Mathematics 2

Mohamed KHALEAOUI

Saturday 4$^{\text{th}}$ May, 2024

Contents

Chapter 1

Matrices and determinants

1.1 Introduction to Matrices

Definition 1.1. A rectangular arrangement of mn numbers, in m rows and n columns and enclosed within a bracket is called a matrix. We shall denote matrices by capital letters as A, B, C etc.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})_{m\times n}$$

A is a matrix of order $m \times n$. The element of the i^{th} row and j^{th} column denoted by a_{ij}.

Remark: A matrix is not just a collection of elements but every element has assigned a definite position in a particular row and column.

1.1.1 Special Types of Matrices (with examples)

a **Square matrix:** A matrix in which numbers of rows are equal to number of columns is called a square matrix.

$$A = \begin{pmatrix} 5 & 3 & -1 \\ 2 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

b **Diagonal matrix:** A square matrix D is called a diagonal matrix if each of its non-diagonal element is zero. That is $a_{ij} = 0$ if $i \neq j$ and at least one element $aii \neq 0$.

$$D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

c **Identity Matrix:** A diagonal matrix whose diagonal elements are equal to 1 is called identity matrix and denoted by I_n. That is $a_{ij} = 0$ if $i \neq j$ and $a_{ii} = 1$.

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(d) **Upper Triangular matrix:** A square matrix said to be a Upper triangular matrix if $a_{ij} = 0$ for $i > j$

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(e) **Lower Triangular matrix:** A square matrix said to be a Lower triangular matrix if $a_{ij} = 0$ for $i < j$

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 9 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & 3 & 2 \end{pmatrix}$$

(f) **Symmetric Matrix:** A square matrix S of order n said to be a symmetric if $a_{ij} = a_{ji}$ for all $i \neq j$.

$$S = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & 6 \\ 3 & 6 & 5 \end{pmatrix}$$

(g) **Skew- Symmetric Matrix:** A square matrix S of order n said to be a skew-symmetric if $a_{ij} = -a_{ji}$ for all $i \neq j$.

$$S = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & 6 \\ -3 & -6 & 5 \end{pmatrix}$$

(h) **Zero Matrix:** A matrix whose all elements are zero is called as Zero Matrix of order $n \times m$. Zero matrix denoted by $O_{n\times m}$.

$$O = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(i) **Row Vector:** A matrix consists a single row is called as a row vector or row matrix.

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 0 & -1 & -1 \end{pmatrix}$$

(j) **Column Vector:** A matrix consists a single column is called a column vector or column matrix.

$$V = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$$

1.2 Matrix Algebra

1.2.1 Equality of two matrices

Two matrices A and B are said to be equal if

(1) They are of same order.

(2) Their corresponding elements are equal.

That is if $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ then $a_{ij} = b_{ij}$ for all i and j.

1.2.2 Scalar multiple of a matrix

Let k be a scalar, then the scalar product of matrix $A = (a_{ij})_{m\times n}$ by k, denoted by kA, and given by $kA = (ka_{ij})_{m\times n}$ or

$$kA = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}$$

1.2.3 Addition of two matrices

Let $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ are two matrices with same order then sum of the two matrices are given by

$$A + B = (a_{ij})_{m\times n} + (b_{ij})_{m\times n} = (a_{ij} + b_{ij})_{m\times n}$$

1.2.4 Multiplication of two matrices

Two matrices A and B are said to be confirmable for product AB if number of columns in A equals to the number of rows in matrix B. Let $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{n\times r}$ be two matrices the product matrix $C = AB = (c_{ij})_{m\times r}$, is matrix of order $m \times r$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots a_{in}b_{nj}$$

Example 1.2.

$$\begin{bmatrix} 2 & 1 & 0 \\ 5 & 7 & 2 \\ 1 & 3 & 1 \\ 11 & 3 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 6 & 7 & 3 \\ 2 & -4 & 5 \end{bmatrix} =$$

$$\begin{bmatrix} 2\cdot 1 + 1\cdot 6 + 0\cdot 2 & 2\cdot 0 + 1\cdot 7 + 0\cdot(-4) & 2\cdot(-1) + 1\cdot 3 + 0\cdot 5 \\ 5\cdot 1 + 7\cdot 6 + 2\cdot 2 & 5\cdot 0 + 7\cdot 7 + 2\cdot(-4) & 5\cdot(-1) + 7\cdot 3 + 2\cdot 5 \\ 1\cdot 1 + 3\cdot 6 + 1\cdot 2 & 1\cdot 0 + 3\cdot 7 + 1\cdot(-4) & 1\cdot(-1) + 3\cdot 3 + 1\cdot 5 \\ 11\cdot 1 + 3\cdot 6 + 3\cdot 2 & 11\cdot 0 + 3\cdot 7 + 3\cdot(-4) & 11\cdot(-1) + 3\cdot 3 + 3\cdot 5 \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 7 & 1 \\ 51 & 41 & 26 \\ 21 & 17 & 13 \\ 35 & 9 & 13 \end{bmatrix}$$

1.2.5 Integral power of Matrices

Let A be a square matrix of order n, and m be positive integer then we define

$$A^m = A \times A \times \cdots \times A \quad \text{(m times multiplication).}$$

1.2.6 Properties of the Matrices

Let A, B and C are three matrices and α and β are scalars then

1. $A + (B + C) + (A + B) + C$.
2. $\beta(A + B) = \beta A + \beta B$.
3. $\alpha(\beta A) = (\alpha\beta)A$.
4. $\alpha(AB) = (\alpha A)B$.
5. $(AB)C = A(BC)$.
6. $A(B + C) = AB + AC$.

1.2.7 Transpose

The transpose of matrix $A = (a_{ij})_{m\times n}$, written $A^t(A'$ or $A^T)$ is the matrix obtained by writing the rows of A in order as columns. That is $A^T = (a_{ji})_{n\times m}$.

Properties of Transpose

(a) $(A + B)^T = A^T + B^T$.

(b) $\left(A^T\right)^T = A$.

(c) $(\beta A)^T = \beta A^T$, for any scalar β.

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

Proof.

(a) Let a_{ij} and b_{ij} are the $(i,j)^{th}$ element of the matrix A and B respectively. Then $a_{ij}+b_{ij}$ is the $(i,j)^{th}$ element of matrix $A+B$ and it is the $(j,i)^{th}$ element of the matrix $(A+B)^T$. Also a_{ij} and b_{ij} are the $(j,i)^{th}$ element of the matrixA^T and B^T respectively. Therefore $a_{ij}+b_{ij}$ is the $(j,i)^{th}$ element of the matrix A^T+B^T.

(b) Let $(i,j)^{th}$ element of the matrix A is a_{ij}, so it is $(j,i)^{th}$ element of A^T then it is $(i,j)^{th}$ element of the matrix$\left(A^T\right)^T$.

(c) Trivial.

(d) $c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots a_{in}b_{nj}$ is the $(i,j)^{th}$ element of the AB. It is result of the multiplication of the i^{th} row and j^{th} column and it is $(j,i)^{th}$ element of the matrix $(AB)^T$. The $(j,i)^{th}$ element of B^TA^T is the multiplication of j^{th} row of with i^{th} column of A^T, That is j^{th} column of B with i^{th} row of A.

□

Definition 1.3. A square matrix A is said to be

(1) Symmetric if $A=A^T$.

(2) Skew- symmetric if $A=-A^T$.

Properties.

(a) AA^T andA^TA are both symmetric.

(b) $A+A^T$ is a symmetric matrix.

(c) $A-A^T$ is a skew-symmetric matrix.

(d) If A is a symmetric matrix and m is any positive integer then A^m is also symmetric.

(e) If A is skew symmetric matrix then odd integral powers of A is skew symmetric, while positive even integral powers of A is symmetric.

(f) $AB+BA$ is symmetric for any symmetric matrices A and B.

(g) $AB-BA$ is skew-symmetric for any symmetric matrices A and B.

Exercise 1.4. Verify the (a) , (b) and (c) using the following matrix

$$A=\begin{pmatrix}1 & 3 & 5\\ -3 & -5 & 10\\ 1 & 8 & 9\end{pmatrix}$$

1.3 Determinant, Minor and Adjoint Matrices

Definition 1.5. Let $A = (a_{ij})_{n\times n}$ be a square matrix of order n, then the number $|A|$ called determinant of the matrix A defined by

1. If A of order 2, then

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

2. If A of order 3, then

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Remark 1.6. The determinant of a matrix of order 3 is

$$\left|\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right. = \begin{array}{l} (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) \\ -(a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}) \end{array}$$

Exercise 1.7. Calculate the determinants of the following matrices

1. $A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 6 & 8 \\ 1 & 9 & 5 \end{pmatrix}$

2. $B = \begin{pmatrix} 2 & -3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix}$

1.3.1 Properties of the Determinant

a. The determinant of a matrix A and its transpose A^T are equal: $|A^T| = |A|$.

b. Let A be a square matrix

   i. If A has a row (column) of zeros then $|A| = 0$.

   ii. If A has two identical rows (or columns) then $|A| = 0$.

c. If A is triangular matrix of order n then $|A|$ is product of the diagonal elements, i.e. $|A| = \prod_{k=1}^{n} a_{kk} = a_{11}a_{22}\cdots a_{nn}$.

d. If A is a square matrix of order n and k is a scalar then $|kA| = k^n|A|$.

1.3.2 Singular Matrix

Let A be a square matrix of order n, A is called singular matrix when $|A| = 0$ and non-singular otherwise.

1.3.3 Minor and Cofactors

Let $A = (a_{ij})_{n\times n}$ be a square matrix. Then M_{ij} denote a sub matrix of A with order $(n-1)\times(n-1)$ obtained by deleting its i^{th} row and j^{th} column. The determinant $|M_{ij}|$ is called the minor of the element a_{ij} of A. The cofactor of a_{ij} denoted by A_{ij} and is equal to $(-1)^{i+j}\,|M_{ij}|$. The matrix of cofactors is of the form

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

Example 1.8. Let

$$A = \begin{bmatrix} 1 & 9 & 2 \\ 5 & 1 & 4 \\ -2 & 1 & 0 \end{bmatrix}.$$

The cofactor matrix:

$$\begin{pmatrix} (-1)^{1+1}\begin{vmatrix} 1 & 4 \\ 1 & 0 \end{vmatrix} & (-1)^{1+2}\begin{vmatrix} 5 & 4 \\ -2 & 0 \end{vmatrix} & (-1)^{1+3}\begin{vmatrix} 5 & 1 \\ -2 & 1 \end{vmatrix} \\ (-1)^{2+1}\begin{vmatrix} 9 & 2 \\ 1 & 0 \end{vmatrix} & (-1)^{2+2}\begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix} & (-1)^{2+3}\begin{vmatrix} 1 & 9 \\ -2 & 1 \end{vmatrix} \\ (-1)^{3+1}\begin{vmatrix} 9 & 2 \\ 1 & 4 \end{vmatrix} & (-1)^{3+2}\begin{vmatrix} 1 & 2 \\ 5 & 4 \end{vmatrix} & (-1)^{3+3}\begin{vmatrix} 1 & 9 \\ 5 & 1 \end{vmatrix} \end{pmatrix} = \begin{bmatrix} -4 & -8 & 7 \\ 2 & 4 & -19 \\ 34 & 6 & -44 \end{bmatrix}$$

1.3.4 Adjoint Matrix

The adjoint of a square matrix A is the transpose of its cofactor matrix and is denoted by $adj(A)$.

Example 1.9. Find the adjoint matrix of the above example.

Theorem 1.10. *For any square matrix A of order n, we have $A\cdot adj(A) = adj(A)\cdot A = |A|\cdot I_n$.*

Proof. Since A is a square matrix of order n, then also $adj(A)$ in same order. Consider

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad then \quad adj(A) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}^T.$$

Now consider the product

$$
\begin{aligned}
A \cdot adj(A) &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}^T \\
&= \begin{pmatrix} \sum_{k=1}^n a_{1k}A_{1k} & \sum_{k=1}^n a_{1k}A_{2k} & \cdots & \sum_{k=1}^n a_{1k}A_{nk} \\ \sum_{k=1}^n a_{2k}A_{1k} & \sum_{k=1}^n a_{2k}A_{2k} & \cdots & \sum_{k=1}^n a_{2k}A_{nk} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{nk}A_{1k} & \sum_{k=1}^n a_{nk}A_{2k} & \cdots & \sum_{k=1}^n a_{nk}A_{nk} \end{pmatrix} \\
&= \begin{pmatrix} |A| & 0 & 0 & \cdots & 0 \\ 0 & |A| & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & |A| \end{pmatrix} = |A| \cdot I_n = adj(A) \cdot A.
\end{aligned}
$$

□

Theorem 1.11. *If A is a non-singular matrix of order n, then $|adj(A)| = |A|^{n-1}$.*

Theorem 1.12. *If A and B are two square matrices of order n, then*

$$adj(AB) = adj(B) \cdot adj(A)$$

Proof. By theorem 1.10 $AB \cdot adj(AB) = adj(AB) \cdot AB = |AB| \cdot I_n$. By the products

$$
\begin{aligned}
(AB)(adj(B) \cdot adj(A)) &= A(B \cdot adj(B)) \cdot adj(A) = A(|B|I_n) \cdot adj(A) \\
&= |B|(A \cdot adj(A)) = |AB|I_n.
\end{aligned}
\tag{1.1}
$$

and

$$
\begin{aligned}
(adj(B) \cdot adj(A))(AB) &= adj(B)(adj(A) \cdot A)B = adj(B)(|A| \cdot I_n)B \\
&= |A|(adj(B) \cdot B) = |AB|I_n.
\end{aligned}
\tag{1.2}
$$

Therefore from 1.1 and 1.2 we conclude the result. □

1.4 Inverse of a Matrix and properties

1.4.1 Inverse of a Matrix

Definition 1.13. If A and B are two matrices of order n such that $AB = BA = I_n$, then each is said to be inverse of the other. The inverse of A is denoted by A^{-1}.

Theorem 1.14. *The necessary and sufficient condition for a square matrix A to have an inverse is that $|A| \neq 0$ (That is A is non singular).*

Proof.

① The necessary condition: Let A be a square matrix of order n and B is an inverse of it, then $AB = I_n$, and $|AB| = |A||B| = 1 \neq 0$. Therefore $|A| \neq 0$.

② The sufficient condition: If $|A| \neq 0$, then we define the matrix B such that

$$B = \frac{1}{|A|} \cdot adj(A).$$

Then

$$AB = A\frac{1}{|A|} \cdot adj(A) = \frac{1}{|A|}A \cdot adj(A) = \frac{1}{|A|}|A|I_n = I_n.$$

Similarly

$$BA = \frac{1}{|A|} \cdot adj(A)A = \frac{1}{|A|}|A|I_n = I_n.$$

So B is an inverse of A.

□

Theorem 1.15. *Inverse of a matrix if it exists is unique.*

Proof. Let B and C are inverses of the matrix A of order n, then

$$AB = BA = CA = AC = I_n$$

So

$$B(AC) = BI_n = B = (BA)C = I_nC = C.$$

□

Example 1.16. Let

$$A = \begin{bmatrix} 1 & 9 & 2 \\ 5 & 1 & 4 \\ -2 & 1 & 0 \end{bmatrix}.$$

$|A| = -62 \neq 0$, so A is invertible. By Example 1.8, the adjoint matrix of A

$$adj(A) = \begin{bmatrix} -4 & -8 & 7 \\ 2 & 4 & -19 \\ 34 & 6 & -44 \end{bmatrix}^T = \begin{bmatrix} -4 & 2 & 34 \\ -8 & 4 & 6 \\ 7 & -19 & -44 \end{bmatrix}.$$

Thus

$$A^{-1} = \frac{1}{-62}adj(A) = \begin{bmatrix} \frac{2}{31} & -\frac{1}{31} & -\frac{17}{31} \\ \frac{4}{31} & -\frac{2}{31} & -\frac{3}{31} \\ -\frac{7}{62} & \frac{19}{62} & \frac{22}{31} \end{bmatrix}.$$

Theorem 1.17. *If A and B are two non-singular matrices of order n, then AB is also non singular and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. Since A and B are non-singular, therefore $|AB| \neq 0$. Consider

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}) = (AI_nA^{-1}) = (AA^{-1}) = I_n. \tag{1.3}$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(AA^{-1})B = (BI_nB^{-1}) = (BB^{-1}) = I_n. \tag{1.4}$$

From 1.3 and 1.4 we conclude $(AB)^{-1} = B^{-1}A^{-1}$. □

Theorem 1.18. *If A is a non-singular matrix of order n, then*

$$(A^T)^{-1} = (A^{-1})^T.$$

Theorem 1.19. *If A is a non-singular matrix of order n, k is non zero scalar, then*

$$(kA)^{-1} = \frac{1}{k}A^{-1}.$$

Theorem 1.20. *If A is a non-singular matrix of order n then*

$$|A^{-1}| = \frac{1}{|A|}.$$

1.4.2 Rank of a Matrix:

Definition 1.21. A positive integer 'r' is said to be the rank of a non- zero matrix A if

1. There exists at least one non-zero minor of order r of A.
2. Every minor of order greater than r of A is zero.

The rank of a matrix A is denoted by $\rho(A)$.

Exercise 1.22. Find the rank of the matrix

$$A = \begin{bmatrix} -4 & 2 & 34 \\ -8 & 4 & 6 \\ 4 & -2 & 28 \end{bmatrix}$$

Chapter 2

System of linear equations

2.1 Solution of the linear system AX= B

Consider the system of equations in unknowns $x_1, x_2, \cdots, x_n$ as

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

is called system of linear equations with n unknowns $x_1, x_2, \cdots, x_n$.
If the constants $b_1, b_2, \cdots, b_m$ are all zero then the system is said to be homogeneous type.
The above system can be put in the matrix form as $AX = B$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

The matrix A is called coefficient matrix, the matrix X is called matrix of unknowns and B is called as matrix of constants (X is a victor of order n and B is a victor of order m).

Definition 2.1. A set of values $x_1, x_2, \cdots, x_n$ which satisfy all these equations simultaneously is called the solution of the system. If the system has at least one solution then the system is said to be consistent otherwise its said to be inconsistent.

Theorem 2.2. *A system of m equations in n unknowns represented by the matrix equation $AX = B$ is said to be consistent if and only if $\rho(A) = \rho(A, B)$. That is the rank of matrix A is equal to rank of augmented matrix (A, B) defined by*

$$(A, B) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Theorem 2.3. *If A be an non-singular matrix of order n, X is a victor of order n and B is a victor of order n, then the system of equations$AX = B$ has a unique solution.*

1 $AX = B$

- Inconsistent if $\rho(A) \neq \rho(A, B)$
- consistent if $\rho(A) = \rho(A, B)$
  - Unique solution if $\rho(A) = n$
  - Infinite solutions if $\rho(A) < n$

2 $AX = 0$ (Always consistent)

- if $\rho(A) = r = n$ trivial solution
- if $\rho(A) = r < n$ infinit solutions

2.2 Elementary Transformations

Some operations on matrices called as elementary transformations. There are six types of elementary transformations, three of then are row transformations and other three of them are column transformations. There are as follows

1. Interchange of any two rows or columns.
2. Multiplication of the elements of any row (or column) by a non zero number k.
3. Multiplication to elements of any row or column by a scalar k and addition of it to the corresponding elements of any other row or column.

We adopt the following notations for above transformations

1. Interchange of i^{th} row and j^{th} row is denoted by $R_i \leftrightarrow R_j$.
2. Multiplication by k to all elements in the i^{th} row $kR_i \to R_i$.
3. Multiplication to elements of j^{th} row by k and adding them to the corresponding elements of i^{th} row is denoted by $R_i + kR_j \to R_i$.

Definition 2.4. Equivalent Matrix: A matrix B is said to be equivalent to a matrix A if B can be obtained from A by forming finitely many successive elementary transformations on a matrix A. Denoted by $A \simeq B$.

2.3 Methods of solving system of linear Equations

2.3.1 Method of inversion or Cramer's rule

Consider the matrix equation $AX = B$ where $|A| \neq 0$. Pre multiplying by A^{-1}, we have

$$A^{-1}(AX) = A^{-1}B = X.$$

Thus $AX = B$ has only one solution given by $X = A^{-1}B$.

Theorem 2.5 (Cramer's rule). *Suppose A is an invertible matrix of order n, and we wish to solve the system $AX = B$. Then x_i can be computed by the rule*

$$x_i = \frac{|A_i|}{|A|},$$

where A_i is the matrix obtained by replacing the i^{th} column of A with B.

Example 2.6. Use Cramer's rule to solve the next system of equations $AX = B$, where

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 1 & 4 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad B = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}.$$

Solution. The matrices A_1, A_2, and A_3 are obtained by respectively replacing the first, second, and third column of A by B. We compute

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 1 & 4 & 1 \end{vmatrix} = 4, \qquad |A_1| = \begin{vmatrix} 3 & 2 & 1 \\ 5 & 2 & 1 \\ 6 & 4 & 1 \end{vmatrix} = 4$$

$$|A_2| = \begin{vmatrix} 1 & 3 & 1 \\ 3 & 5 & 1 \\ 1 & 6 & 1 \end{vmatrix} = 6, \qquad |A_3| = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 5 \\ 1 & 4 & 6 \end{vmatrix} = -4.$$

Then by Cramer's rule,

$$x = \frac{|A_1|}{|A|} = \frac{4}{4} = 1, \qquad y = \frac{|A_2|}{|A|} = \frac{6}{4}, \qquad z = \frac{|A_3|}{|A|} = \frac{-4}{4} = -1.$$

Thus, the solution is $(x, y, z) = (1, \frac{3}{2}, -1)$. ■

2.3.2 Gaussian Elimination

Definition 2.7. An entry of an augmented matrix is called a **leading entry** or **pivot entry** if it is the leftmost nonzero entry of a row. A column containing a pivot entry is called a **pivot column**.

Definition 2.8. An augmented matrix is in **echelon form** (also called **row echelon form**) if

1. All rows of zeros are below all non-zero rows.
2. Each leading entry of a row is in a column to the right of the leading entry of any row above it.

Example 2.9. The following augmented matrices are in echelon form. We have circled the pivot entries for clarity.

Example 2.10. The following augmented matrices are not in echelon form.
In the first matrix, a row of zeros is above a non-zero row. In the second and third matrix, the leading entries of some rows are not to the right of the leading entries of previous rows.

$$\left[\begin{array}{ccccc|c} 0 & 5 & 2 & 1 & 3 & 5 \\ 0 & 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{ccccc|c} 1 & 4 & 0 & 0 & 5 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right], \quad \left[\begin{array}{ccc|c} 3 & 0 & 6 & 2 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & -6 \\ 4 & 0 & 7 \end{array}\right], \quad \left[\begin{array}{ccc|c} 0 & 2 & 3 & 3 \\ 1 & 5 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Gaussian elimination algorithm

This algorithm provides a method for using row operations to take a matrix to its echelon form. We begin with the matrix in its original form.

1. Starting from the left, find the first non-zero column. This is the first pivot column, and the position at the top of this column will be the position of the first pivot entry. Switch rows if necessary to place a non-zero number in the first pivot position.

2. Use row operations to make the entries below the first pivot entry (in the first pivot column) equal to zero.

3. Ignoring the row containing the first pivot entry, repeat steps 1 and 2 with the remaining rows. Repeat the process until there are no more non-zero rows left.

Example 2.11. Solve the following system of equations

$$\begin{cases} x + 4y + 3z = 11 \\ 2x + 10y + 7z = 27 \\ x + y + 2z = 5 \end{cases}.$$

The augmented matrix for this system is

$$\left(\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 2 & 10 & 7 & 27 \\ 1 & 1 & 2 & 5 \end{array}\right)$$

We will use row operations to create zeros in the entries below the 1.

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 2 & 10 & 7 & 27 \\ 1 & 1 & 2 & 5 \end{array}\right] \overset{R_2 \leftarrow R_2 - 2R_1}{\simeq} \left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 0 & 2 & 1 & 5 \\ 1 & 1 & 2 & 5 \end{array}\right] \overset{R_3 \leftarrow R_3 - R_1}{\simeq} \left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 0 & 2 & 1 & 5 \\ 0 & -3 & -1 & -6 \end{array}\right]$$

We now look for the second pivot column, which in this case is column two. Here, the 2 in the second row and second column is in the pivot entry.

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 0 & 2 & 1 & 5 \\ 0 & -3 & -1 & -6 \end{array}\right] \overset{R_3 \leftarrow 2R_3}{\simeq} \left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 0 & 2 & 1 & 5 \\ 0 & -6 & -2 & -12 \end{array}\right] \overset{R_3 \leftarrow R_3+3R_2}{\simeq} \left[\begin{array}{ccc|c} 1 & 4 & 3 & 11 \\ 0 & 2 & 1 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Therefore we have the solution $(x, y, z) = (-2, 1, 3)$.

Proposition 2.12. *The rank of the matrix is equal to the number of non-zero rows in the matrix after reducing it to the row echelon form using elementary transformations over the rows of the matrix*

Example 2.13.

1. The rank of

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$$

Subtract row 1 multiplied by 2 from row 2: $R_2 = R_2 - 2R_1$.

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 3 & 6 \end{bmatrix}$$

Subtract row 1 multiplied by 3 from row 3: $R_3 = R_3 - 3R_1$.

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

so the rank is 1.

2. Let

$$B = \begin{bmatrix} 1 & -2 & 3 & -1 \\ 2 & -1 & 2 & 2 \\ 3 & 1 & 2 & 3 \end{bmatrix}.$$

Subtract row 1 multiplied by 2 from row 2: $R_2 = R_2 - 2R_1$.

$$\begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -4 & 4 \\ 3 & 1 & 2 & 3 \end{bmatrix}$$

Subtract row 1 multiplied by 3 from row 3: $R_3 = R_3 - 3R_1$.

$$\begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -4 & 4 \\ 0 & 7 & -7 & 6 \end{bmatrix}$$

Subtract row 2 multiplied by $\frac{7}{3}$ from row 3: $R_3 = R_3 - \frac{7R_2}{3}$.

$$\begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -4 & 4 \\ 0 & 0 & \frac{7}{3} & -\frac{10}{3} \end{bmatrix}$$

so $\rho(B) = 3$.

Exercise 2.14. Solve the following system of linear equations using Gaussian Elimination.

1 $\begin{cases} 2x + \ \ y - \ \ 2z = 10 \\ \qquad y + 10z = -28 \\ \quad 3y + 16z = -42 \end{cases}$, 2 $\begin{cases} x + 2y - 3z = -1 \\ 3x - \ \ y + 2z = 7 \\ 5x + 3y - 4z = 2 \end{cases}$.

2.4 Eigenvectors and Eigenvalues

When we multiply a square matrix A by a non-zero vector V, we obtain another vector AV. Most of the time, the vectors AV and V are unrelated, they could point in completely different directions. However, sometimes it can happen that AV is a scalar multiple of V. In that case, V is called an **eigenvector** of A.

Definition 2.15. Let A be an $n \times n$-matrix. Suppose that $V \in \mathbb{R}^n$ is a non-zero vector such that AV is a scalar multiple of V. In other words, suppose that there exists a scalarλ such that

$$AV = \lambda V.$$

Then V is called an **eigenvector** of A, and λ is called the corresponding **eigenvalue.**

Theorem 2.16. *Let A be a square matrix of order n, and let λ be a scalar. Then λ is an eigenvalue of A if and only if*

$$|A - \lambda I_n| = 0$$

Example 2.17. Let

$$A = \begin{pmatrix} -5 & 2 \\ -7 & 4 \end{pmatrix}$$

By Theorem 2.16, a scalar λ is an eigenvalue of A if and only if $|A - \lambda I_n| = 0$. We calculate the determinant:

$$|A - \lambda I_n| = \begin{vmatrix} -5-\lambda & 2 \\ -7 & 4-\lambda \end{vmatrix} = (-5-\lambda)(4-\lambda) + 14 = \lambda^2 + \lambda - 6 = (\lambda+3)(\lambda-2) = 0$$

Therefore, the eigenvalues are $\lambda_1 = -3$ and $\lambda_2 = 2$.

Definition 2.18. Let A be a square matrix. The expression

$$p(\lambda) = |A - \lambda I_n|$$

is called the **characteristic polynomial** of A.

Procedure 2.19 (Finding eigenvalues and eigenvectors). Let A be a matrix of order n. To find the eigenvalues and eigenvectors of A

1. Calculate the characteristic polynomial $|A - \lambda I_n|$.

2. The eigenvalues are the roots of the characteristic polynomial.

3. For each eigenvalue λ, find a basis for the eigenvectors by solving the homogeneous system
$$(A - \lambda I_n)V = 0.$$

To double-check your work, make sure that $AV = \lambda V$ for each eigenvalue and associated eigenvector.

Example 2.20. Let
$$A = \begin{pmatrix} 3 & 0 & 2 \\ 6 & 4 & 3 \\ -4 & 0 & -3 \end{pmatrix}$$
The characteristic polynomial is
$$p(\lambda) = |A - \lambda I_n| = -\lambda^3 + 4\lambda^2 + \lambda - 4$$
If $p(\lambda)$ has a integer root k, then k divide the coefficient of $\lambda^0 : (-4)$. So $k \in \{\pm 1, \pm 2, \pm 4\}$. Therefore we find the eigenvalues of A are $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 4$, then
$$p(\lambda) = -(\lambda - 1)(\lambda + 1)(\lambda - 4),$$
We now find the eigenvectors for each eigenvalue.

1. **For** $\lambda = 1$. We must solve $(A - I_n)V = 0$, i.e.
$$\begin{pmatrix} 2 & 0 & 2 \\ 6 & 3 & 3 \\ -4 & 0 & -4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$
By Gaussian elimination
$$\left(\begin{array}{ccc|c} 2 & 0 & 2 & 0 \\ 6 & 3 & 3 & 0 \\ -4 & 0 & -4 & 0 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 2 & 0 & 2 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$
so $2v_1 + 2v_3 = 0$ and $3v_2 - 3v_3 = 0$, that's mean $v_1 = -v_3$ and $v_2 = v_3$, therefore
$$V = \begin{pmatrix} -v_3 \\ v_3 \\ v_3 \end{pmatrix} = v_3 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$
So the eigenvector is
$$V_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

2 **For** $\lambda = -1$. We must solve $(A + I_n)V = 0$, i.e.

$$\begin{pmatrix} 4 & 0 & 2 \\ 6 & 5 & 3 \\ -4 & 0 & -2 \end{pmatrix} V = 0.$$

The eigenvector is

$$V_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}.$$

3 **For** $\lambda = 4$. We must solve $(A + I_n)V = 0$, i.e.

$$\begin{pmatrix} -1 & 0 & 2 \\ 6 & 0 & 3 \\ -4 & 0 & -7 \end{pmatrix} V = 0.$$

The eigenvector is

$$V_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Proposition 2.21. *Let A be an upper or lower triangular matrix. Then the eigenvalues of A are the entries on the main diagonal.*

2.4.1 Diagonalization

Definition 2.22. Let A be an matrix of order n. Then A is said to be **diagonalizable** if there exists an invertible matrix P and a diagonal matrix D such that

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

Definition 2.23. We say that the sequence of vectors $u_1, \cdots, u_k$ is linearly dependent if it contains at least one vector u_j of the form

$$u_j = a_1u_1 + a_2u_2 + \cdots + a_{j-1}u_{j-1}$$

for some scalars $a_1, \cdots, a_{j-1}$. Otherwise, we say that the vectors are linearly independent.

Theorem 2.24. *A matrix A of order n is diagonalizable if and only if A has n linearly independent eigenvectors. Moreover, in this case, let P be the invertible matrix whose columns are n linearly independent eigenvectors of A, and let D be the diagonal matrix whose diagonal entries are the corresponding eigenvalues. Then*

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

Proof. Assume that A has n linearly independent eigenvectors $V_1, \cdots, V_n$. Let $\lambda_1, \cdots, \lambda_n$. be the corresponding eigenvalues, so that

$$AV_i = \lambda_i V_i \tag{2.1}$$

for all$i = 1, \cdots, n$. Let P be the matrix that has $V_1, \cdots, V_n$. as its columns. Then P is invertible because $V_1, \cdots, V_n$. are linearly independent. Let D be the diagonal matrix that has $\lambda_1, \cdots, \lambda_n$ as its diagonal entries. By the column method of matrix multiplication, the i^{th} column of AP is AV_i. Also by the column method of matrix multiplication, the i^{th} column of PD is $\lambda_i V_i$. Therefore, by 2.1, the matrices AP and PD have the same columns, i.e.,

$$AP = PD.$$

It follows that $P^{-1}AP = D$, as desired.

Conversely, assume that A is diagonalizable. Then there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$, or equivalently, $AP = PD$. Let $V_1, \cdots, V_n$ be the columns of P and let $\lambda_1, \cdots, \lambda_n$ be the diagonal entries of D. Again we find that the i^{th} column of AP is AV_i and the i^{th} column of PD is $\lambda_i V_i$, and therefore $AV_i = \lambda_i V_i$ holds for all i. It follows that $V_1, \cdots, V_n$. are eigenvectors of A. Since P is invertible, $V_1, \cdots, V_n$. are linearly independent, so A has n linearly independent eigenvectors. ☐

Example 2.25. Let

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 6 & 4 & 3 \\ -4 & 0 & -3 \end{pmatrix}$$

to Diagonalize A, we shall find the matrices P of eigenvectors and D of eigenvalues. By Example 2.20 we have $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 4$ are the eigenvalues of A, with corresponding eigenvectors

$$V_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \qquad V_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \qquad V_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Therefore we can use

$$P = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} \qquad \text{and} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

To double-check that $P^{-1}AP = D$, we need to calculate $AP = PD$. The last thing is to find P^{-1}.

$$P^{-1} = \frac{1}{|P|} adj(P) = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

2.4.2 Matrix powers

Suppose A is diagonalizable matrix of order n, so that $P^{-1}AP = D$. We can rearrange this equation to write $A = PDP^{-1}$. Now, consider A^2, it follows that

$$A^2 = (PDP^{-1})^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}.$$

Similarly,

$$A^2 = (PDP^{-1})^3 = PDP^{-1}PDP^{-1}PDP^{-1} = PD^3P^{-1}.$$

In general, for $k \in \mathbb{Z}$:

$$A^k = (PDP^{-1})^k = PD^kP^{-1}.$$

Where

$$D^k = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}.$$

Remark 2.26. The last result stay true if k is a rational number (i.e. $k \in \mathbb{Q}$).

Exercise 2.27. Let

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -1 & 5 & 3 \\ 1 & -1 & 1 \end{bmatrix}.$$

Find a square root of A, i.e. find a matrix B such that $A = B^2$.

Solution. We first diagonalize A. The characteristic polynomial is

$$\begin{vmatrix} 1-\lambda & 3 & 3 \\ -1 & 5-\lambda & 3 \\ 1 & -1 & 1-\lambda \end{vmatrix} = -\lambda^3 + 7\lambda^2 - 14\lambda + 8,$$

with roots $\lambda_1 = 1, \lambda_2 = 2$, and $\lambda_3 = 4$. The corresponding eigenvectors are

$$V_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \qquad V_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad V_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

respectively. Therefore we have $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}, \qquad P^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad \text{and} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

The a square root of a diagonal matrix D is easy:

$$\sqrt{D} = D^{\frac{1}{2}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

If we now define $B = PD^{\frac{1}{2}}P^{-1}$, we clearly have $B^2 = PD^{\frac{1}{2}}P^{-1}PD^{\frac{1}{2}}P^{-1} = PDP^{-1} = A$. So the desired square root of A is

$$B = PD^{\frac{1}{2}}P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1-\sqrt{2} & 1+\sqrt{2} & 1 \\ -1+\sqrt{2} & 1-\sqrt{2} & 1 \end{bmatrix}.$$

It's easy to verify that we have computed B correctly by squaring it and double-checking that we really get A. We note that the square root of a matrix is not unique. In fact, D has 8 different square roots, all of the form

$$\begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm\sqrt{2} & 0 \\ 0 & 0 & \pm 2 \end{bmatrix}.$$

It follows that A has 8 different square roots as well. ■

Chapter 3

Introduction to Integration

3.1 Indefinite Integrals

Definition 3.1. Let F be a differentiable real function with $F' = f$. Then F is called *antiderivalive* of f. and any other anti-derivative of f differs from F by a constant. This is expressed by

$$\int f(x)\ dx = F(x) + c.$$

we call $f(x)$ the integrand and c the constant of integration.

Properties 3.2. Let F and G be differentiable real functions with $F' = f$ and $G' = g$. Let $\lambda, p, q \in \mathbb{R}$ with $p \neq 0$. Then

1. $\displaystyle\int \lambda f(x)\ dx = \lambda F(x) + c.$

2. $\displaystyle\int [f(x) + g(x)]\ dx = F(x) + G(x) + c.$

3. $\displaystyle\int f(px + q)\ dx = \frac{1}{p} F(px + q) + c.$

Example 3.3. Find $\int f(x)\ dx$ when $f(x)$ is

a. $x^2 + 8x - 3,$ b. $\cos(5x + 2)$ c. $\dfrac{1}{(6x + 7)^2}.$

Solution.

a. Using Property (2) then Property (1) we have

$$\begin{aligned}
\int (x^2 + 8x - 3)\ dx &= \int x^2\ dx + \int 8x\ dx + \int (-3)\ dx \\
&= \int x^2\ dx + 8 \int x\ dx - 3 \int 1\ dx \\
&= \frac{x^3}{3} + 8\frac{x^2}{2} - 3x + c = \frac{x^3}{3} + 4x^2 - 3x + c.
\end{aligned}$$

(b) Property (3) gives

$$\int \cos(5x+2)\ dx = \frac{1}{5}\sin(5x+2) + c.$$

(c) Again using Property (3),

$$\begin{aligned}\int \frac{1}{(6x+7)^2}\ dx &= \int (6x+7)^{-2}\ dx = \frac{1}{6}\frac{(6x+7)^{-1}}{-1} + c \\ &= \frac{-1}{6(6x+7)} + c.\end{aligned}$$

■

3.2 Table of Integrals

The table 3.1 contain the antiderivates of common functions.

$f(x)$	$F(x)$	$f(x)$	$F(x)$
$x^r, \quad r \neq -1$	$\dfrac{x^{r+1}}{r+1}$	$\dfrac{1}{x}$	$\log\lvert x\rvert$
$\dfrac{1}{x^2+a^2} \quad (a \neq 0)$	$\frac{1}{a}\tan^{-1}\frac{x}{a}$	e^x	e^x
$\cos^{-1} x$	$x\cos^{-1} x - \sqrt{1-x^2}$	$\tan x$	$-\log\lvert\cos x\rvert$
$\sin^{-1} x$	$x\sin^{-1} x + \sqrt{1-x^2}$	$\cos x$	$\sin x$
$\tan^{-1} x$	$x\tan^{-1} x - \frac{1}{2}\log(x^2+1)$	$\log x$	$x\log x - x$
$\dfrac{1}{\sqrt{a^2-x^2}} \quad (a > 0)$	$\sin^{-1}\frac{x}{a}, \quad -\cos^{-1}\frac{x}{a}$		

Table 3.1: Table of Integrals

3.3 Definite Integrals

Definition 3.4. Let f be a real function, continuous on $[a,b]$. Roughly speaking, the *definite integral of* f *over* $[a,b]$, denoted by

$$\int_a^b f(x)\ dx,$$

is the area of the finite region bounded by the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$.

Remark 3.5. The areas of regions above and below the x-axis are regarded as positive and negative, respectively. Thus, for example, Figure 3.1 illustrates that $\int_0^2 (3-x)\ dx = 4$ and $\int_2^5 (3-x)\ dx = (1/2) + (-2) = -3/2$.

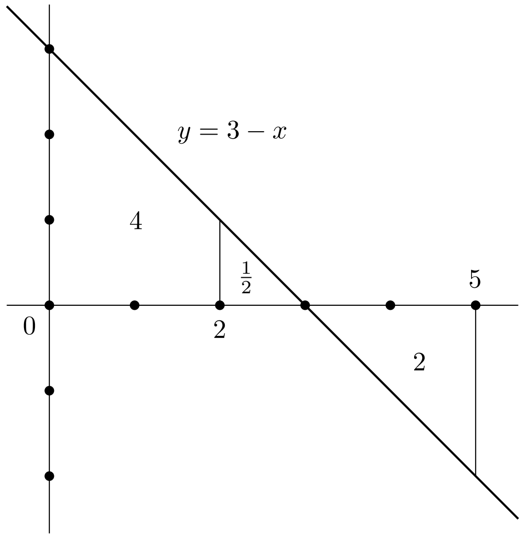

Figure 3.1:

Remark 3.6. In the expression $\int_a^b f(x)\ dx, x$ is a 'dummy' variable. Any symbol will do. Thus

$$\int_a^b f(x)\ dx = \int_a^b f(t)\ dt = \int_a^b f(\theta) d\theta = \cdots$$

Definition 3.7. Let f be a real function, continuous on $[a, b]$, let $n \in \mathbb{N}$ and let $x_0, x_1, \ldots$, $x_n \in \mathbb{R}$, with

$$a = x_0 < x_1 < \cdots < x_n = b$$

We call $\mathcal{P} = (x_0, x_1, \ldots, x_n)$ a *partition* of $[a, b]$. The *size* of $\mathcal{P}$, written $|\mathcal{P}|$, is the length of the longest of the subintervals $[x_{r-1}, x_{r-1}] \quad (r = 1, \ldots, n)$, i.e.

$$|\mathcal{P}| = \max\{|x_r - x_{r-1}| : \ r = 1, \ldots, n\}.$$

Definition 3.8. Let f be a real function, defined on $[a, b]$, and suppose we can find a partition $(x_0, x_1, \ldots, x_n)$ of $[a, b]$ such that, for each $r = 1, \ldots, n$ there exists a real function f_r, continuous on $[x_{r-1}, x_r]$ with $f(x) = f_r(x)$ for all $x \in [x_{r-1}, x_r]$. Then f is said to be *piecewise continuous on* $[a, b]$. Then f is said to be *Riemann integrable over* $[a, b]$ with

$$\int_a^b f(x)\ dx = \int_{x_0}^{x_1} f_1(x)\ dx + \int_{x_1}^{x_2} f_2(x)\ dx + \cdots + \int_{x_{n-1}}^{x_n} f_n(x)\ dx.$$

Theorem 3.9. *Let f be a real function. We say that f is Riemann integrable over $[a, b]$ iff f is piecewise continuous on $[a, b]$.*

Example 3.10. To illustrate the definition of piecewise continuous, consider

1. $f : [-1, 2] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 3/2 & \text{if } -1 \leq x \leq 0, \\ (2-x)^2 & \text{if } 0 < x \leq 2. \end{cases}$$

We can take $f_1 : [-1, 0] \to \mathbb{R}$ defined by $f_1(x) = 3/2$ and $f_2 : [0, 2] \to \mathbb{R}$ defined by $f_2(x) = (2-x)^2$. This shows that f is piecewise continuous on $[-1, 2]$ and

$$\int_{-1}^{2} f(x)\ dx = \int_{-1}^{0} (3/2)\ dx + \int_{0}^{2} (2-x)^2\ dx.$$

2 On the other hand, $g : [-1, 2] \to \mathbb{R}$ defined by

$$g(x) = \begin{cases} 3/2 & \text{if } -1 \leq x \leq 0, \\ (2-x)/x & \text{if } 0 < x \leq 2, \end{cases}$$

can be shown to be neither piecewise continuous on $[-1, 2]$ nor Riemann integrable over $[-1, 2]$.

Theorem 3.11 (Properties of Definite Integrals). *Let $k, \lambda \in \mathbb{R}$ and let f and g are real functions, Riemann integrable over $[a, b]$.*

1 $\displaystyle\int_a^b f(x)\ dx = F(b) - F(a)$, *where F is an antiderivative of f.*

2 *The real (constant) function $x \mapsto k$ is Riemann integrable over $[a, b]$ with*

$$\int_a^b k\ dx = k(b-a).$$

3 *The functions λf and $f + g$ are Riemann integrable over $[a, b]$ with*

a $\displaystyle\int_a^b \lambda f(x)\ dx = \lambda \int_a^b f(x)\ dx,$

b $\displaystyle\int_a^b [f(x) + g(x)]\ dx = \int_a^b f(x)\ dx + \int_a^b g(x)\ dx.$

4 *For $c \in [a, b]$,* $\displaystyle\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx.$

5 *Let $f(x) \leq g(x)$ for all $x \in]a, b[$. Then*

$$\int_a^b f(x)\ dx \leq \int_a^b g(x)\ dx.$$

In particular, if $m \leq f(x) \leq M$ for all $x \in]a, b[$ then

$$m(b-a) \leq \int_a^b f(x)\ dx \leq M(b-a).$$

6 *The function $|f|$ is Riemann integrable over $[a, b]$ and*

$$\left|\int_a^b f(x)\ dx\right| \leq \int_a^b |f(x)|\ dx.$$

7 *Let h be a real function, Riemann integrable over $[-k, k]$. Then*

a *h odd* $\Rightarrow \displaystyle\int_{-k}^{k} h(x)\ dx = 0.$

b *h even* $\Rightarrow \displaystyle\int_{-k}^{k} h(x)\ dx = 2\int_{0}^{k} h(x)\ dx,$

Example 3.12. Evaluate $\mathcal{I} = \displaystyle\int_{-2}^{2} (3 - 2\sin x)\ dx.$

Solution. The properties referred to are those of previous Theorem. We have

$$\begin{aligned}
\mathcal{I} &= \int_{-2}^{2} 3\ dx - 2\int_{-2}^{2} \sin x\ dx && \text{(by Property (2))}\\
&= 2\int_{0}^{2} 3\ dx - 2(0) && \text{(by Property (6))}\\
&= 2(3)(2-0) + 0 = 12 && \text{(by Property (1)).}
\end{aligned}$$

■

Example 3.13. Use the fact that $1 - x^2 \leq \sqrt{1-x^2} \leq 1 (0 \leq x \leq 1)$ to show that

$$2/3 \leq \int_0^1 \sqrt{1-x^2}\ dx \leq 1.$$

Solution. Using Property (4) of the previous Theorem we have

$$\int_0^1 (1-x^2)\ dx \leq \int_0^1 \sqrt{1-x^2}\ dx \leq \int_0^1 1\ dx.$$

Hence, using a previous Example and Property (1) of the previous Theorem,

$$2/3 \leq \int_0^1 \sqrt{1-x^2}\ dx \leq 1$$

as required. ■

3.4 Integration Methods

3.4.1 Substitution

Let f be a continuous function with antiderivative F, and let g be a differentiable function such that g' is continuous and $F \circ g$ is defined. Writing $F \circ g = H$ and $h = H'$, the chain rule givae

$$h(x) = H'(x) = (F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x)$$

So $\displaystyle\int h(x)\ dx = H(x) + c$ can be written

$$\int f(g(x))g'(x)\ dx = F(g(x)) + c \tag{3.1}$$

and we also have

$$\int f(u)\ du = F(u) + c. \tag{3.2}$$

We may transfer between equations 3.1 and 3.2 by substituting

$$g(x) = u, \qquad g'(x)\ dx = \ du.$$

Example 3.14. Consider

$$\mathcal{I} = \int (1+\sin x)^3 \cos x\ dx.$$

Substituting

$$1+\sin x = u, \qquad \cos x\ dx = \ du$$

we get

$$\mathcal{I} = \int u^3\ du = \frac{u^4}{4} + c = \frac{1}{4}(1+\sin x)^4 + c.$$

Remark 3.15. Definite integrals may also be found by substitution. With the functions f, F and g as above, and assuming that g is defined on $[a,b]$ and f on $g([a,b])$, we have

$$\int_a^b f(g(x))g'(x)\ dx = [F\{g(x))]_a^b = F(g(b)) - F(g(a)) \tag{3.3}$$

and

$$\int_{g(a)}^{g(b)} f(u)\ du = [F(u)]_{g(a)}^{g(b)} = F(g(b)) - F(g(a)). \tag{3.4}$$

We may transfer between equations 3.3 and 3.4 by substituting

$$g(x) = u, \qquad g'(x)\ dx = \ du,$$

x	a	b
u	$g(a)$	$g(b)$

where, as before, $g(x) = u$ is the *basic substitution*.

Example 3.16. To evaluate

$$J = \int_0^{\pi/2} (1+\sin x)^3 \cos x\ dx$$

we may substitute

$$1+\sin x = n, \qquad \cos x\ dx = \ du,$$

x	0	$\pi/2$
u	1	2

to get

$$J = \int_1^2 u^3 du = \left[\frac{u^4}{4}\right]_1^2 = \frac{16}{4} - \frac{1}{4} = \frac{15}{4}.$$

Hazard. When substituting in an indefinite integral, remember to return to the original variable after integrating.

Example 3.17. Evaluate

$$\int_0^{\sqrt{3}} \frac{x^2}{x^6+9}\, dx.$$

Solution. Here we substitute

$$x^3 = u, \qquad 3x^2\, dx = du,$$

x	0	$\sqrt{3}$
u	0	$3\sqrt{3}$

to get

$$\begin{aligned} \int_0^{\sqrt{3}} \frac{x^2}{x^6+9}\, dx &= \frac{1}{3}\int_0^{\sqrt{3}} \frac{1}{x^6+9} 3x^2\, dx = \int_0^{3\sqrt{3}} \frac{1}{u^2+9}\, du \\ &= \frac{1}{3}\left[\frac{1}{3}\tan^{-1}\frac{u}{3}\right]_0^{3\sqrt{3}} = \frac{1}{9}\tan^{-1}\sqrt{3} = \frac{\pi}{27}. \end{aligned}$$

■

3.4.2 Rational Integrals

Definition 3.18. Let $f(x), g(x)$ be a polynomials with $g(x)$ non-zero. We can define a real function h by

$$h(x) = \frac{f(x)}{g(x)}.$$

We call h rational function. h is said to be *proper* if h is zero function or $deg(f) < deg(g)$. Otherwise h is improper.

Remark 3.19.

(a) When $h \neq 0$ is proper, we can express $h(x)$ as a sum of terms called *partial fractions*. This sum is the *partial fraction decomposition* of $h(x)$.

(b) When h is improper we can use long division to find $q(x), r(x)$ polynomials, such that

$$f(x) = q(x)g(x) + r(x),$$

where $r(x)$ is not the zero polynomial and $deg(r) < deg(g)$.
For $x \in D_h$ so that $g(x) \neq 0$, so

$$h(x) = \frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)}.$$

Example 3.20. The factorisation of the polynomial

$$g(x) = x^4 + 3x^3 + 4x^2 + 3x + 1 = (x+1)^2(x^2+x+1).$$

Then the partial fraction decomposition of the function

$$h(x) = \frac{4x^2+3x+2}{x^4+3x^3+4x^2+3x+1}$$

is of the form

$$\frac{4x^2+3x+2}{x^4+3x^3+4x^2+3x+1} = \frac{a}{x+1} + \frac{b}{(x+1)^2} + \frac{\alpha x+\beta}{x^2+x+1}.$$

Multiplying both sides by $g(x)$ gives

$$4x^2+3x+2 = (x+1)(x^2+x+1)a + (x^2+x+1)b + (x+1)^2(\alpha x+\beta). \tag{3.5}$$

We consider two ways of determining a, b, α and β.

Method 1. (Substituting Values)

(a) We substitute x by -1 on (3.5), this gives

$$4(-1)^2+3(-1)+2 = 0\times a + ((-1)2+(-1)+1)b + 0\times(\alpha x+\beta),$$

giving $b = 3$.

(b) However, we can substitute 3 for b in (3.5) to get

$$x^2-1 = (x-1)(x+1) = (x+1)(x^2+x+1)a + (x+1)^2(\alpha x+\beta). \tag{3.6}$$

Dividing 3.6 by $(x+1)$ to get

$$(x-1) = (x^2+x+1)a + (x+1)(\alpha x+\beta),$$

and set $x = -1$ again, then $a = -2$. We substitute a by -2 in the last equation to get

$$2x^2+3x+1 = (2x+1)(x+1) = (x+1)(\alpha x+\beta) \Leftrightarrow 2x+1 = (\alpha x+\beta).$$

Thus, $\alpha = 2$ and $\beta = 1$. Thus

$$h(x) = \frac{-2}{x+1} + \frac{3}{(x+1)^2} + \frac{2x+1}{x^2+x+1}.$$

This is the required partial fraction decomposition.

Method 2. (Equating Coefficients)

Here, we multiply out the terms on the right-hand side of $h(x)$ and regroup them to obtain coefficients for the powers of x. Thus

$$\begin{aligned} 4x^2+3x+2 &= (x^3+2x^2+2x+1)a + (x^2+x+1)b + (x^2+2x+1)(\alpha x+\beta) \\ &= (a+\alpha)x^3 + (2a+b+2\alpha+\beta)x^2 + (2a+b+\alpha+2\beta)x + (a+b+\beta). \end{aligned}$$

Then, equating coefficients, we get

$$\begin{cases} a + \alpha = 0 \\ 2a+b+2\alpha+\beta = 4 \\ 2a+b+\alpha+2\beta = 3 \\ a+b+\beta = 2 \end{cases}$$

After solving the last system, we obtain $(a, b, \alpha, \beta) = (-2, 3, 2, 1)$.

Procedure 3.21. For a rational function h the partial fraction decomposition of $h(x)$ is the sum of a polynomial in x together with terms of the form

$$\frac{a}{(x-\alpha)^m} \quad \textit{and} \quad \frac{bx+d}{(x^2+\beta x+\gamma)^n},$$

where $m, n \in \mathbb{N}$.

(a) For $m = 1$:

$$\int \frac{a}{x-\alpha}\, dx = a \log|x-\alpha| + c$$

and for $m \geq 2$:

$$\int \frac{a}{(x-\alpha)^m}\, dx = \frac{-a}{(m-1)(x-\alpha)^{m-1}} + c.$$

(b) It therefore remains to consider integrals of the form

$$\mathcal{I} = \int \frac{bx+d}{(x^2+\beta x+\gamma)^n}\, dx$$

where $x^2+\beta x+\gamma$ is irreducible over $\mathbb{R}$ (i.e. $\Delta < 0$). First we choose constants p and q such that

$$bx + d = p\frac{d}{dx}(x^2+\beta x+\gamma) + q,$$

the constant p being chosen first, to get the coefficient of x correct. Then we write the integral in the form

$$\int \frac{p(2x+\beta)}{(x^2+\beta x+\gamma)^n}\, dx + \int \frac{q}{(x^2+\beta x+\gamma)^n}\, dx = \mathcal{I}_1 + \mathcal{I}_2 \quad \text{(respectively)}.$$

(i) For $\mathcal{I}_1$ we substitute

$$x^2+\beta x+\gamma = u, \quad (2x+\beta)\, dx = \; du.$$

(ii) For $\mathcal{I}_2$ we complete the square of the quadratic to get

$$x^2+\beta x+\gamma = (x+\xi)^2 + \eta^2$$

then substitute

$$x+\xi = \eta\tan\theta, \qquad dx = \eta(\tan^2\theta+1)\; d\theta$$

and

$$(x+\xi)^2+\eta^2 = \eta^2\tan^2\theta + \eta^2 = \eta^2(\tan^2+1)\theta.$$

Exercise 3.22. Find/evaluate

(a) $\mathcal{I} = \displaystyle\int \frac{x^4+3x^3-1}{x^4+3x^3+4x^2+3x+1}\, dx,$

(b) $J = \displaystyle\int_2^5 \frac{x+7}{(x^2-4x+7)^3}\, dx,$

c) $\mathcal{K} = \displaystyle\int_0^5 \frac{25 - x^2}{x^4 + 6x^2 + 25}\, dx.$

Solution. a) The partial fraction decomposition of the integrand is

$$\begin{aligned}\frac{x^4 + 3x^3 - 1}{x^4 + 3x^3 + 4x^2 + 3x + 1} &= \frac{x^4 + 3x^3 + 4x^2 + 3x + 1 - (4x^2 + 3x + 2)}{x^4 + 3x^3 + 4x^2 + 3x + 1} \\ &= 1 - \frac{4x^2 + 3x + 2}{x^4 + 3x^3 + 4x^2 + 3x + 1}.\end{aligned}$$

Hence,

$$\begin{aligned}\mathcal{I} &= \int \left(1 + \frac{2}{x+1} - \frac{3}{(x+1)^2} - \frac{2x+1}{x^2+x+1}\right) dx \\ &= x + 2\log|x+1| + \frac{3}{x+1} - \log\left|x^2 + x + 1\right| + c.\end{aligned}$$

b) Since $\frac{d}{dx}(x^2 - 4x + 7) = 2x - 4$ and $x + 7 = \frac{1}{2}(2x - 4) + 9$, we have

$$J = \int_2^5 \frac{\frac{1}{2}(2x-4)}{(x^2 - 4x + 7)^3}\, dx + \int_2^5 \frac{9}{(x^2 - 4x + 7)^3}\, dx = J_1 + J_2 \text{ (respectively).}$$

i) For J_1, substitute

$$x^2 - 4x + 7 = u, \qquad (2x - 4)\, dx = du,$$

x	2	5
u	3	12

to get

$$J_1 = \frac{1}{2}\int_3^{12} \frac{du}{u^3} = \frac{1}{2}\left[-\frac{1}{2u^2}\right]_3^{12} = \frac{5}{192}.$$

ii) For J_2, complete the square of the quadratic,

$$x^2 - 4x + 7 = (x-2)^2 + 3,$$

and substitute

$$x^2 - 4x + 7 = \sqrt{3}\tan t, \qquad dx = \sqrt{3}(1 + \tan^2 t)\, dt,$$

x	2	5
t	0	$\frac{\pi}{3}$

to get

$$\begin{aligned}J_2 &= 9\int_2^5 \frac{dx}{\left((x-2)^2 + 3\right)^3} = 9\int_0^{\frac{\pi}{3}} \frac{\sqrt{3}(\tan^2 t + 1)}{(3\tan^2 t + 3)^3}\, dt \\ &= 9\int_0^{\frac{\pi}{3}} \frac{\sqrt{3}(\tan^2 t + 1)}{3^3(\tan^2 t + 1)^3}\, dt = \frac{1}{\sqrt{3}}\int_0^{\frac{\pi}{3}} \cos^4 t\, dt, \\ &\qquad \text{(using } (2\cos t)^n = (z + z^{-1})^n, \text{ where } z = e^{it}). \\ &= \frac{1}{\sqrt{3}}\int_0^{\frac{\pi}{3}} \left(\frac{1}{8}\cos 4t + \frac{1}{2}\cos 2t + \frac{3}{8}\right) dt \\ &= \frac{1}{\sqrt{3}}\left[\frac{1}{32}\sin 4t + \frac{1}{4}\sin 2t + \frac{3t}{8}\right]_0^{\pi/3} \\ &= \frac{1}{\sqrt{3}}\left(\frac{1}{32}\left(-\frac{\sqrt{3}}{2}\right) + \frac{1}{4}\frac{\sqrt{3}}{2} + \frac{3}{8}\frac{\pi}{3}\right) = \frac{7}{64} + \frac{\pi}{8\sqrt{3}}.\end{aligned}$$

Hence,

$$J = \frac{5}{192} + \frac{7}{64} + \frac{\pi}{8\sqrt{3}} = \frac{13}{96} + \frac{\pi}{8\sqrt{3}}.$$

(c) The partial fraction decomposition of the integrand. Hence

$$\begin{aligned} \mathcal{K} &= \int_0^5 \frac{1}{2}(\frac{3x+5}{x^2+2x+5})\,dx + \int_0^5 \frac{1}{2}(\frac{-3x+5}{x^2-2x+5})\,dx \\ &= \mathcal{K}_1 + \mathcal{K}_2 \text{ (respectively).} \end{aligned}$$

Then $\frac{d}{dx}(x^2+2x+5) = 2x+2$ and $\frac{1}{2}(3x+5) = \frac{3}{4}(2x+2)+1$. So

$$\begin{aligned} \mathcal{K}_1 &= \frac{3}{4}\int_0^5 \frac{2x+2}{x^2+2x+5}\,dx + \int_0^5 \frac{dx}{(x+1)^2+4} \\ &= \left[\frac{3}{4}\log|x^2+2x+5|\right]_0^5 + \left[\frac{1}{2}\tan^{-1}\frac{(x+1)}{2}\right]_0^5 \\ &= \frac{3}{4}\log 8 + \frac{1}{2}\tan^{-1}3 - \frac{1}{2}\tan^{-1}\left(\frac{1}{2}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} \mathcal{K}_2 &= -\frac{3}{4}\int_0^5 \frac{2x-2}{x^2-2x+5}\,dx + \int_0^5 \frac{dx}{(x-1)^2+4} \\ &= \left[-\frac{3}{4}\log|x^2-2x+5|\right]_0^5 + \left[\frac{1}{2}\tan^{-1}\frac{(x-1)}{2}\right]_0^5 \\ &= -\frac{3}{4}\log 4 + \frac{1}{2}\tan^{-1}2 - \frac{1}{2}\tan^{-1}\left(-\frac{1}{2}\right). \end{aligned}$$

Then, since $\tan^{-1}(-1/2) = -\tan^{-1}(1/2)$, we have

$$\begin{aligned} \mathcal{K} &= \frac{3}{4}(\log 8 - \log 4) + \frac{1}{2}(\tan^{-1}3 + \tan^{-1}2) \\ &= \frac{3}{4}\log 2 + \frac{1}{2}\left(\frac{3\pi}{4}\right) = \frac{3}{4}\left(\log 2 + \frac{\pi}{2}\right). \end{aligned}$$

■

3.4.3 Trigonometric Integrals

Products

Simple products of sines and cosines can be converted to sums of sines and cosines.

Corollary 3.23. *Let $a, b \in \mathbb{R}$. Then*

(a) $\sin a \cos b = \frac{1}{2}(\sin(a+b) + \sin(a-b))$

(b) $\cos a \sin b = \frac{1}{2}(\sin(a+b) - \sin(a-b))$

(c) $\cos a \cos b = \frac{1}{2}(\cos(a+b) + \cos(a-b))$

(d) $\sin a \sin b = \frac{1}{2}(\cos(a-b) - \cos(a+b))$

(e) $\sin a + \sin b = 2\sin\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right)$ (g) $\cos a + \cos b = 2\cos\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right)$

(f) $\sin a - \sin b = 2\cos\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)$ (h) $\cos a - \cos b = -2\sin\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)$

Example 3.24. Evaluate/find

(a) $\displaystyle\int_0^{\pi/2} \sin 7x \cos 5x \; dx$, (b) $\displaystyle\int \sin^6 x \; dx$, (c) $\mathcal{I} = \int \sin^2 x \cos^3 x \; dx$.

Solution. (a) We have

$$\begin{aligned}\int_0^{\pi/2} \sin 7x \cos 5x \; dx &= \frac{1}{2}\int_0^{\pi/2} (\sin 12x + \sin 2x) \; dx \\ &= \frac{1}{2}\left[-\frac{1}{12}\cos 12x - \frac{1}{2}\cos 2x\right]_0^{\pi/2} = \frac{1}{2}.\end{aligned}$$

(b) Let $z = e^{it} = \cos t + i\sin t$, also $z^{-1} = e^{-it} = \cos t - i \sin t$, so we have $(2\cos t)^n = (z + z^{-1})^n$ and $(i2\sin t)^n = (z - z^{-1})^n$. Then

$$\sin^6 x = -\frac{1}{32}\cos 6x + \frac{3}{16}\cos 4x - \frac{15}{32}\cos 2x + \frac{5}{16}.$$

Hence,

$$\int \sin^6 x \; dx = -\frac{1}{192}\sin 6x + \frac{3}{64}\sin 4x - \frac{15}{64}\sin 2x + \frac{5x}{16} + c.$$

(c) We can write

$$\mathcal{I} = \int \sin^2 x \cos^2 x \cos x \; dx$$

so that substituting

$$\sin x = u, \quad \cos x \; dx = \; du, \quad \text{and} \quad \cos^2 x = 1 - \sin^2 x = 1 - u^2$$

gives

$$\mathcal{I} = \int u^2(1-u^2) \; du = \int (u^2 - u^4) \; du = \frac{u^3}{3} - \frac{u^5}{5} + c = \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + c.$$

■

Proposition 3.25. *In the table below, $k \in \mathbb{Z}$, $r \in \mathbb{R}$, and define* $\sec x = \frac{1}{\cos x}$.

Integrand	k	*Substitution*
$(\sin x)^r(\cos x)^k$	*odd*	$\sin x = u$
$(\sin x)^k(\cos x)^r$	*odd*	$\cos x = u$
$(\tan x)^r(\sec x)^k$	*even*	$\tan x = u$
$(\tan x)^k(\sec x)^r$	*odd*	$\sec x = u$

Example 3.26. Find

(a) $\displaystyle\int \frac{\sin^3 x}{\cos^2 x}\,dx$, (b) $\displaystyle\int_0^{\pi/4} \sec^6 x\,dx$, (c) $\displaystyle\int_0^{\frac{\pi}{3}} \tan^3 x(\sec x)^{\frac{1}{3}}\,dx$.

Solution. (a) Here we substitute

$$\cos x = u, \quad -\sin x\,dx = du, \quad \text{and} \quad \sin^2 x = 1 - u^2$$

to get

$$\begin{aligned}\int \frac{\sin^3 x}{\cos^2 x}\,dx &= \int \frac{-\sin^2 x}{\cos^2 x}(-\sin x)\,dx = \int \frac{u^2-1}{u^2}du = \int (1-u^{-2})du \\ &= u + u^{-1} + c = \cos x + \sec x + c.\end{aligned}$$

(b) Here we substitute

$$\tan x = u, \qquad \sec^2 x\,dx = du, \qquad \begin{array}{|c|c|c|}\hline x & 0 & \pi/4 \\ \hline u & 0 & 1 \\ \hline\end{array} \qquad \text{and} \qquad \sec^2 x = u^2 + 1$$

to get

$$\begin{aligned}\int_0^{\pi/4} \sec^6 x\,dx &= \int_0^{\pi/4} \sec^4 x \sec^2 x\,dx = \int_0^1 (u^2+1)^2 du \\ &= \int_0^1 (u^4 + 2u^2 + 1)du = \left[\frac{u^5}{5} + \frac{2u^3}{3} + u\right]_0^1 = \frac{28}{15}.\end{aligned}$$

(c) Here we substitute

$$\sec x = u, \qquad \sec x \tan x\,dx = du, \qquad \begin{array}{|c|c|c|}\hline x & 0 & \pi/3 \\ \hline u & 1 & 2 \\ \hline\end{array} \text{ and } \tan^2 x = u^2 - 1$$

to get

$$\begin{aligned}\int_0^{\pi/3} \tan^3 x \sec x)^{1/3}\,dx &= \int_0^{\pi/3} \tan^2 x(\sec x)^{-2/3}(\sec x \tan x)\,dx \\ &= \int_1^2 (u^2-1)u^{-2/3}du = \int_1^2 \left(u^{4/3} - u^{-2/3}\right)\,du \\ &= \left[\frac{3}{7}u^{7/3} - 3u^{1/3}\right]_1^2 = \frac{9}{7}(2 - 2^{1/3}).\end{aligned}$$

■

Exercise 3.27. For $n \in \mathbb{N}$, define

$$\mathcal{I}_n = \int \tan^n x\,dx.$$

Show that for $n \geq 2$

$$\mathcal{I}_n = \frac{1}{n-1}\tan^{n-1} x - \mathcal{I}_{n-2}$$

Hence find $\mathcal{I}_4$.

The t-formulae

By the next theorem, we can express $\sin x, \cos x$ and $\tan x$ in terms of $t = \tan \frac{x}{2}$. These can be used to convert certain trigonometric integral into integrals of rational functions. In particular, they can be used to find integrals of the form

$$\int \frac{dx}{a + b\sin x + c\cos x},$$

where b and c are not both 0.

Theorem 3.28. *Let $x \in \mathbb{R}$ such that $t = \tan \frac{x}{2}$ is defined. Then*

$$\text{(a)}\ \sin x = \frac{2t}{1+t^2}, \qquad \text{(b)}\ \cos x = \frac{1-t^2}{1+t^2}, \qquad \text{and} \qquad \text{(c)}\ \tan x = \frac{2t}{1-t^2}$$

whenever $\tan x$ *is defined (so that $t^2 \neq 1$).*

The basic substitution is $t = \tan \frac{x}{2}$. Then

$$\frac{dt}{dx} = \frac{1}{2}\sec^2 \frac{x}{2} = \frac{1}{2}\left(1 + \tan^2 \frac{x}{2}\right) = \frac{1}{2}(1+t^2)\ .$$

As necessary, we substitute

$$t = \tan \frac{x}{2}, \qquad dx = \frac{2}{1+t^2}dt$$

and, from last Theorem

$$\sin x = \frac{2t}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad \tan x = \frac{2t}{1-t^2}.$$

Example 3.29. Evaluate

$$\text{(a)}\ \mathcal{I} = \int_0^{\pi/2} \frac{dx}{2+\sin x}, \qquad \text{(b)}\ \mathcal{J} = \int_0^{\pi/2} \frac{dx}{3 + 3\sin x - 2\cos x}.$$

Solution. Here, the basic substitution $t = \tan \frac{x}{2}$, leads to the change of limits shown in the table.

x	0	$\pi/2$
u	0	1

(a) Hence,

$$\begin{aligned}
\mathcal{I} &= \int_0^1 \frac{[2/(1+t^2)]\ dt}{2+[2t/(1+t^2)]} = \int_0^1 \frac{2\ dt}{2(1+t^2)+2t} \\
&\quad \text{(multiplying numerator and denominator by } 1+t^2) \\
&= \int_0^1 \frac{dt}{t^2+t+1} = \int_0^1 \frac{dt}{(t+1/2)^2+3/4} = \left[\frac{2}{\sqrt{3}}\tan^{-1}(\frac{t+1/2}{\sqrt{3}/2})\right]_0^1 \\
&= \frac{2}{\sqrt{3}}\left(\tan^{-1}\sqrt{3} - \tan^{-1}\frac{1}{\sqrt{3}}\right) = \frac{\pi}{3\sqrt{3}}.
\end{aligned}$$

(b) And

$$\begin{aligned}
\mathcal{J} &= \int_0^1 \frac{2\ dt}{3(1+t^2)+3(2t)-2(1-t^2)} = \int_0^1 \frac{2\ dt}{5t^2+6t+1} = \int_0^1 \frac{2\ dt}{(5t+1)(t+1)} \\
&= \int_0^1 \left(\frac{5/2}{5t+1} - \frac{1/2}{t+1}\right)\ dt = \left[\frac{\log|5t+1| - \log|t+1|}{2}\right]_0^1 = \log\sqrt{3}.
\end{aligned}$$

■

3.4.4 Integration by Parts

Let f and G be differentiable functions such that fG is defined. Then the product rule gives

$$\frac{d}{dx}[f(x)G(x)] = f'(x)G(x) + f(x)G'(x),$$

i.e.

$$f(x)G'(x) = \frac{d}{dx}\left[f(x)G(x)\right] - f'(x)G(x).$$

Writing $G' = g$ and integrating, this becomes

$$\int f(x)g(x)\ dx = f(x)G(x) - \int f'(x)G(x)\ dx.$$

This is the formula for *integration by part.* For definite integrals it becomes

$$\int_a^b f(x)g(x)\ dx = [f(x)G(x)]_a^b - \int_a^b f'(x)G(x)\ dx,$$

assuming that f and g are defined on $[a, b]$.

Example 3.30. Find

(a) $\displaystyle\int x\sqrt{x+1}\ dx.$ (b) $\displaystyle\int \sin^{-1} x\ dx.$

Solution. (a) Putting $f(x) = x$ and $g(x) = \sqrt{x+1}$ gives

$$\begin{aligned}
\int x\sqrt{x+1}\ dx &= x\frac{(x+1)^{3/2}}{3/2} - \int 1\frac{(x+1)^{3/2}}{3/2}\ dx \\
&= \frac{2}{3}x(x+1)^{3/2} - \frac{2}{3}\cdot\frac{2}{5}(x+1)^{5/2} + c \\
&= \frac{2}{15}(x+1)^{3/2}(3x-2) + c.
\end{aligned}$$

(b) The 'trick' here is to think of $\sin^{-1} x$ as $(\sin^{-1} x)\cdot 1$ so that putting $f(x) = \sin^{-1} x$ and $g(x) = 1$ gives

$$\begin{aligned}
\int \sin^{-1} x\ dx &= \int (\sin^{-1} x)\cdot 1\ dx = (\sin^{-1} x)x - \int \frac{1}{\sqrt{1-x^2}}x\ dx \\
&= x\sin^{-1} x + \int \frac{(-2x)\ dx}{2\sqrt{1-x^2}}.
\end{aligned}$$

Then, substituting $1 - x^2 = u, -2x\ dx = \ du$,

$$\begin{aligned}\int \sin^{-1} x\ dx &= x \sin^{-1} x + \int \frac{du}{2\sqrt{u}} = x \sin^{-1} x + \sqrt{u} + c \\ &= x \sin^{-1} x + \sqrt{1 - x^2} + c.\end{aligned}$$

■

Chapter 4

Introduction to differential equations

4.1 Introduction

Definition 4.1. An equation involving an *independent* variable x, a *dependent* variable y and derivatives of y with respect to x up to and including the n^{th} derivative, is called an *ordinary differential equation* or ODE of *order* n.

For example,

$$y\frac{dy}{dx} = 2x(y+1) \quad \text{and} \quad \frac{d^2y}{dx^2} - \frac{dy}{dx} - 6y = 0$$

or equivalently,

$$yy^{t} = 2x(y+1) \quad \text{and} \quad y^{t/} - y^{t} - 6y = 0,$$

are first and second order ODEs, respectively.

4.2 First Order Separable Equations

Definition 4.2. A first order ODE is called *separable* if it can be written in the form

$$\frac{dy}{dx} = g(x)h(y) \tag{4.1}$$

where g and h are real functions.

Method. We assume that g and $1/h$ can be integrated. Then using the substitution $y = y(x)$ and equation 4.1 we have

$$\int \frac{1}{h(y)}dy = \int \frac{1}{h(y(x))}\frac{dy}{dx}dx = \int g(x)dx.$$

In effect, we have *separated the variables* x and y in 4.1 so that the left hand side involves the variable y only and the right hand side involves the variable x only. Thus

$$\int \frac{dy}{h(y)} = \int g(x)dx$$

is the *general solution* of equation 4.1, valid for x in any interval where $g(x)$ and $h(y)$ are defined and $h(y) \neq 0$.

If $k \in \mathbb{R}$ and $h(k) = 0$ then $y = k$ will also be a solution of equation 4.1, valid for $x \in \mathbb{R}$.

Example 4.3. Find the general solutions of

a $\dfrac{dy}{dx} = e^{x-2y}$, b $x\dfrac{dy}{dx} = \dfrac{1+x}{2-\sin y}$, c $\dfrac{dy}{dx} - py = 0$,

where $p \in \mathbb{R}$. Find, also, the *particular solution* of b for $x > 0$ which satisfies $y = 2\pi$ when $x = 1$ (denoted $y(1) = 2\pi$).

Solution. a Here,

$$\frac{dy}{dx} = e^{x-2y} = e^x e^{-2y} = e^x \frac{1}{e^{2y}}$$

so that

$$\int e^{2y} dy = \int e^x dx,$$

i.e.

$$\frac{e^{2y}}{2} = e^x + C,$$

i.e.

$$y = \frac{1}{2} \log 2(e^x + C)\ ,$$

where C is a parameter. This is the general solution.

b Here,

$$\frac{dy}{dx} = (\frac{1+x}{x})\frac{1}{2-\sin y}$$

so that

$$\int (2 - \sin y) dy = \int (\frac{1}{x} + 1) dx,$$

i.e.

$$2y + \cos y = \log|x| + x + C,$$

where C is a parameter. This is the general solution. For $x > 0$ this becomes

$$2y + \cos y = \log x + x + C.$$

Substituting $x = 1$ and $y = 2\pi$ gives $C = 4\pi$. Hence, the required particular solution is

$$2y + \cos y = \log x + x + 4\pi \quad (x > 0).$$

c Here,

$$\frac{dy}{dx} = py$$

so that, for $y \neq 0$,

$$\int \frac{dy}{y} = \int p\ dx,$$

i.e.

$$\log |y| = px + C,$$

i.e.

$$|y| = e^{px+C} = e^{px}e^{C},$$

i.e.

$$|y| = De^{px},$$

where $D > 0$ is a parameter. This together with $y = 0$ is the general solution. Since $|y| = \pm y$, we can write the general solution in the form

$$y = Ee^{px},$$

where E (any real number) is a parameter. This includes the solution $y = 0$.

■

Remark 4.4. If the differential equation has an initial conditions for y and there derivatives, then the equation has *particular solution.*

Example 4.5. Find a particular solution of

$$\frac{dy}{dx} = x\sqrt{1-y^2},$$

valid on an interval containing 0 and satisfying $y(0) = 0$.

Solution. For $y^2 \neq 1$, separating the variables gives

$$\int \frac{dy}{\sqrt{1-y^2}} = \int x dx,$$

i.e.

$$\sin^{-1} y = \frac{x^2}{2} + C,$$

where C is a parameter. For $y(0) = 0$ we must take $C = 0$, giving the particular solution

$$\sin^{-1} y = \frac{x^2}{2}.$$

Recalling the definition of $\sin^{-1}$ and remembering that $y^2 \neq 1$, this is equivalent to

$$y = \sin\left(\frac{x^2}{2}\right) \qquad \left(-\frac{\pi}{2} < \frac{x^2}{2} < \frac{\pi}{2}\right)$$

i.e.

$$y = \sin\left(\frac{x^2}{2}\right) \qquad (-\sqrt{\pi} < x < \sqrt{\pi})$$

which is a suitable particular solution. ■

4.3 First Order Homogeneous Equations

Definition 4.6. A first order ODE is called *homogeneous* if it can be written in the form

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

where f is a real function and F is a real valued function of two real variables. When setting $y = vx$, the function $F(x, y) = f(v)$.

Example 4.7. To check that the equation

$$\frac{dy}{dx} = \frac{x^4 + y^4}{xy^3}$$

is homogeneous, set $y = vx$, then for $x, v \neq 0$

$$\frac{x^4 + (vx)^4}{x(vx)^3} = \frac{x^4 + v^4x^4}{xv^3x^3} = \frac{x^4(1 + v^4)}{x^4v^3} = \frac{1 + v^4}{v^3}$$

Method. To solve FOHE we change the variables from (x, y) to (x, v) by making the substitution $y = vx$. This converts the given equation into a separable ODE. The complete substitution is

$$y = vx \qquad \text{and} \frac{dy}{dx} = v + x\frac{dv}{dx}.$$

Example 4.8. Find the general solution of

$$y + \sqrt{x^2 + y^2} - x\frac{dy}{dx} = 0 \qquad (x > 0).$$

Solution. Here,

$$\frac{dy}{dx} = \frac{y + \sqrt{x^2 + y^2}}{x}.$$

Substituting $y = vx$ gives

$$v + x\frac{dv}{dx} = \frac{vx + \sqrt{x^2 + v^2x^2}}{x} = v + \sqrt{1 + v^2}$$

(since $x > 0 \Rightarrow \sqrt{x^2} = x$). Hence,

$$\frac{dv}{dx} = \frac{1}{x}\sqrt{1 + v}.$$

Separating the variables leads to

$$\int \frac{dv}{\sqrt{1 + v^2}} = \int \frac{dx}{x},$$

i.e.

$$\log\left|v + \sqrt{1 + v^2}\right| = \log|x| + C,$$

i.e.

$$\log\left(v + \sqrt{1 + v^2}\right) = \log x + C,$$

where C is a parameter. Replacing v with y/x we get

$$\log\left(\frac{y}{x}+\sqrt{1+\frac{y^2}{x^2}}\right)=\log x+C.$$

This is the required general solution. We can simplify it as follows. Let $D=e^C$ so that $D>0$ and $C=\log D$. Then the general solution becomes

$$\log\left(\frac{y}{x}+\sqrt{1+\frac{y^2}{x^2}}\right)=\log Dx,$$

i.e. (since log is injective)

$$\frac{y}{x}+\sqrt{1+\frac{y^2}{x^2}}=Dx,$$

i.e. (since $x>0$)

$$y+\sqrt{x^2+y^2}=Dx^2.$$

Hence we can write the general solution in the form

$$\sqrt{x^2+y^2}=Dx^2-y\ (x>0)$$

where $D>0$ is a parameter. ■

4.4 First Order Linear Equations

Definition 4.9. An ODE is called n^{th} *order linear* if it can be written in the form

$$a_n(x)\frac{d^ny}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=f(x),$$

where a_n, a_{n-1} . . ., a_1, a_0 and f are real functions and a_n is not the zero function.

Thus a first order linear ODE can be written in the form

$$a_1(x)\frac{dy}{dx}+a_0(x)y=f(x),$$

where a_1, a_0 and f are real function, and a_1 is not the zero function. In fact, the *standard form* of such an equation is

$$\frac{dy}{dx}+p(x)y=q(x)$$

where p and q are real functions, continuous on an interval I.

Method. Let P be any antiderivative of p, and let μ be a real function, continuous on I and such that

$$|\mu(x)|=e^{P(x)}\quad(x\in I).$$

Since μ is non-zero on I, it follows from the intermediate value theorem that either

$$\mu(x)=e^{P(x)}\quad(x\in I)\quad\text{or}\quad\mu(x)=-e^{P(x)}\quad(x\in I).$$

In both cases, $\mu'(x) = \mu(x)p(x)$ so that

$$\frac{d}{dx}\left(\mu(x)y\right) = \mu(x)\frac{dy}{dx} + \mu'(x)y = \mu(x)\frac{dy}{dx} + \mu(x)p(x)y = \mu(x)\left(\frac{dy}{dx} + p(x)y\right).$$

and

$$\frac{d}{dx}\left(\mu(x)y\right) = \mu(x)q(x).$$

Integrating gives

$$\mu(x)y = \int \mu(x)q(x)dx,$$

which is the general solution.

Example 4.10. Find the general solutions of

(a) $\dfrac{dy}{dx} + 2xy = e^{-x^2}$,

(b) $(9-x^2)\dfrac{dy}{dx} + 6y = \left(9-x^2\right)^2$,

(c) $\dfrac{dy}{dx} - y\tan x = \sin x\cos x$.

Solution. We shall use the above notation.

(a) Here, $p(x) = 2x$ and $q(x) = e^{-x^2}$ Then

$$P(x) = \int p(x)dx = \int 2xdx = x^2 + C = x^2 \qquad \text{(taking } C = 0\text{)}$$

and

$$e^{P(x)} = e^{x^2} = |e^{x^2}|.$$

So we can take $\mu(x) = e^{x^2}$. Hence the required general solution is

$$\mu(x)y = \int \mu(x)q(x)dx,$$

i.e.

$$e^{x^2}y = \int e^{x^2}e^{-x^2}dx = \int dx = x + C,$$

i.e.

$$y = e^{-x^2}(x + C),$$

where C is a parameter.

(b) Here,

$$\frac{dy}{dx} + \frac{6}{9-x^2}y = 9 - x^2$$

so that $p(x) = 6/(9-x^2)$ and $q(x) = 9 - x^2$. Then

$$\begin{aligned} P(x) &= \int p(x)dx = \int \frac{6}{9-x^2}dx = \int \left[\frac{1}{3+x} + \frac{1}{3-x}\right]dx \\ &= \log|3+x| - \log|3-x| + C = \log\left|\frac{3+x}{3-x}\right| \qquad \text{(taking } C = 0\text{).} \end{aligned}$$

and

$$e^{P(x)} = \exp\left(\log\left|\frac{3+x}{3-x}\right|\right) = \left|\frac{3+x}{3-x}\right|.$$

So we can take

$$\mu(x) = \frac{3+x}{3-x}.$$

Hence, the required general solution is

$$\mu(x)y = \int \mu(x)q(x)dx,$$

i.e.

$$\left(\frac{3+x}{3-x}\right)y = \int\left(\frac{3+x}{3-x}\right)(9-x^2)dx = \int(3+x)^2dx = \frac{1}{3}(3+x)^3 + C,$$

i.e.

$$y = \frac{1}{3}(3+x)^2(3-x) + C\left(\frac{3-x}{3+x}\right),$$

where C is a parameter.

c Here, $p(x) = -\tan x$ so we can take $P(x) = \log|\cos x|$ and $\mu(x) = \cos x$. Hence, the required general solution is

$$\begin{aligned}(\cos x)y &= \int(\cos x)(\sin x\cos x)dx\\ &= \int\cos^2 x\sin x dx = -\frac{1}{3}\cos^3 x + C,\end{aligned}$$

i.e.

$$y = -\frac{1}{3}\cos^2 x + C\sec x,$$

where C is a parameter.

■

4.5 Second Order Linear Equations

Definition 4.11. The second order linear equations with *constant coefficients* is an ODE of the form

$$ay'' + by' + cy = f(x)$$

where $a, b, c \in \mathbb{R}$ with $a \neq 0$ and f is a real function.

Method. To solve the last equation we begin by solving the corresponding *homogeneous equation* (HE),

$$ay'' + by' + cy = 0.$$

The general solution of the homogeneous equation is called the *complementary function.* Its depends on the roots of the quadric equation

$$am^2 + bm + c = 0$$

which is called the *auxiliary equation.* There are three possibilities: two real roots, one real root or two conjugate complex roots.

Theorem 4.12. *The roots of the auxiliary equation have three possibilities*

1 *Two distinct real roots p and q , then the complementary function is*

$$y = Ae^{px} + Be^{qx}.$$

2 *One real root r, then the complementary function is*

$$y = (A + Bx)e^{rx}.$$

3 *Two conjugate complex roots $s+it$ and $s-it$ where $(s,t \in \mathbb{R})$, then the complementary function is*

$$y = (A\cos tx + B\sin tx)e^{sx}.$$

In each case, A and B are parameters.

Proof. Let $am^2 + bm + c = 0$ be the auxiliary equation of the (HE).

1 Since p and q are the roots of the auxiliary equation, it follows that $p+q = -b/a$ and $pq = c/a$ So the (HE)

$$ay'' + by' + cy = 0$$

is equivalent to

$$y'' - (p+q)y' + pqy = 0,$$

i.e.

$$(y' - qy)' - p(y' - qy) = 0.$$

Let $y' - qy = u$. Then

$$u' - pu = 0.$$

This is a separable equation with general solution $u = De^{px}$, where D is a parameter. Hence,

$$y' - qy = De^{px}.$$

This is a linear equation. We can take $\mu(x) = e^{-qx}$ as an integrating factor to get the general solution

$$e^{-qx}y = \int De^{px}e^{-qx}dx$$

i.e.

$$e^{-qx}y = \int De^{(p-q)x}dx = \frac{D}{p-q}e^{(p-q)x} + B,$$

i.e.

$$y = Ae^{px} + Be^{qx},$$

where A and B are parameters.

(2) The argument for (1) applies with $p = q = r$, we get

$$e^{-rx}y = \int De^{rx}e^{-rx}dx = \int Ddx = Dx + E,$$

i.e.

$$y = (Dx + E)e^{rx},$$

where D and E are parameters.

(3) The argument for (1) applies with $p = s + it$ and $q = s - it$ giving

$$y = Ae^{(s+it)x} + Be^{(s-it)x} = Ae^{sx}e^{itx} + Be^{sx}e^{-itx},$$

i.e.

$$y = Ae^{sx}(\cos tx + i \sin tx) + Be^{sx}(\cos tx - i \sin tx)$$

i.e.

$$y = e^{sx}\left((A + B)\cos tx + i(A - B)\sin tx\right),$$

i.e.

$$y = e^{sx}(M \cos tx + N \sin tx)\ ,$$

where M and N are parameters.

□

Example 4.13. Find the general solution of each of the following homogeneous equations.

(a) $y'' + y' - 6y = 0$, (b) $9y'' - 12y' + 4y = 0$, (c) $y'' + y' + y = 0$.

Solution. (a) The auxiliary equation, $m^2 + m - 6 = 0$, has roots -3 and 2. So the general solution is

$$y = Ae^{-3a} + Be^{2x},$$

where A and B are parameters.

(b) The auxiliary equation, $9m^2 - 12m + 4 = 0$, has root $2/3$ (only). So the general solution is

$$y = (A + Bx)e^{2x/3},$$

where A and B are parameters.

(c) The auxiliary equation, $m^2 + m + 1 = 0$, has roots $-(1/2) \pm i(\sqrt{3}/2)$. So the general solution is

$$y = e^{-x/2}\left(A \cos \frac{\sqrt{3}\,x}{2} + B \sin \frac{\sqrt{3}\,x}{2}\right),$$

where A and B are parameters.

■

Lemma 4.14. *Let $y = u$ be a solution of*

$$ay'' + by' + cy = f(x) \ ,$$

let $y = v$ be a solution of

$$ay'' + by' + w = g(x)$$

and let $\alpha, \beta \in \mathbb{R}$. Then $y = \alpha u + \beta v$ is a solution of

$$ay'' + by' + c = \alpha f(x) + \beta g(x) \ .$$

Proof. Let $y = \alpha u + \beta v$, then

$$\begin{aligned} a(\alpha u + \beta v)'' &+ b(\alpha u + \beta v)' + c(\alpha u + \beta v) \\ &= a(\alpha u'' + \beta v'' + b(\alpha u' + \beta v') + c(\alpha u + \beta v) \\ &= \alpha(au'' + bu' + cu) + \beta(av'' + bv' + cv) \\ &= \alpha f(x) + \beta g(x) \end{aligned}$$

and the proof is complete. □

Theorem 4.15. *Let $y = y_c$ be the complementary function and let $y = y_i$ be a particular integral of the general equation*

$$ay'' + by' + c = f(x).$$

Then the general solution is

$$y = y_c + y_i.$$

Method. To solve the general equation

$$ay'' + by' + w = f(x)$$

for certain types of function f. The problem is to find a particular integral.

1. If $f(x) = he^{kx}$, where $h, k \in \mathbb{R}$, then try, as a particular integral,
$$\begin{cases} He^{kx} & \text{if k is not a root of the auxiliary equation} \\ Hxe^{kx} & \text{if k is one of tow roots of the auxiliary equation} \\ Hx^2e^{kx} & \text{if k is the only root of the auxiliary equation} \end{cases}.$$

2. If $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, where $a_n, a_{n-1}, \ldots, a_1, a_0 \in \mathrm{R}$, then try, as a particular integral,
$$\begin{cases} P(x) & \text{if 0 is not a root of the auxiliary equation} \\ xP(x) & \text{if 0 is one of tow roots of the auxiliary equation} \\ x^2P(x) & \text{if 0 is the only root of the auxiliary equation} \end{cases},$$
where $P(x) = A_nx^n + A_{n-1}x^{n-1} + \cdots + A_1x + A_0$.

③ If $f(x) = u\cos vx + w\sin vx$, where $u, v, w \in \mathbb{R}$, then try, as a particular integral,
$$\begin{cases} Q(x) & \text{if } iv \text{ is not a root of the auxiliary equation} \\ xQ(x) & \text{if } iv \text{ is a root of the auxiliary equation} \end{cases},$$
where $Q(x) = U\cos vx + W\sin vx$.

The constants $H, A_n, A_{n-1}, \ldots, A_1, A_0, U$ and W are found by substituting the proposed particular integral in the general equation.

Example 4.16. Find the general solution of each of the following equations. Find also the particular solution when additional conditions are given.

ⓐ $y'' + 5y' + 6y = 3e^{-2x}$,

ⓑ $y'' + 3y' + 2y = 2x^2 + 1$, where $y(0) = 1$ and $y'(0) = 0$.

Solution. ⓐ The auxiliary equation, $m^2 + 5m + 6 = 0$, has roots -3 and -2. So the complementary function is

$$y = Ae^{-3x} + Be^{-2x}.$$

For a particular integral we try $y = Hxe^{-2x}$. Then

$$y' = H(-2xe^{-2x} + e^{-2x}) \quad \text{and} \quad y'' = H(4xe^{-2x} - 4e^{-2x}).$$

Substituting in the given ODE,

$$H(4xe^{-2x} - 4e^{-2x}) + 5H(-2xe^{-2x} + e^{-2x}) + 6Hxe^{-2x} = 3e^{-2x},$$

i.e.

$$He^{-2x} = 3e^{-2x}.$$

Hence $H = 3$ and the required general solution is

$$y = Ae^{-3x} + Be^{-2x} + 3xe^{-2x},$$

where A and B are parameters.

ⓑ The auxiliary equation, $m^2 + 3m + 2 = 0$, has roots -2 and -1, So the complementary function is

$$y = Ae^{-2x} + Be^{-x}.$$

For a particular integral we try

$$y = Px^2 + Qx + R.$$

Then

$$y' = 2Px + Q \quad \text{and} \quad y'' = 2P.$$

Substituting in the given ODE,

$$(2P) + 3(2Px + Q) + 2(Px^2 + Qx + R) = 2x^2 + 1,$$

i.e.
$$2Px^2 + (6P + 2Q)x + (2P + 3Q + 2R) = 2x^2 + 1,$$
so that, equating coefficients,
$$2P = 2, \quad 6P + 2Q = 0, \quad 2P + 3Q + 2R = 1.$$
Hence $P = 1, Q = -3, R = 4$ and the required general solution is
$$y = Ae^{-2x} + Be^{-x} + x^2 - 3x + 4,$$
where A and B are parameters.

Differentiating the general solution we have
$$y' = -2Ae^{-2x} - Be^{-x} + 2x - 3.$$
Using the additional conditions, $y(0) = 1$ and $y'(0) = 0$, gives
$$A + B + 4 = 1 \quad \text{and} \quad -2A - B - 3 = 0,$$
i.e.
$$A + B = -3 \quad \text{and} \quad 2A + B = -3.$$
Hence $A = 0, B = -3$ and the required particular solution is
$$y = -3e^{-x} + x^2 - 3x + 4.$$

■