MATH2019 ENGINEERING MATHEMATICS 2E

ADDITIONAL LECTURE NOTES

These notes are intended to give a brief outline of the course to be used as an aid in learning. They are not intended to be a replacement for attendance at lectures, problem classes or tutorials. In particular, they contain few examples. Since examination questions in this course consist mainly of examples, you will seriously compromise your chances of passing by not attending lectures, problem classes and tutorials where many examples will be worked out in detail.

TOPIC 1 – PARTIAL DIFFERENTIATION

Partial derivatives are the derivatives we obtain when we hold constant all but one of the independent variables in a function and differentiate with respect to that variable.

Functions of Two Variables

Suppose $z = f(x, y)$. Define

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \longrightarrow 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$
$$\frac{\partial f}{\partial y} = \lim_{\Delta y \longrightarrow 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}$$

These are both functions of x and y and the usual differentiation rules (product, quotient etc) apply.

Notation

$$\frac{\partial f}{\partial x} = f_x = z_x, \qquad \frac{\partial f}{\partial y} = f_y = z_y$$

i.e. subscripts are used to denote differentiation with respect to the indicated variable. Further

$$\frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0) \qquad \text{means} \quad \frac{\partial f}{\partial x} \quad \text{evaluated at the point} \quad (x_0, y_0).$$

Higher-Order Derivatives

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = (f_x)_x = f_{xx}$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = (f_y)_x = f_{yx}$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = (f_x)_y = f_{xy}$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = (f_y)_y = f_{yy}$$

Mixed Derivatives Theorem (M.D.T.)

The functions f_{xy} (the y derivative of f_x) and f_{yx} (the x derivative of f_y) are obtained by different procedures and so would appear to be different functions. In fact, for almost all functions we meet in practical applications, they are identical because of the M.D.T. which says

If $f(x, y)$ and its partial derivatives f_x, f_y, f_{xy} and f_{yx} are all defined and continuous at all points in a region surrounding the point (a, b) then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

This readily extends to higher order derivatives. In particular, if all derivatives are continuous then $\frac{\partial^{n+m} f}{\partial x^n \partial y^m}$ can be used to denote the partial derivative of f n times with respect to x and m times with respect to y in any order whatsoever.

Chain Rule

Recall that if $u = f(x)$ and $x = g(t)$ then

$$\frac{du}{dt} = \frac{du}{dx}\frac{dx}{dt} = \frac{df}{dx}\frac{dg}{dt} = f'(x)g'(t).$$

This readily generalises. If $w = f(x, y)$ and x and y are themselves differentiable functions of t (e.g. x and y are the coordinates of a moving point and t is time), then

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

Functions of Three or more Variables

If $z = f(x_1, x_2, x_3, \ldots)$ then the partial derivative of z with respect to any one variable (call it x_i) is obtained by holding all the other variables constant and then differentiating with respect to x_i. The mixed derivatives theorem extends to these cases.

Chain Rule

This readily extends to functions of three or more variables. For example, if $w = f(x, y, z)$ and x, y, z are themselves functions of t, then

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

Chain Rules for Functions defined on Surfaces

Suppose $w = f(x, y, z)$ and $x = x(r, s),\ y = y(r, s),\ z = z(r, s)$ (the last three define a surface in 3D space) then

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r}$$

and

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}$$

where $\frac{\partial x}{\partial r}$ etc are taken holding s constant and $\frac{\partial x}{\partial s}$ etc are taken holding r constant.

Multivariable Taylor Series

From first year, we know that if f is a function of a single variable x, then

$$\begin{aligned} f(x) &= f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x-a)^n \end{aligned}$$

This extends to functions of 2 or more variables. We consider only $f(x, y)$. The Taylor Series of $f(x, y)$ about the point (a, b) is

$$f(x,y) = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

$$+\frac{1}{2!}\left\{(x-a)^2\frac{\partial^2 f}{\partial x^2}\,(a,b) + 2(x-a)(y-b)\frac{\partial^2 f}{\partial x \partial y}(a,b)\right.$$

$$\left.+(y-b)^2\,\frac{\partial^2 f}{\partial y^2}(a,b)\right\} + \text{higher-order terms.}$$

Standard Linear Approximation

If $y = f(x)$ then a reasonable approximation when x is close to x_0 is

$$f(x) \simeq f(x_0) + (x - x_0)f'(x_0)$$

obtained by truncating the Taylor series after the linear term. Geometrically, we are approximating the curve $y = f(x)$ for x near x_0 by the tangent to the curve at $(x_0, f(x_0))$.

This idea readily extends to functions of two or more variables. All we do is truncate the Taylor Series after the linear terms. The standard linear approximation of $f(x, y)$ near (x_0, y_0) is therefore $f(x, y) \simeq L(x, y)$ where

$$L(x,y) = f(x_0,y_0) + (x-x_0)f_x(x_0,y_0) + (y-y_0)f_y(x_0,y_0).$$

Geometrically, we are approximating the curved surface $z = f(x, y)$ near (x_0, y_0) by the tangent plane at $(x_0, y_0, f(x_0, y_0))$.

Differentials

The expression

$$df = \frac{\partial f}{\partial x}(x_0,y_0)dx + \frac{\partial f}{\partial y}(x_0,y_0)dy$$

is called the **differential.** You can think of it as the "infinitesimal" change df produced in f by "infinitesimal" changes dx in x and dy in y. It is obtained from $L(x, y)$ by replacing $\Delta x = (x - x_0)$ by dx and $\Delta y = (y - y_0)$ by dy.

Error Estimation

The differential can be used to **estimate** changes in f due to small changes in its arguments. If $\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ then

$$\Delta f \simeq \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y.$$

where $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ are evaluated at (x_0, y_0).

If Δx and Δy are known, we just substitute them in. Usually, however, all we know are **bounds** on Δx and Δy. For example, we may only be able to measure temperature to $\pm 0.01°C$. In that case we have, approximately,

$$|\Delta f| \leq \left|\frac{\partial f}{\partial x}\right| |\Delta x| + \left|\frac{\partial f}{\partial y}\right| |\Delta y|.$$

For functions of several variables $f(x_1, x_2, \cdots, x_n)$

$$\Delta f \simeq \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \Delta x_k$$

and

$$|\Delta f| \leq \sum_{k=1}^{n} \left|\frac{\partial f}{\partial x_k}\right| |\Delta x_k|.$$

Leibniz Rule

$$\frac{d}{dx} \int_{u(x)}^{v(x)} f(x,t)dt = \int_{u(x)}^{v(x)} \frac{\partial f}{\partial x} dt + f(x, v(x)) \frac{dv}{dx} - f(x, u(x)) \frac{du}{dx}.$$

TOPIC 2 – EXTREME VALUES

Extrema for functions of two variables

Suppose we have $f(x, y)$ continuous on some region $\mathcal{R}$. What are the extreme values of $f(x, y)$ (*i.e.* the maxima and minima) and how do we find them?

Definition: The function $f(x, y)$ has a **global maximum** or **absolute maximum** at (x_0, y_0) if $f(x, y) \leq f(x_0, y_0)$ for all (x, y) ε $\mathcal{R}$.

Definition: The function $f(x, y)$ has a **global minimum** or **absolute minimum** at (x_0, y_0) if $f(x, y) \geq f(x_0, y_0)$ for all (x, y) ε $\mathcal{R}$.

Definition: The function $f(x, y)$ has a **local maximum** or **relative maximum** at (x_0, y_0) if $f(x, y) \leq f(x_0, y_0)$ for all (x, y) in some neighbourhood of (x_0, y_0).

Definition: The function $f(x, y)$ has a **local minimum** or **relative minimum** at (x_0, y_0) if $f(x, y) \geq f(x_0, y_0)$ for all (x, y) in some neighbourhood of (x_0, y_0).

Definition: A point (x_0, y_0) ε $\mathcal{R}$ is called a **critical point** of f if

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0,$$

or if f is not differentiable at (x_0, y_0).

Definition: A local maximum or minimum is called an **extreme point** of f. These can only occur at

(i) boundary points of $\mathcal{R}$

(ii) critical points of f

Second Derivative Test

If f and all its first and second partial derivatives are continuous in the neighbourhood of (a, b) and $f_x(a, b) = f_y(a, b) = 0$ then

(i) f has a **local maximum** at (a, b) if $f_{xx} < 0$ and $\mathcal{D} = f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(ii) f has a **local minimum** at (a, b) if $f_{xx} > 0$ and $\mathcal{D} = f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(iii) f has a **saddle point** at (a, b) if $\mathcal{D} = f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a, b).

(iv) If $\mathcal{D} = f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b) the second derivative test is **inconclusive**.

The application of these ideas to practical problems will be illustrated in the lectures. The candidates for maxima and minima are found by looking at i) boundary points, ii) points where one or more of the first partial derivatives fail to exist and iii) points where all the first partial derivatives vanish.

The ideas readily generalise to functions of 3 or more variables although the second derivatives test becomes quite messy.

Extreme values for parameterised curves

To find the extreme values of a function $f(x, y)$ on a curve $x = x(t), y = y(t)$ we find where

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

is **zero**. The extreme values are found at

(i) Critical points (where $f' = 0$ or f' does not exist).

(ii) The endpoints of the parameter domain.

Constrained extrema and Lagrange multipliers

Motivation: Suppose we are asked to find the minimum (or maximum) of a function subject to a constraint.

Example: Find the point $P(x, y, z)$ on the plane $2x + y - z - 5 = 0$ that lies closest to the origin.

This involves finding the minimum of the function

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

subject to the constraint that x, y and z satisfy

$$g(x, y, z) = 2x + y - z - 5 = 0$$

In this simple case, it is easy to use the constraint equation to find an explicit expression for one of the variables (say z) in terms of the other two and to then substitute this into f which thus becomes a function of two variables only and then to find the extrema of f as a function of x and y. For a more complicated constraint, it may not be possible to use the constraint equation to obtain an explicit expression for one of the variables in terms of the others so a more general procedure is required.

The method of Lagrange multipliers

To start off, suppose that $f(x, y)$ and $g(x, y)$ and their first partial derivatives are continuous. To find the local minima and maxima of f subject to the constraint $g(x, y) = 0$ we find the values of x, y and λ that simultaneously satisfy the equations

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0, \quad \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0, \quad \text{together with} \quad g(x, y) = 0 \qquad (1)$$

Justification: We can, in principle, use the equation $g(x, y) = 0$ to write y as a function of x although, as indicated above, this may not be possible in practice. Hence, we may consider f to be a function of a single variable x and look for points where $df/dx = 0$. Let $(x, y) = (a, b)$ be such a point. But, by the chain rule

$$\frac{d\,f}{d\,x} = \frac{\partial f}{\partial x}\frac{d\,x}{d\,x} + \frac{\partial f}{\partial y}\frac{d\,y}{d\,x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{d\,y}{d\,x}$$

Thus

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{d\,y}{d\,x} = 0 \qquad \text{at} \qquad (x, y) = (a, b) \qquad (2)$$

However, since $g(x, y) = 0$, $dg/dx = 0$ everywhere (including (a, b)). Thus

$$\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{d\,y}{d\,x} = 0 \qquad \text{at} \qquad (x, y) = (a, b) \qquad (3)$$

Thus, eliminating dy/dx from (2) and (3) we obtain

$$\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} = 0 \qquad \text{at} \qquad (x, y) = (a, b)$$

which can also be written

$$\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix} = 0 \qquad \text{at} \qquad (x, y) = (a, b)$$

Hence, the rows of this determinant must be linearly dependent, Thus there exists a real number λ such that

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$$

These equations, together with $g(x, y) = 0$, are just (1).

N. B. The quantity λ is called a **Lagrange multiplier** and the method also works if f and g are also functions of z. In that case we have the additional equation $\partial f/\partial z = \lambda \partial g/\partial z$ to solve. It is also possible to introduce the so-called **Lagrangian function**

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

The equations (1) and the constraint $g(x, y) = 0$ are obtained by setting to zero the first partial derivatives of $L(x, y, \lambda)$ with respect to x, y and λ.

Lagrange multipliers with two constraints

Suppose we now want to find the maxima and minima of $f(x, y, z)$ subject to

$$g_1(x, y, z) = 0 \qquad \textbf{and} \qquad g_2(x, y, z) = 0.$$

To do this, we introduce two Lagrange multipliers (one for each constraint) and the **Lagrangian function** for this situation

$$L(x, y, z, \lambda, \mu) = f(x, y, z) - \lambda g_1(x, y, z) - \mu g_2(x, y, z)$$

We now need to find the values of x, y, z, λ and μ which simultaneously satisfy the five equations obtained by setting to zero the partial derivatives of L with respect to x, y, z, λ and μ.

TOPIC 3 - VECTOR FIELD THEORY

Quick Revision of Vector Algebra

Scalars are quantities which have only a magnitude (and sign in some cases) such as temperature, mass, time and speed. Vectors have a magnitude and a direction. We will work only in 3D physical space and use the usual right-handed xyz coordinate system. We denote vector quantities by using **bold** symbols, *e.g.* $\mathbf{a}$. We let $\mathbf{i}, \mathbf{j}, \mathbf{k}$ be the three unit vectors parallel to the x, y and z axes respectively. If a point P has coordinates (p_1, p_2, p_3) and Q has coordinates (q_1, q_2, q_3) then the vector $\overrightarrow{PQ}$ from P to Q has components

$$a_1 = q_1 - p_1, \;\; a_2 = q_2 - p_2, \;\; a_3 = q_3 - p_3$$

and $\mathbf{a} = \overrightarrow{PQ} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = \overrightarrow{OQ} - \overrightarrow{OP}$. The length of $\mathbf{a}$ is $|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$. The position vector of a typical point with coordinates (x, y, z) is usually written $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

Addition etc.

Define $\mathbf{0} = 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}$. This is the vector all of whose components are zero, and is not to be confused with the scalar 0. All the usual rules apply, for example

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

$$
\begin{aligned}
\mathbf{a}+\mathbf{0} &= \mathbf{a} \\
c\mathbf{a} &= ca_1\mathbf{i}+ca_2\mathbf{j}+ca_3\mathbf{k} \\
-\mathbf{a} &= (-1)\mathbf{a} = -a_1\mathbf{i}-a_2\mathbf{j}-a_3\mathbf{k} \\
\mathbf{a}+\mathbf{b} &= \mathbf{b}+\mathbf{a} \\
(\mathbf{a}+\mathbf{b})+\mathbf{c} &= \mathbf{a}+(\mathbf{b}+\mathbf{c}) = \mathbf{a}+\mathbf{b}+\mathbf{c} \\
\mathbf{a}+(-\mathbf{a}) &= \mathbf{0}. \\
c(\mathbf{a}+\mathbf{b}) &= c\mathbf{a}+c\mathbf{b}
\end{aligned}
$$

Inner or Dot or Scalar Product of Vectors

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1+a_2b_2+a_3b_3 = |\mathbf{a}||\mathbf{b}|\cos\gamma$$

where γ $(0 \le \gamma \le \pi)$ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

Then $\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}|^2$ and the dot product of two (non-zero) vectors is 0 if and only if they are orthogonal ($\gamma = \frac{\pi}{2}$).

Observe that $\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1$ and

$$\cos\gamma = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{a_1b_1+a_2b_2+a_3b_3}{\sqrt{a_2^2+a_2^2+a_3^2}\sqrt{b_1^2+b_2^2+b_3^2}}$$

The component of a vector $\mathbf{a}$ in the direction of $\mathbf{b}$ (otherwise known as the projection of $\mathbf{a}$ onto $\mathbf{b}$) is

$$p = |\mathbf{a}|\cos\gamma = \frac{|\mathbf{a}|\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{b}|}.$$

Vector or Cross Product of Vectors

$\mathbf{v} = \mathbf{a}\times\mathbf{b}$ is a vector whose magnitude is $|\mathbf{v}| = |\mathbf{a}||\mathbf{b}|\sin\gamma$ (where γ is the angle $(0 \le \gamma \le \pi)$) between $\mathbf{a}$ and $\mathbf{b}$. The vector $\mathbf{v}$ is perpendicular to the plane defined by $\mathbf{a}$ and $\mathbf{b}$, in such a way that a right-handed screw turn in the direction of $\mathbf{v}$ turns $\mathbf{a}$ into $\mathbf{b}$ through an angle of less than π.

Properties

$$
\begin{aligned}
\mathbf{a}\times\mathbf{b} &= -\mathbf{b}\times\mathbf{a} \\
\mathbf{a}\times\mathbf{a} &= \mathbf{0} \\
\mathbf{a}\times\mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}
\end{aligned}
$$

Triple Scalar Product

$$
\begin{aligned}
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \\
&= \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) \stackrel{\text{def}}{=} [\mathbf{a}\ \mathbf{b}\ \mathbf{c}]
\end{aligned}
$$

Also, $|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$ is the volume of the parallelepiped defined by $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$.

Scalar and Vector Fields

Consider some region Ω of 3-dimensional space. Let a typical point in Ω have coordinates (x, y, z). A **scalar field** is a scalar quantity $f(x, y, z)$ defined on Ω. It often depends on time t as well. The temperature or density in the atmosphere are examples of scalar fields.

A **vector field** is a vector each of whose components is a scalar field. Thus

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

where v_1, v_2 and v_3 all depend on x, y, z (and t usually) is a vector field. Velocity and acceleration in a fluid are good examples. If $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ we sometimes write $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$ to indicate that $\mathbf{v}$ depends on position and time.

Differentiation of Vectors

Suppose $\mathbf{v}$ is a vector field which depends on a single quantity ξ (e.g. $\xi =$ time t). Define

$$\frac{d\mathbf{v}}{d\xi} = \lim_{\Delta\xi \longrightarrow 0} \frac{\mathbf{v}(\xi + \Delta\xi) - \mathbf{v}(\xi)}{\Delta\xi}$$

Thus

$$\frac{d\mathbf{v}}{d\xi} = \mathbf{i}\frac{dv_1}{d\xi} + \mathbf{j}\frac{dv_2}{d\xi} + \mathbf{k}\frac{dv_3}{d\xi}.$$

By applying the product rule to each component, we readily derive:

$$\begin{aligned}
\frac{d}{d\xi}(\rho\mathbf{v}) &= \frac{d\rho}{d\xi}\mathbf{v} \;+\; \rho\frac{d\mathbf{v}}{d\xi} \\
\frac{d}{d\xi}(\mathbf{u} \cdot \mathbf{v}) &= \frac{d\mathbf{u}}{d\xi} \cdot \mathbf{v} \;+\; \mathbf{u} \cdot \frac{d\mathbf{v}}{d\xi} \\
\frac{d}{d\xi}(\mathbf{u} \times \mathbf{v}) &= \frac{d\mathbf{u}}{d\xi} \times \mathbf{v} \;+\; \mathbf{u} \times \frac{d\mathbf{v}}{d\xi}.
\end{aligned}$$

where ρ is a scalar.

Partial Derivatives

If $\mathbf{v}$ depends on several independent variables, the partial derivative of $\mathbf{v}$ with respect to any one of these is obtained by holding all other independent variables constant and differentiating with respect to the nominated variable.

Velocity and Acceleration

Consider a point moving through space. Let its coordinates at time t be $(x(t), y(t), z(t))$. Then its position vector is

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

The **velocity** of the point is

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}.$$

The speed is $|\mathbf{v}| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$ and the acceleration is

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}.$$

Gradient of a Scalar Field

$$\nabla\phi = \text{grad}\,\phi = \frac{\partial\phi}{\partial x}\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j} + \frac{\partial\phi}{\partial z}\mathbf{k}.$$

Directional Derivative

Consider a scalar field ϕ. What is the change in ϕ as we move from $P(x, y, z)$ to $Q(x + \Delta x, y + \Delta y, z + \Delta z)$ keeping t constant?

If Δs is the distance from P to Q then

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2.$$

So, letting $\Delta s \longrightarrow 0$, the vector

$$\hat{\mathbf{u}} = \frac{dx}{ds}\mathbf{i} + \frac{dy}{ds}\mathbf{j} + \frac{dz}{ds}\mathbf{k}$$

is seen to be a **unit vector** in the direction from P to Q.
Now, by our earlier work on increment estimation.

$$\begin{aligned}\Delta\phi = \phi_Q - \phi_P &= \frac{\partial\phi}{\partial x}\Delta x + \frac{\partial\phi}{\partial y}\Delta y + \frac{\partial\phi}{\partial z}\Delta z + \text{smaller terms} \\ &= \left(\frac{\partial\phi}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial\phi}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial\phi}{\partial z}\frac{\Delta z}{\Delta s}\right)\Delta s + \text{smaller terms}\end{aligned}$$

Hence, letting $\Delta s \longrightarrow 0$,

$$\begin{aligned}\frac{d\phi}{ds} &= \frac{\partial\phi}{\partial x}\frac{dx}{ds} + \frac{\partial\phi}{\partial y}\frac{dy}{ds} + \frac{\partial\phi}{\partial z}\frac{dz}{ds} \\ &= \nabla\phi \cdot \hat{\mathbf{u}}\end{aligned}$$

Now, the rate of change with respect to distance in the direction specified by the unit vector $\hat{\mathbf{u}}$ is called the **directional derivative** and is denoted by $D_{\hat{\mathbf{u}}}\phi$. We have shown that

$$D_{\hat{\mathbf{u}}}\phi = \nabla\phi \cdot \hat{\mathbf{u}}.$$

(N.B. $\hat{\mathbf{u}}$ is a vector of unit length).

Now if $\theta (0 \leq \theta \leq \pi)$ is the angle between $\nabla\phi$ and $\hat{\mathbf{u}}$, $\frac{d\phi}{ds} = |\nabla\phi| \cos\theta$ since $|\hat{\mathbf{u}}| = 1$. Thus, $\frac{d\phi}{ds}$ has the maximum value $|\nabla\phi|$ when $\theta = 0$ (i.e. $\hat{\mathbf{u}}$ is in the direction of $\nabla\phi$) and the minimum value $-|\nabla\phi|$ (when $\hat{\mathbf{u}}$ is in the direction of $-\nabla\phi$.)

Normal to a Surface

Next, consider a level surface $\phi = C$. This defines a surface $\mathcal{S}$ in space. For example, meteorologists talk about surfaces of constant pressure such as the 500 millibar surface. Let P and Q be any two nearby points on $\mathcal{S}$. Then $\phi_P = \phi_Q = C$, i.e. $\frac{d\phi}{ds} = 0$ at P in any direction tangential to $\mathcal{S}$ at P. Thus $\nabla\phi$ at P is orthogonal to $\overrightarrow{PQ}$.

Since this holds for any point Q close to P, i.e. is independent of the direction from P to Q, it follows that $\nabla\phi$ at P must be orthogonal to the level surface $\phi = C$. A **unit** normal is $\frac{\nabla\phi}{|\nabla\phi|}$.

Equation of Tangent Plane

If P has coordinates (x_0, y_0, z_0) and (x, y, z) is any point in the plane tangent to $\mathcal{S}$ at P then $\nabla\phi$ is normal to this tangent plane which therefore has equation

$$\nabla\phi \cdot [(x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}] = 0$$

where $\nabla\phi$ is evaluated at P.

Divergence of a Vector Field

If $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ then $\nabla \cdot \mathbf{F} = \operatorname{div} \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$. It may be regarded at the dot product of the vector differential operator $\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$ and the vector $\mathbf{F}$. It is just a scalar.

N.B. $\nabla \cdot \mathbf{F} \neq \mathbf{F} \cdot \nabla$. The latter is the differential operator

$$\mathbf{F} \cdot \nabla = F_1\frac{\partial}{\partial x} + F_2\frac{\partial}{\partial y} + F_3\frac{\partial}{\partial z}$$

Theorem $\nabla \cdot (\phi\mathbf{v}) = \phi(\nabla \cdot \mathbf{v}) + \mathbf{v} \cdot (\nabla\phi)$.

Proof

$$\begin{aligned} LHS &= \frac{\partial}{\partial x}(\phi v_1) + \frac{\partial}{\partial y}(\phi v_2) + \frac{\partial}{\partial z}(\phi v_3) \\ &= \phi\frac{\partial v_1}{\partial x} + \frac{\partial \phi}{\partial x}v_1 + \phi\frac{\partial v_2}{\partial y} + \frac{\partial \phi}{\partial y}v_2 + \phi\frac{\partial v_3}{\partial z} + \frac{\partial \phi}{\partial z}v_3 \\ &= \phi\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right) + v_1\frac{\partial \phi}{\partial x} + v_2\frac{\partial \phi}{\partial y} + v_3\frac{\partial \phi}{\partial z} \\ &= \phi\nabla \cdot \mathbf{v} + \mathbf{v} \cdot \nabla\phi = RHS \end{aligned}$$

Q.E.D.

Laplacian

$$\nabla \cdot (\nabla\phi) = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} \stackrel{\text{def}}{=} \nabla^2\phi.$$

Curl of a Vector Field

$$\nabla \times \mathbf{F} = \operatorname{curl} \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$$

$$= \mathbf{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

Theorem $\nabla \times (\nabla\phi) = \mathbf{0}$.

Proof

$$L.H.S = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\phi}{\partial x} & \frac{\partial\phi}{\partial y} & \frac{\partial\phi}{\partial z} \end{vmatrix}$$

$$= \mathbf{i}\left(\frac{\partial^2\phi}{\partial y\partial z} - \frac{\partial^2\phi}{\partial z\partial y}\right) + \mathbf{j}\left(\frac{\partial^2\phi}{\partial z\partial x} - \frac{\partial^2\phi}{\partial x\partial z}\right) + \mathbf{k}\left(\frac{\partial^2\phi}{\partial x\partial y} - \frac{\partial^2\phi}{\partial y\partial x}\right)$$

$$= 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \mathbf{0} = R.H.S.$$

Vector fields $\mathbf{F}$ for which $\nabla \times \mathbf{F} = \mathbf{0}$ are called **irrotational** or **conservative** .

Line Integrals

These are used for calculating, for example, the work done in moving a particle in a force field.

Consider a vector field $\mathbf{F}(\mathbf{r})$ and a curve $\mathcal{C}$ from point A to point B. Let the equation of $\mathcal{C}$ be $\mathbf{r} = \mathbf{r}(t)$ where t is parameter. Let $t = a$ at point A and $t = b$ at point B. We define

$$\int_{\mathcal{C}} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \frac{d\mathbf{r}}{dt} dt$$

In terms of components, this can be written

$$\int_{\mathcal{C}} (F_1 dx + F_2 dy + F_3 dz) = \int_a^b \left(F_1\frac{dx}{dt} + F_2\frac{dy}{dt} + F_3\frac{dz}{dt}\right) dt$$

where dx, dy, dz are displacements measured along $\mathcal{C}$ and $\mathbf{F}$ is evaluated on $\mathcal{C}$. In general, this integral depends not only on $\mathbf{F}$ but also on the **path** we take between A and B. If A and B coincide, we are integrating around a **closed** curve. This is denoted by

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}.$$

Work If $\mathbf{F}$ is a force field, $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$ is the work done in moving from A to B along $\mathcal{C}$.

Simple properties

(i) If k is a constant $\int_{\mathcal{C}} (k\mathbf{F}) \cdot d\mathbf{r} = k \int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$

(ii) $\int_{\mathcal{C}} (\mathbf{F} + \mathbf{G}) \cdot d\mathbf{r} = \int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} + \int_{\mathcal{C}} \mathbf{G} \cdot d\mathbf{r}$.

(iii) $\int_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} = -\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$.

where $\mathcal{C}_1$ is the same curve as $\mathcal{C}$ except that we start at B and finish at A, i.e. reversing the order of integration changes the sign of a line integral.

(iv) $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \int_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} + \int_{\mathcal{C}_2} \mathbf{F} \cdot d\mathbf{r}$

where $\mathcal{C}_1$ is the curve from A to B, $\mathcal{C}_2$ the curve from B to C and $\mathcal{C}$ the curve from A to C following the same path.

TOPIC 4 – DOUBLE INTEGRALS

Reminders on Coordinate Systems

1. Cartesian Coordinates: (x, y, z)

We often use $3D$ Cartesian coordinates xyz. When we do so, the system is always taken as being **right-handed**. By this we mean that a right-handed turn through $90°$ along Ox turns Oy into Oz.

2. Cylindrical Polar Coordinates: (r, θ, z)

$$\left.\begin{array}{l} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{array}\right\} \begin{array}{l} r \geq 0 \\ 0 \leq \theta < 2\pi \end{array}$$

Definition of a Double Integral

Let Ω be some sub-region of the xy plane and $f(x, y)$ be a function defined at all points of Ω. We divide Ω up into N non-overlapping sub-regions $\Delta\Omega_1, \Delta\Omega_2, \cdots, \Delta\Omega_N$ (so $\Omega = \bigcup_{i=1}^{N} \Delta\Omega_i$). Let $\Delta\Omega_j$ have area ΔA_j and let (ξ_j, η_j) be a typical point in $\Delta\Omega_j$. Form the sum

$$I_N = \sum_{j=1}^{N} f(\xi_j, \eta_j)\Delta A_j.$$

Now let $N \longrightarrow \infty$ in such a way that the largest linear dimension of each $\Delta\Omega_j \longrightarrow 0$ as $N \longrightarrow \infty$. Then if I_N tends to some limit as $N \longrightarrow \infty$ we define

$$I = \int_{\Omega} f(x, y)dA = \lim_{N \longrightarrow \infty} I_N.$$

If f is continuous on Ω, this limit will exist.

Interpretations

1. If $f(x, y) = 1$ then $\int_{\Omega} dA$ is the area of Ω.
2. If $f \geq 0$ on Ω, then I is the volume of the solid whose base is Ω in the $x - y$ plane and whose top has equation $z = f(x, y)$

Simple Properties

1. If k is a constant
$$\int_\Omega kf(x,y)dA = k\int_\Omega f(x,y)dA.$$

2. $\int_\Omega (f(x,y)+g(x,y))dA = \int_\Omega f(x,y)dA + \int_\Omega g(x,y)dA.$

3. If $f(x,y) \geq 0$ on Ω then
$$\int_\Omega f(x,y)dA \geq 0$$

4. If Ω_1 and Ω_2 are non-overlapping regions and $\Omega = \Omega_1 \cup \Omega_2$ then
$$\int_\Omega f(x,y)dA = \int_{\Omega_1} f(x,y)dA + \int_{\Omega_2} f(x,y)dA.$$

Evaluation as Repeated Single Integrals

The above definition is not useful for practical evaluation. Instead we almost always evaluate double integrals as repeated single integrals. In xy cartesian coordinates
$$dA = dx\,dy = dy\,dx.$$
The simplest case is when Ω is the rectangle defined by $a \leq x \leq b,\ c \leq y \leq d$. Then
$$\int_\Omega f(x,y)dA = \int_c^d \int_a^b f(x,y)dx\,dy = \int_a^b \int_c^d f(x,y)dy\,dx.$$
The integral can be evaluated in either way.

In the first way, we fix y between c and d and integrate with respect to x from a to b. In this inner integral, $\int_a^b f(x,y)dx$, y is held constant and the result is a function of y only. We then integrate this from $y = c$ to $y = d$. In the second form, we do analogous things with the rôles of x and y interchanged. The two results **must** give the same answer, so it is only necessary to do it one way.

Non Rectangular Regions

1. If Ω is defined by $a \leq x \leq b,\ f_1(x) \leq y \leq f_2(x)$ then
$$\int_\Omega f(x,y)dA = \int_a^b \int_{f_1(x)}^{f_2(x)} f(x,y)dy\,dx$$

2. If Ω is defined by $c \leq y \leq d,\ g_1(y) \leq x \leq g_2(y)$ then

$$\int_\Omega f(x,y)dA = \int_c^d \int_{g_1(y)}^{g_2(y)} f(x,y)dx\,dy$$

The evaluation of such integrals is best understood by examples as will be given in lectures. It is important that you always draw a sketch of the region of integration in order to get the limits correct. Regions of more complicated shape may need to be partitioned into a collection of subregions of type 1 or 2 in order to evaluate the integral.

Reversal of Order

Since the value of a double integral is independent of the order in which we do the integration, it is sometimes easier to reverse the originally specified order of integration. However, we must then be careful to choose the new limits of integration so as to cover the same region Ω as the original integral. Again, this technique is best illustrated by the examples which will be given in lectures.

Density, Mass, Centre of Mass

Consider a lamina in the $x-y$ plane (e.g. a piece of sheet metal). If a small element of area ΔA has mass Δm, we define the surface density (areal density or superficial density are terms also used) as

$$\delta(x,y) = \lim_{\Delta A \longrightarrow 0} \frac{\Delta m}{\Delta A}.$$

Due to varying composition, this may be a function of x and y. A small element $d\Omega$ with area dA then has mass $dm = \delta(x,y)dA$.

If the lamina occupies the region Ω then its total mass will be

$$M = \int_\Omega dm = \int_\Omega \delta(x,y)dA.$$

Consider a small element $d\Omega$ with area dA located at (x,y). Its distance from the y axis is x and its distance from the x axis is y. The first moment of the lamina about the y axis is

$$M_y = \int_\Omega x\delta(x,y)dA$$

and the first moment about the x axis is

$$M_x = \int_\Omega y\delta(x,y)dA.$$

The centre of mass has coordinates (x_m, y_m) defined by

$$x_m = \frac{M_y}{M} \qquad\qquad y_m = \frac{M_x}{M}$$

If δ is constant, it will cancel and the centre of mass then coincides with the centroid or centre of area of Ω which has coordinates $(\overline{x}, \overline{y})$ defined by

$$\overline{x} = \frac{\int\limits_\Omega x dA}{\int\limits_\Omega dA} \qquad\qquad \overline{y} = \frac{\int\limits_\Omega y dA}{\int\limits_\Omega dA}$$

Moments of Inertia

The moments of inertia of the above lamina about the x and y axes are

$$\begin{aligned} I_x &= \int\limits_\Omega y^2 \delta(x,y) dA \\ I_y &= \int\limits_\Omega x^2 \delta(x,y) dA. \end{aligned}$$

The polar moment of inertia about the origin is defined by

$$I_0 = I_x + I_y = \int\limits_\Omega (x^2 + y^2)\delta(x,y) dA.$$

Polar Coordinates

If the region Ω is easily described using polar coordinates ($x = r\cos\theta,\ y = r\sin\theta$) it is often better to evaluate the double integral in polar coordinates. The main task is to express dA in polar coordinates.
If r increases by dr and θ by $d\theta$, the little element of area so generated is

$$dA = (r d\theta) \times dr = r\, dr\, d\theta$$

so

$$\begin{aligned} \int\limits_\Omega f(x,y) dA &= \int\limits_\Omega\int f(r\cos\theta,\ r\sin\theta) r\, dr\, d\theta \\ &= \int\limits_\Omega\int f(r\cos\theta,\ r\sin\theta) r d\,\theta\, dr \end{aligned}$$

In either form the r and θ limits are chosen to cover the region Ω. This will be illustrated by examples in lectures. It is essential to draw a diagram of Ω.

Jacobian Transformation

The evaluation of $\int_a^b f(x)dx$ is often facilitated by the substitution $x = x(u)$ to give

$$\int_a^b f(x)dx = \int_\alpha^\beta f\left(x(u)\right) \frac{dx}{du} du$$

where $x(\alpha) = a,\ x(\beta) = b$. For double integrals, if $x = x(u,v)$, $y = y(u,v)$ then

$$\int\int_{\Omega} f(x,y)dx\,dy = \int\int_{\Omega^*} f\left(x(u,v), y(u,v)\right)|J|du\,dv$$

where Ω^* is the region in the (u,v) plane corresponding to Ω in the (x,y) plane and J is the Jacobian Determinant

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

(Thus $dA = |J|du\,dv$).

(N.B We take the absolute value of J)

This is valid provided J does not change sign on Ω^*. Further $x(u,v)$ and $y(u,v)$ must be continuous functions of u and v with continuous first partial derivatives. The point (x,y) corresponding to any (u,v) in Ω^* lies in Ω and to every point (x,y) in Ω there corresponds one and only one (u,v) in Ω^*.

TOPIC 5 – ORDINARY DIFFERENTIAL EQUATIONS

Differential equations arise naturally in science, engineering, biology and economics. Typically, they relate the rate at which a quantity changes (with respect to time or distance) to the quantity itself as well as to time or distance.

First Order Equations

These are of the form

$$\frac{dy}{dx} = F(x,y).$$

There is no general method of solution but certain equations fall into classes which can be solved. You must learn to recognise these.

1. Separable

These are of the form

$$\frac{dy}{dx} = f(x)g(y).$$

This may be written $\frac{dy}{g(y)} = f(x)dx$.

So $\int \frac{dy}{g(y)} = \int f(x)dx + c.$

This is an implicit relation between y and x and involves only one integration constant. If we were given some initial condition such as $y(0) = 1$, we would impose it now to find c.

2. Linear

The general first-order linear o.d.e. is

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

Multiplication by the **integrating factor** $I(x) = \exp\{\int P(x)dx\}$ reduces this to

$$\frac{d}{dx}(Iy) = IQ$$

which can immediately be solved by integration to give

$$y = \frac{(\int IQdx + c)}{I}.$$

Second-Order Linear Homogeneous O.D.E. with Constant Coefficients

These are of the form

$$y'' + ay' + by = 0 \tag{4}$$

where a and b are real constants. They arise in many engineering applications. The general solution will contain two arbitrary constants α_1 and α_2

$$y = \alpha_1 y_1(x) + \alpha_2 y_2(x)$$

where y_1 and y_2 are linearly independent functions.

How to solve

We look for a trial solution $y = \alpha e^{\lambda x}$. Substituting into (4) gives the characteristic equation

$$\lambda^2 + a\lambda + b = 0 \tag{5}$$

which has the roots

$$\begin{aligned} \lambda_1 &= \tfrac{1}{2}\left[-a + \sqrt{a^2 - 4b}\right] \\ \lambda_2 &= \tfrac{1}{2}\left[-a - \sqrt{a^2 - 4b}\right] \end{aligned}$$

Three Cases Arise

1. $a^2 > 4b$

   Then λ_1 and λ_2 are both real and the general solution is

   $$y = \alpha_1 e^{\lambda_1 x} + \alpha_2 e^{\lambda_2 x}.$$

2. $a^2 = 4b$. Then $\lambda_1 = \lambda_2 = -\frac{1}{2}a$ and the above procedure produces only one solution. In this case only, it can be shown that another solution is $y = xe^{\lambda_1 x}$ so the general solution is

   $$\begin{aligned} y &= \alpha_1 e^{\lambda_1 x} + \alpha_2 x e^{\lambda_1 x} \\ &= (\alpha_1 + \alpha_2 x)e^{-ax/2} \end{aligned}$$

3. $a^2 < 4b$. Then the roots are complex conjugates

$$\lambda_1 = -\frac{1}{2}a + iw,\ \lambda_2 = -\frac{1}{2}a - iw$$

where

$$w = \frac{1}{2}\sqrt{4b - a^2} = \sqrt{b - a^2/4}\ .$$

Then

$$\begin{aligned} y &= \alpha_1 e^{(-\frac{1}{2}a+iw)x} + \alpha_2 e^{(-\frac{1}{2}a-iw)x} \\ &= e^{-ax/2}\{\alpha_1 e^{iwx} + \alpha_2 e^{-iwx}\} \\ &= e^{-ax/2}\{\beta_1 \cos wx + \beta_2 \sin wx\} \end{aligned}$$

Free Oscillations

Several systems (e.g. a mass oscillating at the end of a spring or an electrical circuit) may be described by the d.e.

$$my'' + cy' + ky = 0$$

where $m > 0, c > 0, k > 0$. In a mechanical system, m is typically mass, c a friction coefficient and k a restoring-force coefficient. The independent variable is t and y is a displacement. Such a system is **unforced** since the term on the right is 0.

Seeking solutions $y = Ae^{\lambda t}$ gives the characteristic equation

$$m\lambda^2 + c\lambda + k = 0,$$

which has the solutions

$$\begin{aligned} \lambda_1 &= \frac{1}{2m}\left[-c + \sqrt{c^2 - 4mk}\right] \\ \lambda_2 &= \frac{1}{2m}\left[-c - \sqrt{c^2 - 4mk}\right]. \end{aligned}$$

Three cases arise

1. $c^2 > 4mk$. This is called "overdamping" since the damping or frictional coefficient c is large compared with $2\sqrt{mk}$. In this case λ_1 and λ_2 are both real and negative. The solution is $y = Ae^{\lambda_1 t} + Be^{\lambda_2 t}$ which decays to zero as $t \longrightarrow \infty$.

2. $c^2 = 4mk$. This is "critical damping" and $\lambda_1 = \lambda_2$ so the solution is

$$y = (A + Bt)e^{-ct/2m}.$$

The solution also decays to 0 as $t \longrightarrow \infty$.

3. $c^2 < 4mk$. This is called "underdamping" as c is smaller than $2\sqrt{mk}$. Then

$$\lambda_1 = -\alpha + i\Omega, \quad \lambda_2 = -\alpha - i\Omega$$

where $\alpha = \frac{c}{2m}, \quad \Omega = \sqrt{\frac{k}{m} - \frac{c^2}{4m^2}}$

Thus

$$\begin{aligned} y &= (A\cos\Omega t + B\sin\Omega t)e^{-\alpha t} \\ &= Re^{-\alpha t}\cos(\Omega t - \delta) \end{aligned}$$

where $R = \sqrt{A^2 + B^2}$ and $\tan\delta = B/A$. This represents decaying oscillations. In the idealised case $c = 0$ (no friction), $y = R\cos(\Omega_0 t - \delta)$ where $\Omega_0 = \sqrt{k/m}$. This has period $\frac{2\pi}{\Omega_0}$. In reality, $c > 0$ and these oscillations are killed off by friction.

Non-Homogeneous Second-Order Linear O.D.E with Constant Coefficients

These are of the form

$$y'' + ay' + by = r(x) \tag{6}$$

From First-Year Calculus, we know that the general solution of (6) is

$$y = y_H + y_P$$

where y_P is **any** solution of (6) and y_H is the most general solution of the homogeneous equation

$$y'' + ay' + by = 0. \tag{7}$$

Thus, we solve (6) in the following way:

1. Find the general solution y_H of (7) by means already discussed. This will contain two arbitrary constants and is sometimes called the complementary function or complementary solution.

2. Produce any solution y_P of (6), no matter how special. If $r(x)$ is simple, the method of undetermined coefficients may be used (see later). This solution is sometimes called the particular solution or particular integral.

3. Add the above two solutions to give the general solution of (6).

4. If initial or boundary conditions are specified, impose them now to determine the constants in y_H.

   Clearly, the crucial step is 2.

Method of undetermined coefficients

If $r(x)$ is sufficiently simple, we can guess what y_P must be as in the following table.

Term in $r(x)$	choice of y_P
$ke^{\sigma x}$	$Ce^{\sigma x}$
$P_n(x)$	$Q_n(x)$
$k\cos\theta x$ or $k\sin\theta x$	$K\cos\theta x + M\sin\theta x$
$ke^{\sigma x}\cos\theta x$ or $ke^{\sigma x}\sin\theta x$	$e^{\sigma x}(K\cos\theta x + M\sin\theta x)$
$P_n(x)e^{\sigma x}\cos\theta x$ or $P_n(x)e^{\sigma x}\sin\theta x$	$e^{\sigma x}(Q_n(x)\cos\theta x + R_n(x)\sin\theta x)$

Rules for Method of Undermined Coefficients

1. If a term in $r(x)$ appears in the first column, choose the y_P from the corresponding second column and determine the unknown coefficients by substituting into (6). In the table, P_n, Q_n and R_n are polynomials of degree n. Even if $P_n(x)$ has only one term (x^n), Q_n and R_n will in general have all $(n+1)$ terms, i.e. $Q_n(x) = \sum_{j=0}^{n} q_j x^j$.

2. If a term in our choice of y_P happens to be a solution of the homogeneous equation (7), we multiply that term by x. If our new guess is **still** a solution of (7) we multiply by a further x. (This last case only happens when the characteristic equation has equal roots).

3. If $r(x)$ is a sum of functions, choose for $y_P(x)$ the sum of the appropriate terms in the table.

Forced Oscillations

When simple periodic forcing is added to the mechanical or electrical system studied earlier, we have to solve an equation like

$$my'' + cy' + ky = F_0\cos wt \tag{8}$$

where, as before, $m > 0, c > 0, k > 0$.

This models a mechanical system driven periodic forces or an electrical system forced by a periodic voltage. The solution of the homogeneous equation has already been discussed. For y_P we try

$$y_P = a\cos wt + b\sin wt$$

and find $a = m(w_0^2 - w^2)\frac{F_0}{\Delta},\ b = wc\frac{F_0}{\Delta}$ where

$$w_0^2 = \frac{k}{m} \quad \text{and} \quad \Delta = m^2(w_0^2 - w^2)^2 + w^2c^2.$$

Thus

$$y = \Theta\cos(wt - \delta)$$

where

$$\Theta^2 = a^2 + b^2 = \frac{F_0^2}{\Delta}$$

and

$$\tan\delta = \frac{b}{a} = \frac{wc}{(m(w_0^2 - w^2))}.$$

Undamped Oscillations ($c = 0$)

In the ideal case of no mechanical friction or electrical resistance, the complete solution of (8) is

$$y = A\cos w_0 t + B\sin w_0 t + \frac{F}{m(w_0^2 - w^2)}\cos wt$$

for $w \neq w_0$. This is the sum of free oscillations with frequency $w_0 = \sqrt{\frac{k}{m}}$ and a forced oscillation with frequency w.

As $w \longrightarrow w_0$, the amplitude of the latter gets larger and larger. For $w = w_0$, the above is not valid and the appropriate solution (using rule 2 above) is

$$y = A\cos w_0 t + B\sin w_0 t + \frac{F_0}{2mw_0} t\sin w_0 t$$

The forced response thus consists of a sinusoid with linearly increasing amplitude. This is the phenomenon of **resonance** and occurs when the frequency w of the forcing is exactly equal to the natural frequency w_0 with which the system likes to oscillate.

Effects of Friction

In reality, $c > 0$ so the above is modified by friction or electrical resistance. As we showed earlier, all solutions of the homogeneous equation decay to zero as $t \longrightarrow \infty$ when $c > 0$. These are called the **transients** of the system. Thus we are ultimately left with the directly forced response

$$y_P = \Theta\cos(wt - \delta)$$

where $\frac{\Theta}{F_0} = M = \frac{1}{\sqrt{m^2(w_0^2 - w^2)^2 + w^2c^2}}$ and $\tan\delta = \frac{wc}{m(w_0^2 - w^2)}$ and $w_0 = \sqrt{\frac{k}{m}}$ is the natural frequency of the undamped ($c = 0$) oscillations. The quantity M is the magnification ratio for the amplitude. In engineering design we might want this to be small so as to avoid damaging resonances or to be large to magnify weak signals (e.g. tuning an AM radio). This will be discussed in lectures.

The Method of Variation of Parameters

This is a general method for finding y_P.

Consider

$$y'' + p(x)y' + q(x)y = f(x)$$

The general solution is

$$y_g = y_h + y_p$$

Complementary solution assumed known

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

For particular solution assume

$$y_p(x) = A(x)y_1(x) + B(x)y_2(x)$$

Then $y_p(x)$ is a solution if

$$\begin{aligned} y_1A' + y_2B' &= 0 \\ y_1'A' + y_2'B' &= f(x) \end{aligned}$$

The variable parameters $A(x), B(x)$ are found by solving these two equations first for $A'(x), B'(x)$.

PROOF:

$$\begin{aligned} y_p &= Ay_1 + By_2 \\ y_p' &= A'y_1 + Ay_1' + B'y_2 + By_2' \end{aligned}$$

But if

$$A'y_1 + B'y_2 = 0 \qquad \otimes$$

Then

$$y' = Ay_1' + By_2'$$

Now

$$y'' = A'y_1' + Ay_1'' + B'y_2' + By_2''$$

Substitute trial $y_p(x)$ into the ODE

$$\begin{aligned} A[y_1'' + py_1 + qy_1] \quad & +B[y_2'' + py_2' + qy_2] \\ & +A'y_1' + B'y_2' = f(x) \end{aligned}$$

But y_1, y_2 are solutions of the homogeneous part and so

$$A'y_1' + B'y_2' = f(x) \qquad \oplus$$

Thus $y_p(x)$ is a solution of the ODE provided that $\otimes$ and $\oplus$ are satisfied. Now solve $\otimes$ and $\oplus$ to find $A(x)$ and $B(x)$ First write in matrix form

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} A' \\ B' \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix}$$

We can find a unique solution if the determinant is non-zero. However the determinant is simply the Wronskian

$$W = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$$

Clearly this is non-zero because y_1, y_2 are linearly independent
EXERCISE Solve the matrix equations and show that

$$y_p(x) = -y_1(x) \int \frac{y_2(x)f(x)}{W(x)} dx + y_2(x) \int \frac{y_1(x)f(x)}{W(x)} dx$$

NOTE This method also works if p and q are constants.

TOPIC 6 – MATRICES

Brief revision (including special matrices)

A matrix is a rectangular array of numbers (real or complex) in the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Here, A is an $m \times n$ matrix which has m rows and n columns. We write

$$A = (a_{jk})$$

in which j is the row suffix and k the column suffix, *e.g.*, a_{32} is the entry in the 3rd row, 2nd column.

If all entries of A are real, we call A a real matrix; otherwise it is a complex matrix.

Row vector : (matrix with one row)

$$\mathbf{a} = (a_1, a_2, \dots, a_n)$$

Column vector : (matrix with one column)

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Square matrix : If $m = n$ (*i.e.* A has the same number of rows and columns) we have a square matrix A and we call n its **order**.

Addition of matrices, multiplication by scalars

Definition: Two matrices $A = (a_{jk})$ and $B = (b_{jk})$ are equal if and only if A and B have the same number of rows and columns and

$$a_{jk} = b_{jk} \quad \text{for all} \quad j, k.$$

We write $A = B$.
Definition: Addition of matrices is defined only for matrices with the same number of **rows** and **columns**. The sum of two $m \times n$ matrices A and B is an $m \times n$ matrix $A + B$ with entries

$$a_{jk} + b_{jk} \quad (j = 1, \ldots, m; k = 1, \ldots n).$$

We define the zero matrix 0 to be the $m \times n$ matrix with all entries zero.
Properties of Addition:

a) $A + B = B + A$

b) $(U + V) + W = U + (V + W)$

c) $A + 0 = A$

d) $A + (-A) = 0$, $-A = (-a_{jk})$.

Definition: Multiplication by a scalar

$$cA = Ac = \begin{pmatrix} ca_{11} & ca_{12} & \ldots & ca_{1n} \\ ca_{21} & ca_{22} & \ldots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \ldots & ca_{mn} \end{pmatrix}$$

or $cA = (ca_{jk})$.

For any $m \times n$ matrices A and B and any scalars c and k

a) $c(A + B) = cA + cB$

b) $(c + k)A = cA + kA$

c) $c(kA) = (ck)A$

d) $1A = A$

Matrix multiplication

Let $A = (a_{jk})$ be an $m \times n$ matrix and $B = (b_{jk})$ an $r \times p$ matrix. Then the product AB (in this order) is defined only when $r = n$. (*i.e.* the number of rows of B = the number of columns of A). Then AB is an $m \times p$ matrix $C = (c_{jk})$ where

$$c_{jk} = (j \text{ th row vector of } A) \cdot (k \text{ th column vector of } B)$$

or

$$c_{jk} = \sum_{l=1}^{n} a_{jl} b_{lk} = a_{j1} b_{1k} + a_{j2} b_{2k} + \cdots + a_{jn} b_{nk} \quad (j = 1, \cdots m, k = 1, \cdots p).$$

Properties of Matrix Multiplication: Assuming matrix multiplication is defined

a) $(kA)B = k(AB) = A(kB)$

b) $A(BC) = (AB)C$

c) $(A + B)C = AC + BC$

d) $C(A + B) = CA + CB$

Important notes:
(1) If A and B are matrices such that AB and BA are defined then, in general,

$$AB \neq BA$$

(2) $AB = 0$ does not necessarily imply $A = 0$ or $B = 0$.

Special Matrices:

(1) *Triangular Matrices*: A square $n \times n$ matrix whose entries above the main diagonal are all zero is called a **lower triangular matrix**. Similarly, a square matrix whose entries below the diagonal are all zero is called a **upper triangular matrix**.

(2) *Diagonal Matrices*: A square matrix $A = a_{jk}$ whose entries above **and** below the main diagonal are all zero (*i.e.* $a_{jk} = 0$ for $j \neq k$) is called a **diagonal matrix**.

A **scalar matrix** is a diagonal matrix whose entries on the main diagonal are all equal.

$$S = \begin{pmatrix} c & 0 & \dots & 0 \\ 0 & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{pmatrix}$$

We then have $AS = SA = cA$ (for any $n \times n$ matrix A).

The **unit matrix**

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

We then have $AI = IA = A$ (for any $n \times n$ matrix A).

Transpose of a matrix: The transpose A^T of an $m \times n$ matrix $A = (a_{jk})$ is the $n \times m$ matrix in which the k th column of A becomes the k th row of A^T (and the j th row of A becomes the j th column of A^T).

$$A^T = (a_{kj}) = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

The transpose of a product of matrices is given by

$$(AB)^T = B^T A^T.$$

Definition: A **real** square matrix $A = (a_{jk})$ is **symmetric** if $A^T = A$.
Definition: A **real** square matrix $A = (a_{jk})$ is **skew–symmetric** if $A^T = -A$.

Note: We can write any real square matrix as $A = S + R$ (where S is skew–symmetric and R is symmetric) where

$$R = \frac{1}{2}(A^T + A), \quad S = \frac{1}{2}(A - A^T).$$

Systems of linear equations, Gaussian Elimination

Consider the system of linear equations

$$\left.\begin{array}{ll} a_{11}x_1 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n & = b_2 \\ & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = b_m \end{array}\right\} \tag{9}$$

where the a_{jk} are the coefficients. The system of equations (9) has m equations with n unknowns. If all $b_i = 0$ the system is **homogeneous**. If at least one of the $b_i \neq 0$ the system is **nonhomogeneous**. We can write the system (9) as

$$A\mathbf{x} = \mathbf{b}$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

We define the **augmented matrix** $\tilde{A}$ or $[A|\boldsymbol{b}]$ as

$$\tilde{A} = [A|\boldsymbol{b}] = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

This completely determines the system (9).

The system (9) is **overdetermined** if it has more equations than unknowns ($m > n$). It is **determined** if $m = n$ and is **undetermined** if $m < n$ (*i.e.* the system has fewer equations than unknowns).

Theorem The system (9) has

1. No solution if the rank of A is not equal to the rank of the augmented matrix $\tilde{A}$
2. A unique solution if the ranks of A and $\tilde{A}$ both equal n.
3. Infinitely many solutions if the ranks of A and $\tilde{A}$ are equal and $< n$.

Recall that the rank of a matrix is the maximum number of linearly independent rows of the matrix. Somewhat surprisingly, this is also the maximum number of linearly independent columns of the matrix or, in other words, row-rank equals column-rank.

The solutions of (9) may be found by **Gaussian elimination** which is a systematic process of elimination to reduce the matrix to its echelon form, followed by back-substitution. Gaussian elimination is done using elementary row operations:

1. interchange two rows

2. multiplication of a row by a nonzero constant

3. addition of a constant multiple of one row to another row.

{*See first-year algebra notes for further details of Gaussian elimination.*}

Inverse of a Matrix: The inverse of an $n \times n$ matrix $A = (a_{jk})$, denoted by A^{-1}, is an $n \times n$ matrix such that

$$AA^{-1} = A^{-1}A = I$$

where I is the $n \times n$ identity (unit) matrix. If A has an inverse it is called **nonsingular**. If A has no inverse it is called **singular**.

Existence of an inverse: For an $n \times n$ matrix A the inverse exists if and only if the rank of A is equal to n. This is equivalent to saying that the determinant of A (written $\det A$ or $|A|$) is non-zero. If the inverse exists it is unique and the solution of (9) when A is a non-singular $n \times n$ matrix is given by $\mathbf{x} = A^{-1}\mathbf{b}$. The inverse is of great theoretical importance. However, in practical problems, we solve the system by Gaussian elimination and back-substitution and not by calculation of the inverse followed by $A^{-1}\mathbf{b}$ since the latter approach involves more work than the former.

Useful formula: For a nonsingular 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

where $\det A = a_{11}a_{22} - a_{12}a_{21}$ is the **determinant** of A. Note that a matrix is nonsingular if $\det A \neq 0$ (this holds for any general $n \times n$ matrix.)

Inverse of a diagonal matrix:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{pmatrix}$$

provided $a_{jj} \neq 0$ for all $j = 1, \dots n$.

Inverse of a product of matrices:

$$(AC)^{-1} = C^{-1}A^{-1}.$$

Eigenvalues and eigenvectors:

Let $A = (a_{jk})$ be an $n \times n$ matrix. Consider the **vector equation**

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{10}$$

where λ is some scalar. The vector $\mathbf{x} = \mathbf{0}$ is a solution of (10) for all λ. A value of λ for which (10) has a nonzero solution is called an **eigenvalue** of A; the corresponding vector $\mathbf{x}$ is called an **eigenvector** of A. Observe that $c\mathbf{x}$ is also an eigenvector for any scalar $c \neq 0$.

Determination of eigenvectors: Any $n \times n$ matrix has at least **one** and at most n distinct (real or complex) eigenvalues. Rewrite (10) as

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

This is a set of homogeneous linear equations and has a nontrivial ($\mathbf{x} \neq \mathbf{0}$) solution if and only if

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{11}$$

Equation (11) then gives the **characteristic equation** (or **polynomial**) of A.

To get the eigenvalues we must evaluate the determinant. This can be done for any order matrix but in practice if $n \geq 4$ (say) it is usually simpler to determine the eigenvalues numerically.

For a 2×2 matrix we get a quadratic in λ:

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}$$

For a 3×3 matrix we get a cubic in λ. For a general $n \times n$ matrix A the characteristic equation is a polynomial of order n in λ. Even if A is real, some of the eigenvalues may be complex and hence so will be the corresponding eigenvectors.

Definition: If λ is an eigenvalue of order M_λ (*i.e.* a root of the characteristic equation of order M_λ) then M_λ is called the **algebraic multiplicity** of λ. The number, m_λ, of linearly independent eigenvectors corresponding to λ is called the **geometric multiplicity**. In general $m_\lambda \leq M_\lambda$ (*i.e.* geometric multiplicity $\leq$ algebraic multiplicity).

Some Properties of Eigenvalues & Eigenvectors

In the following list of properties, λ is an eigenvalue of A with $\mathbf{x}$ being the corresponding eigenvector.

1. Zero is an eigenvalue of A if and only if A is singular.

2. If k is a constant, the matrix kA has eigenvectors identical with those of A and eigenvalues $k\lambda$

3. If m is a positive integer, the matrix A^m has eigenvectors identical with those of A and eigenvalues λ^m

4. A^{-1} has eigenvectors identical with those of A and eigenvalues λ^{-1}

5. The transposed matrix A^T has the same eigenvalues as A but, in general, different eigenvectors

6. If A is a real symmetric matrix, ($A^T = A$), the eigenvalues of A are all real and hence the eigenvectors may be taken to be real also.

7. The eigenvectors associated with different eigenvalues are linearly independent

8. If A is a real $n \times n$ symmetric matrix, we can always find n linearly independent eigenvectors of A, even if some of the eigenvalues are repeated.

9. If A is a real symmetric matrix, the eigenvectors associated with different eigenvalues of A are orthogonal to one another.

Reminder : Orthogonality of Vectors

Two $n \times 1$ column vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be **orthogonal** if $\mathbf{a}^T\mathbf{b} = 0$. Since $\mathbf{a}^T\mathbf{b} = \sum_{i=1}^{n} a_i b_i$, this is the n-dimensional version of the dot product of two vectors in 3-dimensional space and so is sometimes written $\mathbf{a} \cdot \mathbf{b}$.

Quadric surfaces

Quadric surfaces are surfaces in space whose equations combine quadratic terms with linear terms and constants.

Examples:

$$\text{ellipsoid :} \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

$$\begin{aligned}
\text{elliptic paraboloid}: &\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z}{c},\\
\text{elliptic cone}: &\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z^2}{c^2},\\
\text{hyperboloid (one sheet)}: &\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1,\\
\text{hyperboloid (two sheet)}: &\quad \frac{z^2}{c^2}-\frac{x^2}{a^2}-\frac{y^2}{b^2}=1.
\end{aligned}$$

Quadratic forms: Quadric surfaces can be written in terms of a quadratic form. Consider a real $n \times n$ matrix A and real vector $\mathbf{x}$. Then

$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \;&= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n\\
&+ a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n\\
&\;\;\vdots\\
&+ a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2,
\end{aligned}$$

and the matrix A may be assumed to symmetric by re-defining the matrix elements as $(a_{ij})_{\text{new}} = (a_{ji})_{\text{new}} = \frac{1}{2}(a_{ij}+a_{ji})_{\text{old}}$. We can then characterise any quadric surface by a corresponding symmetric matrix $A = A^T$.

Orthogonal matrices and diagonalisation

An **orthogonal** matrix is one for which

$$A^T = A^{-1}$$

An $n \times n$ matrix $\hat{A}$ is said to be **similar** to an $n \times n$ matrix A if

$$\hat{A} = T^{-1}AT$$

for some nonsingular matrix T. The transformation, which gives $\hat{A}$ from A, is called a **similarity transformation**.
Theorem: If $\hat{A}$ is similar to A, then $\hat{A}$ has the same eigenvalues as A. Also, if $\mathbf{x}$ is an eigenvector of A then $\mathbf{y} = T^{-1}\mathbf{x}$ is an eigenvector of $\hat{A}$ corresponding to the same eigenvalue.
Proof:

$$\begin{aligned}
A\mathbf{x} &= \lambda\mathbf{x}\\
\Rightarrow T^{-1}A\mathbf{x} &= \lambda T^{-1}\mathbf{x}\\
\Rightarrow T^{-1}ATT^{-1}\mathbf{x} &= \lambda T^{-1}\mathbf{x}\\
\Rightarrow \hat{A}(T^{-1}\mathbf{x}) &= \lambda(T^{-1}\mathbf{x})
\end{aligned}$$

Therefore λ is an eigenvalue of $\hat{A}$ with eigenvector $T^{-1}\mathbf{x}$. ■

Theorem: Let $\lambda_1, \lambda_2, \ldots \lambda_k$ be distinct eigenvalues of an $n \times n$ matrix. Then the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ form a linearly independent set. This is basically just Property 7 above.
Theorem: If an $n \times n$ matrix A has n **distinct** eigenvalues, then A has n linearly independent eigenvectors which therefore constitute a basis for R^n (or C^n). This follows from Property 7 above.
Orthonormal Sets of Vectors: Consider the set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ This set is said to be **orthonormal** if $\mathbf{v}_j^T \mathbf{v}_j = 1$ for $j = 1, \ldots, n$ and $\mathbf{v}_j^T \mathbf{v}_k = 0$ for $j \neq k$. The matrix with these vectors as column vectors is orthogonal *i.e.*. if $P = (\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n})$ then $P^{-1} = P^T$.

Theorem: The eigenvalues of a real, symmetric matrix are all real. Hence, we may take the eigenvectors to be real also.
Theorem: A real, symmetric matrix has an **orthonormal** basis of eigenvectors in R^n.

Hence, even though some of the eigenvalues of a real, symmetric matrix may be equal, we can always find n orthogonal eigenvectors. By normalising each of these to have length 1, we can construct an orthonormal basis for R^n.
Diagonalisation of a Matrix
Theorem: If an $n \times n$ matrix A has a basis of eigenvectors (*i.e.* n linearly independent eigenvectors) then

$$D = P^{-1}AP$$

is diagonal, with the eigenvalues of A as the entries on main diagonal. P is the matrix with the eigenvectors of A as column vectors. Also

$$D^m = P^{-1}A^m P.$$

Diagonalisation of a Symmetric Matrix

Not all matrices can be diagonalised - it relies upon us being able to find n linearly independent eigenvectors. Fortunately, a real symmetric matrix can always be diagonalised (because of Property 8 above). Even better, because of Property 9, the columns of P are orthogonal to one another. If we further normalise each eigenvector by dividing it by its length (the square root of the sum of the squares of its n components), the matrix P is then an orthogonal matrix and so its inverse is just its transpose. This saves a lot of work.

Transformation of a quadratic form to principal axes

Suppose we have a quadratic form

$$Q = \mathbf{x}^T A \mathbf{x} \tag{12}$$

where A is real–symmetric. Then by a previous theorem, A has an **orthonormal basis** of n eigenvectors. Hence the P matrix with eigenvectors of A as column vectors is orthogonal. Now

$$D = P^{-1}AP \Rightarrow A = PDP^{-1} \Rightarrow A = PDP^T$$

Substitution of this into (12) gives

$$Q = \mathbf{x}^T P D P^T \mathbf{x}$$

Set $P^T\mathbf{x} = \mathbf{y} \Rightarrow \mathbf{x} = P\mathbf{y}$ (using $P^{-1} = P^T$). Then

$$Q = \mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{13}$$

where λ_j are the eigenvalues of A. This proves the following
Theorem: The substitution $\mathbf{x} = P\mathbf{y}$ transforms the quadratic form

$$Q = \mathbf{x}^T A \mathbf{x} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k$$

to the **principal axis** form (13) where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues (not necessarily distinct) of the symmetric matrix A and P is the orthogonal matrix with the corresponding eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n$ as column vectors.

Systems of linear ordinary differential equations

Consider the first order system of equations

$$\frac{dy_j}{dt} = \sum_{k=1}^{n} a_{jk} y_k + h_j(t), \quad j = 1, \ldots n. \tag{14}$$

(for simplicity we will assume that the coefficients a_{jk} are constant).
Note: any higher (nth) order o.d.e can be reduced to a system of n first order o.d.e's.
Example:

$$\frac{d^3x}{dt^3} + 2\frac{d^2x}{dt^2} + x = 0.$$

Define

$$\begin{aligned} x &= y_1, \\ \frac{dx}{dt} &= \frac{dy_1}{dt} = y_2, \\ \frac{d^2x}{dt^2} &= \frac{dy_2}{dt} = y_3, \\ \frac{d^3x}{dt^3} &= \frac{dy_3}{dt} = -2\frac{d^2x}{dt^2} - x = -2y_3 - y_1. \end{aligned}$$

Then

$$\frac{dy_1}{dt} = y_2, \quad \frac{dy_2}{dt} = y_3, \quad \frac{dy_3}{dt} = -2y_3 - y_1. \qquad \blacksquare$$

Returning to (14): Rewrite this in matrix form

$$\frac{d\mathbf{y}}{dt} = A\mathbf{y} + \mathbf{h}, \qquad A = (a_{jk}). \tag{15}$$

Assume A has a basis of eigenvectors $\mathbf{p}_1, \ldots \mathbf{p}_n$. Then $P = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$ is nonsingular. To diagonalise (15) we write

$$\mathbf{y} = P\mathbf{z}$$

Then assuming that all the a_{jk} are constant ($\Rightarrow$ P is constant) we have $\mathbf{y}' = P\mathbf{z}'$ and (15) becomes

$$\begin{aligned} P\mathbf{z}' &= AP\mathbf{z} + \mathbf{h} \\ \Rightarrow \mathbf{z}' &= P^{-1}AP\mathbf{z} + P^{-1}\mathbf{h}, \\ \Rightarrow \mathbf{z}' &= D\mathbf{z} + P^{-1}\mathbf{h}, \end{aligned}$$

where D is the diagonal matrix with the eigenvalues of A along the diagonal. Writing the last expression in component form gives

$$z_j' - \lambda_j z_j = r_j(t),$$

where r_j is the jth component of $P^{-1}\mathbf{h}$. This is then a first order, constant coefficient o.d.e whose general solution is

$$z_j(t) = e^{\lambda_j t}\left(\int e^{-\lambda_j t} r_j(t)\, dt + C_j\right)$$

with C_j arbitrary constants. These can be determined by supplying initial conditions

$$\mathbf{y}(t_0) = \mathbf{y}_0 \quad \Rightarrow \quad \mathbf{z}(t_0) = \mathbf{z}_0 = P^{-1}\mathbf{y}_0. \tag{16}$$

Existence and Uniqueness: Let $\mathbf{h}(t)$ in (15) be continuous in the interval $\alpha < t < \beta$ and let t_0 be any given point in that interval. Assume that A has a set of n linearly independent eigenvectors. Then the initial value problem (15), (16) has a unique solution in the interval.
Stability: All solutions of (15) with $\mathbf{h} = \mathbf{0}$ approach zero as $t \longrightarrow \infty$ if and only if all the eigenvalues of A have **negative real parts**.

TOPIC 7 – THE LAPLACE TRANSFORM

Laplace Transforms

Let $f(t)$ be a given function defined for all $t \geq 0$. If the integral

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt$$

exists it is called the **Laplace transform** of $f(t)$. We denote it by $\mathcal{L}(f)$ *i.e.*

$$\mathcal{L}(f) = F(s) = \int_0^\infty e^{-st} f(t)\, dt \tag{17}$$

The original function $f(t)$ in (17) is called the inverse transform of $F(s)$; we denote it by $\mathcal{L}^{-1}(F)$, so

$$f(t) = \mathcal{L}^{-1}(F).$$

Linearity: The Laplace transformation is a linear operation. For any functions $f(t)$ and $g(t)$ whose transforms exist and any constants a and b

$$\mathcal{L}\left\{af(t) + bg(t)\right\} = a\mathcal{L}\left\{f(t)\right\} + b\mathcal{L}\left\{g(t)\right\}. \tag{18}$$

Tables of Laplace Transforms: A short table of Laplace Transforms is at the end of these notes.

Existence theorem: Let $f(t)$ be a function that is piecewise continuous (*i.e.* a function which is continuous except at a finite number of points) on every finite interval in the range $t \geq 0$ and satisfies

$$|f(t)| \leq Me^{\gamma t} \qquad \text{for all } t \geq 0, \tag{19}$$

for some constants γ and M. Then the Laplace transform of $f(t)$ exists for all $s > \gamma$.

Transform of derivatives

Theorem: Suppose that $f(t)$ is continuous for all $t \geq 0$, satisfies (19) (for some γ and M) and $f'(t)$ is piecewise continuous on every finite interval in the range $t \geq 0$. Then the Laplace transform of $f'(t)$ exists when $s > \gamma$ and

$$\mathcal{L}(f') = s\mathcal{L}(f) - f(0) \qquad (\text{for} \quad s > \gamma) \tag{20}$$

Higher derivatives

$$\begin{aligned} \mathcal{L}(f'') &= s\mathcal{L}(f') - f'(0), \\ &= s\left[s\mathcal{L}(f) - f(0)\right] - f'(0) \\ &= s^2\mathcal{L}(f) - sf(0) - f'(0) \end{aligned}$$

Continuing gives us the general result

$$\mathcal{L}(f^{(n)}) = s^n\mathcal{L}(f) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0).$$

Transform of an integral of a function

If $f(t)$ is piecewise continuous and satisfies our earlier inequality then

$$\mathcal{L}\left(\int_0^t f(\tau)d\tau\right) = \frac{1}{s}\mathcal{L}(f(t)) \qquad (\text{for} \quad s > \gamma).$$

Hence if we write $\mathcal{L}(f(t)) = F(s)$ then

$$\mathcal{L}\left(\int_0^t f(\tau)d\tau\right) = \frac{F(s)}{s}$$

Taking the inverse Laplace transform gives

$$\mathcal{L}^{-1}\left(\frac{F(s)}{s}\right) = \int_0^t f(\tau)d\tau.$$

Shifting on the s–axis

Suppose $f(t)$ has the transform $F(s)$ where $s > \gamma$ then $e^{at}f(t)$ has the transform $F(s-a)$ where $s - a > \gamma$, *i.e.* if

$$\mathcal{L}(f(t)) = F(s)$$

then

$$\mathcal{L}(e^{at}f(t)) = F(s-a), \tag{21}$$

So if we know the transform $F(s)$ of a function $f(t)$, (21) gives us the transform of $\exp(at)f(t)$ just by **shifting on the s axis** *i.e.* replacing s with $s-a$ to give $F(s-a)$. Taking the inverse transform in (21) gives

$$\mathcal{L}^{-1}(F(s-a)) = e^{at}f(t).$$

An example: Find $\mathcal{L}(e^{at}\cos\omega t)$.
We know that

$$\mathcal{L}(\cos\omega t) = \frac{s}{s^2+\omega^2}$$

$$(21) \quad \Rightarrow \quad \mathcal{L}(e^{at}\cos\omega t) = \frac{s-a}{(s-a)^2+\omega^2}$$

Shifting on the t–axis

If $f(t)$ has the transform $F(s)$ and $a \geq 0$ then the function

$$\tilde{f}(t) = \begin{cases} 0 & \text{if } t < a \\ f(t-a) & \text{if } t > a \end{cases} \tag{22}$$

has the transform

$$e^{-as}F(s).$$

Thus if we know $F(s)$ is the transform of $f(t)$ then we get the transform of (22) by multiplying $F(s)$ by e^{-as}.

Unit step function $u(t)$

Define

$$u(t) = \begin{cases} 0, & \text{if } t < 0; \\ \frac{1}{2}, & \text{if } t = 0; \\ 1, & \text{if } t > 0. \end{cases}$$

(also called the **Heaviside function**). Note: our previous function

$$\tilde{f}(t) = \begin{cases} 0, & \text{if } t < a; \\ f(t-a), & \text{if } t > a \end{cases}$$

can now be written as

$$\tilde{f}(t) = f(t-a)u(t-a)$$

(strictly, we should define $\tilde{f}(a) = f(0)/2$) and our shifting on the t–axis result is

$$\mathcal{L}\{f(t-a)u(t-a)\} = e^{-as}F(s) \quad \text{for} \quad a \geq 0$$

and the inverse

$$\mathcal{L}^{-1}\left\{e^{-as}F(s)\right\} = f(t-a)u(t-a) \quad \text{for} \quad a \geq 0$$

Another useful formula is

$$\mathcal{L}\{u(t-a)\} = \frac{e^{-as}}{s} \quad \text{for} \quad a \geq 0$$

Partial Fractions

In most applications we obtain $Y(s) = \mathcal{L}(y)$, the transform of the function y, in the form

$$Y(s) = \frac{F(s)}{G(s)},$$

where $F(s)$ and $G(s)$ are polynomials in s. For example in the solution of an ordinary differential equation we would expect a transform of this form.

Assume $F(s)$ and $G(s)$ have real coefficients and no common factors. The degree of F is lower than the degree of G. In practice we express F/G in terms of partial fractions

Example:

$$\begin{aligned}\frac{F(s)}{G(s)} &= \frac{1}{(s^2-3s-4)} = \frac{1}{(s-4)(s+1)} \\ &= \frac{A}{(s-4)} + \frac{B}{(s+1)}\end{aligned}$$

The form of the partial fraction depends on the type of factor in F/G. The four common ones are

Case 1: unrepeated factor $(s-a)$
Case 2: repeated factor $(s-a)^m$
Case 3: complex factors $(s-a)(s-\bar{a})$ ($\bar{a}$ complex conjugate of a)
Case 4: repeated complex factors $[(s-a)(s-\bar{a})]^2$

Cases 1–4 lead to a general form of partial fraction for $Y = F/G$ which then have a corresponding inverse Laplace transform.

Case 1: In $Y = F/G$ a fraction

$$\frac{A}{(s-a)}$$

has inverse transform

$$Ae^{at}$$

with

$$A = F(a)/G'(a).$$

Case 2: In $Y = F/G$ a sum of m fractions

$$\frac{A_m}{(s-a)^m} + \frac{A_{m-1}}{(s-a)^{m-1}} + \cdots + \frac{A_1}{(s-a)}.$$

Using tables and s–shifting this has inverse transform

$$e^{at}\left\{\frac{A_m}{(m-1)!}t^{m-1} + \frac{A_{m-1}}{(m-2)!}t^{m-2} + \cdots + \frac{A_2}{1!}t^1 + A_1\right\},$$

where

$$\begin{aligned} A_m &= \lim_{s\to a} \frac{(s-a)^m F(s)}{G(s)} \\ A_k &= \frac{1}{(m-k)!} \lim_{s\to a} \frac{d^{m-k}}{ds^{m-k}} \left[\frac{(s-a)^m F(s)}{G(s)}\right] \qquad (k = 1, \ldots, m-1) \end{aligned}$$

Case 3: unrepeated complex factors $(s-a)(s-\bar{a})$. Define $a = \alpha + i\beta$. Then

$$(s-a)(s-\bar{a}) = (s-\alpha)^2 + \beta^2$$

This corresponds to the partial fraction

$$\frac{As+B}{(s-\alpha)^2+\beta^2} \qquad \text{or} \qquad \frac{A(s-\alpha)+\alpha A+B}{(s-\alpha)^2+\beta^2}$$

Using the table and s–shifting theorem gives the inverse transform

$$e^{\alpha t}\left(A\cos\beta t + \frac{\alpha A + B}{\beta}\sin\beta t\right)$$

Here A is the imaginary part and $(\alpha A + B)/\beta$ is the real part of

$$Q_a = \frac{1}{\beta}\lim_{s\to a}\frac{[(s-\alpha)^2+\beta^2]F(s)}{G(s)}$$

Case 4: repeated complex factors $[(s-a)(s-\bar{a})]^2$. These correspond to partial fractions in $Y = F/G$

$$\frac{As+B}{[(s-\alpha)^2+\beta^2]^2} + \frac{Cs+D}{(s-\alpha)^2+\beta^2}$$

which has the inverse transform

$$e^{\alpha t}\left[\frac{A}{2\beta}t\sin\beta t+\frac{\alpha A+B}{2\beta^3}(\sin\beta t-\beta t\cos\beta t)\right]$$
$$+\ e^{\alpha t}\left[C\cos\beta t+\frac{\alpha C+D}{\beta}\sin\beta t\right]$$

The constants are given by formulas similar to those above.

Solving Ordinary Differential Equations

One of the main uses of Laplace transforms is in solving differential equations, both ordinary and partial.

Example: Using Laplace transforms solve

$$y''+y=2\cos t,\qquad y(0)=2,\quad y'(0)=0.$$

Solution: Take Laplace transform of the differential equation. Define $Y(s)=\mathcal{L}(y)$.

$$\begin{aligned}\Big[s^2\mathcal{L}(y)-sy(0)-y'(0)\Big]+\mathcal{L}(y)\ &=\ \mathcal{L}(2\cos t),\\ \Rightarrow\quad (s^2+1)\mathcal{L}(y)-2s\ &=\ \frac{2s}{s^2+1}\\ \Rightarrow\quad \mathcal{L}(y)\ &=\ \frac{2s}{s^2+1}+\frac{2s}{(s^2+1)^2}\end{aligned}$$

We have a complex and repeated complex factor.

$$\mathcal{L}^{-1}\left(\frac{2s}{s^2+1}\right)=2\cos t.$$

$$\mathcal{L}^{-1}\left(\frac{2s}{(s^2+1)^2}\right)=t\sin t.$$

Therefore

$$y(t)=2\cos t+t\sin t.$$

TOPIC 8 – FOURIER SERIES

Periodic functions, trigonometric series

Definition: A function $f(x)$ is periodic if it is defined for all real x and if there is some positive number p such that

$$f(x+p)=f(x)\qquad\text{for all } x\ . \tag{23}$$

The number p is then called the period. From (23) we then have

$$f(x+2p)=f(x+p+p)=f(x+p)=f(x)$$

Thus for any integer n

$$f(x + np) = f(x) \qquad \text{for all } x.$$

If $f(x)$ and $g(x)$ have period p then so does

$$h(x) = af(x) + bg(x) \qquad (a, b \text{ constants}).$$

Simple examples:

$$\begin{aligned} f(x) &= \text{constant is periodic} \\ f(x) &= \sin x \text{ is periodic with period } 2\pi \\ f(x) &= x \text{ is } \textbf{not} \text{ periodic} \end{aligned}$$

Observe that a function which is constant has period p for any $p > 0$. If f is periodic but not constant, the smallest positive p for which (23) holds is called the **primitive period** of f.

We will want to represent functions of period $p = 2\pi$ in terms of simple trigonometric functions $1, \cos x, \sin x, \cos 2x, \sin 2x$, etc. Such a representation will give rise to a **trigonometric series**.

$$a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots \tag{24}$$

where a_0, a_1, b_1 etc are real constants—the a_n and b_n are called the **coefficients** of the series. We can rewrite (24) as

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

Each term in the series has period 2π. Hence if the series converges, its sum will be a function with period 2π.

Fourier series, Euler formulæ

Assume $f(x)$ is a periodic function of period 2π that can be represented by the trigonometric series

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{25}$$

i.e. assume the series converges and has $f(x)$ as its sum.

Question: How do we determine the coefficients a_n and b_n in (25) ?

First a_0. Integrate both sides of (25) from $-\pi$ to π.

$$\int_{-\pi}^{\pi} f(x)\, dx = \int_{-\pi}^{\pi} \left[a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \right] dx.$$

Assuming that term-by-term integration is valid, we have

$$\int_{-\pi}^{\pi} f(x)\,dx = a_0 \int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos nx\,dx + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \sin nx\,dx$$

The last two integrals in this expression are zero. Therefore

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\,dx \tag{26}$$

To determine a_m multiply (25) by $\cos mx$, where m is any positive integer, and integrate from $-\pi$ to π.

$$\int_{-\pi}^{\pi} f(x)\cos mx\,dx = \int_{-\pi}^{\pi} \left[a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)\right] \cos mx\,dx.$$

Integrating term by term gives

$$\int_{-\pi}^{\pi} f(x)\cos mx\,dx = a_0 \int_{-\pi}^{\pi} \cos mx\,dx + \sum_{n=1}^{\infty} \left\{ a_n \int_{-\pi}^{\pi} \cos nx \cos mx\,dx + b_n \int_{-\pi}^{\pi} \sin nx \cos mx\,dx \right\}$$

If we now make use of

$$\int_{-\pi}^{\pi} \cos nx \cos mx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} \cos(n+m)x\,dx + \frac{1}{2}\int_{-\pi}^{\pi} \cos(n-m)x\,dx$$

$$\int_{-\pi}^{\pi} \sin nx \cos mx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} \sin(n+m)x\,dx + \frac{1}{2}\int_{-\pi}^{\pi} \sin(n-m)x\,dx$$

and evaluating these integrals shows that all four integrals are zero except for

$$\frac{1}{2}\int_{-\pi}^{\pi} \cos(n-m)x\,dx = \begin{cases} 0, & \text{if } m \neq n; \\ \pi, & \text{if } m = n. \end{cases}$$

We then have

$$\int_{-\pi}^{\pi} f(x)\cos mx\,dx = \pi a_m$$

or

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos mx\,dx \qquad (m = 1, 2, \ldots) \tag{27}$$

Finally to obtain the coefficients b_m we multiply (25) by $\sin mx$ (where m is a positive integer) and integrate from $-\pi$ to π. Following the above procedure we obtain

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx \, dx \qquad (m = 1, 2, \ldots) \tag{28}$$

To summarise

$$\left.\begin{aligned} a_0 &= \tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx \\ a_m &= \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx \, dx \qquad (m = 1, 2, \ldots) \\ b_m &= \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx \, dx \qquad (m = 1, 2, \ldots) \end{aligned}\right\} \tag{29}$$

The expressions (29) are known as the **Euler formulæ**. In these integrals, the limits need not be $-\pi$ and π. Since everything has period 2π, we may take the limits to be α and $\alpha + 2\pi$ for any real number α. In particular, it is often convenient to take them to be 0 and 2π. The a_0, a_n, b_n are called the **Fourier coefficients** of $f(x)$ and the series

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \tag{30}$$

is called the **Fourier series** of $f(x)$.

Functions of arbitrary period

Suppose $f(x) = f(x+2L)$ for all x, *i.e.*, f has period $2L$. Then its Fourier series is of the form

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L})$$

with the Fourier coefficients given by the Euler formulæ

$$\begin{aligned} a_0 &= \frac{1}{2L} \int_{-L}^{L} f(x)\, dx \\ a_m &= \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{m\pi x}{L} \, dx \qquad (m = 1, 2, \ldots) \\ b_m &= \frac{1}{L} \int_{-L}^{L} f(x) \sin \frac{m\pi x}{L} \, dx \qquad (m = 1, 2, \ldots) \end{aligned}$$

The proof follows our earlier derivation of the Euler formulæ for 2π–periodic functions. We can also prove these results by a change of scale. Set $v = \pi x/L$.

Then $x = \pm L \Rightarrow v = \pm\pi$ and $f(x) = g(v)$ where $g(v)$ has period of 2π in v. So $g(v)$ has Fourier series

$$g(v) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nv + b_n \sin nv)$$

with coefficients given by Euler formulæ (29). Changing variables back to x gives the required result.

Convergence of Fourier Series

There are a great many theorems on the convergence of Fourier Series. One of the most straightforward is:
Theorem: If f has period $2L$ and is piecewise continuous and has a left-hand and a right-hand derivative at each point then the Fourier series converges. The sum of the series is $f(x_0)$ if f is continuous at $x = x_0$ and $(f(x_0+) + f(x_0-))/2$ if f is discontinuous at $x = x_0$.

Sums of functions

(1) The Fourier coefficients of a sum $f_1 + f_2$ are the sums of the corresponding Fourier coefficients of f_1 and f_2.
(2) The Fourier coefficients of cf are c times the Fourier coefficients of f.

Even and Odd functions

The above Fourier series simplify somewhat if $f(x)$ is either an even function or an odd function. Remember that $y = g(x)$ is **even** if $g(-x) = g(x)$ for all x and $y = h(x)$ is **odd** if $h(-x) = -h(x)$ for all x.

If $g(x)$ is an even function then

$$\int_{-L}^{L} g(x)\,dx = \int_{0}^{L} g(x)\,dx + \int_{-L}^{0} g(x)\,dx$$

Put $v = -x$ in the second integral to give

$$\begin{aligned}\int_{-L}^{L} g(x)\,dx &= \int_{0}^{L} g(x)\,dx + \int_{L}^{0} (-g(-v))\,dv \\ &= \int_{0}^{L} g(x)\,dx + \int_{0}^{L} g(v)\,dv \\ &= 2\int_{0}^{L} g(x)\,dx\end{aligned}$$

Similarly if $h(x)$ is an odd function

$$\int_{-L}^{L} h(x)\,dx = \int_{0}^{L} h(x)\,dx + \int_{-L}^{0} h(x)\,dx$$

$$
\begin{aligned}
&= \int_0^L h(x)\,dx + \int_L^0 (-h(-v))\,dv \\
&= \int_0^L h(x)\,dx - \int_0^L h(v)\,dv \\
&= 0
\end{aligned}
$$

Fourier series for even and odd functions

Using the above results, we may readily show that the Fourier series of an even function $f(x)$ of period $2L$ is a **Fourier cosine series**

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L}$$

with coefficients

$$a_0 = \frac{1}{L}\int_0^L f(x)\,dx \qquad a_m = \frac{2}{L}\int_0^L f(x)\cos\frac{m\pi x}{L}x \qquad (m = 1, 2, \ldots)$$

The Fourier series of an odd function $f(x)$ of period $2L$ is a **Fourier sine series**

$$f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}$$

with coefficients

$$b_m = \frac{2}{L}\int_0^L f(x)\sin\frac{m\pi x}{L}dx \qquad (m = 1, 2, \ldots)$$

Half–Range expansions

Suppose we have a function $f(x)$ defined on some interval $0 \le x \le L$ and we want to find the Fourier series of this function. By taking period $2L$ we can get

(i) A Fourier cosine series for $f(x)$ by extending it to an even function on the interval $-L \le x \le L$. The cosine half–range expansion for $f(x)$ is thus

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L}$$

where

$$a_0 = \frac{1}{L}\int_0^L f(x)\,dx \qquad a_m = \frac{2}{L}\int_0^L f(x)\cos\frac{m\pi x}{L}dx \qquad (m = 1, 2, \ldots)$$

(ii) A Fourier sine series for $f(x)$ by extending it to an odd function on the interval $-L \le x \le L$. The sine half–range expansion for $f(x)$ is thus

$$f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}$$

where

$$b_m = \frac{2}{L} \int_0^L f(x) \sin \frac{m\pi x}{L} dx \qquad (m = 1, 2, \ldots)$$

In $(0, L)$ both these series represent $f(x)$. Outside $(0, L)$ they represent different functions – **the even and odd periodic extensions** of f respectively.

Complex form of the Fourier series

Recall that

$$e^{i\theta} = \cos\theta + i \sin\theta \tag{31}$$

Taking the complex conjugate gives

$$e^{-i\theta} = \cos\theta - i \sin\theta \tag{32}$$

Then (31)+(32) and (31)-(32) give

$$\cos\theta = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right) \qquad \sin\theta = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right).$$

Hence

$$\cos nx = \frac{1}{2}\left(e^{inx} + e^{-inx}\right) \qquad \sin nx = \frac{1}{2i}\left(e^{inx} - e^{-inx}\right).$$

Now consider the Fourier series

$$\begin{aligned} f(x) &= a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \\ &= a_0 + \sum_{n=1}^{\infty} \frac{a_n}{2}\left(e^{inx} + e^{-inx}\right) + \frac{b_n}{2i}\left(e^{inx} - e^{-inx}\right) \\ &= a_0 + \sum_{n=1}^{\infty} \frac{1}{2}(a_n - ib_n)e^{inx} + \frac{1}{2}(a_n + ib_n)e^{-inx} \\ \Rightarrow f(x) &= c_0 + \sum_{n=1}^{\infty} \left(c_n e^{inx} + k_n e^{-inx}\right) \end{aligned}$$

where

$$c_n = \frac{1}{2}(a_n - ib_n), \qquad k_n = c_n^*.$$

Remember that

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \end{aligned}$$

Hence

$$\begin{aligned} c_n &= \frac{1}{2}\left\{\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx - \frac{i}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx\right\} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(\cos nx - i\sin nx)\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx \end{aligned}$$

Similarly

$$k_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}dx$$

and note that $k_n = c_{-n}$. Then we have

$$\begin{aligned} f(x) &= c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=1}^{\infty} c_{-n}e^{-inx} \\ &= c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=-1}^{-\infty} c_n e^{inx} \end{aligned}$$

Finally noting that $e^{i(0)x} = 1$ we have

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

with

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

This is then the **complex form** of the Fourier series for $f(x)$. The c_n are the complex Fourier coefficients.

Forced Oscillations

Earlier we saw that forced oscillations of a body of mass m on a spring are governed by the equation

$$my'' + cy' + ky = r(t)$$

where k is the spring modulus, c the damping constant. If $r(t)$ is a cos or sin we can solve this equation by the **method of undetermined coefficients**.

What if $r(t)$ is some other periodic function? We can represent $r(t)$ by a Fourier series; we should then be able to find a particular solution as a similar Fourier series. To fix our ideas we consider a simple example

Example: Find the general solution to

$$y'' + \omega^2 y = r(t)$$

where

$$r(t) = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \sin(2n-1)t.$$

Solution: We have

$$y'' + \omega^2 y = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \sin(2n-1)t \tag{33}$$

Consider the equation

$$y_n'' + \omega^2 y_n = \frac{1}{n^2} \sin nt \quad (n = 1, 3, 5, \ldots) \tag{34}$$

Find the general solution of this equation. The general solution of the homogeneous form of (33) is

$$y_h = A \cos \omega t + B \sin \omega t.$$

For a particular solution of (34) try

$$y_n = A_n \cos nt + B_n \sin nt.$$

Differentiating and substituting into (34) gives

$$(-n^2 + \omega^2) A_n \cos nt + (-n^2 + \omega^2) B_n \sin nt = \frac{1}{n^2} \sin nt$$

(assuming $\omega \neq n$ for n odd) we have

$$A_n = 0, \quad B_n = \frac{1}{n^2(\omega^2 - n^2)}.$$

Thus the particular solution of (34) is

$$y_n = \frac{1}{n^2(\omega^2 - n^2)} \sin nt.$$

Since (33) is **linear** the general solution is then a superposition

$$y_1 + y_3 + y_5 + \cdots + y_h$$

Therefore

$$y = A \cos \omega t + B \sin \omega t + \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2[\omega^2 - (2n-1)^2]} \sin(2n-1)t.$$

TOPIC 9 – PARTIAL DIFFERENTIAL EQUATIONS

Basic concepts

An equation involving one or more partial derivatives of an (unknown) function of two or more independent variables is called a **partial differential equation**.

- The order of the highest derivative is called the **order** of the equation.
- A p.d.e (partial differential equation) is **linear** if it is of first degree in the dependent variable (unknown function) and its partial derivatives.
- A partial differential equation is **homogeneous** if each term contains either the dependent variable or one of its derivatives; otherwise it is **inhomogeneous**.
- A **solution** of a p.d.e.. in some region R of the space of the independent variables is a function that has all the partial derivatives appearing in the equation in some domain containing R and satisfies the equation everywhere in R.

Some examples

(1) $\dfrac{\partial^2 u}{\partial t^2} = c^2 \dfrac{\partial^2 u}{\partial x^2}$ one dimensional wave equation

(2) $\dfrac{\partial u}{\partial t} = c^2 \dfrac{\partial^2 u}{\partial x^2}$ one dimensional heat equation

(3) $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = 0$ two dimensional Laplace equation

(4) $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = f(x, y)$ two dimensional Poisson equation

(1)-(3) are homogeneous, linear. (4) is inhomogeneous, linear. All are second order.

We may, depending on the problem, have **boundary conditions** (the solution has some given value on the boundary of some domain) or **initial conditions** (where the value of the solution will be given at some initial time, *e.g.* $t = 0$).
Superposition Theorem: If u_1 and u_2 are any solutions of a linear homogeneous p.d.e in some region, then

$$u = c_1 u_1 + c_2 u_2$$

is also a solution of the equation in the region.

Vibrating Strings

As shown in Section 2 of Chapter 11 of the book by Kreyszig, the small transverse oscillations of a tightly stretched string satisfy

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad \text{where} \quad c^2 = T/\rho,$$

T is the tension in the string, ρ is the constant mass per unit length of the string and $u(x, t)$ is the displacement of the string.

D'Alembert's solution of the wave equation

The equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{35}$$

is called the one-dimensional wave equation.

Introduce new variables

$$v = x + ct, \quad z = x - ct$$

then u in (35) becomes a function of v and z. By the chain rule we have

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x}$$

Thus

$$u_x = u_v + u_z$$

Then

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial v}\left(\frac{\partial u}{\partial v} + \frac{\partial u}{\partial z}\right)\frac{\partial v}{\partial x} + \frac{\partial}{\partial z}\left(\frac{\partial u}{\partial v} + \frac{\partial u}{\partial z}\right)\frac{\partial z}{\partial x}$$
$$\Rightarrow u_{xx} = u_{vv} + 2u_{vz} + u_{zz}.$$

Similarly

$$u_{tt} = c^2(u_{vv} - 2u_{vz} + u_{zz}).$$

Combining these in (35) gives

$$c^2(u_{vv} - 2u_{vz} + u_{zz}) = c^2(u_{vv} + 2u_{vz} + u_{zz}).$$

Thus

$$u_{vz} = 0. \tag{36}$$

Integrating (36) gives

$$\begin{aligned} \frac{\partial u}{\partial v} &= h(v) \\ \Rightarrow u &= \int h(v)dv + \psi(z) \end{aligned}$$

i.e.

$$u = \phi(v) + \psi(z)$$

where $\phi = \int h(v)dv$. Thus

$$u = \phi(x + ct) + \psi(x - ct) \tag{37}$$

where ϕ and ψ are arbitrary functions. This is D'Alembert's solution of the wave equation and is valid for c constant.

The functions ϕ and ψ can be determined from the initial conditions. Suppose we have to solve (35) for $-\infty < x < \infty$ and $t > 0$ subject to the initial conditions

$$u(x,0) = f(x), \qquad \frac{\partial u(x,0)}{\partial t} = g(x) \qquad \text{for} \quad -\infty < x < \infty$$

From (37) we have

$$\frac{\partial u}{\partial t} = c\phi'(x+ct) - c\psi'(x-ct)$$

where the $'$ denotes differentiation with respect to the entire argument $x+ct$ and $x-ct$ respectively. Then

$$c\phi'(x) - c\psi'(x) = g(x) \tag{38}$$
$$\phi(x) + \psi(x) = f(x) \tag{39}$$

We consider only the case $g(x) = 0$. From (39) we have

$$c\phi'(x) - c\psi'(x) = 0 \quad \Rightarrow \quad \phi(x) = \psi(x) + k \text{ (constant)}$$

And from (39)

$$2\psi(x) = f(x) - k$$
$$\Rightarrow \quad \psi = \frac{1}{2}\left[f(x) - k\right], \quad \phi = \frac{1}{2}\left[f(x) + k\right]$$

Hence

$$u(x,t) = \frac{1}{2}\left[f(x+ct) + f(x-ct)\right].$$

If $g(x) \neq 0$ then

$$u(x,t) = \frac{1}{2}\left[f(x+ct) + f(x-ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)ds.$$

Note: In the solution (37), $\psi(x-ct)$ represents a **disturbance travelling to the right with speed c without changing its shape** and $\phi(x+ct)$ represents a **disturbance travelling to the left with speed c without changing its shape**.

SEPARATION OF VARIABLES

This is an important technique which can be used to find the solutions of many linear partial differential equations and is best illustrated by examples.

Wave equation

The one–dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{40}$$

(*e.g.* a model of the deflection $u(x,t)$ of an elastic string) can be solved by this technique. Physically, conditions will be imposed on the solution of (40).
Boundary conditions (string fixed at ends $x = 0$ and $x = L$). Then

$$u(0,t) = 0, \quad u(L,t) = 0 \qquad \text{for all time } t \tag{41}$$

Initial conditions (the motion of the string will depend upon its **initial deflection** and its **initial velocity** *i.e.* the deflection and velocity at $t = 0$).

$$\text{initial deflection} \qquad u(x,0) = f(x) \tag{42}$$

$$\text{initial velocity} \qquad \left.\frac{\partial u}{\partial t}\right|_{t=0} = g(x) \tag{43}$$

The problem is to find the solution of (40) satisfying the conditions (41), (42) and (43) .
This a 3 step process.

Step 1. Apply the **method of separation of variables** to obtain two ordinary differential equations.
Step 2. Determine the solution of these two equations that satisfy the boundary conditions.
Step 3. Combine these solutions so that the result will be a solution of (40) which also satisfies the initial conditions (42) and (43).
Step 1: Separation of variables.

We look for a solution of (40) in the form

$$u(x,t) = F(x)G(t).$$

Differentiating gives

$$\frac{\partial^2 u}{\partial t^2} = F\frac{d^2G}{dt^2}, \quad \frac{\partial^2 u}{\partial x^2} = G\frac{d^2F}{dx^2}$$

Substituting into equation (40) gives

$$F\frac{d^2G}{dt^2} = c^2 G\frac{d^2F}{dx^2}$$

rearranging (divide by c^2FG)

$$\frac{1}{c^2G}\frac{d^2G}{dt^2} = \frac{1}{F}\frac{d^2F}{dx^2}$$

$$\Rightarrow \frac{1}{c^2G}\frac{d^2G}{dt^2} = \frac{1}{F}\frac{d^2F}{dx^2} = k \quad (k \text{ constant}).$$

This then gives us two ordinary differential equations:

$$\frac{d^2F}{dx^2} - kF = 0, \quad \frac{d^2G}{dt^2} - kc^2G = 0\ . \tag{44}$$

Step 2: Satisfy the boundary conditions. With $u = FG$

$$u(0,t) = F(0)G(t) = 0 \quad u(L,t) = F(L)G(t) = 0 \qquad \text{for all } t$$

There are then two possibilities $G(t) = 0$ for all $t \Rightarrow u \equiv 0$. Or

$$F(0) = F(L) = 0 \tag{45}$$

We then need to solve the first of (44) subject to (45).
Cases: (**A**) $k = 0 \Rightarrow F = ax + b$ but (45) $\Rightarrow a = b = 0$. Hence $F \equiv 0 \Rightarrow u \equiv 0$ (trivial solution).

(**B**) $k > 0$. Take $k = \mu^2$. Then

$$F = Ae^{\mu x} + Be^{-\mu x}$$

But (45) $\Rightarrow A = B = 0$. Hence $F \equiv 0 \Rightarrow u \equiv 0$ (trivial solution).

(**C**) $k < 0$. Take $k = -p^2$ where ,without loss of generality, $p > 0$. Then

$$F = C\cos px + D\sin px$$

Boundary conditions

$$F(0) = 0 \quad \Rightarrow \quad C = 0$$
$$F(L) = 0 \quad \Rightarrow \quad D\sin pL = 0 \quad \Rightarrow \quad \sin pL = 0 \quad \Rightarrow \quad p = \frac{n\pi}{L} \quad (n \text{ positive integer})$$

So $p = n\pi/L$ and D is arbitrary (take $D = 1$). We then have an infinite number of solutions

$$F_n(x) = \sin\frac{n\pi x}{L} \qquad \text{for} \qquad n = 1, 2, 3, \cdots$$

Now with $k = -p^2 = -(n\pi/L)^2$ we have

$$\frac{d^2G}{dt^2} + \left(\frac{cn\pi}{L}\right)^2 G = 0.$$

which has the general solution

$$G_n(t) = B_n\cos\lambda_n t + C_n\sin\lambda_n t \quad \text{for} \quad n = 1, 2, 3, \cdots$$

where $\lambda_n = cn\pi/L$. Hence

$$u_n(x,t) = F_n(x)G_n(t)$$

or

$$u_n(x,t) = (B_n\cos\lambda_n t + C_n\sin\lambda_n t)\sin\frac{n\pi x}{L} \quad \text{for} \quad n = 1, 2, 3, \cdots$$

are solutions of (40) satisfying the boundary conditions (41) . We call the u_n's the **eigenfunctions** and the λ_n's the **eigenvalues**. The set $\{\lambda_1, \lambda_2, \cdots\}$ is called the **spectrum**.

Step 3: Satisfy the initial conditions. The initial conditions are given by (42) and (43) . Since the wave equation is linear and homogeneous, a sum of solutions is also a solution. Consider

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} (B_n \cos \lambda_n t + C_n \sin \lambda_n t) \sin \frac{n\pi x}{L}$$

with $\lambda_n = cn\pi/L$. From (42) we have

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L}$$

This is just the Fourier–sine half–range expansion for $f(x)$. Thus

$$B_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} dx \quad (\text{for} \quad n = 1, 2, \ldots) \tag{46}$$

We have

$$\frac{\partial u}{\partial t} = \sum_{n=1}^{\infty} \lambda_n (-B_n \sin \lambda_n t + C_n \cos \lambda_n t) \sin \frac{n\pi x}{L}$$

From (43)

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = g(x) = \sum_{n=1}^{\infty} \lambda_n C_n \sin \frac{n\pi x}{L}$$

Again, this is the Fourier–sine half–range expansion for $g(x)$. Thus

$$\lambda_n C_n = \frac{2}{L} \int_0^L g(x) \sin \frac{n\pi x}{L} dx$$

or

$$C_n = \frac{2}{cn\pi} \int_0^L g(x) \sin \frac{n\pi x}{L} dx \quad (\text{for} \quad n = 1, 2, \ldots) \tag{47}$$

Summary

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

$$u(x,0) = f(x) \qquad \left.\frac{\partial u}{\partial t}\right|_{t=0} = g(x)$$

has solution (see note below)

$$u(x,t) = \sum_{n=1}^{\infty} (B_n \cos \lambda_n t + C_n \sin \lambda_n t) \sin \frac{n\pi x}{L}$$

where

$$B_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}dx \quad (\text{for} \quad n = 1, 2, \ldots)$$

$$C_n = \frac{2}{cn\pi}\int_0^L g(x)\sin\frac{n\pi x}{L}dx \quad (\text{for} \quad n = 1, 2, \ldots)$$

Note: it is a solution if the infinite series converges and the series obtained by differentiating twice with respect to x and t converges with sums u_{xx} and u_{tt} (which are continuous).

For the case $g(x) = 0$, we have $C_n = 0$ and

$$u(x,t) = \sum_{n=1}^{\infty} B_n \cos\lambda_n t \sin\frac{n\pi x}{L}$$

We can write

$$\cos\frac{cn\pi t}{L}\sin\frac{n\pi x}{L} = \frac{1}{2}\left\{\sin\left[\frac{n\pi}{L}(x-ct)\right] + \sin\left[\frac{n\pi}{L}(x+ct)\right]\right\}$$

Then

$$u(x,t) = \frac{1}{2}\sum_{n=1}^{\infty} B_n \sin\left[\frac{n\pi}{L}(x-ct)\right] + \frac{1}{2}\sum_{n=1}^{\infty} B_n \sin\left[\frac{n\pi}{L}(x+ct)\right]$$

but

$$\sum_{n=1}^{\infty} B_n \sin\frac{n\pi x}{L} = f(x)$$

(the Fourier sine series for $f(x)$). Thus

$$u(x,t) = \frac{1}{2}\left\{f^*(x-ct) + f^*(x+ct)\right\}$$

where f^* is the odd periodic extension of f with period $2L$. Since $f(x)$ is continuous on $0 \le x \le L$ and zero at endpoints $\Rightarrow u(x,t)$ is continuous for all x, t. Differentiating $\Rightarrow u(x,t)$ is a solution of (40) (provided f'' exists for $0 < x < L$ and f has one–sided derivatives at $x = 0, L$). Under these conditions $u(x,t)$ is a solution of (40) satisfying the conditions (41), (42) and (43) .

Note: compare the above solution with D'Alembert's solution.

Heat equation

The heat flow in a body of homogeneous material is governed by the heat equation

$$\frac{\partial u}{\partial t} = c^2\nabla^2 u \qquad c^2 = \frac{\kappa}{\sigma\rho}.$$

where $u(x,y,z,t)$ is the temperature in the body, κ =thermal conductivity, σ =specific heat and ρ =density. Here

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

Consider the temperature in a long thin wire of constant cross section (homogeneous material) that is perfectly insulated laterally (so that heat flows only in the x direction). Then u depends only on x and t. The heat equation becomes

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{48}$$

Consider the case when the ends of the bar are kept at zero temperature and suppose the initial temperature is $f(x)$. This then gives us the conditions
boundary conditions:

$$u(0,t) = u(L,t) = 0 \quad \text{for} \quad t > 0 \tag{49}$$

initial conditions:

$$u(x,0) = f(x) \quad \text{for} \quad 0 < x < L \tag{50}$$

We want to determine the solution of (48) subject to (49) and (50). Use separation of variables.
Step 1: Separation of variables. Look for a solution in the form

$$u(x,t) = F(x)G(t) \tag{51}$$

Substituting into (48) gives

$$\frac{1}{c^2 G}\frac{dG}{dt} = \frac{1}{F}\frac{d^2F}{dx^2} = -p^2 = \text{ constant.}$$

Thus

$$\frac{dG}{dt} + c^2 p^2 G = 0 \tag{52}$$

$$\frac{d^2F}{dx^2} + p^2 F = 0 \tag{53}$$

Step 2: Satisfy the boundary conditions (49).

$$(49), (51) \Rightarrow F(0) = F(L) = 0. \tag{54}$$

Solving (53) subject to (54) gives

$$\begin{aligned} F(x) &= A\cos px + B\sin px \\ F(0) = 0 \quad &\Rightarrow \quad A = 0 \\ F(L) = 0 \quad &\Rightarrow \quad B\sin pL = 0 \quad \Rightarrow \quad p = \frac{n\pi}{L} \end{aligned}$$

(setting $B = 1$ we have)

$$F_n = \sin \frac{n\pi x}{L} \quad (n = 1, 2, \ldots).$$

Now equation (53) becomes

$$\frac{dG}{dt} + \frac{c^2 n^2 \pi^2}{L^2} G = 0$$

which has the general solution

$$G_n = B_n e^{-\lambda_n^2 t}, \qquad \lambda_n = cn\pi/L.$$

So the functions

$$u_n = B_n \sin \frac{n\pi x}{L} e^{-\lambda_n^2 t}.$$

are solutions of (48) satisfying (49).
Step 3: Satisfy the initial conditions (50). As the heat equation is linear

$$u(x,t) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} e^{-\lambda_n^2 t}, \tag{55}$$

is also a solution. But

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L}$$

which is just the half–range expansion for $f(x)$ (*i.e.* the Fourier sine series). Hence

$$B_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} dx \tag{56}$$

So (55) with coefficients (56) is the solution to the heat equation (48) satisfying (49) and (50) (provided the series converges). Note: The presence of the $e^{-\lambda_n^2 t}$ terms in (55) means that all terms in (55) approach zero as $t \longrightarrow \infty$ (*i.e.* the temperature tends to zero as $t \longrightarrow \infty$).

Steady state two–dimensional heat flow

If we consider a steady flow of heat, the 2D heat equation reduces to Laplace's equation

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{57}$$

Solution involves solving (57) in some region R of the xy-plane with a given boundary condition of the boundary curve of R—a boundary value problem. Such a problem is called
Dirichlet problem if u is prescribed on C (the boundary curve of R).
Neumann problem if the normal derivative $u_n = \frac{\partial u}{\partial n}$ is prescribed on C.

Mixed problem if u is prescribed on a part of C and u_n on the rest of C.

We will consider only the Dirichlet problem on the rectangle $0 \le x \le a$, $0 \le y \le b$ with $u = 0$ along the boundary except for the edge $y = b$ where $u(x, b) = f(x)$ for $0 \le x \le a$.

To solve (57) we use separation of variables. Substituting $u(x, y) = F(x)G(y)$ into (57) gives

$$\frac{1}{F}\frac{d^2F}{dx^2} = -\frac{1}{G}\frac{d^2G}{dy^2} = -k = (\text{constant}).$$

Our boundary conditions give

$$F(0) = 0, \quad F(a) = 0, \quad G(0) = 0$$

Solving for F gives

$$F_n(x) = \sin\frac{n\pi x}{a} \qquad k = (n\pi/a)^2$$

Solving for G subject to $G(0) = 0$ gives

$$G_n = 2A_n \sinh\frac{n\pi y}{a}$$

Finally we need to satisfy the boundary condition

$$u(x, b) = f(x).$$

Write

$$u(x, y) = 2\sum_{n=1}^{\infty} A_n \sin\frac{n\pi x}{a}\sinh\frac{n\pi y}{a}$$

Then

$$\begin{aligned} f(x) &= 2\sum_{n=1}^{\infty} A_n \sin\frac{n\pi x}{a}\sinh\frac{n\pi b}{a} \\ &= \sum_{n=1}^{\infty} C_n \sin\frac{n\pi x}{a} \end{aligned}$$

where

$$C_n = 2A_n \sinh\frac{n\pi b}{a}.$$

These are the Fourier coefficients of $f(x)$

$$2A_n \sinh\frac{n\pi b}{a} = \frac{2}{a}\int_0^a f(x)\sin\frac{n\pi x}{a}dx$$

Thus

$$A_n = \frac{1}{a\sinh\frac{n\pi b}{a}}\int_0^a f(x)\sin\frac{n\pi x}{a}dx$$

Therefore the solution of the problem is

$$u(x, y) = 2\sum_{n=1}^{\infty} A_n \sin\frac{n\pi x}{a}\sinh\frac{n\pi y}{a}$$

with the A_n's given above.