Chapter 4

Directional Derivative and the Gradient Operator

{chap:4}

Swokowski chapter 12.6

4.1 Vectors: Revision

Before starting the new material in this chapter, we revise some important topics in vectors that you must be familiar with.

Vectors in 3-d space

A *vector* is a quantity which possesses both magnitude and direction. A *scalar* possesses magnitude only.

In the Cartesian coordinate system the unit vectors are $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, all of unit length, directed along the positive x, y and z axes respectively.

Throughout let $\mathbf{a}$=$a_1\mathbf{i}$+$a_2\mathbf{j}$+$a_3\mathbf{k}$ $= (a_1, a_2, a_3)$, $\mathbf{b}$=$b_1\mathbf{i}$+$b_2\mathbf{j}$+$b_3\mathbf{k}$ $= (b_1, b_2, b_3)$ and $\mathbf{c}$=$c_1\mathbf{i}$+$c_2\mathbf{j}$+$c_3\mathbf{k}$ $= (c_1, c_2, c_3)$.

The *magnitude* of $\mathbf{a}$ is: $|\mathbf{a}|$=$\sqrt{a_1^2 + a_2^2 + a_3^2}$.

The *unit vector* in the direction of $\mathbf{a}$ is

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}.$$

Scalar multiplication: $\lambda\mathbf{a} = \lambda a_1\mathbf{i} + \lambda a_2\mathbf{j} + \lambda a_3\mathbf{k}$ (where λ – "*lambda*" is a scalar).

Vector addition: $\mathbf{a}$+$\mathbf{b}$=$(a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}$.

Vectors can be viewed as directed displacements. If $\mathbf{a}$ and $\mathbf{b}$ are displacements then the net result of these two displacements is $\mathbf{a}$+$\mathbf{b}$. Vector addition is illustrated in Figure 4.1.

{fig:vector1}

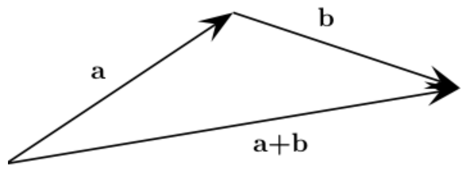

Figure 4.1: Vector addition of two vectors.

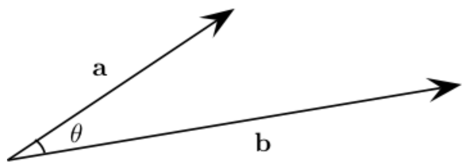

Figure 4.2: Two vectors **a** and **b** and the smaller angle θ are illustrated. This is the angle used in calculating the scalar product $\mathbf{a} \cdot \mathbf{b}$. tor2}

Scalar / dot product

The *scalar* or *dot product* is defined as

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta \tag{4.1}$$ {eq:scalar1}

where θ denotes the smaller angle between the two vectors. The situation is illustrated in Figure 4.2.

Equivalently,

lar2}
$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \ . \tag{4.2}$$

If **a** and **b** are *perpendicular* or *orthogonal*b then **a**.**b**=0 (since $\theta = \pi/2$).

A simple rearrangement gives a formula for the angle between two vectors:

$$\cos\theta = \frac{\mathbf{a}.\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}.$$

Note that $\mathbf{i}.\mathbf{j} = \mathbf{j}.\mathbf{k}$=$\mathbf{k}.\mathbf{i}$=0. *(why?)*

Vector / cross product

The *vector* or *cross product* is defined as

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}$$

where

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

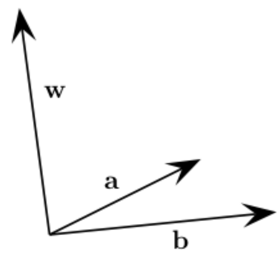

Figure 4.3: The vector product $\mathbf{a} \times \mathbf{b}$ results in a vector $\mathbf{w}$ that is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. {fig:vector3a}

So

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2)\mathbf{i} + (a_3 b_1 - a_1 b_3)\mathbf{j} + (a_1 b_2 - a_2 b_1)\mathbf{k}.$$

Note that if you remember the expression for the $\mathbf{i}$ component, the others can be determined by cyclic rotation of the subscripts.

If $\mathbf{w}$=$\mathbf{a}\times\mathbf{b}$ then $\mathbf{w}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ (see Figure 4.3). *(how would you check this?)* Note that $\mathbf{a}\times\mathbf{b}$=-$\mathbf{b}\times\mathbf{a}$.

It can be easily seen that $\mathbf{i}\times\mathbf{j}$=$\mathbf{k}$, $\mathbf{j}\times\mathbf{k}$=$\mathbf{i}$ and $\mathbf{k}\times\mathbf{i}$=$\mathbf{j}$.

Another way of defining the cross product is $\mathbf{a}\times\mathbf{b}=|\mathbf{a}||\mathbf{b}|\sin\theta\hat{\mathbf{n}}$ where $\hat{\mathbf{n}}$ is the unit vector normal to both $\mathbf{a}$ and $\mathbf{b}$. It follows that $\mathbf{a}\times\mathbf{b}=\mathbf{0}$ if and only if $\mathbf{a}$ is parallel to $\mathbf{b}$. *(why?)*

Triple scalar product

$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} =$ etc

Once the cyclic order $\mathbf{a} \to \mathbf{b} \to \mathbf{c} \to \mathbf{a}$ etc is maintained, $\times$ and . can be interchanged. The brackets are not really necessary because the vector product only produces a vector and the scalar product a scalar. Hence, $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ can only mean that $\mathbf{b} \times \mathbf{c}$ is done first (to produce a vector). Then the scalar product with $\mathbf{a}$ can be done. If one tried to do $\mathbf{a} \cdot \mathbf{b}$ first, you would generate a scalar and then it is not possible to take a vector product with a scalar.

Triple vector product

$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ = ($\mathbf{a}.\mathbf{c}$)$\mathbf{b}$-($\mathbf{a}.\mathbf{b}$)$\mathbf{c}$

Notice that the order of the vectors *is* important here.

Vector equation of a line

Given two points A and B on a line. Let P be any point on the line. Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{r}$ be the position vectors for A, B and P respectively. Then

$$\mathbf{r} = \mathbf{a} + s(\mathbf{b} - \mathbf{a})$$

for suitable choice of the parameter s. $\mathbf{b} - \mathbf{a}$ is a vector joining the points A and B. The vector joining A and P is a multiple of this vector, $s(\mathbf{b} - \mathbf{a})$. See the set up in Figure 4.4.

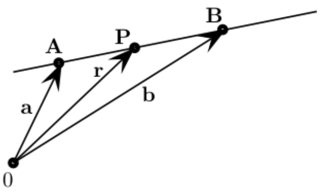

Figure 4.4: The line joins the points A and B. P is a general point lying on the line with position vector $\mathbf{r} = \mathbf{a} + \mathbf{s}(\mathbf{b} - \mathbf{a})$. {fig:vector4}

Cartesian equation of a line

From above **r**=**a**+s(**b**-**a**). If $\mathbf{r}{=}x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, then equating coefficients yields

$$x = a_1 + s(b_1 - a_1), y = a_2 + t(b_2 - a_2), z = a_3 + s(b_3 - a_3),$$

which gives us the Cartesian equation of the line.

Equation of a plane

Let (x_0, y_0, z_0) be a given point on a plane and $\mathbf{n}{=}n_1\mathbf{i}+n_2\mathbf{j}+n_3\mathbf{k}$ the normal to the plane at that point. If (x, y, z) is any point on the plane then

$$(x - x_0)n_1 + (y - y_0)n_2 + (z - z_0)n_3 = 0$$

or equivalently

$$n_1x + n_2y + n_3z = d$$

where

$$d = n_1x_0 + n_2y_0 + n_3z_0.$$

Note Given 3 points on a plane, the normal can be constructed using the cross product. *(how?)*

Notation Vectors are denoted by bold letters in the lecture notes but you must underline letters to indicate that it is a vector. This is important as we frequently use the *same* letters, without the bold or underline, to stand for the magnitude of the vector. So the vector is

$$\mathbf{a} = \underline{a} \qquad \text{this is the vector.}$$

$$|\mathbf{a}| = a \qquad \text{this is the magnitude of the vector } \mathbf{a}$$

Forget to underline a vector and *you will be marked wrong and lose marks*!

4.2 Surfaces

:4.1}

We have already seen that the equation $z = f(x, y)$ defines a *surface* in 3 dimensions. We can write this as

$$z - f(x, y) = 0,$$

or

$$g(x, y, z) = 0, \qquad \text{where } g(x, y, z) = z - f(x, y).$$

The more general equation of a surface is

{eq:4.1a}
$$g(x, y, z) = c, \tag{4.3}$$

where c is a parameter. Each value of c labels one member of the family of surfaces. g has a magnitude but no direction. Thus, g is a *scalar function* of x, y and z.

Example 4.28

Consider $g(x, y, z) = x^2 + y^2 + z^2 = a^2$, where, in the notation above, $c = a^2$ is a constant. This describes the family of concentric spheres centred at the origin and with radius a. An example is shown in Figure 4.5.

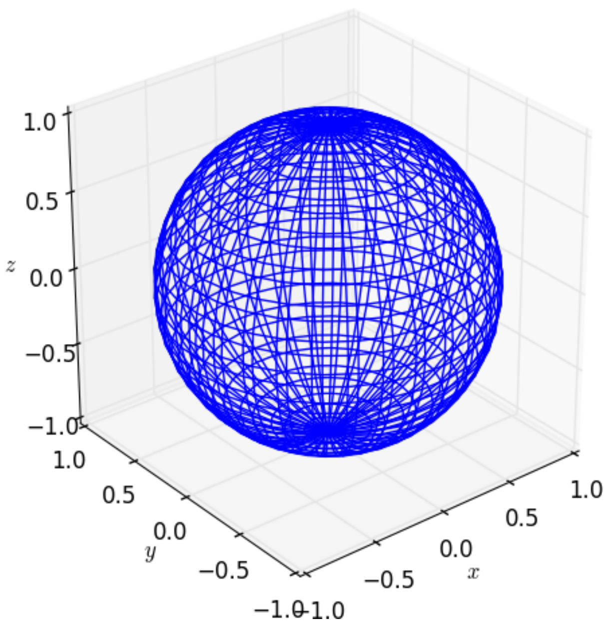

Figure 4.5: The sphere $x^2 + y^2 + z^2 = 1$. {fig:4.11}

Example End

Consider two surfaces (S_1 and S_2) on which g is equal to c_1 and c_2 respectively. This is illustrated in Figure 4.6. Suppose the point P lies on surface S_1 and Q on S_2. At P, $g = c_1$ and at Q $g = c_2$. Thus, the value of g changes from c_1 to c_2 as we move along the path PQ. For general P and Q, we can calculate the *rate of change* of g along the line PQ. This means we calculate the *directional derivative*. Suppose the path is the straight line joining P and Q. Assume that the unit vector $\hat{\mathbf{u}}$, which is parallel to $\vec{PQ}$, has cartesian components

$$\hat{\mathbf{u}} = (l, m, n) \equiv l\mathbf{i} + m\mathbf{j} + n\mathbf{k},$$

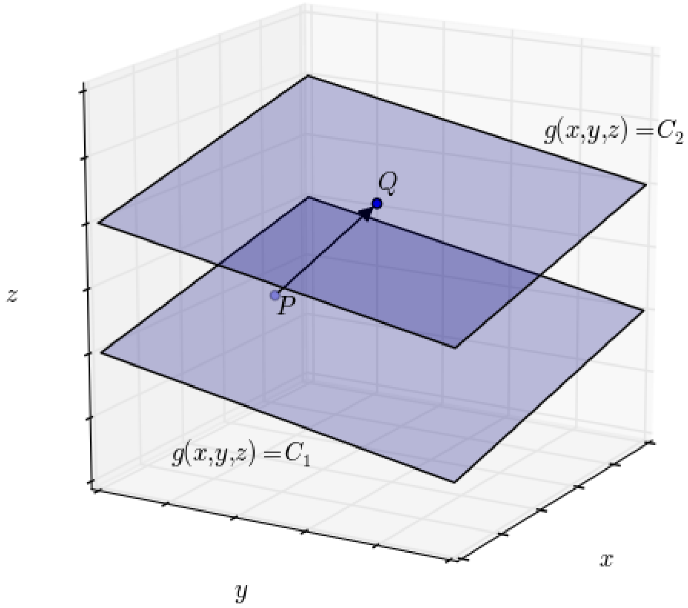

Figure 4.6: Two surfaces labelled by the constants c_1 and c_2. The path between the points P and Q is indicated. {fig:4.2}

and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are the unit vectors along the x, y and z axes. As $\hat{\mathbf{u}}$ is a unit vector $|\hat{\mathbf{u}}| = 1$. This means that

$$|\hat{\mathbf{u}}|^2 = \hat{\mathbf{u}} \cdot \hat{\mathbf{u}} = l^2 + m^2 + n^2 = 1.$$

This result follows directly from the scalar (or dot) product of vectors.

$$\begin{aligned} \mathbf{A} &= (A_x, A_y, A_z) = A_x\mathbf{i} + A_y\mathbf{j} + A_z\mathbf{k}, \\ \mathbf{B} &= (B_x, B_y, B_z) = B_x\mathbf{i} + B_y\mathbf{j} + B_z\mathbf{k}, \end{aligned}$$

then

$$\mathbf{A} \cdot \mathbf{B} = A_xB_x + A_yB_y + A_zB_z.$$

Define the coordinates of P as (x_0, y_0, z_0) and Q as (x, y, z). If Q is a distance s from P in the direction of $\hat{\mathbf{u}}$, the coordinates of Q are

$$\vec{OQ} = \vec{OP} + s\hat{\mathbf{u}},$$

or

{eq:4.1} $$x = x_0 + ls, \qquad y = y_0 + ms, \qquad z = z_0 ns. \tag{4.4}$$

This is written in vector form as

$$\mathbf{r} = \mathbf{r}_0 + s\hat{\mathbf{u}}.$$

As s is varied $(-\infty < s < +\infty)$ then any point of the line may be reached. This is called the *parametric equation* for the line. The coordinates of Q may be written as $(x(s), y(s), z(s))$. Note that the vector $\vec{PQ}$ is

$$\vec{PQ} = s\hat{\mathbf{u}}.$$

Since $\hat{\mathbf{u}}$ is a unit vector, s represents the distance from P to Q.

The variation of g along the line is

$$g(x, y, z) = g(x(s), y(s), z(s)),$$

on using (4.4). Using the Chain Rule

$$\left(\frac{dg}{ds}\right)_P = \left(\frac{\partial g}{\partial x} \cdot \frac{dx}{ds} + \frac{\partial g}{\partial y} \cdot \frac{dy}{ds} + \frac{\partial g}{\partial z} \cdot \frac{dz}{ds}\right)_P .$$

Here the subscript P is used to indicate that the derivatives are evaluated at the point P. Using (4.4), this may be rearranged to give

$$\left(\frac{dg}{ds}\right)_P = \left(\frac{\partial g}{\partial x} l + \frac{\partial g}{\partial y} m + \frac{\partial g}{\partial z} n\right)_P .$$

Note that the right hand side is equivalent to the scalar product of the two vectors

$$\left(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k}\right) \cdot (l\mathbf{i} + m\mathbf{j} + n\mathbf{k}) = \left(\frac{\partial g}{\partial x} l + \frac{\partial g}{\partial y} m + \frac{\partial g}{\partial z} n\right)$$

Thus,

$$\left(\frac{dg}{ds}\right)_P = \left(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k}\right) \cdot \hat{\mathbf{u}}, \tag{4.5}$$ {eq:4.2}

which is called the *directional derivative* of g along the direction $\hat{\mathbf{u}}$ at the point P.

The vector

$$\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k} \equiv \nabla g, \tag{4.6}$$ {eq:4.3}

is so important in mathematics that it is given the special name of *the gradient of the scalar function* $g(x, y, z)$. It is denoted by

$$\nabla g,$$

and is also called either grad g or the gradient of g. Note that ∇ is a *vector* operator. [It converts a *scalar* function into a *vector* function.] We can think of ∇ as the vector operator

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

The symbol ∇ is called *grad, del* or *nabla*. Thus, the directional derivative of $g(x, y, z)$ along $\hat{\mathbf{u}}$ at (x_0, y_0, z_0) is

$$\frac{dg}{ds} = (\nabla g \cdot \hat{\mathbf{u}})_{x_0,y_0,z_0} = (\hat{\mathbf{u}} \cdot \nabla g)_{x_0,y_0,z_0} .$$

Note that both terms on the right hand side are vectors and we take the scalar product of two vectors to produce the directional derivative.

Example 4.29

Find the directional derivative of

$$g = xy^2z^3,$$

in the direction $\mathbf{u} = 2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}$ at the point $P = (1, 1, 1)$.

Solution 4.29

First we need

$$\nabla g = \mathbf{i}\frac{\partial g}{\partial x} + \mathbf{j}\frac{\partial g}{\partial y} + \mathbf{k}\frac{\partial g}{\partial z} = \mathbf{i}(y^2z^3) + \mathbf{j}(2xz^3) + \mathbf{k}(3xy^2z^2).$$

Note that ∇g is a vector. At $(1, 1, 1)$, we have

$$(\nabla g)_P = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}.$$

Next we need to calculate the *unit vector* so that

$$\hat{\mathbf{u}} \equiv \frac{\mathbf{u}}{|\mathbf{u}|} = \frac{(2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k})}{\sqrt{4 + 36 + 9}} = \frac{2}{7}\mathbf{i} + \frac{6}{7}\mathbf{j} + \frac{3}{7}\mathbf{k}.$$

Thus, the directional derivative we require is

$$\frac{dg}{ds} = (\nabla g)_P \cdot \hat{\mathbf{u}} = \frac{1}{7}\left(\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}\right) \cdot \left(2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}\right).$$

Evaluating the scalar product gives the final answer as

$$\frac{dg}{ds} = \frac{1}{7}(2 + 12 + 9) = \frac{23}{7}.$$

Example End

Note that the rates of change of $g(x, y, z)$ along the x, y and z axes are just $\partial g/\partial x$, $\partial g/\partial y$ and $\partial g/\partial z$, from before. To confirm that the directional derivative gives this result, we set $\hat{\mathbf{u}} = \mathbf{i}$. Thus,

$$\nabla g \cdot \mathbf{i} = \left(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k}\right) \cdot \mathbf{i} = \frac{\partial g}{\partial x}.$$

Similarly $\hat{\mathbf{u}} = \mathbf{j}$ gives $\partial g/\partial y$ and $\hat{\mathbf{u}} = \mathbf{k}$ gives $\partial g/\partial z$.

4.3 Normals to surfaces and tangent planes

4.1a}

Given a surface $f(x, y, z) = c$, and a point P on it, the *tangent plane* (T) to the surface at P is the plane which just touches the surface at P. (This is analogous to the tangent to a curve, $y = f(x)$.)

The *normal* vector, ($\mathbf{n}$), to the surface at P is defined as the vector which is orthogonal (perpendicular) to every vector $\mathbf{t}$ in T through P (so that $\mathbf{n} \cdot \mathbf{t} = 0$). This is illustrated in Figure 4.7.

Note 1: Since $f(x, y, z)$ is constant on the surface, the directional derivative, evaluated at P, along *any* $\mathbf{t}$ will be zero. Thus,

:4.4}

$$\left(\frac{df}{ds}\right)_P = (\nabla f)_P \cdot \mathbf{t} = 0, \qquad \text{for any } \mathbf{t}. \tag{4.7}$$

Thus, $(\nabla f)_P$ is normal to both the surface (at P) and the tangent plane T. $(\nabla f)_P$ is parallel to the normal at P called $\mathbf{n}_P$. (Here the subscript just indicates that the vector is evaluated at the point P.)

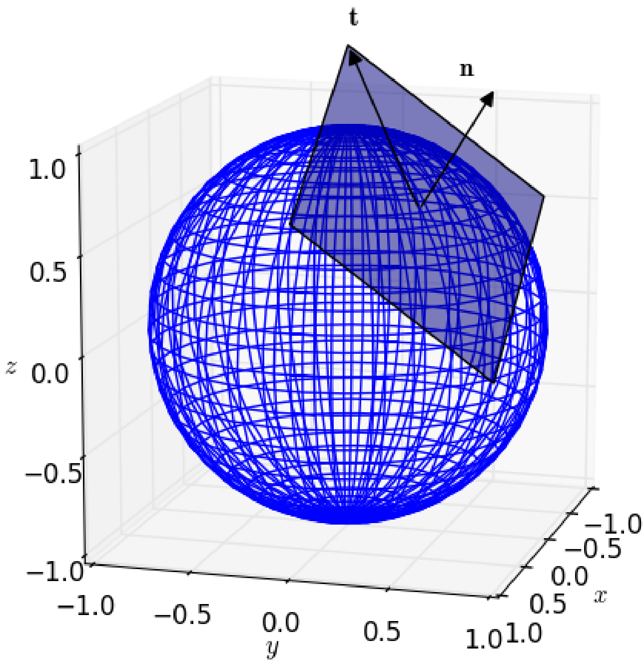

Figure 4.7: The surface $f(x, y, z) = c$ is shown. The tangent plane is labelled by T and a typical vector {fig:4.3} $\mathbf{t}$ lying in the tangent plane passing through P is shown.

Example 4.30

Let $f(x, y, z) = x - y^2 + xz$. The surface $f = -1$ contains the point $P = (1, 2, 2)$, (check to see that $f(1, 2, 2) = -1$). Find a vector parallel to $\mathbf{n}$ at P.

Solution 4.30

$$\nabla f = \mathbf{i}(1 + z) + \mathbf{j}(-2y) + \mathbf{k}(x),$$

and so, evaluating this at $(1, 2, 2)$ gives the vector

$$(\nabla f)_P = 3\mathbf{i} - 4\mathbf{j} + \mathbf{k},$$

and is parallel to $\mathbf{n}$ at P.

Example End

Note 2: Consider the rate of change of $f(x, y, z)$ at P along different directions defined by $\hat{\mathbf{u}}$ (see Figure 4.8). At P,

$$\left(\frac{df}{ds}\right)_{\hat{\mathbf{u}}} = (\nabla f) \cdot \hat{\mathbf{u}} = |\nabla f| \cos \gamma,$$

(from $\mathbf{A} \cdot \mathbf{B} = AB \cos \theta$). When $\gamma = \pi/2$ we find that $(df/ds)_{\gamma=\pi/2} = 0$. This is to be expected since $\hat{\mathbf{u}}$ coincides with some $\mathbf{t}$ in the tangent plane. Thus, $\hat{\mathbf{u}}$ is in the tangent plane and f is constant at P,

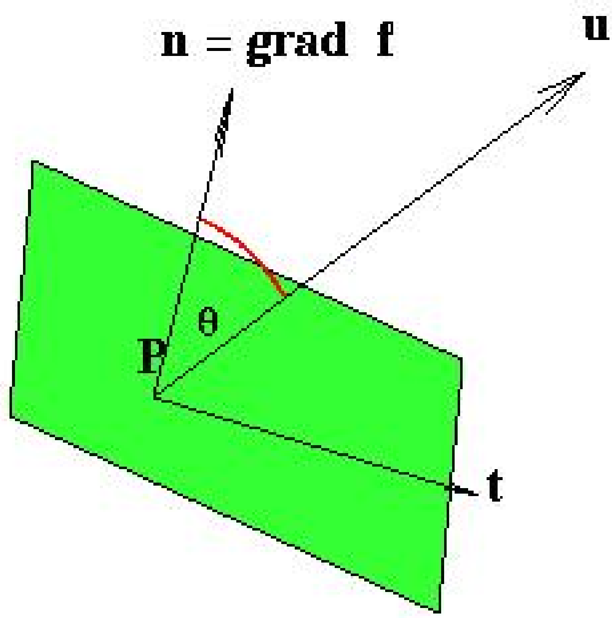

Figure 4.8: The direction of the normal to the surface, $\mathbf{n} = \nabla f$ makes an angle γ to the direction defined by $\mathbf{u}$. {fig:4.4}

see (4.7). Evidently, (df/ds) has its maximum value when $\cos\gamma = 1$, namely when $\gamma = 0$. Hence, $\hat{\mathbf{u}}$ coincides with the normal direction ($\mathbf{n}$ or ∇f). In this case, the maximum value of $|df/ds|$ is given by

$$|\frac{df}{ds}| = |\nabla f|.$$

Note 3: We will calculate the equation of the plane T, through $P = (x_0, y_0, z_0)$, where $\vec{OP} = \mathbf{r}_0$ is the position vector of the point P on the tangent plane T. If $Q(x, y, z)$ is a general point on the tangent plane T, with position vector $\vec{OQ} = \mathbf{r} = x\mathbf{i} + y\mathbf{j} + \mathbf{zk}$, then the vector $(\mathbf{r} - \mathbf{r}_0)$ that lies on the tangent plane must be perpendicular to the normal vector $\mathbf{n}$. Thus,

$$(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n}_P = 0.$$

Therefore, if $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $\mathbf{r}_0 = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}$, then using the definition of ∇f and expanding the scalar product gives the equation of the plane as

$$(x - x_0)\left(\frac{\partial f}{\partial x}\right)_P + (y - y_0)\left(\frac{\partial f}{\partial y}\right)_P + (z - z_0)\left(\frac{\partial f}{\partial z}\right)_P = 0.$$

This is of the form $ax + by + cz = d$, and is the equation of the tangent plane T.

Example 4.31

Find the tangent plane to

$$xy^2 + x^2z = 7,$$

at the point $(1, 2, 3)$.

Solution 4.31

Thus, $f = xy^2 + x^2z$ and $f = 7$. The normal vector is $\mathbf{n}$ and may be taken as $(\nabla f)_{(1,2,3)}$. Thus,

$$\nabla f = \mathbf{i}(y^2 + 2xz) + \mathbf{j}(2xy) + \mathbf{k}(x^2),$$

at the point $(1, 2, 3)$ we have

$$\mathbf{n} = \nabla f = 10\mathbf{i} + 4\mathbf{j} + \mathbf{k}.$$

With $\mathbf{r}_0 = (1, 2, 3)$, so $(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n} = 0$ gives

$$(x-1) \cdot 10 + (y-2) \cdot 4 + (z-3) \cdot 1 = 0, \qquad \Rightarrow \qquad 10x + 4y + z = 21.$$

Example End