

Differential Calculus of Vector Fields

2-1 Understanding physics

The physicist needs a facility in looking at problems from several points of view. The exact analysis of real physical problems is usually quite complicated, and any particular physical situation may be too complicated to analyze directly by solving the differential equation. But one can still get a very good idea of the behavior of a system if one has some feel for the character of the solution in different circumstances. Ideas such as the field lines, capacitance, resistance, and inductance are, for such purposes, very useful. So we will spend much of our time analyzing them. In this way we will get a feel as to what should happen in different electromagnetic situations. On the other hand, none of the heuristic models, such as field lines, is really adequate and accurate for all situations. There is only one precise way of presenting the laws, and that is by means of differential equations. They have the advantage of being fundamental and, so far as we know, precise. If you have learned the differential equations you can always go back to them. There is nothing to unlearn.

It will take you some time to understand what should happen in different circumstances. You will have to solve the equations. Each time you solve the equations, you will learn something about the character of the solutions. To keep these solutions in mind, it will be useful also to study their meaning in terms of field lines and of other concepts. This is the way you will really “understand” the equations. That is the difference between mathematics and physics. Mathematicians, or people who have very mathematical minds, are often led astray when “studying” physics because they lose sight of the physics. They say: “Look, these differential equations—the Maxwell equations—are all there is to electrodynamics; it is admitted by the physicists that there is nothing which is not contained in the equations. The equations are complicated, but after all they are only mathematical equations and if I understand them mathematically inside out, I will understand the physics inside out.” Only it doesn’t work that way. Mathematicians who study physics with that point of view—and there have been many of them—usually make little contribution to physics and, in fact, little to mathematics. They fail because the actual physical situations in the real world are so complicated that it is necessary to have a much broader understanding of the equations.

What it means really to understand an equation—that is, in more than a strictly mathematical sense—was described by Dirac. He said: “I understand what an equation means if I have a way of figuring out the characteristics of its solution without actually solving it.” So if we have a way of knowing what should happen in given circumstances without actually solving the equations, then we “understand” the equations, as applied to these circumstances. A physical understanding is a completely unmathematical, imprecise, and inexact thing, but absolutely necessary for a physicist.

Ordinarily, a course like this is given by developing gradually the physical ideas—by starting with simple situations and going on to more and more complicated situations. This requires that you continuously forget things you previously learned—things that are true in certain situations, but which are not true in general. For example, the “law” that the electrical force depends on the square of the distance is not *always* true. We prefer the opposite approach. We prefer to take first the *complete* laws, and then to step back and apply them to simple situations, developing the physical ideas as we go along. And that is what we are going to do.

2-1 Understanding physics

2-2 Scalar and vector fields— T and h

2-3 Derivatives of fields—the gradient

2-4 The operator ∇

2-5 Operations with ∇

2-6 The differential equation of heat flow

2-7 Second derivatives of vector fields

2-8 Pitfalls

Review: Chapter 11, Vol. I, Vectors

Our approach is completely opposite to the historical approach in which one develops the subject in terms of the experiments by which the information was obtained. But the subject of physics has been developed over the past 200 years by some very ingenious people, and as we have only a limited time to acquire our knowledge, we cannot possibly cover everything they did. Unfortunately one of the things that we shall have a tendency to lose in these lectures is the historical, experimental development. It is hoped that in the laboratory some of this lack can be corrected. You can also fill in what we must leave out by reading the Encyclopedia Britannica, which has excellent historical articles on electricity and on other parts of physics. You will also find historical information in many textbooks on electricity and magnetism.

2-2 Scalar and vector fields— T and h

We begin now with the abstract, mathematical view of the theory of electricity and magnetism. The ultimate idea is to explain the meaning of the laws given in Chapter 1. But to do this we must first explain a new and peculiar notation that we want to use. So let us forget electromagnetism for the moment and discuss the mathematics of vector fields. It is of very great importance, not only for electromagnetism, but for all kinds of physical circumstances. Just as ordinary differential and integral calculus is so important to all branches of physics, so also is the differential calculus of vectors. We turn to that subject.

Listed below are a few facts from the algebra of vectors. It is assumed that you already know them.

$$\mathbf{A} \cdot \mathbf{B} = \text{scalar} = A_x B_x + A_y B_y + A_z B_z \quad (2.1)$$

$$\mathbf{A} \times \mathbf{B} = \text{vector} \quad (2.2)$$

$$(\mathbf{A} \times \mathbf{B})_z = A_x B_y - A_y B_x$$

$$(\mathbf{A} \times \mathbf{B})_x = A_y B_z - A_z B_y$$

$$(\mathbf{A} \times \mathbf{B})_y = A_z B_x - A_x B_z$$

$$\mathbf{A} \times \mathbf{A} = 0 \quad (2.3)$$

$$\mathbf{A} \cdot (\mathbf{A} \times \mathbf{B}) = 0 \quad (2.4)$$

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} \quad (2.5)$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \quad (2.6)$$

Also we will want to use the two following equalities from the calculus:

$$\Delta f(x, y, z) = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial z} \Delta z, \quad (2.7)$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}. \quad (2.8)$$

The first equation (2.7) is, of course, true only in the limit that Δx , Δy , and Δz go toward zero.

The simplest possible physical field is a scalar field. By a field, you remember, we mean a quantity which depends upon position in space. By a *scalar field* we merely mean a field which is characterized at each point by a single number—a scalar. Of course the number may change in time, but we need not worry about that for the moment. We will talk about what the field looks like at a given instant. As an example of a scalar field, consider a solid block of material which has been heated at some places and cooled at others, so that the temperature of the body varies from point to point in a complicated way. Then the temperature will be a function of x , y , and z , the position in space measured in a rectangular coordinate system. Temperature is a scalar field.

Writing vectors by hand.

Some people use

\vec{E} or \bar{E} or just \bar{E} .

Others prefer

\underline{E} .

We like the following way:

A B C D E F G

H I J K L M N

O P Q R S T U

V W X Y Z

Small letters are harder:

a b c d e f g

h i j k l m n

o p q r s t u

v w x y z

You can invent your own.

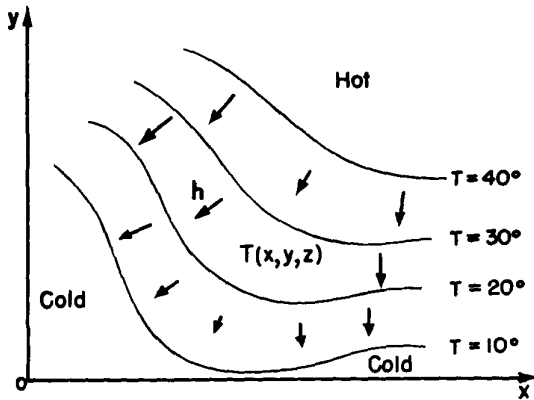


Fig. 2-1. Temperature T is an example of a scalar field. With each point (x, y, z) in space there is associated a number $T(x, y, z)$. All points on the surface marked $T = 20^\circ$ (shown as a curve at $z = 0$) are at the same temperature. The arrows are samples of the heat flow vector h .

One way of thinking about scalar fields is to imagine “contours” which are imaginary surfaces drawn through all points for which the field has the same value, just as contour lines on a map connect points with the same height. For a temperature field the contours are called “isothermal surfaces” or isotherms. Figure 2-1 illustrates a temperature field and shows the dependence of T on x and y when $z = 0$. Several isotherms are drawn.

There are also vector fields. The idea is very simple. A vector is given for each point in space. The vector varies from point to point. As an example, consider a rotating body. The velocity of the material of the body at any point is a vector which is a function of position (Fig. 2-2). As a second example, consider the flow of heat in a block of material. If the temperature in the block is high at one place and low at another, there will be a flow of heat from the hotter places to the colder. The heat will be flowing in different directions in different parts of the block. The heat flow is a directional quantity which we call h . Its magnitude is a measure of how much heat is flowing. Examples of the heat flow vector are also shown in Fig. 2-1.

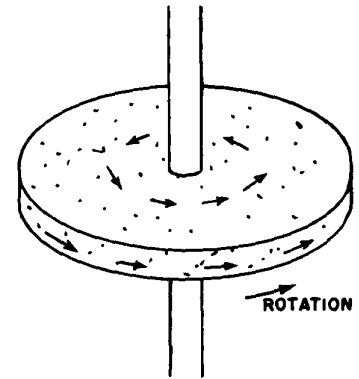


Fig. 2-2. The velocity of the atoms in a rotating object is an example of a vector field.

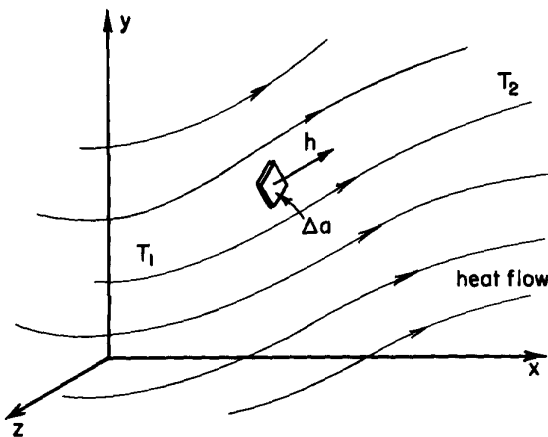


Fig. 2-3. Heat flow is a vector field. The vector h points along the direction of the flow. Its magnitude is the energy transported per unit time across a surface element oriented perpendicular to the flow, divided by the area of the surface element.

Let's make a more precise definition of h : The magnitude of the vector heat flow at a point is the amount of thermal energy that passes, per unit time and per unit area, through an infinitesimal surface element, at right angles to the direction of flow. The vector points in the direction of flow (see Fig. 2-3). In symbols: If ΔJ is the thermal energy that passes per unit time through the surface element Δa , then

$$h = \frac{\Delta J}{\Delta a} e_f, \quad (2.9)$$

where e_f is a *unit vector* in the direction of flow.

The vector h can be defined in another way—in terms of its components. We ask how much heat flows through a small surface at *any* angle with respect to the flow. In Fig. 2-4 we show a small surface Δa_2 inclined with respect to Δa_1 , which is perpendicular to the flow. The *unit vector* n is normal to the surface Δa_2 . The

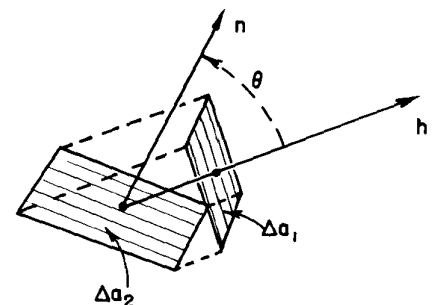


Fig. 2-4. The heat flow through Δa_2 is the same as through Δa_1 .

angle θ between \mathbf{n} and \mathbf{h} is the same as the angle between the surfaces (since \mathbf{h} is normal to Δa_1). Now what is the heat flow *per unit area* through Δa_2 ? The flow through Δa_2 is the same as through Δa_1 ; only the areas are different. In fact, $\Delta a_1 = \Delta a_2 \cos \theta$. The heat flow through Δa_2 is

$$\frac{\Delta J}{\Delta a_2} = \frac{\Delta J}{\Delta a_1} \cos \theta = \mathbf{h} \cdot \mathbf{n}. \quad (2.10)$$

We interpret this equation: the heat flow (per unit time and per unit area) through *any* surface element whose unit normal is \mathbf{n} , is given by $\mathbf{h} \cdot \mathbf{n}$. Equally, we could say: the component of the heat flow perpendicular to the surface element Δa_2 is $\mathbf{h} \cdot \mathbf{n}$. We can, if we wish, consider that these statements *define* \mathbf{h} . We will be applying the same ideas to other vector fields.

2-3 Derivatives of fields—the gradient

When fields vary in time, we can describe the variation by giving their derivatives with respect to t . We want to describe the variations with position in a similar way, because we are interested in the relationship between, say, the temperature in one place and the temperature at a nearby place. How shall we take the derivative of the temperature with respect to position? Do we differentiate the temperature with respect to x ? Or with respect to y , or z ?

Useful physical laws do not depend upon the orientation of the coordinate system. They should, therefore, be written in a form in which either both sides are scalars or both sides are vectors. What is the derivative of a scalar field, say $\partial T/\partial x$? Is it a scalar, or a vector, or what? It is neither a scalar nor a vector, as you can easily appreciate, because if we took a different x -axis, $\partial T/\partial x$ would certainly be different. But notice: We have three possible derivatives: $\partial T/\partial x$, $\partial T/\partial y$, and $\partial T/\partial z$. Since there are three kinds of derivatives and we know that it takes three numbers to form a vector, perhaps these three derivatives are the components of a vector:

$$\left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right) \stackrel{?}{=} \text{a vector}. \quad (2.11)$$

Of course it is not generally true that *any* three numbers form a vector. It is true only if, when we rotate the coordinate system, the components of the vector transform among themselves in the correct way. So it is necessary to analyze how these derivatives are changed by a rotation of the coordinate system. We shall show that (2.11) is indeed a vector. The derivatives do transform in the correct way when the coordinate system is rotated.

We can see this in several ways. One way is to ask a question whose answer is independent of the coordinate system, and try to express the answer in an “invariant” form. For instance, if $S = \mathbf{A} \cdot \mathbf{B}$, and if \mathbf{A} and \mathbf{B} are vectors, we know—because we proved it in Chapter 11 of Vol. I—that S is a scalar. We *know* that S is a scalar without investigating whether it changes with changes in coordinate systems. It *can't*, because it's a dot product of two vectors. Similarly, if we *know* that \mathbf{A} is a vector, and we have three numbers B_1 , B_2 , and B_3 , and we find out that

$$A_x B_1 + A_y B_2 + A_z B_3 = S, \quad (2.12)$$

where S is the same for any coordinate system, then it *must* be that the three numbers B_1 , B_2 , B_3 are the components B_x , B_y , B_z of some vector \mathbf{B} .

Now let's think of the temperature field. Suppose we take two points P_1 and P_2 , separated by the small interval ΔR . The temperature at P_1 is T_1 and at P_2 is T_2 , and the difference $\Delta T = T_2 - T_1$. The temperatures at these real, physical points certainly do not depend on what axis we choose for measuring the coordinates. In particular, ΔT is a number independent of the coordinate system. It is a scalar.

If we choose some convenient set of axes, we could write $T_1 = T(x, y, z)$ and $T_2 = T(x + \Delta x, y + \Delta y, z + \Delta z)$, where Δx , Δy , and Δz are the components of the vector ΔR (Fig. 2-5). Remembering Eq. (2.7), we can write

$$\Delta T = \frac{\partial T}{\partial x} \Delta x + \frac{\partial T}{\partial y} \Delta y + \frac{\partial T}{\partial z} \Delta z. \quad (2.13)$$

The left side of Eq. (2.13) is a scalar. The right side is the sum of three products with Δx , Δy , and Δz , which are the components of a vector. It follows that the three numbers

$$\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}$$

are also the x -, y -, and z -components of a vector. We write this new vector with the symbol ∇T . The symbol ∇ (called "del") is an upside-down Δ , and is supposed to remind us of differentiation. People read ∇T in various ways: "del- T ," or "gradient of T ," or "grad T ;"

$$\text{grad } T = \nabla T = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right)^*. \quad (2.14)$$

Using this notation, we can rewrite Eq. (2.13) in the more compact form

$$\Delta T = \nabla T \cdot \Delta R. \quad (2.15)$$

In words, this equation says that the difference in temperature between two nearby points is the dot product of the gradient of T and the vector displacement between the points. The form of Eq. (2.15) also illustrates clearly our proof above that ∇T is indeed a vector.

Perhaps you are still not convinced? Let's prove it in a different way. (Although if you look carefully, you may be able to see that it's really the same proof in a longer-winded form!) We shall show that the components of ∇T transform in just the same way that components of R do. If they do, ∇T is a vector according to our original definition of a vector in Chapter 11 of Vol. I. We take a new coordinate system x' , y' , z' , and in this new system we calculate $\partial T/\partial x'$, $\partial T/\partial y'$, and $\partial T/\partial z'$. To make things a little simpler, we let $z = z'$, so that we can forget about the z -coordinate. (You can check out the more general case for yourself.)

We take an $x'y'$ -system rotated an angle θ with respect to the xy -system, as in Fig. 2-6(a). For a point (x, y) the coordinates in the prime system are

$$x' = x \cos \theta + y \sin \theta, \quad (2.16)$$

$$y' = -x \sin \theta + y \cos \theta. \quad (2.17)$$

Or, solving for x and y ,

$$x = x' \cos \theta - y' \sin \theta, \quad (2.18)$$

$$y = x' \sin \theta + y' \cos \theta. \quad (2.19)$$

If any pair of numbers transforms with these equations in the same way that x and y do, they are the components of a vector.

Now let's look at the difference in temperature between the two nearby points P_1 and P_2 , chosen as in Fig. 2-6(b). If we calculate with the x - and y -coordinates, we would write

$$\Delta T = \frac{\partial T}{\partial x} \Delta x \quad (2.20)$$

—since Δy is zero.

* In our notation, the expression (a, b, c) represents a vector with components a , b , and c . If you like to use the unit vectors i, j , and k , you may write

$$\nabla T = i \frac{\partial T}{\partial x} + j \frac{\partial T}{\partial y} + k \frac{\partial T}{\partial z}.$$

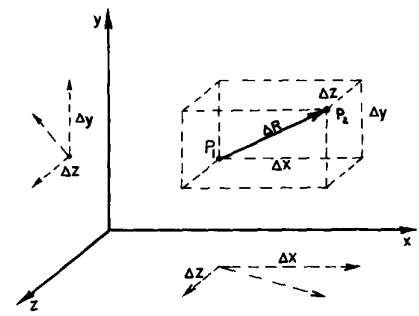


Fig. 2-5. The vector ΔR , whose components are Δx , Δy , and Δz .

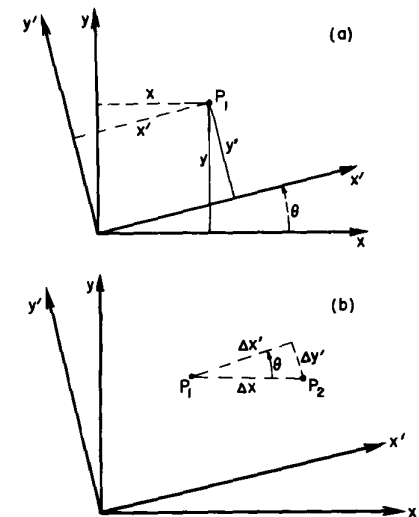


Fig. 2-6. (a) Transformation to a rotated coordinate system. (b) Special case of an interval ΔR parallel to the x -axis.

If we choose some convenient set of axes, we could write $T_1 = T(x, y, z)$ and $T_2 = T(x + \Delta x, y + \Delta y, z + \Delta z)$, where Δx , Δy , and Δz are the components of the vector ΔR (Fig. 2-5). Remembering Eq. (2.7), we can write

$$\Delta T = \frac{\partial T}{\partial x} \Delta x + \frac{\partial T}{\partial y} \Delta y + \frac{\partial T}{\partial z} \Delta z. \quad (2.13)$$

The left side of Eq. (2.13) is a scalar. The right side is the sum of three products with Δx , Δy , and Δz , which are the components of a vector. It follows that the three numbers

$$\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}$$

are also the x -, y -, and z -components of a vector. We write this new vector with the symbol ∇T . The symbol ∇ (called "del") is an upside-down Δ , and is supposed to remind us of differentiation. People read ∇T in various ways: "del- T ," or "gradient of T ," or "grad T ;"

$$\text{grad } T = \nabla T = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right)^*. \quad (2.14)$$

Using this notation, we can rewrite Eq. (2.13) in the more compact form

$$\Delta T = \nabla T \cdot \Delta R. \quad (2.15)$$

In words, this equation says that the difference in temperature between two nearby points is the dot product of the gradient of T and the vector displacement between the points. The form of Eq. (2.15) also illustrates clearly our proof above that ∇T is indeed a vector.

Perhaps you are still not convinced? Let's prove it in a different way. (Although if you look carefully, you may be able to see that it's really the same proof in a longer-winded form!) We shall show that the components of ∇T transform in just the same way that components of R do. If they do, ∇T is a vector according to our original definition of a vector in Chapter 11 of Vol. I. We take a new coordinate system x' , y' , z' , and in this new system we calculate $\partial T/\partial x'$, $\partial T/\partial y'$, and $\partial T/\partial z'$. To make things a little simpler, we let $z = z'$, so that we can forget about the z -coordinate. (You can check out the more general case for yourself.)

We take an $x'y'$ -system rotated an angle θ with respect to the xy -system, as in Fig. 2-6(a). For a point (x, y) the coordinates in the prime system are

$$x' = x \cos \theta + y \sin \theta, \quad (2.16)$$

$$y' = -x \sin \theta + y \cos \theta. \quad (2.17)$$

Or, solving for x and y ,

$$x = x' \cos \theta - y' \sin \theta, \quad (2.18)$$

$$y = x' \sin \theta + y' \cos \theta. \quad (2.19)$$

If any pair of numbers transforms with these equations in the same way that x and y do, they are the components of a vector.

Now let's look at the difference in temperature between the two nearby points P_1 and P_2 , chosen as in Fig. 2-6(b). If we calculate with the x - and y -coordinates, we would write

$$\Delta T = \frac{\partial T}{\partial x} \Delta x \quad (2.20)$$

—since Δy is zero.

* In our notation, the expression (a, b, c) represents a vector with components a , b , and c . If you like to use the unit vectors i, j , and k , you may write

$$\nabla T = i \frac{\partial T}{\partial x} + j \frac{\partial T}{\partial y} + k \frac{\partial T}{\partial z}.$$

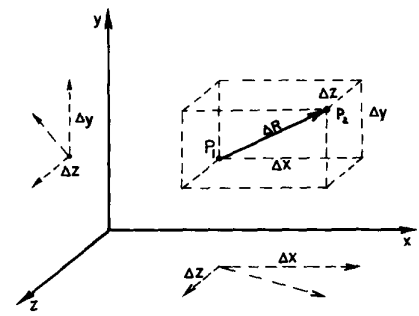


Fig. 2-5. The vector ΔR , whose components are Δx , Δy , and Δz .

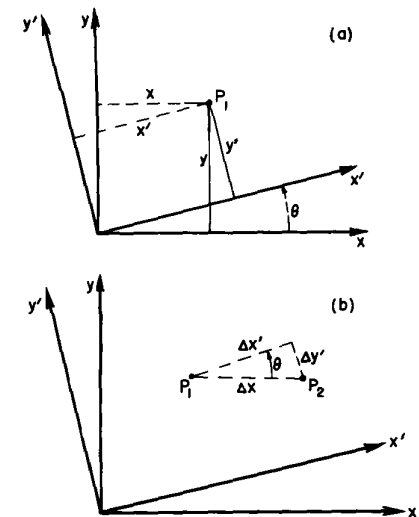


Fig. 2-6. (a) Transformation to a rotated coordinate system. (b) Special case of an interval ΔR parallel to the x -axis.

What would a computation in the prime system give? We would have written

$$\Delta T = \frac{\partial T}{\partial x'} \Delta x' + \frac{\partial T}{\partial y'} \Delta y'. \quad (2.21)$$

Looking at Fig. 2-6(b), we see that

$$\Delta x' = \Delta x \cos \theta \quad (2.22)$$

and

$$\Delta y' = -\Delta x \sin \theta, \quad (2.23)$$

since Δy is negative when Δx is positive. Substituting these in Eq. (2.21), we find that

$$\Delta T = \frac{\partial T}{\partial x'} \Delta x \cos \theta - \frac{\partial T}{\partial y'} \Delta x \sin \theta \quad (2.24)$$

$$= \left(\frac{\partial T}{\partial x'} \cos \theta - \frac{\partial T}{\partial y'} \sin \theta \right) \Delta x. \quad (2.25)$$

Comparing Eq. (2.25) with (2.20), we see that

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial x'} \cos \theta - \frac{\partial T}{\partial y'} \sin \theta. \quad (2.26)$$

This equation says that $\partial T/\partial x$ is obtained from $\partial T/\partial x'$ and $\partial T/\partial y'$, just as x is obtained from x' and y' in Eq. (2.18). So $\partial T/\partial x$ is the x -component of a vector. The same kind of arguments would show that $\partial T/\partial y$ and $\partial T/\partial z$ are y - and z -components. So ∇T is definitely a vector. It is a vector field derived from the scalar field T .

2-4 The operator ∇

Now we can do something that is extremely amusing and ingenious—and characteristic of the things that make mathematics beautiful. The argument that $\text{grad } T$, or ∇T , is a vector did not depend upon *what* scalar field we were differentiating. All the arguments would go the same if T were replaced by *any scalar field*. Since the transformation equations are the same no matter what we differentiate, we could just as well omit the T and replace Eq. (2.26) by the operator equation

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'} \cos \theta - \frac{\partial}{\partial y'} \sin \theta. \quad (2.27)$$

We leave the operators, as Jeans said, “hungry for something to differentiate.”

Since the differential operators themselves transform as the components of a vector should, we can call them components of a *vector operator*. We can write

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right), \quad (2.28)$$

which means, of course,

$$\nabla_x = \frac{\partial}{\partial x}, \quad \nabla_y = \frac{\partial}{\partial y}, \quad \nabla_z = \frac{\partial}{\partial z}. \quad (2.29)$$

We have abstracted the gradient away from the T —that is the wonderful idea.

You must always remember, of course, that ∇ is an operator. Alone, it means nothing. If ∇ by itself means nothing, what does it mean if we multiply it by a scalar—say T —to get the product $T\nabla$? (One can always multiply a vector by a scalar.) It still does not mean anything. Its x -component is

$$T \frac{\partial}{\partial x}, \quad (2.30)$$

which is not a number, but is still some kind of operator. However, according to the algebra of vectors we would still call $T\nabla$ a vector.

Now let's multiply ∇ by a scalar on the other side, so that we have the product (∇T) . In ordinary algebra

$$TA = AT, \quad (2.31)$$

but we have to remember that operator algebra is a little different from ordinary vector algebra. With operators we must always keep the sequence right, so that the operations make proper sense. You will have no difficulty if you just remember that the operator ∇ obeys the same convention as the derivative notation. What is to be differentiated must be placed on the right of the ∇ . The order is important.

Keeping in mind this problem of order, we understand that $T\nabla$ is an operator, but the product ∇T is no longer a hungry operator; the operator is completely satisfied. It is indeed a physical vector having a meaning. It represents the spatial rate of change of T . The x -component of ∇T is how fast T changes in the x -direction. What is the direction of the vector ∇T ? We know that the rate of change of T in any direction is the component of ∇T in that direction (see Eq. 2.15). It follows that the direction of ∇T is that in which it has the largest possible component—in other words, the direction in which T changes the fastest. The gradient of T has the direction of the steepest uphill slope (in T).

2-5 Operations with ∇

Can we do any other algebra with the vector operator ∇ ? Let us try combining it with a vector. We can combine two vectors by making a dot product. We could make the products

$$(\text{a vector}) \cdot \nabla, \quad \text{or} \quad \nabla \cdot (\text{a vector}).$$

The first one doesn't mean anything yet, because it is still an operator. What it might ultimately mean would depend on what it is made to operate on. The second product is some scalar field. ($A \cdot B$ is always a scalar.)

Let's try the dot product of ∇ with a vector field we know, say h . We write out the components:

$$\nabla \cdot h = \nabla_x h_x + \nabla_y h_y + \nabla_z h_z \quad (2.32)$$

or

$$\nabla \cdot h = \frac{\partial h_x}{\partial x} + \frac{\partial h_y}{\partial y} + \frac{\partial h_z}{\partial z}. \quad (2.33)$$

The sum is invariant under a coordinate transformation. If we were to choose a different system (indicated by primes), we would have*

$$\nabla' \cdot h = \frac{\partial h_{x'}}{\partial x'} + \frac{\partial h_{y'}}{\partial y'} + \frac{\partial h_{z'}}{\partial z'}, \quad (2.34)$$

which is the *same* number as would be gotten from Eq. (2.33), even though it looks different. That is,

$$\nabla' \cdot h = \nabla \cdot h \quad (2.35)$$

for every point in space. So $\nabla \cdot h$ is a scalar field, which must represent some physical quantity. You should realize that the combination of derivatives in $\nabla \cdot h$ is rather special. There are all sorts of other combinations like $\partial h_y / \partial x$, which are neither scalars nor components of vectors.

The scalar quantity $\nabla \cdot (\text{a vector})$ is extremely useful in physics. It has been given the name the *divergence*. For example,

$$\nabla \cdot h = \text{div } h = \text{"divergence of } h\text{"} \quad (2.36)$$

As we did for ∇T , we can ascribe a physical significance to $\nabla \cdot h$. We shall, however, postpone that until later.

* We think of h as a *physical* quantity that depends on position in space, and not strictly as a mathematical function of three variables. When h is "differentiated" with respect to x , y , and z , or with respect to x' , y' , and z' , the mathematical expression for h must first be expressed as a function of the appropriate variables.

First, we wish to see what else we can cook up with the vector operator ∇ . What about a cross product? We must expect that

$$\nabla \times \mathbf{h} = \text{a vector.} \quad (2.37)$$

It is a vector whose components we can write by the usual rule for cross products (see Eq. 2.2):

$$(\nabla \times \mathbf{h})_z = \nabla_x h_y - \nabla_y h_x = \frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y}. \quad (2.38)$$

Similarly,

$$(\nabla \times \mathbf{h})_x = \nabla_y h_z - \nabla_z h_y = \frac{\partial h_z}{\partial y} - \frac{\partial h_y}{\partial z} \quad (2.39)$$

and

$$(\nabla \times \mathbf{h})_y = \nabla_z h_x - \nabla_x h_z = \frac{\partial h_x}{\partial z} - \frac{\partial h_z}{\partial x}. \quad (2.40)$$

The combination $\nabla \times \mathbf{h}$ is called “the *curl* of \mathbf{h} .” The reason for the name and the physical meaning of the combination will be discussed later.

Summarizing, we have three kinds of combinations with ∇ :

$$\begin{aligned} \nabla T &= \text{grad } T = \text{a vector,} \\ \nabla \cdot \mathbf{h} &= \text{div } \mathbf{h} = \text{a scalar,} \\ \nabla \times \mathbf{h} &= \text{curl } \mathbf{h} = \text{a vector.} \end{aligned}$$

Using these combinations, we can write about the spatial variations of fields in a convenient way—in a way that is general, in that it doesn’t depend on any particular set of axes.

As an example of the use of our vector differential operator ∇ , we write a set of vector equations which contain the same laws of electromagnetism that we gave in words in Chapter 1. They are called Maxwell’s equations.

Maxwell’s Equations

$$\begin{aligned} (1) \quad \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ (2) \quad \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ (3) \quad \nabla \cdot \mathbf{B} &= 0 \\ (4) \quad c^2 \nabla \times \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t} + \frac{\mathbf{j}}{\epsilon_0} \end{aligned} \quad (2.41)$$

where ρ (rho), the “electric charge density,” is the amount of charge per unit volume, and \mathbf{j} , the “electric current density,” is the rate at which charge flows through a unit area per second. These four equations contain the complete classical theory of the electromagnetic field. You see what an elegantly simple form we can get with our new notation!

2-6 The differential equation of heat flow

Let us give another example of a law of physics written in vector notation. The law is not a precise one, but for many metals and a number of other substances that conduct heat it is quite accurate. You know that if you take a slab of material and heat one face to temperature T_2 and cool the other to a different temperature T_1 , the heat will flow through the material from T_2 to T_1 [Fig. 2-7(a)]. The heat flow is proportional to the area A of the faces, and to the temperature difference. It is also inversely proportional to d , the distance between the plates. (For a given temperature difference, the thinner the slab the greater the heat flow.) Letting J be the thermal energy that passes per unit time through the slab, we write

$$J = \kappa(T_2 - T_1) \frac{A}{d}. \quad (2.42)$$

The constant of proportionality κ (kappa) is called the *thermal conductivity*.

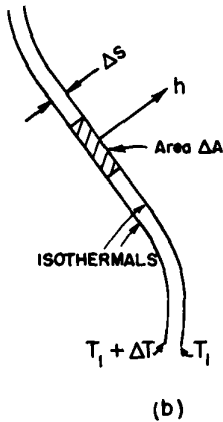
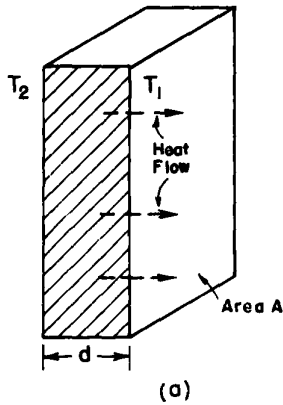


Fig. 2-7. (a) Heat flow through a slab. (b) An infinitesimal slab parallel to an isothermal surface in a large block.

What will happen in a more complicated case? Say in an odd-shaped block of material in which the temperature varies in peculiar ways? Suppose we look at a tiny piece of the block and imagine a slab like that of Fig. 2-7(a) on a miniature scale. We orient the faces parallel to the isothermal surfaces, as in Fig. 2-7(b), so that Eq. (2.42) is correct for the small slab.

If the area of the small slab is ΔA , the heat flow per unit time is

$$\Delta J = \kappa \Delta T \frac{\Delta A}{\Delta s}, \quad (2.43)$$

where Δs is the thickness of the slab. Now $\Delta J/\Delta A$ we have defined earlier as the magnitude of \mathbf{h} , whose direction is the heat flow. The heat flow will be from $T_1 + \Delta T$ toward T_1 , and so it will be perpendicular to the isotherms, as drawn in Fig. 2-7(b). Also, $\Delta T/\Delta s$ is just the rate of change of T with position. And since the position change is perpendicular to the isotherms, our $\Delta T/\Delta s$ is the maximum rate of change. It is, therefore, just the magnitude of ∇T . Now since the direction of ∇T is opposite to that of \mathbf{h} , we can write (2.43) as a vector equation:

$$\mathbf{h} = -\kappa \nabla T. \quad (2.44)$$

(The minus sign is necessary because heat flows “downhill” in temperature.) Equation (2.44) is the differential equation of heat conduction in bulk materials. You see that it is a proper vector equation. Each side is a vector if κ is just a number. It is the generalization to arbitrary cases of the special relation (2.42) for rectangular slabs. Later we should learn to write all sorts of elementary physics relations like (2.42) in the more sophisticated vector notation. This notation is useful not only because it makes the equations *look* simpler. It also shows most clearly the *physical content* of the equations without reference to any arbitrarily chosen coordinate system.

2-7 Second derivatives of vector fields

So far we have had only first derivatives. Why not second derivatives? We could have several combinations:

$$\begin{aligned} \text{(a)} \quad & \nabla \cdot (\nabla T) \\ \text{(b)} \quad & \nabla \times (\nabla T) \\ \text{(c)} \quad & \nabla(\nabla \cdot \mathbf{h}) \\ \text{(d)} \quad & \nabla \cdot (\nabla \times \mathbf{h}) \\ \text{(e)} \quad & \nabla \times (\nabla \times \mathbf{h}) \end{aligned} \quad (2.45)$$

You can check that these are all the possible combinations.

Let's look first at the second one, (b). It has the same form as

$$\mathbf{A} \times (\mathbf{A}T) = (\mathbf{A} \times \mathbf{A})T = 0,$$

since $\mathbf{A} \times \mathbf{A}$ is always zero. So we should have

$$\text{curl}(\text{grad } T) = \nabla \times (\nabla T) = 0. \quad (2.46)$$

We can see how this equation comes about if we go through once with the components:

$$\begin{aligned} [\nabla \times (\nabla T)]_z &= \nabla_x(\nabla T)_y - \nabla_y(\nabla T)_x \\ &= \frac{\partial}{\partial x} \left(\frac{\partial T}{\partial y} \right) - \frac{\partial}{\partial y} \left(\frac{\partial T}{\partial x} \right), \end{aligned} \quad (2.47)$$

which is zero (by Eq. 2.8). It goes the same for the other components. So $\nabla \times (\nabla T) = 0$, for any temperature distribution—in fact, for *any* scalar function.

Now let us take another example. Let us see whether we can find another zero. The dot product of a vector with a cross product which contains that vector is zero:

$$\mathbf{A} \cdot (\mathbf{A} \times \mathbf{B}) = 0. \quad (2.48)$$

because $\mathbf{A} \times \mathbf{B}$ is perpendicular to \mathbf{A} , and so has no components in the direction \mathbf{A} . The same combination appears in (d) of (2.45), so we have

$$\nabla \cdot (\nabla \times \mathbf{h}) = \text{div}(\text{curl } \mathbf{h}) = 0. \quad (2.49)$$

Again, it is easy to show that it is zero by carrying through the operations with components.

Now we are going to state two mathematical theorems that we will not prove. They are very interesting and useful theorems for physicists to know.

In a physical problem we frequently find that the curl of some quantity—say of the vector field \mathbf{A} —is zero. Now we have seen (Eq. 2.46) that the curl of a gradient is zero, which is easy to remember because of the way the vectors work. It could certainly be, then, that \mathbf{A} is the gradient of some quantity, because then its curl would necessarily be zero. The interesting theorem is that if the curl \mathbf{A} is zero, then \mathbf{A} is *always* the gradient of *something*—there is some scalar field ψ (psi) such that \mathbf{A} is equal to $\text{grad } \psi$. In other words, we have the

THEOREM:

$$\begin{array}{ll} \text{If} & \nabla \times \mathbf{A} = 0 \\ \text{there is a} & \psi \\ \text{such that} & \mathbf{A} = \nabla\psi. \end{array} \quad (2.50)$$

There is a similar theorem if the divergence of \mathbf{A} is zero. We have seen in Eq. (2.49) that the divergence of a curl of something is always zero. If you come across a vector field \mathbf{D} for which $\text{div } \mathbf{D}$ is zero, then you can conclude that \mathbf{D} is the curl of some vector field \mathbf{C} .

THEOREM:

$$\begin{array}{ll} \text{If} & \nabla \cdot \mathbf{D} = 0 \\ \text{there is a} & \mathbf{C} \\ \text{such that} & \mathbf{D} = \nabla \times \mathbf{C}. \end{array} \quad (2.51)$$

In looking at the possible combinations of two ∇ operators, we have found that two of them always give zero. Now we look at the ones that are *not* zero. Take the combination $\nabla \cdot (\nabla T)$, which was first on our list. It is not, in general, zero. We write out the components:

$$\nabla T = \nabla_x T + \nabla_y T + \nabla_z T.$$

Then

$$\begin{aligned} \nabla \cdot (\nabla T) &= \nabla_x(\nabla_x T) + \nabla_y(\nabla_y T) + \nabla_z(\nabla_z T) \\ &= \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}, \end{aligned} \quad (2.52)$$

which would, in general, come out to be some number. It is a scalar field.

You see that we do not need to keep the parentheses, but can write, without any chance of confusion,

$$\nabla \cdot (\nabla T) = \nabla \cdot \nabla T = (\nabla \cdot \nabla)T = \nabla^2 T. \quad (2.53)$$

We look at ∇^2 as a new operator. It is a scalar operator. Because it appears often in physics, it has been given a special name—the *Laplacian*.

$$\text{Laplacian} = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (2.54)$$

Since the Laplacian is a scalar operator, we may operate with it on a vector—by which we mean the same operation on each component in rectangular coordinates:

$$\nabla^2 \mathbf{h} = (\nabla^2 h_x, \nabla^2 h_y, \nabla^2 h_z).$$

Let's look at one more possibility: $\nabla \times (\nabla \times \mathbf{h})$, which was (e) in the list (2.45). Now the curl of the curl can be written differently if we use the vector equality (2.6):

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \quad (2.55)$$

In order to use this formula, we should replace \mathbf{A} and \mathbf{B} by the operator ∇ and put $\mathbf{C} = \mathbf{h}$. If we do that, we get

$$\nabla \times (\nabla \times \mathbf{h}) = \nabla(\nabla \cdot \mathbf{h}) - \mathbf{h}(\nabla \cdot \nabla) \dots ???$$

Wait a minute! Something is wrong. The first two terms are vectors all right (the operators are satisfied), but the last term doesn't come out to anything. It's still an operator. The trouble is that we haven't been careful enough about keeping the order of our terms straight. If you look again at Eq. (2.55), however, you see that we could equally well have written it as

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}. \quad (2.56)$$

The order of terms looks better. Now let's make our substitution in (2.56). We get

$$\nabla \times (\nabla \times \mathbf{h}) = \nabla(\nabla \cdot \mathbf{h}) - (\nabla \cdot \nabla)\mathbf{h}. \quad (2.57)$$

This form looks all right. It is, in fact, correct, as you can verify by computing the components. The last term is the Laplacian, so we can equally well write

$$\nabla \times (\nabla \times \mathbf{h}) = \nabla(\nabla \cdot \mathbf{h}) - \nabla^2 \mathbf{h}. \quad (2.58)$$

We have had something to say about all of the combinations in our list of double ∇ 's, except for (c), $\nabla(\nabla \cdot \mathbf{h})$. It is a possible vector field, but there is nothing special to say about it. It's just some vector field which may occasionally come up.

It will be convenient to have a table of our conclusions:

- (a) $\nabla \cdot (\nabla T) = \nabla^2 T =$ a scalar field
- (b) $\nabla \times (\nabla T) = 0$
- (c) $\nabla(\nabla \cdot \mathbf{h}) =$ a vector field
- (d) $\nabla \cdot (\nabla \times \mathbf{h}) = 0$
- (e) $\nabla \times (\nabla \times \mathbf{h}) = \nabla(\nabla \cdot \mathbf{h}) - \nabla^2 \mathbf{h}$
- (f) $(\nabla \cdot \nabla)\mathbf{h} = \nabla^2 \mathbf{h} =$ a vector field

You may notice that we haven't tried to invent a new vector operator ($\nabla \times \nabla$). Do you see why?

2-8 Pitfalls

We have been applying our knowledge of ordinary vector algebra to the algebra of the operator ∇ . We have to be careful, though, because it is possible to go astray. There are two pitfalls which we will mention, although they will not come up in this course. What would you say about the following expression, that involves the two scalar functions ψ and ϕ (phi):

$$(\nabla\psi) \times (\nabla\phi)?$$

You might want to say: it must be zero because it's just like

$$(Aa) \times (Ab),$$

which is zero because the cross product of two *equal* vectors $A \times A$ is always zero. But in our example the two operators ∇ are not equal! The first one operates on one function, ψ ; the other operates on a different function, ϕ . So although we represent them by the same symbol ∇ , they must be considered as different operators. Clearly, the direction of $\nabla\psi$ depends on the function ψ , so it is not likely to be parallel to $\nabla\phi$.

$$(\nabla\psi) \times (\nabla\phi) \neq 0 \quad (\text{generally}).$$

Fortunately, we won't have to use such expressions. (What we have said doesn't change the fact that $\nabla \times \nabla\psi = 0$ for any scalar field, because here both ∇ 's operate on the same function.)

Pitfall number two (which, again, we need not get into in our course) is the following: The rules that we have outlined here are simple and nice when we use rectangular coordinates. For example, if we have $\nabla^2\mathbf{h}$ and we want the x -component, it is

$$(\nabla^2\mathbf{h})_x = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) h_x = \nabla^2 h_x. \quad (2.60)$$

The same expression would *not* work if we were to ask for the *radial* component of $\nabla^2\mathbf{h}$. The radial component of $\nabla^2\mathbf{h}$ is not equal to $\nabla^2 h_r$. The reason is that when we are dealing with the algebra of vectors, the directions of the vectors are all quite definite. But when we are dealing with vector fields, their directions are different at different places. If we try to describe a vector field in, say, polar coordinates, what we call the "radial" direction varies from point to point. So we can get into a lot of trouble when we start to differentiate the components. For example, even for a *constant* vector field, the radial component changes from point to point.

It is usually safest and simplest just to stick to rectangular coordinates and avoid trouble, but there is one exception worth mentioning: Since the Laplacian ∇^2 , is a scalar, we can write it in any coordinate system we want to (for example, in polar coordinates). But since it is a differential operator, we should use it only on vectors whose components are in a fixed direction—that means rectangular coordinates. So we shall express all of our vector fields in terms of their x -, y -, and z -components when we write our vector differential equations out in components.