

Calculus

Haydn Cheng

June 5, 2025

Contents

1	Partial Differentiation	3
1.1	Ordinary Differentiation	3
1.1.1	The Leibnitz' Theorem	3
1.1.2	Special Points of Functions	3
1.2	Independent and Dependent Variables	4
1.3	1 Constriant	5
1.3.1	Partial Derivatives	5
1.3.2	Total Derivatives	5
1.4	More than 1 Constriant	5
1.4.1	Partial Derivatives	6
1.4.2	Formal Partial Derivatives	6
1.4.3	Total Derivatives	7
1.4.4	Chain Rule	7
1.5	Exact Differentials	8
1.6	Elimination, Chain Rule and Differentials	9
1.6.1	Elimination	9
1.6.2	Chain Rule	9
1.6.3	Differentials	10
1.6.4	Intepretation	11
1.7	Symmetry of Mixed Derivatives, Reciprocity Relation and Cyclic Relation	14
1.8	Change of Variables	14
1.8.1	Change of Gradients	14
1.8.2	Change of Areas and Volumes	16
1.8.3	Multiple Jacobians	16
1.9	Taylor's Theorem	19
1.10	Speical Points of a Function	19
2	Vector Calculus	23
2.1	Curvilinear coordinates	23
2.1.1	General Curvilinear Coordinates	23
2.1.2	Spherical Coordinates	25
2.1.3	Cylindrical Coordinates	26

2.2	Gradient, Divergence and Curl	27
2.2.1	Gradient	28
2.2.2	Divergence	31
2.2.3	Curl	32
2.3	Product Rules	34
2.4	Second Derivatives	36
2.5	Index Notation	37
2.6	Dirac Delta Function	39
2.7	Helmholtz Theorem	42
3	Integration	45
3.1	Multiple Integration	45
3.2	Line, Surface and Volume Integration	48
3.2.1	Curves and Surfaces	48
3.2.2	Line Integrals	57
3.3	Surface Integrals	60
3.3.1	Vector Areas	65
3.3.2	Solid Angle	65
3.4	Volume Integrals	66
3.5	Integral Theorems	67
3.5.1	The Divergence Theorem	67
3.5.2	Stokes' Theorem	70
	Appendices	76
A	Rigorous Proofs	77
A.1	Leibnitz' Theorem	77

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 functions $f(x) = u(x)v(x)$ is¹*

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

1.1.2 Special Points of Functions

A stationary point 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 function changes.

Fundamental Theorem of Calculus

The fundamental theorem of calculus states that if function $F(x)$ can be written as

$$F(x) = \int_a^x f(t)dt, \quad (1.2)$$

then its derivative $F'(x)$ is

¹Proof given in section A.1.

$$F'(x) = f(x). \quad (1.3)$$

Example: Fundamental Theorem of Calculus.

Question: Find the partial derivatives F_x and F_y , $F(x, y)$ can be written as

$$F(x, y) = \int_0^{xy^2} e^{t^2} dt. \quad (1.4)$$

Solution: Treating $u \equiv xy^2$ as a single variable we apply the fundamental theorem of calculus to get

$$\frac{\partial F(x, y)}{\partial(xy^2)} = e^{(xy^2)^2}. \quad (1.5)$$

The partial derivatives F_x and F_y can then be obtained by

$$F_x = \frac{\partial F(x, y)}{\partial(xy^2)} \frac{\partial(xy^2)}{x} = y^2 e^{(xy^2)^2} \quad \text{and} \quad F_y = \frac{\partial F(x, y)}{\partial(xy^2)} \frac{\partial(xy^2)}{\partial y} = 2xy e^{(xy^2)^2}. \quad (1.6)$$

1.2 Independent and Dependent Variables

Variables are some quantities that can change. Relations or constraints are equations that relate the variables and put constraints on how they can change.

Independent Variables are variables that can change independently without affecting other independent variables. If there are no constraints then all the variables are independent variables and they can change freely without one affecting another.

Dependent Variables are variables that would change when the independent variables change due to constraints. If there are, say, 3 variables x, y, z and 1 constraint then there would be 1 dependent variables and $3 - 1 = 2$ independent variables. The constraint can be written in the form $z = z(x, y)$, meaning that z is the dependent variable that depends on x and y , which are the independent variables. In other words, z is a function of x and y .

It is important to note that the assignation of independent and dependent variables is entirely arbitrary. As in the case of the above example, we can call any 2 of the 3 variables x, y, z independent, and the rest change according to the constraint. That is to say that the constraint may also be written in the form $x(y, z)$ or $y(x, z)$.

That said, there is usually an obvious dependent variable, such as the price of an orange z , with obvious independent variables, such as the supply x and the demand y .

1.3 1 Constraint

Since the whole point of multivariable calculus is to find how a variable change with another when there are more than the two usual variables x and y , there is at least one constraint imposed on the variables.

We start our discussion by considering the bare-minimum – 3 variables x, y, f with 1 constraint $f = f(x, y)$. Here we use the variable f instead of z to emphasize that f is chosen (arbitrary) to be our dependent variable and x and y can change independently, *i.e.*, $f(x, y)$ is a function of x and y .

1.3.1 Partial Derivatives

The rate in which f changes as one of the independent variables, say x , changes while keeping the other, say y , is found by the partial derivative of f with respect to x as

$$\left(\frac{\partial f}{\partial x}\right)_y \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}. \quad (1.7)$$

1.3.2 Total Derivatives

To consider how $f(x, y)$ changes when both x and y change, we can calculate the infinitesimal difference of f as

$$\begin{aligned} \Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= f(x + \Delta x, y + \Delta y) - f(x, y) + f(x, y + \Delta y) - f(x, y + \Delta y) \\ &= \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x} \Delta x + \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \Delta y \\ &\approx \left(\frac{\partial f}{\partial x}\right)_y \Delta x + \left(\frac{\partial f}{\partial y}\right)_x \Delta y. \end{aligned} \quad (1.8)$$

As $\Delta x \rightarrow 0$ and $\Delta y \rightarrow 0$, we have the total differential df as

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy. \quad (1.9)$$

1.4 More than 1 Constraint

Suppose apart from the constraint (relation to be more precise) that determines $f = f(x, y)$, we introduce an extra dependent variable $z = z(x, y)$ as one of the argument of f , *i.e.*, $f = f(x, y, z(x, y))$, so there are now 4 variables and 2 constraints in total.

Again, we can also, say, treat y and z as the independent variables and x as the dependent variable, so it would be misleading to write $f = f(x, y, z(x, y))$. We would simply bear in mind that x, y, z related by 1 constraint and write $f = f(x, y, z)$.

1.4.1 Partial Derivatives

In the 2 arguments case discussed above, since x and y are independent of each other so to find the rate of change in f with respect to x only we fix y to be constant.

Now however, out of the 3 arguments, only two are independent, and the one remaining dependent variable changes as the independent variables change. We therefore have to specifying, when finding the rate of change in f with respect to x , which other argument, y or z are we keeping constant.²

If we keep y to be the independent variable which is kept constant, then the partial derivative of f with respect to x is

$$\left(\frac{\partial f}{\partial x}\right)_y \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z(x + \Delta x, y)) - f(x, y, z(x, y))}{\Delta x}. \quad (1.10)$$

On the other hand if we keep z to be the independent variable which is kept constant, then the partial derivative of f with respect to x is

$$\left(\frac{\partial f}{\partial x}\right)_z \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y(x + \Delta x, z), z) - f(x, y(x, z), z)}{\Delta x}. \quad (1.11)$$

In general, if there are n variables with m constraints, then we have to specify which $n - m - 1$ variables are kept constant while finding the derivative with respect to one of the n variables.

For instance if we have the function and constraint as

$$f = f(x, y, z, t), \quad xy = zt, \quad (1.12)$$

then only 3 out of the 5 variables x, y, z, t, f can be independent; the fourth and fifth are then determined through the equation $xy = zt$ and $f = f(x, y, z, t)$. Thus we would write expressions like

$$\left(\frac{\partial f}{\partial x}\right)_{y,t} = \text{partial derivative of } f \text{ with respect to } x \text{ while keeping } y \text{ and } t \text{ constant.} \quad (1.13)$$

1.4.2 Formal Partial Derivatives

The formal partial derivative is to carry out a partial derivative assuming that all the arguments are independent of each other and keeping all of them constant except for the variable under differentiation. For example, the formal partial derivative of $f(x, y, z)$ with respect to x is

²Note that we can instead keep things like $\sqrt{y+z} \sin(y-z)$ to be constant, and y and z change as x and $\sqrt{y+z} \sin(y-z)$ change independently. However, this is equivalent to introducing yet another variable and we shall refrain from complicating simple thing. So we will assume only y or z can be kept constant (when differentiating with respect to x).

$$\left(\frac{\partial f}{\partial x}\right)_{y,z} \equiv f_x \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}. \quad (1.14)$$

This form of partial derivative becomes handy when discussing the total derivative when the arguments are related.

1.4.3 Total Derivatives

To consider how $f(x, y, z(x, y))$ changes when x, y and z change, we can calculate the infinitesimal difference of f as

$$\begin{aligned} \Delta f &= f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) \\ &= f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) \\ &\quad + f(x, y + \Delta y, z + \Delta z) - f(x, y + \Delta y, z + \Delta z) \\ &\quad + f(x, y, z + \Delta z) - f(x, y, z + \Delta z) \\ &= \frac{f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y + \Delta y, z + \Delta z)}{\Delta x} \Delta x \\ &\quad + \frac{f(x, y + \Delta y, z + \Delta z) - f(x, y, z + \Delta z)}{\Delta y} \Delta y \\ &\quad + \frac{f(x, y, z + \Delta z) - f(x, y, z)}{\Delta z} \Delta z \\ &\approx \left(\frac{\partial f}{\partial x}\right)_{y,z} \Delta x + \left(\frac{\partial f}{\partial y}\right)_{x,z} \Delta y + \left(\frac{\partial f}{\partial z}\right)_{x,y} \Delta z. \end{aligned} \quad (1.15)$$

As $\Delta x \rightarrow 0, \Delta y \rightarrow 0$ and $\Delta z \rightarrow 0$, we have the total differential df as

$$df = \left(\frac{\partial f}{\partial x}\right)_{y,z} dx + \left(\frac{\partial f}{\partial y}\right)_{x,z} dy + \left(\frac{\partial f}{\partial z}\right)_{x,y} dz = f_x dx + f_y dy + f_z dz. \quad (1.16)$$

It is important to note that the partial derivatives involved in the total differential df are formal partial derivatives but not actual partial derivatives, although x, y, z might depend on each other.

1.4.4 Chain Rule

We are too close to introducing the chain rule in multivariable calculus that it make no sense to jump to the next topic.

If we want to evaluate the partial derivative of $f(x, y, z)$ with respect to x while keeping y constant (and z be whatever $z = z(x, y)$ be), one can in principle use eq. (1.10). However, a faster and more practical method is to use the chain rule, which is when we divide both sides in eq. (1.16) by dx while keeping y constant

$$\left(\frac{\partial f}{\partial x}\right)_y = \left(\frac{\partial f}{\partial x}\right)_{y,z} + \left(\frac{\partial f}{\partial z}\right)_{x,y} \left(\frac{\partial z}{\partial x}\right)_y = f_x + f_z \left(\frac{\partial z}{\partial x}\right)_y. \quad (1.17)$$

If we use the form $f(x, y, z) = f(x(z), y(z), z) = f(z)$, then $f_x = 0$ and the above equation becomes reminiscent to the ordinary chain rule

$$\left(\frac{\partial f}{\partial x}\right)_y = \left(\frac{\partial f}{\partial z}\right)_{x,y} \left(\frac{\partial z}{\partial x}\right)_y. \quad (1.18)$$

The same formula can be more easily obtain if we first write the differential of f , *i.e.*, df in terms of z , *i.e.*, $df = f_z dz$.

If for example, instead, we have 3 constraints for the 4 variables, namely $f = (x, y, z)$, $x = x(z)$ and $y = y(z)$, then we can find the partial (now ordinary) derivative of f with respect to z as

$$\frac{df}{dz} = \left(\frac{\partial f}{\partial x}\right)_{y,z} \frac{dx}{dz} + \left(\frac{\partial f}{\partial y}\right)_{x,z} \frac{dy}{dz} + \left(\frac{\partial f}{\partial z}\right)_{x,y} = f_x \frac{dx}{dz} + f_y \frac{dy}{dz} + f_z. \quad (1.19)$$

1.5 Exact Differentials

Before discussing how to evaluate partial derivatives when the arguments are dependent, we mention the concept of exact differentials, which is closely related to the concept of total derivatives.

An arbitrary differential

$$A(x, y)dx + B(x, y)dy \quad (1.20)$$

is exact if it is the total derivative of a function $f(x, y)$

$$df(x, y) = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy. \quad (1.21)$$

Therefore, we have

$$A(x, y) = \frac{\partial f}{\partial x} \quad \text{and} \quad B(x, y) = \frac{\partial f}{\partial y} \quad (1.22)$$

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

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}.^3 \quad (1.23)$$

³Determining whether a differential containing many variables x_1, x_2, \dots, x_n is exact is a simple extension of the above: a differential $df = \sum_{i=1}^n g_i(x_1, x_2, \dots, x_n) dx_i$ is exact if $\partial g_i / \partial x_j = \partial g_j / \partial x_i$ for all pairs i, j .

1.6 Elimination, Chain Rule and Differentials

There are three ways to carry out partial derivatives when the arguments are dependent, namely elimination, chain rule and differentials.

This is best illustrated by an example. Suppose we have the function

$$f(x, y) = x^2 + y^2 + (x^2 + y^2)^2 \quad (1.24)$$

Now we introduce an extra variable $z(x, y) = x^2 + y^2$, so that f now can be rewritten in virtually infinite ways

$$f(x, y, z) = x^2 + y^2 + z^2 = z + z^2 = x^2 + y^2 + z(x^2 + y^2) \quad \text{etc.} \quad (1.25)$$

We, in the next few subsections are going to find the partial derivatives of f with respect to x and z ⁴ using the three methods mentioned.

1.6.1 Elimination

To evaluate the partial derivative of f with respect to x while keeping y constant (and let z be whatever $z = x^2 + y^2$ is), we eliminate z from $z = z(x, y)$ and differentiate the function $f = f(x, y)$

$$\left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial}{\partial x}(x^2 + y^2 + (x^2 + y^2)^2) = 2x + 4x^3 + 4xy^2. \quad (1.26)$$

To evaluate the partial derivative of f with respect to x while keeping z constant (and let y be whatever $y = \pm\sqrt{z - x^2}$ is), we eliminate y from $y = y(x, z)$ and differentiate the function $f = f(x, z)$

$$\left(\frac{\partial f}{\partial x}\right)_z = \frac{\partial}{\partial x}(x^2 + (z - x^2) + z^2) = 0. \quad (1.27)$$

To evaluate the partial derivative of f with respect to z while keeping x (or y) constant (and let y (or x) be whatever $y = \pm\sqrt{z - x^2}$ is), we eliminate y from $y = y(x, z)$ and differentiate the function $f = f(x, z)$

$$\left(\frac{\partial f}{\partial z}\right)_x = \frac{\partial}{\partial z}(x^2 + (z - x^2) + z^2) = 1 + 2z. \quad (1.28)$$

1.6.2 Chain Rule

To use the chain rule we first write down the total differential of $z(x, y)$, *i.e.*, dz as

$$dz = z_x dx + z_y dy = 2x dx + 2y dy. \quad (1.29)$$

⁴ x and y are symmetric so the partial derivatives of f respect to y is neglected.

Now we divide the whole equation by dx while keeping y constant to get

$$\left(\frac{\partial z}{\partial x}\right)_y = 2x. \quad (1.30)$$

If we instead divide the whole equation by dx while keeping z constant we get

$$\left(\frac{\partial y}{\partial x}\right)_z = -\frac{x}{y}. \quad (1.31)$$

Dividing the whole equation by dz while keeping x constant yields

$$\left(\frac{\partial y}{\partial z}\right)_x = \frac{1}{2y}. \quad (1.32)$$

We can now find the partial derivative of f with respect to x while keeping y constant

$$\left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial}{\partial x}(x^2 + y^2 + z^2) = 2x + 2z \left(\frac{\partial z}{\partial x}\right)_y = 2x + 2(x^2 + y^2)(2x) = 2x + 4x^3 + 4xy^2. \quad (1.33)$$

The partial derivative of f with respect to x while keeping z constant is

$$\left(\frac{\partial f}{\partial x}\right)_z = \frac{\partial}{\partial x}(x^2 + y^2 + z^2) = 2x + 2y \left(\frac{\partial y}{\partial x}\right)_z = 2x + 2\left(-\frac{x}{y}\right) = 0. \quad (1.34)$$

The partial derivative of f with respect to z while keeping x constant is

$$\left(\frac{\partial f}{\partial z}\right)_x = \frac{\partial}{\partial z}(x^2 + y^2 + z^2) = 2y \left(\frac{\partial y}{\partial z}\right)_x + 2z = 2y \left(\frac{1}{2y}\right) + 2z = 2 + 2z. \quad (1.35)$$

1.6.3 Differentials

To use the differentials method we write down the differential forms of the constraints

$$df = 2xdx + 2ydy + 2zdz, \quad (1.36)$$

and

$$dz = 2xdx + 2ydy. \quad (1.37)$$

To find the partial derivative of f with respect to x while keeping y constant we eliminate the differential dz to get

$$df = (2x + 4x(x^2 + y^2))dx + (2y + 4y(x^2 + y^2))dy \implies \left(\frac{\partial f}{\partial x}\right)_y = 2x + 4x^3 + 4xy^2. \quad (1.38)$$

To find the partial derivative of f with respect to x while keeping z constant we eliminate the differential dy to get

$$df = (2z + 1)dz \implies \left(\frac{\partial f}{\partial x} \right)_z = 0. \quad (1.39)$$

To find the partial derivative of f with respect to z while keeping x constant we eliminate the differential dy to get

$$df = (2z + 1)dz \implies \left(\frac{\partial f}{\partial z} \right)_x = 1 + 2z. \quad (1.40)$$

1.6.4 Interpretation

To visualize the results, we refer to fig. 1.1, which shows the relation $z = x^2 + y^2$, and the function $f(x, y, z)$ 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.

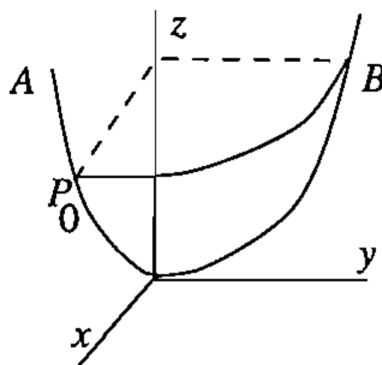


Figure 1.1

Example: Partial and Total Derivatives (1).

Question: Find $(\partial w / \partial y)_{x,t}$, where

$$w(x, y, z, t) = x^3 y - z^2 t \text{ and } xy = zt. \quad (1.41)$$

Solution: With chain rule, we have

$$\left(\frac{\partial w}{\partial y} \right)_{x,t} = x^3 - 2zt \left(\frac{\partial z}{\partial y} \right)_{x,t} = x^3 - 2zx. \quad (1.42)$$

Alternatively, we have

$$dw = 2x^2ydx + x^3dy + 2ztdz - z^2dt \quad \text{and} \quad ydx + xdy = zdt + tdz. \quad (1.43)$$

Eliminating dz , we have

$$dw = (3x^2y - 2zy)dx + (x^3 - 2zx)dy + z^2dt. \quad (1.44)$$

Comparing with the total derivative of $w(x, y, t)$

$$dw = \left(\frac{\partial w}{\partial x} \right)_{y,t} dx + \left(\frac{\partial w}{\partial y} \right)_{x,t} dy + \left(\frac{\partial w}{\partial t} \right)_{x,y} dt, \quad (1.45)$$

we obtain the same answer.

Example: Partial and Total Derivatives (2).

Question: Find $(\partial w / \partial x)_y$ and $(\partial w / \partial y)_x$, where

$$w = x^2 - yz + t^2, \quad z^2 = x + y^2 \quad \text{and} \quad xy = zt. \quad (1.46)$$

Solution: Taking the differential of the three above equations we have

$$dw = 2xdx - zdy - ydz + 2tdt, \quad xdy + ydx = zdt + tdz \quad \text{and} \quad 2zdz = dx + 2ydy. \quad (1.47)$$

Eliminating dt and dz , we have

$$dw = \left(2x - \frac{y}{2z} + \frac{2ty}{z} - \frac{t^2}{z^2} \right) dx + \left(-z - \frac{y^2}{z} + \frac{2xt}{z} - \frac{2t^2y}{z^2} \right) dy. \quad (1.48)$$

Comparing with the total derivative of $w(x, y)$

$$dw = \left(\frac{\partial w}{\partial x} \right)_y dx + \left(\frac{\partial w}{\partial y} \right)_x dy, \quad (1.49)$$

we obtain the desired solutions.

It would be very messy if we try to compute the partial derivatives directly like the first method in the above example.

Example: Partial and Total Derivatives (3).

Question: Suppose the variables x, y, z satisfy an equation $g(x, y, z) = 0$ and the point $P = (1, 1, 1)$ lies on the surface $g = 0$ with $(\nabla g)_P = (-1, 1, 2)$.

Let $f(x, y, z)$ be another function and assume that $(\nabla f)_P = (1, 2, 1)$.

Find the gradient of the function $w = f(x, y, z(x, y))$ of the two independent variables x and y at the point $x = 1$ and $y = 1$.

Solution: In differential forms we have

$$(dw)_P = dx + 2dy + dz, \quad (dg)_P = -dx + dy + 2dz = 0. \quad (1.50)$$

Eliminating dz we have

$$(dw)_P = \frac{3}{2}dx + \frac{3}{2}dy \implies (\nabla w)_P = \frac{3}{2}\hat{\mathbf{i}} + \frac{3}{2}\hat{\mathbf{j}}. \quad (1.51)$$

Example: Formal and Actual Partial Derivatives (1).

Question: Suppose $w = w(x, r)$, with $r = r(x, \theta)$. Give an expression for $(\partial w / \partial r)_\theta$ in terms of formal partial derivatives of w and r .

Solution: The total derivative of w with r and θ being the independent variables is

$$dw = w_x dx + w_r dr. \quad (1.52)$$

Dividing the equation by dr while keeping θ constant, we have

$$\left(\frac{\partial w}{\partial r}\right)_\theta = w_x \left(\frac{\partial x}{\partial r}\right)_\theta + w_r = w_x \left(\frac{\partial r}{\partial x}\right)_\theta^{-1} + w_r = \frac{w_x}{r_x} + w_r. \quad (1.53)$$

Example: Formal and Actual Partial Derivatives (2).

Question: If $f(x, y, z) = xy^2z^4$, where $z = 2x + 3y$, find the three formal derivatives f_x, f_y and f_z and the three of the many possible actual partial derivatives $(\partial f / \partial x)_y, (\partial f / \partial y)_x$ and $(\partial f / \partial z)_x$.

Solution: The three formal derivatives are

$$f_x = y^2z^4, \quad f_y = 2xyz^4 \quad \text{and} \quad f_z = 4xy^2z^3. \quad (1.54)$$

The three actual partial derivatives are

$$\begin{cases} \left(\frac{\partial f}{\partial x}\right)_y = f_x + f_z \left(\frac{\partial z}{\partial x}\right)_y = y^2z^4 + 8xy^2z^3, \\ \left(\frac{\partial f}{\partial y}\right)_x = f_y + f_z \left(\frac{\partial z}{\partial y}\right)_x = 2xyz^4 + 12xy^2z^3, \\ \left(\frac{\partial f}{\partial z}\right)_x = f_y \left(\frac{\partial y}{\partial z}\right)_x + f_z = \frac{2}{3}xyz^4 + 4xy^2z^3, \end{cases} \quad (1.55)$$

where we have used the chain rule at each line.

1.7 Symmetry of Mixed Derivatives, Reciprocity Relation and Cyclic Relation

A very useful fact about the second partial derivative which the proof is not given here (but is intuitive to understand) is the symmetry of mixed derivatives

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = f_{xy} = f_{yx} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x}. \quad (1.56)$$

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

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial y} \right)_z dy + \left(\frac{\partial x}{\partial z} \right)_y dz \quad \text{and} \quad dy = \left(\frac{\partial y}{\partial x} \right)_z dx + \left(\frac{\partial y}{\partial z} \right)_x dz \\ \implies dx &= \left(\frac{\partial x}{\partial y} \right)_z \left(\frac{\partial y}{\partial x} \right)_z dx + \left(\left(\frac{\partial x}{\partial y} \right)_z \left(\frac{\partial y}{\partial z} \right)_x + \left(\frac{\partial x}{\partial z} \right)_y \right) dz. \end{aligned} \quad (1.57)$$

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

$$\left(\frac{\partial x}{\partial y} \right)_z = \left(\frac{\partial y}{\partial x} \right)_z^{-1} \quad \text{and} \quad \left(\frac{\partial y}{\partial z} \right)_x \left(\frac{\partial z}{\partial x} \right)_y + \left(\frac{\partial x}{\partial y} \right)_z = 0. \quad (1.58)$$

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.*,

$$\left(\frac{\partial f}{\partial x} \right)_y \left(\frac{\partial x}{\partial z} \right)_y \left(\frac{\partial z}{\partial f} \right)_y = 1 \implies \left(\frac{\partial f}{\partial x} \right)_y = \left(\frac{\partial f}{\partial z} \right)_y \left(\frac{\partial z}{\partial x} \right)_y. \quad (1.59)$$

Some call this the chain rule in multivariable calculus.

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

$$\left(\frac{\partial f}{\partial y} \right)_{x,z} \left(\frac{\partial y}{\partial x} \right)_{f,z} \left(\frac{\partial x}{\partial f} \right)_{y,z} = -1. \quad (1.60)$$

1.8 Change of Variables

1.8.1 Change of Gradients

In section 1.6, we have discussed the case for 3 arguments with 2 constraints. We, however, would mostly deal with the case for 4 arguments with 2 constraints, when we try to change the whole set of independent variables from (x, y) to (u, v) . The two constraints in this

case can be expressed as $x = x(u, v)$ and $y = y(u, v)$. To find the partial derivatives of f with respect to u and v are, we first write down the total derivative of f , *i.e.*, df , with respect to all the arguments x, y, u and v

$$df = f_x dx + f_y dy + f_u du + f_v dv. \quad (1.61)$$

Now we divide the whole equation by du while keeping v constant

$$\left(\frac{\partial f}{\partial u}\right)_v = f_x \left(\frac{\partial x}{\partial u}\right)_v + f_y \left(\frac{\partial y}{\partial u}\right)_v, \quad (1.62)$$

where the term $(\partial f/\partial u)_v$ vanishes since $f(x, y, u, v) = x^2 + y^2$ has no explicit u -dependence. One might argue that $f(x, y, u, v) = u + v$ has u -dependence. However, we have to use the same form of f throughout. Since we have already assumed $f_x \neq 0$ and $f_y \neq 0$, we have implicitly assumed that we are using $f(x, y, u, v) = x^2 + y^2$ instead of $f(x, y, u, v) = u + v$. If instead we use the latter expression then we would get something like $(\partial f/\partial u)_v = f_u$, which is trivial.

Similarly we divide the whole equation by dv while keeping u constant

$$\left(\frac{\partial f}{\partial v}\right)_u = f_x \left(\frac{\partial x}{\partial v}\right)_u + f_y \left(\frac{\partial y}{\partial v}\right)_u, \quad (1.63)$$

where $(\partial f/\partial v)_u$ vanishes since $f(x, y, u, v) = x^2 + y^2$ has no explicit v -dependence.

To generalize the above discussion, suppose we want to change the set of n variables from $\mathbf{x} = (x_1, x_2, \dots, x_n)$ to $\mathbf{u} = (u_1, u_2, \dots, u_n)$, then the n equations are

$$\frac{\partial f}{\partial u_i} = f_{x_1} \frac{\partial x_1}{\partial u_i} + f_{x_2} \frac{\partial x_2}{\partial u_i} + \dots + f_{x_n} \frac{\partial x_n}{\partial u_i} = \sum_{j=1}^n f_{x_j} \frac{\partial x_j}{\partial u_i}. \quad (1.64)$$

In matrix form this is

$$\nabla_{\mathbf{u}} = \left(\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}, \dots, \frac{\partial}{\partial u_n}\right) = J^T \nabla_{\mathbf{x}} = J^T \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right), \quad (1.65)$$

where J is the Jacobian matrix

$$J \equiv \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{pmatrix} \partial x/\partial u & \partial x/\partial v & \partial x/\partial w \\ \partial y/\partial u & \partial y/\partial v & \partial y/\partial w \\ \partial z/\partial u & \partial z/\partial v & \partial z/\partial w \end{pmatrix} = \begin{pmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{pmatrix}, \quad (1.66)$$

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

In three dimensions, the 3 equations are

$$\begin{pmatrix} \partial/\partial u \\ \partial/\partial v \\ \partial/\partial w \end{pmatrix} = \begin{pmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{pmatrix} \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix}. \quad (1.67)$$

1.8.2 Change of Areas and Volumes

From $x = x(u, v)$ and $y = y(u, v)$ we have

$$\begin{cases} dx = x_u du + x_v dv \\ dy = y_u du + y_v dv. \end{cases} \quad (1.68)$$

In matrix form,

$$d\mathbf{r} = \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \quad (1.69)$$

On the x - y plane, the infinitesimal displacement vector which keeps v and u constant are, respectively

$$d\mathbf{r}_u = \begin{pmatrix} x_u \\ y_u \end{pmatrix} du \quad \text{and} \quad d\mathbf{r}_v = \begin{pmatrix} x_v \\ y_v \end{pmatrix} dv. \quad (1.70)$$

To relate the area $dA_{x,y} = dxdy = dA_{u,v}$,⁵ we have

$$dA_{u,v} = \det(d\mathbf{r}_u, d\mathbf{r}_v) = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} dudv = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv. \quad (1.71)$$

For volume we have

$$dV_{u,v,w} = \left| \frac{\partial(u, v, w)}{\partial(x, y, z)} \right| dudvdw. \quad (1.72)$$

When changing variables in multiple integral, the limits of integration has to change accordingly by finding the maximum and minimum values for u and v in the same region.

1.8.3 Multiple Jacobians

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

$$\frac{\partial x_i}{\partial z_j} = \sum_{k=1}^n \frac{\partial x_i}{\partial y_k} \frac{\partial y_k}{\partial z_j}. \quad (1.73)$$

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

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \implies C = AB \implies \det(C) = \det(A) \det(B) \implies J_{xz} = J_{xy} J_{yz}. \quad (1.74)$$

⁵Note that the $dA_{u,v} = dA_{x,y}$ since both refers to an infinitesimal area in the x - y plane. Formally, we have $\det(d\mathbf{r}_u, d\mathbf{r}_v) = \det(\mathbb{I})dxdy = dxdy$ if we use the former expression for $d\mathbf{r}$ in eq. (1.70) instead of the latter.

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

Example: Jacobian Rule (1).

Question: Prove the first Jacobian rule

$$\left(\frac{\partial u}{\partial x}\right)_v \left(\frac{\partial v}{\partial y}\right)_x = \left|\frac{\partial(u, v)}{\partial(x, y)}\right|. \quad (1.75)$$

Solution: We start with the total derivative of u in terms of x and y

$$du = \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_x dy. \quad (1.76)$$

Dividing the whole equation by dx while keeping v constant, we have

$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_v. \quad (1.77)$$

We then write down the total derivative of v in terms of x and y

$$dv = \left(\frac{\partial v}{\partial x}\right)_y dx + \left(\frac{\partial v}{\partial y}\right)_x dy. \quad (1.78)$$

Dividing the whole equation by dx while keeping v constant, we have

$$\left(\frac{\partial v}{\partial x}\right)_v = 0 = \left(\frac{\partial v}{\partial x}\right)_y + \left(\frac{\partial v}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_v. \quad (1.79)$$

Solving for $(\partial y/\partial x)_v$ gives

$$\left(\frac{\partial y}{\partial x}\right)_v = -\left(\frac{\partial v}{\partial x}\right)_y \left(\frac{\partial v}{\partial y}\right)_x^{-1}. \quad (1.80)$$

Substituting back into $(\partial u/\partial x)_v$, we have

$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial v}{\partial x}\right)_y \left(\frac{\partial v}{\partial y}\right)_x^{-1}. \quad (1.81)$$

Rearranging we get

$$\left(\frac{\partial u}{\partial x}\right)_v \left(\frac{\partial v}{\partial y}\right)_x = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial v}{\partial y}\right)_x - \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial v}{\partial x}\right)_y, \quad (1.82)$$

as desired.

Example: Jacobian Rule (2).

Question: Prove the second Jacobian rule

$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \left(\frac{\partial u}{\partial v}\right)_x \left(\frac{\partial v}{\partial x}\right)_y. \quad (1.83)$$

Solution: We start an intermediate step (eq. (1.81)) from the above example

$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial v}{\partial x}\right)_y \left(\frac{\partial v}{\partial y}\right)_x^{-1}. \quad (1.84)$$

To relate $(\partial u/\partial y)_x$ to $(\partial u/\partial v)_x$, we have the chain rule

$$\left(\frac{\partial u}{\partial v}\right)_x = \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_x \implies \left(\frac{\partial u}{\partial y}\right)_x = \left(\frac{\partial u}{\partial v}\right)_x \left(\frac{\partial v}{\partial y}\right)_x. \quad (1.85)$$

Substitute this into the intermediate step, we have

$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \left(\left(\frac{\partial u}{\partial v}\right)_x \left(\frac{\partial v}{\partial y}\right)_x\right) \left(\frac{\partial v}{\partial x}\right)_y \left(\frac{\partial v}{\partial y}\right)_x^{-1} = \left(\frac{\partial u}{\partial x}\right)_y - \left(\frac{\partial u}{\partial v}\right)_x \left(\frac{\partial v}{\partial x}\right)_y, \quad (1.86)$$

as desired.

Example: Jacobian Rule (3).

Question: Prove the third Jacobian rule

$$\left(\frac{\partial u}{\partial v}\right)_w = \left| \frac{\partial(u, w)}{\partial(x, y)} \right| \left| \frac{\partial(v, w)}{\partial(x, y)} \right|^{-1}. \quad (1.87)$$

Solution: When w is held constant, we have

$$dw = w_x dx + w_y dy = 0 \implies \frac{dy}{dx} = -\frac{w_x}{w_y}. \quad (1.88)$$

Therefore we have

$$\left(\frac{\partial u}{\partial x}\right)_w = u_x + u_y \frac{dy}{dx} = u_x - u_y \frac{w_x}{w_y}, \quad (1.89)$$

and similarly

$$\left(\frac{\partial v}{\partial x}\right)_w = v_x + v_y \frac{dy}{dx} = v_x - v_y \frac{w_x}{w_y}. \quad (1.90)$$

From chain rule we therefore get

$$\left(\frac{\partial u}{\partial v}\right)_w = \left(\frac{\partial u}{\partial x}\right)_w \left(\frac{\partial v}{\partial x}\right)_w^{-1} = \frac{u_x w_y - u_y w_x}{v_x w_y - v_y w_x} = \left| \frac{\partial(u, w)}{\partial(x, y)} \right| \left| \frac{\partial(v, w)}{\partial(x, y)} \right|^{-1}, \quad (1.91)$$

as desired.

Example: Thermodynamics.

Question: In thermodynamics, some of the variables associated with a confined gas are p (pressure), V (volume), T (temperature), U (internal energy), S (entropy) and H (enthalpy). Any two are independent and their values then determine all the others.

Suppose there is a thermodynamics law

$$\left(\frac{\partial U}{\partial p}\right)_T + T \left(\frac{\partial V}{\partial T}\right)_p + p \left(\frac{\partial V}{\partial p}\right)_T = 0, \quad (1.92)$$

where p and T are the independent variables in this equation since they appear at the denominators.

What is the equation for this law when V and T are the independent variables?

Solution: Using chain rule, cyclic relation and reciprocity relation on the three terms respectively we have

$$\left(\frac{\partial U}{\partial V}\right)_T \left(\frac{\partial V}{\partial p}\right)_T - T \left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial V}{\partial p}\right)_T + p \left(\frac{\partial V}{\partial p}\right)_T = 0 = \left(\frac{\partial U}{\partial V}\right)_T - T \left(\frac{\partial p}{\partial T}\right)_V + p. \quad (1.93)$$

1.9 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. (1.16), instead, we get the Taylor series

$$\begin{aligned} f(x, y) = f(x_0, y_0) &+ \left.\frac{\partial f}{\partial x}\right|_{(x_0, y_0)} \Delta x + \left.\frac{\partial f}{\partial y}\right|_{(x_0, y_0)} \Delta y \\ &+ \frac{1}{2!} \left(\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0, y_0)} (\Delta x)^2 + 2 \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0, y_0)} (\Delta x)(\Delta y) + \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0, y_0)} (\Delta y)^2 \right) \\ &+ \mathcal{O}((\Delta x)^3) + \mathcal{O}((\Delta y)^3). \end{aligned} \quad (1.94)$$

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

$$f(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{1}{n!} [(\Delta \mathbf{x} \cdot \nabla)^n f(\mathbf{x})] \Big|_{\mathbf{x}=\mathbf{x}_0}. \quad (1.95)$$

1.10 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

$$\left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)} = \left. \frac{\partial f}{\partial y} \right|_{(x_0, y_0)} = 0. \quad (1.96)$$

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

$$df = f(x, y) - f(x_0, y_0) \approx \frac{1}{2} \left[f_{xx} \left(\Delta x + \frac{f_{xy} \Delta y}{f_{xx}} \right)^2 + (\Delta y)^2 \left(f_{yy} - \frac{f_{xy}^2}{f_{xx}} \right) \right]. \quad (1.97)$$

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 positive, 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: $f_{yy} - f_{xy}^2/f_{xx} > 0$ if $f_{xx} < 0$ and $f_{yy} - f_{xy}^2/f_{xx} < 0$ if $f_{xx} > 0 \implies f_{xy}^2 > f_{xx}f_{yy}$.

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

$$\frac{\partial f}{\partial x_i} = 0 \text{ for all } x_i, \quad (1.98)$$

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

$$df = f(\mathbf{x}) - f(\mathbf{x}_0) \approx \frac{1}{2} \sum_i \sum_j \frac{\partial^2 f}{\partial x_i \partial x_j} \Delta x_i \Delta x_j, \quad (1.99)$$

which must be positive for all Δx_i .

Example: Contour Plot (1).

Question: Sketch the contour plot for the function

$$f(x, y) = xy e^{(x-y)}, \quad (1.100)$$

with the given maximum and saddle points $(-1, 1)$ and $(0, 0)$, respectively.

Solution: The contour plot is shown in fig. 1.2

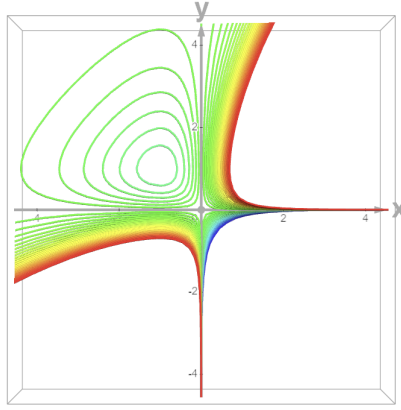


Figure 1.2

Example: Contour Plot (2).

Question: For the function

$$f(x, y) = 4xy - x^4 - y^4, \quad (1.101)$$

Determine the stationary points and its nature. Find the contour line that passes through the saddle point. Sketch the contour plot.

Solution: First we calculate all the useful derivatives

$$f_x = 4y - 4x^3, \quad f_y = 4x - 4y^3, \quad f_{xx} = -12x^2, \quad f_{yy} = -12y^2 \quad \text{and} \quad f_{xy} = 4. \quad (1.102)$$

For the stationary points we require that

$$f_x = f_y = 0 \implies (x, y) = (0, 0), (1, 1) \text{ or } (-1, 1). \quad (1.103)$$

Evaluating f_{xx} , f_{yy} and $D = f_{xy}^2 - f_{xx}f_{yy}$ for each of the stationary points we find that $(0, 0)$ is a saddle point since $D > 0$ while $(1, 1)$ and $(-1, -1)$ are maximum points since $D < 0$ and $f_{xx} < 0$ and $f_{yy} < 0$.

Since $f(0, 0) = 0$, the equation of the contour line that passes through the saddle point $(0, 0)$ is

$$f(x, y) = 4xy - x^4 - y^4 = 0 \implies r(\theta) = \frac{4 \cos \theta \sin \theta}{\cos^4 \theta + \sin^4 \theta}, \quad x = r \cos \theta \text{ and } y = r \sin \theta. \quad (1.104)$$

The contour plot is shown in fig. 1.3. The red thick line represents the contour line that passes through $(0, 0)$. Note that the contour lines do not extend off to infinity but form closed loops instead.

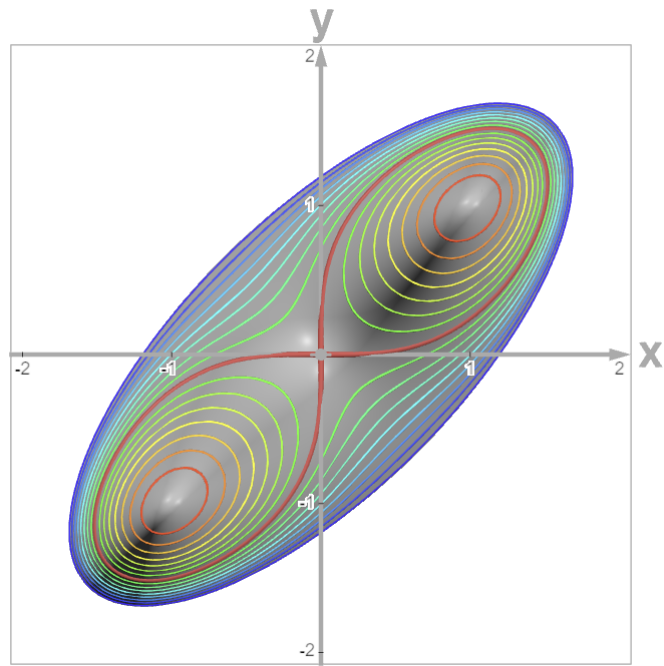


Figure 1.3

2.1 Curvilinear coordinates

2.1.1 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 cylindrical coordinates, $(u, v, w) = (\rho, \phi, z)$.

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

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u} du + \frac{\partial \mathbf{r}}{\partial v} dv + \frac{\partial \mathbf{r}}{\partial w} dw. \quad (2.1)$$

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 increasing u, v and w respectively, then we have

$$\frac{\partial \mathbf{r}}{\partial u} = f\hat{\mathbf{u}}, \frac{\partial \mathbf{r}}{\partial v} = g\hat{\mathbf{v}} \text{ and } \frac{\partial \mathbf{r}}{\partial w} = h\hat{\mathbf{w}}, \quad (2.2)$$

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 cylindrical coordinates, $(f, g, h) = (1, \rho, 1)$.

The infinitesimal displacement vector is now

$$d\mathbf{r} = f(du\hat{\mathbf{u}}) + g(dv\hat{\mathbf{v}}) + h(dw\hat{\mathbf{w}}) \quad (2.3)$$

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

$$ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}} = \sqrt{(fdu)^2 + (gdv)^2 + (hdw)^2}. \quad (2.4)$$

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

$$dS = (fg)dudv \quad (2.5)$$

as shown in fig. 2.1.

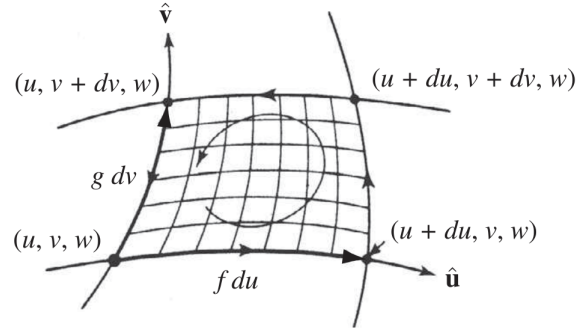


Figure 2.1

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

$$d\tau = |f d\hat{\mathbf{u}} \cdot (g d\hat{\mathbf{v}} \times h d\hat{\mathbf{w}})| dudvdw = (fgh)dudvdw \quad (2.6)$$

as shown in fig. 2.2.

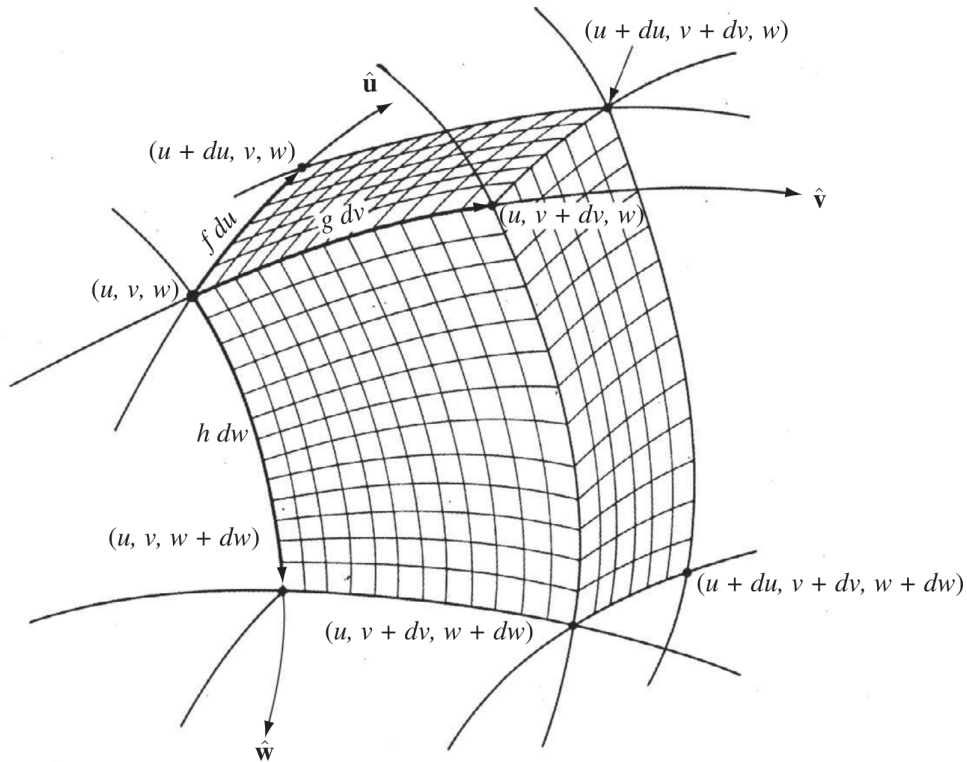


Figure 2.2

2.1.2 Spherical Coordinates

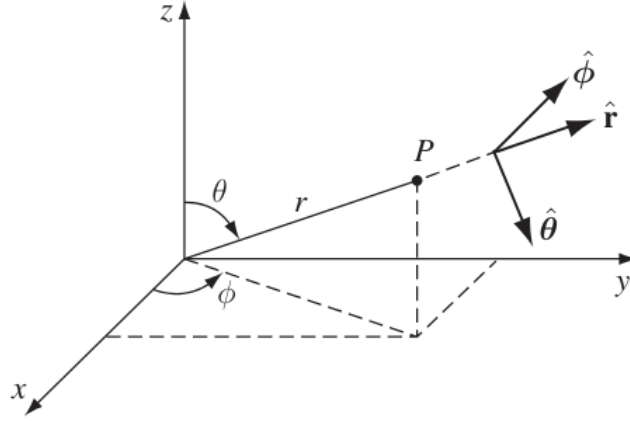


Figure 2.3

From fig. 2.3, the relations of the two set of variables (x, y, z) and (r, θ, ϕ) are

$$\begin{cases} x = r \sin \theta \cos \phi \\ y = r \sin \theta \sin \phi \\ z = r \cos \theta \end{cases} \quad \text{or} \quad \begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arctan \left(\sqrt{x^2 + y^2} / z \right) \\ \phi = \arctan (y/x) \end{cases} . \quad (2.7)$$

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

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = r \sin \theta \cos \phi \hat{\mathbf{i}} + r \sin \theta \sin \phi \hat{\mathbf{j}} + r \cos \theta \hat{\mathbf{k}} = r\hat{\mathbf{r}} = \sqrt{x^2 + y^2 + z^2}\hat{\mathbf{r}}, \quad (2.8)$$

where the last one is seldom used.

A general infinitesimal displacement vector can be written as

$$\begin{aligned} d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial x} dx + \frac{\partial \mathbf{r}}{\partial y} dy + \frac{\partial \mathbf{r}}{\partial z} dz = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}} \\ &= \frac{\partial \mathbf{r}}{\partial r} dr + \frac{\partial \mathbf{r}}{\partial \theta} d\theta + \frac{\partial \mathbf{r}}{\partial \phi} d\phi = f dr\hat{\mathbf{r}} + g d\theta\hat{\boldsymbol{\theta}} + h d\phi\hat{\boldsymbol{\phi}}. \end{aligned} \quad (2.9)$$

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

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial r} &= f\hat{\mathbf{r}} = \sin \theta \cos \phi \hat{\mathbf{i}} + \sin \theta \sin \phi \hat{\mathbf{j}} + \cos \theta \hat{\mathbf{k}}, \\ \frac{\partial \mathbf{r}}{\partial \theta} &= g\hat{\boldsymbol{\theta}} = r(\cos \theta \cos \phi \hat{\mathbf{i}} + \cos \theta \sin \phi \hat{\mathbf{j}} - \sin \theta \hat{\mathbf{k}}) \quad \text{and} \\ \frac{\partial \mathbf{r}}{\partial \phi} &= h\hat{\boldsymbol{\phi}} = r \sin \theta (-\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}). \end{aligned} \quad (2.10)$$

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

$$\begin{cases} \hat{\mathbf{r}} = \sin \theta \cos \phi \hat{\mathbf{i}} + \sin \theta \sin \phi \hat{\mathbf{j}} + \cos \theta \hat{\mathbf{k}}, \\ \hat{\boldsymbol{\theta}} = \cos \theta \cos \phi \hat{\mathbf{i}} + \cos \theta \sin \phi \hat{\mathbf{j}} - \sin \theta \hat{\mathbf{k}}, \\ \hat{\boldsymbol{\phi}} = -\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}. \end{cases} \quad (2.11)$$

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

$$\begin{cases} \hat{\mathbf{i}} = \sin \theta \cos \phi \hat{\mathbf{r}} + \cos \theta \cos \phi \hat{\boldsymbol{\theta}} - \sin \phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{j}} = \sin \theta \sin \phi \hat{\mathbf{r}} + \cos \theta \sin \phi \hat{\boldsymbol{\theta}} + \cos \phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{k}} = \cos \theta \hat{\mathbf{r}} - \sin \theta \hat{\boldsymbol{\theta}}, \end{cases} \quad (2.12)$$

which are seldomly used.

2.1.3 Cylindrical Coordinates

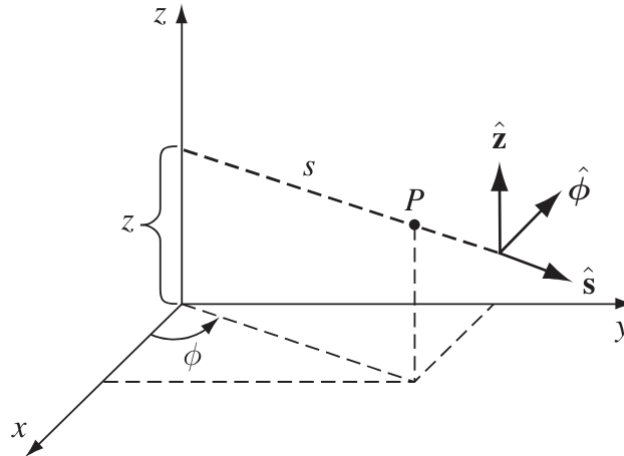


Figure 2.4

From fig. 2.4, the relations of the two set of variables (x, y, z) and (ρ, ϕ, z) are

$$\begin{cases} x = \rho \cos \phi \\ y = \rho \sin \phi \\ z = z \end{cases} \quad \text{or} \quad \begin{cases} \rho = \sqrt{x^2 + y^2} \\ \phi = \arctan(y/x) \\ z = z \end{cases} \quad (2.13)$$

A general position vector can be written in 4 ways as

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = \rho \cos \phi \hat{\mathbf{i}} + \rho \sin \phi \hat{\mathbf{j}} + \hat{\mathbf{k}} = \rho \hat{\boldsymbol{\rho}} + z\hat{\mathbf{z}} = \sqrt{x^2 + y^2} \hat{\boldsymbol{\rho}} + z\hat{\mathbf{z}}, \quad (2.14)$$

where the last one is seldom used.

A general infinitesimal displacement vector can be written as

$$\begin{aligned}
d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial x} dx + \frac{\partial \mathbf{r}}{\partial y} dy + \frac{\partial \mathbf{r}}{\partial z} dz = dx \hat{\mathbf{i}} + dy \hat{\mathbf{j}} + dz \hat{\mathbf{k}} \\
&= \frac{\partial \mathbf{r}}{\partial \rho} d\rho + \frac{\partial \mathbf{r}}{\partial \phi} d\phi + \frac{\partial \mathbf{r}}{\partial z} dz = f d\rho \hat{\boldsymbol{\rho}} + g d\phi \hat{\boldsymbol{\phi}} + h dz \hat{\mathbf{k}}.
\end{aligned} \tag{2.15}$$

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

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial \rho} &= f \hat{\boldsymbol{\rho}} = \cos \phi \hat{\mathbf{i}} + \sin \phi \hat{\mathbf{j}}, \\
\frac{\partial \mathbf{r}}{\partial \phi} &= g \hat{\boldsymbol{\phi}} = \rho(-\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}) \text{ and} \\
\frac{\partial \mathbf{r}}{\partial z} &= h \hat{\mathbf{k}} = \hat{\mathbf{k}}.
\end{aligned} \tag{2.16}$$

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

$$\begin{cases} \hat{\boldsymbol{\rho}} = \cos \phi \hat{\mathbf{i}} + \sin \phi \hat{\mathbf{j}}, \\ \hat{\boldsymbol{\phi}} = -\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}, \\ \hat{\mathbf{k}} = \hat{\mathbf{k}}. \end{cases} \tag{2.17}$$

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

$$\begin{cases} \hat{\mathbf{i}} &= \cos \phi \hat{\boldsymbol{\rho}} - \sin \phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{j}} &= \sin \phi \hat{\boldsymbol{\rho}} + \cos \phi \hat{\boldsymbol{\phi}}, \\ \hat{\mathbf{k}} &= \hat{\mathbf{k}}, \end{cases} \tag{2.18}$$

which are seldomly used.

2.2 Gradient, Divergence and Curl

We start by stating the general form of del operator, gradient, divergence, curl, and Laplacian are defined as¹

¹The 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.

$$\begin{aligned}
\nabla &= \frac{1}{f} \hat{\mathbf{u}} \frac{\partial}{\partial u} + \frac{1}{g} \hat{\mathbf{v}} \frac{\partial}{\partial v} + \frac{1}{h} \hat{\mathbf{w}} \frac{\partial}{\partial w}, \\
\nabla t &= \frac{1}{f} \hat{\mathbf{u}} \frac{\partial t}{\partial u} + \frac{1}{g} \hat{\mathbf{v}} \frac{\partial t}{\partial v} + \frac{1}{h} \hat{\mathbf{w}} \frac{\partial t}{\partial w}, \\
\nabla \cdot \mathbf{T} &= \frac{1}{fgh} \left[\frac{\partial}{\partial u} (ghT_u) + \frac{\partial}{\partial v} (fhT_v) + \frac{\partial}{\partial w} (fgT_h) \right], \\
\nabla \times \mathbf{T} &= \frac{1}{fgh} \begin{vmatrix} f\hat{\mathbf{u}} & g\hat{\mathbf{v}} & h\hat{\mathbf{w}} \\ \partial/\partial u & \partial/\partial v & \partial/\partial w \\ fT_u & gT_v & hT_w \end{vmatrix}, \\
\nabla^2 t &= \nabla \cdot (\nabla t) = (1/fgh) \left[\frac{\partial}{\partial u} \left(\frac{gh}{f} \frac{\partial t}{\partial u} \right) + \frac{\partial}{\partial v} \left(\frac{fh}{g} \frac{\partial t}{\partial v} \right) + \frac{\partial}{\partial w} \left(\frac{fg}{h} \frac{\partial t}{\partial w} \right) \right],
\end{aligned} \tag{2.19}$$

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.

2.2.1 Gradient

Using the del operator defined in eq. (2.19), we can write the infinitesimal change of a scalar function $t(u, v, w)$ as

$$dt = \frac{\partial t}{\partial u} du + \frac{\partial t}{\partial v} dv + \frac{\partial t}{\partial w} dw = \nabla t \cdot d\mathbf{r} = |\nabla t| |d\mathbf{r}| \cos \theta. \tag{2.20}$$

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

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

Integrating, we get the fundamental theorem for gradients

$$t(\mathbf{b}) - t(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{b}} dt = \int_{\mathbf{a}}^{\mathbf{b}} \nabla t \cdot d\mathbf{r}. \tag{2.21}$$

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 ∇t , then the line integral

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = \int_{\mathbf{a}}^{\mathbf{b}} \nabla t \cdot d\mathbf{r} = \int_{\mathbf{a}}^{\mathbf{b}} dt = t(\mathbf{b}) - t(\mathbf{a}) \tag{2.22}$$

problem
1.14,
1.17,
1.1.5

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 $t(\mathbf{b}) - t(\mathbf{a}) = 0$.²

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 $\nabla \times \mathbf{F}(\mathbf{r}) = \nabla \times (\nabla t(\mathbf{r})) = 0$.

Note that if $\mathbf{T}(\mathbf{r})$ is conservative then its total differential $\mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = \nabla t(\mathbf{r}) \cdot d\mathbf{r} = dt$ is exact.

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

$$\mathbf{T}(\mathbf{r}) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) \quad (2.23)$$

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

$$t(\mathbf{r}) = \tan^{-1} \left(\frac{y}{x} \right). \quad (2.24)$$

However, the loop integral

$$\oint_C \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = 2\pi \neq 0 \quad (2.25)$$

when C is a circle centered at the origin. More generally the loop integral is $2\pi n$, where n is the number of times the loop winds around the origin.

The inconsistency arises from the fact that $t(\mathbf{r})$ is not continuous at the origin, or in other words, $\mathbf{T}(\mathbf{r})$ is not continuously differentiable and fails to be defined at the origin, so $\mathbf{T}(\mathbf{r})$ is only conservative on the plane excluding the 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.

In complex analysis point of view,

$$\oint_C \mathbf{T}(\mathbf{r}) \cdot d\mathbf{r} = \oint_C d\theta, \quad (2.26)$$

so if the contour includes the origin then the integral is 2π and if it does not then it is just zero.

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

$$\nabla t = \lim_{V \rightarrow 0} \left(\frac{1}{V} \oint_S t d\mathbf{S} \right). \quad (2.27)$$

However the proof is complicated and is not presented here.

²The 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.

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

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

Solution:

$$\begin{aligned}
 \nabla r^{-1} &= \hat{\mathbf{i}} \frac{\partial}{\partial x} ((x - x')^2 + (y - y')^2 + (z - z')^2)^{-\frac{1}{2}} \\
 &\quad + \hat{\mathbf{j}} \frac{\partial}{\partial y} ((x - x')^2 + (y - y')^2 + (z - z')^2)^{-\frac{1}{2}} \\
 &\quad + \hat{\mathbf{k}} \frac{\partial}{\partial z} ((x - x')^2 + (y - y')^2 + (z - z')^2)^{-\frac{1}{2}} \\
 &= (-\frac{1}{2})(2)((x - x')^2 + (y - y')^2 + (z - z')^2)^{-\frac{3}{2}} \\
 &\quad ((x - x')\hat{\mathbf{i}} + (y - y')\hat{\mathbf{j}} + (z - z')\hat{\mathbf{k}}) \\
 &= -\mathbf{r} r^{-1}.
 \end{aligned} \tag{2.28}$$

This result can be easily generalised to get

$$\nabla(r^n) = n r^{n-1} \hat{\mathbf{r}}. \tag{2.29}$$

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 $\nabla f(x, y) = (2x, 2y)$. 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. (1.67), as

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix}. \tag{2.30}$$

Inverting the equations, we have

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \\ -\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \end{pmatrix}. \tag{2.31}$$

Therefore, we have

$$\begin{aligned}
 \nabla f(x, y) &= \left(\cos \theta \frac{\partial f}{\partial x} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) \hat{\mathbf{i}} + \left(\sin \theta \frac{\partial f}{\partial x} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \right) \hat{\mathbf{j}} = 2r(\cos \theta, \sin \theta) \\
 &= \frac{\partial f}{\partial r} \hat{\mathbf{r}} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\boldsymbol{\theta}} = 2r \hat{\mathbf{r}} = 2r(1, 0).
 \end{aligned} \tag{2.32}$$

Note that the answer $2r(\cos \theta, \sin \theta)$ is still in Cartesian coordinates, 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 coordinates.

Example: Gradient to Compute Changes.

Question: Consider the function T in the x - y plane, given by $T = \ln(x^2 + y^2)$. At point $(1, 2)$,

1. in which direction is the increase in T most rapid?
2. what distance in the direction found in (a) gives an increase of 0.2?
3. what distance in the direction of $\hat{\mathbf{i}} + \hat{\mathbf{j}}$ gives an increase of 0.12?
4. in which directions will T be stationary?

Why is ∇T perpendicular to the curve $T = 0$?

Solution:

1. The direction is given by its gradient $\nabla T = (2/5, 4/5)$.
2. The change in T is given by $dT = \nabla T \cdot d\mathbf{r} = |\nabla T|s = 0.2$, so $s = \sqrt{5}/10$.
3. The change in T is given by $dT = \nabla T \cdot d\mathbf{r} = \nabla T \cdot (1, 1)s/\sqrt{2} = 0.12$, so $s = 4\sqrt{5}/10$.
4. We require $dT = 0$, so $d\mathbf{r} \perp \nabla T$, so the direction is $\pm(2, -1)$.

On the curve $T = 0$, we have $dT = 0$. But in general we also have $dT = \nabla T \cdot d\mathbf{r}$, so ∇T must be perpendicular to $d\mathbf{r}$.

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

$$\nabla t = \frac{\partial t}{\partial x}\hat{\mathbf{i}} + \frac{\partial t}{\partial y}\hat{\mathbf{j}} + \frac{\partial t}{\partial z}\hat{\mathbf{k}} = \frac{\partial t}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial t}{\partial \theta}\hat{\boldsymbol{\theta}} + \frac{1}{r \sin \theta}\frac{\partial t}{\partial \phi}\hat{\boldsymbol{\phi}}. \quad (2.33)$$

2.2.2 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. 2.2 (here we adopt the sign convention of \mathbf{S} is positive if \mathbf{S} is pointing outwards from the interior of the volume)

$$\begin{aligned} \oint_S \mathbf{T} \cdot d\mathbf{S} &= ((T_u)|_{u+du} - (T_u)|_u)(ghdv dw) \\ &\quad + ((T_v)|_{v+dv} - (T_v)|_v)(fhdu dw) \\ &\quad + ((T_w)|_{w+dw} - (T_w)|_w)(fgdu dv) \\ &= \frac{1}{fgh} \left(\frac{\partial}{\partial u}(ghT_u) + \frac{\partial}{\partial v}(fhT_v) + \frac{\partial}{\partial w}(fgT_w) \right) d\tau \\ &= (\nabla \cdot \mathbf{T})d\tau \end{aligned} \quad (2.34)$$

This result can be extended easily. As any arbitrary 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,

$$\oint_S \mathbf{T} \cdot d\mathbf{S} = \int_V (\nabla \cdot \mathbf{T}) d\tau. \quad (2.35)$$

This is the divergence theorem (also known as the Gauss's theorem or the Green's theorem). From the LHS of eq. (2.35), it is evident that the divergence of a vector function is a measure of how much the function spreads out and diverges from a given point in space.

If \mathbf{T} satisfies $\nabla \cdot \mathbf{T} = 0$, then $\mathbf{T} = \nabla \times \mathbf{T}' + \nabla t + \mathbf{C}$ can be written as a curl of a vector \mathbf{T}' , which satisfies $\nabla \times \mathbf{T}' = 0$ 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:

$$\nabla \cdot \mathbf{T} = \frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y} + \frac{\partial T_z}{\partial z} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 T_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta T_\theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} (r T_\phi). \quad (2.36)$$

2.2.3 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. 2.1 (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{S} points in the same direction as $\hat{\mathbf{w}}$)

$$\begin{aligned} \oint_C \mathbf{T} \cdot d\mathbf{r} &= (T_u)|_v f du + (T_v)|_{u+du} g dv + (T_u)|_{v+dv} (-f du) + (T_v)|_u (-g dv) \\ &= \frac{1}{fg} \left(\frac{\partial}{\partial u} (T_v g) - \frac{\partial}{\partial v} (T_u f) \right) (\hat{\mathbf{w}} \cdot d\mathbf{S}) \\ &= (\nabla \times \mathbf{T}) \cdot d\mathbf{S}. \end{aligned} \quad (2.37)$$

With the same argument as eq. (2.34) to eq. (2.35), the above equation can be extended to

$$\oint_C \mathbf{T} \cdot d\mathbf{r} = \int_S (\nabla \times \mathbf{T}) \cdot d\mathbf{S}. \quad (2.38)$$

This is the Stoke's theorem. From the LHS of eq. (2.38), 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 $\int_S (\nabla \times \mathbf{T}) \cdot d\mathbf{S}$ is independent of the surface used but only depends on the boundary line. Hence, $\oint_S (\nabla \times \mathbf{T}) \cdot d\mathbf{S} = 0$ 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 $\nabla \times \mathbf{T} = 0$ and $\mathbf{T} = \nabla t$ 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:

$$\begin{aligned}\nabla \times \mathbf{T} &= \left(\frac{\partial T_z}{\partial y} - \frac{\partial T_y}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial T_x}{\partial z} - \frac{\partial T_z}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y}\right)\hat{\mathbf{k}} \\ &= \frac{1}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (\sin \theta T_\phi) - \frac{\partial T_\theta}{\partial \phi} \right) \hat{\mathbf{r}} + \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial T_r}{\partial \phi} - \frac{\partial}{\partial r} (r T_\phi) \right) \hat{\boldsymbol{\theta}} + \frac{1}{r} \left(\frac{\partial}{\partial r} (r T_\theta) - \frac{\partial T_r}{\partial \theta} \right) \hat{\boldsymbol{\phi}}.\end{aligned}\tag{2.39}$$

Example: Curl.

Question: Find the vector function \mathbf{G} such that $\nabla \times \mathbf{G} = \mathbf{F}$, given that $\mathbf{F} = (xz^2, -yz^2, 0)$.

Solution: Since $F_z = 0$, a simple approach is to set $G_x = G_y = 0$, then the curl reduces to

$$\nabla \times \mathbf{G} = (\partial_y G_z, -\partial_x G_z, 0).\tag{2.40}$$

So we need

$$\partial_y G_z = xz^2 \quad \text{and} \quad \partial_x G_z = yz^2.\tag{2.41}$$

Integrating the first equation gives

$$G_z = xyz^2 + h(x, z).\tag{2.42}$$

Substituting into the second equation gives

$$\partial_x G_z = yz^2 + \frac{\partial h}{\partial x} = yz^2 \implies \frac{\partial h}{\partial x} = 0 \implies h = \text{constant}.\tag{2.43}$$

Therefore the $\mathbf{G} = (0, 0, xyz^2 + C)$.

Example: Conservative Fields.

Question: Prove that the vector field

$$\mathbf{F}(\mathbf{r}) = (\hat{\mathbf{i}} - \hat{\mathbf{j}} - 2(x - y)\mathbf{r})e^{-r^2}\tag{2.44}$$

is conservative. Then find its potential function $\phi(\mathbf{r})$.

Solution: We first write $\mathbf{F}(\mathbf{r})$ in its three Cartesian components

$$F_x = e^{-r^2}(1 - 2x(x - y)), \quad F_y = e^{-r^2}(-1 - 2y(x - y)) \quad \text{and} \quad F_z = e^{-r^2}(-2z(x - y)).\tag{2.45}$$

We then compute the components of $\nabla \times \mathbf{F}$

$$(\nabla \times \mathbf{F})_x = \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \quad (\nabla \times \mathbf{F})_y = \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \quad (\nabla \times \mathbf{F})_z = \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y},\tag{2.46}$$

and find that each of these three expression vanishes. For example we have $(\nabla \times \mathbf{F})_x = 0$ since

$$\begin{aligned}\frac{\partial F_z}{\partial y} &= \frac{\partial}{\partial y} \left[-2z(x-y)e^{-r^2} \right] = -2z \frac{\partial}{\partial y} \left[(x-y)e^{-r^2} \right] = 2ze^{-r^2} [1 + 2y(x-y)], \\ \frac{\partial F_y}{\partial z} &= \frac{\partial}{\partial z} \left[(-1 - 2y(x-y))e^{-r^2} \right] = (-1 - 2y(x-y))(-2z)e^{-r^2} = 2ze^{-r^2} [1 + 2y(x-y)],\end{aligned}\tag{2.47}$$

Therefore $\nabla \times \mathbf{F} = 0$ and thus \mathbf{F} is conservative.

Focusing on the z component of $\mathbf{F}(\mathbf{r})$ we can easily spot the potential function

$$\phi(\mathbf{r}) = (x-y)e^{-r^2}.\tag{2.48}$$

2.3 Product Rules

The two product rules for gradient are

$$\begin{cases} \nabla(fg) = f\nabla g + g\nabla f, \\ \nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A}. \end{cases}\tag{2.49}$$

The two product rules for divergence are

$$\begin{cases} \nabla \cdot (f\mathbf{A}) = f(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\nabla f), \\ \nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}). \end{cases}\tag{2.50}$$

The two product rules for curl are

$$\begin{cases} \nabla \times (f\mathbf{A}) = f(\nabla \times \mathbf{A}) - (\mathbf{A} \times (\nabla f)), \\ \nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A}(\nabla \cdot \mathbf{A}) - \mathbf{B}(\nabla \cdot \mathbf{A}). \end{cases}\tag{2.51}$$

Here note that we can consturct another operator $\mathbf{A} \cdot \nabla$ called the directional derivative, defined by

$$\mathbf{A} \cdot \nabla = A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y} + A_z \frac{\partial}{\partial z},\tag{2.52}$$

which can act on both a sclar function

$$(\mathbf{A} \cdot \nabla)f = A_x \frac{\partial f}{\partial x} + A_y \frac{\partial f}{\partial y} + A_z \frac{\partial f}{\partial z},\tag{2.53}$$

and a vector function

$$\begin{aligned}
(\mathbf{A} \cdot \nabla) \mathbf{B} &= \left(A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y} + A_z \frac{\partial}{\partial z} \right) (B_x \hat{\mathbf{i}} + B_y \hat{\mathbf{j}} + B_z \hat{\mathbf{k}}) \\
&= \left(A_x \frac{\partial B_x}{\partial x} + A_y \frac{\partial B_x}{\partial y} + A_z \frac{\partial B_x}{\partial z} \right) \hat{\mathbf{i}} \\
&\quad + \left(A_x \frac{\partial B_y}{\partial x} + A_y \frac{\partial B_y}{\partial y} + A_z \frac{\partial B_y}{\partial z} \right) \hat{\mathbf{j}} \\
&\quad + \left(A_x \frac{\partial B_z}{\partial x} + A_y \frac{\partial B_z}{\partial y} + A_z \frac{\partial B_z}{\partial z} \right) \hat{\mathbf{k}}.
\end{aligned} \tag{2.54}$$

In component form, we have

$$((\mathbf{A} \cdot \nabla) \mathbf{B})_i = A_j \partial_j B_i. \tag{2.55}$$

Another differential operator we can construct is $(\mathbf{A} \times \nabla)_k = \epsilon_{ijk} A_i \partial_j$ 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

$$\begin{aligned}
\int_V \nabla \cdot (f \mathbf{T}) d\tau &= \int_V f (\nabla \cdot \mathbf{T}) d\tau + \int_V \mathbf{T} \cdot (\nabla f) d\tau = \oint_S f \mathbf{T} \cdot d\mathbf{A} \\
\Rightarrow \int_V f (\nabla \cdot \mathbf{T}) d\tau &= \oint_S (f \mathbf{T}) \cdot d\mathbf{A} - \int_V \mathbf{T} \cdot (\nabla f) d\tau.
\end{aligned} \tag{2.56}$$

Here we transform the integrand from the product of one function (f) and one derivative $\nabla \cdot \mathbf{T}$ to another integrand of the product of one function that is originally the derivative (\mathbf{T}) and one derivative which is originally the function (∇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

$$\int_a^b f \left(\frac{dg}{dx} \right) dx = - \int_a^b g \left(\frac{df}{dx} \right) dx + (fg) \Big|_a^b \tag{2.57}$$

comes from the product rule

$$\frac{d}{dx}(fg) = f \left(\frac{dg}{dx} \right) + g \left(\frac{df}{dx} \right). \tag{2.58}$$

Similarly, we can show that

$$\int_S f (\nabla \times \mathbf{T}) \cdot d\mathbf{A} = \int_S (\mathbf{T} \times (\nabla f)) \cdot d\mathbf{A} + \oint_C f \mathbf{T} \cdot d\mathbf{r} \tag{2.59}$$

and

$$\int_V \mathbf{B} \cdot (\nabla \times \mathbf{T}) d\tau = \int_V \mathbf{T} \cdot (\nabla \times \mathbf{B}) d\tau + \oint_S (\mathbf{T} \times \mathbf{B}) \cdot d\mathbf{A}. \quad (2.60)$$

2.4 Second Derivatives

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

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

$$\nabla \cdot (\nabla t) = \left(\hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \right) \cdot \left(\hat{\mathbf{i}} \frac{\partial t}{\partial x} + \hat{\mathbf{j}} \frac{\partial t}{\partial y} + \hat{\mathbf{k}} \frac{\partial t}{\partial z} \right) = \frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2}, \quad (2.61)$$

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:

$$\nabla^2 t = \nabla \cdot (\nabla t). \quad (2.62)$$

With reference to eq. (2.62), the Laplacian of a scalar function in spherical coordinates is listed here for reference:

$$\nabla^2 t = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial t}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial t}{\partial \theta} \right) + \frac{1}{r^2 \sin \phi} \frac{\partial^2 t}{\partial \phi^2}. \quad (2.63)$$

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

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial t}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} (rt). \quad (2.64)$$

2. Curl of gradient:

$$\nabla \times (\nabla t) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} & \frac{\partial t}{\partial z} \end{vmatrix} = \left(\frac{\partial^2 t}{\partial z \partial y} - \frac{\partial^2 t}{\partial y \partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial^2 t}{\partial x \partial z} - \frac{\partial^2 t}{\partial z \partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial^2 t}{\partial y \partial x} - \frac{\partial^2 t}{\partial x \partial y} \right) \hat{\mathbf{k}} = 0. \quad (2.65)$$

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

$$\nabla (\nabla \cdot \mathbf{T}) \neq (\nabla \cdot \nabla) \mathbf{T}. \quad (2.66)$$

4. Divergence of curl:

$$\nabla \cdot (\nabla \times \mathbf{T}) = \frac{\partial}{\partial x} \left(\frac{\partial T_z}{\partial y} - \frac{\partial T_y}{\partial z} \right) + \frac{\partial}{\partial z} \left(\frac{\partial T_x}{\partial z} - \frac{\partial T_z}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y} \right) = 0. \quad (2.67)$$

5. Curl of curl:

$$\nabla \times (\nabla \times \mathbf{T}) = \nabla(\nabla \cdot \mathbf{T}) - \nabla^2 \mathbf{T}. \quad (2.68)$$

Here, $\nabla^2 \mathbf{T}$ is the Laplacian of a vector function defined as³

$$\nabla^2 \mathbf{T} = (\nabla^2 T_x) \hat{\mathbf{i}} + (\nabla^2 T_y) \hat{\mathbf{j}} + (\nabla^2 T_z) \hat{\mathbf{k}}. \quad (2.69)$$

For the 5 second derivatives listed above, only eqs. (2.61) and (2.68) are used regularly in physics enough that they are worth remembering.

2.5 Index Notation

All of the identities can be easily proven in index notation, as long as one remembers the basic rules

$$(\nabla t)_i = \partial_i t, \quad (\nabla \cdot \mathbf{T})_i = \partial_i T_i \quad \text{and} \quad (\nabla \times \mathbf{T})_i = \epsilon_{ijk} \partial_j T_k. \quad (2.70)$$

There are also some identities worth noting:

$$\partial_i r_i = 3 \quad \text{and} \quad \partial_j r_i = \delta_{ij}. \quad (2.71)$$

To find $\nabla(r^p)$, for example, we have

$$\partial_i r^p = \partial_i (x_j x_j)^{\frac{p}{2}} = \frac{p}{2} (x_j x_j)^{\frac{p}{2}-1} \partial_i (x_j x_j) = p r^{p-2} x_i \implies \nabla(r^p) = p r^{p-1} \hat{\mathbf{r}}. \quad (2.72)$$

Note that it is wrong to write $r = r_i r_i$ because i is already the free index.

Specifically, if $p = 1$ we have

$$\partial_i r = \frac{x_i}{r} \implies \nabla r = \hat{\mathbf{r}}. \quad (2.73)$$

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

$$(\nabla \times \nabla t)_i = \epsilon_{ijk} \partial_j \partial_k t = 0, \quad (2.74)$$

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.

Example: Index Notation (1).

Question: Simplify $\nabla \times (\varphi \mathbf{F})$.

³In fact, eq. (2.68) is often used to define the Laplacian of a vector, since eq. (2.62) makes explicit reference to Cartesian coordinates.

Solution: We have

$$(\nabla \times \varphi \mathbf{F})_i = \epsilon_{ijk} \partial_j (\varphi \mathbf{F})_k = \epsilon_{ijk} (F_k (\partial_j \varphi) + \varphi (\partial_j F_k)) = (\nabla \varphi \times \mathbf{F})_i + \varphi (\nabla \times \mathbf{F})_i. \quad (2.75)$$

Example: Index Notation (2).

Question: Simplify $\nabla \times (\boldsymbol{\omega} \times \mathbf{r})$, where $\boldsymbol{\omega} = \omega \hat{\mathbf{z}}$.

Solution: We have

$$\begin{aligned} (\nabla \times (\boldsymbol{\omega} \times \mathbf{r}))_i &= \epsilon_{ilm} \partial_l \epsilon_{mjk} \omega_j r_k = \epsilon_{ilm} \epsilon_{mjk} \omega_j \partial_l r_k \\ &= \epsilon_{ilm} \epsilon_{mjk} \omega_j \delta_{lk} = (\delta_{ij} \delta_{kk} - \delta_{ik} \delta_{jj}) \omega_j = (3\delta_{ij} - \delta_{ij}) \omega_j = 2\omega_i. \end{aligned} \quad (2.76)$$

Note that from the first equality to the second we have pulled ω_j out of ∂_l since it is a constant and does not depend on position.

Example: Index Notation (3).

Question: Evaluate $\nabla((\mathbf{M} \cdot \hat{\mathbf{r}})/r^2)$, where \mathbf{M} is a constant vector in space.

Solution: Using the identity $\nabla(fg) = f\nabla g + g\nabla f$, we rewrite the expression as

$$\nabla \left(\frac{\mathbf{M} \cdot \hat{\mathbf{r}}}{r^2} \right) = \nabla \left(\frac{\mathbf{M} \cdot \mathbf{r}}{r^3} \right) = (\mathbf{M} \cdot \mathbf{r}) \nabla \left(\frac{1}{r^3} \right) + \frac{1}{r^2} \nabla(\mathbf{M} \cdot \mathbf{r}) = -3 \frac{(\mathbf{M} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}}{r^3} + \frac{\mathbf{M}}{r^3}. \quad (2.77)$$

Alternatively, we can use the identity $\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A}$ to rewrite the expression as

$$\nabla \left(\frac{\mathbf{M} \cdot \hat{\mathbf{r}}}{r^2} \right) = (\mathbf{M} \cdot \nabla) \frac{\hat{\mathbf{r}}}{r^2} + \mathbf{M} \times \left(\nabla \times \left(\frac{\hat{\mathbf{r}}}{r^2} \right) \right). \quad (2.78)$$

The second term is zero since $\nabla \times (\hat{\mathbf{r}}/r^2) = 0$, and the first term is most easily evaluated using index notation

$$\nabla \left(\frac{\mathbf{M} \cdot \hat{\mathbf{r}}}{r^2} \right) = (\mathbf{M} \cdot \nabla) \frac{\hat{\mathbf{r}}}{r^2} = M_j \partial_j (r_i r^{-3}) = M_i r^{-3} + M_j r_i (-3r^{-4} \partial_j r) = \frac{M_i}{r^3} - 3 \frac{(M_j r_j) r_i}{r^5}, \quad (2.79)$$

which is what we have found previously.

2.6 Dirac Delta Function

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

$$\nabla \cdot \left(\frac{\hat{\mathbf{r}}}{r^2} \right) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \left(\frac{1}{r^2} \right) \right) = 0, \quad (2.80)$$

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

$$\oint_S \left(\frac{\hat{\mathbf{r}}}{R^2} \right) \cdot (R^2 \sin \theta d\theta d\phi \hat{\mathbf{r}}) = 4\pi. \quad (2.81)$$

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 origin.

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

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(x) dx = 1.^4 \quad (2.82)$$

The appearance of this function is shown in fig. 2.5.

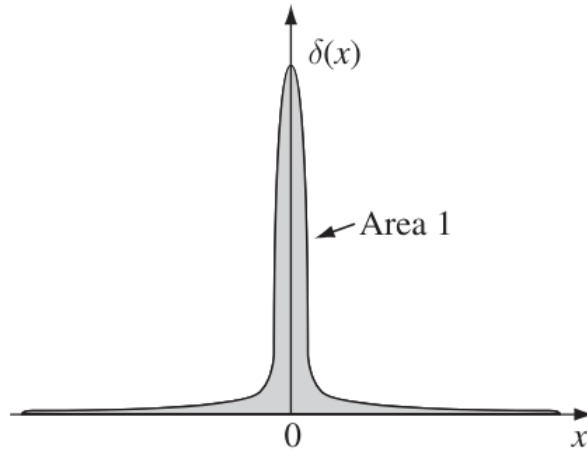


Figure 2.5

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

$$f(x)\delta(x) = f(0)\delta(x), \quad (2.83)$$

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

⁴Therefore, $\delta(x)$ has the dimension of the inverse of its argument

Integrating the above equation,

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = \int_{-\infty}^{+\infty} f(0)\delta(x)dx = f(0) \int_{-\infty}^{+\infty} \delta(x)dx = f(0). \quad (2.84)$$

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

By changing the variable in eq. (2.84), we can pick out the value of $f(x)$ at any arbitrary point $x = a$

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a) \quad (2.85)$$

Example: Dirac Delta Function (1).

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

Solution: Let $u = kx$, then

$$\int_{-\infty}^{+\infty} f(x)\delta(kx)dx = \frac{1}{|k|} \int_{-\infty}^{+\infty} f\left(\frac{u}{k}\right) \delta(u)du = \frac{1}{|k|} f(0) = \int_{-\infty}^{+\infty} f(x) \frac{\delta(x)}{|k|} dx. \quad (2.86)$$

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 $-\delta(x) = x \frac{d}{dx}(\delta(x))$

Solution:

Example: Dirac Delta Function (3).

Question: Show that $\delta(x) = d\theta/dx$, 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:

$$\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z) \quad (2.87)$$

The characteristic equations the three-dimensional delta function are then

$$\int_{\text{all space}} \delta^3(\mathbf{r}) d\tau = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x)\delta(y)\delta(z) dx dy dz = 1 \quad (2.88)$$

and

$$\int_{\text{all space}} f(\mathbf{r}) \delta^3(\mathbf{r} - \mathbf{a}) = f(\mathbf{a}). \quad (2.89)$$

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 $\mathbf{T} = \hat{\mathbf{r}}/r^2$ is in fact

$$\nabla \cdot \left(\frac{\hat{\mathbf{r}}}{r^2} \right) = 4\pi \delta^3(\mathbf{r}). \quad (2.90)$$

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

In general,

$$\nabla \cdot \left(\frac{\hat{\mathbf{z}}}{z^2} \right) = 4\pi \delta^3(\mathbf{z}). \quad (2.91)$$

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

From eq. (2.29), we have also

$$\nabla^2 \left(\frac{1}{z} \right) = -4\pi \delta^3(\mathbf{z}). \quad (2.92)$$

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

$$\begin{aligned} \nabla^2 V(\mathbf{r}) &= \nabla^2 \left(\frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{z} d\tau' \right) = \frac{1}{4\pi\epsilon_0} \int \rho(\mathbf{r}') \nabla^2 \left(\frac{1}{z} \right) d\tau' \\ &= \frac{1}{4\pi\epsilon_0} \int \rho(\mathbf{r}') (-4\pi \delta^3(\mathbf{z})) d\tau' = -\frac{\rho(\mathbf{r})}{\epsilon_0}. \end{aligned} \quad (2.93)$$

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 $A = q/4\pi a$. Therefore, $\rho(r) = q\delta(r - a)/4\pi a$.

Example: Vortex Flow.

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

$$\mathbf{v} = \frac{1}{\rho} \hat{\phi}. \quad (2.94)$$

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

Solution: From this velocity field $\nabla \times \mathbf{v} = 0$ 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

$$\nabla \times \mathbf{v} = 2\pi\delta(\rho). \quad (2.95)$$

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

2.7 Helmholtz Theorem

Theorem 2.7.1 (Helmholtz 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 $r \rightarrow \infty$, where Φ and \mathbf{A} are given by*

$$\Phi(\mathbf{r}) = \frac{1}{4\pi} \int_{all\ space} \frac{\nabla' \cdot \mathbf{T}(\mathbf{r}')}{r} d\tau' + C \quad and \quad \mathbf{A}(\mathbf{r}) = \frac{1}{4\pi} \int_{all\ space} \frac{\nabla' \times \mathbf{T}(\mathbf{r}')}{r} d\tau' + \mathbf{C} + \nabla D. \quad (2.96)$$

Here C and \mathbf{C} are some constant scalar and vector functions respectively, which does not change $\mathbf{T} = \nabla\Phi + \nabla \times \mathbf{A}$, since the gradient and curl of some constant function is zero. 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 equation⁵ and states that

⁵The proof of this theorem is mathematically advanced thus is omitted here but somewhat trivial

there exists an unique solution to the equation

$$-\nabla^2\Phi = -\nabla \cdot (\nabla\Phi) = \nabla \cdot \mathbf{T} \implies \nabla \cdot (\mathbf{T} + \nabla\Phi) = 0. \quad (2.97)$$

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

$$\mathbf{T} = -\nabla\Phi + \nabla \times \mathbf{A} \quad (2.98)$$

Testing the divergence of \mathbf{T} , we have⁶

$$\begin{aligned} \nabla \cdot \mathbf{T} &= -\nabla \cdot (\nabla\Phi) + \nabla \cdot (\nabla \times \mathbf{A}) = -\nabla^2\Phi \\ &= -\frac{1}{4\pi} \int \nabla' \cdot \mathbf{T}(\mathbf{r}') \nabla^2 \left(\frac{1}{\mathcal{Z}} \right) d\tau' \\ &= \int \nabla' \cdot \mathbf{T}(\mathbf{r}') \delta^3(\mathbf{r} - \mathbf{r}') d\tau' = \nabla \cdot \mathbf{T}(\mathbf{r}). \end{aligned} \quad (2.99)$$

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

Testing the curl,

$$\nabla \times \mathbf{T} = -\nabla \times \nabla\Phi + \nabla \times (\nabla \times \mathbf{A}) = -\nabla^2\mathbf{A} + \nabla(\nabla \cdot \mathbf{A}). \quad (2.100)$$

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

$$\begin{aligned} 4\pi(\nabla \cdot \mathbf{A}) &= \int \nabla \cdot \left(\frac{\nabla' \times \mathbf{T}(\mathbf{r}')}{\mathcal{Z}} \right) d\tau' \\ &= \int (\nabla' \times \mathbf{T}(\mathbf{r}')) \cdot \nabla \left(\frac{1}{\mathcal{Z}} \right) d\tau' = - \int (\nabla' \times \mathbf{T}(\mathbf{r}')) \cdot \nabla' \left(\frac{1}{\mathcal{Z}} \right) d\tau' \\ &= \int \frac{1}{\mathcal{Z}} \nabla' \cdot (\nabla' \times \mathbf{T}(\mathbf{r}')) d\tau - \oint \frac{1}{\mathcal{Z}} (\nabla' \times \mathbf{T}(\mathbf{r}')) \cdot d\mathbf{A} \\ &= - \int \frac{1}{\mathcal{Z}} \nabla \cdot (\nabla \times \mathbf{T}(\mathbf{r}')) d\tau = 0. \end{aligned} \quad (2.101)$$

where we used $\nabla(1/\mathcal{Z}) = -\nabla'(1/\mathcal{Z})$ since the derivative of $\mathcal{Z} = |\mathbf{r} - \mathbf{r}'|$ with respect to the primed coordinates differ by a sign from that with respect to unprimed coordinates.⁷ Also, we assumed that $\nabla \times \mathbf{T}$ goes to zero faster than $1/r^2$ as $r \rightarrow \infty$ 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

because it simply implies that we can assign each point in space an electric potential for any arbitrary charge distribution.

⁶Here and after, we will omit the limit of the integrals.

⁷More precisely, $\partial f(x - x')/\partial x = -\partial f(x - x')/\partial x'$.

$$\int \frac{X(r')}{r'} r'^2 dr', \quad (2.102)$$

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

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

□

3.1 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 $f(x, y, z)dxdydz$.¹²³

Referring to fig. 3.1, we can see that a double integral can be evaluated in two different ways.

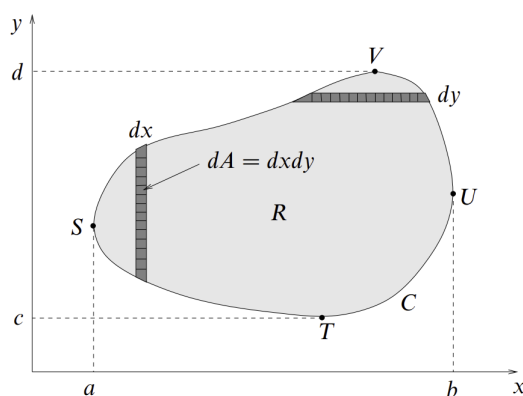


Figure 3.1

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

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

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

³Here the argument (x, y, z) can be unambiguously 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.

$$I = \int_c^d \left(\int_{x_1(y)}^{x_2(y)} f(x, y) dx \right) dy, \quad (3.1)$$

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

$$I = \int_a^b \int_{y_1(x)}^{y_2(x)} f(x, y) dy dx, \quad (3.2)$$

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 $(\int f(x)dx)(\int f(y)dy)$.⁴

The above analysis can be extended to triple integration easily.

When we do a change in variables the limit of integrations have to change accordingly as well. For example, if we are integrating over $x \in (1, 1/2)$ and $y \in (x, 1 - x)$, then when you do the change of variables $(r, s) = (x + y, y - x)$, you would be integrating over $r \in (0, 1)$ and $s \in (0, r)$. This is because the three vertices of the triangular region in the x - y plane $(0, 0)$, $(0, 1)$, $(1/2, 1/2)$ transform to $(0, 0)$, $(1, 1)$, $(0, 1)$ in the r - s plane, respectively. Normally this procedure of changing the limit of integration can be more easily done by moving the x and y around in the equalities defining the boundary of integration region in the x - y plane, especially in the case where the limit of integrations are constants.

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. 3.2.

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

$$V = \int_0^a \int_0^{b-\frac{bx}{a}} \int_0^{c(1-\frac{y}{b}-\frac{x}{a})} dx dy dz, \quad (3.3)$$

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

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

z as

$$V = \int_0^a \int_0^{b-\frac{bx}{a}} c \left(a - \frac{y}{b} - \frac{x}{a} \right) dx dy \quad (3.4)$$

The result is $V = abc/6$.

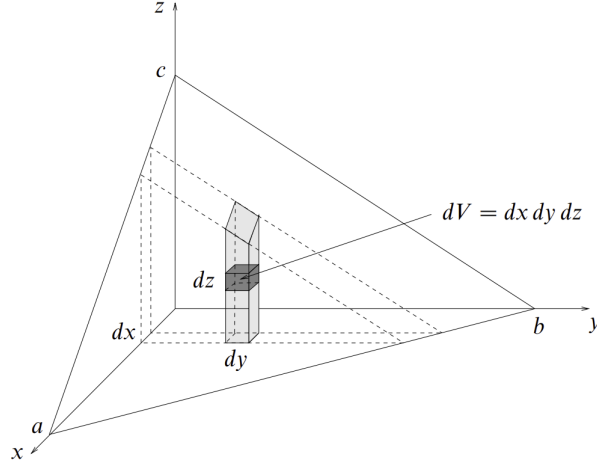


Figure 3.2

Example: Volume of a Cut Cylinder.

Question: A tube with an elliptical cross-section, defined by $x^2 + 4z^2 = 4$, is intersected by the planes $y + z = 1$ and $z = y + 1$. Find the volume of the enclosed region, by transforming to coordinates (r, θ, y) given by $x = 2r \cos \theta$ and $z = r \sin \theta$.

Solution: The Jacobian for this coordinate transformation is

$$J = \left| \frac{\partial(x, y, z)}{\partial(r, \theta, y)} \right| = \det \begin{pmatrix} 2 \cos \theta & -2r \sin \theta & 0 \\ 0 & 0 & 1 \\ \sin \theta & r \cos \theta & 0 \end{pmatrix} = 2r. \quad (3.5)$$

The general jacobian for an elliptic transformation is $J = abr$, where a and b are the semi-major and semi-minor axes.

The volume is limited by

$$y \leq 1 - z, \quad y \geq z - 1 \quad \text{and} \quad x^2 + 4z^2 < 4. \quad (3.6)$$

Substituting $y = y, z = r \sin \theta$ and $x = r \cos \theta$, we get the limits

$$y \leq 1 - r \sin \theta, \quad y \geq r \sin \theta - 1 \quad \text{and} \quad r \leq 1. \quad (3.7)$$

Therefore the volume of the solid is

$$V = \int dV = \int_{r \sin \theta - 1}^{1 - r \sin \theta} \int_0^{2\pi} \int_0^1 2r dr d\theta dy = 4\pi. \quad (3.8)$$

It is important to note that when finding the limit of integrations it is always best to consider the inequalities which define the area we are integrating over.

Example: Elliptic-Paraboloid Sandwich.

Question: Find the volume enclosed by the curves $z = ax^2 + by^2$ and $z = 6 - cx^2 - dy^2$.

Solution: First we find the projection region E in the x - y plane by setting

$$z = ax^2 + by^2 = 6 - cx^2 - dy^2 \implies (a + c)x^2 + (b + d)y^2 = 6. \quad (3.9)$$

For each point (x, y) on the x - y plane we scale the infinitesimal area element by the height

$$\Delta z = (6 - cx^2 - dy^2) - (ax^2 + by^2) = 6 - (a + c)x^2 - (b + d)y^2. \quad (3.10)$$

Therefore the volume is

$$\int_E (6 - (a + c)x^2 - (b + d)y^2) dx dy. \quad (3.11)$$

Using the standard elliptic substitution $x = r \cos \theta / \sqrt{a + c}$ and $y = r \sin \theta / \sqrt{b + d}$ we get

$$V = \int_0^{2\pi} \int_0^{\sqrt{6}} (6 - r^2) \frac{r dr d\theta}{\sqrt{a + c} \sqrt{b + d}} = \frac{18\pi}{\sqrt{(a + c)(b + d)}}. \quad (3.12)$$

3.2 Line, Surface and Volume Integration

Line, surface and volume sum up scalar or vector field times an infinitesimal quantity. The infinitesimal 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

3.2.1 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 systematic 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 algebraically.

⁵The 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.

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

Curves

A curve in space⁸ 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 $f : \mathbb{R} \rightarrow \mathbb{R}^3$, 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

$$\mathbf{r}(t) = (\cos t, \sin t, t), \quad (3.13)$$

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

$$\mathbf{r}(t) = (\cos t, \sin t) \quad \text{or} \quad \left(\frac{2t}{t^2 + 1}, \frac{t^2 - 1}{t^2 + 1} \right).^{11} \quad (3.14)$$

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

But the most common parameterization is a straight line, given a point $\mathbf{r}_0 = (x_0, y_0, z_0)$ on the line and a vector $\mathbf{d} = (a, b, c)$ along the line, which has the form

$$\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{d} = (x_0 + at, y_0 + bt, z_0 + ct), \quad (3.15)$$

then the infinitesimal displacement vector is simply given by

$$d\mathbf{r}(t) = \mathbf{d}dt = (a, b, c)dt. \quad (3.16)$$

Alternatively, a space curve in three-dimensional space can be represented by two simultaneous equations algebraically 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 evolving through space, and it is not a convenient form to work in integrals.

⁸In 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. Examples of non-orientable surfaces include the Möbius strip and the Klein bottle.

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

¹⁰Sometimes, 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.

¹¹With 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.

Surfaces

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

$$\mathbf{r}(\theta, \phi) = R(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (3.17)$$

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

There are generally two ways to compute an infinitesimal surface dS . If the parametric form of the surface is given, then refer to fig. 3.3, since v is constant along KL , the line element KL can be written as

$$d\mathbf{r}_{KL} = \frac{\partial \mathbf{r}}{\partial u} du + \frac{\partial \mathbf{r}}{\partial v} dv = \frac{\partial \mathbf{r}}{\partial u} du = \frac{\partial x}{\partial u} du \hat{\mathbf{i}} + \frac{\partial y}{\partial u} du \hat{\mathbf{j}}. \quad (3.18)$$

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

$$dS_{u,v} = |d\mathbf{r}_{KL} \times d\mathbf{r}_{KN}| = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| dudv = \left| \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right| dudv = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv. \quad (3.19)$$

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

$$d\mathbf{S}_{u,v} = \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) dudv. \quad (3.20)$$

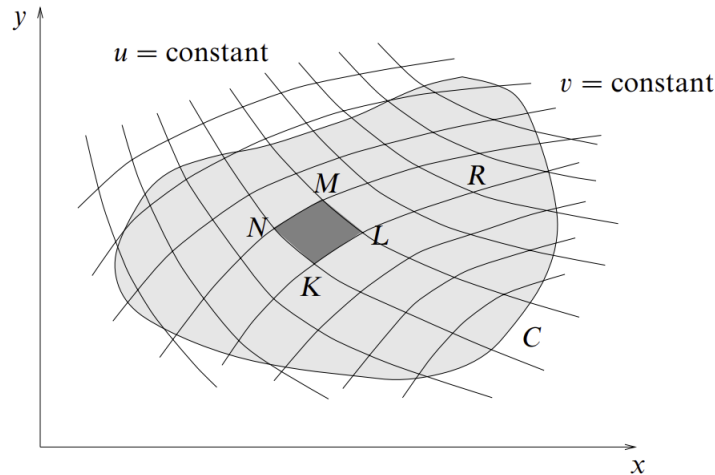


Figure 3.3

If the algebraic form of the surface is given, then we usually project the surface S onto the xy -plane (sometimes the y - z or x - z planes are more useful). From fig. 3.4, we see that

$$dA = |\cos \alpha| dS \implies dS = \frac{dA}{|\hat{\mathbf{n}} \cdot \mathbf{k}|} = \frac{dA}{|\hat{\mathbf{k}} \cdot \nabla f / |\nabla f||} = \frac{|\nabla f|}{|\partial f / \partial z|} dA \implies dS = \frac{|\nabla f|}{|\partial f / \partial z|} dA, \quad (3.21)$$

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.

Note that ∇f points to the side of the surface with $f(x, y, z) > 0$, so occasionally a negative sign has to be added for dS oriente properly. For example, for the paraboloid $z = x^2 + y^2$, if we define $f = z - x^2 - y^2$, then ∇f points to the inside of the paraboloid. Note however that dS is a strictly positive quantity due to the absolute sign in its denominator in the equation above.

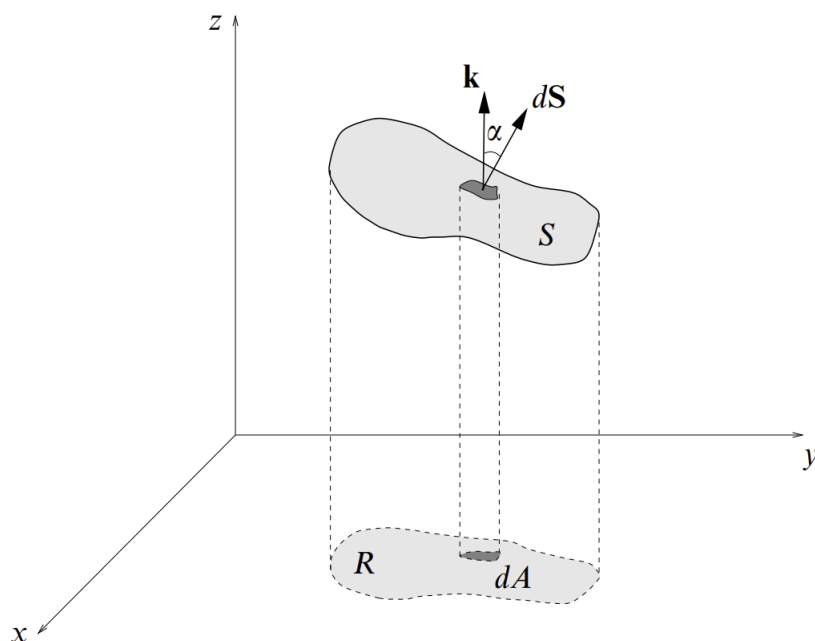


Figure 3.4

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 z - x or the y - z plane.

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

$$\mathbf{r}(\theta, \phi) = R\hat{\mathbf{r}} = R(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (3.22)$$

so we have

$$\begin{aligned} d\mathbf{S} &= \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) d\theta d\phi = R^2 \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix} \times \begin{pmatrix} -\sin \theta \sin \phi \\ \sin \theta \cos \phi \\ 0 \end{pmatrix} d\theta d\phi \\ &= R^2 \sin \theta d\theta d\phi (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) = R^2 \sin \theta d\theta d\phi \hat{\mathbf{r}}. \end{aligned} \quad (3.23)$$

Algebraically, the surface is described as

$$f(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0, \quad (3.24)$$

so we have

$$\begin{aligned} d\mathbf{S} &= \frac{\partial z}{\partial f} dA \nabla f = \frac{1}{2z} dx dy 2(x, y, z) = \left| \frac{\partial(x, y)}{\partial(\theta, \phi)} \right| d\theta d\phi \left(\frac{x}{z}, \frac{y}{z}, 1 \right) \\ &= (R^2 \cos \theta \sin \phi) (\tan \theta \cos \phi, \tan \theta \sin \phi, 1) d\theta d\phi = r^2 \sin \theta d\theta d\phi \hat{\mathbf{r}}. \end{aligned} \quad (3.25)$$

In most cases the projection area on the x - y plane is trivial but in some cases it has to be found algebraically. For example, the intersection surface between the cone $z = \sqrt{x^2 + y^2}$ and the sphere $x^2 + (y - 1)^2 + z^2 = 1$ is founded by substituting z in the first equation into the second.

Example: Space Curve.

Question: Sketch the Folium of Descartes $x^3 + y^3 = 3axy$ and calculate the area enclosed by the loop of this folium.

Solution: We parametrize this curve by

$$x = \frac{3t}{1+t^3} \quad \text{and} \quad y = \frac{3t^2}{1+t^2}, \quad t \geq 0. \quad (3.26)$$

As $t \rightarrow 0$, $(x, y) \rightarrow (0, 0)$. As $t \rightarrow \infty$, $(x, y) \rightarrow (0, 0)$ again. In between, x and y are positive. One therefore finds a loop that starts at the origin, rises to a maximum around $x = y = 1$, then returns to the origin.

The area is therefore

$$A = \oint_C x dy - y dx = \int_0^\infty \frac{3t}{1+t^3} \frac{6t-3t^4}{(1+t^3)^2} dt - \int_0^\infty \frac{3t^2}{1+t^2} \frac{3-6t^3}{(1+t^3)^2} dt = \frac{1}{2} \int_0^\infty \frac{9t^2}{(1+t^3)^2} dt = \frac{3}{2}. \quad (3.27)$$

Example: Space Surface.

Question: Consider the surface $\phi(x, y, z) = c$,

1. Show that $\nabla \phi$ is a vector perpendicular to the surface ϕ .

2. Let R be the distance from a fixed point A to any point P on the surface ϕ . Show that ∇R is a unit vector in the direction AP .
3. Let P be any point on an ellipse whose foci are at points A and B . Given that the distances AP and BP obeys $AP + BP = s$, prove that lines AP and BP make equal angles with the tangent T to the ellipse at P .
4. Imagine a source is emitting sound waves at focus A of the ellipse. Give a physical interpretation of the result in (c).

Solution:

1. From

$$d\phi = \nabla\phi \cdot d\mathbf{r} \quad (3.28)$$

where \mathbf{r} is the position vector of an arbitrary point on the surface and $d\mathbf{r}$ is an arbitrary tangent vector that lies on the surface, we have $\nabla\phi \perp$ every tangent vectors.

2. We calculate the gradient of $R = ((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2)^{1/2}$ to be

$$(\nabla R)_i = \partial_i ((x_i - c_i)(x_i - c_i))^{1/2} = ((x_i - c_i)(x_i - c_i))^{-1/2} (x_i - c_i) = \frac{\mathbf{R}_i}{|\mathbf{R}|}. \quad (3.29)$$

3. Using the fact that $AP + BP = s$, the ellipse can be described by the equation $\phi(x, y) = AP + BP = s$. From (a) and (b), we have the normal vector \mathbf{n}

$$\mathbf{n} = \nabla\phi(x, y) = \nabla AP + \nabla BP = \frac{\mathbf{AP}}{AP} + \frac{\mathbf{BP}}{BP} \quad (3.30)$$

The equal angles can be shown by

$$\frac{\mathbf{AP}}{AP} \cdot \mathbf{n} = 1 + \frac{\mathbf{AP} \cdot \mathbf{BP}}{(AP)(BP)} = \frac{\mathbf{BP}}{BP} \cdot \mathbf{n} = \frac{\mathbf{BP} \cdot \mathbf{AP}}{(BP)(AP)} + 1. \quad (3.31)$$

Alternatively, we can describe the ellipse using

$$\phi(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \implies \nabla\phi(x, y) = \left(\frac{2x}{a^2}, \frac{2y}{b^2} \right). \quad (3.32)$$

Explicitly, we have

$$AP = ((x + c)^2 + y^2)^{1/2} \quad \text{and} \quad BP = ((x - c)^2 + y^2)^{1/2}, \quad (3.33)$$

with their gradients

$$\nabla AP = \frac{(x + c)\hat{\mathbf{i}} + y\hat{\mathbf{j}}}{AP} \quad \text{and} \quad \nabla BP = \frac{(x - c)\hat{\mathbf{i}} + y\hat{\mathbf{j}}}{BP}. \quad (3.34)$$

The equal angles can be shown by

$$\nabla AP \cdot \nabla\phi = \nabla BP \cdot \nabla\phi. \quad (3.35)$$

Substituting and expanding we get

$$a^2y^2 + a^2x^2 + a^2c^2 - a^4 - x^2c^2 = 0, \quad (3.36)$$

which is indeed true once we substitute $c^2 = a^2 - b^2$ and the ellipse equation $x^2/a^2 + y^2/b^2 = 1$.

4. If a point source at A emits sound, each ray travels to the ellipse and reflects according to the law of reflection. By (c), every sound-ray leaving A and hitting the inner boundary of the ellipse will be reflected straight to the other focus B . Thus a listener at B hears sound from A clearly no matter which direction it initially left A .

Example: Orthogonality of Two Space Surfaces.

Question: Calculate all values of the constant a for which the surfaces

$$ay = x \quad \text{and} \quad 5x^2 + y^2 + 3xy + z^2 = 25 \quad (3.37)$$

cut orthogonally. Sketch and describe a geometric interpretation for each of these values of a .

Solution: Rewriting the surfaces in the form

$$v(x, y, z) = ay - x = 0 \quad \text{and} \quad w(x, y, z) = 5x^2 + y^2 + 3xy + z^2 - 25 = 0, \quad (3.38)$$

we require

$$\nabla v \cdot \nabla w = (3a - 10)x + (2a - 3)y = 0. \quad (3.39)$$

This equation does not uniquely determine a , since x and y are still unknown. We combine this condition with the condition that $v(x, y, z) = w(x, y, z) = 0$ to get

$$3a^2 - 8a - 3 = 0 \implies a = 3 \quad \text{or} \quad -\frac{1}{3}. \quad (3.40)$$

The cases for $a = 3$ and $-1/3$ are illustrated in fig. 3.5.

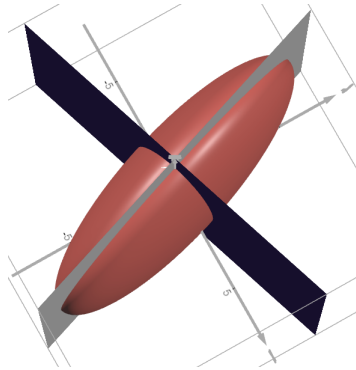


Figure 3.5

Example: Particle on a Surface.

Question: Consider a particle which moves in three dimensions with velocity $\mathbf{v}(t)$ such that

$$\dot{\mathbf{v}}(t) = \gamma t \hat{\mathbf{k}} \times \mathbf{v}(t). \quad (3.41)$$

At time $t = 0$ the particle is located at the origin and moves with velocity $\mathbf{v}(0) = w\hat{\mathbf{i}}$. Determine $\mathbf{v}(t)$ for $t \geq 0$. Find the angular velocity $\omega(t)$ and the radius of curvature $R(t)$.

Solution: In components form, the equation of motion reads

$$\begin{cases} \dot{v}_x &= -\gamma t v_y, \\ \dot{v}_y &= \gamma t v_x. \end{cases} \quad (3.42)$$

Since the force is perpendicular to the velocity the speed remains constant. From the equation of motions we also see that the motion of the particle is confined in the x - y plane. Therefore we can parametrize the velocity of the particle by

$$\mathbf{v}(t) = w(\cos \phi, \sin \phi, 0). \quad (3.43)$$

Substituting into the equations of motion we find

$$\phi = \frac{1}{2}\gamma t^2. \quad (3.44)$$

If ϕ is the polar angle specifying the location of the particle then the angular velocity is

$$\omega(t) = \frac{d\phi(t)}{dt} = \gamma t. \quad (3.45)$$

Since the speed is constant the radius of curvature is simply

$$R = \frac{w}{\omega(t)} = \frac{w}{\gamma t}. \quad (3.46)$$

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

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{|d\mathbf{r}|}. \quad (3.47)$$

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

$$\kappa \hat{\mathbf{n}} = \frac{d\hat{\mathbf{t}}}{ds} = \frac{d^2 \mathbf{r}}{ds^2} \quad (3.48)$$

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 segments around \mathbf{r} form.

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

$$\hat{\mathbf{b}} = \hat{\mathbf{t}} \times \hat{\mathbf{n}}. \quad (3.49)$$

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

$$-\tau \hat{\mathbf{n}} = \frac{d\hat{\mathbf{b}}}{ds}, \quad (3.50)$$

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)

$$\frac{d\hat{\mathbf{t}}}{ds} = \kappa \hat{\mathbf{n}}, \quad \frac{d\hat{\mathbf{n}}}{ds} = \tau \hat{\mathbf{b}} - \kappa \hat{\mathbf{t}} \quad \text{and} \quad \frac{d\hat{\mathbf{b}}}{ds} = -\tau \hat{\mathbf{n}}. \quad (3.51)$$

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

$$\mathbf{a}(t) = \frac{dv}{dt} \hat{\mathbf{t}} + \frac{v^2}{\rho} \hat{\mathbf{n}}. \quad (3.52)$$

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

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{dv}{dt} \hat{\mathbf{t}} + v \frac{d\hat{\mathbf{t}}}{dt} = \frac{dv}{dt} \hat{\mathbf{t}} + v \frac{ds}{dt} \frac{d\hat{\mathbf{t}}}{ds} = \frac{dv}{dt} \hat{\mathbf{t}} + \frac{v^2}{\rho} \hat{\mathbf{n}}. \quad (3.53)$$

3.2.2 Line Integrals

Integrals along a line can involve vector and scalar fields. There are four kinds of line integrals,¹² namely

$$\int_C f(\mathbf{r})ds, \quad \int_C f(\mathbf{r})d\mathbf{r}, \quad \int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} \quad \text{and} \quad \int_C \mathbf{F}(\mathbf{r}) \times d\mathbf{r}. \quad (3.54)$$

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

$$\begin{aligned} s &= \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{r})ds = \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{r})|d\mathbf{r}(u)| = \int_{u_{\mathbf{a}}}^{u_{\mathbf{b}}} f(\mathbf{r}) \left| \frac{d\mathbf{r}}{du} \right| |du| \\ &= \int_{u_{\mathbf{a}}}^{u_{\mathbf{b}}} f(\mathbf{r}) \sqrt{\left(\frac{dx}{du} \right)^2 + \left(\frac{dy}{du} \right)^2 + \left(\frac{dz}{du} \right)^2} |du|, \end{aligned} \quad (3.55)$$

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

$$\mathbf{r}_1(u) = (u, 0) \quad \text{and} \quad \mathbf{r}_2(u) = (1 - u, 0). \quad (3.56)$$

The forward line integral is

$$I = \int_C \phi ds = \int_0^1 2 \left| \frac{d\mathbf{r}_1}{du} \right| du = 2 \int_0^1 \sqrt{1^2 + 0^2} du = 2. \quad (3.57)$$

The backward line integral is

$$I = \int_C \phi ds = \int_0^1 2 \left| \frac{d\mathbf{r}_2}{du} \right| du = 2 \int_0^1 \sqrt{1^2 + 0^2} du = 2. \quad (3.58)$$

¹²However usually only the first and the third integrals are used often enough to be recognized as line integrals.

The limits of the line integral in both cases are 0 to 1, due to the different parameterization 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

$$I = \int_{(1,1)}^{(4,2)} (x + y)dx + (y - x)dy. \quad (3.59)$$

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

$$I = \int_1^2 ((y^2 + y)2y + (y - y^2))dy = \frac{34}{3}. \quad (3.60)$$

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

$$I = \int_0^1 ((3u^2 + u + 2)(4u + 1) - (u^2 + u)(2u))du = \frac{32}{3}. \quad (3.61)$$

3. We split the integral into two parts, then

$$\begin{aligned} I &= \int_{(1,1)}^{(4,1)} ((x + y)dx + (y - x)dy) + \int_{(4,1)}^{(4,2)} ((x + y)dx + (y - x)dy) \\ &= \int_1^4 (x + 1)dx + \int_1^2 (y - 4)dy = 8. \end{aligned} \quad (3.62)$$

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

$$I = \int_{-a}^a \sqrt{a^2 - y^2} dy + \int_a^{-a} \left(-\sqrt{a^2 - y^2}\right) dy = \pi a^2. \quad (3.63)$$

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

$$I = a^2 \int_0^{2\pi} \cos^2 \phi d\phi = \pi a^2. \quad (3.64)$$

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

$$\mathbf{r}(\phi) = a \cos \phi \hat{\mathbf{i}} + a \sin \phi \hat{\mathbf{j}} \quad \text{and} \quad ds = \frac{d\mathbf{r}}{d\phi} d\phi = a d\phi. \quad (3.65)$$

Thus

$$I = \int_0^\pi a^3 (1 - \sin 2\phi) d\phi = \pi a^3. \quad (3.66)$$

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

$$\oint_C (dx + xydz) = \int_0^{2\pi} (-\sin \theta + \cos \theta \sin \theta (-2 \cos \theta \sin \theta - 2 \sin \theta \cos \theta)) d\theta = -\pi. \quad (3.67)$$

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

$$\oint_C (dx + xydz) = \int_S \nabla \times \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}, \quad (3.68)$$

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

$$d\mathbf{S} = \nabla g(\mathbf{r}) = (-2x, 2y, 1), \quad (3.69)$$

where $g(\mathbf{r}) = z - f(x, y) = z - x^2 + y^2$.

Inserting these expressions into the integral we have

$$\int_S \nabla \times \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = -2 \int_{S'} (x^2 + y^2) dx dy = -\pi. \quad (3.70)$$

3.3 Surface Integrals

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

$$\int_S \phi(\mathbf{r}) dS, \quad \int_S \phi(\mathbf{r}) d\mathbf{S}, \quad \int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} \quad \text{and} \quad \int_S \mathbf{F}(\mathbf{r}) \times d\mathbf{S}, \quad (3.71)$$

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 first 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

$$\mathbf{r} = a \sin \theta \cos \phi \hat{\mathbf{i}} + a \sin \theta \sin \phi \hat{\mathbf{j}} + a \cos \theta \hat{\mathbf{k}}. \quad (3.72)$$

An infinitesimal area is given by

$$dS = \left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right| = a^2 \sin \theta d\theta d\phi \quad (3.73)$$

And the total surface area is then

$$S = \int_0^\pi \int_0^{2\pi} a^2 \sin \theta d\theta d\phi = 4\pi a^2. \quad (3.74)$$

Example: Surface Integral (2).

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

Solution: Parameterizing the surface by θ and ϕ ,

$$\mathbf{F} \cdot d\mathbf{S} = x(\hat{\mathbf{i}} \cdot \hat{\mathbf{r}})dS = (a \sin \theta \cos \phi)(\sin \theta \cos \phi)(a^2 \sin \theta d\theta d\phi), \quad (3.75)$$

where we have used

$$\hat{\mathbf{i}} \cdot \hat{\mathbf{r}} = \hat{\mathbf{i}} \cdot \frac{\mathbf{r}}{|\mathbf{r}|} = \hat{\mathbf{i}} \cdot \frac{x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}}{a} = \frac{x}{a}. \quad (3.76)$$

So the integral becomes

$$I = a^3 \int_0^{\frac{\pi}{2}} \sin^3 \theta d\theta \int_0^{2\pi} \cos^2 \phi d\phi = \frac{2\pi a^3}{3}. \quad (3.77)$$

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

$$|\nabla f| = 2|\mathbf{r}| = 2a, \quad \frac{\partial f}{\partial z} = 2z = 2\sqrt{a^2 - x^2 - y^2}. \quad (3.78)$$

Therefore the integral becomes

$$I = \int \int_R \frac{x^2}{\sqrt{a^2 - x^2 - y^2}} dx dy = \frac{2\pi a^3}{3}, \quad (3.79)$$

where the integral is evaluated by changing to polar coordinates.

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 $\nabla g(\mathbf{r}) = (-2x, 1, -2z)$ (and $\partial g / \partial y = 1$). 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

$$\int_{S'} \mathbf{F}(\mathbf{r}) \cdot \nabla g(\mathbf{r}) dx dz = \int_{S'} (-x^2 - 3z^2) dx dz, \quad (3.80)$$

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

$$\int_{S'} (-x^2 - 3z^2) dx dz = \int_0^{2\pi} d\theta \int_0^1 r(-r^2 \cos^2 \theta - 3r^2 \sin^2 \theta) dr = -\pi. \quad (3.81)$$

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

$$\mathbf{r}(\theta, \phi) = 3(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad \theta \in \left[0, \frac{\pi}{2}\right], \quad \phi \in [0, 2\pi]. \quad (3.82)$$

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

$$d\mathbf{S} = \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = 9(\sin^2 \theta \cos \phi, \sin^2 \theta \sin \phi, \sin \theta \cos \theta). \quad (3.83)$$

Then the integral can be carried out straightforwardly as

$$\int_S \mathbf{F}(\mathbf{r}) \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) d\theta d\phi = 279\pi. \quad (3.84)$$

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. 3.6a). 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 .

Since the surface is inherently a sphere, the algebraic form is $f(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$. An infinitesimal surface area is then given by

$$dS = \frac{|\nabla f(x, y, z)|}{\partial f / \partial z} dx dy = \frac{R}{\sqrt{R^2 - x^2 - y^2}} dx dy. \quad (3.85)$$

Changing to polar coordinates with the substitution $x = R/2 + \rho \cos \phi$ and $y = \rho \sin \phi$,

$$\begin{aligned} S &= \int_S dS = \int_S \frac{R}{\sqrt{R^2 - x^2 - y^2}} dx dy \\ &= \int_0^{2\pi} \int_0^{R/2} \frac{R}{\sqrt{R^2 - (R/2 + \rho \cos \phi)^2 - (\rho \sin \phi)^2}} \rho d\rho d\theta, \end{aligned} \quad (3.86)$$

which has no hope to be integrated easily.

Instead, we can parameterize the sphere with the usual θ and ϕ , where $\phi \in [-\pi/2, \pi/2]$ and use the cylinder's equation to find the condition of θ

$$\left(x - \frac{R}{2}\right)^2 + y^2 \leq \left(\frac{R}{2}\right)^2 \implies \sin \theta \leq \cos \phi \implies \theta \leq \sin^{-1}(\cos \phi) = \frac{\pi}{2} - |\phi|. \quad (3.87)$$

The integral becomes

$$S = \int_S dS = \int_{-\pi/2}^{\pi/2} \int_0^{\pi/2-|\phi|} R^2 \sin \theta d\theta d\phi = R^2(\pi - 2). \quad (3.88)$$

Alternatively, we can parameterize the surface with r, θ according to fig. 3.6b, where

$$x = r \cos \theta, \quad y = r \sin \theta \quad \text{and} \quad z = \sqrt{R^2 - x^2 - y^2} = \sqrt{R^2 - r^2} \quad (3.89)$$

for $r \in [0, R \cos \theta], \theta \in [-\pi/2, \pi/2]$

The infinitesimal surface element is then

$$dS = \left| \frac{\partial \mathbf{r}}{\partial r} \times \frac{\partial \mathbf{r}}{\partial \theta} \right| = \frac{rR}{\sqrt{R^2 - r^2}}. \quad (3.90)$$

Thus the integral

$$S = \int_{-\pi/2}^{\pi/2} d\theta \int_0^{R \cos \theta} \frac{rR}{\sqrt{R^2 - r^2}} dr = R^2(\pi - 2). \quad (3.91)$$

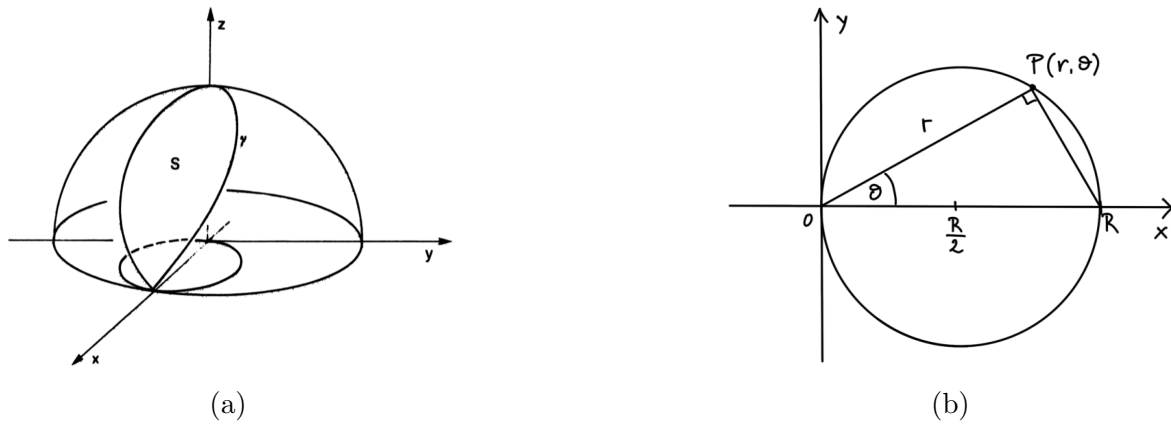


Figure 3.6

Example: Surface Integral (6).

Question: Find the surface area of a cone defined by the equation $4x^2 + y^2 = 9z^2$, and $0 \leq z \leq 1$.

Solution: The cone surface can be parameterized by the ρ and θ as

$$\mathbf{r}(\rho, \theta) = \left(\rho \cos \theta, \rho \sin \theta, \frac{2}{3}\rho \right). \quad (3.92)$$

Therefore an infinitesimal surface area can be found by

$$\begin{aligned} dS &= \left| \frac{\partial \mathbf{r}}{\partial \rho} \times \frac{\partial \mathbf{r}}{\partial \theta} \right| d\rho d\theta = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 2/3 \end{pmatrix} \times \begin{pmatrix} -\rho \sin \theta \\ \rho \cos \theta \\ 0 \end{pmatrix} d\rho d\theta \\ &= \left| \begin{pmatrix} 2\rho \cos \theta/3 \\ 2\rho \sin \theta/3 \\ \rho \end{pmatrix} \right| d\rho d\theta = \frac{\sqrt{13}}{3} \rho d\rho d\theta. \end{aligned} \quad (3.93)$$

The total surface area is therefore

$$S = \int_S dS = \int_0^{2\pi} \int_0^{\frac{3}{2}} \frac{4\sqrt{13}}{3} \rho d\rho d\theta = \frac{3\pi}{4} \sqrt{13}. \quad (3.94)$$

Example: Surface Integral (7).

Question: Evaluate

$$\int_S (x - z) dS, \quad (3.95)$$

where S is the surface of the solid bounded by $x^2 + y^2 = 4$, $z = x - 3$ and $z = x + 2$.

Solution: For the side surface, we use cylindrical coordinates

$$x = 2 \cos \theta, \quad y = 2 \sin \theta, \quad z \in [z_1, z_2], \quad z_1 = 2 \cos \theta - 3 \quad \text{and} \quad z_2 = 2 \cos \theta + 2, \quad (3.96)$$

where the limit in z is bound by the inequality $z \geq x - 3$ and $z \leq x + 2$, which is the formal way to finding the integration limits.

For the bottom and the top cap we have $dS = |\nabla f(x, y, z)| dx dy / (\partial f(x, y, z) / \partial z)$, where $f(x, y, z) = x - z + 3$, so the whole surface integral becomes

$$I = \int_0^{2\pi} \int_{2 \cos \theta - 3}^{2 \cos \theta + 2} (2 \cos \theta - z) 2 dz d\theta + \int 3\sqrt{2} dx dy + \int -2\sqrt{2} dx dy = \pi(10 + 4\sqrt{2}). \quad (3.97)$$

In writing this integral we assume we sum over θ from 0 to 2π , of the vertical strips $(x - z)2(z_{\max} - z_{\min})d\theta$, which would be much harder to do if we sum over z from -5 to 4 , of the circular strips $(x - z)2(\theta_{\max} - \theta_{\min})dz$, since $\theta(z)$ is much harder to find than $z(\theta)$.

3.3.1 Vector Areas

The vector area of a surface S is defined as

$$\mathbf{S} = \int_S d\mathbf{S} = \oint_S \hat{\mathbf{n}} dS \quad (3.98)$$

A closed surface will always has a zero vector area, since when projecting onto the xy -plane, from eq. (3.21) we see that every $d\mathbf{S}_+ = dA/|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|$ will have an opposite contribution $d\mathbf{S}_- = dA/|-\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}| = -d\mathbf{S}_+$.

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

$$\mathbf{S} = \frac{1}{2} \oint_C \mathbf{r} \times d\mathbf{r}, \quad (3.99)$$

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

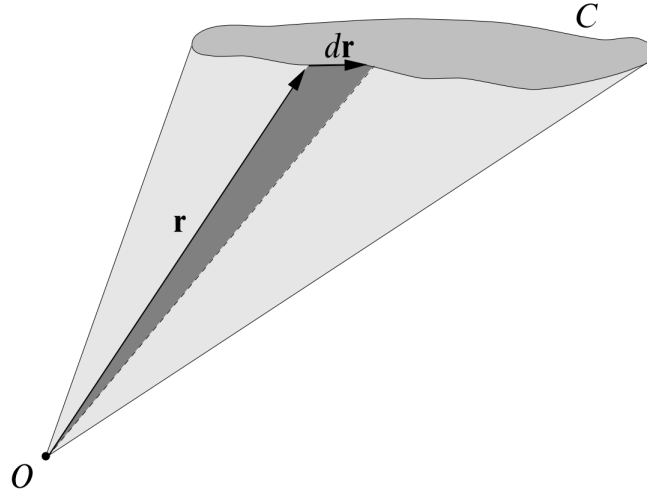


Figure 3.7

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. (3.136).

3.3.2 Solid Angle

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

$$\Omega = \int_S \frac{\hat{\mathbf{r}} \cdot d\mathbf{S}}{r^2}. \quad (3.100)$$

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

3.4 Volume Integrals

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

$$\int_V f(x, y, z) dV \quad \text{and} \quad \int_V \mathbf{F}(x, y, z) dV. \quad (3.101)$$

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. 3.8, we have

$$V = \int_V dV = \frac{1}{3} \oint_S \mathbf{r} \cdot d\mathbf{S}, \quad (3.102)$$

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

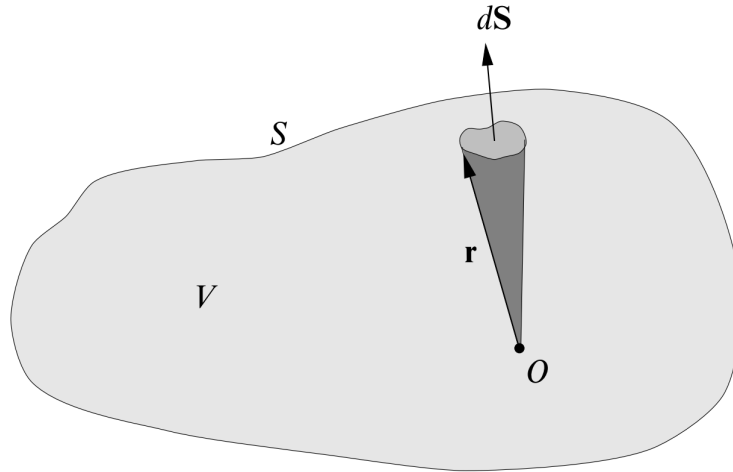


Figure 3.8

Example: Volume Integrals (1).

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

Solution: The most natural choice of coordinates is the cylindrical coordinates, 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 coordinates, these become

$$s \leq \rho \leq R \quad \text{and} \quad -\sqrt{R^2 - \rho^2} \leq z \leq \sqrt{R^2 - \rho^2}. \quad (3.103)$$

So the volume is

$$V = \int_{-\sqrt{R^2 - \rho^2}}^{\sqrt{R^2 - \rho^2}} \int_0^{2\pi} \int_s^R \rho d\rho d\phi dz = \frac{4\pi}{3}(R^2 - s^2)^{\frac{3}{2}}. \quad (3.104)$$

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.

3.5 Integral Theorems

3.5.1 The Divergence Theorem

The divergence theorem states that

$$\int_V (\nabla \cdot \mathbf{F}) dV = \oint_S \mathbf{F} \cdot d\mathbf{S}. \quad (3.105)$$

The corresponding 2D-version is

$$\int_S (\nabla \cdot \mathbf{F}) dS = \oint_C \mathbf{F} \cdot d\mathbf{r}. \quad (3.106)$$

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 $\mathbf{F} = (y - x)\hat{\mathbf{i}} + x^2z\hat{\mathbf{j}} + (z + x^2)\hat{\mathbf{k}}$ 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

$$\int_V (\nabla \cdot \mathbf{F}) dV = 0 = \int_S \mathbf{F} \cdot d\mathbf{S} + \int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1. \quad (3.107)$$

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

$$I = - \int_{S_1} \mathbf{F} \cdot d\mathbf{S}_1 = \int \int_R x^2 dx dy = \frac{\pi a^4}{4}. \quad (3.108)$$

Example: Divergence Theorem.

Question: Show that for any closed surface S in a vector field \mathbf{F} we have

$$\oint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = 0. \quad (3.109)$$

Solution: We apply the divergence theorem to the vector field $\nabla \times \mathbf{F}$ and we get directly

$$\oint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \int_V (\nabla \cdot (\nabla \times \mathbf{F})) dV = 0. \quad (3.110)$$

Example: The Continuity Equation.

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

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (3.111)$$

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

$$\frac{dM}{dt} = \frac{d}{dt} \int_V \rho dV = - \oint_S \rho \mathbf{v} \cdot d\mathbf{S} \quad (3.112)$$

$$\int_V \frac{\partial \rho}{\partial t} dV + \int_V \nabla \cdot (\rho \mathbf{v}) dV = \int_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right) dV = 0. \quad (3.113)$$

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 $\rho = \text{constant}$ and the continuity equation becomes simply $\nabla \cdot \mathbf{v} = 0$.

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 system is

$$\begin{cases} \frac{dw}{dt} &= w(-\alpha + c - \mu w), \\ \frac{dc}{dt} &= c(\beta - w - \nu c), \end{cases} \quad (3.114)$$

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

$$\frac{d\mathbf{a}}{dt} = \mathbf{p}, \quad \mathbf{a} = \begin{pmatrix} w \\ c \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} w(-\alpha + c - \mu w) \\ c(\beta - w - \nu c) \end{pmatrix}, \quad (3.115)$$

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

$$\int_S (\nabla \cdot (b(w, c))\mathbf{p}) dS = \oint_C b(w, c)\mathbf{p} \cdot \hat{\mathbf{n}} dt, \quad (3.116)$$

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

$$\nabla \cdot \left(\frac{\mathbf{p}}{wc} \right) = -\frac{\mu}{c} - \frac{\nu}{w}, \quad (3.117)$$

which is strictly negative. However, if $\nabla \cdot (\mathbf{p}/wc)$ 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

$$\begin{aligned} 4\pi(r + dr)^2 E(r + dr) - 4\pi r^2 E(r) &= \frac{\rho(4\pi r^2 dr)}{\epsilon_0} \\ (r^2 + 2rdr)(E(r) + dE) - r^2 E(r) &= \frac{\rho r^2 dr}{\epsilon_0} \\ 2rE(r)dr + r^2 dE &= \frac{\rho r^2 dr}{\epsilon_0} \\ \frac{2E}{r} + \frac{dE}{dr} &= \frac{1}{r^2} \frac{d}{dr}(r^2 E(r)) = \frac{\rho}{\epsilon_0} \end{aligned} \quad (3.118)$$

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

$$\nabla \cdot \mathbf{E} = \frac{1}{r^2} \frac{d}{dr}(r^2 E(r)) = \frac{\rho}{\epsilon_0}, \quad (3.119)$$

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 $\phi \nabla \psi$, we get

$$\oint_S \phi \nabla \psi \cdot d\mathbf{S} = \int_V \nabla \cdot (\phi \nabla \psi) dV = \int_V (\phi \nabla^2 \psi + (\nabla \phi) \cdot (\nabla \psi)) dV. \quad (3.120)$$

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

$$\oint_S (\phi \nabla \psi - \psi \nabla \phi) \cdot d\mathbf{S} = \int_V (\phi \nabla^2 \psi - \psi \nabla^2 \phi) dV. \quad (3.121)$$

This is known as the Green's second theorem.

Two Other Theorems

Letting \mathbf{F} in eq. (3.105) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_V \nabla f dV = \oint_S f d\mathbf{S} \text{ and } \int_V (\nabla \times \mathbf{F}) dV = \oint_S d\mathbf{S} \times \mathbf{F}. \quad (3.122)$$

3.5.2 Stokes' Theorem

The Stokes' theorem states that

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_C \mathbf{F} \cdot d\mathbf{r}. \quad (3.123)$$

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

$$\int_S \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dS = \oint_C (F_x dx + F_y dy). \quad (3.124)$$

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.

A surface without holes is called simply-connected region, defined by the property that any closed path on the region can be contracted to a point without leaving the surface. In layman's term, it has more than one boundaries, or more simply, it has holes.

A surface that is orientable is one on which you can choose a continuous normal vector everywhere so that as you move around the surface, the normal never flips discontinuously. More simply put, you can distinguish its two-sides globally. A Möbius strip or a Klein bottle are examples of non-orientable surfaces. If you start with a chosen normal vector

at one point and carry it continuously around certain loops, you return to the original point with the normal reversed.

The Stokes' theorem only holds for smooth orientable surfaces. If the region is not simply-connected, 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 continuous boundary as shown in fig. 3.9, with an infinitesimal gap. The two contributions across the gap then cancel. The same technique can be used to apply Stokes' theorem to closed surface.

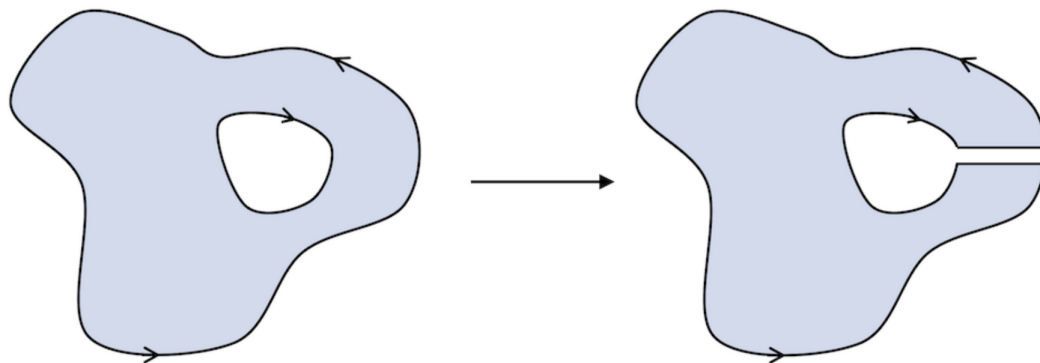


Figure 3.9

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

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^\alpha R^2 \sin \theta d\theta d\phi \begin{pmatrix} -x \\ 0 \\ z \end{pmatrix} \cdot \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} = \pi R^3 \cos \alpha \sin^2 \alpha. \quad (3.125)$$

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

$$\mathbf{r}(\phi) = R(\sin \alpha \cos \phi, \sin \alpha \sin \phi, \cos \alpha) \implies d\mathbf{r} = R(-\sin \alpha \sin \phi, \sin \alpha \cos \phi, 0)d\phi. \quad (3.126)$$

So the line integral is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} R x z \sin \alpha \cos \phi d\phi = \pi R^3 \sin^2 \alpha \cos \alpha. \quad (3.127)$$

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

$$\mathbf{r}(\rho, \phi) = (\rho \cos \phi, \rho \sin \phi, z) \quad (3.128)$$

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

$$d\mathbf{S} = \frac{\partial \mathbf{r}}{\partial \rho} \times \frac{\partial \mathbf{r}}{\partial \phi} = (-\cos \phi, -\sin \phi, 1)\rho d\rho d\phi. \quad (3.129)$$

The surface integral is

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \int_0^{2\pi} \int_a^b \begin{pmatrix} -x \\ 0 \\ z \end{pmatrix} \cdot \begin{pmatrix} -\cos \phi \\ -\sin \phi \\ 1 \end{pmatrix} \rho d\rho d\phi = \pi(b^3 - a^3). \quad (3.130)$$

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

$$\mathbf{r}(\phi) = R(\cos \phi, \sin \phi, 1) \implies d\mathbf{x} = R(-\sin \phi, \cos \phi, 0)d\phi, \quad (3.131)$$

so the line integral of one of the circumference is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} R^3 \cos^2 \phi d\phi = \pi R^3. \quad (3.132)$$

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

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{S} = \mu_0 I = \mu_0 \int_S \mathbf{J} \cdot d\mathbf{S}. \quad (3.133)$$

Hence

$$\int_S (\nabla \times \mathbf{B} - \mu_0 \mathbf{J}) \cdot d\mathbf{S} = 0. \quad (3.134)$$

Example: Area using Stokes' theorem

Question: Show that the area of a region R enclosed by a simple closed curve C is given by

$$A = \frac{1}{2} \oint_C (x dy - y dx) = \oint_C x dy = - \oint_C y dx. \quad (3.135)$$

Hence calculate the area of the ellipse with the parameterization $x = a \cos \phi, y = b \sin \phi$.

Solution: By Green's theorem we have

$$\oint_C (x dy - y dx) = \int_S (1 + 1) dx dy = 2A. \quad (3.136)$$

Similarly,

$$\oint_C x dy = \int_S (1) dx dy = \oint_C -y dx = A. \quad (3.137)$$

Therefore the area of an ellipse is

$$A = \frac{1}{2} \int_0^{2\pi} ab(\cos^2 \phi + \sin^2 \phi) d\phi = \pi ab. \quad (3.138)$$

Example: Green's Theorem in a Plane

Question: A closed path C is described by the upper half ($y \geq 0$) of the circle of radius 1 in the x - y plane centered at the origin, and the portion of $y = x^2 - 1$ between $-1 \leq x \leq 1$. Evaluate the line integral

$$\oint_C \mathbf{F} \cdot d\mathbf{r}, \quad \mathbf{F} = 3y\hat{\mathbf{i}} + (x^2 - y)\hat{\mathbf{j}}. \quad (3.139)$$

Solution: The upper semicircle is parameterized by the polar angle θ as

$$\mathbf{r}_1(\theta) = (\cos \theta, \sin \theta), \quad d\mathbf{r}_1(\theta) = (-\sin \theta, \cos \theta) d\theta. \quad (3.140)$$

The integrand is therefore

$$\mathbf{F}(\mathbf{r}_1(\theta)) \cdot d\mathbf{r}_1 = (3 \sin \theta, \cos^2 \theta - \sin \theta) \cdot (-\sin \theta, \cos \theta) d\theta = -3 \sin^2 \theta + \cos^3 \theta - \sin \theta \cos \theta. \quad (3.141)$$

Integrating over θ from 0 to π gives $I_1 = -3\pi/2$.

The lower parabola is parameterized by x since its algebraic form is given

$$\mathbf{r}_2(x) = (x, x^2 - 1). \quad (3.142)$$

The integrand is therefore

$$\mathbf{F}(\mathbf{r}_2(x)) \cdot d\mathbf{r}_2 = (3(x^2 - 1), 1) \cdot (1, 2x)dx. \quad (3.143)$$

Integrating over x from -1 to 1 gives $I_2 = -4$.

So the final answer is $I = I_1 + I_2 = -3\pi/2 - 4$.

A much faster method is to use the Green's theorem in a plane, which gives

$$\begin{aligned} \oint_C \mathbf{F} \cdot d\mathbf{r} &= \oint_C (3y\hat{\mathbf{i}} + (x^2 - y)\hat{\mathbf{j}}) \cdot (dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}}) = \oint_C (3ydx + (x^2 - y)dy) \\ &= \int_S (2x - 3)dS = \int_{-1}^1 \int_{x^2-1}^{\sqrt{1-x^2}} (2x - 3)dydx = -\frac{3\pi}{2} - 4. \end{aligned} \quad (3.144)$$

Example: Stokes' Theorem of a Closed Surface.

Question: Prove that for a closed surface S

$$\oint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = 0. \quad (3.145)$$

Solution: From the Stokes' theorem

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_C \mathbf{F} \cdot d\mathbf{r}, \quad (3.146)$$

RHS is trivially zero since a closed surface has no boundary curve, so LHS is also zero.

From the divergence theorem

$$\int_V \nabla \cdot (\nabla \times \mathbf{F}) dV = \oint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}, \quad (3.147)$$

LHS is trivially zero due to the vector identity $\nabla \cdot (\nabla \times \mathbf{F}) = 0$, so RHS is also zero.

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 $\partial P/\partial y = \partial Q/\partial x$, 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. (3.123) to be a gradient of another scalar function, or the cross product of two other vector functions, we get

$$\int_S d\mathbf{S} \times \nabla f = \oint_C f d\mathbf{r} \text{ and } \int_S (d\mathbf{S} \times \nabla) \times \mathbf{F} = \oint_C d\mathbf{r} \times \mathbf{F}. \quad (3.148)$$

Appendices

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 product rule, which is trivial, by induction we have proved eq. (1.1) holds for all positive integers n .

□