Vector Calculus

You don't really have to know this stuff to use my *HyperReference*. However, if you are mathematically inclined you will surely enjoy the elegance and economy of vector notation when applied to calculus; if nothing else this is an æsthetic treat — read it just for fun!

Functions of Several Variables

Suppose we go beyond $f(x)$ and talk about $F(x, y, z)$ — *e.g.* a function of the *exact position in space*. This is just an example, of course; the abstract idea of a function of several variables can have "several" be as many as you like and "variables" be anything you choose. Another place where we encounter lots of functions of "several" variables is in THERMODYNAMICS, but for the time being we will focus our attention on the three *spatial* variables x (left-right), y (back-forth) and z (up-down).

How can we tackle *derivatives* of this function?

Partial Derivatives

Well, we do the obvious: we say, "Hold all the *other* variables *fixed* except [for instance] x and then treat $F(x, y, z)$ as a function only of x, with y and z as fixed *parameters*." Then we know just how to define the derivative with respect to x. The short name for this derivative is the PARTIAL DERIVATIVE *with respect to* x, written symbolically

$$\frac{\partial F}{\partial x}$$

where the fact that there are other variables being held fixed is implied by the use of the symbol ∂ instead of just d.

Similarly for $\dfrac{\partial F}{\partial y}$ and $\dfrac{\partial F}{\partial z}$.

Operators

The foregoing description applies for *any* function of (x, y, z); the concept of "taking partial derivatives" is independent of what function we are taking the derivatives *of*. It is therefore practical to learn to think of

$$\frac{\partial}{\partial x} \quad \text{and} \quad \frac{\partial}{\partial y} \quad \text{and} \quad \frac{\partial}{\partial y}$$

as OPERATORS that can be applied to *any* function (like F). Put the operator on the left of a function, perform the operation and you get a partial derivative — a new function of (x, y, z). In general, such "operators" *change one function into another*. Physics is loaded with operators like these.

The GRADIENT Operator

The GRADIENT operator is a *vector* operator, written $\vec{\boldsymbol{\nabla}}$ and called "grad" or "del." It is defined (in Cartesian coordinates x, y, z) as[1]

$$\vec{\boldsymbol{\nabla}} \equiv \hat{\boldsymbol{\imath}} \frac{\partial}{\partial x} + \hat{\boldsymbol{\jmath}} \frac{\partial}{\partial y} + \hat{\boldsymbol{k}} \frac{\partial}{\partial z}$$

and can be applied directly to any *scalar* function of (x, y, z) — say, $\phi(x, y, z)$ — to turn it into a *vector* function, $\vec{\boldsymbol{\nabla}}\phi = \hat{\boldsymbol{\imath}}\frac{\partial \phi}{\partial x} + \hat{\boldsymbol{\jmath}}\frac{\partial \phi}{\partial y} + \hat{\boldsymbol{k}}\frac{\partial \phi}{\partial z}$.

GRADIENTS of Scalar Functions

It is instructive to work up to this "one dimension at a time." For simplicity we will stick to using ϕ as the symbol for the function of which we are taking derivatives.

The GRADIENT in One Dimension

[1] I am using the conventional notation for $\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}$ as the UNIT VECTORS in the x, y, z directions, respectively.

Let the dimension be x. Then we have no "extra" variables to hold constant and the gradient of $\phi(x)$ is nothing but $\hat{\boldsymbol{\imath}}\frac{d\phi}{dx}$. We can illustrate the "meaning" of $\vec{\boldsymbol{\nabla}}\phi$ by an example: let $\phi(x)$ be the mass of an object times the acceleration of gravity times the height h of a hill at horizontal position x. That is, $\phi(x)$ is the *gravitational potential energy* of the object when it is at horizontal position x. Then

$$\vec{\boldsymbol{\nabla}}\phi \;=\; \hat{\boldsymbol{\imath}}\,\frac{d\phi}{dx} \;=\; \hat{\boldsymbol{\imath}}\,\frac{d}{dx}(mgh) \;=\; mg\left(\frac{dh}{dx}\right)\hat{\boldsymbol{\imath}}.$$

Note that $\frac{dh}{dx}$ is the *slope* of the hill and $-\vec{\boldsymbol{\nabla}}\phi$ is the *horizontal component of the net force* (gravity plus the normal force from the hill's surface) on the object. That is, $-\vec{\boldsymbol{\nabla}}\phi$ is the *downhill force*.

The Gradient in Two Dimensions

In the previous example we disregarded the fact that most hills extend in *two* horizontal directions, say x = East and y = North. [If we stick to small distances we won't notice the curvature of the Earth's surface.] In this case there are two *components* to the slope: the Eastward slope $\frac{\partial h}{\partial x}$ and the Northward slope $\frac{\partial h}{\partial y}$. The former is a measure of how steep the hill will seem if you head due East and the latter is a measure of how steep it will seem if you head due North. If you put these together to form a *vector* "steepness" (gradient)

$$\vec{\boldsymbol{\nabla}}h \;=\; \hat{\boldsymbol{\imath}}\,\frac{\partial h}{\partial x} \;+\; \hat{\boldsymbol{\jmath}}\,\frac{\partial h}{\partial y}$$

then the vector $\vec{\boldsymbol{\nabla}}h$ points *uphill* — *i.e.* in the direction of the *steepest ascent*. Moreover, the gravitational potential energy $\phi = mgh$ as before [only now ϕ is a function of *2* variables, $\phi(x,y)$] so that $-\vec{\boldsymbol{\nabla}}\phi$ is once again the *downhill force* on the object.

The Gradient in Three Dimensions

If the potential ϕ is a function of *3* variables, $\phi(x,y,z)$ [such as the three *spatial coordinates* x, y and z — in which case we can write it a little more compactly as $\phi(\vec{\boldsymbol{r}})$ where $\vec{\boldsymbol{r}} \equiv x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}$, the vector distance from the origin of our coordinate system to the point in space where ϕ is being evaluated], then it is a little more difficult to make up a "hill" analogy — try imagining a topographical map in the form of a 3-dimensional hologram where instead of *lines* of constant *altitude* the "equipotentials" are *surfaces* of constant ϕ. (This is just what Physicists do picture!) Fortunately the *math* extends easily to 3 dimensions (or any larger number, if that has any meaning in the context we choose).

In general, any time there is a *potential energy* function $\phi(\vec{\boldsymbol{r}})$ we can immediately write down the *force* $\vec{\boldsymbol{F}}$ associated with it as

$$\vec{\boldsymbol{F}} \;\equiv\; -\vec{\boldsymbol{\nabla}}\phi$$

A perfectly analogous expression holds for the *electric field* $\vec{\boldsymbol{E}}$ [force per unit charge] in terms of the *electrostatic potential* ϕ [potential energy per unit charge]:[2]

$$\vec{\boldsymbol{E}} \;\equiv\; -\vec{\boldsymbol{\nabla}}\phi$$

The Gradient in N Dimensions

Although we won't be needing to go beyond 3 dimensions very often in Physics, you might want to borrow this metaphor for application in other realms of human endeavour where there are more than 3 variables of which your scalar field is a function. You could have ϕ be a measure of *happiness*, for instance [though it is hard to take reliable measurements on such a subjective quantity]; then ϕ might be a function of lots of factors, such as x_1 = freedom from violence, x_2 = freedom from hunger, x_3 = freedom from poverty, x_4 = freedom from

[2] I know, I know, I am using the ϕ symbol for two different things. Well, I *said* it was the preferred symbol for a scalar field, so you shouldn't be surprised to see it "recycled" many times. This won't be the last!

oppression, and so on.[3] Note that with an arbitrary number of variables we get away from thinking up different names for each one and just call the $i^{\rm th}$ variable "x_i."

Then we can define the GRADIENT in N dimensions as

$$\vec{\boldsymbol{\nabla}}\phi \;=\; \hat{\boldsymbol{\imath}}_1 \frac{\partial \phi}{\partial x_1} \;+\; \hat{\boldsymbol{\imath}}_2 \frac{\partial \phi}{\partial x_2} \;+\; \cdots \;+\; \hat{\boldsymbol{\imath}}_N \frac{\partial \phi}{\partial x_N}$$

$$\text{or} \quad \vec{\boldsymbol{\nabla}}\phi \;=\; \sum_{i=1}^{N} \hat{\boldsymbol{\imath}}_i \frac{\partial \phi}{\partial x_i}$$

where $\hat{\boldsymbol{\imath}}_i$ is a UNIT VECTOR in the x_i direction.

DIVERGENCE of a Vector Field

If we form the scalar ("dot") product of $\vec{\boldsymbol{\nabla}}$ with a *vector* function $\vec{\boldsymbol{A}}(x, y, z)$ we get a *scalar* result called the DIVERGENCE of $\vec{\boldsymbol{A}}$:

$$\text{div}\vec{\boldsymbol{A}} \;\equiv\; \vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{A}} \;\equiv\; \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

This name is actually quite mnemonic: the DIVERGENCE of a vector field is a *local* measure of its "outgoingness" — *i.e.* the extent to which there is more *exiting* an infinitesimal region of space than *entering* it. If the field is represented as "flux lines" of some indestructible "stuff" being emitted by "sources" and absorbed by "sinks," then a nonzero DIVERGENCE at some point means there must be a *source* or *sink* at that position. That is to say,

"What leaves a region is no longer in it."

For example, consider the divergence of the CURRENT DENSITY $\vec{\boldsymbol{J}}$, which describes the FLUX of a CONSERVED QUANTITY such as electric charge Q. (Mass, as in the current of a river, would do just as well.)

[3] These are rotten examples, of course — the first practical criterion for the variables of which any ϕ is a function is that they should be *linearly independent* [*i.e.* *orthogonal*] so that the dependence on one is not all mixed up with the dependence on another!

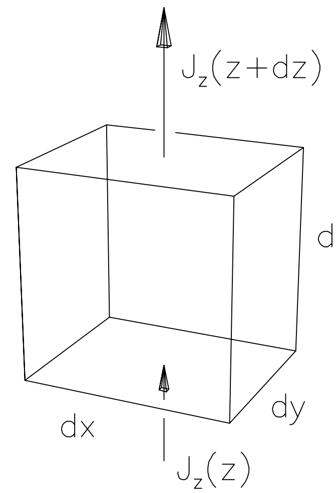

Figure 1 Flux into and out of a volume element $dV = dx\, dy\, dz$.

To make this as easy as possible, let's picture a *cubical* volume element $dV = dx\, dy\, dz$. In general, $\vec{\boldsymbol{J}}$ will (like any vector) have three components (J_x, J_y, J_z), each of which may be a function of position (x, y, z). If we take the lower left front corner of the cube to have coordinates (x, y, z) then the upper right back corner has coordinates $(x + dx,\, y + dy,\, z + dz)$. Let's concentrate first on J_z and how it depends on z.

It may not depend on z at all, of course. In this case, the amount of Q coming into the cube through the bottom surface (per unit time) will be the same as the amount of Q going out through the top surface and there will be no net gain or loss of Q in the volume — at least not due to J_z.

If J_z is bigger at the top, however, there will be a net loss of Q within the volume dV due to the "divergence" of J_z. Let's see how much: the difference between $J_z(z)$ at the bottom and $J_z(z + dz)$ at the top is, by definition, $dJ_z = \left(\frac{\partial J_z}{\partial z}\right) dz$. The flux is over the same area at top and bottom, namely $dx\, dy$, so the total *rate of loss* of Q due to the z-dependence of J_z is given by

$$\dot{Q}_z = -dx\,dy \left(\frac{\partial J_z}{\partial z}\right) dz = -\left(\frac{\partial J_z}{\partial z}\right) dx\,dy\,dz$$

$$\text{or} \quad \dot{Q} = -\left(\frac{\partial J_z}{\partial z}\right) dV\,.$$

A perfectly analogous argument holds for the x-dependence if J_x and the y-dependence of J_y, giving a total rate of change of Q

$$\dot{Q} = -\left(\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z}\right) dV$$

$$\text{or} \quad \dot{Q} = -\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{J}}\, dV$$

The total amount of Q *in* our volume element dV at a given instant is just $\rho\, dV$, of course, so the rate of change of the enclosed Q is just

$$\dot{Q} = \dot{\rho}\, dV$$

which means that we can write

$$\frac{\partial \rho}{\partial t}\, dV = -\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{J}}\, dV$$

or, just cancelling out the common factor dV on both sides of the equation,

$$\boxed{\frac{\partial \rho}{\partial t} = -\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{J}}}$$

which is the compact and elegant "differential form" of the EQUATION OF CONTINUITY.

This equation tells us that the "Q sourciness" of *each point* in space is given by the degree to which flux "lines" of $\vec{\boldsymbol{J}}$ tend to radiate *away* from that point more than they converge *toward* that point — namely, the DIVERGENCE of $\vec{\boldsymbol{J}}$ at the point in question. This esoteric-looking mathematical expression is, remember, just a formal way of expressing our original dumb tautology!

CURL of a Vector Field

If we form the vector ("cross") product of $\vec{\boldsymbol{\nabla}}$ with a *vector* function $\vec{\boldsymbol{A}}(x, y, z)$ we get a *vector* result called the **curl** of $\vec{\boldsymbol{A}}$:

$$\mathbf{curl}\,\vec{\boldsymbol{A}} \equiv \vec{\boldsymbol{\nabla}} \times \vec{\boldsymbol{A}} \equiv \hat{\boldsymbol{\imath}}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \hat{\boldsymbol{\jmath}}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \hat{\boldsymbol{k}}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)$$

This is a lot harder to visualize than the DIVERGENCE, but not impossible. Suppose you are in a boat in a huge river (or Pass) where the current flows mainly in the x direction but where the speed of the current (flux of water) varies with y. Then if we call the current $\vec{\boldsymbol{J}}$, we have a nonzero value for the derivative $\frac{\partial J_x}{\partial y}$, which you will recognize as one of the terms in the formula for $\vec{\boldsymbol{\nabla}} \times \vec{\boldsymbol{J}}$. What does this imply? Well, if you are sitting in the boat, moving with the current, it means the current on your port side moves faster — *i.e.* forward relative to the boat — and the current on your starboard side moves slower — *i.e.* backward relative to the boat — and this implies a *circulation* of the water around the boat — *i.e.* a *whirlpool!* So $\vec{\boldsymbol{\nabla}} \times \vec{\boldsymbol{J}}$ is a measure of the local "swirliness" of the current $\vec{\boldsymbol{J}}$, which means "**curl**" is not a bad name after all!

The LAPLACIAN Operator

If we form the scalar ("dot") product of $\vec{\boldsymbol{\nabla}}$ with *itself* we get a scalar *second derivative* operator called the LAPLACIAN:

$$\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{\nabla}} \equiv \nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

What does the ∇^2 operator "mean?" It is the three-dimensional generalization of the one-dimensional CURVATURE operator d^2/dx^2.

Consider the familiar one-dimensional function $h(x)$ where h is the height of a hill at horizontal position x. Then dh/dx is the *slope* of the hill and d^2h/dx^2 is its *curvature* (the rate of change of the slope with position). This property appears in every form of the WAVE EQUATION. In three dimensions, a nice visualization is harder (there is no extra dimension "into which to curve") but $\nabla^2\phi$ represents the equivalent property of a scalar function $\phi(x, y, z)$.

Gauss' Law

The Equation of Continuity (see above) describes the conservation of "actual physical stuff" entering or leaving an infinitesimal region of space dV. For example, $\vec{\boldsymbol{J}}$ may be the *current density* (charge flow per unit time per unit area normal to the direction of flow) in which case ρ is the *charge density* (charge per unit volume); in that example the conserved "stuff" is electric charge itself. Many other examples exist, such as FLUID DYNAMICS (in which *mass* is the conserved stuff) or HEAT FLOW (in which *energy* is the conserved quantity). In Electromagnetism, however, we deal not only with the conservation of *charge* but also with the *continuity* of abstract *vector fields* like $\vec{\boldsymbol{E}}$ and $\vec{\boldsymbol{B}}$. In order to visualize $\vec{\boldsymbol{E}}$, we have developed the notion of "electric field lines" that cannot be broken except where they originate (from positive charges) and terminate (on negative charges). [This description only holds for *static* electric fields; when things *move* or otherwise change with time, things get a lot more complicated … and interesting!] Thus a positive charge is a "source of electric field lines" and a negative charge is a "sink" — the charges themselves stay put, but the lines of $\vec{\boldsymbol{E}}$ *diverge* out of or into them. You can probably see where this is heading.

Gauss' Law states that the net flux of electric field "lines" *out* of a closed surface $\mathcal{S}$ is proportional to the net electric charge enclosed *within* that surface. The constant of proportionality depends on which system of units one is using; in SI units it is $1/\epsilon_\circ$. In mathematical shorthand, this reads

$$\epsilon_\circ \oiint_{\mathcal{S}} \vec{\boldsymbol{E}} \cdot d\vec{\boldsymbol{A}} \;=\; Q_{\rm encl}\,.$$

Recalling our earlier discussion of DIVERGENCE, we can think of $\vec{\boldsymbol{E}}$ as being a sort of flux density of conserved "stuff" *emitted* by positive electric charges. Remember, in this case the charges themselves do not go anywhere; they simply emit (or absorb) the electric field "lines" which emerge from (or disappear into) the enclosed region. The rate of generation of this "stuff" is $Q_{\rm encl}/\epsilon_\circ$. We can then apply Gauss' Law to an infinitesimal volume element using Fig. 1 with $\epsilon_\circ\vec{\boldsymbol{E}}$ in place of $\vec{\boldsymbol{J}}$. Except for the "fudge factor" $\epsilon_\circ$ and the replacement of $\dot{Q}$ by $Q_{\rm encl}$, the same arguments used to derive the Equation of Continuity lead in this case to a formula relating the divergence of $\vec{\boldsymbol{E}}$ to the electric charge density ρ *at any point in space*, namely

$$\boxed{\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{E}} \;=\; \frac{1}{\epsilon_\circ}\,\rho}\,.$$

This is the *differential form* of Gauss' Law.

Poisson and Laplace

Even in its differential form, Gauss' Law is a little tricky to solve analytically, since it is a *vector* differential equation. Generally we have an easier time solving *scalar* differential equations, even though they may involve higher order partial derivatives. Fortunately, we can convert the former into the latter: recall that the vector electric *field* can always be obtained from the scalar electrostatic *potential* using

$$\vec{\boldsymbol{E}} \;\equiv\; -\vec{\boldsymbol{\nabla}}\phi\,.$$

Thus $\text{div}\vec{\boldsymbol{E}} \equiv \vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{E}} = -\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{\nabla}}\phi$ or

$$\boxed{\nabla^2\phi \;=\; -\,\frac{1}{\epsilon_\circ}\,\rho}\,.$$

This relation is known as POISSON'S EQUATION. Its simplified cousin, LAPLACE'S EQUATION, applies in regions of space where there are *no free charges*:

$$\boxed{\nabla^2 \phi \ = \ 0}.$$

Each of these equations finds much use in real electrostatics problems. Advanced students of electromagnetism learn many types of functions that satisfy LAPLACE'S EQUATION, with different *symmetries*; since a *conductor* is always an *equipotential* (every point in a given conductor must have the same ϕ, otherwise there would be an electric field in the conductor that would cause charges to move until they cancelled out the differences in ϕ), empty regions surrounded by conductors of certain shapes must have ϕ with a spatial dependence satisfying those BOUNDARY CONDITIONS as well as LAPLACE'S EQUATION. One can often write down a complicated-looking formula for ϕ almost by inspection, using this favourite method of Physicists and Mathematicians, namely ... KNOWING THE ANSWER.