MATH2001 & MATH7000

Lecture Workbook

Contents

1 Solutions of first order ODEs

By the end of this section, you should be able to answer the following questions about first order ODEs:

- How do you solve an IVP associated with directly integrable, separable or linear ODEs? (Revision)
- Under what conditions does a solution to an IVP problem exist?
- Under what conditions is a solution to an IVP problem unique?

In MATH1052, you were introduced to Ordinary Differential Equations (ODEs) and Initial Value Problems (IVPs) and saw how to find solutions to some special types of first order equations. In particular, there should be three types of first order ODEs that you are familiar with solving.

- Directly integrable: $\dfrac{dy}{dx} = f(x)$.

$\rightarrow y(x) = \int f(x)\,dx + c$

- Separable: $\dfrac{dy}{dx} = f(x)g(y)$.

$\rightarrow \dfrac{1}{g(y)}\dfrac{dy}{dx} = f(x) \Rightarrow \int \dfrac{1}{g(y)}\dfrac{dy}{dx}\,dx = \int f(x)\,dx$

$\Rightarrow \int \dfrac{dy}{g(y)} = \int f(x)\,dx + c$

- Linear: $\dfrac{dy}{dx} = q(x) - p(x)y$.

$\rightarrow \dfrac{dy}{dx} + p(x)y = q(x) \Rightarrow I(x)\dfrac{dy}{dx} + I(x)p(x)y = I(x)q(x)$

s.t. $\rightarrow \dfrac{d}{dx}(Iy) = Iq$ then integrate

$(I = e^{\int p\,dx}$ "integrating factor")

In most applications involving first order ODEs, we are required to solve an IVP. Generally, this is a problem of the form

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0.$$

In otherwords, we seek to find solutions of the ODE which pass through the point (x_0, y_0) in the x-y plane.

Consider the following three examples.

1.1 Example: $\frac{dy}{dx} = x$, $y(0) = 1$ has a unique solution

$\Rightarrow y = \frac{1}{2}x^2 + c$

$y(0) = 1 \Rightarrow 1 = 0 + c \Rightarrow y(x) = \frac{1}{2}x^2 + 1$

1.2 Example: $\frac{dy}{dx} = 3xy^{1/3}$, $y(0) = 0$ has more than one solution

$\Rightarrow \int y^{-1/3}\,dy = \int 3x\,dx$

$\Rightarrow \frac{3}{2}y^{2/3} = \frac{3}{2}x^2 + c$

but $x = 0, y = 0 \Rightarrow c = 0$

$\Rightarrow y^{2/3} = x^2 \Rightarrow y = x^3$ or $y = -x^3$

(another is the equilibrium solution $y = 0$, are there more?)

1.3 Example: $\dfrac{dy}{dx} = \dfrac{x-y}{x}$, $y(0) = 1$ has no solution

$\Rightarrow \dfrac{dy}{dx} + \dfrac{1}{x} y = 1$ (linear, integrating factor $e^{\int \frac{1}{x} dx} = x$)

$\Rightarrow x\dfrac{dy}{dx} + y = x$

$\Rightarrow \dfrac{d}{dx}(xy) = x$

$\Rightarrow xy = \frac{1}{2}x^2 + C$ (family of hyperbolas? $\rightarrow$ CH19)

solution satisfying $y(0) = 1$ $\Rightarrow C = 0 \Rightarrow y = \frac{1}{2}x$

1.4 Existence and uniqueness criteria

Here we consider the initial value problem of the form

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0.$$

The main result concerns the conditions under which we have existence and uniqueness of a solution.

Picard's Theorem

If f and f_y are continuous in some rectangle

$$R = \{(x, y) |\ |x - x_0| \le a, |y - y_0| \le b\},$$

then there is some interval $|x - x_0| \le h \le a$ which contains a unique solution $y = \phi(x)$ of the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0.$$

The above theorem only tell us that a solution exists or is unique locally (i.e., in the rectangle R). Beyond R, we simply don't know. Let's look at the previous three examples in the context of the theorem. In the next lecture, we look closer at understanding these conditions and how they arise.

1.5 Example: $\dfrac{dy}{dx} = x,\ y(0) = 1$

$f(x,y) = x$ continuous

$f_y = 0$ cts

$\left.\right\}$ $\rightarrow$ Picard's Theorem $\Rightarrow$ unique solution

1.6 Example: $\dfrac{dy}{dx} = 3xy^{1/3},\ y(0) = 0$

$f(x,y) = 3xy^{1/3}$ continuous

$f_y = \dfrac{x}{y^{2/3}}$ not cts in any rectangle containing $(0,0)$

$\rightarrow$ Theorem says solutions exist, but we can't comment on uniqueness.

1.7 Example: $\dfrac{dy}{dx} = \dfrac{x-y}{x},\ y(0) = 1$

$f(x,y) = \dfrac{x-y}{x}$ not cts in any rectangle containing $(0,1)$

Theorem $\rightarrow$ cannot comment on existence or uniqueness.

Notes.

2 Method of successive approximations

By the end of this section, you should be able to answer the following questions:

- How do you convert a first order ODE to an integral equation?
- How do you generate the sequence of functions for method of successive approximations?
- Under what conditions does this iterative method work? (Or: What can go wrong?)

Here we look at a proof of the existence and uniqueness theorem from the previous chapter. The presentation is based on that of the book *Elementary Differential Equations and Boundary Value Problems* by W.E. Boyce and R.C. DiPrima (ed. 10, Wiley, 2012).

2.1 A modified theorem

To start, we note that it is always possible to apply a variable shift so that the initial value problem can be written as

$$\frac{dy}{dx} = f(x, y), \quad y(0) = 0. \tag{1}$$

2.1.1 Example: $y' = 2(x-1)(y-1)$, $y(1) = 2$

set $\left.\begin{array}{l}\bar{x} = x-1 \\ \bar{y} = y-2\end{array}\right\}$ check $\frac{dy}{dx} = \frac{d\bar{y}}{d\bar{x}} \Rightarrow \frac{d\bar{y}}{d\bar{x}} = 2\bar{x}(\bar{y}+1)$, $\bar{y}(0) = 0$

Without loss of generality, we will consider this problem where the initial point is at the origin. We can restate the result as follows.

If f and f_y are continuous in some rectangle

$$R = \{(x, y) \mid |x| \leq a, |y| \leq b\},$$

then there is some interval $|x| \leq h \leq a$ which contains a unique solution $y = \phi(x)$ of the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(0) = 0.$$

2.2 Equivalence with integral equation

Let $y = \phi(x)$ be a solution to the IVP (1), and note that the function $F(x) = f(x, \phi(x))$ is a continuous function of x only. We then have

$$\phi(x) = \int_0^x F(t)\ dt = \int_0^x f(t, \phi(t))\ dt. \tag{2}$$

Note that $\phi(0) = 0$. This is an example of an *integral equation*. Conversely, let $\phi(x)$ satisfy the integral equation (2). By the Fundamental Theorem of Integral Calculus, $\phi'(x) = f(x, \phi(x))$, which implies that $y = \phi(x)$ is a solution of the IVP (1). In other words, the IVP (1) and the integral equation (2) are *equivalent*, meaning that a solution of one is a solution of the other. Herein we work with (2).

2.3 Method of successive approximations

The goal of the approach is to generate a sequence of functions $\{\phi_0, \phi_1, \ldots, \phi_n, \ldots\}$. Starting with the initial function $\phi_0(x) = 0$ (satisfying the initial condition of (1)), the sequence is generated iteratively by

$$\phi_{n+1}(x) = \int_0^x f(t, \phi_n(t))\ dt. \tag{3}$$

Note that each ϕ_n satisfies $\phi_n(0) = 0$, but generally not the integral equation (2) itself. If there is a k, however, such that $\phi_{k+1}(x) = \phi_k(x)$, then $\phi_k(x)$ *is* a solution of the integral equation (2) and hence the IVP (1). Generally this does not occur, but we may instead consider *limit functions*.

There are four key points to consider:

1. Do all members of the sequence *exist*?

2. Does the sequence *converge* to a limit function ϕ?

3. What are the properties of ϕ?

4. If ϕ satisfies the IVP (1), are there other solutions?

2.4 Example: $y' = 2x(y+1)$, $y(0) = 0$, $\phi_0(x) = 0$, $f(x,y) = 2x(y+1)$

$\phi_1(x) = \int_0^x f(t, \phi_0(t))\,dt = \int_0^x 2t\,dt = x^2$

$\phi_2(x) = \int_0^x f(t, \phi_1(t))\,dt = \int_0^x 2t(1+t^2)\,dt = \frac{1}{2}x^4 + x^2$

$\phi_3(x) = \frac{1}{6}x^6 + \frac{1}{2}x^4 + x^2$, $\phi_4(x) = x^2 + \frac{1}{2}x^4 + \frac{1}{6}x^6 + \frac{1}{24}x^8$

<u>Proposition:</u> $\phi_n(x) = x^2 + \frac{1}{2}x^4 + \ldots + \frac{1}{n!}x^{2n}$

<u>Proof</u>: induction: True for $n=1$, assume true for $n=k$

$\Rightarrow \phi_{k+1}(x) = \int_0^x 2t(1+\phi_k(t))\,dt = \int_0^x 2t\left(1 + t^2 + \frac{1}{2}t^4 + \ldots + \frac{1}{k!}t^{2k}\right)dt$

$= x^2 + \frac{1}{2}x^4 + \ldots + \frac{2}{k!}\frac{x^{2k+2}}{(2k+2)}$ $\leftarrow (k+1)!$

$\Rightarrow$ true for $n = k+1$

$\Rightarrow$ true by induction QED

See that $\phi_n(x)$ is the nth partial sum of $\sum_{j=1}^{\infty} \frac{x^{2j}}{j!}$

$\Rightarrow \lim_{n\to\infty} \phi_n(x)$ exists iff this series converges

Applying the <u>ratio test</u>, for each x, we have

$$\left| \frac{x^{2j+2}}{(j+1)!} \cdot \frac{j!}{x^{2j}} \right| = \frac{x^2}{j+1} \longrightarrow 0 \text{ as } j \to \infty$$

$\therefore$ series converges $\forall x$

2.5 Discussion

Even though the example demonstrates many of the features that need to be considered, here we elaborate more on the four points outlined earlier.

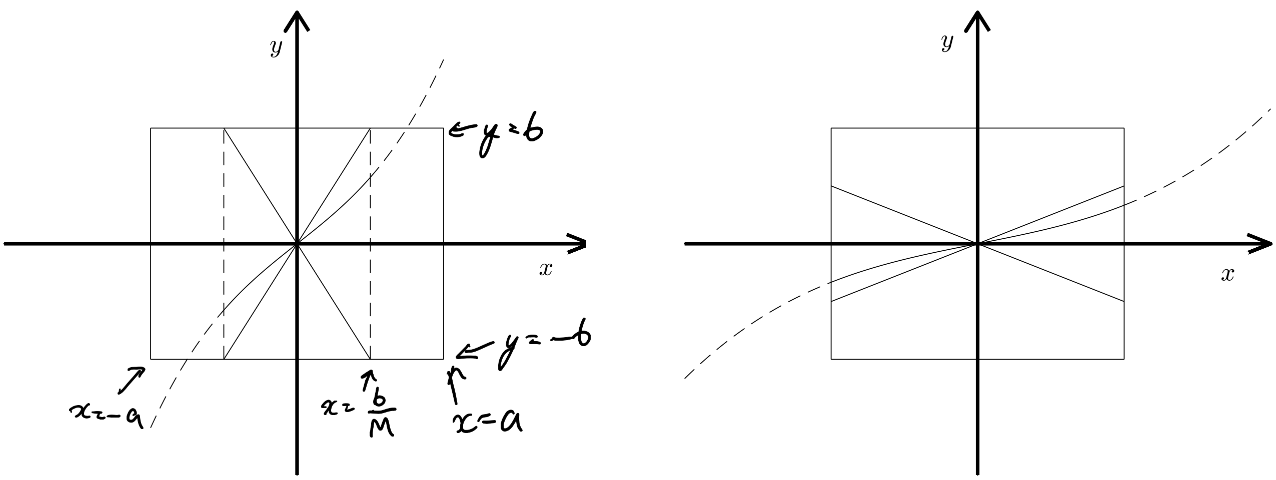
1. Existence of $\{\phi_0, \phi_1, \ldots, \phi_n, \ldots\}$

The conditions of the theorem state that f and f_y are continuous in the rectangle. Even still, it may be that there is a j such that the curve $y = \phi_j(x)$ contains points outside R. In such a case, to apply the iteration and determine $\phi_{j+1}(x)$, we would have to evaluate $f(x, y)$ at points where f may not be continuous or even defined! We can, however, restrict x to a smaller interval if necessary.

To find such an interval, we note that if f is continuous on a closed region, then f is bounded. In other words, there is an M such that

$$|f(x, y)| \leq M, \quad \forall (x, y) \in R.$$

Also, $\phi_j(0) = 0$. Since $f(x, \phi_j(x)) = \phi'_{j+1}(x)$, the largest slope of the curve $y = \phi_{j+1}(x)$ is M. There are two situations which can arise, as depicted in the following diagrams:



We can therefore say that $(x, \phi_{k+1}(x)) \in R$ provided $|x| \leq h$, where $h = \min(\frac{b}{M}, a)$. We then consider the values of (x, y) in the (possibly smaller) rectangle

$$R' = \{(x, y)|\ |x| \leq h, |y| \leq b\}.$$

In this case, all members of the sequence exist.

2. Convergence?

Note that we can write

$$\phi_n(x) = \phi_1(x) + [\phi_2(x) - \phi_1(x)] + [\phi_3(x) - \phi_2(x)] + \cdots + [\phi_n(x) - \phi_{n-1}(x)], \qquad (4)$$

which is the nth partial sum of the series

$$S(x) = \phi_1(x) + \sum_{k=1}^{\infty} [\phi_{k+1}(x) - \phi_k(x)].$$

Convergence of $S(x)$ implies convergence of the sequence $\{\phi_n(x)\}$.

Using the Mean Value Theorem, it is possible to show that if f_y is continuous in R', then there exists a $K > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \leq K|y_1 - y_2|,$$

where $(x, y_1), (x, y_2) \in R'$. In particular, we have

$$|f(x, \phi_n(x)) - f(x, \phi_{n-1}(x))| \leq K|\phi_n(x) - \phi_{n-1}(x)|. \qquad (5)$$

Set $n = 1$ in (5), and show that $\exists M > 0$ such that $|\phi_2(x) - \phi_1(x)| \leq MK|x|^2/2$.

It is then possible to show by induction that

$$|\phi_n(x) - \phi_{n-1}(x)| \leq \frac{MK^{n-1}|x|^n}{n!} \leq \frac{MK^{n-1}h^n}{n!}.$$

Given the expression (4) above for $\phi_n(x)$, we have

$$\begin{aligned} |\phi_n(x)| &\leq |\phi_1(x)| + |\phi_2(x) - \phi_1(x)| + \cdots + |\phi_n(x) - \phi_{n-1}(x)| \\ &\leq \frac{M}{K}\left(Kh + \frac{(Kh)^2}{2!} + \cdots + \frac{(Kh)^n}{n!}\right) \end{aligned} \tag{6}$$

In the limit as $n \to \infty$, this expression becomes

$$\frac{M}{K}(e^{Kh} - 1).$$

Therefore the sum in (6) converges, and so the sum in (4) is absolutely convergent. We conclude that the sequence $\{\phi_n(x)\}$ converges, since it is the sequence of partial sums of a convergent series.

3. Properties of $\phi(x) = \lim_{n\to\infty} \phi_n(x)$

We can make use of the previous discussion in Point 2 to establish *uniform convergence* (see MATH2400/MATH2401) of the sequence $\{\phi_n\}$. In particular, the functions ϕ_n are continuous for all n, so uniform convergence implies the limit function is also continuous.

Now we take the limit of both sides of equation (3):

$$\begin{aligned}\phi(x) &= \lim_{n\to\infty}\int_0^x f(s,\phi_n(s))ds \\ &= \int_0^x \lim_{n\to\infty} f(s,\phi_n(s))ds.\end{aligned}$$

This last step is allowed only because the sequence $\{\phi_n\}$ converges uniformly. Also, since f is continuous, we can develop this as

$$\begin{aligned}\phi(x) &= \int_0^x f(s,\lim_{n\to\infty}\phi_n(s))ds \\ &= \int_0^x f(s,\phi(s))ds.\end{aligned}$$

That is, ϕ solves the integral equation (2) and hence the IVP (1).

4. Uniqueness?

Here we proceed as in the earlier example. That is, we assume $y = \psi(x)$ is another solution. It is possible to show that

$$|\phi(x) - \psi(x)| \le A\int_0^x |\phi(s) - \psi(s)|ds$$

for $0 \le x \le h$ and suitable $A > 0$. Then apply the same argument as in the example.

Notes. From example 2.4:

Uniqueness? Suppose ϕ and ψ satisfy the IVP (& hence the integral eq)

$$\Rightarrow \phi(x) - \psi(x) = \int_0^x 2s(\phi(s) - \psi(s))\,ds$$

$$\Rightarrow |\phi(x) - \psi(x)| \leq \int_0^x 2s\,|\phi(s) - \psi(s)|\,ds$$

$$\leq A\int_0^x |\phi(s) - \psi(s)|\,ds$$

for $0 \leq x \leq \frac{A}{2}$ (i.e. $2x \leq A$) (arbitrary A)

Set $U(x) = \int_0^x |\phi(s) - \psi(s)|\,ds$

$\Rightarrow U(0) = 0$ and $U(x) \geq 0 \quad \forall x \geq 0$ $\qquad (*)$

U is differentiable, $U'(x) = |\phi(x) - \psi(x)|$

$\Rightarrow U'(x) - AU(x) \leq 0$ for $0 \leq x \leq \frac{A}{2}$

$\Rightarrow (e^{-Ax}U(x))' \leq 0$

$\Rightarrow e^{-Ax}U(x) \leq 0$ $\qquad$ (note $e^{-Ax} > 0$ & $U(0) = 0$

Since A is arbitrary, this holds for $\forall x > 0$

Only $U(x) = 0$ works with $(*)$ above

$\Rightarrow U'(x) = 0 \Rightarrow \phi(x) = \psi(x) \quad \forall x \geq 0$

3 Exact first order ODEs

By the end of this section, you should be able to answer the following questions about first order ODEs:

- How do you identify an exact ODE?
- How do you solve an exact ODE?

3.1 Definition

$f(x(t), y(t)) \approx f(t)$

First recall that if $z = f(x, y)$ is a differentiable function of x and y, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t, then z is a differentiable function of t whose derivative is given by the chain rule:

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Now suppose the equation

$$f(x, y) = C$$

defines y implicitly as a function of x (here C is a constant). Then $y = y(x)$ can be shown to satisfy a first order ODE obtained by using the chain rule above. In this case, $z = f(x, y(x)) = C$, so

$$\begin{aligned} (0 = \ )\ \frac{dz}{dx} &= \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} \\ \Rightarrow\ f_x + f_y y' &= 0. \end{aligned} \tag{7}$$

A first order ODE of the form

$$P(x, y) + Q(x, y)\frac{dy}{dx} = 0 \tag{8}$$

is called exact if there is a function $f(x, y)$ (compare (8) with (7) above) such that

$$f_x(x, y) = P(x, y) \text{ and } f_y(x, y) = Q(x, y).$$

The solution is then given implicitly by the equation

$$f(x, y) = C.$$

The constant C can usually be determined by some kind of "initial condition".

Given an equation of the form (8), how do we determine whether or not it is exact? There is a simple test.

3.2 Test for exactness

$\nearrow$ consider $\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial y \partial x}$, $\frac{\partial Q}{\partial x} = \frac{\partial^2 f}{\partial x \partial y}$ $\Big\}$ equal if cts (Clairaut's Thm)

> Let P, Q, $\frac{\partial P}{\partial y}$, and $\frac{\partial Q}{\partial x}$ be continuous over some region of interest. Then
>
> $$P(x,y) + Q(x,y)\frac{dy}{dx} = 0$$
>
> is an exact ODE iff
>
> $$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$
>
> everywhere in the region.

The problem of actually determining $f(x,y)$ is still outstanding. Consider the following example.

3.3 Example: $\overbrace{2x + e^y}^{P} + \overbrace{xe^y}^{Q} y' = 0$

$\nearrow y' = \frac{-(2x+e^y)}{xe^y}$ $\times$

Exact? $\frac{\partial P}{\partial y} = e^y$, $\frac{\partial Q}{\partial x} = e^y \Rightarrow \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} = e^y$

$\Rightarrow$ this ODE is exact

$\Rightarrow \exists f(x,y)$ such that $P = \frac{\partial f}{\partial x}$ and $Q = \frac{\partial f}{\partial y}$

with $f(x,y) = k$ (constant k) giving an implicit solution $y(x)$ of the ODE

$\Rightarrow \frac{\partial f}{\partial x} = 2x + e^y$

Integrate partially w.r.t. x (i.e. treat y as a constant)

$\Rightarrow f(x,y) = x^2 + xe^y + \underline{g(y)}$

this is similar to a constant of integration

$\Rightarrow \frac{\partial f}{\partial y} = 0 + xe^y + g'(y) = Q = \underline{xe^y}$

$\Rightarrow g'(y) = 0 \Rightarrow g(y) = C$ (constant)

$\Rightarrow f(x,y) = x^2 + xe^y + C$

$\therefore$ the implicit solution of the ODE is

$x^2 + xe^y = k \quad \longrightarrow y(x) = \ln\left(\frac{k - x^2}{x}\right)$

$\uparrow$ explicit solution

3.4 Almost exact ODEs and integrating factors

Let's say that we have an equation

$$P(x,y) + Q(x,y)\frac{dy}{dx} = 0$$

such that

$$\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}.$$

The test we have just seen tells us that the ODE is not exact. Are we still able to do anything with it? Here we consider using an "integrating factor", which is different to the one introduced to solve linear ODEs.

The idea is to multiply the ODE by a function $h(x,y)$ and then see if it is possible to choose $h(x,y)$ such that the resulting equation

$$h(x,y)P(x,y) + h(x,y)Q(x,y)\frac{dy}{dx} = 0$$

is exact. We know from the test that this new equation is exact if and only if

$$\frac{\partial}{\partial y}(hP) = \frac{\partial}{\partial x}(hQ).$$

Let's see if we can find such a function:

test $\Rightarrow$ $h_y P + h P_y = h_x Q + h Q_x$

$\Rightarrow$ $h_y P - h_x Q = h(Q_x - P_y)$ $\quad *$

(first order PDE in h)

In general, the equation for $h(x,y)$ is usually just as difficult to solve as the original ODE. In some cases, however, we may be able to find an integrating factor which is a function of only one of the variables x or y. Let's try $h \equiv h(x)$:

$$\frac{dh}{dx} = h\left(\frac{P_y - Q_x}{Q}\right)$$

If $\frac{P_y - Q_x}{Q}$ is a function of x only, then we have a $\boxed{\text{separable ODE}}$ and can be solved for $h(x)$

3.5 Example: $\overbrace{(3xy+y^2)}^{P} + \overbrace{(x^2+xy)}^{Q}\dfrac{dy}{dx} = 0$

$\dfrac{\partial P}{\partial y} = 3x + 2y \qquad \dfrac{\partial Q}{\partial x} = 2x + y$

so $\dfrac{\partial P}{\partial y} \neq \dfrac{\partial Q}{\partial x} \Rightarrow$ ODE is <u>not</u> exact

---

$h = h(x)$

$\Rightarrow \underbrace{h(3xy+y^2)}_{\bar{P}} + \underbrace{h(x^2+xy)}_{\bar{Q}}\dfrac{dy}{dx} = 0$

this equation is exact iff $\dfrac{\partial \bar{P}}{\partial y} = \dfrac{\partial \bar{Q}}{\partial x}$

$\Rightarrow h(3x+2y) = h'(x^2+xy) + h(2x+y)$

$\Rightarrow h(x+y) = h'x(x+y) \Rightarrow \underbrace{h' = \dfrac{h}{x}}$ separable ODE

$\Rightarrow h = \alpha x$ for any α. Choose $\alpha = 1 \Rightarrow h = x$

$\Rightarrow$ new ODE is

$3x^2y + xy^2 + (x^3 + x^2y)\dfrac{dy}{dx} = 0$

This ODE is exact: $\dfrac{\partial P}{\partial y} = 3x^2 + 2xy = \dfrac{\partial Q}{\partial x}$

$\rightarrow x^3y + \frac{1}{2}x^2y^2 = k$

Note: Using $h = h(x)$ doesn't always work!

- $h = h(x)$? If not, try $h = h(y)$?

  If not, try $h = h(x,y) \longrightarrow$ possibly not helpful

Notes.

4 Hyperbolic functions

By the end of this section, you should be able to answer the following questions:

- What is the definition of the sinh and cosh functions?
- What is the definition of the inverse hyperbolic functions?
- What are the derivatives and anti-derivatives of these functions?
- How are hyperbolic functions used in the catenary problem?

4.1 Properties of hyperbolic functions

We define the functions $\cosh(x)$ and $\sinh(x)$ by

$$\begin{aligned}\cosh(x) &= \frac{e^x + e^{-x}}{2}, \\ \sinh(x) &= \frac{e^x - e^{-x}}{2}.\end{aligned}$$

We can check by direct calculation that

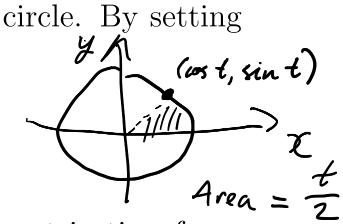
$$\cosh^2(x) - \sinh^2(x) = 1.$$

Compare this with the identity

$$\cos^2(\theta) + \sin^2(\theta) = 1 \tag{9}$$

for trig functions. The identity (9) allows us to parametrise a unit circle. By setting $x(t) = \cos(t)$, $y(t) = \sin(t)$, we have

$$\cos^2(t) + \sin^2(t) = x^2 + y^2 = 1,$$



which is the equation of the unit circle.

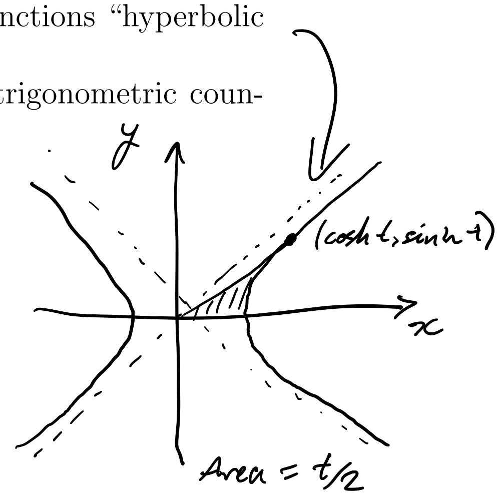
If we set $x(t) = \cosh(t)$ and $y(t) = \sinh(t)$, this gives a parametrisation for a hyperbola (only the right branch), since

$$\cosh^2(t) - \sinh^2(t) = x^2 - y^2 = 1,$$

which is the equation of a hyperbola. This is why we call these functions "hyperbolic functions".

These hyperbolic functions satisfy properties similar to their trigonometric counterparts. For example

$$\begin{aligned}\frac{d}{dx}(\cosh(x)) &= \frac{e^x - e^{-x}}{2} = \sinh(x), \\ \frac{d}{dx}(\sinh(x)) &= \frac{e^x + e^{-x}}{2} = \cosh(x).\end{aligned}$$

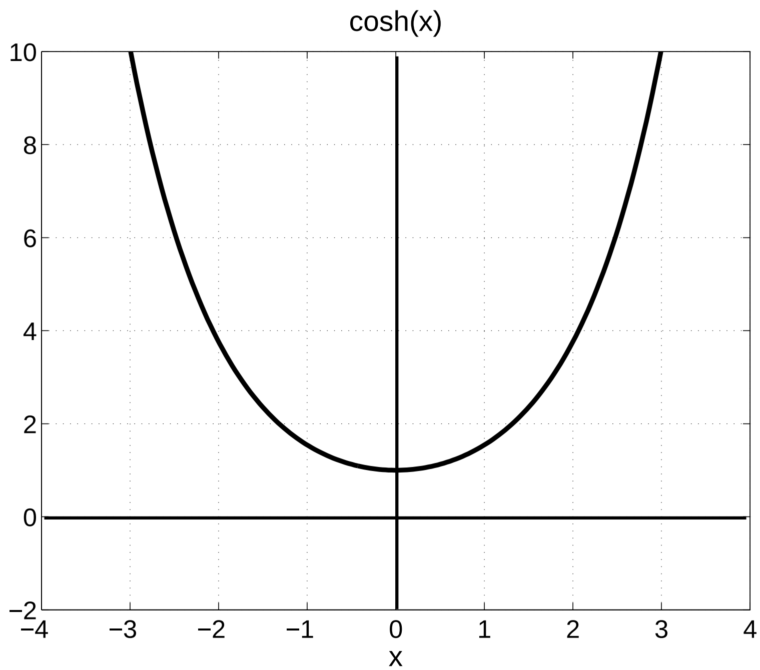


$\cosh(0) = 1$, $\cosh(x) \geq 1$, $\cosh(x)$ is an even function. $\to \cosh(-x) = \cosh(x)$
$\sinh(0) = 0$, $\sinh(x)$ is an odd function. $\to \sinh(-x) = -\sinh(x)$
We also define

$$
\begin{aligned}
\tanh(x) &= \frac{\sinh(x)}{\cosh(x)} = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \quad |\tanh(x)| < 1, \\
\coth(x) &= \frac{\cosh(x)}{\sinh(x)}.
\end{aligned}
$$

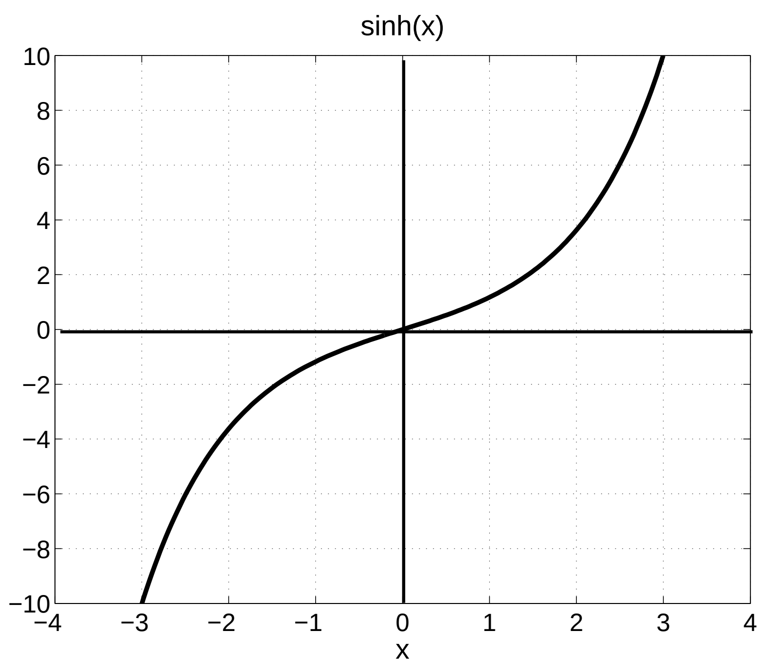
$$\operatorname{sech}(x) = \frac{1}{\cosh(x)}$$

$$\operatorname{cosech}(x) = \frac{1}{\sinh(x)}$$



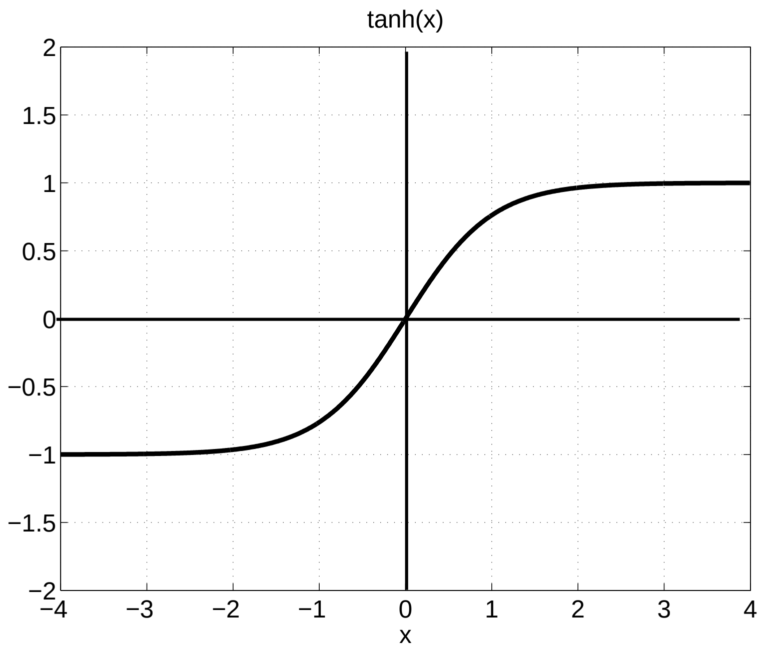
Domain: $\mathbb{R}$
Range: $[1, \infty)$

Figure 1: Graph of $\cosh(x)$



Domain: $\mathbb{R}$
Range: $\mathbb{R}$

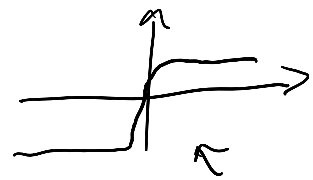
Figure 2: Graph of $\sinh(x)$



Domain: $\mathbb{R}$

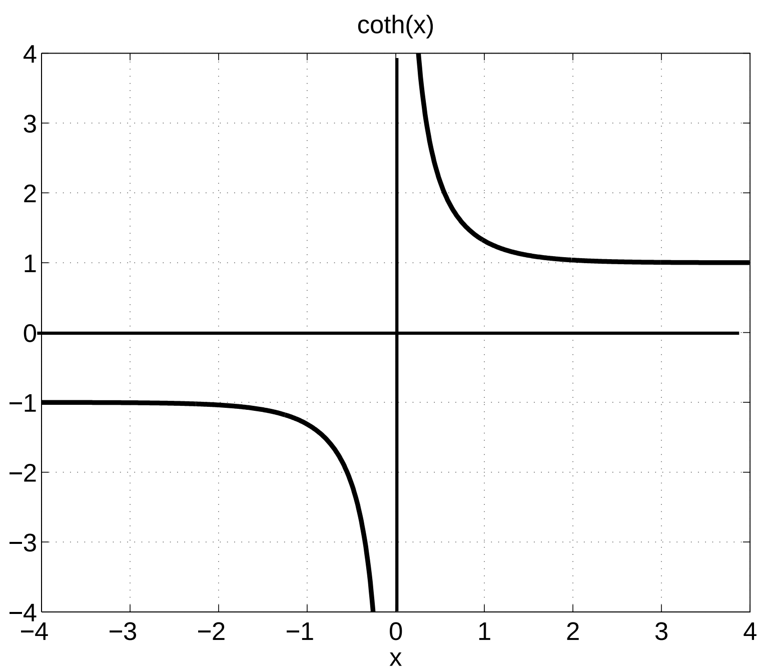
Range: $(-1, 1)$

for $\tanh(\alpha x)$, as α becomes large,



becomes steeper

Figure 3: Graph of $\tanh(x)$



Domain: $\mathbb{R} \setminus \{0\}$

Range: $\mathbb{R} \setminus \{0\}$

Figure 4: Graph of $\coth(x)$

4.2 Inverse hyperbolic functions

The inverse function of cosh is denoted arcosh. $\to \cosh^{-1}$

The inverse function of sinh is denoted arsinh. $\to \sinh^{-1}$

The inverse function of tanh is denoted artanh. $\to \tanh^{-1}$

$\left.\begin{matrix} \\ \\ \end{matrix}\right\}$ not the inverse! $\sinh^{-1} \neq \frac{1}{\sinh}$

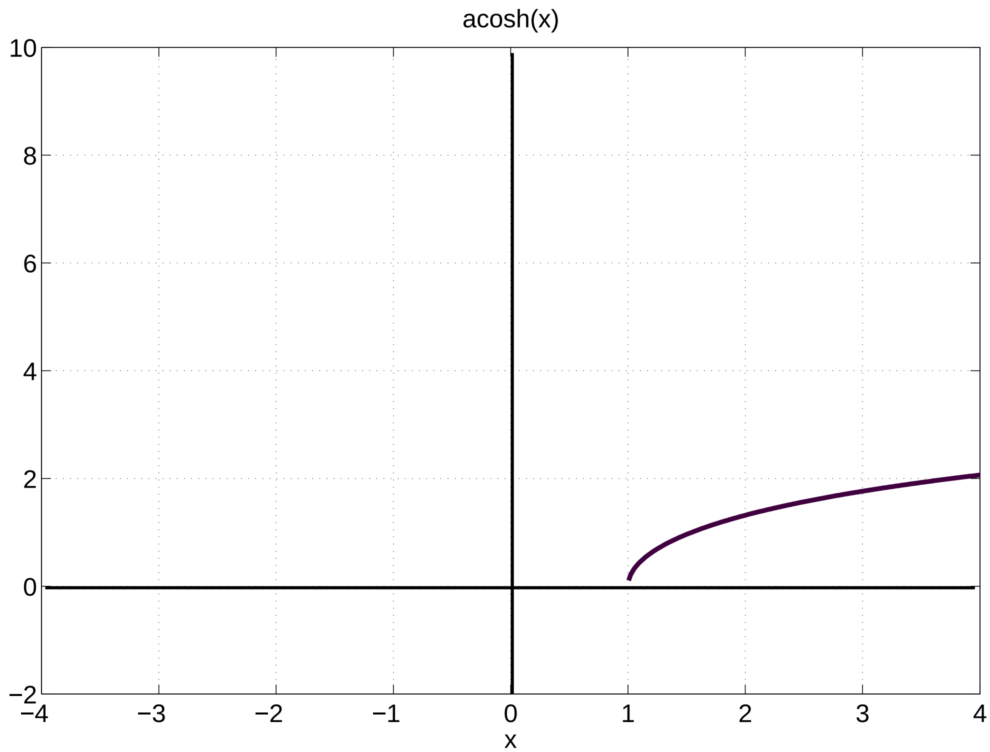


Figure 5: Graph of arcosh (x)

restrict domain of cosh to $[0,\infty)$

$\to$ range of $\cosh^{-1}$ is $[0,\infty)$

and domain of $\cosh^{-1}$ is $[1,\infty)$

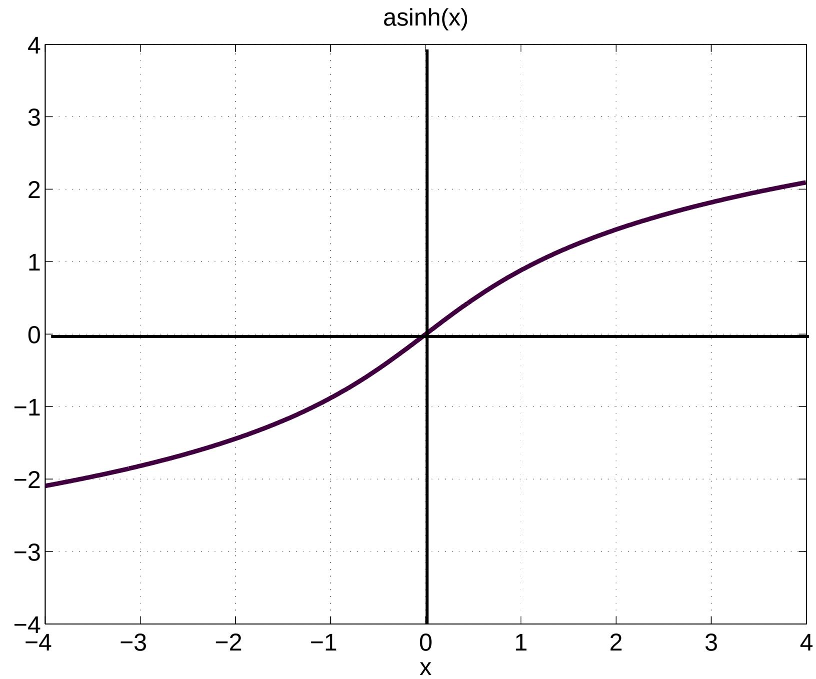


Figure 6: Graph of $\operatorname{arsinh}(x)$

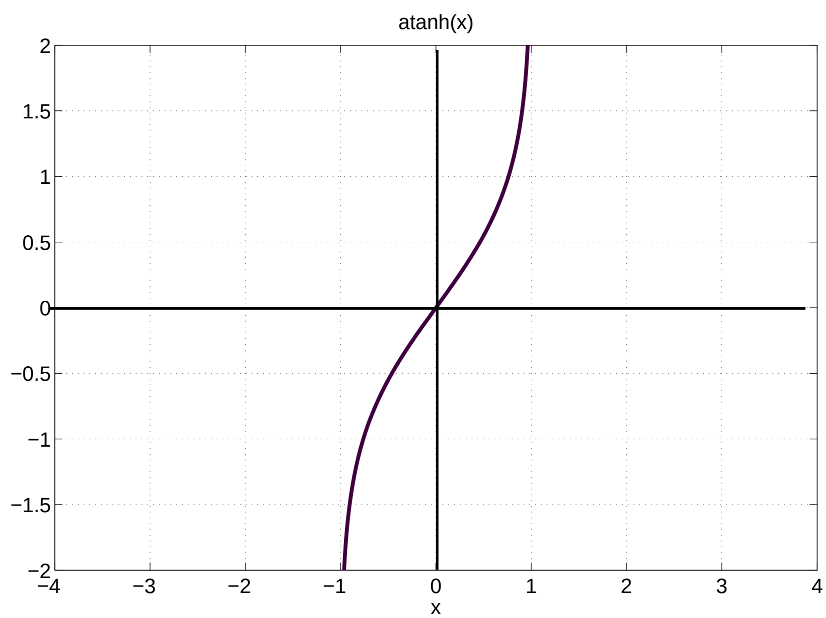


Figure 7: Graph of $\operatorname{artanh}(x)$

We have the following:

$$\int \frac{dx}{\sqrt{1+x^2}} = \operatorname{arsinh}(x) + c$$
$$\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{arcosh}(x) + c, \ x > 1.$$

4.2.1 Show that $\frac{d}{dx}(\operatorname{arsinh}(x)) = \frac{1}{\sqrt{1+x^2}}$

set $y = \sinh^{-1}(x) \Rightarrow x = \sinh(y)$ $\quad (*)$

Differentiate both sides of $(*)$ w.r.t. x:

$$\frac{d}{dx}(x) = \frac{d}{dx}(\sinh(y)) \qquad (\text{assume } y = y(x))$$

$$\Rightarrow 1 = \frac{d}{dy}(\sinh(y))\frac{dy}{dx}$$

$$= \cosh(y)\frac{dy}{dx}$$

$$\Rightarrow \frac{dy}{dx} = \frac{1}{\cosh(y)}$$

But $\cosh^2 y - \sinh^2 y = 1$

$$\Rightarrow \cosh y = \sqrt{1+\sinh^2 y}$$

but $x = \sinh y \Rightarrow x^2 = \sinh^2$ from $(*)$

$$\Rightarrow \cosh y = \sqrt{1+x^2}$$

$$\Rightarrow \frac{d}{dx}\sinh^{-1}(x) = \frac{1}{\sqrt{1+x^2}}$$

(a) (b)

4.2.2 Evaluate the integrals $\int \frac{dx}{\sqrt{1+x^2}}$ **and** $\int \frac{dx}{\sqrt{x^2-1}}$

(a) set $x = \sinh t \Rightarrow$ in the integral, $dx = \cosh t\, dt$ and $t = \sinh^{-1} x$

$$\Rightarrow \int \frac{dx}{\sqrt{1+x^2}} = \int \frac{\cosh t}{\sqrt{1+\sinh^2 t}}\, dt = \int \frac{\cosh t}{\sqrt{\cosh^2 t}}\, dt$$

$$= \int 1\, dt = t + C = \sinh^{-1} x + C$$

(b) Assume $x > 1$ (always be mindful of domain and range in substitutions)

set $x = \cosh t$, $t > 0$

$\Rightarrow$ in the integral, $dx = \sinh t\, dt$

$$\Rightarrow \int \frac{dx}{\sqrt{x^2-1}} = \int \frac{\sinh t\, dt}{\sqrt{\cosh^2 t - 1}} = \int \frac{\sinh t}{\sqrt{\sinh^2 t}}\, dt$$

$$= \int 1\, dt = t + C = \cosh^{-1} x + C$$

(for $x > 1$)

Convention: Square root always positive

$\sqrt{\sinh^2 t} = \sinh t$? only if $t \geq 0$

for $t < 0$, $\sinh t < 0 \Rightarrow \underbrace{\sqrt{\sinh^2 t} = -\sinh t}_{\text{for } t<0}$

4.2.3 Show that $\dfrac{d}{dx}\left(\textbf{artanh}\,(x)\right) = \dfrac{1}{1-x^2}, \quad |x| < 1$

consider $\displaystyle\int \frac{dx}{1-x^2} = \tan^{-1}(x) + C$

left as exercise → similar to 4.2.1

Using partial fractions, we can also evaluate the integral

$$\int \frac{dx}{1-x^2} = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) + C.$$

$$\frac{1}{1-x^2} = \frac{1}{(1+x)(1-x)} = \frac{1/2}{1+x} + \frac{1/2}{1-x}$$

In fact, we have the following identities

$$\begin{aligned}
\operatorname{artanh}(x) &= \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right), \\
\operatorname{arsinh}(x) &= \ln\left(x+\sqrt{x^2+1}\right), \\
\operatorname{arcosh}(x) &= \ln\left(x+\sqrt{x^2-1}\right).
\end{aligned}$$

4.2.4 Show that arsinh $(x) = \ln\left(x + \sqrt{x^2+1}\right)$

Set $y = \sinh^{-1}(x) \Rightarrow x = \sinh(y)$

$$\Rightarrow x = \frac{e^y - e^{-y}}{2}$$

$$\Rightarrow 2x = e^y - e^{-y}$$

multiplying through by $e^y \Rightarrow (e^y)^2 - 2xe^y - 1 = 0$

solve for e^y (quadratic in e^y)

$$\Rightarrow e^y = \frac{2x \pm \sqrt{4x^2 - 4(-1)}}{2}$$

$$= x \pm \sqrt{x^2+1}$$

Note $e^y > 0$, but $x - \sqrt{x^2+1} < 0$

$\therefore e^y = x + \sqrt{x^2+1}$

$\Rightarrow y = \ln\left(x + \sqrt{x^2+1}\right)$

Euler's formula $e^{\pm ix} = \cos x \pm i \sin x \quad (x \in \mathbb{R})$

$$\cosh(ix) = \frac{e^{ix} + e^{-ix}}{2} = \cos x$$

also $\sinh(ix) = i\sin(x)$

OR : $\cosh(x) = \cos(ix)$

$\sinh(x) = -i\sin(ix)$

4.3 Further reading: The catenary problem

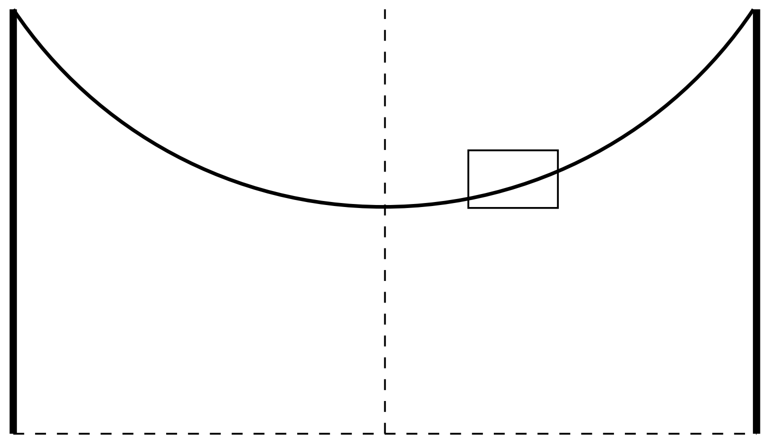


Figure 8: Profile of a heavy chain hanging under gravity.

One of the most famous problems where hyperbolic functions are used is in determining the profile of a heavy chain (of constant density ρ) suspended from two points of equal height (known as a catenary curve).

To derive the differential equation satisfied by the profile $y(x)$, we look at the forces acting on a small element of arc (inside the rectangular box in figure 8).

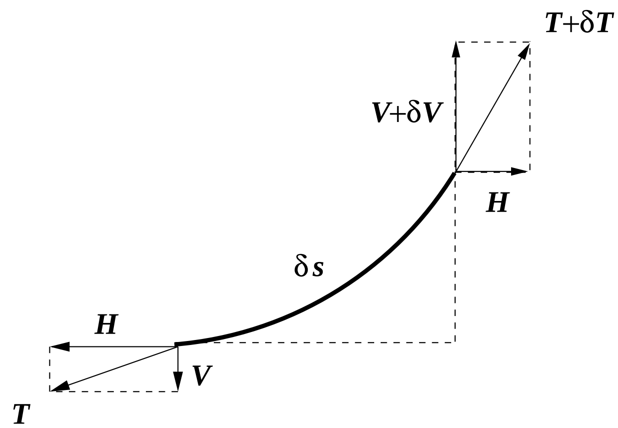


Figure 9: A small arc of heavy chain of length δs.

Let $T(x)$ be the tensile force in the chain with constant horizontal component H (since the load has no x component) and vertical component $V(x)$. In figure 9 the vertical components of the tensile force at either end of the arc are V and $V + \delta V$.

The mass of the arc will be $\rho(\delta s)$, so that the force due to gravity is $\rho g(\delta s)$.

The horizontal equilibrium is the trivial relation $H = H$, whereas the vertical equilibrium is the more informative

$$(V + \delta V) = V + \rho g(\delta s).$$

Dividing both sides by δx gives

$$\frac{\delta V}{\delta x} = \rho g \frac{\delta s}{\delta x}.$$

From geometry, we also have the approximation

$$\frac{\delta y}{\delta x} \approx \frac{V}{H}.$$

We also have the approximation to the arclength δs

$$(\delta s)^2 \approx (\delta x)^2 + (\delta y)^2 \quad \Rightarrow \quad \frac{\delta s}{\delta x} \approx \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}$$

Finally we take the limit $\delta x \to 0$ so that $\delta y \to 0$ and $\delta s \to 0$ simultaneously. We then have the following equations

$$\begin{aligned} \frac{dV}{dx} &= \rho g \frac{ds}{dx}, \\ V &= H \frac{dy}{dx}, \\ \frac{ds}{dx} &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \end{aligned}$$

Putting these equations together gives the ODE satisfied by the profile $y(x)$,

$$\frac{d^2 y}{dx^2} = \frac{\rho g}{H} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}.$$

Notes.

5 Linear second order nonhomogeneous ODEs, the Wronskian

By the end of this section, you should be able to answer the following questions:

- What is the form of the general solution to a linear, second order nonhomogeneous ODE?
- How do you define the Wronskian of an nth order linear second order homogeneous ODE?
- What is the relationship between the Wronskian and linearly independent solutions to a linear, second order homogeneous ODE?

ODEs can be split into two classes: linear and non-linear. Non-linear ODEs are generally very difficult to solve. Linear ODEs are simpler because their solutions have general properties which facilitate working with them. There are also well established methods for solving many linear ODEs of practical significance.

A second order ODE is called linear if it can be written in the form

$$y'' + p(x)y' + q(x)y = r(x). \tag{10}$$

Any second order ODE which cannot be written in this form is called non-linear. Note that y and its derivatives appear linearly and throughout we assume that p, q and r are continuous functions on some open interval I.

Over the next few sections we study linear second order ODEs. The motivation for studying second order ODEs is twofold. Firstly they have applications in mechanics and electric circuit theory, so anyone studying either of these fields will most likely come across second order ODEs. Secondly, the theory of linear second order ODEs is very similar to that of higher order linear ODEs, so that the transition to studying higher order linear ODEs would not require too many new ideas.

A great deal of this discussion can be found in the book *Elementary Differential Equations and Boundary Value Problems* by W.E. Boyce and R.C. DiPrima (ed. 10, Wiley, 2012).

5.1 Existence and uniqueness theorem

The key result for initial value problems related to second order linear ODEs is given as follows.

Consider the IVP

$$y'' + p(x)y' + q(x)y = r(x), \quad y(x_0) = y_0, \quad y'(x_0) = y_0',$$

where p, q and r are continuous functions on an open interval I that contains x_0. Then there is exactly one solution of the IVP which exists throughout I.

The proof of this theorem is beyond the scope of the course. Second order linear ODEs were introduced in MATH1052/MATH1072, and we first recall some important results and concepts, before looking more closely at certain aspects.

5.2 The superposition principle

Recall that if $r(x) = 0$ in equation (10), then we call the ODE *homogeneous*. If $r(x) \neq 0$, then the ODE is *nonhomogeneous*.

For any homogeneous linear equation, if y_1 and y_2 are solutions, so too is the linear combination $Ay_1 + By_2$. It is important to note that the superposition principle is not true for nonlinear equations and nonhomogeneous.

5.3 Linear independence

The functions $f_1, f_2, \ldots, f_n$ are said to be *linearly dependent on an interval I* if there is a set of constants $k_1, k_2, \ldots, k_n$, not all zero, such that

$$k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x) = 0$$

for all $x \in I$. The functions are said to be *linearly independent on I* if they are not linearly dependent there.

5.3.1 Example: linear independence of $g_1(x) = 1$, $g_2(x) = x$, $g_3(x) = x^2$ on the interval $-\infty < x < \infty$ $= I$

Set $k_1 + k_2 x + k_3 x^2 = 0$

for this equation to hold $\forall x$, then it must be true for three distinct values in I,

e.g. $x = 0, x = 1, x = 2$

$$\Rightarrow \quad k_1 = 0 \qquad (1)$$

$$k_1 + k_2 + k_3 = 0 \qquad (2)$$

$$k_1 + 2k_2 + 4k_3 = 0 \qquad (3)$$

(1), (2) $\Rightarrow k_2 = -k_3$

then, (3) $\Rightarrow k_3 = 0 \Rightarrow k_1 = k_2 = k_3 = 0$

$\Rightarrow g_1, g_2, g_3$ are linearly independent

5.4 General solution (homogeneous case), Wronskian of solutions

From the superposition principle, we can see that given two solutions, y_1 and y_2 of the ODE, we are able to construct an infinite family by taking the linear combination Ay_1+By_2. The question that arises is whether or not this family contains all solutions.

A first step is to understand if we can choose constants A and B to satisfy given initial conditions. Let f_1 and f_2 be two differentiable functions on the interval I. The *Wronskian* of f_1 and f_2 at $x \in I$ is defined as

$$W(f_1, f_2)(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix} = f_1(x)f_2'(x) - f_1'(x)f_2(x).$$

Now let y_1 and y_2 be solutions of the ODE

$$y'' + p(x)y' + q(x)y = 0.$$

Is it possible to find A and B so that the initial conditions $y(x_0) = y_0$ and $y'(x_0) = y_0'$ are satisfied by $y(x) = Ay_1(x) + By_2(x)$?

$$\left.\begin{aligned} y(x_0) &= Ay_1(x_0) + By_2(x_0) = y_0 \\ y'(x_0) &= Ay_1'(x_0) + By_2'(x_0) = y_0' \end{aligned}\right\} (*)$$

$$\Rightarrow \underbrace{\begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix}}_{M} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}$$

$$\det M = W(y_1, y_2)(x_0) = W$$

If $W \neq 0 \Rightarrow \begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{W}\begin{pmatrix} y_2'(x_0) & -y_2(x_0) \\ -y_1'(x_0) & y_1(x_0) \end{pmatrix}\begin{pmatrix} y_0 \\ y_0' \end{pmatrix}$

If $W = 0$, either no solution, or there is an infinite family of solutions i.e. there are many initial conditions that cannot be satisfied no matter how we choose A and B

We have the following results involving the Wronskian.

Suppose that y_1 and y_2 are two solutions of

$$y'' + p(x)y' + q(x)y = 0,$$

and consider the associated IVP with initial conditions

$$y(x_0) = y_0, \quad y'(x_0) = y'_0.$$

It is always possible to choose A and B so that

$$y = Ay_1(x) + By_2(x)$$

satisfies the IVP if and only if

$$W(y_1, y_2)(x_0) \neq 0.$$

It is the next result that allows us to use the phrase *general solution* that was introduced in MATH1052/MATH1072. The proof makes use of the previous result.

Let y_1 and y_2 be two solutions of

$$y'' + p(x)y' + q(x)y = 0.$$

The family of solutions

$$y = Ay_1(x) + By_2(x)$$

with arbitrary coefficients A, B includes every solution of this ODE if and only if there is a value x_0 such that $W(y_1, y_2)(x_0) \neq 0$. In such a case we call y_1 and y_2 a *fundamental set of solutions.*

Proof: $A \Leftrightarrow B$

Assume $\exists x_0$ s.t. $W(y_1,y_2)(x_0) \neq 0$ and let ϕ be any solution of the ODE. Can we find A,B s.t. $\phi = Ay_1 + By_2$?
Set $y_0 = \phi(x_0)$, $y_0' = \phi'(x_0)$
Previous result $\Rightarrow$ can choose A,B so that $y = Ay_1 + By_2$ solves the IVP. Uniqueness implies this must be ϕ (which is arbitrary)
This shows $B \Rightarrow A$, now we need $A \Rightarrow B$ (by showing $\neg B \Rightarrow \neg A$)

Assume there is no x_0 s.t. $W(y_1,y_2)(x_0) \neq 0$, i.e. $W = 0$ $\forall x_0$
Previous result $\Rightarrow \exists y_0, y_0'$ s.t. there is no solution of the lower system $(*)$ on p45, but $\exists$ a unique solution of IVP
$\Rightarrow \phi$ is not included in $y = Ay_1 + By_2$ $\rightarrow$ result

5.5 Abel's Theorem

Let y_1 and y_2 be solutions of

$$y'' + p(x)y' + q(x)y = 0,$$

where p and q are continuous on an open interval I. The Wronskian is given by

$$W(y_1, y_2)(x) = Ce^{-\int p(x)\ dx},$$

where C is a constant that depends on y_1 and y_2, but not x. Moreover $W(y_1, y_2)(x)$ either is zero for all $x \in I$, or else is never zero in I.

Set $W = y_1 y_2' - y_1' y_2$

$\Rightarrow W' = y_1' y_2' + y_1 y_2'' - y_1'' y_2 - y_1' y_2'$

$= y_1 y_2'' - y_1'' y_2$

$\Rightarrow W' + pW = y_1 y_2'' - y_1'' y_2 + p y_1 y_2' - p y_1' y_2 + q y_1 y_2 - q y_1 y_2$

$= y_1 \underbrace{(y_2'' + p y_2' + q y_2)}_{=0} - y_2 \underbrace{(y_1'' + p y_1' + q y_1)}_{=0}$

$= 0$

First order linear ODE in W: $W' + pW = 0$

$\Rightarrow$ general solution is $W = Ce^{-\int p\, dx}$

5.6 Nonhomogenous linear ODEs

Now we consider equations of the form

$$y'' + p(x)y' + q(x)y = r(x), \quad r(x) \neq 0. \tag{11}$$

You should know from MATH1052/MATH1072 that the general solution on an open interval I is of the form

$$y = y_H + y_P,$$

where y_H is the general solution of the homogeneous equation (with $r(x) = 0$) on I and y_P is a particular solution of (11) on I containing no arbitrary constants.

One strategy for finding the general solution is to first find y_H, then find y_P. See if you can work through the following example (review from MATH1052/MATH1072).

5.6.1 Example: $y'' - 2y' + y = e^x + x$

for y_H: $\lambda^2 - 2\lambda + 1 = 0$

$(y_H = e^{\lambda x})$ $\Rightarrow (\lambda - 1)^2 = 0 \Rightarrow \lambda = 1$

$y_H = Ae^x + Bxe^x$ (set $y_1 = e^x$, $y_2 = ue^x$)

For y_P: guess $y_P = \underbrace{ax^2e^x}_{\text{contributes to } e^x \text{ on RHS}} + \underbrace{bx + c}_{\text{contributes to } x \text{ on RHS}}$ ← method of undetermined coefficients

sub guess into LHS and set = RHS

(check $a = \frac{1}{2}, b = 1, c = 2$)

$y = Ae^x + Bxe^x + \frac{1}{2}x^2e^x + x + 2$

5.7 Further reading: The method of undetermined coefficients

In what follows, we determine a solution to the homogenous equation, and then *try* a form (with *undetermined coefficients*) for the particular solution which looks like it will result in the function on the right hand side.

The method of undetermined coefficients, as presented here, only works for the constant coefficient case:

$$y'' + ay' + by = r(x),$$

and $r(x)$ contains exponentials, polynomials, sines and cosines, or sums and certain products of these functions.

Choose for y_P a form similar to $r(x)$, involving unknown coefficients. The coefficients are then determined by substituting y_P into the ODE.

$r_i(x)$	$g_i(x)$	$r_i(x)$	$g_i(x)$
$ke^{\gamma x}$	$ae^{\gamma x}$	$k\cos\omega x$, $k\sin\omega x$	$a\cos\omega x + b\sin\omega x$
$\sum_{i=0}^{N} k_i x^i$, $N = 0, 1, 2, \ldots$	$\sum_{i=0}^{N} a_i x^i$	$ke^{\alpha x}\cos\omega x$, $ke^{\alpha x}\sin\omega x$	$e^{\alpha x}(a\cos\omega x + b\sin\omega x)$

We follow these basic steps.

1. Find a solution y_H to the corresponding homogeneous equation.
2. For $r(x) = r_1(x) + r_2(x) + \ldots + r_n(x)$, we first make a guess $g(x) = g_1(x) + g_2(x) + \ldots + g_n(x)$ for y_P, where the $g_i(x)$ correspond to the $r_i(x)$ entries in the table above.
3. If a term $g_i(x)$ appears in y_H, replace $g_i(x)$ in the initial guess by $xg_i(x)$.
4. If any of the $xg_i(x)$ from step 3 appear in y_H, replace $xg_i(x)$ by $x^2 g_i(x)$.
5. Substitute the modified guess $g(x)$ into the left hand side of the ODE and equate coefficients on both sides. Once you have worked out the coefficients, the guess $g(x)$ becomes y_P.

Notes.

This section for solving $y'' + p(x)y' + q(x)y = r(x)$
General solution $y = y_H + y_P$

6 Variation of parameters

By the end of this section, you should be able to answer the following questions:

- Under what conditions does the method work?
- What functions need to be determined first before using the method?
- How do you use the variation of parameters method to solve a nonhomogeneous linear second order ODE?

The method of undetermined coefficients is very easy to apply, but only works for constant coefficients with certain $r(x)$. For other cases, the variation of parameters works well. The process is the following:

- Solve $y'' + p(x)y' + q(x)y = 0$, obtain a fundamental set of solutions y_1, y_2 and calculate the Wronskian, $W(y_1, y_2)(x) \equiv W$. $= y_1 y_2' - y_1' y_2$ (see CH5)
- Set $y_P = u(x)y_1(x) + v(x)y_2(x)$ and substitute into the ODE. We also impose the condition $u'y_1 + v'y_2 = 0$. We have the freedom to impose this extra arbitrary condition because we have two functions (u and v) and only one equation they need to satisfy arising from the ODE.
- We obtain
$$u(x) = -\int \frac{y_2 r}{W} dx, \quad v(x) = \int \frac{y_1 r}{W} dx.$$

This approach is a variant of the method of Reduction of Order, which prescribes that we take a solution, say y_1 of the associated homogeneous equation and seek a particular solution of the form $y_p = U(x)y_1$.

6.1 Derive the formulae for $u(x)$ and $v(x)$ in the variation of parameters

General solution: $y = y_H + y_P$

We know $\boxed{y_H = Ay_1 + By_2}$ $\Rightarrow$ i.e. y_1 & y_2 satisfy $y'' + py' + qy = 0$

Set $y_P = uy_1 + vy_2$

$$\Rightarrow \quad y_P'' + py_P' + qy_P = q(uy_1 + vy_2) + p(u'y_1 + uy_1' + v'y_2 + vy_2') + (u''y_1 + 2u'y_1' + uy_1'' + v''y_2 + 2v'y_2' + vy_2'')$$

$$= u''y_1 + u'(2y_1 + py_1) + u\underbrace{(y_1'' + py_1' + qy_1)}_{=0} + v''y_2 + v'(2y_2 + py_2) + v\underbrace{(y_2'' + py_2' + qy_2)}_{=0}$$

Impose $u'y_1 + v'y_2 = 0$ ①

$$\Rightarrow \quad u''y_1 + u'y_1 + v''y_2 + v'y_2 = 0$$

$$\Rightarrow \quad y_P'' + py_P' + qy_P = u'y_1' + v'y_2' = \text{RHS} = r \qquad ②$$

① & ② $\Rightarrow$

$$A = \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 \\ r \end{pmatrix}$$

Note $W = y_1y_2' - y_1'y_2 = \det(A)$

$$\Rightarrow \begin{pmatrix} u' \\ v' \end{pmatrix} = \frac{1}{W}\begin{pmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{pmatrix}\begin{pmatrix} 0 \\ r \end{pmatrix}$$

row 1 $\Rightarrow u' = \dfrac{-y_2 r}{W} \quad \Rightarrow \quad u = \displaystyle\int \frac{-y_2 r}{W}\,dx$

row 2 $\Rightarrow v' = \dfrac{y_1 r}{W} \quad \Rightarrow \quad v = \displaystyle\int \frac{y_1 r}{W}\,dx$

6.2 Example: $y'' - 4y' + 5y = 2e^{2x}/\sin x$

For y_H: $\lambda^2 - 4\lambda + 5 = 0$ $\quad (y_H = e^{\lambda x})$

$$\lambda^2 - 4\lambda + 4 + 1 = 0 \Rightarrow (\lambda - 2)^2 + 1 = 0 \Rightarrow \lambda = 2 \pm i$$

$$\Rightarrow y_H = A\underbrace{e^{2x}\cos x}_{y_1} + B\underbrace{e^{2x}\sin x}_{y_2}$$

For y_P: use variation of parameters

set $y_P = uy_1 + vy_2$

where $u = -\int \frac{y_2 r}{W}\,dx$, $v = \int \frac{y_1 r}{W}\,dx$

$$W = y_1 y_2' - y_1' y_2 = (e^{2x}\cos x)(2e^{2x}\sin x + e^{2x}\cos x) - (2e^{2x}\cos x - e^{2x}\sin x)(e^{2x}\sin x)$$
$$= e^{4x}(\cos^2 x + \sin^2 x) = e^{4x}$$

$$\Rightarrow u = -\int e^{2x}\sin x \cdot \frac{2e^{2x}}{\sin x} \cdot \frac{1}{e^{4x}}\,dx = -2\int 1\,dx = -2x + c_1$$

$$v = \int e^{2x}\cos x \cdot \frac{2e^{2x}}{\sin x} \cdot \frac{1}{e^{4x}}\,dx = 2\int \frac{\cos x}{\sin x}\,dx$$
$$= 2\ln|\sin x| + c_2$$

$$\Rightarrow y_P = -2xe^{2x}\cos x + 2\ln|\sin x|e^{2x}\sin x$$

$\Rightarrow$ general solution is $y = y_H + y_P$

contains the arbitrary constants

6.3 Repeat the previous example using reduction of order

Set $y_p = Ue^{2x}\sin x \quad (U(x)y_2)$

$y_p' = U'e^{2x}\sin x + 2Ue^{2x}\sin x + Ue^{2x}\cos x$

$y_p'' =$ exercise ☺

& substitute into

$$y_p'' - 4y_p' + 5y_p = \frac{2e^{2x}}{\sin x}$$

6.4 Summary of ODE techniques and types of equations you should know

- First order, directly integrable
- First order, separable
- First order, linear, integrating factor
- First order existence and uniqueness criteria
- First order, exact
- Second order homogeneous, linear, constant coefficients
- Second order nonhomogeneous, constant coefficients, method of undetermined coefficients for certain cases
- Second order nonhomogeneous, variation of parameters.
- Reduction of order

Notes.

7 Vector spaces

By the end of this section, you should be able to understand:

- The definition of a vector space.
- Many new examples of vector spaces.
- The significance of a basis.
- How to find the transition matrix from one basis to another.

Notation. For the next set of lectures on linear algebra, $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$. Thus, if a statement holds for or applies to both number sets, we may simply state it for $\mathbb{F}$. Elements of $\mathbb{F}$ are often called *scalars*.

7.1 Definition

Let V be a nonempty set on which are defined operations "+" (called addition) and "·" (called scalar multiplication). V is a **vector space** (over $\mathbb{F}$) if the following hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $k, \ell \in \mathbb{F}$:

* **(V1)** $\mathbf{u} + \mathbf{v} \in V$ (closure)

(V2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (additive commutativity)

(V3) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (additive associativity)

(V4) $\exists\, \mathbf{0} \in V$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (zero vector, or additive identity)

(V5) For each $\mathbf{u} \in V$, $\exists\, (-\mathbf{u}) \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ (additive inverse)

* **(V6)** $k \cdot \mathbf{u} \in V$

(V7) $k \cdot (\mathbf{u} + \mathbf{v}) = k \cdot \mathbf{u} + k \cdot \mathbf{v}$ (multiplicative-additive distributivity)

(V8) $(k + \ell) \cdot \mathbf{u} = k \cdot \mathbf{u} + \ell \cdot \mathbf{u}$ (additive-multiplicative distributivity)

(V9) $k \cdot (\ell \cdot \mathbf{u}) = (k\ell) \cdot \mathbf{u}$ (multiplicative-multiplicative distributivity)

(V10) $1 \cdot \mathbf{u} = \mathbf{u}$ (multiplicative identity)

The scalar multiplication symbol is often omitted. Elements of a vector space are usually called vectors.

Steps to verify a vector space V:

1. identify the set V
2. identify the operations of addition and scalar multiplication
3. Verify closure under these operations (check V1, V6)
4. Check remaining axioms.

7.2 Example: $\mathbb{F}^n$ – set of n-tuples

e.g. $\mathbb{R}^n = \{(u_1, u_2, \dots, u_n) \mid u_1, u_2, \dots, u_n \in \mathbb{R}\}$

$\underline{u} + \underline{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

$k\underline{u} = (ku_1, ku_2, \dots, ku_n)$ where $k \in \mathbb{R}$

7.3 Example: $M_{m,n}(\mathbb{F})$ – set of $m \times n$ matrices

$$M_{m,n}(\mathbb{R}) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \;\middle|\; \begin{array}{l} a_{ij} \in \mathbb{R}, \\ 1 \le i \le m, \\ 1 \le j \le n \end{array} \right\}$$

with the usual operations of addition and scalar multiplication of matrices

7.4 Example: $C[a,b]$ – set of continuous real-valued functions on $[a,b]$

$f, g \in C[a,b]$ often represented by $f(x)$ & $g(x)$ with operations

$$(f+g)(x) = f(x) + g(x)$$

for $k \in \mathbb{R}$, $(kf)(x) = kf(x)$

7.5 Example: $P_n(\mathbb{F})$ – set of polynomials of degree at most n

e.g. $p : \mathbb{R} \to \mathbb{R}$, $x \longmapsto p(x) = p_0 + p_1 x + \dots + p_n x^n$

where $p_0, p_1, \dots, p_n \in \mathbb{R}$

often use $p(x)$ to represent $p \in P_n(\mathbb{R})$

& same operations as in (7.4)

7.6 Example: Set of solutions to a homogeneous linear ODE

e.g. $y'' + p(x)y' + q(x)y = 0 \qquad (y = y(x))$

If y_1, y_2 are solutions, then so are $y_1 + y_2$, ky_1, ky_2 with $k \in \mathbb{R}$

(operations as in 7.4)

Read this page carefully!

7.7 Familiar concepts in linear algebra

Here we lists several concepts with which you should already be familiar.

- **Linear combination**

  For $v_1, v_2, \ldots, v_n \in V$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$, we call

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

  a *linear combination* of the vectors $v_1, v_2, \ldots, v_n$.

- **Linear independence**

  A non-empty set of vectors $S = \{v_1, v_2, \ldots, v_n\}$ in V is said to be *linearly dependent* if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ not all zero such that

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0. \leftarrow \text{zero vector}$$

  Otherwise, S is called *linearly independent*, i.e. S is linearly independent if

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

- **Subspace**

  A subset $W \subseteq V$ is called a *subspace* if W is also a real vector space with the same addition and scalar multiplication. In particular, W is required to close under addition and scalar multiplication.

- **Span**

  The *span* of a non-empty set of vectors $S = \{v_1, v_2, \ldots, v_n\}$ in V is the set of all linear combinations of vectors in S, denoted $\text{span}(S)$. The set $\text{span}(S)$ is a subspace of V.

7.8 Basis

Let $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in the vector space V. β is a **basis** for V if

(B1) β is linearly independent;

(B2) β spans V.

Note that the notion of a basis is only defined here for finite sets. A nonzero vector space is **finite-dimensional** if it contains a finite set of vectors that forms a basis. If no such set exists, the vector space is **infinite-dimensional**.

Let V be a finite-dimensional vector space. The number of vectors in any basis for V is the same, and this number is known as the **dimension** of V.

An **ordered basis** for a vector space is a basis endowed with a specific order. For some vector spaces, there is a canonical ordered basis, called a **standard basis**.

(standard bases)

e.g. $\mathbb{R}^3$: $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$P_3(\mathbb{R})$: $\{1, x, x^2, x^3\}$

$M_{2,3}(\mathbb{R})$: $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$

7.9 Decomposition theorem

Let $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in the vector space V. Then, β is a basis for V iff each $\mathbf{w} \in V$ can be uniquely expressed as a linear combination of vectors in β.

Proof: "$\Leftarrow$" Since every $\vec{w} \in V$ is a linear combination of vectors in β, we have $\text{span}(\beta) = V$ (B2)

Consider $\vec{w} = \vec{0} \Rightarrow \vec{0} = \sum_{i=1}^{n} a_i \vec{v}_i$

Since the linear combination is unique, it must be the one where all $a_i = 0$

$\Rightarrow \beta$ is linearly independent (B1)

"$\Rightarrow$" Suppose β is a basis & consider

$\vec{w} = \sum_{i=1}^{n} a_i \vec{v}_i$ & $\vec{w} = \sum_{i=1}^{n} b_i \vec{v}_i$

Then $\vec{0} = \vec{w} + (-\vec{w}) = \sum_{i=1}^{n} (a_i - b_i) \vec{v}_i$

Since β is a basis, it's also linearly independent $\Rightarrow a_i - b_i = 0 \quad \forall i$

i.e. $a_i = b_i \quad \forall i$

$\Rightarrow$ expression for $\vec{w}$ is unique.

7.10 Transition matrix

Let $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an ordered basis for the vector space V. For $\mathbf{u} \in V$, let $a_1, \ldots, a_n$ be (the unique) scalars such that

$$\mathbf{u} = \sum_{i=1}^{n} a_i \mathbf{v}_i.$$

The **coordinate vector** of $\mathbf{u}$ relative to β is given by

$$[\mathbf{u}]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Another common notation for this is $[\mathbf{u}]^\beta$.

Let β' be another ordered basis for V. The coordinate vector of $\mathbf{u}$ relative to β' is thus denoted by $[\mathbf{u}]_{\beta'}$. The **transition matrix** from β to β', denoted by $P_{\beta\to\beta'}$, relates the two coordinate vectors of $\mathbf{u}$ as

$$\boxed{[\mathbf{u}]_{\beta'} = P_{\beta\to\beta'}[\mathbf{u}]_\beta. \qquad \text{Definition}}$$

If β'' is yet another ordered basis for V, then

$$P_{\beta'\to\beta''}P_{\beta\to\beta'} = P_{\beta\to\beta''} \qquad \Longrightarrow \qquad P_{\beta'\to\beta}P_{\beta\to\beta'} = P_{\beta\to\beta} = I,$$

where I is the $n \times n$ identity matrix.

To illustrate, let us consider the two ordered bases $\beta = \{1, x\}$ and $\beta' = \{1 + x, 2x\}$ for $P_1(\mathbb{F})$. As the vector (or polynomial) $\mathbf{u} = a + bx$ also can be written as

$$\mathbf{u} = a(1 + x) + \tfrac{1}{2}(b - a)(2x),$$

we have

$$[\mathbf{u}]_\beta = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{and} \qquad [\mathbf{u}]_{\beta'} = \begin{pmatrix} a \\ \frac{1}{2}(b-a) \end{pmatrix}.$$

The corresponding transition matrix $P_{\beta\to\beta'}$ is given as follows:

$$[\vec{u}]_{\beta'} = P_{\beta\to\beta'}[\vec{u}]_\beta$$

$$\begin{pmatrix} a \\ \frac{1}{2}(b-a) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}}_{P_{\beta\to\beta'}} \begin{pmatrix} a \\ b \end{pmatrix}$$

In general, $\beta = \{\vec{v}_1, \ldots, \vec{v}_n\}$ & $\beta' = \{\vec{v}_1{}', \ldots, \vec{v}_n{}'\}$

Then, $P_{\beta\to\beta'} = \left([\vec{v}_1]_{\beta'} \,\middle|\, [\vec{v}_2]_{\beta'} \,\middle|\, \ldots \,\middle|\, [\vec{v}_3]_{\beta'} \right)$

Notes.

8 Real inner product spaces

By the end of this section, you should understand the definition of inner product space, and be aware of many new examples.

Every vector space in this section is real.

8.1 Dot product (or the Euclidean Inner Product)

The familiar **dot product**, $\mathbf{u} \cdot \mathbf{v}$, of the two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ is given by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3,$$

where $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$. This is readily generalised to $\mathbb{R}^n$ by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \ldots + u_n v_n, \quad \text{where} \quad \mathbf{u} = (u_1, \ldots, u_n),\ \mathbf{v} = (v_1, \ldots, v_n).$$

This dot product has the following key properties:

$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$	$\vec{u} \cdot \vec{u} \geq 0$
$(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$	$\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$
$(k\vec{u}) \cdot \vec{v} = k(\vec{u} \cdot \vec{v})$	

8.2 Inner product

Inspired by the dot product on $\mathbb{R}^n$, we define a so-called inner product on a general real vector space by elevating the key properties of the dot product to axioms.

Defn

Accordingly, an **inner product** on V is a function that takes each ordered pair $(\mathbf{u}, \mathbf{v})$ of elements of V to a real number, denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$, such that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $k \in \mathbb{R}$:

(I1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(I2) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

(I3) $\langle k \cdot \mathbf{u}, \mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$

(I4) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$

(I5) $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$ (where $\mathbf{0}$ is the unique zero vector)

Defn

A vector space with an inner product associated to it is called an **inner product space**. As we are assuming that all vector spaces are real in this section, we have thus introduced the notion of a *real* inner product space. One can also define inner products on complex vector spaces, thereby introducing *complex* inner product spaces, but they are beyond the scope of these lectures.

8.3 Example: Weighted dot product (on $\mathbb{R}^n$)

Let $\vec{u}, \vec{v} \in \mathbb{R}^n$ be given by

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}; \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

let $w_1, w_2, \ldots, w_n > 0$, $u_i, v_i, w_i \in \mathbb{R}$

Then define inner product by

$$\langle \vec{u}, \vec{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \ldots + w_n u_n v_n$$

(check that the five axioms are satisfied)

Note: The scalars $w_1, \ldots, w_n$ are called <u>weights</u>

Note that $w_1 = w_2 = \ldots = w_n = 1$ gives the Euclidean inner product

8.4 Example: Inner product generated by a matrix (on $\mathbb{R}^n$)

Let $\vec{u}, \vec{v} \in \mathbb{R}^n$, let $A \in M_{n,n}(\mathbb{R})$ be invertible. Then,

$$\boxed{\langle \vec{u}, \vec{v} \rangle = (A\vec{u}) \cdot (A\vec{v})}$$

regular dot product

defines an inner product

Note that $(A\vec{u}) \cdot (A\vec{v}) = (A\vec{u})^T A\vec{v}$ since $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$

$$= \vec{u}^T A^T A \vec{v}$$

So the weighted dot product corresponds to

$$A = \begin{pmatrix} \sqrt{w_1} & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{w_n} \end{pmatrix}$$

8.5 Example: Inner product on $M_{n,n}(\mathbb{R})$

$A, B \in M_{n,n}(\mathbb{R})$

$\boxed{\langle A, B \rangle = tr(B^T A)}$ is an inner product

Defn where the trace is: $tr\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix} = a_{11} + a_{22} + \ldots + a_{nn}$ i.e. sum of diagonal entries

e.g. $M_{2,2}(\mathbb{R})$, $U = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$, $V = \begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix}$

$$\langle U, V \rangle = tr\left(\begin{pmatrix} v_1 & v_3 \\ v_2 & v_4 \end{pmatrix} \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} \right)$$

$$= tr\left(\begin{pmatrix} u_1 v_1 + u_3 v_3 & * \\ * & u_2 v_2 + u_4 v_4 \end{pmatrix} \right)$$

$$= u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

standard basis; $\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$

$$U = u_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + u_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + u_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow [U]_\beta = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \quad [V]_\beta = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$$

$$\langle U, V \rangle = [U]_\beta \cdot [V]_\beta$$

8.6 Example: Standard inner product on $P_n(\mathbb{R})$

$p, q \in P_n(\mathbb{R})$ write $p(x) = p_0 + p_1 x + \dots + p_n x^n$

$q(x) = q_0 + q_1 x + \dots + q_n x^n$

$\boxed{\langle p, q \rangle = p_0 q_0 + p_1 q_1 + \dots + p_n q_n}$ is an inner product

(consider $\beta = \{1, x, \dots, x^n\}$)

8.7 Example: Evaluation inner product on $P_n(\mathbb{R})$

Let $x_0, x_1, \dots, x_n \in \mathbb{R}$ be fixed and distinct

Let $p, q \in P_n(\mathbb{R})$, then

$\boxed{\langle p, q \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \dots + p(x_n)q(x_n)}$

is an inner product

(I1) → (I3) trivial

(I4) $\langle p, p \rangle = \left(p(x_0)^2 + \dots + p(x_n)^2\right) \geq 0$

(I5) Equality to 0 requires $p(x_0) = p(x_1) = \dots = p(x_n) = 0$

Since a non-zero polynomial of degree n has at most n distinct roots, the polynomial p must be the zero polynomial and

$\langle p, p \rangle = 0$ iff $p = 0$

8.8 Example: Inner product on $C[a,b]$

Let $f, g \in C[a,b]$. Then

$\boxed{\langle f, g \rangle = \int_a^b f(x) g(x)\, dx}$ defines an inner product space

Regarding (I4), $\langle f, f \rangle = \int_a^b (f(x))^2 dx \geq 0$

since $(f(x))^2 \geq$ and f is continuous on the interval by definition

(I5) Equality to 0 requires $f(x) = 0$ on $[a,b]$

i.e. $f = 0$

Notes.

9 Solving linear equations

By the end of this section, you should be able to answer the following questions:

- How do you use Gaussian elimination to find a row echelon form of a matrix?
- What are the three cases for solutions to systems of linear equations?
- How do you solve a system of linear equations?
- What is the rank of a matrix?, column space, row space

Say we have m linear equations in n variables:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}$$

We can write these equations in matrix form: $A\boldsymbol{x} = \boldsymbol{b}$.

$A = [a_{ij}]$ is the $m \times n$ coefficient matrix.

$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is the column vector of unknowns, and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ is the column vector of the right hand side.

Note: a_{ij}, $b_j \in \mathbb{R}$ or $\mathbb{C}$.

9.1 Gaussian Elimination

To solve $A\boldsymbol{x} = \boldsymbol{b}$:
write *augmented matrix*: $[A|\boldsymbol{b}]$.

1. Find the left-most non-zero column, say column j.
2. Interchange top row with another row if necessary, so top element of column j is non-zero. (The **pivot**.)
3. Subtract multiples of row 1 from all other rows so all entries in column j below the top are then 0.
4. Cover top row; repeat **1** above on rest of rows.
Continue until all rows are covered, or until only $00\ldots0$ rows remain.

The result is a triangular system, easily solved by *back substitution*: solve the last equation first, then 2nd last equation and so on.

9.1.1 Example

Use Gaussian elimination to solve:

$$\begin{aligned} x_3 - \quad x_4 &= 2 \\ -9x_1 - 2x_2 + 6x_3 - 12x_4 &= -7 \\ 3x_1 + \ x_2 - 2x_3 + \ 4x_4 &= 2 \\ 2x_3 \qquad\quad &= 6 \end{aligned}$$

$$\left(\begin{array}{cccc|c} 0 & 0 & 1 & -1 & 2 \\ -9 & -2 & 6 & -12 & -7 \\ 3 & 1 & -2 & 4 & 2 \\ 0 & 0 & 2 & 0 & 6 \end{array}\right) \qquad R_1 \Leftrightarrow R_3$$

$$\to \left(\begin{array}{cccc|c} 3 & 1 & -2 & 4 & 2 \\ -9 & -2 & 6 & -12 & -7 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 2 & 0 & 6 \end{array}\right) \qquad R_2 \to R_2 + 3R_1$$

$$\to \left(\begin{array}{cccc|c} 3 & 1 & -2 & 4 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 2 & 0 & 6 \end{array}\right) \qquad R_4 \to R_4 - 2R_3$$

$$\to \left(\begin{array}{cccc|c} 3 & 1 & -2 & 4 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 2 & 2 \end{array}\right)$$

R_4: $2x_4 = 2 \Rightarrow x_4 = 1$

R_3: $x_3 - x_4 = 2 \Rightarrow x_3 = 3$

R_2: $x_2 = -1$

R_1: $3x_1 + x_2 - 2x_3 + 4x_4 = 2$

$\Rightarrow 3x_1 - 1 - 6 + 4 = 2$

$\Rightarrow x_1 = \frac{5}{3}$

9.1.2 Definition (row echelon form)

A matrix is in *row echelon form* (r.e.f.) if each row after the first starts with *more* zeros than the previous row (or else rows at bottom of matrix are all zeros).

The Gauss algorithm converts any matrix to one in row echelon form. The 2 matrices are *equivalent*, that is, they have the same solution set.

9.1.3 Elementary row operations

1. $r_i \leftrightarrow r_j$: swap rows i and j.
2. $r_i \to r_i - cr_j$: replace row i with (row i minus c times row j).
3. $r_i \to cr_i$:
replace row i with c times row i, where $c \neq 0$.

The Gauss algorithm uses only **1** and **2**.

$\underbrace{A\underline{x} = \underline{b}}$

$\downarrow$ row operations

$A' \underline{x} = \underline{b}'$
(-> same solution(s) as orig. system)

9.2 Possible solutions for $A\boldsymbol{x} = \boldsymbol{b}$

Consider the r.e.f. of $[A|\boldsymbol{b}]$. Then we have three possibilities:

(1) *Exactly one* solution; here the r.e.f. gives each variable a single value, so the number of variables, n, equals the number of non-zero rows in the r.e.f.

(2) *Infinitely many* solutions; here the number of non-zero rows of the r.e.f. is *less* than the number of variables.

For cases (1) and (2), the system is said to be *consistent*.

(3) *No* solution; when one row of r.e.f. is $(0\,0\,\ldots\,d)$ with $d \neq 0$. We can't solve $0x_1 + 0x_2 + \cdots + 0x_m = d$ if $d \neq 0$; it says $0 = d$. In this case the system is said to be *inconsistent*.

Note that a *homogeneous* system has $\boldsymbol{b} = \boldsymbol{0}$, i.e., all zero RHS. Then we always have at least the trivial solution, $x_i = 0$, $1 \leq i \leq n$.

Let A be an $m \times n$ matrix:

- The subspace of $\mathbb{R}^n$ spanned by the rows of A is called the *row space* of A, denoted $\mathrm{Row}(A)$.
- The subspace of $\mathbb{R}^m$ spanned by the columns of A is called the *column space* of A, denoted $\mathrm{Col}(A)$.
- The subspace of $\mathbb{R}^n$ that is the solution space of the homogeneous equation $A\boldsymbol{x} = \boldsymbol{0}$ is called the *nullspace* of A (denoted $N(A)$), and its dimension is often referred to as the *nullity*, denoted $\mathrm{nullity}(A)$.
- The *rank* of A, denoted $\underbrace{\mathrm{rank}(A)}$, is the dimension of the column space of A.

↳ number of non-zero rows in a r.e.f. matrix

9.2.1 Examples

$$\begin{aligned} x_1 + x_2 - \ \ x_3 &= 0 \\ 2x_1 - x_2 \qquad &= 0 \\ 4x_1 + x_2 - 2x_3 &= 1 \end{aligned}$$

$$A = \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ 4 & 1 & -2 & 1 \end{array}\right) \quad \begin{array}{l} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 4R_1 \end{array}$$

$$\to \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 2 & 0 \\ 0 & -3 & 2 & 1 \end{array}\right) \quad R_3 \to R_3 - R_2$$

$$\to \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \quad \text{(r.e.f.)} \quad \leftarrow \text{problem!}$$

R_3: $0x_1 + 0x_2 + 0x_3 = 1 \Rightarrow$ no solution

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & a \\ 2 & -1 & 0 & b \\ 4 & 1 & -2 & c \end{array}\right) = \begin{pmatrix} a + b - c \\ 2a - b \\ 4a + b - 2c \end{pmatrix} = a\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + b\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + c\begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix}$$

i.e. span of columns $\Rightarrow$ this product is in $Col(A)$

This example shows that $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is NOT in $Col(A)$

Also check $-\frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$\Rightarrow$ columns are linearly dependent

but $\dim(Col(A)) = rank(A) = 2$ in this case

$\boxed{\text{r.e.f. detects this}}$

$$\begin{aligned} x_1 - 2x_2 + 4x_3 &= 2 \\ 2x_1 - 3x_2 + 7x_3 &= 6 \\ x_2 - \ x_3 &= 2 \end{aligned}$$

$A = \left(\begin{array}{ccc|c} 1 & -2 & 4 & 2 \\ 2 & -3 & 7 & 6 \\ 0 & 1 & -1 & 2 \end{array}\right)$ $\quad R_2 \to R_2 - 2R_1$

$\to \left(\begin{array}{ccc|c} 1 & -2 & 4 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & -1 & 2 \end{array}\right)$ $\quad R_3 \to R_3 - R_2$

$\to \left(\begin{array}{ccc|c} 1 & -2 & 4 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$ $\quad$ (r.e.f.)

$\leftarrow *$ row of zeros

$\Rightarrow$ infinitely many solutions

$R_2: \ x_2 - x_3 = 2 \Rightarrow x_2 = 2 + x_3$

$R_1: \ x_1 - 2x_2 + 4x_3 = 2$

$\Rightarrow x_1 - 2(2 + x_3) + 4x_3 = 2 \Rightarrow x_1 = 6 - 2x_3$

set $x_3 = t \in \mathbb{R} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

These examples highlight some important results for linear systems.

A system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ is in the column space of A.

If $\boldsymbol{x}_0$ is any solution of a consistent linear system $A\boldsymbol{x} = \boldsymbol{b}$, and if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is a basis for the nullspace of A, then every solution of $A\boldsymbol{x} = \boldsymbol{b}$ can be expressed in the form

$$\boldsymbol{x} = \boldsymbol{x}_0 + c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k.$$

In other words, if $A\boldsymbol{x} = \boldsymbol{b}$ corresponds to a consistent linear system of m equations in n unknowns (so A is understood to be $m \times n$), and if $\text{rank}(A) = r$, then the solution contains $n - r$ free parameters.

Notes.

from previous example,

rank$(A) = 2$ (from inspecting r.e.f.)

$$\begin{pmatrix} 1 & -2 & 4 \\ 2 & -3 & 7 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + b\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} + c\begin{pmatrix} 4 \\ 7 \\ -1 \end{pmatrix}$$

choose $a = -2$, $b = 1$, $c = 1$

$$\Rightarrow -2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 1\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} + 1\begin{pmatrix} 4 \\ 7 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

i.e. $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \in \mathcal{N}(A)$ (nullspace of A)

10 Orthogonality

The goal of this section is to develop our understanding of orthogonality in the context of inner product spaces.

10.1 Norm

The norm (or magnitude or length) of an element $\mathbf{v} = (v_1, \dots, v_n)$ of $\mathbb{R}^n$ is given by the familiar expression

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \ldots + v_n^2}.$$

There is a similar notion for any real inner product space V. The **norm** of a vector $\mathbf{v} \in V$, denoted by $||\mathbf{v}||$, is thus defined by

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

A vector with norm 1 is called a **unit vector**.

How would we define the *distance*, $d(\mathbf{u}, \mathbf{v})$, between two vectors $\mathbf{u}, \mathbf{v} \in V$? A natural notion of distance between two vectors should be independent of the order we happen to be viewing them. That is, we want the distance measure to be symmetric: $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$. Again using $\mathbb{R}^n$ as inspiration, we now define the **distance** between two vectors $\mathbf{u}, \mathbf{v} \in V$ as

$$d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||.$$

The symmetry $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$ follows from

$$\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u} \rangle.$$

Note that the notions of norm and distance are relative to the inner product used! For example, with the inner product given in Example 8.8, the norm of a real-valued continuous function on $[-1, 1]$ is given as follows: for $f, g \in [-1, 1]$,

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, dx$$

$$\Rightarrow \quad ||f|| = \sqrt{\int_{-1}^{1} (f(x))^2 \, dx}$$

As in $\mathbb{R}^n$ with inner product given by the usual dot product, we say that two vectors $\mathbf{u}, \mathbf{v} \in V$ are **orthogonal** if

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

We need a bit of preparation before we can talk more generally about the *angle* between two vectors, see Section 10.7.

10.2 Pythagorean theorem

Let V be a real inner product space, and let $\mathbf{u}, \mathbf{v} \in V$. Then,

$$||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 \qquad \iff \qquad \langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

we have $||\vec{u}+\vec{v}||^2 = \langle \vec{u}+\vec{v} \,|\, \vec{u}+\vec{v} \rangle$

$\overset{(I2)}{=} \langle \vec{u}, \vec{u}+\vec{v} \rangle + \langle \vec{v}, \vec{u}+\vec{v} \rangle$

$\overset{(I1)}{=} \langle \vec{u}+\vec{v}, \vec{u} \rangle + \langle \vec{u}+\vec{v}, \vec{v} \rangle$

$\overset{(I2)}{=} \langle \vec{u}, \vec{u} \rangle + \langle \vec{v}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle$

$= ||\vec{u}||^2 + 2\langle \vec{u}, \vec{v} \rangle + ||\vec{v}||^2$ & the result follows

10.3 Cauchy-Schwarz inequality

Let V be a real inner product space, and let $\mathbf{u}, \mathbf{v} \in V$. Then,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}||\, ||\mathbf{v}||.$$

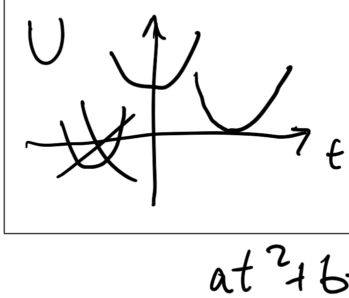
Moreover, this inequality is an equality if and only if $\mathbf{u}$ or $\mathbf{v}$ is a scalar multiple of the other vector.

For $\vec{u} = \vec{0} \Rightarrow$ ok!

For $\vec{u} \neq \vec{0}$, let $t \in \mathbb{R}$. Then,

$0 \leq ||t\vec{u} + \vec{v}||^2 = ||\vec{u}||^2 t^2 + 2\langle \vec{u}, \vec{v} \rangle t + ||\vec{v}||^2$

This is a quadratic expression in t, which must have either no real roots, or a single root.



$\Rightarrow$ discriminant must satisfy

$\left(2\langle \vec{u}, \vec{v} \rangle\right)^2 - 4||\vec{u}||^2||\vec{v}||^2 \leq 0$

& the result follows

$at^2 + bt + c = 0 \Rightarrow t = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

10.4 Triangle inequality

Let V be a real inner product space, and let $\mathbf{u}, \mathbf{v} \in V$. Then,

$$||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||.$$

$$\begin{aligned}
||\vec{u}+\vec{v}||^2 &= ||\vec{u}||^2 + 2\langle \vec{u}, \vec{v}\rangle + ||\vec{v}||^2 \\
&\leq ||\vec{u}||^2 + 2|\langle \vec{u}, \vec{v}\rangle| + ||\vec{v}||^2 \\
&\leq ||\vec{u}||^2 + 2||\vec{u}||\,||\vec{v}|| + ||\vec{v}||^2 \\
&= (||\vec{u}|| + ||\vec{v}||)^2 \quad \to \text{result}
\end{aligned}$$

$k \leq |k|,\ k \in \mathbb{R}$

Cauchy-Schwarz

$$\Rightarrow ||\vec{u}+\vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$$

10.5 Example: $C[-1,1]$

use $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$

For $f, g \in C[-1,1]$

Cauchy-Schwarz

$$\Rightarrow \left| \int_{-1}^{1} f(x)g(x)\,dx \right|^2 \leq \left(\int_{-1}^{1} (f(x))^2 dx\right)\left(\int_{-1}^{1} (g(x))^2 dx\right)$$

10.6 Example: $P_2(\mathbb{R})$

use $\langle u, v\rangle = \int_{-1}^{1} u(x)v(x)\,dx$

For $u, v \in P_2(\mathbb{R})$

e.g. $u(x) = x$, $v(x) = x^2$

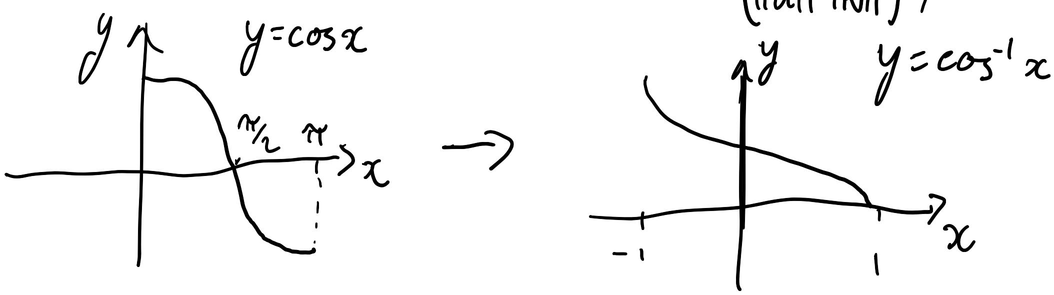
$$\Rightarrow \langle x, x^2\rangle = \int_{-1}^{1} x \cdot x^2\,dx = \left[\tfrac{1}{4}x^4\right]_{-1}^{1} = 0$$

$\Rightarrow$ x & x^2 are orthogonal with respect to the given inner product

10.7 Angle between two vectors

In $\mathbb{R}^n$, the angle θ between $\vec{u}, \vec{v}$ is given by

$$\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\| \Rightarrow \theta = \cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right), \quad \theta\in[0,\pi]$$



For a (general) inner product space, define

$$\theta = \cos^{-1}\left(\frac{\langle\vec{u},\vec{v}\rangle}{\underbrace{\|\vec{u}\|\,\|\vec{v}\|}}\right), \quad \theta\in[0,\pi]$$

Cauchy-Schwarz $\Rightarrow$ the argument of $\cos^{-1}(\cdot)$ is bounded $|\cdot|\le 1$

10.8 Orthogonal complement

Let U be a subset of the real inner product space V. The **orthogonal complement** of U, denoted by $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U. That is,

$$U^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for every } \mathbf{u} \in U\}.$$

$U^\perp$ = "U perp"

This is a vector space with addition and scalar multiplication inherited from V.

10.8.1 For $A \in M_{m,n}(\mathbb{R})$, $\mathbf{Row}(A)^\perp = N(A)$ $\subseteq \mathbb{R}^n$ with respect to the Euclidean inner product

Consider $A\underline{x} = \underline{0}$ $\Rightarrow$ $\underbrace{\begin{pmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_m^T \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \underline{0}$

$\Rightarrow$ $\underline{r}_j \cdot \underline{x} = 0 \quad \forall j \in \{1, 2, \ldots, m\}$, column vectors $\underline{r}_j \in \mathbb{R}^n \Rightarrow \underline{r}_j^T$ is a row vector

In fact, $U^\perp$ is an example of a *subspace*. Indeed, a nonempty subset W of a vector space V is a **subspace** of V if it is a vector space under the addition and scalar multiplication defined on V. To verify that a subset is a subspace, one checks the following:

For $\underline{u}, \underline{v} \in U^\perp$, $\underline{w} \in U$, $k \in \mathbb{R}$

$\underline{u} + \underline{v} \in U^\perp$ since $\langle \underline{u} + \underline{v}, \underline{w} \rangle = \langle \underline{u}, \underline{w} \rangle + \langle \underline{v}, \underline{w} \rangle$
$= 0 + 0 = 0$

$k\underline{u} \in U^\perp$ since $\langle k\underline{u}, \underline{w} \rangle = k\langle \underline{u}, \underline{w} \rangle = 0$

(Also $\underline{0} \in U^\perp$)

Notes.

11 Gram-Schmidt process

The goal of this section is to discuss orthogonal projections, and develop a way of constructing an orthonormal basis for an inner product space.

11.1 Orthogonal set

Let V be a real inner product space. A nonempty set of vectors in V is **orthogonal** if each vector in the set is orthogonal to all the other vectors in the set. That is, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ is orthogonal if

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \qquad i \neq j.$$

Let S be a finite set of vectors in V such that $\mathbf{0} \notin S$. Then,

$$S \text{ orthogonal} \qquad \Longrightarrow \qquad S \text{ linearly independent.}$$

Let $S = \{\underline{v}_1, \ldots, \underline{v}_n\}$ be orthogonal & $\underline{v}_i \neq \underline{0}$ for $i = 1, \ldots, n$

Consider $k_1 \underline{v}_1 + k_2 \underline{v}_2 + \ldots + k_n \underline{v}_n = \underline{0}$

Then for any $\underline{v}_i \in S$

$$\begin{aligned} 0 = \langle \underline{0}, \underline{v}_i \rangle &= \langle k_1 \underline{v}_1 + k_2 \underline{v}_2 + \ldots + k_n \underline{v}_n, \underline{v}_i \rangle \\ &= k_1 \langle \underline{v}_1, \underline{v}_i \rangle + \ldots + k_n \langle \underline{v}_n, \underline{v}_i \rangle \\ &= k_i \langle \underline{v}_i, \underline{v}_i \rangle \end{aligned}$$

Since $\underline{0} \notin S$, we have $\underline{v}_i \neq 0$

so, $\langle \underline{v}_i, \underline{v}_i \rangle \neq 0 \Rightarrow k_i = 0$

This is true for every $i \Rightarrow S$ is linearly independent

11.2 Orthonormal basis

An orthogonal set of vectors in V is called **orthonormal** if all the vectors in the set are unit vectors. That is, the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \subset V$ is orthonormal if

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{i,j},$$

where the **Kronecker delta** is defined by

$$\delta_{i,j} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Examples in $\mathbb{R}^n$ endowed with the usual dot product are given as follows:

(1) Standard basis of $\mathbb{R}^3$ $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

(2) in $\mathbb{R}^3$: $\left\{ \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \right\} =: S$

(3) $S \cup \left\{ \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix} \right\}$

An **orthonormal basis** for V is a basis for V that is also an orthonormal set.

Examples (1) and (3) above are orthonormal bases.

11.3 Decomposition theorem

↗ inner product space

Let $S = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be an orthonormal basis for V, and let $\mathbf{u} \in V$. Then,

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \ldots + \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n$$

and

$$||\mathbf{u}||^2 = \langle \mathbf{u}, \mathbf{e}_1 \rangle^2 + \ldots + \langle \mathbf{u}, \mathbf{e}_n \rangle^2.$$

We can uniquely write

$\underline{u} = a_1\underline{e}_1 + a_2\underline{e}_2 + \ldots + a_n\underline{e}_n$

$\Rightarrow$ for $i = 1, \ldots, n$

$\langle \underline{u}, \underline{e}_i \rangle = \langle a_1\underline{e}_1 + \ldots + a_n\underline{e}_n, \underline{e}_i \rangle$

$= a_1 \langle \underline{e}_1, \underline{e}_i \rangle + \ldots + a_n \langle \underline{e}_n, \underline{e}_i \rangle$

$= a_i \langle \underline{e}_i, \underline{e}_i \rangle = a_i$ $\quad \Rightarrow \langle \underline{e}_i, \underline{e}_i \rangle = 1$

<u>AND</u>

$||\underline{u}||^2 = \langle \underline{u}, \underline{u} \rangle$

$= \langle a_1\underline{e}_1 + \ldots + a_n\underline{e}_n, a_1\underline{e}_1 + \ldots + a_n\underline{e}_n \rangle$

$= a_1^2 + a_2^2 + \ldots + a_n^2$

& the result follows

11.4 Orthogonal projection

Let U be a finite-dimensional subspace of the real inner product space V. Then, each $\mathbf{v} \in V$ can be written in a unique way as

$$\mathbf{v} = \mathbf{u} + \mathbf{w}, \qquad \mathbf{u} \in U, \quad \mathbf{w} \in U^{\perp}.$$

In the proof, we will assume that U has an orthonormal basis $S = \{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$. As we will see in Section 11.6, this assumption is redundant since every finite-dimensional inner product space *has* such a basis.

Let $\underline{v} \in V$. Then $\underline{v} = \underline{u} + (\underline{v} - \underline{u})$ for any $\underline{u} \in U$

Set $\underline{u} = a_1 \underline{e}_1 + \ldots + a_k \underline{e}_k \in U$ & set $\underline{w} = \underline{v} - \underline{u}$

For $i = 1, \ldots, k$

$$\langle \underline{w}, \underline{e}_i \rangle = \langle \underline{v} - (a_1 \underline{e}_1 + \ldots + a_k \underline{e}_k), \underline{e}_i \rangle$$
$$= \langle \underline{v}, \underline{e}_i \rangle - a_i \langle \underline{e}_i, \underline{e}_i \rangle$$
$$= \langle \underline{v}, \underline{e}_i \rangle - a_i \qquad \text{(we want } = 0)$$

choose $a_i = \langle \underline{v}, \underline{e}_i \rangle$ so that $\langle \underline{w}, \underline{e}_i \rangle = 0$

$\Rightarrow$ $\underline{w}$ is orthogonal to every vector in U

$\Rightarrow$ $\underline{w} \in U^{\perp}$

Uniqueness? Let $\underline{v} = \underline{u} + \underline{w}$, $\underline{v} = \underline{u}' + \underline{w}'$, $\underline{u}, \underline{u}' \in U$, $\underline{w}, \underline{w}' \in U^{\perp}$

$\Rightarrow \underbrace{\underline{u} - \underline{u}'}_{\in U} = \underbrace{\underline{w}' - \underline{w}}_{\in U^{\perp}}$, but $U \cap U^{\perp} = \{\underline{0}\} \to \underline{u} - \underline{u}' = \underline{0} = \underline{w}' - \underline{w}$

The vector $\mathbf{u} \in U$ is called the **orthogonal projection of v onto** U and is given by

$$\boxed{\mathrm{Proj}_U(\mathbf{v}) = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \ldots + \langle \mathbf{v}, \mathbf{e}_k \rangle \mathbf{e}_k.}$$

$\{\underline{e}_1, \ldots, \underline{e}_k\}$ is an orthonormal basis for U

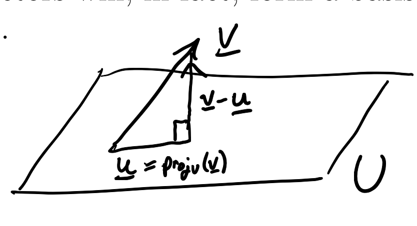
Likewise, the vector $\mathbf{w} \in U^{\perp}$ is called the **orthogonal projection of v onto** $U^{\perp}$ and is given by

$$\mathrm{Proj}_{U^{\perp}}(\mathbf{v}) = \mathbf{v} - \mathrm{Proj}_U(\mathbf{v}).$$

One can show that

$$\dim V = \dim U + \dim U^{\perp}.$$

This can be very helpful when determining the orthogonal complement of a subspace U. Indeed, suppose you have managed to find $\dim V - \dim U$ linearly independent vectors that are all orthogonal to U. Then, these vectors will, in fact, form a basis for $U^{\perp}$. This property will be used in the next example.



← $\underline{u}$ is a projection of $\underline{v}$ onto U

11.5 Example: Orthogonal projection in $\mathbb{R}^3$

Let $\mathbb{R}^3$ be endowed with the usual dot product, and let

$$U = \mathrm{span}(\{\overset{\underline{e}_1}{(0,1,0)}, \overset{\underline{e}_2}{(-\tfrac{4}{5}, 0, \tfrac{3}{5})}\}), \qquad \mathbf{v} = (1,1,1).$$

Find the orthogonal projections of $\mathbf{v}$ onto U and $U^\perp$.

$S = \{\underline{e}_1, \underline{e}_2\}$ is orthonormal (check)

$$\begin{aligned}
\mathrm{Proj}_U(\underline{v}) &= \langle \underline{v}, \underline{e}_1 \rangle \underline{e}_1 + \langle \underline{v}, \underline{e}_2 \rangle \underline{e}_2 \\
&= 1(0,1,0) + \left(-\tfrac{1}{5}\right)\left(-\tfrac{4}{5}, 0, \tfrac{3}{5}\right) \\
&= \left(\tfrac{4}{25}, 1, \tfrac{-3}{25}\right)
\end{aligned}$$

Also
$$\begin{aligned}
\mathrm{Proj}_{U^\perp}(\underline{v}) &= \underline{v} - \mathrm{Proj}_U(\underline{v}) \\
&= \left(\tfrac{21}{25}, 0, \tfrac{28}{25}\right)
\end{aligned}$$

11.6 Construction of orthonormal basis

It is often convenient to have an orthonormal basis for a given finite-dimensional inner product space. The following algorithm turns a linearly independent set of vectors into an orthonormal set of vectors with the same span as the original set. Applying the algorithm to a basis thus turns the basis into an orthonormal basis. Hence:

> Every finite-dimensional real inner product space has an orthonormal basis.

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a linearly independent set of vectors in the real inner product space V. The corresponding **Gram-Schmidt process** is the following algorithm:

Step 1: Set $\underline{e}_1 = \dfrac{\underline{v}_1}{\|\underline{v}_1\|}$ (i.e. $\underline{e}_1$ is now a unit vector)

Step 2: Let $U_1 = \operatorname{span}(\{\underline{e}_1\})$

Set $\underline{w}_2 = \underline{v}_2 - \operatorname{Proj}_{U_1}(\underline{v}_2) \Rightarrow \underline{w}_2 \in U_1^{\perp}$

& $\underline{w}_2 \neq \underline{0}$ & set $\underline{e}_2 = \dfrac{\underline{w}_2}{\|\underline{w}_2\|}$ (i.e. $\underline{e}_2$ is a unit vector)

Step 3: Let $U_2 = \operatorname{span}(\underbrace{\{\underline{e}_1, \underline{e}_2\}}_{\text{orthonormal set}})$

Set $\underline{w}_3 = \underline{v}_3 - \operatorname{Proj}_{U_2}(\underline{v}_3) \Rightarrow \underline{w}_3 \in U_2^{\perp}$

& $\underline{w}_3 \neq 0$ & set $\underline{e}_3 = \dfrac{\underline{w}_3}{\|\underline{w}_3\|}$

$\vdots$

Step $i+1$: Let $U_i = \operatorname{span}(\{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_i\})$

Set $\underline{w}_{i+1} = \underline{v}_{i+1} - \operatorname{Proj}_{U_i}(\underline{v}_{i+1}) \Rightarrow \underline{w}_{i+1} \in U_i^{\perp}$

& $\underline{w}_{i+1} \neq 0$ & $\underline{e}_{i+1} = \dfrac{\underline{w}_{i+1}}{\|\underline{w}_{i+1}\|}$

(Algorithm finishes after step n for vector $\underline{v}_n$)

Result: $S = \{\underline{e}_1, \ldots, \underline{e}_n\}$ is orthonormal and

$\operatorname{span}(S) = \operatorname{span}(\{\underline{v}_1, \ldots, \underline{v}_n\})$

11.7 Example: Orthonormal basis for $P_1(\mathbb{R})$

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$$

Basis: $\{\underbrace{1+x}_{\underline{v}_1}, \underbrace{1-2x}_{\underline{v}_2}\}$

Gram-Schmidt

Step 1: $\|\underline{v}\|^2 = \int_{-1}^{1}(1+x)^2\,dx = \frac{8}{3} \Rightarrow \underline{e}_1 = \frac{1+x}{\sqrt{8/3}} = \frac{\sqrt{3}}{2\sqrt{2}}(1+x)$

Step 2: $\underline{w}_2 = \underline{v}_2 - \underbrace{\langle \underline{v}_2, \underline{e}_1 \rangle \underline{e}_1}_{\text{Proj}_{U_1}(\underline{v}_2),\text{ where } U_1 = \text{span}\{\underline{e}_1\}}$

$$\langle \underline{v}_2, \underline{e}_1 \rangle = \int_{-1}^{1}(1-2x)\frac{\sqrt{3}}{2\sqrt{2}}(1+x)\,dx = \frac{1}{\sqrt{6}}$$

$$\Rightarrow \underline{w}_2 = 1-2x - \frac{1}{\sqrt{6}}\cdot\frac{\sqrt{3}}{2\sqrt{2}}(1+x) = \frac{3}{4} - \frac{9}{4}x$$

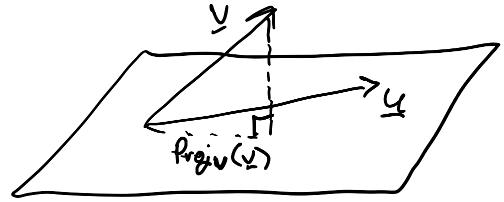
$$\& \ \|\underline{w}_2\|^2 = \int_{-1}^{1}\left(\frac{3}{4} - \frac{9}{4}x\right)^2 dx = \frac{9}{2}$$

$$\Rightarrow \underline{e}_2 = \frac{\underline{w}_2}{\|\underline{w}_2\|} = \frac{\sqrt{2}}{3}\left(\frac{3}{4} - \frac{9}{4}x\right) = \frac{1-3x}{2\sqrt{2}}$$

an orthonormal basis of $P_1(\mathbb{R})$ is

$$\left\{\frac{\sqrt{3}}{2\sqrt{2}}(1+x), \frac{1}{2\sqrt{2}}(1-3x)\right\}$$

Notes.



12 Least squares approximation

12.1 Least squares problem - minimising distance to a subspace

A recurring problem in linear algebra, and in its myriad of applications, is the following:

- Given a vector $\mathbf{v}$ in a real inner product space V, give the best approximation to $\mathbf{v}$ in a finite-dimensional subspace U of V.

Question: What do we mean by "best approximation"?

Answer: Seek $\mathbf{u} \in U$ that minimises $||\mathbf{v} - \mathbf{u}||$. Equivalently, find a vector in a subspace (for example, corresponding to a point on a plane in $\mathbb{R}^3$), of minimal distance to a given vector in the ambient vector space (in this example, corresponding to a point in $\mathbb{R}^3$). Concretely, let $\mathbf{v} \in V$. Then, the problem is to

find $\mathbf{u} \in U$ *such that* $d(\mathbf{u}, \mathbf{v})$ *is as small as possible.*

This problem is called the "least squares problem."

Theorem (Best Approximation Theorem). If U is a finite-dimensional subspace of a real inner product space V, and if $\mathbf{v} \in V$, then $\mathrm{Proj}_U(\mathbf{v})$ is the best approximation to $\mathbf{v}$ from U in the sense that

$$||\mathbf{v} - \mathrm{Proj}_U(\mathbf{v})|| < ||\mathbf{v} - \mathbf{u}|| \;\; \forall \mathbf{u} \in U \;\;:\quad \mathbf{u} \neq \mathrm{Proj}_U(\mathbf{v}).$$

Let $\underline{v} \in V$, $\underline{u} \in U$

$$||\underline{v} - \mathrm{Proj}_U(\underline{v})||^2 \leq ||\underbrace{\underline{v} - \mathrm{Proj}_U(\underline{v})}_{\in U^\perp}||^2 + ||\mathrm{Proj}_U(\underline{v}) - \underline{u}||^2$$

$$= ||\underline{v} - \mathrm{Proj}_U(\underline{v}) + \mathrm{Proj}_U(\underline{v}) - \underline{u}||^2 \quad \text{(pythagoras)}$$

$$= ||\underline{v} - \underline{u}||^2$$

$$\Rightarrow ||\underline{v} - \mathrm{Proj}_U(\underline{v})||^2 \leq ||\underline{v} - \underline{u}||^2 \;\; \forall \underline{u} \in U$$

& result follows

In practice, rather than work with minimising $||\mathbf{v}-\mathbf{u}||$, we minimise $||\mathbf{v}-\mathbf{u}||^2$ (same outcome, avoid square root). Then Best Approximation Theorem $\implies$ $\text{Proj}_U(\mathbf{v})$ is the best approximation

$\iff$ $\mathbf{u} = \text{Proj}_U(\mathbf{v})$ is the vector that minimises $||\mathbf{v} - \mathbf{u}||^2$.

How to find $\text{Proj}_U(\mathbf{v})$*?*

Solution 1. Use Gram-Schmidt process to construct an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ of U. Then

$$\text{Proj}_U(\mathbf{v}) = \langle \mathbf{v}, \mathbf{e}_1\rangle \mathbf{e}_1 + \ldots + \langle \mathbf{v}, \mathbf{e}_n\rangle \mathbf{e}_n.$$

Solution 2.

Let $\gamma = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a basis of the subspace U. We seek coefficients $\alpha_1, \alpha_2, \ldots \alpha_n$ that minimise $||\mathbf{v} - (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n)||^2$.

$$\begin{aligned}
&||\mathbf{v} - (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n)||^2\\
=&\langle \mathbf{v} - (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n), \mathbf{v} - (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n)\rangle\\
=&\langle \mathbf{v}, \mathbf{v}\rangle - 2\alpha_1\langle \mathbf{v}, \mathbf{u}_1\rangle - 2\alpha_2\langle \mathbf{v}, \mathbf{u}_2\rangle - \cdots - 2\alpha_n\langle \mathbf{v}, \mathbf{u}_n\rangle + \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j\langle \mathbf{u}_i, \mathbf{u}_j\rangle\\
=&E(\alpha_1, \alpha_2, \ldots, \alpha_n).
\end{aligned}$$

E is the "error (real valued function)."

From MATH1052/MATH1072: we minimise E. Set $\nabla E = \mathbf{0}$, which gives

$$\frac{\partial E}{\partial \alpha_k} = -2\langle \mathbf{v}, \mathbf{u}_k\rangle + 2\sum_{l=1}^{n}\alpha_l\langle \mathbf{u}_k, \mathbf{u}_l\rangle = 0. \qquad k = 1, \ldots, n$$

This is n equations in n unknowns, which may be expressed in matrix form;

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1\rangle & \langle \mathbf{u}_1, \mathbf{u}_2\rangle & \ldots & \langle \mathbf{u}_1, \mathbf{u}_n\rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1\rangle & \langle \mathbf{u}_2, \mathbf{u}_2\rangle & \ldots & \langle \mathbf{u}_2, \mathbf{u}_n\rangle \\ \vdots & & \vdots & \\ \langle \mathbf{u}_n, \mathbf{u}_1\rangle & \langle \mathbf{u}_n, \mathbf{u}_2\rangle & \ldots & \langle \mathbf{u}_n, \mathbf{u}_n\rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \langle \mathbf{v}, \mathbf{u}_1\rangle \\ \langle \mathbf{v}, \mathbf{u}_2\rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{u}_n\rangle \end{pmatrix}.$$

Note: If γ is orthonormal $\implies$ matrix on LHS$= I$ (the identity matrix) $\implies$ $\alpha_i = \langle \mathbf{v}, \mathbf{u}_i\rangle$, which coincides with Solution 1, as expected.

12.2 Inconsistent linear systems $A\underline{x} = \underline{b}$

In applications, one is often faced with over-determined linear systems. For example, we may have a bunch of data points that we have reasons to believe should fit on a straight line. But real-life data points rarely match predictions exactly. The goal in this section is to develop a method for obtaining the best fit or approximation of a specified kind to a given set of data points.

Considering $A\underline{x} = \underline{b}$

To this end, let $A \in M_{m,n}(\mathbb{R})$ and $\mathbf{b} \in M_{m,1}(\mathbb{R})$. For $m > n$, the linear system described by $A\mathbf{x} = \mathbf{b}$ is over-determined and does not in general have a solution. However, we may be satisfied if we can find $\hat{\mathbf{x}} \in M_{n,1}(\mathbb{R})$ such that $A\hat{\mathbf{x}}$ is as 'close' to $\mathbf{b}$ as possible. That is, we seek to minimise $||\mathbf{b} - A\mathbf{x}||$ for given A and $\mathbf{b}$. For computational reasons, one usually minimises $||\mathbf{b} - A\mathbf{x}||^2$ instead. A solution $\hat{\mathbf{x}}$ to this minimisation problem is referred to as a **least squares solution**.

First, let us recall the **column space** of the matrix A as

$$\mathrm{Col}(A) = \mathrm{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}),$$

where $\mathbf{a}_1, \ldots, \mathbf{a}_n \in M_{m,1}(\mathbb{R})$ are the columns of A. Since

$$\mathbf{a}_1 = A\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \quad \mathbf{a}_2 = A\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, \quad \mathbf{a}_n = A\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix},$$

we have

$$\begin{aligned}\mathrm{Col}(A) &= \mathrm{span}\Bigg(\{A\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, A\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, A\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\}\Bigg) = \{A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} \,|\, x_1, x_2, \ldots, x_n \in \mathbb{R}\}\\ &= \{A\mathbf{x} \,|\, \mathbf{x} \in M_{n,1}(\mathbb{R})\}.\end{aligned}$$

It follows that $A\hat{\mathbf{x}} \in \mathrm{Col}(A)$. We are thus in the situation described in Section 12.1.

For a finite dimensional subspace $U \subseteq V$ & $\underline{v} \in V$

$\underline{u} = \mathrm{Proj}_U(\underline{v})$ minimises $d(\underline{u}, \underline{v})$

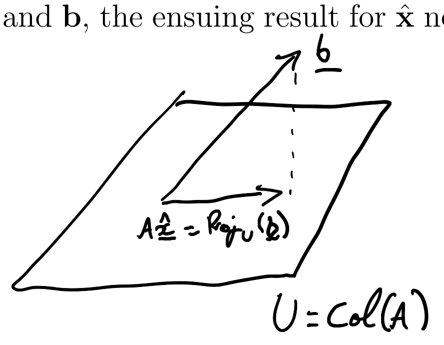
Here $U = \mathrm{Col}(A)$, $\underline{v} = \underline{b}$ with all vectors in U of the form $A\underline{x}$

$(A\underline{x} = \underline{b})$

Consequently, we know that

$$A\hat{\mathbf{x}} = \mathrm{Proj}_{\mathrm{Col}(A)}(\mathbf{b}),$$

but we are still faced with the task of disentangling $\hat{\mathbf{x}}$ from $A\hat{\mathbf{x}}$. Note that, although $A\hat{\mathbf{x}} = \mathrm{Proj}_{\mathrm{Col}(A)}(\mathbf{b})$ is *uniquely* given in terms of A and $\mathbf{b}$, the ensuing result for $\hat{\mathbf{x}}$ need *not* be.



Solution 3. For the current special application, i.e. least squares solutions of linear systems, we have a more direct (and simpler) method. In the following we describe this method and give some examples.

(c.f. $\mathrm{Row}(A)^\perp = N(A)$)

12.3 $\mathrm{Col}(A)^\perp = N(A^T)$

Note: $\underline{a} \cdot \underline{b} = \underline{a}^T \underline{b}$

The orthogonal complement of the column space of A is the null space of A^T.

Want to show equality by establishing "$\subseteq$" and "$\supseteq$"

First consider $(A\underline{v}) \cdot \underline{u} = (A\underline{v})^T \underline{u}$

$$= \underline{v}^T (A^T \underline{u})$$

$$\Rightarrow (A\underline{v}) \cdot \underline{u} = \underline{v} \cdot (A^T \underline{u}) \qquad (*)$$

"$\subseteq$" Let $\underline{u} \in \mathrm{Col}(A)^\perp$. Then $\forall \underline{v} \in \mathbb{R}^n$

$$0 = (A\underline{v}) \cdot \underline{u} \overset{*}{=} \underline{v} \cdot (A^T \underline{u})$$

Now set $\underline{v} = A^T \underline{u} \Rightarrow 0 = (A^T \underline{u}) \cdot (A^T \underline{u})$

$$\Rightarrow A^T \underline{u} = \underline{0} \Rightarrow \underline{u} \in N(A^T)$$

"$\supseteq$" Let $\underline{u} \in N(A^T) \Rightarrow A^T \underline{u} = \underline{0}$

Then $\forall \underline{v} \in \mathbb{R}^n$

$$(A\underline{v}) \cdot \underline{u} \overset{*}{=} \underline{v} \cdot (A^T \underline{u})$$

$$= \underline{v} \cdot \underline{0} = 0$$

$\Rightarrow \underline{u}$ is orthogonal to all vectors of the form $A\underline{v} \Rightarrow \underline{u} \in \mathrm{Col}(A)^\perp$

12.4 Solving for $\hat{\mathbf{x}}$

Recall from Section 11.4 that $\mathbf{v}-\mathrm{Proj}_U(\mathbf{v}) = \mathrm{Proj}_{U^\perp}(\mathbf{v}) \in U^\perp$ for any finite-dimensional subspace U of V and $\mathbf{v} \in V$. Since $A\hat{\mathbf{x}} = \mathrm{Proj}_{\mathrm{Col}(A)}(\mathbf{b})$, we thus have

$$\mathbf{b} - A\hat{\mathbf{x}} \in \mathrm{Col}(A)^\perp.$$

Because $\mathrm{Col}(A)^\perp = N(A^T)$, it follows that $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, so

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}. \quad \longrightarrow \text{"Normal Equation"}$$

Since $A \in M_{m,n}(\mathbb{R}) \;\Rightarrow\; A^TA \in M_{n,n}(\mathbb{R})$, the equation $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ can always be solved for $\hat{\mathbf{x}}$, for example by Gaussian elimination. However, the solution need not be unique. Indeed, the solution is unique if and only if A^TA is invertible, in which case

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}.$$

A key to determining whether A^TA is invertible is the following result:

$$\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \text{ is linearly independent} \quad \Longleftrightarrow \quad A^TA \text{ is invertible.}$$

<u>Lemma</u>: $N(A) = N(A^TA)$

Proof: "$\subseteq$" $\underline{v} \in N(A) \Rightarrow A\underline{v} = \underline{0} \Rightarrow A^TA\underline{v} = \underline{0} \Rightarrow \underline{v} \in N(A^TA)$

"$\supseteq$" $\underline{v} \in N(A^TA)$

Consider $(A\underline{v})\cdot(A\underline{v}) \overset{(*)\,p96}{=} \underline{v}\cdot(A^TA\underline{v}) = \underline{v}\cdot\underline{0} = 0$

$\Rightarrow A\underline{v} = \underline{0} \Rightarrow \underline{v} \in N(A)$

$\{\underline{a}_1, \ldots, \underline{a}_n\}$ is linearly indep. iff $N(A) = \{\underline{0}\}$

iff $N(A^TA) = \{\underline{0}\}$

iff A^TA is invertible

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a\begin{pmatrix} \cdot \\ \vdots \\ \cdot \end{pmatrix} + b\begin{pmatrix} \cdot \\ \vdots \\ \cdot \end{pmatrix} + c\begin{pmatrix} \cdot \\ \vdots \\ \cdot \end{pmatrix} = \underline{0}$$

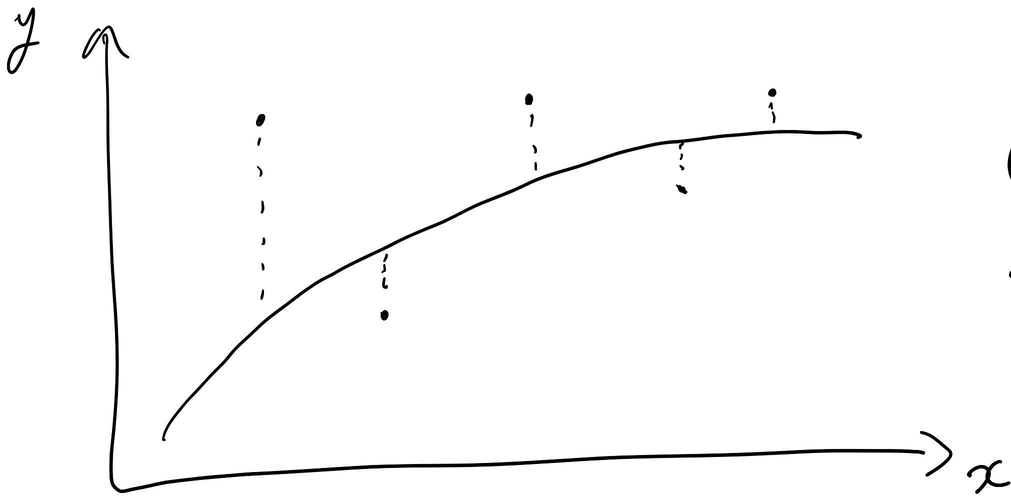
But how does that help? Well, recall that $m > n$ so that A may be describing an over-determined linear system. In fact, in case the linear system encodes experimental or empirical data, m is likely to be much larger than n. The columns of A are then very likely to be linearly independent, and A^TA would indeed be invertible.

12.5 Fitting a curve to data

Experiments yield data (assume x_i distinct and exact)

$$(x_1, y_1),\ (x_2, y_2),\ \cdots,\ (x_n, y_n)$$

which include measurement error. Theory may predict polynomial relation between x and y. We seek *a least squares polynomial function of best fit* (e.g. least squares line of best fit or regression line).



(want to minimise the dotted lines)

Example: Quadratic fit. Suppose some physical system is modeled by a quadratic function $p(t)$. Data in the form $(t, p(t))$ have been recorded as

$$(1, 5),\ (2, 2),\ (4, 7),\ (5, 10).$$

Find the least squares approximation for $p(t)$.

Quadratic $p(t) = a_0 + a_1 t + a_2 t^2$

Linear system:

$t=1 \Rightarrow a_0 + a_1 + a_2 = 5$

$t=2 \Rightarrow a_0 + 2a_1 + 4a_2 = 2$

$t=4 \Rightarrow a_0 + 4a_1 + 16a_2 = 7$

$t=5 \Rightarrow a_0 + 5a_1 + 25a_2 = 10$

$\Rightarrow$ $A\underline{x} = \underline{b}$ we have

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{pmatrix} \quad \underline{b} = \begin{pmatrix} 5 \\ 2 \\ 7 \\ 10 \end{pmatrix}$$

$\rightarrow$ overdetermined and inconsistent

but columns of A are linearly independent

$\rightarrow$ $\exists$ unique least squares solution

$A^T A \hat{\underline{x}} = A^T \underline{b}$ (Normal Equation)

$$\Rightarrow \hat{\underline{x}} = (A^T A)^{-1} A^T \underline{b} = \begin{pmatrix} 8 \\ -9/2 \\ 1 \end{pmatrix}$$

$$\Rightarrow p(t) = 8 - \frac{9}{2}t + t^2$$

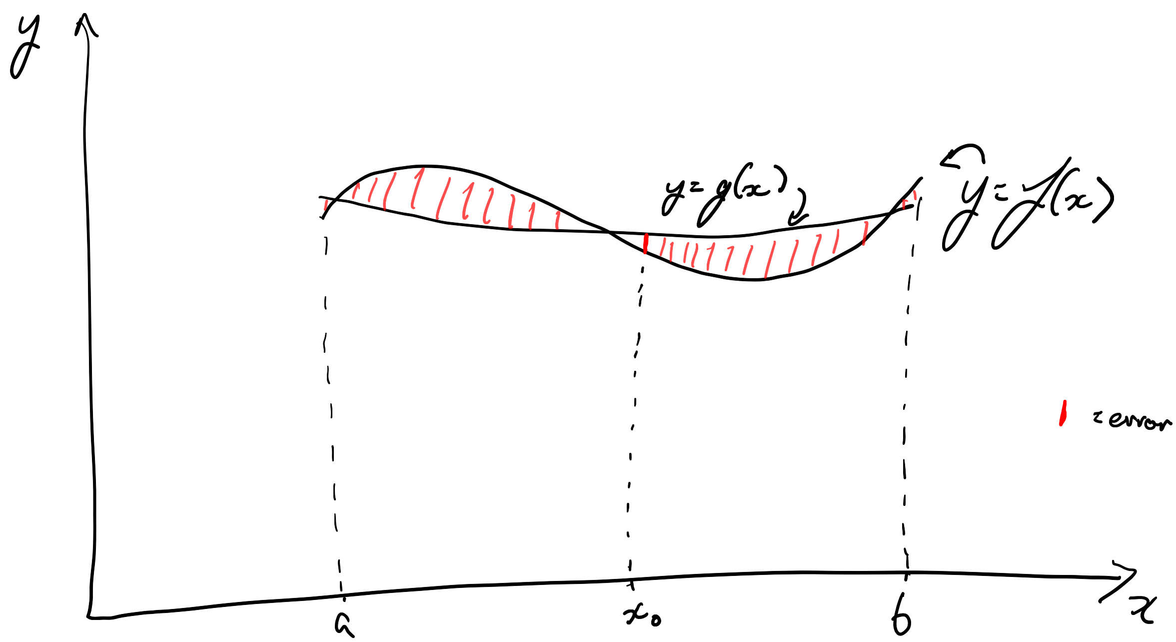
(see MATLAB m-file)

Notes.

13 Least squares function approximation

Given a function $f \in C[a,b]$, find the best approximation to f using only functions from a specified subspace U of $C[a,b]$.

Interpret "best possible" in the sense of least squares.



Consider g as an approximation to f.

At point x_0 the error is $|f(x_0) - g(x_0)|$. For the entire interval, define error as $\int_a^b |f(x) - g(x)|dx$.

This is *area between curves.*

An easier definition (and one more amenable to calculations) is the *mean squared error* (MSE)

$$\text{MSE} = \int_a^b (f(x) - g(x))^2 \, dx.$$

Recall the integral inner product on $C[a,b]$;

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_a^b p(x)q(x) \, dx$$

$$\implies \text{MSE } = ||\mathbf{f} - \mathbf{g}||^2 = \langle \mathbf{f} - \mathbf{g}, \mathbf{f} - \mathbf{g} \rangle = \int_a^b (f(x) - g(x))^2 \, dx.$$

13.1 Example: $\sin(x)$

$U = \text{span}(\gamma)$

Find the least squares approximation for $\sin x$ in the subspace of $C[0, \pi]$ spanned by $\gamma = \{1, x, x^2\}$. Use the inner product

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_0^\pi p(x)q(x)\ dx.$$

Use method of "solution 2" from p94

i.e. For $y = \alpha_1 + \alpha_2 x + \alpha_3 x^2$, solve

$$\begin{pmatrix} \langle 1,1\rangle & \langle 1,x\rangle & \langle 1,x^2\rangle \\ \langle x,1\rangle & \langle x,x\rangle & \langle x,x^2\rangle \\ \langle x^2,1\rangle & \langle x^2,x\rangle & \langle x^2,x^2\rangle \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \langle \sin x, 1\rangle \\ \langle \sin x, x\rangle \\ \langle \sin x, x^2\rangle \end{pmatrix}$$

Note $\displaystyle\int_0^\pi x^n\,dx = \frac{\pi^{n+1}}{n+1}$

& $\langle \sin x, 1\rangle = 2$, $\langle \sin x, x\rangle = \pi$, $\langle \sin x, x^2\rangle = \pi^2 - 4$

$$\Rightarrow \begin{pmatrix} \pi & \pi^2/2 & \pi^3/3 \\ \pi^2/2 & \pi^3/3 & \pi^4/4 \\ \pi^3/3 & \pi^4/4 & \pi^5/5 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 2 \\ \pi \\ \pi^2 - 4 \end{pmatrix}$$

$$\Rightarrow \alpha_1 = \frac{12(\pi^2 - 10)}{\pi^3}, \quad \alpha_2 = \frac{-60(\pi^2 - 12)}{\pi^4}$$

$$\alpha_3 = \frac{60(\pi^2 - 12)}{\pi^5}$$

13.2 Fourier coefficients

In $C[0, 2\pi]$, the set

$$\beta_n = \{\underbrace{\tfrac{1}{\sqrt{2\pi}}}_{g_0}\} \cup \{\underbrace{\tfrac{1}{\sqrt{\pi}} \cos kx \,|\, k = 1, \ldots, n}_{g_1, g_2, \ldots, g_n}\} \cup \{\underbrace{\tfrac{1}{\sqrt{\pi}} \sin kx \,|\, k = 1, \ldots, n}_{g_{n+1}, \ldots, g_{2n}}\},$$

where $n \in \mathbb{N}_0$, is orthonormal with respect to the inner product

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^{2\pi} f(x)g(x)dx.$$

Recall trig identities $\cos A \cos B = \frac{1}{2}(\cos(A-B) + \cos(A+B))$ etc.

Using these, we show, for example with $k \neq k'$

$$\int_0^{2\pi} \cos(kx)\cos(k'x)\,dx = \int_0^{2\pi} \tfrac{1}{2}\cos((k-k')x) + \cos((k+k')x))\,dx$$

$$= \frac{1}{2}\left[\frac{\sin((k-k')x)}{k-k'} + \frac{\sin((k+k')x)}{k+k'}\right]_0^{2\pi} = 0 \text{ etc}$$

Defn: A function of the form

$$p(x) = a_0 + a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx$$
$$+ b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx$$

is called a trigonometric polynomial of degree n if a_n and b_n are not both zero

Problem: Find a least squares approximation of $f \in C[0, 2\pi]$ by a trig. poly. of degree n w.r.t. the given inner product

Solution: Let $W_n = \operatorname{span}(\beta_n)$. The solution is

$$\operatorname{Proj}_{W_n}(f) = \langle f, g_0 \rangle g_0 + \langle f, g_1 \rangle g_1 + \cdots + \langle f, g_{2n} \rangle g_{2n}$$

It follows that β_n is an orthonormal basis for the $(2n+1)$-dimensional subspace $W_n = \text{span}(\beta_n)$ of $C[0, 2\pi]$. The orthogonal projection of $\mathbf{f} \in C[0, 2\pi]$ onto W_n is given by $\text{Proj}_{W_n}(\mathbf{f})$. In the limit $n \to \infty$, the corresponding approximation of $f(x)$ yields the **Fourier series** of $f(x)$ over the interval $[0, 2\pi]$:

(It is possible to show that MSE $\to 0$ as $n \to \infty$)

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

where

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx, \qquad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx,$$

are the associated **Fourier coefficients**.

Find $\langle f, g_m \rangle g_m$ terms: $\quad m = 0, 1, 2, \ldots, n, n+1, \ldots, 2n$

$$\left\langle f(x), \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} = \frac{1}{2\pi} \int_0^{2\pi} f(x)\, dx = \frac{a_0}{2}$$

$$\left\langle f(x), \frac{1}{\sqrt{\pi}} \cos kx \right\rangle \frac{1}{\sqrt{\pi}} \cos kx = \frac{1}{\pi} \left(\int_0^{2\pi} f(x) \cos kx \, dx \right) \cos kx$$

$$= a_k \cos kx$$

& similarly for $\sin kx$ terms

14 Understanding determinants

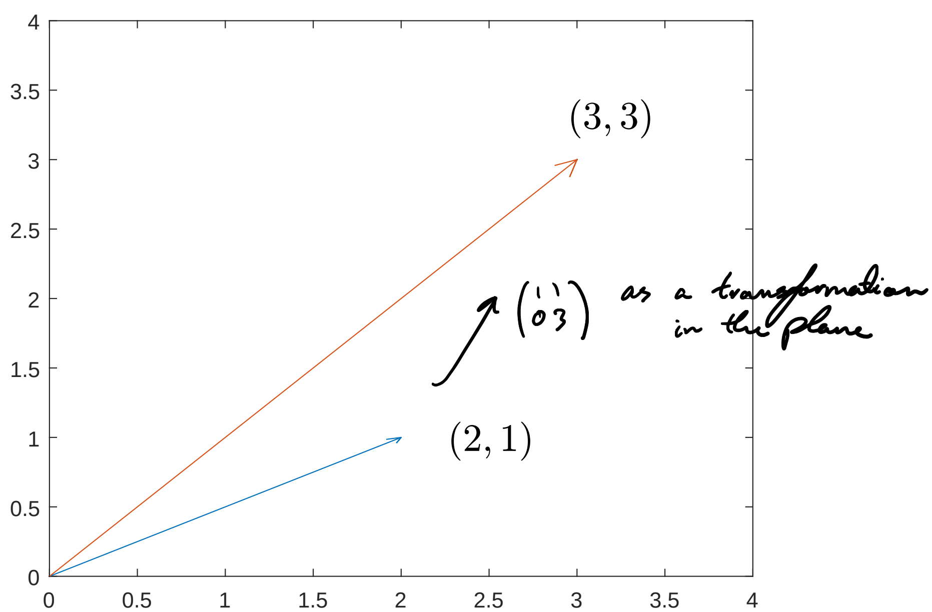
By the end of this section, you should be able to answer the following questions:

- What is the motivation for the definition of the determinant of a 2×2 matrix?
- What does the 2×2 determinant tell us about linear transformations in the plane?

Recall from MATH1051/MATH1071 that a 2×2 matrix can be viewed as a transformation in the plane. For example, the equation

$$\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

can be intepreted as in the following diagram.



A matrix can be viewed as either transforming vectors or *points* in the plane. We also consider what happens to a set of points under such a transformation. For example, a curve, or a region in the plane.

It is also interesting to consider general features of linear transformations. There are two fundamental properties of linear transformations:

1. The origin remains fixed under a linear transformation.
2. Straight lines map to straight lines under a linear transformation.

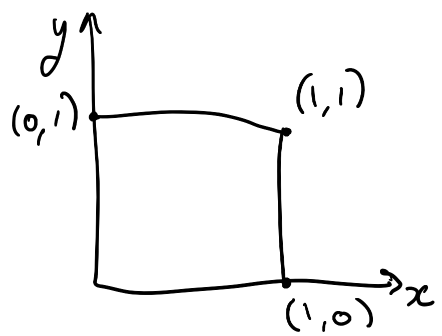
Given the second property, we may consider how to quantify the amount by which a linear transformation stretches or contracts a region in the plane by first consider the image of a unit square.

14.1 The effect of a linear transformation on a grid

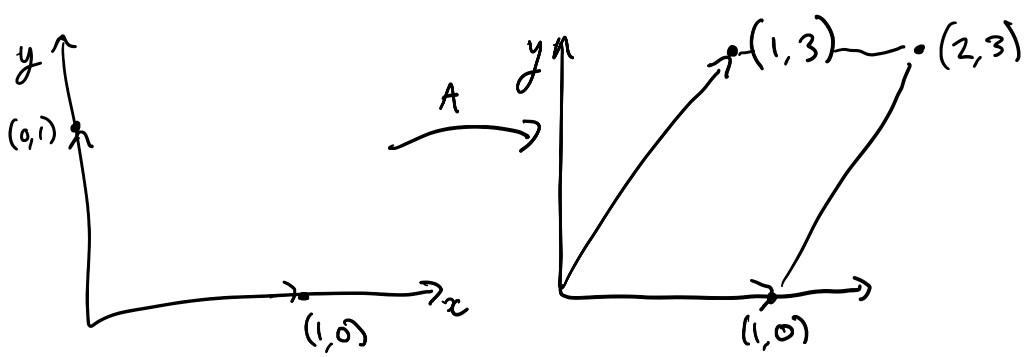
We start with a straightforward example: Describe the effect of the linear transformation $A = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$ on the unit square $S = \{(x, y) \mid 0 \le x, y \le 1\}$.

Every point in S can be represented as a position vector

$\underline{v} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad 0 \le a, b \le 1$

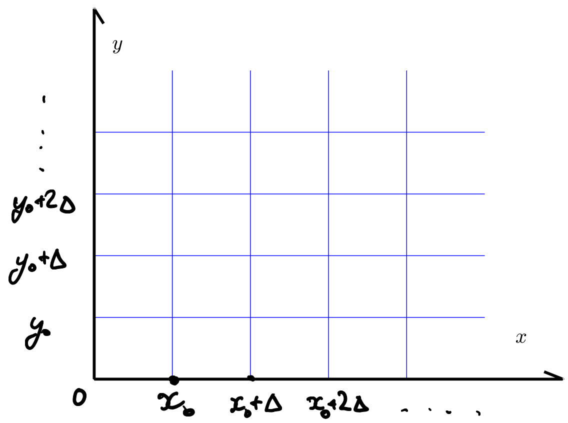


$$A\underline{v} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$



Key is the effect that matrix multiplication has on $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ & $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Now consider a grid of lines parallel to the x and y axes.



These grid lines remain parallel and evenly spaced under a linear transformation.

Vertical gridlines (constant x), $x = x_0 + n\Delta$, $n = 0, 1, 2, \ldots$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_0 + n\Delta \\ y \end{pmatrix} = \begin{pmatrix} ax_0 + an\Delta + by \\ cx_0 + cn\Delta + dy \end{pmatrix}$$

(y varies over $D \subseteq \mathbb{R}$)

$d \neq 0, b = 0 \rightarrow$ lines remain vertical, evenly spaced

$d = 0, b \neq 0 \rightarrow$ lines become horizontal, evenly spaced

$b = 0, d = 0 \rightarrow$ each line contracts to a point

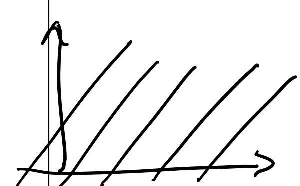
$d, b \neq 0 \rightarrow$ (want an equation relating x' and y' $\rightarrow$ remove y)

$\Rightarrow dx' = adx_0 + adn\Delta + bdy$

$\quad by' = bcx_0 + bcn\Delta + bdy$

$\Rightarrow dx' - by' = (ad - bc)(x_0 + n\Delta)$

$ad - bc \neq 0 \rightarrow$ lines are distinct with a common slope & evenly spaced

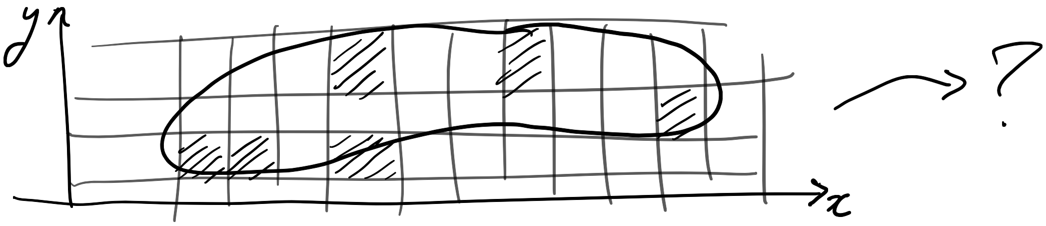


$ad - bc = 0 \rightarrow$ all vertical gridlines map to a common line

Similarly for the <u>horizontal grid lines</u>

14.2 The determinant

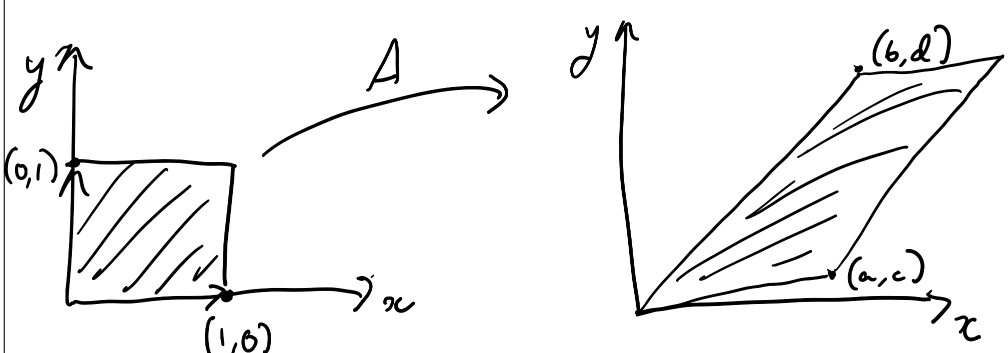
It follows that the effect of a linear transformation on any grid square, regardless of the size, must be the same. It is then helpful to approximate any region in the x-y plane by arbitrarily small grid squares and arrive at the same conclusion.



Under a linear transformation A, the area of any region in the x-y plane scales by the same amount. This amount (up to a sign) is called the *determinant of* A, denoted $|A|$ or $\det(A)$. In the case

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad \det(A) = ad - bc.$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x\begin{pmatrix} a \\ c \end{pmatrix} + y\begin{pmatrix} b \\ d \end{pmatrix}$$



assumed $ad-bc>0$

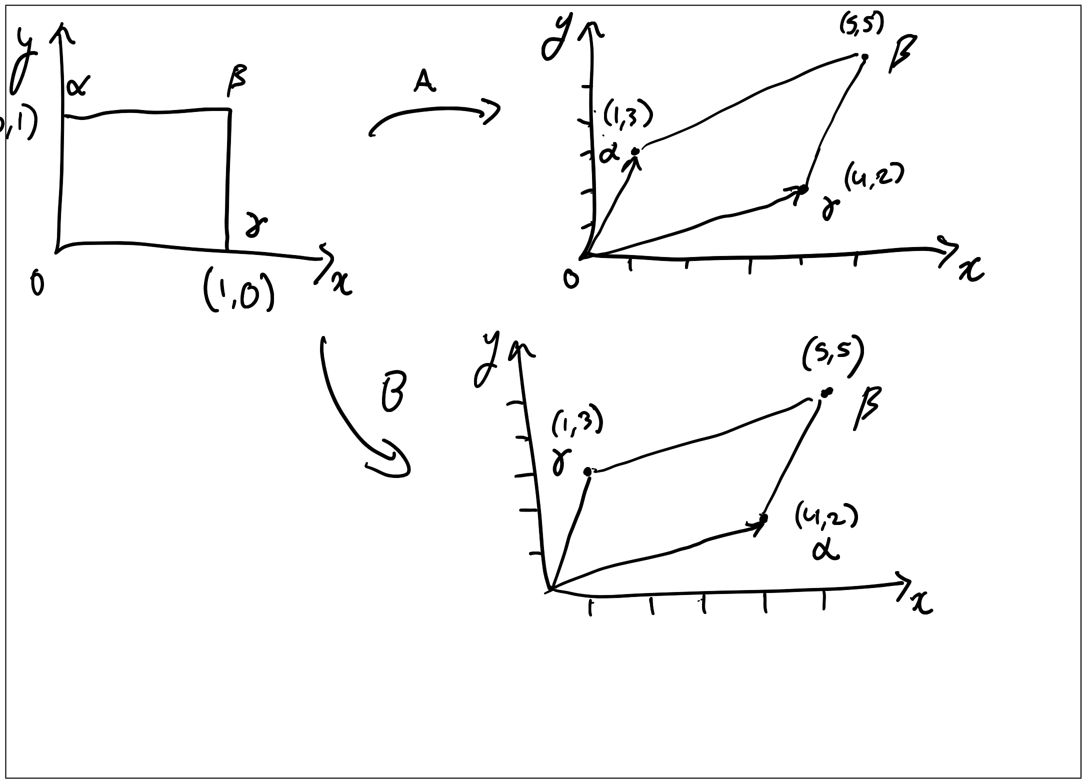
(from MATH1051)

Area of parallelogram $= \left\| \begin{pmatrix} a \\ c \\ 0 \end{pmatrix} \times \begin{pmatrix} b \\ d \\ 0 \end{pmatrix} \right\| = |ad - bc|$

14.3 Sign of the determinant and orientation

For a linear transformation in the plane, A, if $\det(A) < 0$, this implies that the region has undergone a "flip" or change in orientation. This is made clear in the following example.

Example: Compare the effect of $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$ on the unit square $\{(x, y) \mid 0 \le x, y \le 1\}$.



14.4 Further examples

Consider the effect on the unit square under the following linear transformations.

1. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
2. $\begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ x_0+2y \end{pmatrix}$	$\begin{pmatrix} 0 \\ x+2y_0 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$	$\begin{pmatrix} x_0 \\ 2x_0 \end{pmatrix}$	$\begin{pmatrix} x \\ 2x \end{pmatrix}$
4. $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$	$\begin{pmatrix} x_0+2y \\ 2x_0+4y \end{pmatrix}$	$\begin{pmatrix} x+2y_0 \\ 2x+4y_0 \end{pmatrix}$
	Vertical $\begin{pmatrix} x_0 \\ y \end{pmatrix}$	horizontal $\begin{pmatrix} x \\ y_0 \end{pmatrix}$

Notes.

15 Eigenvalues and eigenvectors

By the end of this section, you should be able to answer the following questions:

- How do you find the eigenvalues and eigenvectors of a given square matrix?
- What are some simple properties of eigenvalues and eigenvectors?
- Prove that the eigenvectors corresponding to distinct eigenvalues are linearly independent.

A great deal of this section should be familiar to you. We start by recalling some results on vector spaces associated with matrices.

15.1 Non-singular matrices

For $n \times n$ square matrix A, we have several conditions for the existence of A^{-1}. The the following statements are equivalent:

$AA^{-1} = I = A^{-1}A$

1. A is non-singular. $\hookrightarrow$ A^{-1} exists
2. $A\boldsymbol{x} = \mathbf{0}$ has only the trivial solution $\boldsymbol{x} = \mathbf{0}$.
3. If U is a r.e.f. for A, then U has no row of all zeros.
4. $A\boldsymbol{x} = \boldsymbol{b}$ has a solution for every n-dimensional column vector $\boldsymbol{b}$.
5. $\det(A) \neq 0$.
6. The columns of A are linearly independent.
7. The rows of A are linearly independent.
8. $\text{nullity}(A) = 0$.
9. $\text{rank}(A) = \text{n}$.

$\nearrow$ $\underline{0}$ is NOT an eigenvector

15.2 Eigenvalues and eigenvectors

Let A be a square matrix. Then an *eigenvector* of A is a vector $\boldsymbol{v} \neq \mathbf{0}$ such that

$$A\boldsymbol{v} = \lambda\boldsymbol{v},$$

for some scalar λ. We call λ the *eigenvalue* corresponding to $\boldsymbol{v}$. If $\boldsymbol{v}$ is an eigenvector of A, then so is $t\boldsymbol{v}$ for any scalar $t \neq 0$.

We have

$$A\boldsymbol{v} = \lambda\boldsymbol{v} = \lambda I\boldsymbol{v} \quad \Rightarrow \quad (A - \lambda I)\boldsymbol{v} = \mathbf{0}.$$

Hence $\boldsymbol{x} = \boldsymbol{v}$ is a non-trivial solution to the homogeneous system of equations $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$, and conversely, if there is a non-trivial solution then λ is an eigenvalue of A. Thus:

negation of statements on previous page

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A \quad &\Leftrightarrow \quad (A - \lambda I)\boldsymbol{x} = \mathbf{0} \text{ has a non-trivial solution} \quad \text{(neg. 2.)} \\ &\Leftrightarrow \quad A - \lambda I \text{ is singular} \quad \text{(neg 1.)} \\ &\Leftrightarrow \quad \det(A - \lambda I) = 0. \quad \text{(neg 5.)} \end{aligned}$$

For an $n \times n$ matrix A, $\det(A - \lambda I)$ is a polynomial of degree n in λ, called the *characteristic polynomial* of A. The equation $\det(A - \lambda I) = 0$ is the *characteristic equation* of A.

Eigenvalues λ may be complex numbers, and the eigenvectors $\boldsymbol{v}$ may have complex components, even for real matrices A.

e.g. Consider $W = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

$W^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$

See that $\det(W - \lambda I) = \det\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$

which $= 0$ when $\lambda = \pm i$

To find the eigenvalues and eigenvectors, do the following:

1. Find the roots of the characteristic polynomial, $\det(A - \lambda I) = 0$. These are the eigenvalues.

2. For each eigenvalue λ, find $N(A - \lambda I)$, known as the *eigenspace* associated to λ.

i.e. solve $(A - \lambda I)\underline{v} = \underline{0}$ $\quad\to$ denoted E_λ

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \times \det\begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \times \det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \times \det\begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

15.2.1 Example

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$.

$$\det(A - \lambda I) = \det\begin{pmatrix} -3-\lambda & 1 & 0 \\ 1 & -2-\lambda & 1 \\ 0 & 1 & -3-\lambda \end{pmatrix}$$

$$= (-3-\lambda)\left[(-2-\lambda)(-3-\lambda) - 1\right] - 1\left[-3-\lambda\right]$$

$$= -(3+\lambda)\left[6 + 5\lambda + \lambda^2 - 1 - 1\right]$$

$$= -(3+\lambda)(4+\lambda)(1+\lambda)$$

$\Rightarrow \lambda_1 = -3,\ \lambda_2 = -4,\ \lambda_3 = -1$ (eigenvalues)

Now solve $(A - \lambda I)\underline{v} = \underline{0}$ for each λ

$\underline{\lambda = -3}$: $(A - (-3)I)\underline{v} = \underline{0} \Rightarrow (A + 3I)\underline{v} = \underline{0}$

$\Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $R_1, R_3 \Rightarrow b = 0$

& $R_2 \Rightarrow a + c = 0$

$\Rightarrow \underline{v} = \begin{pmatrix} a \\ 0 \\ -a \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ i.e. $E_{-3} = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right\}$

$\underline{\lambda = -4}$: $(A + 4I)\underline{v} = \underline{0}$

$\Rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\left.\begin{array}{l} R_1 \Rightarrow a + b = 0 \\ R_3 \Rightarrow b + c = 0 \end{array}\right\} \Rightarrow a = -b = c$

$R_2 = R_1 + R_3$

$\Rightarrow \underline{v} = \begin{pmatrix} c \\ -c \\ c \end{pmatrix} = c\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $E_{-4} = \text{span}\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right\}$

Check $\underline{\lambda = -1}$: $E_{-1} = \text{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right\}$

(Properties of det: $\det(AB) = \det(A)\det(B)$
$\det(A^T) = \det(A)$)

15.3 Simple properties

For a square matrix A:

1. A and A^T have the same eigenvalues.

$$\det(A^T - \underbrace{\lambda I}_{\substack{\uparrow \\ \text{diagonal} \\ \text{matrix}}}) = \det((A - \lambda I)^T)$$
$$= \det(A - \lambda I)$$

2. A is singular if and only if $\lambda = 0$ is an eigenvalue of A.

(From 15.1)

A is singular $\Leftrightarrow A\underline{x} = \underline{0}$ has a non-trivial solution
$\Leftrightarrow \lambda = 0$ is an eigenvalue
$(A\underline{v} = 0\underline{v})$

3. If λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2, and $1/\lambda$ is an eigenvalue of A^{-1} when A is non-singular.

$A^2\underline{v} = A(A\underline{v}) = A(\lambda\underline{v}) = \lambda A\underline{v} = \lambda^2\underline{v}$

$A\underline{v} = \lambda\underline{v} \Rightarrow A^{-1}A\underline{v} = A^{-1}(\lambda\underline{v}) \Rightarrow I\underline{v} = \lambda A^{-1}\underline{v} \rightarrow$ result

4. If λ is an eigenvalue of A, then $\lambda - m$ is an eigenvalue of $A - mI$, for any scalar m.

More generally, let $\underline{v}$ be an eigenvector of A & B corresponding to eigenvalues λ_A & λ_B respectively.
For any $k \in \mathbb{R}$, $(A + kB)\underline{v} = A\underline{v} + kB\underline{v}$
$= \lambda_A\underline{v} + k\lambda_B\underline{v}$
$= (\lambda_A + k\lambda_B)\underline{v}$

"LI" = "linearly independent"

15.4 Eigenvectors corresponding to distinct eigenvalues are linearly independent

If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A, with corresponding eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ (such that $\boldsymbol{v}_i$ corresponds to λ_i), then the set of eigenvectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is linearly independent.

Use induction · $k=1$. $\{\underline{v}_1\}$ is LI

since $\underline{v}_1 \neq 0$ $(a_1\underline{v}_1 = 0 \to a_1 = 0)$

assume true for $k = n \geq 1$

(i.e. assume $\{\underline{v}_1, \cdots, \underline{v}_n\}$ is LI) & aim to prove the statement is true for $k = n+1$

i.e. show that $\{\underline{v}_1, \cdots, \underline{v}_n, \underline{v}_{n+1}\}$ is LI

Set $a_1\underline{v}_1 + a_2\underline{v}_2 + \ldots + a_n\underline{v}_n + a_{n+1}\underline{v}_{n+1} = \underline{0}$ $(*)$

$\to$ (mult $(*)$ by A from left)

$\Rightarrow a_1A\underline{v}_1 + a_2A\underline{v}_2 + \cdots + a_nA\underline{v}_n + a_{n+1}A\underline{v}_{n+1} = A\underline{0} = \underline{0}$

$\Rightarrow a_1\lambda_1\underline{v}_1 + a_2\lambda_2\underline{v}_2 + \cdots + a_n\lambda_n\underline{v}_n + a_{n+1}\lambda_{n+1}\underline{v}_{n+1} = \underline{0}$ (1)

$\Rightarrow$ (multiply $(*)$ by $\lambda_{n+1} \neq 0$)

$\Rightarrow a_1\lambda_{n+1}\underline{v}_1 + \cdots + a_n\lambda_{n+1}\underline{v}_n + a_{n+1}\lambda_{n+1}\underline{v}_{n+1} = \underline{0}$ (2)

(2) $-$ (1) $\Rightarrow a_1(\lambda_{n+1}-\lambda_1)\underline{v}_1 + a_2(\lambda_{n+1}-\lambda_2)\underline{v}_2 + \ldots + a_n(\lambda_{n+1}-\lambda_n)\underline{v}_n = \underline{0}$

but $\{\underline{v}_1, \cdots, \underline{v}_n\}$ is LI

$\Rightarrow a_1(\lambda_{n+1}-\lambda_1) = 0 = a_2(\lambda_{n+1}-\lambda_2) = \ldots = a_n(\lambda_{n+1}-\lambda_n)$

$\Rightarrow a_1 = 0 = a_2 = \ldots = a_n$ since λ_i are distinct

$\Rightarrow$ $(*)$ becomes $a_{n+1}\underline{v}_{n+1} = \underline{0} \Rightarrow a_{n+1} = 0$

since $\underline{v}_{n+1} \neq 0$

$\Rightarrow \{\underline{v}_1, \cdots, \underline{v}_{n+1}\}$ is LI $\to$ result by induction

Notes. Induction

Statement S_k, $k=1,2,\ldots$

Suppose the following hold

1. S_1 is true
2. $\forall n \geq 1$, if S_n is true then S_{n+1} is true

Then the statement "$\forall k \geq 1, S_k$" is true".

In proof: 1 show S_1 is true

2 Show that is S_n is true, then S_{n+1} is true $\forall n \geq 1$

$\rightarrow$ result follows.

16 Diagonalisation

By the end of this section, you should be able to answer the following questions:

- How do you find a matrix P which diagonalises a given matrix A?
- How do you determine if A is diagonalisable?
- What are two applications of diagonalisation?

$\to P^{-1}$ exists

A square matrix A is *diagonalisable* if there is a non-singular matrix P such that $P^{-1}AP$ is a diagonal matrix. Here we consider the question: given a matrix, is it diagonalisable? If so, how do we find P?

The secret to constructing such a P is to let the columns of P be the eigenvectors of A. We immediately have that $AP = PD$, where D is a diagonal matrix with eigenvalues on the diagonal. We know from section 15.1 on page 111 that P is invertible if and only if the columns of P are linearly independent. Hence, we have the following result:

15.1 (1) $\leftrightarrow$ (6)

The $n \times n$ matrix A is diagonalisable if and only if A has n linearly independent eigenvectors.

Is the matrix $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ diagonalisable?

From p113

Eigenvalues are $-3, -4, -1$ which are distinct

This 3×3 matrix has 3 distinct eigenvalues

$\Rightarrow$ 3 linearly independent eigenvectors

Form P with eigenvectors as columns

$\Rightarrow P^{-1}$ exists $\Rightarrow$ A is diagonalisable

$\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \leftrightarrow -3$, $\underline{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \leftrightarrow -4$, $\underline{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \leftrightarrow -1$

$\Rightarrow P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ -1 & 1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

such that $P^{-1}AP = D$

16.1 Similar matrices (reading)

Two matrices A and B are *similar* if there is a non-singular matrix P such that $B = P^{-1}AP$.

The statements "A is diagonalisable" and "A is similar to a diagonal matrix" are equivalent.

16.1.1 Theorem (similar matrices)

Similar matrices have the same eigenvalues.

In fact, if $B = P^{-1}AP$ and $\boldsymbol{v}$ is an eigenvector of A corresponding to eigenvalue λ, then $P^{-1}\boldsymbol{v}$ is an eigenvector of B corresponding to eigenvalue λ. This is because

$$\begin{aligned} B(P^{-1}\boldsymbol{v}) &= (P^{-1}AP)P^{-1}\boldsymbol{v} \\ &= P^{-1}(A\boldsymbol{v}) \\ &= P^{-1}(\lambda\boldsymbol{v}) \\ &= \lambda(P^{-1}\boldsymbol{v}) \end{aligned}$$

16.2 A closer look at the diagonal matrix

Let the matrix A be $n \times n$ with n linearly independent eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Let

$$P = (\boldsymbol{v}_1|\ldots|\boldsymbol{v}_n)$$

be the $n \times n$ matrix whose columns are the eigenvectors. Then

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

the diagonal matrix with the eigenvalues down the main diagonal. The important point here is the order in which the eigenvalues appear. They correspond to the order in which the associated eigenvectors appear in the columns of P.

e.g. 3x3 case

$\underline{v}_1, \underline{v}_2, \underline{v}_3$ eigenvectors corr. to $\lambda_1, \lambda_2, \lambda_3$ respectively

$$\begin{aligned} AP &= A(\underline{v}_1|\underline{v}_2|\underline{v}_3) \quad \text{matrix mult.} \\ &= (A\underline{v}_1|A\underline{v}_2|A\underline{v}_3) \\ &= (\lambda_1\underline{v}_1|\lambda_2\underline{v}_2|\lambda_3\underline{v}_3) = (\underline{v}_1|\underline{v}_2|\underline{v}_3)\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = PD \end{aligned}$$

i.e. $AP = PD$ $\Rightarrow P^{-1}AP = D$

$A = PDP^{-1}$

16.3 Diagonalisability

We know that an $n \times n$ matrix A is diagonalisable if and only if A has n linearly independent eigenvectors.

Now say $\lambda_1, \ldots, \lambda_m$ are *distinct* eigenvalues of A, with corresponding eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$. Then we have also seen that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ are linearly independent.

Hence if A is $n \times n$ with n distinct eigenvalues, then A is diagonalisable.

The question remains, if A has fewer than n distinct eigenvalues, how do we know if A is diagonalisable?

16.3.1 Example

Let $A = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Easy to see the characteristic equation of both A and B is $(2-\lambda)(1-\lambda)^2 = 0$, so $\lambda = 2, 1, 1$.

Solve $(A - \lambda I)\underline{v} = 0$

$\underline{\lambda = 2}$: $\begin{pmatrix} 0 & 1 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $R_2 \Rightarrow b = 0$, $R_3 \Rightarrow c = 0$, no constraint for a

$\Rightarrow \underline{v}_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$\underline{\lambda = 1}$: $\begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $R_1 \Rightarrow a + b + 3c = 0 \Rightarrow a = -b - 3c$

$\Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b-3c \\ b \\ c \end{pmatrix} = b\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$

$\uparrow \underline{v}_2$ $\uparrow \underline{v}_3$

linearly independent $\rightarrow$ $\underline{v}_2$, $\underline{v}_3$

$\Rightarrow$ A is diagonalisable

Matrix B solve $(B-\lambda I)\underline{v}=0$

$\underline{\lambda=2}$: $\begin{pmatrix} 0 & 1 & 3 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $R_3 \to c=0$

$R_2 \Rightarrow -b+c=0 \to b=0$

a is free

$\Rightarrow \underline{v}_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$\underline{\lambda=1}$: $\begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $R_2 \to c=0$

$R_1 \to a+b+3c=0$

$\Rightarrow a+b=0 \Rightarrow a=-b$

$\Rightarrow \underline{v}_2 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b \\ b \\ 0 \end{pmatrix} = b\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$

$\Rightarrow B$ only has two LI eigenvectors

$\Rightarrow B$ is <u>NOT</u> diagonalisable

16.4 Algebraic and geometric multiplicity

If we are only interested in finding out whether or not a matrix is diagonalisable, then we need to know the dimension of each eigenspace. There is one theorem (which we will not prove!) that states:

If λ_i is an eigenvalue, then the dimension of the corresponding eigenspace cannot be greater than the number of times $(\lambda - \lambda_i)$ appears as a factor in the characteristic polynomial.

We often use the following terminology:

- The *geometric multiplicity* of the eigenvalue λ_i is the dimension of the eigenspace corresponding to λ_i.
- The *algebraic multiplicity* of the eigenvalue λ_i is the number of times $(\lambda - \lambda_i)$ appears as a factor in the characteristic polynomial.

The main result is the following:

> A square matrix is diagonalisable if and only if the geometric and algebraic multiplicities are equal for every eigenvalue.

From previous example: Characteristic polynomial is $(1-\lambda)^2(2-\lambda)$

	eigenvalues	algebraic mult.	geometric mult	Diagonalisable?
A	$\lambda=1$	2	2	Yes
	$\lambda=2$	1	1	
B	$\lambda=1$	2	1 *	No
	$\lambda=2$	1	1	

e.g. 5×5 matrix

Char. poly. $(\lambda-2)^3(\lambda-1)(\lambda-5)$

$\rightarrow$ only need to inspect values with alge. mult. >1

16.5 Applications of diagonalisability

16.5.1 Systems of differential equations

For a system of coupled differential equations which can be written in matrix form as

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}$$

(where $\boldsymbol{x} = (x_1, \ldots, x_n)^T$, $\dot{\boldsymbol{x}} = (\dot{x}_1, \ldots, \dot{x}_n)^T$),
if A can be diagonalised, say $P^{-1}AP = D$ with D diagonal, then make the substitution $\boldsymbol{x} = P\boldsymbol{y}$. This yields

$$\dot{\boldsymbol{y}} = D\boldsymbol{y}$$

which is easily solved.

$$\left.\begin{aligned} \dot{x}_1 &= x_1 + 2x_2 \\ \dot{x}_2 &= 2x_1 + x_2 \end{aligned}\right\} \quad \text{coupled system of ODEs}$$

Write $\underline{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$, $\dot{\underline{x}}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{or} \quad \dot{\underline{x}} = A\underline{x}$$

Q. Is A diagonalisable? Yes!

$$AP = PD \Rightarrow A = PDP^{-1}$$

$$\dot{\underline{x}} = A\underline{x} \Rightarrow \dot{\underline{x}} = PDP^{-1}\underline{x} \Rightarrow P^{-1}\dot{\underline{x}} = D(P^{-1}\underline{x}) \quad *$$

suggests introducing $\underline{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = P^{-1}\underline{x} \Rightarrow \dot{\underline{y}} = P^{-1}\dot{\underline{x}}$

$$(*) \text{ gives } \dot{\underline{y}} = D\underline{y} \Rightarrow \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\left.\begin{aligned} \Rightarrow \quad & \dot{y}_1 = \lambda_1 y_1 \\ \& \quad & \dot{y}_2 = \lambda_2 y_2 \end{aligned}\right\} \text{ solve for } y_1, y_2 \ \& \ \underline{y} = P^{-1}\underline{x}$$

$$\Rightarrow \underline{x} = P\underline{y}$$

16.5.2 Matrix powers

If A is diagonalisable, say $P^{-1}AP = D$ with D diagonal, then

$$A^n = PD^nP^{-1}.$$

This gives an easy way to calculate A^n.

$$A^n = \underbrace{PDP^{-1} \cdot PDP^{-1} \ldots PDP^{-1}}_{n \text{ times}}$$

$$= PD^nP^{-1}$$

$$D^n = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \lambda_2 & \vdots \\ & \ddots & \\ 0 & \cdots & \lambda_m \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & \cdots & 0 \\ \vdots & \lambda_2^n & \vdots \\ & \ddots & \\ 0 & \cdots & \lambda_m^n \end{pmatrix}$$

16.5.3 Linear systems

What can we say about the linear system (corresponding to a square matrix A)

$$A\boldsymbol{x} = \boldsymbol{b},$$

if we know how to diagonalise A?

$$A = PDP^{-1}$$

$$\Rightarrow A\underline{x} = \underline{b}$$

$$\Rightarrow PDP^{-1}\underline{x} = \underline{b}$$

$$\Rightarrow D(P^{-1}\underline{x}) = P^{-1}\underline{b} = \underline{b}'$$

set $\underline{y} = P^{-1}\underline{x} \Rightarrow D\underline{y} = \underline{b}'$

Notes.

17 Orthogonal matrices

The goal of this section is to investigate orthogonal matrices and to see their relationship with isometries, i.e. functions that preserve distance.

17.1 Transpose of a transition matrix

Using the inner product on $M_{2,2}(\mathbb{R})$ discussed in Section 8.5, we noted that

$$\langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_\beta \cdot [\mathbf{v}]_\beta,$$

where β is the standard ordered basis for $M_{2,2}(\mathbb{R})$. This generalises to any finite-dimensional real inner product space V with orthonormal basis β.

Let $\beta = \{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ be an orthonormal basis for V and let $\underline{u}, \underline{v} \in V$. Then,

$$\underline{u} = \sum_{i=1}^{n} u_i \underline{e}_i, \quad \underline{v} = \sum_{i=1}^{n} v_i \underline{e}_i$$

(note that we know more! e.g. $u_i = \langle \underline{u}, \underline{e}_i \rangle$)

with coordinate vectors

$$[\underline{u}]_\beta = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad [\underline{v}]_\beta = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Then

$$\begin{aligned} \langle \underline{u}, \underline{v} \rangle &= \left\langle \sum_{i=1}^{n} u_i \underline{e}_i, \sum_{j=1}^{n} v_j \underline{e}_j \right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} u_i v_j \langle \underline{e}_i, \underline{e}_j \rangle \\ &= \sum_{i=1}^{n} u_i v_i \qquad \text{since } \langle e_i, e_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \\ &= [\underline{u}]_\beta \cdot [\underline{v}]_\beta \end{aligned}$$

If β, β' are orthonormal bases for V, we then have

$$[\mathbf{u}]_\beta \cdot [\mathbf{v}]_\beta = \langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_{\beta'} \cdot [\mathbf{v}]_{\beta'}.$$

$\nearrow$ Defin

Since the transition matrix $P_{\beta\to\beta'}$ relates the coordinate vectors as $[\mathbf{u}]_{\beta'} = P_{\beta\to\beta'}[\mathbf{u}]_\beta$, it follows that

$$\begin{aligned}[\mathbf{u}]_\beta \cdot [\mathbf{v}]_\beta = [\mathbf{u}]_{\beta'} \cdot [\mathbf{v}]_{\beta'} &= (P_{\beta\to\beta'}[\mathbf{u}]_\beta) \cdot (P_{\beta\to\beta'}[\mathbf{v}]_\beta) = (P_{\beta\to\beta'}[\mathbf{u}]_\beta)^T (P_{\beta\to\beta'}[\mathbf{v}]_\beta) \\ &= ([\mathbf{u}]_\beta)^T (P_{\beta\to\beta'})^T P_{\beta\to\beta'}[\mathbf{v}]_\beta = [\mathbf{u}]_\beta \cdot \big((P_{\beta\to\beta'})^T P_{\beta\to\beta'}[\mathbf{v}]_\beta\big).\end{aligned}$$

Now, could it be that $(P_{\beta\to\beta'})^T P_{\beta\to\beta'}$ is simply the identity matrix? It is! Since transition matrices are invertible, this means that

$$(P_{\beta\to\beta'})^T = (P_{\beta\to\beta'})^{-1}.$$

$\ast$ True for orthonormal bases, β & β'

Let $\beta = \{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}$, $\beta' = \{\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_n\}$ be orthonormal bases. Then,

$$P_{\beta\to\beta'} = \left([\underline{e}_1]_{\beta'} \,\middle|\, [\underline{e}_2]_{\beta'} \,\middle|\, \cdots \,\middle|\, [\underline{e}_n]_{\beta'} \right)$$

$$= \begin{pmatrix} \langle \underline{e}_1, \underline{f}_1 \rangle & \langle \underline{e}_2, \underline{f}_1 \rangle & \cdots & \langle \underline{e}_n, \underline{f}_1 \rangle \\ \langle \underline{e}_1, \underline{f}_2 \rangle & \langle \underline{e}_2, \underline{f}_2 \rangle & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{e}_1, \underline{f}_n \rangle & \langle \underline{e}_2, \underline{f}_n \rangle & \cdots & \langle \underline{e}_n, \underline{f}_n \rangle \end{pmatrix}$$

$$\Rightarrow (P_{\beta\to\beta'})^T = \begin{pmatrix} \langle \underline{f}_1, \underline{e}_1 \rangle & \langle \underline{f}_2, \underline{e}_1 \rangle & \cdots & \langle \underline{f}_n, \underline{e}_1 \rangle \\ \langle \underline{f}_1, \underline{e}_2 \rangle & \langle \underline{f}_2, \underline{e}_2 \rangle & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{f}_1, \underline{e}_n \rangle & \langle \underline{f}_2, \underline{e}_n \rangle & \cdots & \langle \underline{f}_n, \underline{e}_n \rangle \end{pmatrix}$$

$$= \left([\underline{f}_1]_{\beta} \,\middle|\, [\underline{f}_2]_{\beta} \,\middle|\, \cdots \,\middle|\, [\underline{f}_n]_{\beta} \right)$$

$$= P_{\beta'\to\beta} = (P_{\beta\to\beta'})^{-1}$$

17.2 Orthogonal matrices

A square matrix Q with real entries is called **orthogonal** if it is invertible and

$$Q^{-1} = Q^T,$$

that is, if $Q^TQ = QQ^T = I$. As shown in Section 17.1, transition matrices between orthonormal bases are examples of orthogonal matrices. Special cases are

$$Q_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \tilde{Q}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \qquad \theta \in \mathbb{R}.$$

Orthogonal matrices have their name because of the following property. Let $(\mathbf{v}_1 \,|\, \dots \,|\, \mathbf{v}_n)$ denote an $n \times n$ matrix with column vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then,

$$(\mathbf{v}_1 \,|\, \dots \,|\, \mathbf{v}_n) \text{ is orthogonal} \quad \Longleftrightarrow \quad \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \text{ is an } \underline{\text{orthonormal set}}. \qquad \underline{v}_i \in \mathbb{R}^n$$

(w.r.t. Euclidean inner product)

Consider $(\underline{v}_1 | \underline{v}_2 | \dots | \underline{v}_n)^T (\underline{v}_1 | \underline{v}_2 | \dots | \underline{v}_n)$

$$= \begin{pmatrix} \underline{v}_1^T \\ \hline \underline{v}_2^T \\ \hline \vdots \\ \hline \underline{v}_n^T \end{pmatrix} (\underline{v}_1 | \underline{v}_2 | \dots | \underline{v}_n)$$

$$= \begin{pmatrix} \underline{v}_1^T\underline{v}_1 & \underline{v}_1^T\underline{v}_2 & \cdots & \underline{v}_1^T\underline{v}_n \\ \underline{v}_2^T\underline{v}_2 & \underline{v}_2^T\underline{v}_2 & & \underline{v}_2^T\underline{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \underline{v}_n^T\underline{v}_n & \underline{v}_n^T\underline{v}_2 & \cdots & \underline{v}_n^T\underline{v}_n \end{pmatrix}$$

$$= \begin{pmatrix} \underline{v}_1\cdot\underline{v}_1 & \underline{v}_1\cdot\underline{v}_2 & \cdots & \underline{v}_1\cdot\underline{v}_n \\ \underline{v}_2\cdot\underline{v}_1 & \underline{v}_2\cdot\underline{v}_2 & & \underline{v}_2\cdot\underline{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \underline{v}_n\cdot\underline{v}_1 & \underline{v}_2\cdot\underline{v}_n & \cdots & \underline{v}_n\cdot\underline{v}_n \end{pmatrix}$$

$\rightsquigarrow$ result

17.3 Isometries

(w.r.t. Euclidean Inner Product

An **isometry** of $\mathbb{R}^n$ is a function $h : \mathbb{R}^n \to \mathbb{R}^n$ that preserves distances:

$$||h(\mathbf{u}) - h(\mathbf{v})|| = ||\mathbf{u} - \mathbf{v}||, \qquad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

As one can show, every isometry of $\mathbb{R}^n$ can be uniquely written as a composition of a translation and an isometry that fixes $\mathbf{0}$ (the origin of $\mathbb{R}^n$). Examples of isometries that fix $\mathbf{0}$ are rotations and reflections. Concretely, the matrix Q_θ above generates rotations by θ about the origin in the plane $\mathbb{R}^2$, while $\hat{Q}_\theta$ generates reflections about the line through the origin, turned the angle $\theta/2$.

For a function $h : \mathbb{R}^n \to \mathbb{R}^n$, the following are equivalent:

(i) h is an isometry and $h(\mathbf{0}) = \mathbf{0}$;

(ii) $h(\mathbf{u}) \cdot h(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}, \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$ $\leftarrow$ preserves angles, distance

(i) $\Rightarrow$ (ii): For any $\underline{u}, \underline{v} \in \mathbb{R}^n$

$$\|h(\underline{u}) - h(\underline{v})\| = \|\underline{u} - \underline{v}\| \qquad (*)$$

For $\underline{v} = \underline{0}$ $\Rightarrow \|h(\underline{u}) - h(\underline{0})\| = \|\underline{u} - \underline{0}\| \Rightarrow \|h(\underline{u})\| = \|\underline{u}\|$

Square $(*)$ $\to$ $(h(\underline{u}) - h(\underline{v})) \cdot (h(\underline{u}) - h(\underline{v})) = (\underline{u} - \underline{v}) \cdot (\underline{u} - \underline{v})$

$\Rightarrow \|h(\underline{u})\|^2 - 2h(\underline{u}) \cdot h(\underline{v}) + \|h(\underline{u})\|^2 = \|\underline{u}\|^2 - 2\underline{u} \cdot \underline{v} + \|\underline{v}\|^2$

$\Rightarrow h(\underline{u}) \cdot h(\underline{v}) = \underline{u} - \underline{v}$

(ii) $\Rightarrow$ (i):

$\Rightarrow \|h(\underline{u}) - h(\underline{v})\|^2 = h(\underline{u}) \cdot h(\underline{u}) - 2h(\underline{u}) \cdot h(\underline{v}) + h(\underline{v}) \cdot h(\underline{v})$

$= \underline{u} \cdot \underline{u} - 2\underline{u} \cdot \underline{v} + \underline{v} \cdot \underline{v}$

$= \|\underline{u} - \underline{v}\|^2$

$\Rightarrow$ so h is an isometry

Set $\underline{u} = \underline{v} = \underline{0}$ in (ii)

$\Rightarrow \|h(\underline{0})\|^2 = \underline{0} \cdot \underline{0} = \underline{0} \Rightarrow h(\underline{0}) = \underline{0}$

17.4 Orthogonal matrices generate isometries

For $Q \in M_{n,n}(\mathbb{R})$, the following are equivalent:

(i) Q is orthogonal;

(ii) $||Q\mathbf{v}|| = ||\mathbf{v}||, \quad \forall\, \mathbf{v} \in \mathbb{R}^n$;

(iii) $(Q\mathbf{u}) \cdot (Q\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}, \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Thus, isometries of $\mathbb{R}^n$ that fix $\mathbf{0}$ are generated by orthogonal matrices, and vice versa.

(i) ⇒ (ii):

$$\|Q\underline{v}\|^2 = (Q\underline{v})\cdot(Q\underline{v}) = (Q\underline{v})^T Q\underline{v} = \underline{v}^T \overbrace{Q^T Q}^{I} \underline{v} = \underline{v}^T\underline{v} = \|\underline{v}\|^2$$

(ii) ⇒ (iii): Let $\underline{u}, \underline{v} \in \mathbb{R}^n$

Note that
$$(\underline{u}+\underline{v})\cdot(\underline{u}+\underline{v}) = \underline{u}\cdot\underline{u} + 2\underline{u}\cdot\underline{v} + \underline{v}\cdot\underline{v}$$
$$(\underline{u}-\underline{v})\cdot(\underline{u}-\underline{v}) = \underline{u}\cdot\underline{u} - 2\underline{u}\cdot\underline{v} + \underline{v}\cdot\underline{v}$$
(difference) ⇒ $4\underline{u}\cdot\underline{v} = \|\underline{u}+\underline{v}\|^2 - \|\underline{u}-\underline{v}\|^2$

$$(Q\underline{u})\cdot(Q\underline{v}) = \frac{1}{4}\left(\|Q\underline{u}+Q\underline{v}\|^2 - \|Q\underline{u}-Q\underline{v}\|^2\right)$$
$$= \frac{1}{4}\left(\|Q(\underline{u}+\underline{v})\|^2 - \|Q(\underline{u}-\underline{v})\|^2\right)$$
$$= \frac{1}{4}\left(\|\underline{u}+\underline{v}\|^2 - \|\underline{u}-\underline{v}\|^2\right) = \underline{u}\cdot\underline{v}$$

(iii) ⇒ (i): Let $\beta = \{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}$ be the standard basis for $\mathbb{R}^n$ (i.e. $\underline{e}_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \underline{e}_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots$)

For any $A \in M_{n,n}(\mathbb{R})$, the entries
$$A_{ij} = \underline{e}_i^T A \underline{e}_j$$
$$\underline{e}_i\cdot\underline{e}_j = (Q\underline{e}_i)\cdot(Q\underline{e}_j) = \underline{e}_i^T Q^T Q \underline{e}_j = (Q^T Q)_{ij}$$
β is orthonormal ⇒ $Q^T Q = I$

Key: Orthogonal matrices leave distances and angles unchanged

Notes.

18 Orthogonal diagonalisation

By the end of this section, you should be able to answer the following questions:

- What is a symmetric matrix?
- How do you diagonalise symmetric matrices?

Given an $n \times n$ matrix A, we call A *orthogonally diagonalisable* if there exists an orthogonal matrix P such that $P^{-1}AP = P^TAP$ is diagonal. To understand this, we first need to know what is meant by an orthogonal matrix.

18.1 Orthogonal matrices

Recall from the previous chapter that an *orthogonal* matrix is a real square matrix Q such that the columns of Q are mutually orthogonal unit vectors with respect to the Euclidean inner product (i.e. $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ if $i \neq j$, and $||\boldsymbol{v}_i|| = 1$).

An orthogonal matrix is then a real square matrix Q such that $Q^{-1} = Q^T$. An immediate consequence of this is that $\det(Q) = \pm 1$.

Q orthogonal $\Rightarrow QQ^T = I$

$\Rightarrow \det(QQ^T) = \det(I)$

$\Rightarrow \det(Q)\det(Q^T) = 1$

$\Rightarrow (\det(Q))^2 = 1 \Rightarrow \det(Q) = \pm 1$

(Aside: in $\mathbb{R}^2$: orthogonal matrix, $\det = 1 \Leftrightarrow$ rotation

" " , $\det = -1 \Leftrightarrow$ reflection)

18.2 Symmetric matrices

A matrix A is *symmetric* if and only if $A = A^T$. Symmetric matrices are easy to identify due to their "mirror symmetry" about the main diagonal. For example, we can tell by inspection that $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ is symmetric.

i.e. entries are real numbers ($A \in M_{n \times n}(\mathbb{R})$)

18.2.1 **If A is real symmetric, then the eigenvectors corresponding to different eigenvalues are orthogonal.** w.r.t. Euclidean Inner product

Proof:

$A\underline{v}_1 = \lambda_1 \underline{v}_1$, $A\underline{v}_2 = \lambda_2 \underline{v}_2$, $\lambda_1 \neq \lambda_2$, $A = A^T$

Want to show $\underline{v}_1 \cdot \underline{v}_2 = 0$ $\quad$ $\underline{v}_1, \underline{v}_2 \in \mathbb{R}^n$ (column vectors)

or equivalently $\underline{v}_1^T \underline{v}_2 = 0$

Assume $\lambda_1 \neq 0$

Consider
$$\begin{aligned}\lambda_1 \underline{v}_1^T \underline{v}_2 &= (\lambda_1 \underline{v}_1)^T \underline{v}_2 \\ &= (A\underline{v}_1)^T \underline{v}_2 \\ &= \underline{v}_1^T A^T \underline{v}_2 \\ &= \underline{v}_1^T A \underline{v}_2 \\ &= \underline{v}_1^T (\lambda_2 \underline{v}_2)\end{aligned}$$

properties of transpose

$$\lambda_1 \underline{v}_1^T \underline{v}_2 = \lambda_2 \underline{v}_1^T \underline{v}_2$$

$\Rightarrow (\lambda_1 - \lambda_2)\underline{v}_1^T \underline{v}_2 = 0$

since $\lambda_1 - \lambda_2 \neq 0$, we must have

$$\underline{v}_1^T \underline{v}_2 = 0$$

$\rightarrow$ result

18.2.2 Real symmetric matrices are orthogonally diagonalisable

It is straightforward to show that if a matrix is orthogonally diagonalisable, then it is symmetric:

$\Rightarrow \exists P$ such that $P^{-1} = P^T$ & $P^TAP = D$ (or $A = PDP^T$) $\Big\}$ $AP = PD$

(Here A is an $n \times n$ matrix, D is diagonal)

$$
\begin{aligned}
A^T &= (PDP^T)^T \\
&= (P^T)^T D^T P^T \\
&= PDP^T
\end{aligned}
$$

$A^T = A$ i.e. A is symmetric

Contrapositive: $x \Rightarrow y \rightarrow$ "Not y" $\Rightarrow$ "Not x"

If A is NOT symmetric, then A is not orthogonally diagonalisable

In fact, the converse is also true (although difficult to prove), giving us the amazing result:

An $n \times n$ real matrix is orthogonally diagonalisable if and only if it symmetric.

The significance of this is that a symmetric matrix is *always* diagonalisable by an orthogonal matrix.

18.2.3 Eigenvectors and eigenvalues

Here we state two results about any symmetric matrix A without proof:

(1) All the eigenvalues of A are real; $\longrightarrow$ see CH20

(2) A has n linearly independent eigenvectors.

18.2.4 Example

Let $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ (see previous examples).

We already know the eigenvalues are $-3, -1, -4$ with corresponding eigenvectors

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Note that A is real symmetric, so $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ should be pairwise orthogonal.

Check $\underline{v}_i \cdot \underline{v}_j = 0 \quad i \neq j$

$$\|\underline{v}_1\| = \sqrt{\underline{v}_1 \cdot \underline{v}_1} = \sqrt{1^2 + 0^2 + (-1)^2} = \sqrt{2}$$

$$\|\underline{v}_2\| = \sqrt{6} \qquad \|\underline{v}_3\| = \sqrt{3}$$

Normalise eigenvectors (i.e. turn them into unit vectors):

$$\hat{\underline{v}}_i = \frac{\underline{v}_i}{\|\underline{v}_i\|} \qquad i = 1, 2, 3$$

$$\hat{\underline{v}}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}, \quad \hat{\underline{v}}_2 = \begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad \hat{\underline{v}}_3 = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \Rightarrow \text{form an orthonormal set}$$

$$\Rightarrow P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \left(= \left(\hat{\underline{v}}_1 \middle| \hat{\underline{v}}_2 \middle| \hat{\underline{v}}_3 \right) \right)$$

such that $P^T A P = D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}$

Notes.

19 Quadratic forms

By the end of this section, you should be able to answer the following questions:

- What is a quadratic form?
- How do you diagonalise quadratic forms?
- How can you use diagonalisation of two variable quadratic forms to identify conic sections?
- What are quadric surfaces?

This section presents a novel application of orthogonal diagonalisation as a way of identifying conic sections. We also mention the generalisation to three dimensions and how, in principle, we could identify quadric surfaces, although the details in this case can become quite messy.

The majority of this section is based on the section on quadratic forms in the MATH2001 recommended text "Elementary Linear Algebra (Applications Version)" by Anton and Rorres, pages 479–502.

19.1 Definition

Consider n real variables $x_1, x_2, \ldots, x_n$. A function of the form $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$ is called a *quadratic form*, where the a_{ij} are real constants.

For example, the most general quadratic form in the variables x and y is

$$Q(x, y) = ax^2 + by^2 + cxy.$$

In the three variables x, y and z, the most general quadratic form is

$$Q(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz,$$

where in both cases a, b, c, d, e, f are all constants. It is possible to express quadratic forms in n variables as a matrix product $\boldsymbol{v}^T A\boldsymbol{v}$, where $\boldsymbol{v}$ is a vector with the n variables as entries and A is a symmetric matrix.

The two variable quadratic form above can be expressed as

$$Q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x\ y)\begin{pmatrix} ax + \frac{c}{2}y \\ \frac{c}{2}x + by \end{pmatrix}$$

$$= ax^2 + \frac{c}{2}xy + \frac{c}{2}xy + by^2$$

The three variable quadratic form given above can be written as

$$Q(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a & \underline{d/2} & \underline{e/2} \\ \underline{d/2} & b & \underline{f/2} \\ \underline{e/2} & \underline{f/2} & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

— xy
— xz
— yz

As an exercise, trying verifying this by expanding out both expressions. Observe that in both cases the diagonal entries of the matrix are the coefficients of the square terms and the off-diagonal entries in the matrix are the coefficients of the cross-terms.

19.1.1 Give the matrix representation of the quadratic form $2x^2 + 6xy - 7y^2$.

$$2x^2 + 6xy - 7y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Could write $\begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 2 & 2 \\ 4 & -7 \end{pmatrix}}_{\text{diagonalisable?}} \begin{pmatrix} x \\ y \end{pmatrix}$

* with a symmetric matrix, we can <u>always</u> diagonalise.

19.2 Diagonalising quadratic forms

Since we know we can always orthogonally diagonalise a symmetric matrix, if we do this to the symmetric matrix in the matrix representation of the quadratic form, we can reduce the quadratic form to a sum of square terms.

We shall demonstrate this by example:

19.2.1 Express $-3x^2 - 2y^2 - 3z^2 + 2xy + 2yz$ exclusively as a sum of square terms.

$$Q(x,y,z) = (x\ y\ z)\underbrace{\begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}}_{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

From previous chapters, $A = PDP^T$ $\quad (PP^T = I)$

with $P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$, $D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}$

$$\Rightarrow Q(x,y,z) = \underline{x}^T A \underline{x} = (\underline{x}^T P) D (P^T \underline{x}) = \underline{u}^T D \underline{u}$$

where $\underline{u} = P^T x$

$$\Rightarrow \underline{u} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{x-z}{\sqrt{2}} \\ \frac{x+2y+z}{\sqrt{6}} \\ \frac{x-y+z}{\sqrt{3}} \end{pmatrix}$$

$$\Rightarrow Q = (u\ v\ w)\begin{pmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}\begin{pmatrix} u \\ v \\ w \end{pmatrix} = -3u^2 - v^2 - 4w^2$$

$$\Rightarrow Q(x,y,z) = -\frac{3}{2}(x-z)^2 - \frac{1}{6}(x+2y+z)^2 - \frac{4}{3}(x-y+z)^2$$

e.g. Show that $Q(x,y,z)$ is never positive."

19.3 Quadratic equations and conic sections

We now restrict our attention to two dimensions, by investigating quadratic equations, which are equations of the form

$$\underbrace{ax^2 + by^2 + cxy}_{\text{Quadratic form}} + \overbrace{dx + ey}^{\text{linear form}} + \underbrace{f}_{\text{const.}} = 0,$$

where $a, b, c, d, e, f \in \mathbb{R}$.

Graphs of quadratic equations are known as *conic sections*, because they can be realised as the intersection of a plane and a double cone in three dimensions. The most interesting of these are the so-called *non-degenerate* conic sections[1]. A non-degenerate conic section is in standard position relative to the coordinate axes if its equation can be expressed in one of the following forms: "Standard form"

- $\dfrac{x^2}{k^2} + \dfrac{y^2}{l^2} = 1; k, l > 0,$ ellipse/circle
- $\dfrac{x^2}{k^2} - \dfrac{y^2}{l^2} = 1$ or $\dfrac{y^2}{l^2} - \dfrac{x^2}{k^2} = 1; k, l > 0,$ hyperbola
- $x^2 = ky$ or $y^2 = kx; k \neq 0.$ parabola

1

The key observation here is that conic sections in standard form have no cross-terms. Given a quadratic equation with cross-terms in the associated quadratic form, we can *change variables* to remove the cross-terms by orthogonal diagonalisation. Due to the defining property of rotation matrices, an orthogonal matrix P always corresponds to a rotation, provided $\det(P) = 1$ (not -1). Hence, we have the following.

> Changing variables by orthogonal diagonalisation corresponds to a rotation of the coordinate axes. If P is the orthogonal (rotation) matrix, then the new coordinates (u, v) can be expressed in terms of the old coordinates (x, y) as
>
> $$\begin{pmatrix} u \\ v \end{pmatrix} = P^T \begin{pmatrix} x \\ y \end{pmatrix}.$$

2

Another important observation is that there is never an occurance of x^2 and x in the standard form (or y^2 and y). As a general rule, given a quadratic equation (even after changing variables from orthogonal diagonalisation), if we have terms such as x^2 and x (or similar terms involving new variables) we can *complete the square* to be left with only a square term. We have the following.

> Completing the square in a quadratic equation corresponds to translating (or shifting) the coordinate axes.

[1]There are also *degenerate* (points, lines) and *imaginary* (without real graphs) conic sections.

In summary, to identify a quadratic equation as a conic section, we follow these steps:

1. Write the quadratic equation (A is a symmetric matrix)

$$ax^2 + by^2 + cxy + dx + ey + f = 0$$

in the matrix form $\boldsymbol{x}^T A\boldsymbol{x} + K\boldsymbol{x} + f = 0$, where $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $K = \begin{pmatrix} d & e \end{pmatrix}$.

2. Find a matrix P that orthogonally diagonalises A, so $A = PDP^T$. You may need to swap columns of P to ensure that $\det(P) = 1$ (and hence corresponds to a rotation).

3. Define new variables u, v such that $\boldsymbol{v} = \begin{pmatrix} u \\ v \end{pmatrix} = P^T\boldsymbol{x} \Rightarrow \boldsymbol{x} = P\boldsymbol{v}$.

4. Substitute $\boldsymbol{v}$ into the matrix form of the equation, giving

$$\boldsymbol{v}^T D\boldsymbol{v} + KP\boldsymbol{v} + f = 0.$$

5. Complete the square if required. This is necessary if u^2 and u are both present (or v^2 and v). This defines a new set of variables s, t by translating u, v. The translations will be of the form $s = \alpha u + \beta$, $t = \gamma v + \delta$.

6. If it is a non-degenerate conic, the final equation in s and t should be a conic section in standard form.

19.3.1 Describe the conic whose equation is $x^2 + y^2 + 2xy - 3x - 5y + 4 = 0$.

$$(x \; y)\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}}_{A}\begin{pmatrix} x \\ y \end{pmatrix} + (-3 \; -5)\begin{pmatrix} x \\ y \end{pmatrix} + 4$$

A is symmetric $\Rightarrow$ orthogonally diagonalizable

Eigenvalues/vectors of A: $\lambda = 0 \longleftrightarrow \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$

$\lambda = 2 \longleftrightarrow \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ (these eigenvectors form an orthonormal set)

$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, $\det P = 1 \Rightarrow P$ is a rotation

$A = PDP^T$, $D = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$

We have $\underline{x}^T A \underline{x} + k\underline{x} + 4 = 0$, $k = (-3 \;\; -5)$

$\underline{x} = \begin{pmatrix} x \\ y \end{pmatrix}$

$\Rightarrow (\underline{x}^T P) D (P^T \underline{x}) + k\underline{x} + 4 = 0$

set $\underline{u} = P^T \underline{x} \Rightarrow u = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x-y}{\sqrt{2}} \\ \frac{x+y}{\sqrt{2}} \end{pmatrix}$

Also $P\underline{u} = \underbrace{PP^T}_{=I} \underline{x} = \underline{x}$

$kP = (-3 \;\; -5) \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{2}} & -\frac{8}{\sqrt{2}} \end{pmatrix}$

$\Rightarrow \underline{u}^T D \underline{u} + kP\underline{u} + 4 = 0$

$\Rightarrow (u \;\; v) \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \sqrt{2} & -4\sqrt{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + 4 = 0$

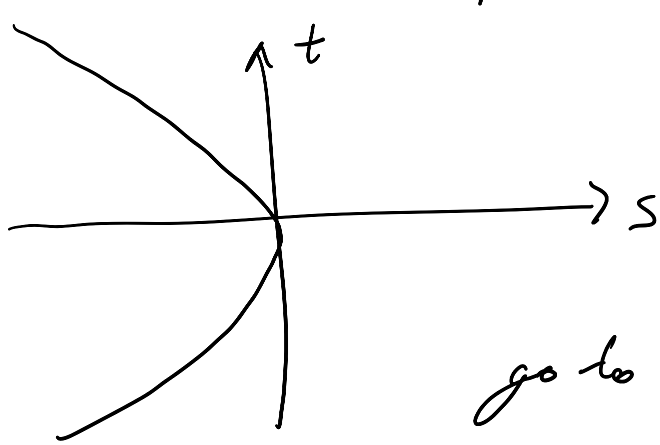
$\Rightarrow 2v^2 + \sqrt{2}u - 4\sqrt{2}v + 4 = 0$ (no cross-terms)

(need to complete the square in v)

$\Rightarrow 2(v^2 - 2\sqrt{2}v + 2) + \sqrt{2}u = 0$

$\Rightarrow 2(v - \sqrt{2})^2 + \sqrt{2}u = 0$

set $\left.\begin{aligned} s &= u \\ t &= v - \sqrt{2} \end{aligned}\right\}$ $\underline{s = -\sqrt{2}t^2}$
parabola



go to p143

19.4 Further reading: quadric surfaces

It turns out we can similarly use orthogonal diagonalisation of 3×3 matrices to simplify and ultimately identify surfaces whose general equation is of the form

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + j = 0,$$

where a, b, c, d, e, f are never all zero. Note that we can rewrite the equation in matrix form as

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + j = 0.$$

If we then orthogonally diagonalise the 3×3 matrix, then complete any squares that are left over, we end up being able to identify the surface as one of the following forms:

- $\dfrac{x^2}{l^2} + \dfrac{y^2}{m^2} + \dfrac{z^2}{n^2} = 1,$
- $z^2 = \dfrac{x^2}{l^2} + \dfrac{y^2}{m^2},$
- $\dfrac{x^2}{l^2} + \dfrac{y^2}{m^2} - \dfrac{z^2}{n^2} = 1,$
- $z = \dfrac{x^2}{l^2} + \dfrac{y^2}{m^2},$
- $\dfrac{z^2}{l^2} - \dfrac{x^2}{m^2} - \dfrac{y^2}{n^2} = 1,$
- $z = \dfrac{y^2}{m^2} - \dfrac{x^2}{l^2}.$

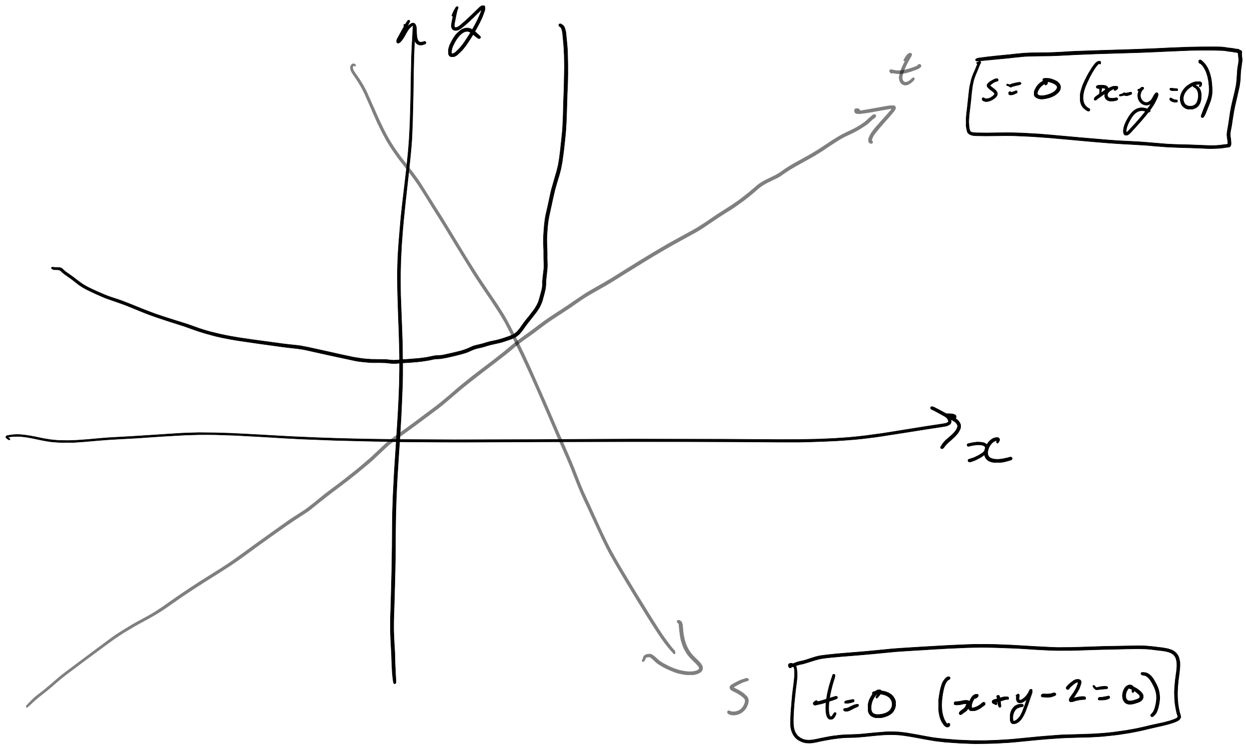
As in the two dimensional case, the orthogonal diagonalisation has the effect of rotating the axes, provided the orthogonal matrix P has $\det P = 1$, which we can choose by carefully ordering the columns. Completing the square has the effect of shifting the axes.

You should be aware that these techniques are available in order to simplify and identify algebraic expressions representing surfaces. The three dimensional case can often become quite complicated. You will not be expected to identify quadric surfaces in an exam.

Notes.

(from p141)

$$s = \frac{x-y}{\sqrt{2}}, \quad t = \frac{x+y-2}{\sqrt{2}}$$



20 Complex matrices

By the end of this section, you should be able to answer the following questions:

- What are unitary, Hermitian and normal matrices?
- Given a complex matrix, determine if it can be unitarily diagonalised, and if so, diagonalise it.

Unitary and Hermitian matrices are complex analogues of orthogonal ($A^{-1} = A^T$) and symmetric ($A = A^T$) real matrices respectively.

In order to define these matrices, we need the following.

20.1 Definition (conjugate transpose)

or $A^\dagger$

Let A be a complex matrix. The conjugate transpose of A, denoted A^*, is given by $(\overline{A})^T$, where $\overline{A}$ is the matrix whose entries are complex conjugates of the corresponding entries of A.

↳ $z = a + ib$, $a, b \in \mathbb{R}$, $i^2 = -1$
complex conjugate of z, denoted $\bar{z}$ is given by $\bar{z} = a - ib$
Note: $z\bar{z} = a^2 + b^2$

Note that if A is real, $A^* = A^T$.

20.1.1 Example

Let $A = \begin{pmatrix} 3+7i & 0 \\ 2i & 4-i \end{pmatrix}$. Write down the conjugate transpose of A.

$$\bar{A} = \begin{pmatrix} 3-7i & 0 \\ -2i & 4+i \end{pmatrix}$$

$$A^* = (\bar{A})^T = \begin{pmatrix} 3-7i & -2i \\ 0 & 4+i \end{pmatrix}$$

20.2 Unitary matrices

A complex matrix A is said to be *unitary* if $A^{-1} = A^*$. Compare this definition with that of real orthogonal matrices. $\rightarrow A^{-1} = A^T$

Recall that a real matrix is orthogonal if and only if its columns form an orthonormal set of vectors. For complex matrices, this property characterises unitary matrices. In this case however, we must use the complex inner product.

20.3 Complex inner product

Recall that in $\mathbb{R}^n$ the inner (or dot) product of two vectors

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

is given by

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \quad = \underline{u}^T \underline{v} \text{ or } \underline{v}^T \underline{u}$$

and the length (a real number!) of $\boldsymbol{u}$ by

$$||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

These definitions are unsuitable for vectors in $\mathbb{C}^n$.

To demonstrate, consider the vector $\boldsymbol{u} = (i, 1)$ in $\mathbb{C}^2$. Using the above expression for length, we would obtain $||\boldsymbol{u}|| = \sqrt{i^2 + 1} = 0$, so $\boldsymbol{u}$ would be a non-zero vector with length 0.

Instead, we introduce the complex inner product

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 \overline{v}_1 + u_2 \overline{v}_2 + \cdots + u_n \overline{v}_n, \quad \in \mathbb{C}$$

where as usual $\overline{v}_i$ denotes the complex conjugate of v_i. In matrix notation, we can write this as $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v}^* \boldsymbol{u}$. Note the length of a complex vector is always a real number.

So now we understand what is meant by the following statement: Columns of a unitary matrix form an orthonormal set with respect to the complex inner product.

- Norm $||\underline{u}|| = \sqrt{\underline{u} \cdot \underline{u}} \in \mathbb{R}$ using complex dot product
- In the case $\underline{v}$ has only real entries, $\underline{v}^* = \underline{v}^T$

More generally,

- Axiom (I1): $\overline{\langle \underline{u}, \underline{v} \rangle} = \langle \underline{v}, \underline{u} \rangle$
  (inner product, CH8), $\langle \underline{u}, \underline{v} \rangle \in \mathbb{C}$

20.4 Hermitian (self-adjoint) matrices

A complex matrix A is called *Hermitian* (or *self-adjoint*) if $A = A^*$.

As with symmetric matrices, we can recognise a Hermitian matrix by inspection. See if you can see the pattern in the following 2×2, 3×3 and 4×4 Hermitian matrices.

$$\begin{pmatrix} a_{11} & a_{12}+ib_{12} \\ a_{12}-ib_{12} & a_{22} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12}+ib_{12} & a_{13}+ib_{13} \\ a_{12}-ib_{12} & a_{22} & a_{23}+ib_{23} \\ a_{13}-ib_{13} & a_{23}-ib_{23} & a_{33} \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12}+ib_{12} & a_{13}+ib_{13} & a_{14}+ib_{14} \\ a_{12}-ib_{12} & a_{22} & a_{23}+ib_{23} & a_{24}+ib_{24} \\ a_{13}-ib_{13} & a_{23}-ib_{23} & a_{33} & a_{34}+ib_{34} \\ a_{14}-ib_{14} & a_{24}-ib_{24} & a_{34}-ib_{34} & a_{44} \end{pmatrix},$$

where $a_{ij}, b_{ij} \in \mathbb{R}$. Note in particular that the diagonal entries are real numbers.

One of the most significant results on Hermitian matrices is that their eigenvalues are real. Note: All real symmetric matrices are Hermitian

20.4.1 Proof that Hermitian matrices have real eigenvalues

Let $\boldsymbol{v} \in \mathbb{C}^n$ be an eigenvector of the Hermitian matrix A, with corresponding eigenvalue λ. In other words,

$$A\boldsymbol{v} = \lambda\boldsymbol{v}. \tag{12}$$

In what follows, we use the fact that $(AB)^* = B^*A^*$ which holds since the same is true for matrix transposition.

We multiply (12) from the left by $\boldsymbol{v}^*$ (treat $\boldsymbol{v}$ as an $n\times 1$ complex matrix) to obtain

$$\boldsymbol{v}^*A\boldsymbol{v} = \boldsymbol{v}^*(\lambda\boldsymbol{v}) = \lambda(\boldsymbol{v}^*\boldsymbol{v}). \tag{13}$$

$(1\times n)(n\times n)(n\times 1)$

Also note that

$$(\boldsymbol{v}^*A\boldsymbol{v})^* = \boldsymbol{v}^*A^*(\boldsymbol{v}^*)^* = \boldsymbol{v}^*A\boldsymbol{v}.$$

In other words, $\boldsymbol{v}^*A\boldsymbol{v}$ is also Hermitian. Since it evaluates to be a 1×1 matrix, and all Hermitian matrices have real numbers on their diagonal, this means that $\boldsymbol{v}^*A\boldsymbol{v}$ is a real number.

The quantity $\boldsymbol{v}^*\boldsymbol{v}$ is precisely the complex inner product of $\boldsymbol{v}$ with itself as we have already seen, which is also a real number.

Therefore equation (13) is of the form

$$x = \lambda y, \quad x, y \in \mathbb{R},$$

from which we must conclude that λ is real.

One consequence of this result is that a real symmetric matrix has real eigenvalues, since every real symetric matrix is Hermitian. This result was stated on page 133 but not proved.

20.5 Unitary diagonalisation

We have seen that real symmetric matrices are orthogonally diagonalisable. There is an analagous concept for complex matrices.

A square matrix A with complex entries is said to be unitarily diagonalisable if there is a unitary matrix P such that P^*AP is diagonal. $P^{-1} = P^*$

It is natural to consider which matrices are unitarily diagonalisable. The answer lies in a more general class of matrix.

20.6 Normal matrices

A square complex matrix is called normal if it commutes with its own conjugate transpose, ie, if $AA^* = A^*A$.

Normal matrices are generally more difficult to identify by inspection. However, we have some classes of matrices which are normal:

- unitary, $AA^* = I = A^*A$
- Hermitian, $A = A^* \Rightarrow AA^* = A^2 = A^*A$
- real skew-symmetric (satisfying $A^T = -A$),
- any diagonal matrix,
- others?

Real orthogonal $\subset$ unitary

real orthogonal $\subset$ Hermitian

We make a note that real normal 2×2 matrices are either symmetric or of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ (which include the skew-symmetric examples).

A class of matrix which is not generally normal is the class of complex symmetric matrices.

20.6.1 Example

Classify the matrix $A = \begin{pmatrix} 1 & 1+i \\ 1+i & -i \end{pmatrix}$.

$A^* = \begin{pmatrix} 1 & 1-i \\ 1-i & i \end{pmatrix} \neq A$ ($\Rightarrow$ not Hermitian)

$A^*A = \begin{pmatrix} 1 & 1-i \\ 1-i & i \end{pmatrix}\begin{pmatrix} 1 & 1+i \\ 1+i & -i \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \neq I$

($\Rightarrow$ NOT unitary)

but we have $A^{-1} = \frac{1}{3}A^*$

$\Rightarrow AA^* = 3I = A^*A$

$\Rightarrow A$ is normal

(c.f. real symmetric = orthogonally diagonalisable)

20.7 Normal = unitarily diagonalisable

The main result we have is completely analagous to the real case of orthogonal diagonalisation and symmetric matrices on page 133. We will not prove this result.

An $n \times n$ complex matrix is unitarily diagonalisable if and only if it normal.

20.7.1 Example

If possible, diagonalise the matrix $\begin{pmatrix} 6 & 2+2i \\ 2-2i & 4 \end{pmatrix} = A$

$A^* = A \Rightarrow A$ is Hermitian $\Rightarrow A$ is normal

$\Rightarrow A$ is unitarily diagonalisable

$$\det\begin{pmatrix} 6-\lambda & 2+2i \\ 2-2i & 4-\lambda \end{pmatrix} = 24 - 10\lambda + \lambda^2 - 8$$

$$= (\lambda-8)(\lambda-2), \quad \lambda_1 = 8, \ \lambda_2 = 2$$

$\underline{\lambda = 8:}$ $\begin{pmatrix} -2 & 2+2i \\ 2-2i & -4 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$R1 \Rightarrow \underline{a = (1+i)b} \Rightarrow \underline{v}_1 = \begin{pmatrix} (1+i)b \\ b \end{pmatrix} = b\begin{pmatrix} 1+i \\ 1 \end{pmatrix}$

$\left(\Rightarrow \frac{a}{1+i} = b \Rightarrow \frac{a}{1+i} \cdot \frac{1-i}{1-i} = b \Rightarrow \frac{a(1-i)}{2} = b \Leftarrow R2 \right)$

Similarly, $\underline{\lambda = 2:}$ $\underline{v}_2 = a\begin{pmatrix} 1 \\ -1+i \end{pmatrix}$

$$\underline{v}_1^* \underline{v}_1 = (1-i)(1+i)+1 = 3$$

$$\Rightarrow \|\underline{v}_1\| = \sqrt{3}$$

$$\& \|\underline{v}_2\| = \sqrt{3}$$

Form the unitary matrix

$$P = \begin{pmatrix} \frac{1+i}{\sqrt{3}} & 1/\sqrt{3} \\ \frac{1}{\sqrt{3}} & \frac{-1+i}{\sqrt{3}} \end{pmatrix}$$

Check $PP^* = I = P^*P$

Then, $P^*AP = D = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$

Notes.

21 Multivariable Taylor series

By the end of this section you should be able to write down Taylor series of smooth, scalar-valued functions of n variables.

Let the function $f : \mathbb{R} \to \mathbb{R}$ be infinitely differentiable at $a \in \mathbb{R}$. Then, the Taylor series of $f(x)$ is the power series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots.$$

A common reason for calculating the Taylor series of a function about a point $x = a$ is that its partial sums provide (polynomial) approximations to the function near the point a. Here, we want to understand how this extends to functions of more than one variable.

Before we continue, it is recalled that Taylor series need not be convergent, and even if the Taylor series of $f(x)$ does converge, its limit need not equal $f(x)$ (in which case f is known as *non-analytic*). To accommodate these situations, we may write

$$f(x) \approx f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots.$$

21.1 Hessian matrix

For $n \in \mathbb{N}$, let $D \subseteq \mathbb{R}^n$ and $f : D \to \mathbb{R}$ such that all second-order partial derivatives of f,

$$f_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \qquad i, j = 1, \dots, n,$$

are continuous on D. The corresponding **Hessian** is then defined as the $n \times n$ matrix

$$H = \begin{pmatrix} f_{x_1x_1} & \cdots & f_{x_1x_n} \\ \vdots & & \vdots \\ f_{x_nx_1} & \cdots & f_{x_nx_n} \end{pmatrix}.$$

According to Clairaut's theorem, $f_{x_ix_j} = f_{x_jx_i}$ for all $i, j = 1, \dots, n$, so $H^T = H$, meaning that H is a *symmetric matrix*. To specify the dependence on f and $\mathbf{x} = (x_1, \dots, x_n)$, one occasionally writes $H(f)(\mathbf{x})$ or $H_f(\mathbf{x})$.

21.2 Example: H of $f(x, y) = x^3y + 2y$ at $(1, 2)$

$f_x = 3x^2y \,, \; f_y = x^3+2$

$$\Rightarrow H = \begin{pmatrix} 6xy & 3x^2 \\ 3x^2 & 0 \end{pmatrix} \Rightarrow H(1,2) = \begin{pmatrix} 12 & 3 \\ 3 & 0 \end{pmatrix}$$

21.3 Taylor series of $f(x, y)$

Let $D \subseteq \mathbb{R}^2$ and $f : D \to \mathbb{R}$ such that f and all its (higher-order) partial derivatives are continuous on D. Such a function is often called **smooth**. Here, we will examine the behaviour of f near $(x, y) = (a, b)$ by reducing the problem to a related one-variable problem we know how to handle.

Let $\mathbf{v} = (h, k) \neq (0, 0)$ be small enough such that $(a+ht, b+kt) \in D$ for $t \in [-1, 1]$. The function

$$F(t) = f(a + ht, b + kt), \qquad t \in [-1, 1],$$

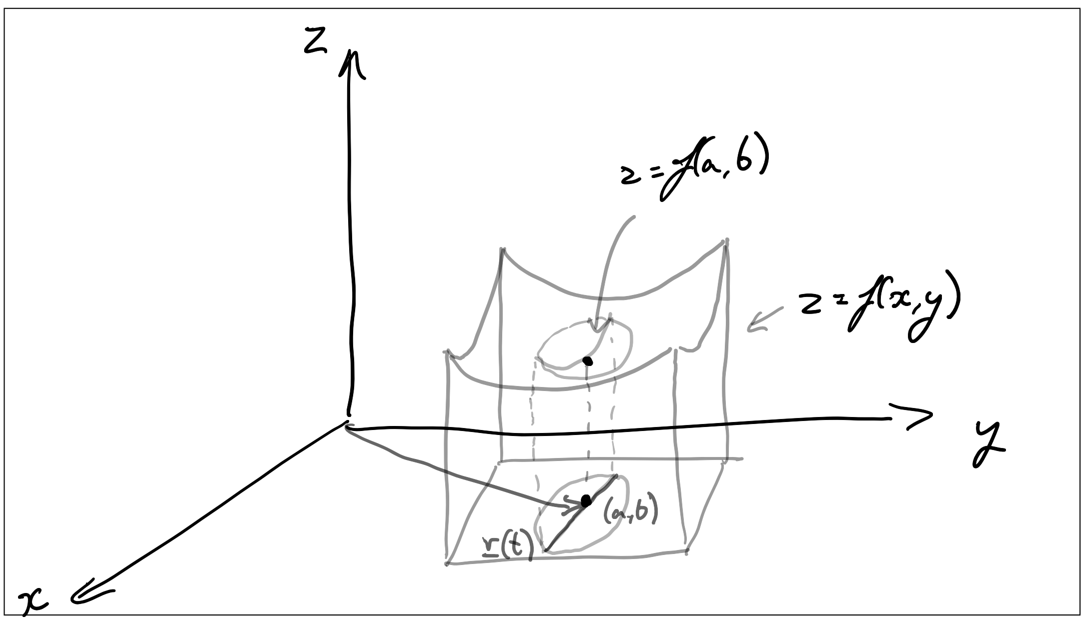
then describes how $f(x, y)$ behaves along the line parameterised as

$$x(t) = a + ht, \qquad y(t) = b + kt. \quad \longrightarrow \quad \underline{r}(t) = x(t)\underline{i} + y(t)\underline{j}$$

$\underline{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\underline{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Understanding this behaviour near $t = 0$ will give us insight into how f behaves near $(x, y) = (a, b)$. Geometrically, we can think of it as follows:



Algebraically, we have the following:

$$F'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = f_x h + f_y k$$

$$F''(t) = \frac{d}{dt}\left(f_x h + f_y h\right)$$

$$= (f_{xx} h + f_{xy} h)h + (f_{yx} h + f_{yy} k)k$$

$$= f_{xx} h^2 + 2 f_{xy} hk + f_{yy} h^2$$

$F(0) = f(a,b)$

Up to higher-order terms in t, we thus have

$$f(a+ht, b+kt) \approx f(a,b) + \big[f_x(a,b)h + f_y(a,b)k\big]t$$
$$+ \big[f_{xx}(a,b)h^2 + 2f_{xy}(a,b)hk + f_{yy}(a,b)k^2\big]\frac{t^2}{2} + \dots.$$

Since $(x,y) = (a+ht, b+kt)$ implies that $ht = x-a$ and $kt = y-b$, we have $h^2t^2 = (x-a)^2$, $hkt^2 = (x-a)(y-b)$ and $k^2t^2 = (y-b)^2$, so

$$f(x,y) \approx f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$
$$+ \frac{1}{2}\underbrace{\big[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\big]} + \dots.$$

Terms corresponding to higher powers of t involve factors of the form $(x-a)^p(y-b)^q$, where $p+q>2$, $p,q \in \mathbb{N}$.

Recalling the notion of *quadratic forms* from ~~the lecture workbook~~ CH19, we can rewrite the first few terms of the Taylor series as follows:

$$f(x,y) \simeq f(a,b) + \begin{pmatrix} f_x(a,b) & f_y(a,b) \end{pmatrix}\begin{pmatrix} x-a \\ y-b \end{pmatrix}$$
$$+ \frac{1}{2}\begin{pmatrix} x-a & y-b \end{pmatrix}\begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{pmatrix}\begin{pmatrix} x-a \\ y-b \end{pmatrix}$$
$$+ \dots$$

Setting

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \mathbf{x}_0 = \begin{pmatrix} a \\ b \end{pmatrix},$$

$$\underset{\sim}{\nabla} f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$$

we have

$$\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \end{pmatrix} = \nabla f(\mathbf{x}_0), \qquad \begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{pmatrix} = H_f(\mathbf{x}_0).$$

This allows us to write the first few terms of the Taylor series in the compact form

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + (\nabla f(\mathbf{x}_0))^T(\mathbf{x}-\mathbf{x}_0) + \tfrac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T H_f(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + \dots.$$

Its vectorial expression indicates quite clearly how it extends to higher dimensions.

$$f_x = \frac{\cos(x-y^2)}{\sqrt{y+1}}, \quad f_y = \frac{-2y\cos(x-y^2) - \frac{1}{2\sqrt{y+1}}\sin(x-y^2)}{y+1}$$

etc

21.4 Example: $f(x, y) = \frac{\sin(x-y^2)}{\sqrt{y+1}}$ about $(0, 0)$

$f(0,0) = 0, \quad f_x(0,0) = 1, \quad f_y(0,0) = 0$

check $f_{xx}(0,0) = 0, \quad f_{xy}(0,0) = -\frac{1}{2}, \quad f_{yy}(0,0) = -2$

$$\Rightarrow f(x,y) \approx 0 + 1(x-0) + 0(y-0)$$
$$+ \frac{1}{2}\left[0(x-0)^2 + 2\left(-\frac{1}{2}\right)(x-0)(y-0) - 2(y-0)^2\right] + \ldots$$
$$= x - \frac{1}{2}xy - y^2$$

21.5 Taylor series in any dimension

$\underline{h} = (h_1, h_2, \ldots, h_n)$

Let f be a smooth scalar-valued function on $D \subseteq \mathbb{R}^n$, and let $\mathbf{h} \in \mathbb{R}^n$ such that $\mathbf{x} + \mathbf{h}t \in D$ for $t \in [-1, 1]$. Then, the Taylor series of $f(\mathbf{x} + \mathbf{h})$ about $\mathbf{x}$ is

$$f(\mathbf{x} + \mathbf{h}) \approx \sum_{\ell=0}^{\infty} \frac{1}{\ell!}(\mathbf{h} \cdot \nabla)^{\ell} f(\mathbf{x}) = f(\mathbf{x}) + \mathbf{h} \cdot \nabla f(\mathbf{x}) + \tfrac{1}{2}(\mathbf{h} \cdot \nabla)^2 f(\mathbf{x}) + \ldots.$$

$\underline{\nabla} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$

$\ell = 0: \quad \frac{1}{0!}(\underline{h} \cdot \underline{\nabla})^0 f(\underline{x}) = f(\underline{x})$

$\ell = 1: \quad \frac{1}{1!}(\underline{h} \cdot \underline{\nabla})^1 f(\underline{x}) = \sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} f(\underline{x})$

$$\left(= \sum_{i=1}^{n} (x_i + h_i - x_i) f_{x_i}(\underline{x})\right)$$

$\ell = 2: \quad \frac{1}{2!}(\underline{h} \cdot \underline{\nabla})^2 f(\underline{x}) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j} f(\underline{x})$

$$\left(= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i + h_i - x_i)(x_j + h_j - x_j) f_{x_j x_i}(\underline{x})\right)$$

22 Critical points in n-dimensions

This chapter brings together a great deal of what we have studied so far in this course. The goal is to be able to classify critical points of functions of any number of variables.

Recall Taylor series of a smooth function f in n variables about a point $\mathbf{x} = \mathbf{x}_0$ is given by;

$$\begin{aligned} f(\mathbf{x}) =& f(\mathbf{x}_0) + (\nabla f(\mathbf{x}_0))^T(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T H(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \\ &+ \langle \text{ higher order terms} \rangle, \end{aligned}$$

$$\text{where } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad H(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}_0) & \dots \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}_0) & \dots \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}_0) & \dots \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{pmatrix} = H(\mathbf{x}_0)^T$$

i.e. $H(\mathbf{x}_0)$ is a real symmetric matrix.

22.1 Classification of critical points in n dimensions

In the following, let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.

Definition 1. A point $\mathbf{x}_0$ is said to be a *critical point* if $\nabla f(\mathbf{x}_0) = \mathbf{0}$ or $\nabla f(\mathbf{x}_0)$ is undefined.

Definition 2. A critical point $\mathbf{x}_0$ satisfying $\nabla f(\mathbf{x}_0) = \mathbf{0}$ is a *local maximum (local minimum)* if there exists some $\epsilon > 0$ such that $f(\mathbf{x}_0) \geq f(\mathbf{x})$ $(f(\mathbf{x}_0) \leq f(\mathbf{x}))$ for all $\mathbf{x}$ such that $||\mathbf{x} - \mathbf{x}_0|| < \epsilon$.

Definition 3. A critical point $\mathbf{x}_0$ satisfying $\nabla f(\mathbf{x}_0) = \mathbf{0}$ is a *saddle point* if it is neither a local maximum nor a local minimum.

$$\underline{\nabla} f(x_1, x_2, \dots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

22.2 Example: In 2d.

① $f(x_1, x_2) = x_1^2 + x_2^2 \Rightarrow \nabla f = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\Rightarrow$ critical point at $(0,0)$ is a minimum

② $f(x_1, x_2) = -x_1^2 - x_2^2 \Rightarrow (0,0)$ is a local max

③ $f(x_1, x_2) = -x_1^2 + x_2^2 \Rightarrow (0,0)$ is a saddle

22.3 Critical points by Taylor series

In MATH1052: We used the"Second derivative test" for functions of two variables.

In MATH2001/7000: We consider a variant of this test that generalises easily to higher dimensions.

Let $\mathbf{x_0}$ be a critical point satisfying $\nabla f(\mathbf{x}_0) = \mathbf{0}$ (in $\mathbb{R}^n$)
$\Longrightarrow$ Taylor series about $\mathbf{x}_0$ is

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T H(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \langle \text{ higher order terms } \rangle.$$

Without loss of generality, we take $\mathbf{x}_0 = \mathbf{0}$ (i.e. by shifting/translating variables if necessary).

We have,

$$f(\mathbf{x}) = f(\mathbf{0}) + \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \langle \text{ higher order terms } \rangle.$$

Here $H = H(\mathbf{0})$. Thus, the behaviour about $\mathbf{0}$ (i.e the critical point) depends on this second order term.

(could also take $f(\underline{0}) = 0$).

Observe that H is real symmetric $\implies$ H is orthogonally diagonalisable, i.e. there exists an orthogonal matrix P such that $P^T H P = D$ with some diagonal matrix D.

Orthogonal matrix P satisfies $PP^T = I = P^T P$

$HP = PD \Rightarrow H = PDP^T$. In n variables, columns of the $n \times n$ matrix P form an orthonormal set of eigenvectors (w.r.t. Euclidean inner product) $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_n\}$

$$P = \left(\underset{\sim}{e}_1 \middle| \underset{\sim}{e}_2 \middle| \dots \middle| \underset{\sim}{e}_n\right), \quad D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & \vdots \\ \vdots & & \ddots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

$$\lambda_i \longleftrightarrow \underset{\sim}{e}_i$$

Also H is real symmetric $\Rightarrow$ all $\lambda_i \in \mathbb{R}$

It follows that

$$\mathbf{x}^T H \mathbf{x} = (\mathbf{x}^T P) D (P^T \mathbf{x}) = \mathbf{y}^T D \mathbf{y}.$$

(i.e. diagonalisation suggests set $P^T \mathbf{x} = \mathbf{y}$).

The critical point is still at $\mathbf{y} = \mathbf{0}$, because $P^T \mathbf{0} = \mathbf{0}$.

Let F denote the function f expressed in this new coordinate system i.e. $F(\mathbf{y}) = f(\mathbf{x}(\mathbf{y}))$.

$$\begin{aligned} \implies \quad F(\mathbf{y}) =& f(\mathbf{0}) + \frac{1}{2}\mathbf{y}^T D \mathbf{y} + \langle \text{ higher order terms } \rangle \\ =& f(\mathbf{0}) + \frac{1}{2}\left(\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2\right) \\ &+ \langle \text{ higher order terms } \rangle, \end{aligned}$$

where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

There are four cases to consider:

① if $\lambda_i > 0 \ \forall i = 1, 2, \dots, n$ then the quadratic form is strictly positive in every direction
i.e. local minimum

② If $\lambda_i < 0 \ \forall i = 1, 2, \dots, n$ then local minimum
(similar to ①

③ If $\exists \, i, j \ (i \neq j)$ such that λ_i and λ_j have opposite sign, then saddle

④ If all non-zero λ_j have the same sign, but there are some $\lambda_j = 0$, then we cannot identify the type of critical point and need to examine those <higher order terms> in the Taylor series

(go to p161)

22.4 Example: $Q = ax^2 + bxy + cy^2$

$a = \frac{1}{2} f_{xx}$, $b = f_{xy}$, $c = \frac{1}{2} f_{yy}$

MATH1052 $\to$ $Q = a\left(\left(x + \frac{by}{2a}\right)^2 + \frac{D}{4a^2} y^2\right)$

with $D = 4ac - b^2 = f_{xx} f_{yy} - f_{xy}{}^2 = \det(H)$

$f_{xx} > 0$, $D > 0 \to$ local min
$f_{xx} < 0$, $D > 0 \to$ local max
$D < 0 \to$ saddle
$D = 0 \to$ inconclusive

$$\text{MATH2001} \rightarrow Q = (x \;\; y) \underbrace{\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}}_{\frac{1}{2}H} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\det \begin{pmatrix} a-\lambda & \frac{b}{2} \\ \frac{b}{2} & c-\lambda \end{pmatrix} = 0 = (a-\lambda)(c-\lambda) - \frac{b^2}{4}$$
$$= \lambda^2 - (a+c)\lambda + ac - \frac{b^2}{4}$$

$$\Rightarrow \lambda_\pm = \frac{a+c \pm \sqrt{(a+c)^2 - D}}{2} \qquad D = 4ac - b^2$$

Look at $(a+c)^2 - D = a^2 + 2ac + c^2 - (4ac - b^2)$
$$= (a-c)^2 + b^2 \geq 0$$

$$\rightarrow Q = \lambda_+ z_+^2 + \lambda_- z_-^2$$

Analyse $\lambda_\pm$: $D = 0 \rightarrow$ one of $\lambda_+, \lambda_- = 0$
$\rightarrow$ inconclusive

$D < 0 \Rightarrow$ saddle (λ_+, λ_- have opposite sign)

$D > 0 \Rightarrow 4ac - b^2 > 0 \Rightarrow 4ac > b^2$ i.e. $ac > 0$

if $a > 0 \Rightarrow c > 0 \Rightarrow \lambda_\pm > 0 \rightarrow$ local min

& if $a < 0 \Rightarrow c < 0 \Rightarrow \lambda_\pm < 0 \rightarrow$ local max

---

in general, $\det(H) = \lambda_1 \lambda_2 \ldots \lambda_n$

in $\mathbb{R}^2$, $\det(H) = \lambda_1 \lambda_2$

Notes. (From p159)

$$Q(x,y,z) = -5x^2 - 2y^2 - 5z^2 - 2xy + 8xz - 2yz$$

$$= \begin{pmatrix} x & y & z \end{pmatrix} \underbrace{\begin{pmatrix} -5 & -1 & 4 \\ -1 & -2 & -1 \\ 4 & -1 & -5 \end{pmatrix}}_{\text{eigenvalues } -9,\,-3,\,0} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$Q(x,y,z) = -\frac{9}{2}(x-z)^2 - \frac{3}{6}(x+2y+z)^2 + \frac{0}{3}(x-y+z)^2$$

Along $\underline{r}(t) = t\underline{i} - t\underline{j} + t\underline{k}$, $\quad t \geq 0$

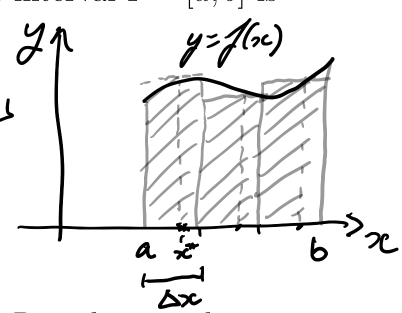
23 Introduction to double integrals, volume below a surface

By the end of this section, you should be able to answer the following questions:

- What is the definition of volume below a surface? $z = f(x,y) > 0$
- What is the definition of a double integral?
- How are the two related?
- What is an iterated integral?

Recall that if $y = f(x)$, the area under the curve over the interval $I = [a, b]$ is

$$\int_I f(x)dx = \underbrace{\lim}_{\substack{n\to\infty \\ \text{or} \\ \Delta x \to 0}} \sum_{i=1}^{n} \overbrace{f(x_i^*)}^{\text{height}} \underbrace{(x_i - x_{i-1})}_{\substack{\Delta x \\ \text{base width}}}$$



where $x_i^* \in [x_i, x_{i-1}]$.

23.1 Double integrals

Suppose we have a surface $z = f(x, y)$ $\underbrace{\text{above}}$ a planar region R in the x-y plane. i.e. $f(x,y) > 0$ for $(x,y) \in R$

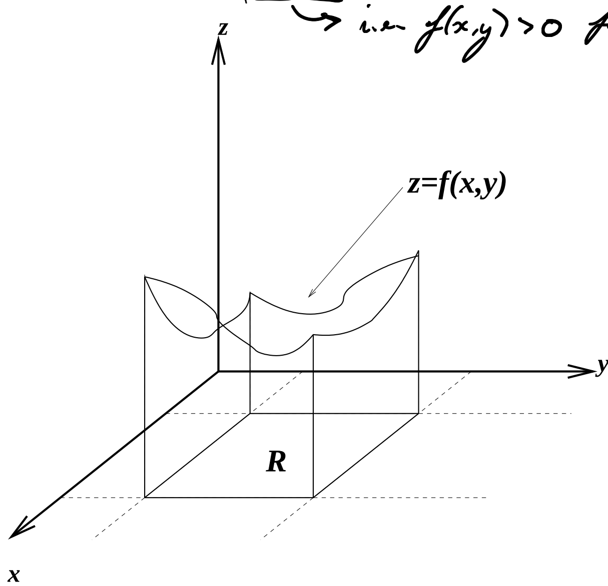
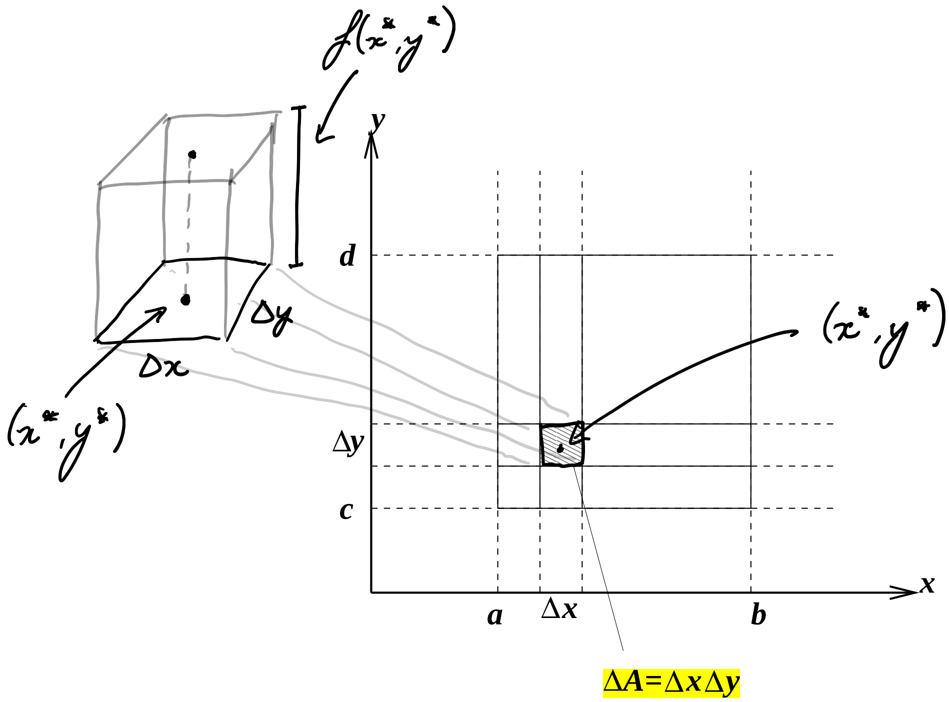


Figure 10: What is the volume V under the surface?

Before moving onto general regions, we start by considering the case where R is a rectangle. That is,

$$R = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, \ c \leq y \leq d\}.$$



Start by dividing R into subrectangles by dividing the interval $[a, b]$ into m subintervals $[x_{i-1}, x_i]$, each of width $\Delta x = \dfrac{b-a}{m}$ and $[c, d]$ into n subintervals $[y_{j-1}, y_j]$ of equal width $\Delta y = \dfrac{d-c}{n}$.

Combining these gives a rectangular grid R_{ij} with subrectangles each of area $\Delta A = \Delta x \Delta y$.

In each subrectangle take any point P_{ij} with co-ordinates (x^*_{ij}, y^*_{ij}).

The volume of the box with base the rectangle ΔA and height the value of the function $f(x, y)$ at the point P_{ij} (so the box touches the surface at a point directly above P_{ij} - see figure 11) is

$$V_{ij} = \overbrace{f(x^*_{ij}, y^*_{ij})}^{\text{height}} \underbrace{\Delta A}_{\text{base area}}.$$

Then for all the subrectangles we have an approximation to the required volume V:

$$V \approx \sum_{i=1}^{m} \sum_{j=1}^{n} f(x^*_{ij}, y^*_{ij}) \Delta A,$$

the double Riemann sum.

(typically just write

$$\sum_{\text{(grid squares)}} f(x^*, y^*) \Delta A$$
)

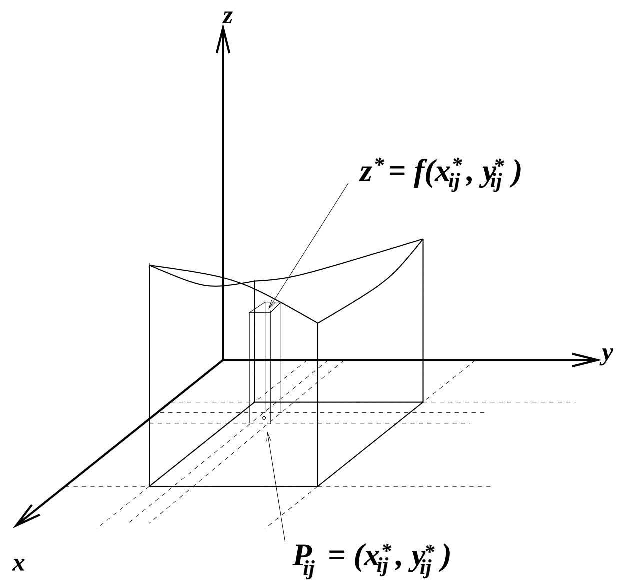


Figure 11: The rectangular box whose volume is $z^*\Delta A$.

Let $\overbrace{\Delta x \to 0 \text{ and } \Delta y \to 0}^{\Delta A \to 0}$, ie $m \to \infty$ and $n \to \infty$, then we *define* the volume to be

$$V = \lim_{m\to\infty} \lim_{n\to\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)\Delta A,$$

if the limits exist and we write this as

$$\iint\limits_R f(x,y)dA.$$

Def'n of double integral $= \lim\limits_{\Delta A \to 0} \sum f(x_i^*, y_j^*)\Delta A$ (grid squares)

We call f integrable if the limits exist. Note that every continuous function is integrable.

23.2 Properties of the double integral

(Think of these properties in terms of volume below a surface)

(i) $\iint\limits_R (f \pm g)dA = \iint\limits_R f dA \pm \iint\limits_R g dA$

$f = f(x,y)$, $g = g(x,y)$

(ii) $\iint\limits_R cf dA = c \iint\limits_R f dA$ $\quad$ c constant

(iii) $\iint\limits_R f dA = \iint\limits_{R_1} f dA + \iint\limits_{R_2} f dA$ $\quad$ $R = R_1 \cup R_2$

(disjoint R_1, R_2 except perhaps on boundary e.g. R1 | R2 $\quad R$)

(iv) If $f(x,y) \geq g(x,y)$ for all $(x,y) \in R$ then

$$\iint\limits_R f dA \geq \iint\limits_R g dA$$

23.3 Iterated integrals

We define $\int_c^d f(x,y)dy$ to mean that x is fixed and $f(x,y)$ is integrated with respect to y from $y=c$ to $y=d$. So

$$A(x) = \int_c^d f(x,y)dy$$

is a function of x only.

If we now integrate $A(x)$ with respect to x from $x=a$ to $x=b$ we have

$$\begin{aligned} \int_a^b A(x)dx &= \int_a^b \left[\int_c^d f(x,y)dy\right] dx \\ &= \int_a^b \int_c^d f(x,y)\ dy\ dx \end{aligned}$$

This is called an iterated integral.

23.3.1 Example: evaluate $\int_0^2 \int_1^3 x^2y\ dy\ dx$

$$= \int_0^2 \left(\int_1^3 x^2 y\ dy \right) dx$$

$$= \int_0^2 \left[\frac{1}{2} x^2 y^2 \right]_1^3 dx$$

$$= \int_0^2 \left(\frac{9}{2}x^2 - \frac{1}{2}x^2 \right) dx$$

$$= \int_0^2 4x^2\ dx = \frac{4}{3}x^3 \Big|_0^2 = \frac{4}{3} \times 8 = \frac{32}{3}$$

Now try integrating the other way around:

23.3.2 Example: evaluate $\int_1^3 \int_0^2 x^2 y \; dx \; dy$

$$= \int_1^3 \left(\int_0^2 x^2 y \, dx \right) dy$$

$$= \int_1^3 \left[\frac{x^3}{3} y \right]_0^2 dy$$

$$= \int_1^3 \left(\frac{8}{3} y \right) dy$$

$$= \frac{8}{6} y^2 \Big|_1^3 = \frac{8}{6}(9-1) = \frac{32}{3}$$

same answer as previous example!

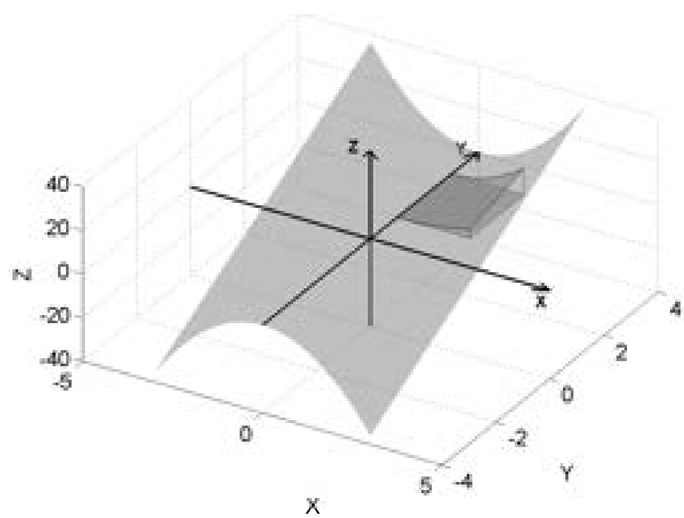


Figure 12: We have just calculated the volume of the solid outlined above.

Notes.

24 Fubini's theorem, volume by slabs

By the end of this section, you should be able to answer the following questions:

- What is Fubini's theorem?
- How is the double integral related to the iterated integral?
- How do you estimate the volume below a surface using slabs?

24.1 Fubini's theorem

If $f(x,y)$ is integrable on the rectangle

$$R = \{(x,y)|a \le x \le b,\ c \le y \le d\},$$

then

$$\begin{aligned}\iint_R f(x,y)dA &= \int_a^b \int_c^d f(x,y)\ dy\ dx \\ &= \int_c^d \int_a^b f(x,y)\ dx\ dy\end{aligned}$$

24.2 Example: evaluate $\iint_R (x^2+y^2)\overbrace{dA}$ where

"dx dy" or "dy dx"

$R = \{(x,y)|0 \le x \le 2,\ 0 \le y \le 1\}$

By Fubini's Theorem,

$$\iint_R (x^2+y^2)dA = \int_0^2\left(\int_0^1 (x^2+y^2)\ dy\right)dx$$

$$= \int_0^2 \left[yx^2 + \frac{y^3}{3}\right]_0^1 dx$$

$$= \int_0^2 x^2 + \frac{1}{3}\ dx$$

$$= \left[\frac{x^3}{3} + \frac{x}{3}\right]_0^2 = \frac{8}{3} + \frac{2}{3} = \frac{10}{3}$$

In the other order of integration

$$\iint_R (x^2+y^2)\,dA = \int_0^1 \left(\int_0^2 (x^2+y^2)\,dx\right) dy$$

$$= \int_0^1 \left[\frac{1}{3}x^3 + xy^2\right]_0^2 dy$$

$$= \int_0^1 \left(\frac{8}{3} + 2y^2\right) dy$$

$$= \left[\frac{8}{3}y + \frac{2}{3}y^3\right]_0^1$$

$$= \frac{8}{3} + \frac{2}{3} = \frac{10}{3}$$

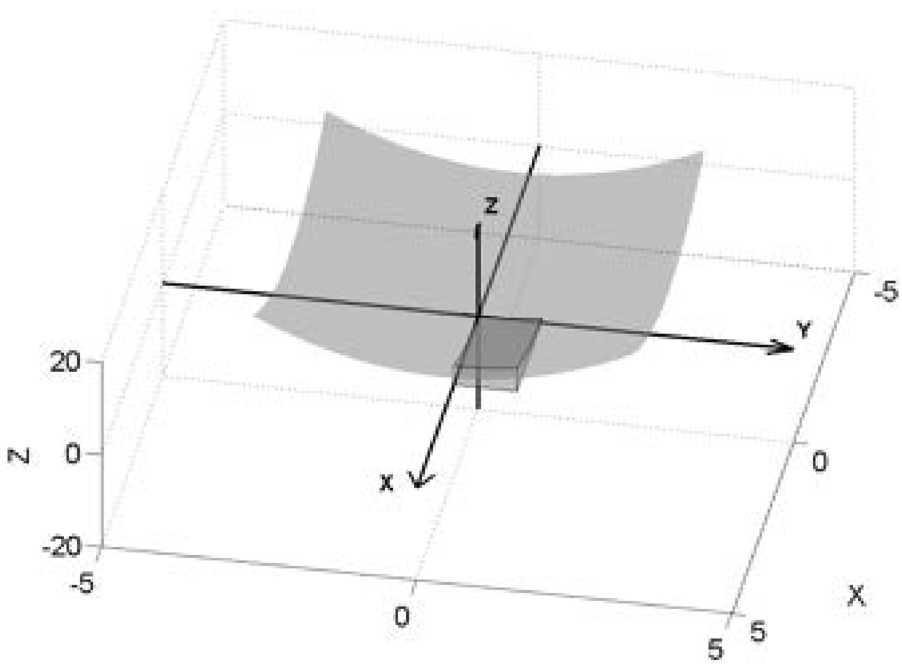


Figure 13: A representation of the volume in example 24.2.

24.3 Interpreting Fubini's theorem in terms of volume

Fubini's theorem is the key result that tells us how to evaluate a double integral. We can see the relation between the iterated integral and the double integral by considering an alternative way of calculating the volume below a surface.

Suppose we want to find the volume below the surface $\underline{z = x^2y}$ above the square region $0 \le x \le 8$ and $0 \le y \le 4$.

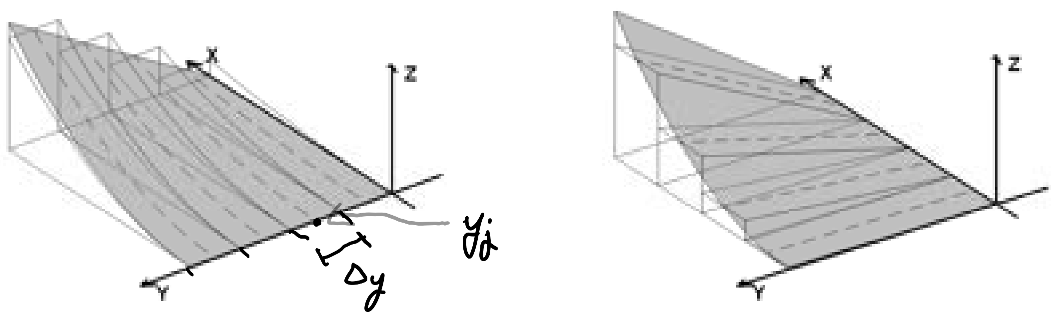
A natural way to solve this problem is to break the region up into slabs of equal depth $\Delta y = y_{j+1} - y_j$ located at y_j, and add up the volume of the slabs

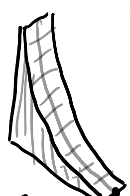
$$V \approx \sum_j \Delta V,$$

where ΔV the volume of the jth slab. Figure 14 below shows two ways of doing this using four slabs in each case. The left diagram follows the method outlined here, taking slabs of thickness Δy.

Four slabs parallel to x-z plane

surface $z = f(x,y) = x^2y$





cross-sectional area $C(y_j)$

Figure 14: Two ways of approximating the volume under $z = x^2y$ using four slabs.

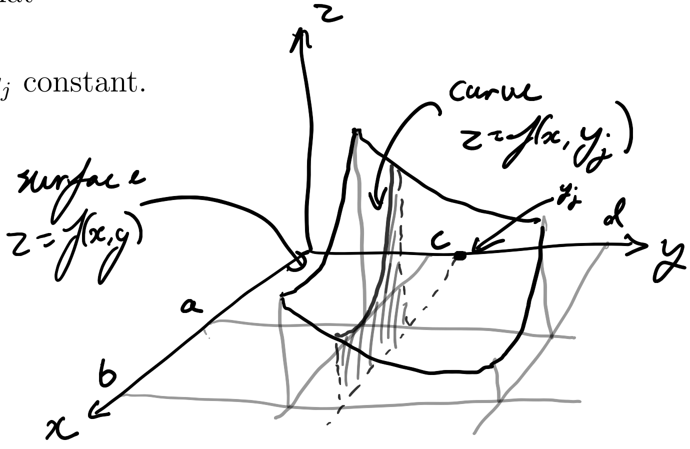
If the slab is very thin (i.e. $\Delta y \ll 1$) then the volume of each slab is

$$\Delta V \approx \boxed{\text{Area of slab}} \times \overbrace{\text{Depth}}^{\Delta y} = C(y_j)\Delta y.$$

cross-sectional area

Here $C(y_j)$ is the area of the slab at the location y_j (and the result will depend on y_j!). From one-dimensional calculus we know exactly that

$$C(y_j) = \int_0^8 f(x, y_j)dx \quad y_j \text{ constant.}$$



It is easy to compute this as a regular integral since y_j does not vary with x. Putting all this together

$$V \approx \sum_j \Delta V_j \approx \sum_j C(y_j)\Delta y.$$

As the slabs become thinner and thinner ($\Delta y \to 0$) the approximation becomes more accurate and we can replace the summation by an integral[2]

$$V = \int_0^4 C(y)dy = \int_0^4 \left(\int_0^8 f(x,y)dx \right) dy$$

Note that the y is held constant in the inner integral.

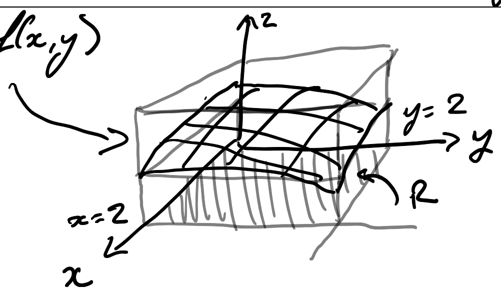
A similar argument can be applied by considering slabs of depth Δx, located at x_j. In other words, take slabs that are parallel to the y-z plane.

Summary: volume defined as $\iint_R f(x,y)\,dA$, but evaluated as an iterated integral

24.4 Example: find the volume of the solid bounded by the elliptic paraboloid $x^2 + 2y^2 + z = 16$, the planes $x = 2$ and $y = 2$, and the three coordinate planes. $\Rightarrow x=0, y=0, \boxed{z=0}$

$\uparrow$ x-y plane

surface $z = 16 - x^2 - 2y^2 = f(x,y)$



vol. below surface & above R in x-y plane $\Bigg\} = \iint_R f(x,y)\,dA$

$$R = \{(x,y) \mid 0 \le x \le 2,\ 0 \le y \le 2\}$$

$$\Rightarrow \text{vol} = \int_0^2 \left(\int_0^2 (16 - x^2 - 2y^2)\,dx \right) dy$$

$$= \int_0^2 \left[16x - \frac{x^3}{3} - 2xy^2 \right]_{x=0}^{x=2} dy$$

---

[2]Recall that is in fact the definition of an integral

$$= \int_0^2 \left(32 - \frac{8}{3} - 4y^2\right) dy = \ldots = 48$$

exercise: try other order of integration

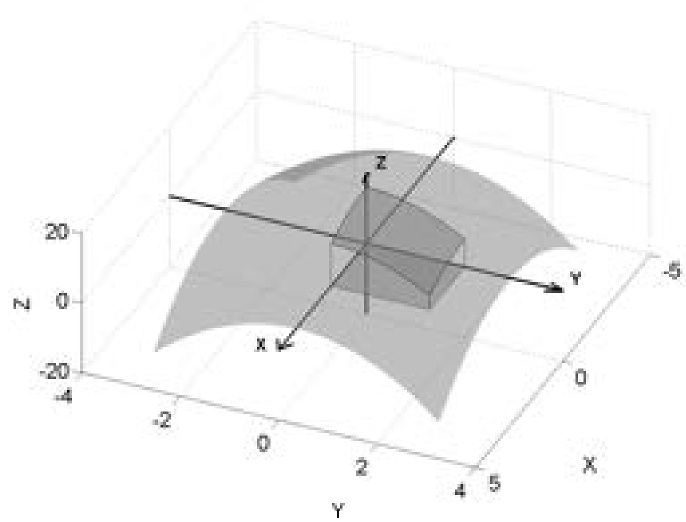


Figure 15: The volume of the solid of example 24.4 is below the surface $z = 16 - x^2 - 2y^2$ and above the x-y plane as shown.

24.5 Special case when $f(x,y) = g(x)h(y)$.

In this case we can separate the integral as follows.

rectangle R $\rightarrow$

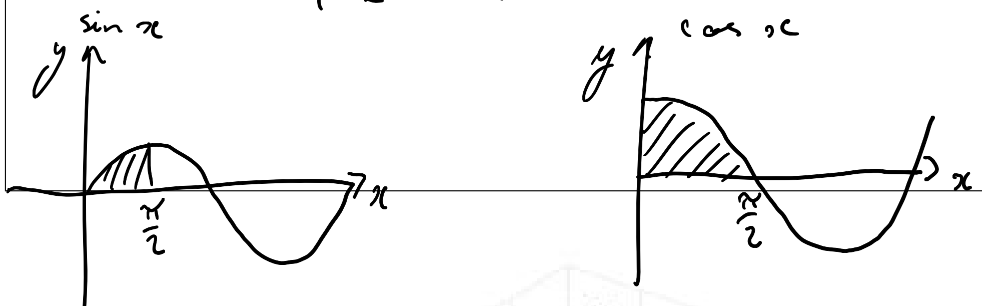
(⇒ bounds of iterated integrals are constant)

$$\iint_R f(x,y)dA = \int_c^d \int_a^b g(x)h(y)\ dx\ dy$$
$$= \int_a^b g(x)dx \int_c^d h(y)dy$$

24.5.1 Example: $\iint_R \sin x \cos y\ dA$ where $R = [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]$

$R = \{(x,y) \mid 0 \le x \le \frac{\pi}{2}, 0 \le y \le \frac{\pi}{2}\}$

$$\Rightarrow \iint_R \sin x \cos y\ dA = \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \sin x \cos y\ dx\ dy$$
$$= \left(\int_0^{\frac{\pi}{2}} \sin x\ dx\right)\left(\int_0^{\frac{\pi}{2}} \cos y\ dy\right)$$
$$= 1 \times 1 = 1$$



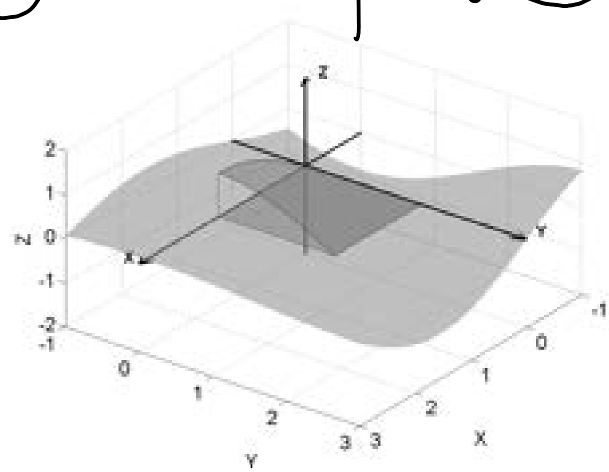


Figure 16: The volume calculated in example 24.5.1 is outlined above.

Notes.

25 Integrals over general regions

By the end of this section, you should be able to answer the following questions:

- How can you identify type I and II regions?
- How do you evaluate a double integral over type I and II regions?
- How can you evaluate a double integral over a more general region comprising finitely many type I and II regions?
- What is meant by net volume below a surface?

To find the double integral over a general region D instead of just a rectangle we consider a rectangle which encloses D and define

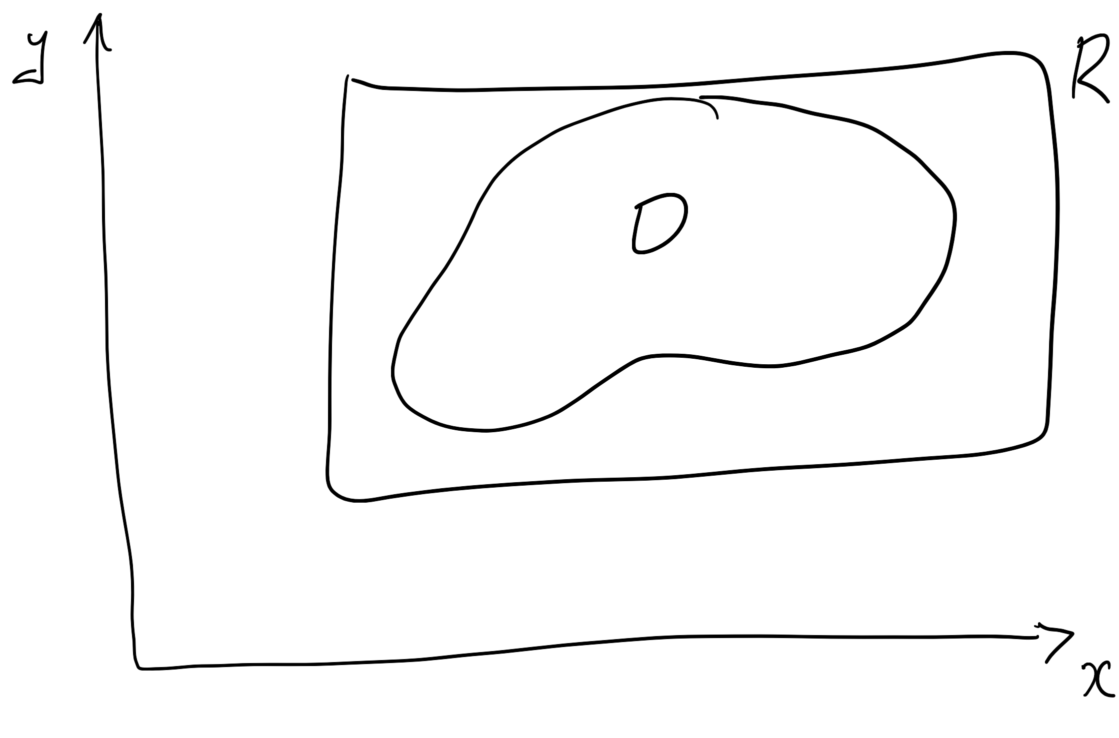
$$F(x,y) = \begin{cases} f(x,y), & \text{if } (x,y) \in D \\ 0, & \text{if } (x,y) \in R \text{ but } \notin D \end{cases}$$

then

Define $$\iint\limits_{D} f(x,y)dA = \iint\limits_{R} F(x,y)dA$$ if F is integrable on R

and we can proceed as before. It is possible to show that F is integrable if the boundary of D is bounded by a finite number of smooth curves of finite length. Note that F may still be discontinuous at the boundary of D.

& f is cts.



25.1 Type I regions

A plane region D is of type I if it lies between the graph of two continuous functions of x. That is $D = \{(x, y) | a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$.

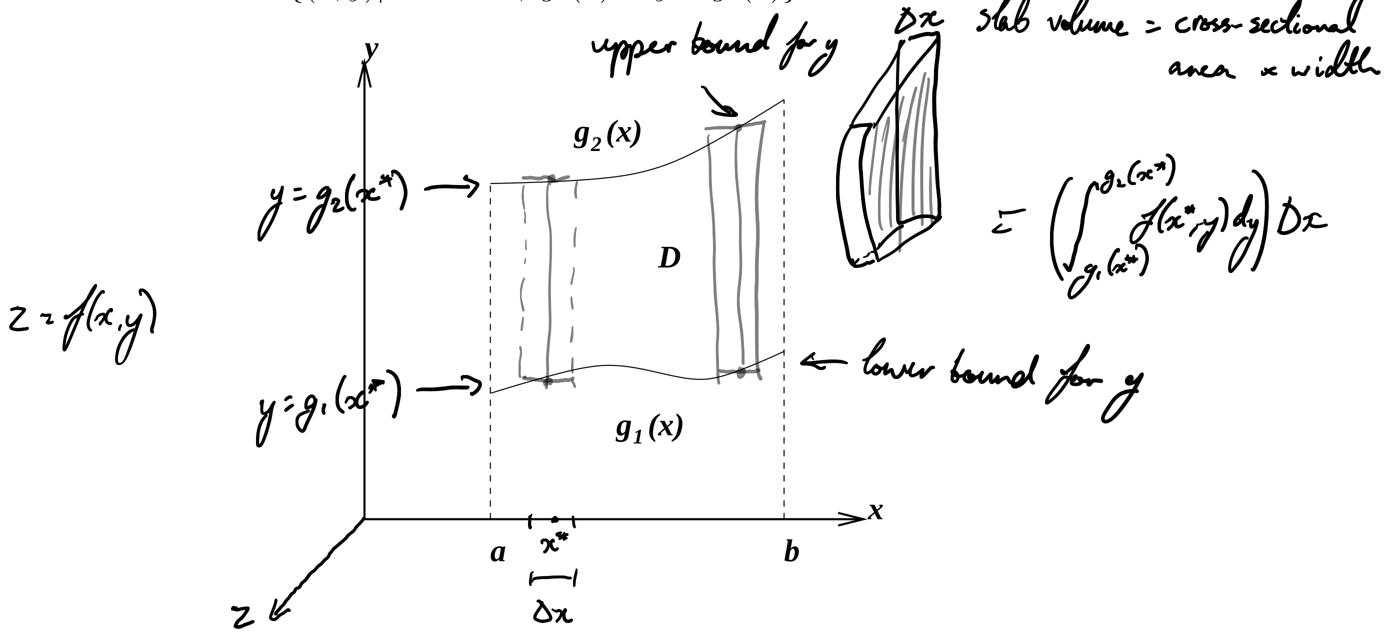
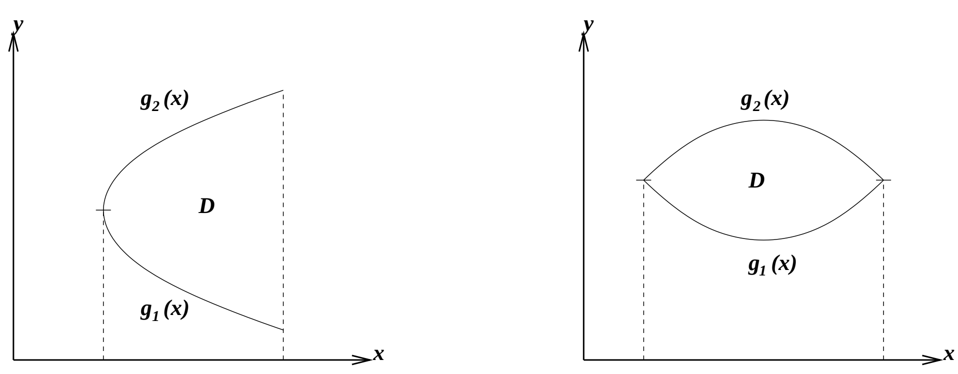


Figure 17: Type I regions are generally bounded by two constant values of x and two functions of x.

In practice, to evaluate $\iint\limits_D f(x, y)dA$ where D is a region of type I we have

$$\iint\limits_D f(x, y)dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\ dy\ dx.$$



(a) (b)

Figure 18: Some more examples of type I regions.

25.1.1 Example: find $\iint\limits_D (4x + 10y)dA$ where D is the region between the parabola $y = x^2$ and the line $y = x + 2$.

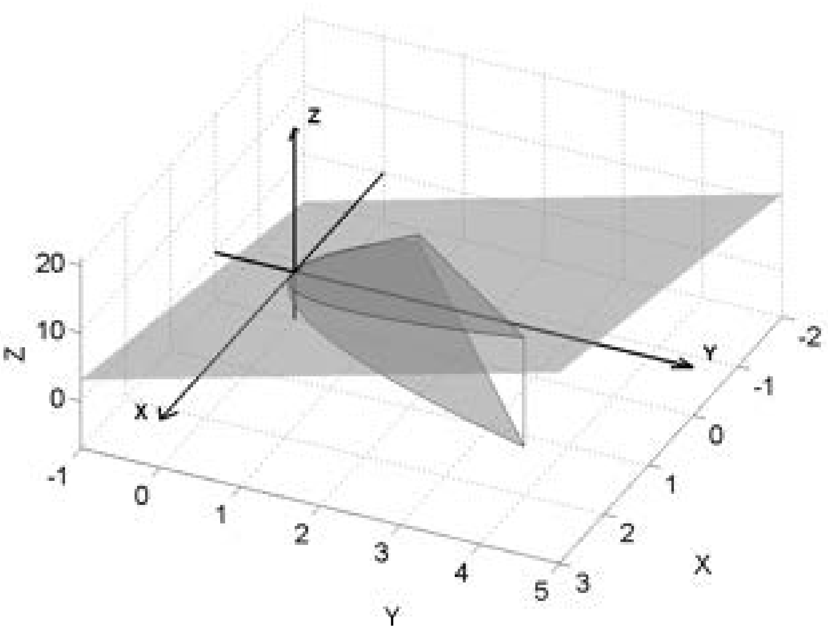
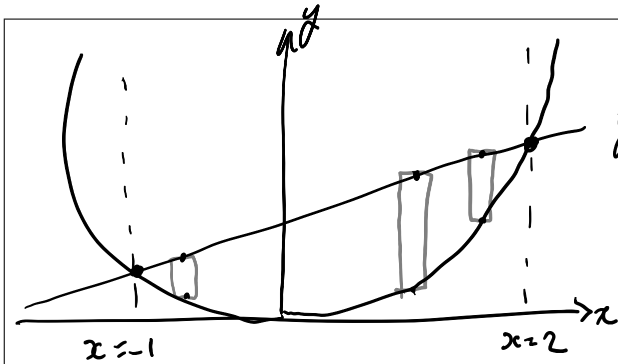


Figure 19: The volume of example 25.1.1 is outlined above.



Points of intersection

$x^2 = x + 2$

$\Rightarrow x = 2, -1$

$D = \{(x,y) \mid -1 \leq x \leq 2,\ x^2 \leq y \leq x+2\}$

$\Rightarrow \iint_D (4x + 10y)\,dA = \int_{-1}^{2} \left(\int_{x^2}^{x+2} (4x + 10y)\,dy \right) dx$

$$= \int_{-1}^{2} \left[4xy + 5y^2 \right]_{y=x^2}^{y=x+2} dx$$

$$= \int_{-1}^{2} \left(4x(x+2) + 5(x+2)^2 - \left(4x^3 + 5x^4 \right) \right) dx$$

$$= \dots = 81$$

25.2 Type II regions

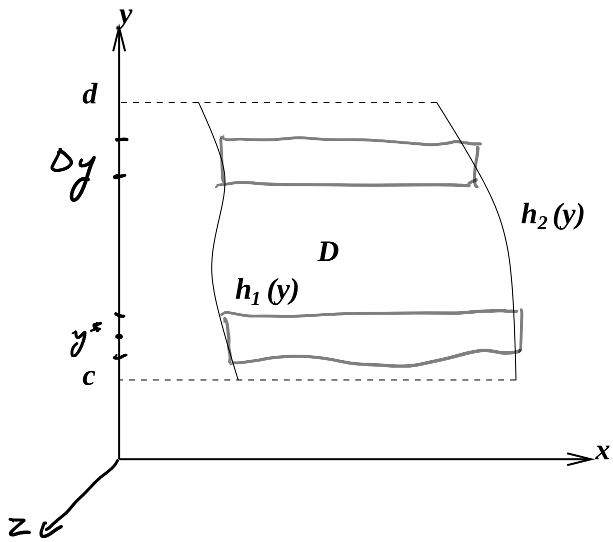


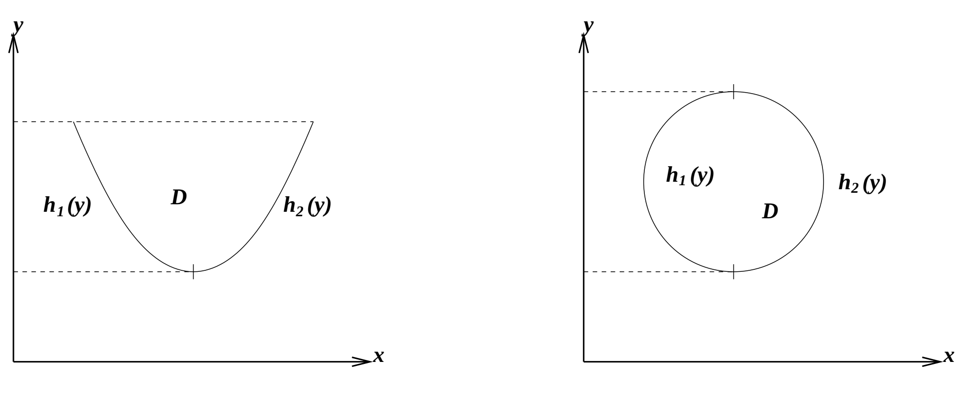
Figure 20: Type II regions are generally bounded by two constant values of y and two functions of y.

A plane region is of type II if it can be expressed by

$$D = \{(x, y) | c \leq y \leq d,\ h_1(y) \leq x \leq h_2(y)\}.$$

In practice, to evaluate $\iint\limits_D f(x, y) dA$ where D is a region of type II we have

$$\iint\limits_D f(x, y) dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\ dx\ dy.$$



(a) **(b)**

Figure 21: Some more examples of type II regions.

25.2.1 Example: evaluate $\iint_D xy\, dA$ where D is the region bounded by the line $y = x - 1$ and the parabola $y^2 = 2x + 6$.

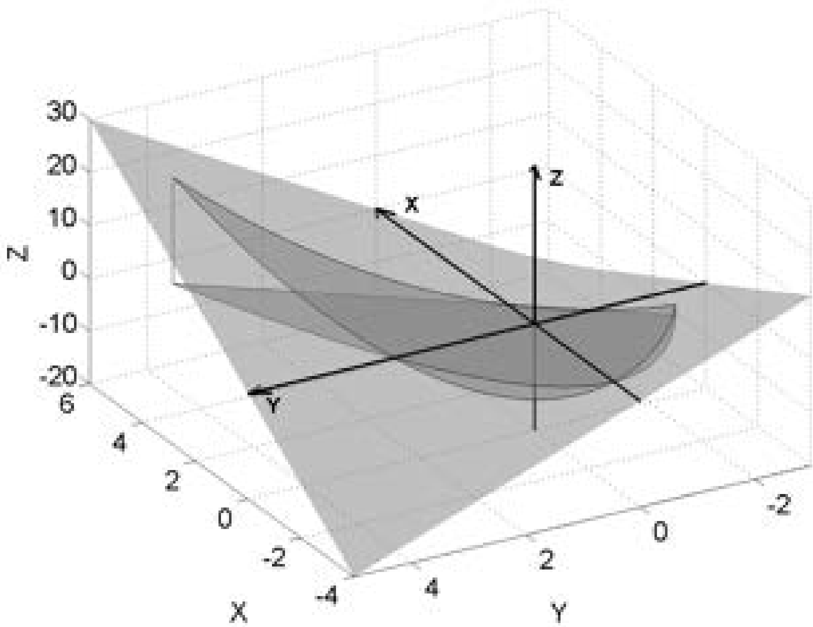


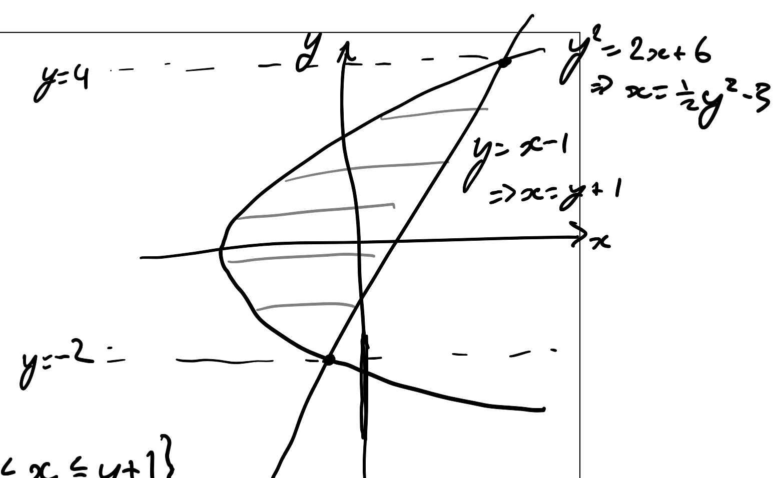
Figure 22: The volume of example 25.2.1 is outlined above. Note carefully that the surface is *above* the x-y plane only in the quadrants where $x, y > 0$ and $x, y < 0$. For x and y values in the other two quadrants, the surface is *below* the x-y plane. Hence in this example we are calculating the "net volume" lying above the x-y plane.

Points of intersection

$$y^2 = 2(y+1) + 6$$

$$\Rightarrow y^2 - 2y - 8 = 0$$

$$\Rightarrow y = 4, -2$$



$$D = \left\{ (x,y) \mid -2 \le y \le 4,\ \tfrac{1}{2}y^2 - 3 \le x \le y+1 \right\}$$

$$f(x,y) = xy$$

$$\Rightarrow \iint_D xy\,dA = \int_{-2}^{4}\left(\int_{\frac{1}{2}y^2-3}^{y+1} xy\,dx\right)dy$$

$$= \int_{-2}^{4}\left[\tfrac{1}{2}x^2 y\right]_{x=\frac{1}{2}y^2-3}^{x=y+1} dy$$

$$= \int_{-1}^{4}\left(\tfrac{1}{2}(y+1)^2 y - \tfrac{1}{2}\left(\tfrac{1}{2}y^2-3\right)^2 y\right)dy$$

$$= \dots = 36$$

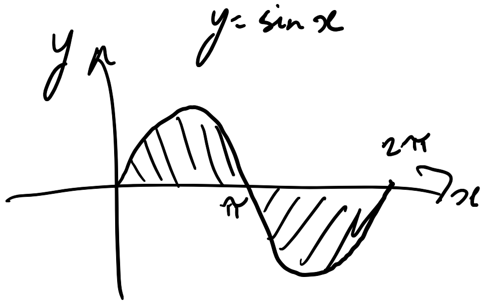
= "net volume" above x-y plane & below surface $z = f(x,y)$

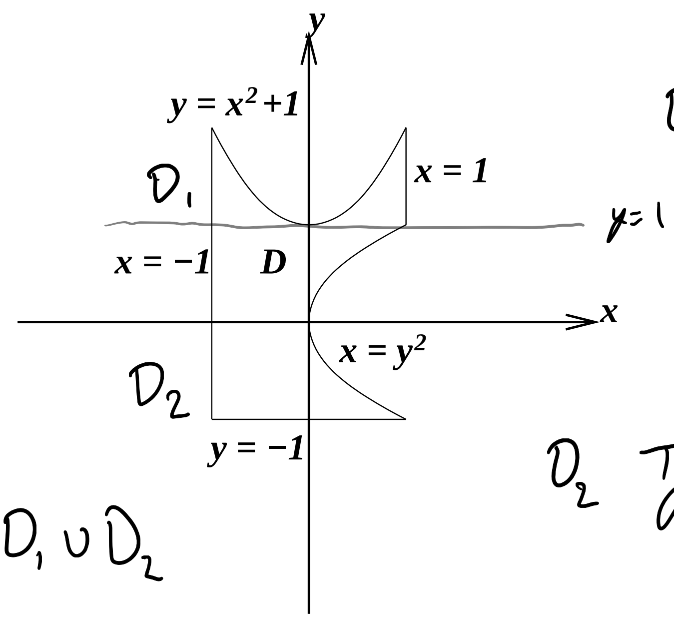
= (vol. above x-y plane) − (vol. below x-y plane)

---

c.f. $$\int_0^{2\pi} \sin x\,dx = 0$$

= "net area" below curve and above x-axis





D_1 Type 1

D_2 Type 2

$D = D_1 \cup D_2$

25.3 **Express** D **as a union of regions of type I or type II and expand the integral** $\iint\limits_D f(x,y)\ dA$**, for some integrable function** f**.**

$$\iint_D f(x,y)\,dA = \iint_{D_1} f(x,y)\,dA + \iint_{D_2} f(x,y)\,dA$$

$$D_1 = \{(x,y) \mid -1 \le x \le 1,\ 1 \le y \le x^2+1\}$$

$$D_2 = \{(x,y) \mid -1 \le y \le 1,\ -1 \le x \le y^2\}$$

$$\Rightarrow \iint_D f(x,y)\,dA = \int_{-1}^{1}\int_{1}^{x^2+1} f(x,y)\,dy\,dx$$

$$+ \int_{-1}^{1}\int_{-1}^{y^2} f(x,y)\,dx\,dy$$

Notes.

26 Interchanging order of integration

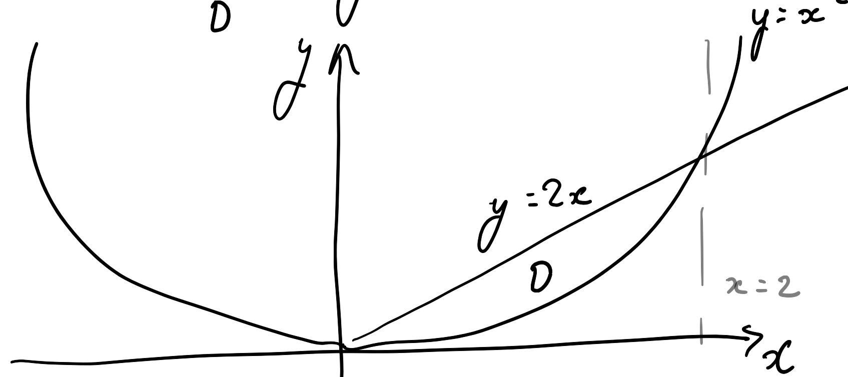
By the end of this section, you should be able to answer the following questions:

- How do you change the order of integration in a double integral?
- when might it be necessary to change the order of integration in a double integral?

It is often possible to represent a type I region as a union of type II regions, or a type II region as a union of type I regions. Why would we want to do that? In some cases, it may only be possible to integrate a function one way but not the other. In this section, we investigate this idea more closely.

26.1 Find the volume under the paraboloid $z = x^2 + y^2$ above the region D, where D is bounded by $y = x^2$ and $y = 2x$. Do the problem twice, first by taking D to be a type I region, then by taking D to be type II.

Volume $= \iint_D (x^2+y^2)\,dA$



Set $x^2 = 2x$
$\to x = 0, 2$

As a type I region, $D = \{(x,y) \mid 0 \le x \le 2,\ x^2 \le y \le 2x\}$

$$\Rightarrow \iint_D (x^2+y^2)\,dA = \int_0^2 \left(\int_{x^2}^{2x} (x^2+y^2)\,dy \right) dx$$

$$= \int_0^2 \left[x^2 y + \frac{1}{3} y^3 \right]_{y=x^2}^{y=2x} dx$$

$$= \int_0^2 \left(2x^3 + \frac{8}{3} x^3 - \left(x^4 + \frac{1}{3} x^6 \right) \right) dx = \dots = \frac{216}{35}$$

$y = 2x \Rightarrow x = \frac{1}{2}y$

$y = x^2 \Rightarrow x = \pm\sqrt{y}$, but not $-\sqrt{y}$, since $x \geq 0$

As a type 2 region, $D = \{(x,y) \mid 0 \leq y \leq 4, \frac{1}{2}y \leq x \leq \sqrt{y}\}$

$$\Rightarrow \iint_D (x^2+y^2)\,dA = \int_0^4 \int_{\frac{1}{2}y}^{\sqrt{y}} (x^2+y^2)\,dx\,dy$$

$$= \int_0^4 \left[\frac{1}{3}x^3 + xy^2\right]_{x=\frac{1}{2}y}^{x=\sqrt{y}} dy$$

$$= \int_0^4 \left(\frac{1}{3}y^{3/2} + y^{5/2} - \left(\frac{1}{3}\left(\frac{1}{2}y\right)^3 + \frac{1}{2}y^3\right)\right) dy$$

$$= \dots = \frac{216}{35}$$

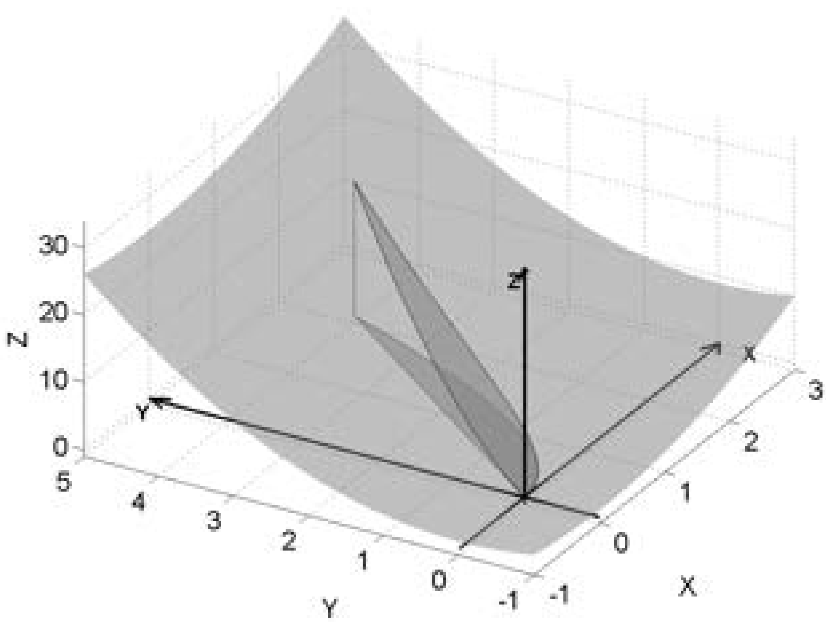


Figure 23: This volume can be calculated by treating the region in the x-y plane as either type I or II as seen in example 26.1.

In the following example, we see how it is sometimes necessary to change the order of integration in order to evaluate the integral.

26.2 Example: Find $\int_0^1 \int_x^1 \sin(y^2)\ dy\ dx$

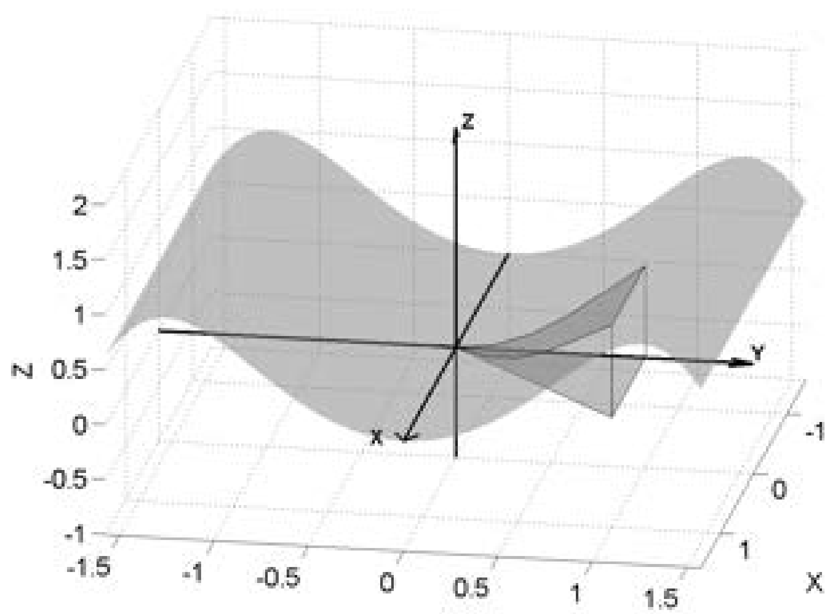
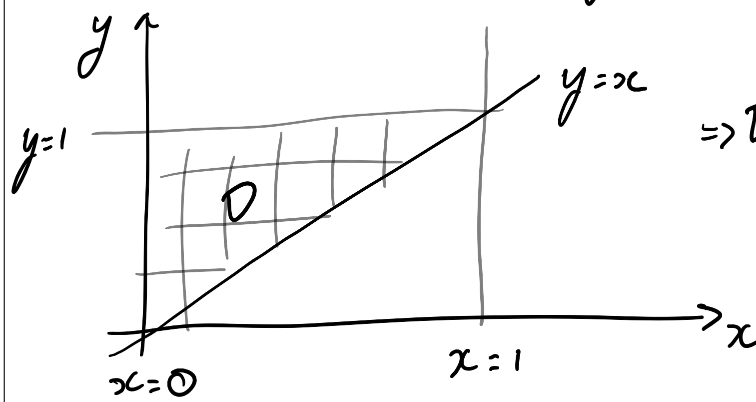


Figure 24: The volume described in example 26.2.

From Stewart: (ed. 8 p546)
"Can we integrate all continuous functions?" (No!)
e.g. $\int \sin(x^2)dx$, $\int \cos(x^2)dx$, $\int e^{\pm x^2}dx$, $\int \frac{e^x}{x}dx$,
$\int \frac{1}{\log(x)}dx$, $\int \sqrt{x^3+1}\,dx$, ...
- elementary function
- non elementary integrals
- Risch algorithm
- Liouville's principle

$\int \sin(y^2)\, dy$ is NOT expressible in terms of elementary functions.

$D = \{(x,y) \mid 0 \le x \le 1,\ x \le y \le 1\}$

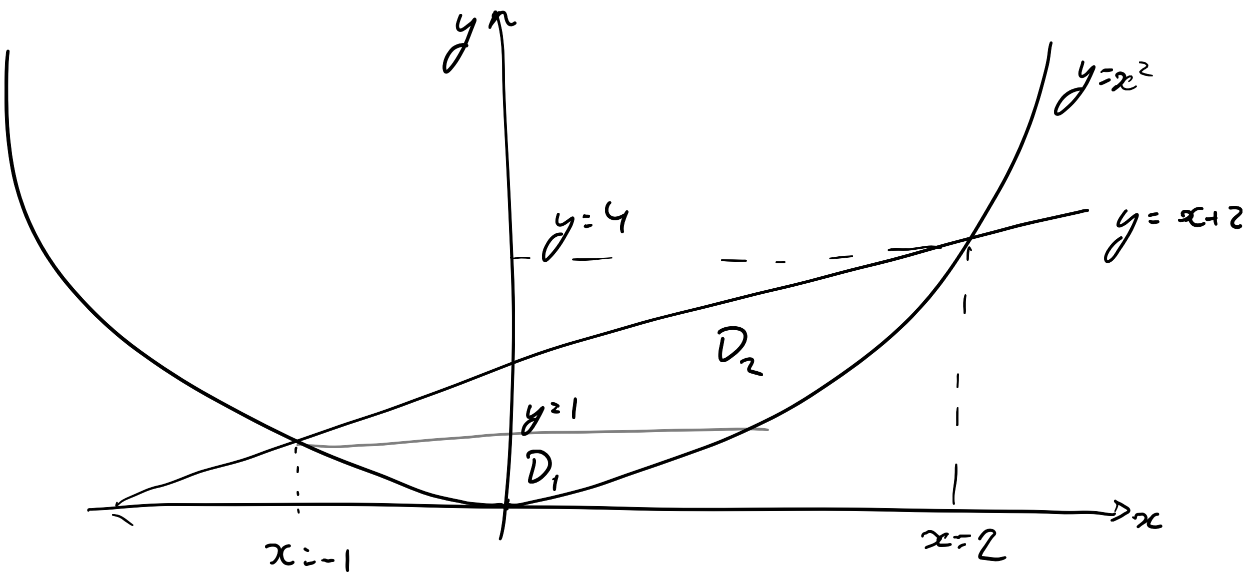


$\Rightarrow D = \{(x,y) \mid 0 \le y \le 1,\ 0 \le x \le y\}$

$$\begin{aligned}
\Rightarrow \text{integral} &= \int_0^1 \int_0^y \sin(y^2)\, dx\, dy \\
&= \int_0^1 \left[x \sin(y^2)\right]_{x=0}^{x=y} dy \\
&= \int_0^1 \left(y \sin(y^2) - 0\right) dy \\
&= -\frac{1}{2}\cos(y^2) \Big|_{y=0}^{y=1} = \frac{1}{2}\left(1 - \cos(1)\right)
\end{aligned}$$

Main point: If you cannot integrate in one order of integration, try changing the order!

Notes. (from p177)



As a type 2 region,

$D = \cancel{\{(x,y) \mid 0 \le y \le 4,\ ?? \le x \le ??\}}$ Wrong approach!

Instead, set $D = D_1 \cup D_2$

$$D_1 = \{(x,y) \mid 0 \le y \le 1,\ -\sqrt{y} \le x \le \sqrt{y}\}$$

$$D_2 = \{(x,y) \mid 1 \le y \le 4,\ y-2 \le x \le \sqrt{y}\}$$

$$\Rightarrow \int_{-1}^{2}\int_{x^2}^{x+2} f(x,y)\,dy\,dx = \int_0^1\int_{-\sqrt{y}}^{\sqrt{y}} f(x,y)\,dx\,dy + \int_1^4\int_{y-2}^{\sqrt{y}} f(x,y)\,dx\,dy$$

$\left(\iint_D f(x,y)\,dA = \iint_{D_1} f(x,y)\,dA + \iint_{D_2} f(x,y)\,dA, \quad D = D_1 \cup D_2 \right.$ (disjoint union as indicated) $\Big)$

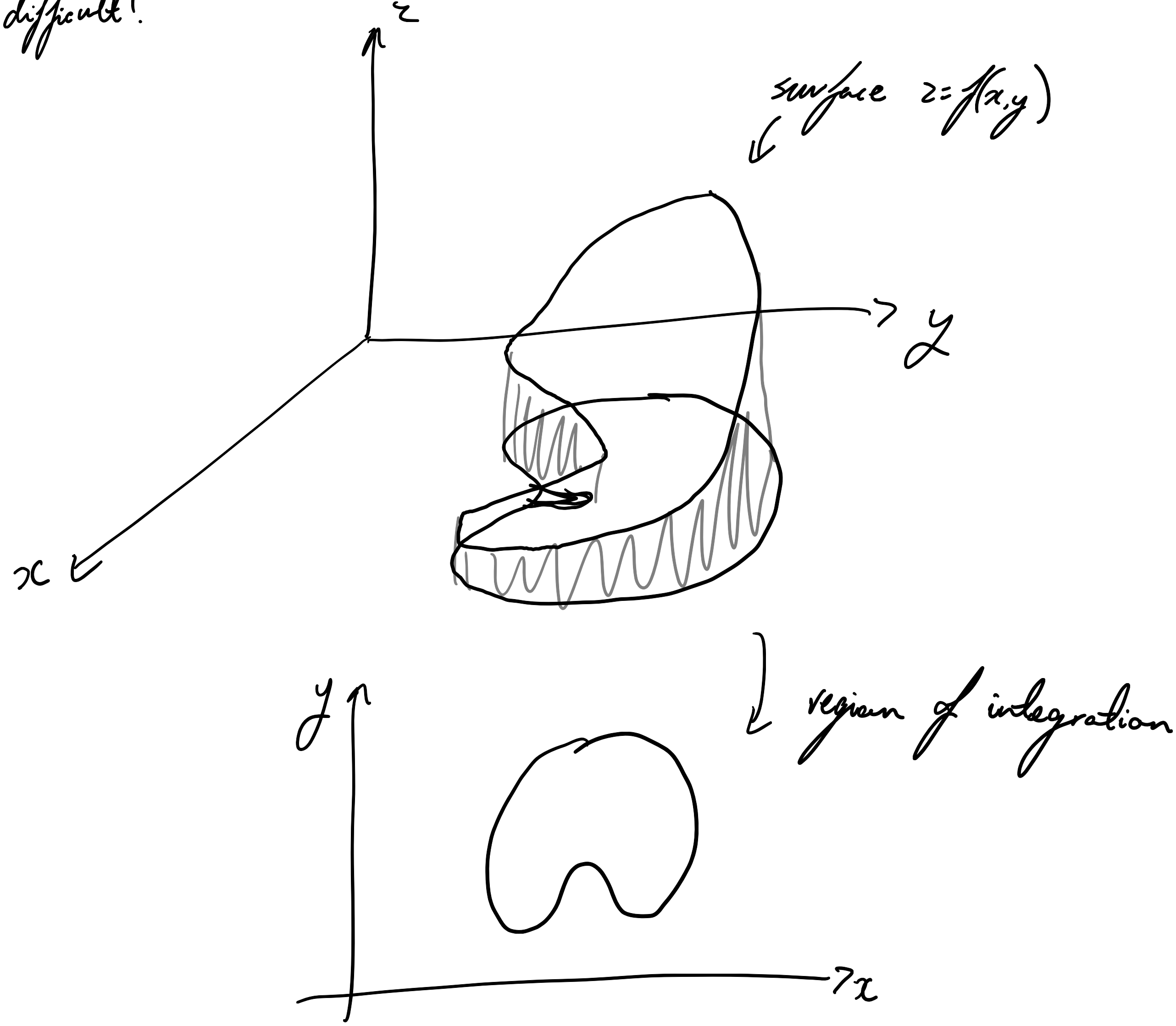
27 Review of applications: volume, area

Main points:

- This section is a review of applications of the double integral such as calculating net volume and area in the plane.
- By this stage you should be comfortable with using a double integral to calculate the net volume below a surface.
- You should know how to find the area of a general region in the plane.

When the regions are more difficult, it is a good idea to draw two diagrams - the 3-D diagram with the x-y-z axes and the 2-D one of the region in the x-y plane.

difficult!



27.1 Example: Find the volume of the tetrahedron bounded by the planes $x + 2y + z = 2$, $x = 2y$, $\underbrace{x = 0}_{y\text{-}z \text{ plane}}$ and $\underbrace{z = 0}_{x\text{-}y \text{ plane}}$.

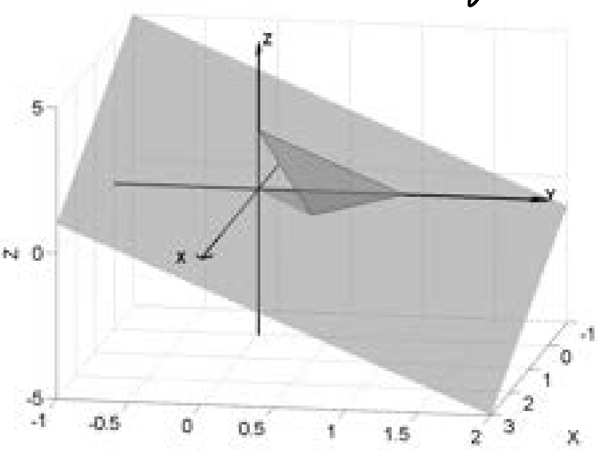
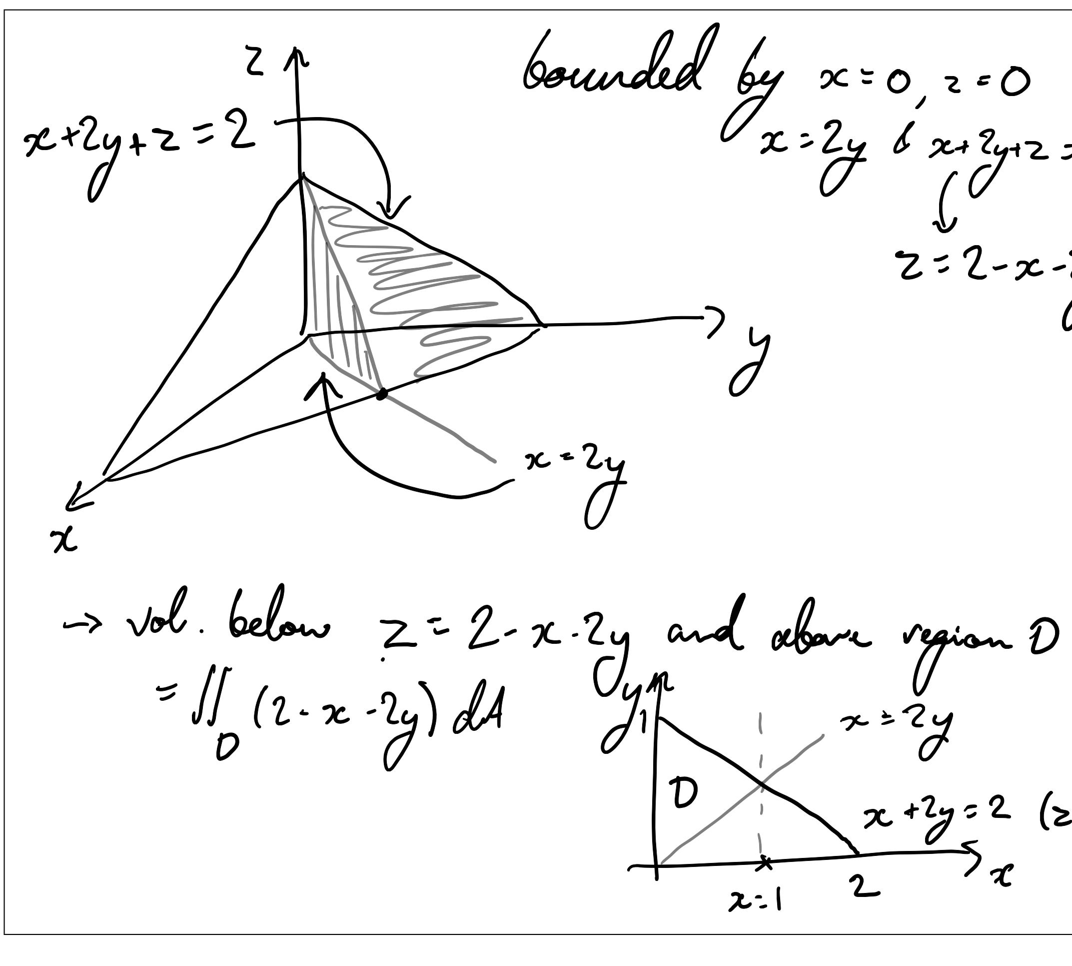


Figure 25: You should be able to reproduce a diagram like this one as an aid to determining the bounds of integration.



As a type 1 region,

$$D = \left\{(x,y) \mid 0 \le x \le 1,\ \tfrac{1}{2}x \le y \le \tfrac{1}{2}(2-x)\right\}$$

$$\Rightarrow \text{vol.} = \int_0^1 \int_{\frac{1}{2}x}^{\frac{1}{2}(2-x)} (2-x-2y)\, dy\, dx$$

$$= \int_0^1 \left[2y - xy - y^2\right]_{y=\frac{1}{2}x}^{y=\frac{1}{2}(2-x)} dx$$

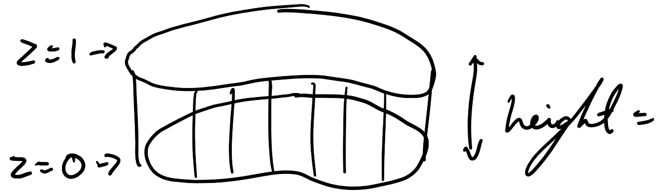
$$= \int_0^1 \left(2 - x - \tfrac{1}{2}x(2-x) - \tfrac{1}{4}(2-x)^2 - \left(x - \tfrac{1}{2}x^2 - \tfrac{1}{4}x^2\right)\right) dx$$

$$= \dots = \tfrac{1}{3}$$

As a type 2 region

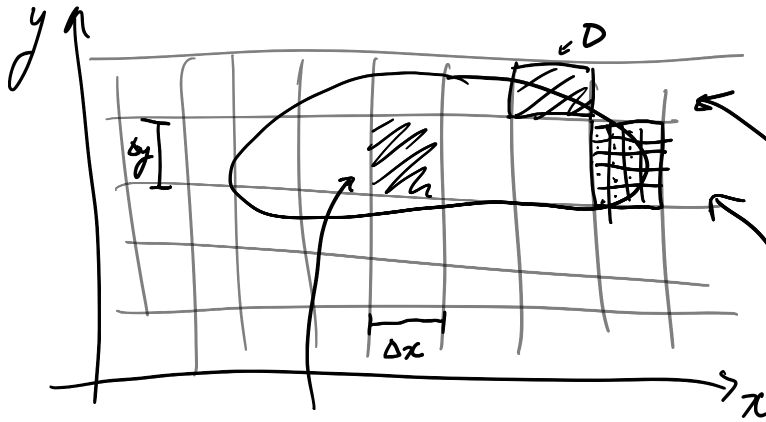
$$D = \left\{(x,y) \mid 0 \le y \le \tfrac{1}{2},\ 0 \le x \le 2y\right\}$$

$$\cup \left\{(x,y) \mid \tfrac{1}{2} \le y \le 1,\ 0 \le x \le 2-2y\right\}$$



27.2 Area

Note that if we take $f(x,y)=1$, we have $\displaystyle\iint_D 1\ dA =$ area of the region D.



area $\Delta A = \Delta x \Delta y$

$\Rightarrow$ area of region $\approx \sum \Delta A$ (..)

not exact on boundary

→ better approximation with smaller grid squares

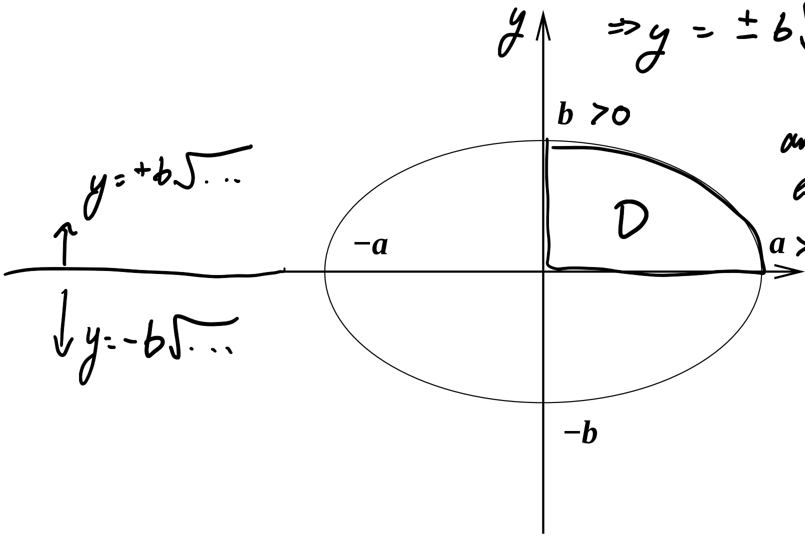
In the limit $\Delta A \to 0$ (i.e as Δx & Δy simultaneously $\to 0$)

$$\Rightarrow \text{area} = \lim_{\Delta A \to 0} \sum \Delta A = \iint_D dA$$

$$\left(\overset{\text{or}}{=} \iint_D 1\, dA\right)$$

27.3 Find the area enclosed by the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$

$\Rightarrow y = \pm b\sqrt{1 - \frac{x^2}{a^2}}$



area of ellipse $= 4 \times \left\{ \text{area of } D \right.$

area of $D = \iint_D 1\, dA$

$D = \left\{ (x,y) \mid 0 \le x \le a,\ 0 \le y \le b\sqrt{1 - \frac{x^2}{a^2}} \right\}$

$$\Rightarrow \iint_D dA = \int_0^a \int_0^{b\sqrt{1-\frac{x^2}{a^2}}} 1\, dy\, dx$$

$$= \int_0^a \Big[y \Big]_{y=0}^{y=b\sqrt{1-\frac{x^2}{a^2}}} dx$$

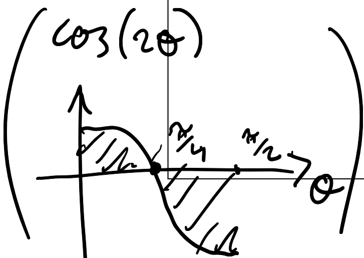
$$= \int_0^a b\sqrt{1 - \frac{x^2}{a^2}}\, dx \rightarrow \text{set } x = \boxed{a \sin\theta} \quad \frac{dx}{d\theta} = a\cos\theta$$

$\theta = 0 \rightarrow x = 0$

$\theta = \frac{\pi}{2} \rightarrow x = a$

$$= \int_0^{\frac{\pi}{2}} b\sqrt{1 - \sin^2\theta}\; a\cos\theta\, d\theta = ab\int_0^{\frac{\pi}{2}} \cos^2\theta\, d\theta$$

$$= ab\int_0^{\pi/2} \left(\frac{1}{2} + \frac{1}{2}\underbrace{\cos(2\theta)}_{=0} \right) d\theta$$



$$= \frac{\pi}{4} ab$$

$\Rightarrow$ area of ellipse $= \pi ab$

Notes.

Area of D} $= \iint_D dA \quad \left(= \iint_D 1\, dA\right)$

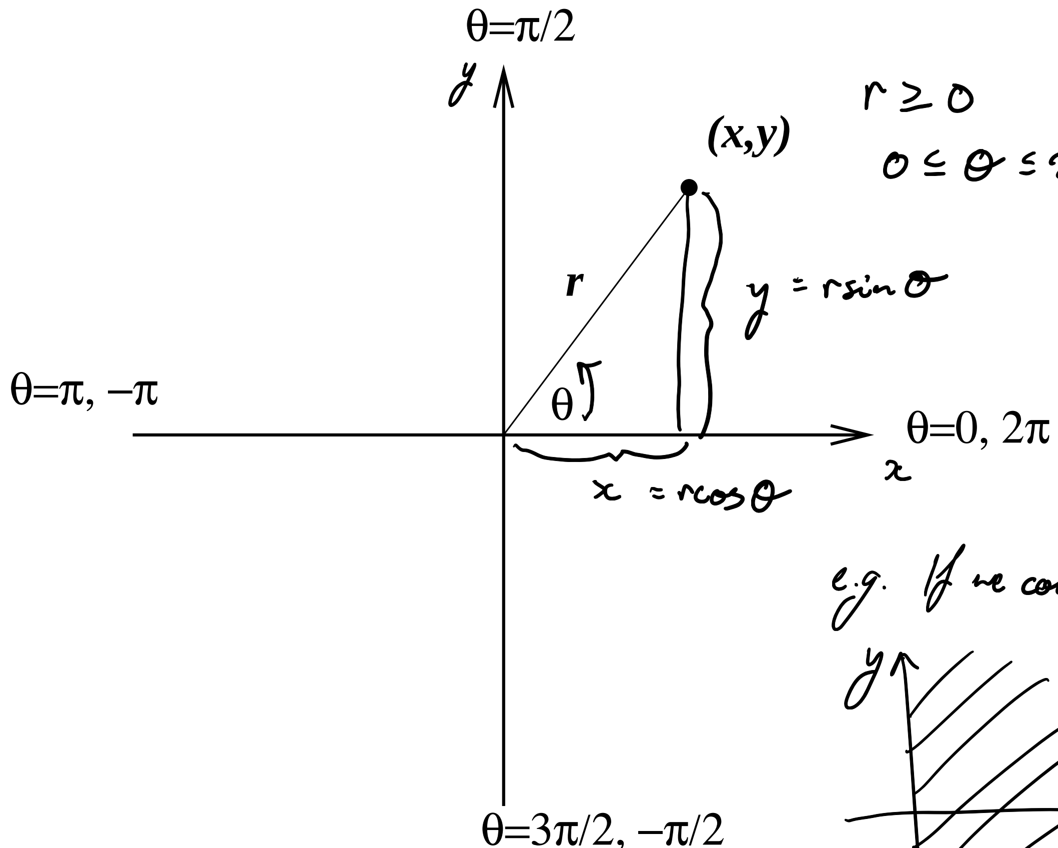
28 Double integrals in polar coordinates

By the end of this section, you should be able to answer the following questions:

- What is the relationship between polar coordinates and rectangular coordinates?
- How do you transform a double integral in rectangular coordinates into one in terms of polar coordinates?
- What is the Jacobian of the transformation?

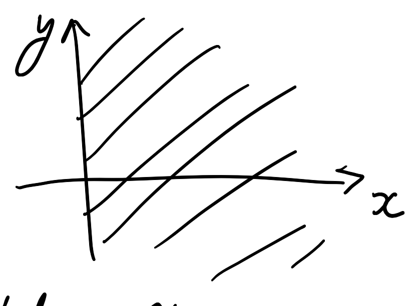
For annular regions with circular symmetry, rectangular coordinates are difficult. It can be more convenient to use *polar coordinates.*

The following diagram explains the relationship between the polar variables r, θ and the usual rectangular ones x, y.



$r \geq 0$
$0 \leq \theta \leq 2\pi$

e.g. If we consider



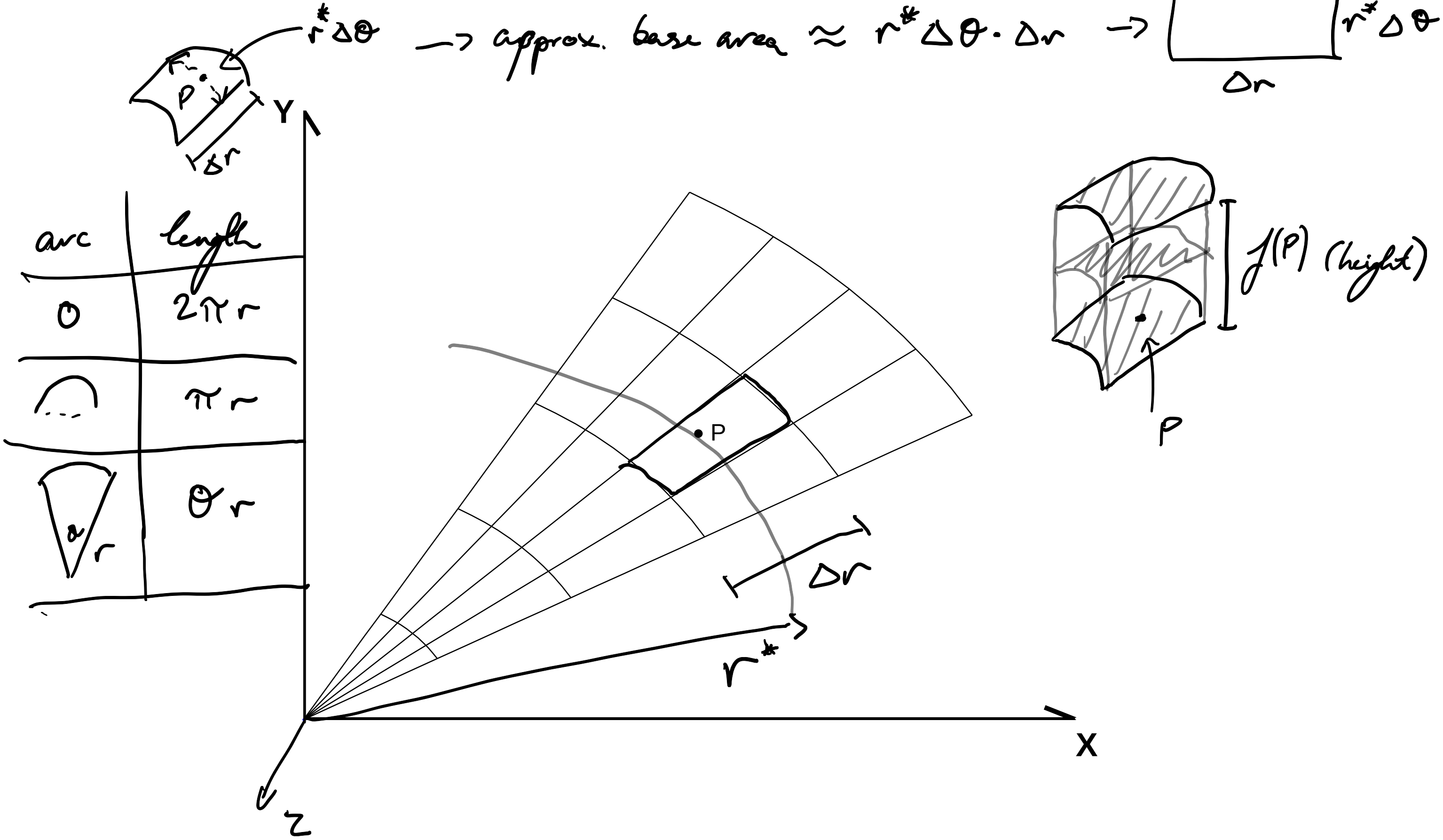
take $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$
rather than
$\left\{0 \leq \theta \leq \frac{\pi}{2}\right\} \cup \left\{\frac{3\pi}{2} \leq \theta \leq 2\pi\right\}$

For polar coordinates, we have

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Consider the volume of a solid beneath a surface $\overset{f(x,y)>0}{\boxed{z = f(x, y)}}$ and above a circular region in the x-y plane.

We divide the region into a polar grid as in the following diagram:



We first approximate the area of each polar rectangle as a regular rectangle. We do this as follows. Choose a point P inside each polar rectangle in the polar grid. Let $\boxed{P = (x^*, y^*)}$ or in polar coordinates $\boxed{P = (r^*, \theta^*)}$, where

$$x^* = r^* \cos\theta^*, \quad y^* = r^* \sin\theta^*.$$

The area of the polar rectangle containing P can be approximated as $r^*\Delta\theta\Delta r$. Therefore the volume under the surface and above each polar rectangle can be approximated as

$$\text{vol. one box } \approx \overbrace{r^*\Delta\theta\Delta r}^{\text{base . area}} \underbrace{f(r^*\cos\theta^*, r^*\sin\theta^*)}_{\text{height}}.$$

Here $f(r^*\cos\theta^*, r^*\sin\theta^*)$ is the value of the function at the point P, which is also the height of the box used in the approximation. To obtain an approximation for the entire

"Jacobian of the polar variable transformation"

$\boxed{r}\, d\theta\, dr$ $\quad x = r\cos\theta$, $y = r\sin\theta$

In general $x = x(u,v)$, $y = y(u,v)$

Jacobian $= \left|\det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}\right|$

volume below the surface, we sum over the entire polar grid:

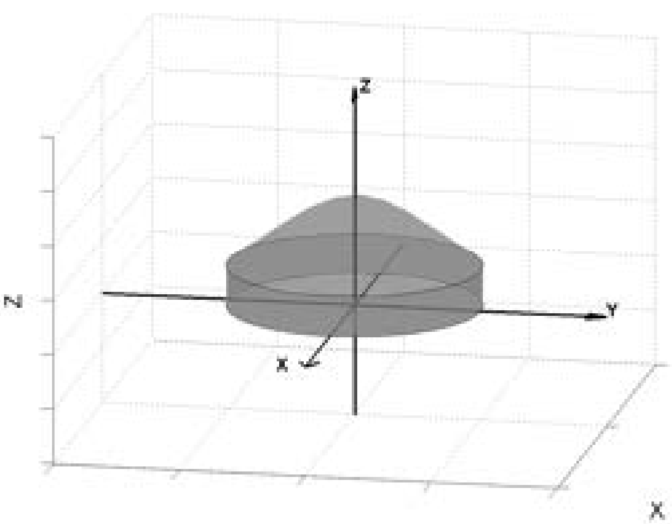
$$\begin{aligned} \text{vol.} &\approx \sum_{\text{(polar grid)}} r^*\Delta\theta\Delta r f(r^*\cos\theta^*, r^*\sin\theta^*) \\ \Rightarrow \text{vol.} &= \lim_{\Delta r,\Delta\theta\to 0} \sum_{\text{(polar grid)}} r^*\Delta\theta\Delta r f(r^*\cos\theta^*, r^*\sin\theta^*) \\ &= \iint_D f(r\cos\theta, r\sin\theta)\underbrace{r\ d\theta\ dr.} \end{aligned}$$

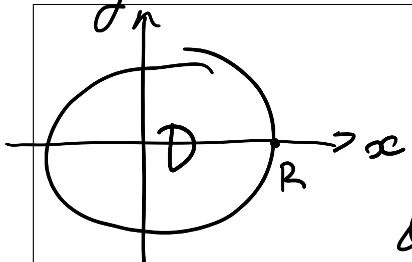
The double integral in rectangular coordinates is then transformed as follows:

$$\iint_R f(x,y)\ dx\ dy = \iint_S f(r\cos\theta, r\sin\theta) r\ dr\ d\theta.$$

$dx\,dy \mapsto r\,d\theta\,dr$ in $\iint$

28.1 Example: Find $\iint_D e^{-(x^2+y^2)}\ dx\ dy$ where D is the region bounded by the circle $x^2 + y^2 = R^2$.





Use polar coordinates

$x = r\cos\theta$, $y = r\sin\theta$

$\Rightarrow x^2 + y^2 = r^2(\cos^2\theta + \sin^2\theta) = r^2$

& $D = \{(r,\theta) \mid 0 \le r \le R,\ 0 \le \theta \le 2\pi\}$

Also in $\iint_D$, $\quad dx\,dy \mapsto r\,d\theta\,dr$

$$\rightarrow \iint_D = \int_0^R \int_0^{2\pi} e^{-r^2}\, r\, d\theta\, dr$$

$$= \left(\int_0^{2\pi} 1\, d\theta\right)\left(\int_0^R r e^{-r^2}\, dr\right)$$

$$= 2\pi \times \frac{1}{2}\left(1 - e^{-R^2}\right)$$

Let $R \to \infty$, $\iint_D$ represents vol. below $z = e^{-(x^2+y^2)}$ & above entire x-y plane

$$\Rightarrow \iint_D = \int_0^{2\pi}\int_0^{\infty} r e^{-r^2}\, dr\, d\theta = \lim_{R\to\infty} \pi\left(1 - e^{-R^2}\right) = \pi$$

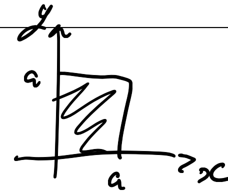
(Aside! $\displaystyle\int_0^\infty f(x)\,dx = \lim_{R\to\infty}\int_0^R f(x)\,dx$)

"improper integral".

Vol. above $x, y \geq 0$ quadrant $= \pi/4$

since $z = e^{-(x^2+y^2)}$ is symmetric about z-axis

Consider $\int_0^a \int_0^a e^{-(x^2+y^2)} \, dx \, dy$



$$= \left(\int_0^a e^{-x^2} dx\right)\left(\int_0^a e^{-y^2} dy\right)$$

$$= \left(\int_0^a e^{-t^2} dt\right)^2$$

Note: $\int e^{-t^2} dt$ is not expressible in terms of elementary functions

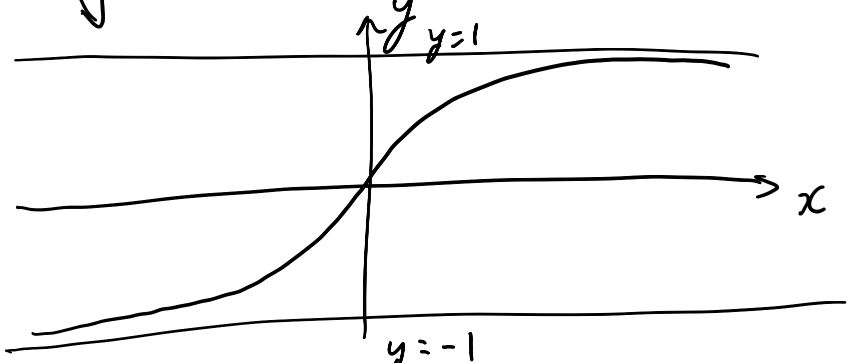
Let $a \to \infty$

$$\Rightarrow \left(\int_0^\infty e^{-t^2} dt\right)^2 = \frac{\pi}{4}$$

$$\Rightarrow \int_0^\infty e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$$

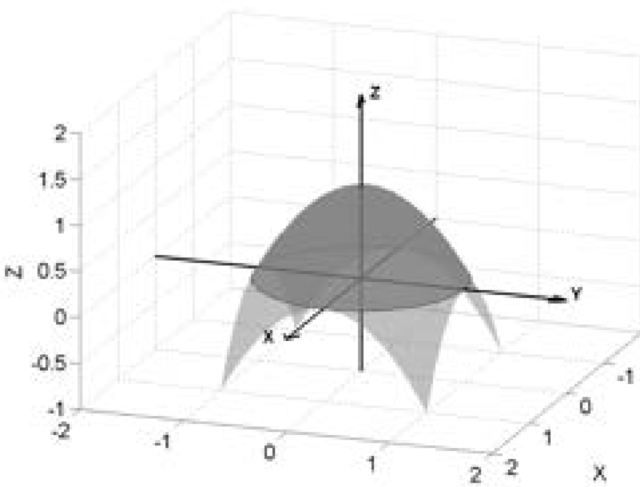
$\longrightarrow$ "error function"

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$



28.2 Example: Find the volume of the solid bounded by the plane $z = 0$ and the paraboloid $\boxed{z = 1 - x^2 - y^2.}$

In $z = 0$, $0 = 1 - x^2 - y^2$
$\Rightarrow x^2 + y^2 = 1$



Find vol. below $z = 1 - x^2 - y^2$ & above circle $x^2 + y^2 = 1$

vol. $= \iint_D (1 - x^2 - y^2)\, dA$

Use polar coordinates!

$D = \{(r, \theta) \mid 0 \le r \le 1,\ 0 \le \theta \le 2\pi\}$

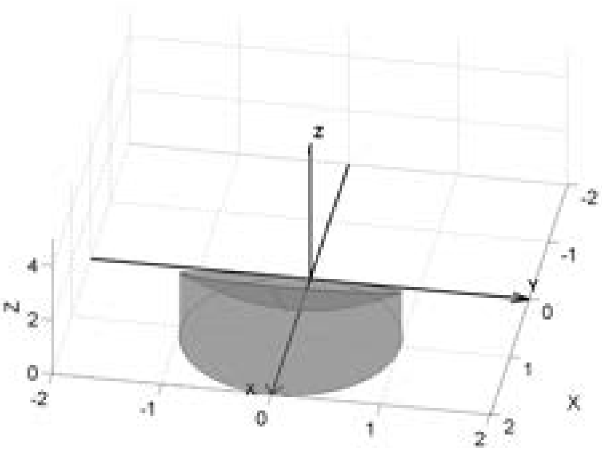
& $\boxed{\text{Jacobian} = r}$

$\Rightarrow$ vol. $= \int_0^1 \int_0^{2\pi} (1 - r^2)\, r \, d\theta\, dr$

$= \ldots = \pi/2$

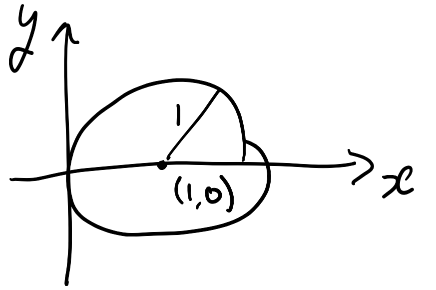
28.3 Find the volume of the solid that lies under the paraboloid $z = x^2 + y^2$ and inside the cylinder $x^2 + y^2 = 2x$, for $z \geq 0$.

i.e. vol. below $z = x^2+y^2$ & above circle $x^2+y^2 = 2x$



$x^2+y^2 = 2x \Rightarrow x^2-2x+1+y^2 = 0+1$

$\Rightarrow (x-1)^2+y^2 = 1$



Use $x = r\cos\theta$, $y = r\sin\theta$?

try it! See Stewart ed. 8 p1054

Instead use "shifted" polar coordinates:

$x-1 = r\cos\theta$, $y = r\sin\theta$ (Jacobian $= r$)

$\Rightarrow x = r\cos\theta + 1$

$\Rightarrow z = x^2+y^2 = (r\cos\theta+1)^2 + r^2\sin^2\theta$

$= r^2\cos^2\theta + 2r\cos\theta + r^2\sin^2\theta + 1$

$= r^2 + 2r\cos\theta + 1$

$D = \{(r,\theta) \mid 0 \leq r \leq 1,\ 0 \leq \theta \leq 2\pi\}$

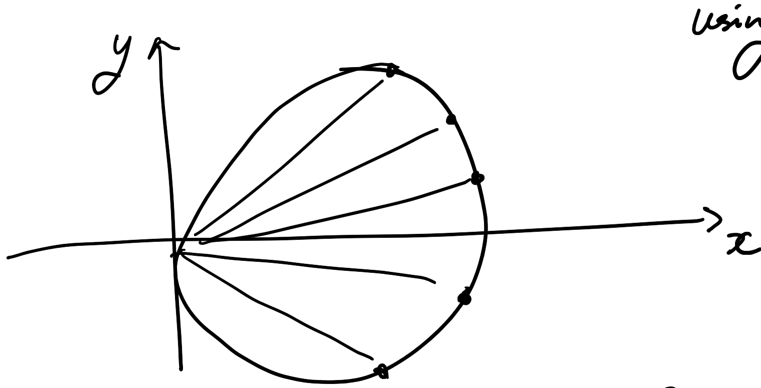
$$\text{Vol.} = \iint_D (x^2+y^2)\,dA$$

$$= \int_0^1 \int_0^{2\pi} (r^2 + 2r\cos\theta + 1)\, r\, d\theta\, dr$$

$$= \left(\int_0^1 r^3\, dr\right)\left(\int_0^{2\pi} d\theta\right) + 2\left(\int_0^1 r^2\, dr\right)\left(\int_0^{2\pi} \cos\theta\, d\theta\right)$$
$$+ \left(\int_0^1 r\, dr\right)\left(\int_0^{2\pi} d\theta\right)$$

$$= \frac{1}{4} \times 2\pi \quad + 0 + \frac{1}{2} \times 2\pi$$

$$= \frac{3\pi}{2}$$



using $x = r\cos\theta$
$y = r\sin\theta$

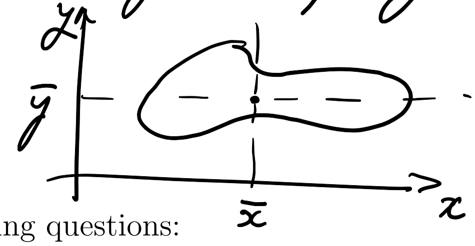
$$x^2 + y^2 = 2x$$
$$r^2 = 2r\cos\theta$$
$$\Rightarrow r = 2\cos\theta$$

$$-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$$

$$0 \le r \le 2\cos\theta$$

Notes.

mass distribution function $\rho(x,y)$

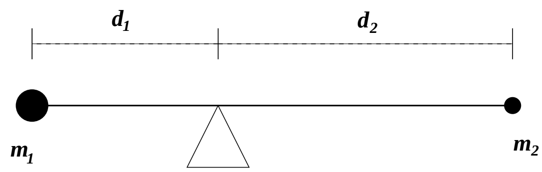


29 Mass, centre of mass and moments

By the end of this section, you should be able to answer the following questions:

- How can we use a double integral to find the mass of a two dimensional object if the density function is known?
- How do we use double integrals to locate the centre of mass of such an object?
- How do we calculate the moments of such an object about the coordinate axes?

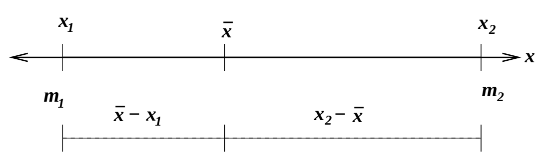
Ultimately we want to find a point P on which a thin plate of any given shape balances horizontally. Such a point is called the centre of mass of the plate.



Consider a rod of negligible mass balanced on a fulcrum. The rod has masses m_1 and m_2 at either end, which are a distance d_1 and d_2 respectively from the fulcrum. Because the rod is balanced, we have (thanks to Archimedes) the relationship

$$\boxed{m_1 d_1 = m_2 d_2.}$$ "law of the lever"

Now suppose the rod lies on the x-axis with m_1 at $x = x_1$, m_2 at $x = x_2$ and the centre of mass at $\overline{x}$.

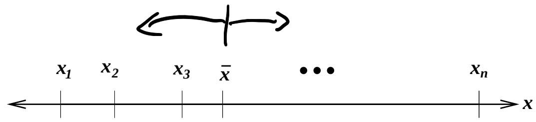


In this case we can write $d_1 = \overline{x} - x_1$ and $d_2 = x_2 - \overline{x}$, so Archimedes' relationship can be expressed

$$m_1(\overline{x} - x_1) = m_2(x_2 - \overline{x}) \;\Rightarrow\; \overline{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}.$$

The numbers $m_1 x_1$ and $m_2 x_2$ are called the *moments* of the masses m_1 and m_2 respectively.

$m_1(\bar{x}-x_1)+m_2(\bar{x}-x_2)+\ldots = \ldots + (x_n-\bar{x})m_n$



In general, a one dimensional system of n "particles" with masses $m_1, \ldots, m_n$ located at $x = x_1, \ldots, x_n$ has its centre of mass located at

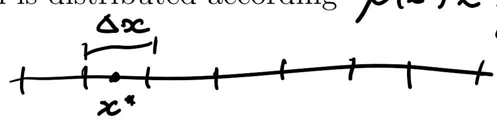
$$\overline{x} = \frac{\displaystyle\sum_{i=1}^{n} m_i x_i}{\displaystyle\sum_{i=1}^{n} m_i} = \frac{M}{m}$$

where $m = \sum m_i$ is the total mass of the system and the sum of the individual moments $M = \sum m_i x_i$ is called the moment of the system (with respect to the origin).

$\rho(x^*) \approx \frac{\Delta m}{\Delta x}$

Now suppose the rod (which has length l) has mass which is distributed according to the (integrable) density function (mass/unit length)

(continuous)



$$\rho(x) = \lim_{\Delta x \to 0} \frac{\Delta m}{\Delta x}.$$

Consider a small strip of width Δx containing the point x^*. The mass of this strip can be approximated by $\boxed{\rho(x^*)\Delta x.}$ Now cut the rod into n strips, and in the same way as above determine (approximately) the mass of each strip. To obtain an approximation for the total mass m of the rod, just add the masses of each n strips:

$$m \approx \sum_{i=1}^{n} \rho(x_i^*)\Delta x_i.$$

To obtain a precise expression for the mass, we take the limit of this sum as $n \to \infty$. In other words,

$$m = \int_0^l \rho(x)dx.$$

We have a similar construction for the moment of the system. Consider the moment of each strip $\approx x_i^* \rho(x_i^*)\Delta x_i$. If we add these, we obtain an approximate expression for the moment of the system:

$$M \approx \sum_{i=1}^{n} x_i^* \rho(x_i^*)\Delta x_i.$$

Taking the limit as $n \to \infty$ we obtain an expression for the moment of the system about the origin:

$$M = \int_0^l x\rho(x)dx.$$

The centre of mass is located at $\overline{x} = M/m$.

Now let's generalize this to two dimensions.

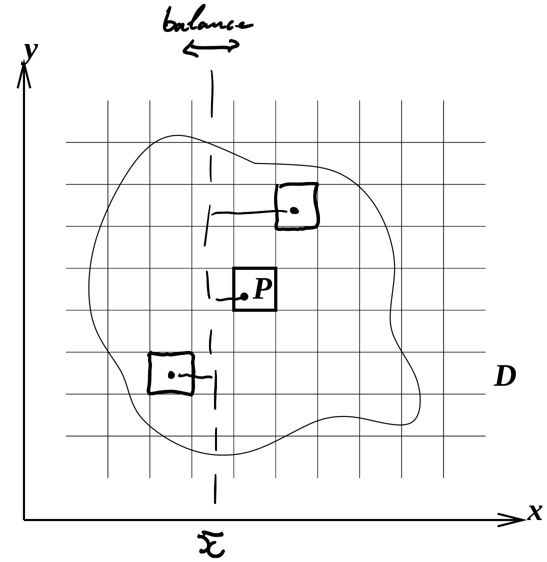
Suppose the lamina occupies a region D in the x-y plane and its density (in units of mass/unit area) is given by an integrable function $\rho(x,y)$. In other words,

2D

(continuous)

$$\rho(x,y) = \lim_{\Delta A \to 0} \frac{\Delta m}{\Delta A},$$

where Δm and ΔA are the mass and area of a small rectangle containing the point (x,y), and the limit is taken as the dimensions of $\Delta A \to 0$.



mass of one grid square $\approx \rho(x^*, y^*)\Delta A$

Figure 26: The point $\boxed{P = (x_i^*, y_j^*)}$ in the rectangle R_{ij}.

To approximate the total mass of the lamina, we partition D into small rectangles (say R_{ij}) and choose a point (x_i^*, y_j^*) inside R_{ij}. The mass of the lamina inside R_{ij} is approximately $\rho(x_i^*, y_j^*)\Delta A_{ij}$, where ΔA_{ij} is the area of R_{ij}. Adding all such masses, we have the approximation

$$m \approx \sum_{i=1}^{m} \sum_{j=1}^{n} \rho(x_i^*, y_j^*)\Delta A_{ij}.$$

If we then take the limit as $m, n \to 0$, we obtain

$$m = \iint_D \rho(x,y)\,dA.$$

x-direction:

$$\sum_{(\text{left})} (\bar{x} - x^*)\,\rho(x^*, y^*)\Delta A \approx \sum_{(\text{right})} (x^* - \bar{x})\,\rho(x^*, y^*)\Delta A$$

$$\Rightarrow \bar{x} \sum_{(\text{all})} \rho(x^*, y^*)\Delta A \approx \sum_{(\text{all})} x^* \rho(x^*, y^*)\Delta A$$

$$\Rightarrow \bar{x} \approx \frac{\sum_{(\text{all})} x^* \rho(x^*, y^*)\Delta A}{\sum_{(\text{all})} \rho(x^*, y^*)\Delta A}$$

In a similar way, we can determine the moment of the lamina about the x-axis to be

$$M_x = \iint_D y\rho(x,y)dA$$

and the moment of the lamina about the y-axis to be

$$M_y = \iint_D x\rho(x,y)dA.$$

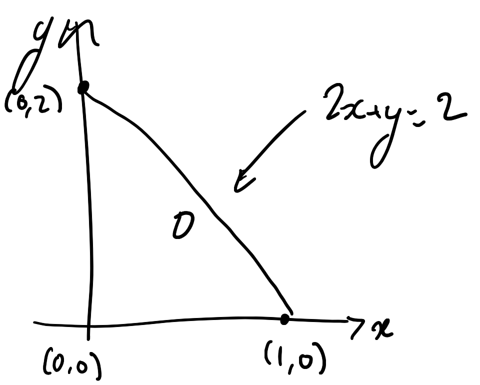
The centre of mass is located at coordinates $(\overline{x}, \overline{y})$, where

$$\overline{x} = \frac{M_y}{m}, \quad \overline{y} = \frac{M_x}{m}.$$

$$\bar{x} = \frac{\iint_D x\rho(x,y)dA}{\iint_D \rho(x,y)dA}$$

$$\bar{y} = \frac{\iint_D y\rho(x,y)dA}{\iint \rho(x,y)dA}$$

29.1 Example: find the centre of mass of a triangular lamina with vertices $(0,0), (1,0)$ and $(0,2)$ with constant density ρ_0.



Need to evaluate three $\iint_D$

$$D = \{(x,y) \mid 0 \le x \le 1,\ 0 \le y \le 2-2x\}$$

$$\text{mass} = \iint_D \rho(x,y)\,dA = \iint_D \rho_0\,dA$$

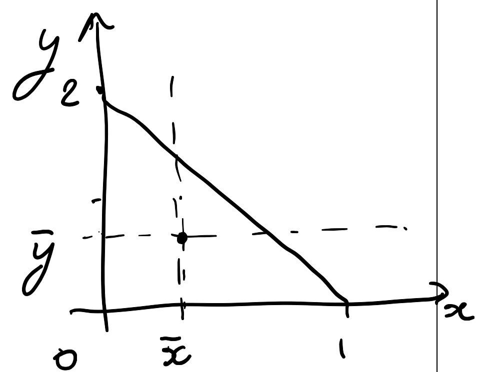
$$= \rho_0 \iint_D dA = \rho_0 \times (\text{area of } D)$$

$$= \rho_0 \times \tfrac{1}{2} \times 2 \times 1 = \underline{\rho_0}$$

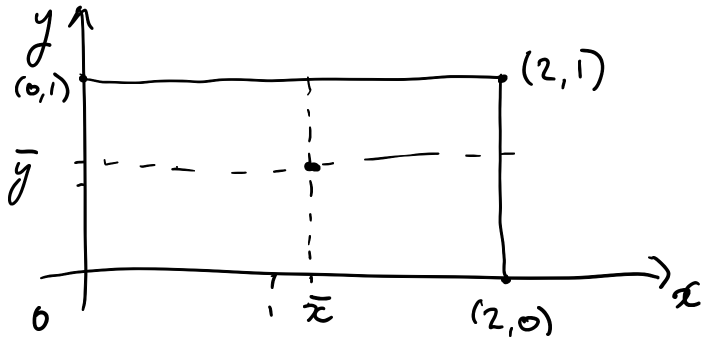
$$\iint_D x\rho(x,y)\,dA = \rho_0\int_0^1\int_0^{2-2x} x\,dy\,dx = \frac{1}{3}\rho_0$$

$$\iint_D y\rho(x,y)\,dA = \rho_0\int_0^1\int_0^{2-2x} y\,dy\,dx = \frac{2}{3}\rho_0$$

$$\Rightarrow \left.\begin{aligned} \bar{x} &= \frac{\frac{1}{3}\rho_0}{\rho_0} = \frac{1}{3} \\ \bar{y} &= \frac{\frac{2}{3}\rho_0}{\rho_0} = \frac{2}{3} \end{aligned}\right\} \Rightarrow$$



29.2 Example: find the centre of mass of a rectangle with vertices $(0,0), (2,0)$, $(2,1)$ and $(0,1)$ with density $\rho(x,y) = 6x + 12y$.



$$\text{mass} = \iint_D \rho(x,y)\,dA = \int_0^1\int_0^2 (6x+12y)\,dx\,dy$$

$$= 6\left(\int_0^1 dy\right)\left(\int_0^2 x\,dx\right) + 12\left(\int_0^1 y\,dy\right)\left(\int_0^2 dx\right)$$

$$= 24$$

$$\iint_D x\rho(x,y)\,dA = \int_0^1\int_0^2 (6x^2+12xy)\,dx\,dy$$

$$= 28$$

$$\Rightarrow \bar{x} = \frac{28}{24} = \frac{7}{6}$$

$$\iint_D y\rho(x,y)\,dA = \int_0^1\left(\int_0^2 (6xy+12y^2)\,dx\right)dy$$

$$= 14$$

$$\Rightarrow \bar{y} = \frac{14}{24} = \frac{7}{12}$$

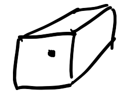
Notes.

CH29: Lamina: 2D object with mass
2D = mass density = mass per unit area
$\to$ density function $\rho(x,y) = \lim_{\Delta A \to 0} \frac{\Delta m}{\Delta A} \to \Delta m \approx \rho(x^*,y^*)\Delta A$

30 Introduction to triple integrals

By the end of this section, you should be able to answer the following questions:

- How do you evaluate a triple integral?
- How do you use a triple integral to find the mass of a solid object with known density?
- How do you change the order of integration in a triple integral?

 (x^*,y^*,z^*)

We can extend the definition of a double integral to a triple integral

(Consider mass density function $f(x,y,z) > 0$ in R) $\to$

$$\iiint_R f(x,y,z)dV, \quad \underline{\underline{\text{def}}} \lim_{\Delta V \to 0} \sum_{\text{(boxes)}} f(x^*,y^*,z^*)\Delta V$$

where R is a region in $\mathbb{R}^3$ and dV is an element of volume.

If R is a region in $\mathbb{R}^3$ specified by

read this in reverse order $\uparrow$

$$\begin{array}{rcl} r(x,y) & \le z \le & s(x,y) \\ p(x) & \le y \le & q(x) \\ a & \le x \le & b \end{array} \tag{14}$$

then

(Fubini's Theorem for $\iiint$ $\to$ see Stewart)

$$\iiint_R f(x,y,z)dV$$

$$= \int_a^b \left\{ \int_{p(x)}^{q(x)} \left[\int_{r(x,y)}^{s(x,y)} f(x,y,z)dz \right] dy \right\} dx.$$

In two dimensions, there are 2 possible orders of integration. In three dimensions, there are 6.

$dx \quad dy \quad dz$

$dx \quad dz \quad dy$

$\boxed{dy \quad dx \quad dz}$ $\to$ $c \le z \le d$, $g_1(z) \le x \le g_2(z)$, $h_1(x,z) \le y \le h_2(x,z)$

$dy \quad dz \quad dx$

$dz \quad dx \quad dy$

$dz \quad dy \quad dx$

30.1 Find the mass of a rectangular block with dimensions $0 \le x \le L$, $0 \le y \le W$ and $0 \le z \le H$ if the density is $\rho = \rho_0 + \alpha xyz$.

in 3D, continuous density function gives mass per unit volume: $\rho(x,y,z) = \lim_{\Delta V \to 0} \frac{\Delta m}{\Delta V}$

mass of one box $\Delta m \approx \rho(x^*, y^*, z^*)\,\Delta V$

(approximate ρ as constant inside ΔV)

$\rightarrow$ total mass $\simeq \sum_{(\text{boxes})} \Delta m = \sum_{(\text{boxes})} \rho(x^*, y^*, z^*)\,\Delta V$

$\lim \Delta V \to 0 \longrightarrow$ mass $= \iiint_R \rho(x,y,z)\,dV$

Example: mass $= \iiint_R (\rho_0 + \alpha xyz)\,dV$

$$= \int_0^H \int_0^W \int_0^L (\rho_0 + \alpha xyz)\,dx\,dy\,dz$$

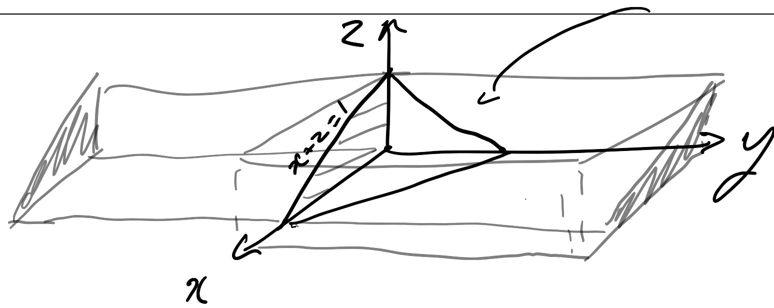
$$= \int_0^H \int_0^W \left[\rho_0 x + \frac{\alpha}{2} x^2 yz\right]_{x=0}^{x=L} dy\,dz$$

$$= \int_0^H \int_0^W \left(\rho_0 L + \frac{\alpha}{2} L^2 yz\right) dy\,dz \quad \text{(double integral}$$

$$= \dots = \underbrace{\rho_0 LWH}_{} + \frac{\alpha}{8} L^2 W^2 H^2$$

constant density term i.e. density & volume

30.2 Evaluate $\iiint\limits_R z\, dV$ over the region R bounded by the surfaces $x = 0$, $y = 0$, $z = 0$ and $x + y + z = 1$.



$$R = \{(x,y,z) \mid 0 \le z \le 1,\ 0 \le x \le 1-z,\ 0 \le y \le 1-x-z\}$$

$$\Rightarrow \iiint\limits_R z\, dV = \int_0^1 \int_0^{1-z} \int_0^{1-x-z} z\, dy\, dx\, dz$$

$$= \int_0^1 \int_0^{1-z} \Big[yz\Big]_{y=0}^{y=1-x-z} dx\, dz$$

$$= \int_0^1 \int_0^{1-z} (z - xz - z^2)\, dx\, dz$$

$$= \dots = 1/24$$

dz dy dx ⟷ dz dx dy ⟷ dx dz dy ⟷ dx dy dz ⟷ dy dx dz ⟷ dy dz dx ⟷ dz dy dx

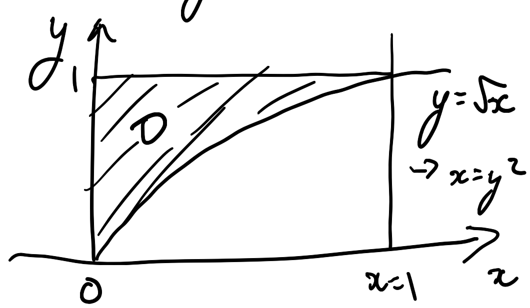
30.3 Changing the order of integration

Express the integral $\int_0^1 \int_{\sqrt{x}}^1 \int_0^{1-y} f(x,y,z)\ dz\ dy\ dx$, in the orders $dz\ dx\ dy$ and $dy\ dz\ dx$.

dz dx dy: Let $g(x,y) = \int_0^{1-y} f(x,y,z)\,dz$

$\Rightarrow I = \int_0^1 \int_{\sqrt{x}}^1 g(x,y)\,dy\,dx$

$D = \{(x,y) \mid 0 \le x \le 1,\ \sqrt{x} \le y \le 1\}$



$= \{(x,y) \mid 0 \le y \le 1,\ 0 \le x \le y^2\}$

$\Rightarrow \int_0^1 \int_0^{y^2} \int_0^{xy} f(x,y,z)\,dz\,dx\,dy$

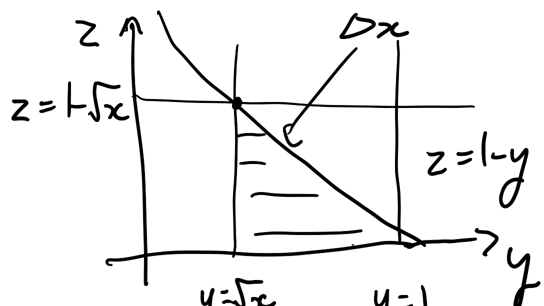
---

dy dz dx: Let $h(x) = \int_{\sqrt{x}}^1 \int_0^{1-y} f(x,y,z)\,dz\,dy$

for $0 \le x \le 1$ $\Rightarrow I = \int_0^1 h(x)\,dx$

Fix $x \in (0,1) \Rightarrow D_x = \{(y,z) \mid \sqrt{x} \le y \le 1,\ 0 \le z \le 1-y\}$

(i.e. treat x as constant)



$\Rightarrow D_x = \{(y,z) \mid 0 \le z \le 1-\sqrt{x},\ \sqrt{x} \le y \le 1-z\}$

$\Rightarrow h(x) = \int_0^{1-\sqrt{x}} \int_{\sqrt{x}}^{1-z} f(x,y,z)\,dy\,dz$

$\Rightarrow I = \int_0^1 \int_0^{1-\sqrt{x}} \int_{\sqrt{x}}^{1-z} f(x,y,z)\,dy\,dz\,dx$

Notes.

Tennille is one

MESSY

bitch

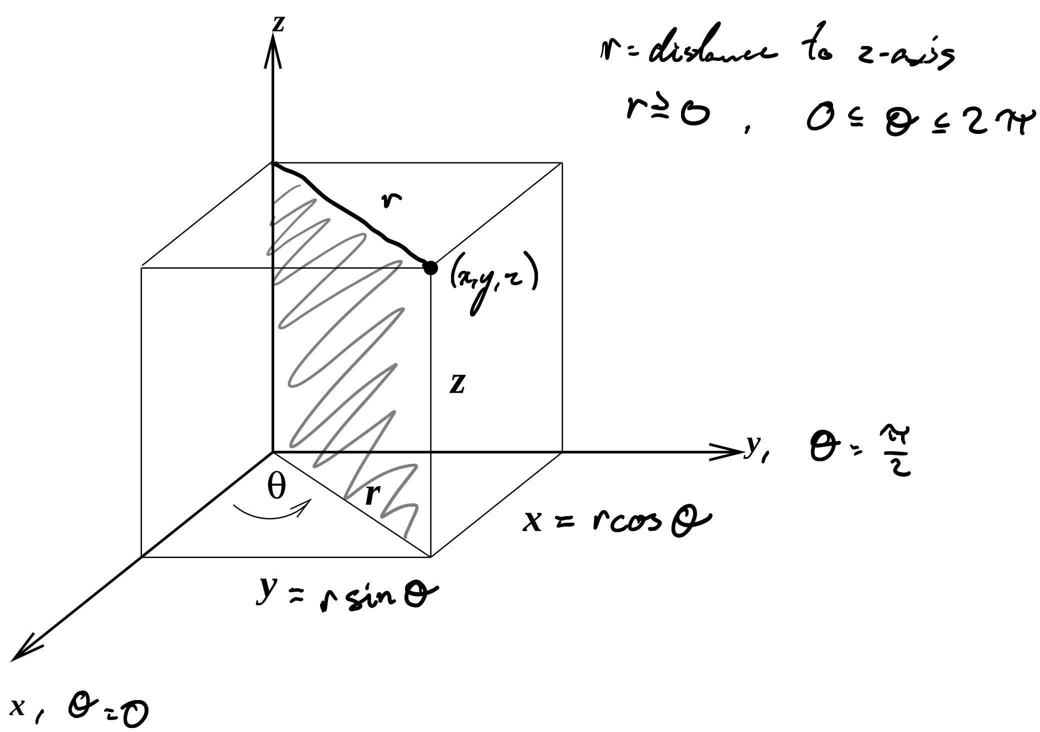
31 Cylindrical coordinates

By the end of this section, you should be able to answer the following questions:

- What is the relationship between rectangular coordinates and cylindrical coordinates?
- How do you transform a triple integral in rectangular coordinates into one in terms of cylindrical coordinates?
- What is the Jacobian of the transformation?

Sometimes it is useful to use cylindrical coordinates in order to simplify the integral. This involves the transformation

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z. \tag{15}$$



We now aim to calculate a small element of volume of a cylindrical shell. This will then show how in a triple integral we can transform from rectangular coordinates to cylindrical coordinates by substituting the transformation (15) and by making the change

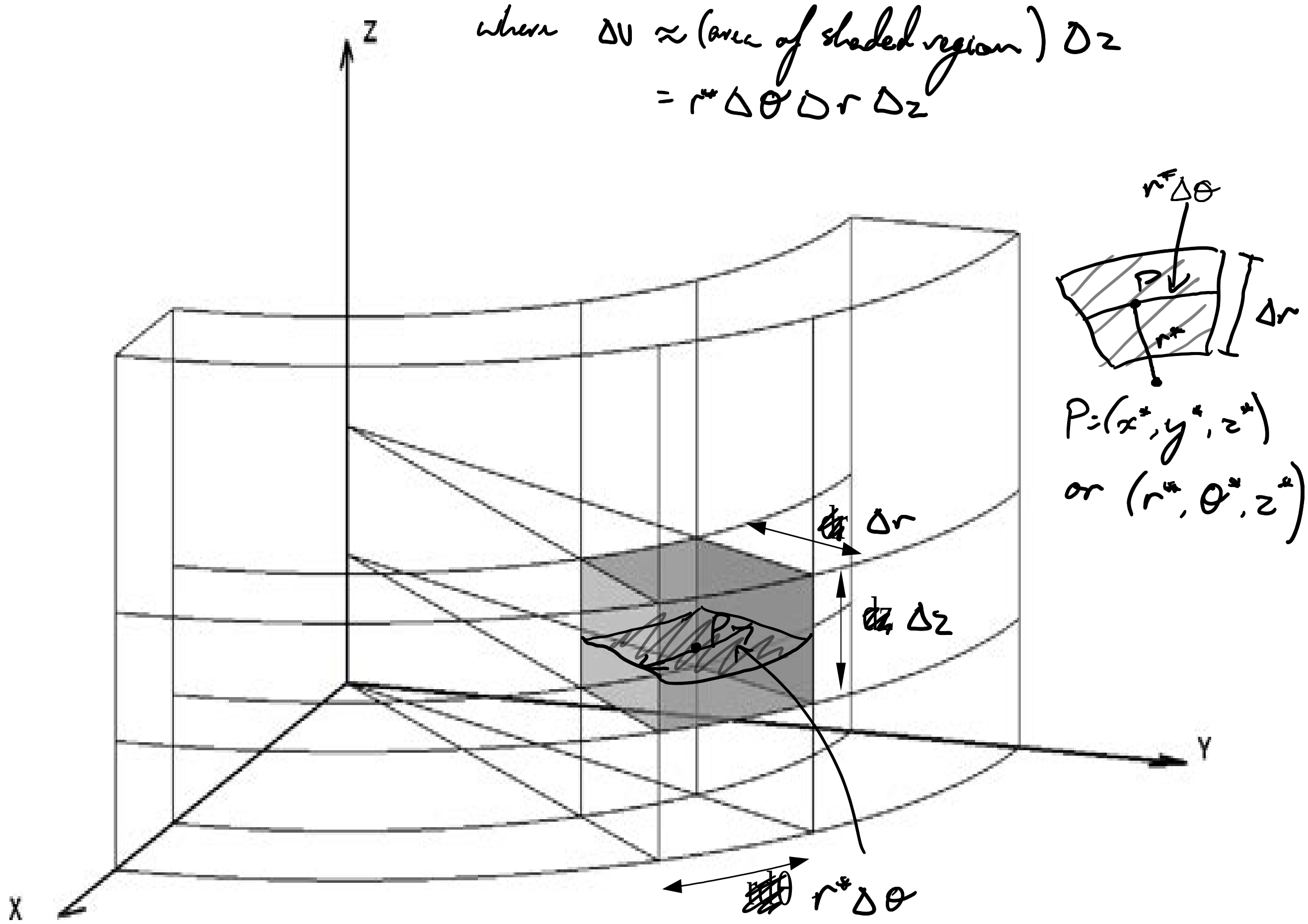
$$dx\ dy\ dz \quad \longrightarrow \quad r\ dr\ d\theta\ dz.$$ (Vol. cyl. box = ΔV)

Consider the following diagram.

Continuous mass density function $f(x,y,z)$

$$\text{mass} \approx \sum_{(\text{boxes})} f(p)\,\Delta V$$

where $\Delta V \approx (\text{area of shaded region})\,\Delta z$
$= r^* \Delta\theta\, \Delta r\, \Delta z$



"Jacobian"

"$dx\ dy\ dz$" $\longmapsto$ "$r\ d\theta\ dr\ dz$"

The important result is that the triple integral in rectangular coordinates transforms as follows:

$\Delta V \to 0$

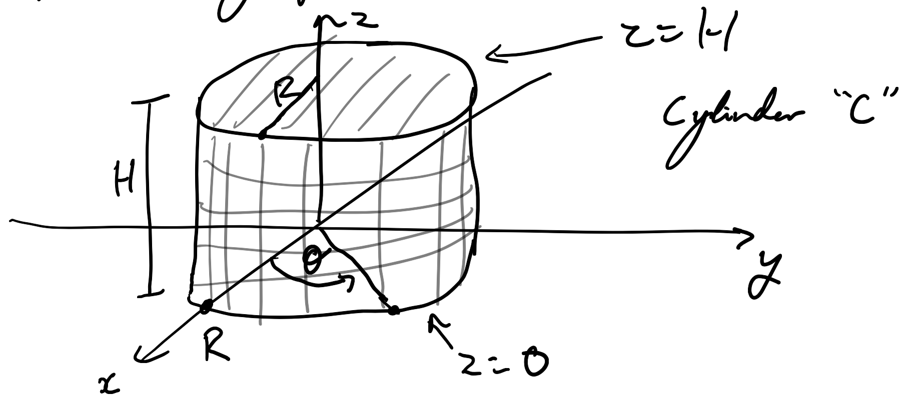
$$\boxed{\iiint_R f(x,y,z)\ dx\ dy\ dz = \iiint_C f(r\cos\theta, r\sin\theta, z)\ r\ dr\ d\theta\ dz.}$$

Generally in 3D

$$\left.\begin{array}{l} x = x(u,v,w) \\ y = y(u,v,w) \\ z = z(u,v,w) \end{array}\right\}\ \text{Jacobian} = \left|\det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}\right|$$

31.1 A simple example: Find the volume of a cylinder of radius R and height H. (Ans. $\pi R^2 H$)

Assume cylinder is centred about z-axis and sits atop the x-y plane.



$$\text{volume} = \lim_{\Delta V \to 0} \sum_{(\text{boxes})} \Delta V = \iiint_C dV \quad \left(= \iiint 1 \, dV \right)$$

$$C = \left\{ (r, \theta, z) \,\middle|\, 0 \le r \le R,\ 0 \le \theta \le 2\pi,\ 0 \le z \le H \right\}$$

$$\to \text{Vol.} = \int_0^R \int_0^{2\pi} \int_0^H r \, dz \, d\theta \, dr$$

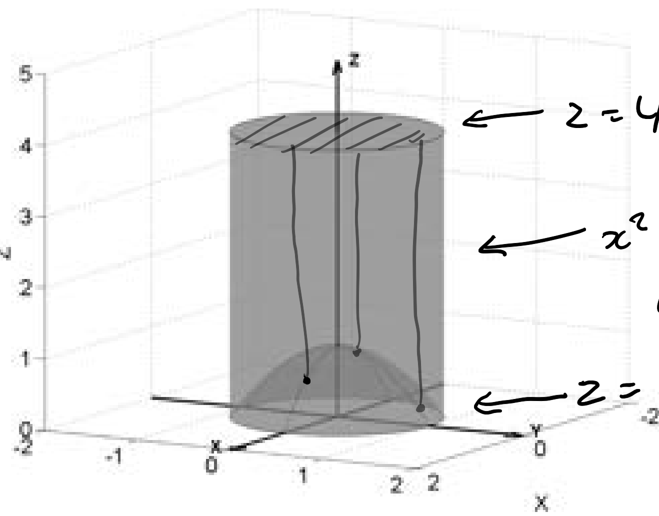
$$= \left(\int_0^R r \, dr \right) \left(\int_0^{2\pi} 1 \, d\theta \right) \left(\int_0^H 1 \, dz \right)$$

$$= \frac{1}{2} R^2 \times 2\pi \times H = \pi R^2 H \qquad \checkmark$$

31.2 Find the mass of the solid defined by the region contained within the cylinder $x^2 + y^2 = 1$ below the plane $z = 4$ and above the paraboloid $z = 1 - x^2 - y^2$. The density at any given point in the region is proportional to the distance from the $\underbrace{\text{axis of the cylinder.}}_{\text{z-axis}}$ $\Rightarrow$ use cylindrical coords

$$\left.\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}\right\} (r, \theta, z)$$

Region R



$z = 4$

$x^2 + y^2 = 1$

$z = 1 - x^2 - y^2$

Density $\rho \propto r$

Take $\rho = kr$ ($k > 0$, constant)

$$\text{Mass} = \iiint_R \rho \, dV$$

$\underline{x^2 + y^2 = 1} \Rightarrow r^2\cos^2\theta + r^2\sin^2\theta = 1 \Rightarrow r^2 = 1 \Rightarrow r = 1$

$\underline{z = 1 - x^2 - y^2} \Rightarrow z = 1 - (x^2 + y^2) = 1 - r^2$

$\underline{\text{Jacobian} = r}$

$R = \{(r, \theta, z) \mid 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi, 1 - r^2 \leq z \leq 4\}$

$$\rightarrow \text{mass} = \int_0^1 \int_{1-r^2}^4 \int_0^{2\pi} \underbrace{kr}_{\text{density}} \cdot \underbrace{r}_{\text{Jacobian}} \, d\theta \, dz \, dr$$

$$= k\left(\int_0^{2\pi} 1 \, d\theta\right) \cdot \left(\int_0^1 \int_{1-r^2}^4 r^2 \, dz \, dr\right)$$

$$= k \times 2\pi \times \int_0^1 \left[r^2 z\right]_{z=1-r^2}^{z=4} dr$$

$$= 2\pi k \times \int_0^1 \left(4r^2 - \left(r^2(1-r^2)\right)\right) dr$$

$$= \ldots = \frac{12\pi k}{5}$$

Notes.

32 Spherical coordinates

By the end of this section, you should be able to answer the following questions:

- What is the relationship between rectangular coordinates and spherical coordinates?
- How do you transform a triple integral in rectangular coordinates into one in terms of spherical coordinates?
- What is the Jacobian of the transformation?

Sometimes it is useful to use spherical coordinates in order to simplify the integral. This involves the transformation

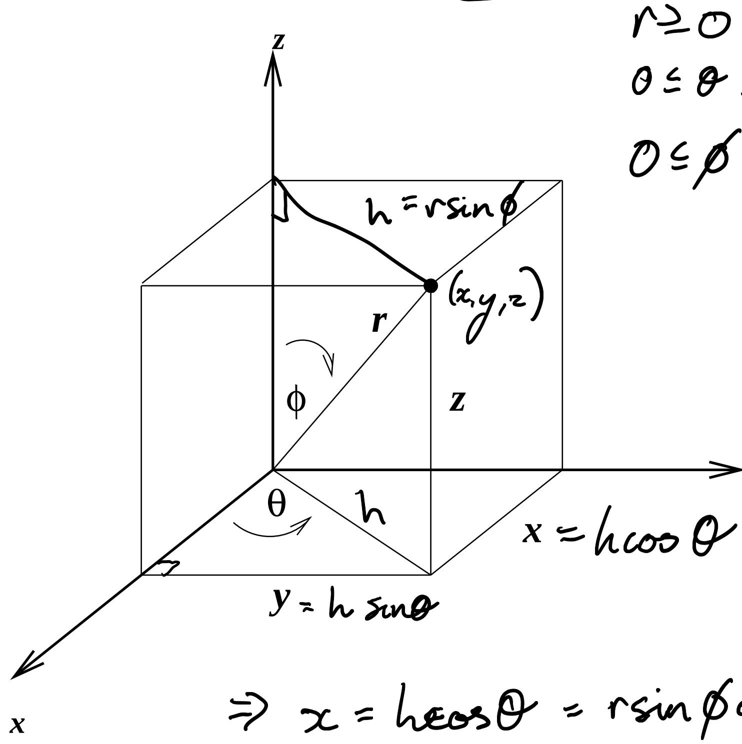
$$x = r\cos\theta\sin\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\phi. \tag{16}$$

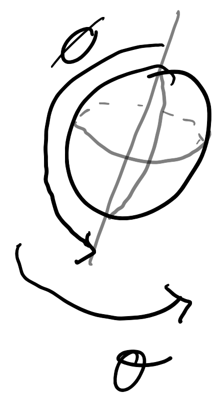
In this case θ is longitude, ϕ is co-latitude, and r the distance from the origin.

$r \geq 0$

$0 \leq \theta \leq 2\pi$

$0 \leq \phi \leq \pi$





$\Rightarrow x = h\cos\theta = r\sin\phi\cos\theta$

$y = h\sin\theta = r\sin\phi\sin\theta$

& $z = r\cos\phi$

We now aim to calculate a small element of volume of a spherical shell. This will then show how in a triple integral we can transform from rectangular coordinates to spherical coordinates by substituting the transformation (16) and by making the change

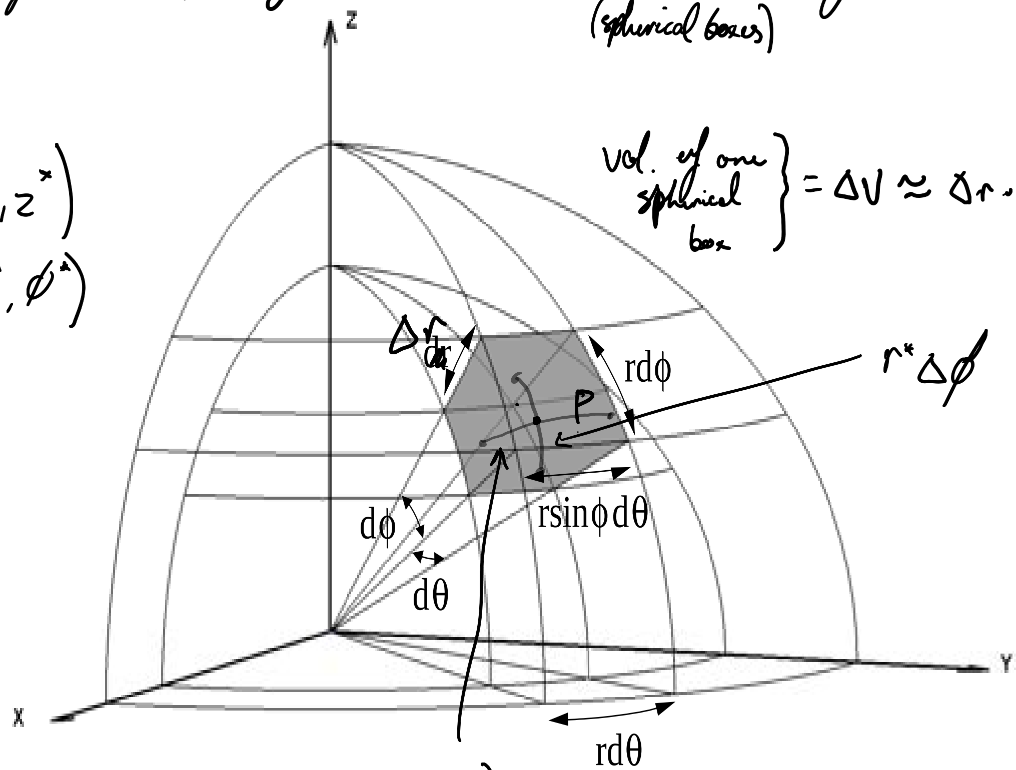
$$dx\ dy\ dz \quad \longrightarrow \quad r^2 \sin\phi\ dr\ d\theta\ d\phi.$$

Consider the following diagram.

Continuous mass density function $\rho(x,y,z) \to$ mass $\approx \sum\limits_{\text{(spherical boxes)}} \rho(x^*,y^*,z^*)\,\Delta V$

$P=(x^*,y^*,z^*)$ or (r^*,θ^*,ϕ^*)

vol. of one spherical box $\Big\} = \Delta V \approx \Delta r \cdot r^*\Delta\phi \cdot r^*\sin\phi^*\Delta\theta$



$(r^*\sin\phi^*)\Delta\theta$

$\Rightarrow$ mass $= \sum\limits_{\text{(spherical boxes)}} \rho(r^*\cos\theta^*\sin\phi^*, r^*\sin\phi^*\sin\theta^*, r^*\cos\phi^*)\,(r^*)^2\sin\phi^*\Delta r\,\Delta\theta\,\Delta\phi$

The important result is that the triple integral in rectangular coordinates transforms as follows:

$\xrightarrow{\lim_{\Delta V\to 0}}$

$$\boxed{\begin{aligned}&\iiint\limits_R f(x,y,z)\ dx\ dy\ dz\\ &= \iiint\limits_S f(r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)\ r^2\sin\phi\ dr\ d\theta\ d\phi.\end{aligned}}$$

$dV \to$ "$dx\ dy\ dz$" $\to$ "$\boxed{r^2\sin\phi^*}\,dr\ d\theta\ d\phi$"

"Jacobian"

32.1 A simple example: Find the volume of a sphere of radius R.

Assume sphere, S, centred at $(0,0,0) = (x,y,z)$

Volume $= \iiint_S dV$

Use spherical coordinates (r, θ, ϕ)

$$S = \{(r,\theta,\phi) \mid 0 \le r \le R,\ 0 \le \theta \le 2\pi,\ 0 \le \phi \le \pi\}$$

$$\Rightarrow \iiint_S dV = \int_0^R \int_0^{2\pi} \int_0^{\pi} r^2 \sin\phi \, d\phi \, d\theta \, dr$$

$$= \left(\int_0^R r^2 \, dr\right)\left(\int_0^{2\pi} d\theta\right)\left(\int_0^{\pi} \sin\phi \, d\phi\right)$$

$$= \frac{1}{3} R^3 \times 2\pi \times 2$$

$$= \frac{4}{3} \pi R^3$$

32.2 Find the mass of a sphere of radius R whose density is given by $\rho(x, y, z) = e^{-(x^2+y^2+z^2)^{1/2}}$.

Problem with wording of the question.
Where is the sphere located?
Take the sphere to be centred at $(x,y,z)=(0,0,0)$

$$\text{Mass} = \iiint_S \rho \, dV$$

Use spherical coords
$$x = r\cos\theta\sin\phi$$
$$y = r\sin\theta\sin\phi$$
$$z = r\cos\phi$$

$$\Rightarrow x^2+y^2+z^2 = r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta\sin^2\phi + r^2\cos^2\phi$$
$$= r^2\underbrace{\left(\sin^2\phi\underbrace{(\cos^2\theta+\sin^2\theta)}_{=1} + \cos^2\phi\right)}_{=1}$$

$$\Rightarrow \rho = e^{-\overset{=r^2}{r}}$$

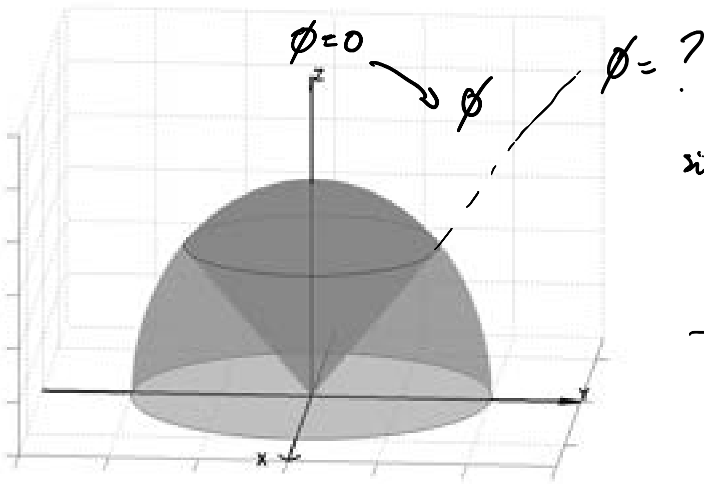
$$\Rightarrow \text{mass} = \int_0^R\int_0^{2\pi}\int_0^{\pi} \underbrace{e^{-r}}_{\rho} \underbrace{r^2\sin\phi}_{\text{Jac}} \, d\phi \, d\theta \, dr$$

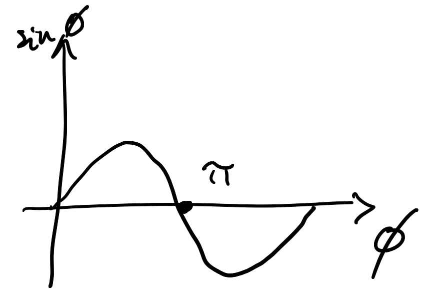
$$= \underbrace{\left(\int_0^R r^2 e^{-r} \, dr\right)}\left(\int_0^{2\pi} d\theta\right)\left(\int_0^{\pi}\sin\phi \, d\phi\right)$$

(use integration by parts twice)

$$= \left(2 - \frac{R^2+2R+2}{e^R}\right) \times 2\pi \times 2$$

32.3 **Find the volume of the "ice cream cone" R between a sphere of radius a (centred at the origin) and the cone $z = \sqrt{x^2 + y^2}$.**



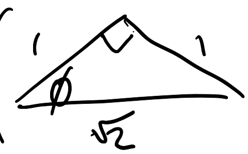


Use of spherical coords $\quad 0 \le r \le a\,,\quad 0 \le \theta \le 2\pi$

Cone $\quad z = \sqrt{x^2+y^2}$

$\Rightarrow r\cos\phi = \sqrt{r^2\sin^2\phi} = r\sin\phi$

$(r = 0 \text{ origin}) \Rightarrow \cos\phi = \sin\phi$ or $\tan\phi = 1$

$\Bigg($  $\Rightarrow \phi = \dfrac{\pi}{4}\Bigg) \Rightarrow 0 \le \phi \le \dfrac{\pi}{4}$

$$\text{volume} = \iint_R dV$$

$$= \int_0^a \int_0^{2\pi} \int_0^{\pi/4} r^2 \sin\phi \, d\phi \, d\theta \, dr$$

$$= \ldots = \frac{\pi}{3} a^3 \left(2 - \sqrt{2}\right)$$

Notes.

33 Moments of inertia (second moments)

By the end of this section, you should be able to answer the following questions:

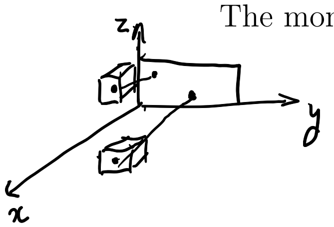
- How do you locate the centre of mass of a solid object using a triple integral?
- How do you calculate the moments of inertia about the three coordinate axes?

The moment of inertia of a particle of mass m about an axis (x, y, or z) is defined to be mr^2 where r is the distance from the particle to the axis.

It is sometimes referred to as rotational inertia and can be thought of as the rotational analogue of mass for linear motion. For example, linear kinetic energy can be expressed as $\frac{1}{2}mv^2$, and the rotational kinetic energy as $\frac{1}{2}I\omega^2$. Linear momentum is determined by the formula $p = mv$, while angular momentum is given by $L = I\omega$. In these examples, I is the moment of inertia and ω the angular velocity.

As we have seen from previous examples, the mass of a solid with density $\rho(x, y, z)$ occupying a region R in $\mathbb{R}^3$ is given by

$$m = \iiint\limits_R \rho(x, y, z)dV.$$



The moments about each of the three coordinate planes are

$$M_{yz} = \iiint\limits_R x\rho(x, y, z)dV, \quad M_{xz} = \iiint\limits_R y\rho(x, y, z)dV,$$

$$M_{xy} = \iiint\limits_R z\rho(x, y, z)dV$$

i.e. $\bar{x} = \dfrac{\iiint_R x\rho \, dV}{\iiint_R \rho \, dV}$

$\bar{y} = \ldots$

$\bar{z} = \ldots$

The centre of mass is then located at the point $(\overline{x}, \overline{y}, \overline{z})$ where

$$\overline{x} = \frac{M_{yz}}{m}, \ \overline{y} = \frac{M_{xz}}{m}, \ \overline{z} = \frac{M_{xy}}{m}.$$

The moments of inertia about each of the three coordinate axes work out to be

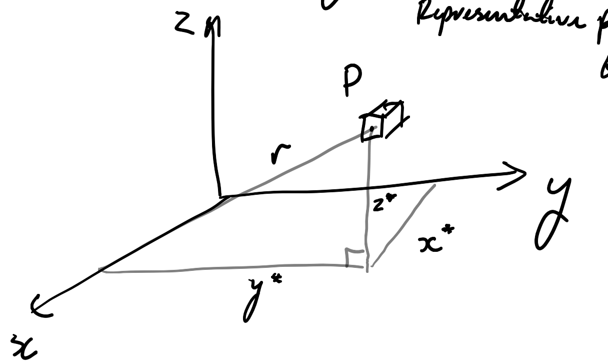
$$I_x = \iiint\limits_R (y^2 + z^2)\rho(x, y, z)dV,$$

$$I_y = \iiint\limits_R (x^2 + z^2)\rho(x, y, z)dV,$$

$$I_z = \iiint\limits_R (x^2 + y^2)\rho(x, y, z)dV.$$

33.1 Derive the integral formula for I_x

Consider a solid occupying a region R in $\mathbb{R}^3$ with continuous density function $\rho(x,y,z)$



$$\begin{Bmatrix}\text{mass of}\\ \text{one box}\end{Bmatrix} = m \simeq \rho(x^*,y^*,z^*)\,\Delta V$$

$$\begin{Bmatrix}\text{Moment of}\\ \text{inertia of box}\\ \text{about } x\text{-axis}\end{Bmatrix} = m\cdot r^2 \simeq \rho(x^*,y^*,z^*)\Delta V\left((y^*)^2+(z^*)^2\right)$$

$$\left(\text{Pythagoras} \Rightarrow r^2 = (y^*)^2+(z^*)^2\right)$$

$$\Rightarrow \begin{Bmatrix}\text{Moment of inertia}\\ \text{of entire region}\\ \text{about } x\text{-axis}\end{Bmatrix} \approx \sum_{(\text{boxes})}\left((y^*)^2+(z^*)^2\right)\rho(x^*,y^*,z^*)\,\Delta V$$

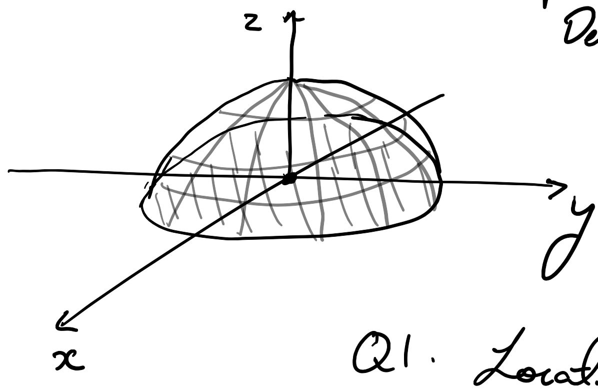
Take limit as $\Delta V \to 0$

$$\Rightarrow \quad I_x = \iiint_R (y^2+z^2)\,\rho(x,y,z)\,dV$$

33.2 Example: locate the centre of mass of a solid hemisphere of radius a with density proportional to the distance from the centre of the base. Find its moment of inertia about the z-axis. ← Problem with wording

(Where is z-axis relative to the hemisphere?)

- Place centre of sphere at $(x,y,z) = (0,0,0)$ and take $z \geq 0$. Use spherical coordinates



Density $\propto r$ (distance to $(0,0,0) \cdot (x,y,z)$)

$\Rightarrow \rho = kr$ (constant $k > 0$)

Q1. Locate $(\bar{x}, \bar{y}, \bar{z})$

$(x = r\cos\theta\sin\phi,\ y = r\sin\theta\sin\phi,\ z = r\cos\phi)$

By symmetry of region and of ρ,

$$\bar{x} = \bar{y} = 0 \quad \& \quad \bar{z} = \frac{\iiint_H z\rho\, dV}{\iiint_H \rho\, dV}$$

$$(\text{mass} =) \iiint_H \rho\, dV = \int_0^a \int_0^{2\pi} \int_0^{\pi/2} kr \cdot r^2 \sin\phi\, d\phi\, d\theta\, dr$$

$$= k\left(\int_0^a r^3\, dr\right)\left(\int_0^{2\pi} 1\, d\theta\right)\left(\int_0^{\pi/2} \sin\phi\, d\phi\right)$$

$$= \ldots = \frac{\pi k a^4}{2}$$

$$\iiint_H z\rho\, dV = \int_0^a \int_0^{2\pi} \int_0^{\pi/2} \underbrace{kr}_{\text{density}} \cdot \underbrace{r\cos\phi}_{z} \cdot \underbrace{r^2\sin\phi}_{\text{Jacobian}} \, d\phi\, d\theta\, dr$$

$$= k\left(\int_0^a r^4\, dr\right)\left(\int_0^{2\pi} 1\, d\theta\right)\left(\int_0^{\frac{\pi}{2}} \cos\phi \sin\phi\, d\phi\right)$$

$$= k \times \frac{1}{5}a^5 \times 2\pi \times \left[\frac{1}{2}\sin^2\phi\right]_0^{\pi/2} = \frac{\pi k a^5}{5}$$

$$\Rightarrow \bar{z} = \frac{\left(\frac{\pi k a^5}{5}\right)}{\left(\frac{\pi k a^4}{2}\right)} = \frac{2}{5}a$$

Q2. $$I_z = \iiint_H (x^2+y^2)\rho\, dV$$

$$x^2+y^2 = r^2\sin^2\phi \qquad (\text{Note: } \sin^2\phi = 1-\cos^2\phi)$$

$$\Rightarrow I_z = \int_0^a \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \underbrace{kr}_{\text{density}} \cdot \underbrace{r^2\sin^2\phi}_{x^2+y^2} \cdot \underbrace{r^2\sin\phi}_{\text{Jacobian}} \, d\phi\, d\theta\, dr$$

$$= k\left(\int_0^a r^5\, dr\right)\left(\int_0^{2\pi} 1\, d\theta\right)\underbrace{\left(\int_0^{\frac{\pi}{2}} (1-\cos^2\phi)\sin\phi\, d\phi\right)}_{\text{sol } u = -\cos\phi \rightarrow \frac{du}{d\phi} = \sin\phi}$$

$$= k \times \frac{1}{6}a^6 \times 2\pi \times \int_{-1}^0 (1-u^2)\, du$$

$$= \dots = \frac{2\pi k a^6}{a}$$

Notes.

34 Conservative vector fields

By the end of this section, you should be able to answer the following questions:

- What is meant by a conservative vector field and a corresponding potential function?
- Given a potential function, how do you determine the corresponding conservative vector field?
- Given a conservative vector field, how do you determine a corresponding potential function?

34.1 Vector fields

In what follows, the notation is always

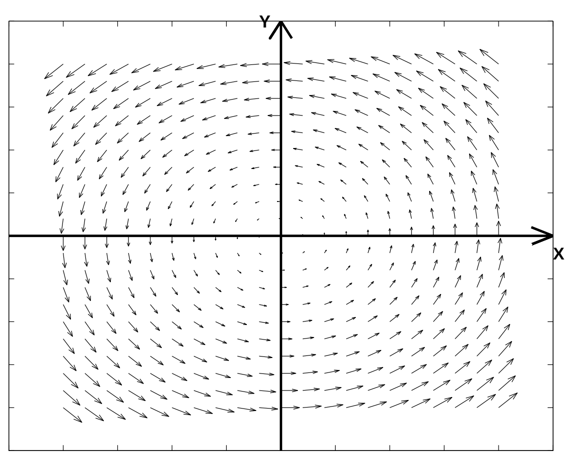
$$\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} \text{ or } \boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}.$$ or $\underline{r}$

A vector field in the x-y plane is a vector function of 2 variables

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{r}) = \boldsymbol{F}(x,y) &= (F_1(x,y), F_2(x,y)) \\ &= F_1(x,y)\boldsymbol{i} + F_2(x,y)\boldsymbol{j}. \end{aligned}$$

That is, associated to a point (x, y) is the vector $\boldsymbol{F}(\boldsymbol{r})$.

34.1.1 Example: $\boldsymbol{F}(\boldsymbol{r}) = (-y, x) = -y\boldsymbol{i} + x\boldsymbol{j}$.

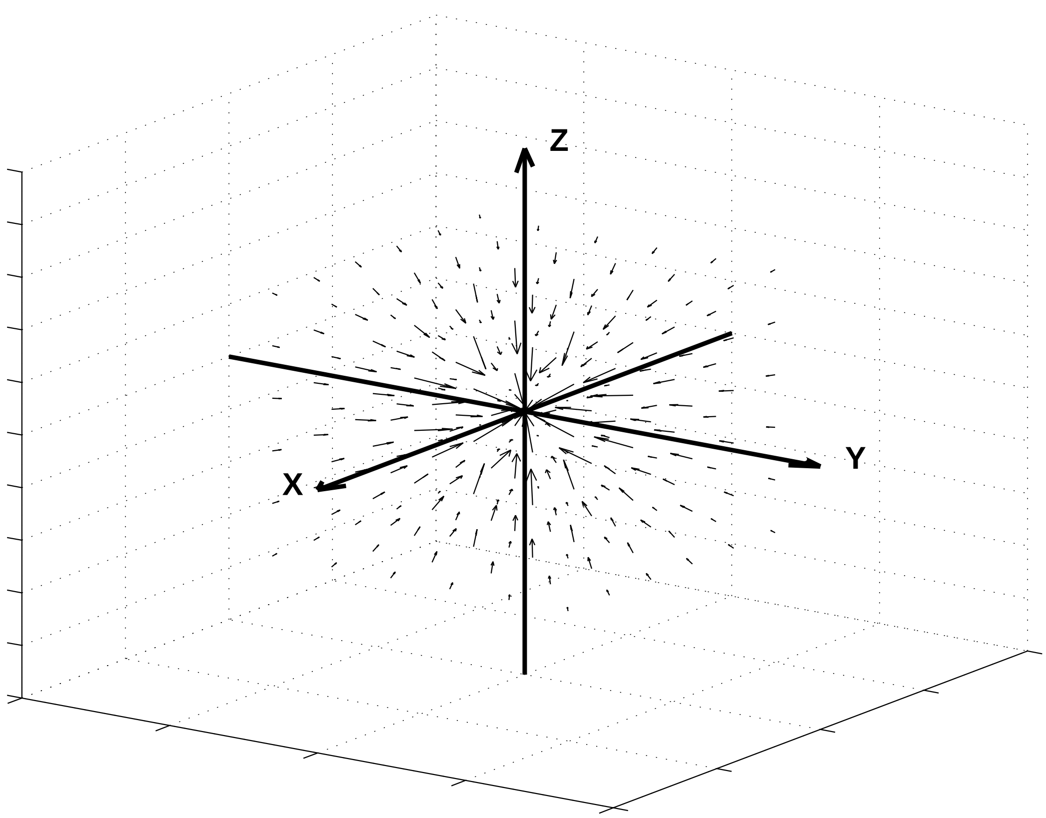


Similarly a vector field in 3-D is a vector function of 3 variables

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{r}) &= \boldsymbol{F}(x,y,z)\\ &= (F_1(x,y,z), F_2(x,y,z), F_3(x,y,z))\\ &= F_1(x,y,z)\boldsymbol{i} + F_2(x,y,z)\boldsymbol{j} + F_3(x,y,z)\boldsymbol{k}\end{aligned}$$

34.1.2 Example: Newtonian gravitational field

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{r}) &= -\frac{mMG}{||\boldsymbol{r}||^3}\boldsymbol{r} = \boldsymbol{F}(x,y,z)\\ &= \frac{-mMGx}{(x^2+y^2+z^2)^{3/2}}\boldsymbol{i} + \frac{-mMGy}{(x^2+y^2+z^2)^{3/2}}\boldsymbol{j}\\ &\quad + \frac{-mMGz}{(x^2+y^2+z^2)^{3/2}}\boldsymbol{k}\end{aligned}$$



34.2 Gradient of a scalar field, conservative vector fields

Recall for a differentiable scalar function $f(x, y)$ in two dimensions, we define

$$\text{grad} f = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j}. \quad = \underline{\nabla} f \qquad \left(2D : \underline{\nabla} = \underline{i}\frac{\partial}{\partial x} + \underline{j}\frac{\partial}{\partial y}\right)$$

For a differentiable scalar function $f(x, y, z)$ in three dimensions, we define

$$\text{grad} f = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j} + \frac{\partial f}{\partial z}\boldsymbol{k}.$$

* Directional derivative

$D_{\underline{\hat{u}}}(f(x,y)) = (\underline{\nabla} f) \cdot \underline{\hat{u}}, \quad \|\underline{\hat{u}}\| = 1$

* Finding critical points of $f(x,y)$

Alternatively we define the differential operator

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

so $\text{grad} f = \nabla f$.

34.2.1 Example: find the gradient of $f(x, y, z) = x^2y^3z^4$.

$$\underline{\nabla} f = \underline{i}\frac{\partial}{\partial x}(x^2y^3z^4) + \underline{j}\frac{\partial}{\partial y}(x^2y^3z^4) + \underline{k}\frac{\partial}{\partial z}(x^2y^3z^4)$$

$$= 2xy^3z^4\underline{i} + 3x^2y^2z^4\underline{j} + 4x^2y^3z^3\underline{k}$$

e.g. at point $(1,1,1)$

$$\underline{\nabla} f(1,1,1) = 2\underline{i} + 3\underline{j} + 4\underline{k}$$

gives the direction of maximum increase of f

& $\|\underline{\nabla} f\|_{(1,1,1)} = \sqrt{4+9+16} = \sqrt{29}$

gives the maximum rate of increase of f

Note ∇f is a vector. It's length and direction are independent of the choice of coordinates. ∇f (evaluated at a given point P) is in the direction of maximum increase of f at P.

You may see the scalar function f referred to as a scalar field. If a vector field $\boldsymbol{v}$ and a scalar field f are related by $\boldsymbol{v} = \nabla f$, we call f a potential function and $\boldsymbol{v}$ a *conservative vector field.*

34.2.2 Verify that the Newtonian gravitational field is conservative with potential function $f(x, y, z) = \dfrac{mMG}{\sqrt{x^2 + y^2 + z^2}}$. $= mMG(x^2+y^2+z^2)^{-1/2}$

$$\frac{\partial f}{\partial x} = -\frac{1}{2} mMG \cdot 2x(x^2+y^2+z^2)^{-3/2}$$

$$\frac{\partial f}{\partial y} = -\frac{1}{2} mMG \cdot 2y(x^2+y^2+z^2)^{-3/2}$$

$$\frac{\partial f}{\partial z} = -\frac{1}{2} mMG \cdot 2z(x^2+y^2+z^2)^{-3/2}$$

$$\Rightarrow \underline{\nabla} f = \frac{-mMG}{(x^2+y^2+z^2)^{3/2}} \left(x\underline{i} + y\underline{j} + z\underline{k}\right)$$

<u>Note</u>: Not all vector fields are conservative

e.g. $\underline{F}(x,y) = -y\underline{i} + x\underline{j}$ is NOT conservative

Given a conservative vector field, how can we determine a corresponding potential function? The next example outlines this procedure.

34.2.3 The vector field $\boldsymbol{F}(x,y) = (3+2xy)\boldsymbol{i} + (x^2-3y^2)\boldsymbol{j}$ is conservative. Find a corresponding potential function.

$\Rightarrow \exists\ f(x,y)$ such that $\underline{F} = \nabla f = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j}$

i.e. Find $f(x,y)$

$\Rightarrow \frac{\partial f}{\partial x} \overset{(set)}{=} 3 + 2xy$ ($\underline{i}$ component of $\underline{F}$)

$\Rightarrow f(x,y) = 3x + x^2 y + \underbrace{g(y)}$

Similar to a constant of integration

$\Rightarrow \frac{\partial f}{\partial y} = 0 + x^2 + g'(y)$

$\overset{(set)}{=} x^2 \underbrace{- 3y^2}$ ($\underline{j}$ component of $\underline{F}$)

$\Rightarrow g'(y) = -3y^2 \Rightarrow g(y) = -y^3 + C$

$\Rightarrow f(x,y) = 3x + x^2 y - y^3 + C$

(for C constant)

$\underline{F} = -y\underline{i} + x\underline{j}$ is NOT conservative

set $\frac{\partial f}{\partial x} = -y \Rightarrow f(x,y) = -xy + g(y)$

$\Rightarrow \frac{\partial f}{\partial y} = \boxed{-x + g'(y) = x \quad ??}$

Can we still determine a potential function when the conservative vector field is in three dimensions?

34.2.4 The vector field $\boldsymbol{F}(x,y,z) = y^2\boldsymbol{i} + (2xy + e^{3z})\boldsymbol{j} + 3ye^{3z}\boldsymbol{k}$ is conservative. Find a corresponding potential function.

$\Rightarrow \exists f(x,y,z)$ s.t. $\underline{F} = \underline{\nabla} f$

$= \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j} + \frac{\partial f}{\partial z}\underline{k}$

$(\underline{i}) \Rightarrow \frac{\partial f}{\partial x} = y^2 \Rightarrow f(x,y,z) = xy^2 + \ldots$

$(\underline{j})\quad \frac{\partial f}{\partial y} = 2xy + e^{3z} \Rightarrow f(x,y,z) = xy^2 + ye^{3z} + \ldots$

$(\underline{k})\quad \frac{\partial f}{\partial z} = 3ye^{3z} \Rightarrow f(x,y,z) = ye^{3z} + \ldots$

$\Rightarrow f(x,y,z) = xy^2 + ye^{3z} + c$

without explicitly integrating to find f such that $\underline{F} = \underline{\nabla} f$

Is there a way of determining whether or not a given vector field is conservative? To answer this question, we need to go back to the study of line integrals.

Notes.

35 The fundamental theorem for line integrals, path independence

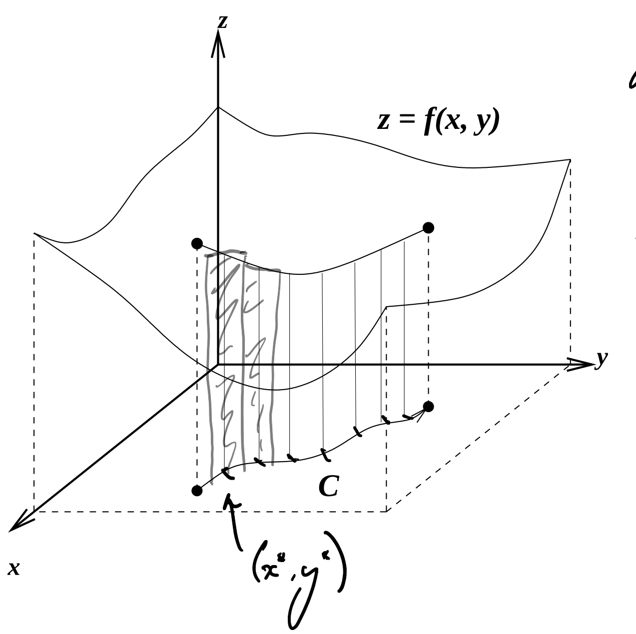
By the end of this section, you should be able to answer the following questions:

- How do you evaluate line integrals?
- What is the fundamental theorem for line integrals and its consequences?
- What is a path independent line integral and what are its connections with conservative vector fields and line integrals over closed curves?

35.1 Line integrals in the plane

Recall the definite integral $\int_a^b f(x)\,dx$ gives the net area above the x-axis and below the curve $y = f(x)$. We can generalise this.

Consider the following problem: How do we calculate the area of the region between the curve C in the x-y plane and its image on the surface $z = f(x, y)$?



$$\text{area} \approx \sum_{\text{(segment)}} \overbrace{f(x^*, y^*)}^{\text{height}}\,\underbrace{\Delta S}_{\text{width}}$$

$$\rightarrow \text{area} = \lim_{\Delta S \to 0} \sum_{\text{(segment)}} f(x^*, y^*)\,\Delta S \overset{\text{(def)}}{=} \int_C f(x,y)\,ds$$

If the curve C can be parametrised by $\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j}$ for $a \leq t \leq b$, then the area in question is given by the formula

$$\text{area} = \int_C f(x, y)\,dS = \int_a^b f(x(t), y(t))\|\boldsymbol{r}'(t)\|\,dt,$$

where dS is the infinitesimal element of arclength of C.

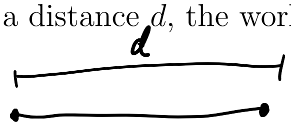
Any expression of this form is called a "line integral"

35.2 Work done by a force, line integrals of vector fields

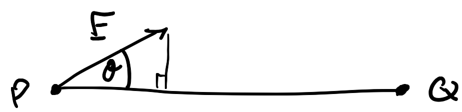
We can also consider integrating a vector field over a curve in the plane.

In the case $\boldsymbol{F}$ is a field of force you should already have an idea (from MATH1052/MATH1072) how to determine the work done by $\boldsymbol{F}$ in moving a particle along a curve C. First recall the cases of:

1. Constant F in 1D over a distance d, the work done is given by $W = Fd$.



2. Constant $\boldsymbol{F}$ in 2D over a straight line between points P and Q, $W = \boldsymbol{F} \cdot \overrightarrow{PQ}$.



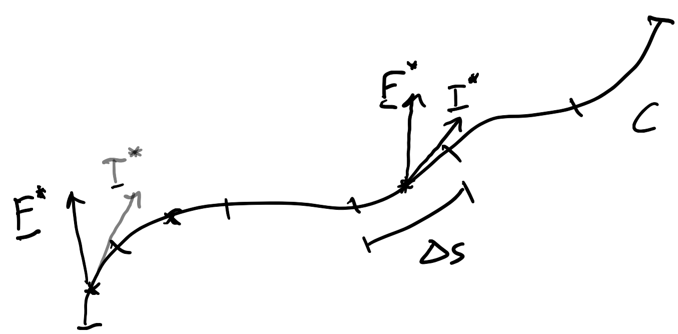
$W = \|\underline{F}\| \cos\theta \, \|\overrightarrow{PQ}\| = \underline{F} \cdot \overrightarrow{PQ}$

We can use these straightforward cases to derive the more general expression for work done by a variable force $\boldsymbol{F}(x, y)$ over a piecewise continuous smooth curve C:

$$W = \int_C \boldsymbol{F}(x, y) \cdot \boldsymbol{T}(x, y)\; dS,$$

where $\boldsymbol{T}(x, y)$ is a unit tangent vector to C at a given point (x, y) on C.

(assume $\underline{F}$ is continuous)



Approx $\underline{F}^*$ as constant over each segment of C, each segment we approximate as a straight line

$$W \approx \sum_{\text{(segments)}} \underline{F}^* \cdot (\underline{T}^* \Delta s) \xrightarrow{\lim \Delta s \to 0} \int_C \underline{F} \cdot \underline{T}\, dS$$

35.3 Evaluating line integrals of vector fields

To evaluate

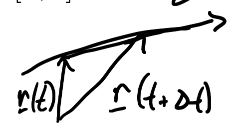
$$\int_C \boldsymbol{F}(x,y) \cdot \boldsymbol{T}(x,y)\; dS, \tag{17}$$

$\Delta t < 1$

the strategy is to parameterise C and express all quantities in the line integral in terms of this parameterisation. Namely,

$\frac{\underline{r}(t+\Delta t) - \underline{r}(t)}{\Delta t}$

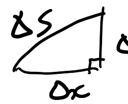
1. Parameterise C by specifiying a vector function $\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j}$ with $t \in [a,b]$ that describes C.



2. Write $\boldsymbol{F}(x,y)$ restricted to C as $\boldsymbol{F}(\boldsymbol{r}(t)) = \boldsymbol{F}(x(t), y(t))$.

3. Write $\boldsymbol{T}(x,y) = \dfrac{\boldsymbol{r}'(t)}{||\boldsymbol{r}'(t)||}$, where $\boldsymbol{r}'(t) = x'(t)\boldsymbol{i} + y'(t)\boldsymbol{j}$ is a tangent vector to C.

4. Write $dS = ||\boldsymbol{r}'(t)||dt$.



$\to \Delta S^2 \approx \Delta x^2 + \Delta y^2 \Rightarrow \Delta S \approx \underbrace{\sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2}}_{\simeq \|\underline{r}'(t)\|} \Delta t$

5. Evaluate the line integral as a definite integral in terms of the parameter t:

$$\begin{aligned}\int_C \boldsymbol{F}(x,y) \cdot \boldsymbol{T}(x,y)\; dS &= \int_a^b \left[\boldsymbol{F}(\boldsymbol{r}(t)) \cdot \frac{\boldsymbol{r}'(t)}{||\boldsymbol{r}'(t)||}\right] ||\boldsymbol{r}'(t)||dt \\ &= \int_a^b \boldsymbol{F}(\boldsymbol{r}(t)) \cdot \boldsymbol{r}'(t)dt.\end{aligned}$$

$\longleftarrow$ evaluation

35.4 Common notation

Let C be a piecewise continuous smooth curve in the x-y plane connecting points A and B. Let $\boldsymbol{F}(x,y) = F_1(x,y)\boldsymbol{i} + F_2(x,y)\boldsymbol{j}$ be a vector field. The line integral in expression (17) is often expressed as

$$\int_C \boldsymbol{F}(\boldsymbol{r}) \cdot \boldsymbol{dr}$$

or

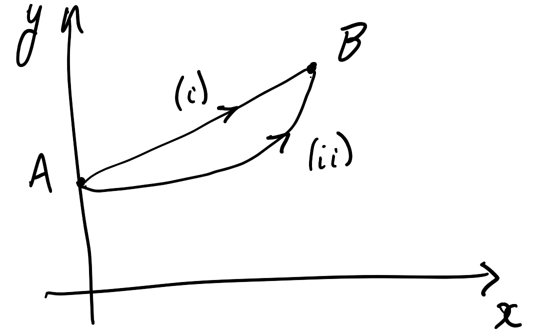
$$\int_C F_1(x,y)dx + F_2(x,y)dy,$$

both of which are *parameter independent* ways of writing the line integral in (17). The expressions use the notation $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$, $\boldsymbol{dr} = dx\boldsymbol{i} + dy\boldsymbol{j}$. Introducing a parameterisation $\boldsymbol{r}(t)$ for $t \in [a,b]$ allows us to evaluate the line integral. The expressions above are useful notations to remind us how to evaluate the line integrals.

$\int \left(\underline{F} \cdot \frac{d\underline{r}}{dt}\right) dt$

$\underset{\sim}{F}(x,y) = (x^2-y)\underset{\sim}{i} + (y^2+x)\underset{\sim}{j}$

35.4.1 Example: let $A = (0,1)$, $B = (1,2)$. Evaluate $\int_C ((x^2 - y)dx + (y^2 + x)dy)$ along the curve C given by: (i) the straight line from A to B; (ii) the parabola $y = x^2 + 1$ from A to B.



(i) Line $y = x+1$

parametrise $\underset{\sim}{r}(t) = t\underset{\sim}{i} + (t+1)\underset{\sim}{j}, \quad 0 \le t \le 1$

i.e. $x = t, \quad y = t+1$

$$\int_C = \int_0^1 \left((x^2-y)\frac{dx}{dt} + (y^2+x)\frac{dy}{dt}\right)dt$$

$$= \int_0^1 \left((t^2-(t+1))\cdot 1 + ((t+1)^2+t)\cdot 1\right)dt$$

$$= \ldots = 5/3$$

(ii) $y = x^2+1 \Rightarrow \underset{\sim}{r}(t) = t\underset{\sim}{i} + (t^2+1)\underset{\sim}{j}, \quad 0 \le t \le 1$

$x = t, \quad y = t^2+1 \Rightarrow \frac{dx}{dt} = 1, \quad \frac{dy}{dt} = 2t$

$$\int_C = \int_0^1 ((t^2-(t^2+1))\cdot 1 + ((t^2+1)^2+t)\cdot 2t)dt$$

$$= \ldots = 2$$

Note the line integrals in the previous example were *path dependent*. In other words, they have different values for different paths.

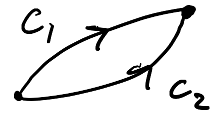
We will now investigate path independent line integrals.

35.5 Line integrals of conservative vector fields, path independence.

$\underline{F} = \underline{\nabla} f$

If $\boldsymbol{F}$ is a continuous vector field with domain D, we say the line integral $\int_C \boldsymbol{F} \cdot \boldsymbol{dr}$ is *path independent* if

$$\int_{C_1} \boldsymbol{F} \cdot \boldsymbol{dr} = \int_{C_2} \boldsymbol{F} \cdot \boldsymbol{dr}$$



for any two paths C_1 and C_2 in D that have the same end points.

35.5.1 The fundamental theorem for line integrals

If C is a smooth curve determined by $\boldsymbol{r}(t)$ for $t \in [a, b]$ and $f(x, y)$ is differentiable with ∇f being continuous on C, then

$f(\underline{r}(t)) = f(x(t), y(t))$

$$\int_C \nabla f \cdot \boldsymbol{dr} = f(\boldsymbol{r}(b)) - f(\boldsymbol{r}(a)).$$

$d\underline{r} = dx\,\underline{i} + dy\,\underline{i}$

Proof:

$$\underline{\nabla} f = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j}$$

$$\Rightarrow \int_C \underline{\nabla} f \cdot d\underline{r} = \int_a^b \underbrace{\left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\right)}_{\text{chain rule}} dt$$

$$= \int_a^b \frac{df}{dt}\, dt$$

$$= \int_{f(\underline{r}(a))}^{f(\underline{r}(b))} df \rightarrow \text{result}$$

One consequence is that for conservative vector fields ∇f, we have

$$\int_{C_1} \nabla f \cdot \boldsymbol{dr} = \int_{C_2} \nabla f \cdot \boldsymbol{dr}.$$

That is, the line integral of a conservative vector field is path independent.

It turns out, the converse is also true. Suppose $\boldsymbol{F}$ is continuous on an open, connected region D. If $\int_C \boldsymbol{F} \cdot \boldsymbol{dr}$ is path independent in D, then $\boldsymbol{F}$ is conservative.

Proof:

Stewart ed. 8 "Theorem 4"

p1129

Open region: every point in the region is the centre of some disc lying entirely in the region (ie. an open region doesn't include the boundary points).

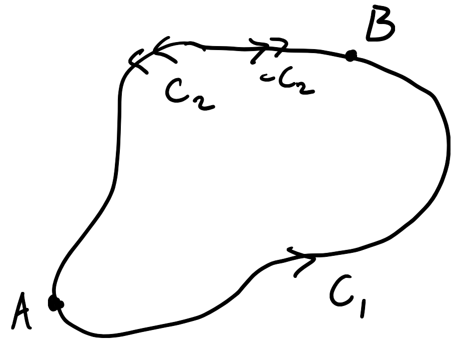
Connected region: Any two points in D can be joined by a path lying entirely in D.

Another interesting result is that if $\displaystyle\int_C \boldsymbol{F}\cdot d\boldsymbol{r}$ is path independent in some region D, then $\displaystyle\oint_{C'} \boldsymbol{F}\cdot d\boldsymbol{r} = 0$ for *every* closed path C' in D. Here the symbol "$\oint$" indicates the integral is over a *closed* curve.

Proof:

Assume path independence

Let C' be a closed curve



choose two points on C', say A and B

Define
$$\left.\begin{aligned} C_1 &: A \to B \\ C_2 &: B \to A \\ -C_2 &: A \to B \end{aligned}\right\} \text{as indicated}$$

$C' = C_1$ followed by C_2

Note: $$\int_{C_2} \underset{\sim}{F}\cdot d\underset{\sim}{r} = -\int_{-C_2} \underset{\sim}{F}\cdot d\underset{\sim}{r} \qquad (*)$$

$$\Rightarrow \oint_{C'} \underset{\sim}{F}\cdot d\underset{\sim}{r} = \int_{C_1} \underset{\sim}{F}\cdot d\underset{\sim}{r} + \int_{C_2} \underset{\sim}{F}\cdot d\underset{\sim}{r} \qquad (**)$$

$$= \int_{C_1} \underset{\sim}{F}\cdot d\underset{\sim}{r} - \int_{-C_2} \underset{\sim}{F}\cdot d\underset{\sim}{r} = 0$$

since C_1 & $-C_2$ have the same end points & line integrals are path independent

Perhaps it is not surprising that the converse is also true. That is, if $\oint_{C'} \boldsymbol{F} \cdot \boldsymbol{dr} = 0$ for *every* closed path C' in some region D, then $\int_C \boldsymbol{F} \cdot \boldsymbol{dr}$ is path independent in D.

Proof:

Stewart ed. 8 p1128-1129

We are looking at these results carefully because we ultimately want a simple way of checking whether or not a vector field is conservative. We are not quite there yet, but in the next section, we will arrive at a surprisingly simple test for a conservative vector field.

Note also that more details of these proofs (with slightly more mathematical rigour) can be found in Stewart, pages 1099 – 1103.

Notes.

36 Green's theorem and a test for conservative fields

By the end of this section, you should be able to answer the following questions:

- What is Green's theorem and under what conditions can it be applied?
- How do you apply Green's theorem?
- Given a vector field in two dimensions, how can we test whether or not it is conservative?

36.1 The story so far (in two dimensions, $\mathbb{R}^2$)

The following diagram summarises the relationships between conservative vector fields, path independent line integrals and closed line integrals we have seen so far.

$$\begin{tikzcd}
& \boldsymbol{F} \text{ conservative } (\text{i.e. } \underline{F} = \underline{\nabla} f) \arrow[dl, shift left] \\
\displaystyle\int_C \boldsymbol{F} \cdot \boldsymbol{dr} \text{ path independent} \arrow[ur, shift left, "\text{connected region } *"] \arrow[dr, shift left] & \\
& \displaystyle\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0, \ \forall \text{ closed } C \arrow[ul, shift left]
\end{tikzcd}$$

36.2 Clairaut's theorem and consequences

Suppose a function of two variables f is defined on a disc D that contains the point (a, b). If the functions $\dfrac{\partial^2 f}{\partial x \partial y}$ and $\dfrac{\partial^2 f}{\partial y \partial x}$ are both continuous on D, then

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b).$$

$F(x,y)$

Say we have a conservative vector field $\boldsymbol{F} = F_1\boldsymbol{i} + F_2\boldsymbol{j}$. This means that there exists an $f(x, y)$ such that

$$F_1 = \frac{\partial f}{\partial x}, \quad F_2 = \frac{\partial f}{\partial y}.$$

An immediate consequence of Clairaut's theorem is that

$$\frac{\partial F_1}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial F_2}{\partial x}.$$

In otherwords, we have the following:

> If $\boldsymbol{F} = F_1\boldsymbol{i} + F_2\boldsymbol{j}$ is a conservative vector field, then
> $$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}.$$

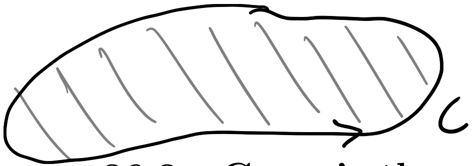
Let's add this to our diagram:

$$\begin{tikzcd}
 & \boldsymbol{F} \text{ conservative} \arrow[dl, shift left] \arrow[dr] & \\
\displaystyle\int_C \boldsymbol{F} \cdot \boldsymbol{dr} \text{ path independent} \arrow[ur, shift left] \arrow[dr, shift left] & & \dfrac{\partial F_1}{\partial y} = \dfrac{\partial F_2}{\partial x} \\
 & \displaystyle\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0, \ \forall \text{ closed } C \arrow[ul, shift left] &
\end{tikzcd}$$

If we can reverse the new arrow, then we would have the criterion that we need! That is, the condition

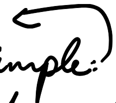
$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$$

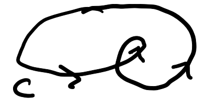
would be a test for a conservative vector field. To do this, we require one more piece of the puzzle. That is Green's theorem.



C

simple closed curve: Does not intersect itself anywhere between end points

e.g. simple: 

not simple: 

C

36.3 Green's theorem

Let D be a region in the x-y plane bounded by a piecewise-smooth, simple closed curve C, which is traversed with D always on the left. Let $F_1(x, y)$, $F_2(x, y)$, $\dfrac{\partial F_1}{\partial y}$ and $\dfrac{\partial F_2}{\partial x}$ be continuous in D. Then

"positively oriented"

$$\boxed{\iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \, dx \, dy = \oint_C (F_1 dx + F_2 dy).} \longrightarrow \oint_C \underline{F} \cdot \underline{T} \, ds$$

This theorem relates a double integral to a line integral over a closed curve. For example, we can use Green's theorem to evaluate complicated line integrals by treating them as double integrals, or vice versa.

Regarding our discussion on conservative vector fields, we have the following *corollary* to Green's theorem:

$$\text{If } \frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \text{ then } \oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0.$$

Note that $\boldsymbol{F} = F_1 \boldsymbol{i} + F_2 \boldsymbol{j}$.

If we add this to our diagram, we can now link any four statements via the arrows. In otherwords all four statements are equivalent.

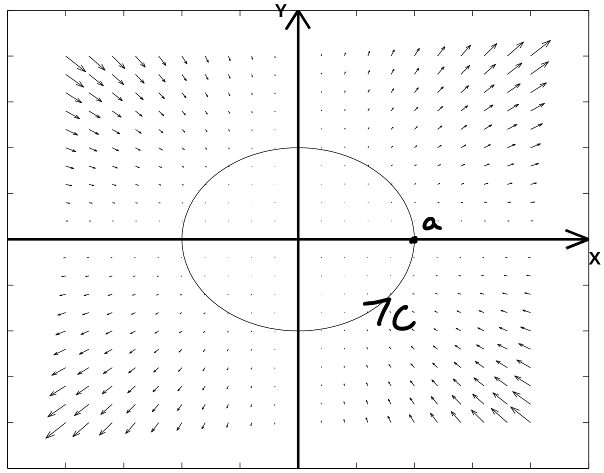
$$\begin{tikzcd}
 & \boldsymbol{F} \text{ conservative} \arrow[dl] \arrow[dr] & \\
\displaystyle\int_C \boldsymbol{F} \cdot \boldsymbol{dr} \text{ path independent} \arrow[ur] \arrow[dr] & & \dfrac{\partial F_1}{\partial y} = \dfrac{\partial F_2}{\partial x} \arrow[dl, "{*\text{Green's Theorem}}"] \\
 & \displaystyle\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0, \ \forall \text{ closed } C \arrow[ul] &
\end{tikzcd}$$

In particular, we now have a test to determine whether or not a given two dimensional vector field is conservative:

$$\text{The vector field } \boldsymbol{F} \text{ is conservative } \underline{\text{if and only}} \text{ if } \frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}.$$

$= F_1 \underline{i} + F_2 \underline{j}$

36.3.1 Find the work done by the force $\boldsymbol{F} = x^2y\boldsymbol{i} + xy^2\boldsymbol{j}$ anticlockwise around the circle with centre at the origin and radius a.



Conditions of Green's Theorem hold

$$\text{Work} = \oint_C \underline{F} \cdot d\underline{r} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA$$

polar coordinates $\Rightarrow$ $x = r\cos\theta$, $y = r\sin\theta$

$$D = \{(r, \theta) \mid 0 \le r \le a,\ 0 \le \theta \le 2\pi\}$$

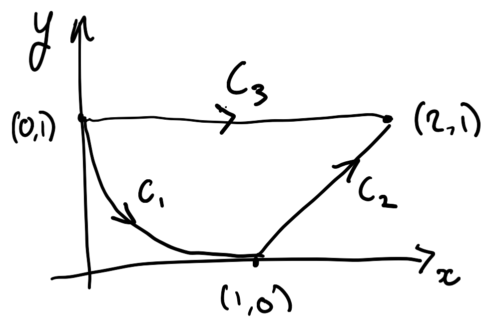
$$\boxed{\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = y^2 - x^2} = \underbrace{r^2(\sin^2\theta - \cos^2\theta) = -r^2\cos 2\theta}_{\text{(trig identity)}}$$

$$\Rightarrow \oint_C \underline{F} \cdot d\underline{r} = -\int_0^a \int_0^{2\pi} r^2 \cos 2\theta \cdot r \, d\theta \, dr = 0$$

Note, however, $\underline{F}$ is NOT conservative

$$\underline{F}(x,y) = 2xy\,\underline{i} + (x^2+3y^2)\,\underline{j}$$

36.3.2 **Evaluate the line integral** $\int_C 2xy\ dx + (x^2+3y^2)\ dy$**, where** C **is the path from** $(0,1)$ **to** $(1,0)$ **along** $y=(x-1)^2$ **and then from** $(1,0)$ **to** $(2,1)$ **along** $y = x-1$**.**



$C = C_1$ then C_2

$$\left.\begin{array}{l} F_1 = 2xy \Rightarrow \dfrac{\partial F_1}{\partial y} = 2x \\ F_2 = x^2+3y^2 \Rightarrow \dfrac{\partial F_2}{\partial x} = 2x \end{array}\right\} \Rightarrow \underline{F}(x,y) \text{ is conservative}$$

$\Rightarrow \int_C \underline{F}\cdot d\underline{r}$ is path independent

$\Rightarrow$ can choose any path from $(0,1)$ to $(2,1)$. Choose C_3 as indicated

C_3: $\underline{r}(t) = t\,\underline{i} + 1\,\underline{j}$, $0 \le t \le 2$

$$\left.\begin{array}{l} \left(\dfrac{d\underline{r}}{dt} = \right) \underline{r}'(t) = \underline{i} \\ \underline{F}(\underline{r}(t)) = 2t\,\underline{i} + (t^2+3)\,\underline{j} \end{array}\right] \rightarrow \underline{F}(\underline{r}(t))\cdot\frac{d\underline{r}}{dt} = 2t + 0$$

$$\Rightarrow \int_{C_3} \underline{F}\cdot d\underline{r} = \int_0^2 \left(\underline{F}(\underline{r}(t))\cdot \frac{d\underline{r}}{dt}\right) dt$$

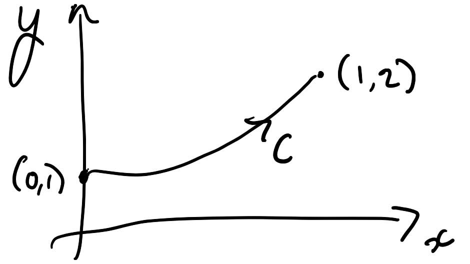
---

or $\underline{F} = \underline{\nabla} f \rightarrow f(x,y) = x^2 y + y^3 + c$

$$\Rightarrow \int_C = f(2,1) - f(0,1) = 4$$

$\underline{F}(x,y) = (3+2xy)\underline{i} + (x^2-3y^2)\underline{j}$

36.3.3 **Evaluate** $\displaystyle\int_C (3+2xy)dx + (x^2-3y^2)dy$ **where** C **is the curve parametrised by** $\boldsymbol{r}(t) = (1-\cos(\pi t))\boldsymbol{i} + (1+\sin^3(\pi t))\boldsymbol{j}$ **for** $0 \le t \le 1/2$.



$\underline{r}(0) = 0\underline{i} + 1\underline{j}$

$\underline{r}(\tfrac{1}{2}) = 1\underline{i} + 2\underline{j}$

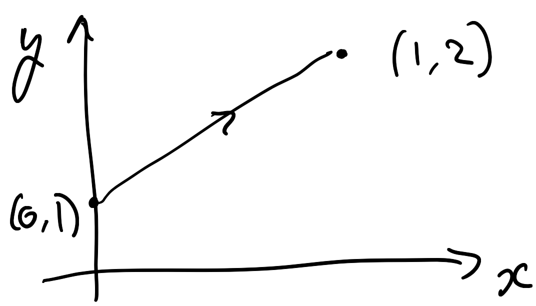
Same $\underline{F}(x,y)$ as ex. 34.2.3 on p239

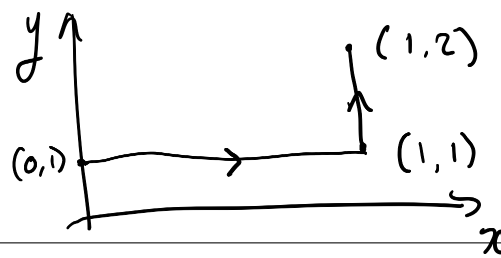
$\underline{F} = \underline{\nabla} f, \quad f(x,y) = 3x + x^2y - y^3 + c$

Fundamental Theorem for line integrals

$$\Rightarrow \int_C \underline{F}\cdot d\underline{r} = f(1,2) - f(0,1)$$
$$= (3+2-8+c) - (0+0-1+c)$$
$$= -2$$

OR





Since path independent!

Notes.

37 On proving Green's theorem

Recall some important properties of double and line integrals.

$$\iint\limits_{D_1 \cup D_2} f(x,y)\ dA = \iint\limits_{D_1} f(x,y)\ dA + \iint\limits_{D_2} f(x,y)\ dA,$$
$$\iint\limits_{D} (f_1(x,y) + f_2(x,y))\ dA = \iint\limits_{D} f_1(x,y)\ dA + \iint\limits_{D} f_2(x,y)\ dA,$$

where D_1 and D_2 are disjoint (apart from sharing a boundary). If C consists of the segments $C_1, \cdots, C_4$ such that $C = C_1 \cup C_2 \cup C_3 \cup C_4$, then

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r} + \int_{C_3} \mathbf{F} \cdot d\mathbf{r} + \int_{C_4} \mathbf{F} \cdot d\mathbf{r}$$

With the understanding that

$$\int_C F_1\, dx = \int_C (F_1\, dx + 0\, dy), \qquad \int_C F_2\, dy = \int_C (0\, dx + F_2\, dy),$$

one can prove (Exercise!) that

$$\int_C (F_1\, dx + F_2\, dy) = \int_C F_1\, dx + \int_C F_2\, dy.$$

Moreover one can prove the following two results:

or $\mathbf{r}(x) = x\mathbf{i} + g(x)\mathbf{j}$, $a \le x \le b$

1. If C can be parametrized by $\underline{\mathbf{r}(t) = t\mathbf{i} + g(t)\mathbf{j} \text{ with } a \le t \le b}$, then

$$\int_C F_1(x,y)\ dx = \int_a^b F_1(x, g(x))\ dx.$$

or $\mathbf{r}(y) = h(y)\mathbf{i} + y\mathbf{j}$, $c \le y \le d$

2. If C can be parametrized by $\underline{\mathbf{r}(t) = h(t)\mathbf{i} + t\mathbf{j} \text{ with } c \le t \le d}$, then

$$\int_C F_2(x,y)\ dy = \int_c^d F_2(h(y), y)\ dy.$$

We will use these results in the following.

Green's Theorem. Let D be a region in the x-y plane bounded by a piecewise-smooth, simple, closed, positively oriented curve C, and let $F_1(x,y)$, $F_2(x,y)$, $\frac{\partial F_1(x,y)}{\partial y}$, $\frac{\partial F_2(x,y)}{\partial x}$ be continuous. Then,

$$\iint\limits_D \Big(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\Big) dA = \oint_C (F_1\, dx + F_2\, dy).$$

In Sections 37.1-37.3, we prove the theorem for a particular class of regions. A general proof is beyond the scope of this workbook.

↳ D is both type I and type II

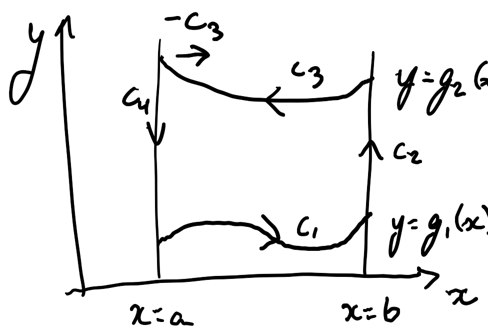
37.1 Region of type I

Recall that a region D is of type I if

$$D = \{(x,y)\,|\, a \le x \le b,\ g_1(x) \le y \le g_2(x)\},$$

where g_1 and g_2 are continuous functions. We prove that for D of type I,

$$\iint\limits_D \frac{\partial F_1}{\partial y}\, dA = -\oint_C F_1\, dx.$$



$C = C_1 \cup C_2 \cup C_3 \cup C_4$

$$\text{LHS} = \iint\limits_D \frac{\partial F_1}{\partial y}\, dA = \int_a^b \Big(\int_{g_1(x)}^{g_2(x)} \frac{\partial F_1}{\partial y}\, dy \Big) dx$$

$$= \int_a^b \Big(F_1(x, g_2(x)) - F_1(x, g_1(x)) \Big) dx$$

From RHS, $\oint_C F_1\,dx = \int_{C_1} F_1\,dx + \int_{C_2} F_1\,dx + \int_{C_3} F_1\,dx + \int_{C_4} F_1\,dx$

C_1: $\underline{r}(x) = x\underline{i} + g_1(x)\underline{j}$, $a \le x \le b$

$$\Rightarrow \int_{C_1} F_1\,dx = \int_a^b F_1(x, g_1(x))\,dx$$

C_2: $\underline{r}(y) = b\underline{i} + y\underline{j}$, $g_1(b) \le y \le g_2(b)$

$$\int_{C_2} F_1\,dx = 0 = \int_{C_4} F_1\,dx \quad \text{(by similarity)}$$

$-C_3$: $\underline{r}(x) = x\underline{i} + g_2(x)\underline{j}$, $a \le x \le b$

$$\int_{C_3} F_1\,dx = -\int_{-C_3} F_1\,dx = -\int_a^b F_1(x, g_2(x))\,dx$$

C_4: $\int_{C_4} F_1\,dx = 0$ as above

$$\Rightarrow \oint F_1\,dx = \int \Big(F_1(x, g_1(x)) - F_1(x, g_2(x))\Big)\,dx$$

$$= -\iint_D \frac{\partial F_1}{\partial y}\,dA$$

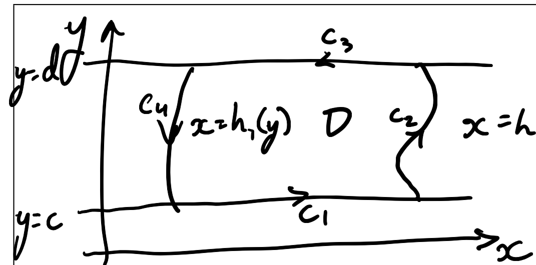
37.2 Region of type II

A region D is of type II if

$$D = \{(x, y) \mid h_1(y) \le x \le h_2(y),\, c \le y \le d\},$$

where h_1 and h_2 are continuous functions. Then similarly to the last subsection, by considering D as a type II region, we find that

$$\iint\limits_D \frac{\partial F_2}{\partial x}\, dA = \oint_C F_2\, dy.$$



$$\iint_D \frac{\partial F_2}{\partial x}\, dA = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} \frac{\partial F_2}{\partial x}\, dx \right) dy$$

$$= \int_c^d \left(F_2(h_2(y), y) - F_1(h_1(y), y) \right) dy$$

$$\oint_C F_2\, dy = \underbrace{\int_{C_1} F_2\, dy}_{=0} + \int_{C_2} F_2\, dy + \underbrace{\int_{C_3} F_2\, dy}_{=0} + \int_{C_4} F_2\, dy$$

C_2: $\underline{r}(y) = h_2(y)\underline{i} + y\underline{j}$, $c \le y \le d$

$-C_4$: $\underline{r}(y) = h_1(y)\underline{i} + y\underline{j}$, $c \le y \le d$ $\Rightarrow \int_{C_4} F_2\, dy = -\int_{-C_4} F_2\, dy$

$\Rightarrow \oint_C F_2\, dy = \int \left(F_2(h_2(y), y) - F_2(h_1(y), y) \right) dy$

$= \iint_D \frac{\partial F_2}{\partial x}\, dA$

37.3 Region of both type I and II

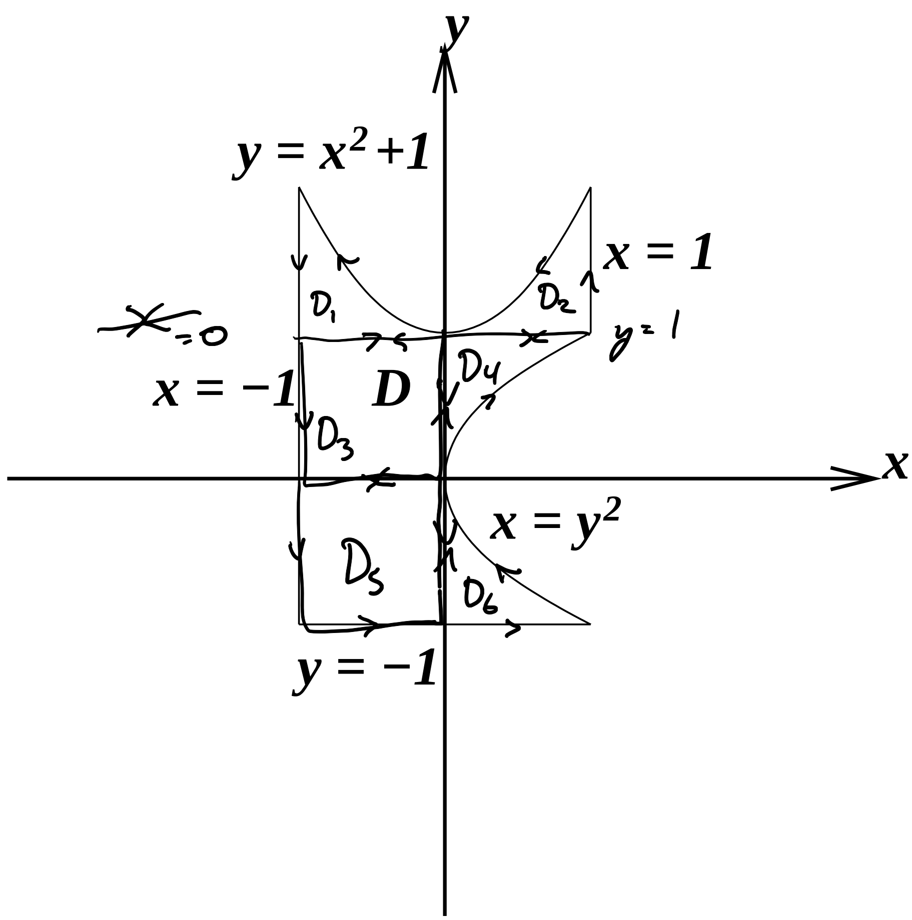
If D is of *both* type I and II, we can combine the results of Sections 37.1 and 37.2, thereby obtaining Green's theorem:

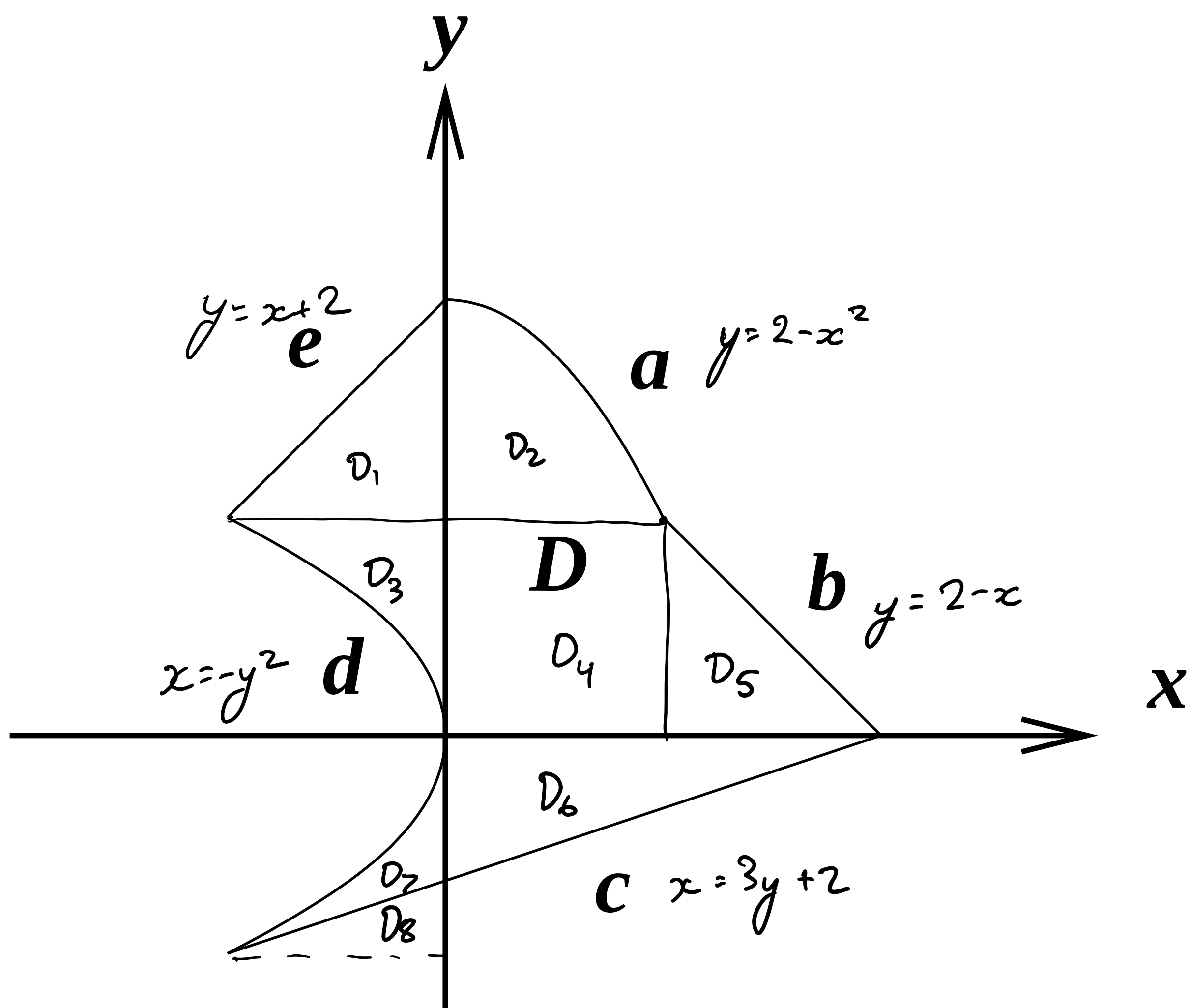
$$\iint\limits_D \Big(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\Big) dA = \iint\limits_D \frac{\partial F_2}{\partial x}\, dA - \iint\limits_D \frac{\partial F_1}{\partial y}\, dA = \oint_C F_2\, dy - \Big(- \oint_C F_1\, dx\Big)$$
$$= \oint_C (F_1\, dx + F_2\, dy).$$

37.4 More general regions

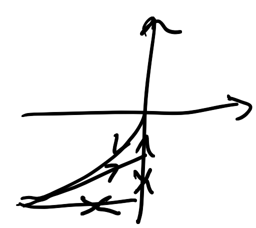
If the planar region is neither type I nor type II, our strategy is to attempt to partition the region into subregions which are both type I and type II, then make use of the results of the previous calculations.

In the following two examples, partition the regions into subregions that are both type I and type II. Could we generalise the previous calculations to deal with subregions that are only either type I or II?





want $\iint_{D_7}$,



$$\iint_{D_7 \cup D_8} = \iint_{D_7} + \iint_{D_8} \Rightarrow \iint_{D_7} = \iint_{D_7 \cup D_8} - \iint_{D_8}$$

Notes.

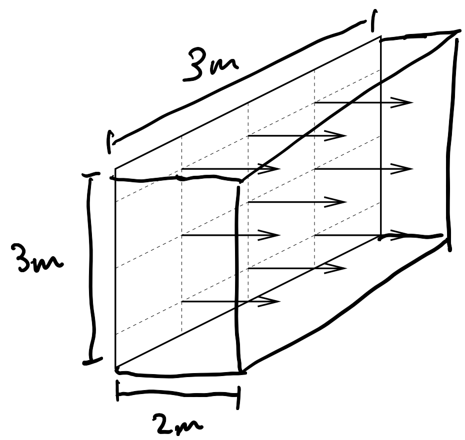
38 Flux of a vector field

By the end of this section, you should be able to answer the following questions:

- What is the flux of a constant vector field across a flat surface in 3D?
- What is the flux of a vector field across a plane curve in 2D?

In this section we introduce the concept of *flux*: In three dimensions, the flux of a vector field across a given surface is defined to be the "flow rate" of the vector field through the surface.

Since many vector fields involve no motion (eg. electric fields, magnetic fields), this definition can be very difficult to comprehend at first. A nice context for working with flux in order to understand its definition is by considering the velocity vector of a fluid (so now we do have motion). In three dimensions, the flux of a fluid across a surface is given in units of volume per unit time. In other words, the flux tells us how much of the fluid (volume) passes through a given surface in one second. (constant vector field $\leftrightarrow$ velocity)



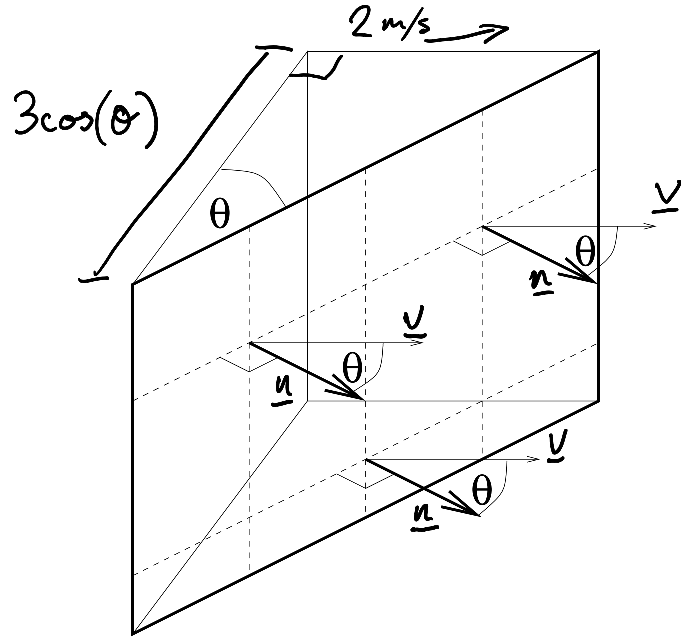
More generally, flux across a surface $\equiv$ instantaneous flow rate on surface

Consider a river flowing at a constant velocity of $2m/s$ in only one direction. Now imagine placing a $3m$ square fishing net into the river so that it somehow stays perpendicular to the flow of the river. What is the flux of the water through the net?

Volume through net after 1 second

$$= 3 \times 3 \times 2 = 18m^3$$

$$\Rightarrow \text{flux} = 18m^3/s$$



Now if we rotate the net through an angle θ, what is the flux through the net?

In 1 second, same volume passes through the net as a net perpendicular to motion with dimensions $3 \times 3\cos\theta$

$\Rightarrow$ volume after 1sec $= 3 \times 3\cos\theta \times 2$
$= 18\cos\theta \ m^3$

$\Rightarrow$ flux $= 18\cos\theta \ m^3/s$

OR/ Component of flow perpendicular to net $\underline{v} \cdot \underline{n}$ ($\underline{n}$ unit vector perpendicular to surface = "unit normal vector")

$\underline{v} \cdot \underline{n} = \|\underline{v}\| \|\underline{n}\| \cos\theta = 2 \times 1 \times \cos\theta$

$\Rightarrow$ flux $= \underline{v} \cdot \underline{n} \times$ (surface area of net)

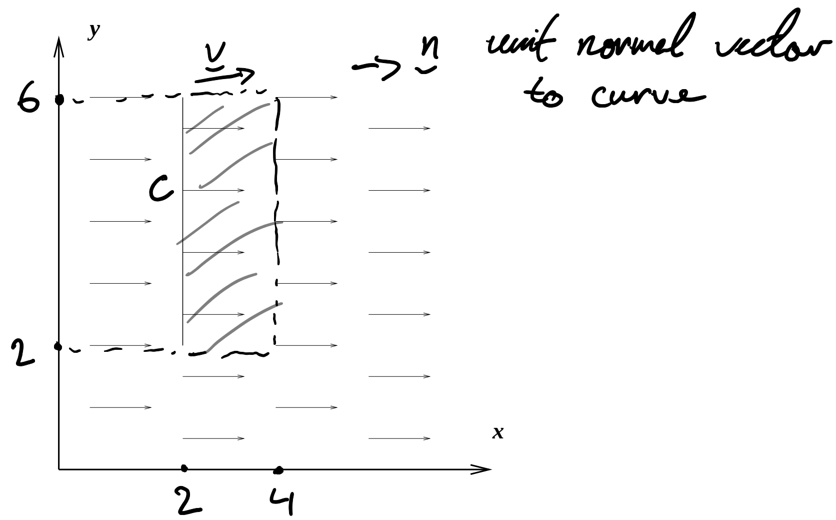
$= 2\cos\theta \times 3 \times 3 = 18\cos\theta$

38.1 Flux in 2D

Before we look at the flux of a vector field through more general surfaces, let's look at flux in two dimensions, by considering the flow of a two dimensional fluid through a curve in the x-y plane. Note that in this context of a fluid in 2D, flux has dimensions *area* per unit time.

To start, consider the problem of calculating the flux of a fluid with constant velocity $\underset{\sim}{\boldsymbol{v} = 2\boldsymbol{i}}$ through a line segment C perpendicular to the flow, where C is given by

$$C = \{(x, y) \mid x = 2,\ 2 \le y \le 6\}.$$



flux $= (\underset{\sim}{v} \cdot \underset{\sim}{n}) \times$ (length of line segment)

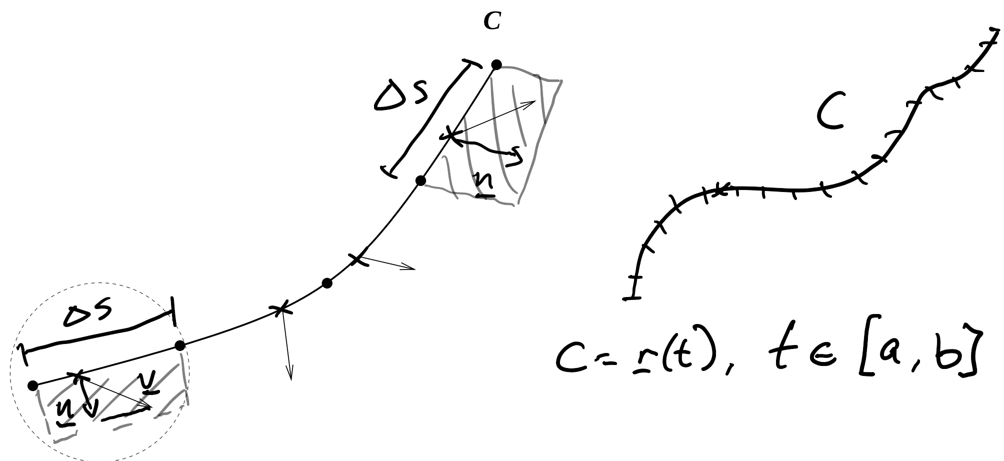
$= 2 \times 4 = 8$

OR/ "area" through line segment (2D volume) after 1 second

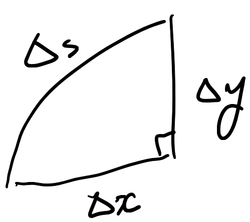
(Assume C is piecewise smooth, $\underline{v}(x,y)$ continuous)

Now consider calculating the flux of the velocity vector $\boldsymbol{v}(x,y)$ in the x-y plane through a curve C.

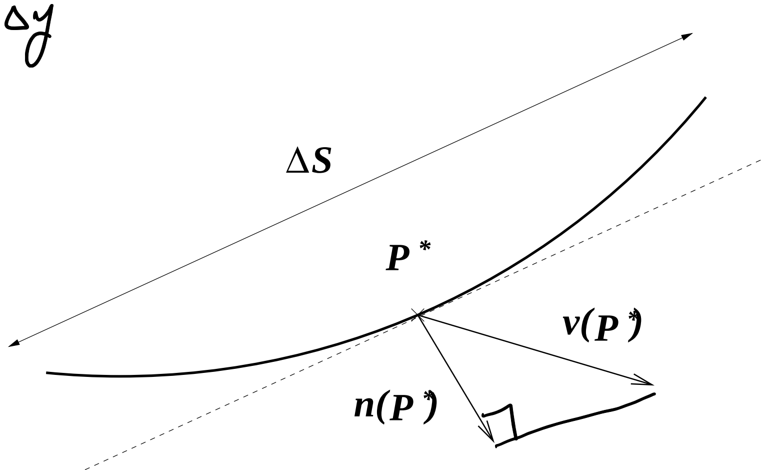
We first divide C up into arcs of length ΔS, and *approximate* $\boldsymbol{v}$ as constant over each arc.



This constant vector over each arc shall be evaluated at a representative point in each arc, say $P^* = (x^*, y^*)$. We also approximate the arc as a straight line, so that



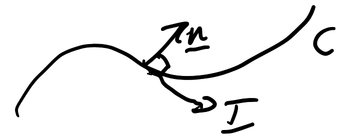
$$\Delta S \approx \sqrt{(\Delta x)^2 + (\Delta y)^2} \approx ||\boldsymbol{r}'(t)||\ \Delta t.$$



The component of $\boldsymbol{v}$ which is perpendicular to C (over ΔS) is $\approx \boldsymbol{v}(P^*) \cdot \boldsymbol{n}(P^*)$. We then have

$$\text{flux through one arc } \approx \boldsymbol{v}(P^*) \cdot \boldsymbol{n}(P^*)\Delta S.$$

$$\Rightarrow \text{ total flux through } C \approx \sum_i \boldsymbol{v}(P_i^*) \cdot \boldsymbol{n}(P_i^*)\Delta S_i.$$



If we take the limit as $\Delta S \to 0$, we obtain an exact expression for the flux over the entire curve C as a line integral:

note $\underline{n}(x,y)$

"flux integral in 2D" →

$$\text{Flux} \;=\; \int_C \boldsymbol{v}\cdot\boldsymbol{n}\; dS,$$

c.f. $\int_C \underline{v}\cdot\underline{T}\; dS$

$\underline{v}(x,y)$

where $\boldsymbol{n}$ is a unit vector normal to C.

We use this expression as a *definition* of flux of any two dimensional vector field $\boldsymbol{v}$ across a plane curve C. Note then that

$$\text{dimensions of flux (in 2D)} \;=\; (\text{dimensions of } \boldsymbol{v}) \;\times\; (\text{distance}).$$

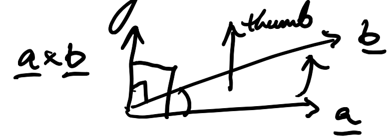
38.1.1 Evaluating flux in 2D

vel. $\frac{m}{s}$ × m

To evaluate the line integral in the definition of flux, we need a parametrisation of C, say $\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j}$ for $a \le t \le b$ (say). We define

$$\boldsymbol{r}'(t) = \dot{x}\;\boldsymbol{i} + \dot{y}\;\boldsymbol{j}.$$

"right hand rule"



A unit tangent vector to C is then given by

$$\boldsymbol{T} = \frac{\boldsymbol{r}'(t)}{||\boldsymbol{r}'(t)||}.$$

By the definition of vector cross product, and since $\boldsymbol{k}$ is a unit vector normal to the x-y plane, being careful of the direction of $\boldsymbol{n}$, we can take

$$\boldsymbol{n} = \boldsymbol{T}\times\boldsymbol{k} = \frac{1}{||\boldsymbol{r}'(t)||}\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dot{x} & \dot{y} & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

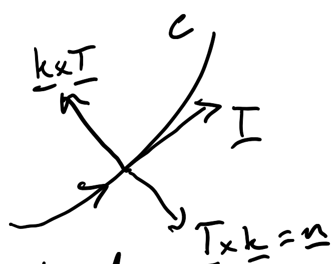
$$= \frac{1}{||\boldsymbol{r}'(t)||}(\dot{y}\;\boldsymbol{i} - \dot{x}\;\boldsymbol{j})$$

$$\Rightarrow\; \boldsymbol{v}\cdot\boldsymbol{n} \;=\; \frac{\boldsymbol{v}\cdot(\dot{y}\;\boldsymbol{i} - \dot{x}\;\boldsymbol{j})}{||\boldsymbol{r}'(t)||}$$

$$= \frac{v_1\dot{y} - v_2\dot{x}}{||\boldsymbol{r}'(t)||},$$

or.. $\underline{n} = \underline{k} \times \underline{T}$

important: $\underline{n}$ determines the direction of the positive flux



unit vector
$\underline{i} \to x$
$\underline{j} \to y$
$\underline{k} \to z$

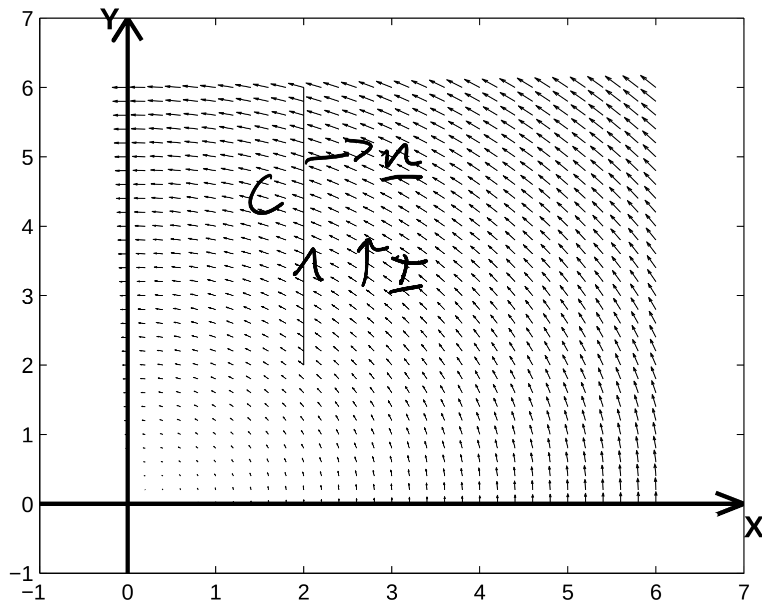
where $\boldsymbol{v}(x,y) = v_1(x,y)\boldsymbol{i} + v_2(x,y)\boldsymbol{j}$. Noting also that in the integral we have $dS = ||\boldsymbol{r}'(t)||\; dt$, we then have a means of evaluating the line integral (2D flux integral) as

$$\int_C \boldsymbol{v}\cdot\boldsymbol{n}\; dS = \int_{t=a}^{t=b} (v_1(t)\dot{y} - v_2(t)\dot{x})\; dt.$$

$||\underline{a} \times \underline{b}|| = ||\underline{a}||\,||\underline{b}|| \sin\theta$

go to p275

direction of positive flux →



take $\underline{T} \times \underline{k}$ from RHR (align thumb with $\underline{n}$, then fingers curl to connect $\underline{T}$ and $\underline{k}$)

38.1.2 Calculate the flux of $\boldsymbol{v} = -y\boldsymbol{i} + x\boldsymbol{j}$ $\overbrace{\text{(in the positive } x \text{ direction)}}^{\text{direction of positive flux}}$ across the line $x = 2$ (for $2 \leq y \leq 6$).

Strategy: flux $= \int_C \underline{v} \cdot \underline{n} \, ds$

$C: \underline{r}(t) = 2\underline{i} + t\underline{j}, \quad 2 \leq t \leq 6$

$\Rightarrow \underline{r}'(t) = \underline{j}$

$\Rightarrow$ flux $= \int \underline{v}(\underline{r}(t)) \cdot \left(\frac{\underline{r}'(t)}{\cancel{\|\underline{r}'(t)\|}} \times \underline{k} \right) \cancel{\|\underline{r}'(t)\|} \, dt$

$(\underline{v}(\underline{r}(t))) = -t\underline{i} + 2\underline{j}, \quad r'(t) \times \underline{k} = \underline{j} \times \underline{k} = \underline{i}$

$\Rightarrow$ flux $= \int_2^6 (-t) \, dt = -16$

Answer is negative. See from graph, $\underline{v}$ is "flowing" from right to left across C, but the direction of positive flux was fixed.

38.2 Outward flux across a closed curve in the plane

Let C be a piecewise-smooth, simple closed curve. Let $v_1(x, y)$, $v_2(x, y)$ be continuous in the region bounded by C. (Note that these are some of the conditions of Green's theorem!)

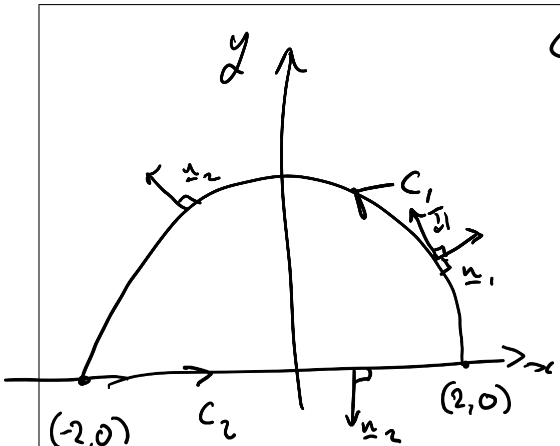
The net outward flux of $\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j}$ across C is given by

$$\text{Net outward flux } = \oint_C \boldsymbol{v} \cdot \boldsymbol{n}\, dS,$$

where $\boldsymbol{n}$ is a unit vector normal to C, directed outward from the region bounded by C.

→ direction of positive flux, i.e. $\underline{n}$

38.2.1 Calculate the outward flux of $\boldsymbol{v} = xy\boldsymbol{i} + xy\boldsymbol{j}$ across the curve from (2,0) to (-2,0) via the semicircle of radius 2 centred at the origin (for $y \geq 0$) followed by the straight line from (-2,0) to (2,0).



$C = C_1 \cup C_2$

$$\text{Flux} = \oint_C \underline{v} \cdot \underline{n}\, dS$$

$$= \int_{C_1} \underline{v} \cdot \underline{n}_1\, dS + \int_{C_2} \underline{v} \cdot \underline{n}_2\, dS$$

C_1: $\underline{r}(t) = 2\cos t\, \underline{i} + 2\sin t\, \underline{j}$, $0 \leq t \leq \pi$

$$\Rightarrow \underline{r}'(t) = -2\sin t\, \underline{i} + 2\cos t\, \underline{j} \quad \left(\underline{T}_1 = \frac{\underline{r}'(t)}{\|\underline{r}'(t)\|}\right)$$

Strategy:

$$\int_{C_1} \underline{v} \cdot \underline{n}_1\, dS = \int \underline{v}(\underline{r}(t)) \cdot \left(\frac{\underline{r}'(t)}{\|\underline{r}'(t)\|} \times \underline{k}\right) \|\underline{r}'(t)\|\, dt$$

$$\underline{r}'(t)\times\underline{k} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ -2\sin t & 2\cos t & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= 2\cos t\,\underline{i} - (-2\sin t)\underline{j} + 0\underline{k}$$

$$\Rightarrow \underline{v}(\underline{r}(t))\cdot(\underline{r}'(t)\times\underline{k})$$

$$= (4\cos t\sin t\,\underline{i} + 4\cos t\sin t\,\underline{j})\cdot(2\cos t\,\underline{i} + 2\sin t\,\underline{j})$$

$$= 8\cos^2 t\sin t + 8\cos t\sin^2 t$$

$$\Rightarrow \int_{C_1} \underline{v}\cdot\underline{n}\,ds = \int_0^{\pi}\left(8\cos^2 t\sin t + 8\cos t\sin^2 t\right)dt = \frac{16}{3}$$

C_2 : on C_2, $\underline{v} = 0$ everywhere,

$$\Rightarrow \int_{C_2} \underline{v}\cdot\underline{n}_2\,ds = 0$$

$$\Rightarrow \oint_C \underline{v}\cdot\underline{n}\,ds = \frac{16}{3} + 0$$

$$= \frac{16}{3}$$

Notes.

(from p270)

Generally, $C : \underline{r}(t)$, $a \le t \le b$

$$\int_C \underline{v} \cdot \underline{n} \, ds = \begin{cases} \int_C \underline{v} \cdot (\underline{T} \times \underline{k}) \, ds = \int \underline{v}(\underline{r}(t)) \cdot \left(\dfrac{\underline{r}'(t)}{\|\underline{r}'(t)\|} \times \underline{k} \right) \|\underline{r}'(t)\| \, dt \\ \text{OR} \\ \int_C \underline{v} \cdot (\underline{k} \times \underline{T}) \, ds = \int \underline{v}(\underline{r}(t)) \cdot (\underline{k} \times \underline{r}'(t)) \, dt \end{cases}$$

depends on direction of positive flux

$\rightarrow$ $\underline{n}$, which should either be understood from the problem, or specified by you

39 Divergence of a vector field (div)

By the end of this section, you should be able to answer the following questions:

- How do you calculate the divergence of a given vector field?
- What is the significance of divergence?
- How does it relate to flux?

In this section we introduce the concept of *divergence* of a vector field.

39.1 Calculating divergence

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

Let

$$\boldsymbol{v}(x,y,z) = v_1(x,y,z)\boldsymbol{i} + v_2(x,y,z)\boldsymbol{j} + v_3(x,y,z)\boldsymbol{k}$$

be a differentiable vector function. Then the function

$$\text{div } \boldsymbol{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} = \nabla \cdot \boldsymbol{v}$$

is called the divergence of $\boldsymbol{v}$. Note div $\boldsymbol{v}$ is a scalar quantity.

Divergence has an analogous definition in two dimensions. For

$$\boldsymbol{F}(x,y) = F_1(x,y)\boldsymbol{i} + F_2(x,y)\boldsymbol{j} \quad \Rightarrow \quad \text{div}\boldsymbol{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y}.$$

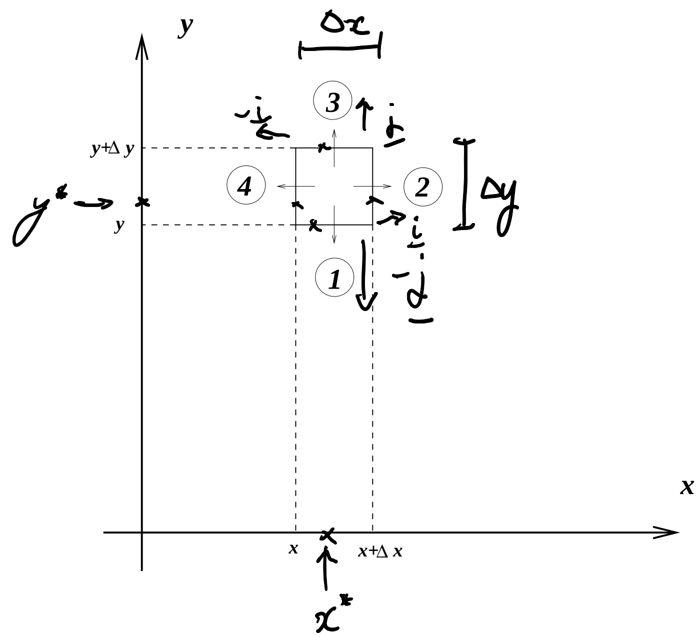
39.1.1 Example: $\boldsymbol{v} = xy^2\boldsymbol{i} + xyz\boldsymbol{j} + yz^2\boldsymbol{k}$. Find div $\boldsymbol{v}$

$$\underline{\nabla}\cdot\underline{v} = \frac{\partial}{\partial x}(xy^2) + \frac{\partial}{\partial y}(xyz) + \frac{\partial}{\partial z}(yz^2)$$

$$= y^2 + xz + 2yz$$

39.2 Understanding div in two dimensions.

Consider the flow of a two dimensional fluid with continuous velocity field $\boldsymbol{v}(x,y) = v_1(x,y)\boldsymbol{i} + v_2(x,y)\boldsymbol{j}$. Our aim is to calculate the outward flux from a small rectangle in the plane of area $\Delta x \Delta y$ as in the diagram below.



We first approximate the flux across each of the four sides of the rectangle. In each case the approximation will be $\boldsymbol{v}\cdot\boldsymbol{n}\Delta S$, where we assume $\boldsymbol{v}$ is constant over each edge. Also let $x^* \in [x, x+\Delta x]$ and $y^* \in [y, y+\Delta y]$ represent chosen points in each interval.

Edge 1: we evaluate $\boldsymbol{v}$ at (x^*, y) and assume it is constant across the entire edge. An outwardly pointing unit normal vector is $-\boldsymbol{j}$.

$$\text{flux } \approx \boldsymbol{v}(x^*, y)\cdot(-\boldsymbol{j})\ \Delta x.$$

Edge 2: we evaluate $\boldsymbol{v}$ at $(x+\Delta x, y^*)$ and assume it is constant across the entire edge. An outwardly pointing unit normal vector is $\boldsymbol{i}$.

$$\text{flux } \approx \boldsymbol{v}(x+\Delta x, y^*)\cdot(\boldsymbol{i})\ \Delta y.$$

Edge 3: we evaluate $\boldsymbol{v}$ at $(x^*, y+\Delta y)$ and assume it is constant across the entire edge. An outwardly pointing unit normal vector is $\boldsymbol{j}$.

$$\text{flux } \approx \boldsymbol{v}(x^*, y+\Delta y)\cdot(\boldsymbol{j})\ \Delta x.$$

Edge 4: we evaluate $\boldsymbol{v}$ at (x, y^*) and assume it is constant across the entire edge. An outwardly pointing unit normal vector is $-\boldsymbol{i}$.

$$\text{flux } \approx \boldsymbol{v}(x, y^*)\cdot(-\boldsymbol{i})\ \Delta y.$$

Combining all four terms gives an approximation to the net outward flux:
net outward flux

$$
\begin{aligned}
&\approx \overbrace{(\boldsymbol{v}(x+\Delta x, y^*) - \boldsymbol{v}(x, y^*)) \cdot \boldsymbol{i}\ \Delta y}^{**} + \overbrace{(\boldsymbol{v}(x^*, y+\Delta y) - \boldsymbol{v}(x^*, y)) \cdot \boldsymbol{j}\ \Delta x} ^{*}\\
&= \left(\frac{\boldsymbol{v}(x+\Delta x, y^*) - \boldsymbol{v}(x, y^*)}{\Delta x}\right) \cdot \boldsymbol{i} \Delta x \Delta y + \left(\frac{\boldsymbol{v}(x^*, y+\Delta y) - \boldsymbol{v}(x^*, y)}{\Delta y}\right) \cdot \boldsymbol{j} \Delta x \Delta y \\
&= \left(\frac{v_1(x+\Delta x, y^*) - v_1(x, y^*)}{\Delta x} + \frac{v_2(x^*, y+\Delta y) - v_2(x^*, y)}{\Delta y}\right) \Delta x \Delta y \\
&\left.\begin{aligned}
&\approx \left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}\right) \Delta x \Delta y. \\
&= \operatorname{div}(\boldsymbol{v}) \Delta x \Delta y.
\end{aligned}\right\} \text{ at } (x^*, y^*)
\end{aligned}
$$

Hence, we have

$$\frac{\text{flux out of a rectangle}}{\text{area of rectangle}} \approx \operatorname{div}(\boldsymbol{v}).$$

If we take the limit as the dimensions of the rectangle appproach 0, we have

$$\operatorname{div}(\boldsymbol{v}) = \lim_{\Delta A \to 0} \frac{\boxed{\text{flux out}} \text{ of } \Delta A}{\Delta A}.$$

In other words, $\operatorname{div}(\boldsymbol{v})$ is the "outward flux density" of $\boldsymbol{v}$ at a given point.

This concept generalises quite naturally to three dimensions:

$$\operatorname{div}(\boldsymbol{v}(x, y, z)) = \lim_{\Delta V \to 0} \frac{\text{flux out of } \Delta V}{\Delta V}.$$

In the context of fluids (our main focus so far) we can say $\operatorname{div}(\boldsymbol{v}(x, y, z))$ measures the tendency of the fluid to "diverge" from the point (x, y, z).

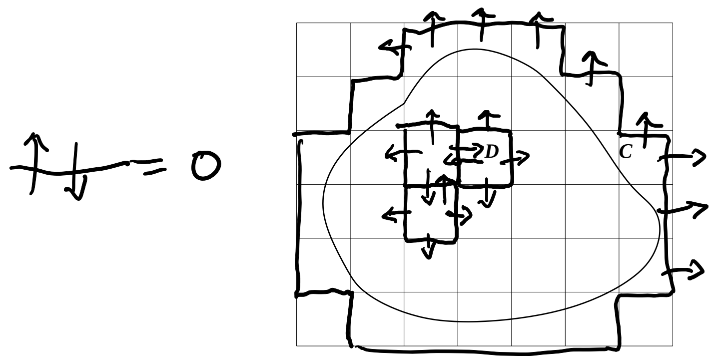
Compare with 2D mass density:

$$\rho(x,y) = \lim_{\Delta A \to 0} \frac{\boxed{\text{mass}} \text{ of } \Delta A}{\Delta A}$$

39.3 Outward flux across a closed curve in the plane (revisited)

One final calculation uses the divergence to calculate the net outward flux of $\boldsymbol{v}$ across a closed curve. We have already seen that we can evaluate this quantity by calculating $\oint_C \boldsymbol{v} \cdot \boldsymbol{n}\, dS$.

Now let D be a region in the x-y plane bounded by a piecewise-smooth, simple closed curve C, which is traversed with D always on the left. Let $v_1(x, y)$, $v_2(x, y)$ have continuous derivatives in D (again the conditions of Green's theorem!).



The only surviving terms are on boundary

By the previous calculation involving divergence, we can also approximate the outward flux from the region by dividing D up into small rectangles and approximating the net outward flux across each rectangle. We know that for one rectangle,

$$\text{outward flux of one rectangle } \approx \operatorname{div}(\boldsymbol{v}(x^*, y^*))\Delta x \Delta y,$$

where (x^*, y^*) is some point inside the rectangle. We repeat this for each rectangle containing part of the region D, so that

$$\text{net outward flux across } C \approx \sum \operatorname{div}(\boldsymbol{v}(x^*, y^*))\Delta x \Delta y.$$

Taking the limit as $\Delta x, \Delta y \to 0$, we have

$$\text{net outward flux across } C = \iint_D \operatorname{div}(\boldsymbol{v}(x, y))\, dA,$$

the double integral of the region D.

To obtain the flux, we integrate the flux density over the region. Compare this with the context of mass density: to obtain the mass, we integrate the mass density over the region.

Finally, the two ways of calculating the same quantity must obviously be equal:

$$\oint_C \boldsymbol{v}(x, y) \cdot \boldsymbol{n}\, dS = \iint_D \operatorname{div}(\boldsymbol{v}(x, y))\, dA.$$

"Flux form of Green's Theorem"

39.4 Relationship to Green's theorem (reading)

We have seen how to evaluate the 2D flux integral:

$$\oint_C \boldsymbol{v} \cdot \boldsymbol{n}\, dS = \int_{t=a}^{t=b} (v_1(t)\dot{y} - v_2(t)\dot{x})\, dt.$$

This can be rewritten as

$$\oint_C \boldsymbol{v} \cdot \boldsymbol{n}\, dS = \oint_C v_1\, dy - v_2\, dx.$$

If we define $F_1(x, y) = -v_2(x, y)$ and $F_2(x, y) = v_1(x, y)$, we then have

$$\oint_C \boldsymbol{v} \cdot \boldsymbol{n}\, dS = \oint_C F_1\, dx + F_2\, dy.$$

We also have

$$\mathrm{div}(\boldsymbol{v}) = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y},$$

so that

$$\iint_D \mathrm{div}(\boldsymbol{v})\, dA = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dA.$$

This tells us that in terms of the new vector field

$$\boldsymbol{F} = -v_2\boldsymbol{i} + v_1\boldsymbol{j} = F_1\boldsymbol{i} + F_2\boldsymbol{j},$$

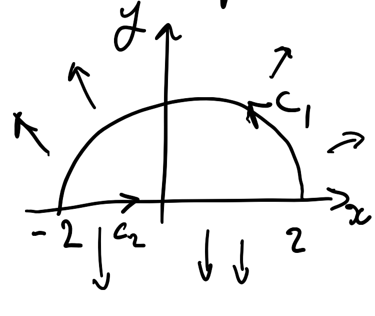
the two ways of calculating flux are given by

$$\iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx\, dy = \oint_C (F_1 dx + F_2 dy).$$

This is none other than Green's theorem. So the flux identity we obtained at the bottom of the previous page is just Green's theorem in disguise. We shall call this the *flux form* of Green's theorem.

39.4.1 **Use the flux form of Green's theorem to calculate the outward flux of $\boldsymbol{v} = xy\boldsymbol{i} + xy\boldsymbol{j}$ across the curve from (2,0) to (-2,0) via the semicircle of radius 2 centred at the origin (for $y \geq 0$) followed by the straight line from (-2,0) to (2,0).**

Same example as p273, 38.2.1



$$\underline{v} = xy\underline{i} + xy\underline{j}$$

$$\Rightarrow \underline{\nabla} \cdot \underline{v} = \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(xy)$$

$$= y + x$$

Net outward flux $= \iint_D \nabla \cdot \underline{v} \, dA$

(by the flux form of Green's Theorem)

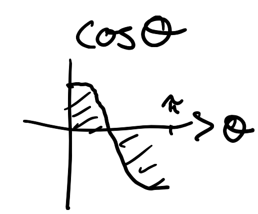
use polar coords $\Rightarrow D = \{(r,\theta) \mid 0 \leq r \leq 2, 0 \leq \theta \leq \pi\}$

$x = r\cos\theta$, $y = r\sin\theta$, Jacobian $= r$

$$\Rightarrow \text{flux} = \int_0^2 \int_0^\pi (r\cos\theta + r\sin\theta) \cdot r \, d\theta \, dr$$

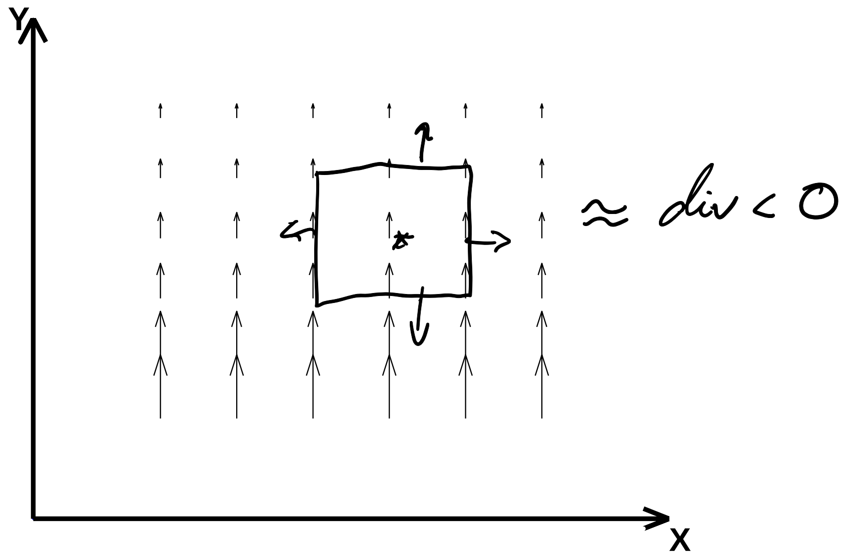
$$= \left(\int_0^2 r^2 \, dr\right)\left(\int_0^\pi (\cos\theta + \sin\theta) \, d\theta\right)$$

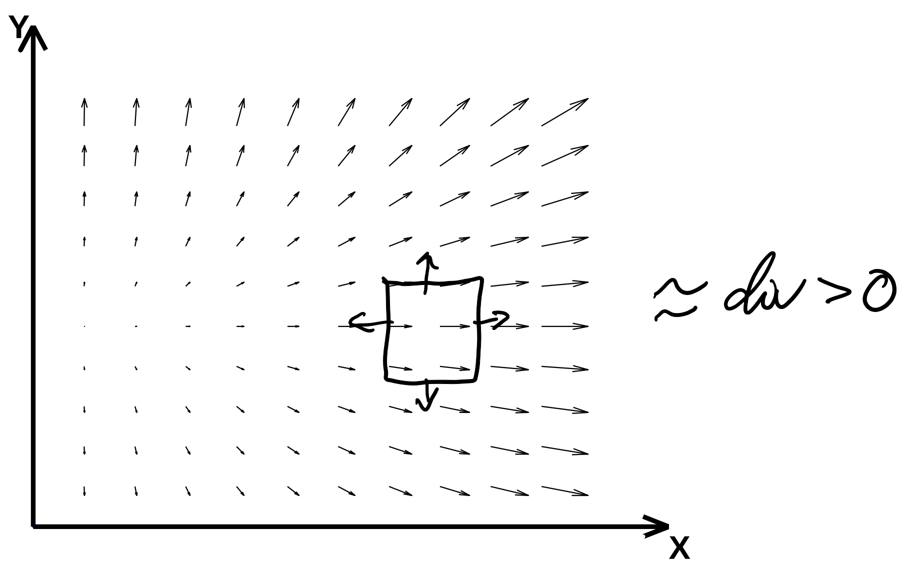
$$= \frac{8}{3} \times 2 = \frac{16}{3}$$

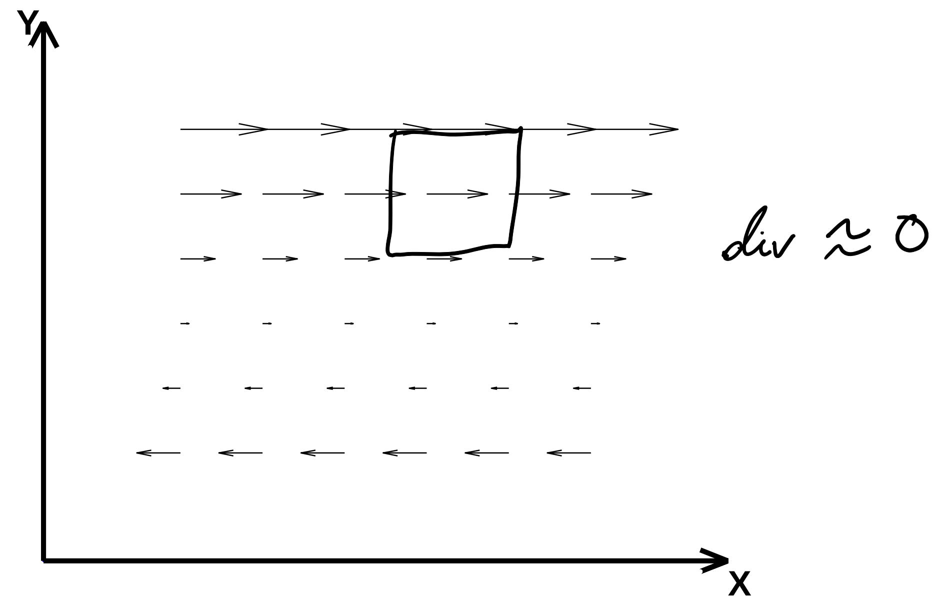


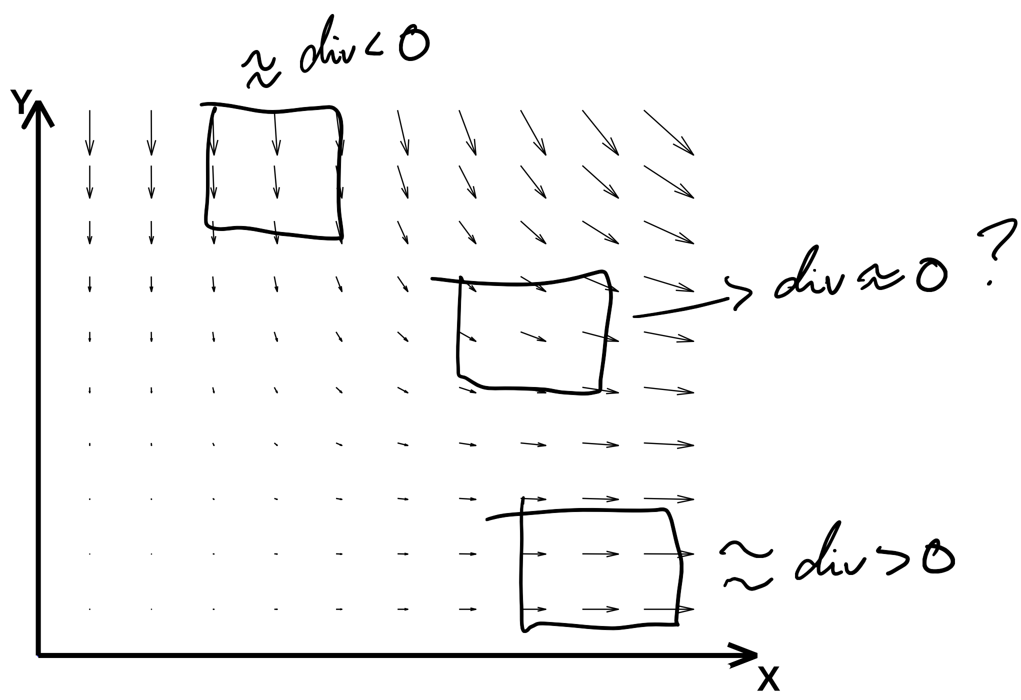
divergence = outward flux density

39.4.2 **For the following graphs of vector fields, determine whether the divergence is positive, negative or zero.**






Y
X
div ≈ 0


≈ div < 0
Y
X
div ≈ 0 ?
≈ div > 0

Notes.

40 Parametrisation of surfaces in $\mathbb{R}^3$

By the end of this section, you should be able to answer the following questions:

- What does it mean to parametrise a surface in $\mathbb{R}^3$?
- How do you parametrise certain surfaces?

40.1 Parametric surfaces

We have already seen two ways of representing a surface in $\mathbb{R}^3$: explicitly as $z = f(x, y)$ or implicitly as $F(x, y, z) = 0$.

Another way of representing a surface S in $\mathbb{R}^3$ is by a parametrisation. This is where the coordinate variables are functions of two parameters u and v:

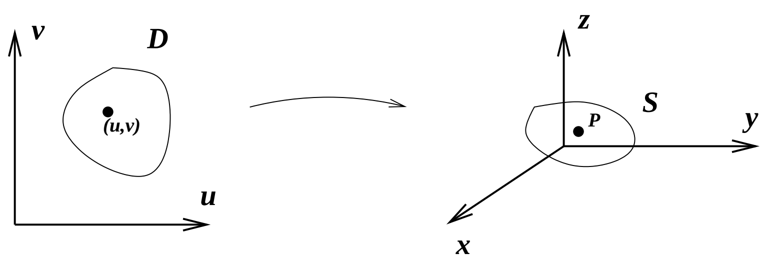
$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)$$

and the vector

$$\boldsymbol{r}(u, v) = x(u, v)\boldsymbol{i} + y(u, v)\boldsymbol{j} + z(u, v)\boldsymbol{k}$$

traces out the surface as u, v vary over some region D in the "u-v plane". So for every point (u, v) in D, there corresponds a point on the surface S.

The following diagram shows the point P on the surface S which corresponds to the point (u, v) in the region D in the u-v plane. As (u, v) moves around all points in D, the point P moves around in S, tracing out the entire surface.



Note that a surface defined explicitly by $z = f(x, y)$ is equivalent to a parametrisation

$$\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + f(x, y)\boldsymbol{k},$$

where we treat the coordinate variables x and y as the parameters. Note that we have not specified any bounds on the variables. Often the challenge is to not only find suitable functions for a parametrisation, but for a finite surface to determine bounds on the parameters.

40.2 Parametrising surfaces using cylindrical and spherical coordinates

We can use our knowledge of cylindrical and spherical coordinates to parametrise certain surfaces with which these coordinates are naturally associated.

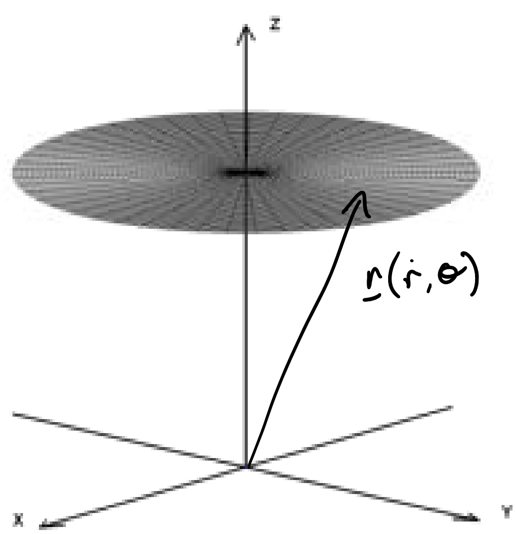
Recall cylindrical coordinates:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z.$$

Setting exactly one of the cylindrical coordinates to a constant value necessarily gives a parametric surface.

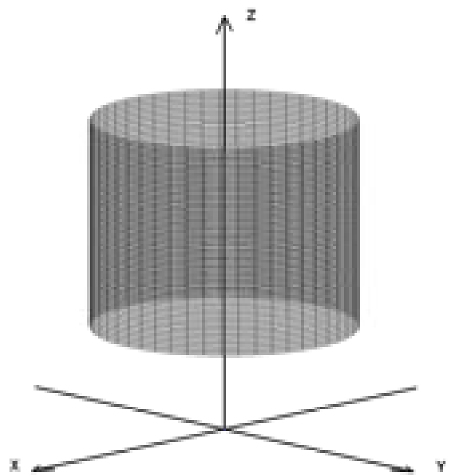
Setting $z = 2$ with $0 \le \theta \le 2\pi$, $0 \le r \le 3$ describes a disc of radius 3, centred at the z axis lying in the plane $z = 2$:

$\underline{r}(r,\theta) = r\cos\theta\,\underline{i} + r\sin\theta\,\underline{j} + 2\,\underline{k}$



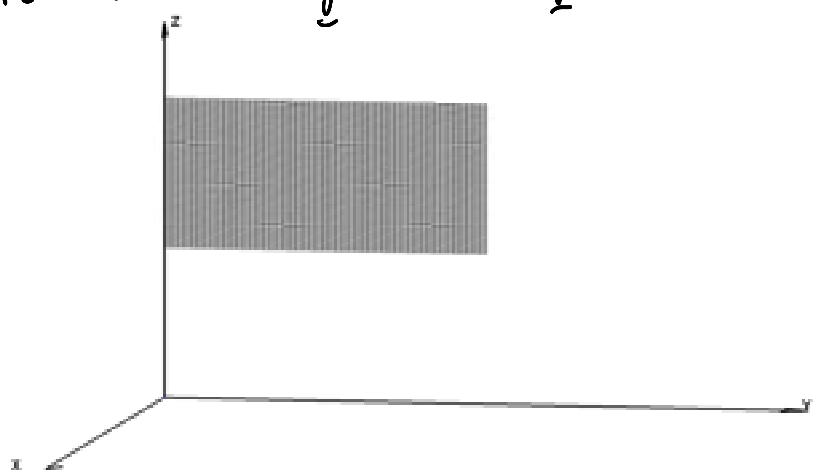
Setting $r = 5$ with $0 \le \theta \le 2\pi$, $1 \le z \le 3$ describes the surface of a cylinder of radius 5 and of height 2 between $z = 1$ and $z = 3$:

$\underline{r}(z,\theta) = 5\cos\theta\,\underline{i} + 5\sin\theta\,\underline{j} + z\,\underline{k}$



Setting $\theta = \pi/2$ with $2 \le z \le 4$, $0 \le r \le 1$ describes a rectangle lying in the y-z plane. Another description of the same surface would be $x = 0$, $\{(y,z) \mid 0 \le y \le 1,\ 2 \le z \le 4\}$:

$\underline{r}(r,z) = r\cos\frac{\pi}{2}\,\underline{i} + r\sin\frac{\pi}{2}\,\underline{j} + z\,\underline{k} = r\,\underline{j} + 2\,\underline{k}$



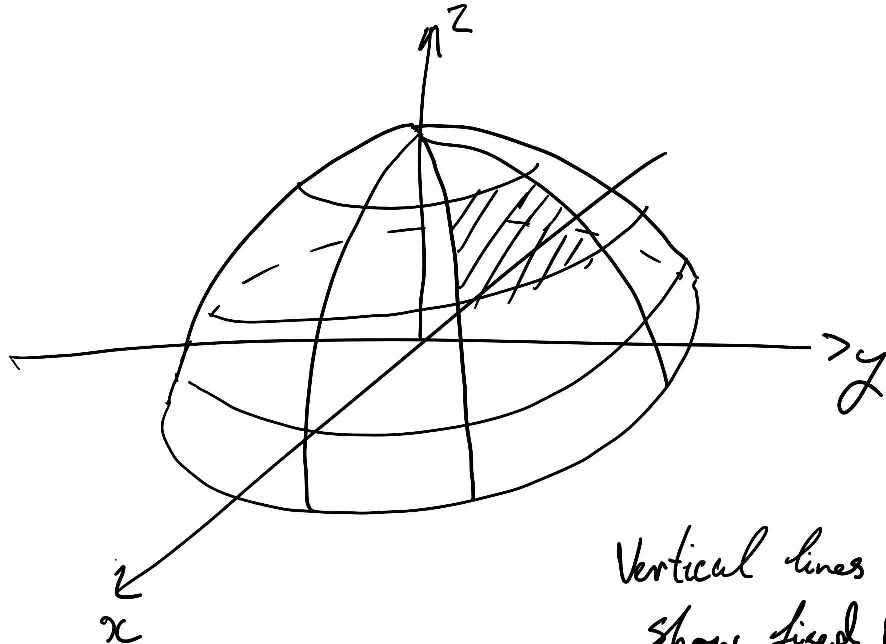
40.2.1 Parametrise the paraboloid $z = 1 - x^2 - y^2$ for $z \geq 0$.

Based on cylindrical coordinates,

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$z = 1 - r^2\cos^2\theta - r^2\sin^2\theta$$
$$= 1 - r^2 \geq 0$$

$$\Rightarrow \quad 0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi$$

$$\underline{r}(r,\theta) = r\cos\theta\,\underline{i} + r\sin\theta\,\underline{j} + (1-r^2)\,\underline{k}$$



Vertical lines show fixed θ

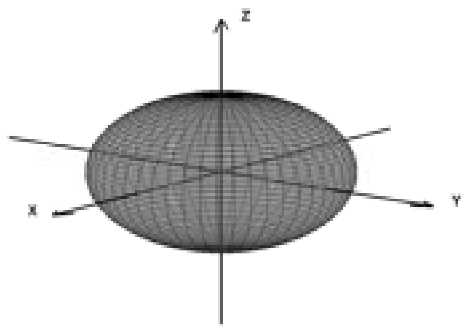
horizontal lines show fixed r (or fixed z)

Recall spherical coordinates: $x = r\cos\theta\sin\phi$, $\quad y = r\sin\theta\sin\phi$, $\quad z = r\cos\phi$.

Setting exactly one of the spherical coordinates to a constant value necessarily gives a parametric surface.

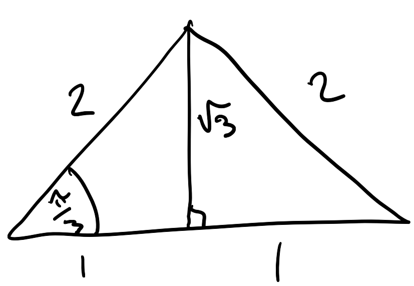
Setting $r = 2$ with $0 \le \theta \le 2\pi$, $0 \le \phi \le \pi$ describes the surface of a sphere of radius 2 centred at the origin:

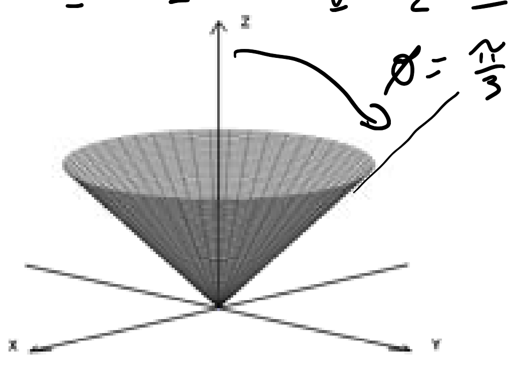
$$\underline{r}(\theta,\phi) = 2\cos\theta\sin\phi\,\underline{i} + 2\sin\theta\sin\phi\,\underline{j} + 2\cos\phi\,\underline{k}$$



Setting $\phi = \pi/3$ with $0 \le r \le 2$, $0 \le \theta \le 2\pi$ describes the open cone with angle $\pi/3$ to the positive z-axis, the "mouth" of which lies on the sphere of radius 2 and with vertex located at the origin:

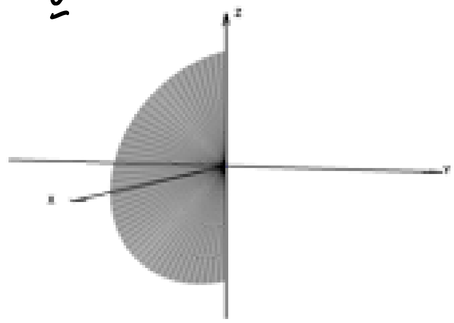
$$\begin{aligned}\underline{r}(r,\theta) &= r\cos\theta\sin\tfrac{\pi}{3}\,\underline{i} + r\sin\theta\sin\tfrac{\pi}{3}\,\underline{j} + r\cos\tfrac{\pi}{3}\,\underline{k} \\ &= \tfrac{\sqrt{3}}{2}r\cos\theta\,\underline{i} + \tfrac{\sqrt{3}}{2}r\sin\theta\,\underline{j} + \tfrac{1}{2}r\,\underline{k}\end{aligned}$$



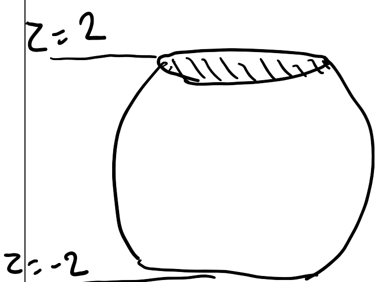


Setting $\theta = 0$ with $0 \le r \le 3$, $0 \le \phi \le \pi$ describes the half disc of radius 3 lying in the x-z plane:

$$\underline{r}(r,\phi) = r\sin\phi\,\underline{i} + 0\,\underline{j} + r\cos\phi\,\underline{k}$$



40.2.2 Parametrise the part of the sphere $x^2+y^2+z^2=16$ that lies between the planes $z=2$ and $z=-2$.



Based on spherical coordinates,

$$\left.\begin{aligned} x &= 4\cos\theta\sin\phi \\ y &= 4\sin\theta\sin\phi \\ z &= 4\cos\phi \end{aligned}\right\} \Rightarrow \underline{r}(\theta,\phi)$$

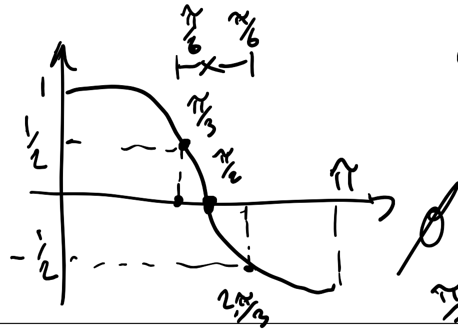
$$\boxed{0 \le \theta \le 2\pi}$$

$$z = 2 \Rightarrow 4\cos\phi = 2 \Rightarrow \cos\phi = \tfrac{1}{2} \Rightarrow \phi = \tfrac{\pi}{3}$$

$$z = -2 \Rightarrow 4\cos\phi = -2 \Rightarrow \cos\phi = -\tfrac{1}{2}$$

$$\cos\left(\tfrac{\pi}{2}+A\right) = -\cos\left(\tfrac{\pi}{2}-A\right)$$

$$\Rightarrow \frac{\pi}{3} \le \phi \le \frac{2\pi}{3}$$



$$\pi/2 + \pi/6 = \frac{2\pi}{3}$$

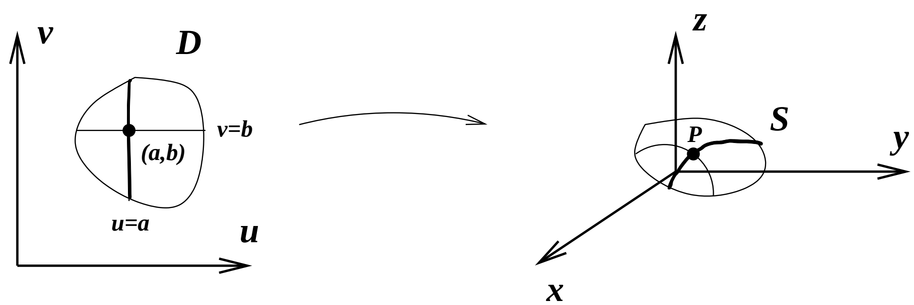
40.3 Tangent planes

Let S be a surface parametrised by

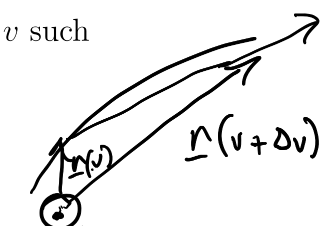
$$\boldsymbol{r}(u, v) = x(u, v)\boldsymbol{i} + y(u, v)\boldsymbol{j} + z(u, v)\boldsymbol{k}.$$

Here we find the tangent plane to S at a point P specified by $\boldsymbol{r}(a, b)$.

There are two important families of curves on S. One where u is a constant, the other where v is a constant. The diagram below shows the relationship between horizontal and vertical lines in D (in the u-v plane) and curves on S.



Setting $u = a$ defines a curve on S parametrised by $\boldsymbol{r}(a, v)$, for all values of v such that (a, v) lies in D. A tangent vector to this curve at P is



$$\boldsymbol{r}_v = \frac{\partial x}{\partial v}(a, b)\boldsymbol{i} + \frac{\partial y}{\partial v}(a, b)\boldsymbol{j} + \frac{\partial z}{\partial v}(a, b)\boldsymbol{k}.$$

Similarly setting $v = b$ defines another curve on S parametrised by $\boldsymbol{r}(u, b)$. A tangent vector to this curve at P is

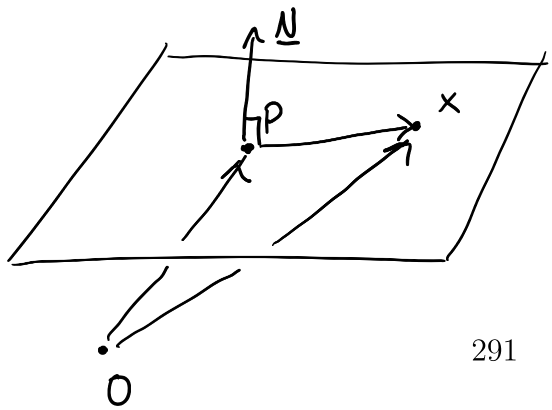
$$\boldsymbol{r}_u = \frac{\partial x}{\partial u}(a, b)\boldsymbol{i} + \frac{\partial y}{\partial u}(a, b)\boldsymbol{j} + \frac{\partial z}{\partial u}(a, b)\boldsymbol{k}.$$

If $\boldsymbol{r}_u$ and $\boldsymbol{r}_v$ are continuous and $\boldsymbol{r}_u \times \boldsymbol{r}_v$ is never $\mathbf{0}$ inside D (we make an exception for points on the boundary of D), we call the surface *smooth* (it has no "kinks").

For a smooth surface, $\boldsymbol{r}_u \times \boldsymbol{r}_v$ is a normal vector at any point inside D. This vector evaluated at $(u, v) = (a, b)$ is also normal to the tangent plane at the point $P = (x(a, b), y(a, b), z(a, b))$.

The equation of the tangent plane at P is given by

$$\overbrace{(\boldsymbol{r}_u(a, b) \times \boldsymbol{r}_v(a, b))}^{\underline{N}} \cdot (\overbrace{(x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})}^{\vec{OX}} - \overbrace{\boldsymbol{r}(a, b)}^{\vec{OP}}) = 0.$$



must have

$\underline{N} \cdot \vec{PX} = 0$

$\Rightarrow \underline{N} \cdot (\vec{OX} - \vec{OP}) = 0$

40.3.1 Find the tangent plane to the surface parametrised by $\boldsymbol{r}(u, v) = u^2\boldsymbol{i} + v^2\boldsymbol{j} + (u + 2v)\boldsymbol{k}$ at the point $(1, 1, 3)$.

$$\underline{r}_u = 2u\underline{i} + 0\underline{j} + 1\underline{k}$$

$$\underline{r}_v = 0\underline{i} + 2v\underline{j} + 2\underline{k}$$

$$\underline{r}_u \times \underline{r}_v = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ 2u & 0 & 1 \\ 0 & 2v & 2 \end{vmatrix}$$

$$= -2v\underline{i} - 4u\underline{j} + 4uv\underline{k}$$

point $(1,1,3)$ corresponds to

$$\left.\begin{array}{l} u^2 = 1 \\ v^2 = 1 \\ u + 2v = 3 \end{array}\right\} \Rightarrow u = 1,\ v = 1$$

$$\underline{r}_u \times \underline{r}_v \Big|_{(1,1)} = -2\underline{i} - 4\underline{j} + 4\underline{k}$$

tangent plane has equ.

$$(\underline{r}_u \times \underline{r}_v) \cdot \left(x\underline{i} + y\underline{j} + z\underline{k} - (\underline{i} + \underline{j} + 3\underline{k})\right) = 0$$

$$\Rightarrow -2(x-1) - 4(y-1) + 4(z-3) = 0$$

Notes.

41 Surface integrals

By the end of this section, you should be able to answer the following questions:

- What is a surface integral?
- How do you calculate the area of a parametric surface?
- How do you use surface integrals in applications such as calculating the mass of a “surface lamina” and finding the average temperature over a surface.

41.1 Area of a parametric surface

Let S be a smooth parametric surface given by

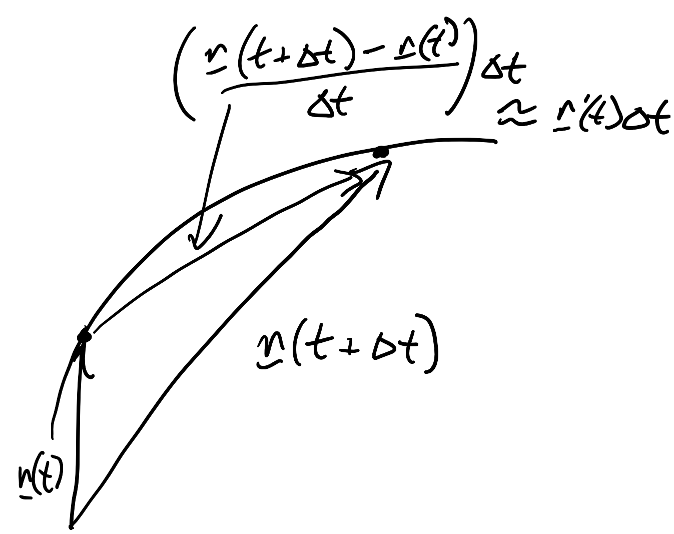
see C-140

$$\boldsymbol{r}(u,v) = x(u,v)\boldsymbol{i} + y(u,v)\boldsymbol{j} + z(u,v)\boldsymbol{k}, \quad , (u,v) \in D$$

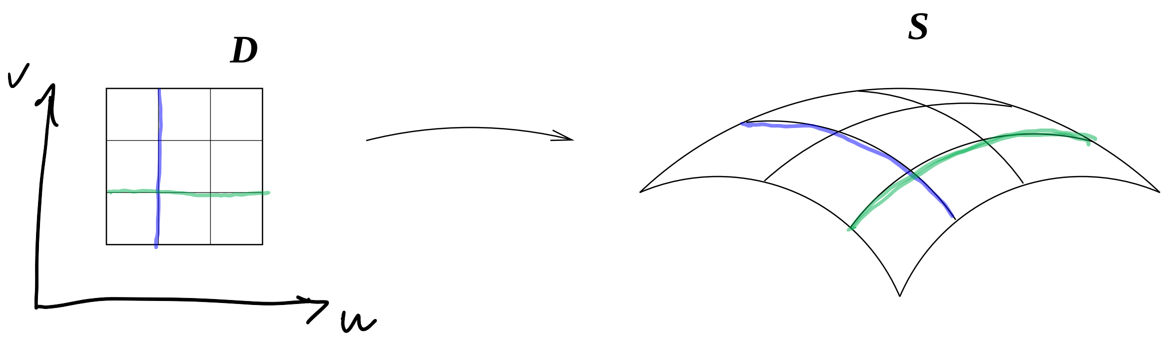
where we assume for simplicity that the parameter domain is a rectangle in the u-v plane. To calculate the area of S, we work through the following steps:

1. Partition S into small patches.
2. Approximate each patch by a parallelogram lying in the tangent plane to the corner of the patch closest to the u-v origin.
3. Calculate the area ΔS of each parallelogram and add them to give an approximation to the area of S.
4. Take the limit as the dimensions of $\Delta S \to 0$ to obtain an exact expression for the area.

Let’s have a closer look at each step.

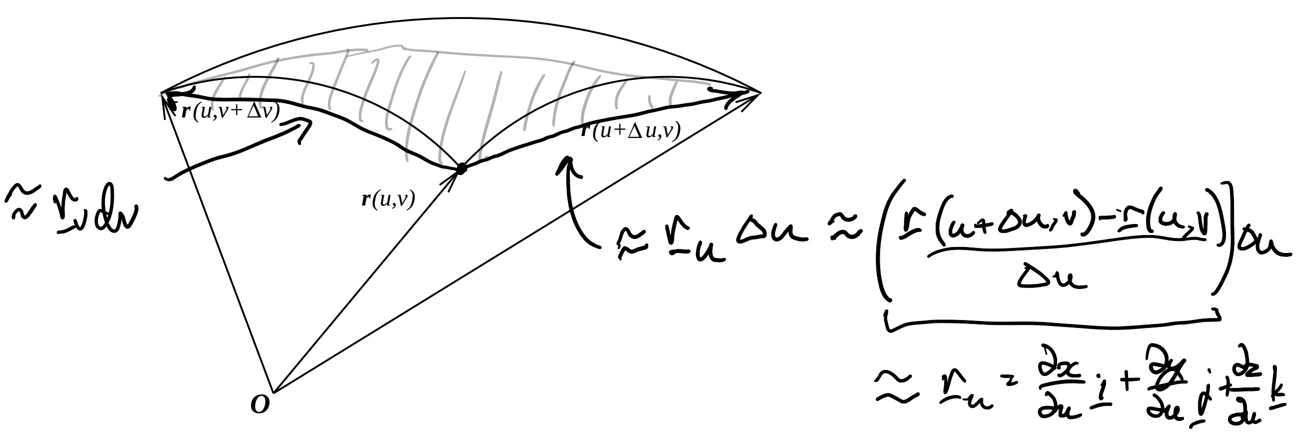


1. A partition of S into patches will correspond to a partition of D (in the u-v plane) into small rectangles.



The dimensions of the rectangles in D will be $\Delta u \Delta v$.

2. Let one of the edges of a single patch be defined from parameter values (u, v) to $(u + \Delta u, v)$.

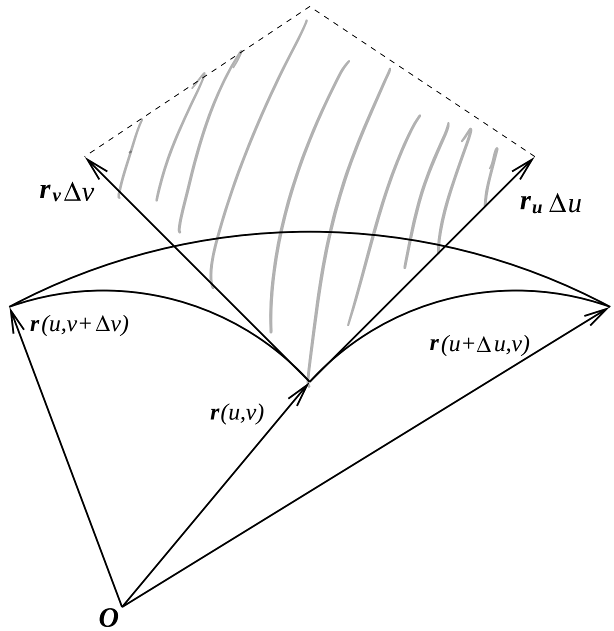


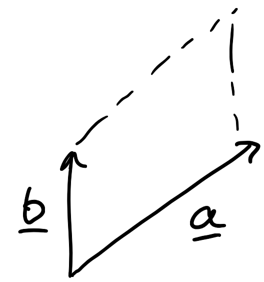
Using Pythagoras' law in three dimensions, we can approximate the length of this edge as

$$
\begin{aligned}
\text{length} &\approx \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \\
&= \left(\sqrt{\left(\frac{\Delta x}{\Delta u}\right)^2 + \left(\frac{\Delta y}{\Delta u}\right)^2 + \left(\frac{\Delta z}{\Delta u}\right)^2} \right) \Delta u \\
&\approx ||\boldsymbol{r}_u|| \Delta u,
\end{aligned}
$$

where in this case we have used $\Delta x = x(u + \Delta u, v) - x(u, v)$ etc (ie. the change is only in u). Similarly, for an edge of patch running from parameter values (u, v) to $(u, v + \Delta v)$ the length of that edge will be approximately $||\boldsymbol{r}_v|| \Delta v$.

At the corner of the patch corresponding to parameter values (u, v), we can define the two vectors $\boldsymbol{r}_u \Delta u$ and $\boldsymbol{r}_v \Delta v$ which form two sides of a parallelogram, the side lengths of which coincide with our approximations to the lengths of the edges of the patch.





area of this parallelogram $= ||\underline{a} \times \underline{b}||$

3. The vector $(\boldsymbol{r}_u \; \Delta u) \times (\boldsymbol{r}_v \; \Delta v)$ is normal to the surface (and hence the tangent plane) at that point. Its magnitude gives the area of the parallelogram we use to approximate the area of the patch ΔS. We then have

$$\Delta S \approx ||\boldsymbol{r}_u \times \boldsymbol{r}_v|| \; \Delta u \; \Delta v.$$

Adding these approximations for each patch in S gives us an approximation to the area of S:

$$\text{area of } S \approx \sum_i \Delta S_i = \sum_i ||\boldsymbol{r}_{u_i} \times \boldsymbol{r}_{v_i}|| \; \Delta u_i \; \Delta v_i.$$

4. Finally taking the limit as $\Delta u, \Delta v \to 0$ we obtain

$$\text{surface area } = \iint_S dS = \iint_D ||\boldsymbol{r}_u \times \boldsymbol{r}_v|| \; du \; dv.$$

defn of surface integral

first example of surface integral

double integral in terms of u, v

41.1.1 Application: find the surface area of the paraboloid $z = 1 - x^2 - y^2$ for $z \geq 0$.

(Strategy: Surface area of $S = \iint_S dS = \iint_D \|\underline{r}_u \times \underline{r}_v\| \, du \, dv$)

From previous chapter ex 40.2.1 on p288

$$\underline{r}(r,\theta) = r\cos\theta\,\underline{i} + r\sin\theta\,\underline{j} + (1-r^2)\,\underline{k}$$

$$D = \{(r,\theta) \mid 0 \leq r \leq 1,\ 0 \leq \theta \leq 2\pi\}$$

$$\underline{r}_r \times \underline{r}_\theta = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \cos\theta & \sin\theta & -2r \\ -r\sin\theta & r\cos\theta & 0 \end{vmatrix} = 2r^2\cos\theta\,\underline{i} - (-2r^2\sin\theta)\,\underline{j} + r\,\underline{k}$$

$$\Rightarrow \|\underline{r}_r \times \underline{r}_\theta\| = \sqrt{4r^4\cos^2\theta + 4r^4\sin^2\theta + r^2}$$

$$= r\sqrt{4r^2+1}$$

$$\Rightarrow \text{surface area} = \int_0^1 \int_0^{2\pi} r\sqrt{4r^2+1}\, d\theta\, dr$$

$$= \ldots = \frac{\pi}{6}(5\sqrt{5} - 1)$$

<u>Note</u>: $\|\underline{r}_r \times \underline{r}_\theta\|$ generalises the Jacobian. In the case that the surface lies in the x-y plane (i.e. $z=0$), it is the Jacobian. (see CH42)

41.2 More on calculating surface integrals, applications

Let $f(x,y,z)$ be a scalar function in $\mathbb{R}^3$. We can define the surface integral of f over a smooth parametric surface S in $\mathbb{R}^3$ as

$$\lim_{\Delta S \to 0} \sum_{(\text{patches})} f(x^*,y^*,z^*)\Delta S = \underbrace{\iint_S f(x,y,z)\ dS}_{\text{definition}} = \underbrace{\iint_D f(\boldsymbol{r}(u,v))||\boldsymbol{r}_u \times \boldsymbol{r}_v||\ du\ dv}_{\text{evaluate}}.$$

Surface integrals and double integrals have similar applications. Indeed, a double integral is merely a special case of a surface integral where the surface lies entirely in the x-y plane.

For example, if a thin sheet has the shape of a surface S and the mass density at the point (x,y,z) is $\rho(x,y,z)$, then the mass of the sheet is given by a surface integral:

$$\text{mass of sheet } = \iint_S \rho(x,y,z)\ dS.$$

Another application is in calculating the average value of a function over a surface. Let S be a smooth surface in $\mathbb{R}^3$. Then the average value of the function $f(x,y,z)$ over that surface is given by

$$\text{average value over surface } = \frac{1}{\text{area of } S} \iint_S f(x,y,z)\ dS.$$

If the surface S is a closed surface, it is convention to write

$$\oiint_S f(x,y,z)\ dS$$

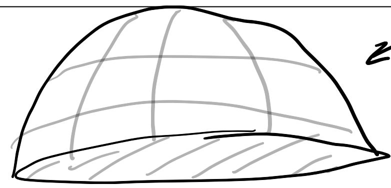
to represent the surface integral.

If S is a finite union of smooth surfaces S_1, S_2,..., S_n that intersect only at their boundaries, then

$$\iint_S f(x,y,z)\ dS = \iint_{S_1} f(x,y,z)\ dS + \iint_{S_2} f(x,y,z)\ dS + \ldots + \iint_{S_n} f(x,y,z)\ dS.$$

Closed surfaces are often unions of smooth surfaces as demonstrated in the following example.

41.2.1 **The function $T(x,y,z) = x^2+y^2+z^2+4$ gives the temperature at any point (x,y,z) on the surface of a solid hemisphere of radius 1 centred at the origin, defined for $z \geq 0$. Find the average temperature over the surface.**



$S = S_1 \cup S_2$

average of $T = \dfrac{\oiint_S T \, dS}{\text{surface area of } S} = \dfrac{\iint_{S_1} T \, dS + \iint_{S_2} T \, dS}{\iint_{S_1} dS + \iint_{S_2} dS}$

S_1: disc in x-y plane

$\underline{r}(r,\theta) = r\cos\theta \, \underline{i} + r\sin\theta \, \underline{j}$, $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$

(polar coords!)

Check $\|\underline{r}_r \times \underline{r}_\theta\| = r$ (which is the Jacobian for polars)

on S_1 $x^2+y^2+z^2+4 = r^2+4$

$$\Rightarrow \iint_{S_1} T \, dS = \int_0^1 \int_0^{2\pi} (r^2+4) \, r \, d\theta \, dr$$

$$= \frac{9}{2}\pi$$

$$\iint_{S_1} dS = \text{area of } S_1 = \pi \times 1^2 = \pi$$

S_2 : half sphere $x^2+y^2+z^2=1$

$$\iint_{S_2} (\underbrace{x^2+y^2+z^2}_{=1}+4)\,dS = \iint 5\,dS$$

$$= 5 \times \iint_{S_2} dS$$

$$= 5 \times (\text{surface area of } S_2)$$

$$= 5 \times \left(\tfrac{1}{2} \times 4\pi \times 1^2\right) = 10\pi$$

& $\iint_{S_2} dS = 2\pi$

$$\Rightarrow \text{average of } T = \frac{\frac{9}{2}\pi + 10\pi}{\pi + 2\pi} = \frac{29}{6}$$

<u>Generally</u>

① Parameterise S with $\underline{r}(u,v)$, $(u,v) \in D$

② Find $\|\underline{r}_u \times \underline{r}_v\|$

③ $$\iint_S f(x,y,z)\,dS = \iint_D f(\underline{r}(u,v))\,\|\underline{r}_u \times \underline{r}_v\|\,du\,dv$$

Notes.

42 Variable transformations in double integrals

Recall that, under a change of variable $x = x(u)$, a definite integral transforms as

$$\int_a^b f(x)\,dx = \int_c^d f(x(u))\frac{dx}{du}\,du,$$

where $a = x(c)$ and $b = x(d)$. We have seen how this generalises to iterated integrals, when changing from cartesian coordinates to polar, cylindrical or spherical coordinates. Here, we discuss how integrals transform under more general coordinate changes. For simplicity, emphasis will be on coordinate changes in two dimensions and on the corresponding iterated (double) integrals.

42.1 Two-variable change of coordinates

(onto)

In two dimensions, a change of coordinates is conveniently described by a surjective transformation

$$T : S \to R, \qquad (u, v) \mapsto (x, y),$$

For every $(x,y) \in R$ there is at least one $(u,v) \in S$ such that $T(u,v) = (x,y)$

where

$$x = x(u, v), \qquad y = y(u, v).$$

We shall assume that T and its first-order partial derivatives are continuous on S. A key property of T is that it maps any boundary of the region S in the u-v plane to a boundary of R in the x-y plane. Such a transformation is particularly useful if we can restrict the coordinates u, v to take values on a rectangle. After possibly applying a second transformation between this rectangle and the **unit square**

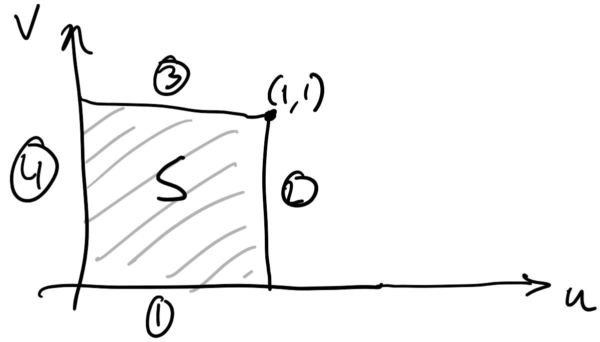
$$\{(u, v) \in \mathbb{R}^2 \,|\, 0 \le u, v \le 1\},$$

this prompts us to focus on the case where S *itself* is the unit square.

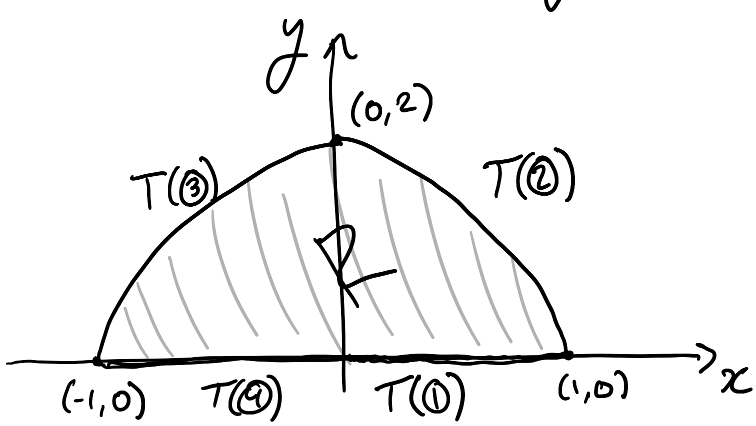
If $T(u, v) = (x, y)$, then the point (x, y) is called the *image* of the point (u, v). We extend this definition to regions.

42.2 **Find the image of the region** $\overbrace{\{(u,v) \mid 0 \le u, v \le 1\}}^{S}$ **under the** $\underbrace{\textbf{transformation}}_{\text{Call it "T"}}$ $x = u^2 - v^2$, $y = 2uv$

$0 \le u \le 1,\ 0 \le v \le 1$



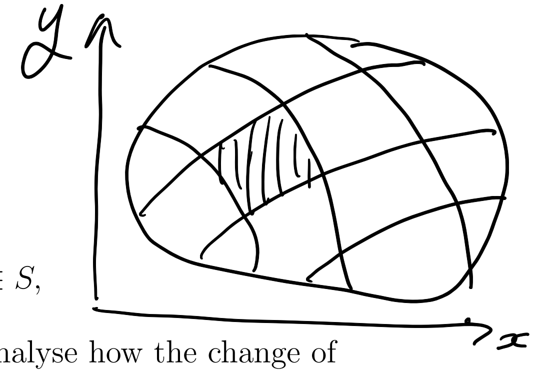
① : $0 \le u \le 1,\ v = 0 \Rightarrow x = u^2,\ y = 0 \Rightarrow 0 \le x \le 1,\ y = 0$



② : $u = 1,\ 0 \le v \le 1 \Rightarrow x = 1 - v^2,\ y = 2v \Rightarrow x = 1 - \frac{y^2}{4},\ 0 \le y \le 2$

③ : $0 \le u \le 1,\ v = 1 \Rightarrow x = u^2 - 1,\ y = 2u \Rightarrow x = \frac{y^2}{4} - 1,\ 0 \le y \le 2$

④ : $u = 0,\ 0 \le v \le 1,\ \Rightarrow x = -v^2,\ y = 0 \Rightarrow -1 \le x \le 0,\ y = 0$



42.3 Jacobian

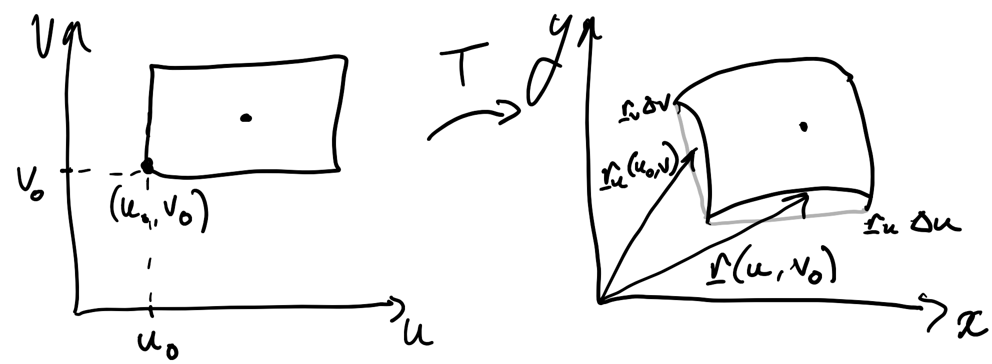
We can view the change of variables as a parameterisation

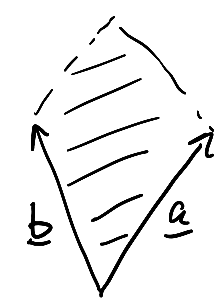
$$\mathbf{r}(u,v) = x(u,v)\mathbf{i} + y(u,v)\mathbf{j}, \qquad (u,v) \in S,$$

of the region R in the x-y plane. This can then be used to analyse how the change of variables affects a double integral over R. To this end, we recall that a double integral over R arises as the limit of a sum over 'larger and larger' families of 'smaller and smaller' patches making up R:

$$\sum_{\text{patches}} f(x^*, y^*)\,\Delta S^* \;\to\; \iint_R f(x,y)\,dS.$$

Here, ΔS^* is the area of the patch containing the point (x^*, y^*).





area $= \|\underline{a} \times \underline{b}\|$

$$\underline{r}_u = \frac{\partial x}{\partial u}\underline{i} + \frac{\partial y}{\partial u}\underline{j}, \quad \underline{r}_v = \frac{\partial x}{\partial v}\underline{i} + \frac{\partial y}{\partial v}\underline{j}$$

Approximate patch as a parallelogram:

$$\Delta S^* \approx \|(\underline{r}_u \Delta u) \times (\underline{r}_v \Delta v)\| = \underbrace{\|\underline{r}_u \times \underline{r}_v\|}_{\text{This factor quantifies the distortion of area between u-v plane and x-y plane}} \Delta u \Delta v$$

$$\underline{r}_u \times \underline{r}_v = \det\begin{pmatrix} \underline{i} & \underline{j} & \underline{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{pmatrix}$$

$$= \det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}\underline{k} \Rightarrow \|\underline{r}_u \times \underline{r}_v\| = \left|\det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}\right|$$

By working out an approximation of the patch area ΔS^*, and expressing it in terms of the u, v variables, we thus arrive at the formula

$$\iint_R f(x,y)\,dx\,dy = \iint_S f(x(u,v), y(u,v))\left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv,$$

where

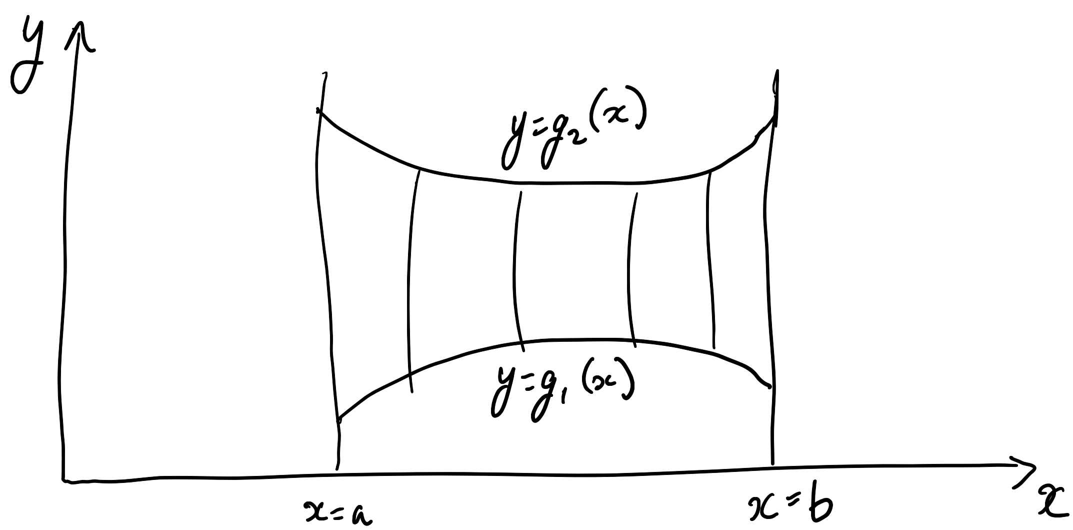
$$\frac{\partial(x,y)}{\partial(u,v)} = \det\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{pmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

is called the **Jacobian** of the transformation T.

Question: Can one always find a variable transformation that maps from a rectangular or a unit square?

Consider a type I region R,

$$R = \{(x,y)\,|\,a \le x \le b,\ g_1(x) \le y \le g_2(x)\}.$$



$$\begin{aligned} \text{Set } x(u,v) &= (1-u)a + ub, && 0 \le u \le 1 \\ y &= (1-v)g_1(x) + vg_2(x), && 0 \le v \le 1 \\ \implies y(u,v) &= (1-v)g_1((1-u)a+ub) + vg_2((1-u)a+ub). \end{aligned}$$

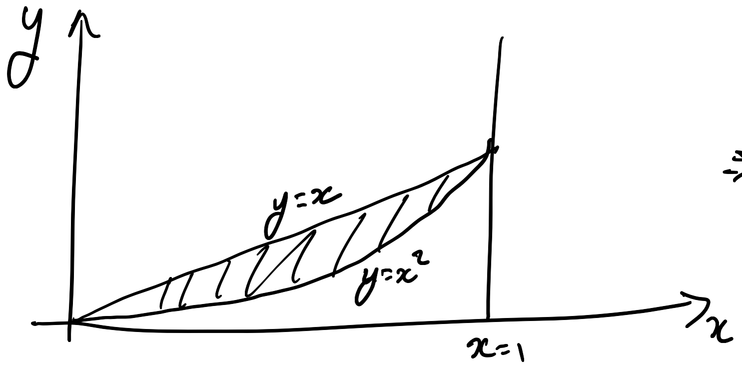
What is the Jacobian then?

Because $\frac{\partial x}{\partial v} = 0$, the corresponding Jacobian is simply given by

$$\begin{aligned} \frac{\partial(x,y)}{\partial(u,v)} &= \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} b-a & 0 \\ \star & g_2((1-u)a+ub) - g_1((1-u)a+ub) \end{vmatrix} \\ &= (b-a)\big[g_2\big(a+(b-a)u\big) - g_1\big(a+(b-a)u\big)\big] \end{aligned}$$

Of course, a similar change of variables applies if R is of type II.

42.4 Example: $R = \{(x,y) \mid 0 \le x \le 1,\, x^2 \le y \le x\}$



$\Rightarrow a = 0$
$b = 1$

set $x = u$ $\quad \boxed{0 \le u \le 1}$

$y = (1-v)x^2 + vx$

$\Rightarrow y = (1-v)u^2 + vu \quad \boxed{0 \le v \le 1}$

$$|\text{Jacobian}| = \left| \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \right|$$

$$= \left| \det \begin{pmatrix} 1 & 0 \\ * & u-u^2 \end{pmatrix} \right|$$

$$= u - u^2 \quad (\ge 0,\ \text{since } 0 \le u \le 1)$$

$$\Rightarrow \iint_R f(x,y)\,dx\,dy = \int_0^1 \int_0^1 f\big(u, (1-v)u^2 + vu\big)(u-u^2)\,du\,dv$$

Notes.

43 Flux integrals and Gauss' divergence theorem

By the end of this section, you should be able to answer the following questions:

- What is a flux integral?
- How do you use a flux integral to calculate the flux of a vector field across a surface?
- What is Gauss' divergence theorem and under what conditions can it be applied?
- How do you apply Gauss' divergence theorem?

We have already been introduced to the idea of flux of a variable vector field across a curve (in $\mathbb{R}^2$) and the flux of a constant vector field across rectangular surfaces (in $\mathbb{R}^3$). In this section we look at calculating the flux across smoothly parametric surfaces.

43.1 Orientable surfaces (unit normal $\underline{n}$)

Let S be a smooth surface. If we can choose a unit vector that is normal to S at every point so that $\boldsymbol{n}$ varies continuously over S, we call S an *orientable* surface. The choice of $\boldsymbol{n}$ provides S with an *orientation*. There are only ever two possible orientations.

An example of an orientable surface is the surface of a sphere. The two possible orientations are out of the sphere or into the sphere.

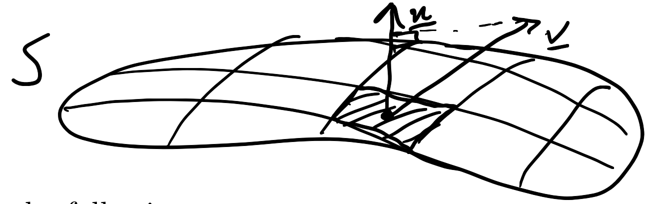
An example of a non-orientable surface is a Möbius strip (see Stewart ed 7. page 1139). (p1167 ed 8)

The orientation of a surface is important when considering flux through that surface. The orientation we choose is always the direction of positive flux.

43.2 The flux integral $\rightarrow$ continuous $\underline{v}$

For a vector field $\boldsymbol{v}(x, y, z)$, we are interested in the flux of $\boldsymbol{v}$ across a smooth orientable parametric surface S in $\mathbb{R}^3$, parametrised by $\boldsymbol{r}(u, v)$, with u and v defined over some domain D. Let $\boldsymbol{n}(u, v)$ be a unit vector normal to the surface S which defines the orientation of the surface (and hence the direction of positive flux).

It would be most convenient to consider the context of fluid flow with $\boldsymbol{v}(x, y, z)$ being the velocity of a fluid at the point (x, y, z).



To calculate the flux through S, we work through the following steps:

1. Partition S into small patches.
2. Approximate each patch by a parallelogram lying in the tangent plane to the corner of the patch closest to the u-v origin.
3. Approximate the flux through each parallelogram of approximate area ΔS and add them to give an approximation to the total flux through S.
4. Take the limit as the dimensions of $\Delta S \to 0$ to obtain an exact expression for the flux.

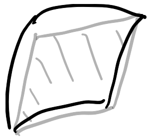
Let's have a closer look at these steps.

1,2. Steps 1 and 2 are exactly the same as steps 1 and 2 on page 295-296 of our calculation of surface area.

3. We approximate the flux through one patch by treating $\boldsymbol{v}$ as constant over the patch (ie. the patch is small enough for this to be a decent approximation). Since we have already approximated the shape of the patch as a parallelogram, we need to work out the flux of a constant vector through a parallelogram.

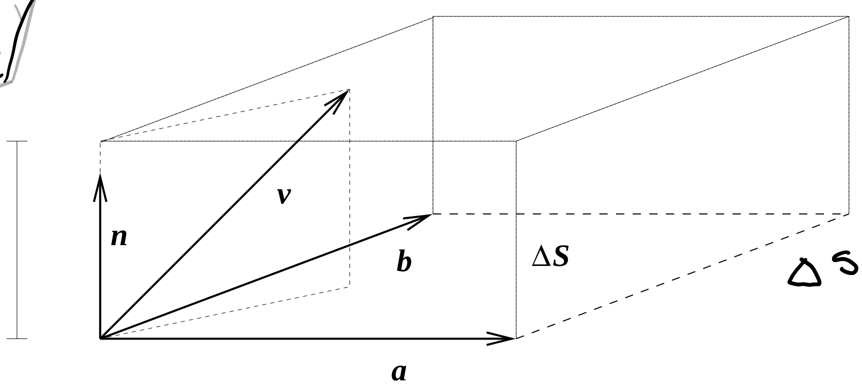
   To this end, consider the parallelogram defined by the two (non-parallel) vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. If we take the area of the patch to be ΔS, it can be seen from the diagram below that the flux (volume per unit time if $\boldsymbol{v}$ is velocity) passing through the parallelogram is

$$\text{flux across parallelogram } \approx \boxed{\boldsymbol{v}\cdot\boldsymbol{n}\,\Delta S.}$$



$\underline{a} = \underline{r}_u \Delta u$

$\underline{b} = \underline{r}_v \Delta v$



$\Delta S \approx \|\underline{r}_u \times \underline{r}_v\| \Delta u \Delta v$

$\underline{n} = \dfrac{\underline{r}_u \times \underline{r}_v}{\|\underline{r}_u \times \underline{r}_v\|}$

c.f. volume of parallelopiped (scalar triple product)

We can take $\boldsymbol{n} = \dfrac{\boldsymbol{a}\times\boldsymbol{b}}{|\boldsymbol{a}\times\boldsymbol{b}|}$, and the area of the parallelogram is $||\boldsymbol{a}\times\boldsymbol{b}|| \approx \Delta S$. We then have

$$\text{flux across parallelogram } \approx \boldsymbol{v}\cdot\frac{\boldsymbol{a}\times\boldsymbol{b}}{||\boldsymbol{a}\times\boldsymbol{b}||}\,||\boldsymbol{a}\times\boldsymbol{b}|| = \boldsymbol{v}\cdot(\boldsymbol{a}\times\boldsymbol{b}).$$

$$\longrightarrow \underline{v}\cdot\underline{n}\,\Delta S = \underline{v}(\underline{r}(u,v))\cdot\left(\frac{\underline{r}_u\times\underline{r}_v}{\|\underline{r}_u\times\underline{r}_v\|}\right)\|\underline{r}_u\times\underline{r}_v\|\,\Delta u\,\Delta v$$

As shown previously, a patch of surface can be approximated by a parallelogram determined by the two vectors $\boldsymbol{r}_u \Delta u$ and $\boldsymbol{r}_v \Delta v$. Hence we have

$$\text{flux across one patch } \approx \boldsymbol{v}\cdot\boldsymbol{n}\ \Delta S = \boldsymbol{v}\cdot(\boldsymbol{r}_u \times \boldsymbol{r}_v)\Delta u \Delta v.$$

* Note that we need to check that the vector $\boldsymbol{r}_u \times \boldsymbol{r}_v$ points in the direction of positive flux. If not, we use $\boldsymbol{r}_v \times \boldsymbol{r}_u$. *

$\underline{r}_v \times \underline{r}_u = -\underline{r}_u \times \underline{r}_v$

Adding these approximations over the entire surface S, we obtain

$$\text{flux across } S \approx \sum_i \boldsymbol{v}_i \cdot \boldsymbol{n}_i\ \Delta S_i = \sum_i \boldsymbol{v}(u_i, v_i)\cdot(\boldsymbol{r}_{u_i} \times \boldsymbol{r}_{v_i})\Delta u_i \Delta v_i.$$

$\Delta S \to 0$

4. To obtain an exact expression for the flux across S we take the limit as $\Delta u, \Delta v \to 0$.

$$\text{flux across } S = \iint_S \boldsymbol{v}\cdot\boldsymbol{n}\ dS = \iint_D \boldsymbol{v}\cdot(\boldsymbol{r}_u \times \boldsymbol{r}_v)\ du\ dv.$$

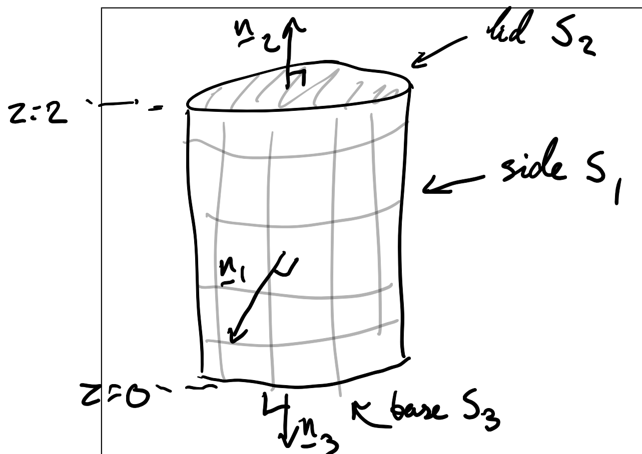
double integral in u, v

defn

This expression is called a *flux integral* and is used to calculate the flux of any vector field across a smooth orientable surface, not just fluids with a given velocity field.

positive direction of flux $\to \underline{n}$

43.2.1 Calculate the net outward flux of $\boldsymbol{F}(x,y,z) = z\boldsymbol{i} + y\boldsymbol{j} + x\boldsymbol{k}$ across the surface of the cylindrical solid given by $\{(x,y,z) \mid x^2 + y^2 \le 1,\ 0 \le z \le 2\}$.



net outward flux

$$= \oiint_S \underline{F}\cdot\underline{n}\ dS$$

$$= \iint_{S_1} \underline{F}\cdot\underline{n}_1\, dS + \iint_{S_2} \underline{F}\cdot\underline{n}_2\, dS + \iint_{S_3} \underline{F}\cdot\underline{n}_3\, dS$$

S_1: Use cylindrical coords,

$$\underline{r}(\theta, z) = \cos\theta\,\underline{i} + \sin\theta\,\underline{j} + z\,\underline{k}\ ,\ 0 \le \theta \le 2\pi,\ 0 \le z \le 2$$

(Strategy: $$\iint_{S_1} \underline{F}\cdot\underline{n}_1\, dS = \iint_{D_1} \underbrace{\underline{F}(\underline{r}[u,v])}_{\underline{F}} \cdot \underbrace{\left(\frac{\underline{r}_u \times \underline{r}_v}{\|\underline{r}_u \times \underline{r}_v\|}\right)}_{\underline{n}} \underbrace{\|\underline{r}_u \times \underline{r}_v\|\, du\, dv}_{dS}$$

$$\underline{r}_z \times \underline{r}_\theta = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ 0 & 0 & 1 \\ -\sin\theta & \cos\theta & 0 \end{vmatrix} = -\cos\theta\,\underline{i} - \sin\theta\,\underline{j}$$

which is directed <u>into</u> the cylinder (take $\theta = 0 \Rightarrow -\cos\theta\,\underline{i} = -\underline{i}$)

$\Rightarrow \underline{r}_\theta \times \underline{r}_z = -(\underline{r}_z \times \underline{r}_\theta) = \cos\theta\,\underline{i} + \sin\theta\,\underline{j}$ is directed out

$$\underline{F}(\underline{r}(\theta, z)) = z\,\underline{i} + \sin\theta\,\underline{j} + \cos\theta\,\underline{k}$$

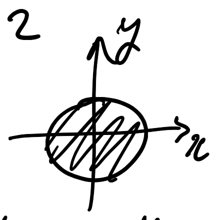
$$\Rightarrow \iint_{S_1} \underline{F}\cdot\underline{n}_1\, dS = \iint_{D_1} \underline{F}(\underline{r}(\theta,z))\cdot(\underline{r}_\theta \times \underline{r}_z)\, d\theta\, dz$$

$$= \int_0^2 \int_0^{2\pi} (z\cos\theta + \sin^2\theta)\, d\theta\, dz = \ldots = 2\pi$$

S_2: disc in plane $z = 2$, $\underline{n}_2 = \underline{k}$

$$\Rightarrow \underline{F}\cdot\underline{n}_2 = x \Rightarrow \iint_{S_2} \underline{F}\cdot\underline{n}_2\, dS = \iint_{S_2} x\, dS = 0 \text{ by symmetry}$$

in plane $z = 2$



(top down view)

flux in = flux out

& S_3: (similar to S_2) $\underline{n}_3 = -\underline{k} \Rightarrow \underline{F}\cdot\underline{n}_3 = -x$

$$\Rightarrow \iint_{S_3} \underline{F}\cdot\underline{n}_3\, dS = 0$$

$\Rightarrow$ net out flux $= 2\pi$

43.3 Gauss' divergence theorem

On page 279 we saw the flux form of Green's theorem:

$$\oint_C \boldsymbol{v}(x,y)\cdot\boldsymbol{n}\; dS = \iint_D \operatorname{div}(\boldsymbol{v}(x,y))\; dA.$$

The left hand side is essentially a flux integral in two dimensions, with $\boldsymbol{n}$ being an outwardly pointing unit normal vector to the curve C. The right hand side was derived from our realisation of the divergence as the "flux density".

It would be natural to ask if it is possible to extend this result to three dimensions.

Given a vector field in three dimensions, $\boldsymbol{F}(x,y,z)$, we have seen that the net outward flux across a closed, smooth, orientable surface S is given by $\displaystyle\oiint_S \boldsymbol{F}\cdot\boldsymbol{n}\; dS$, where $\boldsymbol{n}$ is an outwardly pointing unit normal.

We have also seen that its divergence ($\operatorname{div}\boldsymbol{F}$) can be viewed as the flux density, so

$$\operatorname{div}\boldsymbol{F} = \lim_{\Delta V\to 0}\frac{\text{flux of } \boldsymbol{F} \text{ out of } \Delta V}{\Delta V}.$$

 ΔV

Hence we expect to be able to calculate the net outward flux across a closed, smooth, orientable surface S as the triple integral of the flux density (ie. $\operatorname{div}\boldsymbol{F}$) over the volume enclosed by S.

Indeed, this is true, with $\boldsymbol{F}$ and S subject to certain conditions. The result is known as *Gauss' divergence theorem*:

> Let S be a piecewise smooth, orientable, closed surface enclosing a region V in $\mathbb{R}^3$. Let $\boldsymbol{F}(x,y,z)$ be a vector field whose component functions are continuous and have continuous partial derivatives in V. Then
>
> $$\oiint_S \boldsymbol{F}\cdot\boldsymbol{n}\; dS = \iiint_V \operatorname{div}(\boldsymbol{F})dV,$$
>
> where $\boldsymbol{n}$ is the outwardly directed unit normal to S.

This theorem connects the flux of a vector field out of a volume with the flux through its surface. It says that we can calculate the net outward flux either as a closed surface integral, or as a triple integral.

43.3.1 **Use Gauss' divergence theorem to calculate the net outward flux of $\boldsymbol{F}(x,y,z) = z\boldsymbol{i} + y\boldsymbol{j} + x\boldsymbol{k}$ across the surface of the cylindrical solid given by $\{(x,y,z) \mid x^2+y^2 \le 1,\ 0 \le z \le 2\}$.**

Same problem as 43.2.1 on p310

$$\underline{\nabla} \cdot \underline{F} = \frac{\partial}{\partial x}(z) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(x)$$
$$= 0 + 1 + 0$$

By Gauss' Divergence Theorem,

$$\text{net outward flux} = \oiint \underline{F} \cdot \underline{n}\, dS$$
$$= \iiint_V \underline{\nabla} \cdot \underline{F}\, dV$$
$$= \iiint_V 1\, dV$$
$$= \text{volume of cylinder (height 2, radius 1)}$$
$$= \pi \times 1^2 \times 2 = 2\pi$$

Notes.

44 On proving Gauss' divergence theorem

Recall triple integrals: 6 possible orders of integration, categorise into 3 depending on the inner *variable* i.e.

1. $$\iint_{D_{xy}} \left(\int_{g(x,y)}^{h(x,y)} f(x,y,z)dz \right) dA_{xy}$$

2. $$\iint_{D_{xz}} \left(\int_{g(x,z)}^{h(x,z)} f(x,y,z)dy \right) dA_{xz}$$

3. $$\iint_{D_{yz}} \left(\int_{g(y,z)}^{h(y,z)} f(x,y,z)dx \right) dA_{yz}$$

44.1 Special case proof

Recall the integral identity stated in **Gauss' Divergence Theorem**:

$$\iiint_R \nabla \cdot \mathbf{F} dV = \oiint_{\partial R} \mathbf{F} \cdot \mathbf{n}\, dS,$$

where ∂R denotes boundary of R and $\mathbf{n}$ is the outward unit normal vector of ∂R.

Note: Writing $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$, we have

Here we prove the theorem in the case R can be represented as either a type 1, 2, or 3 region in $\mathbb{R}^3$. This is a special case proof on building block regions in $\mathbb{R}^3$. See Stewart (7 ed) page 1153.

We intend to equate the following terms

$$\iiint_R \frac{\partial F_1}{\partial x}\, dV = \oiint_{\partial R} F_1\, \mathbf{i} \cdot \mathbf{n}\, dS,$$

$$\iiint_R \frac{\partial F_2}{\partial y}\, dV = \oiint_{\partial R} F_2\, \mathbf{j} \cdot \mathbf{n}\, dS,$$

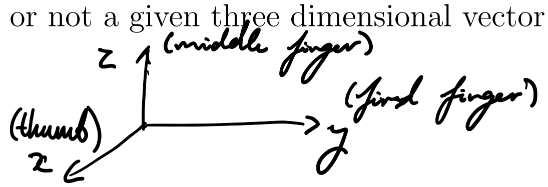
$$\iiint_R \frac{\partial F_3}{\partial z}\, dV = \oiint_{\partial R} F_3\, \mathbf{k} \cdot \mathbf{n}\, dS.$$

Notes.

45 Curl of a vector field

By the end of this section, you should be able to answer the following questions:

- How do you calculate the curl of a given vector field?
- What is the significance of curl?
- How do you test whether or not a given three dimensional vector field is conservative?



45.1 Calculating curl

If (x, y, z) is a right handed Cartesian coordinate system and $\boldsymbol{v}(x, y, z) = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}$ is a differentiable vector field, then define

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

$$\operatorname{curl}(\boldsymbol{v}) = \nabla \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\boldsymbol{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\boldsymbol{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\boldsymbol{k}.$$

* defn

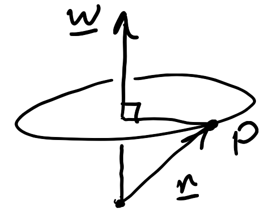
Note that $\operatorname{curl}(\boldsymbol{v})$ is a vector field.

45.1.1 Example: let $\boldsymbol{v} = yz^2\boldsymbol{i} + zx^2\boldsymbol{j} + xy^2\boldsymbol{k}$. Find $\operatorname{curl}(\boldsymbol{v})$.

$$\underline{\nabla} \times \underline{v} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ yz^2 & zx^2 & xy^2 \end{vmatrix}$$

$$= \left(\frac{\partial xy^2}{\partial y} - \frac{\partial zx^2}{\partial z}\right)\underline{i} - \left(\frac{\partial xy^2}{\partial x} - \frac{\partial yz^2}{\partial z}\right)\underline{j} + \left(\frac{\partial zx^2}{\partial x} - \frac{\partial yz^2}{\partial y}\right)\underline{k}$$

$$= (2xy - x^2)\underline{i} + (2yz - y^2)\underline{j} + (2xz - z^2)\underline{k}$$



45.2 Understanding curl

For the rotation of a rigid body about a fixed axis with angular velocity $\boldsymbol{w}$, the velocity at a point P, whose position vector is $\boldsymbol{r}$, is given by $\boldsymbol{v} = \boldsymbol{w} \times \boldsymbol{r}$.

If we choose the axis of rotation to be the z-axis, then $\boldsymbol{w} = \omega \boldsymbol{k}$. Calculate curl($\boldsymbol{v}$).

$$\underline{r} = x\underline{i} + y\underline{j} + z\underline{k}$$

$$\underline{w} = \omega \underline{k}$$

$$\underline{v} = \underline{\omega} \times \underline{r} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix}$$

$$= -\omega y \underline{i} + \omega x \underline{j}$$

$$\Rightarrow \quad \underline{\nabla} \times \underline{v} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ -\omega y & \omega x & 0 \end{vmatrix}$$

$$= 0\underline{i} - 0\underline{j} + (\omega - (-\omega))\underline{k} = 2\omega\underline{k}$$

$$= 2\underline{\omega}$$

$\Rightarrow$ curl($\underline{v}$) is proportional to the angular velocity.

In general, curl($\boldsymbol{v}$) characterises the rotation of a vector field. We will investigate this further in the next section.

(paddle wheel experiment)
go to p283

45.3 Conservative fields revisited

It turns out that the curl of a vector field is exactly what we need to generalise the result at the bottom of page 254 to three dimensions.

Show that if $\boldsymbol{F}$ is a conservative vector field, then curl$\boldsymbol{F} = \mathbf{0}$.

$$\Rightarrow \exists\, f(x,y,z) \text{ s.t. } \underline{F} = \underline{\nabla} f = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j} + \frac{\partial f}{\partial z}\underline{k}$$

$$\Rightarrow \underline{\nabla} \times \underline{F} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix}$$

$$= \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right)\right)\underline{i} - \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right)\right)\underline{j}$$

$$+ \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)\underline{k}$$

$= \underline{0}$ by Clairaut's Theorem

(assuming these partial derivatives are cont.)

i.e. $\underline{\nabla} \times (\underline{\nabla} f) = \underline{0}$

i.e. $\text{curl}(\text{grad}(f)) = \underline{0}$

Indeed, the diagram on page 254 that outlines our logic can be extended directly to the three dimensional case. The only difference is the condition which will serve as our test for conservative fields, namely curl$\boldsymbol{F} = \mathbf{0}$.

The proofs of the links in the diagram for the three dimensional case below are very similar to those used in the two dimensional case. The only detail that is significantly different is showing that if curl$\boldsymbol{F} = \boldsymbol{0}$ then $\displaystyle\oint_C \boldsymbol{F}\cdot d\boldsymbol{r} = 0$. Note also that $\boldsymbol{F}$ must be a vector field defined everywhere in $\mathbb{R}^3$ with continuous partial derivatives. The proof of that part of the diagram requires a generalisation of Green's theorem known as *Stokes' theorem*, which we will investigate in the next section.

$$
\begin{tikzcd}
 & \boldsymbol{F} \text{ conservative} \arrow[dl, shift left] \arrow[dr, "* \checkmark"] & \\
\displaystyle\int_C \boldsymbol{F}\cdot d\boldsymbol{r} \text{ path independent} \arrow[ur, shift left] \arrow[dr, shift left] & & \text{curl}\boldsymbol{F} = \boldsymbol{0} \arrow[dl, "* \text{ Stokes Theorem}"] \\
 & \displaystyle\oint_C \boldsymbol{F}\cdot d\boldsymbol{r} = 0,\ \forall \text{ closed } C \arrow[ul, shift left] &
\end{tikzcd}
$$

The main consequence of this diagram is that we have the following test for a conservative vector field in three dimensions:

$$\boxed{\text{A vector field } \boldsymbol{F} \text{ is conservative if and only if curl}\boldsymbol{F} = \boldsymbol{0}.}$$

45.3.1 Determine whether or not the vector field $\boldsymbol{F} = (1+yz)\boldsymbol{i}+(1+xz)\boldsymbol{j}+xy\boldsymbol{k}$ is conservative.

$$\underline{\nabla} \times \underline{F} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ 1+yz & 1+xz & xy \end{vmatrix}$$

$$= \underline{i}\,(x-x) - \underline{j}\,(y-y) + \underline{k}\,(z-z) = \underline{0}$$

$\therefore$ $\underline{F}$ is conservative.

(in this case $\underline{F} = \underline{\nabla} f$, $f(x,y,z) = x + y + xyz + c$)

Notes.

46 Stokes' theorem

By the end of this section, you should be able to answer the following questions:

- What is Stokes' theorem and under what conditions can it be applied?
- How do you apply Stokes' theorem?
- What is the circulation of a vector field?

46.1 Summary of surfaces and curves

Here we summarise the different types of curves and surfaces which we need to understand Stokes' theorem. Although most of these definitions have already been given, you may find it useful to have all of this information in one place so you can review at a glance.

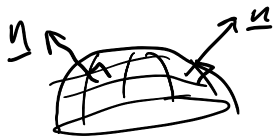
- closed surface → surface of a solid/boundary of a region in $\mathbb{R}^3$
- open surface → not closed.

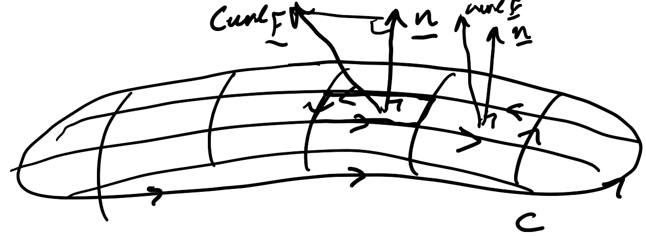
46.1.1 Surfaces

- *Smooth*: the surface normal vector depends continuously on the points on the surface. Parametric surface: $\mathbf{r}_u \times \mathbf{r}_v \neq 0$ inside u-v domain
- *Piecewise smooth*: the surface consists of finitely many smooth surfaces intersecting only at their boundaries.
- *Oriented* (or *orientable*): the direction of the positive normal vector can be continued uniquely and continuously across the whole surface (especially if the surface is piecewise smooth).

46.1.2 Curves

- *Smooth*: the tangent at each point on the curve is unique and varies continuously.
- *Piecewise smooth*: the curve consists of finitely many smooth curves.
- *Simple*: the curve never intersects itself anywhere between its endpoints.
- Boundary curve of an open parametric surface depends on boundary of parameter domain.



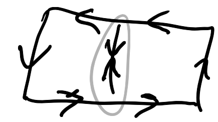


46.2 Stokes' theorem

Let S be a piecewise smooth, orientable surface in $\mathbb{R}^3$ and let the boundary of S be a piecewise smooth, simple, closed curve C. Let $\boldsymbol{F}(x, y, z)$ be a continuous vector function with continuous first partial derivatives in some domain containing S. Then

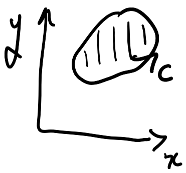
$$\iint_S (\text{curl}\boldsymbol{F}) \cdot \boldsymbol{n} \, dA = \oint_C \boldsymbol{F} \cdot d\boldsymbol{r}, = \oint_C \underline{F} \cdot \underline{T} \, ds,$$

where $\boldsymbol{n}$ is a unit normal vector of S, and the integration around C is taken in the direction using the "right hand rule" with $\boldsymbol{n}$.



T unit tangent vector to C

46.2.1 Relation to Green's theorem

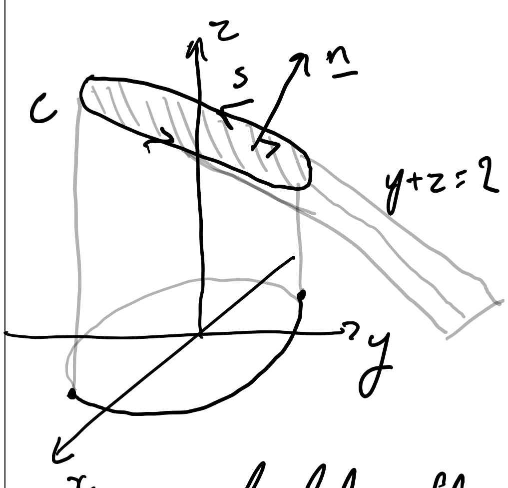


Recall Green's theorem in the plane. It relates a line integral on a boundary to a double integral over a region in the plane. Roughly speaking, Stokes' theorem is a 3-D version of this: it relates a surface integral on a piece of surface (in 3-D) to a line integral on the boundary of the surface.

In fact, note that if the surface is in the x-y plane with $\boldsymbol{n} = \boldsymbol{k}$, Stokes' theorem reduces to Green's theorem, since the $\boldsymbol{k}$ component of curl$\boldsymbol{F}$ is just $\dfrac{\partial F_2}{\partial x} - \dfrac{\partial F_1}{\partial y}$.

46.2.2 Verify Stokes' theorem where C is the curve of intersection of the plane $y+z = 2$ and the cylinder $x^2+y^2 = 1$, oriented counterclockwise when looking from above, and $\boldsymbol{F} = [-y^2, x, z^2]$.

Calculate $\oint_C \underline{F} \cdot \underline{T} \, ds$



In this example, we are free to choose any S with C as its ~~boundary~~.

→ choose S as part of surface $y+z=2$ inside cylinder (shaded ellipse)

calculate $\iint_S (\underline{\nabla} \times \underline{F}) \cdot \underline{n} \, dS$

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

$\boxed{\oint_C}$: Parameterise C $\quad \{(x,y,z) \mid x^2+y^2=1,\ y+z=2\} \to z = 2-y$

$$\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{i} + (2-\sin t)\,\mathbf{k}, \quad 0 \le t \le 2\pi$$

$$\mathbf{r}'(t) = -\sin t\,\mathbf{i} + \cos t\,\mathbf{i} - \cos t\,\mathbf{k}$$

$$\mathbf{F}(\mathbf{r}(t)) = -\sin^2 t\,\mathbf{i} + \cos t\,\mathbf{j} + (4 - 4\sin t + \sin^2 t)\,\mathbf{k}$$

$$\Rightarrow \oint_C \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt$$

$$= \int_0^{2\pi} \left(\cancelto{0}{\sin^3 t} + \cos^2 t - \cancelto{0}{4\cos t} + \cancelto{0}{4\cos t\sin t} - \cancelto{0}{\cos t \sin^2 t}\right) dt$$

(If m or n is odd, then $\int_0^{2\pi} \sin^m t \cos^n t\,dt = 0$)

$$= \int_0^{2\pi} \left(\frac{1}{2} + \frac{1}{2}\cancelto{0}{\cos 2t}\right) dt = \pi$$

$\boxed{\iint_S}$: $\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ -y^2 & x & z^2 \end{vmatrix} = (1+2y)\,\mathbf{k}$

Parameterise S : $\mathbf{r}(r,\theta) = r\cos\theta\,\mathbf{i} + r\sin\theta\,\mathbf{j} + (2 - r\sin\theta)\,\mathbf{k}$

$$0 \le r \le 1, \quad 0 \le \theta \le 2\pi$$

$$\mathbf{r}_r \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos\theta & \sin\theta & -\sin\theta \\ -r\sin\theta & r\cos\theta & -r\cos\theta \end{vmatrix} = r\,\mathbf{j} + r\,\mathbf{k}$$

(orientation is ok!)

$$(\nabla\times\mathbf{F})\cdot(\mathbf{r}_r\times\mathbf{r}_\theta) = r(1+2r\sin\theta)$$

$$\Rightarrow \iint_S (\nabla\times\mathbf{F})\cdot\mathbf{n}\,dS = \iint_D \left((\nabla\times\mathbf{F})(r,\theta)\cdot(\mathbf{r}_r\times\mathbf{r}_\theta)\right) dr\,d\theta$$

$$= \int_0^1\int_0^{2\pi} r\left(1 + 2r\cancelto{0}{\sin\theta}\right) d\theta\,dr = \pi$$

46.3 Further reading: Circulation

Let $\boldsymbol{v}$ represent the velocity field of a fluid and C is a piecewise smooth, simple, closed curve. We have

$$\oint_C \boldsymbol{v} \cdot \boldsymbol{dr} = \oint_C \boldsymbol{v} \cdot \boldsymbol{T} \, dS,$$

where $\boldsymbol{T}$ is a unit tangent vector in the direction of the orientation of the curve. The dot product $\boldsymbol{v} \cdot \boldsymbol{T}$ is the component of $\boldsymbol{v}$ in the direction of $\boldsymbol{T}$ (and hence the curve), so we can interpret $\oint_C \boldsymbol{v} \cdot \boldsymbol{T} \, dS$ as a measure of the tendency of the fluid to move around the curve C. We call this quantity the *circulation* of $\boldsymbol{v}$ around C.

Now define a small circle C_a of radius a about a point P_0, such that the disc S_a enclosed by C_a is normal to the vector $\boldsymbol{n}(P_0)$. Our aim here is to better understand $\mathrm{curl}\boldsymbol{v}$.

Since $\mathrm{curl}\boldsymbol{v}$ is continuous, we approximate $\mathrm{curl}\boldsymbol{v}$ over S_a as $\mathrm{curl}\boldsymbol{v}(P_0)$. Stokes theorem then gives us

$$\begin{aligned}
\oint_{C_a} \boldsymbol{v} \cdot \boldsymbol{dr} &= \iint\limits_{S_a} \mathrm{curl}\boldsymbol{v} \cdot \boldsymbol{n} \, dS \\
&\approx \iint\limits_{S_a} \mathrm{curl}\boldsymbol{v}(P_0) \cdot \boldsymbol{n}(P_0) \, dS \\
&= \mathrm{curl}\boldsymbol{v}(P_0) \cdot \boldsymbol{n}(P_0) \iint\limits_{S_a} dS \\
&= \mathrm{curl}\boldsymbol{v}(P_0) \cdot \boldsymbol{n}(P_0)(\pi a^2) \\
\Rightarrow \quad \mathrm{curl}\boldsymbol{v}(P_0) \cdot \boldsymbol{n}(P_0) &\approx \frac{1}{\pi a^2} \oint_{C_a} \boldsymbol{v} \cdot \boldsymbol{dr} \\
&\approx \frac{\text{circulation around disc}}{\text{area of disc}}.
\end{aligned}$$

This approximation improves as $a \to 0$. Indeed

$$\mathrm{curl}\boldsymbol{v}(P_0) \cdot \boldsymbol{n}(P_0) = \lim_{a \to 0} \frac{1}{\pi a^2} \oint_{C_a} \boldsymbol{v} \cdot \boldsymbol{dr}.$$

Note that this has a maximum value when $\mathrm{curl}\boldsymbol{v}(P_0)$ and $\boldsymbol{n}(P_0)$ have the same direction.

In particular, if we take $\boldsymbol{n}(P_0)$ to be each of the coordinate unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, we have the following: The $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ components of $\mathrm{curl}\boldsymbol{v}(P_0)$ give the *circulation density* at P_0 in planes normal to each of the $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ respectively. The magnitude of $\mathrm{curl}\boldsymbol{v}(P_0)$ gives the maximum circulation density about P_0 in a plane normal to $\mathrm{curl}\boldsymbol{v}(P_0)$.

46.4 Further reading: Curl fields and vector potentials

One immediate consequence is that if there are two different surfaces S_1 and S_2 satisfying the criteria of Stokes' theorem, both with the same boundary curve C, then

$$\iint_{S_1} \mathrm{curl}\boldsymbol{F} \cdot \boldsymbol{n}_1 \; dS = \oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = \iint_{S_2} \mathrm{curl}\boldsymbol{F} \cdot \boldsymbol{n}_2 \; dS.$$

We have that if S is a closed surface satisfying all of the other criteria of Stokes' theorem, and if we define C to be any closed curve lying on S, so that S_1 and S_2 are two open surfaces whose union makes up S and whose common boundary is C, then

$$\begin{aligned} \oint\!\!\oint_S \mathrm{curl}\boldsymbol{F} \cdot \boldsymbol{n} \; dS &= \iint_{S_1} \mathrm{curl}\boldsymbol{F} \cdot \boldsymbol{n}_1 \; dS + \iint_{S_2} \mathrm{curl}\boldsymbol{F} \cdot \boldsymbol{n}_2 \; dS \\ &= \oint_C \boldsymbol{F} \cdot \boldsymbol{dr} + \oint_{-C} \boldsymbol{F} \cdot \boldsymbol{dr} \\ &= \oint_C \boldsymbol{F} \cdot \boldsymbol{dr} - \oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0, \end{aligned}$$

since the orientation of C as a boundary to S_1 will be in the opposite direction to that of S_2.

Let $\boldsymbol{F}$ be a vector field satisfying $\boldsymbol{F} = \mathrm{curl}\boldsymbol{G}$ for some vector field $\boldsymbol{G}$. We call $\boldsymbol{F}$ a *curl field* and $\boldsymbol{G}$ a corresponding *vector potential*.

The above result says that the net outward flux of a curl field across any closed surface is zero.

We can verify that $\mathrm{div}(\mathrm{curl}\boldsymbol{G}) = 0$ for any vector field $\boldsymbol{G}$. Consequently we should not be too surprised by the above result, since Gauss' divergence theorem says that

$$\oint\!\!\oint_S (\mathrm{curl}\boldsymbol{G}) \cdot \boldsymbol{n} \; dS = \iiint_V \mathrm{div}(\mathrm{curl}\boldsymbol{G}) \; dV = 0.$$

In fact, it turns out that we have the following test for curl fields:

Let $\boldsymbol{F}$ be a vector field whose components and their partial derivatives are continuous. If every closed surface in the domain of $\boldsymbol{F}$ only encloses points which are also in the domain of $\boldsymbol{F}$, and if $\mathrm{div}\boldsymbol{F} = 0$, then there exists some $\boldsymbol{G}$ such that $\boldsymbol{F} = \mathrm{curl}\boldsymbol{G}$. That is, $\boldsymbol{F}$ is a curl field.

Notes.