Physics 464/511   Lecture B   Fall 2016

1   Fields

In elementary Physics courses many problems involve a point particle, and one is asked to solve, say, for the motion of the particle in response to a force. There is then only one point, with a few degrees of freedom. Most physical problems, however, involve not the classical motion of a point particle but the motion of a medium spread throughout some spatial region. That is, the properties we need describe some physical quantity at each point in some space.

As an example, consider the position of a drumhead. To a good approximation, one can ignore the motion of the skin in the plane of the face of the drum. Thus the position of the skin is given by a function $z(x, y)$, and solving for the motion involves a field equation not for a variable but for a function $z(x, y)$ as a function also of t. That, of course, will involve partial differential equations.

A set of physics degrees of freedom describing properties individually at each point in space at a given time is called a *field*. The height of the drumhead, $z(x, y, t)$ is a field in two spatial dimensions (and one time).

Other fields are usually defined on a three dimensional region. For example, the pressure within the air inside the body of a violin $p(x, y, z, t)$ is a physical pressure associated with each point within a 3-D volume. Pressure is called a scalar field because the pressure itself has no direction associated with it, only a magnitude. Of course, in a given situation there may be a direction associated with, say, the variation of pressure, which is to say that the gradient of pressure is not a scalar field, but instead is a vector field.

There are other fundamental fields which are intrinsically vectors. One example with which you are very familiar is the electric field $\vec{E}(x, y, z, t)$. At each point in space, at a given time, there is an electric field which has both magnitude and direction.

Before we continue with fields, we should step back and clarify two things — what a vector space is, and what position and a spatial region are.

1.1   Vector space

Many objects in physics live in a *vector space*. Physics students usually meet the concept of a vector first in terms of position in three dimensional space. Position not really a good example for the reason given in the third bullet below, but displacements in ordinary three dimensional Euclidean space are vectors. We realize that we need the concepts of

- addition of two vectors: successive displacements produce another displacement. This means the operation of addition is defined and the vector space is *closed under addition*.

- associativity of addition: $\vec{v}_1 + (\vec{v}_2 + \vec{v}_3) = (\vec{v}_1 + \vec{v}_2) + \vec{v}_3$.

  In fact, with identity as the zero displacement and additive inverse $-\vec{V}$, displacements form a group under addition.

- multiplication by a scalar: For displacements, this means multiplication by a real number, like moving twice as far in the same direction. Note this is also why position is not really a vector — 42nd street is not three times 14th street. More generally, a vector space is defined "over a field".

Here we are using the word "field" in the mathematician's sense, not the physics definition. As we mentioned last time, what a mathematician means by a field F is a set of elements together with two binary operators (that is, functions of two elements $F \times F \to F$) under which the set is closed. The two operators are called *addition* and *multiplication*, written $x + y$ and xy. *Closure* means that for any $x \in F$ and $y \in F$, $x + y \in F$ and $xy \in F$. Both the addition and multiplication are required to be *commutative* ($x + y = y + x$ and $xy = yx$) and *associative* $(x + y) + z = x + (y + z)$ and $(xy)z = x(yz)$, and multiplication must be *distributive* over addition: $x(y + z) = xy + xz$. There are also identity elements 0 and 1 for addition and multiplication respectively, ($0 + x = x$, $1x = x$ for all $x \in F$), and every element x has an additive inverse (called $-x$, with $-x + x = 0$), and every element x except 0 has a multiplicative inverse x^{-1} with $x^{-1}x = 1$.

These formal properties which define a field are satisfied by the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, and the rational numbers $\mathbb{Q}$. There are also fields of characterist p where p is a prime number, basically arithmetic mod p. We will almost always take our field to be either the reals or the complex numbers.

Getting back to a vector space:

A *vector space* V over a field F is a set of elements $\vec{u}, \vec{v}, \dots$ together with a commutative associative binary operator called *addition* and a *scalar multiplication* $F \times V \to V$ with, for all $a \in F, b \in F, \vec{u} \in V, \vec{v} \in V$,

- $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$

- $(a + b)\vec{u} = a\vec{u} + b\vec{u}$

- $a(b\vec{u}) = (ab)\vec{u}$

- $1\,\vec{v} = \vec{v}$

These are only a formal statement of properties with which we are very familiar. Displacements add without regard to order or grouping, and adding vectors twice as big gives a result twice as big, *etc.*.

Note these formal requirements are stated without any mention of components or basis vectors, which we introduce only to have ways of making concrete statements about particular vectors or transformations.

If we think about displacements in ordinary space, we usually describe a vector in terms of three basis vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$, or $\hat{e}_1$, $\hat{e}_2$, $\hat{e}_3$, multiplied by three real coefficients (coming from $\mathbb{R}$), v_x, v_y, v_z, or v_1, v_2, v_3. The use of x, y, z does not generalize well to vector spaces in other than three dimensions, which we will want to discuss, so I prefer to use subscripts 1, 2, $\dots D$ for a D dimensional space. It is important to realize, however, that the vector itself is the fundamental object, the same regardless of the basis vectors (and components) used to describe it.

This formal acknowledgment that an object has an identity beyond its description in terms of components applies even more to ordinary space, which we usually take to be three-dimensional Euclidean space. A position in space is an object which we usually describe in terms of some coordinate system involving a choice of an origin and a choice of Cartesian basis vectors, but the position is the same position even when described by different choices of coordinates. In particular, positions are not vectors, because there is no intrinsic meaning to multiplication by a scalar before one chooses an origin, and in fact one cannot even add two positions, as one can two displacements. This may seem like nit-picking, but when we start to consider the possibility that space is not a flat Euclidean one, these distinctions will become crucial.

In describing vector fields, the way we picture
them may be arrows in space, but our equations
usually have to deal with them in terms of com-
ponents. Note, however, that a vector itself does
not have a position — two arrows of the same
length and parallel are the same vector regardless
of where they are drawn. However when we draw
a vector *field*, we expect the arrows are drawn at
the position the vector is defined, $\vec{V}(\mathbf{r})$ is drawn
at the point $\mathbf{r}$. Though we should probably distinguish the position $\mathbf{r}$ from
the displacement $\vec{r}$ from the arbitrary origin, we generally don't. So fields
will mostly be written as $\Phi(\vec{r}, t)$ or $\vec{V}(\vec{r}, t)$.
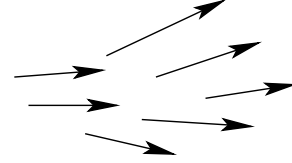
We may write a vector field $\vec{E}(\vec{r}, t)$ in terms of cartesian coordinates
$\vec{E} = E_x \hat{\imath} + E_y \hat{\jmath} + E_z \hat{k}$, where $\hat{\imath}, \hat{\jmath}$ and $\hat{k}$ are standard names for unit vectors
pointing in the x, y and z directions. The component fields, *e.g.* $E_x(\vec{r}, t)$ are
multiplied by unit vectors (*e.g.* $\hat{\imath}$) as defined at the same physical point.

In the case of cartesian coordinates, $\hat{\imath}$ is a fixed vector, the same regardless
of $\vec{r}$. This is because two vectors of the same length are the same if they are
parallel. One may also express the fields in terms of other unit vectors. In
cylindrical coordinates, for example,

$$\vec{E}(\vec{r}) = E_r(\vec{r})\, \hat{e}_r + E_\theta(\vec{r})\, \hat{e}_\theta + E_z(\vec{r})\, \hat{e}_z,$$

where $\hat{e}_r$ is a unit vector at the point $\vec{r}$ pointing
away from the origin. The $\hat{e}_r$ at one point is not
equal to the $\hat{e}_r$ at another (in general), so we really
mean
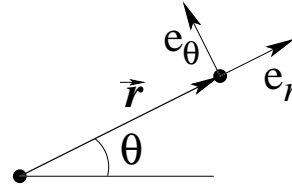
$$\vec{E}(\vec{r}) = E_r(\vec{r})\, \hat{e}_r(\vec{r}) + E_\theta(\vec{r})\, \hat{e}_\theta(\vec{r}) + E_z(\vec{r})\, \hat{e}_z(\vec{r}).$$

Tensor Product, Direct Sum and Direct Product

Before proceeding to differential operators, I need to define two more things
one can make from vectors. One is the *tensor product.* If we have two vector
spaces U and V, with basis vectors u_i and v_j, the tensor product $U \otimes V$ is
defined as all linear combinations of basis vectors $\Phi_{ij} := u_i \otimes v_j$. This should
be familiar from quantum mechanics, where a quantum state of an electron
with a given orbital angular momentum ℓ is described by $|\ell, m; s\rangle = |\ell m\rangle |s\rangle$,
the first ket a vector describing ordinary angular momentum and the other,
with $s = \pm 1/2$ the state of the spin. The tensor product differs from the

Cartesian product $U \times V$ in two important ways. Recall $U \times V$ consists of pairs of vectors, one from U and one from V, so $(2u, v)$ is distinct from $(u, 2v)$, while for the tensor product they are the same. Secondly, the tensor product is a vector space and consists of all linear combinations, including ones which are not simply one vector from U and one from V.

Our quantum state of an electron brings up another concept, that of *direct sum* or *direct product*, which are almost the same, written as $U \oplus V$, but are different from the tensor product. For two vector spaces, $U \oplus V$ is just the sum of vectors in U and vectors in V, not in the same space but, if the spaces are unitary vector spaces, orthogonal to each other. Thus the basis just consists of the basis of U together and independent of the basis of V. If U and V have dimensions D_U and D_V respectively, $D_{U \otimes V} = D_U D_V$ and $D_{U \oplus V} = D_U + D_V$. If we ask about the state of an electron, ignoring spin, in general it will be a sum $\psi = \sum_{n \ell m} a_{n \ell m} R_n(r) Y_m^\ell(\theta, \phi)$, which includes a direct sum of the spaces with different ℓ. We will discuss $R_n(r)$ and $Y_m^\ell(\theta, \phi)$ later.

Of course we recognize that neither $U \otimes V$ nor $U \oplus V$ have anything to do with the scalar or vector product of vectors, $\vec{u} \times \vec{v}$ or $\vec{u} \cdot \vec{v}$.

1.2   The gradient

Suppose we have a scalar field $\lambda(\vec{r})$. Using Cartesian coordinates, the gradient of λ is a vector field

$$\vec{\nabla} \lambda(\vec{r}) = \hat{\imath}\, \frac{\partial \lambda}{\partial x} + \hat{\jmath}\, \frac{\partial \lambda}{\partial y} + \hat{k}\, \frac{\partial \lambda}{\partial z}.$$

The use of a gradient is to dot it into a small change in position $d\vec{r} = \hat{\imath}\, dx + \hat{\jmath}\, dy + \hat{k}\, dz$. Then

$$d\lambda = \vec{\nabla} \lambda \cdot d\vec{r} = \frac{\partial \lambda}{\partial x}\, dx + \frac{\partial \lambda}{\partial y}\, dy + \frac{\partial \lambda}{\partial z}\, dz,$$

where $d\lambda$ means the change in the **value** of λ as one moves from $\vec{r}$ to $\vec{r} + d\vec{r}$. $d\lambda$ is called a differential, rather than a numerical quantity (*e.g.* a real number times a unit) because it depends on the unspecified change in position $d\vec{r}$. More highfalutinly, this is a differential 1-form, but more on that later.

If $\vec{\nabla} \lambda \cdot d\vec{r} = 0$ then $d\lambda = 0$, so a small move in the direction $d\vec{r}$ produces no change in λ if $\vec{\nabla} \lambda \perp d\vec{r}$. Thus $\vec{\nabla} \lambda$ points perpendicular to a surface of constant λ.

Notice that the components $\left(\vec{\nabla}\lambda\right)_x$ corresponds to the change in λ when I move one unit of distance in the x direction. (Well, only if the unit is infinitesimal, ignoring second derivative terms.)

The simple correspondence between the components of $\vec{\nabla}\lambda$ and partial derivatives only holds for cartesian coordinates. A field $\lambda(x, y, z)$ can also be described by a function of the cylindrical polar coordinates

$$
\begin{aligned}
r &= \sqrt{x^2 + y^2} \\
\theta &= \tan^{-1} y/x, \qquad \text{with} \;\; x = r\cos\theta, \quad y = r\sin\theta
\end{aligned}
$$

and $\quad z$

When we write $\lambda(r, \theta, z)$ we mean a different function of three variables than the function $\lambda(x, y, z)$ of its three variables, but the same function of position. We should probably give them different names $\lambda_{\mathrm{P}}(r, \theta, z)$ and $\lambda_{\mathrm{C}}(x, y, z)$, where

$$
\lambda_{\mathrm{C}}(x, y, z) = \lambda_{\mathrm{P}}(r, \theta, z) \quad \text{if} \quad x = r\cos\theta, \; y = r\sin\theta.
$$

But we don't do that, because we think of λ as really being a function from *Points in space* rather then on a combination of three meaningless continuous parameters. So we think of λ_{C} and λ_{P} as the same field, even though as functions of three real variables they are different.

We also want to think of $\vec{\nabla}\lambda$ as a vector field which has a reality independent of the parameters used to describe space-time. If, however, we are using polar coordinates, we would like to express that vector in terms of its coordinates along radial and tangential directions $\hat{e}_r$ and $\hat{e}_\theta$, as well as $\hat{e}_z$. Here $\hat{e}_r$ is a unit vector $\parallel \vec{r}$, so we see $\hat{e}_r = \hat{\imath}\cos\theta + \hat{\jmath}\sin\theta$, and $\hat{e}_\theta$ is perpendicular to it, $\hat{e}_\theta = -\hat{\imath}\sin\theta + \hat{\jmath}\cos\theta$.
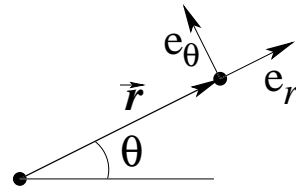
Now we can invert

$$
\begin{aligned}
\hat{\imath} &= \hat{e}_r \cos\theta - \hat{e}_\theta \sin\theta \\
\hat{\jmath} &= \hat{e}_r \sin\theta + \hat{e}_\theta \cos\theta
\end{aligned}
$$

and make use of the chain rule

$$
\begin{aligned}
\left.\frac{\partial\lambda}{\partial x}\right|_{y,z} &= \left.\frac{\partial\lambda}{\partial r}\right|_{\theta,z}\left.\frac{\partial r}{\partial x}\right|_{y,z} + \left.\frac{\partial\lambda}{\partial\theta}\right|_{r,z}\left.\frac{\partial\theta}{\partial x}\right|_{y,z} \\
&= \frac{\partial\lambda}{\partial r}\cos\theta + \frac{\partial\lambda}{\partial\theta}\left(-\frac{1}{r}\sin\theta\right).
\end{aligned}
$$

Here I have evaluated $\partial r/\partial x$ and $\partial\theta/\partial x$ and expressed them as functions of the polar coordinates. The first follows from $r^2 = x^2 + y^2$ so

$$
r\,dr = x\,dx + y\,dy, \quad \text{or} \quad \left.\frac{\partial r}{\partial x}\right|_{dy=0} = \frac{x}{r} = \cos\theta,
$$

and the second from

$$
\tan\theta = \frac{y}{x} \implies \sec^2\theta\,d\theta = -\frac{y}{x^2}\,dx \text{ if } dy=0, \quad \text{so} \quad \left.\frac{\partial\theta}{\partial x}\right|_y = -\frac{1}{r}\sin\theta.
$$

Similarly $\left.\dfrac{\partial\lambda}{\partial y}\right|_{x,z} = \dfrac{\partial\lambda}{\partial r}\sin\theta + \dfrac{\partial\lambda}{\partial\theta}\dfrac{\cos\theta}{r}$. Now

$$
\begin{aligned}
\vec{\nabla}\lambda &= \hat{i}\frac{\partial\lambda}{\partial x} + \hat{j}\frac{\partial\lambda}{\partial y} + \hat{k}\frac{\partial\lambda}{\partial z} \\
&= (\hat{e}_r\cos\theta - \hat{e}_\theta\sin\theta)\left(\cos\theta\frac{\partial\lambda}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial\lambda}{\partial\theta}\right) \\
&\quad + (\hat{e}_r\sin\theta + \hat{e}_\theta\cos\theta)\left(\sin\theta\frac{\partial\lambda}{\partial r} + \frac{1}{r}\cos\theta\frac{\partial\lambda}{\partial\theta}\right) + \hat{k}\frac{\partial\lambda}{\partial z} \\
&= \hat{e}_r\frac{\partial\lambda}{\partial r} + \hat{e}_\theta\frac{1}{r}\frac{\partial\lambda}{\partial\theta} + \hat{e}_z\frac{\partial\lambda}{\partial z}.
\end{aligned}
$$

Note that the $\hat{e}_\theta$ term has a coefficient $\dfrac{1}{r}\dfrac{\partial\lambda}{\partial\theta}$, not $\dfrac{\partial\lambda}{\partial\theta}$.

The gradient gives the rate at which a function changes. More particularly, if $\vec{\beta}$ is a unit vector, $\vec{\beta}\cdot\vec{\nabla}\lambda$ gives the rate of change of λ per unit distance as we move in the direction $\vec{\beta}$. If $\vec{\beta}$ points in the tangential direction, the distance moved is $r\,d\theta$, so the rate at which λ changes is $\vec{\beta}\cdot\vec{\nabla}\lambda = \frac{1}{r}\frac{\partial\lambda}{\partial\theta}$, in agreement with the discussion above.
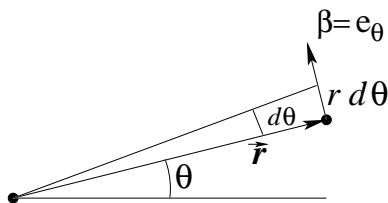
We will return to other "curvilinear coordinate" systems after we have introduced the other necessary calculus. Later we will get more general, in the language of differential geometry.

1.3   Derivatives of Vector Fields

We have seen that the derivative of a scalar field λ is a vector field, with the direction coming from the fact that, to specify how the field varies, one must specify the direction in space that one is moving.

Now consider a vector field $\vec{V}(\vec{r})$. This might be the gradient of a scalar field or it might be some other kind of vector field such as the magnetic field $\vec{B}$. If we ask for the derivative of $\vec{V}$ there are two directions which come into play: the direction of differentiation and the direction of ΔV. Thus the derivative is a tensor object, with components that have two indices to specify these two directionalities.

One particularly important piece of this derivative is the divergence

$$\vec{\nabla} \cdot \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z},$$

which we get from identifying the direction of differentiation with the component in that direction and adding all directions. That is, the trace of the tensor. The resulting field has no directionality left — it is a scalar.

A very important example: Suppose we have a fluid of possibly varying density $\rho(\vec{r}, t)$ with a velocity field $\vec{v}(\vec{r}, t)$ which gives the velocity of the fluid at each point. Consider a small volume of the fluid. The rate at which fluid is leaving the box through the face BDHF is $\rho v_y dx dz$, where ρv_y has to be evaluated over the BDHF face, or more precisely averaged over this face. This will be mostly compensated by a flow into the box through face ACGE, also given by $\rho v_y dx dz$, but where ρv_y is evaluated on ACGE. Now the only difference is that

ρv_y on BDHF is evaluated at a y value dy greater than at ACGE, so
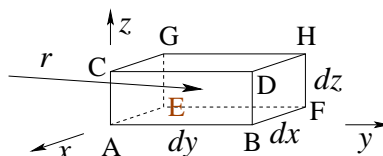
$$\rho v_y|_{\text{BDHF}} - \rho v_y|_{\text{ACGE}} = dy \frac{\partial}{\partial y} \left(\rho v_y \right),$$

and the net flow out of the box through these two faces is $dx\, dy\, dz\, \dfrac{\partial}{\partial y} (\rho v_y)$.

By the same argument for faces ABDC and EFHG, the net flow out is $dx\,dy\,dz\,\dfrac{\partial}{\partial x}\left(\rho v_x\right)$, and the contribution of ABFE and CDHG is $dx\,dy\,dz\,\dfrac{\partial}{\partial z}\left(\rho v_z\right)$, so all together the rate of flow out of the box, per unit volume, is $\vec{\nabla}\cdot\left(\rho\vec{V}\right)$, the divergence of the vector field $\rho\vec{V}$ formed by multiplying a scalar field by a vector field.

The net rate at which the fluid is flowing out of the box is clearly the sum of

- $g(\vec{r})\,dx\,dy\,dz$, the rate at which matter is being spontaneously generated (usually assumed to be zero!), and

- $-\dfrac{\partial\rho}{\partial t}\,dx\,dy\,dz$, the rate at which the total matter in the box is decreasing with time.

Thus

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\left(\rho\vec{V}\right) = g.$$

If ρ represents the density of all matter, and we generally believe matter is not generated, then $g = 0$. Then we say matter is *conserved*, and

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\left(\rho\vec{V}\right) = 0,$$

which is called the *continuity equation*. My g is not completely crazy, however. If ρ represents the density of carbon monoxide in a system which also contains carbon and oxygen, g might not be zero if there are chemical interactions occurring.

Note from the definition

$$\begin{aligned}
\vec{\nabla}\cdot(\rho\vec{v}) &= \frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z}\\[2mm]
&= \frac{\partial\rho}{\partial x}v_x + \frac{\partial\rho}{\partial y}v_y + \frac{\partial\rho}{\partial z}v_z + \rho\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right)\\[2mm]
&= \left(\vec{\nabla}\rho\right)\cdot\vec{v} + \rho\vec{\nabla}\cdot\vec{v},
\end{aligned}$$

a sort of product (or Leibnitz) rule.

Consider the electrostatic potential of a point charge $\phi = \dfrac{q}{4\pi\epsilon_0}\dfrac{1}{r}$. In fact, just to add generality, consider $\phi = kr^{2n}$, which reduces to electrostatics if $k = q/4\pi\epsilon_0$ and $n = -\frac{1}{2}$. Then

$$
\begin{aligned}
\vec{\nabla}\phi &= k\vec{\nabla}\left(r^2\right)^n = kn\left(r^2\right)^{n-1}\vec{\nabla}r^2 \\
&= knr^{2n-2}\left[\hat{\imath}\,2x + \hat{\jmath}\,2y + \hat{k}\,2z\right] \\
&= 2knr^{2n-2}\vec{r}.
\end{aligned}
$$

For electrostatics,

$$
\vec{E} = -\vec{\nabla}\phi = -\frac{q}{4\pi\epsilon_0}\left(2\cdot-\frac{1}{2}\right)\frac{\vec{r}}{r^3} = \frac{q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3}.
$$

Now consider $\vec{\nabla}\cdot\vec{E}$. From our product rule,

$$
\vec{\nabla}\cdot\vec{E} = \frac{q}{4\pi\epsilon_0}\left[\vec{\nabla}\left(r^{-3}\right)\right]\cdot\vec{r} + \frac{q}{4\pi\epsilon_0}r^{-3}\vec{\nabla}\cdot\vec{r}.
$$

$\vec{\nabla}\left(r^{-3}\right) = -3r^{-5}\vec{r}$ from our formula above with $n = -3/2$, while $\vec{\nabla}\cdot\vec{r} = \dfrac{\partial x}{\partial x} + \dfrac{\partial y}{\partial y} + \dfrac{\partial z}{\partial z} = 3$, so

$$
\vec{\nabla}\cdot\vec{E} = \frac{q}{4\pi\epsilon_0}\left(-3\frac{\vec{r}\cdot\vec{r}}{r^5} + r^{-3}\cdot3\right) = 0,
$$

except when $r = 0$ where everything is undefined. We will see later how to describe $\vec{\nabla}\cdot\vec{E}$ at $r = 0$.

This, of course, is what we expect from Maxwell's equation $\vec{\nabla}\cdot\epsilon_0\vec{E} = \rho$, as the charge density $= 0$ except at $r = 0$, where the point charge is.

A useful abstract notation

In 3-D space we are used to writing the variables x, y, z and the components of a vector as v_x, v_y, v_z. Not all mathematical spaces of interest are three-dimensional, and this notation does not generalize easily. We then use a notation with $\vec{r}$ expressed in terms of cartesian coordinates $r_1, , r_2, \ldots r_d$ (for a d dimensional space) and components of the vector as $v_1, v_2, \ldots v_d$, and unit

vectors $\hat{e}_1, \hat{e}_2 \ldots \hat{e}_d$. Then

$$\vec{\nabla}\lambda \;=\; \sum_{i=1}^{d} \hat{e}_i \frac{\partial \lambda}{\partial r_i}$$

$$\vec{\nabla} \cdot \vec{V} \;=\; \sum_{i=1}^{d} \frac{\partial}{\partial r_i} V_i.$$

While all the vectors we meet with early in our courses have their directions coming from space, that is not always the case. In quantum field theory the vector fields may point in directions in isotopic spin space, for example. A more concrete example is the light emission from a flat-panel display. At each position in two dimensions (x, y), we have an intensity I_r, I_g, I_b of light from the red, green and blue emitters respectively. Thus $\vec{I}$ is a vector in three dimensional color space, and the state of the screen is a vector field mapping two dimensional physical space into three dimensional color space, $\vec{I}(\vec{r}) = (I_r(x, y), I_g(x, y), I_b(x, y))$. Of course if the vector space is not the space on which the fields are defined, it makes no sense to examine the trace of the matrix of deriviatives, $\partial I_c / \partial r_i$.

1.4   The Cross Product

In 3D space only, we can define a peculiar product of two vectors, $\vec{C} = \vec{A} \times \vec{B}$, $C_x = A_y B_z - A_z B_y$, *etc.*, which can also be written in a formal fashion as the determinant of

$$\vec{C} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}.$$

I don't much like this way of writing things. A notation which is more compact and powerful (once you learn the tricks of manipulating it) is to define the Levi-Civita symbol, a 3 index tensor

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if} \quad (i, j, k) = (1, 2, 3), (2, 3, 1) \text{ or } (3, 1, 2) \\ -1 & \text{if} \quad (i, j, k) = (1, 3, 2), (2, 1, 3) \text{ or } (3, 2, 1) \\ 0 & \text{if} \quad i = j \text{ or } j = k \text{ or } i = k \end{cases}.$$

Then $\vec{C} = \vec{A} \times \vec{B}$ means

$$C_i = \sum_{j,k=1}^{3} \epsilon_{ijk} A_j B_k.$$

From the definition we see immediately that ϵ_{ijk} changes sign under interchanging any two indices, and is unaffected by a cyclic permutation of them $((i, j, k) \rightarrow (j, k, i))$. An immediate consequence of the former is $\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$, while the latter shows

$$\vec{V} \cdot \left(\vec{A} \times \vec{B} \right) = \sum_{ijk} V_i \epsilon_{ijk} A_j B_k = \sum \epsilon_{jki} A_j B_k V_i = \vec{A} \cdot \left(\vec{B} \times \vec{V} \right).$$

A very useful property of the ϵ's is what happens if you sum the product of two of them over one index:

$$\sum_{i=1}^{3} \epsilon_{ijk} \epsilon_{i\ell m} = \delta_{j\ell} \delta_{km} - \delta_{jm} \delta_{k\ell},$$

where $\delta_{j\ell} = 1$ if $j = \ell$ and otherwise is 0, and is called the *Kronecker delta*.

$$\sum_{j\ell} A_j \delta_{j\ell} B_\ell = \sum_j A_j B_j = \vec{A} \cdot \vec{B}.$$

Thus

$$\begin{aligned}
\left(\vec{A} \times \vec{B} \right) \cdot \left(\vec{C} \times \vec{D} \right) &= \sum \epsilon_{ijk} A_j B_k \epsilon_{i\ell m} C_\ell D_m \\
&= \sum A_j B_k C_\ell D_m \left(\delta_{j\ell} \delta_{km} - \delta_{jm} \delta_{k\ell} \right) \\
&= \vec{A} \cdot \vec{C} \, \vec{B} \cdot \vec{D} - \vec{A} \cdot \vec{D} \, \vec{B} \cdot \vec{C}.
\end{aligned}$$

Many other arcane identities can be easily proved once you learn to jiggle the ϵ's.

Homework:

1. Using the Levi-Civita ϵ's, show that[1]

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}).$$

2. Consider a parallelopiped whose edges at one vertex are given by three vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$. Show that it has a volume given by $\vec{A} \cdot (\vec{B} \times \vec{C})$. Note that the volume is the area of one face times the distance (measured perpendicular to that face) between that face and the one opposite.

---

[1]Note that this implies $\vec{A} \times (\vec{B} \times \vec{C}) - (\vec{A} \times \vec{B}) \times \vec{C} = \vec{A} \times (\vec{B} \times \vec{C}) - \vec{C} \times (\vec{B} \times \vec{C}) = \vec{A}(\vec{B} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}) \neq 0$, so the cross product is *not* associative.

1.5   Curl

Another way of combining derivatives and vectors is the curl, which is a formal cross product of the derivative operator with a vector field

$$\vec{C} = \vec{\nabla} \times \vec{V} \quad \text{which means} \quad C_i = \sum_{jk} \epsilon_{ijk} \frac{\partial}{\partial x_j} V_k.$$

The word comes from the fact that if $\vec{v}$ is the velocity field, $\vec{\nabla} \times \vec{v}$ is a measure of the rotation of a small volume of the fluid. If we think of a small volume, then the motion will have $v_i(\vec{r}) = v_i(\vec{r}_0) + (\vec{r} - \vec{r}_0)_j \frac{\partial}{\partial x_j} v_i + \mathcal{O}(\vec{r} - \vec{r}_0)^2$.

The first term is the overall velocity. The second term is $\sum_j (\vec{r} - \vec{r}_0)_j M_{ji}$, with $M_{ji} = \frac{\partial v_i}{\partial x_j}$. M can be though of as a sum of

- $D\,\delta_{ij}$, a term which describes how the volume is expanding

- $\sum_k L_k \epsilon_{jik}$, a rotational term given by the curl, $\vec{L} = \vec{\nabla} \times \vec{v}$

- $\frac{1}{2}\left(M_{ij} + M_{ji}\right) - \frac{1}{3}\delta_{ij} \sum_k M_{kk}$, a traceless symmetric tensor describing deformations in which dimensions in one dimension are expanded and another contracted, or a shear. This is called the *deviatoric* part.

Irrotational flow is defined by $\vec{\nabla} \times \vec{v} = 0$ everywhere. A flow is called solenoidal if $\vec{\nabla} \cdot \vec{v} = 0$.

Homework: Show that $\vec{u} \times \vec{v}$ is solenoidal (*i.e.* $\vec{\nabla} \cdot (\vec{u} \times \vec{v}) = 0$) if $\vec{u}$ and $\vec{v}$ are each irrotational (*i.e.* $\vec{\nabla} \times \vec{u} = \vec{\nabla} \times \vec{v} = 0$).

1.6   Repeated Derivatives

Consider a scalar field $\lambda(\vec{x})$. Its gradient $\vec{\nabla}\lambda$ is a vector field, so we can ask what the divergence and curl are.

$$\vec{\nabla} \cdot \vec{\nabla}\lambda = \sum_i \frac{\partial}{\partial x_i} \frac{\partial \lambda}{\partial x_i} = \sum_i \frac{\partial^2 \lambda}{\partial x_i^2} = \nabla^2 \lambda,$$

the laplacian.

The curl is

$$\left(\vec{\nabla} \times \vec{\nabla}\lambda\right)_i = \sum_{jk} \epsilon_{ijk} \frac{\partial}{\partial x_j} \frac{\partial \lambda}{\partial x_k}$$

$$= \sum_{jk} \underbrace{\epsilon_{ijk}}_{\text{antisym under } j \leftrightarrow k} \underbrace{\frac{\partial^2 \lambda}{\partial x_j \partial x_k}}_{\text{sym under } j \leftrightarrow k} = 0$$

Thus in general, the curl of a gradient vanishes.

Now consider a vector field $\vec{V}$. Its divergence $\vec{\nabla} \cdot \vec{V}$ is a scalar with a gradient

$$\vec{\nabla}\left(\vec{\nabla} \cdot \vec{V}\right) = \sum_j \hat{e}_j \frac{\partial}{\partial x_j} \frac{\partial V_i}{\partial x_i} = \sum_j \hat{e}_j \frac{\partial^2 V_i}{\partial x_j \partial x_i}.$$

The curl of $\vec{V}$, $\left(\vec{\nabla} \times \vec{V}\right)_i = \sum_{jk} \epsilon_{ijk} \frac{\partial V_k}{\partial x_j}$, has a divergence

$$\vec{\nabla} \cdot \left(\vec{\nabla} \times \vec{V}\right) = \sum_i \frac{\partial}{\partial x_i} \left(\sum_{jk} \epsilon_{ijk} \frac{\partial V_k}{\partial x_j}\right) = \sum_{ijk} \epsilon_{ijk} \frac{\partial^2 V_k}{\partial x_i \partial x_j} = 0$$

as for $\vec{\nabla} \times \vec{\nabla}\lambda$. So the divergence of a curl vanishes in general. The curl of the curl,

$$\left[\vec{\nabla} \times \left(\vec{\nabla} \times \vec{V}\right)\right]_i = \sum_{ijk} \epsilon_{ijk} \frac{\partial}{\partial x_j} \left(\vec{\nabla} \times \vec{V}\right)_k = \sum_{ijk\ell m} \epsilon_{ijk}\epsilon_{k\ell m} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_\ell} \vec{V}_m.$$

Noting that $\epsilon_{ijk} = \epsilon_{kij}$ and recalling that $\sum_k \epsilon_{kij}\epsilon_{k\ell m} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}$, we have

$$\left[\vec{\nabla} \times \left(\vec{\nabla} \times \vec{V}\right)\right]_i = \sum_{j\ell m} (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}) \frac{\partial^2}{\partial x_j \partial x_\ell} V_m$$

$$= \sum_j \frac{\partial^2}{\partial x_i \partial x_j} V_j - \left(\sum_j \frac{\partial^2}{\partial x_j^2}\right) V_i$$

The first term is just $\left[\vec{\nabla}\left(\vec{\nabla} \cdot \vec{V}\right)\right]_i$, while the second is $-\nabla^2 V_i$, so

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{V}\right) = \vec{\nabla}\left(\vec{\nabla} \cdot \vec{V}\right) - \nabla^2\vec{V}.$$

2   Integration

In integration in a 3 dimensional space, we are involved with integrals along paths: $\int_C d\vec{r}$; integrals over surfaces: $\int_S d\vec{\sigma}$; and integrals over volumes: $\int_V d\tau$.

Integrals over paths: the most important one is $\int \vec{V} \cdot d\vec{r}$, for some vector field $\vec{V}$. For example, the work done by a force $\vec{F}$ is $W = \int_C \vec{F} \cdot d\vec{r}$, where the path C is the path the particle acted on moves along. The vector $d\vec{r}$ points along the path. If we parameterize the path with the real variable λ, so that $\vec{r}(\lambda)$ describes the path, the integral is then $W = \int d\lambda \vec{F} \cdot \dfrac{\partial \vec{r}}{\partial \lambda}(\lambda)$.

For surface integrals, $d\vec{\sigma}$ is thought of as a vector[2] perpendicular to the surface element, with a length equal to the area of the element. It can be thought of as $d\vec{\sigma} = d\vec{u} \times d\vec{v}$, where $d\vec{u}$ and $d\vec{v}$ are the sides of an infinitesimal parallelogram on the surface. There is an ambiguity in the direction. For a closed surface (*e.g.* a sphere's surface), it is by convention taken outwards. For a surface with a boundary the direction taken for the surface elements imposes a direction on the boundary, by the right hand rule, fingers curling in the direction of the boundary, thumb pointing in the direction of the surface $d\vec{\sigma}$.

Finally, we have volume integrals, where $d\tau = dx\, dy\, dz$ acts like a scalar.

There are some very important relations on these integrals which are essentially statements which are generalizations to more dimensions of the fundamental theorem of calculus: $\displaystyle\int_a^b dx\, \frac{df}{dx} = f(b) - f(a)$ for an integral of a derivative on a real line segment. These are

$$\text{Gauss' Theorem:}\quad \int_V \vec{\nabla} \cdot \vec{B}\, d\tau = \int_S \vec{B} \cdot d\vec{\sigma}.$$

Here S is a closed surface bounding the three dimensional region V, $\vec{B}$ any arbitrary differentiable vector field.

$$\text{Stokes' Theorem:}\quad \int_S \left(\vec{\nabla} \times \vec{B} \right) \cdot d\vec{\sigma} = \oint_C \vec{B} \cdot d\vec{r}.$$
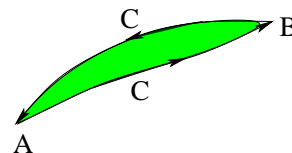
Here S is a bounded two-dimensional region with boundary C. The direction

---

[2]Note this is only in three dimensional space. In a space with more than three dimensions, the surface element must be considered an antisymmetric tensor or 2-form.

of S and C agree as discussed above. Note if S is closed, C may be considered a point, or empty, and the integral must be zero.

Examples of the usefulness of these concepts

a) Suppose we have a force $\vec{F}$ given by $\vec{F} = -\vec{\nabla}\phi$, where ϕ is the potential for this force. We expect $\vec{F}$ is conservative, so that the work done in going from A to B is independent of the path, and minus the work done in going from B to A. Therefore the work done along any closed path C should be zero. Let S be the surface bounded by C.

$$\oint_C \vec{F} \cdot d\vec{r} = \int_S \left(\vec{\nabla} \times \vec{F} \right) \cdot d\vec{\sigma} = -\int_S \left(\underbrace{\vec{\nabla} \times \vec{\nabla}\phi}_{0} \right) \cdot d\vec{\sigma} = 0,$$

So $\vec{F} = -\vec{\nabla}\phi$ implies $\vec{F}$ is conservative.

b) Maxwell's law $\vec{\nabla} \cdot \vec{D} = \rho$ means the total charge within a volume

$$\int_V d\tau\, \rho = \int_V d\tau\, \vec{\nabla} \cdot \vec{D} = \int_S \vec{D} \cdot d\vec{\sigma} \tag{1}$$

is the flux of $\vec{D}$ emerging from the surface which bounds this volume. For a point charge at the center, we have

$$q = \int_V d\tau\, \vec{\nabla} \cdot \left(\frac{q}{4\pi} \frac{\vec{r}}{r^3} \right) = -\int_V d\tau\, \nabla^2 \frac{q}{4\pi r},$$

where we know, from last time, that the integrand is zero except at the origin. Thus we learn

$$\nabla^2 \frac{1}{r} = -4\pi \delta^3(\vec{r}),$$

where $\delta^3(\vec{r} - \vec{r}_0)$ is a Dirac delta function in 3 dimensions, a "generalized function" zero everywhere except at the origin, but so singular there that

$$\int f(\vec{r})\, \delta^3(\vec{r} - \vec{r}_0)\, d^3r = f(\vec{r}_0)$$

for any function f.

The statement Eq.(1) that the total charge within a volume can be determined by the flux of electric displacement $\vec{D}$ on the surface has a profound consequence when considered together with the relativistic notion that no signal can travel faster than light. It says charge is conserved. For it says that if charge near the center of a large sphere were not conserved, but changed at time t_0, the flux measured as a function of time t would have to change at time t_0, before any signal that the charge had changed could reach the surface. This is impossible, so charge must be conserved!

2.1   Helmholtz' Theorem

A vector field $\vec{V}$ whose curl is zero is called *irrotational*. It can be written, at least locally[3], as a gradient of a scalar field.

A vector field $\vec{V}$ whose divergence is zero is called *solenoidal*. It can be written as a curl of another vector field.

Any vector field $\vec{V}$ can be written as a sum of an irrotational field and a solenoidal field, so there exist ϕ and $\vec{A}$ such that

$$\vec{V} = -\vec{\nabla}\phi + \vec{\nabla} \times \vec{A} \,.$$

ϕ and $\vec{A}$ can be found by noting

$$\begin{aligned}
\vec{\nabla} \cdot \vec{V} &= -\nabla^2 \phi =: s(\vec{r}) \\
\vec{\nabla} \times \vec{V} &= \vec{\nabla} \times \left(\vec{\nabla} \times \vec{A}\right) = \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A}\right) - \nabla^2 \vec{A} =: \vec{c}(\vec{r}).
\end{aligned}$$

We can find a solution to the Poisson equation

$$\nabla^2 \lambda = -\rho(\vec{r})$$

by setting

$$\lambda(\vec{r}) = \int \frac{d^3 r'}{4\pi} \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|}.$$

---

[3]On a region which is not simply connected, the scalar field may not be single-valued. An example is a vortex $|\vec{V}| = 1/r$ in the tangential direction in the plane, which is only defined in a region not including the origin, and thus not simply connected. Then $\vec{V} = \vec{\nabla}\theta$, but θ is not uniquely defined, as $(x = -1, y = 0)$ can be $\theta = \pi$ or $\theta = -\pi$. For more complicated regions this ambiguity can be more interesting, leading to *cohomology*.

Then

$$\nabla^2 \lambda(\vec{r}) = \int \frac{d^3 r'}{4\pi} \rho(\vec{r}') \underbrace{\nabla^2 \frac{1}{|\vec{r} - \vec{r}'|}}_{-4\pi \delta^3 (\vec{r} - \vec{r}')} = -\rho(\vec{r}).$$

For $\vec{A}$, there is an ambiguity in the solution, for if $\vec{A} \to \vec{A} + \vec{\nabla}\sigma$, with σ a scalar field (gauge transformation), then $\vec{\nabla} \times \vec{A}$ is unchanged. This means that we can choose $\vec{\nabla} \cdot \vec{A} = 0$. Then $\nabla^2 A_i = -c_i$, so[4]

$$A_i(\vec{r}) = \int \frac{d^3 r'}{4\pi} \frac{c_i(\vec{r}')}{|\vec{r} - \vec{r}'|}.$$

Of course we need to check that $\vec{\nabla} \cdot \vec{A} = 0$.

$$\vec{\nabla} \cdot \vec{A} = \int \frac{d^3 r'}{4\pi} \vec{c}(\vec{r}') \cdot \vec{\nabla}_{\vec{r}} \frac{1}{|\vec{r} - \vec{r}'|}.$$

But $\vec{\nabla}_{\vec{r}} \frac{1}{|\vec{r} - \vec{r}'|} = -\vec{\nabla}_{\vec{r}'} \frac{1}{|\vec{r} - \vec{r}'|}$ because the thing being differentiated is only a function of $\vec{r} - \vec{r}'$. Thus

$$\begin{aligned}
\vec{\nabla} \cdot \vec{A} &= -\int \frac{d^3 r'}{4\pi} \vec{c}(\vec{r}') \cdot \vec{\nabla}_{\vec{r}'} \frac{1}{|\vec{r} - \vec{r}'|} \\
&= -\int \frac{d^3 r'}{4\pi} \vec{\nabla}_{\vec{r}'} \cdot \frac{\vec{c}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int \frac{d^3 r'}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|} \vec{\nabla}_{\vec{r}'} \cdot \vec{c}(\vec{r}')
\end{aligned}$$

by the Leibniz rule for $\vec{\nabla}\cdot$. The first term can be reduced to a surface integral which vanishes if we take the surface to infinity and assume $\vec{c}$ dies off quickly enough, while the second term is linear in $\vec{\nabla} \cdot \vec{c} = \vec{\nabla} \cdot \vec{\nabla} \times \vec{V} = 0$, so vanishes.

Thus the $\vec{A}$ and ϕ we have constructed satisfy

$$\vec{V} = -\vec{\nabla}\phi + \vec{\nabla} \times \vec{A}$$

provided the field $\vec{V}$ falls off fast enough at ∞ so that our integrals are defined and the surface contribution goes to zero.

You may recognize this procedure as the Green's function for the Laplace equation. We will discuss Green's functions later on.

Note the implications:

_____

[4]Note if the region is not simply connected, this is not a unique solution.

1. Any force whose work is path independent throughout space has zero curl and therefore can be written as $-\vec{\nabla}\phi$ for some potential ϕ

2. Maxwell's equations for the magnetic field (assuming the non-existence of magnetic monopoles) $\vec{\nabla}\cdot\vec{B} = 0$ implies there exists a vector potential $\vec{A}$ such that $\vec{B} = \vec{\nabla} \times \vec{A}$.