspondence with $\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$.

Proposition 4.1 *The vectors* $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ *are written in the Cartesian basis as*

$$\hat{\boldsymbol{h}}_i = \sum_j \frac{1}{h_i} \frac{\partial x_j}{\partial w_i} \hat{\boldsymbol{\delta}}_j.$$

Proof: By definition we have

$$\hat{\boldsymbol{h}}_i = \frac{\boldsymbol{h}_i}{h_i} = \frac{1}{h_i} \frac{\partial \boldsymbol{R}}{\partial w_i} = \frac{1}{h_i} \frac{\partial}{\partial w_i} \left(\sum_j x_j \hat{\boldsymbol{\delta}}_j \right).$$

Since the Cartesian basis vectors $\hat{\boldsymbol{\delta}}_j$ are constant

$$\frac{\partial}{\partial w_i} \left(\sum_j x_j \hat{\boldsymbol{\delta}}_j \right) = \sum_j \frac{\partial x_j}{\partial w_i} \hat{\boldsymbol{\delta}}_j.$$

The desired result follows. ■

Proposition 4.2 *The Cartesian basis vectors are written in the basis* $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ *as*

$$\hat{\boldsymbol{\delta}}_i = \sum_j \frac{1}{h_j} \frac{\partial x_i}{\partial w_j} \hat{\boldsymbol{h}}_j.$$

Proof: We write

$$\hat{\boldsymbol{\delta}}_i = \sum_j v_j \hat{\boldsymbol{h}}_j$$

and solve for v_j. Since $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ is an orthonormal basis,

$$v_k = \hat{\boldsymbol{\delta}}_i \cdot \hat{\boldsymbol{h}}_k.$$

Using Proposition 4.1, it follows that

$$\begin{aligned} \hat{\boldsymbol{\delta}}_i \cdot \hat{\boldsymbol{h}}_k &= \hat{\boldsymbol{\delta}}_i \cdot \left(\sum_j \frac{1}{h_k} \frac{\partial x_j}{\partial w_k} \hat{\boldsymbol{\delta}}_j \right) \\ &= \sum_j \frac{1}{h_k} \frac{\partial x_j}{\partial w_k} \hat{\boldsymbol{\delta}}_i \cdot \hat{\boldsymbol{\delta}}_j \\ &= \sum_j \frac{1}{h_k} \frac{\partial x_j}{\partial w_k} \delta_{ij} \\ &= \frac{1}{h_k} \frac{\partial x_i}{\partial w_k}. \end{aligned}$$

The desired result follows. ■

Remark 4.3 *Propositions 4.1 and 4.2 look very similar. The difference lies in the indexing. One should carefully keep track of indices when doing these calculations.*

5 Partial Derivative Relationships

5.1 Introduction

Since the frame $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ varies from point to point, we need to understand how these and related quantities change. Our ultimate goal is to express the partial derivatives of $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ in terms of $\{\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3\}$ and the partial derivatives of the Lamé coefficients. This will be necessary to obtain the expressions for differential operators in orthogonal curvilinear coefficients. This is where the derivations for specific coordinate systems, such as spherical coordinates, becomes potentially painful. Here all of the pain will be done once, then we can use the results for **any** orthogonal curvilinear coordinates.

5.2 Partial Derivatives of Lamé Coefficients

The first step is to compute the partial derivatives of the Lamé coefficients.

Lemma 5.1 *Suppose $\{h_1, h_2, h_3\}$ are Lamé coefficients for orthogonal curvilinear coordinates $\{w_1, w_2, w_3\}$ in $\mathbb{R}^3$ where $\{x_1, x_2, x_3\}$ are the usual Cartesian coordinates. Then*

$$\frac{\partial h_i}{\partial w_j} = \sum_k \frac{1}{h_i} \left(\frac{\partial x_k}{\partial w_i}\right) \frac{\partial^2 x_k}{\partial w_j \partial w_i}.$$

Proof: Using the definition:

$$h_i = \left|\frac{\partial \boldsymbol{R}}{\partial w_i}\right|$$

we compute:

$$\begin{aligned}
\frac{\partial h_i}{\partial w_j} &= \frac{\partial}{\partial w_j}\left(\left|\frac{\partial \boldsymbol{R}}{\partial w_i}\right|\right) \\
&= \frac{\partial}{\partial w_j}\left(\left|\sum_k \frac{\partial x_k}{\partial w_i} \hat{\boldsymbol{\partial}}_k\right|\right) \\
&= \frac{\partial}{\partial w_j}\left[\left(\sum_k \left(\frac{\partial x_k}{\partial w_i}\right)^2\right)^{1/2}\right] \\
&= \frac{1}{2}\left[\sum_k \left(\frac{\partial x_k}{\partial w_i}\right)^2\right]^{-1/2} \frac{\partial}{\partial w_j}\left[\sum_k \left(\frac{\partial x_k}{\partial w_i}\right)^2\right] \\
&= \frac{1}{2\left[\sum_k \left(\frac{\partial x_i}{\partial w_k}\right)^2\right]^{1/2}} \sum_k \frac{\partial}{\partial w_k}\left[\left(\frac{\partial x_k}{\partial w_j}\right)^2\right] \\
&= \frac{1}{2h_i}\sum_k 2\left(\frac{\partial x_k}{\partial w_i}\right)\frac{\partial^2 x_k}{\partial w_j \partial w_i} \\
&= \frac{1}{h_i}\sum_k \left(\frac{\partial x_k}{\partial w_i}\right)\frac{\partial^2 x_k}{\partial w_j \partial w_i}
\end{aligned}$$

as claimed. ∎

5.3 Cyclic Partial Relation

Lemma 5.2 *If i, j, k are distinct indices:*

$$\boldsymbol{h}_i \cdot \frac{\partial \boldsymbol{h}_j}{\partial w_k} = 0.$$

Proof: Since the indices are all distinct:

$$\boldsymbol{h}_i \cdot \boldsymbol{h}_j = \boldsymbol{h}_i \cdot \boldsymbol{h}_k = \boldsymbol{h}_j \cdot \boldsymbol{h}_k = 0.$$

Differentiating we get

$$\begin{aligned}
\frac{\partial}{\partial w_k}(\boldsymbol{h}_i \cdot \boldsymbol{h}_j) &= 0. \\
\frac{\partial}{\partial w_j}(\boldsymbol{h}_i \cdot \boldsymbol{h}_k) &= 0. \\
\frac{\partial}{\partial w_i}(\boldsymbol{h}_j \cdot \boldsymbol{h}_k) &= 0.
\end{aligned}$$

Using the definitions, this means

$$\begin{aligned}
\frac{\partial}{\partial w_k}\left(\frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial \boldsymbol{R}}{\partial w_j}\right) &= 0. \\
\frac{\partial}{\partial w_j}\left(\frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial \boldsymbol{R}}{\partial w_k}\right) &= 0. \\
\frac{\partial}{\partial w_i}\left(\frac{\partial \boldsymbol{R}}{\partial w_j} \cdot \frac{\partial \boldsymbol{R}}{\partial w_k}\right) &= 0.
\end{aligned}$$

Using the product rule (which holds over dot products) we get the following system of equations:

$$\begin{aligned}
\frac{\partial^2 \boldsymbol{R}}{\partial w_k \partial w_i} \cdot \frac{\partial \boldsymbol{R}}{\partial w_j} + \frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_k \partial w_j} &= 0. \quad (5.1) \\
\frac{\partial^2 \boldsymbol{R}}{\partial w_j \partial w_i} \cdot \frac{\partial \boldsymbol{R}}{\partial w_k} + \frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_j \partial w_k} &= 0. \quad (5.2) \\
\frac{\partial^2 \boldsymbol{R}}{\partial w_i \partial w_j} \cdot \frac{\partial \boldsymbol{R}}{\partial w_k} + \frac{\partial \boldsymbol{R}}{\partial w_j} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_i \partial w_k} &= 0. \quad (5.3)
\end{aligned}$$

Since we have sufficient smoothness, equality of mixed partial derivatives and equation (5.3) gives us

$$\frac{\partial^2 \boldsymbol{R}}{\partial w_j \partial w_i} \cdot \frac{\partial \boldsymbol{R}}{\partial w_k} = -\frac{\partial \boldsymbol{R}}{\partial w_j} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_i \partial w_k}. \quad (5.4)$$

Substituting equation (5.4) into (5.2) gives us

$$-\frac{\partial \boldsymbol{R}}{\partial w_j} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_i \partial w_k} + \frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_k \partial w_j} = 0. \quad (5.5)$$

Adding equations (5.1) and (5.5), invoking equality of mixed partial derivatives again, we get

$$2\frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial^2 \boldsymbol{R}}{\partial w_k \partial w_j} = 0.$$

Rewriting we get

$$\frac{\partial \boldsymbol{R}}{\partial w_i} \cdot \frac{\partial}{\partial w_k}\left(\frac{\partial \boldsymbol{R}}{\partial w_j}\right) = 0.$$

Unraveling the definitions, this says that

$$\boldsymbol{h}_i \cdot \frac{\partial \boldsymbol{h}_j}{\partial w_k} = 0$$

as desired. ■

5.4 Partial Derivatives of Unit Vectors

We now have the necessary tools to satisfy our short term goal. Recall that we wish to obtain expressions for the partial derivatives of the unit vectors $\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3$ in terms of $\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \hat{\boldsymbol{h}}_3$ and the partial derivatives of the Lamé coefficients. In the following proposition, we will state the desired expression, and the proof will be a very painful calculation. In other words, ITS GO TIME!

Proposition 5.3 *For each $i, j = 1, 2, 3$:*

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_j - \delta_{ij}\sum_k \frac{1}{h_k}\frac{\partial h_i}{\partial w_k}\hat{\boldsymbol{h}}_k.$$

PROOF: Bring it on!! We start by doing basic computations using the definition of $\boldsymbol{h}_i$:

$$\begin{aligned}
\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= \frac{\partial}{\partial w_j}\left(\frac{\boldsymbol{h}_i}{h_i}\right) \\
&= \frac{\partial}{\partial w_j}\left(\frac{1}{h_i}\right)\boldsymbol{h_i} + \frac{1}{h_i}\frac{\partial \boldsymbol{h}_i}{\partial w_j} \\
&= -\frac{1}{h_i^2}\frac{\partial h_i}{\partial w_j}\boldsymbol{h}_i + \frac{1}{h_i}\frac{\partial^2 \boldsymbol{R}}{\partial w_j \partial w_i} \\
&= -\frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k \frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{\delta}}_k.
\end{aligned}$$

By Proposition 4.2:

$$\hat{\boldsymbol{\delta}}_k = \sum_l \frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\hat{\boldsymbol{h}}_l.$$

Therefore we have

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = -\frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k \sum_l \frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_l.$$

Add and subtract the Kronecker delta δ_{il} in the second term to obtain:

$$\begin{aligned}
\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= -\frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k\sum_l \delta_{il}\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_l + \frac{1}{h_i}\sum_k\sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_l \\
&= -\frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k \frac{1}{h_i}\frac{\partial x_k}{\partial w_i}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k\sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j w_i}\hat{\boldsymbol{h}}_l.
\end{aligned}$$

By Lemma 5.1, this gives us

$$\begin{aligned}\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= -\frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i + \frac{1}{h_i}\sum_k\sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j w_i}\hat{\boldsymbol{h}}_l \\ &= \frac{1}{h_i}\sum_k\sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j w_i}\hat{\boldsymbol{h}}_l.\end{aligned}$$

Therefore:

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\sum_k\sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j w_i}\hat{\boldsymbol{h}}_l. \tag{5.6}$$

There are now two cases: $i = j$ and $i \neq j$.

Case 1: $i \neq j$

This is the easy case. When $i = l$, $1 - \delta_{il} = 0$ and when $i \neq l$, $1 - \delta_{il} = 1$. Therefore equation (5.6) gives us

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\sum_k\sum_{l\neq i} \frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_l.$$

Since $i \neq j$, there are two values of l that are not i, one of which is j. Call the other value m. Then we get

$$\begin{aligned}\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= \frac{1}{h_i}\sum_k\left(\frac{1}{h_j}\frac{\partial x_k}{\partial w_j}\frac{\partial^2 x_k}{\partial w_i \partial w_j}\hat{\boldsymbol{h}}_j + \frac{1}{h_m}\frac{\partial x_k}{\partial w_m}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_m\right) \\ &= \frac{1}{h_i}\left(\sum_k \frac{1}{h_j}\frac{\partial x_k}{\partial w_j}\frac{\partial^2 x_k}{\partial w_i \partial w_j}\right)\hat{\boldsymbol{h}}_j + \frac{1}{h_i h_m}\left(\sum_k \frac{\partial x_k}{\partial w_m}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\right)\hat{\boldsymbol{h}}_m.\end{aligned}$$

By Lemma 5.1, the first term is

$$\frac{1}{h_i}\left(\sum_k \frac{1}{h_j}\frac{\partial x_k}{\partial w_j}\frac{\partial^2 x_k}{\partial w_i \partial w_j}\right)\hat{\boldsymbol{h}}_j = \frac{1}{h_i}\frac{\partial h_j}{\partial w_i}.$$

By Lemma 5.2, since i,j and m are all distinct, the second term is

$$\begin{aligned}\frac{1}{h_i hm}\left(\sum_k \frac{\partial x_k}{\partial w_m}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\right)\hat{\boldsymbol{h}}_m &= \frac{1}{h_i h_m}\left[\frac{\partial \boldsymbol{R}}{\partial w_m}\cdot\frac{\partial}{\partial w_j}\left(\frac{\partial \boldsymbol{R}}{\partial w_i}\right)\right]\hat{\boldsymbol{h}}_m \\ &= \frac{1}{h_i h_m}\left(\boldsymbol{h}_m \cdot \frac{\partial \boldsymbol{h}_i}{\partial w_j}\right)\hat{\boldsymbol{h}}_m \\ &= 0.\end{aligned}$$

Therefore, when $i \neq j$:

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_j. \tag{5.7}$$

Case 2: $i = j$

We start from equation (5.6):

$$\begin{aligned}
\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= \frac{1}{h_i}\sum_k \sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_j \partial w_i}\hat{\boldsymbol{h}}_l \\
&= \frac{1}{h_i}\sum_k \sum_l (1-\delta_{il})\frac{1}{h_l}\frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_i^2}\hat{\boldsymbol{h}}_l \\
&= \frac{1}{h_i}\sum_l (1-\delta_{il})\frac{1}{h_l}\left[\sum_k \frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_i^2}\right]\hat{\boldsymbol{h}}_l.
\end{aligned}$$

Observe that the product rule gives us

$$\frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right) = \frac{\partial^2 x_k}{\partial w_i \partial w_l}\frac{\partial x_k}{\partial w_i} + \frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial x_i^2}.$$

Accordingly:

$$\begin{aligned}
\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= \frac{1}{h_i}\sum_l (1-\delta_{il})\frac{1}{h_l}\left[\sum_k \frac{\partial x_k}{\partial w_l}\frac{\partial^2 x_k}{\partial w_i^2}\right]\hat{\boldsymbol{h}}_l \\
&= \frac{1}{h_i}\sum_l (1-\delta_{il})\frac{1}{h_l}\left[\sum_k \left(\frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right) - \frac{\partial^2 x_k}{\partial w_i \partial w_l}\frac{\partial x_k}{\partial w_i}\right)\right]\hat{\boldsymbol{h}}_l \\
&= \sum_l (1-\delta_{il})\frac{1}{h_l}\left[\frac{1}{h_i}\sum_k \frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right) - \sum_k \frac{1}{h_i}\frac{\partial x_k}{\partial w_i}\frac{\partial^2 x_k}{\partial w_l \partial w_i}\right]\hat{\boldsymbol{h}}_l \\
&= \sum_l (1-\delta_{il})\frac{1}{h_i h_l}\left[\sum_k \frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right)\right]\hat{\boldsymbol{h}}_l + \sum_l (\delta_{il}-1)\frac{1}{h_l}\left[\sum_k \frac{1}{h_i}\frac{\partial x_k}{\partial w_i}\frac{\partial^2 x_k}{\partial w_l \partial w_i}\right]\hat{\boldsymbol{h}}_l.
\end{aligned}$$

I claim that the first term is zero. Observe that:

$$\sum_k \frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right) = \frac{\partial}{\partial w_i}\sum_k \frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i} = \frac{\partial}{\partial w_i}\left(\frac{\partial \boldsymbol{R}}{\partial w_l}\cdot\frac{\partial \boldsymbol{R}}{\partial w_i}\right).$$

Since we are working in orthogonal curvilinear coordinates, if $i \neq l$

$$\frac{\partial \boldsymbol{R}}{\partial w_l}\cdot\frac{\partial \boldsymbol{R}}{\partial w_i} = 0.$$

In particular, since for every index i

$$\frac{\partial \boldsymbol{R}}{\partial w_i} = \boldsymbol{h}_i = h_i \hat{\boldsymbol{h}}_i,$$

it follows that

$$\frac{\partial \boldsymbol{R}}{\partial w_l}\cdot\frac{\partial \boldsymbol{R}}{\partial w_i} = h_i h_l \delta_{il}.$$

Therefore the first term is given by

$$\begin{aligned}\sum_l (1-\delta_{il})\frac{1}{h_i h_l}\left[\sum_k \frac{\partial}{\partial w_i}\left(\frac{\partial x_k}{\partial w_l}\frac{\partial x_k}{\partial w_i}\right)\right]\hat{\boldsymbol{h}}_l &= \sum_l (1-\delta_{il})\frac{1}{h_i h_l}\left[\frac{\partial}{\partial w_i}(h_i h_l \delta_{il})\right]\hat{\boldsymbol{h}}_l \\ &= \sum_l (1-\delta_{il})\delta_{il}\frac{1}{h_i h_l}\left[\frac{\partial}{\partial w_i}(h_i h_l)\right]\hat{\boldsymbol{h}}_l.\end{aligned}$$

Since $1-\delta_{il}=0$ when $i=l$ and $\delta_{il}=0$ when $i\neq l$, the first term is zero as claimed.

This leaves us with

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \sum_l (\delta_{il}-1)\frac{1}{h_l}\left[\sum_k \frac{1}{h_i}\frac{\partial x_k}{\partial w_i}\frac{\partial^2 x_k}{\partial w_l \partial w_i}\right]\hat{\boldsymbol{h}}_l.$$

By Lemma 5.1:

$$\sum_k \frac{1}{h_i}\frac{\partial x_k}{\partial w_i}\frac{\partial^2 x_k}{\partial w_l \partial w_i} = \frac{\partial h_i}{\partial w_l}.$$

Therefore

$$\begin{aligned}\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} &= \sum_l (\delta_{il}-1)\frac{1}{h_l}\frac{\partial h_i}{\partial w_l}\hat{\boldsymbol{h}}_l \\ &= \sum_l \delta il \frac{1}{h_l}\frac{\partial h_i}{\partial w_l}\hat{\boldsymbol{h}}_l - \sum_l \frac{1}{h_l}\frac{\partial h_i}{\partial w_l} \\ &= \frac{1}{h_i}\frac{\partial h_i}{\partial w_i}\hat{\boldsymbol{h}}_i - \sum_l \frac{1}{h_l}\frac{\partial h_i}{\partial w_l}\hat{\boldsymbol{h}}_l.\end{aligned}$$

Change the index over which we are summing in the second term to k, and change a couple of i's in the first term to j's (since $i=j$). This gives us

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_j - \sum_k \frac{1}{h_k}\frac{\partial h_i}{\partial w_k}\hat{\boldsymbol{h}}_k. \tag{5.8}$$

If we compare equations (5.7) and (5.8), the first term on the right hand side of equation (5.8) is exactly the same as the only term on the right hand side of equation (5.7), and we only get the second term on the right hand side of equation (5.8) when $i=j$. It follows that we can summarize both cases by the equation

$$\frac{\partial \hat{\boldsymbol{h}}_i}{\partial w_j} = \frac{1}{h_i}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_j - \delta_{ij}\sum_k \frac{1}{h_k}\frac{\partial h_i}{\partial w_k}\hat{\boldsymbol{h}}_k$$

which was what we set out to prove. ■

We have now established an expression for the partial derivatives of the unit vectors that only depends on the unit vectors themselves and the partial derivatives of the Lamé coefficients. We can now use this result to establish the forms of differential operators in orthogonal curvilinear coordinates.

6 Differential Calculus of Vectors in Orthogonal Curvilinear Coordinates

We are now prepared to tackle the job set out for us. We will compute expressions for the gradient, divergence, curl and Laplacian in orthogonal curvilinear coordinates, where the expressions will be in terms of the Lamé coefficients and their derivatives. At each step, we will compute the appropriate quantity in spherical coordinates.

6.1 The Del Operator

The Del operator will be crucial to our ability to obtain the gradient, divergence, curl and Laplacian. It is therefore fitting that we express the Del operator in orthogonal curvilinear coordinates. In Cartesian coordinates, we have

$$\boldsymbol{\nabla} = \sum_i \hat{\boldsymbol{\delta}}_i \frac{\partial}{\partial x_i}.$$

By Proposition 4.2, we can write $\hat{\boldsymbol{\delta}}_i$ in orthogonal curvilinear coordinates giving us

$$\begin{aligned} \boldsymbol{\nabla} &= \sum_i \left[\sum_j \frac{1}{h_j} \frac{\partial x_i}{\partial w_j} \hat{\boldsymbol{h}}_j \right] \frac{\partial}{\partial x_i} \\ &= \sum_i \sum_j \frac{\hat{\boldsymbol{h}}_j}{h_j} \frac{\partial x_i}{\partial w_j} \frac{\partial}{\partial x_i} \\ &= \sum_j \frac{\hat{\boldsymbol{h}}_j}{h_j} \left(\sum_i \frac{\partial x_i}{\partial w_j} \frac{\partial}{\partial x_i} \right). \end{aligned}$$

By the chain rule:

$$\frac{\partial}{\partial w_j} = \sum_i \frac{\partial x_i}{\partial w_j} \frac{\partial}{\partial x_i}$$

and therefore, changing indices from j to i we get

$$\boldsymbol{\nabla} = \sum_i \frac{\hat{\boldsymbol{h}}_i}{h_i} \frac{\partial}{\partial w_i}.$$

6.2 The Gradient of a Scalar Function

Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is a smooth function. We will compute $\boldsymbol{\nabla} f$ in orthogonal curvilinear coordinates. Using the expression for $\boldsymbol{\nabla}$ in orthogonal curvilinear coordinates:

$$\boldsymbol{\nabla} f = \left(\sum_i \frac{\hat{\boldsymbol{h}}_i}{h_i} \frac{\partial}{\partial w_i} \right) f = \sum_i \frac{1}{h_i} \frac{\partial f}{\partial w_i} \hat{\boldsymbol{h}}_i.$$

As an example, consider spherical coordinates. Using the Lamé coefficients $h_1 = 1, h_2 = r$ and $h_3 = r \sin\theta$ we get:

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial r} \hat{\boldsymbol{e}}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\boldsymbol{e}}_\theta + \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \hat{\boldsymbol{e}}_\phi.$$

This is the same expression given in the introduction.

6.3 Coordinate Surface Normals

In order to handle the divergence of a vector field, we need a brief detour about coordinate surface normals. In order to obtain the tools we need, we have to be able to use the gradient in orthogonal curvilinear coordinates, which we just established.

6.3.1 What are the Coordinate Surface Normals?

Recall that the orthogonal curvilinear coordinates (w_1, w_2, w_3) are each scalar functions of x, y and z. In particular, they define the level surfaces we called the orthogonal coordinate surfaces. Since each w_i is a scalar function, we can compute the gradient:

$$\begin{aligned}\boldsymbol{\nabla} w_i &= \sum_j \frac{1}{h_j}\frac{\partial w_i}{\partial w_j}\hat{\boldsymbol{h}}_j \\ &= \sum_j \frac{1}{h_j}\delta_{ij}\hat{\boldsymbol{h}}_j \\ &= \frac{\hat{\boldsymbol{h}}_i}{h_i}.\end{aligned}$$

The gradient $\boldsymbol{\nabla} w_i$ is normal to the level surface $w_i = c_i$, so we call the collection $\{\boldsymbol{\nabla} w_i\}$ the coordinate surface normals. The computation above establishes that

$$\boldsymbol{\nabla} w_i = \frac{\hat{\boldsymbol{h}}_{\boldsymbol{i}}}{h_i}$$

and

$$|\boldsymbol{\nabla} w_i| = \frac{1}{h_i}.$$

The coordinate surface normals are a mutually orthogonal set of vectors, each of which lies in the same direction as the tangent vectors $\boldsymbol{h}_i$ to the orthogonal coordinate curves.

This last point deserves to be underlined. It is a special feature of the fact that our coordinates are orthogonal. In general, the coordinate surface normals and the tangent vectors to the coordinate curves need not be in the same direction. Instead, these sets of vectors are reciprocal to one another, and this concept simultaneously leads to the notion of dual bases and covariant and contravariant tensors.

6.3.2 Cross Products of Coordinate Surface Normals

We compute the cross products of the coordinate surface normals using the definition:

$$\begin{aligned}\boldsymbol{\nabla} w_1 \times \boldsymbol{\nabla} w_2 &= \frac{\hat{\boldsymbol{h}}_1}{h_1} \times \frac{\hat{\boldsymbol{h}}_2}{h_2} \\ &= \frac{1}{h_1 h_2}\hat{\boldsymbol{h}}_3.\end{aligned}$$

Similarly:

$$\begin{aligned}\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3 &= \frac{1}{h_2 h_3}\hat{\boldsymbol{h}}_1 \\ \boldsymbol{\nabla} w_3 \times \boldsymbol{\nabla} w_1 &= \frac{1}{h_3 h_1}\hat{\boldsymbol{h}}_2.\end{aligned}$$

Another fact we will use, which I will not prove (to keep our detour brief), is that for any distinct i and j, the vector field $\boldsymbol{\nabla} w_i \times \boldsymbol{\nabla} w_j$ is solenoidal. This means that

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} w_i \times \boldsymbol{\nabla} w_j) = 0.$$

6.3.3 Unit Tangent Vector Relations

From the cross product calculations, we get the following useful relations:

$$\begin{aligned}
\hat{\boldsymbol{h}}_1 &= h_2 h_3 \boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3 \\
\hat{\boldsymbol{h}}_2 &= h_3 h_1 \boldsymbol{\nabla} w_3 \times \boldsymbol{\nabla} w_1 \\
\hat{\boldsymbol{h}}_3 &= h_1 h_2 \boldsymbol{\nabla} w_1 \times \boldsymbol{\nabla} w_2.
\end{aligned}$$

6.4 Divergence of a Vector Field

We are now ready to resume our quest. The next object of desire is the divergence of a vector field. Let $\boldsymbol{V} : \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field given in orthogonal curvilinear coordinates by

$$\boldsymbol{V} = \sum_i v_i \hat{\boldsymbol{h}}_i$$

where each v_i is a scalar function. Then:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{V} &= \boldsymbol{\nabla} \cdot \left(\sum_i v_i \hat{\boldsymbol{h}}_i \right) \\
&= \sum_i \boldsymbol{\nabla} \cdot (v_i \hat{\boldsymbol{h}}_i) \\
&= \boldsymbol{\nabla} \cdot (v_1 \hat{\boldsymbol{h}}_1) + \boldsymbol{\nabla} \cdot (v_2 \hat{\boldsymbol{h}}_2) + \boldsymbol{\nabla} \cdot (v_3 \hat{\boldsymbol{h}}_3).
\end{aligned}$$

Using the unit tangent vector relations, and a version of the product rule:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot (v_1 \hat{\boldsymbol{h}}_1) &= \boldsymbol{\nabla} \cdot [v_1 h_2 h_3 (\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3)] \\
&= \boldsymbol{\nabla}(v_1 h_2 h_3) \cdot (\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3) + v_1 h_2 h_3 \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3).
\end{aligned}$$

Since $\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3$ is solenoidal, its divergence is zero and thus the second term is zero. Since

$$\boldsymbol{\nabla} w_2 \times \boldsymbol{\nabla} w_3 = \frac{1}{h_2 h_3} \hat{\boldsymbol{h}}_1$$

we get

$$\boldsymbol{\nabla} \cdot (v_1 \hat{\boldsymbol{h}}_1) = \boldsymbol{\nabla}(v_1 h_2 h_3) \cdot \left(\frac{1}{h_2 h_3} \hat{\boldsymbol{h}}_1 \right).$$

Since $v_1 h_2 h_3$ is a scalar function, using our formula for the gradient in orthogonal curvilinear coordinates:

$$\begin{aligned}
\boldsymbol{\nabla}(v_1 h_2 h_3) \cdot \left(\frac{1}{h_2 h_3} \hat{\boldsymbol{h}}_1 \right) &= \left(\sum_i \frac{1}{h_i} \frac{\partial (v_1 h_2 h_3)}{\partial w_i} \hat{\boldsymbol{h}}_i \right) \cdot \left(\frac{1}{h_2 h_3} \hat{\boldsymbol{h}}_1 \right) \\
&= \frac{1}{h_2 h_3} \sum_i \frac{1}{h_i} \frac{\partial (v_1 h_2 h_3)}{\partial w_i} \hat{\boldsymbol{h}}_i \cdot \hat{\boldsymbol{h}}_1.
\end{aligned}$$

Since the coordinates are orthogonal, $\hat{\boldsymbol{h}}_i \cdot \hat{\boldsymbol{h}}_1 = \delta_{i1}$. Therefore:

$$\begin{aligned}
\frac{1}{h_2 h_3} \sum_i \frac{1}{h_i} \frac{\partial(v_1 h_2 h_3)}{\partial w_i} \hat{\boldsymbol{h}}_i \cdot \hat{\boldsymbol{h}}_1 &= \frac{1}{h_2 h_3} \sum_i \frac{1}{h_i} \frac{\partial(v_1 h_2 h_3)}{\partial w_i} \delta_{i1} \\
&= \frac{1}{h_1 h_2 h_3} \frac{\partial(v_1 h_2 h_3)}{\partial w_1}.
\end{aligned}$$

Thus:

$$\boldsymbol{\nabla} \cdot (v_1 \hat{\boldsymbol{h}}_1) = \frac{1}{h_1 h_2 h_3} \frac{\partial(v_1 h_2 h_3)}{\partial w_1}.$$

Similarly:

$$\boldsymbol{\nabla} \cdot (v_2 \hat{\boldsymbol{h}}_2) = \frac{1}{h_1 h_2 h_3} \frac{\partial(v_2 h_3 h_1)}{\partial w_2}$$

and

$$\boldsymbol{\nabla} \cdot (v_3 \hat{\boldsymbol{h}}_3) = \frac{1}{h_1 h_2 h_3} \frac{\partial(v_3 h_1 h_2)}{\partial w_3}$$

Therefore, our expression for the divergence of a vector field is:

$$\boldsymbol{\nabla} \cdot \boldsymbol{V} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial(v_1 h_2 h_3)}{\partial w_1} + \frac{\partial(v_2 h_3 h_1)}{\partial w_2} + \frac{\partial(v_3 h_1 h_2)}{\partial w_3} \right].$$

Returning to our spherical coordinates:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{V} &= \frac{1}{r^2 \sin\theta} \left[\frac{\partial(r^2 \sin\theta v_r)}{\partial r} + \frac{\partial(r \sin\theta v_\theta)}{\partial \theta} + \frac{\partial(r v_\phi)}{\partial \phi} \right] \\
&= \frac{1}{r^2 \sin\theta} \left[\sin\theta \frac{\partial(r^2 v_r)}{\partial r} + r \frac{\partial(\sin\theta v_\theta)}{\partial \theta} + r \frac{\partial v_\phi}{\partial \phi} \right] \\
&= \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 v_r) + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta}(\sin\theta v_\theta) + \frac{1}{r \sin\theta} \frac{\partial v_\phi}{\partial \phi}.
\end{aligned}$$

This is the same as the expression given in the introduction.

6.5 Curl of a Vector Field

We now discuss the curl of a vector field in orthogonal curvilinear coordinates. Using the definitions:

$$\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{V} &= \left(\sum_i \frac{\hat{\boldsymbol{h}}_i}{h_i} \frac{\partial}{\partial w_i} \right) \times \left(\sum_j v_j \hat{\boldsymbol{h}}_j \right) \\
&= \sum_i \sum_j \frac{\hat{\boldsymbol{h}}_i}{h_i} \times \frac{\partial}{\partial w_i}(v_j \hat{\boldsymbol{h}}_j) \\
&= \sum_i \sum_j \frac{\hat{\boldsymbol{h}}_i}{h_i} \times \left[\frac{\partial v_j}{\partial w_i} \hat{\boldsymbol{h}}_j + v_j \frac{\partial \hat{\boldsymbol{h}}_j}{\partial w_i} \right] \\
&= \sum_i \sum_j \frac{1}{h_i} \frac{\partial v_j}{\partial w_i} \hat{\boldsymbol{h}}_i \times \hat{\boldsymbol{h}}_j + \sum_i \sum_j \frac{v_j}{h_i} \hat{\boldsymbol{h}}_i \times \frac{\partial \hat{\boldsymbol{h}}_j}{\partial w_i}.
\end{aligned}$$

The first term is given by

$$\begin{aligned}\sum_i\sum_j \frac{1}{h_i}\frac{\partial v_j}{\partial w_i}\hat{\boldsymbol{h}}_i\times\hat{\boldsymbol{h}}_j &= \sum_i\sum_j \frac{1}{h_i}\frac{\partial v_j}{\partial w_i}\left[\sum_k \epsilon_{ijk}\hat{\boldsymbol{h}}_k\right] \\ &= \sum_i\sum_j\sum_k \epsilon_{ijk}\frac{1}{h_i}\frac{\partial v_j}{\partial w_i}\hat{\boldsymbol{h}}_k.\end{aligned}$$

Using Proposition 5.3, the second term is:

$$\begin{aligned}\sum_i\sum_j \frac{v_j}{h_i}\hat{\boldsymbol{h}}_i\times\frac{\partial\hat{\boldsymbol{h}}_j}{\partial w_i} &= \sum_i\sum_j\frac{v_j}{h_i}\left[\hat{\boldsymbol{h}}_i\times\left(\frac{1}{h_j}\frac{\partial h_i}{\partial w_j}\hat{\boldsymbol{h}}_i-\delta_{ij}\sum_k\frac{1}{h_k}\frac{\partial h_j}{\partial w_k}\hat{\boldsymbol{h}}_k\right)\right] \\ &= -\sum_i\sum_j\sum_k\delta_{ij}\frac{v_j}{h_kh_i}\frac{\partial h_j}{\partial w_k}\hat{\boldsymbol{h}}_i\times\hat{\boldsymbol{h}}_k \\ &= -\sum_j\sum_k\frac{v_j}{h_kh_j}\frac{\partial h_j}{\partial w_k}\hat{\boldsymbol{h}}_j\times\hat{\boldsymbol{h}}_k \\ &= -\sum_k\sum_j\frac{v_j}{h_kh_j}\frac{\partial h_j}{\partial w_k}\hat{\boldsymbol{h}}_j\times\hat{\boldsymbol{h}}_k.\end{aligned}$$

Exchange the dummy index k with an i. Then we get:

$$\begin{aligned}-\sum_i\sum_j\frac{v_j}{h_ih_j}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_j\times\hat{\boldsymbol{h}}_i &= \sum_i\sum_j\frac{v_j}{h_ih_j}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_i\times\hat{\boldsymbol{h}}_j \\ &= \sum_i\sum_j\frac{v_j}{h_ih_j}\frac{\partial h_j}{\partial w_i}\left(\sum_k\epsilon_{ijk}\hat{\boldsymbol{h}}_k\right) \\ &= \sum_i\sum_j\sum_k\epsilon_{ijk}\frac{v_j}{h_ih_j}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_k.\end{aligned}$$

We can now add the first and second terms:

$$\begin{aligned}\sum_i\sum_j\frac{1}{h_i}\frac{\partial v_j}{\partial w_i}\hat{\boldsymbol{h}}_i\times\hat{\boldsymbol{h}}_j+\sum_i\sum_j\frac{v_j}{h_i}\hat{\boldsymbol{h}}_i\times\frac{\partial\hat{\boldsymbol{h}}_j}{\partial w_i} &= \sum_i\sum_j\sum_k\epsilon_{ijk}\frac{1}{h_i}\frac{\partial v_j}{\partial w_i}\hat{\boldsymbol{h}}_k+\sum_i\sum_j\sum_k\epsilon_{ijk}\frac{v_j}{h_ih_j}\frac{\partial h_j}{\partial w_i}\hat{\boldsymbol{h}}_k \\ &= \sum_i\sum_j\sum_k\epsilon_{ijk}\frac{1}{h_ih_j}\left(h_j\frac{\partial v_j}{\partial w_i}+v_j\frac{\partial h_j}{\partial w_i}\right)\hat{\boldsymbol{h}}_k \\ &= \sum_i\sum_j\sum_k\epsilon_{ijk}\frac{1}{h_ih_j}\frac{\partial(v_jh_j)}{\partial w_i}\hat{\boldsymbol{h}}_k \\ &= \sum_i\sum_j\sum_k\epsilon_{ijk}\frac{1}{h_ih_jh_k}\frac{\partial(v_jh_j)}{\partial w_i}h_k\hat{\boldsymbol{h}}_k.\end{aligned}$$

Writing this out explicitly yields the following:

$$\boldsymbol{\nabla}\times\boldsymbol{V} = \frac{1}{h_1h_2h_3}\left[\left(\frac{\partial(v_3h_3)}{\partial w_2}-\frac{\partial(v_wh_2)}{\partial w_3}\right)h_1\hat{\boldsymbol{h}}_1+\left(\frac{\partial(v_1h_1)}{\partial w_3}-\frac{\partial(v_3h_3)}{\partial w_1}\right)h_2\hat{\boldsymbol{h}}_2+\left(\frac{\partial(v_2h_2)}{\partial w_1}-\frac{\partial(v_1h_1)}{\partial w_2}\right)h_3\hat{\boldsymbol{h}}_3\right]$$

As with the cross product in Cartesian coordinates, this can be written as the determinant of a matrix as follows:

$$\boldsymbol{\nabla} \times \boldsymbol{V} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \hat{\boldsymbol{h}}_1 & h_2 \hat{\boldsymbol{h}}_2 & h_3 \hat{\boldsymbol{h}}_3 \\ \dfrac{\partial}{\partial w_1} & \dfrac{\partial}{\partial w_2} & \dfrac{\partial}{\partial w_3} \\ v_1 h_1 & v_2 h_2 & v_3 h_3 \end{vmatrix}$$

Let us perform the calculation in spherical coordinates. Since $h_1 = 1, h_2 = r$ and $h_3 = r\sin\theta$, while $w_1 = r$, $w_2 = \theta$ and $w_3 = \phi$, we get:

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{V} &= \frac{1}{r^2 \sin\theta} \left[\left(\frac{\partial(r\sin\theta v_\phi)}{\partial\theta} - \frac{\partial(r v_\theta)}{\partial\phi} \right) \hat{\boldsymbol{e}}_r + \left(\frac{\partial v_r}{\partial\phi} - \frac{\partial(r\sin\theta v_\phi)}{\partial r} \right) r\hat{\boldsymbol{e}}_\theta + \left(\frac{\partial(r v_\theta)}{\partial r} - \frac{\partial v_r}{\partial\theta} \right) r\sin\theta \hat{\boldsymbol{e}}_\phi \right] \\ &= \frac{1}{r\sin\theta} \left[\frac{\partial}{\partial\theta}(\sin\theta v_\phi) - \frac{\partial v_\theta}{\partial\phi} \right] \hat{\boldsymbol{e}}_r + \frac{1}{r} \left[\frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{\partial}{\partial r}(r v_\phi) \right] \hat{\boldsymbol{e}}_\theta + \frac{1}{r} \left[\frac{\partial}{\partial r}(r v_\theta) - \frac{\partial v_r}{\partial\theta} \right] \hat{\boldsymbol{e}}_\phi. \end{aligned}$$

This is the same expression stated in the introduction.

6.6 Laplacian of a Scalar Function

As our last differential operation, we will compute the Laplacian of a scalar function in orthogonal curvilinear coordinates. The Laplacian is the divergence of the gradient. Thus if $f : \mathbb{R}^3 \to \mathbb{R}$ is a scalar function, the Laplacian of f in orthogonal curvilinear coordinates is given by:

$$\begin{aligned} \nabla^2 f &= \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f \\ &= \frac{1}{h_1 h_2 h_3} \left[\frac{\partial((\boldsymbol{\nabla} f)_1 h_2 h_3)}{\partial w_1} + \frac{\partial((\boldsymbol{\nabla} f)_2 h_3 h_1)}{\partial w_2} + \frac{\partial((\boldsymbol{\nabla} f)_3 h_1 h_2)}{\partial w_3} \right] \\ &= \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial w_1} \left(\frac{1}{h_1}\frac{\partial f}{\partial w_1} h_2 h_3 \right) + \frac{\partial}{\partial w_2} \left(\frac{1}{h_2}\frac{\partial f}{\partial w_2} h_3 h_1 \right) + \frac{\partial}{\partial w_3} \left(\frac{1}{h_3}\frac{\partial f}{\partial w_3} h_1 h_2 \right) \right]. \end{aligned}$$

Therefore the Laplacian of f is given by:

$$\nabla^2 f = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial w_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial f}{\partial w_1} \right) + \frac{\partial}{\partial w_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial f}{\partial w_2} \right) + \frac{\partial}{\partial w_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial f}{\partial w_3} \right) \right].$$

In spherical coordinates, we get

$$\begin{aligned} \nabla^2 f &= \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial r} \left(r^2 \sin\theta \frac{\partial f}{\partial r} \right) + \frac{\partial}{\partial\theta} \left(\frac{r\sin\theta}{r} \frac{\partial f}{\partial\theta} \right) + \frac{\partial}{\partial\phi} \left(\frac{r}{r\sin\theta} \frac{\partial f}{\partial\phi} \right) \right] \\ &= \frac{1}{r^2}\frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial f}{\partial\theta} \right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial f}{\partial\phi}. \end{aligned}$$

As before, we have recovered the expression in the introduction.

We can repeat this process now for any orthogonal curvilinear coordinate system, such as cylindrical coordinates, paraboloidal coordinates, elliptic coordinates, toroidal coordinates, or oblate spheroidal coordinates. All one needs are the Lamé coefficients. This resolves the mystery behind these expressions for differential operators in orthogonal curvilinear coordinates.

7 General Curvilinear Coordinates

What happens if we have curvilinear coordinates that are not orthogonal? A lot our results so far were manageable because we had $\boldsymbol{h}_i \cdot \boldsymbol{h}_j = \delta_{ij}$. Now we lose this property. An important consequence is that the tangent vectors to the coordinate curves and the coordinate surface normals are no longer proportional. In order to be able to distinguish between these situations, some vectors will be indexed with subscripts and others will be indexed with superscripts. This section is intended to be a brief tour of general curvilinear coordinates, so statements will be made without proof.

7.1 Curvilinear Coordinates

First we define curvilinear coordinates in much the same way we defined orthogonal curvilinear coordinates. We suppose that we have three level surfaces

$$q^i = f^i(x, y, z) = c^i$$

for each i and that

$$[\boldsymbol{\nabla} q^1, \boldsymbol{\nabla} q^2, \boldsymbol{\nabla} q^3] > 0.$$

As before, we say that these level surfaces are coordinate surfaces. Any two coordinate surfaces intersect in a coordinate curve. All three coordinate curves intersect in a single point P. Since P is uniquely specified by (c^1, c^2, c^3), then a point P is uniquely specified by prescribing values for (q^1, q^2, q^3). Thus (q^1, q^2, q^3) represents P in curvilinear coordinates. All the ancillary definitions that were made for orthogonal curvilinear coordinates apply to the general case.

We also assume that each function can be inverted. Hence we can assume that we can write

$$\begin{aligned} x &= k^1(q^1, q^2, q^3) \\ y &= k^2(q^1, q^2, q^3) \\ z &= k^3(q^1, q^2, q^3). \end{aligned}$$

In particular, then, we can write the vector field of position $\boldsymbol{R}$ in curvilinear coordinates.

7.2 Tangent Vectors

As in the orthogonal case, $\dfrac{\partial \boldsymbol{R}}{\partial q^i}$ is a vector tangent to any q^i coordinate curve. Therefore we define

$$\boldsymbol{g}_i = \frac{\partial \boldsymbol{R}}{\partial q^i}$$

as the coordinate tangent vectors. Note that if (q^1, q^2, q^3) are orthogonal curvilinear coordinates, $\boldsymbol{g}_i = \boldsymbol{h}_i$.

7.3 Coordinate Surface Normals

For each i we define

$$\boldsymbol{g}^i = \boldsymbol{\nabla} q^i.$$

These are the coordinate surface normals. In general, $\boldsymbol{g}^i$ is **not** proportional to $\boldsymbol{g}_i$. We saw that in the orthogonal case, these vectors coincide, up to scaling. This does not hold in general.

7.4 Bases

We can expand both sets of vectors in the Cartesian basis. We end up with the following:

$$\boldsymbol{g}_i = \sum_j \frac{\partial x_j}{\partial q^i} \hat{\boldsymbol{\delta}}_j$$

and

$$\boldsymbol{g}^i = \sum_j \frac{\partial q^i}{\partial x_j} \hat{\boldsymbol{\delta}}_j.$$

It turns out that each triplet of vectors is linearly independent and hence at each point P forms a basis for $\mathbb{R}^3$ footed at P. Since $\boldsymbol{g}_i$ and $\boldsymbol{g}_j$ (respectively $\boldsymbol{g}^i, \boldsymbol{g}^j$) are not orthogonal, there is no point in normalizing them.

Since $\{\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3\}$ form a basis for $\mathbb{R}^3$, we can expand any vector $\boldsymbol{v}$ in this basis. Hence we can write

$$\boldsymbol{v} = \sum_i v_i \boldsymbol{g}^i.$$

It turns out that $v_i = \boldsymbol{v} \cdot \boldsymbol{g}_i$. This is a consequence of what follows in the next section. Since the vectors $\boldsymbol{g}^i$ are not unit vectors, the numbers v_i are not physical components of the vector. They are called the **covariant** components of $\boldsymbol{v}$.

Similarly we can expand $\boldsymbol{v}$ in the basis $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3\}$ to write

$$\boldsymbol{v} = \sum_i v^i \boldsymbol{g}_i.$$

It also turns out that $v^i = \boldsymbol{v} \cdot \boldsymbol{g}^i$. Since the vectors $\boldsymbol{g}_i$ do not have unit length, the real numbers v^i are not physical components of $\boldsymbol{v}$. They are called the **contravariant** components of $\boldsymbol{v}$.

Tensors can also be written in each of these bases (or in mixed bases) and components may be contravariant, covariant, or mixed. We have thus arrived at the classical origins of covariant, contravariant and mixed tensors.

7.5 Relationship Between Tangent Vectors and Coordinate Surface Normals

The sets of vectors $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3\}$ and $\{\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3\}$ are reciprocal sets of vectors. This means that for all i and j,

$$\boldsymbol{g}_i \cdot \boldsymbol{g}^j = \delta_i^{\ j}.$$

Since these reciprocal sets of vectors are both bases for $\mathbb{R}^3$, we say that the sets are **dual bases**. We have now arrived at the classical origin of dualization.

If we define A to be the matrix whose rows are the vectors $\boldsymbol{g}^i$ and B to be the matrix whose columns are the vectors $\boldsymbol{g}_i$, it follows that A and B are inverses. Since the determinant of A is the triple scalar product $[\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3]$ (which is positive from our assumptions), we get the relation:

$$[\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3] = \frac{1}{[\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3]}.$$

7.6 Metric Coefficients

The **metric coefficients** $\{g_{ij}\}$ (for $i, j = 1, 2, 3$) are defined by

$$g_{ij} = \boldsymbol{g}_i \cdot \boldsymbol{g}_j.$$

The reason for the name will be explained shortly. Note that if the curvilinear coordinates are orthogonal, the metric coefficients are zero unless $i = j$ in which case the metric coefficient is the square of the i^{th} Lamé coefficient.

We can also define reciprocal metric coefficients $\{g^{ij}\}$. These are the quantities given by

$$g^{ij} = \boldsymbol{g}^i \cdot \boldsymbol{g}^j.$$

Observe that commutativity of the dot product implies that for all i and j, $g_{ij} = g_{ji}$ and $g^{ij} = g^{ji}$.

Using the metric coefficients and the reciprocal metric coefficients, we can determine the contravariant components of $\boldsymbol{g}^i$ and the covariant components of $\boldsymbol{g}_i$. Specifically:

$$\boldsymbol{g}^i = \sum_j g^{ij} \boldsymbol{g}_j$$

and

$$\boldsymbol{g}_i = \sum_j g_{ij} \boldsymbol{g}^j.$$

We can use these relations to determine how to transform between covariant and contravariant components of a vector. To wit, if

$$\boldsymbol{v} = \sum_i v_i \boldsymbol{g}^i = \sum_i v^i \boldsymbol{g}_i$$

then:

$$v_i = \sum_j g_{ij} v^j$$

and

$$v^i = \sum_j g^{ij} v_j.$$

Similar rules define how contravariant and covariant components of tensors can be transformed. This is the germ of the notion that a tensor can be defined as a collection of quantities that transform according to certain rules.

Notice that there is a nice pattern to the transformations between covariant and contravariant components. Multiplying a quantity with a superscript of j by a g_{ij} and summing over j lowers the superscript j and changes it to i. Similarly multiplying a quantity with a subscript of j by g^{ij} and summing over j raises the j to an i. Because of this pattern, the metric coefficients and reciprocal metric coefficients are sometimes referred to as lowering and raising operators, respectively.

7.7 Integration in General Curvilinear Coordinates

Arrange the metric coefficients in a matrix G in the obvious way. Denote

$$g = \det G.$$

We can similarly arrange the reciprocal metric coefficients in a matrix, and the result is the inverse G^{-1}. If the coordinates are orthogonal, then in terms of the Lamé coefficients we have

$$g = h_1^2 h_2^2 h_3^2.$$

Observe that

$$[\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3] = \sqrt{g}.$$

It turns out that this holds true in general curvilinear coordinates as well, except we have to be careful about which triple scalar product is $\sqrt{g}$. The following is the correct result:

$$[\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3] = \sqrt{g}.$$

Therefore, to integrate in general curvilinear coordinates, the integrand must be multiplied by $\sqrt{g}$.

Since the coordinate surface normals are reciprocal to the coordinate tangent vectors, we get

$$[\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3] = \frac{1}{\sqrt{g}}.$$

One condition that is equivalent to a collection $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ and $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ being reciprocal sets of vectors is that

$$\boldsymbol{v}_i = \frac{\boldsymbol{w}_j \times \boldsymbol{w}_k}{[\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3]}$$

whenever i, j, k are a cyclic permutation of $1, 2$ and 3. Accordingly, when i, j, k are cyclic permutations of $1, 2, 3$ we get:

$$\boldsymbol{g}^i = \frac{\boldsymbol{g}_j \times \boldsymbol{g}_k}{\sqrt{g}}$$

and

$$\boldsymbol{g}_i = \sqrt{g}(\boldsymbol{g}^j \times \boldsymbol{g}^k).$$

This gives us cross-product relationships:

$$\boldsymbol{g}_i \times \boldsymbol{g}_j = \sum_k \epsilon_{ijk} \sqrt{g} \boldsymbol{g}^k$$

and

$$\boldsymbol{g}^i \times \boldsymbol{g}^j = \sum_k \frac{\epsilon^{ijk}}{\sqrt{g}} \boldsymbol{g}_k.$$

Sometimes the following notation is used:

$$\mathcal{E}_{ijk} = \epsilon_{ijk}\sqrt{g} \quad \text{and} \quad \mathcal{E}^{ijk} = \frac{\epsilon^{ijk}}{\sqrt{g}}.$$

This allows us to write

$$\boldsymbol{g}_i \times \boldsymbol{g}_j = \sum_k \mathcal{E}_{ijk} \boldsymbol{g}_k$$

and

$$\boldsymbol{g}^i \times \boldsymbol{g}^j = \sum_k \mathcal{E}^{ijk} \boldsymbol{g}_k.$$

This also makes the Levi-Cevita symbol into a tensor.

7.8 The Metric Tensor

Using the metric coefficients, we can define a special dyadic tensor. In order to make sense of this concept, I will give a brief description of what a dyadic tensor is.

7.8.1 Dyadic Tensors

We discuss dyadic tensors in Cartesian coordinates. Suppose you have two vectors in $\mathbb{R}^3$ given by:

$$\boldsymbol{v} = 3\hat{\boldsymbol{\imath}} - 4\hat{\boldsymbol{k}}, \quad \boldsymbol{w} = -\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}}.$$

Suppose, moreover, you are teaching vector calculus and a student asks why you can't multiply $\boldsymbol{v}$ and $\boldsymbol{w}$ by writing

$$\boldsymbol{v}\boldsymbol{w} = (3\hat{\boldsymbol{\imath}} - 4\hat{\boldsymbol{k}})(-\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}})$$

and using FOIL. The answer is that you can, but what you get back is no longer a vector. Rather it is a quantity that looks like

$$\boldsymbol{v}\boldsymbol{w} = -3\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\imath}} + 6\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} + 4\hat{\boldsymbol{k}}\hat{\boldsymbol{\imath}} - 8\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}}.$$

This quantity is a dyadic tensor, and the symbols such as $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\imath}}$ represent new basis vectors for a nine-dimensional vector space. They are called the unit dyads, and the product $\boldsymbol{v}\boldsymbol{w}$ is called the dyadic product. In modern terminology, if we view $\mathbb{R}^3$ as a three dimensional vector space V over $\mathbb{R}$, the nine-dimensional vector space $\boldsymbol{v}\boldsymbol{w}$ lives in is the tensor product $V \otimes V$. A quantity such as $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\imath}}$ is now denoted $\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$, and the dyadic product is denoted $\boldsymbol{v} \otimes \boldsymbol{w}$. The fact that not every element of a tensor product can be written as a simple tensor $\boldsymbol{v} \otimes \boldsymbol{w}$ corresponds to the notion that not every dyadic tensor can be decomposed into the dyadic product of two vectors.

Note that sometimes the dyadic product is called an outer product. The reason is as follows. The nine double-indexed components of a dyadic tensor can be arranged into a three-by-three matrix in an obvious way. If we represent a vector by a column matrix, the inner product of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ can be described as the matrix product

$$\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^T \boldsymbol{w}.$$

The other product, given by

$$\boldsymbol{v}\boldsymbol{w} = \boldsymbol{v}\boldsymbol{w}^T$$

where the right hand side is matrix multiplication, gives a three by three matrix whose components are the same as their dyadic product. This gives a concrete way of viewing dyadic tensors.

7.8.2 The Metric Dyadic Tensor

The dyadic product can be rewritten in orthogonal or general curvilinear coordinates. Accordingly, we can consider the dyadic tensor

$$\boldsymbol{g}^i \boldsymbol{g}^j.$$

Since we can arrange the metric coefficients in a matrix the same way we can arrange the components of a dyadic tensor, it is not surprising that we can write down a dyadic tensor

$$\overline{\overline{\boldsymbol{G}}} = \sum_i \sum_j g_{ij} \boldsymbol{g}^i \boldsymbol{g}^j.$$

This is called the metric dyadic tensor. In modern language, this is a Riemannian metric. The action on pairs of vectors would be obtained using an operation called the double-dot product, which I will not discuss. I will only mention that this action corresponds to the ordinary inner product when $g_{ij} = \delta_{ij}$ and the coordinates correspond to the Cartesian coordinates. In fact, if we define the change of coordinates map Φ from curvilinear coordinates to Cartesian coordinates, the metric $\overline{\overline{\boldsymbol{G}}}$ acts as the pull-back of the usual inner product. Hence, in modern language, $\mathbb{R}^3$ with curvilinear coordinates and the metric $\overline{\overline{\boldsymbol{G}}}$ is isometric to $\mathbb{E}^3$.

One should pause and note at this point that we have arrived at the foundation of Riemannian geometry, using only basic vector calculus and linear algebra!

7.9 Christoffel Symbols

Just as in the orthogonal curvilinear case, $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3\}$ and $\{\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3\}$ vary from point to point. Consequently, we need to understand how these change from point to point in order to do differential operations. As a matter of notation, when we differentiate a vector field $\boldsymbol{V}$ with respect to q^j, we will write a subscript $/j$. For example, $\boldsymbol{g}_i = \boldsymbol{R}_{/i}$.

As horrific as the derivations in the orthogonal curvilinear case, orthogonality helped a great deal. Hence some symbols are used to pack up the ugliness into something more compact. These symbols are the Christoffel symbols. We define Chirstoffel symbols of the first kind by

$$[ij, k] = \boldsymbol{g}_{i/j} \cdot \boldsymbol{g}_k.$$

Then we can write

$$\boldsymbol{g}_{i/j} = \sum_k [ij, k] \boldsymbol{g}^k.$$

This gives the covariant components of $\boldsymbol{g}_{i/j}$. Similarly, we define Christoffel symbols of the second kind by

$$\begin{Bmatrix} k \\ ij \end{Bmatrix} = \boldsymbol{g}_{i/j} \cdot \boldsymbol{g}^k.$$

Then we can write:

$$\boldsymbol{g}_{i/j} = \sum_k \begin{Bmatrix} k \\ ij \end{Bmatrix} \boldsymbol{g}_k.$$

This gives the contravariant components of $\boldsymbol{g}_{i/j}$. Note that by equality of mixed partial derivatives, since we have sufficient smoothness,

$$\boldsymbol{g}_{i/j} = \boldsymbol{R}_{/i/j} = \boldsymbol{R}_{/j/i} = \boldsymbol{g}_{j/i}.$$

Hence

$$[ij, k] = [ji, k] \text{ and } \begin{Bmatrix} k \\ ij \end{Bmatrix} = \begin{Bmatrix} k \\ ji \end{Bmatrix}$$

showing that the Christoffel symbols have a certain type of symmetry.

All of this begs the question, how on earth do we compute the Christoffel symbols? It turns out that one can prove that the Christoffel symbols can be computed using the metric coefficients. The formula for the Christoffel symbols of the first kind is

$$[ij,k] = \frac{1}{2}(g_{ki/j} + g_{jk/i} - g_{ij/k}).$$

The formula for the Christoffel symbols of the second kind, which will look familiar to some readers, is given by

$$\left\{ \begin{matrix} k \\ ij \end{matrix} \right\} = \frac{1}{2}\sum_l g^{kl}(g_{il/j} + g_{jl/i} - g_{ij/l}).$$

This shows that the Christoffel symbols of the second kind are precisely the same as the modern notion of the Christoffel symbols Γ^k_{ij} used in defining the Levi-Civita connection with respect to the Riemannian metric $\overline{\overline{\boldsymbol{G}}}$ defined above.

Last, we can use the Christoffel symbols of the second kind to describe the partial derivatives of the coordinate surface normals. Namely:

$$\boldsymbol{g}^i{}_{/j} = -\sum_k \left\{ \begin{matrix} i \\ kj \end{matrix} \right\} \boldsymbol{g}^k.$$

Now that we know how to differentiate our basis vectors without reference to Cartesian coordinates, we are prepared to make the last step in our journey and do differential calculus in general curvilinear coordinates.

7.10 Differential Calculus of Vectors in General Curvilinear Coordinates

As with the orthogonal case, we begin with the Del operator. This turns out to be quite simple, although one should note that the coordinate surface normals are used, and not the coordinate tangent vectors. With that remark, the Del operator in curvilinear coordinates is given by

$$\boldsymbol{\nabla} = \sum_j \boldsymbol{g}^i \frac{\partial}{\partial q^i}.$$

We then get the gradient of a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$ by

$$\boldsymbol{\nabla} f = \sum_i f_{/i}\boldsymbol{g}^i.$$

The divergence of a vector field $\boldsymbol{V}$ is given by

$$\boldsymbol{\nabla} \cdot \boldsymbol{V} = \sum_i \left(v^i{}_{/i} + \sum_j v^j \left\{ \begin{matrix} i \\ ji \end{matrix} \right\} \right).$$

This can be simplified, using various vector identities, to

$$\boldsymbol{\nabla} \cdot \boldsymbol{V} = \frac{1}{\sqrt{g}} \sum_i (v^i \sqrt{g})_{/i}.$$

As an amusing corollary, we get

$$\sum_j \begin{Bmatrix} j \\ ij \end{Bmatrix} = (\ln \sqrt{g})_{/i}.$$

The curl of a vector field is given by

$$\boldsymbol{\nabla} \times \boldsymbol{V} = \sum_i \sum_j \sum_k \frac{\epsilon^{ijk}}{\sqrt{g}} \left(v_{j/i} - \sum_l \begin{Bmatrix} l \\ ji \end{Bmatrix} v_l \right) \boldsymbol{g}_k.$$

In this form, the curl of a vector field is useful in covariant differentiation. A simpler form of the curl of a vector field is given by

$$\boldsymbol{\nabla} \times \boldsymbol{V} = \frac{1}{\sqrt{g}} \sum_i \sum_j \sum_k \epsilon^{ijk} v_{i/j} \boldsymbol{g}_k.$$

This can be arranged as the determinant of a matrix:

$$\boldsymbol{\nabla} \times \boldsymbol{V} = \frac{1}{\sqrt{g}} \begin{vmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \boldsymbol{g}_3 \\ \dfrac{\partial}{\partial q^1} & \dfrac{\partial}{\partial q^2} & \dfrac{\partial}{\partial q^3} \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

Finally, the Laplacian of a scalar function is given by

$$\nabla^2 f = \frac{1}{\sqrt{g}} \sum_i \sum_j (\sqrt{g} g^{ij} f_{/j})_{/i}.$$

8 Conclusion

After a long road, we have learned how to handle differential vector calculus in general curvilinear coordinates in $\mathbb{R}^3$. When the coordinates are orthogonal, the expressions for all of our differential operations are in terms of the Lamé coefficients. When the coordinates are not orthogonal, the expressions for the differential operations involve only the metric coefficients (and possibly the Christoffel symbols, which in turn depend on the metric coefficients). Along the way, we have witnessed the development of basic elements of Riemannian geometry, including Riemannian metrics and Christoffel symbols. We have also witnessed the classical origins of concepts such as tensor products, dual bases, and covariant/contravariant tensors. One can proceed to define covariant differentiation, the curvature tensor and other such players from differential geometry using the tools we have constructed. The interested reader is referred to [Spa03] or [Kre91]. One can proceed in a different direction and define differential calculus with tensors, starting with dyadic tensors. This gets complicated very quickly, since the number of components and indices as well as the types of operations one can perform get out of control. At the level of vectors (one-tensors), however, the classical background adds depth to the formalities one encounters using the modern notions.

9 Exercises

Determine expressions for the Del operator, the gradient of a scalar function, the divergence of a vector field, the curl of a vector field, and the Laplacian of a scalar function in each of the following coordinate systems.

1. Cylindrical coordinates defined by:

$$\begin{aligned} x &= \rho \cos \phi \\ y &= \rho \sin \phi \\ z &= z \end{aligned}$$

where $\rho > 0$ and $0 < \phi < 2\pi$.

2. Parabolic coordinates defined by:

$$\begin{aligned} x &= \frac{1}{2}u^2 - \frac{1}{2}v^2 \\ y &= uv \\ z &= z \end{aligned}$$

where $v \geq 0$.

3. Paraboloidal coordinates defined by:

$$\begin{aligned} x &= uv \cos \phi \\ y &= uv \sin \phi \\ z &= \frac{1}{2}u^2 - \frac{1}{2}v^2 \end{aligned}$$

where $u, v \geq 0$ and $0 < \phi < 2\pi$.

4. Elliptic coordinates defined by:

$$\begin{aligned} x &= \cosh u \cos v \\ y &= \sinh u \sin v \\ z &= z \end{aligned}$$

where $u \geq 0$ and $0 < v < 2\pi$.

5. Prolate spheroidal coordinates defined by:

$$\begin{aligned} x &= \sinh \xi \sin \eta \cos \phi \\ y &= \sinh \xi \sin \eta \sin \phi \\ z &= \cosh \xi \cos \eta \end{aligned}$$

where $\xi \geq 0$, $0 < \eta < \pi$ and $0 < \phi < 2\pi$.

6. Oblate spheroidal coordinates defined by:

$$\begin{aligned} x &= \cosh \xi \cos \eta \cos \phi \\ y &= \cosh \xi \cos \eta \sin \phi \\ z &= \sinh \xi \sin \eta \end{aligned}$$

where $\xi \geq 0$, $-\pi/2 < \eta < \pi/2$ and $0 < \phi < 2\pi$.

7. Bipolar cylindrical coordinates defined by:

$$\begin{aligned} x &= \frac{\sinh v}{\cosh v - \cos u} \\ y &= \frac{\sin u}{\cosh v - \cos u} \\ z &= z \end{aligned}$$

where $0 < u < 2\pi$.

8. Bispherical coordinates defined by:

$$\begin{aligned} x &= \frac{\sin \xi \cos \phi}{\cosh \eta - \cos \xi} \\ y &= \frac{\sin \xi \sin \phi}{\cosh \eta - \cos \xi} \\ z &= \frac{\sinh \eta}{\cosh \eta - \cos \xi}. \end{aligned}$$

9. Toroidal coordinates defined by:

$$\begin{aligned} x &= \frac{\sinh v \cos \phi}{\cosh v - \cos u} \\ y &= \frac{\sinh v \sin \phi}{\cosh v - \cos u} \\ z &= \frac{\sin u}{\cosh v - \cos u} \end{aligned}$$

where $v \geq 0$ and $0 < \phi, u < 2\pi$.

10. Conical coordinates defined with parameters a and b by:

$$\begin{aligned} x &= \frac{\lambda \mu \nu}{ab} \\ y &= \frac{\lambda}{a}\sqrt{\frac{(\mu^2 - a^2)(\nu^2 - a^2)}{a^2 - b^2}} \\ z &= \frac{\lambda}{b}\sqrt{\frac{(\mu^2 - b^2)(\nu^2 - b^2)}{b^2 - a^2}}. \end{aligned}$$

References

[Kre91] Erwin Kreyszig. *Differential Geometry.* Dover, 1991.

[PS02] Gerald Pollack and Daniel Stump. *Electromagnetism.* Addison Wesley, 2002.

[Spa03] Brian Spain. *Tensor Calculus: A Concise Course.* Dover, 2003.