Mathematical Preliminaries
(Marion & Thornton Chapter One)

Physics A300*

Fall 2002

Much of this will be familiar from past physics & math courses, but it's useful to review it & to formulate things in ways appropriate to a careful treatment of physics. All along, we will have two arenas in mind in which these quantities live: an abstract mathematical arena, where properties can be obtained by calculation & manipulation, and a geometrical/physical one, where properties are related to a picture of the real world.

Let's start with a few basic geometric objects:

1   Scalars

A scalar is just a single number, with an intrinsic value which can be defined without reference to any coördinate system or vector basis. (As an aside, this "number" may also have physical units attached to it, so it may be 68.1 kilograms or equivalently 6.81×10^4 grams, or may be a dimensionless number like 2π.) A physical example would be the mass of a particular particle or the length of a given curve through space.

2   Scalar Fields

A scalar field consists of a scalar value associated with each point in space. (Physical examples would be the temperature inside a room or the density of a fluid.) The value of the scalar field at a given point in space will again be an invariant, but since a coördinate system is needed to label each point for quantitative calculations, the actual function of those coördinate values will change. Mathematically, if the scalar field expressed in terms of the old coördinates is written as $\varphi(x, y, z)$, there will be a new functional form $\varphi'(x', y', z')$ used to describe the field as a function of the new coördinates.

This is most clearly illustrated by an example. Suppose the scalar field is such that in a particular Cartesian coördinate system $\{x, y, z\}$, it can be written as $\varphi(x, y, z) = axy$ where

---

*Copyright 2002, John T. Whelan, and all that

a is some constant. And suppose we choose new coördinates rotated 45° so that

$$x' = \frac{1}{\sqrt{2}}(x + y) \tag{1a}$$

$$y' = \frac{1}{\sqrt{2}}(-x + y) \tag{1b}$$

and thus

$$x = \frac{1}{\sqrt{2}}(x' - y') \tag{2a}$$

$$y = \frac{1}{\sqrt{2}}(x' + y') . \tag{2b}$$

Then by substitution $\varphi(x, y, z) = axy = \frac{1}{2}a(x'^2 - y'^2) = \varphi'(x', y', z')$.

3 Vectors

Lots of different definitions out there; focus on two:

- Mathematical: ordered set of numbers

- Physical: magnitude & direction

3.1 Mathematical Picture

Mathematically, we can think about 3-dimensional vectors as column vectors, i.e., 3×1 matrices. We know we can add & subtract these objects, and multiply them by a constant, e.g.:

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = \begin{pmatrix} A_1 + B_1 \\ A_2 + B_2 \\ A_3 + B_3 \end{pmatrix} \tag{3}$$

or

$$a\mathbf{A} = a \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} aA_1 \\ aA_2 \\ aA_3 \end{pmatrix} \tag{4}$$

Matrix algebra also lets us associate a row vector (1×3 matrix) $\mathbf{A}^{\mathrm{T}} = (A_1, A_2, A_3)$ with the column vector $\mathbf{A}$. We can then use matrix multiplication to define, for any two vectors $\mathbf{A}$ and $\mathbf{B}$, the "inner product"

$$\mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{pmatrix} A_1 & A_2 & A_3 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = A_1 B_1 + A_2 B_2 + A_3 B_3 = \sum_{i=1}^{3} A_i B_i \tag{5}$$

(1×1 matrix, i.e., a number) and the "outer product"

$$\mathbf{A}\mathbf{B}^{\mathrm{T}} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} = \begin{pmatrix} A_1 B_1 & A_1 B_2 & A_1 B_3 \\ A_2 B_1 & A_2 B_2 & A_2 B_3 \\ A_3 B_1 & A_3 B_2 & A_3 B_3 \end{pmatrix} \tag{6}$$

(3×3) i.e.,

$$(\mathbf{AB}^{\mathrm{T}})_{ij} = A_i B_j \tag{7}$$

Both the inner and outer product are special cases of matrix multiplication, so they obey a slew of nice properties usually associated with multiplication, in particular what could be called bilinearity:

$$(a\mathbf{A} + b\mathbf{B})^{\mathrm{T}}(c\mathbf{C} + d\mathbf{D}) = ac\mathbf{A}^{\mathrm{T}}\mathbf{C} + bc\mathbf{B}^{\mathrm{T}}\mathbf{C} + ad\mathbf{A}^{\mathrm{T}}\mathbf{D} + bd\mathbf{B}^{\mathrm{T}}\mathbf{D} \tag{8}$$

3.1.1 Basis Vectors

This can be further abstracted by introducing the *basis*

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{9}$$

Then a vector is related to its components by

$$\mathbf{A} = \sum_{i=1}^{3} A_i \mathbf{e}_i \ . \tag{10}$$

Since the inner products of the basis vectors with each other are

$$\mathbf{e}_i^{\mathrm{T}} \mathbf{e}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{11}$$

we can also calculate the components of $\mathbf{A}$ as

$$\mathbf{e}_i^{\mathrm{T}} \mathbf{A} = \mathbf{e}_i^{\mathrm{T}} \sum_{j=1}^{3} A_j \mathbf{e}_j = \sum_{j=1}^{3} A_j \delta_{ij} = A_i \tag{12}$$

3.1.2 Change of Basis

In an abstract sense, there's nothing all that special about the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ we've been using. We can replace them with some new set of three linearly independent vectors $\{\mathbf{e}_{\bar{1}}, \mathbf{e}_{\bar{2}}, \mathbf{e}_{\bar{3}}\}$. These vectors, like any vectors, can be expressed in terms of their components in the old basis:

$$\mathbf{e}_{\bar{k}} = \sum_{i=1}^{3} \left(\Lambda^{-1} \right)_{i\bar{k}} \mathbf{e}_i \tag{13}$$

where the components are

$$\left(\Lambda^{-1} \right)_{i\bar{k}} = \mathbf{e}_i^{\mathrm{T}} \mathbf{e}_{\bar{k}} \ . \tag{14}$$

The somewhat odd choice of notation will become clear in a moment.

We can resolve any vector in the new basis:

$$\mathbf{A} = \sum_{\overline{k}=1}^{3} A_{\overline{k}} \mathbf{e}_{\overline{k}} \tag{15}$$

which shows us that

$$A_i = \mathbf{e}_i^{\mathrm{T}} \mathbf{A} = \sum_{\overline{k}=1}^{3} A_{\overline{k}} \mathbf{e}_i^{\mathrm{T}} \mathbf{e}_{\overline{k}} = \sum_{\overline{k}=1}^{3} \left(\Lambda^{-1} \right)_{i\overline{k}} A_{\overline{k}} \tag{16}$$

Or, equivalently

$$A_{\overline{k}} = \sum_{i=1}^{3} \Lambda_{\overline{k}i} A_i \tag{17}$$

Of course, for this new set of basis vectors to be on equal footing with the original ones, they need to obey a relation analogous to (11), i.e.,

$$\begin{aligned}
\mathbf{e}_{\overline{k}}^{\mathrm{T}} \mathbf{e}_{\overline{\ell}} &= \left(\sum_{i=1}^{3} \left(\Lambda^{-1} \right)_{i\overline{k}} \mathbf{e}_i \right)^{\mathrm{T}} \left(\sum_{j=1}^{3} \left(\Lambda^{-1} \right)_{j\overline{\ell}} \mathbf{e}_j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\Lambda^{-1} \right)_{i\overline{k}} \left(\Lambda^{-1} \right)_{j\overline{\ell}} \delta_{ij} \\
&= \sum_{i=1}^{3} \left(\Lambda^{-1} \right)_{i\overline{k}} \left(\Lambda^{-1} \right)_{i\overline{\ell}} = \delta_{\overline{k}\overline{\ell}}
\end{aligned} \tag{18}$$

which limits the set of transformations we consider to take us between "equivalent" bases. In matrix notation, the last equality becomes

$$\left(\Lambda^{-1} \right)^{\mathrm{T}} \Lambda^{-1} = \mathbf{1} \tag{19}$$

or

$$\Lambda = \left(\Lambda^{-1} \right)^{\mathrm{T}} \tag{20}$$

which in terms of components means

$$\left(\Lambda^{-1} \right)_{i\overline{k}} = \Lambda_{\overline{k}i} \tag{21}$$

Such a transformation matrix Λ is called an *orthogonal matrix*. The rotation matrices described in section 1.4 are an example of orthogonal matrices.

3.2 Geometrical Picture

We can perform the manipulations of matrix algebra to our hearts' content, but what makes vectors so interesting and useful for physics is the second idea, that they represent quantities with a magnitude and direction in three-dimensional space. To emphasize the geometric significance of this picture we talk about a geometrical vector $\vec{A}$ to which the both the row vector $\mathbf{A}$ and the column vector $\mathbf{A}^{\mathrm{T}}$ correspond. The key to the correspondence will be the orthonormal basis vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$, which each have a length of 1 and which point in three perpendicular directions. (This basis also obeys the right-hand rule.)

Geometrically, we think of the magnitude and direction of $\vec{A}$ as being fundamental, rather than its components A_1, A_2, A_3 in a particular basis. So, if we need to deal with the components, we're free to rotate the trio of basis vectors however we like, to obtain any of the set of equivalent right-handed orthonormal bases.

Now, in an arbitrary orthonormal basis, the geometrical version of (10) is

$$\vec{A} = A_1\vec{e}_1 + A_2\vec{e}_2 + A_3\vec{e}_3 \tag{22}$$

But geometrically, this is just the resolution of $\vec{A}$ into its projections along the three basis vectors.

This leads naturally to familiar geometric constructions: First, that multiplying a vector $\vec{A}$ by a constant a gives a new vector $a\vec{A}$ with the same direction and a length multiplied by a. Second, that the sum $\vec{A} + \vec{B}$ of two vectors can be constructed by placing the tail of the first on the tip of the second.

3.2.1   Scalar Product

Now we're ready to talk about the geometrical meaning of the inner product. Geometrically, we call this the "dot product":

$$\vec{A} \cdot \vec{B} = \mathbf{A}^{\mathrm{T}}\mathbf{B} = \sum_{i=1}^{3} A_i B_i \tag{23}$$

But now let's use our freedom to rotate our basis vectors to choose a basis where $\vec{e}_1$ is parallel to $\vec{A}$ and where $\vec{B}$ lies in the plane spanned by $\vec{e}_1$ and $\vec{e}_2$. Then the vector $\vec{A}$ has magnitude $A = \left|\vec{A}\right|$ and a direction parallel to $\vec{e}_1$, which means that $\vec{A} = A\vec{e}_1$. The inner product is then

$$\vec{A} \cdot \vec{B} = A\vec{e}_1 \cdot \vec{B} = AB_1 \tag{24}$$

But now recall that B_1 is just the projection of $\vec{B}$ onto the direction $\vec{e}_1$, which is in this choice of basis parallel to $\vec{A}$. If the angle between $\vec{A}$ and $\vec{B}$ is θ, the projection of $\vec{B}$ parallel to $\vec{e}_1$ is (defining $B = \left|\vec{B}\right|$ as the magnitude of $\vec{B}$) $B_1 = B\cos\theta$ and the projection perpendicular is $B_2 = B\sin\theta$. This means

$$\vec{A} \cdot \vec{B} = \left|\vec{A}\right|\left|\vec{B}\right|\cos\theta \tag{25}$$

but this expression makes no reference to the basis we chose, only to three geometrical quantities: the magnitudes $\left|\vec{A}\right|$ and $\left|\vec{B}\right|$ of the two vectors and the angle θ between them. Thus the dot product of two vectors is a geometrical quantity whose value is independent of the basis used to calculate it, i.e., a scalar, which is why we also call it the "scalar product".

3.2.2   Tensor Product

The outer product of two vectors produces a 3×3 matrix $\mathbf{C} = \mathbf{A}\mathbf{B}^{\mathrm{T}}$. Geometrically, we represent this as $\overleftrightarrow{C} = \vec{A} \otimes \vec{B}$. $\overleftrightarrow{C}$ is a type of object known as a tensor, which is more difficult to visualize, and which we will put aside until we consider rigid body motion next semester.

3.2.3 Vector Product

So far, nothing we've done has been particularly special to three dimensions. The scalar product (and the tensor product) is fundamentally defined in linear algebra, and our treatment there could just as easily have been done in an arbitrary number N of dimensions. In contrast, something which can be defined in three dimensions, but which doesn't have a simple analog in any other number of dimensions, is an operation, called the "cross product" which multiplies two vectors $\vec{A}$ and $\vec{B}$ and gives a third vector $\vec{C}$. The underlying reason for this is that there are three vectors involved in that operation, and so it has a natural definition in three-dimensional space.

If we can derive the cross products $\vec{e}_i \times \vec{e}_j$ of the basis vectors with themselves, we'll have enough information to deduce the cross product of any two vectors. In particular, we'll want to work out the components

$$\epsilon_{ijk} = \vec{e}_i \times \vec{e}_j \cdot \vec{e}_k \tag{26}$$

for the $3^3 = 27$ different combinations of i, j, and k.

Derivation of the Vector Product (Extra!) Almost all of these values can be filled in from the geometrical requirement that the answer be the same for any orthonormal right-handed basis. We only need one definition to start off, and for that we use

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3 \tag{27}$$

(Here's where the fact that we have three basis vectors came in handy; $\vec{e}_3$ is the obvious thing to put on the right-hand side.) So we've now got

$$\epsilon_{121} = 0 \qquad \epsilon_{122} = 0 \qquad \epsilon_{123} = 1 \tag{28}$$

just from the definition.

First, let's note that we only need to fill in the $3^2 = 9$ elements $\{\epsilon_{1jk}\}$ and we should be able to deduce the other 18 elements $\{\epsilon_{2jk}\}$ and $\{\epsilon_{3jk}\}$. This is because we can go from one right-handed orthonormal basis to another by rotating through an angle of $120°$ about a line parallel to $\vec{e}_1 + \vec{e}_2 + \vec{e}_3$. This operation will permute the basis vectors:

$$\vec{e}_{\bar{1}} = \vec{e}_2 \qquad \vec{e}_{\bar{2}} = \vec{e}_3 \qquad \vec{e}_{\bar{3}} = \vec{e}_1 \tag{29}$$

Since any right-handed orthonormal basis should be as good as any other, $\epsilon_{\overline{k\ell m}} = \epsilon_{ijk}$ which in this case means we can permute 1, 2, and 3 when they appear as indices in ϵ_{ijk}. So this means, for example, (28) implies

$$\epsilon_{232} = 0 \qquad \epsilon_{233} = 0 \qquad \epsilon_{231} = 1 \tag{30}$$

and similarly, applying the permutation again allows us to deduce $\{\epsilon_{3jk}\}$ from $\{\epsilon_{2jk}\}$.

We next have a look at the product $\vec{e}_2 \times \vec{e}_1$ by considering a basis which has been rotated through $90°$ around $\vec{e}_3$ so that

$$\vec{e}_{\bar{1}} = \vec{e}_2 \qquad \vec{e}_{\bar{2}} = -\vec{e}_1 \qquad \vec{e}_{\bar{3}} = \vec{e}_3 \tag{31}$$

Now

$$\vec{e_2} \times \vec{e_1} = (-\vec{e_1}) \times \vec{e_2} = -(\vec{e_1} \times \vec{e_2}) = -\vec{e_3} = -\vec{e_3} \tag{32}$$

where a crucial step is to assume that the cross product is bilinear. But of course the two bases are equivalent so

$$\vec{e_2} \times \vec{e_1} = -\vec{e_3} \tag{33}$$

and

$$\epsilon_{211} = 0 \qquad \epsilon_{212} = 0 \qquad \epsilon_{213} = -1 \tag{34}$$

(Note that this means the cross product is not commutative, since $\vec{e_2} \times \vec{e_1} = -\vec{e_1} \times \vec{e_2}$.) We can perform the same permutation as before to obtain

$$\epsilon_{133} = 0 \qquad \epsilon_{131} = 0 \qquad \epsilon_{132} = -1 \tag{35}$$

Finally, we ask what happens when we take the vector product of a unit vector with itself. By definition

$$\vec{e_1} \times \vec{e_1} = \epsilon_{111}\vec{e_1} + \epsilon_{112}\vec{e_2} + \epsilon_{113}\vec{e_3} \tag{36}$$

This splits into a projection $\epsilon_{111}\vec{e_1}$ parallel to $\vec{e_1}$ and a projection $\epsilon_{112}\vec{e_2} + \epsilon_{113}\vec{e_3}$ perpendicular to $\vec{e_1}$. Consider first the piece perpendicular to $\vec{e_1}$; if we rotate the basis 180° about $\vec{e_1}$, we get

$$\vec{e_{\bar{1}}} = \vec{e_1} \qquad \vec{e_{\bar{2}}} = -\vec{e_2} \qquad \vec{e_{\bar{3}}} = -\vec{e_3} \tag{37}$$

Working in this basis,

$$\vec{e_1} \times \vec{e_1} = \vec{e_{\bar{1}}} \times \vec{e_{\bar{1}}} = \epsilon_{111}\vec{e_{\bar{1}}} + \epsilon_{112}\vec{e_{\bar{2}}} + \epsilon_{113}\vec{e_{\bar{3}}} = \epsilon_{111}\vec{e_1} - \epsilon_{112}\vec{e_2} - \epsilon_{113}\vec{e_3} \tag{38}$$

but comparing this to (36), we can read off

$$\epsilon_{112} = -\epsilon_{112} \qquad \epsilon_{113} = -\epsilon_{113} \tag{39}$$

which means

$$\epsilon_{112} = 0 = \epsilon_{113} \tag{40}$$

leaving only the parallel term

$$\vec{e_1} \times \vec{e_1} = \epsilon_{111}\vec{e_1} \tag{41}$$

But now consider a different rotation, 180° about $\vec{e_3}$, so that

$$\vec{e_{\bar{1}}} = -\vec{e_1} \qquad \vec{e_{\bar{2}}} = -\vec{e_2} \qquad \vec{e_{\bar{3}}} = \vec{e_3} \tag{42}$$

Now

$$\vec{e_{\bar{1}}} \times \vec{e_{\bar{1}}} = \epsilon_{111}\vec{e_{\bar{1}}} = \epsilon_{111}(-\vec{e_1}) = -\epsilon_{111}\vec{e_1} = (-\vec{e_1}) \times (-\vec{e_1}) = \vec{e_1} \times \vec{e_1} = \epsilon_{111}\vec{e_1} \tag{43}$$

which means that also $\epsilon_{111} = 0$ so $\vec{e_1} \times \vec{e_1} = 0$

Definition of the Vector Product Putting all of the properties together, we find

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \tag{44a}$$

$$\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \tag{44b}$$

$$\epsilon_{ijk} = 0 \qquad \text{for all other } i,\, j,\, k \tag{44c}$$

This ϵ_{ijk} is called the Levi-Civita symbol or the alternating symbol. Many more properties can be found in the textbook.

Geometrical Interpretation of the Vector Product Now we can associate a geometrical meaning to the cross product $\vec{A} \times \vec{B}$, looking again in the special basis where $\vec{A} = A\vec{e}_1$ and $\vec{B} = B\cos\theta\vec{e}_1 + B\sin\theta\vec{e}_2$ (where θ is once again the angle between the two vectors, now measured counter-clockwise from $\vec{A}$ to $\vec{B}$). The cross products of the unit vectors tell us that

$$\vec{A} \times \vec{B} = AB\sin\theta\vec{e}_3 \tag{45}$$

But geometrically, $AB\sin\theta$ is the area of the parallelogram spanned by $\vec{A}$ and $\vec{B}$, while $\vec{e}_3$ is the unit vector perpendicular to the plane in which it lies.

4   Vector Fields

A vector field consists of a vector at every point in space; examples include the gravitational, electric, or magnetic field, or the velocity field of a fluid flow. To resolve the field in components, we need to have a full set of basis vectors at every point in space.

4.1   Relationship Between Cartesian Coördinate Systems and Orthonormal Bases

Given a Cartesian coördinate system, there is a natural orthonormal basis to choose. If the coördinates are x, y, and z, it's convenient to call the basis vectors $\vec{e}_x$, $\vec{e}_y$, and $\vec{e}_z$. The vector $\vec{e}_i$ "points in the x_i direction", which means in the direction where x_i is increasing and the other two coördinates remain constant. Note that there are other names used for these basis vectors, such as $(\hat{x}, \hat{y}, \hat{z})$ or $(\hat{\imath}, \hat{\jmath}, \hat{k})$, or even $(\vec{\imath}, \vec{\jmath}, \vec{k})$. (The book calls them $(\mathbf{i}, \mathbf{j}, \mathbf{k})$.)

Now, the most tempting way to formalize this correspondence is to introduce a "position vector"

$$\vec{x} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z \tag{46}$$

(referred to in the book as $\vec{r}$ or actually $\mathbf{r}$). After all, if we rotate our coördinates using a rotation matrix $R_{\overline{k}i}$, the transformation is

$$x_{\overline{k}} = R_{\overline{k}i}x_i \tag{47}$$

which is just the way a vector transforms under a rotation. Unfortunately, (47) is not the only transformation which takes us from a Cartesian coördinate system to another Cartesian

coördinate system. We can also *translate* the origin by a constant distance (X_1, X_2, and X_3, respectively) in each direction, and get another perfectly good Cartesian coördinate system:

$$x_{\overline{k}} = x_i - X_i \tag{48}$$

However, this new coördinate system still has its axes pointed in the same direction, i.e.,

$$\vec{e}_{\overline{x}} = \vec{e}_x \tag{49a}$$
$$\vec{e}_{\overline{y}} = \vec{e}_y \tag{49b}$$
$$\vec{e}_{\overline{z}} = \vec{e}_z \tag{49c}$$

which would mean that, if $\vec{x}$ were a true vector, its components in the new basis would be the same as in the old one.

The most general transformation from one right-handed Cartesian coördinate system to another is

$$x_{\overline{k}} = R_{\overline{k}i}x_i - X_{\overline{k}} \tag{50}$$

where $\{R_{\overline{k}i}\}$ are the components of a constant rotation matrix (orthogonal and with unit determinant) and $\{X_{\overline{k}}\}$ are three constant lengths. The change in basis under this transformation is

$$\vec{e}_{\overline{k}} = R_{\overline{k}i}\vec{e}_i \tag{51}$$

and the transformation of the components of a vector $\vec{A}$ is

$$A_{\overline{k}} = R_{\overline{k}i}A_i \tag{52}$$

It is often convenient to pretend that there is such a thing as a "position vector $\vec{x}$", and it usually doesn't get us into trouble, since the difference between the position vectors describing two points does act like a vector (at least when we confine attention to Cartesian coördinate systems). Explicitly, if we write

$$\vec{x}_{PQ} = \vec{x}_Q - \vec{x}_P = (x_Q - x_P)\vec{e}_x + (y_Q - y_P)\vec{e}_y + (z_Q - z_P)\vec{e}_z \tag{53}$$

then the transformation properties of $\vec{x}_{PQ}$ can be derived from the coordinate transformations (50):

$$x_{PQ\overline{k}} = x_{Q\overline{k}} - x_{P\overline{k}} = \left(R_{\overline{k}i}x_{Qi} - X_{\overline{k}}\right) - \left(R_{\overline{k}i}x_{Pi} - X_{\overline{k}}\right) = R_{\overline{k}i}\left(x_{Qi} - x_{Pi}\right) = R_{\overline{k}i}x_{PQi} \tag{54}$$

Since differentiation is just a limiting case of subtraction, this means that given a particle trajectory $\vec{x}(t)$, we can define a velocity vector

$$\dot{\vec{x}}(t) = \frac{d\vec{x}}{dt} = \lim_{\delta t \to 0} \frac{\vec{x}(t + \delta t) - \vec{x}(t)}{\delta t} \tag{55}$$

which is a *bona fide*, honest-to-goodness, full-fledged vector.

4.2 Non-Cartesian Bases

The orthonormal bases associated with Cartesian coördinates have the advantage that the basis vectors are the same at all points in space. However, if the geometry of a problem makes it more natural to use non-Cartesian cördinates such as spherical coördinates (r, θ, ϕ) defined by

$$x = r \sin\theta \cos\phi \tag{56a}$$
$$y = r \sin\theta \sin\phi \tag{56b}$$
$$z = r \cos\theta \tag{56c}$$

or cylindrical coördinates (ρ, ϕ, z) defined by

$$x = \rho \cos\phi \tag{57a}$$
$$y = \rho \sin\phi \tag{57b}$$
$$z = z \tag{57c}$$

or, in two dimensions, plane polar coördinates (r, ϕ) defined by

$$x = r \cos\phi \tag{58a}$$
$$y = r \sin\phi \tag{58b}$$

it's natural to define an orthonormal basis at each point whose basis vectors point in the directions in which our chosen coördinates are locally increasing. Non-Cartesian coördinate systems can be studied in all generality, which is not only an elegant process, but also essential for the study of General Relativity. But within the scope of this course, we'll take the same approach as Marion and Thornton and just look at a few special cases.

So for plane polar coördinates, we find geometrically that

$$\vec{e}_r = \cos\phi \vec{e}_x + \sin\phi \vec{e}_y \tag{59a}$$
$$\vec{e}_\phi = -\sin\phi \vec{e}_x + \cos\phi \vec{e}_y \tag{59b}$$

writing this in matrix form

$$\begin{pmatrix} \vec{e}_r \\ \vec{e}_\phi \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \vec{e}_x \\ \vec{e}_y \end{pmatrix} \tag{60}$$

makes it clear that the inverse transformation is

$$\begin{pmatrix} \vec{e}_x \\ \vec{e}_y \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \vec{e}_r \\ \vec{e}_\phi \end{pmatrix} \tag{61}$$

Now, although the vectors $\vec{e}_x$ and $\vec{e}_y$ are constant, the vectors $\vec{e}_r$ and $\vec{e}_\phi$ change from point to point. Specifically, differentiating (60) tells us

$$\begin{pmatrix} d\vec{e}_r \\ d\vec{e}_\phi \end{pmatrix} = \begin{pmatrix} -\sin\phi \, d\phi & \cos\phi \, d\phi \\ -\cos\phi \, d\phi & -\sin\phi \, d\phi \end{pmatrix} \begin{pmatrix} \vec{e}_x \\ \vec{e}_y \end{pmatrix} = \begin{pmatrix} -\sin\phi \, d\phi & \cos\phi \, d\phi \\ -\cos\phi \, d\phi & -\sin\phi \, d\phi \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \vec{e}_r \\ \vec{e}_\phi \end{pmatrix}$$
$$= \begin{pmatrix} 0 & d\phi \\ -d\phi & 0 \end{pmatrix} \begin{pmatrix} \vec{e}_r \\ \vec{e}_\phi \end{pmatrix} \tag{62}$$

or in other words

$$d\vec{e}_r = \vec{e}_\phi d\phi \tag{63a}$$

$$d\vec{e}_\phi = -\vec{e}_r d\phi \tag{63b}$$

4.3 Vector Calculus

In addition to an infinitesimal displacement $d\vec{x}$ being a vector, derivatives with respect to coördinates behave in a convenient way thanks to the relation

$$\frac{\partial x_{\overline{k}}}{\partial x_i} = \Lambda_{\overline{k}i} \tag{64}$$

That means that if we take the derivative of a scalar field φ with respect to the coördinate x_i, the chain rule tells us

$$\frac{\partial \varphi}{\partial x_i} = \sum_{\overline{k}=1}^{3} \frac{\partial x_{\overline{k}}}{\partial x_i} \partial\varphi \partial x_{\overline{k}} = \sum_{\overline{k}=1}^{3} \Lambda_{\overline{k}i} \frac{\partial \varphi}{\partial x_{\overline{k}}} \tag{65}$$

But since $\boldsymbol{\Lambda}$ is an orthogonal matrix, $\Lambda_{\overline{k}i} = \Lambda_{i\overline{k}}^{-1}$, so

$$\frac{\partial \varphi}{\partial x_i} = \sum_{\overline{k}=1}^{3} \Lambda_{i\overline{k}}^{-1} \partial\varphi \partial x_{\overline{k}} \tag{66}$$

or

$$\frac{\partial \varphi}{\partial x_{\overline{k}}} = \sum_{i=1}^{3} \Lambda_{\overline{k}i} \partial\varphi \partial x_i \tag{67}$$

but this is just the transformation law for a vector. This vector is called the *gradient* of φ

$$\vec{\nabla}\varphi = \sum_{i=1}^{3} \frac{\partial \varphi}{\partial x_i} \vec{e}_i \tag{68}$$

Note that this demonstration only works in Cartesian coördinate systems, and the gradient is only defined by (68) in Cartesian coördinates. Most of a course could be spent on how to do vector calculus in an arbitrary coördinate system, but for our purposes, the definition of the gradient can be extended to non-Cartesian coördinate systems by using (68) as the starting point and transforming into the non-Cartesian system. The specific forms in the most common coördinate systems are given in Appendix F of Marion & Thornton.

There are two other interesting vector derivatives, which basically amount to taking the dot and cross products of the gradient operator with a vector field $\vec{A}$. They are the *curl*, defined by

$$\vec{\nabla} \times \vec{A} = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon_{ijk} \frac{\partial A_j}{\partial x_i} \vec{e}_k \tag{69}$$

and the *divergence*, defined by

$$\vec{\nabla} \cdot \vec{A} = \sum_{i=1}^{3} \frac{\partial A_i}{\partial x_i} \tag{70}$$

the demonstrations that these are a vector and a scalar, respectively, rely on the fact that the transformation matrix elements $\Lambda_{\bar{k}i}$ are constants.

4.3.1 Vector Integrals

We can perform single, double, or triple integrals involving vectors. They are summarized in this table:

#Dim	Type of Integral	Form
1	line integral	$\int_{\mathcal{C}} \vec{A} \cdot d\vec{\ell} = \int_{s_P}^{s_Q} \left(\vec{A} \cdot \frac{d\vec{\ell}}{ds} \right) ds$
2	surface integral	$\iint_{S} \vec{A} \cdot d^2\vec{a} = \iint_{S} (\vec{A} \cdot \vec{n}) \, d^2 a$
3	volume integral	$\iiint_{V} \varphi \, d^3 v$

Integration by Parts   Recall the second fundamental theorem of calculus, which says that if you integrate the derivative of a function, the result is given by the difference between the values of the function at the boundaries:

$$\int_a^b F'(x) \, dx = F(b) - F(a) \tag{71}$$

Each of the integrals in the table above satisfies an identity which can ultimately be derived from this. In each case, one replaces the n-dimensional integral of a vector derivative with an $n - 1$-dimensional integral, without the derivative, over the boundary. Basically the derivative allows you to get rid of one of the integrals.

Explicitly, consider the line integral of a gradient:

$$\int_{\mathcal{C}} \vec{\nabla}\varphi \cdot d\vec{\ell} \tag{72}$$

where $\mathcal{C}$ is a curve with endpoints P and Q. We can *parametrize* this curve with functions $x(s), y(s), z(s)$, or abusing the vector notation slightly, $\vec{x}(s)$. The parameter ranges from s_P to s_Q. The line integral can then be written as

$$\int_{\mathcal{C}} \vec{\nabla}\varphi \cdot d\vec{\ell} = \int_{s_P}^{s_Q} \left((\vec{\nabla}\varphi) \cdot \frac{d\vec{x}}{ds} \right) ds = \int_{s_P}^{s_Q} \sum_{i=1}^{3} \frac{\partial \varphi}{\partial x_i} \frac{dx_i}{ds} ds = \int_{s_P}^{s_Q} \frac{d\varphi(\vec{x}(s))}{ds} ds$$
$$= \varphi(\vec{x}(s_Q)) - \varphi(\vec{x}(s_P)) = \varphi(Q) - \varphi(P) \tag{73}$$

Where we have used the chain rule as well as the second fundamental theorem of calculus.

The pair of points Q and P can be thought of as the zero-dimensional boundary of the one-dimensional curve $\mathcal{C}$. The standard notation is to use ∂ to indicate the $n-1$ dimensional boundary of an n-dimensional object. So we refer to these two points as $\partial\mathcal{C}$.

Note that no matter what curve we use, as it begins at P and ends at Q (i.e., has boundary $\partial\mathcal{C}$), if we integrate a gradient along it, we get the same answer.

The higher-dimensional versions of this have somewhat more involved derivations, but all come down to the second fundamental theorem of calculus, and can all be summarized as "the integral of the appropriate vector derivative over an n-dimensional region is equal to an integral, without the derivative, over the $n-1$-dimensional boundary". The identities are summarized below:

#Dim	Name	Identity
$1 \to 0$		$\int_{\mathcal{C}}(\vec{\nabla}\varphi)\cdot d\vec{\ell} = \varphi(Q) - \varphi(P)$
$2 \to 1$	Stokes's Theorem	$\iint_{S}(\vec{\nabla}\times\vec{A})\cdot d^2\vec{a} = \oint_{\partial S}\vec{A}\cdot d\vec{\ell}$
$3 \to 2$	Gauss's Theorem (or Divergence Theorem)	$\iiint_{V}\vec{\nabla}\cdot\vec{A}\,d^3v = \oiint_{\partial V}\vec{A}\cdot d^2\vec{a}$

Note that in the case of Stokes's theorem, as with the integral of a gradient, it doesn't matter over which surface you integrate a curl, as long as it has the same boundary.

Also, note that the boundary of an n-dimensional region is in each case a closed $n-1$-dimensional region, i.e., it does not have a ($n-2$-dimensional) boundary of its own. Specifically, consider the boundary ∂V of a volume V. It is a closed surface, and has no boundary $\partial(\partial V)$. The same thing is true of the curve ∂S which forms the boundary of the surface S. It has no boundary $\partial(\partial S)$. The deeper underlying identity is sometimes stated as "the boundary of a boundary is zero": $\partial^2 = 0$.

A   Appendix: Correspondence to class lectures

Date	Sections	Pages
2002 August 26	1–2	1–2
2002 August 28	3–3.1.2	2–3
2002 August 30	3.1.2–3.1.2	3–4
2002 September 4	3.2–3.2.3	4–8
2002 September 6	4–4.1	8–9
2002 September 9	Review	
2002 September 11	Review	
2002 September 13	4.1–4.2	9–10
2002 September 16	4.2–4.2	10–11
2002 September 18	4.3–4.3.1	11–12
2002 September 20	4.3.1–4.3.1	12–13