LAST TIME: Vector operators, divergence, curl, examples of line integrals and surface integrals, divergence theorem, Stokes' theorem, index notation, Kronecker delta, and Levi-Civita symbol.

What about derivatives and operators in index notation?

$$grad\ \varphi = \frac{\partial \varphi}{\partial x_i}, \quad div\ \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{\partial A_i}{\partial x_i}, \quad (curl\ \mathbf{A})_i = (\nabla \times \mathbf{A})_i = \varepsilon_{ijk}\frac{\partial A_j}{\partial x_k},$$

i runs from 1 to 3 in each case.

Just as a check, let's write Maxwell's equations in index notation for 3 dimensions.

$$\frac{\partial E_i}{\partial x_i} = \frac{\rho}{\epsilon_o}; \quad \frac{\partial B_i}{\partial x_i} = 0; \quad \varepsilon_{ijk}\frac{\partial E_j}{\partial x_k} = \frac{\partial B_i}{\partial t}; \quad \varepsilon_{ijk}\frac{\partial B_j}{\partial x_k} = \mu_o J_i + \epsilon_o \mu_o \frac{\partial E_i}{\partial t}.$$

Now, let's consider the definitions of a vector: an object that has both magnitude and direction, an object that satisfies the properties of a vector space, and an object that transforms according to a rotation matrix – an orthogonal matrix.

Vector space:

1. $\mathbf{A} + \mathbf{B} = \mathbf{C}$
2. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
3. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
4. $\mathbf{A} + \mathbf{0} = \mathbf{A} = \mathbf{0} + \mathbf{A}$
5. $c\mathbf{A} = \mathbf{B}$; $\mathrm{b}(\mathrm{c}\mathbf{A}) = (\mathrm{bc})\mathbf{A}$
6. $(b + c)\mathbf{A} = b\mathbf{A} + c\mathbf{A}$; $\quad c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\boldsymbol{B}$
7. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$
8. $1\mathbf{A} = \mathbf{A}$

Rotating the rectangular Cartesian coordinate system:

One definition of a vector is an object that transforms the same way as a point transforms under a rotation. I will show the general 3 – D rotation, but specialize to two dimensions to keep the mathematics simpler. The physics must be the same in any coordinate system. We start with a vector **A** represented in two different coordinate systems, one rotated through an angle with respect to the other.

Then

$$\mathbf{A} = A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}} = A_{x'}\hat{\mathbf{x}}' + A_{y'}\hat{\mathbf{y}}' + A_{z'}\hat{\mathbf{z}}'.$$

In general,

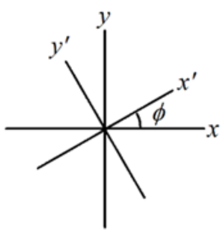
$$A_{x'} = \mathbf{A} \cdot \hat{\mathbf{x}}' = (\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}')A_x + (\hat{\mathbf{y}} \cdot \hat{\mathbf{x}}')A_y + (\hat{\mathbf{z}} \cdot \hat{\mathbf{x}}')A_z.$$

You obtain similar results for $A_{y'}$ and $A_{z'}$. You can see that this will all fit nicely into a matrix multiplication given by

$$\begin{bmatrix} A_{x'} \\ A_{y'} \\ A_{z'} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}\cdot\hat{\mathbf{x}}' & \hat{\mathbf{y}}\cdot\hat{\mathbf{x}}' & \hat{\mathbf{z}}\cdot\hat{\mathbf{x}}' \\ \hat{\mathbf{x}}\cdot\hat{\mathbf{y}}' & \hat{\mathbf{y}}\cdot\hat{\mathbf{y}}' & \hat{\mathbf{z}}\cdot\hat{\mathbf{y}}' \\ \hat{\mathbf{x}}\cdot\hat{\mathbf{z}}' & \hat{\mathbf{y}}\cdot\hat{\mathbf{z}}' & \hat{\mathbf{z}}\cdot\hat{\mathbf{z}}' \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix}.$$

You can also see how the transformation from $A_{x'}, A_{y'}, A_{z'}$ to A_x, A_y, A_z would go.

Here is the picture for two dimensions. We may now fill in the actual values for the scalar products based on this specific rotation.



$$\mathcal{R} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Successive rotations would be done in exactly the same way, but of course, they have to be done in the correct order because matrix multiplication does not commute. What about successive rotations about the same axis? Notice that if you take the scalar product of any row with any other row, it is orthogonal, meaning the scalar product is zero. Similarly, taking the scalar product of any column with any other column gives zero as well. Because of this feature, such rotation matrices are said to be orthogonal matrices and orthogonal transformations. They are length preserving transformations. (you will prove this in a homework assignment.) Vectors that satisfy these properties are said to be Cartesian vectors, although you will seldom hear anyone use that terminology. This property is easily summarized as follows.

$\sum_j \mathcal{R}_{ij}\,\mathcal{R}_{kj} = \delta_{ik}$; $\sum_i \mathcal{R}_{ij}\,\mathcal{R}_{ik} = \delta_{jk}$ are the orthogonal conditions for the matrix.

It is not difficult to show that one very special property is the $\mathcal{R}^{-1} = \mathcal{R}^T$.

So why bother with these ideas. Remember that we must have the laws of physics valid in all reference frames. Let's see how this works for Newton's second law of motion given by

$$F_i = ma_i.$$

Both F_i and a_i are vectors so we should be able to transform them to another rotated coordinate system where they have exactly the same form. Let's see how this works. Consider transforming each vector to the new system as follows.

$$F_i' = \mathcal{R}_{ij}F_j \;\; and \;\; a_i' = \mathcal{R}_{ij}a_j$$

We want to show explicitly that $F_i' = ma_i'$. To do this, we invert each of the equations to get

$$(\mathcal{R}^{-1})_{ki}F_i' = (\mathcal{R}^{-1})_{ki}\mathcal{R}_{ij}F_j \;\; and \;\; (\mathcal{R}^T)_{ki}F_i' = \delta_{kj}F_j \;\; so \; \mathcal{R}_{ik}F_i' = F_k.$$

Both a_i and F_i transform in the same way so we get

$$\mathcal{R}_{ji}F_j' = m\mathcal{R}_{ji}a_j' \ \ so\ \mathcal{R}_{ji}\left(F_j' - ma_j'\right) = 0.$$

Now multiply by $\mathcal{R}_{ki}$ and use the orthogonal condition to get $\mathcal{R}_{ki}\mathcal{R}_{ji}\left(F_j' - ma_j'\right) = 0$. Therefore,

$$\delta_{kj}\left(F_j' - ma_j'\right) = 0 \text{ and } F_k' = ma_k' \text{ so the equation is true in S}'\text{if it is true in S.}$$

Curvilinear coordinate systems

As you know, the operators grad, div, and curl appear in many fundamental physical laws. The symmetry of a problem dictates the form of the operator to use. Usually, we need plane polar, cylindrical, or spherical coordinates, but many other coordinate systems exist that might be appropriate for some cases. Among these are parabolic, elliptical, and oblate and prolate spheroidal systems. All of these, however, are orthogonal, so they can be quite useful and not too difficult to deal with.

The simplest operator is the gradient operator, and the easiest way to deal with it is to consider the line element in an orthogonal curvilinear coordinate system. Recall that we are dealing with systems have transformation rules of the following form given by

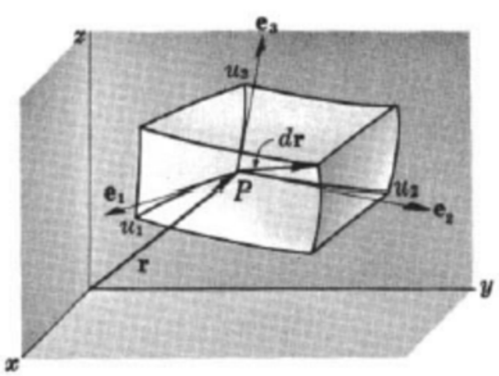
$$x = x(u_1, u_2, u_3);\ y = y(u_1, u_2, u_3); \text{and } z = z(u_1, u_2, u_3).$$

We need to be able to invert these to get

$$u_1 = u_1(x, y, z,);\quad u_2 = u_2(x, y, z,); \text{and } u_3 = u_3(x, y, z,).$$

These equations are just generalizations of converting back and forth from *x*, *y*, and *z* to *r*, ϕ, and θ, or to *r*, ϕ, and *z*, something you have done multiple times in your career so far.

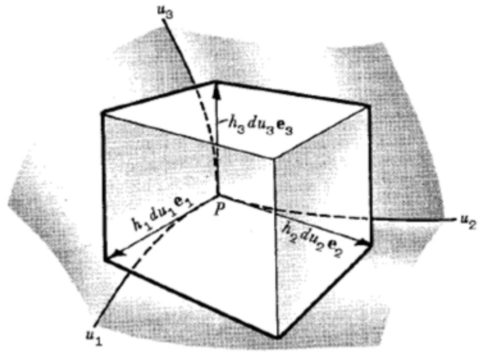
Referring to a figure similar one of our earlier ones, we introduce



$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u_1}du_1 + \frac{\partial \mathbf{r}}{\partial u_2}du_2 + \frac{\partial \mathbf{r}}{\partial u_3}du_3$$

$$= h_1\, du_1\hat{\mathbf{e}}_1 + h_2\, du_2\hat{\mathbf{e}}_2 + h_3\, du_3\hat{\mathbf{e}}_3,$$

where you can easily identify $h_i\ \hat{\mathbf{e}}_i$ as $\frac{\partial \mathbf{r}}{\partial u_i}$ with h_i known as scale factors $h_i = \left|\frac{\partial \mathbf{r}}{\partial u_i}\right|$. Note that *i* has values of 1, 2, and 3 here. Because this is an orthogonal system, all scalar products of $\hat{\mathbf{e}}_i$ with each other are zero except for equal values. The differential arc length is determined by $ds^2 = d\mathbf{r}\cdot d\mathbf{r} = (h_1\, du_1)^2 + (h_2\, du_2)^2 + (h_3\, du_3)^2$.
The figure on the following page shows more detail.



Now write $\boldsymbol{\nabla}\Phi = f_1\hat{\mathbf{e}}_1 + f_2\hat{\mathbf{e}}_2 + f_3\hat{\mathbf{e}}_3$ with $f_1, f_2,$ and f_3 to be determined. Next, use two different expressions for dΦ to get the values for $f_1, f_2,$ and f_3.

$$d\Phi = \boldsymbol{\nabla}\Phi \cdot d\mathbf{r} = h_1 f_1 du_1 + h_2 f_2 du_2 + h_3 f_3 du_3$$

and

$$d\Phi = \frac{\partial\Phi}{\partial u_1}du_1 + \frac{\partial\Phi}{\partial u_2}du_2 + \frac{\partial\Phi}{\partial u_3}du_3.$$

Comparing these two expressions shows that $h_1 f_1 = \frac{\partial\Phi}{\partial u_1}, h_2 f_2 = \frac{\partial\Phi}{\partial u_2}, and\ h_3 f_3 = \frac{\partial\Phi}{\partial u_3}$.

Solving for the values of *f* and inserting these values in the top equation gives the gradient operator as

$$\boldsymbol{\nabla} = \frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial}{\partial u_1} + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial}{\partial u_2} + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial}{\partial u_3}.$$

Once you know the line element, you have the values for the h_i's, and they are usually not hard to get. We will look at an example next Monday. We would also like to know the divergence and curl in general curvilinear coordinates. There are two general ways of determining these. One is a geometrical method using the curvilinear coordinates to get the volume and area. The other is an algebraic method that relies on being able to determine the effects of the del operator on a vector written in curvilinear coordinates. Your text shows the geometric method, so I will do the problem algebraically because it gives some insight into general curvilinear coordinates.

Consider a vector $\mathbf{A}$ given by $\mathbf{A} = A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2 + A_3\hat{\mathbf{e}}_3$. You might think that applying the divergence operator $\boldsymbol{\nabla}\cdot$ is easy, but you need to remember that $\hat{\mathbf{e}}_i$ is a function of position, so things must be handled with some care. The key point in what I will show you is that the orthogonality of the system means that there is no need to distinguish between the basis vectors for the covariant and contravariant components. Consider the gradient of u_1 given by

$$\boldsymbol{\nabla}u_1 = \frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial u_1}{\partial u_1} + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial u_1}{\partial u_2} + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial u_1}{\partial u_3} = \frac{\hat{\mathbf{e}}_1}{h_1}.$$

Why is this true? There are similar equations for u_2 and u_3. This means that

$$\boldsymbol{\nabla}u_2 \times \boldsymbol{\nabla}u_3 = \frac{\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3}{h_2 h_3} = \frac{\hat{\mathbf{e}}_1}{h_2 h_3} \text{ or } \hat{\mathbf{e}}_1 = h_2 h_3 \boldsymbol{\nabla}u_2 \times \boldsymbol{\nabla}u_3.$$

To show just one term of the calculation, consider $\boldsymbol{\nabla}\cdot(A_1\hat{\mathbf{e}}_1) = \boldsymbol{\nabla}\cdot(A_1 h_2 h_3 \boldsymbol{\nabla}u_2 \times \boldsymbol{\nabla}u_3)$.
Now use a vector identity to get

$$\nabla \cdot (A_1 \hat{\mathbf{e}}_1) = \nabla(A_1 h_2 h_3) \cdot (\nabla u_2 \times \nabla u_3) + (A_1 h_2 h_3) \nabla \cdot (\nabla u_2 \times \nabla u_3) =$$

$$\nabla(A_1 h_2 h_3) \cdot \left(\frac{\hat{\mathbf{e}}_2}{h_2} \times \frac{\hat{\mathbf{e}}_3}{h_3}\right) + \nabla u_3 \cdot \nabla \times \nabla u_2 - \nabla u_2 \cdot \nabla \times \nabla u_3 = \nabla(A_1 h_2 h_3) \cdot \frac{\hat{\mathbf{e}}_1}{h_2 h_3}.$$

The last step follows because the curl of the gradient is zero and because the cross product of $\hat{\mathbf{e}}_2$ with $\hat{\mathbf{e}}_3$ gives $\hat{\mathbf{e}}_1$. Now we only need to write out $\nabla(A_1 h_2 h_3)$. This is given by

$$\left[\frac{\hat{\mathbf{e}}_1}{h_1} \frac{\partial}{\partial u_1}(A_1 h_2 h_3) + \frac{\hat{\mathbf{e}}_2}{h_2} \frac{\partial}{\partial u_2}(A_1 h_2 h_3) + \frac{\hat{\mathbf{e}}_3}{h_3} \frac{\partial}{\partial u_3}(A_1 h_2 h_3)\right] \cdot \frac{\hat{\mathbf{e}}_1}{h_2 h_3} = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial u_1}(A_1 h_2 h_3).$$

Using the same procedure for each of the other two terms gives the divergence for any orthogonal, curvilinear coordinate system. Finally, the divergence is given by

$$\frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}(A_1 h_2 h_3) + \frac{\partial}{\partial u_2}(A_2 h_3 h_1) + \frac{\partial}{\partial u_3}(A_3 h_1 h_2)\right] = \nabla \cdot \mathbf{A}.$$

I will let you work out the curl in a similar way for homework.

NEXT TIME: Some examples, reciprocal basis vectors, more on covariant and contravariant vector components, introduce the metric, a few comments on 4-vectors.