Calculus

Haydn Cheng

March 21, 2025

Contents

1 Differentiation

1.1 Ordinary Differentiation

1.1.1 The Leibnitz' Theorem

Theorem 1.1.1 (The Leibnitz' Theorem)**.** *The n^{th} order derivative of the product of two fucntions $f(x) = u(x)v(x)$ is*[1]

$$f^{(n)} = \sum_{r=0}^{n} \binom{n}{r} u^{(r)} v^{(n-r)}. \tag{1.1}$$

1.1.2 Special Points of Functions

A stationary piont is characterised by $df/dx = 0$, and can be further classified into

1. Minimum point: $d^2f/dx^2 > 0$,
2. Maximum point: $d^2f/dx^2 < 0$ and
3. Inflection point (which is also stationary): $d^2f/dx^2 = 0$ and d^2f/dx^2 changes sign.

If $d^2f/dx^2 = 0$ but it does not change sign before and after the stationary point, it could be either of the three cases; we would have to check higher derivatives to verify its nature.

An inflection point (which is not stationary) is when $d^2f/dx^2 = 0$ but $df/dx \neq 0$ where the concavity of the fucntion changes.

1.2 Partial Differentiation

1.2.1 Partial Derivatives

Suppose we have three variables x, y and f, with $f = f(x, y)$ being assinged (arbitraily) as the dependent variable, and x, y being the independent variables (which are independent of each other, *i.e.*, x and y can vary independently). The rate in which f changes as x changes while keeping y fixed can be calculated by the partial derivative of f with respect to x

[1] Proof given in section A.1.

$$\left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial f}{\partial x} = f_x = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}, \tag{1.2}$$

where the variable which is kept constant is omitted at the first equality. In most cases, omitting the constant variable is conventional as they are obvious. However, care must be taken when the independent variables are not independent of each other.

The situation is best illustrated by an example. Suppose

$$f(x,y) = x^2 + y^2 + (x^2+y^2)^2, \tag{1.3}$$

and we introduce another independent variable $z = x^2 + y^2$, which is independent of x and y. Then there would be infinite ways to express $f(x,y,z)$ in terms of the independent variables x, y and z, such as $f(x,y,z) = x^2 + y^2 + z^2 = z + z^2 = x^2 + y^2 + z(x^2+y^2)$ *etc.*

If we want to calculate the partial derivative of $f(x,y,z)$ with respect to x, then we must write down explicitly the variable we would like to keep constant. In fact, the variable that is kept constant need not be y or z, it can be any combination of x, y and z, such as $w = y^2 \sin(x + \sqrt{z})$.

If we keep y to be constant (and z be whatever $z = x^2 + y^2$ is), then the partial derivative is

$$\left(\frac{\partial f(x,y,z)}{\partial x}\right)_y = \frac{\partial}{\partial x}(x^2 + y^2 + (x^2+y^2)^2) = 2x + 4x^3 + 4xy^2. \tag{1.4}$$

On the other hand, if we keep z to be constant (and y be whatever $y = \pm\sqrt{z - x^2}$ is), then the partial derivative is

$$\left(\frac{\partial f(x,y,z)}{\partial x}\right)_z = \frac{\partial}{\partial x}(z + z^2) = 0. \tag{1.5}$$

To see why is this so, refer to fig. 1.1, which shows the relation $z = x^2 + y^2$, and the variable $f(x,y,z)$ calculates the square of the distance from the origin.

Suppose we want to calculate the partial derivative of $f(x,y,z)$ with respect to x at the point $P_0 = (1,0,1)$. If we hold y fixed and let x varies, then the point P moves in the xz-plane towards A, and the distance increases as x increases. However, if we hold z fixed and let x varies, then point P moves in the plane $z = 1$, along the circular path towards B and the distance is fixed.

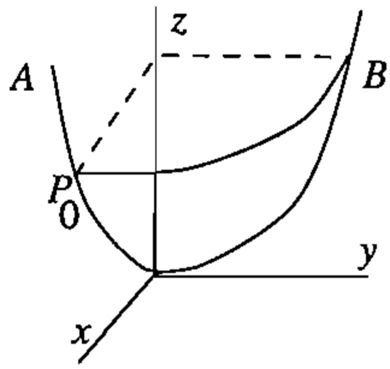


Figure 1.1

A very useful fact about the second partial derivative which the proof is not given here is that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{xy} = f_{yx} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}. \tag{1.6}$$

1.2.2 Total Derivatives

Now we consider how $f(x, y)$ changes when both x and y change. We have

$$\begin{aligned}
\Delta f &= f(x+\Delta x, y+\Delta y) - f(x, y) \\
&= f(x+\Delta x, y+\Delta y) - f(x, y) + f(x, y+\Delta y) - f(x, y+\Delta y) \\
&= \frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)}{\Delta x}\Delta x + \frac{f(x, y+\Delta y) - f(x, y)}{\Delta y}\Delta y \\
&\approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y.
\end{aligned} \tag{1.7}$$

Note that the partial deriavtives are carried out as if all the independent variables are independent of each other (although they might be dependent). The approximation becomes exact as $\Delta x \to 0$ and $\Delta y \to 0$.

This result shows that if the changes in x and y are small enough, we can consider the change in $f(x, y)$ due to x and y separately, as any change in $f(x, y)$ that is not linear to Δx or Δy is negligble as $\Delta x \to 0$ and $\Delta y \to 0$.

Example: Partial and Total Derivatives.

Question: Find $(\partial w/\partial y)_{x,t}$, where $w = x^3 y - z^2 t$ and $xy = zt$.

Solution: Straightforwardly, we have

$$\left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zt\left(\frac{\partial z}{\partial y}\right)_{x,t} = x^3 - 2zx. \tag{1.8}$$

Alternatively, we have

$$dw = 2x^2ydx + x^3dy + 2ztdz - z^2dt \quad \text{and} \quad ydx + xdy = zdt + tdz. \tag{1.9}$$

Eliminating dz, we have

$$dw = (3x^2y - 2zy)dx + (x^3 2zx)dy + z^2dt \implies \left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zx \tag{1.10}$$

1.2.3 Exact Differentials

An arbitrary differential

$$A(x,y)dx + B(x,y)dy \tag{1.11}$$

is exact if it is the differential of a function

$$df(x,y) = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \tag{1.12}$$

Therefore, we have

$$A(x,y) = \frac{\partial f}{\partial x} \text{ and } B(x,y) = \frac{\partial f}{\partial y} \tag{1.13}$$

Since $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$, we obtain a necessary (and a sufficient) condition for the differential to be exact, which is

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}.^2 \tag{1.14}$$

1.2.4 Reciprocity Relation and Cyclic Relation

So far our discussion has centred on a function $f(x,y)$ dependent on two variables, x and y. However, $f(x,y)$ is not superior, $x(f,y)$ and $y(x,f)$ are equally valid and are expressing the identical relation between x, y and f. To emphasise the point that all the variables are of equal standing we now replace f by z. Then we have

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \quad \text{and} \quad dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \\ \implies dx &= \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right) dz. \end{aligned} \tag{1.15}$$

Since dx and dz are independent, we the reciprocity and the cyclic relations

[2] Determining whether a differential containing many variables $x_1, x_2, \ldots, x_n$ is exact is a simple extension of the above: a differential $df = \sum_{i=1}^n g_i(x_1, x_2, \ldots, x_n)dx_i$ is exact if $\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i}$ for all pairs i, j.

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial y}{\partial x}\right)_z^{-1} \quad \text{and} \quad \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y + \left(\frac{\partial x}{\partial y}\right)_z = -1 \tag{1.16}$$

The same relationships can be found by dividing the total derivative by df or dx while keeping y or f constant, respectively.

It is important to note that in the reciprocity relationship, the same thing is held constant in both partial derivatives.

The reciprocity and the cyclic relations can be generalized to higher number of variables. For the reciprocity relation, we simply keep linking each numerator with the previous denominator while keeping the same constraint, *e.g.,*

$$\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial f}\right)_y = 1. \tag{1.17}$$

For the cyclic relation, the further variables simply appear as "spectators" which are held constant in all the terms, *e.g.,*

$$\left(\frac{\partial f}{\partial y}\right)_{x,z} \left(\frac{\partial y}{\partial x}\right)_{f,z} \left(\frac{\partial x}{\partial f}\right)_{y,z} = -1. \tag{1.18}$$

1.2.5 Change of Variables

Again, we emphasize that it is completely fine if the variables are related by some separate relations or cosntriant. Suppose x and y are parameterized by some other variable u, *i.e.,* $x = x(u)$ and $y = y(u)$. Then to find out how $f(x, y)$ changes with u, we simplify divide the total derivative of f with respect to x and y by du, so

$$\frac{df}{du} = \left(\frac{\partial f}{\partial x}\right)_y \frac{dx}{du} + \left(\frac{\partial f}{\partial y}\right)_x \frac{dy}{du}, \tag{1.19}$$

which is analogous to chain rule in ordinary differentiation.

Suppose now instead of having a constriant on x and y, we would like to change the whole set of variables from (x, y) to (u, v). Then we have $x = x(u, v)$ and $y = y(u, v)$. So the derivatives become

$$\left(\frac{\partial f}{\partial u}\right)_v = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v \quad \text{and} \quad \left(\frac{\partial f}{\partial v}\right)_u = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u . \tag{1.20}$$

To generalize the above discussion, suppose we want to change the set of n variables from $\mathbf{x}$ to $\mathbf{u}$, then the n equations are

$$\frac{\partial f}{\partial u_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_i}\frac{\partial f}{\partial x_j}, \tag{1.21}$$

or written in matrix form $\boldsymbol{\nabla}_{\mathbf{u}} = J^T \boldsymbol{\nabla}_{\mathbf{x}}$, where J is the Jacobian matrix

$$J \equiv \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}, \tag{1.22}$$

where we have assumed, for example, that v and w are kept constant in $\frac{\partial x}{\partial u}$.

In three dimensions, the 3 equations are

$$\begin{pmatrix} \frac{\partial}{\partial u} \\ \frac{\partial}{\partial v} \\ \frac{\partial}{\partial w} \end{pmatrix} f = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}^T \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} f. \tag{1.23}$$

For three sets of variables $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$, we have

$$\frac{\partial x_i}{\partial z_j} = \sum_{k=1}^{n} \frac{\partial x_i}{\partial y_k} \frac{\partial y_k}{\partial z_j}. \tag{1.24}$$

Now let A, B and C as the matrices whose ij^{th} elements are $\partial x_i/\partial y_j, \partial y_i/\partial z_j$ and $\partial x_i/\partial z_j$ respectively. We can then rewrite the above equation as

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \implies C = AB \implies \det(C) = \det(A)\det(B) \implies J_{xz} = J_{xy} J_{yz}. \tag{1.25}$$

In the special case where $z_i = x_i$, we get $J_{xy} J_{yx} = 1$.

1.2.6 Taylor's Theorem

When Δx and Δy are finite, we can no longer neglect the terms which are not linear in Δx or Δy in eq. (1.7), instead, we get the taylor series

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} \Delta x + \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} \Delta y \\ &+ \frac{1}{2!}\left(\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} (\Delta x)^2 + 2 \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)} (\Delta x)(\Delta y) + \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)} (\Delta y)^2 \right) \\ &+ \mathcal{O}((\Delta x)^3) + \mathcal{O}((\Delta y)^3). \end{aligned} \tag{1.26}$$

It can be shown that the general Taylor's theorem can be written as

$$f(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[(\Delta \mathbf{x} \cdot \nabla)^n f(\mathbf{x}) \right] \Big|_{\mathbf{x}=\mathbf{x}_0}. \tag{1.27}$$

1.2.7 Speical Points of a Function

From the Taylor's series above, we can see that a necessary and sufficient condition for a stationary point is that both partial derivatives vanish

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = 0. \tag{1.28}$$

To find the natures of the sationary points, we first complete the square so that

$$df = f(x,y) - f(x_0,y_0) \approx \frac{1}{2}\left[f_{xx}\left(\Delta x + \frac{f_{xy}\Delta y}{f_{xx}}\right)^2 + (\Delta y)^2\left(f_{yy} - \frac{f_{xy}^2}{f_{xx}}\right)\right]. \tag{1.29}$$

For a minimum point, we require that $df > 0$ for arbitrary Δx and Δy. This implies that $f_{xx} > 0$ and $f_{xx}f_{yy} > f_{xy}^2$. Due to symmetry of x and y, $f_{yy} > 0$ is also necessary. For saddle point, df can be positve, negative or zero depending on the choice of Δx and Δy. Therefore,

1. Minimum point: $f_{xx} > 0, f_{yy} > 0$ and $f_{xy}^2 < f_{xx}f_{yy}$.
2. Maximum point: $f_{xx} < 0, f_{yy} < 0$ and $f_{xy}^2 < f_{xx}f_{yy}$.
3. Saddle point: f_{xx} and f_{yy} have opposite signs, or $f_{xy}^2 > f_{xx}f_{yy}$.

If $f_{xy}^2 = f_{xx}f_{yy}$, then df must be one of the four forms $\pm(|f_{xx}|^{\frac{1}{2}}\Delta x/2 \pm |f_{yy}|^{\frac{1}{2}}\Delta y)^2$, then for some choice of the ratio $\Delta y/\Delta x$, $df = 0$ so higher order terms are needed to find the nature of the stationary point.

For functions with more than 2 variables, the conditions for stationary points are

$$\frac{\partial f}{\partial x_i} = 0 \text{ for all } x_i, \tag{1.30}$$

where x_i are the variables.

To investigate the nature of the stationary points, we again use the second order term of the Taylor's series

$$df = f(\mathbf{x}) - f(\mathbf{x}_0) \approx \frac{1}{2}\sum_i\sum_j \frac{\partial^2 f}{\partial x \partial y}\Delta x_i \Delta x_j, \tag{1.31}$$

which must be positive for all Δx_i.

1.3 Curvilinear coordinates

1.3.1 General Curvilinear Coordinates

A point in three-dimensional space can be specified by three coordinates (u, v, w). In Cartesian coordinates, $(u, v, w) = (x, y, z)$; In spherical coordinates, $(u, v, w) = (r, \theta, \phi)$; In cyilndrical coordintes, $(u, v, w) = (\rho, \phi, z)$.

The infinitesimal displacement vector $d\mathbf{r}$ from (u, v, w) to $(u + du, v + dv, w + dw)$ can be written as

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u} du + \frac{\partial \mathbf{r}}{\partial v} dv + \frac{\partial \mathbf{r}}{\partial w} dw. \tag{1.32}$$

If the coordinate system is orthogonal *i.e.,* $\hat{\mathbf{u}} \perp \hat{\mathbf{v}} \perp \hat{\mathbf{w}}$, where $\hat{\mathbf{u}}, \hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$ are the unit vectors whose direction are directed along increaseing u, v and w respectively, then we have

$$\frac{\partial \mathbf{r}}{\partial u} = f\hat{\mathbf{u}}, \frac{\partial \mathbf{r}}{\partial v} = g\hat{\mathbf{v}} \text{ and } \frac{\partial \mathbf{r}}{\partial w} = h\hat{\mathbf{w}}, \tag{1.33}$$

where f, g and h are characteristic constants of a coordinates system which scale the unit vectors. In Cartesian coordinates, $(f, g, h) = (1, 1, 1)$; In spherical coordinates, $(f, g, h) = (1, r, r\sin\theta)$; In cyilndrical coordinates, $(f, g, h) = (1, \rho, 1)$.

The infinitesimal displacement vector is now

$$d\mathbf{r} = f(du\hat{\mathbf{u}}) + g(dv\hat{\mathbf{v}}) + h(dw\hat{\mathbf{w}}) \tag{1.34}$$

and the arc length is the norm of $d\mathbf{r}$, which is

$$ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}} = \sqrt{(fdu)^2 + (gdv)^2 + (hdw)^2}. \tag{1.35}$$

The infinitesimal area perpendicular to $\hat{\mathbf{w}}$ will be a rectangle with area

$$dS = (fg)dudv \tag{1.36}$$

as shown in fig. 1.2.

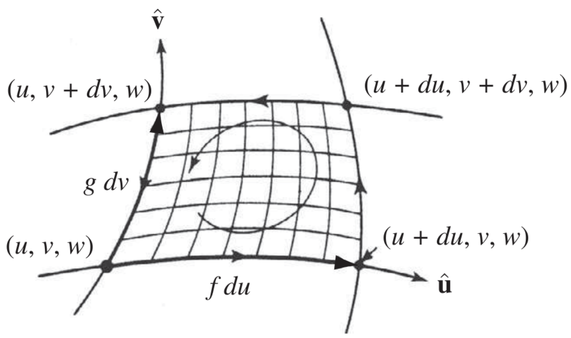


Figure 1.2

The infinitesimal volume is a parallelepiped (rectangular solid if the system is orthogonal) with volume

$$d\tau = |fd\hat{\mathbf{u}} \cdot (gd\hat{\mathbf{v}} \times hd\hat{\mathbf{w}})|dudvdw = (fgh)dudvdw \tag{1.37}$$

as shown in fig. 1.3.

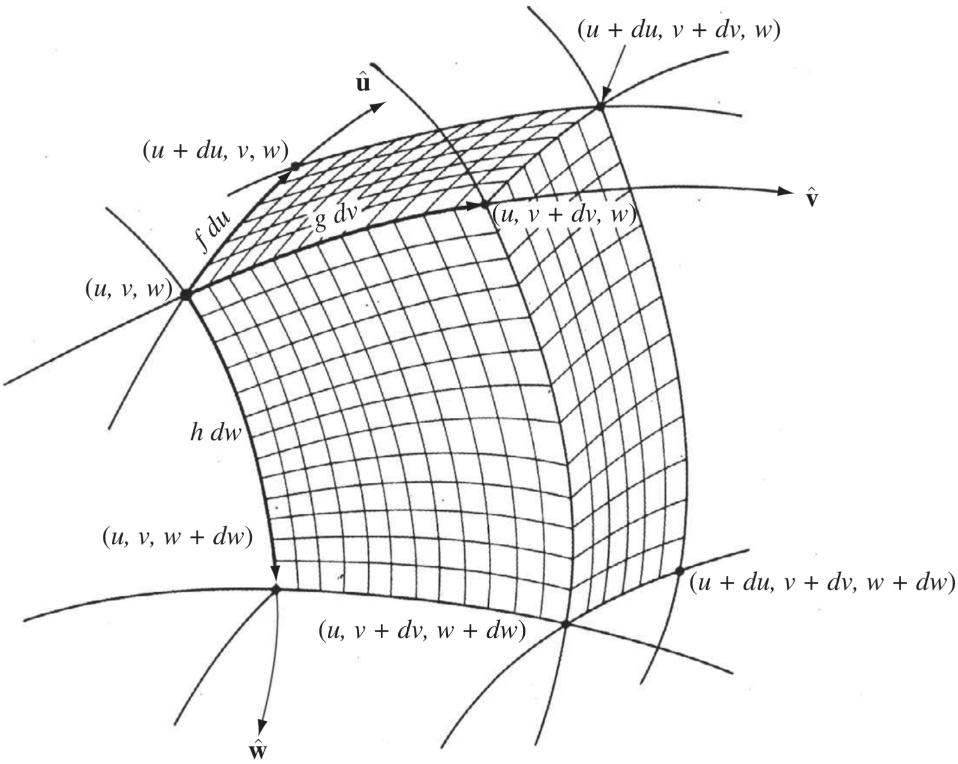


Figure 1.3

1.3.2 Spherical Coordinates

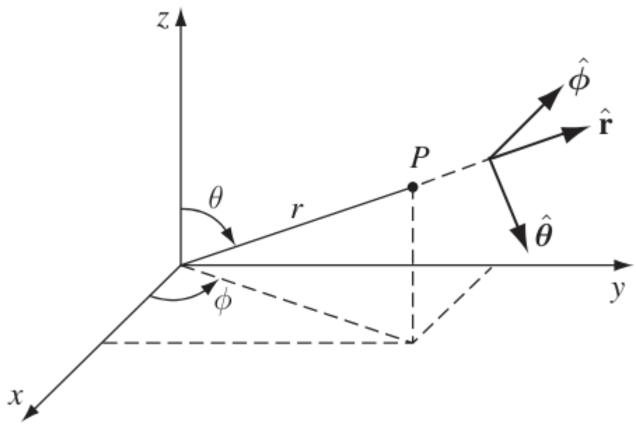


Figure 1.4

From fig. 1.4, the relations of the two set of variables (x, y, z) and (r, θ, ϕ) are

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases} \quad \text{or} \quad \begin{cases} r = \sqrt{x^2+y^2+z^2} \\ \theta = \arctan\left(\sqrt{x^2+y^2}/z\right) \\ \phi = \arctan(y/x) \end{cases} \quad . \tag{1.38}$$

A general position vector can then be written in 4 different ways as

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = r\sin\theta\cos\phi\hat{\mathbf{i}} + r\sin\theta\sin\phi\hat{\mathbf{j}} + r\cos\theta\hat{\mathbf{k}} = r\hat{\mathbf{r}} = \sqrt{x^2+y^2+z^2}\hat{\mathbf{r}}, \tag{1.39}$$

where the last one is seldom used.

A general infinitismal displacement vector can be written as

$$\begin{aligned} d\mathbf{r} &= \frac{\partial\mathbf{r}}{\partial x}dx + \frac{\partial\mathbf{r}}{\partial y}dy + \frac{\partial\mathbf{r}}{\partial z}dz = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}} \\ &= \frac{\partial\mathbf{r}}{\partial r}dr + \frac{\partial\mathbf{r}}{\partial\theta}d\theta + \frac{\partial\mathbf{r}}{\partial\phi}d\phi = fdr\hat{\mathbf{r}} + gd\theta\hat{\boldsymbol{\theta}} + hd\phi\hat{\boldsymbol{\phi}}. \end{aligned} \tag{1.40}$$

where $\partial\mathbf{r}/\partial r, \partial\mathbf{r}/\partial\theta$ and $\partial\mathbf{r}/\partial\phi$ can be found by direct differentiaion as

$$\begin{aligned} \frac{\partial\mathbf{r}}{\partial r} &= f\hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{i}} + \sin\theta\sin\phi\hat{\mathbf{j}} + \cos\theta\hat{\mathbf{k}}, \\ \frac{\partial\mathbf{r}}{\partial\theta} &= g\hat{\boldsymbol{\theta}} = r(\cos\theta\cos\phi\hat{\mathbf{i}} + \cos\theta\sin\phi\hat{\mathbf{j}} - \sin\theta\hat{\mathbf{k}}) \text{ and} \\ \frac{\partial\mathbf{r}}{\partial\phi} &= r\hat{\boldsymbol{\phi}} = \sin\theta(-\sin\phi\hat{\mathbf{i}} + \cos\phi\hat{\mathbf{j}}). \end{aligned} \tag{1.41}$$

Thus $f = 1, g = r$ and $h = r\sin\theta$.

We can thus solve for $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ in terms of $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ as

$$\begin{cases} \hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{i}} + \sin\theta\sin\phi\hat{\mathbf{j}} + \cos\theta\hat{\mathbf{k}}, \\ \hat{\boldsymbol{\theta}} = \cos\theta\cos\phi\hat{\mathbf{i}} + \cos\theta\sin\phi\hat{\mathbf{j}}0\sin\theta\hat{\mathbf{k}}, \\ \hat{\boldsymbol{\phi}} = -\sin\phi\hat{\mathbf{i}} + \cos\phi\hat{\mathbf{j}}. \end{cases} \tag{1.42}$$

We can also solve for $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ in terms of $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ as

$$\begin{cases} \hat{\mathbf{i}} = \sin\theta\cos\phi\hat{\mathbf{r}} + \cos\theta\cos\phi\hat{\boldsymbol{\theta}} - \sin\phi\hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{j}} = \sin\theta\sin\phi\hat{\mathbf{r}} + \cos\theta\sin\phi\hat{\boldsymbol{\theta}} + \cos\phi\hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{k}} = \cos\theta\hat{\mathbf{r}} - \sin\theta\hat{\boldsymbol{\theta}}, \end{cases} \tag{1.43}$$

which are seldomly used.

1.3.3 Cylindrical Coordinates

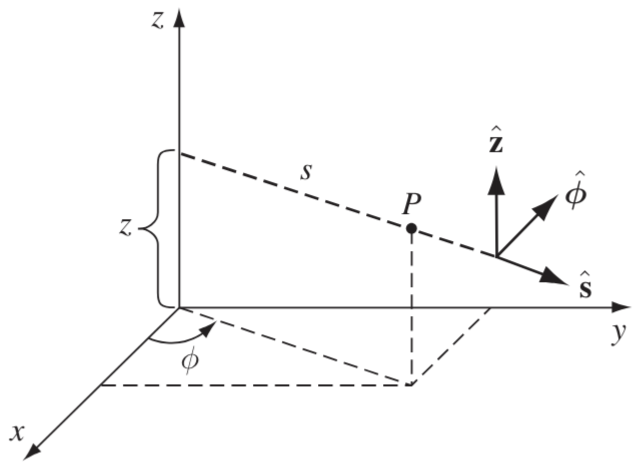


Figure 1.5

From fig. 1.5, the relations of the two set of variables (x, y, z) and (ρ, ϕ, z) are

$$\begin{cases} x = \rho\cos\phi \\ y = \rho\sin\phi \\ z = z \end{cases} \quad \text{or} \quad \begin{cases} \rho = \sqrt{x^2+y^2} \\ \phi = \arctan\,(y/x) \\ z = z \end{cases} \quad . \tag{1.44}$$

A general position vector can be written in 4 ways as

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = \rho\cos\phi\hat{\mathbf{i}} + \rho\sin\phi\hat{\mathbf{j}} + \hat{\mathbf{k}} = \rho\hat{\boldsymbol{\rho}} + z\hat{\mathbf{z}} = \sqrt{x^2+y^2}\hat{\boldsymbol{\rho}} + z\hat{\mathbf{z}}, \tag{1.45}$$

where the last one is seldom used.

A general infinitismal displacement vector can be written as

$$\begin{aligned} d\mathbf{r} &= \frac{\partial\mathbf{r}}{\partial x}dx + \frac{\partial\mathbf{r}}{\partial y}dy + \frac{\partial\mathbf{r}}{\partial z}dz = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}} \\ &= \frac{\partial\mathbf{r}}{\partial\rho}d\rho + \frac{\partial\mathbf{r}}{\partial\phi}d\phi + \frac{\partial\mathbf{r}}{\partial z}dz = fd\rho\hat{\boldsymbol{\rho}} + gd\phi\hat{\boldsymbol{\phi}} + hdz\hat{\mathbf{z}}. \end{aligned} \tag{1.46}$$

where $\partial\mathbf{r}/\partial\rho, \partial\mathbf{r}/\partial\phi$ and $\partial\mathbf{r}/\partial z$ can be found by direct differentiaion as

$$\begin{aligned} \frac{\partial\mathbf{r}}{\partial\rho} &= f\hat{\boldsymbol{\rho}} = \cos\phi\hat{\mathbf{i}} + \sin\phi\hat{\mathbf{j}}, \\ \frac{\partial\mathbf{r}}{\partial\phi} &= g\hat{\boldsymbol{\phi}} = \rho(-\sin\phi\hat{\mathbf{i}} + \cos\phi\hat{\mathbf{j}}) \text{ and} \\ \frac{\partial\mathbf{r}}{\partial z} &= h\hat{\mathbf{k}} = \hat{\mathbf{k}}. \end{aligned} \tag{1.47}$$

Thus $f = 1, g = \rho$ and $h = 1$.

We can thus solve for $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{k}})$ in terms of $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ as

$$\begin{cases} \hat{\boldsymbol{\rho}} = \cos\phi \hat{\mathbf{i}} + \sin\phi \hat{\mathbf{j}}, \\ \hat{\boldsymbol{\phi}} = -\sin\phi \hat{\mathbf{i}} + \cos\phi \hat{\mathbf{j}}, \\ \hat{\mathbf{k}} = \hat{\mathbf{k}}. \end{cases} \tag{1.48}$$

We can also solve for $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ in terms of $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{k}})$ as

$$\begin{cases} \hat{\mathbf{i}} &= \cos\phi \hat{\boldsymbol{\rho}} - \sin\phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{j}} &= \sin\phi \hat{\boldsymbol{\rho}} + \cos\phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{k}} &= \hat{\mathbf{k}}, \end{cases} \tag{1.49}$$

which are seldomly used.

1.4 Gradient, Divergence and Curl

We start by stating the gerneral form of del operator, gradient, divergence, curl, and Laplacian are defined as[3]

$$\begin{aligned} \boldsymbol{\nabla} &= \frac{1}{f}\hat{\mathbf{u}}\frac{\partial}{\partial u} + \frac{1}{g}\hat{\mathbf{v}}\frac{\partial}{\partial v} + \frac{1}{h}\hat{\mathbf{w}}\frac{\partial}{\partial w}, \\ \boldsymbol{\nabla} t &= \frac{1}{f}\hat{\mathbf{u}}\frac{\partial t}{\partial u} + \frac{1}{g}\hat{\mathbf{v}}\frac{\partial t}{\partial v} + \frac{1}{h}\hat{\mathbf{w}}\frac{\partial t}{\partial w}, \\ \boldsymbol{\nabla} \cdot \mathbf{T} &= \frac{1}{fgh}\left[\frac{\partial}{\partial u}(ghT_u) + \frac{\partial}{\partial v}(fhT_v) + \frac{\partial}{\partial w}(fgT_h)\right], \\ \boldsymbol{\nabla} \times \mathbf{T} &= \frac{1}{fgh}\begin{vmatrix} f\hat{\mathbf{u}} & g\hat{\mathbf{v}} & h\hat{\mathbf{w}} \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ fT_u & gT_v & hT_w \end{vmatrix}, \\ \nabla^2 t = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} t) &= (1/fgh)\left[\frac{\partial}{\partial u}\left(\frac{gh}{f}\frac{\partial t}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{fh}{g}\frac{\partial t}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{fg}{h}\frac{\partial t}{\partial w}\right)\right], \end{aligned} \tag{1.50}$$

where $t(u, v, w)$ is a scalar field, which appoints each point in space a scalar, and $\mathbf{T} = T_u\hat{\mathbf{u}} + T_v\hat{\mathbf{v}} + T_w\hat{\mathbf{w}}$ is a vector field which appoints each point in space a vector.

1.4.1 Gradient

Using the del operator defined in eq. (1.50), we can write the infinitesimal change of a scalar function $t(u, v, w)$ as

[3]The del operator is not written as $\frac{1}{f}\frac{\partial}{\partial u}\hat{\mathbf{u}} + \frac{1}{g}\frac{\partial}{\partial v}\hat{\mathbf{v}} + \frac{1}{h}\frac{\partial}{\partial w}\hat{\mathbf{w}}$ because in a general coordinates system, unit vector is not a constant but depends on the coordinates of the point in space, so we take the unit vectors out of the partial derivatives to avoid confusion since we are not differentiating them.

$$dt = \frac{\partial t}{\partial u}du + \frac{\partial t}{\partial v}dv + \frac{\partial t}{\partial w}dw = \boldsymbol{\nabla} t \cdot d\mathbf{r} = |\boldsymbol{\nabla} t||d\mathbf{r}|\cos\theta. \tag{1.51}$$

where θ is the angle between $\boldsymbol{\nabla} t$ and $d\mathbf{r}$.

From the above equation, it is evident that dt attains maximum when $\theta = 0$, *i.e.* $d\mathbf{r} /\!/ \boldsymbol{\nabla} t$. Thus, the gradient $\boldsymbol{\nabla} t$ points in the direction of maximum increase of the function t and $|\boldsymbol{\nabla} t|$ gives the slope along this maximal direction.

problem 1.14, 1.17, 1.1.5

Integrating, we get the fundamental theorem for gradients

$$t(\mathbf{b}) - t(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{b}} dt = \int_{\mathbf{a}}^{\mathbf{b}} \boldsymbol{\nabla} t \cdot d\mathbf{r}. \tag{1.52}$$

This equation shows that if one would like to determine the height of Mountain Everest, one could place altimeters at the top and the bottom and subtract the two readings, or climb the mountain and measure the rise at each step.

A quick corollary is that if the vector field $\mathbf{T}(\mathbf{r})$ can be written as the gradient of a scalar field $\boldsymbol{\nabla} t$, then the line integral

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = \int_{\mathbf{a}}^{\mathbf{b}} \boldsymbol{\nabla} t \cdot d\mathbf{r} = \int_{\mathbf{a}}^{\mathbf{b}} dt = t(\mathbf{b}) - t(\mathbf{a}) \tag{1.53}$$

is independent of $\mathbf{a}$ and $\mathbf{b}$ and the vector field $\mathbf{T}(\mathbf{r})$ is known as a conservative field. Then also, $\oint_C \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = 0$ for any closed loop if $\mathbf{T}(\mathbf{r})$ is conservative, since $t(\mathbf{b}) - t(\mathbf{a}) = 0$.[4]

A quick way to check whether $\mathbf{T}(\mathbf{r})$ is conservative is to check if its curl is zero, since the curl of a gradient is always zero, so $\boldsymbol{\nabla} \times \mathbf{F}(\mathbf{r}) = \boldsymbol{\nabla} \times (\boldsymbol{\nabla} t(\mathbf{r})) = 0$.

An additional condition for us to safely say that $\mathbf{T}(\mathbf{r})$ is conservative, which is usually assumed to be met, is that $t(\mathbf{r})$ should be continuous. For example, for the field

$$\mathbf{T}(\mathbf{r}) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2} \right) \tag{1.54}$$

zero curl, so the it is conservative and the corresonding potential function is

$$t(\mathbf{r}) = \tan^{-1}\left(\frac{y}{x}\right). \tag{1.55}$$

However, the loop integral

$$\oint_C \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = 2\pi \neq 0 \tag{1.56}$$

when C is a circle centered at the origin. The inconsistency arises from the fact that $t(\mathbf{r})$ is not continuous at the origin, so $\mathbf{T}(\mathbf{r})$ is only conservative on the plane excluding the

[4] The converse, that we can always find a potential function $t(\mathbf{r})$ for $\oint_C \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = 0$ is also true, but involve moderate algebra, so we will take this obvious fact as granted.

origin, *i.e.,* $\mathbb{R}^2 - 0, 0$. If the loop integral of $\mathbf{T}(\mathbf{r})$ includes the origin then it would not be zero, since $\mathbf{T}(\mathbf{r})$ is no longer conservative.

It can also be shown that the gradient at a point can be written in terms of a surface integral over an infinitesimal surface surrounding the point

$$\boldsymbol{\nabla} t = \lim_{V\to 0}\left(\frac{1}{V}\oint_S t d\mathbf{S}\right). \tag{1.57}$$

However the proof is complicated and is not presented here.

Example: Gradient of $\mathscr{r}^{-1}$ in Cartesian Coordinates.

Question: Evaluate $\boldsymbol{\nabla}(\mathscr{r}^{-1})$, where $\mathscr{r} = |\boldsymbol{\mathscr{r}}| = |\mathbf{r} - \mathbf{r}'| = \left|(x-x')\hat{\mathbf{i}} + (y-y')\hat{\mathbf{j}} + (z-z')\hat{\mathbf{k}}\right| = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$. Generalize the result to obtain $\boldsymbol{\nabla}(\mathscr{r}^n)$.

Solution:

$$\begin{aligned}
\boldsymbol{\nabla}\mathscr{r}^{-1} &= \hat{\mathbf{i}}\frac{\partial}{\partial x}\left((x-x')^2 + (y-y')^2 + (z-z')^2\right)^{-\frac{1}{2}} \\
&+ \hat{\mathbf{j}}\frac{\partial}{\partial y}\left((x-x')^2 + (y-y')^2 + (z-z')^2\right)^{-\frac{1}{2}} \\
&+ \hat{\mathbf{k}}\frac{\partial}{\partial z}\left((x-x')^2 + (y-y')^2 + (z-z')^2\right)^{-\frac{1}{2}} \\
&= (-\frac{1}{2})(2)(((x-x')^2 + (y-y')^2 + (z-z')^2))^{-\frac{3}{2}} \\
&((x-x')\hat{\mathbf{i}} + (y-y')\hat{\mathbf{j}} + (z-z')\hat{\mathbf{k}}) \\
&= -\boldsymbol{\mathscr{r}}\mathscr{r}^{-1}.
\end{aligned} \tag{1.58}$$

This result can be easily generalised to get

$$\boldsymbol{\nabla}(\mathscr{r}^n) = n\mathscr{r}^{n-1}\hat{\boldsymbol{\mathscr{r}}}. \tag{1.59}$$

Example: Gradient in Cartesian and Polar Coordinates.

Question: Calculate the gradient of the function $f(x, y) = x^2 + y^2$ in both Cartesian and the polar coordinates.

Solution: In Cartesian coordinates, the gradient is $\boldsymbol{\nabla} f(x, y) = (2x, 2y)$.
The two sets of variables $\mathbf{x} = (x, y)$ and $\mathbf{u} = (r, \theta)$ are related by $x = r\cos\theta$ and $y = r\sin\theta$. The gradients are related by eq. (1.23), as

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}. \tag{1.60}$$

Inverting the equations, we have

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\frac{\sin\theta}{r} \\ \sin\theta & \frac{\cos\theta}{r} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \\ \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta} \end{pmatrix}. \tag{1.61}$$

Therefore, we have

$$\begin{aligned} \boldsymbol{\nabla} f(x,y) &= \left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\right)\hat{\mathbf{i}} + \left(\sin\theta\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)\hat{\mathbf{j}} = 2r(\cos\theta, \sin\theta) \\ &= \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial\theta}\hat{\boldsymbol{\theta}} = 2r\hat{\mathbf{r}} = 2r(1,0). \end{aligned} \tag{1.62}$$

Note that the answer $2r(\cos\theta, \sin\theta)$ is still in Cartesian coordiantes, just that we used variables in polar coordinates to represent it. The answer $2r(1,0)$ on the other hand, is the gradient in polar coordiantes.

The gradient in Cartesian and spherical coordinates are stated here for reference:

$$\boldsymbol{\nabla} t = \frac{\partial t}{\partial x}\hat{\mathbf{i}} + \frac{\partial t}{\partial y}\hat{\mathbf{j}} + \frac{\partial t}{\partial z}\hat{\mathbf{k}} = \frac{\partial t}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial t}{\partial\theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial t}{\partial\phi}\hat{\boldsymbol{\phi}}. \tag{1.63}$$

1.4.2 Divergence

In order to seek geometrical interpretation of the divergence operator, we consider the closed surface integral of $\mathbf{T}$ over the surface of an infinitesimal volume depicted in fig. 1.3 (here we adopt the sign convention of $\mathbf{S}$ is positive if $\mathbf{S}$ is pointing outwards from the interior of the volume)

$$\begin{aligned} \oint_S \mathbf{T}\cdot d\mathbf{S} &= ((T_u)|_{u+du} - (T_u)|_u)(ghdvdw) \\ &\quad + ((T_v)|_{v+dv} - (T_v)|_v)(fhdudw) \\ &\quad + ((T_w)|_{w+dw} - (T_w)|_w)(fgdudv) \\ &= \frac{1}{fgh}\left(\frac{\partial}{\partial u}(ghT_u) + \frac{\partial}{\partial v}(fhT_v) + \frac{\partial}{\partial w}(fgT_h)\right)d\tau \\ &= (\boldsymbol{\nabla}\cdot\mathbf{T})d\tau \end{aligned} \tag{1.64}$$

This result can be extended easily. As any arbitary volume can be divided infinitely into infinitesimal pieces, and the surface integral of each individual pieces cancel in pairs, the remaining part is only the surface integral of the surface of the whole volume. Therefore,

$$\oint_S \mathbf{T}\cdot d\mathbf{S} = \int_V (\boldsymbol{\nabla}\cdot\mathbf{T})d\tau. \tag{1.65}$$

This is the divergence theorem (also known as the Gauss's theorem or the Green's theorem). From the LHS of eq. (1.65), it is evident that the divergence of a vector function is a measure of how much the function spreds out and diverges from a given point in space.

If $\mathbf{T}$ satisfies $\boldsymbol{\nabla}\cdot\mathbf{T} = 0$, then $\mathbf{T} = \boldsymbol{\nabla}\times\mathbf{T}' + \boldsymbol{\nabla} t + \mathbf{C}$ can be written as a curl of a vector $\mathbf{T}'$, which statisfies $\boldsymbol{\nabla}\times\mathbf{T}' = 0$ plus a gradient of a scalar function t plus a constant vector $\mathbf{C}$.

The divergence in Cartesian and spherical coordinates are stated here for reference:

$$\boldsymbol{\nabla}\cdot\mathbf{T} = \frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y} + \frac{\partial T_z}{\partial z} = \frac{1}{r^2}\frac{\partial}{\partial r}\big(r^2 T_r\big) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta T_\theta) + \frac{1}{r}\frac{\partial}{\partial\phi}(rT_\phi). \tag{1.66}$$

1.4.3 Curl

In order to seek geometrical interpretation of the curl operator, we consider the loop integral of $\mathbf{T}$ over an infinitesimal loop depicted in fig. 1.2 (here since the coordinates system is right handed, $\hat{\mathbf{w}}$ points out of the page, hence we are obliged by the right-hand rule to run the line integral counterclockwise such that $\mathbf{S}$ points in the same direction as $\hat{\mathbf{w}}$)

$$\begin{aligned}\oint_C \mathbf{T}\cdot d\mathbf{r} &= (T_u)|_v f du + (T_v)|_{u+du} g dv + (T_u)|_{v+dv}(-f du) + (T_v)|_u(-g dv)\\ &= \frac{1}{fg}\left(\frac{\partial}{\partial u}(T_v g) - \frac{\partial}{\partial v}(T_u f)\right)(\hat{\mathbf{w}}\cdot d\mathbf{S})\\ &= (\boldsymbol{\nabla}\times\mathbf{T})\cdot d\mathbf{S}.\end{aligned} \tag{1.67}$$

With the same argument as eq. (1.64) to eq. (1.65), the above equation can be extended to

$$\oint_C \mathbf{T}\cdot d\mathbf{r} = \int_S (\boldsymbol{\nabla}\times\mathbf{T})\cdot d\mathbf{S}. \tag{1.68}$$

This is the Stoke's theorem. From the LHS of eq. (1.68), it is evident that the curl of a vector function is a measure of how much the function rotates and curls around a given point in space.

A quick corollary is that the integral $\int_S(\boldsymbol{\nabla}\times\mathbf{T})\cdot d\mathbf{S}$ is independent of the surface used but only depends on the boundary line. Hence, $\oint_S(\boldsymbol{\nabla}\times\mathbf{T})\cdot d\mathbf{S} = 0$ for any closed surface, since the boundary line, like the mouth of a ballon, shrinks down to a point, and thus $\oint_P \mathbf{T}\cdot d\mathbf{r} = 0$.

$\mathbf{T}$ is called conservative if $\int_A^B \mathbf{T}\cdot d\mathbf{r}$ is independent of the path taken from A to B. This implies that $\oint_C \mathbf{T}\cdot d\mathbf{r} = 0$ and thus $\boldsymbol{\nabla}\times\mathbf{T} = 0$ and $\mathbf{T} = \boldsymbol{\nabla}t$ can be written as a gradient of some scalar function t. Also, $\mathbf{T}\cdot d\mathbf{r}$ is an exact differential.

The curl in Cartesian and spherical coordinates are stated here for reference:

$$\begin{aligned}\boldsymbol{\nabla}\times\mathbf{T} &= (\frac{\partial T_z}{\partial y} - \frac{\partial T_y}{\partial z})\hat{\mathbf{i}} + (\frac{\partial T_x}{\partial z} - \frac{\partial T_z}{\partial x})\hat{\mathbf{j}} + (\frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y})\hat{\mathbf{k}}\\ &= \frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}(\sin\theta T_\phi) - \frac{\partial T_\theta}{\partial\phi}\right)\hat{\mathbf{r}} + \frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial T_r}{\partial\phi} - \frac{\partial}{\partial r}(rT_\phi)\right)\hat{\boldsymbol{\theta}} + \frac{1}{r}\left(\frac{\partial}{\partial r}(rT_\theta) - \frac{\partial T_r}{\partial\theta}\right)\hat{\boldsymbol{\phi}}.\end{aligned} \tag{1.69}$$

1.4.4 Product Rules

The two product rules for gradient are

$$
\begin{cases}
\boldsymbol{\nabla}(fg) = f\boldsymbol{\nabla}g + g\boldsymbol{\nabla}f, \\
\boldsymbol{\nabla}(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\times(\boldsymbol{\nabla}\times\mathbf{B}) + \mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{A}) + (\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{B} + (\mathbf{B}\cdot\boldsymbol{\nabla})\mathbf{A}.
\end{cases}
\tag{1.70}
$$

The two product rules for divergence are

$$
\begin{cases}
\boldsymbol{\nabla}\cdot(f\mathbf{A}) = f(\boldsymbol{\nabla}\cdot\mathbf{A}) + \mathbf{A}\cdot(\boldsymbol{\nabla}f), \\
\boldsymbol{\nabla}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) - \mathbf{A}\cdot(\boldsymbol{\nabla}\times\mathbf{B}).
\end{cases}
\tag{1.71}
$$

The two product rules for curl are

$$
\begin{cases}
\boldsymbol{\nabla}\times(f\mathbf{A}) = f(\boldsymbol{\nabla}\times\mathbf{A}) - (\mathbf{A}\times(\boldsymbol{\nabla}f)), \\
\boldsymbol{\nabla}\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\cdot\boldsymbol{\nabla})\mathbf{A} - (\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{B} + \mathbf{A}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \mathbf{B}(\boldsymbol{\nabla}\cdot\mathbf{A}).
\end{cases}
\tag{1.72}
$$

Here note that

$$
\begin{aligned}
(\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{B} &= \left(A_x\frac{\partial}{\partial x} + A_y\frac{\partial}{\partial y} + A_z\frac{\partial}{\partial z}\right)\left(B_x\hat{\mathbf{i}} + B_y\hat{\mathbf{j}} + B_z\hat{\mathbf{k}}\right) \\
&= \left(A_x\frac{\partial B_x}{\partial x} + A_y\frac{\partial B_x}{\partial y} + A_z\frac{\partial B_x}{\partial z}\right)\hat{\mathbf{i}} \\
&+ \left(A_x\frac{\partial B_y}{\partial x} + A_y\frac{\partial B_y}{\partial y} + A_z\frac{\partial B_y}{\partial z}\right)\hat{\mathbf{j}} \\
&+ \left(A_x\frac{\partial B_z}{\partial x} + A_y\frac{\partial B_z}{\partial y} + A_z\frac{\partial B_z}{\partial z}\right)\hat{\mathbf{k}} \\
&\neq \mathbf{A}\cdot(\boldsymbol{\nabla}\mathbf{B}) \\
&= A_x\frac{\partial B}{\partial x} + A_y\frac{\partial B}{\partial y} + A_z\frac{\partial B}{\partial z}.
\end{aligned}
\tag{1.73}
$$

In component form, this is $(\mathbf{A}\cdot\boldsymbol{\nabla})_i = A_i\partial_i$. Another differential operator we can construct is $(\mathbf{A}\times\boldsymbol{\nabla})_k = \epsilon_{ijk}A_i\partial_j$ but it is seen much less often.

With the product rules in hand, we can perform the so-called "Integration by part" trick. For example, by integrating the first product rule of divergence and using the divergence theorem, we have

$$
\begin{aligned}
&\int_V \boldsymbol{\nabla}\cdot(f\mathbf{T})d\tau = \int_V f(\boldsymbol{\nabla}\cdot\mathbf{T})d\tau + \int_V \mathbf{T}\cdot(\boldsymbol{\nabla}f)d\tau = \oint_S f\mathbf{T}\cdot d\mathbf{A} \\
\Longrightarrow &\int_V f(\boldsymbol{\nabla}\cdot\mathbf{T})d\tau = \oint_S (f\mathbf{T})\cdot d\mathbf{A} - \int_V \mathbf{T}\cdot(\boldsymbol{\nabla}f)d\tau.
\end{aligned}
\tag{1.74}
$$

Here we transform the integrand from the product of one function (f) and one derivative $\boldsymbol{\nabla}\cdot\mathbf{T}$ to another integrand of the product of one function that is orginally the derivative ($\mathbf{T}$) and one derivative which is orginally the function ($\boldsymbol{\nabla}f$), at a cost of a minus sign and a boundary term $\left(\oint_S (f\mathbf{T}) \cdot d\mathbf{A}\right)$, just like integration by part in ordinary derivatives, where

$$\int_a^b f(\frac{\mathrm{d}g}{\mathrm{d}x})dx = -\int_a^b g\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)dx + (fg)\Big|_a^b \tag{1.75}$$

comes from the product rule

$$\frac{\mathrm{d}}{\mathrm{d}x}(fg) = f\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right) + g\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right). \tag{1.76}$$

Similarly, we can show that

$$\int_S f(\boldsymbol{\nabla} \times \mathbf{T}) \cdot d\mathbf{A} = \int_S (\mathbf{T} \times (\boldsymbol{\nabla} f)) \cdot d\mathbf{A} + \oint_C f\mathbf{T} \cdot d\mathbf{r} \tag{1.77}$$

and

$$\int_V \mathbf{B} \cdot (\boldsymbol{\nabla} \times \mathbf{T}) d\tau = \int_V \mathbf{T} \cdot (\boldsymbol{\nabla} \times \mathbf{B}) d\tau + \oint_S (\mathbf{T} \times \mathbf{B}) \cdot d\mathbf{A}. \tag{1.78}$$

1.4.5 Second Derivatives

From the nature of gradient, divergence and curl, we can construct five species of second derivatives. They are

1. Divergence of gradient $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} t)$:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} t) = \left(\hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z}\right) \cdot \left(\hat{\mathbf{i}}\frac{\partial t}{\partial x} + \hat{\mathbf{j}}\frac{\partial t}{\partial y} + \hat{\mathbf{k}}\frac{\partial t}{\partial z}\right) = \frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2}, \tag{1.79}$$

In fact, the divergence of gradient operator is so frequently used in physics that it gets its own name and symbol known as the Laplacian:

$$\nabla^2 t = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} t). \tag{1.80}$$

With reference to eq. (1.80), the Laplacian of a scalar function in spherical coordinates is listed here for reference:

$$\nabla^2 t = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial t}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial t}{\partial \theta}\right) + \frac{1}{r^2 \sin\phi}\frac{\partial^2 t}{\partial \phi^2}. \tag{1.81}$$

In some case, we can simplify the expression by rewriting the fisrt term on the RHS as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial t}{\partial r}\right) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rt). \tag{1.82}$$

2. Curl of gradient:

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}t) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} & \frac{\partial t}{\partial z} \end{vmatrix} = \left(\frac{\partial^2 t}{\partial z \partial y} - \frac{\partial^2 t}{\partial y \partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial^2 t}{\partial x \partial z} - \frac{\partial^2 t}{\partial z \partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial^2 t}{\partial y \partial x} - \frac{\partial^2 t}{\partial x \partial y}\right)\hat{\mathbf{k}} = 0. \tag{1.83}$$

3. Gradient of divergence (not identical to the divergence of gradient):

$$\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{T}) \neq (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\mathbf{T}. \tag{1.84}$$

4. Divergence of curl:

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{T}) = \frac{\partial}{\partial x}\left(\frac{\partial T_z}{\partial y} - \frac{\partial T_y}{\partial z}\right) + \frac{\partial}{\partial z}\left(\frac{\partial T_x}{\partial z} - \frac{\partial T_z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y}\right) = 0. \tag{1.85}$$

5. Curl of curl:

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{T}) = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{T}) - \nabla^2\mathbf{T}. \tag{1.86}$$

Here, $\nabla^2\mathbf{T}$ is the Laplacian of a vector function defined as[5]

$$\nabla^2\mathbf{T} = (\nabla^2 T_x)\hat{\mathbf{i}} + (\nabla^2 T_y)\hat{\mathbf{j}} + (\nabla^2 T_z)\hat{\mathbf{k}}. \tag{1.87}$$

For the 5 second derivatives listed above, only eqs. (1.79) and (1.86) are used regularly in physics enough that they are worth remembering. All of the identities can be easily proven in index notation, as long as one remembers the basic rules

$$(\boldsymbol{\nabla}t)_i = \partial_i t, \quad (\boldsymbol{\nabla}\cdot\mathbf{T})_i = \partial_i T_i \text{ and } (\boldsymbol{\nabla}\times\mathbf{T})_i = \epsilon_{ijk}\partial_j T_k. \tag{1.88}$$

To find $\boldsymbol{\nabla}r^p$, for example, we have

$$\partial_i r^p = \partial_i (x_j x_j)^{\frac{p}{2}} = \frac{p}{2}(x_j x_j)^{\frac{p}{2}-1}\partial_i(x_j x_j) = p r^{p-2} x_i \implies \boldsymbol{\nabla}r^p = p r^{p-1}\hat{\mathbf{r}}. \tag{1.89}$$

To prove that $\boldsymbol{\nabla}\times\boldsymbol{\nabla}t = 0$, for example, we can show that

$$(\boldsymbol{\nabla}\times\boldsymbol{\nabla}t)_i = \epsilon_{ijk}\partial_i\partial_j t = 0, \tag{1.90}$$

since the ϵ_{ijk} is antisymmetric over ij, but the partial derivatives $\partial_i\partial_j$ are symmetric, so the terms like $\partial_1\partial_2 - \partial_2\partial_1$ cancel.

[5]In fact, eq. (1.86) is often used to define the Laplacian of a vector, since eq. (1.80) makes explicit reference to Cartesian coordinates.

1.4.6 Dirac Delta Function

The motivation of the dirac delta function comes from the result of the divergence of the vector function $\mathbf{T} = \hat{\mathbf{r}}/r^2$:

$$\boldsymbol{\nabla} \cdot \left(\frac{\hat{\mathbf{r}}}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left(\frac{1}{r^2}\right)\right) = 0, \tag{1.91}$$

which is not consistent with the result we would expect from the divergence theorem, since the suface integral over a hypothetical sphere with raidus R is

$$\oint_S \left(\frac{\hat{\mathbf{r}}}{R^2}\right) \cdot (R^2 \sin\theta d\theta d\phi \hat{\mathbf{r}}) = 4\pi. \tag{1.92}$$

The root of this problem lies on the fact that the function itself blows up at the origin, so while it is true that the divergence of this function equals to 0 at every point, it does not apply to the origin. The 4π contribution in comes entirely from the orgin.

To describe this behaviour, we introduce the dirac delta function which is defined by

$$\delta(x) = \begin{cases} \infty & \text{if} \quad x = 0 \\ 0 & \text{if} \quad x \neq 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(x)dx = 1.^6 \tag{1.93}$$

The appearance of this function is shown in fig. 1.6.

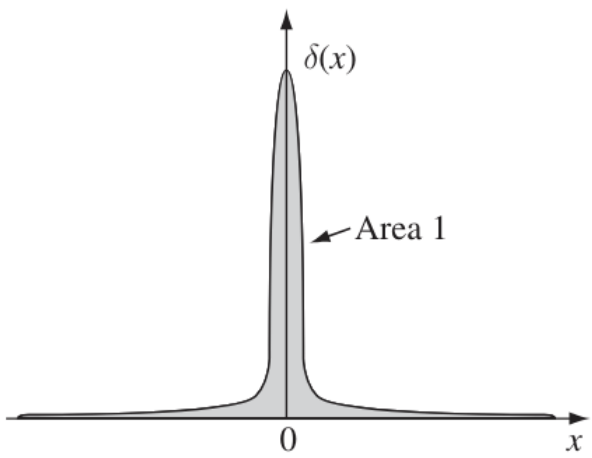


Figure 1.6

Let $f(x)$ to be an ordinary (continuous) function, then it follows that

$$f(x)\delta(x) = f(0)\delta(x), \tag{1.94}$$

since $f(x)\delta(x) \neq 0$ only if $x = 0$.

[6] Therefore, $\delta(x)$ has the dimension of the inverse of its argument

Integrating the above equation,

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = \int_{-\infty}^{+\infty} f(0)\delta(x)dx = f(0)\int_{-\infty}^{+\infty} \delta(x)dx = f(0). \tag{1.95}$$

Using the integral, the dirac delta function picks out the value of $f(x)$ at the origin.

By changing the variable in eq. (1.95), we can pick out the value of $f(x)$ at any arbitary point $x = a$

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a) \tag{1.96}$$

Example: Dirac Delta Function (1).

Question: Show that $\delta(kx) = \frac{1}{|k|}\delta(x)$.

Solution: Let $u = kx$, then

$$\int_{-\infty}^{+\infty} f(x)\delta(kx)dx = \frac{1}{|k|}\int_{-\infty}^{+\infty} f\left(\frac{u}{k}\right)\delta(u)du = \frac{1}{|k|}f(0) = \int_{-\infty}^{+\infty} f(x)\frac{\delta(x)}{|k|}dx. \tag{1.97}$$

The absolute signs are necessary since if k is negative then the lower and the upper limit of integration will be reversed. By comparing the integrands, we yield the desired result.

Example: Dirac Delta Function (2).

Question: Show that $-\delta(x) = x\frac{\mathrm{d}}{\mathrm{d}x}(\delta(x))$

Solution:

Example: Dirac Delta Function (3).

Question: Show that $\delta(x) = \frac{\mathrm{d}\theta}{\mathrm{d}x}$, where $\theta(x)$ is the Heaviside step function defined as $\theta(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0. \end{cases}$

Solution:

It is natural to generalize $\delta(x)$ to three dimensions as follows:

$$\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z) \tag{1.98}$$

The characteristic equations the three-dimensional delta function are then

$$\int_{all\ space} \delta^3(\mathbf{r})d\tau = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \delta(x)\delta(y)\delta(z)\ dxdydz = 1 \tag{1.99}$$

and

$$\int_{all\ space} f(\mathbf{r})\delta^3(\mathbf{r}-\mathbf{a}) = f(\mathbf{a}). \tag{1.100}$$

As in the one-dimensional case, the $\delta^3(\mathbf{r}-\mathbf{a})$ picks out the value of $f(\mathbf{r})$ at $\mathbf{a}$.

The inconsistency can now be resolved as the divergence of the vector function $\mathbf{T} = \frac{\hat{\mathbf{r}}}{r^2}$ is in fact

$$\boldsymbol{\nabla}\cdot\frac{\hat{\mathbf{r}}}{r^2} = 4\pi\delta^3(\mathbf{r}). \tag{1.101}$$

which equals to 0 except in the origin where it equals to 4π.

In general,

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2}\right) = 4\pi\delta^3(\boldsymbol{\mathscr{r}}). \tag{1.102}$$

Here, the derivative is evaluated with respect to $\mathbf{r}$ and $\mathbf{r}'$ is fixed.

From eq. (1.59), we have also

$$\nabla^2\left(\frac{1}{\mathscr{r}}\right) = -4\pi\delta^3(\boldsymbol{\mathscr{r}}). \tag{1.103}$$

With this equation in hand, we can find prove that the potential created by the charge distribution $\rho(\mathbf{r}')$ satisfies the Poisson's equation

$$\begin{aligned}\nabla^2 V(\mathbf{r}) &= \nabla^2\left(\frac{1}{4\pi\epsilon_0}\int\frac{\rho(\mathbf{r}')}{\mathscr{r}}d\tau'\right) = \frac{1}{4\pi\epsilon_0}\int\rho(\mathbf{r}')\nabla^2\left(\frac{1}{\mathscr{r}}\right)d\tau' \\ &= \frac{1}{4\pi\epsilon_0}\int\rho(\mathbf{r}')(-4\pi\delta^3(\boldsymbol{\mathscr{r}}))d\tau' = -\frac{\rho(\mathbf{r})}{\epsilon_0}.\end{aligned} \tag{1.104}$$

Example: Griffith (5th ed.) Problem 1.47

Question: Find the charge density $\rho(\mathbf{r})$ for

1. a point charge q located at $\mathbf{r}'$,
2. an electric dipole consisting of a point charge $-q$ at the origin and $+q$ at $\mathbf{a}$ and
3. an infinitesimal sphere of charge q and radius a centered at the origin

Solution:

1. $\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{r}')$

2. $\rho(\mathbf{r}) = -q\delta(\mathbf{r}) + q\delta(\mathbf{r} - \mathbf{a})$

3. $\rho(r) = A\delta(r - a)$, where A can be determined by $\int_V \rho(r)d\tau = \int A\delta(r - a)4\pi r^2 dr = 4\pi Aa = q$, thus $A = \frac{q}{4\pi a}$. Therefore, $\rho(r) = \frac{q}{4\pi a}\delta(r - a)$.

Example: Vortex Flow.

Question: Consider vortex flow in an incompressible fluid with a velocity field

$$\mathbf{v} = \frac{1}{\rho}\hat{\boldsymbol{\phi}}. \tag{1.105}$$

Explain the discrepancy in Stokes' theorem using Dirac delta function.

Solution: From this velocity field $\boldsymbol{\nabla} \times \mathbf{v} = 0$ everywhere except on the axis $\rho = 0$ where $\mathbf{v}$ has a singularity. Therefore $\oint_C \mathbf{v} \cdot d\mathbf{r} = 0$ for any path C that does not enclose the vortex line on the axis and 2π if C does enclose the axis.
In order for Stokes' theorem to be valid for all paths C, we therefore set

$$\boldsymbol{\nabla} \times \mathbf{v} = 2\pi\delta(\rho). \tag{1.106}$$

So $\mathbf{v}$ is locally conservative over any region that excludess the axis, but is globally non-conservative.

1.4.7 Helmholtz Theorem

Theorem 1.4.1 (Helmoholtz Theorem)**.** *Given the divergence and curl of a differentiable vector function $\mathbf{T}(\mathbf{r})$, it can be written as some gradient of some scalar function $\Phi(\mathbf{r})$ plus some vector function $\mathbf{A}(\mathbf{r})$ if $\mathbf{T}$ goes to zero faster than $1/r$ as $r \to \infty$, where Φ and $\mathbf{A}$ are given by*

$$\Phi(\mathbf{r}) = \frac{1}{4\pi}\int_{all\ space} \frac{\boldsymbol{\nabla}' \cdot \mathbf{T}(\mathbf{r}')}{\mathscr{r}} d\tau' + C \quad and \quad \mathbf{A}(\mathbf{r}) = \frac{1}{4\pi}\int_{all\ space} \frac{\boldsymbol{\nabla}' \times \mathbf{T}(\mathbf{r}')}{\mathscr{r}} d\tau' + \mathbf{C} + \boldsymbol{\nabla} D. \tag{1.107}$$

Here C and $\mathbf{C}$ are some constant scalar and vector functions respectively, which does not change $\mathbf{T} = \boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times \mathbf{A}$, since the gradient and curl of some constant function is zero. $\mathbf{D}$, on the other hand, can be any function since the curl of the gradient is always zero.

Proof. We start from the existence theorem of the Poisson's equation[7] and states that

[7] The proof of this theorem is mathematically advanced thus is omitted here but somewhat trivial because it simply implies that we can assign each point in space an electric potential for any arbitrary charge distribution.

there exists an unique solution to the equation

$$-\nabla^2\Phi = -\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\Phi) = \boldsymbol{\nabla}\cdot\mathbf{T} \implies \boldsymbol{\nabla}\cdot(\mathbf{T}+\boldsymbol{\nabla}\Phi) = 0. \tag{1.108}$$

But since the divergence of a curl is always zero, we have

$$\mathbf{T} = -\boldsymbol{\nabla}\Phi + \boldsymbol{\nabla}\times\mathbf{A} \tag{1.109}$$

Testing the divergence of $\mathbf{T}$, we have[8]

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\mathbf{T} &= -\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\Phi) + \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) = -\nabla^2\Phi \\
&= -\frac{1}{4\pi}\int \boldsymbol{\nabla}'\cdot\mathbf{T}(\mathbf{r}')\nabla^2\left(\frac{1}{\mathscr{r}}\right)d\tau' \\
&= \int \boldsymbol{\nabla}'\cdot\mathbf{T}(\mathbf{r}')\delta^3(\mathbf{r}-\mathbf{r}')d\tau' = \boldsymbol{\nabla}\cdot\mathbf{T}(\mathbf{r}).
\end{aligned} \tag{1.110}$$

where $\boldsymbol{\nabla}'\cdot\mathbf{T}(\mathbf{r}')d\tau'$ is taken out of the Laplacian since it depends only on $\mathbf{r}'$ but not $\mathbf{r}$.

Testing the curl,

$$\boldsymbol{\nabla}\times\mathbf{T} = -\boldsymbol{\nabla}\times\boldsymbol{\nabla}\Phi + \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) = -\nabla^2\mathbf{A} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}). \tag{1.111}$$

Now, the first term is simply $\boldsymbol{\nabla}\times\mathbf{T}$ in a similar fashion as the divergence, and the second term is zero, since

$$\begin{aligned}
4\pi(\boldsymbol{\nabla}\cdot\mathbf{A}) &= \int \boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}')}{\mathscr{r}}\right)d\tau' \\
&= \int(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot\boldsymbol{\nabla}\left(\frac{1}{\mathscr{r}}\right)d\tau' = -\int(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot\boldsymbol{\nabla}'\left(\frac{1}{\mathscr{r}}\right)d\tau' \\
&= \int\frac{1}{\mathscr{r}}\boldsymbol{\nabla}'\cdot(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))d\tau - \oint\frac{1}{\mathscr{r}}(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot d\mathbf{A} \\
&= -\int\frac{1}{\mathscr{r}}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{T}(r'))d\tau = 0.
\end{aligned} \tag{1.112}$$

where we used $\boldsymbol{\nabla}(1/\mathscr{r}) = -\boldsymbol{\nabla}'(1/\mathscr{r})$ since the derivative of $\mathscr{r} = |\mathbf{r}-\mathbf{r}'|$ with respect to the primed coordinates differ by a sign from that with respect to unprimed coordinates.[9] Also, we assumed that $\boldsymbol{\nabla}\times\mathbf{T}$ goes to zero faster than $1/r^2$ as $r\to\infty$ so the surface integral goes to zero.

Of course we still have to prove that the given forms for the divergence and curl of $\mathbf{T}$ converge. At the large r' limit, they have the form

$$\int\frac{X(r')}{r'}r'^2dr', \tag{1.113}$$

[8] Here and after, we will omit the limit of the integrals.

[9] More precisely, $\frac{\partial}{\partial x}f(x-x') = -\frac{\partial}{\partial x'}f(x-x')$.

where $X(r')$ stands for $\boldsymbol{\nabla} \cdot \mathbf{T}$ or $\boldsymbol{\nabla} \times \mathbf{T}$.

Therefore, we also require that the divergence and curl of $\mathbf{T}$ goes to zero more repaidly than $1/r^2$ as $r \to \infty$ for the proof to hold.

The solution is unique in a sense that no function that has zero divergence and zero curl everywhere goes to zero at infinity. Therefore no constant function can be added to $\mathbf{T}(\mathbf{r})$ which doesn't change the divergence and curl of $\mathbf{T}$.

As a matter of fact, any differentiable vector function regardless of its behavior at infinity can be written as a gradient and a curl, as proved above, just that there solution is not given by the Helmholtz theorem.

$\square$

2

Integration

2.1 Multiple Integration

Multiple integration includes single, double and triple integrations, and they sum up the scalar field times an infinitesimal quantity, namely $f(x)dx$, $f(x,y)dA = f(x,y)dxdy$ and $f(x,y,z)dxdydz$.[1][2][3]

Refering to fig. 2.1, we can see that a double integral can be evaluated in two different ways.

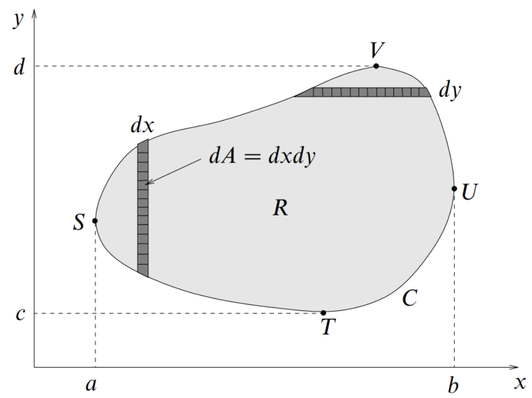


Figure 2.1

The first is to sum up all the horizontal strips, then

[1]This pattern can obviously be generalized to higher dimensions, but in physics we are only concerned with the three cases.

[2]Depending on whether $dA = dxdy$ and $dV = dxdydz$ are written explicitly, two or three integral signs will be used.

[3]Here the argument (x,y,z) can be unambigously interchanged by $(\mathbf{r})$, where $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ is the position vector. The relation between the two forms can be given explicitly as $x = r_x, y = r_y, z = r_z$. The only difference is that the latter notation emphasize that the (scalar or vector field) depends only on the location in space and does not specify the usage of Cartesian coordinates and is thus more abstract but general.

$$I = \int_c^d \left(\int_{x_1(y)}^{x_2(y)} f(x,y)dx \right) dy, \tag{2.1}$$

where $x_1(y)$ and $x_2(y)$ are the equations of the curves TSV and TUV respectively, and the parenthesis is usually omitted.

The second way is to sum up all the vertical strips, then

$$I = \int_a^b \int_{y_1(x)}^{y_2(x)} f(x,y)dydx, \tag{2.2}$$

where $y_1(x)$ and $y_2(x)$ are the equations of the curves STU and SVU respectively. Thus the order of integration does not matter (given that $f(x,y)$ behaves properly) as long as the limits of integrals are correct.

Double integration can be regarded as calculating a volume which the base is flat, and the height is $f(x,y)$, as the same way as how a single integration is almost always used to find the area, where the side is straight, and the height being $f(x)$.

If the integrand $f(x,y)$ can be written as a product of functions of x and y separately, *i.e.*, $f(x,y) = f(x)f(y)$, then we can evaluate each integral independently as $\left(\int f(x)dx\right)\left(\int f(y)dy\right)$.[4]

The above analysis can be extended to triple integration easily.

Example: Volume of a Tetrahedron.

Question: Find the volume of the tetrahedron bounded by the three coordinate surfaces $x = 0, y = 0$ and $z = 0$ and the plane $x/a + y/b + z/c = 1$ as shown in fig. 2.2.

Solution: The volume can be written as a double or triple integral, depending on how one inteprets the summing process. In triple integral,

$$V = \int_0^a \int_0^{b-\frac{bx}{a}} \int_0^{c(1-\frac{y}{b}-\frac{x}{a})} dxdydz, \tag{2.3}$$

which can be easily reduced to a double integral by evaluating the integral involving z as

$$V = \int_0^a \int_0^{b-\frac{bx}{a}} c\left(a - \frac{y}{b} - \frac{x}{a}\right) dxdy \tag{2.4}$$

The result is $V = abc/6$.

[4]The same applies to summation, as integration and summation are essentially the same process.

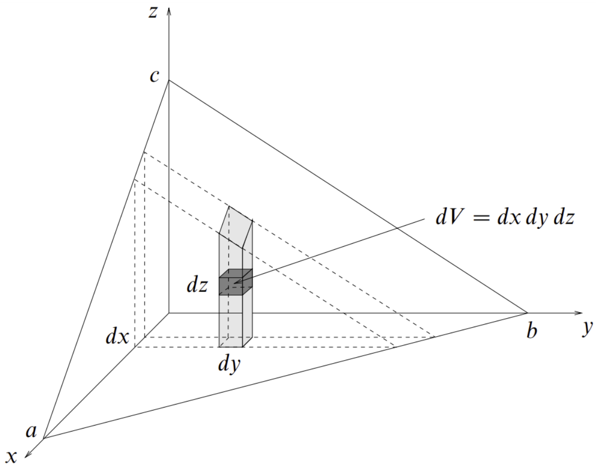


Figure 2.2

2.2 Line, Surface and Volume Integration

Line, surface and volume sum up scalar or vector field times an infinitesimal quantity. The infinitisimal quantities are (in Cartesian coordinates), $d\mathbf{r} = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}}$, $d\mathbf{S}$[5] and $dV = dxdydz$,[6] out of which the first two are vectors and the last one is scalar.[7]

2.2.1 Curves and Surfaces

In evaluating line and surface integrals, we always have to specify what curve and surface are we integrating over, and we therefore need to devise a way to describe these curves and surfaces. For both curves and surfaces, they can be represented in parameterized form, where an extra variable(s) is(are) introduced to specify a point on the curve or surface. Alternatively, they can be represented algebrically.

Curves

A curve in space[8] can be described by the vector $\mathbf{r}(u)$ joining the origin O of a coordinate system to a point on the curve. It can be viewed as a map$f : \mathbb{R} \to \mathbb{R}^3$, where we assign each parameter u in $\mathbb{R}$ a corresponding vector in $\mathbb{R}^3$. As the parameter u varies, the end-point on the curve moves along the curve.[9] Some common examples of parameterized curves include

$$\mathbf{r}(t) = (\cos t, \sin t, t), \tag{2.5}$$

[5]The general form of $d\mathbf{S}$ is not trivial. It is exactly this property which makes surface integral the hardest type to deal with.

[6]dV is a scalar since an infinitesimal volume has no preferred direction in three-dimensional space.

[7]Note that if the final form of integrands contain vectors then we must convert it into Cartesian coordinates, since it is the only coordiante system where the unit vectors are fixed (and so can be brought out of the integral sign).

[8]In the following discussion we assume that the surfaces are orientable, *i.e.,* there is a consistent choice for the normal vector $\hat{\mathbf{n}}$ which varies smoothly over the surface. Exmples of non-orientable surfaces include the Möbius strip and the klein bottle.

[9]A useful visualization is imagine sliding the sidebar for u on Desmos and the point corresponding point will move along the curve accordingly.

which describes a helical path of a particle if t is time. We also note that the choice of parameterization is not unique. The parameterized curve $\mathbf{r}(t) = (\cos\lambda t, \sin\lambda t, \lambda t)$ describes the same helix for all $\lambda \neq 0$.[10]

Another example is the unit circle, which can be parameterized by

$$\mathbf{r}(t) = (\cos t, \sin t) \quad \text{or} \quad \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right).^{11} \tag{2.6}$$

Often $x = t$ so we can simply treat x as the parameter and express the integral in terms of x.

Alternatively, a space curve in three-dimensional space can be represented by two simultaneous equations algebrically as $F(x, y, z) = G(x, y, z) = 0$. For example, the first example above can be described equivalently as $x - \cos z = y - \sin z = 0$. However, we lost the picture of the particle's position evloving through space, and it is not a convenient form to work in integrals.

Surfaces

A point on a surface, on the other hand, requires two parmeters u and v, and it can be viewe as a map $f : \mathbb{R}^2 \to \mathbb{R}^3$. A common example of parameterized curves is

$$\mathbf{r}(\theta, \phi) = R(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{2.7}$$

which describes the surface of a sphere with radius R.

Alternatively, a space surface in three-dimensional space can be represented by a single equation algebrically as $F(x, y, z) = 0$.

There are generally two ways to compute an infinitesimal surface dS. If the parametric form of the surface is given, then refer to fig. 2.3, since v is constant along KL, the line element KL can be written as

$$d\mathbf{r}_{KL} = \frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv = \frac{\partial \mathbf{r}}{\partial u}du = \frac{\partial x}{\partial u}du\hat{\mathbf{i}} + \frac{\partial y}{\partial u}du\hat{\mathbf{j}}. \tag{2.8}$$

Similarly, $d\mathbf{r}_{KN} = (\partial x/\partial v)dv\hat{\mathbf{i}} + (\partial y/\partial v)dv\hat{\mathbf{j}}$, thus the area of the parallelogram $KLMN$ is given by

[10]Sometimes, one may ponder upon an unwise choice of parameterization in which the derivative $\dot{\mathbf{r}}$ vanishes at some point. For example, consider the curve $\mathbf{r}(t) = (t^3, t^3)$, which is just the straight line $x = y$. However, $\dot{\mathbf{r}}$ vanishes at $t = 0$. Using the parameterization $\mathbf{r}(t) = (t, t)$ which describes the same curve, we find that this vanishing is not a property of the curve but of our choice of parameterization. We will assume that the curves we encounter are always regularly parameterized, which ensures proper definition of relevant parameters such as the tangent vector and the arc length.

[11]With the exception of the point $(0, 1)$ since we have $(t^2-1)/(t^2+1) < 0$ for all t, so the y-coordinate would not be equal to 1 unless $t = \infty$, but t in reality only tends to infinity.

$$dS_{u,v} = |d\mathbf{r}_{KL} \times d\mathbf{r}_{KN}| = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| dudv = \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right| dudv = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dudv. \tag{2.9}$$

Since both $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ lie in the tangent plane to the surface at $\mathbf{r}$, the direction of the infinitesimal surface is also encoded in the cross product

$$d\mathbf{S}_{u,v} = \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) dudv. \tag{2.10}$$

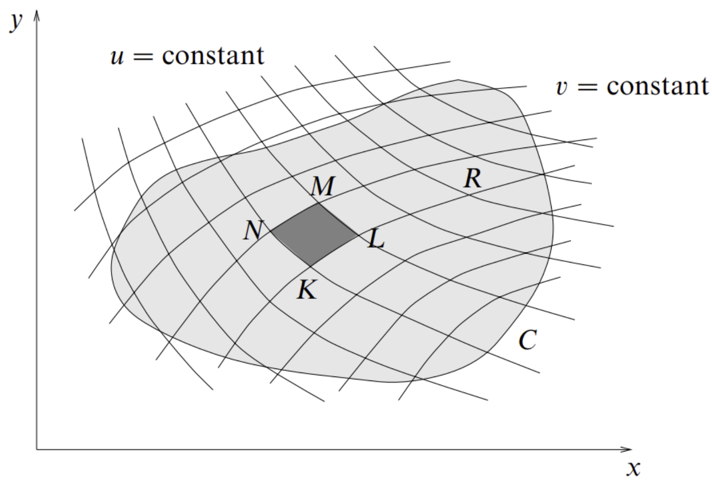


Figure 2.3

If the algebraic form of the surface is given, then we usually project the surface S onto the xy-plane (sometimes the yz- or xz- planes are more useful). From fig. 2.4, we see that

$$dA = |\cos\alpha| dS \implies dS = \frac{dA}{\hat{\mathbf{n}} \cdot \mathbf{k}} = \frac{dA}{\frac{\boldsymbol{\nabla} f}{|\boldsymbol{\nabla} f|} \cdot \hat{\mathbf{k}}} = \frac{\partial z}{\partial f} dA |\boldsymbol{\nabla} f| \implies d\mathbf{S} = \frac{\partial z}{\partial f} dA \boldsymbol{\nabla} f, \tag{2.11}$$

where α is the angle between the unit vector $\hat{\mathbf{k}}$ in the z-direction and the unit normal $\hat{\mathbf{n}}$ to the surface, and $f(x, y, z) = 0$ is the equation which describe the surface.[12]

[12] If the equation is in the form $z = g(x, y) \implies f(x, y, z) = z - g(x, y) = 0$, then $\frac{\partial f}{\partial z} = 1$ and the relation between surface and area becomes $dS = |\boldsymbol{\nabla} f| dA$.

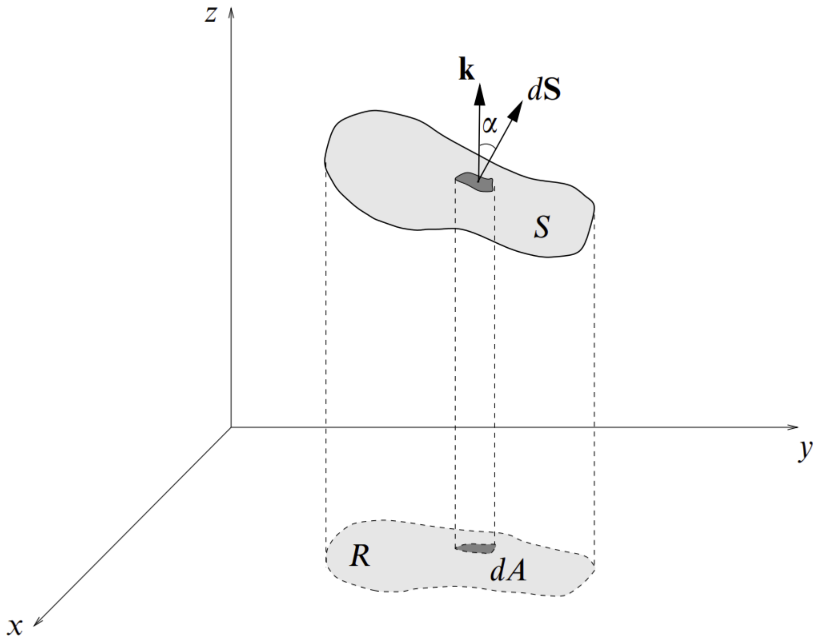


Figure 2.4

Using the above equation, we can convert any surface integral over S as a double integral over the region R in the xy-plane.

Note that in the above discussion, however, tht we assumed any line parallel to the z-axis only intersects S once. If this is not the case, we must split up the surface into smaller surfaces S_1, S_2 *etc.* Also, sometimes instead of projecting the surface onto the xy-plane, it might be easier to project it onto the zx-plane or the yz-plane.

A surface that comes up all the time is a spherical surface with raidus R. Parametrically, the surface is described as

$$\mathbf{r}(\theta, \phi) = R\hat{\mathbf{r}} = R(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{2.12}$$

so we have

$$\begin{aligned} d\mathbf{S} &= \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right) d\theta d\phi = R^2 \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix} \times \begin{pmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{pmatrix} d\theta d\phi \\ &= R^2 \sin\theta d\theta d\phi (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) = R^2 \sin\theta d\theta d\phi \hat{\mathbf{r}}. \end{aligned} \tag{2.13}$$

Algebrically, the surface is described as

$$f(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0, \tag{2.14}$$

so we have

$$
\begin{aligned}
d\mathbf{S} = \frac{\partial z}{\partial f} dA \boldsymbol{\nabla} f = \frac{1}{2z} dx dy 2(x,y,z) = \left| \frac{\partial(x,y)}{\partial(\theta,\phi)} \right| d\theta d\phi (\frac{x}{z}, \frac{y}{z}, 1) \\
= (R^2 \cos\theta \sin\phi)(\tan\theta\cos\phi, \tan\theta\sin\phi, 1) d\theta d\phi = r^2 \sin\theta d\theta d\phi \hat{\mathbf{r}}.
\end{aligned}
\tag{2.15}
$$

Frenet-Serret Equations

We can also define an orthogonal coordinate system that is intrinsic to the curve. For the three axes, an obvious choice would be the tangential direction, whose unit vector can be defined by

$$
\hat{\mathbf{t}} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{|d\mathbf{r}|}. \tag{2.16}
$$

The curvature κ and the pricinpal normal unit vector are defined together from the rate of change of $\hat{\mathbf{t}}$ with respect to s

$$
\kappa \hat{\mathbf{n}} = \frac{d\hat{\mathbf{t}}}{ds} = \frac{d^2\mathbf{r}}{ds^2} \tag{2.17}
$$

and the radius of curvature is defined as $\rho = 1/\kappa$. We see that the radius of curvature measures the second derivative, capturing how much the curve fails to be a straight line. Both $\hat{\mathbf{n}}$ and $\hat{\mathbf{t}}$ lie on the plane that the small segements around $\mathbf{r}$ form.

The binormal unit vector is then defined as the unit vector perpendicular to both $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$

$$
\hat{\mathbf{b}} = \hat{\mathbf{t}} \times \hat{\mathbf{n}}. \tag{2.18}
$$

We can similarly define the torsion τ as the rate of change of $\hat{\mathbf{b}}$ with respect to s

$$
-\tau \hat{\mathbf{n}} = \frac{d\hat{\mathbf{b}}}{ds}, \tag{2.19}
$$

and radius of torsion as $\sigma = 1/\tau$. We see that the radius of torsion measures the third derivative, capturing how much the curve fails to be planar.

In summary, $\hat{\mathbf{t}}, \hat{\mathbf{n}}$ and $\hat{\mathbf{b}}$ and their derivatives with respect to s are related to one another by the relations (called the Frenet-Serret formulae)

$$
\frac{d\hat{\mathbf{t}}}{ds} = \kappa \hat{\mathbf{n}}, \quad \frac{d\hat{\mathbf{n}}}{ds} = \tau \hat{\mathbf{b}} - \kappa \hat{\mathbf{t}} \quad \text{and} \quad \frac{d\hat{\mathbf{b}}}{ds} = -\tau \hat{\mathbf{n}}. \tag{2.20}
$$

Note that the shape of a curve is completely determined by κ and τ, for example, a curve is planar if $\tau = 0$, and a curve in which $\hat{\mathbf{t}}$ makes a constant angle with a fixed direction, *i.e.,* when either κ or τ is independent of s.

Example: Acceleration of a Particle.

Question: Show that the acceleration of a particle travelling along a trajectory $\mathbf{r}(t)$ is given by

$$\mathbf{a}(t) = \frac{dv}{dt}\hat{\mathbf{t}} + \frac{v^2}{\rho}\hat{\mathbf{n}}. \tag{2.21}$$

Solution: The velocity of the particle is $\mathbf{v}(t) = v\hat{\mathbf{t}}$, so the acceleration is

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{dv}{dt}\hat{\mathbf{t}} + v\frac{d\hat{\mathbf{t}}}{dt} = \frac{dv}{dt}\hat{\mathbf{t}} + v\frac{ds}{dt}\frac{d\hat{\mathbf{t}}}{ds} = \frac{dv}{dt}\hat{\mathbf{t}} + \frac{v^2}{\rho}\hat{\mathbf{n}}. \tag{2.22}$$

2.2.2 Line Integrals

Integrals along a line can involve vector and scalar fields. There are four kinds of line integrals, namely

$$\int_C f(\mathbf{r})ds, \quad \int_C f(\mathbf{r})d\mathbf{r}, \quad \int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} \text{ and } \int_C \mathbf{F}(\mathbf{r}) \times d\mathbf{r}. \tag{2.23}$$

The curve C can be open, *i.e.,* the beginning and end point are not the same; or closed, where C is a closed loop and we will add a circle to the integral sign as $\oint_C$. For a closed curve the direction of integration is conventionally taken to be anticlockwise.

Note that for the first type of the four line integrals above $ds = |d\mathbf{r}|$ is an intrinsically positive quantity, which represent an infinitesimal arc length in $\mathbb{R}^3$. On the other hand, dx in a single integral represents a straight infinitesimal interval element in $\mathbb{R}$, so it carries sign. The integral is typically carried out as

$$s = \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{r})ds = \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{r})|d\mathbf{r}(u)| = \int_{u_{\mathbf{a}}}^{u_{\mathbf{b}}} f(\mathbf{r})\left|\frac{d\mathbf{r}}{du}\right||du| = \int_{u_{\mathbf{a}}}^{u_{\mathbf{b}}} f(\mathbf{r})\sqrt{\left(\frac{dx}{du}\right)^2 + \left(\frac{dy}{du}\right)^2 + \left(\frac{dz}{du}\right)^2}|du|, \tag{2.24}$$

If $u_{\mathbf{b}} > u_{\mathbf{a}}$, then du is positive and we can omit the absolute sign of du. If $u_{\mathbf{b}} < u_{\mathbf{a}}$, then du is negative, and we can omit the absolute sign of du if we add a negative sign in front. This is equivalent of reversing the limit of the integrations ($u_{\mathbf{a}}$ and $u_{\mathbf{b}}$). Therefore, the limit of integration does not matter when evaluating the first type of the four line integrals above.

Example: Line Integral (1).

Question: Consider the scalar function $\phi(\mathbf{r}) = 2$. Evaluate the line integral $I = \int_C \phi ds$, along a straight line from $(0, 1)$ to $(1, 0)$ and in the reverse direction.

Solution: The forward and backward parameterizations are

$$\mathbf{r}(u) = (u, 0) \quad \text{and} \quad \mathbf{r}'(u) = (1-u, 0). \tag{2.25}$$

The forward line integral is

$$I = \int_C \phi ds = \int_0^1 2\left|\frac{d\mathbf{r}}{du}\right| du = 2\int_0^1 \sqrt{1^2 + 0^2} du = 2. \tag{2.26}$$

The backward line integral is

$$I = \int_C \phi ds = \int_0^1 2\left|\frac{d\mathbf{r}}{du}\right| du = 2\int_0^1 \sqrt{1^2 + 0^2} du = 2. \tag{2.27}$$

The limits of the line integral in both cases are 0 to 1, due to the different paratermization used. Of course, we can use the same parameterization in the second line integral as the first one, where the limits of integration will be from 1 to 0, but then we should not replace $|du|$ by du, but should add a negative sign, which gives the same answer.

Example: Line Integral (2).

Question: Evaluate the line integral $I = \int_C \mathbf{F} \cdot d\mathbf{r}$ from $(1,1)$ to $(4,2)$, where $\mathbf{F} = (x+y)\hat{\mathbf{i}} + (y-x)\hat{\mathbf{j}}$, along

1. the parabola $y^2 = x$,
2. the curve $x = 2u^2 + u + 1, y = 1 + u^2$, and
3. the line $y = 1$ from $(1,1)$ to $(4,1)$, followed by the line $x = 4$ from $(4,1)$ to $(4,2)$.

Solution: Evaluating the dot product explicitly, we have

$$I = \int_{(1,1)}^{(4,2)} (x+y)dx + (y-x)dy. \tag{2.28}$$

1. Along the parabola $y^2 = x$, we have $2ydy = dx$, so

$$I = \int_1^2 ((y^2 + y)2y + (y - y^2))dy = \frac{34}{3}. \tag{2.29}$$

2. We have $dx = (4u+1)du$ and $dy = 2udu$, so

$$I = \int_0^1 ((3u^2 + u + 2)(4u + 1) - (u^2 + u)(2u))du = \frac{32}{3}. \tag{2.30}$$

3. We split the integral into two parts, then

$$
\begin{aligned}
I &= \int_{(1,1)}^{(4,1)} ((x+y)dx + (y-x)dy) + \int_{(4,1)}^{(4,2)} ((x+y)dx + (y-x)dy) \\
&= \int_1^4 (x+1)dx + \int_1^2 (y-4)dy = 8.
\end{aligned}
\tag{2.31}
$$

Example: Line Integrals (3).

Question: Evaluate the line integral $I = \oint_C xdy$, where C is the circle in the xy-plane defined by $x^2 + y^2 = a^2, z = 0$.

Solution: Since x is not a single-valued function of y, we must divide the path into two parts with $x = +\sqrt{a^2 - y^2}$ for $x \geq 0$ and $x = -\sqrt{a^2 - y^2}$ for $x \leq 0$. So

$$
I = \int_{-a}^{a} \sqrt{a^2 - y^2} dy + \int_{a}^{-a} (-\sqrt{a^2 - y^2}) dy = \pi a^2. \tag{2.32}
$$

Alternatively, we can represent the entire circle parametrically, by $x = a\cos\phi, y = a\sin\phi$ with ϕ running from 0 to 2π and we have

$$
I = a^2 \int_0^{2\pi} \cos^2\phi d\phi = \pi a^2. \tag{2.33}
$$

Example: Line Integral (4).

Question: Evaluate the line integral $I = \int_C (x - y)^2 ds$, where C is the semicircle of radius a running from $A = (a, 0)$ to $B = (-a, 0)$ and for which $y \geq 0$.

Solution: Introducing a parametric variable ϕ running from 0 to 2π, we have

$$
\mathbf{r}(\phi) = a\cos\phi\hat{\mathbf{i}} + a\sin\phi\hat{\mathbf{j}} \quad \text{and} \quad ds = \frac{d\mathbf{r}}{d\phi} d\phi = a d\phi. \tag{2.34}
$$

Thus

$$
I = \int_0^{\pi} a^3 (1 - \sin 2\phi) d\phi = \pi a^3. \tag{2.35}
$$

Example: Line Integral (5).

Question: Evaluate the integral $\oint_C (dx + xydz)$, where C is the closed curve forming the boundary of the surface given by the equation $z = x^2 - y^2$, for $x^2 + y^2 \leq 1$.

Solution: Parametrizing the curve in terms of θ, the integral becomes

$$\oint_C (dx + xydz) = \int_0^{2\pi} (-\sin\theta + \sin\theta\cos\theta(-2\cos\theta\sin\theta - 2\sin\theta\cos\theta))d\theta = -\pi. \tag{2.36}$$

Alternatively, we can use Stokes' theorem to rewrite the integral as

$$\oint_C (dx + xydz) = \int_S \boldsymbol{\nabla} \times \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}, \tag{2.37}$$

where $\mathbf{F}(\mathbf{r}) = (1, 0, xy) \implies \boldsymbol{\nabla} \times \mathbf{F}(\mathbf{r}) = (x, -y, 0)$. The infinitesimal surface is

$$d\mathbf{S} = \boldsymbol{\nabla} g(\mathbf{r}) = (-2x, 2y, 1), \tag{2.38}$$

where $g(\mathbf{r}) = z - f(x, y) = z - x^2 + y^2$.
Inserting these expressions into the integral we have

$$\int_S \boldsymbol{\nabla} \times \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = -2 \int_{S'} (x^2 + y^2)dxdy = -\pi. \tag{2.39}$$

2.3 Surface Integrals

As with line integrals, integrals over a surface can involve vector and scalar fields. There are four kinds of surface integrals, namely

$$\int_S \phi(\mathbf{r})dS, \quad \int_S \phi(\mathbf{r})d\mathbf{S}, \quad \int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} \quad \text{and} \quad \int_S \mathbf{F}(\mathbf{r}) \times d\mathbf{S}, \tag{2.40}$$

where $d\mathbf{S} = \hat{\mathbf{n}}dS$ is the infinitesimal vector area element. The direction of $\hat{\mathbf{n}}$ is conventionally assumed to be directed outwards from the closed volume if the surface is closed; or given by the right-hand rule if the surface is open and spans some perimeter curve C.

dA is used in double integral to represent a flat infinitesimal area element $\in \mathbb{R}^2$. dS is used in surface integral to represent an arbitrary infinitesimal surface element $\in \mathbb{R}^3$. So the main difference between the first integral and an ordinary double integral is that the surface S here can span through three dimensions while the area A in double integral is confined to two dimensions.

Entirely analogous to the fact that the first type of line integrals in the previous subsection is independent of the limits of integration (*i.e.,* the orientation of the curve), the firsty type of the four surface integrals above is also independent of the orientation of the surface (*i.e.,* the limits of integration).

Example: Surface Integral(1).

Question: Find the element of area on the surface of a sphere of radius a, and hence calculate the total surface area of the sphere.

Solution: We can represent a point $\mathbf{r}$ on the surface of the sphere in terms of the two parameters θ and ϕ as

$$\mathbf{r} = a\sin\theta\cos\phi\hat{\mathbf{i}} + a\sin\theta\sin\phi\hat{\mathbf{j}} + a\cos\theta\hat{\mathbf{k}}. \tag{2.41}$$

An infinitesimal area is given by

$$dS = \left|\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right| = a^2 \sin\theta d\theta d\phi \tag{2.42}$$

And the total surface area is then

$$S = \int_0^{\pi}\int_0^{2\pi} a^2 \sin\theta d\theta d\phi = 4\pi a^2. \tag{2.43}$$

Example: Surface Integral (2).

Question: Evaluate the surface integral $I = \int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$, where $F = x\hat{\mathbf{i}}$ and S is the surface of the hemisphere $x^2 + y^2 + z^2 = a^2$.

Solution: Parameterizing the surface by θ and ϕ,

$$\mathbf{F} \cdot d\mathbf{S} = x(\hat{\mathbf{i}} \cdot \hat{\mathbf{r}})dS = (a\sin\theta\cos\phi)(\sin\theta\cos\phi)(a^2\sin\theta d\theta d\phi), \tag{2.44}$$

we have

$$I = a^3 \int_0^{\frac{\pi}{2}} \sin^3\theta d\theta \int_0^{2\pi} \cos^2\phi d\phi = \frac{2\pi a^3}{3}. \tag{2.45}$$

Alternatively, we can describe the surface of the hemisphere as $f(\mathbf{r}) = x^2 + y^2 + z^2 - a^2 = 0$, so we have

$$|\boldsymbol{\nabla} f| = 2|\mathbf{r}| = 2a, \quad \frac{\partial f}{\partial z} = 2z = 2\sqrt{a^2 - x^2 - y^2} \text{ and } \hat{\mathbf{i}} \cdot \hat{\mathbf{r}} = \hat{\mathbf{i}} \cdot \frac{\mathbf{r}}{|\mathbf{r}|} = \frac{x}{a}. \tag{2.46}$$

Therfore the integral becomes

$$I = \int\int_R \frac{x^2}{\sqrt{a^2 - x^2 - y^2}} dxdy = \frac{2\pi a^3}{3}. \tag{2.47}$$

Example: Surface Integral (3).

Question: Evaluate $\int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$ over the surface bounded by the paraboloid $y = x^2 + z^2$ for $y \in [0, 1]$, and the disc $x^2 + z^2 \leq 1$ at $y = 1$, for $\mathbf{F} = (0, y, -z)$.

Solution: The algebraic form of the curved surface is $g(\mathbf{r}) = y - f(x,z) = y - x^2 - z^2 = 0$, whose gradient is $\boldsymbol{\nabla} g(\mathbf{r}) = (-2x, 1, -2z)$ (and $\frac{\partial g}{\partial y} = 1$). This vector points inward to the surface, so we should change its sign to adhere to the normal convention of outward-pointing normals. The integral over the curved surface then becomes

$$\int_{S'} \mathbf{F}(\mathbf{r}) \cdot \boldsymbol{\nabla} g(\mathbf{r}) dxdz = \int_{S'} (-x^2 - 3z^2) dxdz, \tag{2.48}$$

where the surface S' is the projection of the paraboloid onto the xz-plane, which is a circle, which is most conveniently parameterized by the polar coordinates (r, θ), so the integral becomes

$$\int_{S'} (-x^2 - 3z^2) dxdz = \int_0^{2\pi} d\theta \int_0^1 r(-r^2 \cos^2\theta - 3r^2 \sin^2\theta) dr = -\pi. \tag{2.49}$$

The integral over the end surface is trivial, which upon integration gives π, so the two contributions sum to zero.

Example: Surface Integral (4).

Question: Evaluate $\int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$ over the upper half of the sphere $x^2 + y^2 + z^2 = 9$ and for $\mathbf{F}(\mathbf{r}) = (x, y, z^4)$.

Solution: The surface is given parametrically by the equation

$$\mathbf{r}(\theta, \phi) = 3(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \quad \theta \in [0, \frac{\pi}{2}], \phi \in [0, 2\pi]. \tag{2.50}$$

The infinitesimal surface vector $d\mathbf{S}$ can be calculated explicitly by

$$d\mathbf{S} = \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = 9(\sin^2\theta\cos\phi, \sin^2\theta\sin\phi, \sin\theta\cos\theta). \tag{2.51}$$

Then the integral can be carried out straightforwardly as

$$\int_S \mathbf{F}(\mathbf{r}) \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) d\theta d\phi = 279\pi. \tag{2.52}$$

Example: Surface Integral (5).

Question: Viviani's window is the part of the surface of the sphere of radius R centered in the origin, contained within the right cylinder with base given by a circumference of radius $R/2$, centered at $(x, y) = (R/2, 0)$, for $z \geq 0$ (see fig. 2.5a). Find the area S of the Viviani's window.

Solution: We are tasked with solving the integral $S = \int_S dS$, and the difficulty lies on parameterizing the surface S. One is to project the surface onto the base and evaluate the double integral, an equally valid way is to parameterize the surface with r, θ according to fig. 2.5b, where

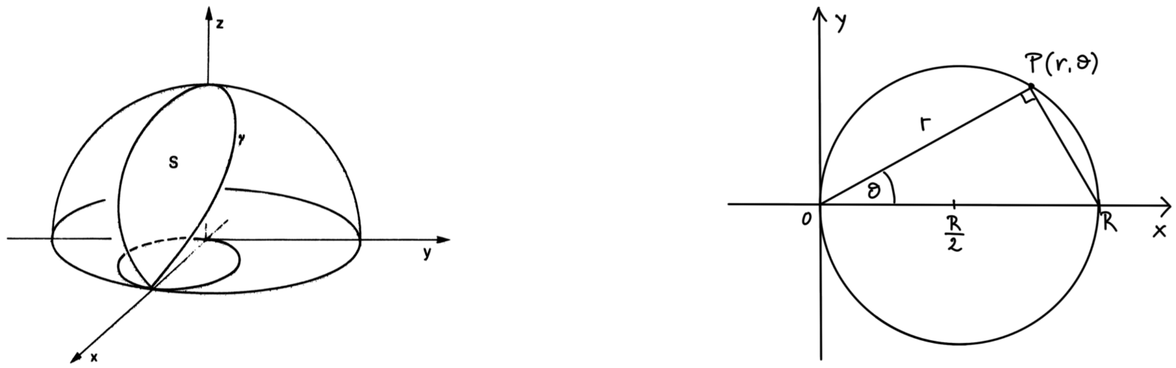
$$x = r\cos\theta, \quad y = r\sin\theta \text{ and } z = \sqrt{R^2 - x^2 - y^2} = \sqrt{R^2 - r^2} \tag{2.53}$$

for $r \in [0, R\cos\theta], \theta \in [-\pi/2, \pi/2]$
The infinitesimal surface element is then

$$dS = \left| \frac{\partial \mathbf{r}}{\partial r} \times \frac{\partial \mathbf{r}}{\partial \theta} \right| = \frac{rR}{\sqrt{R^2 - r^2}}. \tag{2.54}$$

Thus the integral

$$S = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \int_0^{R\cos\theta} \frac{rR}{\sqrt{R^2 - r^2}} dr = R^2(\pi - 2). \tag{2.55}$$



(a) (b)

Figure 2.5

2.3.1 Vector Areas

The vector area of a surface S is defined as

$$\mathbf{S} = \int_S d\mathbf{S} = \oint_S \hat{\mathbf{n}} dS \tag{2.56}$$

A closed surface will always has a zero vector area, since when projecting onto the xy-plane, from eq. (2.11) we see that every $d\mathbf{S}_+ = dA/\left|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}\right|$ will have an opposite contribution $d\mathbf{S}_- = dA/\left|-\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}\right| = -d\mathbf{S}_+$.

This fact implies that the vector area of any open surface S only depends on its perimeter curve C, since we can construct a closed surface with arbitrary upper and lower surface and their contribution must be equals in magnitude so that the sum is zero. We know that their directions are the same due to right hand rule.

Specifically, the vector area can be represented by the line integral

$$\mathbf{S} = \frac{1}{2} \oint_C \mathbf{r} \times d\mathbf{r}, \tag{2.57}$$

since one of the possible surface spanned by the perimeter C is a cone with its vertex at the origin with the perimeter of the base C as shown in fig. 2.6 and its area is the sum of all the infinitesimal triangle, each with vector area $d\mathbf{S} = \mathbf{r} \times d\mathbf{r}/2$.

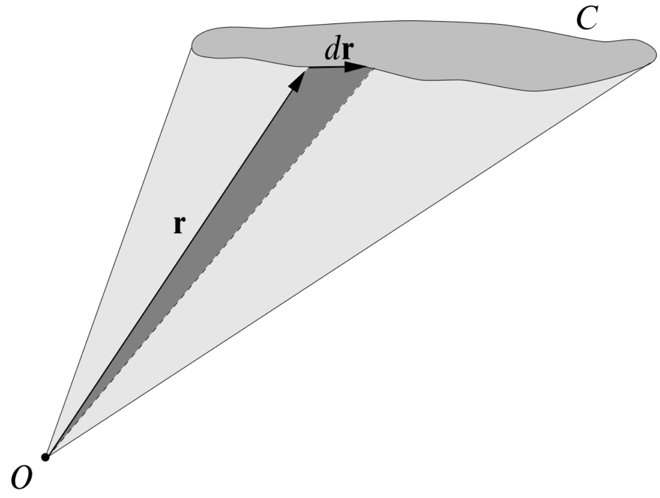


Figure 2.6

For a surface confined to the xy-plane, $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ and $d\mathbf{r} = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}}$, thus $\mathbf{r} \times d\mathbf{r} = (xdy - ydx)\hat{\mathbf{k}}$, so the area is what we have found earlier in eq. (2.91).

2.3.2 Solid Angle

The solid angle Ω subtended at a point O by a surface S is defined as

$$\Omega = \int_S \frac{\hat{\mathbf{r}} \cdot d\mathbf{S}}{r^2}. \tag{2.58}$$

In particular, when the surface is clossed $\Omega = 0$ if O is outside S and $\Omega = 4\pi$ if O is an interior point.

2.4 Volume Integrals

Since dV is a scalar, there are only two kinds of volume integrals

$$\int_V f(x,y,z)dV \text{ and } \int_V \mathbf{F}(x,y,z)dV. \tag{2.59}$$

Similar to how the vector area of a surface S can be represented by a line integral along its perimeter C, the volume of a volume V can be represented by a surface integral over the surface S that bounds it.

Referring to fig. 2.7, we have

$$V = \int_V dV = \frac{1}{3} \oint_S \mathbf{r} \cdot d\mathbf{S}, \tag{2.60}$$

as the volume of each cone is $dV = \mathbf{r} \cdot d\mathbf{S}/3$.

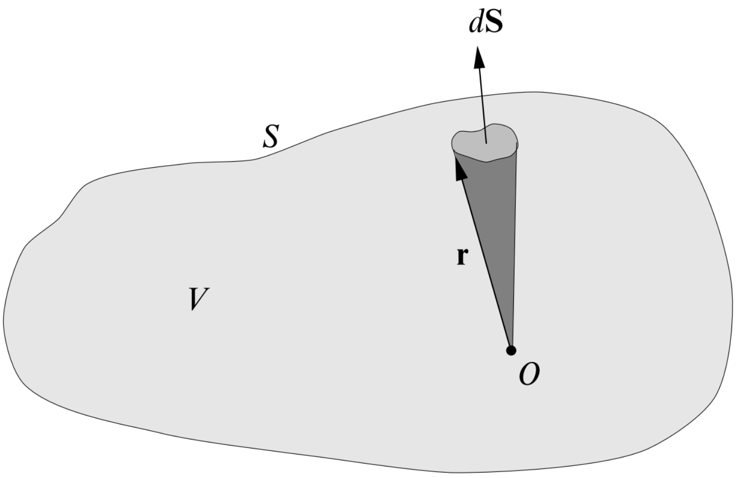


Figure 2.7

Example: Volume Integrals (1).

Question: Find the volume of a sphere with a cylinder of raidus $s < R$ removed from the middle.

Solution: The most natural choice of coordinates is the cylindrical corrdinates, and the region we want to integrate over is $x^2 + y^2 + z^2 \leq R^2$, together with $x^2 + y^2 \geq s^2$. In cylindrical corrdinates, these become

$$s \leq \rho \leq R \quad \text{and} \quad -\sqrt{R^2 - \rho^2} \leq z \leq \sqrt{R^2 - \rho^2}. \tag{2.61}$$

So the volume is

$$V = \int_{-\sqrt{R^2-\rho^2}}^{\sqrt{R^2-\rho^2}} \int_0^{2\pi} \int_s^R \rho d\rho d\phi dz = \frac{4\pi}{3}(R^2 - s^2)^{\frac{3}{2}}. \tag{2.62}$$

As we have noted, sometimes it is easier to find the limits of integration by writing them in Cartesian coordinates first and then express them in the coordinate system one is using.

2.5 Integral Theorems

2.5.1 The Divergence Theorem

The divergence theorem states that

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{F}) dV = \oint_S \mathbf{F} \cdot d\mathbf{S}. \tag{2.63}$$

The corresponding 2D-version is

$$\int_S (\boldsymbol{\nabla} \cdot \mathbf{F}) dS = \oint_C \mathbf{F} \cdot d\mathbf{r}. \tag{2.64}$$

With this pattern we can easily generalize the divergence theorem into higher dimensions.

Example: Surface Integral by the Divergence Theorem.

Question: Evaluate the surface integral $I = \int_S \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = (y - x)\hat{\mathbf{i}} + x^2 z\hat{\mathbf{j}} + (z + x^2)\hat{\mathbf{k}}$ and S is the open surface of the hemisphere $x^2 + y^2 + z^2 = a^2, z \geq 0$.

Solution: Consider the closed surface $S' = S + S_1$, where S_1 is the circular area in the xy-plane given by $x^2 + y^2 \leq a^2, z = 0$. By the divergence theorem we have

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{F}) dV = 0 = \int_S \mathbf{F} \cdot d\mathbf{S} + \int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1. \tag{2.65}$$

Therefore we can simly evaluate the surface integral over S_1 and add a negative sign to get the desired result. Thus

$$I = -\int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1 = \int\int_R x^2 dx dy = \frac{\pi a^4}{4}. \tag{2.66}$$

Example: The Continuity Equation.

Question: For a compressible fluid with time-varing position-dependent density $\rho(\mathbf{r}, t)$ and velocity field $v(\mathbf{r}, t)$, in which fluid is neither being created nor destroyed, show that

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0. \tag{2.67}$$

Solution: Consider an arbitrary volume V in the fluid bounded by S. From conservation of mass, we have

$$\frac{dM}{dt} = \frac{d}{dt} \int_V \rho dV = -\oint_S \rho \mathbf{v} \cdot d\mathbf{S} \tag{2.68}$$

$$\int_V \frac{\partial \rho}{\partial t} dV + \int_V \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) dV = \int_V \left(\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) \right) dV = 0. \tag{2.69}$$

But since the volume V is arbitrary the integrand must be identically zero, arriving at the desired result.
For the flow of an incompressible fluid $\rho =$ constant and the continuity equation becomes simply $\boldsymbol{\nabla} \cdot v = 0$.

Example: Predator-Prey Systems.

Question: Suppose the population of the predator and prey are $w(t)$ and $c(t)$ respectively, then the set of differential equations governing this sytem is

$$\begin{cases} \frac{dw}{dt} &= w(-\alpha + c - \mu w), \\ \frac{dc}{dt} &= c(\beta - w - \nu c), \end{cases} \tag{2.70}$$

where the meaning of the coefficients $(\alpha, \beta, \mu, \nu)$ can be easily interpreted. Does there exist a periodic orbit in the c-w phase space?

Solution: We first change the form of the differential equation as

$$\frac{d\mathbf{a}}{dt} = \mathbf{p}, \quad \mathbf{a} = \begin{pmatrix} w \\ c \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} w(-\alpha + c - \mu w) \\ c(\beta - w - \nu c) \end{pmatrix}, \tag{2.71}$$

which makes it clear that $\mathbf{p}$ is the tangent to any path $\mathbf{a}(t)$ in the animal phase plane. Using the divergence theorem in 2 dimensions, we have

$$\int_S (\boldsymbol{\nabla} \cdot (b(w,c))\mathbf{p}) dS = \oint_C b(w,c)\mathbf{p} \cdot \hat{\mathbf{n}} dt, \tag{2.72}$$

where $b(c,w)$ can be any function of our choice. If $b(w,c) = 1/wc$, then we have

$$\boldsymbol{\nabla} \cdot \left(\frac{\mathbf{p}}{wc}\right) = -\frac{\mu}{c} - \frac{\nu}{w}, \tag{2.73}$$

which is strictly negative. However, if $\boldsymbol{\nabla} \cdot (\mathbf{p}/wc)$ is always negative then there's no way to integrate it over a region and get zero. There for by contradiction we conclude that there is no closed orbit in the animal phse plane. In other words, therei is no nice periodic solutions.

Example: Integral Form and Differntial Form of Gauss's Law.

Question: Asuuming spherical symmetry, prove that the integral form and the differential form of the Gauss's law are equivalent by considering the volume enclosed by two spherical shells of radius r and $r + dr$.

Solution: Since there is spherical symmetry, we have $\mathbf{E}(\mathbf{r}) = \mathbf{E}(r)$. From the integral from of the Gauss's law we have

$$\begin{aligned} 4\pi(r+dr)^2 E(r+dr) - 4\pi r^2 E(r) &= \frac{\rho(4\pi r^2 dr)}{\epsilon_0} \\ (r^2 + 2rdr)(E(r) + dE) - r^2 E(r) &= \frac{\rho r^2 dr}{\epsilon_0} \\ 2rE(r)dr + r^2 dE &= \frac{\rho r^2 dr}{\epsilon_0} \\ \frac{2E}{r} + \frac{dE}{dr} = \frac{1}{r^2}\frac{d}{dr}(r^2 E(r)) &= \frac{\rho}{\epsilon_0} \end{aligned} \tag{2.74}$$

The differential form of the Gauss's law, on the other hand, reads

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{1}{r^2}\frac{d}{dr}(r^2 E(r)) = \frac{\rho}{\epsilon_0}, \tag{2.75}$$

which is equivalent to the integral form.

Green's Theorems

Consider two scalar functions ϕ and ψ in some volume V bounded by a surface S. Applying the divergence theorem to the vector field $\phi\boldsymbol{\nabla}\psi$, we get

$$\oint_S \phi\boldsymbol{\nabla}\psi \cdot d\mathbf{S} = \int_V \boldsymbol{\nabla} \cdot (\phi\boldsymbol{\nabla}\psi) dV = \int_V (\phi\nabla^2\psi + (\boldsymbol{\nabla}\phi) \cdot (\boldsymbol{\nabla}\psi)) dV. \tag{2.76}$$

This is known as the Green's first theorem.

Reversing the roles of ϕ and ψ in the above equation and subtracting the two equations gives

$$\oint_S (\phi\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\phi) \cdot d\mathbf{S} = \int_V (\phi\nabla^2\psi - \psi\nabla^2\phi) dV. \tag{2.77}$$

This is known as the Green's second theorem.

Two Other Theorems

Letting $\mathbf{F}$ in eq. (2.63) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_V \boldsymbol{\nabla} f dV = \oint_S f d\mathbf{S} \text{ and } \int_V (\boldsymbol{\nabla} \times \mathbf{F}) dV = \oint_S d\mathbf{S} \times \mathbf{F}. \tag{2.78}$$

2.5.2 Stokes' Theorem

The Stokes' theorem states that

$$\int_S (\boldsymbol{\nabla} \times \mathbf{F}) \cdot d\mathbf{S} = \oint_C \mathbf{F} \cdot d\mathbf{r}. \tag{2.79}$$

In two dimensions, curl is not defiend, but the main idea stands:

$$\int_S \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dS = \oint_C (F_x dx + F_y dy). \tag{2.80}$$

This is known as the Green's theorem in a plane, which is equivalent to the two-dimensional divergence theorem. Apply to a rectangular region, Green's theorem in a plane reduces to the fundamental theorem of calculus.

If an area has a number of disconnected componenets, then the integral should be done in an anti-clockwise direction around the exterior boundary, and in a clockwise direction on any interior boundary. The quickest way to see this is to do the integration around a continouous boundary as shown in fig. 2.8, with an infinitesimal gap. The two contributions across the gap then cancel.

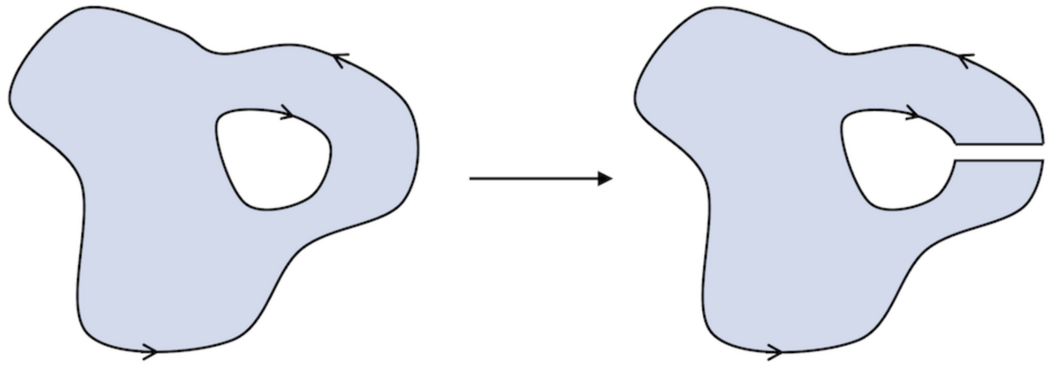


Figure 2.8

Example: Stokes' Theorem (1).

Question: Let S be the cap of a sphere of radius R that is covered by the angle $0 \leq \theta \leq \alpha$. Confirm the Stokes' theorem with the vector field $\mathbf{F} = (0, xz, 0)$

Solution: The surface integral is

$$\int_S (\boldsymbol{\nabla} \times \mathbf{F}) \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^{\alpha} R^2 \sin\theta d\theta d\phi \begin{pmatrix} -x \\ 0 \\ z \end{pmatrix} \cdot \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} = \pi R^3 \cos\alpha \sin^2\alpha. \tag{2.81}$$

The line integral is performed around the rim C parameterized by the angle ϕ, given by

$$\mathbf{r}(\phi) = R(\sin\alpha\cos\phi, \sin\alpha\sin\phi, \cos\alpha) \implies d\mathbf{r} = R(-\sin\alpha\sin\phi, \sin\alpha\cos\phi, 0)d\phi. \tag{2.82}$$

So the line integral is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} Rxz \sin\alpha\cos\phi d\phi = \pi R^3 \sin^2\alpha\cos\alpha. \tag{2.83}$$

Example: Stokes' Theorem (2).

Question: Consider the conical surface S defined by $z^2 = x^2 + y^2$ with $0 < a \leq z \leq b$. Confirm the Stoke's theorem with the vector field $\mathbf{F} = (0, xz, 0)$.

Solution: The surface is parameterized, in cylindrical polar coordinates, by

$$\mathbf{r}(\rho, \phi) = (\rho\cos\phi, \rho\sin\phi, z) \tag{2.84}$$

with $a \leq \rho \leq b$ and $0 \leq \phi < 2\pi$.
So the infinitesimal surface vector area $d\mathbf{S}$ can be found by

$$d\mathbf{S} = \frac{\partial\mathbf{r}}{\partial\rho} \times \frac{\partial\mathbf{r}}{\partial\phi} = (-\cos\phi, -\sin\phi, 1)\rho d\rho d\phi. \tag{2.85}$$

The surface integral is

$$\int_S (\boldsymbol{\nabla} \times \mathbf{F}) \cdot d\mathbf{S} = \int_0^{2\pi}\int_a^b \begin{pmatrix} -x \\ 0 \\ z \end{pmatrix} \cdot \begin{pmatrix} -\cos\phi \\ -\sin\phi \\ 1 \end{pmatrix} \rho d\rho d\phi = \pi(b^3 - a^3). \tag{2.86}$$

Now the line integral is performed around the two circumferences C_1 and C_2 with radius $R = a$ and b, which are parameterized by the angle ϕ, given by

$$\mathbf{r}(\phi) = R(\cos\phi, \sin\phi, 1) \implies d\mathbf{x} = R(-\sin\phi, \cos\phi, 0)d\phi, \tag{2.87}$$

so the line integral of one of the circumference is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} R^3 \cos^2\phi d\phi = \pi R^3. \tag{2.88}$$

Since the orientation of C_1 and C_2 are in opposite direction, substituting $R = a$ and b and subtracting the results give the same result as the surface integral.

Example: The Ampere's Law.

Question: Convert the integral form of Ampere's Law into differential form

Solution: Amere's law for any circuit C bounding a surface S is given by

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = \int_S (\boldsymbol{\nabla} \times \mathbf{B}) \cdot d\mathbf{S} = \mu_0 I = \mu_0 \int_S \mathbf{J} \cdot d\mathbf{S}. \tag{2.89}$$

Hence

$$\int_S (\boldsymbol{\nabla} \times \mathbf{B} - \mu_0\mathbf{J}) \cdot d\mathbf{S} = 0. \tag{2.90}$$

Example: Area of an Ellipse.

Question: Show that the area of a region R enclosed by a simle closed curve C is given by $A = \frac{1}{2}\oint_C (xdy - ydx) = \oint_C xdy = -\oint_C ydx$. Hence calculate the area of the ellipse $x = a\cos\phi, y = b\sin\phi$.

Solution: By Green's theorem we have

$$\oint_C (xdy - ydx) = \int\int_R (1+1)dxdy = 2A. \tag{2.91}$$

Therefore the area of an ellipse is

$$A = \frac{1}{2}\int_0^{2\pi} ab(\cos^2\phi + \sin^2\phi)d\phi = \pi ab. \tag{2.92}$$

The Green's theorem is also valid for region with holes, however, the line integral must be carry out in the direction that a person travelling along the boundaries always has the region R on their left.

We also see that if the line integral around a closed loop is zero, Green's theorem implies that $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, which is equivalent to saying that $P(x,y)dx + Q(x,y)dy$ is an exact differential such that it equals to the differential for some function $\phi(x,y)$ and for a closed loop the beginning and the end points are the same thus we evaluate ϕ at the same point and thus the result is zero.

The Two Other Theorems

Letting $\mathbf{F}$ in eq. (2.79) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_S d\mathbf{S} \times \boldsymbol{\nabla} f = \oint_C f d\mathbf{r} \text{ and } \int_S (d\mathbf{S} \times \boldsymbol{\nabla}) \times \mathbf{F} = \oint_C d\mathbf{r} \times \mathbf{F}. \tag{2.93}$$

Appendices

A Rigorous Proofs

A.1 Leibnitz' Theorem

Here we provide a proof for section 1.1.1.

Proof. Suppose eq. (1.1) is valid for n equals to some integer N, then

$$\begin{aligned}
f^{(N+1)} &= \sum_{r=0}^{N} \binom{n}{r} \frac{d}{dx}(u^{(r)}v^{(N-r)}) \\
&= \sum_{r=0}^{N} \binom{N}{r} (u^{(r)}v^{(N-r+1)} + u^{(r+1)}v^{(N-r)}) \\
&= \sum_{s=0}^{N} \binom{N}{s} u^{(s)}v^{(N+1-s)} + \sum_{s=1}^{N+1} \binom{N}{s-1} u^{(s)}v^{(N+1-s)} \\
&= \binom{N}{0} u^{(0)}v^{(N+1)} + \sum_{s=1}^{N} \binom{N+1}{s} u^{(s)}v^{(N+1-s)} + \binom{N}{N} u^{(N+1)}v^{(0)} \\
&= \binom{N+1}{0} u^{(0)}v^{(N+1)} + \sum_{s=1}^{N} \binom{N+1}{s} u^{(s)}v^{(N+1-s)} + \binom{N+1}{N+1} u^{(N+1)}v^{(0)} \\
&= \sum_{s=0}^{N+1} \binom{N+1}{s} u^{(s)}v^{(N+1-s)}.
\end{aligned} \tag{A.1}$$

Since $N = 1$ corresponds to prodouct rule, which is trivial, by induction we have proved eq. (1.1) holds for all positive integers n.

□