Chapter 15 ∇ in other coordinates

On a number of occasions we have noticed that del is geometrically determined — it does not depend on a choice of coordinates for $\mathbb{R}^n$. This was shown to be true for ∇f, the *gradient* of a function from $\mathbb{R}^n$ to $\mathbb{R}$ (Section 2H). It was also verified for $\nabla \bullet F$, the *divergence* of a function from $\mathbb{R}^n$ to $\mathbb{R}^n$ (Section 14B). And in the case $n = 3$, we saw in Section 13G that it is true for $\nabla \times F$, the *curl* of a function from $\mathbb{R}^3$ to $\mathbb{R}^3$.

These three instances beg the question of how we might express ∇ in other coordinate systems for $\mathbb{R}^n$. A recent example of this is found in Section 13G, where a formula is given for $\nabla \times F$ in terms of an arbitrary right-handed orthonormal frame for $\mathbb{R}^3$. We shall accomplish much more in this chapter.

A very interesting book about ∇, by Harry Moritz Schey, has the interesting title *Div, Grad, Curl, And All That.*

A. **Biorthogonal systems**

We begin with some elementary linear algebra. Consider an arbitrary frame $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ for $\mathbb{R}^n$. We know of course that the Gram matrix is of great interest:

$$G = (\varphi_i \bullet \varphi_j).$$

This is a symmetric positive definite matrix, and we shall denote its entries as

$$g_{ij} = \varphi_i \bullet \varphi_j.$$

Of course, G is the identity matrix $\iff$ we have an *orthonormal* frame.

We also form the matrix Φ whose columns are the vectors φ_i. Symbolically we write

$$\Phi = (\varphi_1 \ \varphi_2 \dots \varphi_n).$$

We know that Φ is an orthogonal matrix $\iff$ we have an orthonormal frame (Problem 4–20).

Now denote by Ψ the transpose of the inverse of Φ, and express this new matrix in terms of its columns as

$$\Psi = (\psi_1 \ \psi_2 \dots \psi_n).$$

We then have the matrix product

$$\begin{aligned} \Psi^t \Phi \ &= \begin{pmatrix} \psi_1^t \\ \vdots \\ \psi_n^t \end{pmatrix} (\varphi_1 \dots \varphi_n) \\ &= (\psi_i \bullet \varphi_j). \end{aligned}$$

But since Ψ^t equals the inverse of Φ, we conclude that

$$\psi_i \bullet \varphi_j = \delta_{ij}.$$

DEFINITION. Two frames $\{\varphi_1, \dots, \varphi_n\}$ and $\{\psi_1, \dots, \psi_n\}$ are called a *biorthogonal system* if

$$\psi_i \bullet \varphi_j = \delta_{ij}.$$

Clearly, $\{\varphi_1, \dots, \varphi_n\}$ and $\{\varphi_1, \dots, \varphi_n\}$ form a biorthogonal system $\Longleftrightarrow$ the frame $\{\varphi_1, \dots, \varphi_n\}$ is an orthonormal one. So the present definition should be viewed as a generalization of the concept of orthonormal basis.

Also it is clear that the relation between the φ_i's and the ψ_i's given here is completely symmetric.

PROBLEM 15–1. Let a frame for $\mathbb{R}^2$ be $\{\hat{\imath}, a\hat{\imath} + b\hat{\jmath}\}$, where of course $b \neq 0$. Compute the corresponding $\{\psi_1, \psi_2\}$ which produces a biorthogonal system. Sketch all four vectors on a copy of $\mathbb{R}^2$.

PROBLEM 15–2. Let $\{\varphi_1, \varphi_2, \varphi_3\}$ be a frame for $\mathbb{R}^3$. Prove that the biorthogonal frame is given by

$$\psi_3 = \frac{\varphi_1 \times \varphi_2}{[\varphi_1, \varphi_2, \varphi_3]} \text{ etc.}$$

PROBLEM 15–3. Suppose $\{\varphi_1, \dots, \varphi_n\}$ is an *orthogonal* frame. Show that the corresponding vectors $\psi_1, \dots, \psi_n$ are given as

$$\psi_i = \frac{\varphi_i}{\|\varphi_i\|^2}.$$

PROBLEM 15–4. Given a frame $\{\varphi_1, \dots, \varphi_n\}$ denote

$$J = \det \Phi.$$

Prove that $\det G = J^2$.

Finally, a little more notation. It has become customary to denote the inverse of $G = (g_{ij})$ by

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

(called "raising the indices"). That is,

$$\sum_{k=1}^{n} g^{ik} g_{kj} = \sum_{k=1}^{n} g_{ik} g^{kj} = \delta_{ij}.$$

Inasmuch as

$$G = \Phi^t \Phi,$$

we conclude that

$$\begin{aligned} \Psi &= (\Phi^t)^{-1} \\ &= \Phi G^{-1}. \end{aligned}$$

In terms of entries of these matrices this equation means that

$$(\Psi)_{ij} = \sum_{k=1}^{n} (\Phi)_{ik} g^{kj}.$$

Thus we find that the columns satisfy

$$\psi_j = \sum_{k=1}^{n} \varphi_k g^{kj}.$$

Rewriting this equation gives

$$\psi_i = \sum_{j=1}^{n} g^{ij} \varphi_j.$$

The corresponding inverse equation is of course

$$\varphi_i = \sum_{j=1}^{n} g_{ij} \psi_j.$$

PROBLEM 15–5. Prove that for all $v \in \mathbb{R}^n$

$$v = \sum_{i=1}^{n} (\psi_i \bullet v) \varphi_i.$$

The result of this problem is that when you wish to express v as a linear combination of the basis vectors φ_i, the corresponding coefficients are given directly in terms of inner products with the ψ_i's. You might think of the formal expression

$$\sum_{i=1}^{n} (\psi_i \bullet (\ ))\varphi_i$$

as representing the identity linear function on $\mathbb{R}^n$. In other words,

PROBLEM 15–6. Prove that

$$I = \sum_{i=1}^{n} \varphi_i \psi_i^t.$$

B. **The gradient**

We continue with an arbitrary biorthogonal system on $\mathbb{R}^n$, and maintain the notation of Section A.

Suppose that $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ is differentiable at a point $x \in \mathbb{R}^n$. Our task is to express the vector $\nabla f(x) \in \mathbb{R}^n$ as a linear combination of the vectors φ_i in the given frame. This is quite an easy task. In fact, Problem 15–5 gives immediately

$$\nabla f(x) = \sum_{i=1}^{n} \nabla f(x) \bullet \psi_i \varphi_i.$$

The inner products in this formula are of course directional derivatives of f in the directions ψ_i. Thus, in the notation of Section 2C

$$\nabla f(x) = \sum_{i=1}^{n} Df(x; \psi_i)\varphi_i.$$

PROBLEM 15–7. Show that everything in this formula can be expressed entirely in terms of the frame $\{\varphi_1, \dots, \varphi_n\}$ by a double sum

$$\nabla f(x) = \sum_{i,j=1}^{n} g^{ij} Df(x; \varphi_j)\varphi_i.$$

C. **The divergence**

We want to discuss a vector field f defined on an open subset of $\mathbb{R}^n$. We can thus regard f as a function from $\mathbb{R}^n$ to $\mathbb{R}^n$, and as such it has a derivative. At a point x in its domain, the derivative $Df(x)$ is a linear transformation of $\mathbb{R}^n$ to $\mathbb{R}^n$, represented in terms of the standard coordinate basis $\hat{e}_1, \dots, \hat{e}_n$, by the $n \times n$ Jacobian matrix

$$\left(\frac{\partial f_i}{\partial x_j} \right).$$

It is the *trace* of this matrix which is $\nabla \bullet f$, the divergence of f. This observation is the key to our representation of $\nabla \bullet f$, and we need a simple fact from linear algebra:

THEOREM. *Suppose A is a real $n \times n$ matrix, and regard the φ_i's and ψ_i's as column vectors. Then*

$$\mathit{trace} A = \sum_{i=1}^{n} A\psi_i \bullet \varphi_i.$$

PROOF. We use the matrices Φ and Ψ from Section A, so that

$$A\psi_i \bullet \varphi_j = ji \text{ entry of the matrix } \Phi^{-t} A\Psi.$$

Thus

$$\begin{aligned} \sum_{i=1}^{n} A\psi_i \bullet \varphi_i &= \text{trace}(\Phi^t A\Psi) \\ &= \text{trace}(A\Psi\Phi^t) \quad \text{(Problem 4–2)} \\ &= \text{trace}(A), \end{aligned}$$

since Ψ and Φ^t are inverses.

QED

As a result we now have

$$\begin{aligned} \nabla \bullet f(x) &= \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} \\ &= \sum_{i=1}^{n} Df(x)\psi_i \bullet \varphi_i \\ &= \sum_{i=1}^{n} Df(x; \psi_i) \bullet \varphi_i. \end{aligned}$$

D. **The curl**

Before going into the representation of curl, we summarize what we have obtained to this point. In terms of a fixed biorthogonal system we have found

$$\nabla f(x) = \nabla f(x) = \sum_{i=1}^{n} Df(x;\psi_i)\varphi_i$$

and

$$\mathrm{div} F(x) = \nabla \bullet F(x) = \sum_{i=1}^{n} DF(x;\psi_i) \bullet \varphi_i.$$

These two formulas can be incorporated as a single expression

$$\nabla = \sum_{i=1}^{n} \varphi_i \ \textit{times}\ D(\ ;\psi_i)$$

where ∇ and the right side are both regarded as *operators*. They operate on scalar-valued functions to produce the gradient and on vector fields to produce the divergence. In the gradient situation "*times*" is scalar multiplication, whereas in the divergence situation it is dot product.

A good way to remember this formula is to replace the directional derivative

$$Df(x;\psi_i)$$

by the symbol

$$\frac{\partial f}{\partial \psi_i}$$

Then we have

$$\boxed{\nabla = \sum_{i=1}^{n} \varphi_i \ \textit{times}\ \frac{\partial}{\partial \psi_i}.}$$

This leads us to an educated guess for curl, where we just use the same formula. Thus we suppose that $\{\varphi_1, \varphi_2, \varphi_3\}$ and $\{\psi_1, \psi_2, \psi_3\}$ form a biorthogonal system for $\mathbb{R}^3$. Then we have

THEOREM. *For a vector field F on $\mathbb{R}^3$,*

$$\textit{curl} F(x) = \nabla \times F(x) = \sum_{i=1}^{3} \varphi_i \times DF(x;\psi_i).$$

PROOF. Every vector field F can be expressed in the given frame in the form

$$F(x) = \sum_{i=1}^{3} f_j(x)\varphi_j.$$

Thus it suffices to prove the formula for the special case of vector fields of the form $f_j(x)\varphi_j$, for $j = 1, 2, 3$. Thus we assume

$$F(x) = f(x)\varphi_j.$$

We have from the product rule of Problem 13–14,

$$\nabla \times F(x) = \nabla f(x) \times \varphi_j \quad \text{(since } \varphi_j \text{ is constant)}$$

and our formula for gradient gives

$$\begin{aligned}
\nabla \times F(x) &= \sum_{i=1}^{3} Df(x;\psi_i)\varphi_i \times \varphi_j \\
&= \sum_{i=1}^{3} \varphi_i \times Df(x;\psi_i)\varphi_j \\
&= \sum_{i=1}^{3} \varphi_i \times D(f\varphi_j)(x;\psi_i) \\
&= \sum_{i=1}^{3} \varphi_i \times DF(x;\psi_i).
\end{aligned}$$

QED

PROBLEM 15–8. Are you surprised that the formula for curl does not require the frame $\{\varphi_1, \varphi_2, \varphi_3\}$ to be right-handed? Explain what happens if φ_3 is replaced with $-\varphi_3$.

PROBLEM 15–9. Prove that

$$\begin{aligned}
\varphi_1 \times \varphi_2 &= J\psi_3, \\
\varphi_2 \times \varphi_3 &= J\psi_1, \\
\varphi_3 \times \varphi_1 &= J\psi_2.
\end{aligned}$$

PROBLEM 15–10. Show how the formulas of the preceding problem behave when each φ_i is replaced with the vector $a_i\varphi_i$, where the a_i's are nonzero scalars.

E. **Curvilinear coordinates**

All that we have done up to now is represent ∇ in terms of a fixed frame for $\mathbb{R}^n$. In practice, however, the coordinates themselves are changed in a perhaps nonlinear fashion. We still need convenient expressions for ∇ in terms of the new coordinates.

We shall call our coordinate system T. What we mean is that T is a C^1 function from $\mathbb{R}^n$ to $\mathbb{R}^n$, and has a C^1 inverse as well. (Usually the domains of T and T^{-1} will not be all of $\mathbb{R}^n$.) The notation we shall use is

$$x = T(u),$$

representing the formula which determines $x = (x_1, \dots, x_n)$ in terms of the *curvilinear* coordinates $u = (u_1, \dots, u_n)$. The inverse function

$$u = T^{-1}(x)$$

gives the curvilinear coordinates of the given point x.

The Jacobian matrix of T is of course

$$T'(u) = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \dots & \frac{\partial x_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial u_1} & \dots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}.$$

This matrix must be invertible and we denote its determinant by

$$J(u) = \det T'(u).$$

The columns of $T'(u)$ represent linearly independent vectors of $\mathbb{R}^n$, and we think of them as "attached" to the point $x = T(u)$. That is, they are tangent vectors to the manifold $\mathbb{R}^n$ at the point x. These vectors play the role of a *moving* frame for $\mathbb{R}^n$:

$$\varphi_i(u) = \frac{\partial T}{\partial u_i}, \quad 1 \le i \le n.$$

There are other natural tangent vectors at x, namely the gradients of the cordinates u_j viewed as functions of x. We denote them as

$$\psi_j(x) = \nabla u_j, \quad 1 \le j \le n.$$

Since u_j is the j^{th} coordinate of $T^{-1}(x)$, the gradient ∇u_j is the j^{th} row of the Jacobian matrix for T^{-1}. Therefore, since the chain rule implies

$$(T^{-1})'(T(u))T'(u) = I,$$

we have the matrix equation

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} (\varphi_1 \ \varphi_2 \dots \varphi_n) = I.$$

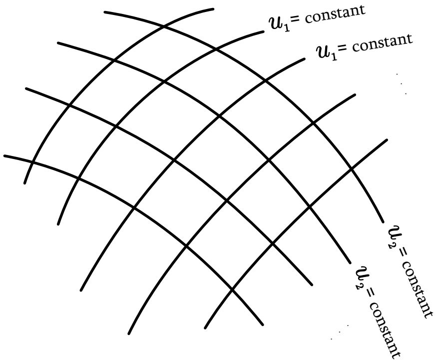
That is,

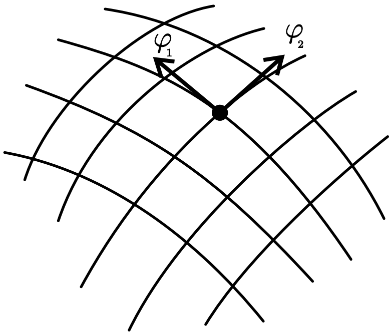
$$\psi_i \bullet \varphi_j = \delta_{ij}.$$

In other words, the frames at x given by

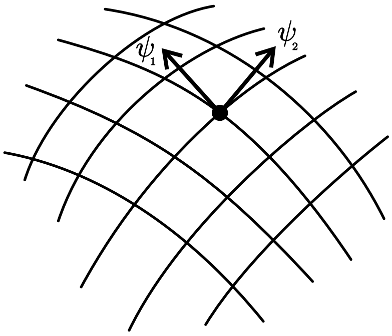
$$\{\varphi_1(T(x)), \dots, \varphi_n(T(x))\} \quad \text{and} \quad \{\psi_1(x), \dots, \psi_n(x)\}$$

are **biorthogonal**. Here's a representative sketch in the x_1—x_2 plane:





The φ_i`s are tangent to the coordinate curves



The ψ_i`s are orthogonal to the coordinate curves

Our point of view is going to be that when we work on $\mathbb{R}^n$ and use the coordinates $u_1, \ldots, u_n$, we are interested in calculating everything in terms of the "natural" vectors $\varphi_i(u)$ and the "natural" derivatives $\partial/\partial u_j$.

Of course, each ψ_j is a vector field on $\mathbb{R}^n$ — in fact, a gradient field. And each $\varphi_i \circ T$ is also a vector field on $\mathbb{R}^n$, but not necessarily a gradient field.

We still maintain the notations introduced in Section A. Thus

$$g_{ij}(u) = \varphi_i(u) \bullet \varphi_j(u),$$

and the matrix inverse to (g_{ij}) has entries

$$g^{ij}(u).$$

Furthermore, it follows that, with $x = T(u)$,

$$\psi_i(x) = \sum_{j=1}^{n} g^{ij}(u)\varphi_j(u)$$

and

$$\varphi_i(u) = \sum_{j=1}^{n} g_{ij}(u)\psi_j(x).$$

F. **The gradient**

The formula for ∇f goes over with no change from Section B. We read it off from Problem 15–7: with $x = T(u)$,

$$\nabla f(x) = \sum_{i,j=1}^{n} g^{ij}(u) Df(x;\varphi_j(u))\varphi_i(u).$$

We notice that the chain rule gives

$$\begin{aligned} Df(x;\varphi_j(u)) &= \nabla f(T(u)) \bullet \frac{\partial T(u)}{\partial u_j} \\ &= \frac{\partial}{\partial u_j} f(T(u)). \end{aligned}$$

NOTATION. In this context we denote the *pull-back* of f by the coordinate transformation $x = T(u)$ to be the function

$$f^*(u) = f(T(u)).$$

Thus we have

$$\boxed{\nabla f(x) = \sum_{i,j=1}^{n} g^{ij}(u) \frac{\partial f^*}{\partial u_j} \varphi_i(u).}$$

This is exactly the sort of formula we want. Everything on the right side of the equation is in terms of the coordinates we are working with.

G. **Spherical coordinates**

We pause for a significant example. We use our standard spherical coordinates for $\mathbb{R}^3$:

$$\begin{aligned} x &= r\sin\varphi\cos\theta, \\ y &= r\sin\varphi\sin\theta, \\ z &= r\cos\varphi. \end{aligned}$$

We then have the frame

$$\begin{aligned} \varphi_1 &= (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi), \\ \varphi_2 &= r(\cos\varphi\cos\theta, \cos\varphi\sin\theta, -\sin\varphi), \\ \varphi_3 &= r\sin\varphi(-\sin\theta, \cos\theta, 0). \end{aligned}$$

Thus $g_{ij} = 0$ if $i \neq j$, and

$$\begin{aligned} g_{11} &= 1, \\ g_{22} &= r^2, \\ g_{33} &= r^2\sin^2\varphi. \end{aligned}$$

Furthermore, $g^{ij} = 0$ if $i \neq j$, and

$$\begin{aligned} g^{11} &= 1, \\ g^{22} &= \frac{1}{r^2}, \\ g^{33} &= \frac{1}{r^2\sin^2\varphi}. \end{aligned}$$

Thus

$$\nabla f(x, y, z) = \frac{\partial f^*}{\partial r}\varphi_1 + \frac{1}{r^2}\frac{\partial f^*}{\partial \varphi}\varphi_2 + \frac{1}{r^2\sin^2\varphi}\frac{\partial f^*}{\partial \theta}\varphi_3.$$

Frequently this formula is displayed in terms of the *orthonormal* frame associated with $\{\varphi_1, \varphi_2, \varphi_3\}$. Namely, define

$$\hat{\varphi}_i = \frac{\varphi_i}{\|\varphi_i\|}.$$

Then it becomes

$$\nabla f = \frac{\partial f^*}{\partial r}\hat{\varphi}_1 + \frac{1}{r}\frac{\partial f^*}{\partial \varphi}\hat{\varphi}_2 + \frac{1}{r\sin\varphi}\frac{\partial f^*}{\partial \theta}\hat{\varphi}_3.$$

In fact, a nice alternative for this notation is to say that in general

$$\hat{u}_i = \frac{\frac{\partial T}{\partial u_i}}{\|\frac{\partial T}{\partial u_i}\|},$$

so the formula becomes

$$\nabla f = \frac{\partial f^*}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f^*}{\partial \varphi}\hat{\varphi} + \frac{1}{r \sin\varphi}\frac{\partial f^*}{\partial \theta}\hat{\theta}.$$

PROBLEM 15–11. Show that $J = r^2 \sin\varphi$.

PROBLEM 15–12. Carry out the same analysis for the case of cylindrical coordinates for $\mathbb{R}^3$:

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta, \\ z &= z. \end{aligned}$$

Conclude that

$$\nabla f = \frac{\partial f^*}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f^*}{\partial \theta}\hat{\theta} + \frac{\partial f^*}{\partial z}\hat{z}.$$

H. **The divergence**

A new situation arises in the curvilinear representation of divergence and curl. Namely we assume from the start that the vector field in question is represented in terms of curvilinear coordinates and the frame $\{\varphi_1, \dots, \varphi_n\}$ associated with them. Thus we treat the vector field $F(x)$ by first expressing it in the form

$$F(x) = \sum_{i=1}^{n} F_i(u)\varphi_i(u), \quad \text{where} \quad x = T(u).$$

We then want to express the scalar $\nabla \bullet F$ in terms of derivatives $\partial/\partial u_i$ involving the coefficients $F_i(u)$. Since the vector $\varphi_i(u)$ comes from T by differentiation and is liable to be subject to differentiation again, we need to assume T is of class C^2. We shall make this assumption for the remainder of the chapter.

It turns out that there is a huge simplification available if we first establish the remarkable

LEMMA. $\nabla \bullet (J^{-1}\varphi_i) = 0$.

PROOF. The vector field we are dealing with in this context is

$$G(x) = J^{-1}(u)\varphi_i(u), \quad \text{where} \;\; x = T(u).$$

We give a proof of this lemma based on the divergence theorem. To set this up, suppose h is any C^1 real-valued function on $\mathbb{R}^n$ which is zero outside a small ball. We shall prove that

$$\int_{\mathbb{R}^n} h\nabla \bullet G dx = 0. \tag{$*$}$$

Since $\nabla \bullet G$ is a continuous function, this will prove the result. For if $\nabla \bullet G(x_0) > 0$ for some x_0, then $\nabla \bullet G(x) > 0$ for all x in some neighborhood of x_0. We can then choose a suitable h to make the integrand in $(*)$ always ≥ 0 and positive in a neighborhood of x_0, so the integral in $(*)$ will be positive, a contradiction. Likewise if $\nabla \bullet G(x_0) < 0$.

The function h in this calculation is frequently called a *test function.* It is unimportant in itself, but is used to "test" the crucial function $\nabla \bullet G$.

We now turn to the proof of $(*)$. First,

$$\int_{\mathbb{R}^n} \nabla \bullet (hG) dx = 0$$

follows from the divergence theorem. Actually, just the fundamental theorem of calculus is needed, as a typical term in the Cartesian formula for divergence is

$$\int_{\mathbb{R}^n} \frac{\partial}{\partial x_j}(hG \bullet \hat{e}_j) dx,$$

and performing the x_j integration first gives the result, since h is zero outside a small ball.

Since the product rule gives

$$\nabla \bullet (hG) = \nabla h \bullet G + h\nabla \bullet G,$$

the proof of $(*)$ now reduces to proving that

$$\int_{\mathbb{R}^n} \nabla h \bullet G dx = 0.$$

In this integral we change variables with $x = T(u)$, and we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} \nabla h \bullet G dx &= \int_{\mathbb{R}^n} \nabla h(T(u)) \bullet G(T(u)) |\det T'(u)| du \\
&= \int_{\mathbb{R}^n} \nabla h(T(u)) \bullet J^{-1}(u)\varphi_i(u)|J(u)|du \\
&= \pm \int_{\mathbb{R}^n} \nabla h(T(u)) \bullet \frac{\partial T}{\partial u_i} du \\
&\overset{\text{chain rule}}{=} \pm \int_{\mathbb{R}^n} \frac{\partial}{\partial u_i}(h(T(u))du \\
&\overset{\text{FTC}}{=} 0.
\end{aligned}$$

QED

Now we compute the divergence of a typical summand in $F(x)$: the product rule gives

$$\begin{aligned}
\nabla \bullet (F_i(u)\varphi_i(u)) &= \nabla \bullet (J(u)F_i(u)\frac{1}{J(u)}\varphi_i(u)) \\
&= \nabla(J(u)F_i(u)) \bullet \frac{1}{J(u)}\varphi_i(u),
\end{aligned}$$

the other term being zero thanks to the lemma. And now we invoke the chain rule to get

$$\begin{aligned}
\frac{1}{J(u)}\nabla(J(u)F_i(u)) \bullet \varphi_i(u) &= \frac{1}{J}\nabla(JF_i) \bullet \frac{\partial T}{\partial u_i} \\
&= \frac{1}{J}\frac{\partial}{\partial u_i}(JF_i).
\end{aligned}$$

Thus we have our final formula: in case

$$F(x) = \sum_{i=1}^{n} F_i(u)\varphi_i(u), \quad x = T(u),$$

then

$$\boxed{\nabla \bullet F = \frac{1}{J(u)} \sum_{i=1}^{n} \frac{\partial (J(u)F_i)}{\partial u_i}.}$$

Notice how closely this resembles the Cartesian case!

For our spherical coordinate example, write

$$F(x, y, z) = F_1(r, \varphi, \theta)\hat{r} + F_2\hat{\varphi} + F_3\hat{\theta},$$

so that the corresponding functions in our formula are actually

$$F_1, F_2/r, \quad \text{and} \quad F_3/r\sin\varphi, \quad \text{respectively.}$$

Thus

$$\begin{aligned}\nabla \bullet F &= \frac{1}{r^2\sin\varphi}\left[\frac{\partial}{\partial r}(r^2\sin\varphi F_1) + \frac{\partial}{\partial\varphi}(r\sin\varphi F_2) + \frac{\partial}{\partial\theta}(rF_3)\right] \\ &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_1) + \frac{1}{r\sin\varphi}\frac{\partial}{\partial\varphi}(\sin\varphi F_2) + \frac{1}{r\sin\varphi}\frac{\partial F_3}{\partial\theta}.\end{aligned}$$

PROBLEM 15–13. For the case of cylindrical coordinates for $\mathbb{R}^3$ let

$$F(x,y,z) = F_1\hat{r} + F_2\hat{\theta} + F_3\hat{z},$$

and prove that

$$\nabla \bullet F = \frac{1}{r}\frac{\partial(rF_1)}{\partial r} + \frac{1}{r}\frac{\partial F_2}{\partial\theta} + \frac{\partial F_3}{\partial z}.$$

PROBLEM 15–14. Suppose that $\{\varphi_1,\dots,\varphi_n\}$ is a frame field for $\mathbb{R}^n$, so that it may vary from point to point. Suppose that the corresponding frame field $\{\psi_1,\dots,\psi_n\}$ yields a biorthogonal system. Prove that the basic formula of Section C is still valid:

$$\nabla \bullet F = \sum_{i=1}^{n} \varphi_i \bullet \frac{\partial F}{\partial\psi_i}.$$

I. **The curl**

We continue the notation of the preceding section, so that we are concerned with a vector field on $\mathbb{R}^3$, but instead of expanding F in terms of the frame $\{\varphi_1,\varphi_2,\varphi_3\}$, it seems better to work with $F(x)$ itself. We then follow the notation of Section F and define the pull-back

$$F^*(u) = F(T(u)).$$

We shall employ the formula for curl as given in Section D, but with the φ_i's and ψ_i's interchanged. So we find

$$\nabla \times F(x) = \sum_{i=1}^{3} \psi_i \times DF(x;\varphi_i).$$

Since $\varphi_i = \partial T/\partial u_i$, the chain rule gives with $x = T(u)$,

$$\nabla \times F(x) = \sum_{i=1}^{3} \psi_i(x) \times \frac{\partial F^*}{\partial u_i}(u).$$

Next we insert the algebra of Problem 15–8:

$$\psi_1 = J^{-1}\varphi_2 \times \varphi_3 \quad \text{etc.},$$

obtaining

$$\nabla \times F = J^{-1}(\varphi_2 \times \varphi_3) \times \frac{\partial F^*}{\partial u_1} + \cdots .$$

The formula for vector triple product of Problem 7–4 gives

$$J\nabla \times F = \varphi_2 \bullet \frac{\partial F^*}{\partial u_1}\varphi_3 - \varphi_3 \bullet \frac{\partial F^*}{\partial u_1}\varphi_2 + \cdots .$$

The product rule gives

$$\begin{aligned} J\nabla \times F = {} & \frac{\partial}{\partial u_1}(\varphi_2 \bullet F^*)\varphi_3 - \frac{\partial \varphi_2}{\partial u_1} \bullet F^* \varphi_3 \\ & - \frac{\partial}{\partial u_1}(\varphi_3 \bullet F^*)\varphi_2 + \frac{\partial \varphi_3}{\partial u_1} \bullet F^* \varphi_2 + \cdots . \end{aligned}$$

The other eight terms are obtained from these by cycling through the indices $1 \to 2 \to 3 \to 1$. Look at the terms in which F^* appears undifferentiated: the coefficient of φ_3 in these terms is

$$-\frac{\partial \varphi_2}{\partial u_1} \bullet F^* + \frac{\partial \varphi_1}{\partial u_2} \bullet F^*.$$

But

$$\frac{\partial \varphi_1}{\partial u_2} = \frac{\partial}{\partial u_2}\frac{\partial T}{\partial u_1} = \frac{\partial}{\partial u_1}\frac{\partial T}{\partial u_2} = \frac{\partial \varphi_2}{\partial u_1},$$

so that the coefficient of φ_3 just obtained is zero. Notice here that we have the assumption that T is of class C^2 so that the mixed partial derivatives are equal. Likewise for φ_1 and φ_2. Thus we have

$$J\nabla \times F = \frac{\partial}{\partial u_1}(\varphi_2 \bullet F^*)\varphi_3 - \frac{\partial}{\partial u_1}(\varphi_3 \bullet F^*)\varphi_2 + \cdots$$

(four more terms). This expression is precisely the formal determinant

$$\det \begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \varphi_1 \bullet F^* & \varphi_2 \bullet F^* & \varphi_3 \bullet F^* \end{pmatrix}.$$

This gives our final formula for curl:

$$\boxed{\nabla \times F = \frac{1}{J} \det \begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \varphi_1 \bullet F^* & \varphi_2 \bullet F^* & \varphi_3 \bullet F^* \end{pmatrix}}.$$

Of course, the pattern is just that of the original definition of curl in terms of Cartesian coordinates.

Now let's examine the spherical coordinate case:

$$F(x, y, z) = F_1(r, \varphi, \theta)\hat{r} + F_2\hat{\varphi} + F_3\hat{\theta}.$$

Then

$$\begin{aligned} \varphi_1 \bullet F^* &= \hat{r} \bullet F^* = F_1, \\ \varphi_2 \bullet F^* &= r\hat{\varphi} \bullet F^* = rF_2, \\ \varphi_3 \bullet F^* &= r \sin\varphi\hat{\theta} \bullet F^* = r\sin\varphi F_3, \end{aligned}$$

so that

$$\nabla \times F = \frac{1}{r^2 \sin\varphi} \det \begin{pmatrix} \hat{r} & r\hat{\varphi} & r\sin\varphi\hat{\theta} \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \varphi} & \dfrac{\partial}{\partial \theta} \\ F_1 & rF_2 & r\sin\varphi F_3 \end{pmatrix}.$$

PROBLEM 15–15. For the case of cylindrical coordinates in the notation of Problem 15–12, show that

$$\nabla \times F = \frac{1}{r} \det \begin{pmatrix} \hat{r} & r\hat{\theta} & \hat{z} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ F_1 & rF_2 & F_3 \end{pmatrix}.$$

In particular, if the vector field is planar,

$$F = F_1(r,\theta)\hat{r} + F_2(r,\theta)\hat{\theta},$$

conclude that

$$\nabla \times F = \left(\frac{\partial F_2}{\partial r} + \frac{1}{r} F_2 - \frac{1}{r} \frac{\partial F_1}{\partial \theta} \right) \hat{k}.$$

It is especially interesting to see the formula for the operator ∇ in terms of these curvilinear coordinates. The directional derivative in the direction $\varphi_i(u)$, as we have noticed, is just $\partial/\partial u_i$. Thus we may rewrite the general formula from Section D in the form

$$\nabla = \sum_{i=1}^{n} \psi_i \textit{ times } \frac{\partial}{\partial u_i},$$

or if we replace ψ by its definition,

$$\boxed{\nabla = \sum_{i=1}^{n} \nabla u_i \textit{ times } \frac{\partial}{\partial u_i}.}$$

As always, *times* means

- scalar multiplication if we are calculating gradient,
- dot product if we are calculating divergence,
- cross product if we are calculating curl.

This version of the expression for ∇ is rather easy to remember. If we think of ∇ loosely as $\partial/\partial x$, then the formula is similar to

$$\frac{\partial}{\partial x} = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x} \textit{ times } \frac{\partial}{\partial u_i},$$

or, even more briefly,

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \ \textit{times}\ \frac{\partial}{\partial u}.$$

J. **Orthogonal curvilinear coordinates**

Many coordinate systems that arise in specific applications have the feature that the curves in $\mathbb{R}^n$ determined by them are orthogonal. This means in our notation that for $i \neq j$

$$\varphi_i \bullet \varphi_j = \frac{\partial T}{\partial u_i} \bullet \frac{\partial T}{\partial u_j} = 0.$$

In other words, $g_{ij} = 0$ if $i \neq j$. Then the numbers of importance are

$$g_{ii}(u), \quad 1 \leq i \leq n,$$

and

$$g^{ij}(u) = \begin{cases} 0 & \text{if } \ i \neq j, \\ \frac{1}{g_{ii}(u)} & \text{if } \ i = j. \end{cases}$$

Moreover,

$$J^2 = g_{11} g_{22} \dots g_{nn}.$$

Notably, spherical coordinates and also cylindrical coordinates both fit this situation.

Our formulas for ∇ all simplify under the assumption of orthogonality. Thus the gradient is given by

$$\nabla f(x) = \sum_{i=1}^{n} \frac{1}{g_{ii}(u)} \frac{\partial f^*}{\partial u_i} \varphi_i(u).$$

The divergence is still the same, no simplification occuring at all.

The formula for curl is much simpler. For in the notation

$$F(x) = \sum_{i=1}^{3} F_i(u) \varphi_i(u)$$

we now have simply

$$\varphi_i \bullet F^* = g_{ii} F_i.$$

Thus

$$\nabla \times F = \frac{1}{g_{11} g_{22} g_{33}} \det \begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 \\ \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \\ g_{11} F_1 & g_{22} F_2 & g_{33} F_3 \end{pmatrix}.$$

PROBLEM 15–16. If you use instead the associated ortho*normal* frame so that

$$F(x) = \sum_{i=1}^{3} F_i(u)\hat{\varphi}_i(u),$$

show that

$$\nabla \times F = \frac{1}{g_{11}g_{22}g_{33}} \det \begin{pmatrix} \sqrt{g_{11}}\hat{\varphi}_1 & \sqrt{g_{22}}\hat{\varphi}_2 & \sqrt{g_{33}}\hat{\varphi}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \sqrt{g_{11}}F_1 & \sqrt{g_{22}}F_2 & \sqrt{g_{33}}F_3 \end{pmatrix}.$$

K. **The Laplacian**

Without a doubt the most important second order partial differential operator on $\mathbb{R}^n$ is the *Laplace operator*, or the *Laplacian*. It operates on a real-valued function by first forming its gradient and then forming the divergence of the resulting vector field. Thus the notation is ∇^2:

$$\nabla^2 f = \nabla \bullet (\nabla f).$$

Another common notation is $\Delta = \nabla^2$.

Thus we have in Cartesian coordinates

$$\nabla^2 f(x) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}.$$

Of course, we well recognize that ∇^2 has an intrinsic geometric meaning apart from any coordinate system, since the same is true of ∇f and of the divergence of a vector field.

Before giving the general formula in curvilinear coordinates we mention the linear case in which $x = T(u)$ is given by matrix multiplication

$$x = Au,$$

where $A = (a_{ij})$ is a real nonsingular $n \times n$ matrix. Then our notation gives the following:

$$\varphi_j = \frac{\partial x}{\partial u_j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = j^{\text{th}} \text{ column of } A;$$

$$\begin{aligned} \psi_i &= \nabla u_i = \nabla (A^{-1}x)_{i^{\text{th}} \text{ entry}} \\ &= \begin{pmatrix} a^{i1} \\ a^{i2} \\ \vdots \\ a^{in} \end{pmatrix} = \text{ transpose of } i^{\text{th}} \text{ row of } A^{-1}; \end{aligned}$$

$$J = \det A;$$

$$g_{ij} = \sum_{k=1}^{n} a_{ki} a_{kj};$$

$$g^{ij} = \sum_{k=1}^{n} a^{ik} a^{jk}.$$

PROBLEM 15–17. Prove that

$$\nabla^2 f(x) = \sum_{i,j=1}^{n} g^{ij} \frac{\partial^2 f^*(u)}{\partial u_i \partial u_j}.$$

PROBLEM 15–18. Prove that ∇^2 is *invariant* with respect to this change of variables, that is,

$$\nabla^2 f(x) = \sum_{i=1}^{n} \frac{\partial^2 f^*}{\partial u_i^2},$$

if and only if $A \in O(n)$.

Now we turn to the general case of curvilinear coordinates. We have from Section F

$$\nabla f(x) = \sum_{i,j=1}^{n} g^{ij}(u) \frac{\partial f^*}{\partial u_j} \varphi_i(u),$$

so the gradient vector field ∇f is represented as in Section H with the corresponding coefficients denoted there as $F_i(u)$ given by

$$F_i(u) = \sum_{j=1}^{n} g^{ij}(u) \frac{\partial f^*}{\partial u_j}(u).$$

Therefore the result of Section H gives immediately

$$\nabla^2 f(x) = \frac{1}{J} \sum_{i,j=1}^{n} \frac{\partial}{\partial u_i} \left(J g^{ij} \frac{\partial f^*}{\partial u_j} \right).$$

If we have the special case of *orthogonal* curvilinear coordinates, then of course

$$\begin{aligned} J^2 &= g_{11} g_{22} \dots g_{nn}, \\ g^{ii} &= \frac{1}{g_{ii}}, \end{aligned}$$

so we have

$$\nabla^2 f = \frac{1}{\sqrt{g_{11} \dots g_{nn}}} \sum_{i=1}^{n} \frac{\partial}{\partial u_i} \left(\frac{\sqrt{g_{11} \dots g_{nn}}}{g_{ii}} \frac{\partial f^*}{\partial u_i} \right).$$

For spherical coordinates on $\mathbb{R}^3$ this gives

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

PROBLEM 15–19. For the case of cylindrical coordinates in $\mathbb{R}^3$, show that

$$\nabla^2 f = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f^*}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f^*}{\partial \theta^2} + \frac{\partial^2 f^*}{\partial z^2}.$$

L. **Parabolic coordinates**

Here's an example that we have not yet mentioned. It's a coordinate system for $\mathbb{R}^2$ based on squaring complex numbers. Namely, in complex notation

$$x + iy = \frac{1}{2}(u + iv)^2.$$

In real notation,

$$\begin{cases} x &= \frac{1}{2}(u^2 - v^2), \\ y &= uv. \end{cases}$$

We assume that $u > 0$, $-\infty < v < \infty$. Then we produce all points of $\mathbb{R}^2$ except the negative x-axis $(-\infty, 0] \times \{0\}$.

PROBLEM 15–20. Write out explicit formulas for u, v as functions of x, y. (Be careful to have $u > 0$.) (See Problem 2–93.)

PROBLEM 15–21. We call u, v *parabolic coordinates* because the curves in the $x - y$ plane on which u is constant and also those along which v is constant are *parabolas*. Prove this, and also prove that the origin is the focus of each of these parabolas.

Our notation from Section E leads us to the frame vectors

$$\varphi_1(u, v) = (u, v),$$
$$\varphi_2(u, v) = (-v, u).$$

These vectors are orthogonal, so we have here orthogonal curvilinear coordinates.

PROBLEM 15–22. Show that

$$\nabla^2 f(x, y) = \frac{1}{u^2 + v^2} \left(\frac{\partial^2 f^*}{\partial u^2} + \frac{\partial^2 f^*}{\partial v^2} \right).$$