Calculus

Haydn Cheng

December 15, 2024

Contents

1 Differentiation

1.1 Ordinary Differentiation

1.1.1 The Leibnitz' Theorem

[The Leibnitz' Theorem]

Theorem 1.1.1. *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 $\dfrac{df}{dx} = 0$, and can be further classified into

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

If $\dfrac{d^2 f}{dx^2} = 0$ 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 $\dfrac{d^2 f}{dx^2} = 0$ but $\dfrac{df}{dx} \neq 0$ where the concavity of the fucntion changes.

[1] Proof given in section A.1.

1.2 Partial Differentiation

In this section we investigate how a function with more than one variable changes when the variables change.[2] We will restrict ourselves most to functions that depend on two variables, since $f(x, y)$ can still be visualized by a surface in a three-dimensional space, similar to how we represent $f(x)$ as a curve in a two-dimensional space. However, it is difficult to visualise functions of more than two variables.

1.2.1 Partial Derivatives

We start by finding how $f(x, y)$ changes when one of the variables, say x, changes, while the other variable (y) is kept constant. But now this becomes an ordinary differentiation since $f(x, y) = f(x)$ when y is kept constant. In analogous to how ordinary derivativce is defined, we have

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

where 3 equivalent notations are used to denote the partial derivative of $f(x, y)$ with respect to x in order of descending formality.

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

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{xy} = f_{yx} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}. \tag{1.3}$$

1.2.2 Total Derivatives

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

$$\begin{aligned}
\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\
&= f(x + \Delta x, y + \Delta y) - f(x, y) + f(x, y + \Delta y) - f(x, y + \Delta y) \\
&= \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x}\Delta x + \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\Delta y \\
&\approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y.
\end{aligned} \tag{1.4}$$

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 $\Delta y \to 0$.

[2] It is fine if the variables are themselves connected with each other, say $y = y(x)$.

1.2.3 Exact Differentials

An arbitrary differential

$$A(x,y)dx + B(x,y)dy \tag{1.5}$$

is exact if it is the differential of a function

$$df(x,y) = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \tag{1.6}$$

Therefore, we have

$$A(x,y) = \frac{\partial f}{\partial x} \text{ and } B(x,y) = \frac{\partial f}{\partial y} \tag{1.7}$$

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

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}.^{3} \tag{1.8}$$

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

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \text{ and } dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \\ \implies dx &= \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right) dz. \end{aligned} \tag{1.9}$$

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

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial y}{\partial x}\right)_z^{-1} \text{ and } \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y + \left(\frac{\partial x}{\partial y}\right)_z = -1 \tag{1.10}$$

[3] 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.

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

$$\frac{df}{du} = \frac{\partial f}{\partial x}\frac{dx}{du} + \frac{\partial f}{\partial y}\frac{dy}{du}, \tag{1.11}$$

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

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \text{ and } \frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}. \tag{1.12}$$

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. (1.4), instead, we get the taylor series

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} \Delta x + \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} \Delta y \\ &+ \frac{1}{2!}\left(\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} (\Delta x)^2 + 2\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)} (\Delta x)(\Delta y) + \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)} (\Delta y)^2 \right) \\ &+ \mathcal{O}((\Delta x)^3) + \mathcal{O}((\Delta y)^3). \end{aligned} \tag{1.13}$$

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

$$f(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[(\Delta \mathbf{x} \cdot \nabla)^n f(\mathbf{x})\right]\Big|_{\mathbf{x}=\mathbf{x}_0}. \tag{1.14}$$

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

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = 0. \tag{1.15}$$

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

$$df = f(x,y) - f(x_0,y_0) \approx \frac{1}{2}\left[f_{xx}\left(\Delta x + \frac{f_{xy}\Delta y}{f_{xx}}\right)^2 + (\Delta y)^2\left(f_{yy} - \frac{f_{xy}^2}{f_{xx}}\right)\right]. \tag{1.16}$$

For a minimum point, we require that $df > 0$ for arbitrary Δx and Δy. This implies that $f_{xx} > 0$ and $\left(f_{xx}f_{yy} > f_{xy}^2\right)$. Due to symmtry of x and y, $f_{\mathbf{y}} > 0$ 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: both $f_{xx} > 0, f_{yy} > 0$ and $f_{xy}^2 < f_{xx}f_{yy}$,
2. Maximum point: both $f_{xx} < 0, f_{yy} < 0$ and $f_{xy}^2 < f_{xx}f_{yy}$ and
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 $\pm\frac{1}{2}(|f_{xx}|^{\frac{1}{2}}\Delta x \pm |f_{yy}|^{\frac{1}{2}}\Delta y)^2$, then for some choice of the ratio $\frac{\Delta y}{\Delta x}$, $df = 0$ 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

$$\frac{\partial f}{\partial x_i} = 0 \quad \forall\ x_i, \tag{1.17}$$

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

$$df = f(\mathbf{x}) - f(\mathbf{x}_0) \approx \frac{1}{2}\sum_i\sum_j \frac{\partial^2 f}{\partial x \partial y}\Delta x_i \Delta x_j, \tag{1.18}$$

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

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv + \frac{\partial \mathbf{r}}{\partial w}dw. \tag{1.19}$$

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

$$\frac{\partial \mathbf{r}}{\partial u} = f\hat{\mathbf{u}}, \frac{\partial \mathbf{r}}{\partial v} = g\hat{\mathbf{v}} \text{ and } \frac{\partial \mathbf{r}}{\partial w} = h\hat{\mathbf{w}}, \tag{1.20}$$

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

$$d\mathbf{r} = f(du\hat{\mathbf{u}}) + g(dv\hat{\mathbf{v}}) + h(dw\hat{\mathbf{w}}) \tag{1.21}$$

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

$$ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}} = \sqrt{(fdu)^2 + (gdv)^2 + (hdw)^2}. \tag{1.22}$$

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

$$da = (fg)dudv \tag{1.23}$$

as shown in fig. 1.1.

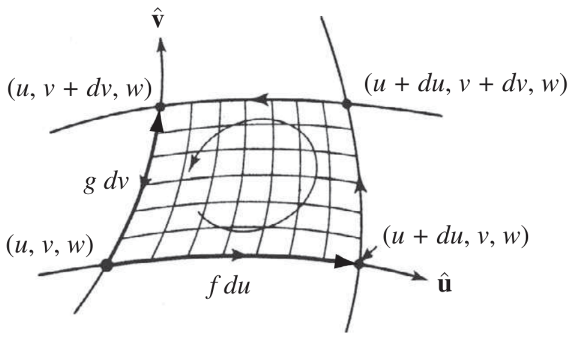

Figure 1.1

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

$$d\tau = |fd\hat{\mathbf{u}} \cdot (gd\hat{\mathbf{v}} \times hd\hat{\mathbf{w}})|dudvdw = (fgh)dudvdw \tag{1.24}$$

as shown in fig. 1.2.

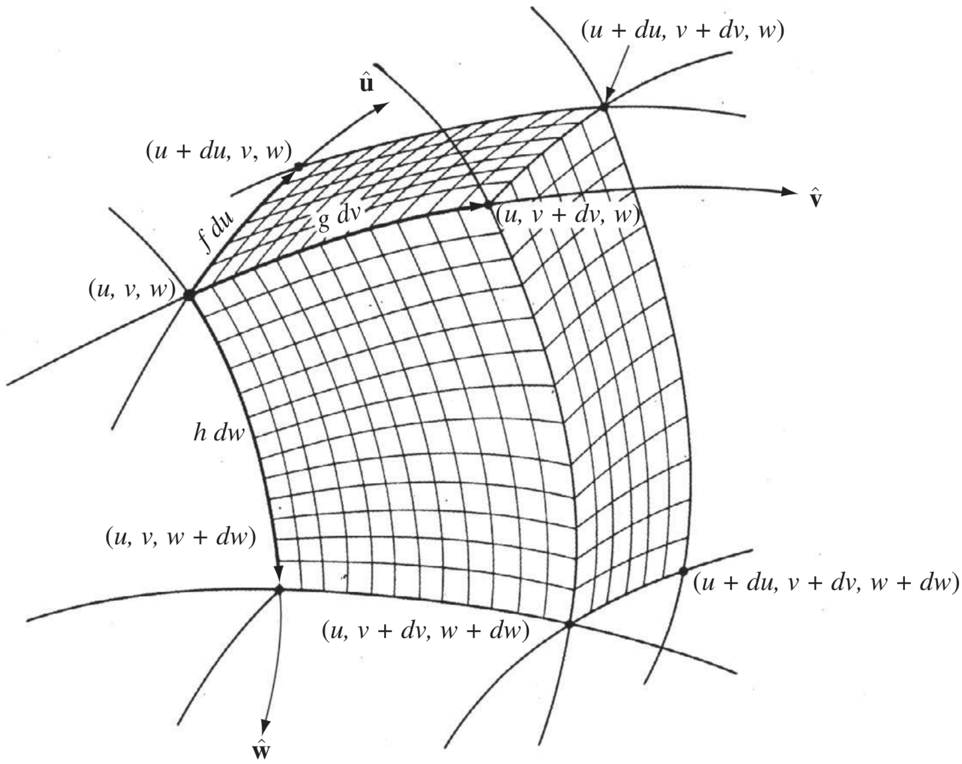

Figure 1.2

1.3.2 Spherical Coordinates

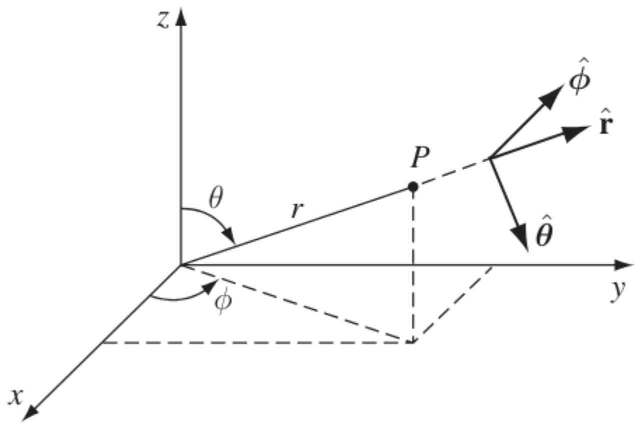

Figure 1.3

From fig. 1.3, the relations of the two set of variables (x, y, z) and (r, θ, ϕ) are

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases} \quad \text{or} \quad \begin{cases} r = \sqrt{x^2+y^2+z^2} \\ \theta = \arctan\left(\dfrac{\sqrt{x^2+y^2}}{z}\right) \\ \phi = \arctan\left(\dfrac{y}{x}\right) \end{cases} . \tag{1.25}$$

A general position vector can then be written as

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} = r\sin\theta\cos\phi\hat{\mathbf{x}} + r\sin\theta\sin\phi\hat{\mathbf{y}} + r\cos\theta\hat{\mathbf{z}}. \tag{1.26}$$

A general infinitismal displacement vector can be written as

$$\begin{aligned} d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial x}dx + \frac{\partial \mathbf{r}}{\partial y}dy + \frac{\partial \mathbf{r}}{\partial z}dz = dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}} + dz\hat{\mathbf{z}} \\ &= \frac{\partial \mathbf{r}}{\partial r}dr + \frac{\partial \mathbf{r}}{\partial \theta}d\theta + \frac{\partial \mathbf{r}}{\partial \phi}d\phi = fdr\hat{\mathbf{r}} + gd\theta\hat{\boldsymbol{\theta}} + hd\phi\hat{\boldsymbol{\phi}}. \end{aligned} \tag{1.27}$$

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

$$\begin{aligned} &\frac{\partial \mathbf{r}}{\partial r} = f\hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}, \\ &\frac{\partial \mathbf{r}}{\partial \theta} = g\hat{\boldsymbol{\theta}} = r(\cos\theta\cos\phi\hat{\mathbf{x}} + \cos\theta\sin\phi\hat{\mathbf{y}} - \sin\theta\hat{\mathbf{z}}) \text{ and} \\ &\frac{\partial \mathbf{r}}{\partial \phi} = r\hat{\boldsymbol{\phi}} = \sin\theta(-\sin\phi\hat{\mathbf{x}} + \cos\phi\hat{\mathbf{y}}). \end{aligned} \tag{1.28}$$

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 $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ as

$$\begin{cases} \hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}, \\ \hat{\boldsymbol{\theta}} = \cos\theta\cos\phi\hat{\mathbf{x}} + \cos\theta\sin\phi\hat{\mathbf{y}}0\sin\theta\hat{\mathbf{z}}, \\ \hat{\boldsymbol{\phi}} = -\sin\phi\hat{\mathbf{x}} + \cos\phi\hat{\mathbf{y}}. \end{cases} \tag{1.29}$$

We can also solve for $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ in terms of $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ as

$$\begin{cases} \hat{\mathbf{x}} = \sin\theta\cos\phi\hat{\mathbf{r}} + \cos\theta\cos\phi\hat{\boldsymbol{\theta}} - \sin\phi\hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{y}} = \sin\theta\sin\phi\hat{\mathbf{r}} + \cos\theta\sin\phi\hat{\boldsymbol{\theta}} + \cos\phi\hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{z}} = \cos\theta\hat{\mathbf{r}} - \sin\theta\hat{\boldsymbol{\theta}}. \end{cases} \tag{1.30}$$

1.3.3 Cylindrical Coordinates

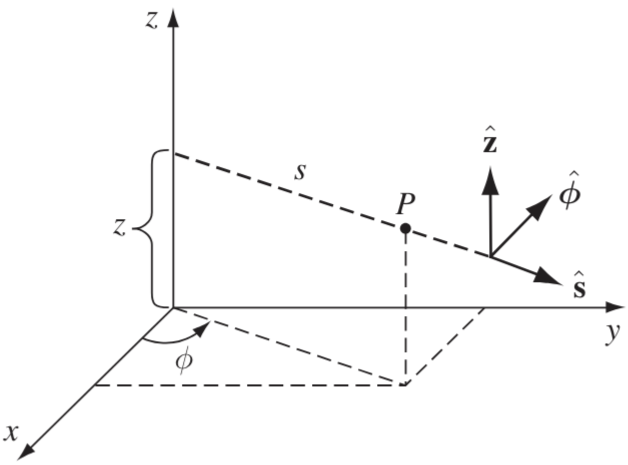

Figure 1.4

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

$$\begin{cases} x = \rho\cos\phi \\ y = \rho\sin\phi \\ z = z \end{cases} \quad \text{or} \quad \begin{cases} \rho = \sqrt{x^2+y^2} \\ \phi = \arctan\left(\dfrac{y}{x}\right) \\ z = z \end{cases} \quad . \tag{1.31}$$

A general position vector can be written as

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} = \rho\cos\phi\hat{\mathbf{x}} + \rho\sin\phi\hat{\mathbf{y}} + \hat{\mathbf{z}}. \tag{1.32}$$

A general infinitismal displacement vector can be written as

$$\begin{aligned} d\mathbf{r} &= \frac{\partial\mathbf{r}}{\partial x}dx + \frac{\partial\mathbf{r}}{\partial y}dy + \frac{\partial\mathbf{r}}{\partial z}dz = dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}} + dz\hat{\mathbf{z}} \\ &= \frac{\partial\mathbf{r}}{\partial\rho}d\rho + \frac{\partial\mathbf{r}}{\partial\phi}d\phi + \frac{\partial\mathbf{r}}{\partial z}dz = fd\rho\hat{\boldsymbol{\rho}} + gd\phi\hat{\boldsymbol{\phi}} + hdz\hat{\mathbf{z}}. \end{aligned} \tag{1.33}$$

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

$$\begin{aligned} \frac{\partial\mathbf{r}}{\partial\rho} &= f\hat{\boldsymbol{\rho}} = \cos\phi\hat{\mathbf{x}} + \sin\phi\hat{\mathbf{y}}, \\ \frac{\partial\mathbf{r}}{\partial\phi} &= g\hat{\boldsymbol{\phi}} = \rho(-\sin\phi\hat{\mathbf{x}} + \cos\phi\hat{\mathbf{y}}) \text{ and} \\ \frac{\partial\mathbf{r}}{\partial z} &= h\hat{\mathbf{z}} = \hat{\mathbf{z}}. \end{aligned} \tag{1.34}$$

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

We can thus solve for $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{z}})$ in terms of $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ as

$$\begin{cases} \hat{\boldsymbol{\rho}} = \cos\phi \hat{\mathbf{x}} + \sin\phi \hat{\mathbf{y}}, \\ \hat{\boldsymbol{\phi}} = -\sin\phi \hat{\mathbf{x}} + \cos\phi \hat{\mathbf{y}}, \\ \hat{\mathbf{z}} = \hat{\mathbf{z}}. \end{cases} \tag{1.35}$$

We can also solve for $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ in terms of $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{z}})$ as

$$\begin{cases} \hat{\mathbf{x}} & = \cos\phi \hat{\boldsymbol{\rho}} - \sin\phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{y}} & = \sin\phi \hat{\boldsymbol{\rho}} + \cos\phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{z}} & = \hat{\mathbf{z}}. \end{cases} \tag{1.36}$$

1.3.4 Space Curves

A curve in space can be described by the vector $\mathbf{r}(t)$ joining the origin O of a coordinate system to a point on the curve. As the parameter t varies, the end-point on the curve moves along the curve. Some common examples of the parameter are time and arclength. In Cartesian coordinates,

$$\mathbf{r}(u) = x(u)\hat{\mathbf{x}} + y(u)\hat{\mathbf{y}} + z(u)\hat{\mathbf{z}}. \tag{1.37}$$

Alternatively, a space curve in three-dimensional space can be represented by two simultaneous equations $F(x, y, z) = G(x, y, z) = 0$.

Imagine two position vectors $\mathbf{r}(u)$ and $\mathbf{r}(u + du)$, the difference $d\mathbf{r}$ is a vector tangent to C at that point in the direction of increasing u. In the special case where the parameter u is the arc length s along the curve

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{|d\mathbf{r}|} \tag{1.38}$$

is the unit tangent vector.

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

$$\kappa\hat{\mathbf{n}} = \frac{d\hat{\mathbf{t}}}{ds} = \frac{d^2\mathbf{r}}{ds^2} \tag{1.39}$$

and the radius of curvature is defined as $\rho = \frac{1}{\kappa}$.

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

$$\hat{\mathbf{b}} = \hat{\mathbf{t}} \times \hat{\mathbf{n}}. \tag{1.40}$$

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

$$-\tau\hat{\mathbf{n}} = \frac{d\hat{\mathbf{b}}}{ds}, \tag{1.41}$$

and radius of torsion as $\sigma = \frac{1}{\tau}$.

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)

$$\frac{d\hat{\mathbf{t}}}{ds} = \kappa\hat{\mathbf{n}}, \quad \frac{d\hat{\mathbf{n}}}{ds} = \tau\hat{\mathbf{b}} - \kappa\hat{\mathbf{t}} \quad \text{and} \quad \frac{d\hat{\mathbf{b}}}{ds} = -\tau\hat{\mathbf{n}}. \tag{1.42}$$

Example: Acceleration of a Particle.

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

$$\mathbf{a}(t) = \frac{dv}{dt}\hat{\mathbf{t}} + \frac{v^2}{\rho}\hat{\mathbf{n}}. \tag{1.43}$$

Solution: The velocity of the particle is

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{ds}\frac{ds}{dt} = \frac{ds}{dt}\hat{\mathbf{t}} = v\hat{\mathbf{t}}. \tag{1.44}$$

So the acceleration is

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{dv}{dt}\hat{\mathbf{t}} + v\frac{d\hat{\mathbf{t}}}{dt}. \tag{1.45}$$

But since

$$\frac{d\hat{\mathbf{t}}}{dt} = \frac{ds}{dt}\frac{d\hat{\mathbf{t}}}{ds} = v\kappa\hat{\mathbf{n}} = \frac{v}{\rho}\hat{\mathbf{n}}, \tag{1.46}$$

we have

$$\mathbf{a}(t) = \frac{dv}{dt}\hat{\mathbf{t}} + \frac{v^2}{\rho}\hat{\mathbf{n}}. \tag{1.47}$$

1.4 Space Surfaces

In Cartesian coordinates a surface is given parametrically by

$$\mathbf{r}(u, v) = x(u, v)\hat{\mathbf{x}} + y(u, v)\hat{\mathbf{y}} + z(u, v)\hat{\mathbf{z}}, \tag{1.48}$$

or algebrically as $F(x, y, z) = 0$.

The infinitesimal displacement vector is

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u} du + \frac{\partial \mathbf{r}}{\partial v} dv \tag{1.49}$$

and thus the infinitesimal area is

$$da = \left| \frac{\partial \mathbf{r}}{\partial u} du \times \frac{\partial \mathbf{r}}{\partial v} dv \right|. \tag{1.50}$$

Example: Area of a Sphere.

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

$$\mathbf{r} = a \sin\theta \cos\phi \hat{\mathbf{x}} + a \sin\theta \sin\phi \hat{\mathbf{y}} + a \cos\theta \hat{\mathbf{z}}. \tag{1.51}$$

An infinitesimal area is given by

$$da = \left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right| = a^2 \sin\theta d\theta d\phi \tag{1.52}$$

And the total surface area is then

$$A = \int_0^{\pi} \int_0^{2\pi} a^2 \sin\theta d\theta d\phi = 4\pi a^2. \tag{1.53}$$

1.5 Gradient, Divergence and Curl

We start by stating the gerneral form of del operator, gradient, divergence, curl, and Laplacian are defined as[4]

[4] 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.

$$
\begin{aligned}
\boldsymbol{\nabla} &= \frac{1}{f}\hat{\mathbf{u}}\frac{\partial}{\partial u} + \frac{1}{g}\hat{\mathbf{v}}\frac{\partial}{\partial v} + \frac{1}{h}\hat{\mathbf{w}}\frac{\partial}{\partial w}, \\
\boldsymbol{\nabla} t &= \frac{1}{f}\hat{\mathbf{u}}\frac{\partial t}{\partial u} + \frac{1}{g}\hat{\mathbf{v}}\frac{\partial t}{\partial v} + \frac{1}{h}\hat{\mathbf{w}}\frac{\partial t}{\partial w}, \\
\boldsymbol{\nabla} \cdot \mathbf{T} &= \frac{1}{fgh}\left[\frac{\partial}{\partial u}(ghT_u) + \frac{\partial}{\partial v}(fhT_v) + \frac{\partial}{\partial w}(fgT_h)\right], \\
\boldsymbol{\nabla} \times \mathbf{T} &= \frac{1}{fgh}\begin{vmatrix} f\hat{\mathbf{u}} & g\hat{\mathbf{v}} & h\hat{\mathbf{w}} \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ fT_u & gT_v & hT_w \end{vmatrix}, \\
\nabla^2 t &= \frac{1}{fgh}\left[\frac{\partial}{\partial u}\left(\frac{gh}{f}\frac{\partial t}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{fh}{g}\frac{\partial t}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{fg}{h}\frac{\partial t}{\partial w}\right)\right],
\end{aligned}
\tag{1.54}
$$

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

1.5.1 Gradient

Using the del operator defined in eq. (1.54), we can write the infinitesimal change of a scalar function $t(u, v, w)$ as

$$
dt = \frac{\partial t}{\partial u}du + \frac{\partial t}{\partial v}dv + \frac{\partial t}{\partial w}dw = \boldsymbol{\nabla} t \cdot dr = |\boldsymbol{\nabla} t||d\mathbf{r}|\cos\theta. \tag{1.55}
$$

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

$$
t(\mathbf{b}) - t(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{b}} \boldsymbol{\nabla} t \cdot d\mathbf{r}. \tag{1.56}
$$

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 the integral $\int_{\mathbf{a}}^{\mathbf{b}} \boldsymbol{\nabla} t \cdot d\mathbf{r}$ is independent of the path from $\mathbf{a}$ to $\mathbf{b}$ but only depends on the beginning and end points. Hence, $\oint_C \boldsymbol{\nabla} t \cdot d\mathbf{r} = 0$ for any closed loop, since $t(\mathbf{b}) - t(\mathbf{a}) = 0$.

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 $\mathscr{r} = |\boldsymbol{\mathscr{r}}| = |\mathbf{r} - \mathbf{r}'| = |(x-x')\hat{\mathbf{x}} + (y-y')\hat{\mathbf{y}} + (z-z')\hat{\mathbf{z}}| = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$. Generalize the result to obtain $\boldsymbol{\nabla}(\mathscr{r}^n)$.

Solution:

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

This result can be easily generalised to get

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

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

$$
\boldsymbol{\nabla}t = \frac{\partial t}{\partial x}\hat{\mathbf{x}} + \frac{\partial t}{\partial y}\hat{\mathbf{y}} + \frac{\partial t}{\partial z}\hat{\mathbf{z}} = \frac{\partial t}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial t}{\partial \theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial t}{\partial \phi}\hat{\boldsymbol{\phi}}. \tag{1.60}
$$

1.5.2 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. 1.2 (here we adopt the sign convention of $\mathbf{A}$ is positive if $\mathbf{A}$ is pointing outwards from the interior of the volume)

$$
\begin{aligned}
\oint_S \mathbf{T}\cdot d\mathbf{A} &= ((T_u)|_{u+du} - (T_u)|_u)(ghdvdw) \\
&+ ((T_v)|_{v+dv} - (T_v)|_v)(fhdudw) \\
&+ ((T_w)|_{w+dw} - (T_w)|_w)(fgdudv) \\
&= \frac{1}{fgh}\left(\frac{\partial}{\partial u}(ghT_u) + \frac{\partial}{\partial v}(fhT_v) + \frac{\partial}{\partial w}(fgT_h)\right)d\tau \\
&= (\boldsymbol{\nabla}\cdot\mathbf{T})d\tau
\end{aligned}
\tag{1.61}
$$

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,

$$\oint_S \mathbf{T} \cdot d\mathbf{A} = \int_V (\boldsymbol{\nabla} \cdot \mathbf{T}) d\tau. \tag{1.62}$$

This is the divergence theorem (also known as the Gauss's theorem or the Green's theorem). From the LHS of eq. (1.62), 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.

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

$$\boldsymbol{\nabla} \cdot \mathbf{T} = \lim_{V \to 0} \left(\frac{1}{V} \oint_S \mathbf{T} \cdot d\mathbf{S} \right). \tag{1.63}$$

However the proof is comlicated and is not presented here.

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:

$$\boldsymbol{\nabla} \cdot \mathbf{T} = \frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y} + \frac{\partial T_z}{\partial z} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 T_r \right) + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} (\sin\theta T_\theta) + \frac{1}{r} \frac{\partial}{\partial \phi} (r T_\phi). \tag{1.64}$$

1.5.3 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. 1.1 (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 $\mathbf{A}$ points in the same direction as $\hat{\mathbf{w}}$)

$$\begin{aligned} \oint_C \mathbf{T} \cdot d\mathbf{r} &= (T_u)|_v f du + (T_v)|_{u+du} g dv + (T_u)|_{v+dv} (-f du) + (T_v)|_u (-g dv) \\ &= \frac{1}{fg} \left(\frac{\partial}{\partial u} (T_v g) - \frac{\partial}{\partial v} (T_u f) \right) (\hat{\mathbf{w}} \cdot d\mathbf{A}) \\ &= (\boldsymbol{\nabla} \times \mathbf{T}) \cdot d\mathbf{A}. \end{aligned} \tag{1.65}$$

With the same argument as eq. (1.61) to eq. (1.62), the above equation can be extended to

$$\oint_C \mathbf{T} \cdot d\mathbf{r} = \int_S (\boldsymbol{\nabla} \times \mathbf{T}) \cdot d\mathbf{A}. \tag{1.66}$$

This is the Stoke's theorem. From the LHS of eq. (1.66), 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.

From the above equation we can write the gradient at a point in terms of a line integral over a infinitesimal curve surrounding the point

$$(\boldsymbol{\nabla} \times \mathbf{T} \cdot \hat{\mathbf{n}}) = \lim_{A\to 0} \left(\frac{1}{A} \oint_C \mathbf{T} \cdot d\mathbf{r} \right). \tag{1.67}$$

An equivalent form is

$$\boldsymbol{\nabla} \times \mathbf{T} = \lim_{V\to 0} \left(\frac{1}{V} \oint_S d\mathbf{S} \times \mathbf{T} \right). \tag{1.68}$$

However the proof is complicated and is not presented here.

A quick corollary is that the integral $\int_S (\boldsymbol{\nabla} \times \mathbf{T}) \cdot d\mathbf{A}$ is independent of the surface used but only depends on the boundary line. Hence, $\oint_S (\boldsymbol{\nabla} \times \mathbf{T}) \cdot d\mathbf{A} = 0$ 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:

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

1.5.4 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

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

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 ($\oint_S (f\mathbf{T}) \cdot d\mathbf{A}$), just like integration by part in ordinary derivatives, where

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

comes from the product rule

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

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

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

1.5.5 Second Derivatives

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

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

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

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:

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

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]

$$\nabla^2\mathbf{T} = (\nabla^2 T_x)\hat{\mathbf{x}} + (\nabla^2 T_y)\hat{\mathbf{y}} + (\nabla^2 T_z)\hat{\mathbf{z}}. \tag{1.87}$$

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.

1.5.6 Dirac Delta Function

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

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{r}}}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left(\frac{1}{r^2}\right)\right) = 0, \tag{1.88}$$

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

$$\oint_S \left(\frac{\hat{\mathbf{r}}}{R^2}\right)\cdot(R^2\sin\theta d\theta d\phi\hat{\mathbf{r}}) = 4\pi. \tag{1.89}$$

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

$$\delta(x) = \begin{cases} \infty & \text{if} \quad x = 0 \\ 0 & \text{if} \quad x \neq 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty}\delta(x)dx = 1.^6 \tag{1.90}$$

The appearance of this function is shown in fig. 1.5.

[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.

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

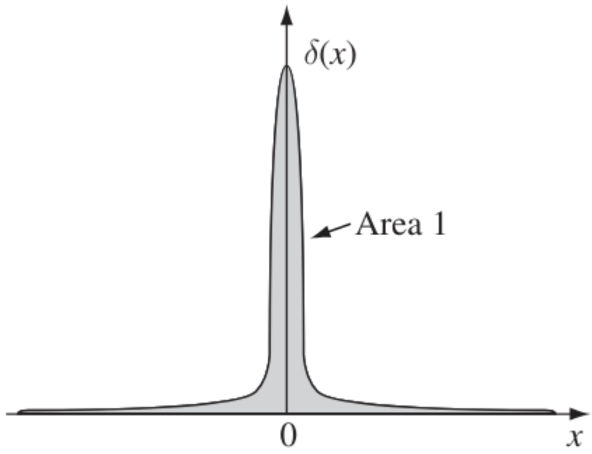

Figure 1.5

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

$$f(x)\delta(x) = f(0)\delta(x), \tag{1.91}$$

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

Integrating the above equation,

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = \int_{-\infty}^{+\infty} f(0)\delta(x)dx = f(0)\int_{-\infty}^{+\infty} \delta(x)dx = f(0). \tag{1.92}$$

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

By changing the variable in eq. (1.92), we can pick out the value of $f(x)$ at any arbitary point $x = a$

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a) \tag{1.93}$$

Example: Dirac Delta Function (1).

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

Solution: Consider the integral

$$\int_{-\infty}^{+\infty} f(x)\delta(kx)dx, \tag{1.94}$$

Making the substitution $u = kx$, we have

$$\int_{-\infty}^{+\infty} f(x)\delta(kx)dx = \pm\frac{1}{k}\int_{-\infty}^{+\infty} f\left(\frac{u}{|k|}\right)\delta(u)du = \frac{1}{|k|}f(0) = \int_{-\infty}^{+\infty} f(x)\frac{\delta(x)}{|k|}dx. \tag{1.95}$$

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:

$$\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z) \tag{1.96}$$

The characteristic equations the three-dimensional delta function are then

$$\int_{all\ space} \delta^3(\mathbf{r})d\tau = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \delta(x)\delta(y)\delta(z)\ dxdydz = 1 \tag{1.97}$$

and

$$\int_{all\ space} f(\mathbf{r})\delta^3(\mathbf{r}-\mathbf{a}) = f(\mathbf{a}). \tag{1.98}$$

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

$$\boldsymbol{\nabla}\cdot\frac{\hat{\mathbf{r}}}{r^2} = 4\pi\delta^3(\mathbf{r}). \tag{1.99}$$

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

In general,

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2}\right) = 4\pi\delta^3(\boldsymbol{\mathscr{r}}). \tag{1.100}$$

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

From eq. (1.59), we have also

$$\nabla^2\left(\frac{1}{\mathscr{r}}\right) = -4\pi\delta^3(\boldsymbol{\mathscr{r}}). \tag{1.101}$$

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 $\int_V \rho(r)d\tau = \int A\delta(r-a)4\pi r^2 dr = 4\pi A a = q$, 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

$$\mathbf{v} = \frac{1}{\rho}\hat{\boldsymbol{\phi}}. \tag{1.102}$$

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

Solution: Fro this velocity field $\boldsymbol{\nabla}\times\mathbf{v} = 0$ everywhere except on the axis $\rho = 0$ where $\mathbf{v}$ has a singularity. Therefore $\displaystyle\oint_C \mathbf{v}\cdot d\mathbf{r} = 0$ 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

$$\boldsymbol{\nabla}\times\mathbf{v} = 2\pi\delta(\rho). \tag{1.103}$$

1.5.7 Helmholtz Theorem

Theorem 1.5.1 (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* $\dfrac{1}{r}$ *as* $r \to \infty$*, where* Φ *and* $\mathbf{A}$ *are given by*

$$\Phi(\mathbf{r}) = \frac{1}{4\pi}\int_{all\ space} \frac{\boldsymbol{\nabla}' \cdot \mathbf{T}(\mathbf{r}')}{\mathscr{r}} d\tau' + C \quad and \quad \mathbf{A}(\mathbf{r}) = \frac{1}{4\pi}\int_{all\ space} \frac{\boldsymbol{\nabla}' \times \mathbf{T}(\mathbf{r}')}{\mathscr{r}} d\tau' + \mathbf{C} + \boldsymbol{\nabla} D. \tag{1.104}$$

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 there exists an unique solution to the equation

$$-\nabla^2\Phi = -\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\Phi) = \boldsymbol{\nabla} \cdot \mathbf{T} \implies \boldsymbol{\nabla} \cdot (\mathbf{T} + \boldsymbol{\nabla}\Phi) = 0. \tag{1.105}$$

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

$$\mathbf{T} = -\boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times \mathbf{A} \tag{1.106}$$

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

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{T} &= -\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\Phi) + \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = -\nabla^2\Phi \\ &= -\frac{1}{4\pi}\int \boldsymbol{\nabla}' \cdot \mathbf{T}(\mathbf{r}')\nabla^2\left(\frac{1}{\mathscr{r}}\right) d\tau' \\ &= \int \boldsymbol{\nabla}' \cdot \mathbf{T}(\mathbf{r}')\delta^3(\mathbf{r} - \mathbf{r}') d\tau' = \boldsymbol{\nabla} \cdot \mathbf{T}(\mathbf{r}). \end{aligned} \tag{1.107}$$

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,

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

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

[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.

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

$$
\begin{aligned}
4\pi(\boldsymbol{\nabla}\cdot\mathbf{A}) &= \int \boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}')}{\eta}\right)d\tau' \\
&= \int(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot\boldsymbol{\nabla}\left(\frac{1}{\eta}\right)d\tau' = -\int(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot\boldsymbol{\nabla}'\left(\frac{1}{\eta}\right)d\tau' \\
&= \int\frac{1}{\eta}\boldsymbol{\nabla}'\cdot(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))d\tau - \oint\frac{1}{\eta}(\boldsymbol{\nabla}'\times\mathbf{T}(\mathbf{r}'))\cdot d\mathbf{A} \\
&= -\int\frac{1}{\eta}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{T}(r'))d\tau = 0.
\end{aligned}
\tag{1.109}
$$

where we used $\boldsymbol{\nabla}\left(\frac{1}{\eta}\right) = -\boldsymbol{\nabla}'\left(\frac{1}{\eta}\right)$ since the derivative of $\eta = |\mathbf{r}-\mathbf{r}'|$ 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 $\frac{1}{r^2}$ 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

$$
\int\frac{X(r')}{r'}r'^2dr', \tag{1.110}
$$

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 $\frac{1}{r^2}$ 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.

□

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

2 Multiple Integration

Single, double and triple integrals describe integrations over an interval in one dimension, an area in two dimensions and a volume in three dimensions, respectively. The infinitisimal quantities are (in Cartesian coordinates), $dx, dA = dxdy$ and $dV = dxdydz$, respectively, which are all scalar quantities.

The most general forms (in Cartesian coordiantes) are

$$\int_a^b f(x)dx, \quad \int_R f(x,y)dA, \quad \text{and} \quad \int_V f(x,y,z)dV. \tag{2.1}$$

Two or three integral signs will be used if $dA = dxdy$ or $dV = dxdydz$ are written out explicitly.

2.1 Double Integration

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

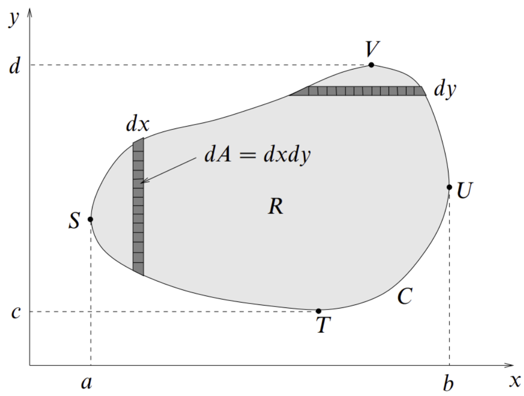

Figure 2.1

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

$$I = \int_c^d \int_{x_1(y)}^{x_2(y)} f(x,y)dxdy, \tag{2.2}$$

where $x_1(y)$ and $x_2(y)$ are the equations of the curves TSV and TUV respectively.

For a specific range $y \to y + dy$, the inner integral calculates the contribution to I by the horizontal strip located from y to $y + dy$. The outer integral then sums up the contributions of these horizontal strips.

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

$$I = \int_a^b \int_{y_1(x)}^{y_2(x)} f(x,y)dydx, \tag{2.3}$$

where $y_1(x)$ and $y_2(x)$ are the equations of the curves STU and SVU respectively.

A useful trick invovle Pappus' second theorem, which states that the surface of revolution of a plance curve is given by the length of the curve L multiplied by the distance moved by its centroid, since the suface area generated is given by $S = \displaystyle\int 2\pi y ds = 2\pi\overline{y}L$, where $\overline{y} = \dfrac{1}{L}\displaystyle\int yds$ is the definition of the centroid.

2.2 Triple Integration

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 $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ as shown in fig. 2.2.

Solution: $V = \displaystyle\int_0^a \int_0^{b-\frac{bx}{a}} \int_0^{c(1-\frac{y}{b}-\frac{x}{a})} dxdydz$. The limit of integral over z can be obtained as integration over z adds up the boxes from the shaded column in the figure. A quicker way is to simply sum up all the vertical columns as $V = \displaystyle\int_0^a \int_0^{b-\frac{bx}{a}dyc\left(a-\frac{y}{b}-\frac{x}{a}\right)}$ which skips the step of integration over z as it is trivial. The result is $V = \dfrac{abc}{6}$.

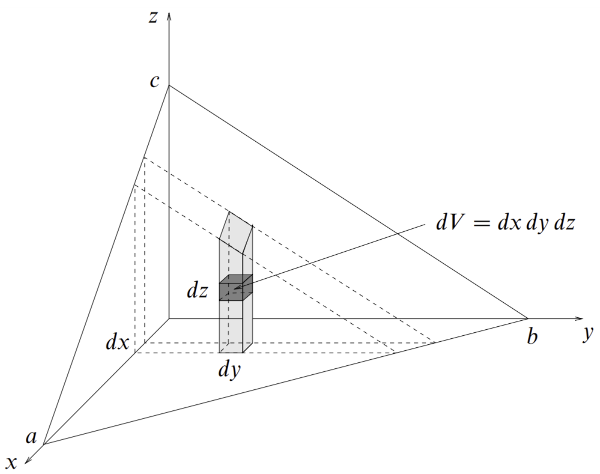

Figure 2.2

A useful trick invovle Pappus' first theorem, which states that the volume of revolution of a plance surface is given by the area of the surface A multiplied by the distance moved by its centroid, since the suface area generated is given by $S = \int 2\pi y dA = 2\pi \overline{y} A$, where $\overline{y} = \frac{1}{A}\int y dA$ is the definition of the centroid.

2.3 Jacobian

To change variables in a double integral from (x, y) to (u, v), we have to find the express the infinitisimal area $dA = dxdy$ in terms of u and v.

Refer to fig. 2.3, since v is constant along KL, the line element KL can be written as $d\mathbf{r}_{KL} = \frac{\partial}{\partial u}(x\hat{\mathbf{x}} + y\hat{\mathbf{y}})du = \frac{\partial x}{\partial u}du\hat{\mathbf{x}} + \frac{\partial y}{\partial u}du\hat{\mathbf{y}}$. Similarly, $d\mathbf{r}_{KN} = \frac{\partial x}{\partial v}dv\hat{\mathbf{x}} + \frac{\partial y}{\partial v}dv\hat{\mathbf{y}}$. Thus the area of the parallelogram $KLMN$ is given by

$$dA_{uv} = \left|\frac{\partial x}{\partial u}du\frac{\partial y}{\partial v}dv - \frac{\partial x}{\partial v}dv\frac{\partial y}{\partial u}du\right| = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|dudv, \tag{2.4}$$

where

$$J = \frac{\partial(x, y)}{\partial(u, v)} \equiv \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}. \tag{2.5}$$

is defined as the Jacobian of (x, y) with respect to (u, v).

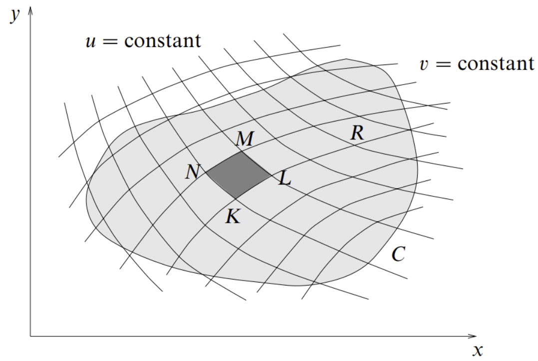

Figure 2.3

Similar to the Jacobian in double integral, the Jacobian in triple integral is defined as

$$J = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{pmatrix}. \tag{2.6}$$

For three sets of variables x_i, y_i and z_i, with i running from 1 to n. We know from eq. (1.12) that

$$\frac{\partial x_i}{\partial z_j} = \sum_{k=1}^{n} \frac{\partial x_i}{\partial y_k} \frac{\partial y_k}{\partial z_j}. \tag{2.7}$$

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

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \implies C = AB \implies \det(C) = \det(A)\det(B) \implies J_{xz} = J_{xy} J_{yz}. \tag{2.8}$$

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

3

Line, Surface and Volume Integration

Line, surface and volume integrals describe integrations over a line, a surface and a volume.[1], respectively, all in three dimensional space. The infinitisimal quantities are (in Cartesian coordinates), $d\mathbf{r} = dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}} + dz\hat{\mathbf{z}}, dS^2$ and $dV = dxdydz$, out of which the first two are vectors and the last one is scalar[3]

3.1 Line Integrals

3.1.1 Evaluation

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

$$\int_C f(x,y,z)dr, \quad \int_C f(x,y)d\mathbf{r}, \quad \int_C \mathbf{F}(x,y,z)\cdot d\mathbf{r} \;\text{ and }\; \int_C \mathbf{F}(x,y,z)\times d\mathbf{r}. \tag{3.1}$$

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.

Here the argument (x, y, z) can be unambigously interchanged by $(\mathbf{r})$, where $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ is the position vector. The relation between the two forms can be given explicityl 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.

[1]Thus a volume integral is equivalent to a triple integral. Since an infinitesimal volume has no preferred direction in three-dimensional space.

[2]The general form of dS is not trivial. It is exactly this property which makes surface integral the hardest type to deal with.

[3]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).

Example: Line Integral (1).

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

$$I = \int_{(1,1)}^{(4,2)} ((x+y)dx + (y-x)dy). \tag{3.2}$$

1. Along the parabola $y^2 = x$, we have $2ydy = dx$, so
$$I = \int_1^2 ((y^2+y)2y + (y-y^2))dy = \frac{34}{3}. \tag{3.3}$$
2. We have $dx = (4u+1)du$ and $dy = 2udu$, so
$$I = \int_0^1 ((3u^2+u+2)(4u+1) - (u^2+u)(2u))du = \frac{32}{3} \tag{3.4}$$
3. We split the integral into two parts, then
$$\begin{aligned} I &= \int_{(1,1)}^{(4,1)} ((x+y)dx + (y-x)dy) + \int_{(4,1)}^{(4,2)} ((x+y)dx + (y-x)dy) \\ &= \int_1^4 (x+1)dx + \int_1^2 (y-4)dy = 8. \end{aligned} \tag{3.5}$$

Example: Line Integrals (2).

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

$$I = \int_{-a}^{a} \sqrt{a^2 - y^2}dy + \int_a^{-a} (-\sqrt{a^2 - y^2}dy = \pi a^2). \tag{3.6}$$

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

$$I = a^2 \int_0^{2\pi} \cos^2\phi d\phi = \pi a^2. \tag{3.7}$$

Example: Line Integral (3).

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

$$\mathbf{r}(\phi) = a\cos\phi\hat{\mathbf{x}} + a\sin\phi\hat{\mathbf{y}} \text{ and } ds = \sqrt{\frac{d\mathbf{r}}{d\phi}\cdot\frac{d\mathbf{r}}{d\phi}} = ad\phi. \tag{3.8}$$

Thus

$$I = \int_0^{\pi} a^3(1-\sin 2\phi)d\phi = \pi a^3. \tag{3.9}$$

3.2 Surface Integrals

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

$$\int_S \phi dS, \quad \int_S \phi d\mathbf{S}, \quad \int_S \mathbf{a}\cdot d\mathbf{S} \quad \text{and} \quad \int_S \mathbf{a}\times d\mathbf{S}, \tag{3.10}$$

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 $\in \mathbb{R}^3$.

To find the genearl form of dS, we project the surface S onto the xy-plane. From fig. 3.1, we see that

$$dA = |\cos\alpha|dS \implies dS = \frac{dA}{\hat{\mathbf{n}}\cdot\mathbf{k}} = \frac{dA}{\frac{\nabla f}{|\nabla f|}\hat{\mathbf{k}}} = \frac{|\nabla f|dA}{\frac{\partial f}{\partial z}}, \tag{3.11}$$

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.

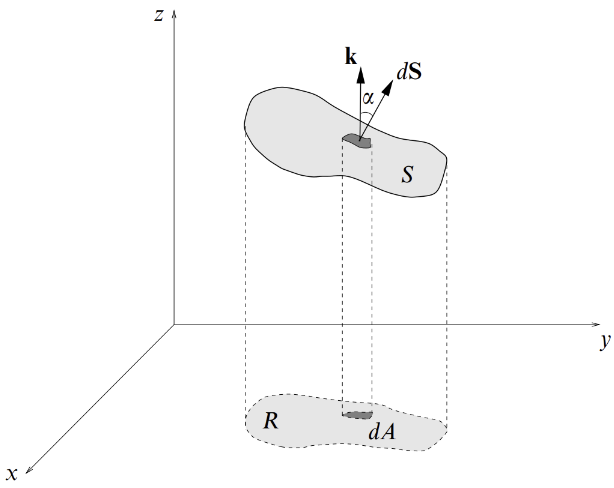

Figure 3.1

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 parallet 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.

Example: Surface Integral.

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

Solution: Since

$$\mathbf{F} \cdot d\mathbf{S} = x(\hat{\mathbf{x}} \cdot \hat{\mathbf{r}})d\mathbf{S} = (a \sin\theta \cos\phi)(\sin\theta\cos\phi)(a^2 \sin\theta d\theta d\phi), \tag{3.12}$$

we have

$$I = a^3 \int_0^{\frac{\pi}{2}} \sin^3\theta d\theta \int_0^{2\pi} \cos^2\phi d\phi = \frac{2\pi a^3}{3}. \tag{3.13}$$

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

$$|\boldsymbol{\nabla} f| = 2|\mathbf{r}| = 2a, \quad \frac{\partial f}{\partial z} = 2z = 2\sqrt{a^2 - x^2 - y^2} \text{ and } \hat{\mathbf{x}} \cdot \hat{\mathbf{r}} = \hat{\mathbf{x}} \cdot \frac{\mathbf{r}}{|\mathbf{r}|} = \frac{x}{a}. \tag{3.14}$$

Therfore the integral becomes

$$I = \int\int_R \frac{x^2}{\sqrt{a^2 - x^2 - y^2}} dxdy = \frac{2\pi a^3}{3}. \tag{3.15}$$

3.2.1 Vector Areas

The vector area of a surface S is defined as

$$\mathbf{S} = \int_S d\mathbf{S}. \tag{3.16}$$

A closed surface will always has a zero vector area, since when projecting onto the xy-plane, from eq. (3.11) we see that every $d\mathbf{S}_+ = \frac{dA}{|\hat{\mathbf{n}}\cdot\hat{\mathbf{k}}|}$ will have an opposite contribution $d\mathbf{S}_- = \frac{dA}{|-\hat{\mathbf{n}}\cdot\hat{\mathbf{k}}|} = -d\mathbf{S}_+$.

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

$$\mathbf{S} = \frac{1}{2}\oint_C \mathbf{r} \times d\mathbf{r}, \tag{3.17}$$

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. 3.2 and its area is the sum of all the infinitesimal triangle, each with vector area $d\mathbf{S} = \frac{1}{2}\mathbf{r} \times d\mathbf{r}$.

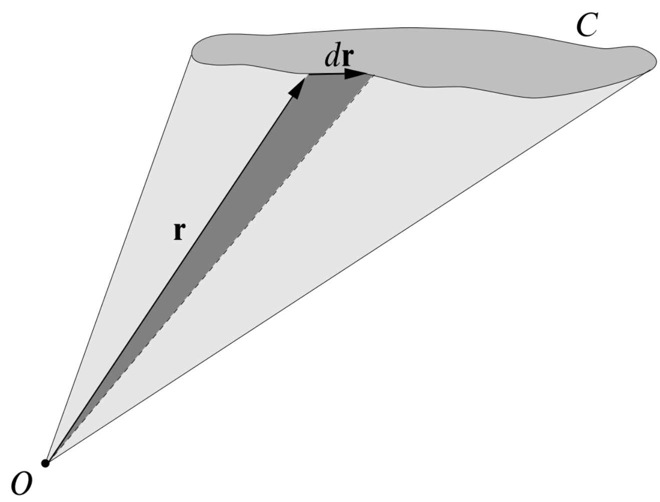

Figure 3.2

For a surface confined to the xy-plane, $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$ and $d\mathbf{r} = dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}}$, thus $\mathbf{r} \times d\mathbf{r} = (xdy - ydx)\hat{\mathbf{z}}$, so the area is what we have found earlier in eq. (3.30).

3.2.2 Solid Angle

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

$$\Omega = \int_S \frac{\hat{\mathbf{r}} \cdot d\mathbf{S}}{r^2}. \tag{3.18}$$

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

3.3 Volume Integrals

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

$$\int_V f(x, y, z) dV \text{ and } \int_V \mathbf{F}(x, y, z) dV. \tag{3.19}$$

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. 3.3, we have

$$V = \int_V dV = \frac{1}{3} \oint_S \mathbf{r} \cdot d\mathbf{S}, \tag{3.20}$$

as the volume of each cone is $dV = \frac{1}{3}\mathbf{r} \cdot d\mathbf{S}$.

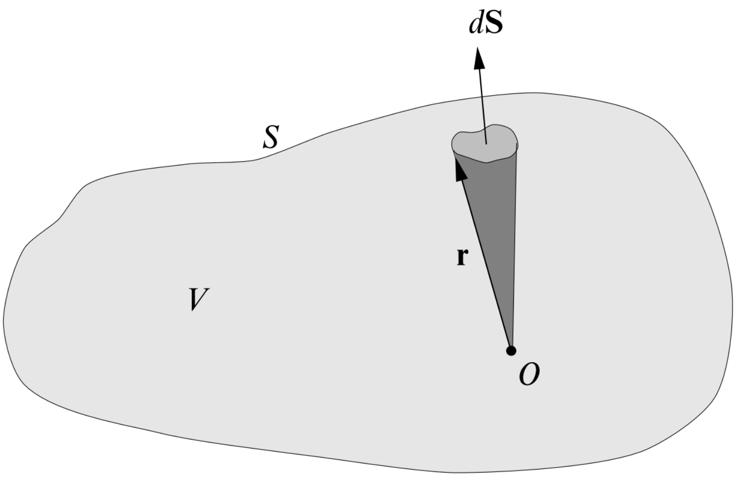

Figure 3.3

3.4 Integral Theorems

3.4.1 The Divergence Theorem

The Divergence Theorem in Three Dimensions

The divergence theorem (in three dimensions) states that

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{F}) dV = \oint_S \mathbf{F} \cdot d\mathbf{S}. \tag{3.21}$$

Example: Surface Integral by the Divergence Theorem.

Question: Evaluate the surface integral $I = \int_S \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = (y - x)\hat{\mathbf{x}} + x^2 z\hat{\mathbf{y}} + (z + x^2)\hat{\mathbf{z}}$ 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

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{F}) dV = 0 = \int_S \mathbf{F} \cdot d\mathbf{S} + \int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1. \tag{3.22}$$

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

$$I = -\int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1 = \int \int_R x^2 dx dy = \frac{\pi a^4}{4}. \tag{3.23}$$

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

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0. \tag{3.24}$$

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

$$\frac{dM}{dt} = \frac{d}{dt} \int_V \rho dV = -\oint_S \rho \mathbf{v} \cdot d\mathbf{S} \tag{3.25}$$

$$\int_V \frac{\partial \rho}{\partial t} dV + \int_V \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) dV = \int_V \left(\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) \right) dV = 0. \tag{3.26}$$

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 $\boldsymbol{\nabla} \cdot v = 0$.

The Divergence Theorem in Two Dimensions

The divergence theorem (in two dimensions) (also known as the Green's theorem in a plane) states that

$$\oint_C (Pdx + Qdy) = \int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy. \tag{3.27}$$

To prove this theorem we refer to fig. 2.1. Let $y = y_1(x)$ and $y = y_2(x)$ be the equations of the curves STU and SVU respectively. We then find

$$\begin{aligned} \int\int_R \frac{\partial P}{\partial y} dxdy &= \int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y} dy = \int_a^b P(x,y)\Big|_{y=y_1(x)}^{y=y_2(x)} dx \\ &= \int_a^b \left[P(x, y_2(x)) - P(x, y_1(x))\right] dx \\ &= -\int_a^b P(x, y_1(x))dx - \int_b^a P(x, y_2(x))dx = -\oint_C Pdx. \end{aligned} \tag{3.28}$$

If we now let $x = x_1(y)$ and $x = x_2(y)$ as the equations of the curves TSV and TUV respectively, then we can similarly show that

$$\int\int_R \frac{\partial Q}{\partial x} dxdy = \oint_C Qdy. \tag{3.29}$$

Subtracting the two equations gives Green's theorem.

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

$$\oint_C (xdy - ydx) = \int\int_R (1+1)dxdy = 2A. \tag{3.30}$$

Therefore the area of an ellipse is

$$A = \frac{1}{2}\int_0^{2\pi} ab(\cos^2\phi + \sin^2\phi)d\phi = \pi ab. \tag{3.31}$$

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 ttravelling 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.

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

$$\oint_S \phi\boldsymbol{\nabla}\psi \cdot d\mathbf{S} = \int_V \boldsymbol{\nabla}\cdot(\phi\boldsymbol{\nabla}\psi)dV = \int_V (\phi\nabla^2\psi + (\boldsymbol{\nabla}\phi)\cdot(\boldsymbol{\nabla}\psi))dV. \tag{3.32}$$

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

$$\oint_S (\phi\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\phi)\cdot d\mathbf{S} = \int_V (\phi\nabla^2\psi - \psi\nabla^2\phi)dV. \tag{3.33}$$

This is known as the Green's second theorem.

Two Other Theorems

Letting $\mathbf{F}$ in eq. (3.21) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_V \boldsymbol{\nabla} f dV = \oint_S f d\mathbf{S} \text{ and } \int_V \boldsymbol{\nabla}\times\mathbf{F}dV = \oint_S d\mathbf{S}\times\mathbf{F}. \tag{3.34}$$

3.4.2 Stokes' Theorem

Stokes' Theorem in Three Dimensions

The Stokes' Theorem (in three dimensions) states that

$$\int_S (\boldsymbol{\nabla}\times\mathbf{F})dS = \oint_C \mathbf{F}\cdot d\mathbf{r}. \tag{3.35}$$

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

$$\oint_C \mathbf{B}\cdot d\mathbf{r} = \int_S (\boldsymbol{\nabla}\times\mathbf{B})\cdot d\mathbf{S} = \mu_0 I = \mu_0 \int_S \mathbf{J}\cdot d\mathbf{S}. \tag{3.36}$$

Hence

$$\int_S (\boldsymbol{\nabla} \times \mathbf{B} - \mu_0 \mathbf{J}) \cdot d\mathbf{S} = 0. \tag{3.37}$$

The Stokes' Theorem in Two Dimensions

The Stokes' Theorem (in two dimensions) is also yields Green's theorem in a plane, just as the divergence theorem did.

The Two Other Theorems

Letting $\mathbf{F}$ in eq. (3.35) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_S d\mathbf{S} \times \boldsymbol{\nabla} f = \oint_C f d\mathbf{r} \text{ and } \int_S (d\mathbf{S} \times \boldsymbol{\nabla}) \times \mathbf{F} = \oint_C d\mathbf{r} \times \mathbf{F}. \tag{3.38}$$

Appendices

$\mathcal{A}$ 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.

□