Mean Value Theorems

THEOREM 2

Mean Value Theorem

Let $f(x, y, z)$ be continuous and have continuous first partial derivatives in a domain D in xyz-space. Let P_0: (x_0, y_0, z_0) and P: $(x_0 + h, y_0 + k, z_0 + l)$ be points in D such that the straight line segment P_0P joining these points lies entirely in D. Then

$$f(x_0 + h, y_0 + k, z_0 + l) - f(x_0, y_0, z_0) = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + l\frac{\partial f}{\partial z}, \tag{7}$$

the partial derivatives being evaluated at a suitable point of that segment.

Special Cases

For a function $f(x, y)$ of two variables (satisfying assumptions as in the theorem), formula (7) reduces to (Fig. 214)

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}, \tag{8}$$

and, for a function $f(x)$ of a single variable, (7) becomes

$$f(x_0 + h) - f(x_0) = h\frac{\partial f}{\partial x}, \tag{9}$$

where in (9), the domain D is a segment of the x-axis and the derivative is taken at a suitable point between x_0 and $x_0 + h$.

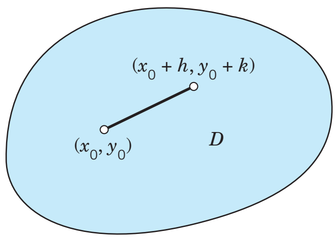

Fig. 214. Mean value theorem for a function of two variables [Formula (8)]

9.7 Gradient of a Scalar Field. Directional Derivative

We shall see that *some* of the vector fields that occur in applications—not all of them!—can be obtained from scalar fields. Using scalar fields instead of vector fields is of a considerable advantage because scalar fields are easier to use than vector fields. It is the "gradient" that allows us to obtain vector fields from scalar fields, and thus the gradient is of great practical importance to the engineer.

DEFINITION 1

Gradient

The setting is that we are given a scalar function $f(x, y, z)$ that is defined and differentiable in a domain in 3-space with Cartesian coordinates x, y, z. We denote the **gradient** of that function by grad f or ∇f (read **nabla** f). Then the qradient of $f(x, y, z)$ is defined as the vector function

(1)
$$\operatorname{grad} f = \nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right] = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}.$$

Remarks. For a definition of the gradient in curvilinear coordinates, see App. 3.4. As a quick example, if $f(x, y, z) = 2y^3 + 4xz + 3x$, then grad $f = [4z + 3, 6y^2, 4x]$. Furthermore, we will show later in this section that (1) actually does define a vector.

The notation ∇f is suggested by the *differential operator* ∇ (read *nabla*) defined by

(1*)
$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}.$$

Gradients are useful in several ways, notably in giving the rate of change of $f(x, y, z)$ in any direction in space, in obtaining surface normal vectors, and in deriving vector fields from scalar fields, as we are going to show in this section.

Directional Derivative

From calculus we know that the partial derivatives in (1) give the rates of change of $f(x, y, z)$ in the directions of the three coordinate axes. It seems natural to extend this and ask for the rate of change of f *in an arbitrary direction* in space. This leads to the following concept.

DEFINITION 2

Directional Derivative

The directional derivative $D_{\mathbf{b}} f$ or df/ds of a function $f(x, y, z)$ at a point P in the direction of a vector $\mathbf{b}$ is defined by (see Fig. 215)

(2)
$$D_{\mathbf{b}} f = \frac{df}{ds} = \lim_{s \to 0} \frac{f(Q) - f(P)}{s}.$$

Here Q is a variable point on the straight line L in the direction of $\mathbf{b}$, and $|s|$ is the distance between P and Q. Also, $s > 0$ if Q lies in the direction of $\mathbf{b}$ (as in Fig. 215), $s < 0$ if Q lies in the direction of $-\mathbf{b}$, and $s = 0$ if $Q = P$.

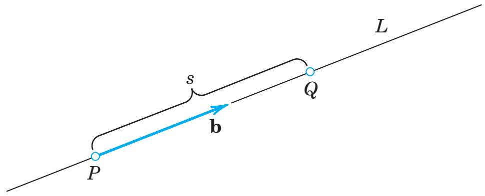

Fig. 215. Directional derivative

The next idea is to use Cartesian xyz-coordinates and for $\mathbf{b}$ a unit vector. Then the line L is given by

$$\mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k} = \mathbf{p}_0 + s\mathbf{b} \qquad (|\mathbf{b}| = 1) \tag{3}$$

where $\mathbf{p}_0$ the position vector of P. Equation (2) now shows that $D_{\mathbf{b}}f = df/ds$ is the derivative of the function $f(x(s), y(s), z(s))$ with respect to the arc length s of L. Hence, assuming that f has continuous partial derivatives and applying the chain rule [formula (3) in the previous section], we obtain

$$D_{\mathbf{b}}f = \frac{df}{ds} = \frac{\partial f}{\partial x}x' + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial z}z' \tag{4}$$

where primes denote derivatives with respect to s (which are taken at $s = 0$). But here, differentiating (3) gives $\mathbf{r}' = x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k} = \mathbf{b}$. Hence (4) is simply the inner product of grad f and $\mathbf{b}$ [see (2), Sec. 9.2]; that is,

$$\boxed{D_{\mathbf{b}}f = \frac{df}{ds} = \mathbf{b} \bullet \operatorname{grad} f} \qquad (|\mathbf{b}| = 1). \tag{5}$$

ATTENTION! If the direction is given by a vector $\mathbf{a}$ of any length ($\neq 0$), then

$$\boxed{D_{\mathbf{a}}f = \frac{df}{ds} = \frac{1}{|\mathbf{a}|}\mathbf{a} \bullet \operatorname{grad} f.} \tag{5*}$$

EXAMPLE 1 Gradient. Directional Derivative

Find the directional derivative of $f(x, y, z) = 2x^2 + 3y^2 + z^2$ at P: (2, 1, 3) in the direction of $\mathbf{a} = [1, 0, -2]$.

Solution. grad $f = [4x, 6y, 2z]$ gives at P the vector grad $f(P) = [8, 6, 6]$. From this and (5*) we obtain, since $|\mathbf{a}| = \sqrt{5}$,

$$D_{\mathbf{a}}f(P) = \frac{1}{\sqrt{5}}[1, 0, -2] \bullet [8, 6, 6] = \frac{1}{\sqrt{5}}(8 + 0 - 12) = -\frac{4}{\sqrt{5}} = -1.789.$$

The minus sign indicates that at P the function f is decreasing in the direction of $\mathbf{a}$. ■

Gradient Is a Vector. Maximum Increase

Here is a finer point of mathematics that concerns the consistency of our theory: grad f in (1) *looks* like a vector—after all, it has three components! But to prove that it *actually is* a vector, since it is defined in terms of components depending on the Cartesian coordinates, we must show that grad f has a length and direction independent of the choice of those coordinates. See proof of Theorem 1. In contrast, $[\partial f/\partial x, 2\partial f/\partial y, \partial f/\partial z]$ also looks like a vector but does not have a length and direction independent of the choice of Cartesian coordinates.

Incidentally, the direction makes the gradient eminently useful: grad f *points in the direction of maximum increase of* f.

THEOREM 1

Use of Gradient: Direction of Maximum Increase

Let $f(P) = f(x, y, z)$ *be a scalar function having continuous first partial derivatives in some domain* B *in space. Then* grad f *exists in* B *and is a vector, that is, its length and direction are independent of the particular choice of Cartesian coordinates. If* grad $f(P) \neq \mathbf{0}$ *at some point* P, *it has the direction of maximum increase of* f *at* P.

PROOF From (5) and the definition of inner product [(1) in Sec. 9.2] we have

$$D_{\mathbf{b}} f = |\mathbf{b}||\text{grad } f| \cos\gamma = |\text{grad } f| \cos\gamma \tag{6}$$

where γ is the angle between $\mathbf{b}$ and grad f. Now f is a scalar function. Hence its value at a point P depends on P but not on the particular choice of coordinates. The same holds for the arc length s of the line L in Fig. 215, hence also for $D_{\mathbf{b}} f$. Now (6) shows that $D_{\mathbf{b}} f$ is maximum when $\cos\gamma = 1$, $\gamma = 0$, and then $D_{\mathbf{b}} f = |\text{grad } f|$. It follows that the length and direction of grad f are independent of the choice of coordinates. Since $\gamma = 0$ if and only if $\mathbf{b}$ has the direction of grad f, the latter is the direction of maximum increase of f at P, provided grad $f \neq \mathbf{0}$ at P. Make sure that you understood the proof to get a good feel for mathematics.

Gradient as Surface Normal Vector

Gradients have an important application in connection with surfaces, namely, as surface normal vectors, as follows. Let S be a surface represented by $f(x, y, z) = c = \text{const}$, where f is differentiable. Such a surface is called a **level surface** of f, and for different c we get different level surfaces. Now let C be a curve on S through a point P of S. As a curve in space, C has a representation $\mathbf{r}(t) = [x(t), y(t), z(t)]$. For C to lie on the surface S, the components of $\mathbf{r}(t)$ must satisfy $f(x, y, z) = c$, that is,

$$f(x(t), y(t), z(t)) = c. \tag{7}$$

Now a tangent vector of C is $\mathbf{r}'(t) = [x'(t), y'(t), z'(t)]$. And the tangent vectors of all curves on S passing through P will generally form a plane, called the **tangent plane** of S at P. (Exceptions occur at edges or cusps of S, for instance, at the apex of the cone in Fig. 217.) The normal of this plane (the straight line through P perpendicular to the tangent plane) is called the **surface normal** to S at P. A vector in the direction of the surface normal is called a **surface normal vector** of S at P. We can obtain such a vector quite simply by differentiating (7) with respect to t. By the chain rule,

$$\frac{\partial f}{\partial x}x' + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial z}z' = (\operatorname{grad} f) \bullet \mathbf{r}' = 0.$$

Hence grad f is orthogonal to all the vectors $\mathbf{r}'$ in the tangent plane, so that it is a normal vector of S at P. Our result is as follows (see Fig. 216).

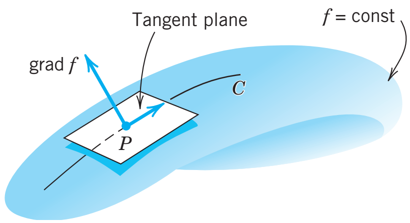

Fig. 216. Gradient as surface normal vector

THEOREM 2

Gradient as Surface Normal Vector

Let f be a differentiable scalar function in space. Let $f(x, y, z) = c = const$ represent a surface S. Then if the gradient of f at a point P of S is not the zero vector, it is a normal vector of S at P.

EXAMPLE 2

Gradient as Surface Normal Vector. Cone

Find a unit normal vector $\mathbf{n}$ of the cone of revolution $z^2 = 4(x^2 + y^2)$ at the point P: (1, 0, 2).

Solution. The cone is the level surface $f = 0$ of $f(x, y, z) = 4(x^2 + y^2) - z^2$. Thus (Fig. 217)

$$\operatorname{grad} f = [8x, \quad 8y, \quad -2z], \qquad \operatorname{grad} f(P) = [8, \quad 0, \quad -4]$$

$$\mathbf{n} = \frac{1}{|\operatorname{grad} f(P)|}\operatorname{grad} f(P) = \left[\frac{2}{\sqrt{5}}, \quad 0, \quad -\frac{1}{\sqrt{5}}\right].$$

$\mathbf{n}$ points downward since it has a negative z-component. The other unit normal vector of the cone at P is $-\mathbf{n}$. ■

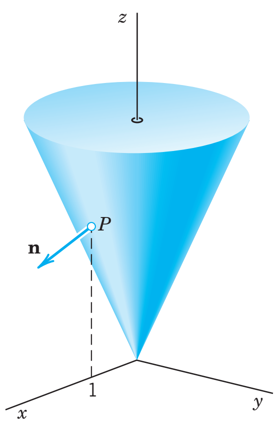

Fig. 217. Cone and unit normal vector **n**

Vector Fields That Are Gradients of Scalar Fields ("Potentials")

At the beginning of this section we mentioned that some vector fields have the advantage that they can be obtained from scalar fields, which can be worked with more easily. Such a vector field is given by a vector function $\mathbf{v}(P)$, which is obtained as the gradient of a scalar function, say, $\mathbf{v}(P) = \text{grad} f(P)$. The function $f(P)$ is called a *potential function* or a **potential** of $\mathbf{v}(P)$. Such a $\mathbf{v}(P)$ and the corresponding vector field are called **conservative** because in such a vector field, energy is conserved; that is, no energy is lost (or gained) in displacing a body (or a charge in the case of an electrical field) from a point P to another point in the field and back to P. We show this in Sec. 10.2.

Conservative fields play a central role in physics and engineering. A basic application concerns the gravitational force (see Example 3 in Sec. 9.4) and we show that it has a potential which satisfies Laplace's equation, the most important partial differential equation in physics and its applications.

THEOREM 3

Gravitational Field. Laplace's Equation

The force of attraction

$$\mathbf{p} = -\frac{c}{r^3}\mathbf{r} = -c\left[\frac{x-x_0}{r^3}, \frac{y-y_0}{r^3}, \frac{z-z_0}{r^3}\right] \tag{8}$$

between two particles at points P_0: (x_0, y_0, z_0) and P: (x, y, z) (as given by Newton's law of gravitation) *has the potential* $f(x, y, z) = c/r$, *where* $r\ (> 0)$ *is the distance between* P_0 *and* P.

Thus $\mathbf{p} = \text{grad} f = \text{grad}(c/r)$. *This potential* f *is a solution of* **Laplace's equation**

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0. \tag{9}$$

[$\nabla^2 f$ (read *nabla squared* f) is called the **Laplacian** of f.]

PROOF That distance is $r = ((x - x_0)^2 + (y - y_0)^2 + (z - z_2)^2)^{1/2}$. The key observation now is that for the components of $\mathbf{p} = [p_1, p_2, p_3]$ we obtain by partial differentiation

$$\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = \frac{-2(x-x_0)}{2[(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2]^{3/2}} = -\frac{x-x_0}{r^3} \tag{10a}$$

and similarly

$$\begin{aligned} \frac{\partial}{\partial y}\left(\frac{1}{r}\right) &= -\frac{y-y_0}{r^3}, \\ \frac{\partial}{\partial z}\left(\frac{1}{r}\right) &= -\frac{z-z_0}{r^3}. \end{aligned} \tag{10b}$$

From this we see that, indeed, **p** is the gradient of the scalar function $f = c/r$. The second statement of the theorem follows by partially differentiating (10), that is,

$$\frac{\partial^2}{\partial x^2}\left(\frac{1}{r}\right) = -\frac{1}{r^3} + \frac{3(x - x_0)^2}{r^5},$$

$$\frac{\partial^2}{\partial y^2}\left(\frac{1}{r}\right) = -\frac{1}{r^3} + \frac{3(y - y_0)^2}{r^5},$$

$$\frac{\partial^2}{\partial z^2}\left(\frac{1}{r}\right) = -\frac{1}{r^3} + \frac{3(z - z_0)^2}{r^5},$$

and then adding these three expressions. Their common denominator is r^5. Hence the three terms $-1/r^3$ contribute $-3r^2$ to the numerator, and the three other terms give the sum

$$3(x - x_0)^2 + 3(y - y_0)^2 + 3(z - z_0)^2 = 3r^2,$$

so that the numerator is 0, and we obtain (9). ■

$\nabla^2 f$ is also denoted by Δf. The differential operator

$$\nabla^2 = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{11}$$

(read "nabla squared" or "delta") is called the **Laplace operator**. It can be shown that the field of force produced by any distribution of masses is given by a vector function that is the gradient of a scalar function f, and f satisfies (9) in any region that is free of matter.

The great importance of the Laplace equation also results from the fact that there are other laws in physics that are of the same form as Newton's law of gravitation. For instance, in electrostatics the force of attraction (or repulsion) between two particles of opposite (or like) charge Q_1 and Q_2 is

$$\mathbf{p} = \frac{k}{r^3}\mathbf{r} \qquad \textbf{(Coulomb's law}^6\textbf{)}. \tag{12}$$

Laplace's equation will be discussed in detail in Chaps. 12 and 18.

A method for finding out whether a given vector field has a potential will be explained in Sec. 9.9.

---

[6] CHARLES AUGUSTIN DE COULOMB (1736–1806), French physicist and engineer. Coulomb's law was derived by him from his own very precise measurements.