Multivariable Calculus Notes ©

David A. Santos

Philadelphia, PA October 22, 2003

Contents

1 Brief Review of Linear Algebra **1**

1.1 $\mathbb{R}^n$ 1
1.2 $\mathbf{M}_{m\times n}(\mathbb{R})$ 7
1.3 Determinants 13
1.4 Linear Transformations 16
1.5 Inner Products 22
1.6 Cross product in $\mathbb{R}^3$ 27
1.7 Lines and Planes in $\mathbb{R}^3$ 31
1.8 Topology of $\mathbb{R}^n$ 37
1.9 Quadratic Forms 39
1.10 Quadratic Surfaces 41
1.11 Canonical Surfaces in $\mathbb{R}^3$ 44
1.12 Parametric Curves and Surfaces 49
1.13 Frenet-Serret Formulæ 54
1.14 Limits 56

2 Differentiation **59**

2.1 Local Study of Functions 59
2.2 Definition of the Derivative 67
2.3 The Jacobi Matrix 73
2.4 Gradients and Directional Derivatives 81
2.5 Extrema 85
2.6 Lagrange Multipliers 93
2.7 Arithmetic Mean-Geometric Mean Inequality 97

3 Integration **105**
3.1 Differential Forms 105
3.2 Integrating in $\wedge^0(\mathbb{R}^n)$ 111
3.3 Integrating in $\wedge^1(\mathbb{R}^n)$ 112
3.4 Closed and Exact Forms 122
3.5 Integrating in $\wedge^2(\mathbb{R}^2)$ 127
3.6 Change of Variables in $\wedge^2(\mathbb{R}^2)$ 141
3.7 Change to Polar Co-ördinates 149
3.8 Integrating in $\wedge^3(\mathbb{R}^3)$ 157
3.9 Change of Variables in $\wedge^3(\mathbb{R}^3)$ 160
3.10 Integration in $\wedge^2(\mathbb{R}^3)$ 165
3.11 Green's, Stokes', and Gauß' Theorems 171

Chapter 1

Brief Review of Linear Algebra

1.1 $\mathbb{R}^n$

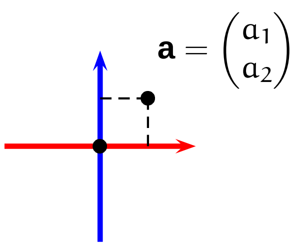

Figure 1.1: A point in $\mathbb{R}^2$.

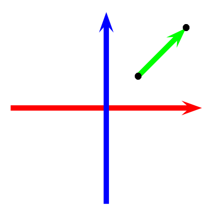

Figure 1.2: A bi-point in $\mathbb{R}^2$.

1 Definition $\mathbb{R}^n$ is the set of real n-tuples

$$\mathbb{R}^n = \left\{ \mathbf{x} : \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, x_k \in \mathbb{R} \right\}.$$

These n-tuples are called *points*.

Thus $\mathbb{R}^2$ is the collection of points on the plane (see figure 1.1) and $\mathbb{R}^3$ is the collection of points in three-dimensional space.

Points on their own are very boring entities. They are devoid of arithmetical properties (you cannot "add" two points), they are simply a list of real numbers.

Suppose now that we are given two points **x** and **y** in $\mathbb{R}^n$. Starting from **x** we move on a straight line towards **y**, and we denote this ordered movement by the notation $[\mathbf{x}, \mathbf{y}]$ (read the "bi-point **x**, **y**"). This movement involves a displacement of each of the n co-ördinates, the k-th co-ördinate being displaced $y_k - x_k$ units. If we let $a_k = y_k - x_k$ record the displacement of the k-th co-ördinate, then we say that $[\mathbf{x}, \mathbf{y}]$ is a representative of the *vector*

$$\overrightarrow{\mathbf{a}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Notice that there are infinitely many bi-points representing the same vector. Thus in figure 1.2 we see a representative bi-point of a vector in $\mathbb{R}^2$. The same vector would represent any parallel displacement of this bi-point.

2 Example Consider the points

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{y}_1 = \begin{pmatrix} 3 \\ -4 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Though the bi-points $[\mathbf{x_1}, \mathbf{y_1}]$ and $[\mathbf{x_2}, \mathbf{y_2}]$ are in different locations on the plane, they represent the same vector

$$\overrightarrow{\mathbf{a}} = \begin{bmatrix} 3-1 \\ -4-2 \end{bmatrix} = \begin{bmatrix} 5-3 \\ -1-5 \end{bmatrix}.$$

These two bi-points are parallel and have the same length, moreover, they are pointing in the same direction.

3 Definition A vector

$$\overrightarrow{\mathbf{a}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$$

is an equivalence class denoting all those bi-points whose displacement in the k-th co-ördinate is a_k.

We write points of $\mathbb{R}^n$ using bold-face letters and parentheses, as in

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

We write vectors in $\mathbb{R}^n$ using bold-face letters with arrows on top and square brackets, as in

$$\overrightarrow{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

4 Definition If $\overrightarrow{\mathbf{a}}$ and $\overrightarrow{\mathbf{b}}$ are two vectors in $\mathbb{R}^n$ their *vector sum* $\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}}$ is defined by the co-ördinatewise addition

$$\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}. \tag{1.1}$$

5 Definition A real number $\alpha \in \mathbb{R}$ will be called a *scalar*. If $\alpha \in \mathbb{R}$ and $\overrightarrow{\mathbf{a}} \in \mathbb{R}^n$ we define *scalar multiplication* of a vector and a scalar by the co-ördinatewise multiplication

$$\alpha\overrightarrow{\mathbf{a}} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{bmatrix}. \tag{1.2}$$

6 Theorem The operations of vector addition 1.1 and scalar multiplication 1.2 make $\mathbb{R}^n$ a *vector space*. That is, these operations satisfy

$$\forall(\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}}, \overrightarrow{\mathbf{c}}) \in (\mathbb{R}^n)^3, \ \forall(\alpha, \beta) \in \mathbb{R}^2,$$

❶ **Closure under vector addition:**

$$\vec{\mathbf{a}} + \vec{\mathbf{b}} \in \mathbb{R}^n \tag{1.3}$$

❷ **Closure under scalar multiplication:**

$$\alpha \vec{\mathbf{a}} \in \mathbb{R}^n \tag{1.4}$$

❸ **Commutativity of addition:**

$$\vec{\mathbf{a}} + \vec{\mathbf{b}} = \vec{\mathbf{b}} + \vec{\mathbf{a}} \tag{1.5}$$

❹ **Associativity:**

$$(\vec{\mathbf{a}} + \vec{\mathbf{b}}) + \vec{\mathbf{c}} = \vec{\mathbf{a}} + (\vec{\mathbf{b}} + \vec{\mathbf{c}}) \tag{1.6}$$

❺ **Existence of additive identity:**

$$\exists \vec{\mathbf{0}} \in \mathbb{R}^n : \vec{\mathbf{a}} + \vec{\mathbf{0}} = \vec{\mathbf{0}} + \vec{\mathbf{a}} = \vec{\mathbf{a}} \tag{1.7}$$

❻ **Existence of additive inverses:**

$$\exists - \vec{\mathbf{a}} \in \mathbb{R}^n : \vec{\mathbf{a}} + (-\vec{\mathbf{a}}) = -\vec{\mathbf{a}} + \vec{\mathbf{a}} = \vec{\mathbf{0}} \tag{1.8}$$

❼ **Distributive Law:**

$$\alpha(\vec{\mathbf{a}} + \vec{\mathbf{b}}) = \alpha \vec{\mathbf{a}} + \alpha \vec{\mathbf{b}} \tag{1.9}$$

❽ **Distributive Law:**

$$(\alpha + \beta) \vec{\mathbf{a}} = \alpha \vec{\mathbf{a}} + \beta \vec{\mathbf{a}} \tag{1.10}$$

❾

$$1 \vec{\mathbf{a}} = \vec{\mathbf{a}} \tag{1.11}$$

❿

$$(\alpha\beta) \vec{\mathbf{a}} = \alpha(\beta \vec{\mathbf{a}}) \tag{1.12}$$

7 Definition A set of vectors $\overrightarrow{\mathbf{a}}_k \in \mathbb{R}^n$, $1 \leq k \leq l$ is said to be *linearly independent* if

$$\sum_{k=1}^{l} \alpha_k \overrightarrow{\mathbf{a}}_k = \overrightarrow{\mathbf{0}} \implies \alpha_1 = \alpha_2 = \cdots = \alpha_l = 0.$$

8 Definition An ordered set

$$\mathscr{A} = \{\overrightarrow{\mathbf{v}}_1, \overrightarrow{\mathbf{v}}_2, \ldots, \overrightarrow{\mathbf{v}}_n\}$$

of n linearly independent vectors in $\mathbb{R}^n$ is called an *ordered basis* for $\mathbb{R}^n$. This means that every vector in $\mathbb{R}^n$ can be uniquely written as a linear combination of the $\overrightarrow{\mathbf{v}}_k$, that is, if $\overrightarrow{\mathbf{u}} \in \mathbb{R}^n$ then there exist unique scalars α_k (called the *co-ördinates of* $\overrightarrow{\mathbf{v}}$ under $\mathscr{A}$) such that

$$\overrightarrow{\mathbf{u}} = \sum_{k=1}^{n} \alpha_k \overrightarrow{\mathbf{v}}_k := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_{\mathscr{A}} .$$

9 Example The family $\mathscr{A} = \{\overrightarrow{\mathbf{i}}, \overrightarrow{\mathbf{j}}, \overrightarrow{\mathbf{k}}\}$ with

$$\overrightarrow{\mathbf{i}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{\mathbf{j}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{\mathbf{k}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

forms an ordered basis for $\mathbb{R}^3$ (these is called the *natural basis for* $\mathbb{R}^3$). Any vector $\overrightarrow{\mathbf{u}}$ can be written uniquely as a linear combination of these vectors, for example

$$\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}_{\mathscr{A}} = 2\overrightarrow{\mathbf{i}} - 3\overrightarrow{\mathbf{j}} + 4\overrightarrow{\mathbf{k}}.$$

The family $\mathscr{B} = \{\overrightarrow{\mathbf{i}}, \overrightarrow{\mathbf{k}}, \overrightarrow{\mathbf{j}}\}$ forms a different ordered basis for $\mathbb{R}^3$. In this case

$$\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}_{\mathscr{B}} = 2\overrightarrow{\mathbf{i}} + 4\overrightarrow{\mathbf{j}} - 3\overrightarrow{\mathbf{k}}.$$

In most cases we will be using the standard ordered basis

$$\mathscr{A} = \{\overrightarrow{\mathbf{e}}_1, \overrightarrow{\mathbf{e}}_2, \ldots, \overrightarrow{\mathbf{e}}_n,\}$$

with

$$\overrightarrow{\mathbf{e}}_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

(a 1 *in the* k *slot and* 0*'s everywhere else). In such cases we will write*

$$\sum_{k=1}^{n} \alpha_k \overrightarrow{\mathbf{e}}_k = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

(without the $\mathscr{A}$ *subscript) rather than*

$$\sum_{k=1}^{n} \alpha_k \overrightarrow{\mathbf{e}}_k = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_{\mathscr{A}} .$$

10 Example Prove that the family $\mathscr{C} = \{\overrightarrow{\mathbf{b}}_1, \overrightarrow{\mathbf{b}}_2, \overrightarrow{\mathbf{b}}_3\}$ with

$$\overrightarrow{\mathbf{b}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{\mathbf{b}}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{\mathbf{b}}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

forms an ordered basis for $\mathbb{R}^3$ and find the co-ördinates of

$$\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$$

under $\mathscr{C}$.

Solution: We have

$$\alpha_1\overrightarrow{\mathbf{b}}_1+\alpha_2\overrightarrow{\mathbf{b}}_2+\alpha_3\overrightarrow{\mathbf{b}}_3=\begin{bmatrix}0\\0\\0\end{bmatrix}\iff\begin{array}{rcl}\alpha_1+\alpha_2+\alpha_3&=&0\\\alpha_2+\alpha_3&=&0\\\alpha_3&=&0.\end{array}$$

Solving this triangular system gives $\alpha_1=\alpha_2=\alpha_3=0$, whence $\mathscr{C}$ is a linearly independent family of 3 vectors and hence an ordered basis for $\mathbb{R}^3$. It is easy to verify that

$$\begin{bmatrix}2\\-3\\4\end{bmatrix}=5\begin{bmatrix}1\\0\\0\end{bmatrix}-7\begin{bmatrix}1\\1\\0\end{bmatrix}+4\begin{bmatrix}1\\1\\1\end{bmatrix}=5\overrightarrow{\mathbf{b}}_1-7\overrightarrow{\mathbf{b}}_2+4\overrightarrow{\mathbf{b}}_3=\begin{bmatrix}5\\-7\\4\end{bmatrix}_{\mathscr{C}}.$$

1.2 $\mathbf{M}_{m\times n}(\mathbb{R})$

11 Definition An $m\times n$ (m by n) *matrix* A with m rows and n columns is a rectangular array of the form

$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\cdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix},$$

where $\forall(i,j)\in\{1,2,\ldots,m\}\times\{1,2,\ldots,n\},\quad a_{ij}\in\mathbb{R}$.

As a shortcut, we often use the notation $A=[a_{ij}]$ to denote the matrix A with entries a_{ij}. Notice that when we refer to the matrix we put square brackets (as in "$[a_{ij}]$"), and when we refer to a specific entry we do not use the surrounding parentheses (as in "a_{ij}").

12 Example

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

is a 2×3 matrix and

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

is a 3×2 matrix

13 Definition We denote by $\mathbf{M}_{m\times n}(\mathbb{R})$ the set of all $m \times n$ matrices with real number entries. If $m = n$ we use the abbreviated notation $\mathbf{M}_n(\mathbb{R}) = \mathbf{M}_{n\times n}(\mathbb{R})$. $\mathbf{M}_n(\mathbb{R})$ is thus the set of all *square* matrices of size n with real entries.

14 Definition The $n \times n$ *zero matrix* $\mathbf{0}_n \in \mathbf{M}_n(\mathbb{R})$ is the matrix with 0's everywhere,

$$\mathbf{0}_n = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

15 Definition The $n \times n$ *identity matrix* $\mathbf{I}_n \in \mathbf{M}_n(\mathbb{R})$ is the matrix with 1's on the main diagonal and 0's everywhere else,

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

16 Definition The *main diagonal* of a matrix matrix $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$ is the set $\{a_{ii} | i \leq \min(m, n)\}$. A matrix is *diagonal* if every entry off its main diagonal is 0.

17 Definition $A \in \mathbf{M}_{m\times n}(\mathbb{R})$ is said to be *upper triangular* if

$$(\forall (i, j) \in \{1, 2, \cdots, n\}^2), (i > j, \ a_{ij} = 0),$$

that is, every element below the main diagonal is 0. Similarly, A is *lower triangular* if

$$(\forall (i, j) \in \{1, 2, \cdots, n\}^2), (i < j, \ a_{ij} = 0),$$

that is, every element above the main diagonal is 0.

18 Example The matrix $A \in \mathbf{M}_{3\times 4}(\mathbb{R})$ shewn is upper triangular and $B \in \mathbf{M}_4(\mathbb{R})$ is lower triangular.

$$A = \begin{bmatrix} 1 & a & b & c \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 1 & 1 & t & 1 \end{bmatrix}$$

19 Definition Let $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$, $B = [b_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$ and $\alpha \in \mathbb{R}$. The matrix $A + \alpha B$ is the matrix $C = [c_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$ with entries $c_{ij} = a_{ij} + \alpha b_{ij}$.

20 Example Let

$$M = \begin{bmatrix} a & -2a & c \\ 0 & -a & b \\ a+b & 0 & -1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 2a & c \\ a & b-a & -b \\ a-b & 0 & -1 \end{bmatrix}.$$

Then

$$M + N = \begin{bmatrix} a+1 & 0 & 2c \\ a & b-2a & 0 \\ 2a & 0 & -2 \end{bmatrix}, \quad 2M = \begin{bmatrix} 2a & -4a & 2c \\ 0 & -2a & 2b \\ 2a+2b & 0 & -2 \end{bmatrix}.$$

21 Definition Let $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$ and $B = [b_{ij}] \in \mathbf{M}_{n\times p}(F)$. Then the matrix product AB is defined as the matrix $C = [c_{ij}]$ with entries $c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}$.

In order to obtain the ij-th entry of the matrix AB we multiply elementwise the i-th row of A by the j-th column of B. Observe that AB is a $m \times p$ matrix.

Observe that we use juxtaposition rather than a special symbol to denote matrix multiplication. This will simplify notation.

22 Example Let

$$M = \begin{bmatrix} a & -2a & c \\ 0 & -a & b \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 2a \\ a-b & 0 \\ 1 & 2 \end{bmatrix}$$

be matrices over $\mathbb{R}$.Then

$$MN = \begin{bmatrix} a - 2a(a-b) + c & 2a^2 + 2c \\ -a(a-b) + b & 2b \end{bmatrix}, \quad NM = \begin{bmatrix} a & -2a - 2a^2 & c + 2ab \\ a(a-b) & -2a(a-b) & (a-b)c \\ a & -4a & c + 2b \end{bmatrix}.$$

Matrix multiplication is not necessarily commutative.

23 Example

Solution: We have

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

over $\mathbb{R}$.

Observe then that the product of two non-zero matrices may be the zero matrix.

24 Theorem If $(A, B, C) \in \mathbf{M}_{m\times n}(F) \times \mathbf{M}_{n\times r}(F) \times \mathbf{M}_{r\times s}(F)$ we have

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

i.e., matrix multiplication is associative.

Proof To shew this we only need to consider the ij-th entry of each side. Both are equal to

$$\sum_{k=1}^{n} \sum_{k'=1}^{r} a_{ik} b_{kk'} c_{k'j}.$$

Even though matrix multiplication is not necessarily commutative, it is associative.

By virtue of associativity, a square matrix commutes with its powers, that is, if $A \in \mathbf{M}_n(\mathbb{R})$, and $(r, s) \in \mathbb{N}^2$, then $(A^r)(A^s) = (A^s)(A^r) = A^{r+s}$.

25 Definition Let $A = [a_{ij}] \in \mathbf{M}_n(\mathbb{R})$. Then the *trace* of A, denoted by $\mathrm{tr}\,(A)$ is the sum of the diagonal elements of A, that is

$$\mathrm{tr}\,(A) = \sum_{k=1}^{n} a_{kk}.$$

26 Theorem Let $A = [a_{ij}] \in \mathbf{M}_n(\mathbb{R}), B = [b_{ij}] \in \mathbf{M}_n(\mathbb{R}), \alpha \in \mathbb{R}$. Then

$$\mathrm{tr}\,(\alpha A) = \alpha \mathrm{tr}\,(A)\,, \tag{1.13}$$

$$\mathrm{tr}\,(A + B) = \mathrm{tr}\,(A) + \mathrm{tr}\,(B)\,, \tag{1.14}$$

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)\,. \tag{1.15}$$

Proof The first assertion is trivial. To prove the second, observe that $AB = [\sum_{k=1}^{n} a_{ik}b_{kj}]$ and $BA = [\sum_{k=1}^{n} b_{ik}a_{kj}]$. Then

$$\mathrm{tr}\,(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \mathrm{tr}\,(BA)\,,$$

whence the theorem follows. ❑

27 Example Write

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \in \mathbf{M}_3(\mathbb{R})$$

as the sum of two 3×3 matrices E_1, E_2, with $\mathrm{tr}\,(E_2) = 10$.

Solution: There are infinitely many solutions. Here is one:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -9 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 10 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

28 Example Given a square matrix $A \in \mathbf{M}_4(\mathbb{R})$ such that $\text{tr}\left(A^2\right) = -4$, and

$$(A - \mathbf{I}_4)^2 = 3\mathbf{I}_4,$$

find $\text{tr}\,(A)$.

Solution:

$$\begin{aligned} \text{tr}\left((A - \mathbf{I}_4)^2\right) &= \text{tr}\left(A^2 - 2A + \mathbf{I}_4\right) \\ &= \text{tr}\left(A^2\right) - 2\text{tr}\,(A) + \text{tr}\,(\mathbf{I}_4) \\ &= -4 - 2\text{tr}\,(A) + 4 \\ &= -2\text{tr}\,(A)\,, \end{aligned}$$

and $\text{tr}\,(3\mathbf{I}_4) = 12$. Hence $-2\text{tr}\,(A) = 12$ or $\text{tr}\,(A) = -6$.

29 Definition The *transpose* of a matrix of a matrix $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R})$ is the matrix $A^T = B = [b_{ij}] \in \mathbf{M}_{n\times m}(F)$, where $b_{ij} = a_{ji}$.

30 Example We have

$$M = \begin{bmatrix} a & -2a & c \\ 0 & -a & b \\ a+b & 0 & -1 \end{bmatrix}, \quad M^T = \begin{bmatrix} a & 0 & a+b \\ -2a & -a & 0 \\ c & b & -1 \end{bmatrix}.$$

31 Theorem Let

$$A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R}),\ B = [b_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{R}),\ C = [c_{ij}] \in \mathbf{M}_{n\times r}(\mathbb{R}),\ \alpha \in \mathbb{R}.$$

Then

$$A^{TT} = A, \tag{1.16}$$

$$(A + \alpha B)^T = A^T + \alpha B^T, \tag{1.17}$$

$$(AC)^T = C^T A^T. \tag{1.18}$$

Proof The first two assertions are obvious. To prove the third put $A^T = (\alpha_{ij})$, $\alpha_{ij} = a_{ji}$, $C^T = (\gamma_{ij})$, $\gamma_{ij} = c_{ji}$, $AC = (u_{ij})$ and $C^TA^T = (v_{ij})$. Then

$$u_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj} = \sum_{k=1}^{n} \alpha_{ki}\gamma_{jk} = \sum_{k=1}^{n} \gamma_{jk}\alpha_{ki} = v_{ji},$$

whence the theorem follows. ❑

32 Definition A matrix $A \in \mathbf{M}_n(\mathbb{R})$ is *symmetric* if $A^T = A$. A matrix $B \in \mathbf{M}_n(\mathbb{R})$ is *skew-symmetric* if $B^T = -B$.

33 Example If

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, B = \begin{bmatrix} 0 & -2 & 3 \\ 2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix}$$

then A is symmetric and B is skew-symmetric.

34 Theorem Any matrix $A \in \mathbf{M}_n(\mathbb{R})$ can be written as the sum of a symmetric and a skew-symmetric matrix.

Proof Observe that

$$(A + A^T)^T = A^T + A^{TT} = A^T + A,$$

and so $A + A^T$ is symmetric. Also,

$$(A - A^T)^T = A^T - A^{TT} = -(A - A^T),$$

and so $A - A^T$ is skew-symmetric. We only need to write A as

$$A = (\frac{1}{2})(A + A^T) + (\frac{1}{2})(A - A^T)$$

to prove the assertion. ❑

1.3 Determinants

We now define the notion of *determinant* of a matrix $A \in \mathbf{M}_n(\mathbb{R})$. We will use an inductive definition, so that we can effect calculations of determinants quickly.

35 Definition The *determinant* of a square matrix $A \in \mathbf{M}_n(\mathbb{R})$, denoted by $\det A$ is defined inductively as follows.

1. If $n = 1$, $A = [a]$, then $\det A = a$.

2. If $n = 2$, $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, then $\det A = ad - bc$.

3. If $n \geq 3$, $A = [a_{ij}]$, let $A_{ij} \in \mathbf{M}_{n-1}(\mathbb{R})$ denote the matrix obtained by deleting the i-th row and the j-th column from A. Then

$$\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij},$$

(the development along the i-th row) or, alternatively,

$$\det A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij}.$$

(the development along the j-th column).

We have assumed that no matter which row or column we choose, we always obtain the same determinant. This seems like an arbitrary assumption, but in linear algebra courses we do see that this is indeed the case. The result is independent of our choice. It is therefore advantageous to choose that row, or column, with a maximal number of 0's. It also follows that $\det A = \det A^T$, and that the determinant of a triangular matrix is the product of its diagonal elements.

36 Example Find

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

by expanding along the first row.

Solution: We have

$$\begin{aligned} \det A &= 1(-1)^{1+1} \det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} + 2(-1)^{1+2} \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 3(-1)^{1+3} \det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \\ &= 1(45 - 48) - 2(36 - 42) + 3(32 - 35) = 0. \end{aligned}$$

37 Example Find

$$\det\begin{bmatrix} 1 & 0 & -1 & 1 \\ 2 & 0 & 0 & 1 \\ 666 & -3 & -1 & 1000000 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Solution: Since the second column has three 0's, it is advantageous to expand along it, and thus we are reduced to calculate

$$-3(-1)^{3+2}\det\begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Expanding this last determinant along the second column, the original determinant is thus

$$-3(-1)^{3+2}(-1)(-1)^{1+2}\det\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = -3(-1)(-1)(-1)(1) = 3.$$

38 Definition Let $A \in \mathbf{M}_n(\mathbb{R})$. A vector $\overrightarrow{\mathbf{v}} \neq \overrightarrow{\mathbf{0}}$ is an *eigenvector* and a scalar λ is an *eigenvalue* if

$$A\overrightarrow{\mathbf{v}} = \lambda\overrightarrow{\mathbf{v}}.$$

To find the eigenvalues of a matrix A it is necessary and sufficient to solve the equation (called the *characteristic equation*)

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

39 Example Since

$$\det\begin{bmatrix} \lambda-1 & 1 & 1 \\ -1 & \lambda-3 & -1 \\ 3 & -1 & \lambda+1 \end{bmatrix} = \lambda^3 - 3\lambda^2 - 4\lambda + 12 = (\lambda-2)(\lambda+2)(\lambda-3),$$

the matrix

$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

has eigenvalues $-2, 2, 3$.

40 Theorem The eigenvalues of a real symmetric matrix are all real.

41 Definition Let $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_k$ be k vectors in $\mathbb{R}^n$. The k-parallelotope spanned by the $\vec{\mathbf{v}}_i$ is the set

$$\left\{\sum_{i=1}^{k} t_i \vec{\mathbf{v}}_i : t_i \in [0;1]\right\}.$$

If

$$A = [\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_k]$$

is the $n \times k$ matrix having the k vectors as columns, then

$$\sqrt{\det A^T A}$$

is the k-*dimensional volume* of the k-parallelotope spanned by the $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_k$.

We will se later that $\det A^T A \geq 0$ *so the square root of it is defined.*

42 Theorem The signed area of a triangle in $\mathbb{R}^2$ spanned by the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$ is

$$\frac{1}{2}\det[\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2].$$

1.4 Linear Transformations

43 Definition Let V, W be two vector spaces over the field of real numbers. A function

$$L : \begin{array}{ccc} V & \to & W \\ \vec{\mathbf{a}} & \mapsto & L(\vec{\mathbf{a}}) \end{array},$$

is called a *linear transformation* if it is

- **Linear:** $L(\vec{\mathbf{a}} + \vec{\mathbf{b}}) = L(\vec{\mathbf{a}}) + L(\vec{\mathbf{b}})$,
- **Homogeneous:** $L(\alpha\vec{\mathbf{a}}) = \alpha L(\vec{\mathbf{a}})$, for $\alpha \in \mathbb{R}$.

These two properties can be condensed by saying that a linear transformation is a function $L : V \to W$ *such that*

$$L(\overrightarrow{\mathbf{a}} + \alpha \overrightarrow{\mathbf{b}}) = L(\overrightarrow{\mathbf{a}}) + \alpha L(\overrightarrow{\mathbf{b}}).$$

44 Example The trace map

$$\mathrm{tr}\,(\cdot) : \begin{array}{ccc} \mathbf{M}_n(\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \mathrm{tr}\,(A) \end{array}$$

is linear, since in view of Theorem 26,

$$\mathrm{tr}\,(A + B) = \mathrm{tr}\,(A) + \mathrm{tr}\,(B)\,, \quad \mathrm{tr}\,(\alpha A) = \alpha \mathrm{tr}\,(A)\,.$$

45 Example The transpose map

$$L : \begin{array}{ccc} \mathbf{M}_n(\mathbb{R}) & \to & \mathbf{M}_n(\mathbb{R}) \\ A & \mapsto & A^T \end{array}$$

is linear, since in view of Theorem 31,

$$L(A + \alpha B) = (A + \alpha B)^T = A^T + (\alpha B)^T = A^T + \alpha B^T = L(A) + \alpha L(B).$$

46 Example Let $X \in \mathbf{M}_n(\mathbb{R})$ be a fixed square matrix. Prove that the map

$$L : \begin{array}{ccc} \mathbf{M}_n(\mathbb{R}) & \to & \mathbf{M}_n(\mathbb{R}) \\ A & \mapsto & XAX \end{array}$$

is linear.

Solution: Let A, B be matrices in $\mathbf{M}_n(\mathbb{R})$ and let α be a scalar. We have

$$\begin{array}{rcl} L(A + \alpha B) & = & X(A + \alpha B)X \\ & = & XAX + X(\alpha B)X \\ & = & XAX + \alpha XBX \\ & = & L(A) + \alpha L(B), \end{array}$$

whence the claim follows.

47 Example Prove that the map

$$f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^3 \\ \begin{bmatrix} x \\ y \end{bmatrix} & \mapsto & \begin{bmatrix} x+2y \\ 2x \\ -y \end{bmatrix} \end{array}$$

is linear.

Solution: Put

$$\overrightarrow{\mathbf{v}}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \overrightarrow{\mathbf{v}}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

and let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} L(\overrightarrow{\mathbf{v}}_1 + \alpha\overrightarrow{\mathbf{v}}_2) &= L\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \alpha \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \\ &= L\left(\begin{bmatrix} x_1 + \alpha x_2 \\ y_1 + \alpha y_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} (x_1 + \alpha x_2) + 2(y_1 + \alpha y_2) \\ 2(x_1 + \alpha x_2) \\ -(y_1 + \alpha y_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 2y_1 \\ 2x_1 \\ -y_1 \end{bmatrix} + \alpha \begin{bmatrix} x_2 + 2y_2 \\ 2x_2 \\ -y_2 \end{bmatrix} \\ &= L(\overrightarrow{\mathbf{v}}_1) + \alpha L(\overrightarrow{\mathbf{v}}_2), \end{aligned}$$

whence L is linear.

If $\{\overrightarrow{\mathbf{v}}_i\}_{i\in[1;n]}$ is an ordered basis for $\mathbb{R}^n$, $\overrightarrow{\mathbf{a}} = \sum_{i=1}^{n} \alpha_i \overrightarrow{\mathbf{v}}_i$, and L is linear, then

$$L(\overrightarrow{\mathbf{a}}) = L\left(\sum_{i=1}^{n} \alpha_i \overrightarrow{\mathbf{v}}_i\right) = \sum_{i=1}^{n} \alpha_i L(\overrightarrow{\mathbf{v}}_i),$$

meaning that the action of L on an arbitrary vector $\overrightarrow{\mathbf{a}} \in \mathbb{R}^n$ is completely determined by the action L has on the given ordered basis of $\mathbb{R}^n$. This in turn gives the following.

48 Theorem Let

$$L: \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m \\ \overrightarrow{\mathbf{a}} & \mapsto & L(\overrightarrow{\mathbf{a}}) \end{array}$$

be a linear transformation. Then there is a unique matrix $A_L \in \mathbf{M}_{m\times n}(\mathbb{R})$ such that for all $\overrightarrow{\mathbf{a}} \in \mathbb{R}^n$,

$$L(\overrightarrow{\mathbf{a}}) = A_L\overrightarrow{\mathbf{a}}.$$

49 Example Find the matrix representation of the linear map in example 47 if

❶ both $\mathbb{R}^2$ and $\mathbb{R}^3$ have as ordered bases their standard ordered bases.

❷ $\mathbb{R}^2$ has the ordered basis

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

and $\mathbb{R}^3$ has the standard basis as ordered basis.

Solution:

❶ We have

$$L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad L\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix},$$

and hence the desired matrix is

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

❷ We have

$$L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix},$$

and hence the desired matrix is

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

50 Example Consider $L : \mathbb{R}^3 \to \mathbb{R}^3$, with

$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x - y - z \\ x + y + z \\ z \end{bmatrix}.$$

❶ Prove that L is a linear transformation.

❷ Find the matrix corresponding to L under the standard basis.

❸ Find the matrix corresponding to L under the ordered basis

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\},$$

for both the domain and the image of L.

Solution:

❶ Let $\alpha \in \mathbb{R}$. Put $\overrightarrow{\mathbf{u}}_1 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\overrightarrow{\mathbf{u}}_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Then

$$\begin{aligned} L(\overrightarrow{\mathbf{u}} + \alpha\overrightarrow{\mathbf{u}}_2) &= L\left(\begin{bmatrix} x + \alpha a \\ y + \alpha b \\ z + \alpha c \end{bmatrix}\right) \\ &= \begin{bmatrix} (x + \alpha a) - (y + \alpha b) - (z + \alpha c) \\ (x + \alpha a) + (y + \alpha b) + (z + \alpha c) \\ z + \alpha c \end{bmatrix} \\ &= \begin{bmatrix} x - y - z \\ x + y + z \\ z \end{bmatrix} + \alpha \begin{bmatrix} a - b - c \\ a + b + c \\ c \end{bmatrix} \\ &= L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) + \alpha L\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) \\ &= L(\overrightarrow{\mathbf{u}}) + \alpha L(\overrightarrow{\mathbf{u}}_2) \end{aligned}$$

proving that L is a linear transformation.

❷ We have $L\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $L\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}-1\\1\\0\end{bmatrix}$, and $L\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}-1\\1\\1\end{bmatrix}$, whence the desired matrix is

$$\begin{bmatrix}1 & -1 & -1\\1 & 1 & 1\\0 & 0 & 1\end{bmatrix}.$$

❸ We have

$$L\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix} = 0\begin{bmatrix}1\\0\\0\end{bmatrix} + 1\begin{bmatrix}1\\1\\0\end{bmatrix} + 0\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\end{bmatrix}_{\mathscr{B}},$$

$$L\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}0\\2\\0\end{bmatrix} = -2\begin{bmatrix}1\\0\\0\end{bmatrix} + 2\begin{bmatrix}1\\1\\0\end{bmatrix} + 0\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}-2\\2\\0\end{bmatrix}_{\mathscr{B}},$$

and

$$L\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}0\\2\\1\end{bmatrix} = -3\begin{bmatrix}1\\0\\0\end{bmatrix} + 2\begin{bmatrix}1\\1\\0\end{bmatrix} + 1\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}-3\\2\\1\end{bmatrix}_{\mathscr{B}},$$

whence the desired matrix is

$$\begin{bmatrix}0 & -2 & -3\\1 & 2 & 2\\0 & 0 & 1\end{bmatrix}_{\mathscr{B}}.$$

We will also use the following result.

51 Theorem Let

$$L_1 : \begin{array}{ccc}\mathbb{R}^n & \to & \mathbb{R}^m\\ \overrightarrow{\mathbf{a}} & \mapsto & L_1(\overrightarrow{\mathbf{a}})\end{array}, \quad L_2 : \begin{array}{ccc}\mathbb{R}^m & \to & \mathbb{R}^l\\ \overrightarrow{\mathbf{a}} & \mapsto & L_2(\overrightarrow{\mathbf{a}})\end{array}$$

be linear transformations with matrix representations $A_{L_1} \in \mathbf{M}_{m\times n}(\mathbb{R})$ and $B_{L_2} \in \mathbf{M}_{l\times m}(\mathbb{R})$ respectively. Then the composition map

$$L_2 \circ L_1 : \begin{array}{ccc}\mathbb{R}^n & \to & \mathbb{R}^l\\ \overrightarrow{\mathbf{a}} & \mapsto & (L_2 \circ L_1)(\overrightarrow{\mathbf{a}})\end{array}$$

has as matrix representation the product of matrices $B_{L_2}A_{L_1} \in \mathbf{M}_{l\times n}(\mathbb{R})$.

1.5 Inner Products

52 Definition Let V be a vector space over $\mathbb{R}$. An *inner product* $\bullet$ is a function

$$\bullet : \begin{array}{ccc} V \times V & \to & \mathbb{R} \\ (\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) & \mapsto & \overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}} \end{array}$$

satisfying

❶ $(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{y}})\bullet\overrightarrow{\mathbf{z}} = \overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{z}} + \overrightarrow{\mathbf{y}}\bullet\overrightarrow{\mathbf{z}}$

❷ $(\alpha\overrightarrow{\mathbf{x}})\bullet\overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{x}}\bullet(\alpha\overrightarrow{\mathbf{y}}) = \alpha(\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}}), \alpha \in \mathbb{R}$.

❸ $\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{y}}\bullet\overrightarrow{\mathbf{x}}$

❹ $\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{x}} \geq 0$

❺ $\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{x}} = 0 \Leftrightarrow \overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{0}}$

53 Definition Given a vector space V over $\mathbb{R}$ with inner product $\bullet$, the *norm* $||\overrightarrow{\mathbf{a}}||$ of a vector $\overrightarrow{\mathbf{a}}$ is

$$||\overrightarrow{\mathbf{a}}|| = \sqrt{\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{a}}}.$$

54 Definition Given a vector space V over $\mathbb{R}$ with inner product $\bullet$ and norm $||\cdot||$, the *distance* $d\left(\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}}\right)$ between $\overrightarrow{\mathbf{a}}$ and $\overrightarrow{\mathbf{b}}$ is

$$d\left(\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}}\right) = \left|\left|\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}\right|\right|.$$

55 Example Given $\overrightarrow{\mathbf{a}}$ and $\overrightarrow{\mathbf{b}}$ in $\mathbb{R}^n$, the usual inner product is the *dot product* defined by

$$\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{b}} = \sum_{k=1}^{n} a_k b_k. \tag{1.19}$$

56 Example Let

$$\overrightarrow{\mathbf{a}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \overrightarrow{\mathbf{b}} = \begin{bmatrix} 1 \\ 4 \\ -3 \end{bmatrix}$$

be vectors in $\mathbb{R}^3$. Their dot product is

$$\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{b}} = (1)(1) + (2)(4) + (3)(-3) = 0,$$

their norms are

$$\left|\left|\overrightarrow{\mathbf{a}}\right|\right| = \sqrt{(1)^2 + (2)^2 + (3)^2} = \sqrt{14}, \quad \left|\left|\overrightarrow{\mathbf{b}}\right|\right| = \sqrt{(1)^2 + (4)^2 + (-3)^2} = \sqrt{26},$$

and their distance is

$$d\left(\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}}\right) = \left|\left|\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}\right|\right| = \left|\left|\begin{bmatrix} 0 \\ -2 \\ 6 \end{bmatrix}\right|\right| = \sqrt{(0)^2 + (-2)^2 + (6)^2} = 2\sqrt{10}.$$

We took for granted the fact that the dot product in $\mathbb{R}^n$ does define an inner product. This does require (a very easy) proof.

57 Example Let A, B in $\mathbf{M}_n(\mathbb{R})$. Prove that the operation

$$A\bullet B = \mathrm{tr}\left(B^T A\right)$$

is an inner product. This is the standard inner product for matrices with real number entries.

Solution: Observe that

$$\begin{aligned} (A_1 + \alpha A_2)\bullet B &= \mathrm{tr}\left(B^T(A_1 + \alpha A_2)\right) \\ &= \mathrm{tr}\left(B^T A_1 + B^T(\alpha A_2)\right) \\ &= \mathrm{tr}\left(B^T A_1\right) + \alpha\mathrm{tr}\left(B^T A_2\right) \\ &= A_1\bullet B + \alpha A_2\bullet B, \end{aligned}$$

and so (1) and (2) in definition 52 are verified. Also (3) follows from Theorem 26. To check (4) and (5) we only need to observe that

$$A\bullet A = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2.$$

58 Example Given

$$A = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

their standard inner product is

$$A\bullet B = \text{tr}\left(\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\right) = \text{tr}\left(\begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}\right) = -2,$$

their standard norms are

$$||A|| = \left|\left|\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\right|\right| = \sqrt{(1)^2 + (1)^2 + (0)^2 + (-1)^2} = \sqrt{3},$$

and

$$||B|| = \left|\left|\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right|\right| = \sqrt{(-1)^2 + (0)^2 + (0)^2 + (1)^2} = \sqrt{2}.$$

59 Theorem (Cauchy-Bunyakovsky-Schwarz Inequality) Let V be a vector space over $\mathbb{R}$ having inner product $\bullet$ and corresponding norm $||\cdot||$. Then for any two vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ we have

$$|\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}}| \leq ||\overrightarrow{\mathbf{x}}||||\overrightarrow{\mathbf{y}}||.$$

Proof Since the norm of any vector is non-negative, we have

$$\begin{aligned} ||\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{y}}|| \geq 0 \quad &\iff \quad (\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{y}})\bullet(\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{y}}) \geq 0 \\ &\iff \quad \overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{x}} + 2t\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}} + t^2\overrightarrow{\mathbf{y}}\bullet\overrightarrow{\mathbf{y}} \geq 0 \\ &\iff \quad ||\overrightarrow{\mathbf{x}}||^2 + 2t\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}} + t^2||\overrightarrow{\mathbf{y}}||^2 \geq 0. \end{aligned}$$

This last expression is a quadratic polynomial in t which is always non-negative. As such its discriminant must be non-positive, that is,

$$(2\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}})^2 - 4(||\overrightarrow{\mathbf{x}}||^2)(||\overrightarrow{\mathbf{y}}||^2) \leq 0 \iff |\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}}| \leq ||\overrightarrow{\mathbf{x}}||||\overrightarrow{\mathbf{y}}||,$$

giving the theorem. ❑

60 Corollary (Triangle Inequality) Let V be a vector space over $\mathbb{R}$ having inner product $\bullet$ and corresponding norm $||\cdot||$. Then for any two vectors $\overrightarrow{\mathbf{a}}$ and $\overrightarrow{\mathbf{b}}$ we have

$$\left|\left|\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}}\right|\right| \leq ||\overrightarrow{\mathbf{a}}|| + \left|\left|\overrightarrow{\mathbf{b}}\right|\right|.$$

Proof

$$\begin{aligned} ||\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}}||^2 &= (\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}})\bullet(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}}) \\ &= \overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{a}} + 2\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{b}} + \overrightarrow{\mathbf{b}}\bullet\overrightarrow{\mathbf{b}} \\ &\leq ||\overrightarrow{\mathbf{a}}||^2 + 2||\overrightarrow{\mathbf{a}}||||\overrightarrow{\mathbf{b}}|| + ||\overrightarrow{\mathbf{b}}||^2 \\ &= (||\overrightarrow{\mathbf{a}}|| + ||\overrightarrow{\mathbf{b}}||)^2, \end{aligned}$$

from where the desired result follows. ❑

61 Definition Let $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ be two non-zero vectors in a vector space over the real numbers. Then the angle $\widehat{(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})}$ between them is given by the relation

$$\cos\widehat{(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})} = \frac{\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}}}{||\overrightarrow{\mathbf{x}}||||\overrightarrow{\mathbf{y}}||}.$$

This expression agrees with the geometry in the case of the dot product for $\mathbb{R}^2$ and $\mathbb{R}^3$.

62 Example Let $\overrightarrow{\mathbf{u}}, \overrightarrow{\mathbf{v}}$ be vectors in a vector space V over $\mathbb{R}$ with inner product $\bullet$. Prove the *polarisation identity:*

$$\overrightarrow{\mathbf{u}} \bullet \overrightarrow{\mathbf{v}} = \frac{1}{4}\left(||\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}||^2 - ||\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{v}}||^2\right).$$

Solution: We have

$$\begin{aligned} ||\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}||^2 - ||\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{v}}||^2 &= (\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}})\bullet(\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}) - (\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{v}})\bullet(\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{v}}) \\ &= \overrightarrow{\mathbf{u}}\bullet\overrightarrow{\mathbf{u}} + 2\overrightarrow{\mathbf{u}}\bullet\overrightarrow{\mathbf{v}} + \overrightarrow{\mathbf{v}}\bullet\overrightarrow{\mathbf{v}} - (\overrightarrow{\mathbf{u}}\bullet\overrightarrow{\mathbf{u}} - 2\overrightarrow{\mathbf{u}}\bullet\overrightarrow{\mathbf{v}} + \overrightarrow{\mathbf{v}}\bullet\overrightarrow{\mathbf{v}}) \\ &= 4\overrightarrow{\mathbf{u}}\bullet\overrightarrow{\mathbf{v}}, \end{aligned}$$

giving the result.

63 Example Let $\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}}$ be fixed vectors in $\mathbb{R}^2$. Prove that if

$$\forall\overrightarrow{\mathbf{v}} \in \mathbb{R}^2, \overrightarrow{\mathbf{v}} \bullet \overrightarrow{\mathbf{a}} = \overrightarrow{\mathbf{v}} \bullet \overrightarrow{\mathbf{b}},$$

then $\overrightarrow{\mathbf{a}} = \overrightarrow{\mathbf{b}}$.

Solution: We have $\forall \overrightarrow{\mathbf{v}} \in \mathbb{R}^2, \overrightarrow{\mathbf{v}} \bullet (\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}) = 0$. In particular, choosing $\overrightarrow{\mathbf{v}} = \overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}$, we gather

$$(\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}) \bullet (\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}) = ||\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}}||^2 = 0.$$

But the norm of a vector is 0 if and only if the vector is the $\overrightarrow{\mathbf{0}}$ vector. Therefore $\overrightarrow{\mathbf{a}} - \overrightarrow{\mathbf{b}} = \overrightarrow{\mathbf{0}}$, i.e., $\overrightarrow{\mathbf{a}} = \overrightarrow{\mathbf{b}}$.

The Cauchy-Bunyakovsky-Schwarz (CBS) Inequality applied to the dot product in $\mathbb{R}^n$ gives

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \left(\sum_{k=1}^{n} x_k^2\right)^{1/2} \left(\sum_{k=1}^{n} y_k^2\right)^{1/2}. \tag{1.20}$$

64 Example Assume that $a_k, b_k, c_k, k = 1, \ldots, n$, are positive real numbers. Shew that

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^4 \leq \left(\sum_{k=1}^{n} a_k^4\right) \left(\sum_{k=1}^{n} b_k^4\right) \left(\sum_{k=1}^{n} c_k^2\right)^2.$$

Solution: Using CBS on $\sum_{k=1}^{n} (a_k b_k) c_k$ once we obtain

$$\sum_{k=1}^{n} a_k b_k c_k \leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2}.$$

Using CBS again on $\left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2}$ we obtain

$$\begin{aligned} \sum_{k=1}^{n} a_k b_k c_k &\leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2} \\ &\leq \left(\sum_{k=1}^{n} a_k^4\right)^{1/4} \left(\sum_{k=1}^{n} b_k^4\right)^{1/4} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2}, \end{aligned}$$

which gives the required inequality.

1.6 Cross product in $\mathbb{R}^3$

We now define the standard cross product in $\mathbb{R}^3$ as a product satisfying the following properties.

65 Definition Let $(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}, \overrightarrow{\mathbf{z}}, \alpha) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$. The cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is a closed binary operation satisfying

❶ **Anti-commutativity:** $\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}} = -(\overrightarrow{\mathbf{y}} \times \overrightarrow{\mathbf{x}})$

❷ **Bilinearity:**

$$(\overrightarrow{\mathbf{x}}+\overrightarrow{\mathbf{z}})\times\overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}}+\overrightarrow{\mathbf{z}}\times\overrightarrow{\mathbf{y}} \quad \text{and} \quad \overrightarrow{\mathbf{x}}\times(\overrightarrow{\mathbf{z}}+\overrightarrow{\mathbf{y}}) = \overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{z}}+\overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}}$$

❸ **Scalar homogeneity:** $(\alpha\overrightarrow{\mathbf{x}}) \times \overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{x}} \times (\alpha\overrightarrow{\mathbf{y}}) = \alpha(\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}})$

❹ $\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{0}}$

❺ **Right-hand Rule:**

$$\overrightarrow{\mathbf{i}} \times \overrightarrow{\mathbf{j}} = \overrightarrow{\mathbf{k}}, \quad \overrightarrow{\mathbf{j}} \times \overrightarrow{\mathbf{k}} = \overrightarrow{\mathbf{i}}, \quad \overrightarrow{\mathbf{k}} \times \overrightarrow{\mathbf{i}} = \overrightarrow{\mathbf{j}}.$$

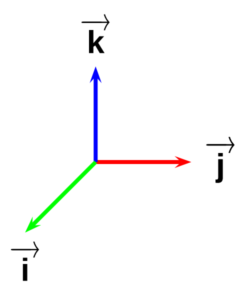

Figure 1.3: Right-handed system.

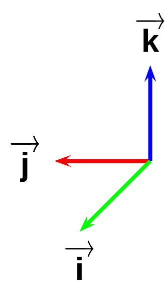

Figure 1.4: Left-handed system.

To study points in space we must first agree on the orientation that we will give our co-ördinate system. We will use, unless otherwise noted, a right-handed orientation, as in figure 1.3.

66 Example Find

$$\begin{bmatrix}1\\0\\-3\end{bmatrix}\times\begin{bmatrix}0\\1\\2\end{bmatrix}.$$

Solution: We have

$$\begin{aligned}(\overrightarrow{\mathbf{i}}-3\overrightarrow{\mathbf{k}})\times(\overrightarrow{\mathbf{j}}+2\overrightarrow{\mathbf{k}}) &= \overrightarrow{\mathbf{i}}\times\overrightarrow{\mathbf{j}}+2\overrightarrow{\mathbf{i}}\times\overrightarrow{\mathbf{k}}-3\overrightarrow{\mathbf{k}}\times\overrightarrow{\mathbf{j}}-6\overrightarrow{\mathbf{k}}\times\overrightarrow{\mathbf{k}}\\ &= \overrightarrow{\mathbf{k}}-2\overrightarrow{\mathbf{j}}+3\overrightarrow{\mathbf{i}}+6\overrightarrow{\mathbf{0}}\\ &= 3\overrightarrow{\mathbf{i}}-2\overrightarrow{\mathbf{j}}+\overrightarrow{\mathbf{k}}.\end{aligned}$$

Hence

$$\begin{bmatrix}1\\0\\-3\end{bmatrix}\times\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}3\\-2\\1\end{bmatrix}.$$

Operating as in example 66 we obtain

67 Theorem Let $\overrightarrow{\mathbf{x}}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ and $\overrightarrow{\mathbf{y}}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$ be vectors in $\mathbb{R}^3$. Then

$$\overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}}=(x_2y_3-x_3y_2)\overrightarrow{\mathbf{i}}+(x_3y_1-x_1y_3)\overrightarrow{\mathbf{j}}+(x_1y_2-x_2y_1)\overrightarrow{\mathbf{k}}.$$

68 Theorem $\overrightarrow{\mathbf{x}}\perp(\overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}})$ and $\overrightarrow{\mathbf{y}}\perp(\overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}})$.

Proof We will only check the first assertion, the second verification is analogous.

$$\begin{aligned}\overrightarrow{\mathbf{x}}\bullet(\overrightarrow{\mathbf{x}}\times\overrightarrow{\mathbf{y}}) &= (x_1\overrightarrow{\mathbf{i}}+x_2\overrightarrow{\mathbf{j}}+x_3\overrightarrow{\mathbf{k}})\bullet((x_2y_3-x_3y_2)\overrightarrow{\mathbf{i}}\\ &\quad+(x_3y_1-x_1y_3)\overrightarrow{\mathbf{j}}+(x_1y_2-x_2y_1)\overrightarrow{\mathbf{k}})\\ &= x_1x_2y_3-x_1x_3y_2+x_2x_3y_1-x_2x_1y_3+x_3x_1y_2-x_3x_2y_1\\ &= 0,\end{aligned}$$

completing the proof. ❑

69 Example Let $a\in\mathbb{R}$. Find a vector of unit length simultaneously perpendicular to $\overrightarrow{\mathbf{v}}=\begin{bmatrix}0\\-a\\a\end{bmatrix}$ and $\overrightarrow{\mathbf{w}}=\begin{bmatrix}1\\a\\0\end{bmatrix}$.

Solution: Either of $\dfrac{\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}}}{||\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}}||}$ or $-\dfrac{\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}}}{||\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}}||}$ will do. Now

$$\begin{aligned}
\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}} &= (-a\overrightarrow{\mathbf{j}} + a\overrightarrow{\mathbf{k}}) \times (\overrightarrow{\mathbf{i}} + a\overrightarrow{\mathbf{j}}) \\
&= -a(\overrightarrow{\mathbf{j}} \times \overrightarrow{\mathbf{i}}) - a^2(\overrightarrow{\mathbf{j}} \times \overrightarrow{\mathbf{j}}) + a(\overrightarrow{\mathbf{k}} \times \overrightarrow{\mathbf{i}}) + a^2(\overrightarrow{\mathbf{k}} \times \overrightarrow{\mathbf{j}}) \\
&= a\overrightarrow{\mathbf{k}} + a\overrightarrow{\mathbf{j}} - a^2\overrightarrow{\mathbf{i}} \\
&= \begin{pmatrix} -a^2 \\ a \\ a \end{pmatrix},
\end{aligned}$$

and $||\overrightarrow{\mathbf{v}} \times \overrightarrow{\mathbf{w}}|| = \sqrt{a^4 + a^2 + a^2} = \sqrt{2a^2 + a^4}$. Hence we may take either

$$\frac{1}{\sqrt{2a^2 + a^4}} \begin{pmatrix} -a^2 \\ a \\ a \end{pmatrix}$$

or

$$-\frac{1}{\sqrt{2a^2 + a^4}} \begin{pmatrix} -a^2 \\ a \\ a \end{pmatrix}.$$

The cross product of vectors in $\mathbb{R}^3$ *is not associative, since*

$$\overrightarrow{\mathbf{i}} \times (\overrightarrow{\mathbf{i}} \times \overrightarrow{\mathbf{j}}) = \overrightarrow{\mathbf{i}} \times \overrightarrow{\mathbf{k}} = -\overrightarrow{\mathbf{j}}$$

but

$$(\overrightarrow{\mathbf{i}} \times \overrightarrow{\mathbf{i}}) \times \overrightarrow{\mathbf{j}} = \overrightarrow{\mathbf{0}} \times \overrightarrow{\mathbf{j}} = \overrightarrow{\mathbf{0}}.$$

We have, however, the following theorem.

70 Theorem

$$\overrightarrow{\mathbf{a}} \times (\overrightarrow{\mathbf{b}} \times \overrightarrow{\mathbf{c}}) = (\overrightarrow{\mathbf{a}} \bullet \overrightarrow{\mathbf{c}})\overrightarrow{\mathbf{b}} - (\overrightarrow{\mathbf{a}} \bullet \overrightarrow{\mathbf{b}})\overrightarrow{\mathbf{c}}.$$

Proof

$$\begin{aligned}
\overrightarrow{\mathbf{a}} \times (\overrightarrow{\mathbf{b}} \times \overrightarrow{\mathbf{c}}) &= (a_1\overrightarrow{\mathbf{i}} + a_2\overrightarrow{\mathbf{j}} + a_3\overrightarrow{\mathbf{k}}) \times ((b_2c_3 - b_3c_2)\overrightarrow{\mathbf{i}} + \\
&\quad +(b_3c_1 - b_1c_3)\overrightarrow{\mathbf{j}} + (b_1c_2 - b_2c_1)\overrightarrow{\mathbf{k}}) \\
&= a_1(b_3c_1 - b_1c_3)\overrightarrow{\mathbf{k}} - a_1(b_1c_2 - b_2c_1)\overrightarrow{\mathbf{j}} - a_2(b_2c_3 - b_3c_2)\overrightarrow{\mathbf{k}} \\
&\quad +a_2(b_1c_2 - b_2c_1)\overrightarrow{\mathbf{i}} + a_3(b_2c_3 - b_3c_2)\overrightarrow{\mathbf{j}} - a_3(b_3c_1 - b_1c_3)\overrightarrow{\mathbf{i}} \\
&= (a_1c_1 + a_2c_2 + a_3c_3)(b_1\overrightarrow{\mathbf{i}} + b_2\overrightarrow{\mathbf{j}} + b_3\overrightarrow{\mathbf{i}}) + \\
&\quad (-a_1b_1 - a_2b_2 - a_3b_3)(c_1\overrightarrow{\mathbf{i}} + c_2\overrightarrow{\mathbf{j}} + c_3\overrightarrow{\mathbf{i}}) \\
&= = (\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{c}})\overrightarrow{\mathbf{b}} - (\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{b}})\overrightarrow{\mathbf{c}},
\end{aligned}$$

completing the proof. ❑

Permuting the vectors in the above theorem and adding, we obtain Jacobi's Identity:

$$\overrightarrow{\mathbf{a}} \times (\overrightarrow{\mathbf{b}} \times \overrightarrow{\mathbf{c}}) + \overrightarrow{\mathbf{b}} \times (\overrightarrow{\mathbf{c}} \times \overrightarrow{\mathbf{a}}) + \overrightarrow{\mathbf{c}} \times (\overrightarrow{\mathbf{a}} \times \overrightarrow{\mathbf{b}}) = \overrightarrow{\mathbf{0}}.$$

71 Theorem Let $(\widehat{\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}}) \in [0;\pi]$ be the convex angle between two vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$. Then

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}}|| = ||\overrightarrow{\mathbf{x}}||||\overrightarrow{\mathbf{y}}|| \sin(\widehat{\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}}).$$

Proof We have

$$\begin{aligned}
||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}}||^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&= x_2^2y_3^2 - 2x_2y_3x_3y_2 + x_3^2y_2^2 + x_3^2y_1^2 - 2x_3y_1x_1y_3 + \\
&\quad +x_1^2y_3^2 + x_1^2y_2^2 - 2x_1y_2x_2y_1 + x_2^2y_1^2 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= ||\overrightarrow{\mathbf{x}}||^2||\overrightarrow{\mathbf{y}}||^2 - (\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{y}})^2 \\
&= ||\overrightarrow{\mathbf{x}}||^2||\overrightarrow{\mathbf{y}}||^2 - ||\overrightarrow{\mathbf{x}}||^2||\overrightarrow{\mathbf{y}}||^2\cos^2(\widehat{\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}}) \\
&= ||\overrightarrow{\mathbf{x}}||^2||\overrightarrow{\mathbf{y}}||^2\sin^2(\widehat{\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}}),
\end{aligned}$$

whence the theorem follows. ❑

The following corollaries are now obvious.

72 Corollary Two non-zero vectors $\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}$ satisfy $\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{0}}$ if and only if they are parallel.

73 Corollary (Lagrange's Identity)

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{y}}||^2 = ||\mathbf{x}||^2||\mathbf{y}||^2 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}})^2.$$

74 Example Let $\overrightarrow{\mathbf{x}} \in \mathbb{R}^3, ||\mathbf{x}|| = 1$. Find

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{i}}||^2 + ||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{j}}||^2 + ||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{k}}||^2.$$

Solution: By Lagrange's Identity,

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{i}}||^2 = \left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2\left|\left|\overrightarrow{\mathbf{i}}\right|\right|^2 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{i}})^2 = 1 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{i}})^2,$$

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{k}}||^2 = \left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2\left|\left|\overrightarrow{\mathbf{j}}\right|\right|^2 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{j}})^2 = 1 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{j}})^2,$$

$$||\overrightarrow{\mathbf{x}} \times \overrightarrow{\mathbf{j}}||^2 = \left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2\left|\left|\overrightarrow{\mathbf{k}}\right|\right|^2 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}})^2 = 1 - (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}})^2,$$

and since $(\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{i}})^2 + (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{j}})^2 + (\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}})^2 = \left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2 = 1$, the desired sum equals $3 - 1 = 2$.

1.7 Lines and Planes in $\mathbb{R}^3$

75 Definition Let $\mathbf{a} \in \mathbb{R}^3$ and $\overrightarrow{\mathbf{v}} \in \mathbb{R}^3 \setminus \{\overrightarrow{\mathbf{0}}\}$. Put $\overrightarrow{\mathbf{r}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. The line passing through $\mathbf{a}$ in the direction of $\overrightarrow{\mathbf{v}}$ is the set

$$\{\overrightarrow{\mathbf{r}} : \overrightarrow{\mathbf{r}} = \overrightarrow{\mathbf{a}} + t\overrightarrow{\mathbf{v}},\ t \in \mathbb{R}\}.$$

76 Example Find the equation of the line passing through $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in the direction of $\begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}$.

Solution: The desired equation is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}.$$

77 Example Find the equation of the line passing through $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$.

Solution: The line follows the direction

$$\begin{bmatrix} 1-(-2) \\ 2-(-1) \\ 3-0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

The desired equation is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

? *Why does*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

represent the same line as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}?$$

STOP *Given two lines in space, one of the following three situations might arise: (i) the lines intersect at a point, (ii) the lines are parallel, (iii) the lines are skew (one over the other, without intersecting).*

78 Definition The set of points $\mathbf{x} \in \mathbb{R}^n$ satisfying an equation of the form

$$\overrightarrow{\mathbf{a}} \bullet \overrightarrow{\mathbf{x}} = c$$

($\overrightarrow{\mathbf{a}} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ fixed) is called a *hyperplane* in $\mathbb{R}^n$.

In three-dimensions a hyperplane is simply a plane. To obtain the equation of a plane in $\mathbb{R}^3$ simply notice that if the vector $\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$ is normal (perpendicular) to the plane passing through the point $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ then form any other point $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ on the plane, the vector $\begin{bmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{bmatrix}$ will be perpendicular to the vector $\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$ and so

$$\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \bullet \begin{bmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{bmatrix} = 0$$

or equivalently,

$$n_1(x - a_1) + n_2(y - a_2) + n_3(z - a_3) = 0, \tag{1.21}$$

gives the equation of a plane in $\mathbb{R}^3$.

79 Example The equation of the plane passing through the point $(1, -1, 2)$ and normal to the vector $\begin{bmatrix} -3 \\ 2 \\ 4 \end{bmatrix}$ is

$$-3(x - 1) + 2(y + 1) + 4(z - 2) = 0.$$

STOP *From* $n_1(x - a_1) + n_2(y - a_2) + n_3(z - a_3) = 0$ *we obtain that the equation of a plane is of the form*

$$n_1x + n_2y + n_3z = d,$$

where $\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$ *is perpendicular to the plane, and* $d \in \mathbb{R}$.

80 Example Find the equation of plane containing the point $(1, 1, 1)$ and perpendicular to the line $x = 1 + t, y = -2t, z = 1 - t$.

Solution: The vectorial form of the equation of the line is

$$\overrightarrow{\mathbf{r}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

Since the line follows the direction of $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$, this means that $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ is normal to the plane, and thus the equation of the desired plane is

$$(x - 1) - 2(y - 1) - (z - 1) = 0.$$

81 Example Find the equation of plane containing the point $(1, -1, -1)$ and containing the line $x = 2y = 3z$.

Solution: Observe that $(0, 0, 0)$ (as $0 = 2(0) = 3(0)$) is on the line, and hence on the plane. Thus the vector

$$\begin{bmatrix} 1 - 0 \\ -1 - 0 \\ -1 - 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

lies on the plane. Now, if $x = 2y = 3z = t$, then $x = t, y = t/2, z = t/3$. Hence, the vectorial form of the equation of the line is

$$\overrightarrow{\mathbf{r}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix} = t \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix}.$$

This means that $\begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix}$ also lies on the plane, and thus

$$\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/6 \\ -4/3 \\ 3/2 \end{bmatrix}$$

is normal to the plane. The desired equation is thus

$$\frac{1}{6}x - \frac{4}{3}y + \frac{3}{2}z = 0.$$

Given three planes in space, they may (i) be parallel (which allows for some of them to coincide), (ii) two may be parallel and the third intersect each of the other two at a line, (iii) intersect at a line, (iv) intersect at a point.

82 Example Find the equation of the plane passing through the points $(a, 0, a)$, $(-a, 1, 0)$, and $(0, 1, 2a)$ in $\mathbb{R}^3$.

The vectors

$$\begin{bmatrix} a-(-a) \\ 0-1 \\ a-0 \end{bmatrix} = \begin{bmatrix} 2a \\ -1 \\ a \end{bmatrix}$$

and

$$\begin{bmatrix} 0-(-a) \\ 1-1 \\ 2a-0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ 2a \end{bmatrix}$$

lie on the plane. A vector normal to the plane is

$$\begin{bmatrix} 2a \\ -1 \\ a \end{bmatrix} \wedge \begin{bmatrix} a \\ 0 \\ 2a \end{bmatrix} = \begin{bmatrix} -2a \\ -3a^2 \\ a \end{bmatrix}.$$

The equation of the plane is thus given by

$$\begin{bmatrix} -2a \\ -3a^2 \\ a \end{bmatrix} \bullet \begin{bmatrix} x-a \\ y-0 \\ z-a \end{bmatrix} = 0,$$

that is,

$$2ax + 3a^2y - az = a^2.$$

83 Example Find the equation of the line perpendicular to the plane $ax + a^2y + a^3z = 0, \quad a \neq 0$ and passing through the point $(0, 0, 1)$.

Solution: A vector normal to the plane is $\begin{bmatrix} a \\ a^2 \\ a^2 \end{bmatrix}$. The line sought has the same direction as this vector, thus the equation of the line is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} a \\ a^2 \\ a^2 \end{bmatrix}, \quad t \in \mathbb{R}.$$

84 Example Find the equation of the plane perpendicular to the line $ax = by = cz, \quad abc \neq 0$ and passing through the point $(1, 1, 1)$ in $\mathbb{R}^3$.

Solution: Put $ax = by = cz = t$, so $x = t/a; y = t/b; z = t/c$. The parametric equation of the line is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix}, \quad t \in \mathbb{R}.$$

Thus the vector $\begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix}$ is perpendicular to the plane. Therefore, the equation of the plane is

$$\begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix} \bullet \begin{bmatrix} x - 1 \\ y - 1 \\ z - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

or

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}.$$

We may also write this as

$$bcx + cay + abz = ab + bc + ca.$$

85 Example (Putnam Exam 1980) Let S be the solid in three-dimensional space consisting of all points (x,y,z) satisfying the following system of six conditions:

$$x \geq 0, \quad y \geq 0, \quad z \geq 0,$$

$$x+y+z \leq 11,$$

$$2x+4y+3z \leq 36,$$

$$2x+3z \leq 24.$$

Determine the number of vertices and the number of edges of S.

Solution: There are 7 vertices ($V_0 = (0,0,0), V_1 = (11,0,0), V_2 = (0,9,0), V_3 = (0,0,8)$, $V_4 = (0,3,8)$, $V_5 = (9,0,2)$, $V_6 = (4,7,0)$) and 11 edges (V_0V_1, V_0V_2, V_0V_3, V_1V_5, V_1V_6, V_2V_4, V_3V_4, V_3V_5, V_4V_5, and V_4V_6).

1.8 Topology of $\mathbb{R}^n$

86 Definition Let $\overrightarrow{\mathbf{a}} \in \mathbb{R}^n$ and let $\varepsilon > 0$. An *open ball* centred at $\overrightarrow{\mathbf{a}}$ of radius ε is the set

$$B_\varepsilon(\overrightarrow{\mathbf{a}}) = \{\mathbf{x} \in \mathbb{R}^n : d(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}}) < \varepsilon\}.$$

87 Example An open ball in $\mathbb{R}$ is an open interval, an open ball in $\mathbb{R}^2$ is an open disk and an open ball in $\mathbb{R}^3$ is an open sphere.

88 Definition A set $S \subseteq \mathbb{R}^n$ is said to be *open* if for every point belonging to it we can surround the point by a sufficiently small open ball so that this balls lies completely within the set. That is, $\forall \overrightarrow{\mathbf{a}} \in S \;\; \exists \varepsilon > 0$ such that $B_\varepsilon(\overrightarrow{\mathbf{a}}) \subseteq S$.

On the real line, an open ball is an interval of length $2r$ centred at a point p, on the plane, an open ball is an open disk centred about $\overrightarrow{\mathbf{p}}$, in 3-dimensional space, an open ball is a sphere excluding its boundary and centred at $\overrightarrow{\mathbf{p}}$.

89 Definition A set $\mathscr{O} \subseteq \mathbb{R}^n$ is said to be *open* in $\mathbb{R}^n$ if $\forall \overrightarrow{\mathbf{x}} \in \mathscr{O} \; \exists r > 0$ such that $B_r(\overrightarrow{\mathbf{x}}) \subseteq \mathscr{O}$.

That is, a set is open if for all its elements, there exist open balls centred at the elements and totally contained in the set.

90 Example The open interval $]-1;1[$ is open in $\mathbb{R}$. The interval $]-1;1]$ is not open, however, as no interval centred at 1 is totally contained in $]-1;1]$.

91 Example The region $]-1;1[\times]0;+\infty[$ is open in $\mathbb{R}^2$.

92 Example The ellipsoidal region $\{(x,y) \in \mathbb{R}^2 : x^2+4y^2 < 4\}$ is open in $\mathbb{R}^2$.

The reader will recognise that open boxes, open ellipsoids and their unions and finite intersections are open sets in $\mathbb{R}^n$.

93 Definition A set $\mathscr{F} \subseteq \mathbb{R}^n$ is said to be *closed* in $\mathbb{R}^n$ if its complement $\mathbb{R}^n \setminus \mathscr{F}$ is open.

94 Example The closed interval $[-1;1]$ is closed in $\mathbb{R}$, as its complement, $\mathbb{R} \setminus [-1;1] =]-\infty;-1[\cup]1;+\infty[$ is open in $\mathbb{R}$. The interval $]-1;1]$ is neither open nor closed, however.

95 Example The region $[-1;1] \times [0;+\infty[\times[0;2]$ is closed in $\mathbb{R}^3$.

96 Example (Putnam Exam 1969) Let $p(x,y)$ be a polynomial with real coefficients in the real variables x and y, defined over the entire plane $\mathbb{R}^2$. What are the possibilities for the image (range) of $p(x,y)$?

Solution: Since polynomials are continuous functions and the image of a connected set is connected for a continuous function, the image must be an interval of some sort. If the image were a finite interval, then $f(x,kx)$ would be bounded for every constant k, and so the image would just be the point $f(0,0)$. The possibilities are thus (i) a single point (take for example, $p(x,y)=0$), (ii) a semi-infinite interval with an endpoint (take for example $p(x,y)=x^2$ whose image is $[0;+\infty[$), (iii) a semi-infinite interval with no endpoint (take for example $p(x,y)=(xy-1)^2+x^2$ whose image is $]0;+\infty[$), (iv) all real numbers (take for example $p(x,y)=x$).

97 Example (Putnam Exam 1984) Let A be a solid $a \times b \times c$ rectangular brick in three dimensions, where $a > 0, b > 0, c > 0$. Let B be the set of all points which are at distance at most 1 from some point of A (in particular, $A \subset B$). Express the volume of B as a polynomial in a, b, c.

Solution: The set B can be decomposed into the following subsets:

❶ The set A itself, of volume abc.

❷ Two $a \times b \times 1$ bricks, two $b \times c \times 1$ bricks, and two $c \times a \times 1$ bricks,

❸ Four quarter-cylinders of length a and radius 1, four quarter-cylinders of length b and radius 1, and four quarter-cylinders of length c and radius 1,

❹ Eight eighth-of-spheres of radius 1.

Thus the required formula for the volume is

$$abc + 2(ab + bc + ca) + \pi(a + b + c) + \frac{4\pi}{3}.$$

1.9 Quadratic Forms

98 Definition A matrix $A \in \mathbf{M}_n(\mathbb{R})$ is called *positive definite* if $\forall \overrightarrow{\mathbf{x}} \in \mathbb{R}^n \setminus \{\overrightarrow{\mathbf{0}}\}$,

$$\overrightarrow{\mathbf{x}}^T A \overrightarrow{\mathbf{x}} > 0.$$

It is *positive semi-definite* if $\forall \overrightarrow{\mathbf{x}} \in \mathbb{R}^n$, $\overrightarrow{\mathbf{x}}^T A \overrightarrow{\mathbf{x}} \geq 0$. It is *negative definite* if $\forall \overrightarrow{\mathbf{x}} \in \mathbb{R}^n \setminus \{\overrightarrow{\mathbf{0}}\}$,

$$\overrightarrow{\mathbf{x}}^T A \overrightarrow{\mathbf{x}} < 0.$$

It is *negative semi-definite* if $\forall \overrightarrow{\mathbf{x}} \in \mathbb{R}^n$, $\overrightarrow{\mathbf{x}}^T A \overrightarrow{\mathbf{x}} \leq 0$. A matrix is *indefinite* if it is neither positive nor negative definite (semi-definite).

99 Example The matrix

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

is positive definite, since

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 + 2y^2 + 3z^2 > 0,$$

for $(x, y, z) \neq (0, 0, 0)$, being a sum of squares.

100 Example The matrix

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

is indefinite, since

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 > 0, \quad \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = -2 < 0.$$

101 Example The matrix

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is negative semi-definite, since

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0, \quad \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -(x^2+2y^2) < 0.$$

102 Definition The n *principal minors* of a matrix $A = [a_{ij}]$ are

$$\Delta_1 = a_{11},$$

$$\Delta_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

$$\Delta_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\vdots$$

$$\Delta_k = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kk} \end{bmatrix},$$

$$\vdots$$

$$\Delta_n = \det A.$$

103 Theorem (Sylvester's Criterion) A matrix $A \in \mathbf{M}_n(\mathbb{R})$ is positive definite if and only if all its principal minors are positive. It is positive semi-definite if all its principal minors are non-negative. It is negative definite if all of its odd order minors are negative and all of its even order minors are positive. It is negative semi-definite if all of its odd order minors are non-positive and all of its even order minors are non-negative. It is indefinite if none of the above cases occur.

1.10 Quadratic Surfaces

104 Definition A *quadric (or quadratic) surface in* $\mathbb{R}^3$ is a surface whose cartesian equation is a polynomial of degree 2. Thus the general form of a quadric surface is

$$Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2 + 2Gx + 2Hy + 2Iz + J = 0$$

where all the coefficients are real. This can be written in the matrix form

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} A & B & C & G \\ B & D & E & H \\ C & E & F & I \\ G & H & I & J \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0, \qquad (1.22)$$

where we have identified a 1×1 matrix with a real number.

The square matrix of coefficients is symmetric, and as such all its eigenvalues are real by virtue of Theorem 40.

105 Example Write the quadric surface

$$2x^2 - 3y^2 + 10xy + 5 + (z-2)^2 = 0$$

in matrix form.

Solution: First observe that

$$2x^2 - 3y^2 + 10xy + (z-2)^2 = 2x^2 - 3y^2 + 10xy + 5 + z^2 - 4z + 9.$$

The desired form is

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 & 0 & 0 \\ 5 & -3 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0.$$

In the case when there are no cross terms, things greatly simplify, and, up to permutations of the letters, we have the following result.

106 Theorem If

$$p\frac{x^2}{a^2} + q\frac{y^2}{b^2} + r\frac{z^2}{c^2} = d,$$

then this equation represents

- an ellipsoid: $p = q = r = d = 1$, a, b, c are the lengths of the semi axes.
- a single-sheet hyperboloid: $p = q = d = 1$, $r = -1$.
- a double-sheet hyperboloid: $r = d = 1$, $p = q = -1$.
- Cone: $p = q = 1$, $r = -1$, $d = 0$.

If

$$p\frac{x^2}{a^2} + q\frac{y^2}{b^2} + r\frac{z}{c^2} = d,$$

then this equation represents

- an elliptic paraboloid: $p = q = 1$, $r = -1$, $d = 0$.
- a hyperbolic paraboloid: $p = r = -1$, $q = 1$, $d = 0$.

- an elliptic cylinder: $p = q = -1$, $r = d = 0$.
- a hyperbolic cylinder: $p = d = 1$, $q = -1$, $r = 0$.
- a pair of planes: $p = 1$, $q = -1$, $d = 0$.

If

$$py^2 + qx = d$$

then this equation represents

- a parabolic cylinder: $p, q > 0$.
- a pair of parallel planes: $d > 0$, $q = 0$, $p \neq 0$.
- two coinciding planes: $p \neq 0$, $q = d = 0$.

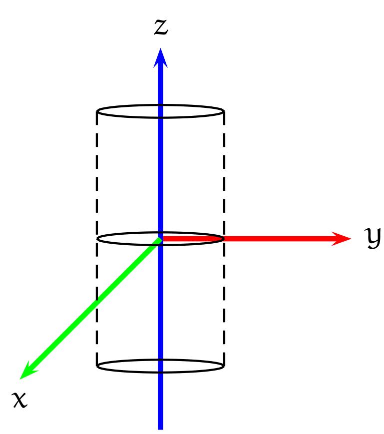

Figure 1.5: Example 107.

107 Example Demonstrate that the surface in $\mathbb{R}^3$

$$S : 4x^2 + y^2 - 4 = 0$$

is a cylinder and draw its graph.

Solution: The variable z is missing. Notice that on the plane, $4x^2+y^2-4=0$ defines an ellipse. Thus the surface is an elliptic cylinder, with the z-axis as directrix. Its graph appears in figure 1.5.

108 Example The surface in $\mathbb{R}^3$ given by

$$S : x^2 - z^2 = 1$$

is a cylinder, as the y-variable is missing. It is a cylindrical hyperboloid.

109 Example The surface in $\mathbb{R}^3$ given by

$$S : y - .05z^2 = 1$$

is a cylinder, as the x-variable is missing. It is a cylindrical paraboloid.

110 Example (Putnam Exam 1970) Determine, with proof, the radius of the largest circle which can lie on the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, a > b > c > 0.$$

Solution: The largest circle has radius b. Parallel cross sections of the ellipsoid are similar ellipses, hence we may increase the size of these by moving towards the centre of the ellipse. Every plane through $(0, 0, 0)$ which makes a circular cross section must intersect the yz-plane, and the diameter of any such cross section must be a diameter of the ellipse $x = 0, \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Therefore, the radius of the circle is at most b. Arguing similarly on the xy-plane shews that the radius of the circle is at least b. To shew that circular cross section of radius b actually exist, one may verify that the two planes given by $a^2(b^2-c^2)z^2 = c^2(a^2-b^2)x^2$ give circular cross sections of radius b.

1.11 Canonical Surfaces in $\mathbb{R}^3$

111 Definition A surface S consisting of all lines parallel to a given line Δ and passing through a given curve Γ is called a *cylinder*. The line Δ is called the *directrix* of the cylinder.

To recognise whether a given surface is a cylinder we look at its Cartesian equation. If it is of the form $f(A, B) = 0$*, where* A, B *are secant planes, then the curve is a cylinder. Under these conditions, the lines generating* S *will be parallel to the line of equation* $A = 0, B = 0$*. In practice, if one of the variables* x, y*, or* z *is missing, then the surface is a cylinder, whose directrix will be the axis of the missing coordinate.*

112 Example Demonstrate that the surface in $\mathbb{R}^3$

$$S : e^{x^2+y^2+z^2} - (x+z)e^{-2xz} = 0,$$

implicitly defined, is a cylinder.

Solution: The planes $A : x + z = 0$ and $B : y = 0$ are secant. The surface has equation of the form $f(A, B) = e^{A^2+B^2} - A = 0$, and it is thus a cylinder. The directrix has direction $\overrightarrow{\mathbf{i}} - \overrightarrow{\mathbf{k}}$.

113 Example Shew that the surface S in $\mathbb{R}^3$ given implicitly by the equation

$$\frac{1}{x-y} + \frac{1}{y-z} + \frac{1}{z-x} = 1$$

is a cylinder and find the direction of its directrix.

Solution: Considering the planes $A : x - y = 0, B : y - z = 0$, the equation takes the form

$$f(A, B) = \frac{1}{A} + \frac{1}{B} - \frac{1}{A+B} - 1 = 0,$$

thus the equation represents a cylinder. To find its directrix, we find the intersection of the planes $x = y$ and $y = z$. This gives $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. The direction vector is thus $\overrightarrow{\mathbf{i}} + \overrightarrow{\mathbf{j}} + \overrightarrow{\mathbf{k}}$.

114 Definition Given a point $\Omega \in \mathbb{R}^3$ (called the *apex*) and a curve Γ (called the generating curve), the surface S obtained by drawing rays from Ω and passing through Γ is called a *cone*.

In practice, if the Cartesian equation of a surface can be put into the form $f(\frac{A}{C}, \frac{B}{C}) = 0$, *where* A, B, C, *are planes secant at exactly one point, then the surface is a cone, and its apex is given by* $A = 0, B = 0, C = 0$.

115 Example Demonstrate that the surface in $\mathbb{R}^3$ given implicitly by

$$z^2 - xy = 2z - 1$$

is a cone

Solution: After rearranging, we obtain

$$(z-1)^2 - xy = 0,$$

or

$$-\frac{x}{z-1}\frac{y}{z-1} + 1 = 0.$$

Considering the planes

$$A : x = 0,\ B : y = 0,\ \ C : z = 1,$$

we see that our surface is a cone, with apex at $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

116 Example The surface in $\mathbb{R}^3$ implicitly by

$$z^2 = x^2 + y^2$$

is a cone, as its equation can be put in the form $(\frac{x}{z})^2 + (\frac{y}{z})^2 - 1 = 0$. Considering the planes $x = 0, y = 0, z = 0$, the apex is located at $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

117 Definition A surface S obtained by making a curve Γ turn around a line Δ is called a *surface of revolution.* We then say that Δ is the axis of revolution. The intersection of S with a half-plane bounded by Δ is called a *meridian*.

If the Cartesian equation of S can be put in the form $f(A, \Sigma) = 0$, where A is a plane and Σ is a sphere, then the surface is of revolution. The axis of S is the line passing through the centre of Σ and perpendicular to the plane A.

118 Example Shew that the surface S in $\mathbb{R}^3$ implicitly defined as

$$xy + yz + zx + x + y + z + 1 = 0$$

is of revolution and find its axis.

Solution: Rearranging,

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

so we may take $A : x + y + z = 0, \Sigma : x^2 + y^2 + z^2 = 0$ as our plane and sphere. The axis of revolution is then in the direction of $\overrightarrow{\mathbf{i}} + \overrightarrow{\mathbf{j}} + \overrightarrow{\mathbf{k}}$.

119 Example Shew that the surface in $\mathbb{R}^3$ implicitly defined by

$$x^4 + y^4 + z^4 - 4xyz(x + y + z) = 1$$

is a surface of revolution, and find its axis of revolution.

Solution: Rearranging,

$$(x^2 + y^2 + z^2)^2 - \frac{1}{2}((x + y + z)^2 - (x^2 + y^2 + z^2)) - 1 = 0,$$

and so we may take $A : x + y + z = 0, \Sigma : x^2 + y^2 + z^2 = 0$, shewing that the surface is of revolution. Its axis is the line in the direction $\overrightarrow{\mathbf{i}} + \overrightarrow{\mathbf{j}} + \overrightarrow{\mathbf{k}}$.

120 Example Find the equation of the surface of revolution generated by revolving the hyperbola $x^2 - 4z^2 = 1$ about the z-axis.

Let (x, y, z) be a point on S. If this point were on the xz plane, it would be on the hyperbola, and its distance to the axis of rotation would be $|x| = \sqrt{1 + 4z^2}$. Anywhere else, the distance of (x, y, z) to the axis of rotation is the same as the distance of (x, y, z) to $(0, 0, z)$, that is $\sqrt{x^2 + y^2}$. We must have

$$\sqrt{x^2 + y^2} = \sqrt{1 + 4z^2},$$

which is to say

$$x^2 + y^2 - 4z^2 = 1.$$

121 Example Find the equation of the surface of revolution generated by revolving the line $3x + 4y = 1$ about the y-axis.

Solution: Let (x, y, z) be a point on S. If this point were on the xy plane, it would be on the line, and its distance to the axis of rotation would be $|x| = \frac{1}{3}|1 - 4y|$. Anywhere else, the distance of (x, y, z) to the axis of rotation is the same as the distance of (x, y, z) to $(0, y, 0)$, that is $\sqrt{x^2 + z^2}$. We must have

$$\sqrt{x^2 + z^2} = \frac{1}{3}|1 - 4y|,$$

which is to say

$$9x^2 + 9z^2 - 16y^2 + 8y - 1 = 0.$$

122 Example Find the equation of the surface of revolution S generated by revolving the ellipse $4x^2 + z^2 = 1$ about the z-axis.

Solution: Let (x, y, z) be a point on S. If this point were on the xz plane, it would be on the ellipse, and its distance to the axis of rotation would be $|x| = \frac{1}{2}\sqrt{1 - z^2}$. Anywhere else, the distance from (x, y, z) to the z-axis is the distance of this point to the point $(0, 0, z)$: $\sqrt{x^2 + y^2}$. This distance is the same as the length of the segment on the xz-plane going from the z-axis. We thus have

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

or

$$4x^2 + 4y^2 + z^2 = 1.$$

123 Example The circle $(y-a)^2+z^2=r^2, (a,r)\in(\mathbb{R}^*)^2$ on the yz plane is revolved around the z-axis, forming a torus T. Find the equation of this torus.

Solution: Let (x,y,z) be a point on T. If this point were on the yz plane, it would be on the circle, and the of the distance to the axis of rotation would be $y = a+\text{sgn}(y-a)\sqrt{r^2-z^2}$, where $\text{sgn}(t)$ (with $\text{sgn}(t)=-1$ if $t<0$, $\text{sgn}(t)=1$ if $t>0$, and $\text{sgn}(0)=0$) is the sign of t. Anywhere else, the distance from (x,y,z) to the z-axis is the distance of this point to the point $(0,0,z)$: $\sqrt{x^2+y^2}$. We must have

$$x^2+y^2=(a+\text{sgn}(y-a)\sqrt{r^2-z^2})^2=a^2+2a\text{sgn}(y-a)\sqrt{r^2-z^2}+r^2-z^2.$$

Rearranging

$$x^2+y^2+z^2-a^2-r^2=2a\text{sgn}(y-a)\sqrt{r^2-z^2},$$

or

$$(x^2+y^2+z^2-(a^2+r^2))^2=4a^2r^2-4a^2z^2$$

since $(\text{sgn}(y-a))^2=1$, (it could not be 0, why?). Rearranging again,

$$(x^2+y^2+z^2)^2-2(a^2+r^2)(x^2+y^2)+2(a^2-r^2)z^2+(a^2-r^2)^2=0.$$

The equation of the torus thus, is of fourth degree.

1.12 Parametric Curves and Surfaces

124 Definition Let $[a;b]\subseteq\mathbb{R}$. A *parametric curve* representation **r** of a curve Γ is a function $\overrightarrow{\mathbf{r}} : [a;b]\to\mathbb{R}^n$, with

$$\overrightarrow{\mathbf{r}}(t)=\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix},$$

and such that $\overrightarrow{\mathbf{r}}([a;b])=\Gamma$. $\mathbf{r}(a)$ is the *initial point* of the curve and $\mathbf{r}(b)$ its *terminal point*. A curve is *closed* if its initial point and its final point coincide.

The *trace* of the curve $\overrightarrow{\mathbf{r}}$ is the set of all images of $\overrightarrow{\mathbf{r}}$, that is, Γ. If there exist $t_1 \neq t_2$ such that $\overrightarrow{\mathbf{r}}(t_1) = \overrightarrow{\mathbf{r}}(t_2) = \mathbf{p}$, then $\overrightarrow{\mathbf{p}}$ is a *multiple point* of the curve. The curve is *simple* if its has no multiple points. A closed curve whose only multiple points are its endpoints is called a *Jordan curve.*

It is important to realise that a curve Γ might have different parametric representations. For example

$$\overrightarrow{\mathbf{r}} : \begin{array}{ccc} [0; 2\pi] & \to & \mathbb{R}^2 \\ t & \mapsto & \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} \end{array}$$

and

$$\overrightarrow{\mathbf{s}} : \begin{array}{ccc} [0; 2\pi] & \to & \mathbb{R}^2 \\ t & \mapsto & \begin{bmatrix} \sin 2t \\ \cos 2t \end{bmatrix} \end{array}$$

are two parametrisations for for the unit circle $\Gamma : \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Notice that **r** travels the unit circle once starting at $(1, 0)$ and going counterclockwise, and so **r** is a simple curve. One the other hand, **s** starts at $(0, 1)$ and travels the unit circle twice going clockwise.

We define

$$\lim_{t \to a} \overrightarrow{\mathbf{r}}(t) = \begin{bmatrix} \lim_{t \to a} x_1(t) \\ \lim_{t \to a} x_2(t) \\ \vdots \\ \lim_{t \to a} x_n(t) \end{bmatrix},$$

provided each of the limits on the right hand side exists and

$$\overrightarrow{\mathbf{r}}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix},$$

provided each of the entries on the right hand side be differentiable at t. In particular, we define the differential of $\overrightarrow{\mathbf{r}}$ as

$$d\overrightarrow{\mathbf{r}} = \begin{bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{bmatrix}.$$

We will keep our practice of writing

$$\overrightarrow{\mathbf{r}}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

in $\mathbb{R}^2$ and

$$\overrightarrow{\mathbf{r}}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}$$

in $\mathbb{R}^3$.

125 Example The trace of

$$\overrightarrow{\mathbf{r}}(t) = \overrightarrow{\mathbf{i}} \cos t + \overrightarrow{\mathbf{j}} \sin t + \overrightarrow{\mathbf{k}} t$$

is known as a *cylindrical helix.*

126 Example Find a parametric representation for the curve resulting by the intersection of the plane $3x + y + z = 1$ and the cylinder $x^2 + 2y^2 = 1$ in $\mathbb{R}^3$.

Solution: The projection of the intersection of the plane $3x + y + z$ and the cylinder is the ellipse $x^2 + 2y^2 = 1$, on the xy-plane. This ellipse can be parametrised as

$$x = \cos t, \ y = \frac{\sqrt{2}}{2} \sin t, \ 0 \leq t \leq 2\pi.$$

From the equation of the plane,

$$z = 1 - 3x - y = 1 - 3\cos t - \frac{\sqrt{2}}{2} \sin t.$$

Thus we may take the parametrisation

$$\overrightarrow{\mathbf{r}}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \cos t \\ \frac{\sqrt{2}}{2} \sin t \\ 1 - 3\cos t - \frac{\sqrt{2}}{2} \sin t \end{bmatrix}.$$

127 Example Let P be a point at a distance d from the centre of a circle of radius r. The curve traced out by P as the circle rolls along a straight line is called a *trochoid*. Find a parametrisation of the trochoid.

Solution: Let θ be the angle (in radians) of rotation of the circle, and let C be the centre of the circle. At $\theta = 0$ the centre of the circle is at $(0, r)$, and $P = (0, r - d)$. Suppose the circle is displaced towards the right, making the point P to rotate an angle of $\theta = \theta_0$ radians. Then the centre of the circle has displaced $r\theta_0$ units horizontally, and so is now located at $(r\theta, r)$. The polar co-ördinates of the point P are $(d \sin \theta_0, d \cos \theta_0)$, in relation to the centre of the circle (notice that the circle moves clockwise). The point P has moved $X = r\theta - d \sin \theta_0$ horizontal units and $Y = r - d \cos \theta_0$ units. This is the desired parametrisation.

128 Example A *hypocycloid* is a curve traced out by a fixed point P on a circle $\mathscr{C}$ of radius r as $\mathscr{C}$ rolls on the inside of a circle with centre at O and radius R. If the initial position of P is $(R, 0)$, and θ is the angle, measured counterclockwise, that a ray starting at O and passing through the centre of $\mathscr{C}$ makes with the x-axis, shew that a parametrisation of the hypocycloid is

$$x = (R - r) \cos \theta + r \cos \left(\frac{(R - r)\theta}{r} \right),$$

$$y = (R - r) \sin \theta - r \sin \left(\frac{(R - r)\theta}{r} \right).$$

Solution: Suppose that starting from $\theta = 0$, the centre O' of the small circle moves counterclockwise inside the larger circle by an angle θ, and the point $P = (x, y)$ moves clockwise an angle ϕ. The arc length travelled by the centre of the small circle is $(R - r)\theta$ radians. At the same time the point P has rotated $r\phi$ radians, and so $(R - r)\theta = r\phi$. See figure 1, where $O'B$ is parallel to the x-axis.

Let A be the projection of P on the x-axis. From figure 2, $\angle OAP = \angle OPO' = \dfrac{\pi}{2}$, $\angle OO'P = \pi - \phi - \theta$, $\angle POA = \dfrac{\pi}{2} - \phi$, and $OP = (R -$

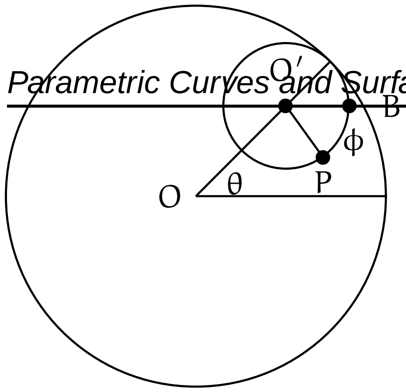

Figure 1.6: Hypocycloid

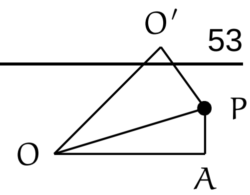

Figure 1.7: Hypocycloid

$r)\sin(\pi-\phi-\theta)$. Hence

$$x = (OP)\cos\angle POA = (R-r)\sin(\pi-\phi-\theta)\cos(\frac{\pi}{2}-\phi),$$

$$y = (R-r)\sin(\pi-\phi-\theta)\sin(\frac{\pi}{2}-\phi).$$

Now

$$\begin{aligned} x &= (R-r)\sin(\pi-\phi-\theta)\cos(\frac{\pi}{2}-\phi) \\ &= (R-r)\sin(\phi+\theta)\sin\phi \\ &= \frac{(R-r)}{2}(\cos\theta-\cos(2\phi+\theta)) \\ &= (R-r)\cos\theta-\frac{(R-r)}{2}(\cos\theta+\cos(2\phi+\theta)) \\ &= (R-r)\cos\theta-(R-r)(\cos(\theta+\phi)\cos\phi). \end{aligned}$$

Now, $\cos(\theta+\phi) = -\cos(\pi-\theta-\phi) = -\frac{r}{OO'} = -\frac{r}{R-r}$ and $\cos\phi = \cos\left(\frac{(R-r)\theta}{r}\right)$ and so

$$x = (R-r)\cos\theta-(R-r)(\cos(\theta+\phi)\cos\phi = (R-r)\cos\theta+r\cos\left(\frac{(R-r)\theta}{r}\right),$$

as required. The identity for y is proved similarly.

129 Definition A *parametric representation* **r** of a surface S in $\mathbb{R}^3$ is a function of the form

$$\overrightarrow{\mathbf{r}}: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^3 \\ (u,v) & \mapsto & \mathbf{r}(u,v) \end{array},$$

where

$$\overrightarrow{\mathbf{r}}(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix},$$

and whose image is S.

130 Example The surface of a sphere $S : \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 = R^2\}$ centred at the origin and of fixed radius $R > 0$ can be parametrised by using spherical co-ördinates, with $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$ and

$$\overrightarrow{\mathbf{r}}(\theta,\phi) = \begin{bmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{bmatrix} = \begin{bmatrix} R\cos\theta\sin\phi \\ R\sin\theta\sin\phi \\ R\cos\phi \end{bmatrix}.$$

For

$$x^2+y^2+z^2 = R^2(\cos^2\theta\sin^2\theta + \sin^2\theta\sin^2\phi + \cos^2\phi) = R^2.$$

131 Example A parametrisation for the torus generated by revolving the circle $(y-a)^2+z^2 = r^2$ around the z-axis is

$$\overrightarrow{\mathbf{r}}(\theta,\alpha) = \begin{bmatrix} a\cos\theta + r\cos\theta\cos\alpha \\ a\sin\theta + r\sin\theta\cos\alpha \\ r\sin\alpha \end{bmatrix},$$

with $(\theta,\alpha) \in [-\pi;\pi]^2$.

1.13 Frenet-Serret Formulæ

132 Definition A curve $\overrightarrow{\mathbf{r}} : [a;b] \to \mathbb{R}^3$ is said to be *smooth* if it is differentiable on $[a;b]$ and $\mathbf{r}'$ is continuous there. The *unit tangent vector* $\overrightarrow{\mathbf{T}}$ of a smooth curve at t is

$$\overrightarrow{\mathbf{T}}(t) = \frac{\mathbf{r}'(t)}{||\mathbf{r}'(t)||}.$$

The *unit normal vector* $\overrightarrow{\mathbf{N}}$ at t is

$$\overrightarrow{\mathbf{N}}(t) = \frac{\overrightarrow{\mathbf{T}}'(t)}{||\overrightarrow{\mathbf{T}}'(t)||},$$

and the *unit binormal vector* at t is

$$\overrightarrow{\mathbf{B}}(t) = \overrightarrow{\mathbf{T}}(t) \times \overrightarrow{\mathbf{N}}(t).$$

The three vectors $\{\overrightarrow{\mathbf{T}}, \overrightarrow{\mathbf{N}}, \overrightarrow{\mathbf{B}}\}$ are called the *Frenet-Serret* frame of the curve **r**.

133 Example Compute the Frenet-Serret frame at t for the cylindrical helix

$$\mathbf{r}(t) = \mathbf{i}\cos t + \mathbf{j}\sin t + \mathbf{k}t.$$

Solution: We have

$$\mathbf{r}'(t) = -\mathbf{i}\sin t + \mathbf{j}\cos t + \mathbf{k},$$

and $||\mathbf{r}'(t)|| = \sqrt{2}$. Thus

$$\overrightarrow{\mathbf{T}}(t) = \frac{1}{\sqrt{2}}\begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix}$$

Also,

$$\overrightarrow{\mathbf{T}}'(t) = \frac{1}{\sqrt{2}}\begin{bmatrix} -\cos t \\ -\sin t \\ 0 \end{bmatrix},$$

and $||\overrightarrow{\mathbf{T}}'(t)|| = \frac{1}{\sqrt{2}}$. Hence

$$\overrightarrow{\mathbf{N}} = \begin{bmatrix} -\cos t \\ -\sin t \\ 0 \end{bmatrix}.$$

Finally,

$$\overrightarrow{\mathbf{B}}(t) = \frac{1}{\sqrt{2}}\begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix} \times \begin{bmatrix} -\cos t \\ -\sin t \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}(\mathbf{i}(\sin t) - \mathbf{j}(\cos t) + \mathbf{k}).$$

1.14 Limits

134 Definition A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to have a limit $\mathbf{L} \in \mathbb{R}^m$ at $\mathbf{a} \in \mathbb{R}^n$ if $\forall \epsilon > 0 \exists \delta > 0$ such that

$$0 < ||\mathbf{x} - \mathbf{a}|| < \delta \Longrightarrow ||f(\mathbf{x}) - \mathbf{L}|| < \epsilon.$$

In such a case we write,

$$\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = \mathbf{L}.$$

The notions of infinite limits, limits at infinity, and continuity at a point, are analogously defined. Limits in more than one dimension are perhaps trickier to find, as one must approach the test point from infinitely many directions.

135 Example Find $\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2 + y^2}$.

Solution: We use the sandwich theorem. Observe that $0 \le x^2 \le x^2 + y^2$, and so $0 \le \frac{x^2}{x^2 + y^2} \le 1$. Thus

$$\lim_{(x,y)\to(0,0)} 0 \le \lim_{(x,y)\to(0,0)} \left| \frac{x^2 y}{x^2 + y^2} \right| \le \lim_{(x,y)\to(0,0)} |y|,$$

and hence

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2 + y^2} = 0.$$

136 Example Find $\lim_{(x,y)\to(0,0)} \frac{x^5 y^3}{x^6 + y^4}$.

Solution: Observe that if $|x| \le |y| \le 0$, then

$$\left| \frac{x^5 y^3}{x^6 + y^4} \right| \le \frac{y^8}{y^4} = y^4.$$

If $|y| \le |x| \le 0$, then

$$\left| \frac{x^5 y^3}{x^6 + y^4} \right| \le \frac{x^8}{x^6} = x^2.$$

Thus
$$\left|\frac{x^5y^3}{x^6+y^4}\right| \leq \max(y^4, x^2) \leq y^4 + x^2 \longrightarrow 0,$$
as $(x, y) \to (0, 0)$.
Aliter: Let $X = x^3, Y = y^2$.
$$\left|\frac{x^5y^3}{x^6+y^4}\right| = \frac{X^{5/3}Y^{3/2}}{X^2+Y^2}.$$

Passing to polar co-ördinates $X = \rho\cos\theta, Y = \rho\sin\theta$, we obtain
$$\left|\frac{x^5y^3}{x^6+y^4}\right| = \frac{X^{5/3}Y^{3/2}}{X^2+Y^2} = \rho^{5/3+3/2-2}|\cos\theta|^{5/3}|\sin\theta|^{3/2} \leq \rho^{7/6} \to 0,$$
as $(x, y) \to (0, 0)$.

137 Example Find $\lim_{(x,y)\to(0,0)} \frac{1+x+y}{x^2-y^2}$.

Solution: When $y = 0$,
$$\frac{1+x}{x^2} \to +\infty,$$
as $x \to 0$. When $x = 0$,
$$\frac{1+y}{-y^2} \to -\infty,$$
as $y \to 0$. The limit does not exist.

138 Example Find $\lim_{(x,y)\to(0,0)} \frac{xy^6}{x^6+y^8}$.

Solution: Putting $x = t^4, y = t^3$, we find
$$\frac{xy^6}{x^6+y^8} = \frac{1}{2t^2} \to +\infty,$$
as $t \to 0$. But when $y = 0$, the function is 0. Thus the limit does not exist.

139 Example Find $\lim_{(x,y)\to(0,0)} \frac{((x-1)^2+y^2)\log_e((x-1)^2+y^2)}{|x|+|y|}$.

Solution: When $y = 0$ we have

$$\frac{2(x-1)^2 \ln(|1-x|)}{|x|} \sim -\frac{2x}{|x|},$$

and so the function does not have a limit at $(0, 0)$.

140 Example Find $\lim_{(x,y)\to(0,0)} \frac{\sin(x^4) + \sin(y^4)}{\sqrt{x^4 + y^4}}$.

Solution: $\sin(x^4) + \sin(y^4) \leq x^4 + y^4$ and so

$$\left|\frac{\sin(x^4) + \sin(y^4)}{\sqrt{x^4 + y^4}}\right| \leq \sqrt{x^4 + y^4} \to 0,$$

as $(x, y) \to (0, 0)$.

141 Example Find $\lim_{(x,y)\to(0,0)} \frac{\sin x - y}{x - \sin y}$.

Solution: When $y = 0$ we obtain

$$\frac{\sin x}{x} \to 1,$$

as $x \to 0$. When $y = x$ the function is identically -1. Thus the limit does not exist.

Chapter 2

Differentiation

2.1 Local Study of Functions

142 Definition A function $\alpha : \mathbb{R} \to \mathbb{R}$ is said to be *infinitesimal* as $x \to a$ if $\lim_{x \to a} \alpha(x) = 0$.

143 Example $\sin : \mathbb{R} \to [-1; 1]$ is infinitesimal as $x \to 0$, since $\lim_{x \to 0} \sin x = 0$.

144 Example $f : \mathbb{R}^* \to \mathbb{R}, x \mapsto \frac{1}{x}$ is infinitesimal as $x \to +\infty$, since $\lim_{x \to +\infty} \frac{1}{x} = 0$.

145 Definition We say that a function $\alpha : \mathbb{R} \to \mathbb{R}$ is *asymptotic* to a function $\beta : \mathbb{R} \to \mathbb{R}$ as $x \to a$, and we write $\alpha \sim \beta$, if

$$\lim_{x \to a} \frac{\alpha(x)}{\beta(x)} = 1.$$

146 Example We have $\sin x \sim x$ as $x \to 0$, since $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

147 Example We have $x^2 + x \sim x$ as $x \to 0$, since $\lim_{x \to 0} \frac{x^2 + x}{x} = 1$.

148 Example We have $x^2 + x \sim x^2$ as $x \to +\infty$, since $\lim_{x \to +\infty} \frac{x^2 + x}{x^2} = 1$.

149 Definition (Du-Bois Raymond-Landau "small oh" notation) We say that $\alpha : \mathbb{R} \to \mathbb{R}$ is *negligible* in relation to $\beta : \mathbb{R} \to \mathbb{R}$ as $x \to a$ or that β is *preponderant* in relation to α as $x \to a$, if

$$\lim_{x \to a} \frac{\alpha(x)}{\beta(x)} = 0.$$

We express the condition above with the notation $\alpha(x) = \underset{x \to a}{\mathbf{o}} (\beta(x))$ (read "α of x is small oh of β of x as x tends to a")

When the limit to which the independent variable tends is understood, we abbreviate $\underset{x \to a}{\mathbf{o}} (\alpha(x))$ *as* $\mathbf{o} (\alpha(x))$

In particular, if α is infinitesimal as $x \to a$, then $\alpha(x) = \mathbf{o}(1)$ as $x \to a$.

150 Example We have $x^2 = \mathbf{o}(x)$ as $x \to 0$ since

$$\lim_{x \to 0} \frac{x^2}{x} = \lim_{x \to 0} x = 0.$$

151 Example We have $x = \mathbf{o}\left(x^2\right)$ as $x \to +\infty$ since

$$\lim_{x \to +\infty} \frac{x}{x^2} = \lim_{x \to 0} \frac{1}{x} = 0.$$

We write $\alpha(x) = \gamma(x) + \mathbf{o}(\beta(x))$ as $x \to a$ if $\alpha(x) - \gamma(x) = \mathbf{o}(\beta(x))$ as $x \to a$.

152 Example We have $\sin x = x + \mathbf{o}(x)$ as $x \to 0$ since

$$\lim_{x \to 0} \frac{\sin x - x}{x} = \lim_{x \to 0} \frac{\sin x}{x} - \lim_{x \to 0} 1 = 1 - 1 = 0.$$

153 Theorem Let α and β be infinitesimal functions as $x \to a$. Then the following hold.

- The sum of two infinitesimals is an infinitesimal:

$$\mathbf{o}\left(\beta(x)\right) + \mathbf{o}\left(\beta(x)\right) = \mathbf{o}\left(\beta(x)\right).$$

- The difference of two infinitesimals is an infinitesimal:

$$\mathbf{o}\left(\beta(x)\right) - \mathbf{o}\left(\beta(x)\right) = \mathbf{o}\left(\beta(x)\right).$$

- $\forall c \in \mathbb{R} \setminus \{0\}, \mathbf{o}\left(c\beta(x)\right) = \mathbf{o}\left(\beta(x)\right).$
- $\forall n \in \mathbb{N}, n \geq 2, 1 \leq k \leq n-1,\ \mathbf{o}\left((\beta(x))^n\right) = \mathbf{o}\left((\beta(x))^k\right).$
- $\mathbf{o}\left(\mathbf{o}\left(\beta(x)\right)\right) = \mathbf{o}\left(\beta(x)\right).$
- $\forall n \in \mathbb{N}, n \geq 1,\ (\beta(x))^n \mathbf{o}\left(\beta(x)\right) = \mathbf{o}\left((\beta(x))^{n+1}\right).$
- $\forall n \in \mathbb{N}, n \geq 2, \dfrac{\mathbf{o}\left((\beta(x))^n\right)}{\beta(x)} = \mathbf{o}\left((\beta(x))^{n-1}\right).$
- $\dfrac{\mathbf{o}\left(\beta(x)\right)}{\beta(x)} = \mathbf{o}\left(1\right).$
- If c_k are real numbers, then $\mathbf{o}\left(\sum_{k=1}^{n} c_k(\beta(x))^k\right) = \mathbf{o}\left(\beta(x)\right).$
- $(\alpha\beta)(x) = \mathbf{o}\left(\alpha(x)\right)$ and $(\alpha\beta)(x) = \mathbf{o}\left(\beta(x)\right).$
- If $\alpha \sim \beta$, then $(\alpha - \beta)(x) = \mathbf{o}\left(\alpha(x)\right)$ and $(\alpha - \beta)(x) = \mathbf{o}\left(\beta(x)\right).$

154 Example In view of theorem 153, we have

$$\mathbf{o}\left(-2x^3 + 8x^2\right) = \mathbf{o}\left(x\right),$$

as $x \to 0$.

155 Example In view of theorem 153, we have

$$-2x^3 + 8x^2 + x \sim x,$$

as $x \to 0$.

156 Example In view of theorem 153, we have

$$\mathbf{o}\left(-2x^3 + 8x^2\right) = \mathbf{o}\left(x^4\right),$$

as $x \to +\infty$.

157 Example In view of theorem 153, we have

$$-2x^3 + 8x^2 + x \sim -2x^3,$$

as $x \to +\infty$.

The following result follows easily from the Maclaurin expansion of the given function.

158 Theorem Let $x \to 0$. Then

- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \mathbf{o}\left(x^{2n+2}\right)$. In particular, $\sin x = x + \mathbf{o}\left(x^2\right)$.

- $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \mathbf{o}\left(x^{2n+1}\right)$. In particular, $\cos x = 1 + \mathbf{o}\left(x\right)$.

- $\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \mathbf{o}\left(x^5\right)$.

- $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \mathbf{o}\left(x^n\right)$

- $\log_e(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n+1}\frac{x^n}{n} + \mathbf{o}\left(x^n\right)$.

- $(1+x)^\tau = 1 + \tau x + \frac{\tau(\tau-1)}{2}x^2 + \cdots + \frac{\tau(\tau-1)(\tau-2)(\tau-3)\cdots(\tau-n+1)}{n!}x^n + \mathbf{o}\left(x^n\right)$.

159 Example Since $\tan x = x + \mathbf{o}(x)$, we have

$$\tan \frac{x^2}{2} = \frac{x^2}{2} + \mathbf{o}\left(\frac{x^2}{2}\right) = \frac{x^2}{2} + \mathbf{o}\left(x^2\right),$$

as $x \to 0$. Also,

$$(\tan x)^3 = (x + \mathbf{o}(x))^3 = x^3 + 3x^2\mathbf{o}(x) + 3x\mathbf{o}\left(x^2\right) + (\mathbf{o}(x))^3 = x^3 + \mathbf{o}\left(x^3\right).$$

160 Example Since $\cos x = 1 - \dfrac{x^2}{2!} + \mathbf{o}\left(x^2\right)$, we have

$$\cos 3x^2 = 1 - \frac{9x^4}{2} + \mathbf{o}\left(x^4\right).$$

161 Example Find an asymptotic expansion of $\cot^2 x$ of type $\mathbf{o}\left(x^{-2}\right)$ as $x \to 0$.

Solution: Since $\tan x \sim x$ we have

$$\cot^2 x \sim \frac{1}{x^2}.$$

We can write this as $\cot^2 x = \dfrac{1}{x^2} + \mathbf{o}\left(\dfrac{1}{x^2}\right)$.

162 Example Calculate

$$\lim_{x \to 0} \frac{\sin \sin \tan \frac{x^2}{2}}{\log_e \cos 3x}.$$

Solution: We use theorems 158 and 153.

$$\begin{aligned}
\sin \sin \tan \frac{x^2}{2} &= \sin \sin\left(\frac{x^2}{2} + \mathbf{o}\left(x^2\right)\right) \\
&= \sin\left(\frac{x^2}{2} + \mathbf{o}\left(x^2\right) + \mathbf{o}\left(\frac{x^2}{2} + \mathbf{o}\left(x^2\right)\right)\right) \\
&= \sin\left(\frac{x^2}{2} + \mathbf{o}\left(x^2\right)\right) \\
&= \frac{x^2}{2} + \mathbf{o}\left(x^2\right),
\end{aligned}$$

and

$$\begin{aligned}\log_e \cos 3x &= \log_e\left(1-\frac{9x^2}{2}+\mathbf{o}\left(x^2\right)\right)\\ &= -\frac{9x^2}{2}+\mathbf{o}\left(x^2\right)+\mathbf{o}\left(-\frac{9x^2}{2!}+\mathbf{o}\left(x^2\right)\right)\\ &= -\frac{9x^2}{2}+\mathbf{o}\left(x^2\right)\end{aligned}$$

The limit is thus equal to

$$\lim_{x\to 0}\frac{\dfrac{x^2}{2}+\mathbf{o}\left(x^2\right)}{-\dfrac{9x^2}{2}+\mathbf{o}\left(x^2\right)}=\lim_{x\to 0}\frac{\dfrac{1}{2}+\mathbf{o}\left(1\right)}{-\dfrac{9}{2}+\mathbf{o}\left(1\right)}=-\frac{1}{9}.$$

163 Example Find $\lim_{x\to 0}(\cos x)^{(\cot^2 x)}$.

Solution: By example 161, we have $\cot^2 x=\dfrac{1}{x^2}+\mathbf{o}\left(\dfrac{1}{x^2}\right)$. Also,

$$\log_e \cos x=\log_e\left(1-\frac{x^2}{2}+\mathbf{o}\left(x^2\right)\right)=-\frac{x^2}{2}+\mathbf{o}\left(x^2\right).$$

Hence

$$\begin{aligned}(\cos x)^{\cot^2 x} &= \exp\left((\cot^2 x)\log_e \cos x\right)\\ &= \exp\left(\left(\frac{1}{x^2}+\mathbf{o}\left(\frac{1}{x^2}\right)\right)\left(-\frac{x^2}{2}+\mathbf{o}\left(x^2\right)\right)\right)\\ &= \exp(-\frac{1}{2}+\mathbf{o}\left(1\right))\\ &\to e^{-\frac{1}{2}},\end{aligned}$$

as $x\to 0$.

164 Example Find an asymptotic development of $\log_e(2\cos x+\sin x)$ around $x=0$ of order $\mathbf{o}\left(x^4\right)$.

Solution: By theorem 158,

$$\begin{aligned}
2\cos x + \sin x &= 2\left(1 - \frac{x^2}{2} + \frac{x^4}{24} + \mathbf{o}\left(x^5\right)\right) + \left(x - \frac{x^3}{6} + \mathbf{o}\left(x^4\right)\right) \\
&= 2 + x - x^2 - \frac{x^3}{6} + \frac{x^4}{12} + \mathbf{o}\left(x^4\right) \\
&= 2\left(1 + \frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right),
\end{aligned}$$

and so,

$$\begin{aligned}
\log_e(2\cos x + \sin x) &= \log_e 2\left(1 + \frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right) \\
&= \log_e 2 + \log_e\left(1 + \frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right) \\
&= \log_e 2 + \left(\frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right) \\
&\quad - \frac{1}{2}\left(\frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right)^2 \\
&\quad + \frac{1}{3}\left(\frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right)^3 \\
&\quad - \frac{1}{4}\left(\frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24} + \mathbf{o}\left(x^4\right)\right)^4 + \mathbf{o}\left(x^4\right) \\
&= \log_e 2 + \left(\frac{x}{2} - \frac{x^2}{2} - \frac{x^3}{12} + \frac{x^4}{24}\right) - \frac{1}{2}\left(\frac{x^2}{4} - \frac{x^3}{2} + \frac{x^4}{6}\right) \\
&\quad + \frac{1}{3}\left(\frac{x^3}{8} - \frac{3x^4}{8}\right) - \frac{1}{4}\cdot\frac{x^4}{16} + \mathbf{o}\left(x^4\right) \\
&= \log_e 2 + \frac{x}{2} - \frac{5x^2}{8} + \frac{5x^3}{24} - \frac{35x^4}{192} + \mathbf{o}\left(x^4\right)
\end{aligned}$$

as $x \to 0$.

Our main concern, as far as asymptotic developments is concerned, will be expansions when the vector tends to $\overrightarrow{\mathbf{0}}$. Let $f, g : \mathbb{R}^n \to \mathbb{R}^m$. By $f(\overrightarrow{\mathbf{v}}) = g(\overrightarrow{\mathbf{v}}) + \mathbf{o}\left(||\overrightarrow{\mathbf{v}}||\right)$, as $\overrightarrow{\mathbf{v}} \to \overrightarrow{\mathbf{0}}$, we mean that

$$\lim_{\overrightarrow{\mathbf{v}} \to \overrightarrow{\mathbf{0}}} \frac{||f(\overrightarrow{\mathbf{v}}) - g(\overrightarrow{\mathbf{v}})||}{||\overrightarrow{\mathbf{v}}||} = 0.$$

In the direction of local behaviour of vector functions, we start with the following theorem.

165 Theorem Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $L(\overrightarrow{\mathbf{0}}) = \overrightarrow{\mathbf{0}}$. Moreover, there exists a constant $M > 0$ such that

$$||L(\overrightarrow{\mathbf{h}})|| \leq M\left|\left|\overrightarrow{\mathbf{h}}\right|\right|.$$

Proof By linearity,

$$L(\overrightarrow{\mathbf{0}}) = L(\overrightarrow{\mathbf{0}} + \overrightarrow{\mathbf{0}}) = L(\overrightarrow{\mathbf{0}}) + L(\overrightarrow{\mathbf{0}}),$$

which gives the first part.

Now, let

$$\overrightarrow{\mathbf{h}} = \sum_{k=1}^{n} h_k \overrightarrow{\mathbf{e}}_k,$$

where the $\overrightarrow{\mathbf{e}}_k$ are the standard basis for $\mathbb{R}^n$. Then

$$L(\overrightarrow{\mathbf{h}}) = \sum_{k=1}^{n} h_k L(\overrightarrow{\mathbf{e}}_k),$$

and hence by the triangle inequality, and by the Cauchy-Bunyakovsky-Schwarz inequality,

$$\begin{aligned} ||L(\overrightarrow{\mathbf{h}})|| &\leq \sum_{k=1}^{n} |h_k|||L(\overrightarrow{\mathbf{e}}_k)|| \\ &\leq (\sum_{k=1}^{n} |h_k|^2)^{1/2}(\sum_{k=1}^{n} ||L(\overrightarrow{\mathbf{e}}_k)||^2)^{1/2} \\ &= \left|\left|\overrightarrow{\mathbf{h}}\right|\right|(\sum_{k=1}^{n} ||L(\overrightarrow{\mathbf{e}}_k)||^2)^{1/2}, \end{aligned}$$

and the theorem follows upon putting $M = (\sum_{k=1}^{n} ||L(\overrightarrow{\mathbf{e}}_k)||^2)^{1/2}$. ❑

166 Corollary With M as in theorem 165,

$$||\overrightarrow{\mathbf{h}} \times L(\overrightarrow{\mathbf{h}})|| \leq M\left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2 = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right),$$

as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$.

2.2 Definition of the Derivative

We now define the derivative in a multidimensional vector space. Recall that in one variable, a function $g : \mathbb{R} \to \mathbb{R}$ is said to be differentiable at $x = a$ if the limit

$$\lim_{x\to a} \frac{g(x) - g(a)}{x - a} = g'(a)$$

exists. The limit condition above is equivalent to saying that

$$\lim_{x\to a} \frac{g(x) - g(a) - g'(a)(x - a)}{x - a} = 0,$$

or equivalently,

$$\lim_{h\to 0} \frac{g(a + h) - g(a) - g'(a)(h)}{h} = 0.$$

This last condition we may write in small oh notation as

$$g(a + h) - g(a) = g'(a)(h) + \mathbf{o}(h),$$

as $h \to 0$. The above analysis provides an analogue definition for the higher-dimensional case. Observe that since we may not divide by vectors, the corresponding definition in higher dimensions involves quotients of norms.

167 Definition Let $A \subseteq \mathbb{R}^n$. A function $f : A \to \mathbb{R}^m$ is said to be *differentiable* at $\overrightarrow{\mathbf{a}} \in A$ if there is a linear transformation, called the *derivative of* f *at* $\overrightarrow{\mathbf{a}}$, $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\overrightarrow{\mathbf{x}}\to\overrightarrow{\mathbf{a}}} \frac{||f(\overrightarrow{\mathbf{x}}) - f(\overrightarrow{\mathbf{a}}) - \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})||}{||\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}||} = 0.$$

Equivalently, f is differentiable at $\overrightarrow{\mathbf{a}}$ if there is a linear transformation $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)$ such that

$$f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{h}}) + \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right),$$

as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$.

The condition for differentiability at $\overrightarrow{\mathbf{a}}$ is equivalent to

$$f(\overrightarrow{\mathbf{x}}) - f(\overrightarrow{\mathbf{a}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + \mathbf{o}\left(||\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}||\right),$$

as $\overrightarrow{\mathbf{x}} \to \overrightarrow{\mathbf{a}}$.

168 Theorem If A is an open set in definition 167, $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)$ is uniquely determined.

Proof Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be another linear transformation satisfying definition 167. We must prove that $\forall \overrightarrow{\mathbf{v}} \in \mathbb{R}^n, L(\overrightarrow{\mathbf{v}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{v}})$. Since A is open, $\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}} \in A$ for sufficiently small $\left|\left|\overrightarrow{\mathbf{h}}\right|\right|$. By definition, as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$, we have

$$f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{h}}) + \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right).$$

and

$$f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}}) = L(\overrightarrow{\mathbf{h}}) + \mathbf{o}\left(||vectorh||\right).$$

Now, observe that

$$\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{v}}) - L(\overrightarrow{\mathbf{v}} = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) + f(\overrightarrow{\mathbf{a}}) + f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}}) - L(\overrightarrow{\mathbf{h}}).$$

By the triangle inequality,

$$\begin{aligned} ||\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{v}}) - L(\overrightarrow{\mathbf{v}}|| &\leq ||\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) + f(\mathbf{a})|| \\ &\quad + ||f(\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{a}}) - L(\mathbf{h})|| \\ &= \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right) + \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right) \\ &= \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right), \end{aligned}$$

as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$ This means that

$$||L(\overrightarrow{\mathbf{v}}) - \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{v}})|| \to 0,$$

i.e., $L(\overrightarrow{\mathbf{v}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{v}})$, completing the proof. ❑

If $A = \{\overrightarrow{\mathbf{a}}\}$, a singleton, then $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)$ is not uniquely determined. For $||\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}|| < \delta$ holds only for $\overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{a}}$, and so $f(\overrightarrow{\mathbf{x}}) = f(\overrightarrow{\mathbf{a}})$. Any linear transformation T will satisfy the definition, as $T(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) = T(\overrightarrow{\mathbf{0}}) = \overrightarrow{\mathbf{0}}$, and

$$||f(\overrightarrow{\mathbf{x}}) - f(\overrightarrow{\mathbf{a}}) - T(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})|| = ||\overrightarrow{\mathbf{0}}|| = 0,$$

identically.

169 Example If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(L) = L$, for any $\mathbf{a} \in \mathbb{R}^n$.

Solution: Since $\mathbb{R}^n$ is an open set, we know that $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(L)$ uniquely determined. Thus if L satisfies definition 167, then the claim is established. But by linearity

$$||L(\overrightarrow{\mathbf{x}}) - L(\overrightarrow{\mathbf{a}}) - L(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})|| = ||L(\overrightarrow{\mathbf{x}}) - L(\overrightarrow{\mathbf{a}}) - L(\overrightarrow{\mathbf{x}}) + L(\overrightarrow{\mathbf{a}})|| = ||0|| = 0,$$

whence the claim follows.

170 Example Let $f : \mathbb{R}^n \to \mathbb{R}, n \geq 1, f(\overrightarrow{\mathbf{x}}) = ||x||$ be the usual norm in $\mathbb{R}^n$, with $\left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2 = \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{x}}$. Prove that

$$\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)(\overrightarrow{\mathbf{v}}) = \frac{\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{v}}}{||x||},$$

for $\overrightarrow{\mathbf{x}} \neq \overrightarrow{\mathbf{0}}$, but that f is not differentiable at $\overrightarrow{\mathbf{0}}$.

Solution: Assume that $\overrightarrow{\mathbf{x}} \neq \overrightarrow{\mathbf{0}}$. We use the fact that $(1+t)^{1/2} = 1 + \frac{t}{2} + \mathbf{o}(t)$ as $t \to 0$, from theorem 153. We have

$$\begin{aligned} f(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{x}}) &= ||\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}|| - ||x|| \\ &= \sqrt{(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) \bullet (\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}})} - ||x|| \\ &= \sqrt{\left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2 + 2\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2} - \left|\left|\overrightarrow{\mathbf{x}}\right|\right| \\ &= \frac{2\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2}{\sqrt{\left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2 + 2\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2} + \left|\left|\overrightarrow{\mathbf{x}}\right|\right|}. \end{aligned}$$

As $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$,

$$\sqrt{\left|\left|\overrightarrow{\mathbf{x}}\right|\right|^2 + 2\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2} + \left|\left|\overrightarrow{\mathbf{x}}\right|\right| \to 2\left|\left|\overrightarrow{\mathbf{x}}\right|\right|.$$

Since $\left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2 = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right)$ as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$, we have

$$\frac{2\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2}{\sqrt{||\overrightarrow{\mathbf{x}}||^2 + 2\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{h}} + \left|\left|\overrightarrow{\mathbf{h}}\right|\right|^2} + ||\overrightarrow{\mathbf{x}}||} \to \frac{\overrightarrow{\mathbf{x}}\bullet\overrightarrow{\mathbf{h}}}{\left|\left|\overrightarrow{\mathbf{h}}\right|\right|} + \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right),$$

proving the first assertion.

To prove the second assertion, assume that there is a linear transformation $\mathscr{D}_{\overrightarrow{\mathbf{0}}}(f) = L$, $L : \mathbb{R}^n \to \mathbb{R}$ such that

$$||f(\overrightarrow{\mathbf{0}} + \overrightarrow{\mathbf{h}}) - f(\overrightarrow{\mathbf{0}}) - L(\overrightarrow{\mathbf{h}})|| = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right),$$

as $\left|\left|\overrightarrow{\mathbf{h}}\right|\right| \to 0$. Recall that by theorem 165, $L(\overrightarrow{\mathbf{0}}) = \overrightarrow{\mathbf{0}}$, and so by example 169, $\mathscr{D}_{\overrightarrow{\mathbf{0}}}(L)(\overrightarrow{\mathbf{0}}) = L(\overrightarrow{\mathbf{0}}) = \overrightarrow{\mathbf{0}}$. This implies that $\dfrac{L(\overrightarrow{\mathbf{h}})}{\left|\left|\overrightarrow{\mathbf{h}}\right|\right|} \to \mathscr{D}_{\overrightarrow{\mathbf{0}}}(L)(\overrightarrow{\mathbf{0}}) = \overrightarrow{\mathbf{0}}$, as $\left|\left|\overrightarrow{\mathbf{h}}\right|\right| \to 0$. Since $f(\overrightarrow{\mathbf{0}}) = ||0|| = 0$, $f(\overrightarrow{\mathbf{h}}) = \left|\left|\overrightarrow{\mathbf{h}}\right|\right|$ this would imply that

$$\left|\left|\left|\left|\overrightarrow{\mathbf{h}}\right|\right| - L(\overrightarrow{\mathbf{h}})\right|\right| = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right),$$

or

$$\left|\left|1 - \frac{L(\overrightarrow{\mathbf{h}})}{\left|\left|\overrightarrow{\mathbf{h}}\right|\right|}\right|\right| = \mathbf{o}(1).$$

But the sinistral side $\to 1$ as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$, and the dextral side $\to 0$ as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$. This is a contradiction, and so, such linear transformation L does not exist at the point $\overrightarrow{\mathbf{0}}$.

171 Example Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation and

$$F : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \overrightarrow{\mathbf{x}} & \mapsto & \overrightarrow{\mathbf{x}} \times L(\overrightarrow{\mathbf{x}}) \end{array}.$$

Shew that F is differentiable and that

$$\mathscr{D}_{\overrightarrow{\mathbf{x}}}(F)(\overrightarrow{\mathbf{h}}) = \overrightarrow{\mathbf{x}} \times L(\overrightarrow{\mathbf{h}}) + \overrightarrow{\mathbf{h}} \times L(\overrightarrow{\mathbf{x}}).$$

Solution: We have

$$\begin{aligned} F(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) - F(\overrightarrow{\mathbf{x}}) &= (\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) \times L(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) - \overrightarrow{\mathbf{x}} \times L(\overrightarrow{\mathbf{x}}) \\ &= (\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) \times (L(\overrightarrow{\mathbf{x}}) + L(\overrightarrow{\mathbf{h}})) - \overrightarrow{\mathbf{x}} \times L(\overrightarrow{\mathbf{x}}) \\ &= \overrightarrow{\mathbf{x}} \times L(\overrightarrow{\mathbf{h}}) + \overrightarrow{\mathbf{h}} \times L(\overrightarrow{\mathbf{x}}) + \overrightarrow{\mathbf{h}} \times L(\overrightarrow{\mathbf{h}}) \end{aligned}$$

Now, $||\overrightarrow{\mathbf{h}} \times L(\overrightarrow{\mathbf{h}})|| = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right)$ as $\overrightarrow{\mathbf{h}} \to \overrightarrow{\mathbf{0}}$ by virtue of corollary 166. This yields the desired result.

The following examples consider derivatives of functions with domains other than $\mathbb{R}^n$.

172 Example Let

$$f: \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ (\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) & \mapsto & \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}} \end{array}$$

be the usual dot product in $\mathbb{R}^3$. Shew that f is differentiable and that

$$\mathscr{D}_{(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})} f(\overrightarrow{\mathbf{h}}, \overrightarrow{\mathbf{k}}) = \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}} + \overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{y}}.$$

Solution: We have

$$\begin{aligned} f(\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}, \overrightarrow{\mathbf{y}} + \overrightarrow{\mathbf{k}}) - f(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) &= (\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{h}}) \bullet (\overrightarrow{\mathbf{y}} + \overrightarrow{\mathbf{k}}) - \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}} \\ &= \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}} + \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}} + \overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{y}} + \overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{k}} - \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}} \\ &= \overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{k}} + \overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{y}} + \overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{k}}. \end{aligned}$$

As $(\overrightarrow{\mathbf{h}}, \overrightarrow{\mathbf{k}}) \to (\overrightarrow{\mathbf{0}}, \overrightarrow{\mathbf{0}})$, we have by the Cauchy-Buniakovskii-Schwarz inequality, $|\overrightarrow{\mathbf{h}} \bullet \overrightarrow{\mathbf{k}}| \leq \left|\left|\overrightarrow{\mathbf{h}}\right|\right| ||k|| = \mathbf{o}\left(\left|\left|\overrightarrow{\mathbf{h}}\right|\right|\right)$, which proves the assertion.

173 Example Shew that

$$f: \begin{array}{ccc} \mathbf{M}_n(\mathbb{R}) & \to & \mathbf{M}_n(\mathbb{R}) \\ X & \mapsto & X^2 \end{array}$$

is differentiable and that

$$\mathscr{D}_X f(H) = XH + HX.$$

Solution: We have

$$\begin{aligned} f(X+H)-f(X) &= (X+H)^2 - X^2 \\ &= X^2+XH+HX+H^2-X^2 \\ &= XH+HX+H^2 \\ &= XH+HX+\mathbf{o}\left(H\right), \end{aligned}$$

proving the assertion.

174 Example Let

$$f: \begin{array}{ccc} \mathbf{M}_2(\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \det A \end{array}.$$

Given $H \in \mathbf{M}_2(\mathbb{R})$, find an expression for $\mathscr{D}_A(f)(H)$

Solution: Put

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad H = \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix}.$$

Then

$$\begin{aligned} f(A+H)-f(A) &= \det \begin{bmatrix} a_1+h_1 & a_2+h_2 \\ a_3+h_3 & a_4+h_4 \end{bmatrix} - \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \\ &= (a_1+h_1)(a_4+h_4)-(a_2+h_2)(a_3+h_3)-a_1a_4+a_2a_3 \\ &= a_1a_4-a_2a_3+a_1h_4+a_4h_1-a_2h_3-a_3h_2 \\ &\quad +h_1h_4-h_2h_3-a_1a_4+a_2a_3 \\ &= a_1h_4+a_4h_1-a_2h_3-a_3h_2+h_1h_4-h_2h_3 \\ &= \operatorname{tr}\left(\begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix}\begin{bmatrix} a_4 & -a_2 \\ -a_3 & a_1 \end{bmatrix}\right)+h_1h_4-h_2h_3 \\ &= \operatorname{tr}\left(H\operatorname{adj}\left(A\right)\right)+h_1h_4-h_2h_3, \end{aligned}$$

where $\operatorname{adj}(A)$ is the (classical) adjoint matrix of A. Now, the Arithmetic-Mean-Geometric-Mean inequality for real numbers a, b says that $|a||b| \leq \frac{a^2+b^2}{2}$ and so

$$|h_1h_4-h_2h_3| \leq |h_1||h_4|+|h_2||h_3| \leq \frac{h_1^2+h_2^2+h_3^2+h_4^2}{2} = \frac{1}{2}||H||^2 = \mathbf{o}\left(H\right),$$

as $H \to \mathbf{0}_2$. This implies that

$$\mathscr{D}_A(f)(H) = \operatorname{tr}\left(H\operatorname{adj}\left(A\right)\right).$$

Just like in the one variable case, differentiability at a point, implies continuity at that point.

175 Theorem Suppose $A \subseteq \mathbb{R}^n$ is open and $f : A \to \mathbb{R}^n$ is differentiable on A. Then f is continuous on A.

Proof Given $\overrightarrow{\mathbf{a}} \in A$, we must shew that

$$\lim_{\overrightarrow{\mathbf{x}} \to \overrightarrow{\mathbf{a}}} f(\overrightarrow{\mathbf{x}}) = f(\overrightarrow{\mathbf{a}}).$$

Since f is differentiable at $\overrightarrow{\mathbf{a}}$ we have

$$f(\overrightarrow{\mathbf{x}}) - f(\overrightarrow{\mathbf{a}}) = \mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + \mathbf{o}\left(||\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}||\right),$$

and so

$$f(\overrightarrow{\mathbf{x}}) - f(\overrightarrow{\mathbf{a}}) \to \overrightarrow{\mathbf{0}},$$

as $\overrightarrow{\mathbf{x}} \to \overrightarrow{\mathbf{a}}$, proving the theorem. ❑

2.3 The Jacobi Matrix

We now establish a way which simplifies the process of finding the derivative of a function at a given point.

176 Definition Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, and put

$$f(\overrightarrow{\mathbf{x}}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}.$$

Here $f_i : \mathbb{R}^n \to \mathbb{R}$. The *partial derivative* $\dfrac{\partial f_i}{\partial x_j}(\mathbf{x})$ is defined as

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}) = \lim_{h \to 0} \frac{f_i(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n)}{h},$$

whenever this limit exists.

To find partial derivatives with respect to the j-th variable, we simply keep the other variables fixed and differentiate with respect to the j-th variable.

177 Example If $f : \mathbb{R}^3 \to \mathbb{R}$, and $f(\overrightarrow{\mathbf{r}}) = f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x + y^2 + z^3 + 3xy^2z^3$ then

$$\mathscr{D}_1(f)(\overrightarrow{\mathbf{r}}) = 1 + 3y^2z^3,$$

$$\mathscr{D}_2(f)(\overrightarrow{\mathbf{r}}) = 2y + 6xyz^3,$$

and

$$\mathscr{D}_3(f)(\overrightarrow{\mathbf{r}}) = 3z^2 + 9xy^2z^2.$$

Sometimes we will use the equivalent notations

$$\mathscr{D}_1(f)(\overrightarrow{\mathbf{r}}) = \frac{\partial f}{\partial x}(x, y, z),$$

$$\mathscr{D}_2(f)(\overrightarrow{\mathbf{r}}) = \frac{\partial f}{\partial y}(x, y, z),$$

etc..

178 Example Suppose $g : \mathbb{R} \to \mathbb{R}$ is continuous and $a \in \mathbb{R}$ is a constant. Find the partial derivatives with respect to x and y of

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f\begin{bmatrix} x \\ y \end{bmatrix} = \int_a^{x^2y} g(t)\, dt.$$

Solution: We have

$$\frac{\partial f}{\partial x}(x, y, z) = 2xyg(x^2y),$$

and

$$\frac{\partial f}{\partial y}(x, y, z) = x^2g(x^2y).$$

Since the derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, it can be represented by aid of matrices. The following theorem will allow us to determine the matrix representation for $\mathscr{D}_{\overrightarrow{\mathbf{a}}}(f)$ under the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

179 Theorem Let

$$f(\overrightarrow{\mathbf{x}}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}.$$

Suppose $A \subseteq \mathbb{R}^n$ is an open set and $f : A \to \mathbb{R}^m$ is differentiable. Then each partial derivative $\dfrac{\partial f_i}{\partial x_j}(\mathbf{x})$ exists, and the matrix representation of $\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)$ with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is the *Jacobi matrix*

$$\mathscr{J}_{\overrightarrow{\mathbf{x}}} f = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{x}) & \dfrac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \dfrac{\partial f_2}{\partial x_1}(\mathbf{x}) & \dfrac{\partial f_2}{\partial x_2}(\mathbf{x}) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(\mathbf{x}) & \dfrac{\partial f_n}{\partial x_2}(\mathbf{x}) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

Proof Let $\overrightarrow{\mathbf{e}}_j, 1 \leq j \leq n$, be the standard basis for $\mathbb{R}^n$. To obtain the Jacobi matrix, we must compute $\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)(\overrightarrow{\mathbf{e}}_j)$, which will give us the j-th column of the Jacobi matrix. Let $\mathscr{J}_{\overrightarrow{\mathbf{x}}} f = (J_{ij})$, and observe that

$$\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)(\overrightarrow{\mathbf{e}}_j) = \begin{bmatrix} J_{1j} \\ J_{2j} \\ \vdots \\ J_{nj} \end{bmatrix}.$$

and put $\overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{x}} + \varepsilon \overrightarrow{\mathbf{e}}_j, \varepsilon \in \mathbb{R}$. Notice that

$$\begin{aligned} & \frac{||f(\overrightarrow{\mathbf{y}}) - f(\overrightarrow{\mathbf{x}}) - \mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)(\overrightarrow{\mathbf{y}} - \overrightarrow{\mathbf{x}})||}{||\overrightarrow{\mathbf{y}} - \overrightarrow{\mathbf{x}}||} \\ = & \frac{||f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_j, \ldots, x_n) - \varepsilon \mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)(\overrightarrow{\mathbf{e}}_j)||}{|\varepsilon|}. \end{aligned}$$

Since the sinistral side $\to 0$ as $\varepsilon \to 0$, the so does the i-th component of the numerator, and so,

$$\frac{|f_i(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n) - \varepsilon J_{ij}|}{|\varepsilon|} \to 0.$$

This entails that

$$J_{ij} = \lim_{\varepsilon\to 0} \frac{f_i(x_1, x_2, \ldots, x_j + \varepsilon, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n)}{\varepsilon} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}).$$

This finishes the proof. ❑

180 Example Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$f(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} xy + yz \\ \log_e xy \end{bmatrix}.$$

Compute the Jacobi matrix of f under the usual basis for $\mathbb{R}^3$ and $\mathbb{R}^2$. Also, compute the matrix representation for $\mathscr{D}_{\overrightarrow{\mathbf{r}}}(f)$ under the the usual basis for $\mathbb{R}^3$ and the basis $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$. Finally, find $\mathscr{D}_{(1,2,3)} f \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

Solution: The Jacobi matrix is the 2×3 matrix

$$\mathscr{J}_{\overrightarrow{\mathbf{r}}} f = \begin{bmatrix} \frac{\partial f_1}{\partial x}(\overrightarrow{\mathbf{r}}) & \frac{\partial f_1}{\partial y}(\overrightarrow{\mathbf{r}}) & \frac{\partial f_1}{\partial z}(\overrightarrow{\mathbf{r}}) \\ \frac{\partial f_2}{\partial x}(\overrightarrow{\mathbf{r}}) & \frac{\partial f_2}{\partial y}(\overrightarrow{\mathbf{r}}) & \frac{\partial f_2}{\partial z}(\overrightarrow{\mathbf{r}}) \end{bmatrix} = \begin{bmatrix} y & x + z & y \\ \frac{1}{x} & \frac{1}{y} & 0 \end{bmatrix}.$$

Each of the column vectors of $\mathscr{J}_{\overrightarrow{\mathbf{r}}} f$ is expressed in the standard basis of $\mathbb{R}^2$. To express them in the basis $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ of $\mathbb{R}^2$ we simply write

$$\begin{bmatrix} y \\ \frac{1}{x} \end{bmatrix} = (y - \frac{1}{x}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{x} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} x + z \\ \frac{1}{y} \end{bmatrix} = (x + z - \frac{1}{y}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{y} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = y \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The desired matrix representation is

$$A_{\overrightarrow{r}} = \begin{bmatrix} y - \dfrac{1}{x} & x + z - \dfrac{1}{y} & y \\ \dfrac{1}{x} & \dfrac{1}{y} & 0 \end{bmatrix}.$$

To compute $\mathscr{D}_{(1,2,3)}f\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, we will use the Jacobi matrix, which will give us the result in the standard basis of $\mathbb{R}^2$.

$$\mathscr{J}_{(1,2,3)}\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 1 & \dfrac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 40 \\ \dfrac{7}{2} \end{bmatrix}.$$

181 Example Let $f(\rho, \theta, z) = (\rho \cos \theta, \rho \sin \theta, z)$ be the function which changes from cylindrical co-ördinates to Cartesian co-ördinates. We have

$$\mathscr{J}_{\overrightarrow{(\rho,\theta,z)}} f = \begin{bmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

182 Example Let $f(\rho, \phi, \theta) = (\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi)$ be the function which changes from spherical co-ördinates to Cartesian co-ördinates. We have

$$\mathscr{J}_{\overrightarrow{(\rho,\phi,\theta)}} f = \begin{bmatrix} \cos\theta\sin\phi & \rho\cos\theta\cos\phi & -\rho\sin\phi\sin\theta \\ \sin\theta\sin\phi & \rho\sin\theta\cos\phi & \rho\cos\theta\sin\phi \\ \cos\phi & -\rho\sin\phi & 0 \end{bmatrix}.$$

The Jacobi matrix provides a convenient computational tool to compute the derivative of a function at a point. Thus differentiability at a point implies that the partial derivatives of the function exist at the point. The converse, however, is not true.

183 Example Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} y & \text{if} \quad x = 0, \\ x & \text{if} \quad y = 0, \\ 1 & \text{if} \quad xy \neq 0. \end{cases}$$

Observe that f is not continuous at $(0, 0)$ ($f(0, 0) = 0$ but $f(x, y) = 1$ for values arbitrarily close to $(0, 0)$), and hence, it is not differentiable there. We have however, $\dfrac{\partial f}{\partial x_1}(\mathbf{0}) = \dfrac{\partial f}{\partial x_2}(\mathbf{0}) = 1$. Thus even if both partial derivatives exist at $(0, 0)$ is no guarantee that the function will be differentiable at $(0, 0)$. You should also notice that both partial derivatives are not continuous at $(0, 0)$.

We have, however, the following.

184 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and let $f : \mathbb{R}^n \to \mathbb{R}^m$. Put $f = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_m \end{bmatrix}$. If each of the partial derivatives $\mathscr{D}_j f_i$ exists and is continuous on A, then f is differentiable on A.

Just like in the one-variable case, we have the following rules of differentiation. Let $A \subseteq \mathbb{R}^n, B \subseteq \mathbb{R}^m$ be open sets $f, g : A \to \mathbb{R}^m, \alpha \in \mathbb{R}$, be differentiable on A, $h : B \to \mathbb{R}^l$ be differentiable on B, and $f(A) \subseteq B$. Then we have

- **Addition Rule:** $\mathscr{D}_{\overrightarrow{\mathbf{x}}}((f + \alpha g)) = \mathscr{D}_{\overrightarrow{\mathbf{x}}}(f) + \alpha \mathscr{D}_{\overrightarrow{\mathbf{x}}}(g)$.

- **Chain Rule:** $\mathscr{D}_{\overrightarrow{\mathbf{x}}}((h \circ f)) = \left(\mathscr{D}_{\overrightarrow{f(\mathbf{x})}}(h)\right) \circ (\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f))$.

Since composition of linear mappings expressed as matrices is matrix multiplication, the Chain Rule takes the alternative form when applied to the Jacobi matrix.

$$\mathscr{J}_{\overrightarrow{\mathbf{x}}}(h \circ f) \quad = \quad (\mathscr{J}_{\overrightarrow{f(\mathbf{x})}} h)(\mathscr{J}_{\overrightarrow{\mathbf{x}}} f). \tag{2.1}$$

185 Example Let

$$f\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}ue^{v}\\u+v\\uv\end{bmatrix},$$

$$h(\overrightarrow{\mathbf{r}})=\begin{bmatrix}x^2+y\\y+z\end{bmatrix}.$$

Find $\mathscr{J}_{\overrightarrow{\mathbf{r}}}(f\circ h)$.

Solution: We have

$$\mathscr{J}_{\overrightarrow{(u,v)}}f=\begin{bmatrix}e^{v} & ue^{v}\\1 & 1\\v & u\end{bmatrix},$$

and

$$\mathscr{J}_{\overrightarrow{\mathbf{r}}}h=\begin{bmatrix}2x & 1 & 0\\0 & 1 & 1\end{bmatrix}.$$

Observe also that

$$\mathscr{J}_{\overrightarrow{h(\mathbf{r})}}f=\begin{bmatrix}e^{y+z} & (x^2+y)e^{y+z}\\1 & 1\\y+z & x^2+y\end{bmatrix}.$$

Hence

$$\begin{aligned}\mathscr{J}_{\overrightarrow{\mathbf{r}}}(f\circ h) &= (\mathscr{J}_{\overrightarrow{h(\mathbf{r})}}f)(\mathscr{J}_{\overrightarrow{\mathbf{r}}}h)\\ &= \begin{bmatrix}e^{y+z} & (x^2+y)e^{y+z}\\1 & 1\\y+z & x^2+y\end{bmatrix}\begin{bmatrix}2x & 1 & 0\\0 & 1 & 1\end{bmatrix}\\ &= \begin{bmatrix}2xe^{y+z} & (1+x^2+y)e^{y+z} & (x^2+y)e^{y+z}\\2x & 2 & 1\\2xy+2xz & x^2+2y+z & x^2+y\end{bmatrix}.\end{aligned}$$

186 Example Let

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f \begin{bmatrix} u \\ v \end{bmatrix} = u^2 + e^v,$$

$$u, v : \mathbb{R}^3 \to \mathbb{R} \quad u(\overrightarrow{\mathbf{r}}) = xz,\ v(\overrightarrow{\mathbf{r}}) = y + z.$$

Put $h(\overrightarrow{\mathbf{r}}) = f \begin{bmatrix} u(x,y,z) \\ v(x,y,z) \end{bmatrix}$. Find $\dfrac{\partial h}{\partial x_i}(\mathbf{r})$ for $i = 1, 2, 3$.

Solution: Put $g : \mathbb{R}^3 \to \mathbb{R}^2, g(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} u(\overrightarrow{\mathbf{r}}) \\ v(\overrightarrow{\mathbf{r}}) \end{bmatrix} = \begin{bmatrix} xz \\ y + z \end{bmatrix}$. Observe that $h = f \circ g$. Now,

$$\mathscr{J}_{\overrightarrow{\mathbf{r}}}\, g = \begin{bmatrix} z & 0 & x \\ 0 & 1 & 1 \end{bmatrix},$$

$$\mathscr{J}_{\overrightarrow{(u,v)}}\, f = \begin{bmatrix} 2u & e^v \end{bmatrix},$$

$$\mathscr{J}_{\overrightarrow{g(\mathbf{r})}}\, f = \begin{bmatrix} 2xz & e^{y+z} \end{bmatrix}.$$

Thus

$$\begin{aligned} \begin{bmatrix} \dfrac{\partial h}{\partial x_1}(\mathbf{r}) & \dfrac{\partial h}{\partial x_2}(\mathbf{r}) & \dfrac{\partial h}{\partial x_3}(\mathbf{r}) \end{bmatrix} &= \mathscr{J}_{\overrightarrow{\mathbf{r}}}\, h \\ &= (\mathscr{J}_{\overrightarrow{g(\mathbf{r})}}\, f)(\mathscr{J}_{\overrightarrow{\mathbf{r}}}\, g) \\ &= \begin{bmatrix} 2xz & e^{y+z} \end{bmatrix} \begin{bmatrix} z & 0 & x \\ 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2xz^2 & e^{y+z} & 2x^2z + e^{y+z} \end{bmatrix} \end{aligned}.$$

Equating components, we obtain

$$\frac{\partial h}{\partial x_1}(\mathbf{r}) = 2xz^2,$$

$$\frac{\partial h}{\partial x_2}(\mathbf{r}) = e^{y+z},$$

$$\frac{\partial h}{\partial x_3}(\mathbf{r}) = 2x^2z + e^{y+z}.$$

2.4 Gradients and Directional Derivatives

A function

$$f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m \\ \overrightarrow{\mathbf{x}} & \mapsto & f(\overrightarrow{\mathbf{x}}) \end{array}$$

is called a *vector field.* If $m = 1$, it is called a *scalar field.*

187 Definition Let

$$f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ \overrightarrow{\mathbf{x}} & \mapsto & f(\overrightarrow{\mathbf{x}}) \end{array}$$

be a scalar field. The *gradient* of f is the vector defined and denoted by

$$\nabla f(\overrightarrow{\mathbf{x}}) = (\mathscr{J}_{\overrightarrow{\mathbf{x}}} f)^T = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) \\ \frac{\partial f}{\partial x_2}(\mathbf{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

The *gradient operator* is the operator

$$\nabla = \begin{bmatrix} \mathscr{D}_1 \\ \mathscr{D}_2 \\ \vdots \\ \mathscr{D}_n \end{bmatrix}.$$

188 Theorem Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $K \in \mathbb{R}$ be a constant. Then $\nabla f(\overrightarrow{\mathbf{x}})$ is orthogonal to the surface implicitly defined by $f(\overrightarrow{\mathbf{x}}) = K$.

Proof Let

$$\overrightarrow{\mathbf{c}} : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^n \\ t & \mapsto & \overrightarrow{\mathbf{c}}(t) \end{array}$$

be a curve lying on this surface. Choose t_0 so that $\overrightarrow{\mathbf{c}}(t_0) = \overrightarrow{\mathbf{x}}$. Then

$$(f \circ \overrightarrow{\mathbf{c}})(t_0) = f(\overrightarrow{\mathbf{c}}(t)) = K,$$

and using the chain rule

$$(\mathscr{J}_{\overrightarrow{\mathbf{c}}(t_0)}f)(\mathscr{J}_{\overrightarrow{t_0}}\overrightarrow{\mathbf{c}}) = 0,$$

which translates to

$$(\nabla f(\overrightarrow{\mathbf{x}}))\bullet(\overrightarrow{\mathbf{c}}'(t_0)) = 0.$$

Since $\overrightarrow{\mathbf{c}}'(t_0)$ is tangent to the surface and its dot product with $\nabla f(\overrightarrow{\mathbf{x}})$ is 0, we conclude that $\nabla f(\overrightarrow{\mathbf{x}})$ is normal to the surface. ❑

Let θ be the angle between $\nabla f(\overrightarrow{\mathbf{x}})$ and $\overrightarrow{\mathbf{c}}'(t_0)$. Since

$$|(\nabla f(\overrightarrow{\mathbf{x}}))\bullet(\overrightarrow{\mathbf{c}}'(t_0))| = ||\nabla f(\overrightarrow{\mathbf{x}})||||\overrightarrow{\mathbf{c}}'(t_0)|| \cos\theta,$$

$\nabla f(\overrightarrow{\mathbf{x}})$ *is the direction in which* f *is changing the fastest.*

189 Example Find a unit vector normal to the surface $x^3 + y^3 + z = 4$ at the point $(1, 1, 2)$.

Solution: Here $f(x, y, z) = x^3 + y^3 + z - 4$ has gradient

$$\nabla f(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} 3x^2 \\ 3y^2 \\ 1 \end{bmatrix}$$

which at $(1, 1, 2)$ is $\begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$. Normalising this vector we obtain

$$\begin{bmatrix} \dfrac{3}{\sqrt{19}} \\ \dfrac{3}{\sqrt{19}} \\ \dfrac{1}{\sqrt{19}} \end{bmatrix}.$$

190 Example Find the direction of the greatest rate of increase of $f(x, y, z) = xye^z$ at the point $(2, 1, 2)$.

Solution: The direction is that of the gradient vector. Here

$$\nabla f(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} ye^z \\ xe^z \\ xye^z \end{bmatrix}$$

which at $(2, 1, 2)$ becomes $\begin{bmatrix} e^2 \\ 2e^2 \\ 2e^2 \end{bmatrix}$. Normalising this vector we obtain

$$\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

191 Example Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = x + y^2 - z^2.$$

Find the equation of the tangent plane to f at $(1, 2, 3)$.

Solution: A vector normal to the plane is $\nabla f(1, 2, 3)$. Now

$$\nabla f(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} 1 \\ 2y \\ -2z \end{bmatrix}$$

which is

$$\begin{bmatrix} 1 \\ 4 \\ -6 \end{bmatrix}$$

at $(1, 2, 3)$. The equation of the tangent plane is thus

$$1(x - 1) + 4(y - 2) - 6(z - 3) = 0,$$

or

$$x + 4y - 6z = -9.$$

192 Definition Let

$$f: \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ \overrightarrow{\mathbf{x}} & \mapsto & f(\overrightarrow{\mathbf{x}}) \end{array}$$

be a vector field with

$$f(\overrightarrow{\mathbf{x}}) = \begin{bmatrix} f_1(\overrightarrow{\mathbf{x}}) \\ f_2(\overrightarrow{\mathbf{x}}) \\ \vdots \\ f_n(\overrightarrow{\mathbf{x}}) \end{bmatrix}.$$

The *divergence* of f is defined and denoted by

$$\operatorname{div} f(\mathbf{x}) = \nabla \bullet f(\overrightarrow{\mathbf{x}}) = \frac{\partial f_1}{\partial x_1}(\mathbf{x}) + \frac{\partial f_2}{\partial x_2}(\mathbf{x}) + \cdots + \frac{\partial f_n}{\partial x_n}(\mathbf{x}).$$

193 Example If $f(x, y, z) = (x^2, y^2, ye^{z^2})$ then

$$\operatorname{div} f(\mathbf{x}) = 2x + 2y + 2yze^{z^2}.$$

194 Definition Let $g_k : \mathbb{R}^n \to \mathbb{R}^n$, $1 \leq k \leq n-2$ be vector fields with $g_i = (g_{i1}, g_{i2}, \ldots, g_{in})$. Then the *curl* of $(g_1, g_2, \ldots, g_{n-2})$

$$\operatorname{curl}(g_1, g_2, \ldots, g_{n-2})(\mathbf{x}) = \det \begin{bmatrix} \overrightarrow{\mathbf{e}}_1 & \overrightarrow{\mathbf{e}}_2 & \cdots & \overrightarrow{\mathbf{e}}_n \\ \mathscr{D}_1 & \mathscr{D}_2 & \cdots & \mathscr{D}_n \\ g_{11}(\overrightarrow{\mathbf{x}}) & g_{12}(\overrightarrow{\mathbf{x}}) & \cdots & g_{1n}(\overrightarrow{\mathbf{x}}) \\ g_{21}(\overrightarrow{\mathbf{x}}) & g_{22}(\overrightarrow{\mathbf{x}}) & \ldots & g_{2n}(\overrightarrow{\mathbf{x}}) \\ \vdots & \vdots & \vdots & \vdots \\ g_{(n-2)1}(\overrightarrow{\mathbf{x}}) & g_{(n-2)2}(\overrightarrow{\mathbf{x}}) & \ldots & g_{(n-2)n}(\overrightarrow{\mathbf{x}}) \end{bmatrix}.$$

195 Example If $f(x, y, z) = (x^2, y^2, ye^{z^2})$ then

$$\operatorname{curl} f(\mathbf{r}) = \det \begin{bmatrix} \overrightarrow{\mathbf{i}} & \overrightarrow{\mathbf{j}} & \overrightarrow{\mathbf{k}} \\ \mathscr{D}_1 & \mathscr{D}_2 & \mathscr{D}_3 \\ x^2 & y^2 & ye^{z^2} \end{bmatrix} = (e^{z^2})\overrightarrow{\mathbf{i}}.$$

196 Example If $f(x, y, z, w) = (e^{xyz}, 0, 0, w^2)$, $g(x, y, z, w) = (0, 0, z, 0)$ then

$$\operatorname{curl}(f, g) \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \det \begin{bmatrix} \overrightarrow{\mathbf{e}}_1 & \overrightarrow{\mathbf{e}}_2 & \overrightarrow{\mathbf{e}}_3 & \overrightarrow{\mathbf{e}}_4 \\ \mathscr{D}_1 & \mathscr{D}_2 & \mathscr{D}_3 & \mathscr{D}_4 \\ e^{xyz} & 0 & 0 & w^2 \\ 0 & 0 & z & 0 \end{bmatrix} = (xz^2e^{xyz})\overrightarrow{\mathbf{e}}_4.$$

197 Definition Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $\overrightarrow{\mathbf{v}} \in \mathbb{R}^n \setminus \{\overrightarrow{\mathbf{0}}\}$ be such that $\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{v}} \in A$ for sufficiently small $t \in \mathbb{R}$. Then the *directional derivative of* f *in the direction of* $\overrightarrow{\mathbf{v}}$ *at the point* $\overrightarrow{\mathbf{x}}$ is defined and denoted by

$$\frac{\partial f}{\partial x_{\overrightarrow{\mathbf{v}}}}(\mathbf{x}) = \lim_{t \to 0} \frac{f(\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{v}}) - f(\overrightarrow{\mathbf{x}})}{t}.$$

Some authors require that the vector $\overrightarrow{\mathbf{v}}$ *in definition 197 be a unit vector.*

198 Theorem Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $\overrightarrow{\mathbf{v}} \in \mathbb{R}^n \setminus \{\overrightarrow{\mathbf{0}}\}$ be such that $\overrightarrow{\mathbf{x}} + t\overrightarrow{\mathbf{v}} \in A$ for sufficiently small $t \in \mathbb{R}$. Then the *directional derivative of* f *in the direction of* $\overrightarrow{\mathbf{v}}$ *at the point* $\overrightarrow{\mathbf{x}}$ is given by

$$\nabla f(\overrightarrow{\mathbf{x}}) \bullet \overrightarrow{\mathbf{v}}.$$

199 Example Find the directional derivative of $f(x, y, z) = x^3 + y^3 - z^2$ in the direction of $\overrightarrow{\mathbf{v}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution: We have

$$\nabla f(\overrightarrow{\mathbf{r}}) = \begin{bmatrix} 3x^2 \\ 3y^2 \\ -2z \end{bmatrix}$$

and so

$$(\nabla f(\overrightarrow{\mathbf{r}})) \bullet \overrightarrow{\mathbf{v}} = 3x^2 + 6y^2 - 6z.$$

2.5 Extrema

We now turn to the problem of finding maxima and minima for vector functions. As in the one-variable case, the derivative will provide us with information about the extrema, and the "second derivative" will provide us with information about the nature of these extreme points.

To define an analogue for the second derivative, let us consider the following. Let $A \subset \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ be differentiable on A. We know that for fixed $\overrightarrow{\mathbf{x}}_0 \in A$, $\mathscr{D}_{\overrightarrow{\mathbf{x}_0}}(f)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. This means that we have a function

$$T : \begin{array}{ccc} A & \to & \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) \\ \overrightarrow{\mathbf{x}} & \mapsto & \mathscr{D}_{\overrightarrow{\mathbf{x}}}(f) \end{array},$$

where $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ denotes the space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. Hence, if we differentiate T at $\overrightarrow{\mathbf{x}}_0$ again, we obtain a linear transformation $\mathscr{D}_{\overrightarrow{\mathbf{x}_0}}(T) = \mathscr{D}_{\overrightarrow{\mathbf{x}_0}}(\mathscr{D}_{\overrightarrow{\mathbf{x}_0}}(f)) = \mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)$ from $\mathbb{R}^n$ to $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. Hence, given $\overrightarrow{\mathbf{x}_1} \in \mathbb{R}^n$, we have $\mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)(\overrightarrow{\mathbf{x}_1}) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. Again, this means that given $\overrightarrow{\mathbf{x}}_2 \in \mathbb{R}^n$, $\mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)(\overrightarrow{\mathbf{x}_1}))(\overrightarrow{\mathbf{x}}_2) \in \mathbb{R}^m$. Thus the function

$$B_{\overrightarrow{\mathbf{x}_0}} : \begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) \\ (\overrightarrow{\mathbf{x}_1}, \overrightarrow{\mathbf{x}_2}) & \mapsto & \mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)(\overrightarrow{\mathbf{x}_1}, \overrightarrow{\mathbf{x}_2}) \end{array}$$

is well defined, and linear in each variable $\overrightarrow{\mathbf{x}_1}$ and $\overrightarrow{\mathbf{x}_2}$, that is, it is a *bilinear* function.

Just as the Jacobi matrix was a handy tool for finding a matrix representation of $\mathscr{D}_{\overrightarrow{\mathbf{x}}}(f)$ in the natural bases, when f maps into $\mathbb{R}$, we have the following analogue representation of the second derivative.

200 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be twice differentiable on A. Then the matrix of $\mathscr{D}^2_{\overrightarrow{\mathbf{x}}}(f) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with respect to the standard basis is given by the *Hessian matrix:*

$$\mathscr{H}_{\overrightarrow{\mathbf{x}}} f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\overrightarrow{\mathbf{x}}) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\overrightarrow{\mathbf{x}}) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\overrightarrow{\mathbf{x}}) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\overrightarrow{\mathbf{x}}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_2}(\overrightarrow{\mathbf{x}}) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\overrightarrow{\mathbf{x}}) \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\overrightarrow{\mathbf{x}}) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(\overrightarrow{\mathbf{x}}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\overrightarrow{\mathbf{x}}) \end{bmatrix}$$

201 Example Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$f(\overrightarrow{\mathbf{r}}) = f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = xy^2z^3.$$

Then

$$\mathscr{H}_{\overrightarrow{\mathbf{r}}} f = \begin{bmatrix} 0 & 2yz^3 & 3y^2z^2 \\ 2yz^3 & 2xz^3 & 6xyz^2 \\ 3y^2z^2 & 6xyz^2 & 6xy^2z \end{bmatrix}$$

From the preceding example, we notice that the mixed partial derivatives $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$, $\frac{\partial^2 f}{\partial x_1 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_1}$, and $\frac{\partial^2 f}{\partial x_2 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_2}$, are equal. This is no coincidence, as guaranteed by the following theorem.

202 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be twice differentiable on A. If $\mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)$ is continuous, then $\mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)$ is symmetric, that is, $\forall(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{x_2}}) \in \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{x_2}}) = \mathscr{D}^2_{\overrightarrow{\mathbf{x}}_0}(f)(\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{x_1}}).$$

We are now ready to study extrema in several variables. The basic theorems resemble those of one-variable calculus. First, we make some analogous definitions.

203 Definition Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$. If there is some open ball $B_{x_0}(\mathbf{r})$ on which $\forall \overrightarrow{\mathbf{x}} \in B_{x_0}(\mathbf{r})$, $f(\overrightarrow{\mathbf{x}}_0) \geq f(\overrightarrow{\mathbf{x}})$, we say that $f(\overrightarrow{\mathbf{x}}_0)$ is a *local maximum* of f. Similarly, if there is some open ball $B_{x_1}(\mathbf{r})$ on which $\forall \overrightarrow{\mathbf{x}} \in B_{x_0}(\mathbf{r}')$, $f(\overrightarrow{\mathbf{x}}_1) \leq f(\overrightarrow{\mathbf{x}})$, we say that $f(\overrightarrow{\mathbf{x}}_1)$ is a *local maximum* of f. A point is called an *extreme point* if it is either a local minimum or local maximum. A point $\overrightarrow{\mathbf{t}}$ is called a *critical point* if f is differentiable at $\overrightarrow{\mathbf{t}}$ and $\mathscr{D}_{\overrightarrow{\mathbf{t}}}(f) = 0$. A critical point which is neither a maxima nor a minima is called a *saddle point*.

204 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be differentiable on A. If $\overrightarrow{\mathbf{x}}_0$ is an extreme point, then $\mathscr{D}_{\overrightarrow{\mathbf{x_0}}}(f) = 0$, that is, $\overrightarrow{\mathbf{x}}_0$ is a critical point. Moreover, if f is twice-differentiable with continuous second derivative and $\overrightarrow{\mathbf{x_0}}$ is a critical point such that $\mathscr{H}_{\overrightarrow{\mathbf{x_0}}} f$ is negative definite, then f has a local maximum at $\overrightarrow{\mathbf{x}}_0$. If $\mathscr{H}_{\overrightarrow{\mathbf{x_0}}} f$ is positive definite, then f has a local minimum at $\overrightarrow{\mathbf{x}}_0$. If $\mathscr{H}_{\overrightarrow{\mathbf{x_0}}} f$ is indefinite, then f has a saddle point. If $\mathscr{H}_{\overrightarrow{\mathbf{x_0}}} f$ is semi-definite (positive or negative), the test is inconclusive.

205 Example Find the critical points of

$$f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2+xy+y^2+2x+3y \end{array}$$

and investigate their nature.

Solution: We have

$$\mathscr{J}_{\vec{r}} f = \begin{bmatrix} 2x+y+2 & x+2y+3 \end{bmatrix},$$

and so to find the critical points we solve

$$2x+y+2=0,$$

$$x+2y+3=0,$$

which yields $x=-\frac{1}{3}, y=-\frac{4}{3}$. Now,

$$\mathscr{H}_{\vec{r}} f = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

which is positive definite, since $\Delta_1 = 2 > 0$ and $\Delta_2 = \det \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = 3 > 0$. Thus $\mathbf{x_0} = (-\frac{1}{3}, -\frac{4}{3})$ is a relative minimum and we have

$$-\frac{7}{3} = f(-\frac{1}{3}, -\frac{4}{3}) \leq f(x,y) = x^2+xy+y^2+2x+3y.$$

206 Example Find the extrema of

$$f: \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & x^2+y^2+3z^2-xy+2xz+yz \end{array}.$$

Solution: We have

$$\mathscr{J}_{\vec{r}} f = \begin{bmatrix} 2x-y+2z & 2y-x+z & 6z+2x+y \end{bmatrix},$$

which vanishes when $x = y = z = 0$. Now,

$$\mathscr{H}_{\vec{r}} f = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 6 \end{bmatrix},$$

which is positive definite, since $\Delta_1 = 2 > 0$, $\Delta_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 3 > 0$, and $\Delta_3 = \det \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 6 \end{bmatrix} = 4 > 0$. Thus f has a relative minimum at $(0, 0, 0)$ and

$$0 = f(0, 0, 0) \leq f(x, y, z) = x^2 + y^2 + 3z^2 - xy + 2xz + yz.$$

207 Example Find the extrema of $f : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \mapsto & x^2 + y^2 + z^2 + xyz \end{array}$.

Solution: We have

$$\mathscr{J}_{\vec{r}} f = \begin{bmatrix} 2x + yz & 2y + xz & 2z + xy \end{bmatrix},$$

and

$$\mathscr{H}_{\vec{r}} f = \begin{bmatrix} 2 & z & y \\ z & 2 & x \\ y & x & 2 \end{bmatrix}.$$

We see that $\Delta_1(x, y, z) = 2$, $\Delta_2(x, y, z) = \det \begin{bmatrix} 2 & z \\ z & 2 \end{bmatrix} = 4 - z^2$ and $\Delta_3(x, y, z) = \det \mathscr{H}_{\vec{r}} f = 8 - 2x^2 - 2y^2 - 2z^2 + 2xyz$.

If $\mathscr{J}_{\vec{r}} f = (0\ 0\ 0)$ then we must have

$$2x = -yz,$$

$$2y = -xz,$$

$$2z = -xy,$$

and upon multiplication of the three equations,

$$8xyz = -x^2y^2z^2,$$

that is,

$$xyz(xyz+8)=0.$$

Clearly, if $xyz=0$, then we must have at least one of the variables equalling 0, in which case, by virtue of the original three equations, all equal 0. Thus $(0,0,0)$ is a critical point. If $xyz=-8$, then none of the variables is 0, and solving for x, say, we must have $x=-\frac{8}{yz}$, and substituting this into $2x+yz=0$ we gather $(yz)^2=16$, meaning that either $yz=4$, in which case $x=-2$, or $zy=-4$, in which case $x=2$. It is easy to see then that either exactly one of the variables is negative, or all three are negative. The other critical points are therefore $(-2,2,2)$, $(2,-2,2)$, $(2,2,-2)$, and $(-2,-2,-2)$.

At $(0,0,0)$, $\Delta_1(0,0,0)=2>0, \Delta_2(0,0,0)=4>0, \Delta_1(0,0,0)=8>0$, and thus $(0,0,0)$ is a minimum point. If $x^2=y^2=z^2=4, xyz=-8$, then $\Delta_2(x,y,z)=0, \Delta_3=-32$, so these points are saddle points.

208 Example Find the extrema of $f: \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & x^2y+y^2z+2x-z \end{array}$.

Solution: We have

$$\mathscr{J}_{\overrightarrow{\mathbf{r}}}f=\begin{bmatrix}2xy+2 & x^2+2yz & y^2-1\end{bmatrix},$$

and

$$\mathscr{H}_{\overrightarrow{\mathbf{r}}}f=\begin{bmatrix}2y & 2x & 0\\ 2x & 2z & 2y\\ 0 & 2y & 0\end{bmatrix}.$$

We see that $\Delta_1(x,y,z)=2y, \Delta_2(x,y,z)=\det\begin{bmatrix}2y & 2x\\ 2x & 2z\end{bmatrix}=4yz-4x^2$ and $\Delta_3(x,y,z)=\det\mathscr{H}_{\overrightarrow{\mathbf{r}}}f=-8y^3$.

If $\mathscr{J}_{\overrightarrow{\mathbf{r}}}f=(0\ 0\ 0)$ then we must have

$$xy=-1,$$

$$x^2=-2yz,$$

$$y=\pm 1,$$

and hence $(1, -1, \frac{1}{2})$, and $(-1, 1, -\frac{1}{2})$ are the critical points. Now, $\Delta_1(1, -1, \frac{1}{2}) = -2$, $\Delta_2(1, -1, \frac{1}{2}) = -6$, and $\Delta_3(1, -1, \frac{1}{2}) = 8$. Thus $(1, -1, \frac{1}{2})$ is a saddle point. Similarly, $\Delta_1(-1, 1, -\frac{1}{2}) = 2$, $\Delta_2(-1, 1, -\frac{1}{2}) = -6$, and $\Delta_3(-1, 1, -\frac{1}{2}) = -8$, shewing that $(1, -1, \frac{1}{2})$ is also a saddle point.

209 Example Determine the nature of the critical points of

$$f(x, y, z) = 4xyz - x^4 - y^4 - z^4.$$

Solution: We find

$$\nabla f(x, y, z) = \begin{bmatrix} 4yz - 4x^3 \\ 4xz - 4y^3 \\ 4xy - 4z^3 \end{bmatrix}.$$

Assume $\nabla f(x, y, z) = \mathbf{0}$. Then

$$4yz = 4x^3,\ 4xz = 4y^3,\ 4xy = 4z^3 \implies xyz = x^4 = y^4 = z^4.$$

Thus $xyz \geq 0$, and if one of the variables is 0 so are the other two. Thus $(0, 0, 0)$ is the only critical point with at least one of the variables 0. Assume now that $xyz \neq 0$. Then

$$(xyz)^3 = x^4y^4z^4 = (xyz)^4 \implies xyz = 1 \implies yz = \frac{1}{x} \implies x^4 = 1 \implies x = \pm 1.$$

Similarly, $y = \pm 1, z = \pm 1$. Since $xyz = 1$, exactly two of the variables can be negative. Thus we find the following critical points:

$$(0, 0, 0), (1, 1, 1), (-1, -1, 1), (-1, 1, -1), (1, -1, -1).$$

The Hessian is

$$\mathscr{H}_{\vec{x}} f = \begin{bmatrix} -12x^2 & 4z & 4y \\ 4z & -12y^2 & 4x \\ 4y & 4x & -12z^2 \end{bmatrix}.$$

If $1 = xyz = x^2 = y^2 = z^2$, we have $\Delta_1 = -12x^2 = -12 < 0$, $\Delta_2 = 144x^2y^2 - 16z^2 = 144 - 16 = 128 > 0$, and

$$\begin{aligned} \Delta_3 &= -1728x^2y^2z^2 + 192x^4 + 192z^4 + 128zyx + 192y^4 \\ &= -1728 + 192 + 192 + 128 + 192 \\ &= -1024 \\ &< 0. \end{aligned}$$

This means that for $xyz \neq 0$ the Hessian is negative definite and the function has a local maximum at each of the four points $(1,1,1),(-1,-1,1)$, $(-1,1,-1)$, $(1,-1,-1)$. Observe that at these critical points $f = 1$. Now $f(0,0,0) = 0$ and $f(-1,1,1) = -7$.

210 Example Determine the nature of the critical points of

$$g(x,y,z) = xyze^{-x^2-y^2-z^2}.$$

Solution: To facilitate differentiation observe that $g(x,y,z) = (xe^{-x^2})(ye^{-y^2})(ze^{-z^2})$. Now

$$\nabla g(x,y,z) = \begin{bmatrix} (1-2x^2)(yz)(e^{-x^2})(e^{-y^2})(e^{-z^2}) \\ (1-2y^2)(xz)(e^{-x^2})(e^{-y^2})(e^{-z^2}) \\ (1-2z^2)(xy)(e^{-x^2})(e^{-y^2})(e^{-z^2}) \end{bmatrix}.$$

The function is 0 if any of the variables is 0. Since the function clearly assumes positive and negative values, we can discard any point with a 0. If $\nabla(x,y,z) = \mathbf{0}$, then $x = \pm\frac{1}{\sqrt{2}}$; $y = \pm\frac{1}{\sqrt{2}}$ $z = \pm\frac{1}{\sqrt{2}}$. We find

$$\mathscr{H}_{\vec{\mathbf{x}}}\, g = t(x,y,z) \begin{bmatrix} (4x^3-6x)(yz) & (1-2x^2)(1-2y^2)z & (1-2x^2)(1-2z^2)y \\ (1-2y^2)(1-2x^2)z & (4y^3-6y)(xz) & (1-2y^2)(1-2z^2)x \\ (1-2z^2)(1-2x^2)y & (1-2z^2)(1-2y^2)x & (4z^3-6z)(xy) \end{bmatrix},$$

with $t(x,y,z) = (e^{-x^2})(e^{-y^2})(e^{-z^2})$. Since at the critical points we have $1-2x^2 = 1-2y^2 = 1-2z^2 = 0$, the Hessian reduces to

$$\mathscr{H}_{\vec{\mathbf{x}}}\, g = (e^{-3/2}) \begin{bmatrix} (4x^3-6x)(yz) & 0 & 0 \\ 0 & (4y^3-6y)(xz) & 0 \\ 0 & 0 & (4z^3-6z)(xy) \end{bmatrix}.$$

We have

$$\Delta_1 = (4x^3-6x)(yz)$$

$$\Delta_2 = (4x^3-6x)(4y^3-6y)(xyz^2)$$

$$\Delta_3 = (4x^3-6x)(4y^3-6y)(4z^3-6z)(x^2y^2z^2).$$

Also,

$$4\left(\frac{1}{\sqrt{2}}\right)^3 - 6\left(\frac{1}{\sqrt{2}}\right) = -2\sqrt{2} < 0, \quad 4\left(-\frac{1}{\sqrt{2}}\right)^3 - 6\left(-\frac{1}{\sqrt{2}}\right) = 2\sqrt{2} > 0.$$

This means that if an even number of the variables is negative (0 or 2), then we the Hessian is negative definite, and if an odd numbers of the variables is positive (1 or 3), the Hessian is positive definite. We conclude that we have local maxima at

$$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$$

and local minima at

$$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}).$$

2.6 Lagrange Multipliers

In some situations we wish to optimise a function given a set of constraints. For such cases, we have the following.

211 Theorem Let $A \subseteq \mathbb{R}^n$ and let $f : A \to \mathbb{R}$, $g : A \to \mathbb{R}$ be functions whose respective derivatives are continuous. Let $g(\mathbf{x}_0) = c_0$ and let $S = g^{-1}(c_0)$ be the level set for g with value c_0, and assume $\nabla g(\mathbf{x}_0) \neq 0$. If the restriction of f to S has an extreme point at $\mathbf{x}_0$, then $\exists \lambda \in \mathbb{R}$ such that

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0).$$

The above theorem only locates extrema, it does not say anything concerning the nature of the critical points found.

212 Example Optimise $f : \mathbb{R}^2 \to \mathbb{R}, f(x, y) = x^2 - y^2$ given that $x^2 + y^2 = 1$.

Solution: Let $g(x, y) = x^2 + y^2 - 1$. We solve

$$\nabla f \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \nabla g \begin{bmatrix} x \\ y \end{bmatrix}$$

for x, y, λ. This requires

$$\begin{bmatrix} 2x \\ -2y \end{bmatrix} = \begin{bmatrix} 2x\lambda \\ 2y\lambda \end{bmatrix}.$$

From $2x = 2x\lambda$ we get either $x = 0$ or $\lambda = 1$. If $x = 0$ then $y = \pm 1$ and $\lambda = -1$. If $\lambda = 1$, then $y = 0, x = \pm 1$. Thus the potential critical points are $(\pm 1, 0)$ and $(0, \pm 1)$. If $x^2 + y^2 = 1$ then

$$f(x, y) = x^2 - (1 - x^2) = 2x^2 - 1 \geq -1,$$

and

$$f(x, y) = 1 - y^2 - y^2 = 1 - 2y^2 \leq 1.$$

Thus $(\pm 1, 0)$ are maximum points and $(0, \pm 1)$ are minimum points.

213 Example Find the maximum and the minimum points of $f(x, y) = 4x + 3y$, subject to the constraint $x^2 + 4y^2 = 4$, using Lagrange multipliers.

Solution: Putting $g(x, y) = x^2 + 4y^2 - 4$ we have

$$\nabla f(x, y) = \lambda \nabla g(x, y) \implies \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 8y \end{bmatrix}.$$

Thus $4 = 2\lambda x, \;\; 3 = 8\lambda y$. Clearly then $\lambda \neq 0$. Upon division we find $\dfrac{x}{y} = \dfrac{16}{3}$. Hence

$$x^2 + 4y^2 = 4 \implies \frac{256}{9}y^2 + 4y^2 = 4 \implies y = \pm\frac{3}{\sqrt{73}}, x = \pm\frac{16}{\sqrt{73}.}$$

The maximum is clearly then

$$4\left(\frac{16}{\sqrt{73}}\right) + 3\left(\frac{3}{\sqrt{73}}\right) = \sqrt{73},$$

and the minimum is $-\sqrt{73}$.

214 Example Let $a > 0, b > 0, c > 0$. Determine the maximum and minimum values of $f(x, y, z) = \dfrac{x}{a} + \dfrac{y}{b} + \dfrac{z}{c}$ on the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

Solution: We use Lagrange multipliers. Put $g(x, y, z) = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} - 1$. Then

$$\nabla f(x,y,z) = \lambda \nabla g(x,y,z) \iff \begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix} = \lambda \begin{bmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{bmatrix}.$$

It follows that $\lambda \neq 0$. Hence $x = \dfrac{a}{2\lambda}, y = \dfrac{b}{2\lambda}, z = \dfrac{c}{2\lambda}$. Since $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$, we deduce $\dfrac{3}{4\lambda^2} = 1$ or $\lambda = \pm\dfrac{\sqrt{3}}{2}$. Since a, b, c are positive, f will have a maximum when all x, y, z are positive and a minimum when all x, y, z are negative. Thus the maximum is when

$$x = \frac{a}{\sqrt{3}}, y = \frac{b}{\sqrt{3}}, z = \frac{c}{\sqrt{3}},$$

and

$$f(x,y,z) \leq \frac{3}{\sqrt{3}}$$

and the minimum is when

$$x = -\frac{a}{\sqrt{3}}, y = -\frac{b}{\sqrt{3}}, z = -\frac{c}{\sqrt{3}},$$

and

$$f(x,y,z) \geq -\frac{3}{\sqrt{3}}.$$

215 Example Let $a > 0, b > 0, p > 1$. Maximise $f(x,y) = ax + by$ subject to the constraint $x^p + y^p = 1$.

Solution: Put $g(x,y) = x^p + y^p - 1$. We need $a = p\lambda x^{p-1}$ and $b = p\lambda y^{p-1}$. Clearly then , $\lambda \neq 0$. We then have

$$x = \left(\frac{a}{\lambda p}\right)^{1/(p-1)}, \quad y = \left(\frac{b}{\lambda p}\right)^{1/(p-1)}.$$

Thus

$$1 = x^p + y^p = \left(\frac{a}{\lambda p}\right)^{p/(p-1)} + \left(\frac{b}{\lambda p}\right)^{p/(p-1)},$$

which gives

$$\lambda = \left(\left(\frac{a}{p}\right)^{p/(p-1)} + \left(\frac{b}{p}\right)^{p/(p-1)}\right)^{(p-1)/p}.$$

This gives

$$x = \frac{a^{1/(p-1)}}{(a^{1/(p-1)} + b^{1/(p-1)})^{1/p}}, \quad y = \frac{b^{1/(p-1)}}{(a^{1/(p-1)} + b^{1/(p-1)})^{1/p}}.$$

Since f is non-negative, these points define a maximum for f and so

$$ax + by \leq \frac{a^{p/(p-1)}}{(a^{1/(p-1)} + b^{1/(p-1)})^{1/p}} + \frac{b^{p/(p-1)}}{(a^{1/(p-1)} + b^{1/(p-1)})^{1/p}}.$$

216 Example Find the extrema of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint $(x-1)^2 + (y-2)^2 + (z-3)^2 = 4$.

Solution: Let $g(x, y, z) = (x-1)^2 + (y-2)^2 + (z-3)^2 - 4$. We solve

$$\nabla f \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \nabla g \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

for x, y, λ. This requires

$$\begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} = \begin{bmatrix} 2(x-1)\lambda \\ 2(y-2)\lambda \\ 2(z-3)\lambda \end{bmatrix}.$$

Clearly, $\lambda \neq 1$. This gives $x = \dfrac{-\lambda}{1-\lambda}$, $y = \dfrac{-2\lambda}{1-\lambda}$, and $z = \dfrac{-3\lambda}{1-\lambda}$. Substituting into $(x-1)^2 + (y-2)^2 + (z-3)^2 = 4$, we gather that

$$\left(\frac{-\lambda}{1-\lambda} - 1\right)^2 + \left(\frac{-2\lambda}{1-\lambda} - 2\right)^2 + \left(\frac{-3\lambda}{1-\lambda} - 3\right)^2 = 4,$$

from where

$$\lambda = 1 \pm \frac{\sqrt{14}}{2}.$$

This gives the two points

$$(x, y, z) = \left(1 + \frac{2}{\sqrt{14}}, 2 + \frac{4}{\sqrt{14}}, 3 + \frac{6}{\sqrt{14}}\right)$$

and

$$(x, y, z) = \left(1 - \frac{2}{\sqrt{14}}, 2 - \frac{4}{\sqrt{14}}, 3 - \frac{6}{\sqrt{14}}\right).$$

The first point gives an absolute maximum of $18 + \frac{12\sqrt{14}}{7}$ and the second an absolute minimum of $18 - \frac{12\sqrt{14}}{7}$.

217 Example Optimise $f(x, y, z) = x + y + z$ subject to $x^2 + y^2 = 2$, and $x + z = 1$.

Solution: Put $g(x, y, z) = x^2 + y^2 - 2, h(x, y, z) = x + z - 1$. We must find λ, δ such that

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \delta \nabla h(x, y, z),$$

which translates into

$$1 = 2\lambda x + \delta,$$
$$1 = 2\lambda y,$$
$$1 = \delta,$$

and

$$x^2 + y^2 = 1,$$
$$x + z = 1.$$

We deduce that $x = 0, y = \pm\sqrt{2}, z = 1$. We may shew that $(0, \sqrt{2}, 1)$ yields a maximum and that $(0, -\sqrt{2}, 1)$ yields a minimum.

2.7 Arithmetic Mean-Geometric Mean Inequality

218 Definition Let $a_1, a_2, \cdots a_n$ be n non-negative real numbers. Their *arithmetic mean* or *average* is

$$\frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Their *geometric mean* is

$$(a_1a_2\cdots a_n)^{1/n}.$$

The *Arithmetic-Mean-Geometric-Mean Inequality* (AMGM) asserts that

$$(a_1a_2\cdots a_n)^{1/n} \leq \frac{a_1+a_2+\cdots+a_n}{n},$$

with equality if and only if

$$a_1 = a_2 = \cdots = a_n.$$

We will see that this simple inequality allows us to solve many extremum problems without using calculus. We will give two proofs of AMGM, one using Lagrange Multipliers, and one using one-variable calculus due to George Pólya.

219 Theorem Let $a_1, a_2, \cdots a_n$ be n non-negative real numbers. Then

$$(a_1a_2\cdots a_n)^{1/n} \leq \frac{a_1+a_2+\cdots+a_n}{n},$$

with equality if and only if

$$a_1 = a_2 = \cdots = a_n.$$

Proof (Using Lagrange Multipliers) Let

$$f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n) = x_1x_2\cdots x_n,$$

$$g(\mathbf{x}) = g(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n,$$

with $x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0$ and $x_1 + x_2 + \cdots + x_n = S$. Observe that if any of the $x_i = 0$, then f vanishes. So we may assume that no x_i vanishes. Observe that

$$\forall i\ 1 \leq i \leq n,\ \ \nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) \implies \frac{f(\mathbf{x})}{x_i} = x_1x_2\cdots x_{i-1}x_{i+1}\cdots x_n = \lambda.$$

Since $f(\mathbf{x}) \neq 0$, the equality $\dfrac{f(\mathbf{x})}{x_i} = \lambda$ imposes $x_1 = x_2 = \cdots = x_n$. This means that $x_i = \dfrac{S}{n}$. Since f is non-negative on the stipulated conditions, the extrema at $x_i = \dfrac{S}{n}$ is a maximum, and so

$$x_1 x_2 \cdots x_n \leq \left(\frac{S}{n}\right)^n = \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)^n,$$

which gives Arithmetic-Mean-Geometric-Mean Inequality. ❑

For Pólya's proof we need the following lemma.

220 Lemma For all real numbers x it is verified that

$$x \leq e^{x-1},$$

with equality only when $x = 1$.

Proof Put $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^{x-1} - x$. Clearly $f(1) = e^0 - 1 = 0$. Now,

$$f'(x) = e^{x-1} - 1,$$

$$f''(x) = e^{x-1}.$$

If $f'(x) = 0$ then $e^{x-1} = 1$ implying that $x = 1$. Thus f has a single minimum point at $x = 1$. Thus for all real numbers x

$$0 = f(1) \leq f(x) = e^{x-1} - x,$$

which gives the desired result. ❑

Proof (George Pólya) Put

$$A_k = \frac{na_k}{a_1 + a_2 + \cdots + a_n},$$

and $G_n = a_1 a_2 \cdots a_n$. Observe that

$$A_1 A_2 \cdots A_n = \frac{n^n G_n}{(a_1 + a_2 + \cdots + a_n)^n},$$

and that

$$A_1 + A_2 + \cdots + A_n = n.$$

By the preceding lemma, we have

$$A_1 \leq \exp(A_1 - 1),$$

$$A_2 \leq \exp(A_2 - 1),$$

$$\vdots$$

$$A_n \leq \exp(A_n - 1).$$

Since all the quantities involved are non-negative, we may multiply all these inequalities together, to obtain,

$$A_1 A_2 \cdots A_n \leq \exp(A_1 + A_2 + \cdots + A_n - n).$$

In view of the observations above, the preceding inequality is equivalent to

$$\frac{n^n G_n}{(a_1 + a_2 + \cdots + a_n)^n} \leq \exp(n - n) = e^0 = 1.$$

We deduce that

$$G_n \leq \left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^n,$$

which is equivalent to

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Now, for equality to occur, we need each of the inequalities $A_k \leq \exp(A_k - 1)$ to hold. This occurs, in view of the preceding lemma, if and only if $A_k = 1$, $\forall k$, which translates into $a_1 = a_2 = \ldots = a_n$. This completes the proof. ❑

221 Example Let $f(x) = (a + x)^5 (a - x)^3,\ x \in [-a; a]$. Find the maximum value of f by means of $AMGM$.

Solution: If $x \in [-a; a]$, then $a + x \geq 0$ and $a - x \geq 0$ thus we may use AMGM with $n = 8$, $a_1 = a_2 = \cdots = a_5 = \frac{a+x}{5}$ and $a_6 = a_7 = a_8 = \frac{a-x}{3}$. We then deduce

$$\left(\frac{a+x}{5}\right)^5 \left(\frac{a-x}{3}\right)^3 \leq \left(\frac{5\left(\frac{a+x}{5}\right) + 3\left(\frac{a-x}{3}\right)}{8}\right)^8 = \left(\frac{a}{4}\right)^8,$$

whence

$$f(x) \leq \frac{5^5 3^3 a^8}{4^8},$$

and equality holds if and only if $\frac{a+x}{5} = \frac{a-x}{3}$.

222 Example For any positive integer $n > 1$ we have

$$1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1) < n^n.$$

For, by AMGM,

$$1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1) < \left(\frac{1+3+5+\cdots+(2n-1)}{n}\right)^n = \left(\frac{n^2}{n}\right)^n = n^n.$$

Notice that since the factors are unequal we have strict inequality.

223 Example The sequence $x_n = \left(1 + \frac{1}{n}\right)^n, n = 1, 2, \ldots$ is increasing. For the set of $n + 1$ numbers

$$1, 1 + \frac{1}{n}, 1 + \frac{1}{n}, \ldots, 1 + \frac{1}{n},$$

has arithmetic mean

$$1 + \frac{1}{n+1}$$

and geometric mean

$$\left(1 + \frac{1}{n}\right)^{n/(n+1)}.$$

Therefore

$$1+\frac{1}{n+1} > \left(1+\frac{1}{n}\right)^{n/(n+1)},$$

that is to say

$$\left(1+\frac{1}{n+1}\right)^{n+1} > \left(1+\frac{1}{n}\right)^{n},$$

or

$$x_{n+1} > x_n,$$

giving the result.

224 Example Find the volume of the largest rectangular box with sides parallel to the coordinate planes which can be inscribed in the ellipsoid

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1.$$

Here $a > 0, b > 0, c > 0$.

Solution: Let $2x, 2y, 2z$ be the dimensions of this box. We seek to maximise $8xyz$. Putting $n = 3, x_1 = \frac{x^2}{a^2}, x_2 = \frac{y^2}{b^2}, x_3 = \frac{z^2}{c^2}$, we have

$$\frac{x^2}{a^2}\frac{y^2}{b^2}\frac{z^2}{c^2} = x_1x_2x_3 \leq \left(\frac{x_1+x_2+x_3}{3}\right)^3 = \frac{1}{27}.$$

Thus

$$8xyz \leq \frac{8abc}{3\sqrt{3}}$$

225 Definition Let $a_1 > 0, a_2 > 0, \ldots, a_n > 0$. Their *harmonic mean* is given by

$$\frac{n}{\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_n}}.$$

As a corollary to AMGM we obtain

226 Corollary (Harmonic Mean-Geometric Mean Inequality) Let $b_1 > 0, b_2 > 0, \ldots, b_n > 0$. Then

$$\frac{n}{\frac{1}{b_1}+\frac{1}{b_2}+\cdots+\frac{1}{b_n}} \leq (b_1 b_2 \cdots b_n)^{1/n}.$$

Proof This follows by putting $a_k = \frac{1}{b_k}$ in Theorem 219. For then

$$\left(\frac{1}{b_1}\frac{1}{b_2}\cdots\frac{1}{b_n}\right)^{1/n} \leq \frac{\frac{1}{b_1}+\frac{1}{b_2}+\cdots+\frac{1}{b_n}}{n}.$$

Combining Theorem 219 and Corollary 226, we deduce

227 Corollary (Harmonic Mean-Arithmetic Mean Inequality) Let $b_1 > 0, b_2 > 0, \ldots, b_n > 0$. Then

$$\frac{n}{\frac{1}{b_1}+\frac{1}{b_2}+\cdots+\frac{1}{b_n}} \leq \frac{b_1+b_2+\cdots+b_n}{n}.$$

228 Example Let $a_k > 0$, and $s = a_1 + a_2 + \cdots + a_n$. Prove that

$$\sum_{k=1}^{n} \frac{s}{s-a_k} \geq \frac{n^2}{n-1}$$

and

$$\sum_{k=1}^{n} \frac{a_k}{s-a_k} \geq \frac{n}{n-1}.$$

Solution: Put $b_k = \frac{s}{s-a_k}$. Then

$$\sum_{k=1}^{n} \frac{1}{b_k} = \sum_{k=1}^{n} \frac{s-a_k}{s} = n-1$$

and from Corollary 227,

$$\frac{n}{n-1} \leq \frac{\sum_{k=1}^{n} \frac{s}{s-a_k}}{n},$$

from where the first inequality is proved. Since $\frac{s}{s-a_k} - 1 = \frac{a_k}{s-a_k}$,we have

$$\begin{aligned}\sum_{k=1}^{n} \frac{a_k}{s-a_k} &= \sum_{k=1}^{n} \left(\frac{s}{s-a_k} - 1\right) \\ &= \sum_{k=1}^{n} \left(\frac{s}{s-a_k}\right) - n \\ &\geq \frac{n^2}{n-1} - n \\ &= \frac{n}{n-1}.\end{aligned}$$

Chapter 3

Integration

3.1 Differential Forms

We will now consider integration in several variables. In order to smooth our discussion, we need to consider the concept of differential forms.

229 Definition Consider n variables

$$x_1, x_2, \ldots, x_n$$

in n-dimensional space (used as the names of the axes), and let

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} \in \mathbb{R}^n, \quad 1 \leq j \leq k,$$

be $k \leq n$ vectors in $\mathbb{R}^n$. Moreover, let $\{j_1, j_2, \ldots, j_k\} \subseteq \{1, 2, \ldots, n\}$ be a collection of k sub-indices. An *elementary* k*-differential form (*$k > 1$*) acting on the vectors* $\mathbf{a}_j$, $1 \leq j \leq k$ is defined and denoted by

$$dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k) = \det \begin{bmatrix} a_{j_1 1} & a_{j_1 2} & \cdots & a_{j_1 k} \\ a_{j_2 1} & a_{j_2 2} & \cdots & a_{j_2 k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j_k 1} & a_{j_k 2} & \cdots & a_{j_k k} \end{bmatrix}.$$

In other words, $dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ is the $x_{j_1}x_{j_2}\ldots x_{j_k}$ component of the signed k-volume of a k-parallelotope in $\mathbb{R}^n$ spanned by $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$.

By virtue of being a determinant, the wedge product $\wedge$ of differential forms has the following properties

❶ **anti-commutativity:** $da \wedge db = -db \wedge da$.

❷ **linearity:** $d(a + b) = da + db$.

❸ **scalar homogeneity:** *if* $\lambda \in \mathbb{R}$, *then* $d\lambda a = \lambda da$.

❹ **associativity:** $(da \wedge db) \wedge dc = da \wedge (db \wedge dc)$.[1]

Anti-commutativity yields

$$da \wedge da = 0.$$

230 Example Consider

$$\mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^3.$$

Then

$$\begin{aligned} dx(\mathbf{a}) &= \det(1) = 1, \\ dy(\mathbf{a}) &= \det(0) = 0, \\ dz(\mathbf{a}) &= \det(-1) = -1, \end{aligned}$$

are the (signed) 1-volumes (that is, the length) of the projections of $\mathbf{a}$ onto the co-ördinate axes.

231 Example Consider the vectors in $\mathbb{R}^4$

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}.$$

[1]Notice that associativity does not hold for the wedge product of *vectors*.

Then, for example,

$$dx_1 \wedge dx_2(\mathbf{a}, \mathbf{b}) = d \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \\ 4 & 2 \end{bmatrix} = \det \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} = 2,$$

$$dx_2 \wedge dx_4(\mathbf{a}, \mathbf{b}) = d \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \\ 4 & 2 \end{bmatrix} = \det \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix} = 4,$$

$$dx_3 \wedge dx_1(\mathbf{a}, \mathbf{b}) = d \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \\ 4 & 2 \end{bmatrix} = \det \begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix} = -4.$$

232 Example Consider the vectors in $\mathbb{R}^4$

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then, for example,

$$dx_1 \wedge dx_2(\mathbf{a}, \mathbf{b}, \mathbf{c})$$

is undefined, since for a 2-form we need two vectors. We have, however,

$$dx_1 \wedge dx_2 \wedge dx_4(\mathbf{a}, \mathbf{b}, \mathbf{c}) = d \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 2 & 0 \end{bmatrix} = \det \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 0 \\ 4 & 2 & 0 \end{bmatrix} = 4.$$

233 Example In $\mathbb{R}^3$ we have $dx \wedge dy \wedge dx = 0$, since we have a repeated variable.

234 Example In $\mathbb{R}^3$ we have

$$dx \wedge dz + 5dz \wedge dx + 4dx \wedge dy - dy \wedge dx + 12dx \wedge dx = -4dx \wedge dz + 5dx \wedge dy.$$

In order to avoid redundancy we will make the convention that if a sum of two or more terms have the same differential form up to permutation of the variables, we will simplify the summands and express the other differential forms in terms of the one differential form whose indices appear in increasing order.

235 Example In $\mathbb{R}^5$ we have

$$dx_5 \wedge dx_4 \wedge dx_2 = -dx_4 \wedge dx_5 \wedge dx_2 = dx_4 \wedge dx_2 \wedge dx_5 = -dx_2 \wedge dx_4 \wedge dx_5,$$

$$dx_4 \wedge dx_2 \wedge dx_5 = -dx_2 \wedge dx_4 \wedge dx_5$$

$$dx_2 \wedge dx_5 \wedge dx_4 = -dx_2 \wedge dx_4 \wedge dx_5.$$

Hence we write

$$dx_2 \wedge dx_5 \wedge dx_4 + 5dx_4 \wedge dx_2 \wedge dx_5 + 2dx_5 \wedge dx_4 \wedge dx_2$$

as

$$-8dx_2 \wedge dx_4 \wedge dx_5.$$

236 Definition A 0-differential form in $\mathbb{R}^n$ is simply a differentiable function in $\mathbb{R}^n$.

237 Definition A k-*differential form field in* $\mathbb{R}^n$ is an expression of the form

$$\omega = \sum_{1 \leq j_1 \leq j_2 \leq \cdots \leq j_k \leq n} a_{j_1 j_2 \ldots j_k} dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_k},$$

where the $a_{j_1 j_2 \ldots j_k}$ are differentiable functions in $\mathbb{R}^n$.

238 Example

$$g(x, y, z, w) = x + y^2 + z^3 + w^4$$

is a 0-form in $\mathbb{R}^4$.

239 Example An example of a 1-form field in $\mathbb{R}^3$ is

$$\omega = xdx + y^2dy + xyz^3dz.$$

240 Example An example of a 2-form field in $\mathbb{R}^3$ is

$$\omega = x^2 dx \wedge dy + y^2 dy \wedge dz + dz \wedge dx.$$

241 Example An example of a 3-form field in $\mathbb{R}^3$ is

$$\omega = (x + y + z) dx \wedge dy \wedge dz.$$

242 Theorem If ω is an l-form and ϕ is a k-form then $\omega \wedge \phi$ is a $l + k$-form and

$$\omega \wedge \phi = (-1)^{lk} \phi \wedge \omega.$$

243 Example The product of the 1-form fields in $\mathbb{R}^3$

$$\omega_1 = y dx + x dy,$$

$$\omega_2 = -2x dx + 2y dy,$$

is

$$\omega_1 \wedge \omega_2 = (2x^2 + 2y^2) dx \wedge dy.$$

244 Theorem Let $0 \leq k \leq n$, and let $\wedge^k(\mathbb{R}^n)$ denote the space of k-differential forms in $\mathbb{R}^n$. Then $\wedge^k(\mathbb{R}^n)$ is a vector space of dimension

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

A basis for this space is given by the $\binom{n}{k}$ elementary forms

$$dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_n}, \quad 1 \leq j_1 < j_2 < \cdots < j_k \leq n.$$

245 Example In $\mathbb{R}^3$, $\wedge^1(\mathbb{R}^3)$ has dimension $\binom{3}{1} = 3$, and a basis for $\wedge^1(\mathbb{R}^3)$ is

$$\{dx, \; dy, \; dz\}.$$

246 Example In $\mathbb{R}^3$, $\wedge^2(\mathbb{R}^3)$ has dimension $\binom{3}{2} = 3$, and a basis for $\wedge^2(\mathbb{R}^3)$ is

$$\{dx \wedge dy, \; dx \wedge dz, \; dy \wedge dz\}.$$

247 Example In $\mathbb{R}^3$, $\bigwedge^3(\mathbb{R}^3)$ has dimension $\binom{3}{3} = 1$, and a basis for $\bigwedge^3(\mathbb{R}^3)$ is

$$\{dx \wedge dy \wedge dz\}.$$

248 Example In $\mathbb{R}^4$, $\bigwedge^2(\mathbb{R}^4)$ has dimension $\binom{4}{2} = 6$, and a basis for $\bigwedge^2(\mathbb{R}^4)$ is

$$\{dx_1 \wedge dx_2,\ dx_1 \wedge dx_3,\ dx_1 \wedge dx_4, dx_2 \wedge dx_3,\ dx_2 \wedge dx_4, dx_3 \wedge dx_4\}.$$

249 Definition Let $f(x_1, x_2, \ldots, x_n)$ be a 0-form in $\mathbb{R}^n$. The *exterior derivative* df of f is

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i.$$

Furthermore, if

$$\omega = f(x_1, x_2, \ldots, x_n) dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}$$

is a k-form in $\mathbb{R}^n$, the *exterior derivative* $d\omega$ of ω is the $(k+1)$-form

$$d\omega = df(x_1, x_2, \ldots, x_n) \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}.$$

250 Example If in $\mathbb{R}^2$, $\omega = x^3y^4$, then

$$d(x^3y^4) = 3x^2y^4dx + 4x^3y^3dy.$$

251 Example If in $\mathbb{R}^2$, $\omega = x^2ydx + x^3y^4dy$ then

$$\begin{aligned} d\omega &= d(x^2ydx + x^3y^4dy) \\ &= (2xydx + x^2dy) \wedge dx + (3x^2y^4dx + 4x^3y^3dy) \wedge dy \\ &= x^2dy \wedge dx + 3x^2y^4dx \wedge dy \\ &= (3x^2y^4 - x^2)dx \wedge dy \end{aligned}$$

252 Example Consider the change of variables $x = u + v, y = uv$. Then

$$dx = du + dv,$$

$$dy = v du + u dv,$$

whence

$$dx \wedge dy = (u - v) du \wedge dv.$$

253 Example Consider the transformation of co-ördinates xyz into uvw co-ördinates given by

$$u = x + y + z,\ v = \frac{z}{y+z},\ w = \frac{y+z}{x+y+z}.$$

Then

$$du = dx + dy + dz,$$

$$dv = -\frac{z}{(y+z)^2}dy + \frac{y}{(y+z)^2}dz,$$

$$dw = -\frac{y+z}{(x+y+z)^2}dx + \frac{x}{(x+y+z)^2}dy + \frac{x}{(x+y+z)^2}dz.$$

Multiplication gives

$$\begin{aligned} du \wedge dv \wedge dw &= \left(-\frac{zx}{(y+z)^2(x+y+z)^2} - \frac{y(y+z)}{(y+z)^2(x+y+z)^2} \right. \\ &\qquad \left. + \frac{z(y+z)}{(y+z)^2(x+y+z)^2} - \frac{xy}{(y+z)^2(x+y+z)^2}\right) dx \wedge dy \wedge dz \\ &= \frac{z^2 - y^2 - zx - xy}{(y+z)^2(x+y+z)^2} dx \wedge dy \wedge dz. \end{aligned}$$

3.2 Integrating in $\bigwedge^0(\mathbb{R}^n)$

254 Definition A 0-*dimensional oriented manifold of* $\mathbb{R}^n$ is simply a point $\mathbf{x} \in \mathbb{R}^n$, with a choice of the $+$ or $-$ sign. A general oriented 0-manifold is a union of oriented points.

255 Definition Let $M = +\{\mathbf{b}\} \cup -\{\mathbf{a}\}$ be an oriented 0-manifold, and let ω be a 0-form. Then

$$\int_M \omega = \omega(\mathbf{b}) - \omega(\mathbf{a}).$$

$-\mathbf{x}$ *has opposite orientation to* $+\mathbf{x}$ *and*

$$\int_{-\mathbf{x}} \omega = -\int_{+\mathbf{x}} \omega.$$

256 Example Let $M = -\{(1,0,0)\} \cup +\{(1,2,3)\} \cup -\{(0,-2,0)\}$[2] be an oriented 0-manifold, and let $\omega = x + 2y + z^2$. Then

$$\int_M \omega = -\omega((1,0,0)) + \omega(1,2,3) - \omega(0,0,3) = -(1) + (14) - (-4) = 17.$$

3.3 Integrating in $\bigwedge^1(\mathbb{R}^n)$

257 Definition A 1-*dimensional oriented manifold of* $\mathbb{R}^n$ is simply an oriented smooth curve $\Gamma \in \mathbb{R}^n$, with a choice of a $+$ orientation if the curve traverses in the direction of increasing t, or with a choice of a $-$ sign if the curve traverses in the direction of decreasing t. A general oriented 1-manifold is a union of oriented curves.

The curve $-\Gamma$ *has opposite orientation to* Γ *and*

$$\int_{-\Gamma} \omega = -\int_{\Gamma} \omega.$$

We now turn to the problem of integrating 1-forms.

258 Example Calculate

$$\int_\Gamma xy\,dx + (x+y)\,dy$$

where Γ is the parabola $y = x^2$, $x \in [-1; 2]$ oriented in the positive direction.

[2] Do not confuse, say, $-\{(1,0,0)\}$ with $-(1,0,0) = (-1,0,0)$. The first one means that the point $(1,0,0)$ is given negative orientation, the second means that $(-1,0,0)$ is the additive inverse of $(1,0,0)$.

Solution: We parametrise the curve as $x = t, y = t^2$. Then

$$xy\mathrm{d}x + (x+y)\mathrm{d}y = t^3\mathrm{d}t + (t+t^2)\mathrm{d}t^2 = (3t^3+2t^2)\mathrm{d}t,$$

whence

$$\begin{aligned}\int_\Gamma \omega &= \int_{-1}^{2} (3t^3+2t^2)\mathrm{d}t \\ &= \left[\frac{2}{3}t^3 + \frac{3}{4}t^4\right]_{-1}^{2} \\ &= \frac{69}{4}.\end{aligned}$$

What would happen if we had given the curve above a different parametrisation? First observe that the curve travels from $(-1,1)$ to $(2,4)$ on the parabola $y = x^2$. These conditions are met with the parametrisation $x = \sqrt{t}-1, y = (\sqrt{t}-1)^2$, $t \in [0;9]$. Then

$$\begin{aligned}xy\mathrm{d}x + (x+y)\mathrm{d}y &= (\sqrt{t}-1)^3\mathrm{d}(\sqrt{t}-1) + ((\sqrt{t}-1)+(\sqrt{t}-1)^2)\mathrm{d}(\sqrt{t}-1)^2 \\ &= (3(\sqrt{t}-1)^3 + 2(\sqrt{t}-1)^2)\mathrm{d}(\sqrt{t}-1) \\ &= \frac{1}{2\sqrt{t}}(3(\sqrt{t}-1)^3 + 2(\sqrt{t}-1)^2)\mathrm{d}t,\end{aligned}$$

whence

$$\begin{aligned}\int_\Gamma \omega &= \int_0^9 \frac{1}{2\sqrt{t}}(3(\sqrt{t}-1)^3 + 2(\sqrt{t}-1)^2)\mathrm{d}t \\ &= \left[\frac{3t^2}{4} - \frac{7t^{3/2}}{3} + \frac{5t}{2} - \sqrt{t}\right]_0^9 \\ &= \frac{69}{4},\end{aligned}$$

as before.

It turns out that if two different parametrisations of the same curve have the same orientation, then their integrals are equal. Hence, we only need to worry about finding a suitable parametrisation.

259 Example Calculate the line integral

$$\int_\Gamma y \sin x\mathrm{d}x + x\cos y\mathrm{d}y,$$

where Γ is the line segment from $(0,0)$ to $(1,1)$ in the positive direction.

Solution: This line has equation $y = x$, so we choose the parametrisation $x = y = t$. The integral is thus

$$\begin{aligned}\int_\Gamma y\sin x dx + x\cos y dy &= \int_0^1 (t\sin t + t\cos t)dt \\ &= [t(\sin x - \cos t)]_0^1 - \int_0^1 (\sin t - \cos t)dt \\ &= 2\sin 1 - 1,\end{aligned}$$

upon integrating by parts.

260 Example Calculate the path integral

$$\int_\Gamma \frac{x+y}{x^2+y^2}\, dy + \frac{x-y}{x^2+y^2}\, dx$$

around the closed square $\Gamma = ABCD$ with $A = (1,1)$, $B = (-1,1)$, $C = (-1,-1)$, and $D = (1,-1)$ in the direction $ABCDA$.

Solution: On AB, $y = 1, dy = 0$, on BC, $x = -1, dx = 0$, on CD, $y = -1, dy = 0$, and on DA, $x = 1, dx = 0$. The integral is thus

$$\begin{aligned}\int_\Gamma \omega &= \int_{AB} \omega + \int_{BC} \omega + \int_{CD} \omega + \int_{DA} \omega \\ &= \int_1^{-1} \frac{x-1}{x^2+1}\, dx + \int_1^{-1} \frac{y-1}{y^2+1}\, dy + \int_{-1}^1 \frac{x+1}{x^2+1}\, dx + \int_{-1}^1 \frac{y+1}{y^2+1}\, dy \\ &= 4\int_{-1}^1 \frac{1}{x^2+1}\, dx \\ &= 4\arctan x|_{-1}^1 \\ &= 2\pi.\end{aligned}$$

When the integral is along a closed path, like in the preceding example, it is customary to use the symbol $\oint_\Gamma$ *rather than* $\int_\Gamma$. *The positive direction of integration is that sense that when traversing the path, the area enclosed by the curve is to the left of the curve.*

261 Example Calculate the path integral

$$\oint_\Gamma x^2 dy + y^2 dx,$$

where Γ is the ellipse $9x^2 + 4y^2 = 36$ traversed once in the positive sense.

Solution: Parametrise the ellipse as $x = 2\cos t, y = 3\sin t, t \in [0; 2\pi]$. Observe that when traversing this closed curve, the area of the ellipse is on the left hand side of the path, so this parametrisation traverses the curve in the positive sense. We have

$$\begin{aligned}\oint_\Gamma \omega &= \int_0^{2\pi} ((4\cos^2 t)(3\cos t) + (9\sin t)(-2\sin t))dt \\ &= \int_0^{2\pi} (12\cos^3 t - 18\sin^3 t)dt \\ &= 0.\end{aligned}$$

262 Example Find $\oint_\Gamma z dx + x dy + y dz$ where Γ is the intersection of the sphere $x^2 + y^2 + z^2 = 1$ and the plane $x + y = 1$, traversed in the positive direction.

Solution: The curve lies on the sphere, and to parametrise this curve, we dispose of one of the variables, y say, from where $y = 1 - x$ and $x^2 + y^2 + z^2 = 1$ give

$$\begin{aligned}x^2 + (1-x)^2 + z^2 = 1 &\implies 2x^2 - 2x + z^2 = 0 \\ &\implies 2\left(x - \frac{1}{2}\right)^2 + z^2 = \frac{1}{2} \\ &\implies 4\left(x - \frac{1}{2}\right)^2 + 2z^2 = 1.\end{aligned}$$

So we now put

$$x = \frac{1}{2} + \frac{\cos t}{2}, \quad z = \frac{\sin t}{\sqrt{2}}, \quad y = 1 - x = \frac{1}{2} - \frac{\cos t}{2}.$$

We must integrate on the side of the plane that can be viewed from the point $(1, 1, 0)$ (observe that the vector $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is normal to the plane). On the zx-plane, $4\left(x - \frac{1}{2}\right)^2 + 2z^2 = 1$ is an ellipse. To obtain a positive parametrisation we must integrate from $t = 2\pi$ to $t = 0$ (this is because when you look at the ellipse from the point $(1, 1, 0)$ the positive x-axis is to your left, and not your right). Thus

$$\begin{aligned}
\oint_{\Gamma} z\mathrm{d}x + x\mathrm{d}y + y\mathrm{d}z &= \int_{2\pi}^{0} \frac{\sin t}{\sqrt{2}}\mathrm{d}\left(\frac{1}{2} + \frac{\cos t}{2}\right) \\
&\quad + \int_{2\pi}^{0} \left(\frac{1}{2} + \frac{\cos t}{2}\right)\mathrm{d}\left(\frac{1}{2} - \frac{\cos t}{2}\right) \\
&\quad + \int_{2\pi}^{0} \left(\frac{1}{2} - \frac{\cos t}{2}\right)\mathrm{d}\left(\frac{\sin t}{\sqrt{2}}\right) \\
&= \int_{2\pi}^{0} \left(\frac{\sin t}{4} + \frac{\cos t}{2\sqrt{2}} + \frac{\cos t \sin t}{4} - \frac{1}{2\sqrt{2}}\right) \mathrm{d}t \\
&= \frac{\pi}{\sqrt{2}}.
\end{aligned}$$

Given a curve Γ how can we find its length? The idea, as seen in figure 3.1 is to consider the projections $\mathrm{d}x_i, 1 \leq i \leq n$ at each point. The length of the vector

$$\mathrm{d}\mathbf{x} = \begin{bmatrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \\ \vdots \\ \mathrm{d}x_n \end{bmatrix}$$

is

$$||\mathrm{d}\mathbf{x}|| = \sqrt{(\mathrm{d}x_1)^2 + (\mathrm{d}x_2)^2 + \cdots + (\mathrm{d}x_n)^2}.$$

Hence the length of Γ is given by

$$\int_\Gamma \|d\mathbf{x}\| = \int_\Gamma \sqrt{(dx_1)^2 + (dx_2)^2 + \cdots + (dx_n)^2}. \tag{3.1}$$

Similarly, suppose that Γ is a simple closed curve in $\mathbb{R}^2$. How do we find the (oriented) area of the region it encloses? The idea, borrowed from finding areas of polygons, is to split the region into triangles, each of area

$$\frac{1}{2}\det\begin{bmatrix} x & dx \\ y & dy \end{bmatrix} = \frac{1}{2}(xdy - ydx),$$

and to sum over the closed curve, obtaining a total oriented area of

$$\frac{1}{2}\oint_\Gamma \det\begin{bmatrix} x & dx \\ y & dy \end{bmatrix} = \frac{1}{2}\oint_\Gamma (xdy - ydx). \tag{3.2}$$

Here $\oint_\Gamma$ denotes integration around the closed curve.

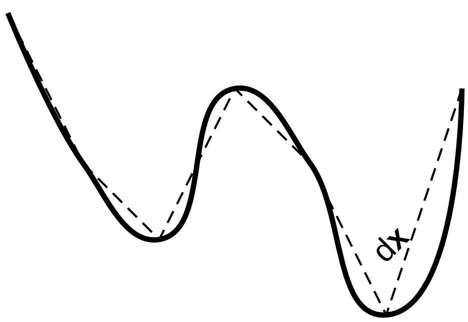

Figure 3.1: Length of a curve.

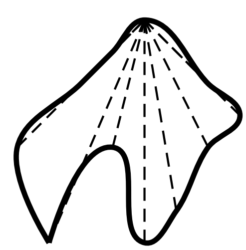

Figure 3.2: Area enclosed by a simple closed curve

263 Example Let $(A, B) \in \mathbb{R}^2, A > 0, B > 0$. Find a parametrisation of the ellipse

$$\Gamma : \{(x, y) \in \mathbb{R}^2 : \frac{x^2}{A^2} + \frac{y^2}{B^2} = 1\}.$$

Furthermore, find an integral expression for the perimeter of this ellipse and find the area it encloses.

Solution: Consider the parametrisation $\Gamma : [0; 2\pi] \to \mathbb{R}^2$, with

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A\cos t \\ B\sin t \end{bmatrix}.$$

This is a parametrisation of the ellipse, for

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = \frac{A^2\cos^2 t}{A^2} + \frac{B^2\sin^2 t}{B^2} = \cos^2 t + \sin^2 t = 1.$$

Notice that this parametrisation goes around once the ellipse counterclockwise. The perimeter of the ellipse is given by

$$\int_\Gamma \|d\mathbf{x}\| = \int_0^{2\pi} \sqrt{A^2\sin^2 t + B^2\cos^2 t}\, dt.$$

The above integral is an *elliptic integral* and quite hard to evaluate. We will have better luck with the area of the ellipse, which is given by

$$\begin{aligned} \frac{1}{2}\oint_\Gamma (x\,dy - y\,dx) &= \frac{1}{2}\oint (A\cos t\, d(B\sin t) - B\sin t\, d(A\cos t)) \\ &= \frac{1}{2}\int_0^{2\pi} (AB\cos^2 t + AB\sin^2 t))\, dt \\ &= \frac{1}{2}\int_0^{2\pi} AB\, dt \\ &= \pi AB. \end{aligned}$$

264 Example Find a parametric representation for the astroid

$$\Gamma : \{(x, y) \in \mathbb{R}^2 : x^{2/3} + y^{2/3} = 1\},$$

in figure 3.3. Find the perimeter of the astroid and the area it encloses.

Solution: Take

$$\begin{bmatrix} x \\ \\ y \end{bmatrix} = \begin{bmatrix} \cos^3 t \\ \\ \sin^3 t \end{bmatrix}$$

with $t \in [0; 2\pi] \to \mathbb{R}^2$. with Then

$$x^{2/3} + y^{2/3} = \cos^2 t + \sin^2 t = 1.$$

The perimeter of the astroid is

$$\begin{aligned} \int_\Gamma \|d\mathbf{x}\| &= \int_0^{2\pi} \sqrt{9\cos^4 t \sin^2 t + 9\sin^4 t \cos^2 t}\, dt \\ &= \int_0^{2\pi} 3|\sin t \cos t|\, dt \\ &= \frac{3}{2}\int_0^{2\pi} |\sin 2t|\, dt \\ &= 6\int_0^{\pi/2} \sin 2t\, dt \\ &= 6. \end{aligned}$$

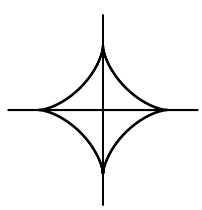

Figure 3.3: Example 264.

The area of the astroid is given by

$$\begin{aligned}
\frac{1}{2}\oint_{\Gamma}(x\mathrm{d}y - y\mathrm{d}x) &= \frac{1}{2}\oint(\cos^3 t\ \mathrm{d}(\sin^3 t) - \sin^3 t\ \mathrm{d}(\cos^3 t)) \\
&= \frac{1}{2}\int_0^{2\pi}(3\cos^4 t\sin^2 t + 3\sin^4 t\cos^2 t))\ \mathrm{d}t \\
&= \frac{3}{2}\int_0^{2\pi}(\sin t\cos t)^2\ \mathrm{d}t \\
&= \frac{3}{8}\int_0^{2\pi}(\sin 2t)^2\ \mathrm{d}t \\
&= \frac{3}{16}\int_0^{2\pi}(1-\cos 4t)\ \mathrm{d}t \\
&= \frac{3\pi}{8}.
\end{aligned}$$

265 Example Find the area enclosed by the curve $x = \sin^3 t, y = (\cos t)(1+\sin^2 t)$.

Solution: Observe that the parametrisation traverses the curve once clockwise if $t \in [0; 2\pi]$. The area is given by

$$\begin{aligned}
\frac{1}{2}\oint_{\Gamma}\det\begin{bmatrix} x & \mathrm{d}x \\ y & \mathrm{d}y \end{bmatrix} &= \frac{1}{2}\oint x\mathrm{d}y - y\mathrm{d}x \\
&= \frac{4}{2}\int_{\pi/2}^{0}(\sin^3 t(-\sin t(1+\sin^2 t) + 2\sin t\cos^2 t) \\
&\qquad -\cos t(1+\sin^2 t)(3\sin^2 t\cos t))\mathrm{d}t \\
&= 2\int_{\pi/2}^{0}(-3\sin^2 t + \sin^4 t)\mathrm{d}t \\
&= 2\int_{\pi/2}^{0}\left(-\frac{9}{8} + \cos 2t + \frac{1}{8}\cos 4t\right)\mathrm{d}t \\
&= \frac{9\pi}{8}.
\end{aligned}$$

266 Definition Let Γ be a smooth curve. The integral

$$\int_\Gamma f(\mathbf{x})||d\mathbf{x}||$$

is called the *path integral of* f *along* Γ*.*

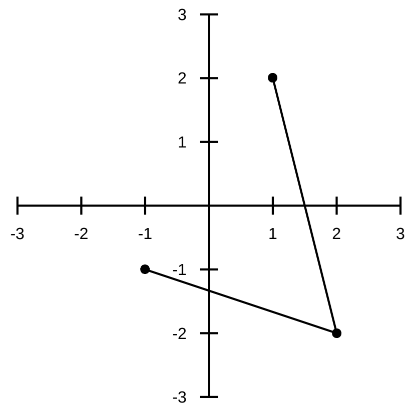

Figure 3.4: Example 267.

267 Example Find $\int_\Gamma x||d\mathbf{x}||$ where Γ is the triangle starting at $A : (-1, -1)$ to $B : (2, -2)$, and ending in $C : (1, 2)$.

Solution: The lines passing through the given points have equations L_{AB} : $y = \dfrac{-x-4}{3}$, and $L_{BC} : y = -4x + 6$. On L_{AB}

$$x||d\mathbf{x}|| = x\sqrt{(dx)^2 + (dy)^2} = x\sqrt{1 + \left(-\frac{1}{3}\right)^2}dx = \frac{x\sqrt{10}dx}{3},$$

and on L_{BC}

$$x||d\mathbf{x}|| = x\sqrt{(dx)^2 + (dy)^2} = x(\sqrt{1 + (-4)^2})dx = x\sqrt{17}dx.$$

Hence

$$\begin{aligned}\int_{\Gamma} x||d\mathbf{x}|| &= \int_{L_{AB}} x||d\mathbf{x}|| + \int_{L_{BC}} x||d\mathbf{x}|| \\ &= \int_{-1}^{2} \frac{x\sqrt{10}dx}{3} + \int_{2}^{1} x\sqrt{17}dx \\ &= \frac{\sqrt{10}}{2} - \frac{3\sqrt{17}}{2}.\end{aligned}$$

3.4 Closed and Exact Forms

268 Lemma (Poincaré Lemma) If ω is a p-differential form of continuously differentiable functions in $\mathbb{R}^n$ then

$$dd\omega = 0.$$

Proof We will prove this by induction on p. For $p = 0$ if

$$\omega = f(x_1, x_2, \ldots, x_n)$$

then

$$d\omega = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} dx_k$$

and

$$\begin{aligned}dd\omega &= \sum_{k=1}^{n} d\left(\frac{\partial f}{\partial x_k}\right) \wedge dx_k \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k} \wedge dx_j\right) \wedge dx_k \\ &= \sum_{1 \leq j \leq k \leq n}^{n} \left(\frac{\partial^2 f}{\partial x_j \partial x_k} - \frac{\partial^2 f}{\partial x_k \partial x_j}\right) dx_j \wedge dx_k \\ &= 0,\end{aligned}$$

since ω is continuously differentiable and so the mixed partial derivatives are equal. Consider now an arbitrary p-form, $p > 0$. Since such a form can be written as

$$\omega = \sum_{1\leq j_1 \leq j_2 \leq \cdots \leq j_p \leq n} a_{j_1 j_2 \ldots j_p} dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p},$$

where the $a_{j_1 j_2 \ldots j_p}$ are continuous differentiable functions in $\mathbb{R}^n$, we have

$$\begin{aligned} d\omega &= \sum_{1\leq j_1 \leq j_2 \leq \cdots \leq j_p \leq n} da_{j_1 j_2 \ldots j_p} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p} \\ &= \sum_{1\leq j_1 \leq j_2 \leq \cdots \leq j_p \leq n} \left(\sum_{i=1}^{n} \frac{\partial a_{j_1 j_2 \ldots j_p}}{\partial x_i} dx_i \right) \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}, \end{aligned}$$

it is enough to prove that for each summand

$$d\left(da \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right) = 0.$$

But

$$\begin{aligned} d\left(da \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right) &= dda \wedge \left(dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right) \\ &\quad + da \wedge d\left(dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right) \\ &= da \wedge d\left(dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right), \end{aligned}$$

since $dda = 0$ from the case $p = 0$. But an independent induction argument proves that

$$d\left(dx_{j_1} \wedge dx_{j_2} \wedge \cdots dx_{j_p}\right) = 0,$$

completing the proof. ❑

269 Definition A differential form ω is said to be *exact* if there is a continuously differentiable function F such that

$$dF = \omega.$$

270 Example The differential form

$$xdx + ydy$$

is exact, since

$$xdx + ydy = d\left(\frac{1}{2}(x^2 + y^2)\right).$$

271 Example The differential form

$$y\,dx + x\,dy$$

is exact, since

$$y\,dx + x\,dy = d\,(xy)\,.$$

272 Example The differential form

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

is exact, since

$$\frac{x}{x^2+y^2}dx + \frac{y}{x^2+y^2}dy = d\left(\frac{1}{2}\log_e(x^2+y^2)\right).$$

Let $\omega = dF$ be an exact form. By the Poincaré Lemma Theorem 268, $d\omega = ddF = 0$. A result of Poincaré says that for certain domains (called star-shaped domains*) the converse is also true, that is, if $d\omega = 0$ on a star-shaped domain then ω is exact.*

273 Example Determine whether the differential form

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

is exact.

Solution: Assume there is a function F such that

$$dF = \omega.$$

By the Chain Rule

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy.$$

This demands that

$$\frac{\partial F}{\partial x} = \frac{2x(1-e^y)}{(1+x^2)^2},$$

$$\frac{\partial F}{\partial y} = \frac{e^y}{1+x^2}.$$

We have a choice here of integrating either the first, or the second expression. Since integrating the second expression (with respect to y) is easier, we find

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

where $\phi(x)$ is a function depending only on x. To find it, we differentiate the obtained expression for F with respect to x and find

$$\frac{\partial F}{\partial x} = -\frac{2xe^y}{(1+x^2)^2} + \phi'(x).$$

Comparing this with our first expression for $\frac{\partial F}{\partial x}$, we find

$$\phi'(x) = \frac{2x}{(1+x^2)^2},$$

that is

$$\phi(x) = -\frac{1}{1+x^2} + c,$$

where c is a constant. We then take

$$F(x,y) = \frac{e^y - 1}{1+x^2} + c.$$

274 Example Is there a continuously differentiable function such that

$$dF = \omega = y^2z^3dx + 2xyz^3dy + 3xy^2z^2dz \ \ ?$$

Solution: We have

$$\begin{aligned} d\omega &= (2yz^3dy + 3y^2z^2dz) \wedge dx \\ &\quad +(2yz^3dx + 2xz^3dy + 6xyz^2dz) \wedge dy \\ &\quad +(3y^2z^2dx + 6xyz^2dy + 6xy^2zdz) \wedge dz \\ &= 0, \end{aligned}$$

so this form is exact in a star-shaped domain. So put

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz = y^2z^3dx + 2xyz^3dy + 3xy^2z^2dz.$$

Then

$$\frac{\partial F}{\partial x} = y^2z^3 \implies F = xy^2z^3 + a(y, z),$$
$$\frac{\partial F}{\partial y} = 2xyz^3 \implies F = xy^2z^3 + b(x, z),$$
$$\frac{\partial F}{\partial z} = 3xy^2z^2 \implies F = xy^2z^3 + c(x, y),$$

Comparing these three expressions for F, we obtain $F(x, y, z) = xy^2z^3$.

We have the following equivalent of the Fundamental Theorem of Calculus.

275 Theorem Let $U \subseteq \mathbb{R}^n$ be an open set. Assume $\omega = dF$ is an exact form, and Γ a path in U with starting point A and endpoint B. Then

$$\int_\Gamma \omega = \int_A^B dF = F(B) - F(A).$$

In particular, if Γ is a simple closed path, then

$$\oint_\Gamma \omega = 0.$$

276 Example Evaluate the integral

$$\oint_\Gamma \frac{2x}{x^2 + y^2}dx + \frac{2y}{x^2 + y^2}\, dy$$

where Γ is the closed polygon with vertices at $A = (0, 0)$, $B = (5, 0)$, $C = (7, 2)$, $D = (3, 2)$, $E = (1, 1)$, traversed in the order $ABCDEA$.

Solution: Observe that

$$d\left(\frac{2x}{x^2 + y^2}dx + \frac{2y}{x^2 + y^2}\, dy\right) = -\frac{4xy}{(x^2 + y^2)^2}dy \wedge dx - \frac{4xy}{(x^2 + y^2)^2}dx \wedge dy = 0,$$

and so the form is exact in a start-shaped domain. By virtue of Theorem 275, the integral is 0.

277 Example Calculate the path integral

$$\oint_\Gamma (x^2 - y)dx + (y^2 - x)dy,$$

where Γ is a loop of $x^3 + y^3 - 2xy = 0$ traversed once in the positive sense.

Solution: Since

$$\frac{\partial}{\partial y}(x^2 - y) = -1 = \frac{\partial}{\partial x}(y^2 - x),$$

the form is exact, and since this is a closed simple path, the integral is 0.

278 Theorem If α is a p-form and β is a q-form on n-dimensional space, then

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta.$$

3.5 Integrating in $\bigwedge^2(\mathbb{R}^2)$

279 Definition A 2-*dimensional oriented manifold of* $\mathbb{R}^2$ is simply an open set (region) $D \in \mathbb{R}^2$, where the $+$ orientation is counter-clockwise and the $-$ orientation is clockwise. A general oriented 2-manifold is a union of open sets.

The region $-D$ *has opposite orientation to* D *and*

$$\int_{-D} \omega = -\int_D \omega.$$

In this section, unless otherwise noticed, we will choose the positive orientation for the regions considered. This corresponds to using the area form $dx \wedge dy$.

Let $D \subseteq \mathbb{R}^2$. Given a function $f : D \to \mathbb{R}$, the integral

$$\iint_D f(x, y) dx \wedge dy$$

is the sum of all the values of f restricted to D. In particular,

$$\iint_D dx \wedge dy$$

is the area of D.

In order to evaluate double integrals, we need the following.

280 Theorem (Fubini's Theorem) Let $D = [a;b] \times [c;d]$, and let $f : A \to \mathbb{R}$ be continuous. Then

$$\iint_D f(x,y)dx \wedge dy = \int_a^b \left(\int_c^d f(x,y)dy\right) dx = \int_c^d \left(\int_a^b f(x,y)dx\right) dy$$

Fubini's Theorem allows us to convert the double integral into iterated (single) integrals.

281 Example

$$\begin{aligned} \iint_{[0;1]\times[2;3]} xy dx \wedge dy &= \int_0^1 \left(\int_2^3 xy dy\right) dx \\ &= \int_0^1 \left(\left[\frac{xy^2}{2}\right]_2^3\right) dx \\ &= \int_0^1 \left(\frac{9x}{2} - 2x\right) dx \\ &= \left[\frac{5x^2}{4}\right]_0^1 \\ &= \frac{5}{4}. \end{aligned}$$

Notice that if we had integrated first with respect to x we would have obtained the same result:

$$\begin{aligned}\int_2^3 \left(\int_0^1 xy\mathrm{d}x\right) \mathrm{d}y &= \int_2^3 \left(\left[\frac{x^2y}{2}\right]_0^1\right) \mathrm{d}y \\ &= \int_2^3 \left(\frac{y}{2}\right) \mathrm{d}x \\ &= \left[\frac{y^2}{4}\right]_2^3 \\ &= \frac{5}{4}.\end{aligned}$$

Also, this integral is "factorable into x and y pieces" meaning that

$$\begin{aligned}\iint_{[0;1]\times[2;3]} xy\mathrm{d}x \wedge \mathrm{d}y &= \left(\int_0^1 x\mathrm{d}x\right)\left(\int_2^3 y\mathrm{d}y\right) \\ &= \left(\frac{1}{2}\right)\left(\frac{5}{2}\right) \\ &= \frac{5}{4}\end{aligned}$$

282 Example We have

$$\begin{aligned}\int_3^4 \int_0^1 (x+2y)(2x+y)\ \mathrm{d}x \wedge \mathrm{d}y &= \int_3^4 \int_0^1 (2x^2+5xy+2y^2)\ \mathrm{d}x \wedge \mathrm{d}y \\ &= \int_3^4 \left(\frac{2}{3}+\frac{5}{2}y+2y^2\right)\ \mathrm{d}y \\ &= \frac{409}{12}.\end{aligned}$$

283 Example Find

$$\iint\limits_{D} (x+y)(\sin x)(\sin y)\mathrm{d}x \wedge \mathrm{d}y$$

where $D = [0;\pi]^2$.

Solution: The integral equals

$$\begin{aligned}\iint\limits_{D} x\sin x\sin y\,dx\wedge dy+\iint\limits_{D} y\sin x\sin y\,dx\wedge dy &= 2\left(\int_0^{\pi} y\sin y\;dy\right)\left(\int_0^{\pi}\sin x\;dx\right)\\ &= 4\pi.\end{aligned}$$

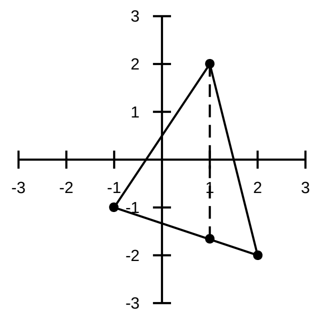

Figure 3.5: Example 284. Integration order $dydx$.

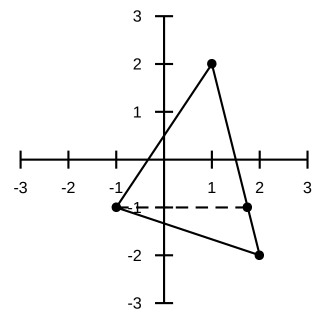

Figure 3.6: Example 284. Integration order $dxdy$.

In the cases when the domain of integration is not a rectangle, we decompose so that, one variable is kept constant.

284 Example Find $\iint\limits_{D} xy\;dxdy$ in the triangle with vertices $A:(-1,-1)$, $B:(2,-2)$, $C:(1,2)$.

The lines passing through the given points have equations $L_{AB}: y=\dfrac{-x-4}{3}$, $L_{BC}: y=-4x+6$, $L_{CA}: y=\dfrac{3x+1}{2}$. Now, we draw the region *carefully*. If we integrate first with respect to y, we must divide the region as in figure 3.5, because there are two upper lines which the upper value of y might be. The lower point of the dashed line is $(1,-5/3)$. The integral is thus

$$\int_{-1}^{1} x\left(\int_{(-x-4)/3}^{(3x+1)/2} y\, dy\right) dx + \int_{1}^{2} x\left(\int_{(-x-4)/3}^{-4x+6} y\, dy\right) dx = -\frac{11}{8}.$$

If we integrate first with respect to x, we must divide the region as in figure 3.6, because there are two left-most lines which the left value of x might be. The right point of the dashed line is $(7/4, -1)$. The integral is thus

$$\int_{-2}^{-1} y\left(\int_{-4-3y}^{(6-y)/4} x\, dx\right) dy + \int_{-1}^{2} y\left(\int_{(2y-1)/3}^{(6-y)/4} x\, dx\right) dy = -\frac{11}{8}.$$

285 Example Consider the region inside the parallelogram P with vertices at $A : (6, 3)$, $B : (8, 4)$, $C : (9, 6)$, $D : (7, 5)$, as in figure 3.7. Find

$$\frac{1}{3}\iint_P xy \;\; dx \wedge dy.$$

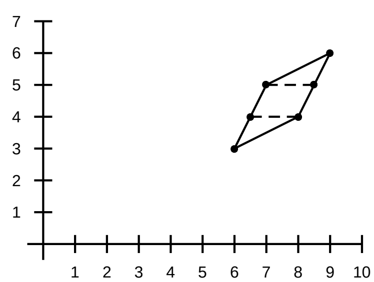

Figure 3.7: Example 285.

Solution: The lines joining the points have equations

$$L_{AB} : \; y = \frac{x}{2},$$

$$L_{BC} : \; y = 2x - 12,$$

$$L_{CD} : \; y = \frac{x}{2} + \frac{3}{2},$$

$$L_{DA}: \ y = 2x - 9.$$

The integral is thus

$$\frac{1}{3}\int_3^4\int_{(y+9)/2}^{2y} xy \, dx{\wedge}dy + \frac{1}{3}\int_4^5\int_{(y+9)/2}^{(y+12)/2} xy \, dx{\wedge}dy + \frac{1}{3}\int_4^5\int_{2y-3}^{(y+12)/2} xy \, dx{\wedge}dy = \frac{409}{12}.$$

286 Example Find $\iint\limits_D (2x+3y+1)\, dxdy$, where D is the triangle with vertices at $A(-1,-1)$, $B(2,-4)$, and $C(1,3)$.

Solution: The line joining A, and B has equation $y = -x-2$, line joining B, and C has equation $y = -7x+10$, and line joining A, and C has equation $y = 2x+1$. We split the triangle along the vertical line $x = 1$, and integrate first with respect to y. The desired integral is then

$$\begin{aligned}
\iint_D (2x+3y+1)\, dxdy &= \int_{-1}^{1}\left(\int_{-x-2}^{2x+1}(2x+3y+1)dy\right)dx \\
&\quad + \int_1^2\left(\int_{-x-2}^{-7x+10}(2x+3y+1)dy\right)dx \\
&= \int_{-1}^{1}\left(\frac{21}{2}x^2+9x-\frac{3}{2}\right)dx \\
&\quad + \int_1^2\left(60x^2-198x+156\right)dx \\
&= 4-1 \\
&= 3.
\end{aligned}$$

287 Example Find

$$\iint\limits_D (xy(x+y))dx\wedge dy$$

where

$$D = \{(x,y)\in\mathbb{R}^2 | x\geq 0, y\geq 0, x+y\leq 1\}.$$

Solution: The domain of integration is a triangle. The integral equals

$$\begin{aligned}\iint\limits_D xy(x+y)dx\wedge dy &= \int_0^1 \left(\int_0^{1-x} xy(x+y)\, dy\right)\, dx \\ &= \int_0^1 x\left[x\frac{y^2}{2}+\frac{y^3}{3}\right]_0^{1-x} dx \\ &= \int_0^1 x\left(\frac{x(1-x)^2}{2}+\frac{(1-x)^3}{3}\right)\, dx \\ &= \frac{1}{30}.\end{aligned}$$

288 Example Find

$$\iint\limits_D \frac{1}{(x+y)^4}dx\wedge dy$$

where

$$D=\{(x,y)\in\mathbb{R}^2|x\geq 1, y\geq 1, x+y\leq 4\}.$$

Solution: The integral equals

$$\begin{aligned}\iint\limits_D \frac{1}{(x+y)^4}dx\wedge dy &= \int_1^3 \left(\int_1^{4-x} \frac{dy}{(x+y)^4}\, dy\right)\, dx \\ &= \int_1^3 \left[-\frac{1}{3}(x+y)^{-3}\right]_1^{4-x} dx \\ &= \frac{1}{3}\int_1^3 \left(\frac{1}{(1+x)^3}-\frac{1}{64}\right)\, dx \\ &= \frac{1}{48}.\end{aligned}$$

289 Example Find

$$\iint\limits_D x dx\wedge dy$$

where

$$D=\{(x,y)\in\mathbb{R}^2|y\geq 0, x-y+1\geq 0, x+2y-4\leq 0\}.$$

Solution: The integral equals

$$\begin{aligned}\iint_D x\mathrm{d}x\wedge \mathrm{d}y &= \int_{-1}^{2/3}\left(\int_0^{x+1}\mathrm{d}y\right)x\,\mathrm{d}x+\int_{2/3}^{4}\left(\int_0^{2-\frac{x}{2}}\mathrm{d}y\right)x\,\mathrm{d}x\\ &= \int_{-1}^{2/3}x(x+1)\,\mathrm{d}x+\int_{2/3}^{4}x\left(2-\frac{x}{2}\right)\,\mathrm{d}x\\ &= \frac{275}{54}.\end{aligned}$$

290 Example Find

$$\iint_D xy\mathrm{d}x\wedge \mathrm{d}y$$

where

$$D=\{(x,y)\in\mathbb{R}^2|y\geq x^2, x\geq y^2\}.$$

Solution: The integral equals

$$\begin{aligned}\iint_D xy\mathrm{d}x\wedge \mathrm{d}y &= \int_0^1 x\left(\int_{x^2}^{\sqrt{x}} y\,\mathrm{d}y\right)\,\mathrm{d}x\\ &= \int_0^1\frac{1}{2}x(x-x^4)\,\mathrm{d}x\\ &= \frac{1}{12}.\end{aligned}$$

291 Example Find

$$\iint_D \log_e(1+x+y)\mathrm{d}x\wedge \mathrm{d}y$$

where

$$D=\{(x,y)\in\mathbb{R}^2|x\geq 0, y\geq 0, x+y\leq 1\}.$$

Solution: Integrating by parts,

$$\begin{aligned}\iint\limits_D \log_e(1+x+y)dx \wedge dy &= \int_0^1 \left(\int_0^{1-x} \log_e(1+x+y)\, dy\right)\, dx \\ &= \int_0^1 [(1+x+y)\log_e(1+x+y) - (1+x+y)]_0^{1-x}\, dx \\ &= \int_0^1 (2\log_e(2) - 1 - \log_e(1+x) - x\log_e(1+x) + x)\, dx \\ &= \frac{1}{4}.\end{aligned}$$

292 Example Find

$$\iint\limits_D xy dx \wedge dy$$

where

$$D = \{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, xy + y + x \leq 1\}.$$

Solution: The integral equals

$$\begin{aligned}\iint\limits_D xy dx \wedge dy &= \int_0^1 \left(\int_0^{\frac{1-x}{1+x}} xy\, dy\right)\, dx \\ &= \int_0^1 \left(\frac{1}{2}x\left(\frac{1-x}{1+x}\right)^2\right)\, dx \\ &= \int_1^2 \frac{(t-1)(t-2)^2}{t^2}\, dt \\ &= 4\log_e 2 - \frac{11}{4}.\end{aligned}$$

293 Example Find

$$\iint\limits_D \log_e(1+x^2+y)dx \wedge dy$$

where

$$D = \{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, x^2 + y \leq 1\}.$$

Solution: Using integration by parts,

$$\begin{aligned}\iint\limits_{D} \log_e(1+x^2+y)dx\wedge dy &= \int_0^1 \left(\int_0^{1-x^2} \log_e(1+x^2+y)\, dy\right) dx \\ &= \int_0^1 (2\log_e(2)-1-\log_e(1+x^2))dx \\ &\quad + \int_0^1 (-x^2\log_e(1+x^2)+x^2)\, dx \\ &= \frac{2}{3}\log_e 2+\frac{8}{9}-\frac{\pi}{3}.\end{aligned}$$

294 Example Find

$$\iint\limits_{D} \frac{y}{x^2+1} dx\wedge dy$$

where

$$D=\{(x,y)\in\mathbb{R}^2 | x\geq 0, x^2+y^2\leq 1\}.$$

Solution: The integral is 0. Observe that if $(x,y)\in D$ then $(x,-y)\in D$. Also, $f(x,-y)=-f(x,y)$.

295 Example Find

$$\iint\limits_{D} 2x(x^2+y^2)dx\wedge dy$$

where

$$D=\{(x,y)\in\mathbb{R}^2 | x^4+y^4+x^2-y^2\leq 1\}.$$

Solution: The integral is 0. Observe that if $(x,y)\in D$ then $(-x,y)\in D$. Also, $f(-x,y)=-f(x,y)$.

296 Example Find

$$\iint\limits_{D} |x-y| dx\wedge dy$$

where

$$D=\{(x,y)\in\mathbb{R}^2 | |x|\leq 1, |y|\leq 1\}.$$

Solution: Let

$$D_1 = \{(x, y) \in \mathbb{R}^2 | -1 \le x \le 1, x \le y\},$$

$$D_2 = \{(x, y) \in \mathbb{R}^2 | -1 \le x \le 1, x > y\}.$$

Then $D = D_1 \cup D_2, D_1 \cap D_2 = \emptyset$ and so

$$\iint\limits_D f(x, y)dx \wedge dy = \iint_{D_1} f(x, y)dx \wedge dy + \iint_{D_2} f(x, y)dx \wedge dy.$$

By symmetry,

$$\iint_{D_1} f(x, y)dx \wedge dy = \iint_{D_2} f(x, y)dx \wedge dy,$$

and so

$$\begin{aligned}
\iint_D f(x, y)dx \wedge dy &= 2\iint_{D_1} f(x, y)dx \wedge dy \\
&= 2\int_{-1}^{1} \left(\int_x^1 (y - x)\ dy\right)\ dx \\
&= \int_{-1}^{1} (1 - 2x + x^2)\ dx \\
&= \frac{8}{3}.
\end{aligned}$$

297 Example Find

$$\int_0^4 \left(\int_{y/2}^{\sqrt{y}} e^{y/x}\quad dx\right) dy.$$

Solution: We have

$$0 \le y \le 4, \quad \frac{y}{2} \le x \le \sqrt{y} \implies 0 \le x \le 2, \quad x^2 \le y \le 2x.$$

We then have

$$\begin{aligned}
\int_0^4 \left(\int_{y/2}^{\sqrt{y}} e^{y/x} dx\right) dy &= \int_0^2 \left(\int_{x^2}^{2x} e^{y/x} dy\right) dx \\
&= \int_0^2 \left(xe^{y/x}\Big|_{x^2}^{2x}\right) dx \\
&= \int_0^2 (xe^2 - xe^x)\, dx \\
&= 2e^2 - (2e^2 - e^2 + 1) \\
&= e^2 - 1
\end{aligned}$$

298 Example Find

$$\int_1^2 \left(\int_{\sqrt{x}}^{x} \sin\frac{\pi x}{2y}\, dy\right)\, dx + \int_2^4 \left(\int_{\sqrt{x}}^{2} \sin\frac{\pi x}{2y}\, dy\right)\, dx.$$

Solution: Upon splitting the domain of integration, we find that the integral equals

$$\begin{aligned}
\int_1^2 \left(\int_y^{y^2} \sin\frac{\pi x}{2y} dx\right) dy &= \int_1^2 \left[-\frac{2y}{\pi}\cos\frac{\pi x}{2y}\right]_y^{y^2} dy \\
&= -\int_1^2 -\frac{2y}{\pi}\cos\frac{\pi y}{2} dy \\
&= \frac{4(\pi+2)}{\pi^3},
\end{aligned}$$

upon integrating by parts.

299 Example Find the area of the region

$$R = \{(x, y) \in \mathbb{R}^2 : \sqrt{x} + \sqrt{y} \geq 1, \quad \sqrt{1-x} + \sqrt{1-y} \geq 1\}.$$

Solution: The area is given by

$$\begin{aligned}\iint_D dx \wedge dy &= \int_0^1 \left(\int_{(1-\sqrt{x})^2}^{1-(1-\sqrt{1-x})^2} dy\right) dx \\ &= 2\int_0^1 (\sqrt{1-x} + \sqrt{x} - 1)dx \\ &= \frac{2}{3}.\end{aligned}$$

300 Example Find $\iint_D \sqrt{xy}\, dx \wedge dy$, where

$$D = \{(x, y) \in \mathbb{R}^2 : y \geq 0, (x+y)^2 \leq 2x\}.$$

Solution: Observe that $x \geq \frac{1}{2}(x+y)^2 \geq 0$. Hence we may take the positive square root giving $y \leq \sqrt{2x} - x$. Since $y \geq 0$, we must have $\sqrt{2x} - x \geq 0$ which means that $x \leq 2$. The integral equals

$$\begin{aligned}\int_0^2 \left(\int_0^{\sqrt{2x}-x} \sqrt{xy}dy\right) dx &= \frac{2}{3}\int_0^2 \sqrt{x}(\sqrt{2x} - x)^{3/2}dx \\ &= \frac{4}{3}\int_0^{\sqrt{2}} u^2(u\sqrt{2} - u^2)^{3/2}du \\ &= \frac{1}{6}\int_{-1}^{1} (1-v^2)^{3/2}(1+v)^2 dv \\ &= \frac{1}{6}\int_{-\pi/2}^{\pi/2} \cos^4\theta(1+\sin^2\theta)d\theta \\ &= \frac{7\pi}{96}.\end{aligned}$$

301 Example A rectangle R on the plane is the disjoint union $R = \cup_{k=1}^{N} R_k$ of rectangles R_k. It is known that at least one side of each of the rectangles R_k is an integer. Shew that at least one side of R is an integer.

Solution: Observe that

$$\int_0^a \sin 2\pi x \, dx = \begin{cases} 0 & \text{if } a \text{ is an integer} \\ \dfrac{1}{2\pi}(1 - \cos 2\pi a) & \text{if } a \text{ is not an integer} \end{cases}$$

Thus

$$\int_0^a \sin 2\pi x \, dx = 0 \iff a \text{ is an integer.}$$

Now

$$\sum_{k=1}^{N} \iint_{R_k} \sin 2\pi x \sin 2\pi y \, dx \wedge dy = 0$$

since at least one of the sides of each R_k is an integer. Since

$$\iint_R \sin 2\pi x \sin 2\pi y \, dx \wedge dy = \sum_{k=1}^{N} \iint_{R_k} \sin 2\pi x \sin 2\pi y \, dx \wedge dy,$$

we deduce that at least one of the sides of R is an integer, finishing the proof.

302 Example Evaluate

$$\int_0^1 \int_0^1 \cdots \int_0^1 (x_1 x_2 \cdots x_n) dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n.$$

Solution: We have

$$\int_0^1 \int_0^1 \cdots \int_0^1 (x_1 x_2 \cdots x_n) dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n = \prod_{k=1}^{n} \left(\int_0^1 x_k dx_k \right) = \prod_{k=1}^{n} \frac{1}{2} = \frac{1}{2^n}.$$

303 Example Evaluate

$$\int_0^1 \int_0^1 \cdots \int_0^1 (x_1 + x_2 + \cdots + x_n) \, dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n.$$

Solution: This is

$$\begin{aligned}\int_0^1\int_0^1\cdots\int_0^1\left(\sum_{k=1}^n x_k\right)dx_1\wedge dx_2\wedge\ldots\wedge dx_n &= \sum_{k=1}^n\int_0^1\int_0^1\cdots\int_0^1 x_k dx_1\wedge dx_2\wedge\ldots\wedge dx_n\\ &= \sum_{k=1}^n\frac{1}{2}\\ &= \frac{n}{2}.\end{aligned}$$

304 Example (Putnam Exam 1965) Evaluate

$$\lim_{n\to+\infty}\int_0^1\int_0^1\cdots\int_0^1\cos^2\left(\frac{\pi}{2n}(x_1+x_2+\cdots+x_n)\right)dx_1\,dx_2\ldots\,dx_n.$$

Solution: Make the change of variables $x_k = 1 - y_k$. Then

$$I=\int_0^1\int_0^1\cdots\int_0^1\cos^2\left(\frac{\pi}{2n}(x_1+x_2+\cdots+x_n)\right)dx_1\,dx_2\ldots\,dx_n$$

equals

$$\int_0^1\int_0^1\cdots\int_0^1\sin^2\left(\frac{\pi}{2n}(y_1+y_2+\cdots+y_n)\right)dy_1\,dy_2\ldots\,dy_n.$$

Since $\sin^2 t+\cos^2 t=1$, we have $2I=1$, and so $I=\dfrac{1}{2}$.

3.6 Change of Variables in $\bigwedge^2(\mathbb{R}^2)$

We now perform a multidimensional analogue of the change of variables theorem in one variable.

305 Theorem Let $(D,\Delta)\in(\mathbb{R}^n)^2$ be open, bounded sets in $\mathbb{R}^n$ with volume and let $g:\Delta\to D$ be a continuously differentiable bijective mapping such that $\det \mathscr{J}_{\vec{u}}g\neq 0$, and both $|\det \mathscr{J}_{\vec{u}}g|$, $\dfrac{1}{|\det \mathscr{J}_{\vec{u}}g|}$ are bounded on Δ. For $f:D\to\mathbb{R}$ bounded and integrable, $f\circ g|\det \mathscr{J}_{\vec{u}}g|$ is integrable on Δ and

$$\int\cdots\int_D f=\int\cdots\int_\Delta f\circ g|\det \mathscr{J}_{\vec{u}}g|,$$

that is

$$\int \cdots \int_D f(x_1, x_2, \ldots, x_n) dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n$$
$$= \int \cdots \int_\Delta f(g(u_1, u_2, \ldots, u_n)) |\det \mathscr{J}_{\vec{u}} g| du_1 \wedge du_2 \wedge \ldots \wedge du_n.$$

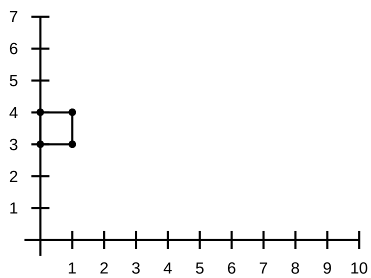

Figure 3.8: Example 306. xy-plane.

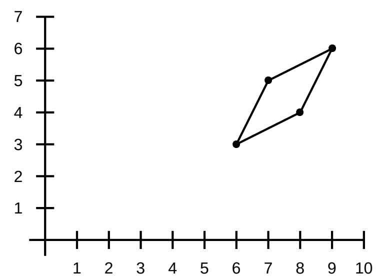

Figure 3.9: Example 306. uv-plane.

One normally chooses changes of variables that map into rectangular regions, or that simplify the integrand. Let us start with a rather trivial example.

306 Example Evaluate the integral

$$\int_3^4 \int_0^1 (x + 2y)(2x + y) dx \wedge dy.$$

Solution: Observe that we have already computed this integral in example 282. Put

$$u = x + 2y \implies du = dx + 2dy,$$
$$v = 2x + y \implies dv = 2dx + dy,$$

giving

$$du \wedge dv = -3dx \wedge dy.$$

Now,

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

is a linear transformation, and hence it maps quadrilaterals into quadrilaterals. The corners of the rectangle in the area of integration in the xy-plane are $(0,3)$, $(1,3)$, $(1,4)$, and $(0,4)$, (traversed counter-clockwise) and they map into $(6,3)$, $(7,5)$, $(9,6)$, and $(8,4)$, respectively, in the uv-plane (see figure 3.9). The form $dx \wedge dy$ has opposite orientation to $du \wedge dv$ so we use

$$dv \wedge du = 3dx \wedge dy$$

instead. The integral sought is

$$\frac{1}{3}\iint\limits_{P} uv \, dv \wedge du = \frac{409}{12},$$

from example 285.

307 Example The integral

$$\iint_{[0;1]^2} (x^4 - y^4)dx \wedge dy = \int_0^1 \left(\frac{1}{5} - y^4\right) dy = 0.$$

Evaluate it using the change of variables $u = x^2 - y^2, v = 2xy$.

Solution: First we find

$$du = 2xdx - 2ydy,$$
$$dv = 2ydx + 2xdy,$$

and so

$$du \wedge dv = (4x^2 + 4y^2)dx \wedge dy.$$

We now determine the region Δ into which the square $D = [0;1]^2$ is mapped. We use the fact that boundaries will be mapped into boundaries. Put

$$AB = \{(x, 0) : 0 \le x \le 1\},$$

$$BC = \{(1, y) : 0 \leq y \leq 1\},$$

$$CD = \{(1 - x, 1) : 0 \leq x \leq 1\},$$

$$DA = \{(0, 1 - y) : 0 \leq y \leq 1\}.$$

On AB we have $u = x, v = 0$. Since $0 \leq x \leq 1$, AB is thus mapped into the line segment $0 \leq u \leq 1, v = 0$.

On BC we have $u = 1 - y^2, v = 2y$. Thus $u = 1 - \frac{v^2}{4}$. Hence BC is mapped to the portion of the parabola $u = 1 - \frac{v^2}{4}, 0 \leq v \leq 2$.

On CD we have $u = (1 - x)^2 - 1, v = 2(1 - x)$. This means that $u = \frac{v^2}{4} - 1, 0 \leq v \leq 2$.

Finally, on DA, we have $u = -(1 - y)^2, v = 0$. Since $0 \leq y \leq 1$, DA is mapped into the line segment $-1 \leq u \leq 0, v = 0$. The region Δ is thus the area in the uv plane enclosed by the parabolas $u \leq \frac{v^2}{4} - 1, u \leq 1 - \frac{v^2}{4}$ with $-1 \leq u \leq 1, 0 \leq v \leq 2$.

We deduce that

$$\begin{aligned}
\iint_{[0;1]^2} (x^4 - y^4) dx \wedge dy &= \iint_\Delta (x^4 - y^4) \frac{1}{4(x^2 + y^2)} du \wedge dv \\
&= \frac{1}{4} \iint_\Delta (x^2 - y^2) du \wedge dv \\
&= \frac{1}{4} \iint_\Delta u du \wedge dv \\
&= \frac{1}{4} \int_0^2 \left(\int_{v^2/4-1}^{1-v^2/4} u du \right) dv \\
&= 0,
\end{aligned}$$

as before.

308 Example Find

$$\iint_D e^{(x^3+y^3)/xy} \quad dx \wedge dy$$

where

$$D = \{(x, y) \in \mathbb{R}^2 | y^2 - 2px \leq 0, x^2 - 2py \leq 0, p \in]0; +\infty[\text{ fixed}\},$$

using the change of variables $x = u^2v,\ y = uv^2$.

Solution: We have

$$dx = 2uv du + u^2 dv,$$

$$dy = v^2 du + 2uv dv,$$

$$dx \wedge dy = 3u^2v^2 du \wedge dv.$$

The region transforms into

$$\Delta = \{(u, v) \in \mathbb{R}^2 | 0 \leq u \leq (2p)^{1/3},\ 0 \leq v \leq (2p)^{1/3}\}.$$

The integral becomes

$$\begin{aligned} \iint_D f(x, y) dx \wedge dy &= \iint_\Delta \exp\left(\frac{u^6v^3 + u^3v^6}{u^3v^3}\right)(3u^2v^2)\, du \wedge dv \\ &= 3\iint_\Delta e^{u^3}e^{v^3}u^2v^2\, du \wedge dv \\ &= \frac{1}{3}\left(\int_0^{(2p)^{1/3}} 3u^2e^{u^3}\, du\right)^2 \\ &= \frac{1}{3}(e^{2p} - 1)^2. \end{aligned}$$

As an exercise, you may try the (more natural) substitution $x^3 = u^2v, y^3 = v^2u$ and verify that the same result is obtained.

309 Example Find $\iint_D f(x, y) dx \wedge dy$ where

$$D = \{(x, y) \in \mathbb{R}^2 | a \leq xy \leq b, y \geq x \geq 0, y^2 - x^2 \leq 1, (a, b) \in \mathbb{R}^2, 0 < a < b\}$$

and $f(x, y) = y^4 - x^4$ by using the change of variables $u = xy, v = y^2 - x^2$.

Solution: Here we argue that

$$du = ydx + xdy,$$

$$dv = -2xdx + 2ydy.$$

Taking the wedge product of differential forms,

$$du \wedge dv = 2(y^2 + x^2)dx \wedge dy.$$

Hence

$$\begin{aligned} f(x,y)dx \wedge dy &= (y^4 - x^4)\frac{1}{2(y^2+x^2)}du \wedge dv \\ &= \frac{1}{2}(y^2 - x^2)du \wedge dv \\ &= \frac{v}{2}du \wedge dv \end{aligned}$$

The region transforms into

$$\Delta = [a; b] \times [0; 1].$$

The integral becomes

$$\begin{aligned} \iint_D f(x,y)dx \wedge dy &= \iint_\Delta v\, du \wedge dv \\ &= \frac{1}{2}\left(\int_a^b du\right)\left(\int_0^1 v\, dv\right) \\ &= \frac{b-a}{4}. \end{aligned}$$

310 Example Let $D' = \{(u,v) \in \mathbb{R}^2 : u \leq 1, -u \leq v \leq u\}$. Consider

$$\Phi : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (u,v) & \mapsto & \left(\dfrac{u+v}{2}, \dfrac{u-v}{2}\right) \end{array}.$$

❶ Find the image of Φ on D', that is, find $D = \Phi(D')$.

❷ Find

$$\iint_D (x+y)^2 e^{x^2-y^2} dx \wedge dy.$$

Solution:

❶ Put $x = \dfrac{u+v}{2}$ and $y = \dfrac{u-v}{2}$. Then $x+y=u$ and $x-y$. Observe that D' is the triangle in the uv plane bounded by the lines $u=0, u=1, v=u, v=-u$. Its image under Φ is the triangle bounded by the equations $x=0, y=0, x+y=1$. Clearly also

$$dx \wedge dy = \frac{1}{2} du \wedge dv.$$

❷ From the above

$$\begin{aligned}\iint_D (x+y)^2 e^{x^2-y^2} dx \wedge dy &= \frac{1}{2}\iint_{D'} u^2 e^{uv} du \wedge dv \\ &= \frac{1}{2}\int_0^1 \int_{-u}^{u} u^2 e^{uv} du dv \\ &= \frac{1}{2}\int_0^1 u(e^{u^2} - e^{-u^2}) du \\ &= \frac{1}{4}(e + e^{-1} - 2).\end{aligned}$$

311 Example Use the following steps (due to Tom Apostol) in order to prove that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

❶ Use the series expansion

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \cdots \quad |t| < 1,$$

in order to prove (formally) that

$$\int_0^1 \int_0^1 \frac{dxdy}{1-xy} = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

❷ Use the change of variables $u = x + y, v = x - y$ to shew that

$$\int_0^1\int_0^1 \frac{dxdy}{1-xy} = 2\int_0^1 \left(\int_{-u}^{u} \frac{dv}{4-u^2+v^2}\right) du + 2\int_1^2 \left(\int_{u-2}^{2-u} \frac{dv}{4-u^2+v^2}\right) du.$$

❸ Shew that the above integral reduces to

$$2\int_0^1 \frac{2}{\sqrt{4-u^2}} \arctan \frac{u}{\sqrt{4-u^2}}\, du + 2\int_1^2 \frac{2}{\sqrt{4-u^2}} \arctan \frac{2-u}{\sqrt{4-u^2}}\, du.$$

❹ Finally, prove that the above integral is $\frac{\pi^2}{6}$ by using the substitution $\theta = \arcsin \frac{u}{2}$.

Solution:

❶ Formally,

$$\begin{aligned}
\int_0^1\int_0^1 \frac{dxdy}{1-xy} &= \int_0^1\int_0^1 (1 + xy + x^2y^2 + x^3y^3 + \cdots)dxdy \\
&= \int_0^1 \left(y + \frac{xy^2}{2} + \frac{x^2y^3}{3} + \frac{x^3y^4}{4} + \cdots\right)_0^1 dx \\
&= \int_0^1 (1 + \frac{x}{2} + \frac{x^2}{3} + \frac{x^3}{4} + \cdots)dx \\
&= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots
\end{aligned}$$

❷ This change of variables transforms the square $[0;1] \times [0;1]$ in the xy plane into the square with vertices at $(0,0)$, $(1,1)$, $(2,0)$, and $(1,-1)$ in the uv plane. We will split this region of integration into two disjoint triangles: T_1 with vertices at $(0,0)$, $(1,1)$, $(1,-1)$, and T_2 with vertices at $(1,-1)$, $(1,1)$, $(2,0)$. Observe that

$$dx \wedge dy = \frac{1}{2} du \wedge dv,$$

and that $u + v = 2x, u - v = 2y$ and so $4xy = u^2 - v^2$. The integral becomes

$$\int_0^1 \int_0^1 \frac{dxdy}{1 - xy} = \frac{1}{2} \iint_{T_1 \cup T_2} \frac{du \wedge dv}{1 - \frac{u^2 - v^2}{4}}$$
$$= 2\int_0^1 \left(\int_{-u}^{u} \frac{dv}{4 - u^2 + v^2} \right) du + 2\int_1^2 \left(\int_{u-2}^{2-u} \frac{dv}{4 - u^2 + v^2} \right) du,$$

as desired.

❸ This follows by using the identity

$$\int_0^t \frac{d\Omega}{1 + \Omega^2} = \arctan t.$$

❹ This is straightforward but tedious!

3.7 Change to Polar Co-ördinates

One of the most common changes of variable is the passage to polar co-ördinates where

$$x = \rho \cos\theta \implies dx = \cos\theta d\rho - \rho \sin\theta d\theta,$$

$$y = \rho \sin\theta \implies dy = \sin\theta d\rho + \rho \cos\theta d\theta,$$

whence

$$dx \wedge dy = (\rho \cos^2\theta + \rho \sin^2\theta) d\rho \wedge d\theta = \rho d\rho \wedge d\theta.$$

312 Example Find

$$\iint_D xy\sqrt{x^2 + y^2} dx \wedge dy$$

where

$$D = \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, y \leq x, x^2 + y^2 \leq 1\}.$$

We use polar co-ördinates. The region D transforms into the region

$$\Delta = [0;1] \times \left[0; \frac{\pi}{4}\right].$$

Therefore the integral becomes

$$\begin{aligned}\iint_{\Delta} \rho^4 \cos\theta \sin\theta \, d\rho \wedge d\theta &= \left(\int_0^{\pi/4} \cos\theta \sin\theta \, d\theta\right)\left(\int_0^1 \rho^4 \, d\rho\right) \\ &= \frac{1}{20}.\end{aligned}$$

313 Example Find

$$\iint_D \sqrt{x^2+y^2} dx \wedge dy$$

where

$$D = \{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, x^2+y^2 \leq 1, x^2+y^2-2y \geq 0\}.$$

Using polar co-ördinates,

$$\begin{aligned}\iint_D f(x,y) dx \wedge dy &= \int_0^{\pi/6} \left(\int_{2\sin\theta}^1 \rho^2 \, d\rho\right) d\theta \\ &= \frac{1}{3}\int_0^{\pi/6} (1 - 8\sin^3\theta) \, d\theta \\ &= \frac{\pi}{18} - \frac{16}{9} + \sqrt{3}.\end{aligned}$$

314 Example Find

$$\iint_D (x^2 - y^2) dx \wedge dy$$

where

$$D = \{(x,y) \in \mathbb{R}^2 | (x-1)^2 + y^2 \leq 1\}.$$

Solution: Using polar co-ördinates,

$$\begin{aligned}\iint_D x^2 - y^2 dx \wedge dy &= \int_{-\pi/2}^{\pi/2} \left(\int_0^{2\cos\theta} \rho^3 \, d\rho\right) (\cos^2\theta - \sin^2\theta) d\theta \\ &= 8\int_0^{\pi/2} \cos^4\theta(\cos^2\theta - \sin^2\theta) \, d\theta \\ &= \pi.\end{aligned}$$

315 Example Find

$$\iint_D \sqrt{xy} dx \wedge dy$$

where

$$D = \{(x, y) \in \mathbb{R}^2 | (x^2 + y^2)^2 \leq 2xy\}.$$

Solution: Using polar co-ördinates,

$$\begin{aligned}\iint_D \sqrt{xy} dx \wedge dy &= 4\int_0^{\pi/4} \left(\int_0^{\sqrt{\sin 2\theta}} \rho\sqrt{\rho^2 \cos\theta \sin\theta} \, d\rho\right) d\theta \\ &= \frac{4}{3}\int_0^{\pi/4} (\sqrt{\sin 2\theta})^3 \sqrt{\cos\theta \sin\theta} \, d\theta \\ &= \frac{4}{3\sqrt{2}}\int_0^{\pi/4} \sin^2 2\theta \, d\theta \\ &= \frac{\pi\sqrt{2}}{12}.\end{aligned}$$

316 Example Find $\iint_D f(x, y) dx \wedge dy$ where

$$D = \{(x, y) \in \mathbb{R}^2 : b^2x^2 + a^2y^2 = a^2b^2, (a, b) \in]0; +\infty[\text{ fixed}\}$$

and $f(x, y) = x^3 + y^3$.

Solution: Using $x = a\rho\cos\theta$, $y = b\rho\sin\theta$, the integral becomes

$$(ab)\left(\int_0^{2\pi} a^3\cos^3\theta + b^3\sin^3\theta d\theta\right)\left(\int_0^1 \rho^4 d\rho\right) = \frac{2}{15}(ab)(a^3 + b^3).$$

317 Example Find $\iint_D f(x,y)dx \wedge dy$ where

$$D = \{(x,y) \in \mathbb{R}^2 | y \geq 0, x^2 + y^2 - 2x \leq 0\}$$

and $f(x,y) = x^2 y$.

Solution: Using polar co-ördinates the integral becomes

$$\int_0^{\pi/2} \left(\int_0^{2\cos\theta} \rho^4 d\rho \right) \cos^2\theta \sin\theta d\theta = \frac{4}{5}.$$

318 Example Find $\iint_D f(x,y)dx \wedge dy$ where

$$D = \{(x,y) \in \mathbb{R}^2 | x \geq 1, x^2 + y^2 - 2x \leq 0\}$$

and $f(x,y) = \frac{1}{(x^2+y^2)^2}$.

Solution: Using polar co-ördinates the integral becomes

$$\int_{-\pi/4}^{\pi/4} \left(\int_{1/\cos\theta}^{2\cos\theta} \frac{1}{\rho^3} d\rho \right) d\theta = \int_0^{\pi/4} \left(\cos^2\theta - \frac{\sec^2\theta}{4} \right) d\theta = \frac{\pi}{8}.$$

319 Example Find $\iint_D e^{-x^2-xy-y^2}\, dxdy$, where

$$D = \{(x,y) \in \mathbb{R}^2 : x^2 + xy + y^2 \leq 1\}.$$

Solution: Completing squares

$$x^2 + xy + y^2 = \left(x + \frac{y}{2}\right)^2 + \left(\frac{\sqrt{3}y}{2}\right)^2.$$

Put $U = x + \frac{y}{2}$, $V = \frac{\sqrt{3}y}{2}$. The integral becomes

$$\iint_{\{x^2+xy+y^2\leq 1\}} e^{-x^2-xy-y^2} dxdy = \frac{2}{\sqrt{3}} \iint_{\{U^2+V^2\leq 1\}} e^{-(U^2+V^2)} dUdV.$$

Passing to polar co-ördinates, the above equals

$$\frac{2}{\sqrt{3}}\int_0^{2\pi}\int_0^1 \rho e^{-\rho^2} d\rho d\theta = \frac{2\pi}{\sqrt{3}}(1-e^{-1}).$$

320 Example Let

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - y \leq 0, x^2 + y^2 - x \leq 0\}.$$

Find the integral

$$\iint_D (x+y)^2 dx \wedge dy.$$

Solution: Put

$$D' = \{(x, y) \in \mathbb{R}^2 : y \geq x, x^2 + y^2 - y \leq 0, x^2 + y^2 - x \leq 0\}.$$

Then the integral equals

$$2\iint_{D'} (x+y)^2 dx \wedge dy.$$

Using polar co-ördinates the integral equals

$$\begin{aligned} 2\int_{\pi/4}^{\pi/2}(\cos\theta + \sin\theta)^2\left(\int_0^{\cos\theta}\rho^3 d\rho\right)d\theta &= \frac{1}{2}\int_{\pi/4}^{\pi/2}\cos^4\theta(1 + 2\sin\theta\cos\theta)d\theta \\ &= \frac{3\pi}{64} - \frac{5}{48}. \end{aligned}$$

321 Example Let $D = \{(x, y) \in \mathbb{R}^2 | y \leq x^2 + y^2 \leq 1\}$. Compute

$$\iint_D \frac{dx \wedge dy}{(1 + x^2 + y^2)^2}.$$

Solution: Observe that $D = D_2 \setminus D_1$ where D_2 is the disk limited by the equation $x^2 + y^2 = 1$ and D_1 is the disk limited by the equation $x^2 + y^2 = y$. Hence

$$\iint_D \frac{dx \wedge dy}{(1 + x^2 + y^2)^2} = \iint_{D_2} \frac{dx \wedge dy}{(1 + x^2 + y^2)^2} - \iint_{D_1} \frac{dx \wedge dy}{(1 + x^2 + y^2)^2}.$$

Using polar co-ördinates we have

$$\iint_{D_2} \frac{dx \wedge dy}{(1+x^2+y^2)^2} = \int_0^{2\pi}\int_0^1 \frac{\rho}{(1+\rho^2)^2} d\rho d\theta = \frac{\pi}{2}$$

and

$$\begin{aligned}\iint_{D_1} \frac{dx \wedge dy}{(1+x^2+y^2)^2} &= 2\int_0^{\pi/2}\int_0^{\sin\theta} \frac{\rho}{(1+\rho^2)^2} d\rho d\theta = \int_0^{\pi/2} \frac{\sin^2\theta d\theta}{1+\sin^2\theta} \\ &= \int_0^{+\infty} \frac{dt}{t^2+1} - \frac{dt}{2t^2+1} = \frac{\pi}{2} - \frac{\pi\sqrt{2}}{4}.\end{aligned}$$

(We evaluated this last integral using $t = \tan\theta$) Finally, the integral equals

$$\frac{\pi}{2} - \left(\frac{\pi}{2} - \frac{\pi\sqrt{2}}{4}\right) = \frac{\pi\sqrt{2}}{4}.$$

322 Example Evaluate

$$\iint_{\{(x,y)\in\mathbb{R}^2 : x\geq 0, y\geq 0, x^4+y^4\leq 1\}} x^3y^3\sqrt{1-x^4-y^4}\, dx \wedge dy$$

using $x^2 = \rho\cos\theta$, $y^2 = \rho\sin\theta$.

Solution: We have

$$2xdx = \cos\theta d\rho - \rho\sin\theta d\theta, \quad 2ydy = \sin\theta d\rho + \rho\cos\theta d\theta,$$

whence

$$4xy dx \wedge dy = \rho d\rho \wedge d\theta.$$

It follows that

$$\begin{aligned}x^3y^3\sqrt{1-x^4-y^4}\, dx \wedge dy &= \frac{1}{4}(x^2y^2)(\sqrt{1-x^4-y^4})(4xy\, dx \wedge dy) \\ &= \frac{1}{4}(\rho^3\cos\theta\sin\theta\sqrt{1-\rho^2}) d\rho \wedge d\theta\end{aligned}$$

Observe that

$$x^4 + y^4 \leq 1 \implies \rho^2\cos^2\theta + \rho^2\sin^2\theta \leq 1 \implies \rho \leq 1.$$

Since the integration takes place on the first quadrant, we have $0 \leq \theta \leq \pi/2$. Hence the integral becomes

$$\begin{aligned}\int_0^{\pi/2}\int_0^1 \frac{1}{4}(\rho^3\cos\theta\sin\theta\sqrt{1-\rho^2})d\rho\wedge d\theta &= \frac{1}{4}\left(\int_0^{\pi/2}\cos\theta\sin\theta d\theta\right)\left(\int_0^1\rho^3\sqrt{1-\rho^2}d\rho\right)\\ &= \frac{1}{4}\cdot\frac{1}{2}\cdot\frac{2}{15}\\ &= \frac{1}{60}.\end{aligned}$$

323 Example William Thompson (Lord Kelvin) is credited to have said: "A mathematician is someone to whom

$$\int_0^{+\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$$

is as obvious as twice two is four to you. Liouville was a mathematician." Prove that

$$\int_0^{+\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$$

by following these steps.

❶ Let $a > 0$ be a real number and put $D_a = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 \leq a^2\}$. Find

$$I_a = \iint_{D_a} e^{-(x^2+y^2)}\, dx \wedge dy.$$

❷ Let $a > 0$ be a real number and put $\Delta_a = \{(x,y) \in \mathbb{R}^2 | |x| \leq a, |y| \leq a\}$. Let

$$J_a = \iint_{\Delta_a} e^{-(x^2+y^2)}\, dx \wedge dy.$$

Prove that

$$I_a \leq J_a \leq I_{a\sqrt{2}}.$$

❸ Deduce that

$$\int_0^{+\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

Solution:

❶ Using polar co-ördinates

$$I_a = \int_0^{2\pi} \left(\int_0^a \rho e^{-\rho^2} d\rho \right) d\theta = \pi(1 - e^{-a^2}).$$

❷ The domain of integration of J_a is a square of side $2a$ centred at the origin. The respective domains of integration of I_a and $I_{a\sqrt{2}}$ are the inscribed and the exscribed circles to the square.

❸ First observe that

$$J_a = \left(\int_{-a}^{a} e^{-x^2} dx \right)^2 .$$

Since both I_a and $I_{a\sqrt{2}}$ tend to π as $a \to +\infty$, we deduce that $J_a \to \pi$. This gives the result.

Recall from formula 3.2 that the area enclosed by a simple closed curve Γ is given by

$$\frac{1}{2} \int_\Gamma x dy - y dx.$$

Using polar co-ördinates

$$\begin{aligned} x dy - y dx &= (\rho \cos\theta)(\sin\theta d\rho + \rho\cos\theta d\theta) - (\rho \sin\theta)(\cos\theta d\rho - \rho \sin\theta d\theta) \\ &= \rho^2 d\theta. \end{aligned}$$

This will be used in the next problem.

324 Example Prove that every closed convex region in the plane of area $\geq \pi$ has two points which are two units apart.

Solution: Parametrise the curve enclosing the region by polar co-ördinates so that the region is tangent to the polar axis at the origin. Let the equation of the curve be $\rho = f(\theta)$. The area of the region is then given by

$$\frac{1}{2} \int_0^{\pi} \rho^2 d\theta = \frac{1}{2} \int_0^{\pi} (f(\theta))^2 d\theta = \frac{1}{2} \int_0^{\pi/2} ((f(\theta))^2 + (f(\theta + \pi/2))^2) d\theta.$$

By the Pythagorean Theorem, the integral above is the integral of the square of the chord in question. If no two points are farther than 2 units, their squares are no farther than 4 units, and so the area

$$< \frac{1}{2}\int_0^{\pi/2} 4\mathrm{d}\theta = \pi,$$

a contradiction.

325 Example (Putnam Exam 1976) In the xy-plane, if R is the set of points inside and on a convex polygon, let $D(x, y)$ be the distance from (x, y) to the nearest point R. Show that

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-D(x,y)}\,\mathrm{d}x\wedge\mathrm{d}y = 2\pi + L + A,$$

where L is the perimeter of R and A is the area of R.

Solution: Let $I(S)$ denote the integral sought over a region S. Since $D(x, y) = 0$ inside R, $I(R) = A$. Let $\mathscr{L}$ be a side of R with length l and let $S(\mathscr{L})$ be the half strip consisting of the points of the plane having a point on $\mathscr{L}$ as nearest point of R. Set up co-ördinates uv so that u is measured parallel to $\mathscr{L}$ and v is measured perpendicular to L. Then

$$I(S(\mathscr{L})) = \int_0^{l}\int_0^{+\infty} e^{-v}\,\mathrm{d}u\wedge\mathrm{d}v = l.$$

The sum of these integrals over all the sides of R is L.

If $\mathscr{V}$ is a vertex of R, the points that have $\mathscr{V}$ as nearest from R lie inside an angle $S(\mathscr{V})$ bounded by the rays from $\mathscr{V}$ perpendicular to the edges meeting at $\mathscr{V}$. If α is the measure of that angle, then using polar co-ördinates

$$I(S(\mathscr{V})) = \int_0^{\alpha}\int_0^{+\infty} \rho e^{-\rho}\,\mathrm{d}\rho\wedge\mathrm{d}\theta = \alpha.$$

The sum of these integrals over all the vertices of R is 2π. Assembling all these integrals we deduce the result.

3.8 Integrating in $\bigwedge^3(\mathbb{R}^3)$

326 Definition A 3-*dimensional oriented manifold of* $\mathbb{R}^3$ is simply an open set (body) $V \in \mathbb{R}^3$, where the $+$ orientation is in the direction of the outward pointing normal to the body, and the $-$ orientation is in the direction of the inward pointing normal to the body. A general oriented 3-manifold is a union of open sets.

The region $-V$ has opposite orientation to D and

$$\int_{-V} \omega = -\int_{V} \omega.$$

In this section, unless otherwise noticed, we will choose the positive orientation for the regions considered. This corresponds to using the volume form $dx \wedge dy \wedge dz$.

Let $V \subseteq \mathbb{R}^3$. Given a function $f : V \to \mathbb{R}$, the integral

$$\iint\limits_{V} f(x, y, z) dx \wedge dy \wedge dz$$

is the sum of all the values of f restricted to V. In particular,

$$\iint\limits_{V} dx \wedge dy \wedge dz$$

is the oriented volume of V.

327 Example Find

$$\iiint\limits_{[0;1]^3} x^2 y e^{xyz}\, dx \wedge dy \wedge dz.$$

Solution: The integral is

$$\begin{aligned}
\int_0^1 \left(\int_0^1 \left(\int_0^1 x^2 y e^{xyz}\, dz\right) dy\right) dx &= \int_0^1 \left(\int_0^1 x(e^{xy} - 1)\, dy\right) dx \\
&= \int_0^1 (e^x - x - 1) dx \\
&= e - \frac{5}{2}.
\end{aligned}$$

328 Example Find $\iiint\limits_R z\,dx \wedge dy \wedge dz$ if

$$R = \{(x, y, z) \in \mathbb{R}^3 | x \geq 0, y \geq 0, z \geq 0, \sqrt{x} + \sqrt{y} + \sqrt{z} \leq 1\}.$$

Solution: The integral is

$$\begin{aligned}
\iiint_R z dx \wedge dy \wedge dz &= \int_0^1 z \left(\int_0^{(1-\sqrt{z})^2} \left(\int_0^{(1-\sqrt{z}-\sqrt{x})^2} dy \right) dx \right) dz \\
&= \int_0^1 z \left(\int_0^{(1-\sqrt{z})^2} (1 - \sqrt{z} - \sqrt{x})^2 dx \right) dz \\
&= \frac{1}{6} \int_0^1 z(1 - \sqrt{z})^4\, dz \\
&= \frac{1}{840}.
\end{aligned}$$

329 Example Prove that

$$\iiint\limits_V x dx dy dz = \frac{a^2bc}{24},$$

where V is the tetrahedron

$$V = \left\{ (x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, \frac{x}{a} + \frac{y}{b} + \frac{z}{c} \leq 1 \right\}.$$

Solution: We have

$$\begin{aligned}
\iiint\limits_V x dx dy dz &= \int_0^c \int_0^{b-bz/c} \int_0^{a-ay/b-az/c} x dx dy dz \\
&= \frac{1}{2} \int_0^c \int_0^{b-bz/c} \left(a - \frac{ay}{b} - \frac{az}{c} \right)^2 dy dz \\
&= \frac{1}{6} \int_0^c \frac{a^2 (-z + c)^3 b}{c^3} dx \\
&= \frac{a^2bc}{24}
\end{aligned}$$

3.9 Change of Variables in $\bigwedge^3(\mathbb{R}^3)$

330 Example Find

$$\iiint_R (x+y+z)(x+y-z)(x-y-z)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z,$$

where R is the tetrahedron bounded by the planes $x+y+z=0$, $x+y-z=0$, $x-y-z=0$, and $2x-z=1$.

Solution: We make the change of variables

$$u = x+y+z \implies \mathrm{d}u = \mathrm{d}x + \mathrm{d}y + \mathrm{d}z,$$

$$v = x+y-z \implies \mathrm{d}v = \mathrm{d}x + \mathrm{d}y - \mathrm{d}z,$$

$$w = x-y-z \implies \mathrm{d}w = \mathrm{d}x - \mathrm{d}y - \mathrm{d}z.$$

This gives

$$\mathrm{d}u \wedge \mathrm{d}v \wedge \mathrm{d}w = -4\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z.$$

These forms have opposite orientations, so we choose, say,

$$\mathrm{d}u \wedge \mathrm{d}w \wedge \mathrm{d}v = 4\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$$

which have the same orientation. Also,

$$2x - z = 1 \implies u + v + 2w = 2.$$

The tetrahedron in the xyz-co-ördinate frame is mapped into a tetrahedron bounded by $u=0$, $v=0$, $u+v+2w=1$ in the uvw-co-ördinate frame. The integral becomes

$$\frac{1}{4}\int_0^2\int_0^{1-v/2}\int_0^{2-v-2w} uvw \ \mathrm{d}u \wedge \mathrm{d}w \wedge \mathrm{d}v = \frac{1}{180}.$$

Consider a transformation to cylindrical co-ördinates

$$(x, y, z) = (\rho\cos\theta, \rho\sin\theta, z).$$

From what we know about polar co-ördinates

$$\mathrm{d}x \wedge \mathrm{d}y = \rho \mathrm{d}\rho \wedge \mathrm{d}\theta.$$

Since the wedge product of forms is associative,

$$dx \wedge dy \wedge dz = \rho d\rho \wedge d\theta \wedge dz.$$

331 Example Find $\iiint_R z^2 dx \wedge dy \wedge dz$ if

$$R = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 \leq 1, 0 \leq z \leq 1\}.$$

Solution: The region of integration is mapped into

$$\Delta = [0; 2\pi] \times [0; 1] \times [0; 1]$$

through a cylindrical co-ördinate change. The integral is therefore

$$\begin{aligned} \iiint_R f(x, y, z) dx \wedge dy \wedge dz &= \left(\int_0^{2\pi} d\theta\right)\left(\int_0^1 \rho \, d\rho\right)\left(\int_0^1 z^2 \, dz\right) \\ &= \frac{\pi}{3}. \end{aligned}$$

332 Example Evaluate $\iiint_D (x^2+y^2) dx \wedge dy \wedge dz$ over the first octant region bounded by the cylinders $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ and the planes $z = 0, z = 1, x = 0, x = y$.

Solution: The integral is

$$\int_0^1 \int_{\pi/4}^{\pi/2} \int_1^2 \rho^3 d\rho \wedge d\theta \wedge dz = \frac{15\pi}{16}.$$

333 Example Three long cylinders of radius R intersect at right angles. Find the volume of their intersection.

Solution: Let V be the desired volume. By symmetry, $V = 2^4 V'$, where

$$V' = \iiint_{D'} dx \wedge dy \wedge dz,$$

$$D' = \{(x, y, z) \in \mathbb{R}^3 : 0 \le y \le x, 0 \le z, x^2+y^2 \le R^2, y^2+z^2 \le R^2, z^2+x^2 \le R^2\}.$$

In this case it is easier to integrate with respect to z first. Using cylindrical co-ördinates

$$\Delta' = \left\{(\theta, \rho, z) \in \left[0; \frac{\pi}{4}\right] \times [0; R] \times [0; +\infty[, 0 \le z \le \sqrt{R^2 - \rho^2 \cos^2\theta}\right\}.$$

Now,

$$\begin{aligned}
V' &= \int_0^{\pi/4} \left(\int_0^R \left(\int_0^{\sqrt{R^2-\rho^2\cos^2\theta}} dz\right) \rho d\rho\right) d\theta \\
&= \int_0^{\pi/4} \left(\int_0^R \rho\sqrt{R^2-\rho^2\cos^2\theta} d\rho\right) d\theta \\
&= \int_0^{\pi/4} -\frac{1}{3\cos^2\theta} \left[(R^2-\rho^2\cos^2\theta)^{3/2}\right]_0^R d\theta \\
&= \frac{R^3}{3}\int_0^{\pi/4} \frac{1-\sin^3\theta}{\cos^2\theta} d\theta \\
u = \overset{=}{\cos}\theta \quad & \frac{R^3}{3}\left([\tan\theta]_0^{\pi/4} + \int_1^{\frac{\sqrt{2}}{2}} \frac{1-u^2}{u^2} du\right) \\
&= \frac{R^3}{3}\left(1 - \left[u^{-1}+u\right]_1^{\frac{\sqrt{2}}{2}}\right) \\
&= \frac{\sqrt{2}-1}{\sqrt{2}} R^3.
\end{aligned}$$

Finally

$$V = 16V' = 8(2-\sqrt{2})R^3.$$

Consider now a change to spherical co-ördinates

$$x = \rho\cos\theta\sin\phi, \quad y = \rho\sin\theta\sin\phi, \quad z = \rho\cos\phi.$$

We have

$$\begin{aligned}
dx &= \cos\theta\sin\phi d\rho - \rho\sin\theta\sin\phi d\theta + \rho\cos\theta\cos\phi d\phi, \\
dy &= \sin\theta\sin\phi d\rho + \rho\cos\theta\sin\phi d\theta + \rho\sin\theta\cos\phi d\phi, \\
dz &= \cos\phi d\rho - \rho\sin\phi d\phi.
\end{aligned}$$

This gives

$$dx \wedge dy \wedge dz = -\rho^2 \sin\theta d\rho \wedge d\theta \wedge d\phi.$$

From this derivation, the form $d\rho \wedge d\theta \wedge d\phi$ is negatively oriented, and so we choose

$$dx \wedge dy \wedge dz = \rho^2 \sin\theta d\rho \wedge d\phi \wedge d\theta$$

instead.

334 Example Let $(a, b, c) \in]0; +\infty[^3$ be fixed. Find $\iiint_R xyz dx \wedge dy \wedge dz$ if

$$R = \left\{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1, x \geq 0, y \geq 0, z \geq 0\right\}.$$

We use spherical co-ördinates, where

$$(x, y, z) = (a\rho \cos\theta \sin\phi, b\rho \sin\theta \sin\phi, c\rho \cos\phi).$$

We have

$$dx \wedge dy \wedge dz = abc\rho^2 \sin\phi d\rho \wedge d\phi \wedge d\rho.$$

The integration region is mapped into

$$\Delta = [0; 1] \times [0; \frac{\pi}{2}] \times [0; \frac{\pi}{2}].$$

The integral becomes

$$(abc)^2 \left(\int_0^{\pi/2} \cos\theta \sin\theta \, d\theta\right) \left(\int_0^1 \rho^5 \, d\rho\right) \left(\int_0^{\pi/2} \cos^3\phi \sin\phi \, d\phi\right) = \frac{(abc)^2}{48}.$$

335 Example Let $V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 9, 1 \leq z \leq 2\}$. Then

$$\begin{aligned} \iiint\limits_V dx \wedge dy \wedge dz &= \int_0^{2\pi} \int_{\pi/2-\arcsin 2/3}^{\pi/2-\arcsin 1/3} \int_{1/\cos\phi}^{2/\cos\phi} \rho^2 \sin\phi \, d\rho \wedge d\phi \wedge d\theta \\ &= \frac{63\pi}{4}. \end{aligned}$$

336 Lemma Let m, n be non-negative integers. Then

$$\int_0^1 x^m(1-x)^n \, dx = \frac{m!n!}{(m+n+1)!}.$$

337 Example (Putnam Exam 1984) Find

$$\iiint_R x^1y^9z^8(1-x-y-z)^4 \, dx \wedge dy \wedge dz,$$

where

$$R = \{(x,y,z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, x+y+z \leq 1\}.$$

Solution: We make the change of variables

$$u = x+y+z \implies du = dx+dy+dz,$$

$$uv = y+z \implies udv + vdu = dy + dz,$$

$$uvw = z \implies uvdw + uwdv + vwdu = dz.$$

This gives

$$x = u(1-v),$$

$$y = uv(1-w),$$

$$z = uvw,$$

$$u^2v \, du \wedge dv \wedge dw = dx \wedge dy \wedge dz.$$

To find the limits of integration we observe that the limits of integration using $dx \wedge dy \wedge dz$ are

$$0 \leq z \leq 1, \; 0 \leq y \leq 1-z, \; 0 \leq x \leq 1-y-z.$$

This translates into

$$0 \leq uvw \leq 1, \; 0 \leq uv - uvw \leq 1 - uvw, \; 0 \leq u - uv \leq 1 - uv + uvw - uvw.$$

Thus

$$0 \leq uvw \leq 1, \; 0 \leq uv \leq 1, \; 0 \leq u \leq 1,$$

which finally give

$$0 \leq u \leq 1,\ 0 \leq v \leq 1,\ 0 \leq w \leq 1.$$

The integral sought is then, using Lemma 336,

$$\int_0^1\int_0^1\int_0^1 u^{20}v^{18}w^8(1-u)^4(1-v)(1-w)^9\,\mathrm{d}u\wedge\mathrm{d}v\wedge\mathrm{d}w,$$

which in turn is

$$\left(\int_0^1 u^{20}(1-u)^4\mathrm{d}u\right)\left(\int_0^1 v^{18}(1-v)\mathrm{d}v\right)\left(\int_0^1 w^8(1-w)^9\mathrm{d}w\right) = \frac{1}{265650}\cdot\frac{1}{380}\cdot\frac{1}{437580}$$

which is

$$= \frac{1}{44172388260000}.$$

3.10 Integration in $\bigwedge^2(\mathbb{R}^3)$

338 Definition A 2-*dimensional oriented manifold of* $\mathbb{R}^3$ is simply a smooth surface $D \in \mathbb{R}^3$, where the $+$ orientation is in the direction of the outward normal pointing away from the origin and the $-$ orientation is in the direction of the inward normal pointing towards the origin. A general oriented 2-manifold in $\mathbb{R}^3$ is a union of surfaces.

STOP *The surface $-\Sigma$ has opposite orientation to Σ and*

$$\int_{-\Sigma}\omega = -\int_{\Sigma}\omega.$$

STOP *In this section, unless otherwise noticed, we will choose the positive orientation for the regions considered. This corresponds to using the ordered basis*

$$\{\mathrm{d}y\wedge\mathrm{d}z,\ \mathrm{d}z\wedge\mathrm{d}x,\ \mathrm{d}x\wedge\mathrm{d}y\}.$$

339 Definition Let $f : \mathbb{R}^3 \to \mathbb{R}$. The integral of f over the smooth surface Σ (oriented in the positive sense) is given by the expression

$$\iint_{\Sigma} f \left|\left| d^2\mathbf{x} \right|\right|.$$

Here

$$\left|\left| d^2\mathbf{x} \right|\right| = \sqrt{(dx \wedge dy)^2 + (dz \wedge dx)^2 + (dy \wedge dz)^2}$$

is the *surface area element*.

340 Example Evaluate $\iint_{\Sigma} z \left|\left| d^2\mathbf{x} \right|\right|$ where Σ is the outer surface of the section of the paraboloid $z = x^2 + y^2, 0 \leq z \leq 1$.

Solution: We parametrise the paraboloid as follows. Let $x = u, y = v, z = u^2 + v^2$. Observe that the domain D of Σ is the unit disk $u^2 + v^2 \leq 1$. We see that

$$dx \wedge dy = du \wedge dv,$$

$$dy \wedge dz = -2u du \wedge dv,$$

$$dz \wedge dx = -2v du \wedge dv,$$

and so

$$\left|\left| d^2\mathbf{x} \right|\right| = \sqrt{1 + 4u^2 + 4v^2} du \wedge dv.$$

Now,

$$\iint_{\Sigma} z \left|\left| d^2\mathbf{x} \right|\right| = \iint_{D} (u^2 + v^2)\sqrt{1 + 4u^2 + 4v^2} du \wedge dv.$$

To evaluate this last integral we use polar co-ördinates, and so

$$\begin{aligned} \iint_{D} (u^2 + v^2)\sqrt{1 + 4u^2 + 4v^2} du \wedge dv &= \int_0^{2\pi}\int_0^1 \rho^3\sqrt{1 + 4\rho^2} d\rho d\theta \\ &= \frac{\pi}{12}(5\sqrt{5} + \frac{1}{5}). \end{aligned}$$

341 Example Evaluate $\iint\limits_{\Sigma} y\left|\left|d^2\mathbf{x}\right|\right|$ where Σ is the surface $z = x + y^2, 0 \leq x \leq 1, 0 \leq y \leq 2$.

Solution: We parametrise the surface by letting $x = u, y = v, z = u + v^2$. Observe that the domain D of Σ is the square $[0;1] \times [0;2]$. Observe that

$$dx \wedge dy = du \wedge dv,$$

$$dy \wedge dz = -du \wedge dv,$$

$$dz \wedge dx = -2v du \wedge dv,$$

and so

$$\left|\left|d^2\mathbf{x}\right|\right| = \sqrt{2 + 4v^2} du \wedge dv.$$

The integral becomes

$$\begin{aligned} \iint\limits_{\Sigma} y\left|\left|d^2\mathbf{x}\right|\right| &= \int_0^2 \int_0^1 v\sqrt{2+4v^2} du \wedge dv \\ &= \left(\int_0^1 du\right)\left(\int_0^2 y\sqrt{2+4v^2} dv\right) \\ &= \frac{13\sqrt{2}}{3}. \end{aligned}$$

342 Example Evaluate $\iint\limits_{\Sigma} x^2\left|\left|d^2\mathbf{x}\right|\right|$ where Σ is the surface of the unit sphere $x^2 + y^2 + z^2 = 1$.

Solution: We use spherical co-ördinates, $(x, y, z) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi)$. Here $\theta \in [0; 2\pi]$ is the latitude and $\phi \in [0; \pi]$ is the longitude. Observe that

$$dx \wedge dy = \sin\phi \cos\phi d\phi \wedge d\theta,$$

$$dy \wedge dz = \cos\theta \sin^2\phi d\phi \wedge d\theta,$$

$$dz \wedge dx = -\sin\theta \sin^2\phi d\phi \wedge d\theta,$$

and so

$$\left|\left|d^2\mathbf{x}\right|\right| = \sin\phi d\phi \wedge d\theta.$$

The integral becomes

$$\begin{aligned}\iint\limits_{\Sigma} x^2 ||d^2\mathbf{x}|| &= \int_0^{2\pi}\int_0^{\pi} \cos^2\theta\sin^3\phi d\phi\wedge d\theta \\ &= \frac{4\pi}{3}.\end{aligned}$$

343 Example Evaluate $\iint_S z||d^2\mathbf{x}||$ over the conical surface $z = \sqrt{x^2+y^2}$ between $z = 0$ and $z = 1$.

Solution: Put $x = u, y = v, z^2 = u^2 + v^2$. Then

$$dx = du,\ dy = dv,\ zdz = udu + vdv,$$

whence

$$dx\wedge dy = du\wedge dv, dy\wedge dz = -\frac{u}{z}du\wedge dv, dz\wedge dx = -\frac{v}{z}du\wedge dv,$$

and so

$$\begin{aligned}||d^2\mathbf{x}|| &= \sqrt{(dx\wedge dy)^2 + (dz\wedge dx)^2 + (dy\wedge dz)^2} \\ &= \sqrt{1+\frac{u^2+v^2}{z^2}}\ du\wedge dv \\ &= \sqrt{2}\ du\wedge dv.\end{aligned}$$

Hence

$$\iint\limits_{\Sigma} z||d^2\mathbf{x}|| = \iint\limits_{u^2+v^2\le 1} \sqrt{u^2+v^2}\ \sqrt{2}\ dudv = \sqrt{2}\int_0^{2\pi}\int_0^1 \rho^2\ d\rho\wedge d\theta = \frac{2\pi\sqrt{2}}{3}.$$

344 Example Find the area of that part of the cylinder $x^2 + y^2 = 2y$ lying inside the sphere $x^2 + y^2 + z^2 = 4$.

Solution: We have

$$x^2 + y^2 = 2y \iff x^2 + (y-1)^2 = 1.$$

We parametrise the cylinder by putting $x = \cos u, y - 1 = \sin u$, and $z = v$. Hence

$$dx = -\sin u du,\ dy = \cos u du,\ dz = dv,$$

whence

$$dx \wedge dy = 0, dy \wedge dz = \cos u du \wedge dv, dz \wedge dx = \sin u du \wedge dv,$$

and so

$$\begin{aligned} \|d^2\mathbf{x}\| &= \sqrt{(dx \wedge dy)^2 + (dz \wedge dx)^2 + (dy \wedge dz)^2} \\ &= \sqrt{\cos^2 u + \sin^2 u}\, du \wedge dv \\ &= du \wedge dv. \end{aligned}$$

The cylinder and the sphere intersect when $x^2+y^2 = 2y$ and $x^2+y^2+z^2 = 4$, that is, when $z^2 = 4-2y$, i.e. $v^2 = 4-2(1+\sin u) = 2-2\sin u$. Also $0 \le u \le 2\pi$. The integral is thus

$$\begin{aligned} \iint_\Sigma \|d^2\mathbf{x}\| &= \int_0^{2\pi}\int_{-\sqrt{2-2\sin u}}^{\sqrt{2-2\sin u}} dv du = \int_0^{2\pi} 2\sqrt{2-2\sin u} du \\ &= 2\sqrt{2}\int_0^{2\pi}\sqrt{1-\sin u}\, du \\ &= 16. \end{aligned}$$

345 Example Evaluate

$$\iint_\Sigma x dy \wedge dz + (z^2 - zx) dz \wedge dx - xy dx \wedge dy,$$

where Σ is the top side of the triangle with vertices at $(2,0,0)$, $(0,2,0)$, $(0,0,4)$.

Solution: Observe that the plane passing through the three given points has equation $2x + 2y + z = 4$. We project this plane onto the co-ördinate axes obtaining

$$\iint_{\Sigma} x\mathrm{d}y \wedge \mathrm{d}z = \int_0^4 \int_0^{2-z/2} (2 - y - z/2)\mathrm{d}y \wedge \mathrm{d}z = \frac{8}{3},$$

$$\iint_{\Sigma} (z^2 - zx)\mathrm{d}z \wedge \mathrm{d}x = \int_0^2 \int_0^{4-2x} (z^2 - zx)\mathrm{d}z \wedge \mathrm{d}x = 8,$$

$$-\iint_{\Sigma} xy\mathrm{d}x \wedge \mathrm{d}y = -\int_0^2 \int_0^{2-y} xy\mathrm{d}x \wedge \mathrm{d}y = -\frac{2}{3},$$

and hence

$$\iint_{\Sigma} x\mathrm{d}y \wedge \mathrm{d}z + (z^2 - zx)\mathrm{d}z \wedge \mathrm{d}x - xy\mathrm{d}x \wedge \mathrm{d}y = 10.$$

346 Example Evaluate

$$\iint_{\Sigma} xy\mathrm{d}y \wedge \mathrm{d}z - x^2\mathrm{d}z \wedge \mathrm{d}x + (x+z)\mathrm{d}x \wedge \mathrm{d}y,$$

where Σ is the top of the triangular region of the plane $2x + 2y + z = 6$ bounded by the first octant.

Solution: We project this plane onto the co-ördinate axes obtaining

$$\iint_{\Sigma} xy\mathrm{d}y \wedge \mathrm{d}z = \int_0^6 \int_0^{3-z/2} (3 - y - z/2)y\mathrm{d}y \wedge \mathrm{d}z = \frac{27}{4},$$

$$-\iint_{\Sigma} x^2\mathrm{d}z \wedge \mathrm{d}x = -\int_0^3 \int_0^{6-2x} x^2\mathrm{d}z \wedge \mathrm{d}x = -\frac{27}{2},$$

$$\iint_{\Sigma} (x+z)\mathrm{d}x \wedge \mathrm{d}y = \int_0^3 \int_0^{3-y} (6 - x - 2y)\mathrm{d}x \wedge \mathrm{d}y = \frac{27}{2},$$

and hence

$$\iint_{\Sigma} xy\mathrm{d}y \wedge \mathrm{d}z - x^2\mathrm{d}z \wedge \mathrm{d}x + (x+z)\mathrm{d}x \wedge \mathrm{d}y = \frac{27}{4}.$$

3.11 Green's, Stokes', and Gauß' Theorems

We are now in position to state the general Stoke's Theorem.

347 Theorem (General Stoke's Theorem) Let M be a smooth oriented manifold, having boundary ∂M. If ω is a differential form, then

$$\int_{\partial M} \omega = \int_M d\omega.$$

In $\mathbb{R}^2$, if ω is a 2-form, this takes the name of *Green's Theorem.*

348 Example Evaluate $\oint_C (x-y^3)dx+x^3dy$ where C is the circle $x^2+y^2=1$.

Solution: We will first use Green's Theorem and then evaluate the integral directly. We have

$$\begin{aligned} d\omega &= d(x-y^3)\wedge dx + d(x^3)\wedge dy \\ &= (dx - 3y^2dy)\wedge dx + (3x^2dx)\wedge dy \\ &= (3y^2+3x^2)dx\wedge dy. \end{aligned}$$

The region M is the area enclosed by the circle $x^2+y^2=1$. Thus by Green's Theorem, and using polar co-ördinates,

$$\begin{aligned} \oint_C (x-y^3)dx + x^3dy &= \int_M (3y^2+3x^2)dx\wedge dy \\ &= \int_0^{2\pi}\int_0^1 3\rho^2\rho d\rho\wedge d\theta \\ &= \frac{3\pi}{2}. \end{aligned}$$

Aliter: We can evaluate this integral directly, again resorting to polar coördinates.

$$\begin{aligned} \oint_C (x-y^3)dx + x^3dy &= \int_0^{2\pi} (\cos\theta - \sin^3\theta)(-\sin\theta)d\theta + (\cos^3\theta)(\cos\theta)d\theta \\ &= \int_0^{2\pi} (\sin^4\theta + \cos^4\theta - \sin\theta\cos\theta)d\theta. \end{aligned}$$

To evaluate the last integral, observe that $1 = (\sin^2\theta + \cos^2\theta)^2 = \sin^4\theta + 2\sin^2\theta\cos^2\theta + \cos^4\theta$, whence the integral equals

$$\begin{aligned}\int_0^{2\pi}(\sin^4\theta + \cos^4\theta - \sin\theta\cos\theta)d\theta &= \int_0^{2\pi}(1 - 2\sin^2\theta\cos^2\theta - \sin\theta\cos\theta)d\theta \\ &= \frac{3\pi}{2}.\end{aligned}$$

349 Example Evaluate $\oint_C x^3y dx + xy dy$ where C is the square with vertices at $(0,0)$, $(2,0)$, $(2,2)$ and $(0,2)$.

Evaluating this directly would result in evaluating four path integrals, one for each side of the square. We will use Green's Theorem. We have

$$\begin{aligned}d\omega &= d(x^3y)\wedge dx + d(xy)\wedge dy \\ &= (3x^2y dx + x^3 dy)\wedge dx + (y dx + x dy)\wedge dy \\ &= (y - x^3)dx\wedge dy.\end{aligned}$$

The region M is the area enclosed by the square. The integral equals

$$\begin{aligned}\oint_C x^3y dx + xy dy &= \int_0^2\int_0^2 (y - x^3)dx\wedge dy \\ &= -4.\end{aligned}$$

350 Example Consider the triangle $\triangle$ with vertices $A : (0,0)$, $B : (1,1)$, $C : (-2,2)$.

❶ If L_{PQ} denotes the equation of the line joining P and Q find L_{AB}, L_{AC}, and L_{BC}.

❷ Evaluate

$$\oint_{\triangle} y^2 dx + x dy.$$

❸ Find

$$\iint_{\mathscr{D}} (1-2y)\mathrm{d}x \wedge \mathrm{d}y$$

where $\mathscr{D}$ is the interior of $\triangle$.

Solution:

❶ L_{AB} is $y = x$; L_{AC} is $y = -x$, and L_{BC} is clearly $y = -\frac{1}{3}x + \frac{4}{3}$.

❷ We have

$$\begin{aligned}
\int_{AB} y^2\mathrm{d}x + x\mathrm{d}y &= \int_0^1 (x^2+x)\mathrm{d}x &&= \frac{5}{6}\\
\int_{BC} y^2\mathrm{d}x + x\mathrm{d}y &= \int_1^{-2} \left(\left(-\frac{1}{3}x+\frac{4}{3}\right)^2 - \frac{1}{3}x\right)\mathrm{d}x &&= -\frac{15}{2}\\
\int_{CA} y^2\mathrm{d}x + x\mathrm{d}y &= \int_{-2}^0 (x^2-x)\mathrm{d}x &&= \frac{14}{3}
\end{aligned}$$

Adding these integrals we find

$$\oint_{\triangle} y^2\mathrm{d}x + x\mathrm{d}y = -2.$$

❸ We have

$$\begin{aligned}
\iint_{\mathscr{D}} (1-2y)\mathrm{d}x\wedge\mathrm{d}y &= \int_{-2}^0 \left(\int_{-x}^{-x/3+4/3} (1-2y)\mathrm{d}y\right)\mathrm{d}x\\
&\quad + \int_0^1 \left(\int_x^{-x/3+4/3} (1-2y)\mathrm{d}y\right)\mathrm{d}x\\
&= -\frac{44}{27} - \frac{10}{27}\\
&= -2.
\end{aligned}$$

351 Example Use Green's Theorem to prove that

$$\int_{\Gamma} (x^2 + 2y^3)\mathrm{d}y = 16\pi,$$

where Γ is the circle $(x-2)^2 + y^2 = 4$. Also, prove this directly by using a path integral.

Solution: Observe that

$$d(x^2 + 2y^3) \wedge dy = 2x dx \wedge dy.$$

Hence by the generalised Stokes' Theorem the integral equals

$$\iint_{\{(x-2)^2+y^2 \le 4\}} 2x dx \wedge dy = \int_{-\pi/2}^{\pi/2} \int_0^{4\cos\theta} 2\rho^2 \cos\theta d\rho \wedge d\theta = 16\pi.$$

To do it directly, put $x - 2 = 2\cos t, y = 2\sin t, 0 \le t \le 2\pi$. Then the integral becomes

$$\begin{aligned}\int_0^{2\pi} ((2 + 2\cos t)^2 + 16\sin^3 t) d2\sin t &= \int_0^{2\pi} (8\cos t + 16\cos^2 t \\ &\quad +8\cos^3 t + 32\cos t \sin^3 t) dt \\ &= 16\pi.\end{aligned}$$

In $\mathbb{R}^3$, if ω is a 2-form, the above theorem takes the name of *Gauß'* or the *Divergence Theorem.*

352 Example Evaluate $\iint_S (x - y) dy \wedge dz + z dz \wedge dx - y dx \wedge dy$ where S is the surface of the sphere $x^2 + y^2 + z^2 = 9$ and the positive direction is the outward normal.

Solution: The region M is the interior of the sphere $x^2 + y^2 + z^2 = 9$. Now,

$$\begin{aligned} d\omega &= (dx - dy) \wedge dy \wedge dz + dz \wedge dz \wedge dx - dy \wedge dx \wedge dy \\ &= dx \wedge dy \wedge dz. \end{aligned}$$

The integral becomes

$$\begin{aligned}\iiint_M dx \wedge dy \wedge dz &= \frac{4\pi}{3}(27) \\ &= 36\pi.\end{aligned}$$

Aliter: We could evaluate this integral directly. We have

$$\iint_{\Sigma}(x-y)dy \wedge dz = \iint_{\Sigma} xdy \wedge dz,$$

since $(x, y, z) \mapsto -y$ is an odd function of y and the domain of integration is symmetric with respect to y. Now,

$$\begin{aligned}\iint_{\Sigma} xdy \wedge dz &= \int_{-3}^{3}\int_{0}^{2\pi} |\rho|\sqrt{9-\rho^2}d\rho d\theta \\ &= 36\pi.\end{aligned}$$

Also

$$\iint_{\Sigma} zdz \wedge dx = 0,$$

since $(x, y, z) \mapsto z$ is an odd function of z and the domain of integration is symmetric with respect to z. Similarly

$$\iint_{\Sigma} -ydx \wedge dy = 0,$$

since $(x, y, z) \mapsto -y$ is an odd function of y and the domain of integration is symmetric with respect to y.

The classical Stokes' Theorem occurs when ω is a 1-form in $\mathbb{R}^3$.

353 Example Evaluate $\oint_C ydx + (2x - z)dy + (z - x)dz$ where C is the intersection of the sphere $x^2 + y^2 + z^2 = 4$ and the plane $z = 1$.

Solution: We have

$$\begin{aligned}d\omega &= (dy) \wedge dx + (2dx - dz) \wedge dy + (dz - dx) \wedge dz \\ &= -dx \wedge dy + 2dx \wedge dy + dy \wedge dz + dz \wedge dx \\ &= dx \wedge dy + dy \wedge dz + dz \wedge dx.\end{aligned}$$

Since on C, $z = 1$, the surface Σ on which we are integrating is the inside of the circle $x^2 + y^2 + 1 = 4$, i.e., $x^2 + y^2 = 3$. Also, $z = 1$ implies $dz = 0$

and so

$$\iint_{\Sigma} d\omega = \iint_{\Sigma} dx \wedge dy.$$

Since this is just the area of the circular region $x^2 + y^2 \leq 3$, the integral evaluates to

$$\iint_{\Sigma} dx \wedge dy = 3\pi.$$

354 Example Let Γ denote the curve of intersection of the plane $x + y = 2$ and the sphere $x^2 - 2x + y^2 - 2y + z^2 = 0$, oriented clockwise when viewed from the origin. Use Stoke's Theorem to prove that

$$\int_{\Gamma} y dx + z dy + x dz = -2\pi\sqrt{2}.$$

Prove this directly by parametrising the boundary of the surface and evaluating the path integral.

Solution: At the intersection path

$$0 = x^2+y^2+z^2-2(x+y) = (2-y)^2+y^2+z^2-4 = 2y^2-4y+z^2 = 2(y-1)^2+z^2-2,$$

which describes an ellipse on the yz-plane. Similarly we get $2(x-1)^2+z^2 = 2$ on the xz-plane. We have

$$d\,(y dx + z dy + x dz) = dy \wedge dx + dz \wedge dy + dx \wedge dz = -dx \wedge dy - dy \wedge dz - dz \wedge dx.$$

Since $dx \wedge dy = 0$, by Stokes' Theorem the integral sought is

$$-\iint_{2(y-1)^2+z^2\leq 2} dy dz - \iint_{2(x-1)^2+z^2\leq 2} dz dx = -2\pi(\sqrt{2}).$$

(To evaluate the integrals you may resort to the fact that the area of the elliptical region $\dfrac{(x-x_0)^2}{a^2} + \dfrac{(y-y_0)^2}{b^2} \leq 1$ is πab).

If we were to evaluate this integral directly, we would set

$$y = 1 + \cos\theta, \; z = \sqrt{2}\sin\theta, x = 2 - y = 1 - \cos\theta.$$

The integral becomes

$$\int_0^{2\pi} (1+\cos\theta)d(1-\cos\theta) + \sqrt{2}\sin\theta d(1+\cos\theta) + (1-\cos\theta)d(\sqrt{2}\sin\theta)$$

which in turn

$$= \int_0^{2\pi} \sin\theta + \sin\theta\cos\theta - \sqrt{2} + \sqrt{2}\cos\theta d\theta = -2\pi\sqrt{2}.$$