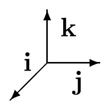
1 Curvilinear Coordinates

1.1 Basis vectors

The usual basis vectors are denoted by $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and are as the following picture describes.

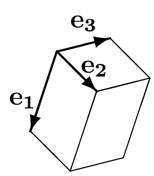


The vectors, $\mathbf{i}, \mathbf{j}, \mathbf{k}$, are fixed. If $\mathbf{v}$ is a vector, there are unique scalars called components such that $\mathbf{v} = v^1\mathbf{i} + v^2\mathbf{j} + v^3\mathbf{k}$. This is what we mean when we say $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is a basis.

Now suppose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are three vectors which satisfy

$$\mathbf{e}_1 \times \mathbf{e}_2 \cdot \mathbf{e}_3 \neq 0.$$

Recall this means the volume of the box spanned by the three vectors is not zero.



Suppose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are as just described. Does it follow that they form a basis? We show this is the case. Thus there are unique scalars, v^1, v^2, and v^3 such that

$$\mathbf{v} = v^i \mathbf{e}_i.$$

This is the content of the following theorem.

Theorem 1 *If $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are three vectors, then they form a basis if and only if*

$$\mathbf{e}_1 \times \mathbf{e}_2 \cdot \mathbf{e}_3 \neq 0.$$

Proof: Suppose first the above condition holds. Let $\mathbf{i}_1 \equiv \mathbf{i}, \mathbf{i}_2 \equiv \mathbf{j}, \mathbf{i}_3 \equiv \mathbf{k}$ and suppose $\mathbf{v} = u^j \mathbf{i}_j$. Therefore, u^1, u^2, u^3 are the components of the vector $\mathbf{v}$ with respect to the usual basis vectors. Also let

$$\mathbf{e}_i = a_i^j \mathbf{i}_j$$

thus writing each $\mathbf{e}_i$ in terms of the vectors $\mathbf{i}_j$. Then from the definition of the box product in terms of the usual basis vectors, we see

$$0 \neq \mathbf{e}_1 \times \mathbf{e}_2 \cdot \mathbf{e}_3 = \det\left(a_i^j\right)$$

$$\equiv \det \begin{pmatrix} a_1^1 & a_1^2 & a_1^3 \\ a_2^1 & a_2^2 & a_2^3 \\ a_3^1 & a_3^2 & a_3^3 \end{pmatrix} = \det \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{pmatrix} \quad (1)$$

and we want to show there exists a unique solution to

$$\mathbf{v} = u^j \mathbf{i}_j = v^i \mathbf{e}_i = a_i^j v^i \mathbf{i}_j. \quad (2)$$

In other words, we want a unique solution to the system of equations

$$u^j = a_i^j v^i, \ j = 1, 2, 3.$$

This is in the form

$$\mathbf{u} = A\mathbf{v} \tag{3}$$

where A is matrix which, by 1, has non zero determinant.

If the box product is equal to zero, then the system which needs to be solved is of the form 3 where $\det(A) = 0$. Therefore, A has an eigenvector, $\mathbf{v} = (v_1, v_2, v_3)^T$ which corresponds to the eigen value $\lambda = 0$. Therefore, 2 shows $\mathbf{0}$ has more than one set of components with respect to the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and this set of vectors is therefore not a basis.

This gives a simple geometric condition which determines whether a list of three vectors forms a basis in $\mathbb{R}^3$. One simply takes the box product. If the box product is not equal to zero, then the vectors form a basis. If not, the list of three vectors does not form a basis. This condition generalizes to $\mathbb{R}^n$ as follows. If $\mathbf{e}_i = a_i^j \mathbf{i}_j$, then $\{\mathbf{e}_i\}_{i=1}^n$ forms a basis if and only if $\det\left(a_i^j\right) \neq 0$.

These vectors may or may not be orthonormal. In any case, it is convenient to define something called the dual basis.

Definition 2 *Let $\{\mathbf{e}_i\}_{i=1}^n$ form a basis for $\mathbb{R}^n$. Then $\left\{\mathbf{e}^i\right\}_{i=1}^n$ is called the dual basis if*

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} . \tag{4}$$

Theorem 3 *If $\{\mathbf{e}_i\}_{i=1}^n$ is a basis then $\left\{\mathbf{e}^i\right\}_{i=1}^n$ is also a basis provided 4 holds.*

Proof: Suppose

$$\mathbf{v} = v_i \mathbf{e}^i. \tag{5}$$

Then taking the dot product of both sides of 5 with $\mathbf{e}_j$,yields

$$v_j = \mathbf{v} \cdot \mathbf{e}_j. \tag{6}$$

Thus there is at most one choice of scalars v_j such that $\mathbf{v} = v_j \mathbf{e}^j$ and it is given by 6.

$$\left(\mathbf{v} - \mathbf{v} \cdot \mathbf{e}_j \mathbf{e}^j\right) \cdot \mathbf{e}_k = 0$$

and so, since $\{\mathbf{e}_i\}_{i=1}^n$ is a basis,

$$\left(\mathbf{v} - \mathbf{v} \cdot \mathbf{e}_j \mathbf{e}^j\right) \cdot \mathbf{w} = 0$$

for all vectors, $\mathbf{w}$. It follows $\mathbf{v} - \mathbf{v} \cdot \mathbf{e}_j \mathbf{e}^j = \mathbf{0}$ and this shows $\left\{\mathbf{e}^i\right\}_{i=1}^n$ is a basis.

In the above argument we obtained formulas for the components of a vector $\mathbf{v}$, v_i, with respect to the dual basis, found to be $v_j = \mathbf{v} \cdot \mathbf{e}_j$. In the same way, we may find the components of a vector with respect to the basis $\{\mathbf{e}_i\}_{i=1}^n$. Let $\mathbf{v}$ be any vector and let

$$\mathbf{v} = v^j \mathbf{e}_j. \tag{7}$$

Then using 4 and taking the dot product of both sides of 7 with $\mathbf{e}^i$ we see $v^i = \mathbf{e}^i \cdot \mathbf{v}$.

Does there exist a dual basis and is it uniquely determined?

Theorem 4 *If $\{\mathbf{e}_i\}_{i=1}^n$ is a basis for $\mathbb{R}^n$, then there exists a unique dual basis, $\left\{\mathbf{e}^j\right\}_{j=1}^n$ satisfying*

$$\mathbf{e}^j \cdot \mathbf{e}_i = \delta_i^j.$$

Proof: First we show the dual basis is unique. Suppose $\left\{\mathbf{f}^j\right\}_{j=1}^n$ is another set of vectors which satisfies $\mathbf{f}^j \cdot \mathbf{e}_i = \delta_i^j$. Then

$$\mathbf{f}^j = \mathbf{f}^j \cdot \mathbf{e}_i \mathbf{e}^i = \delta_i^j \mathbf{e}^i = \mathbf{e}^j.$$

Note that from the definition, the dual basis to $\{\mathbf{i}_j\}_{j=1}^n$ is just $\mathbf{i}^j = \mathbf{i}_j$. Letting

$$\mathbf{e}_i = a_i^j \mathbf{i}_j$$

where the vectors, $\{\mathbf{i}_j\}_{j=1}^n$ are the standard basis vectors, it follows that since the $\mathbf{e}_i$ form a basis that the matrix whose ij th entry is a_i^j is an invertible matrix. Letting

$$\mathbf{e}^k = b_r^k \mathbf{i}^r,$$

we need to choose b_r^k such that

$$b_r^k \mathbf{i}^r \cdot a_i^j \mathbf{i}_j = b_r^k a_i^j \delta_j^r = b_j^k a_i^j = \delta_i^k.$$

But this is nothing more than the matrix equation for B which is of the form

$$AB = I$$

where $A = \left(a_i^j\right)$ and has an inverse. There exists a unique solution to this equation given by $B = A^{-1}$ and this proves the existence of the dual basis.

Summarizing what has been shown so far, we know that $\{\mathbf{e}_i\}_{i=1}^n$ is a basis for $\mathbb{R}^n$ if and only if when $\mathbf{e}_i = a_i^j \mathbf{i}_j$,

$$\det\left(a_i^j\right) \neq 0. \tag{8}$$

If $\{\mathbf{e}_i\}_{i=1}^n$ is a basis, then there exists a unique dual basis, $\left\{\mathbf{e}^j\right\}_{j=1}^n$ satisfying

$$\mathbf{e}^j \cdot \mathbf{e}_i = \delta_i^j, \tag{9}$$

and that if $\mathbf{v}$ is any vector,

$$\mathbf{v} = v_j \mathbf{e}^j,\ \mathbf{v} = v^j \mathbf{e}_j. \tag{10}$$

The components of $\mathbf{v}$ which have the index on the top are called the contravariant components of the vector while the components which have the index on the bottom are called the covariant components. In general $v_i \neq v^j$! We also have formulae for these components in terms of the dot product.

$$v_j = \mathbf{v} \cdot \mathbf{e}_j,\ v^j = \mathbf{v} \cdot \mathbf{e}^j. \tag{11}$$

We define $g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j$ and $g^{ij} \equiv \mathbf{e}^i \cdot \mathbf{e}^j$. The next theorem describes the process of raising or lowering an index.

Theorem 5 *The following hold.*

$$g^{ij} \mathbf{e}_j = \mathbf{e}^i,\ g_{ij} \mathbf{e}^j = \mathbf{e}_i, \tag{12}$$

$$g^{ij} v_j = v^i,\ g_{ij} v^j = v_i, \tag{13}$$

$$g^{ij} g_{jk} = \delta_k^i, \tag{14}$$

$$\det\left(g_{ij}\right) > 0,\ \det\left(g^{ij}\right) > 0. \tag{15}$$

Proof: First,

$$\mathbf{e}^i = \mathbf{e}^i \cdot \mathbf{e}^j \mathbf{e}_j = g^{ij}\mathbf{e}_j$$

by 10 and 11. Similarly, by 10 and 11,

$$\mathbf{e}_i = \mathbf{e}_i \cdot \mathbf{e}_j \mathbf{e}^j = g_{ij}\mathbf{e}^j.$$

This verifies 12. To verify 13,

$$v^i = \mathbf{e}^i \cdot \mathbf{v} = g^{ij}\mathbf{e}_j \cdot \mathbf{v} = g^{ij} v_j.$$

The proof of the remaining formula in 13 is similar.

To verify 14,

$$g^{ij} g_{jk} = \mathbf{e}^i \cdot \mathbf{e}^j \mathbf{e}_j \cdot \mathbf{e}_k = \left(\left(\mathbf{e}^i \cdot \mathbf{e}^j\right) \mathbf{e}_j\right) \cdot \mathbf{e}_k = \mathbf{e}^i \cdot \mathbf{e}_k = \delta^i_k.$$

This shows the two determinants in 15 are non zero because the two matrices are inverses of each other. It only remains to verify that one of these is greater than zero. Letting $\mathbf{e}_i = a^j_i \mathbf{i}_j = b^i_j \mathbf{i}^j$, we see that since $\mathbf{i}_j = \mathbf{i}^j$, $a^j_i = b^i_j$. Therefore,

$$\mathbf{e}_i \cdot \mathbf{e}_j = a^r_i \mathbf{i}_r \cdot b^j_k \mathbf{i}^k = a^r_i b^j_k \delta^k_r = a^k_i b^j_k = a^k_i a^k_j.$$

It follows that for G the matrix whose ij th entry is $\mathbf{e}_i \cdot \mathbf{e}_j$, $G = AA^T$ where the ik th entry of A is a^k_i. Therefore, $\det(G) = \det(A)\det\left(A^T\right) = \det(A)^2 > 0$. It follows from 14 that if H is the matrix whose ij th entry is g^{ij}, then $GH = I$ and so $H = G^{-1}$ and

$$\det(G)\det\left(G^{-1}\right) = \det\left(g^{ij}\right)\det(G) = 1.$$

Therefore, $\det\left(G^{-1}\right) > 0$ also. This proves the theorem.

Definition 6 *The matrix* $(g_{ij}) = G$ *is called the metric tensor.*

1.2 Exercises

1. Let $\mathbf{e}_1 = \mathbf{i} + \mathbf{j}, \mathbf{e}_2 = \mathbf{i} - \mathbf{j}, \mathbf{e}_3 = \mathbf{j} + \mathbf{k}$. Find $\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3$, $(g_{ij}), \left(g^{ij}\right)$. If $\mathbf{v} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$, find v^i and v_j, the contravariant and covariant components of the vector.

2. Let $\mathbf{e}^1 = 2\mathbf{i} + \mathbf{j}, \mathbf{e}^2 = \mathbf{i} - 2\mathbf{j}, \mathbf{e}^3 = \mathbf{k}$. Find $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, $(g_{ij}), \left(g^{ij}\right)$. If $\mathbf{v} = 2\mathbf{i} - 2\mathbf{j} + \mathbf{k}$, find v^i and v_j, the contravariant and covariant components of the vector.

3. Suppose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ have the property that $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ whenever $i \neq j$. Show the same is true of the dual basis and that in fact, $\mathbf{e}^i$ is a multiple of $\mathbf{e}_i$.

4. Let $\mathbf{e}_1, \cdots, \mathbf{e}_3$ be a basis for $\mathbb{R}^n$ and let $\mathbf{v} = v^i \mathbf{e}_i = v_i \mathbf{e}^i, \mathbf{w} = w^j \mathbf{e}_j = w_j \mathbf{e}^j$ be two vectors. Show

$$\mathbf{v} \cdot \mathbf{w} = g_{ij} v^i w^j = g^{ij} v_i w_j.$$

5. Show if $\{\mathbf{e}_i\}_{i=1}^3$ is a basis in $\mathbb{R}^3$

$$\mathbf{e}^1 = \frac{\mathbf{e}_2 \times \mathbf{e}_3}{\mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1},\ \mathbf{e}^2 = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\mathbf{e}_1 \times \mathbf{e}_3 \cdot \mathbf{e}_2},\ \mathbf{e}^3 = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\mathbf{e}_1 \times \mathbf{e}_2 \cdot \mathbf{e}_3}.$$

6. Let $\{\mathbf{e}_i\}_{i=1}^n$ be a basis and define

$$\mathbf{e}_i^* \equiv \frac{\mathbf{e}_i}{|\mathbf{e}_i|},\ \mathbf{e}^{*i} \equiv \mathbf{e}^i\,|\mathbf{e}_i|\,.$$

Show $\mathbf{e}^{*i} \cdot \mathbf{e}_j^* = \delta_j^i$.

7. If $\mathbf{v}$ is a vector, v_i^* and v^{*i}, are defined by

$$\mathbf{v} \equiv v_i^* \mathbf{e}^{*i} \equiv v^{*i} \mathbf{e}_i^*.$$

These are called the physical components of $\mathbf{v}$. Show

$$v_i^* = \frac{v_i}{|\mathbf{e}_i|},\ v^{*i} = v^i\,|\mathbf{e}_i|\ \text{(No summation on } i \text{).}$$

1.3 Curvilinear Coordinates

With the algebraic preparation of the last section, we are ready to consider curvilinear coordinates. Let $D \subseteq \mathbb{R}^n$ be an open set and let $\mathbf{M} : D \to \mathbb{R}^n$ satisfy

$$\mathbf{M} \text{ is } C^2, \tag{16}$$

$$\mathbf{M} \text{ is one to one.} \tag{17}$$

Letting $\mathbf{x} \in D$, we can write

$$\mathbf{M}(\mathbf{x}) = M^k(\mathbf{x})\,\mathbf{i}_k$$

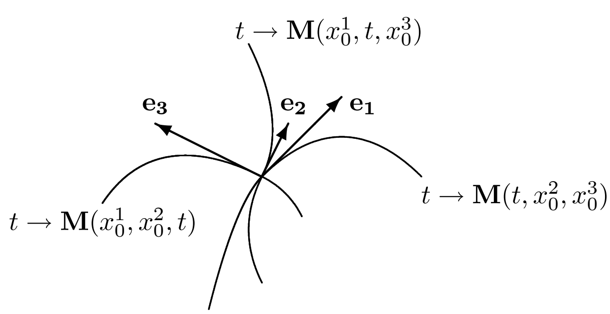
where, as usual, $\mathbf{i}_k$ are the standard basis vectors for $\mathbb{R}^n$, $\mathbf{i}_k$ being the vector in $\mathbb{R}^n$ which has a one in the kth coordinate and a 0 in every other spot. For a fixed $\mathbf{x} \in D$, we can consider the curves,

$$t \to \mathbf{M}(\mathbf{x}+t\mathbf{i}_k)$$

for $t \in I$, some open interval containing 0. Then for the point $\mathbf{x}$,we let

$$\mathbf{e}_k \equiv \frac{\partial \mathbf{M}}{\partial x^k}(\mathbf{x}) \equiv \frac{d}{dt}\left(\mathbf{M}(\mathbf{x}+t\mathbf{i}_k)\right)|_{t=0}.$$

We will denote this vector as $\mathbf{e}_k(\mathbf{x})$ to emphasize its dependence on $\mathbf{x}$. The following picture illustrates the situation in $\mathbb{R}^3$.



We want $\{\mathbf{e}_k\}_{k=1}^n$ to be a basis. Thus we need

$$\det\left(\frac{\partial M^i}{\partial x^k}\right) \neq 0. \tag{18}$$

Let

$$y^i = M^i(\mathbf{x}) \ \ i = 1, \cdots, n \tag{19}$$

so that the y^i are the usual coordinates with respect to the usual basis vectors $\{\mathbf{i}_k\}_{k=1}^n$ of the point $\mathbf{M}(\mathbf{x})$. Letting $\mathbf{x} \equiv \left(x^1, \cdots, x^n\right)$, it follows from the inverse function theorem of advanced calculus that $\mathbf{M}(D)$ is open, and that 18, 16, and 17 imply the equations 19 define each x^i as a C^2 function of $\mathbf{y} \equiv \left(y^1, \cdots, y^n\right)^T$. Thus, abusing notation slightly, the equations 19 are equivalent to

$$x^i = x^i(\mathbf{y}), \ i = 1, \cdots, n$$

where x^i is a C^2 function. Thus

$$\nabla x^k(\mathbf{y}) = \frac{\partial x^k(\mathbf{y})}{\partial y^j}\mathbf{i}^j.$$

Then

$$\nabla x^k(\mathbf{y}) \cdot \mathbf{e}_j = \frac{\partial x^k}{\partial y^s}\mathbf{i}^s \cdot \frac{\partial y^r}{\partial x^j}\mathbf{i}_r = \frac{\partial x^k}{\partial y^s}\frac{\partial y^s}{\partial x^j} = \delta_j^k$$

by the chain rule. Therefore, the dual basis is given by

$$\mathbf{e}^k(\mathbf{x}) = \nabla x^k(\mathbf{y}). \tag{20}$$

Notice that it might be hard or even impossible to solve algebraically for x^i in terms of the y^j. Thus the straight forward approach to finding $\mathbf{e}^k$ by 20 might be impossible. Also, this approach leads to an expression in terms of the $\mathbf{y}$ coordinates rather than the desired $\mathbf{x}$ coordinates. Therefore, it is expedient to use another method to obtain these vectors. The vectors, $\mathbf{e}^k(\mathbf{x})$ may always be found by using formula 12 and the result is in terms of the curvilinear coordinates, $\mathbf{x}$. We illustrate in the following example.

Example 7 *$D \equiv (0, \infty) \times (0, \pi) \times (0, 2\pi)$ and*

$$\begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} x^1 \sin\left(x^2\right)\cos\left(x^3\right) \\ x^1 \sin\left(x^2\right)\sin\left(x^3\right) \\ x^1\cos\left(x^2\right) \end{pmatrix}$$

(We usually write this as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \rho\sin(\phi)\cos(\theta) \\ \rho\sin(\phi)\sin(\theta) \\ \rho\cos(\phi) \end{pmatrix}$$

where (ρ, ϕ, θ) are the spherical coordinates. We are calling them x^1, x^2, and x^3 to preserve the notation just discussed.) Thus

$$\mathbf{e}_1(\mathbf{x}) = \sin\left(x^2\right)\cos\left(x^3\right)\mathbf{i}_1 + \sin\left(x^2\right)\sin\left(x^3\right)\mathbf{i}_2 + \cos\left(x^2\right)\mathbf{i}_3,$$

$$\mathbf{e}_2(\mathbf{x}) = x^1\cos\left(x^2\right)\cos\left(x^3\right)\mathbf{i}_1$$

$$+ x^1\cos\left(x^2\right)\sin\left(x^3\right)\mathbf{i}_2 - x^1\sin\left(x^2\right)\mathbf{i}_3,$$

$$\mathbf{e}_3(\mathbf{x}) = -x^1\sin\left(x^2\right)\sin\left(x^3\right)\mathbf{i}_1 + x^1\sin\left(x^2\right)\cos\left(x^3\right)\mathbf{i}_2 + 0\mathbf{i}_3.$$

It follows the metric tensor is

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(x^1\right)^2 & 0 \\ 0 & 0 & \left(x^1\right)^2 \sin^2\left(x^2\right) \end{pmatrix} = \left(g_{ij}\right) = \left(\mathbf{e}_i \cdot \mathbf{e}_j\right). \tag{21}$$

Therefore, by Theorem 5

$$G^{-1} = \left(g^{ij}\right)$$

$$= \left(\mathbf{e}^i, \mathbf{e}^j\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(x^1\right)^{-2} & 0 \\ 0 & 0 & \left(x^1\right)^{-2} \sin^{-2}\left(x^2\right) \end{pmatrix}.$$

To obtain the dual basis, use Theorem 5 to write

$$\mathbf{e}^1 = g^{1j}\mathbf{e}_j = \mathbf{e}_1$$

$$\mathbf{e}^2 = g^{2j}\mathbf{e}_j = \left(x^1\right)^{-2}\mathbf{e}_2$$

$$\mathbf{e}^3 = g^{3j}\mathbf{e}_j = \left(x^1\right)^{-2}\sin^{-2}\left(x^2\right)\mathbf{e}_3.$$

It is natural to ask if we can get a transformation $\mathbf{M}$ such that

$$\frac{\partial \mathbf{M}}{\partial x^1} = \mathbf{i} = \mathbf{i}_1, \frac{\partial \mathbf{M}}{\partial x^2} = \mathbf{j} = \mathbf{i}_2, \frac{\partial \mathbf{M}}{\partial x^3} = \mathbf{k} = \mathbf{i}_3. \tag{22}$$

The answer is that we can. Let

$$\mathbf{M}\left(x^1, x^2, x^3\right) \equiv x^1\mathbf{i} + x^2\mathbf{j} + x^3\mathbf{k}.$$

Then 22 holds for this transformation.

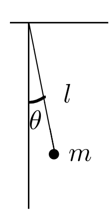
1.4 Exercises

1. Let

$$\begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} x^1 + 2x^2 \\ x^2 + x^3 \\ x^1 - 2x^2 \end{pmatrix}$$

where the y^i are the rectangular coordinates of the point. Find $\mathbf{e}^i, \mathbf{e}_i, i = 1, 2, 3$, and find $\left(g_{ij}\right)(\mathbf{x})$ and $\left(g^{ij}(\mathbf{x})\right)$.

2. Let $\mathbf{y} = \mathbf{y}(\mathbf{x}, t)$ where t signifies time and $\mathbf{x} \in U \subseteq \mathbb{R}^m$ for U an open set, while $\mathbf{y} \in \mathbb{R}^n$ and suppose $\mathbf{x}$ is a function of t. Physically, this corresponds to an object moving over a surface in $\mathbb{R}^n$ which may be changing as a function of t. The point $\mathbf{y} = \mathbf{y}(\mathbf{x}(t), t)$ is the point in $\mathbb{R}^n$ corresponding to t. For example, consider the pendulum



in which $n = 2, l$ is fixed and $y^1 = l\sin\theta, y^2 = l - l\cos\theta$. Thus, in this simple example, $m = 1$. If l were changing in a known way with respect to t, then this would be of the form $\mathbf{y} = \mathbf{y}(\mathbf{x}, t)$. The kinetic energy is defined as

$$T \equiv \frac{1}{2} m \dot{\mathbf{y}} \cdot \dot{\mathbf{y}} \qquad (*)$$

where the dot on the top signifies differentiation with respect to t. Show

$$\frac{\partial T}{\partial \dot{x}^k} = m\dot{\mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial x^k}.$$

Hint: First show

$$\dot{\mathbf{y}} = \frac{\partial \mathbf{y}}{\partial x^j} \dot{x}^j + \frac{\partial \mathbf{y}}{\partial t} \qquad (**)$$

and so

$$\frac{\partial \dot{\mathbf{y}}}{\partial \dot{x}^j} = \frac{\partial \mathbf{y}}{\partial x^j}.$$

3. ↑ Show

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}^k}\right) = m\ddot{\mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial x^k} + m\dot{\mathbf{y}} \cdot \frac{\partial^2 \mathbf{y}}{\partial x^k \partial x^r} \dot{x}^r + m\dot{\mathbf{y}} \cdot \frac{\partial^2 \mathbf{y}}{\partial t \partial x^k}.$$

4. ↑ Show

$$\frac{\partial T}{\partial x^k} = m\dot{\mathbf{y}} \cdot \left(\frac{\partial^2 \mathbf{y}}{\partial x^r \partial x^k} \dot{x}^r + \frac{\partial^2 \mathbf{y}}{\partial t \partial x^k}\right).$$

Hint: Use $*$ and $**$.

5. ↑ Now show from Newton's second law (mass times acceleration equals force) that for $\mathbf{F}$ the force,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}^k}\right) - \frac{\partial T}{\partial x^k} = m\ddot{\mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial x^k} = \mathbf{F} \cdot \frac{\partial \mathbf{y}}{\partial x^k}. \qquad (***)$$

6. ↑ In the example of the simple pendulum above,

$$\mathbf{y} = \begin{pmatrix} l \sin \theta \\ l - l \cos \theta \end{pmatrix} = l \sin \theta \mathbf{i} + (l - l \cos \theta) \mathbf{j}.$$

Use $***$ to find a differential equation which describes the vibrations of the pendulum in terms of θ. First write the kinetic energy and then consider the force acting on the mass which is

$$-mg\mathbf{j}.$$

7. The above problem is fairly easy to do without the formalism developed. Now consider the case where $\mathbf{x} = (\rho, \theta, \phi)$, spherical coordinates, and write differential equations for $\rho, \theta,$ and ϕ to describe the motion of an object in terms of these coordinates given a force, $\mathbf{F}$.

8. Suppose the pendulum is not assumed to vibrate in a plane. Let it be suspended at the origin and consider spherical coordinates. Find differential equations for θ and ϕ.

9. If there are many masses, $m_\alpha, \alpha = 1, \cdots, R$, the kinetic energy is the sum of the kinetic energies of the individual masses. Thus,

$$T \equiv \frac{1}{2} \sum_{\alpha=1}^{R} m_\alpha \left| \dot{\mathbf{y}}_\alpha \right|^2 .$$

Generalize the above problems to show that, assuming

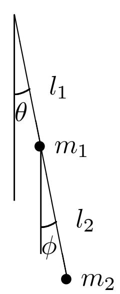
$$\mathbf{y}_\alpha = \mathbf{y}_\alpha (\mathbf{x}, t) ,$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}^k}\right) - \frac{\partial T}{\partial x^k} = \sum_{\alpha=1}^{R} \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{y}_\alpha}{\partial x^k}$$

where $\mathbf{F}_\alpha$ is the force acting on m_α.

10. Discuss the equivalence of these formulae with Newton's second law, force equals mass times acceleration. What is gained from the above so called Lagrangian formalism?

11. The double pendulum has two masses instead of only one.

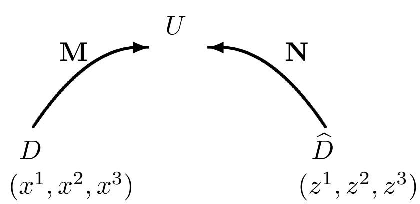


Write differential equations for θ and ϕ to describe the motion of the double pendulum.

1.5 Transformation of coordinates.

The scalars $\{x^i\}$ are called cuvilinear coordinates. Note they can be used to identify a point in $\mathbb{R}^n$ and $\mathbf{x} = (x^1, \cdots, x^n)$ is a point in $\mathbb{R}^n$. The basis vectors associated with this particular set of curvilinear coordinates at a point identified by $\mathbf{x}$ are denoted by $\mathbf{e}_i(\mathbf{x})$ and the dual basis vectors at this point are denoted by $\mathbf{e}^j(\mathbf{x})$. What if other curvilinear coordinates are used? How do we write $\mathbf{e}^k(\mathbf{x})$ in terms of the vectors, $\mathbf{e}^j(\mathbf{z})$ where $\mathbf{z}$ is some other type of curvilinear coordinates? We consider this topic next.

Consider the following picture in which U is an open set in $\mathbb{R}^n$, D, and $\widehat{D}$ are open sets in $\mathbb{R}^n$, and $\mathbf{M}, \mathbf{N}$ are C^2 mappings which are one to one from D and $\widehat{D}$ respectively. We will suppose that a point in U is identified by the curvilinear coordinates $\mathbf{x}$ in D and $\mathbf{z}$ in $\widehat{D}$.



Thus $\mathbf{M}(\mathbf{x}) = \mathbf{N}(\mathbf{z})$. Now by the chain rule,

$$\mathbf{e}_i(\mathbf{z}) \equiv \frac{\partial \mathbf{N}}{\partial z^i} = \frac{\partial \mathbf{M}}{\partial x^j}\frac{\partial x^j}{\partial z^i} = \frac{\partial x^j}{\partial z^i}\mathbf{e}_j(\mathbf{x}) . \tag{23}$$

We define the covariant and contravariant coordinates for the various curvilinear coordinates in the obvious way. Thus,

$$\mathbf{v} = v_i(\mathbf{x})\,\mathbf{e}^i(\mathbf{x}) = v^i(\mathbf{x})\,\mathbf{e}_i(\mathbf{x}) = v_j(\mathbf{z})\,\mathbf{e}^j(\mathbf{z}) = v^j(\mathbf{z})\,\mathbf{e}_j(\mathbf{z})\,.$$

Then we have the following theorem about transforming the vectors and coordinates.

Theorem 8 *The following transformation rules hold for pairs of curvilinear coordinates.*

$$v_i(\mathbf{z}) = \frac{\partial x^j}{\partial z_i} v_j(\mathbf{x})\,,\ v^i(\mathbf{z}) = \frac{\partial z^i}{\partial x^j} v^j(\mathbf{x})\,, \tag{24}$$

$$\mathbf{e}_i(\mathbf{z}) = \frac{\partial x^j}{\partial z_i}\mathbf{e}_j(\mathbf{x})\,,\ \mathbf{e}^i(\mathbf{z}) = \frac{\partial z^i}{\partial x^j}\mathbf{e}^j(\mathbf{x})\,, \tag{25}$$

$$g_{ij}(\mathbf{z}) = \frac{\partial x^r}{\partial z^i}\frac{\partial x^s}{\partial z^j} g_{rs}(\mathbf{x})\,,\ g^{ij}(\mathbf{z}) = \frac{\partial z^i}{\partial x^r}\frac{\partial z^j}{\partial x^s} g^{rs}(\mathbf{x})\,. \tag{26}$$

Proof: We already have shown the first part of 25 in 23. Then, from 23,

$$\begin{aligned}
\mathbf{e}^i(\mathbf{z}) &= \mathbf{e}^i(\mathbf{z})\cdot\mathbf{e}_j(\mathbf{x})\,\mathbf{e}^j(\mathbf{x}) = \mathbf{e}^i(\mathbf{z})\cdot\frac{\partial z^k}{\partial x^j}\mathbf{e}_k(\mathbf{z})\,\mathbf{e}^j(\mathbf{x}) \\
&= \delta^i_k \frac{\partial z^k}{\partial x^j}\mathbf{e}^j(\mathbf{x}) = \frac{\partial z^i}{\partial x^j}\mathbf{e}^j(\mathbf{x})
\end{aligned}$$

and this proves the second part of 25. Now to show 24,

$$v_i(\mathbf{z}) = \mathbf{v}\cdot\mathbf{e}_i(\mathbf{z}) = \mathbf{v}\cdot\frac{\partial x^j}{\partial z_i}\mathbf{e}_j(\mathbf{x}) = \frac{\partial x^j}{\partial z_i} v_j(\mathbf{x})$$

and

$$v^i(\mathbf{z}) = \mathbf{v}\cdot\mathbf{e}^i(\mathbf{z}) = \mathbf{v}\cdot\frac{\partial z^i}{\partial x^j}\mathbf{e}^j(\mathbf{x}) = \frac{\partial z^i}{\partial x^j} v^j(\mathbf{x})\,.$$

To verify 26,

$$g_{ij}(\mathbf{z}) = \mathbf{e}_i(\mathbf{z})\cdot\mathbf{e}_j(\mathbf{z}) = \mathbf{e}_r(\mathbf{x})\frac{\partial x^r}{\partial z^i}\cdot\mathbf{e}_s(\mathbf{x})\frac{\partial x^s}{\partial z^j} = g_{rs}(\mathbf{x})\frac{\partial x^r}{\partial z^i}\frac{\partial x^s}{\partial z^j}.$$

This proves the theorem.

We will denote by $\mathbf{y}$ the curvilinear coordinates with the property that

$$\mathbf{e}^k(\mathbf{y}) = \mathbf{i}_k = \mathbf{e}_k(\mathbf{y})\,.$$

1.6 Differentiation and Christoffel Symbols

Let $\mathbf{F} : U \to \mathbb{R}^n$ be differentiable. We call $\mathbf{F}$ a vector field and it is used to model force, velocity, acceleration, or any other vector quantity which may change from point to point in U. Then

$$\frac{\partial \mathbf{F}(\mathbf{x})}{\partial x^j}$$

is a vector and so there exist scalars, $F^i_{,j}(\mathbf{x})$ and $F_{i,j}(\mathbf{x})$ such that

$$\frac{\partial \mathbf{F}(\mathbf{x})}{\partial x^j} = F^i_{,j}(\mathbf{x})\,\mathbf{e}_i(\mathbf{x}) = F_{i,j}(\mathbf{x})\,\mathbf{e}^j(\mathbf{x})\,. \tag{27}$$

We will see how these scalars transform when the coordinates are changed.

Theorem 9 *If* $\mathbf{x}$ *and* $\mathbf{z}$ *are curvilinear coordinates,*

$$F^r_{,s}(\mathbf{x}) = F^i_{,j}(\mathbf{z}) \frac{\partial x^r}{\partial z^i} \frac{\partial z^j}{\partial x^s}, \; F_{r,s}(\mathbf{x}) \frac{\partial x^r}{\partial z^i} \frac{\partial x^s}{\partial z^j} = F_{i,j}(\mathbf{z}). \tag{28}$$

Proof:

$$F^r_{,s}(\mathbf{x}) \mathbf{e}_r(\mathbf{x}) \equiv \frac{\partial \mathbf{F}(\mathbf{x})}{\partial x^s} = \frac{\partial \mathbf{F}(\mathbf{z})}{\partial z^j} \frac{\partial z^j}{\partial x^s} \equiv$$

$$F^i_{,j}(\mathbf{z}) \mathbf{e}_i(\mathbf{z}) \frac{\partial z^j}{\partial x^s} = F^i_{,j}(\mathbf{z}) \frac{\partial x^r}{\partial z^i} \frac{\partial z^j}{\partial x^s} \mathbf{e}_r(\mathbf{x})$$

which shows the first formula of 27. To show the other formula,

$$F_{i,j}(\mathbf{z}) \mathbf{e}^i(\mathbf{z}) \equiv \frac{\partial \mathbf{F}(\mathbf{z})}{\partial z^j} = \frac{\partial \mathbf{F}(\mathbf{x})}{\partial x^s} \frac{\partial x^s}{\partial z^j} \equiv$$

$$F_{r,s}(\mathbf{x}) \mathbf{e}^r(\mathbf{x}) \frac{\partial x^s}{\partial z^j} = F_{r,s}(\mathbf{x}) \frac{\partial x^r}{\partial z^i} \frac{\partial x^s}{\partial z^j} \mathbf{e}^i(\mathbf{z}),$$

and this shows the second formula for transforming these scalars.

Now $\mathbf{F}(\mathbf{x}) = F^i(\mathbf{x}) \mathbf{e}_i(\mathbf{x})$ and so by the product rule,

$$\frac{\partial \mathbf{F}}{\partial x^j} = \frac{\partial F^i}{\partial x^j} \mathbf{e}_i(\mathbf{x}) + F^i(\mathbf{x}) \frac{\partial \mathbf{e}_i(\mathbf{x})}{\partial x^j}.$$

Now $\frac{\partial \mathbf{e}_i(\mathbf{x})}{\partial x^j}$ is a vector and so there exist scalars, $\left\{ {k \atop ij} \right\}$ such that

$$\frac{\partial \mathbf{e}_i(\mathbf{x})}{\partial x^j} = \left\{ {k \atop ij} \right\} \mathbf{e}_k(\mathbf{x}).$$

Therefore,

$$\frac{\partial \mathbf{F}}{\partial x^j} = \frac{\partial F^k}{\partial x^j} \mathbf{e}_k(\mathbf{x}) + F^i(\mathbf{x}) \left\{ {k \atop ij} \right\} \mathbf{e}_k(\mathbf{x})$$

which shows

$$F^k_{,j}(\mathbf{x}) = \frac{\partial F^k}{\partial x^j} + F^i(\mathbf{x}) \left\{ {k \atop ij} \right\}.$$

This is sometimes called the covariant derivative.

These scalars are called the Christoffel symbols of the second kind. The next theorem is devoted to properties of these Christoffel symbols. Before stating the theorem, we recall that the mapping, $\mathbf{M}$ which defines the curvilinear coordinates is C^2. The reason for this is that we want to be able to assert the mixed partial derivatives are equal.

Theorem 10 *The Christoffel symbols of the second kind satisfy the following*

$$\frac{\partial \mathbf{e}_i(\mathbf{x})}{\partial x^j} = \left\{ {k \atop ij} \right\} \mathbf{e}_k(\mathbf{x}), \tag{29}$$

$$\frac{\partial \mathbf{e}^i(\mathbf{x})}{\partial x^j} = -\left\{ {i \atop kj} \right\} \mathbf{e}^k(\mathbf{x}), \tag{30}$$

$$\left\{ {k \atop ij} \right\} = \left\{ {k \atop ji} \right\}, \tag{31}$$

$$\left\{ {m \atop ik} \right\} = \frac{g^{jm}}{2} \left[\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} \right]. \tag{32}$$

Proof: Formula 29 is the definition of the Christoffel symbols. We verify 30 next. To do so, note

$$\mathbf{e}^i(\mathbf{x}) \cdot \mathbf{e}_k(\mathbf{x}) = \delta^i_k.$$

Then from the product rule,

$$\frac{\partial \mathbf{e}^i(\mathbf{x})}{\partial x^j} \cdot \mathbf{e}_k(\mathbf{x}) + \mathbf{e}^i(\mathbf{x}) \cdot \frac{\partial \mathbf{e}_k(\mathbf{x})}{\partial x^j} = 0.$$

Now from the definition,

$$\frac{\partial \mathbf{e}^i(\mathbf{x})}{\partial x^j} \cdot \mathbf{e}_k(\mathbf{x}) = -\mathbf{e}^i(\mathbf{x}) \cdot \begin{Bmatrix} r \\ kj \end{Bmatrix} \mathbf{e}_r(\mathbf{x}) = -\begin{Bmatrix} i \\ kj \end{Bmatrix}.$$

But also,

$$\frac{\partial \mathbf{e}^i(\mathbf{x})}{\partial x^j} = \frac{\partial \mathbf{e}^i(\mathbf{x})}{\partial x^j} \cdot \mathbf{e}_k(\mathbf{x}) \, \mathbf{e}^k(\mathbf{x}) = -\begin{Bmatrix} i \\ kj \end{Bmatrix} \mathbf{e}^k(\mathbf{x}).$$

This verifies 30.

Letting $\frac{\partial \mathbf{M}(\mathbf{x})}{\partial x^j} = \mathbf{e}_j(\mathbf{x})$, it follows from equality of mixed partial derivatives,

$$\begin{Bmatrix} k \\ ij \end{Bmatrix} \mathbf{e}_k(\mathbf{x}) = \frac{\partial \mathbf{e}_i}{\partial x^j} \equiv \frac{\partial^2 \mathbf{M}}{\partial x^j \partial x^i} = \frac{\partial^2 \mathbf{M}}{\partial x^i \partial x^j} = \frac{\partial \mathbf{e}_j}{\partial x^i} = \begin{Bmatrix} k \\ ji \end{Bmatrix} \mathbf{e}_k(\mathbf{x}),$$

which shows 31. It remains to show 32.

$$\frac{\partial g_{ij}}{\partial x^k} = \frac{\partial \mathbf{e}_i}{\partial x^k} \cdot \mathbf{e}_j + \mathbf{e}_i \cdot \frac{\partial \mathbf{e}_j}{\partial x^k} = \begin{Bmatrix} r \\ ik \end{Bmatrix} \mathbf{e}_r \cdot \mathbf{e}_j + \mathbf{e}_i \cdot \mathbf{e}_r \begin{Bmatrix} r \\ jk \end{Bmatrix}.$$

Therefore,

$$\frac{\partial g_{ij}}{\partial x^k} = \begin{Bmatrix} r \\ ik \end{Bmatrix} g_{rj} + \begin{Bmatrix} r \\ jk \end{Bmatrix} g_{ri}. \tag{33}$$

Switching i and k while remembering 31 yields

$$\frac{\partial g_{kj}}{\partial x^i} = \begin{Bmatrix} r \\ ik \end{Bmatrix} g_{rj} + \begin{Bmatrix} r \\ ji \end{Bmatrix} g_{rk}. \tag{34}$$

Now switching j and k in 33, we obtain,

$$\frac{\partial g_{ik}}{\partial x^j} = \begin{Bmatrix} r \\ ij \end{Bmatrix} g_{rk} + \begin{Bmatrix} r \\ jk \end{Bmatrix} g_{ri}. \tag{35}$$

Adding 33 to 34 and subtracting 35 yields

$$\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} = 2 \begin{Bmatrix} r \\ ik \end{Bmatrix} g_{rj}.$$

Now multiplying both sides by g^{jm} and using the fact shown earlier in Theorem 5 that

$$g_{rj} g^{jm} = \delta^m_r,$$

we obtain

$$2 \begin{Bmatrix} m \\ ik \end{Bmatrix} = g^{jm} \left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} \right)$$

which proves 32.

This is a very interesting formula because it shows the Christoffel symbols are completely determined by the metric tensor and its derivatives.

1.7 Gradients and divergence

In this section we express the gradient and the divergence of a vector field in general curvilinear coordinates. As before, $\mathbf{y}$ will denote the standard coordinates with respect to the usual basis vectors. Thus

$$\mathbf{N}(\mathbf{y}) \equiv y^k \mathbf{i}_k,\ \mathbf{e}_k(\mathbf{y}) = \mathbf{i}_k = \mathbf{e}^k(\mathbf{y}).$$

Let $\phi : U \to \mathbb{R}$ be a differentiable scalar function, sometimes called a "scalar field" in this subject. We write $\phi(\mathbf{x})$ to denote the value of ϕ at the point whose coordinates are $\mathbf{x}$. In general, we follow this practice for any field, vector or scalar. Thus $\mathbf{F}(\mathbf{x})$ is the value of a vector field at the point of U determined by the coordinates $\mathbf{x}$. If we are using the standard coordinates, we know what we mean by the gradient of ϕ. It is given by the following formula.

$$\nabla\phi(\mathbf{y}) = \frac{\partial\phi(\mathbf{y})}{\partial y^k}\mathbf{e}^k(\mathbf{y}).$$

Therefore, using the chain rule, if the coordinates of the point of U are given as $\mathbf{x}$,

$$\nabla\phi(\mathbf{x}) = \nabla\phi(\mathbf{y})$$

$$= \frac{\partial\phi(\mathbf{x})}{\partial x^r}\frac{\partial x^r}{\partial y^k}\frac{\partial y^k}{\partial x^s}\mathbf{e}^s(\mathbf{x}) = \frac{\partial\phi(\mathbf{x})}{\partial x^r}\delta^r_s\mathbf{e}^s(\mathbf{x}) = \frac{\partial\phi(\mathbf{x})}{\partial x^r}\mathbf{e}^r(\mathbf{x}).$$

This shows the covariant components of $\nabla\phi(\mathbf{x})$ are

$$(\nabla\phi(\mathbf{x}))_r = \frac{\partial\phi(\mathbf{x})}{\partial x^r}. \tag{36}$$

To find the contravariant components, we "raise the index" in the usual way. Thus

$$(\nabla\phi(\mathbf{x}))^r = g^{rk}(\mathbf{x})(\nabla\phi(\mathbf{x}))_k = g^{rk}(\mathbf{x})\frac{\partial\phi(\mathbf{x})}{\partial x^k}. \tag{37}$$

What about the divergence of a vector field? The divergence of a vector field, $\mathbf{F}$ defined on U is a scalar field, $\operatorname{div}(\mathbf{F})$ which we know from calculus to be

$$\frac{\partial F^k}{\partial y^k}(\mathbf{y}) = F^k_{,k}(\mathbf{y})$$

in terms of the usual coordinates $\mathbf{y}$. The reason the above equation holds in this case is that $\mathbf{e}_k(\mathbf{y})$ is a constant and so the Christoffel symbols are zero. We want an expression for the divergence in an arbitrary coordinate system. From Theorem 9,

$$F^i_{,j}(\mathbf{y}) = F^r_{,s}(\mathbf{x})\frac{\partial x^s}{\partial y^j}\frac{\partial y^i}{\partial x^r}$$

$$= \left(\frac{\partial F^r(\mathbf{x})}{\partial x^s} + F^k(\mathbf{x})\left\{\begin{matrix} r \\ ks \end{matrix}\right\}(\mathbf{x})\right)\frac{\partial x^s}{\partial y^j}\frac{\partial y^i}{\partial x^r}.$$

Letting $j = i$ yields

$$\begin{aligned}
\operatorname{div}(\mathbf{F}) &= \left(\frac{\partial F^r(\mathbf{x})}{\partial x^s} + F^k(\mathbf{x})\left\{\begin{matrix} r \\ ks \end{matrix}\right\}(\mathbf{x})\right)\frac{\partial x^s}{\partial y^i}\frac{\partial y^i}{\partial x^r} \\
&= \left(\frac{\partial F^r(\mathbf{x})}{\partial x^s} + F^k(\mathbf{x})\left\{\begin{matrix} r \\ ks \end{matrix}\right\}(\mathbf{x})\right)\delta^s_r \\
&= \left(\frac{\partial F^r(\mathbf{x})}{\partial x^r} + F^k(\mathbf{x})\left\{\begin{matrix} r \\ kr \end{matrix}\right\}(\mathbf{x})\right).
\end{aligned} \tag{38}$$

We will simplify $\left\{ {r \atop kr} \right\}$ using the description of it in Theorem 10. Thus, from this theorem,

$$\left\{ \begin{matrix} r \\ rk \end{matrix} \right\} = \frac{g^{jr}}{2} \left[\frac{\partial g_{rj}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^r} - \frac{\partial g_{rk}}{\partial x^j} \right]$$

Now consider $\frac{g^{jr}}{2}$ times the last two terms in $[\cdot]$. Relabeling the indices r and j in the second term implies

$$\frac{g^{jr}}{2} \frac{\partial g_{kj}}{\partial x^r} - \frac{g^{jr}}{2} \frac{\partial g_{rk}}{\partial x^j} = \frac{g^{jr}}{2} \frac{\partial g_{kj}}{\partial x^r} - \frac{g^{rj}}{2} \frac{\partial g_{jk}}{\partial x^r} = 0.$$

Therefore,

$$\left\{ \begin{matrix} r \\ rk \end{matrix} \right\} = \frac{g^{jr}}{2} \frac{\partial g_{rj}}{\partial x^k}. \tag{39}$$

Now recall $g \equiv \det(g_{ij}) = \det(G) > 0$ from Theorem 5. Also from the formula for the inverse of a matrix and this theorem,

$$g^{jr} = A^{rj} (\det G)^{-1} = A^{jr} (\det G)^{-1}$$

where A^{rj} is the $rjth$ cofactor of the matrix (g_{ij}). Also recall that

$$g = \sum_{r=1}^{n} g_{rj} A^{rj} \text{ no sum on } j.$$

Therefore, g is a function of the variables $\{g_{rj}\}$ and

$$\frac{\partial g}{\partial g_{rj}} = A^{rj}.$$

From 39,

$$\left\{ \begin{matrix} r \\ rk \end{matrix} \right\} = \frac{g^{jr}}{2} \frac{\partial g_{rj}}{\partial x^k} = \frac{1}{2g} \frac{\partial g_{rj}}{\partial x^k} A^{jr} = \frac{1}{2g} \frac{\partial g}{\partial g_{rj}} \frac{\partial g_{rj}}{\partial x^k} = \frac{1}{2g} \frac{\partial g}{\partial x^k}$$

and so from 38,

$$\operatorname{div}(\mathbf{F}) = \frac{\partial F^k(\mathbf{x})}{\partial x^k} +$$

$$+F^k(\mathbf{x}) \frac{1}{2g(\mathbf{x})} \frac{\partial g(\mathbf{x})}{\partial x^k} = \frac{1}{\sqrt{g(\mathbf{x})}} \frac{\partial}{\partial x^i} \left(F^i(\mathbf{x}) \sqrt{g(\mathbf{x})} \right). \tag{40}$$

This is our formula for the divergence of a vector field in general curvilinear coordinates.

The Laplacian of a scalar field is nothing more than the divergence of the gradient. In symbols,

$$\Delta \phi \equiv \nabla \cdot \nabla \phi$$

From 40 and 37 it follows

$$\Delta \phi(\mathbf{x}) = \frac{1}{\sqrt{g(\mathbf{x})}} \frac{\partial}{\partial x^i} \left(g^{ik}(\mathbf{x}) \frac{\partial \phi(\mathbf{x})}{\partial x^k} \sqrt{g(\mathbf{x})} \right). \tag{41}$$

We summarize the conclusions of this section in the following theorem.

Theorem 11 *The following formulas hold for the gradient, divergence and Laplacian in general curvilinear coordinates.*

$$(\nabla\phi(\mathbf{x}))_r = \frac{\partial\phi(\mathbf{x})}{\partial x^r}, \tag{42}$$

$$(\nabla\phi(\mathbf{x}))^r = g^{rk}(\mathbf{x})\frac{\partial\phi(\mathbf{x})}{\partial x^k}, \tag{43}$$

$$\operatorname{div}(\mathbf{F}) = \frac{1}{\sqrt{g(\mathbf{x})}}\frac{\partial}{\partial x^i}\left(F^i(\mathbf{x})\sqrt{g(\mathbf{x})}\right), \tag{44}$$

$$\Delta\phi(\mathbf{x}) = \frac{1}{\sqrt{g(\mathbf{x})}}\frac{\partial}{\partial x^i}\left(g^{ik}(\mathbf{x})\frac{\partial\phi(\mathbf{x})}{\partial x^k}\sqrt{g(\mathbf{x})}\right). \tag{45}$$

1.8 Exercises

1. Let $y^1 = x^1 + 2x^2, y^2 = x^2 + 3x^3, y^3 = x^1 + x^3$. Let

   $$\mathbf{F}(\mathbf{x}) = x^1\mathbf{e}_1(\mathbf{x}) + x^2\mathbf{e}_2(\mathbf{x}) + \left(x^3\right)^2\mathbf{e}(\mathbf{x}).$$

   Find $\operatorname{div}(\mathbf{F})(\mathbf{x})$.

2. For the coordinates of the preceding problem, and ϕ a scalar field, find

   $$(\nabla\phi(\mathbf{x}))^3$$

   in terms of the partial derivatives of ϕ taken with respect to the variables x^i.

3. Let $y^1 = 7x^1 + 2x^2, y^2 = x^2 + 3x^3, y^3 = x^1 + x^3$. Let ϕ be a scalar field. Find $\nabla^2\phi(\mathbf{x})$.

4. Derive $\nabla^2 u$ in cylindrical coordinates, r, θ, z, where u is a scalar field on $\mathbb{R}^3$.

   $$x = r\cos\theta,\ y = r\sin\theta,\ z = z.$$

5. ↑ Find all solutions to $\nabla^2 u = 0$ which depend only on r where $r \equiv \sqrt{x^2+y^2}$.

6. Let u be a scalar field on $\mathbb{R}^3$. Find all solutions to $\nabla^2 u = 0$ which depend only on

   $$\rho \equiv \sqrt{x^2+y^2+z^2}.$$

7. The temperature, u, in a solid satisfies $\nabla^2 u = 0$ after a long time. Suppose in a long pipe of inner radius 9 and outer radius 10 the exterior surface is held at 100° while the inner surface is held at 200° find the temperature in the solid part of the pipe.

8. Show

   $$\left\{\begin{matrix} l \\ ij \end{matrix}\right\} = \frac{\partial\mathbf{e}_i}{\partial x^j}\cdot\mathbf{e}^l.$$

   Find the Christoffel symbols of the second kind for spherical coordinates in which $x^1 = \phi$, $x^2 = \theta$, and $x^3 = \rho$. Do the same for cylindrical coordinates letting $x^1 = r$, $x^2 = \theta$, $x^3 = z$.

9. Show velocity can be expressed as $\mathbf{v} = v_i(\mathbf{x})\,\mathbf{e}^i(\mathbf{x})$, where

$$v_i(\mathbf{x}) = \frac{\partial r_i}{\partial x^j}\frac{dx^j}{dt} - r_p(\mathbf{x})\begin{Bmatrix} p \\ ik \end{Bmatrix}\frac{dx^k}{dt}$$

and $r_i(\mathbf{x})$ are the covariant components of the displacement vector,

$$\mathbf{r} = r_i(\mathbf{x})\,\mathbf{e}^i(\mathbf{x}).$$

10. ↑ Using problem 8 and 9, show the covariant components of velocity in spherical coordinates are

$$v_1 = \rho^2\frac{d\phi}{dt},\ v_2 = \rho^2\sin^2(\phi)\frac{d\theta}{dt},\ v_3 = \frac{d\rho}{dt}.$$

Hint: First observe that if $\mathbf{r}$ is the position vector from the origin, then $\mathbf{r} = \rho\mathbf{e}_3$ so $r_1 = 0 = r_2$, and $r_3 = \rho$. Now use 9.

1.9 Curl and cross products

In this section we consider the curl and cross product in general curvilinear coordinates in $\mathbb{R}^3$. We will always assume that for $\mathbf{x}$ a set of curvilinear coordinates,

$$\det\left(\frac{\partial y^i}{\partial x^j}\right) > 0 \tag{46}$$

Where the $\mathbf{y}_i$ are the usual coordinates in which $\mathbf{e}_k(\mathbf{y}) = \mathbf{i}_k$.

Theorem 12 *Let 46 hold. Then*

$$\det\left(\frac{\partial y^i}{\partial x^j}\right) = \sqrt{g(\mathbf{x})} \tag{47}$$

and

$$\det\left(\frac{\partial x^i}{\partial y^j}\right) = \frac{1}{\sqrt{g(\mathbf{x})}}. \tag{48}$$

Proof:

$$\mathbf{e}_i(\mathbf{x}) = \frac{\partial y^k}{\partial x^i}\mathbf{i}_k$$

and so

$$g_{ij}(\mathbf{x}) = \frac{\partial y^k}{\partial x^i}\mathbf{i}_k \cdot \frac{\partial y^l}{\partial x^j}\mathbf{i}_l = \frac{\partial y^k}{\partial x^i}\frac{\partial y^k}{\partial x^j}.$$

Therefore, $g = \det(g_{ij}(\mathbf{x})) = \left(\det\left(\frac{\partial y^k}{\partial x^i}\right)\right)^2$. By 46, $\sqrt{g} = \det\left(\frac{\partial y^k}{\partial x^i}\right)$ as claimed. Now

$$\frac{\partial y^k}{\partial x^i}\frac{\partial x^i}{\partial y^r} = \delta^k_r$$

and so

$$\det\left(\frac{\partial x^i}{\partial y^r}\right) = \frac{1}{\sqrt{g(\mathbf{x})}}.$$

This proves the theorem.

To get the curl and cross product in curvilinear coordinates, let ϵ^{ijk} be the usual permutation symbol. Thus,

$$\epsilon^{123} = 1$$

and when any two indices in ϵ^{ijk} are switched, the sign changes. Thus

$$\epsilon^{132} = -1, \epsilon^{312} = 1, \text{ etc.}$$

Now define

$$\varepsilon^{ijk}(\mathbf{x}) \equiv \epsilon^{ijk} \frac{1}{\sqrt{g(\mathbf{x})}}.$$

Then for $\mathbf{x}$ and $\mathbf{z}$ satisfying 46,

$$\varepsilon^{ijk}(\mathbf{x}) \frac{\partial z^r}{\partial x^i} \frac{\partial z^s}{\partial x^j} \frac{\partial z^t}{\partial x^k} = \epsilon^{ijk} \det\left(\frac{\partial x^p}{\partial y^q}\right) \frac{\partial z^r}{\partial x^i} \frac{\partial z^s}{\partial x^j} \frac{\partial z^t}{\partial x^k}$$

$$= \epsilon^{rst} \det\left(\frac{\partial x^p}{\partial y^q}\right) \det\left(\frac{\partial z^i}{\partial x^k}\right) = \epsilon^{rst} \det(MN)$$

where N is the matrix whose pq th entry is $\frac{\partial x^p}{\partial y^q}$ and M is the matrix whose ik th entry is $\frac{\partial z^i}{\partial x^k}$. Therefore, from the definition of matrix multiplication and the chain rule, this equals

$$= \epsilon^{rst} \det\left(\frac{\partial z^i}{\partial y^p}\right) \equiv \varepsilon^{rst}(\mathbf{z})$$

from the above discussion.

Now $\varepsilon^{ijk}(\mathbf{y}) = \epsilon^{ijk}$ and for a vector field, $\mathbf{F}$,

$$\operatorname{curl}(\mathbf{F}) \equiv \varepsilon^{ijk}(\mathbf{y}) F_{k,j}(\mathbf{y}) \mathbf{e}_i(\mathbf{y}).$$

Therefore, since we know how everything transforms assuming 46, it is routine to write this in terms of $\mathbf{x}$.

$$\operatorname{curl}(\mathbf{F}) = \varepsilon^{rst}(\mathbf{x}) \frac{\partial y^i}{\partial x^r} \frac{\partial y^j}{\partial x^s} \frac{\partial y^k}{\partial x^t} F_{p,q}(\mathbf{x}) \frac{\partial x^p}{\partial y^k} \frac{\partial x^q}{\partial y^j} \mathbf{e}_m(\mathbf{x}) \frac{\partial x^m}{\partial y^i}$$

$$= \varepsilon^{rst}(\mathbf{x}) \delta_r^m \delta_s^q \delta_t^p F_{p,q}(\mathbf{x}) \mathbf{e}_m(\mathbf{x}) = \varepsilon^{mqp}(\mathbf{x}) F_{p,q}(\mathbf{x}) \mathbf{e}_m(\mathbf{x}). \tag{49}$$

More simplification is possible. Recalling the definition of $F_{p,q}(\mathbf{x})$,

$$\frac{\partial \mathbf{F}}{\partial x^q} \equiv F_{p,q}(\mathbf{x}) \mathbf{e}^p(\mathbf{x}) = \frac{\partial}{\partial x^q}\left[F_p(\mathbf{x}) \mathbf{e}^p(\mathbf{x})\right]$$

$$= \frac{\partial F_p(\mathbf{x})}{\partial x^q} \mathbf{e}^p(\mathbf{x}) + F_p(\mathbf{x}) \frac{\partial \mathbf{e}^p}{\partial x^q} = \frac{\partial F_p(\mathbf{x})}{\partial x^q} \mathbf{e}^p(\mathbf{x}) - F_r(\mathbf{x}) \left\{ \begin{matrix} r \\ pq \end{matrix} \right\} \mathbf{e}^p(\mathbf{x})$$

by Theorem 10. Therefore,

$$F_{p,q}(\mathbf{x}) = \frac{\partial F_p(\mathbf{x})}{\partial x^q} - F_r(\mathbf{x}) \left\{ \begin{matrix} r \\ pq \end{matrix} \right\}$$

and so

$$\operatorname{curl}(\mathbf{F}) = \varepsilon^{mqp}(\mathbf{x}) \frac{\partial F_p(\mathbf{x})}{\partial x^q} \mathbf{e}_m(\mathbf{x}) - \varepsilon^{mqp}(\mathbf{x}) F_r(\mathbf{x}) \begin{Bmatrix} r \\ pq \end{Bmatrix} \mathbf{e}_m(\mathbf{x}).$$

However, because $\left\{ {r \atop pq} \right\} = \left\{ {r \atop qp} \right\}$, the second term in this expression equals 0. To see this,

$$\varepsilon^{mqp}(\mathbf{x}) \begin{Bmatrix} r \\ pq \end{Bmatrix} = \varepsilon^{mpq}(\mathbf{x}) \begin{Bmatrix} r \\ qp \end{Bmatrix} = -\varepsilon^{mqp}(\mathbf{x}) \begin{Bmatrix} r \\ pq \end{Bmatrix}.$$

Therefore, by 49,

$$\mathit{curl}(\mathbf{F}) = \varepsilon^{mqp}(\mathbf{x}) \frac{\partial F_p(\mathbf{x})}{\partial x^q} \mathbf{e}_m(\mathbf{x}). \tag{50}$$

What about the cross product of two vector fields? Let $\mathbf{F}$ and $\mathbf{G}$ be two vector fields. Then in terms of standard coordinates, $\mathbf{y}$,

$$\mathbf{F} \times \mathbf{G} = \varepsilon^{ijk}(\mathbf{y}) F_j(\mathbf{y}) G_k(\mathbf{y}) \mathbf{e}_i(\mathbf{y})$$

$$= \varepsilon^{rst}(\mathbf{x}) \frac{\partial y^i}{\partial x^r} \frac{\partial y^j}{\partial x^s} \frac{\partial y^k}{\partial x^t} F_p(\mathbf{x}) \frac{\partial x^p}{\partial y^j} G_q(\mathbf{x}) \frac{\partial x^q}{\partial y^k} \mathbf{e}_l(\mathbf{x}) \frac{\partial x^l}{\partial y^i}$$

$$= \varepsilon^{rst}(\mathbf{x}) \delta_s^p \delta_t^q \delta_r^l F_p(\mathbf{x}) G_q(\mathbf{x}) \mathbf{e}_l(\mathbf{x}) = \varepsilon^{lpq}(\mathbf{x}) F_p(\mathbf{x}) G_q(\mathbf{x}) \mathbf{e}_l(\mathbf{x}). \tag{51}$$

We summarize these results in the following theorem.

Theorem 13 *Suppose* $\mathbf{x}$ *is a system of curvilinear coordinates in* $\mathbb{R}^3$ *such that*

$$\det\left(\frac{\partial y^i}{\partial x^j}\right) > 0.$$

Let

$$\varepsilon^{ijk}(\mathbf{x}) \equiv \epsilon^{ijk} \frac{1}{\sqrt{g(\mathbf{x})}}.$$

Then the following formulas for curl and cross product hold in this system of coordinates.

$$\operatorname{curl}(\mathbf{F}) = \varepsilon^{mqp}(\mathbf{x}) \frac{\partial F_p(\mathbf{x})}{\partial x^q} \mathbf{e}_m(\mathbf{x}),$$

and

$$\mathbf{F} \times \mathbf{G} = \varepsilon^{lpq}(\mathbf{x}) F_p(\mathbf{x}) G_q(\mathbf{x}) \mathbf{e}_l(\mathbf{x}).$$