

# Chapter 2

## Mathematical preliminaries

### 2.1. Introduction

In this chapter we introduce a number of basic mathematical concepts that will be used heavily throughout the course. In particular, we review basic set theory, linear algebra, analysis, and partial differential equations. We will also introduce a convenient short-hand notation known as *indicial notation*.

### 2.2. Sets

A *set*  $\mathcal{X}$  is a collection of objects, usually referred to as elements or points. An element  $x$  is a member of a set  $\mathcal{X}$ , denoted  $x \in \mathcal{X}$ . The set  $\mathcal{Y}$  is a subset of a set  $\mathcal{X}$ , denoted  $\mathcal{Y} \subset \mathcal{X}$  if  $y \in \mathcal{Y}$  implies  $y \in \mathcal{X}$ . In this course we will make extensive use of various numeric sets, e.g., the set of real  $\mathbb{R}$  and complex  $\mathbb{C}$  numbers, natural numbers  $\mathbb{N} = \{1, 2, \dots\}$ , whole numbers  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ , and integers  $\mathbb{Z}$ . We will also use  $\mathbb{R}_{>0}$  to denote the set of positive real numbers and  $\mathbb{R}_{\geq 0}$  as the set of non-negative real numbers. The set of real and complex numbers, endowed with the standard operations of addition and multiplication, are algebraic *fields*. Many of the results in this course hold for both the real and complex numbers and in these cases  $\mathbb{F}$  is used to denote either  $\mathbb{R}$  or  $\mathbb{C}$ . Elements of a field  $\mathbb{F}$  are called *scalars*.

Binary set operations map two sets into a set. Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two sets. The *union* of sets, denoted  $\mathcal{X} \cup \mathcal{Y}$ , is the set of all points belonging to either  $\mathcal{X}$  or  $\mathcal{Y}$ . The *intersection* of sets, denoted  $\mathcal{X} \cap \mathcal{Y}$ , is the set of all points belonging to both  $\mathcal{X}$  and  $\mathcal{Y}$ . The *difference* between  $\mathcal{X}$  and  $\mathcal{Y}$ , denoted  $\mathcal{X} \setminus \mathcal{Y}$ , is the set of all points belonging to  $\mathcal{X}$ , but not  $\mathcal{Y}$  (Figure 2.1). The *empty set*  $\emptyset$  is the set containing no elements and, by convention, is a member of every set:  $\emptyset \in \mathcal{X} \cup \emptyset = \mathcal{X}$ ,  $\mathcal{X} \cap \emptyset = \emptyset$ . Two sets  $\mathcal{X}$ ,  $\mathcal{Y}$  are equal if they are subset of each other  $\mathcal{X} \subset \mathcal{Y}$  and  $\mathcal{Y} \subset \mathcal{X}$ , or their difference is the empty set  $\mathcal{X} \setminus \mathcal{Y} = \mathcal{Y} \setminus \mathcal{X} = \emptyset$ . The union and intersection operations are commutative, i.e.,  $\mathcal{X} \cup \mathcal{Y} = \mathcal{Y} \cup \mathcal{X}$  and, in general, the difference operation is not  $\mathcal{X} \setminus \mathcal{Y} \neq \mathcal{Y} \setminus \mathcal{X}$ . Operations involving more than two sets are built up by recursively applying binary set operations.

The *Cartesian product* between two set  $\mathcal{X}$ ,  $\mathcal{Y}$  is defined as the set  $\mathcal{X} \times \mathcal{Y}$

$$\mathcal{X} \times \mathcal{Y} := \{(x, y) \mid x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

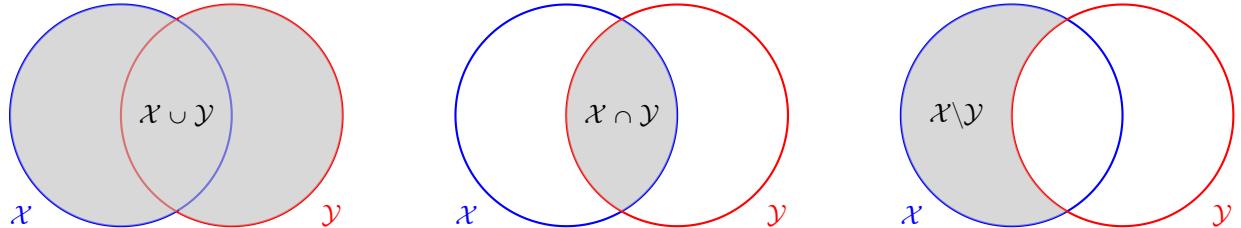


Figure 2.1: Illustration of set operations: union (left), intersection (center), and set difference (right).

The pair  $(x, y)$  is ordered, i.e.,  $(x, y) \neq (y, x)$ . We use  $\mathcal{X}^n$  to denote the Cartesian product of the set  $\mathcal{X}$  with itself  $n$  times, i.e.,

$$\mathcal{X}^n = \underbrace{\mathcal{X} \times \cdots \times \mathcal{X}}_{n \text{ times}}.$$

In this course, we will make extensive use of the Cartesian product of numeric sets, most notably the  $n$ -product of fields  $\mathbb{F}^n$ .

The set of  $m \times n$  matrices with entries belonging to the numeric set  $\mathbb{S}$  will be denoted  $M_{m,n}(\mathbb{S})$ . Similar notation will be used to denote multidimensional arrays, i.e.,  $M_{m,n,k}(\mathbb{S})$  is the set of  $m \times n \times k$  arrays with entries belonging to  $\mathbb{S}$  and  $M_{i_1, \dots, i_d}$  is the set of  $i_1 \times \cdots \times i_d$  arrays with entries belonging to  $\mathbb{S}$ . In this course, we will mostly consider real-valued matrices and arrays  $\mathbb{S} = \mathbb{R}$ .

### 2.3. Linear spaces

A linear space is a key mathematical concept upon which most of linear algebra is constructed. While we will mostly consider real linear spaces, we introduce them over a general field  $\mathbb{F}$ .

**Definition 2.3.1** (Linear space). A linear (vector) space is a nonempty set  $\mathcal{V}$  and a field  $\mathbb{F}$  combined with two operations

$$\begin{aligned} + : \mathcal{V} \times \mathcal{V} &\rightarrow \mathcal{V}, \quad (x, y) \mapsto x + y && \text{(addition)} \\ \cdot : \mathbb{F} \times \mathcal{V} &\rightarrow \mathcal{V}, \quad (\lambda, x) \mapsto \lambda \cdot x = \lambda x && \text{(multiplication)} \end{aligned}$$

such that the following properties hold for all  $x, y, z \in \mathcal{V}$  and  $\alpha, \beta \in \mathbb{F}$

- (a) commutativity w.r.t. addition:  $x + y = y + x$
- (b) associativity w.r.t. addition:  $(x + y) + z = x + (y + z)$
- (c) identity element w.r.t. addition: there exists  $0 \in \mathcal{V}$  such that  $x + 0 = x$
- (d) inverse element w.r.t. addition: there exists  $-x \in \mathcal{V}$  (additive inverse) such that  $x + (-x) = 0$
- (e) compatibility of scalar/field multiplication:  $\alpha(\beta x) = (\alpha\beta)x$
- (f) identity element w.r.t. scalar multiplication:  $1x = x$ , where  $1 \in \mathbb{F}$  is the multiplicative identity
- (g) distributivity of scalar multiplication, vector addition:  $\alpha(x + y) = \alpha x + \alpha y$
- (h) distributivity of scalar multiplication, field addition:  $(\alpha + \beta)x = \alpha x + \beta x$ .

To completely specify a linear space, the set  $\mathcal{V}$ , field  $\mathbb{F}$ , and two operations  $(+, \cdot)$  must be specified, which leads to the lengthy description:  $(\mathcal{V}, \mathbb{F}, +, \cdot)$ . A linear space will be denoted by its set  $\mathcal{V}$  or set and field  $(\mathcal{V}, \mathbb{F})$  when there is no risk of confusion.

A subset  $\mathcal{W} \subset \mathcal{V}$ , where  $(\mathcal{V}, \mathbb{F}, +, \cdot)$  is a linear space, is called a *linear subspace*, or *subspace*, if  $(\mathcal{W}, \mathbb{F}, +, \cdot)$  is a linear space, i.e., a linear space over the same field and operations of the base space. Since many of the linear space properties are inherited from the original space,  $\mathcal{W}$  is a linear subspace provided  $\alpha x + \beta y \in \mathcal{W}$  for any  $x, y \in \mathcal{W}$  and  $\alpha, \beta \in \mathbb{F}$  (closed under addition and scalar multiplication).

**Definition 2.3.2** (Linear combination). Let  $x_1, \dots, x_n$  be elements of a linear space  $\mathcal{V}$ . A vector  $x \in \mathcal{V}$  is a *linear combination* of vectors  $\{x_k\}_{k=1}^n$  if there exist scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that

$$x = \sum_{k=1}^n \alpha_k x_k. \tag{2.1}$$

**Definition 2.3.3** (Linear independence). A finite collection of vectors  $\{x_k\}_{k=1}^n$  is *linearly independent* if

$$\sum_{k=1}^n \alpha_k x_k = 0 \iff \alpha_1 = \cdots = \alpha_n = 0. \tag{2.2}$$

An infinite collection of vectors  $\mathcal{W}$  is linearly independent if every finite subset of  $\mathcal{W}$  is linearly independent. If a collection of vectors is not linearly independent, a vector in the set can be written as a linear combination of others in the set and is called *linearly dependent*.

Let  $\mathcal{A}$  be a subset of a linear space  $\mathcal{V}$ . The span of  $\mathcal{A}$ , denoted  $\text{span } \mathcal{A}$ , is the set of all finite linear combinations of vectors from  $\mathcal{A}$ , i.e.,

$$\text{span } \mathcal{A} = \left\{ \sum_{k=1}^n \alpha_k x_k \mid x_1, \dots, x_n \in \mathcal{A}, \alpha_1, \dots, \alpha_n \in \mathbb{F}, n \in \mathbb{N} \right\}. \quad (2.3)$$

Note that we explicitly include the summation symbol to clearly specify the range of the summation. The span of  $\mathcal{A}$  is the smallest linear subspace of  $\mathcal{V}$  containing  $\mathcal{A}$ .

**Definition 2.3.4** (Basis). A set of vectors  $\mathcal{B} \subset \mathcal{V}$ , where  $\mathcal{V}$  is a linear space is a *basis* of  $\mathcal{V}$  if  $\mathcal{B}$  is linearly independent and  $\text{span } \mathcal{B} = \mathcal{V}$ .

If there exists a finite basis for  $\mathcal{V}$ , the linear space is *finite-dimensional*. Otherwise it is *infinite-dimensional*. There are many possible bases of a given linear space; it can be proven that all bases of a given space must possess the same number of vectors. Thus the number of vectors needed to define a basis for a linear space is a fundamental property of the space called its *dimension*.

### 2.3.1 Scalar spaces

The set of real numbers over the field of real numbers with the usual operations of addition and multiplication  $(\mathbb{R}, \mathbb{R})$  is a linear space of dimension 1. Properties (a)-(h) in Definition 2.3.1 follow trivially from the properties of real numbers. Similarly, the set of complex numbers over the field of complex numbers  $(\mathbb{C}, \mathbb{C})$  under the usual operations of addition and multiplication is a linear space of dimension 1. Since these spaces are one-dimensional, any single non-zero element of the space is a basis. In contrast, the set of complex numbers over the field of real numbers  $(\mathbb{C}, \mathbb{R})$  under the usual operations of addition and multiplication is a linear space of dimension 2 and a basis is  $\{1, i\}$  ( $i = \sqrt{-1}$ ). In this course, any field  $\mathbb{F}$  will implicitly represent the linear space  $(\mathbb{F}, \mathbb{F})$  under the field operations of addition and multiplication unless otherwise specified.

### 2.3.2 Array spaces

The space of  $n$  vectors with entries belonging to the field  $\mathbb{F}$ , denoted  $\mathbb{F}^n$ , i.e., a Cartesian product of  $\mathbb{F}$ , is a linear space over the field  $\mathbb{F}$  under component-wise addition and scalar multiplication  $(\mathbb{F}^n, \mathbb{F})$ , i.e., let  $v, w \in \mathbb{F}^n$  and  $\lambda \in \mathbb{F}$

$$(v + w)_i = v_i + w_i, \quad (\lambda v)_i = \lambda v_i. \quad (2.4)$$

Due to the component-wise nature of addition and scalar multiplication, properties (a)-(h) follow trivially from the properties of a field with the additive identity  $z \in \mathbb{F}$  (c) and inverse of  $x \in \mathbb{F}$  (d)

$$z = (0, \dots, 0), \quad -x = (-x_1, \dots, x_n), \quad (2.5)$$

respectively. This space can be trivially extended to higher dimensional arrays:  $M_{m,n}(\mathbb{F})$  and  $M_{i_1, \dots, i_d}(\mathbb{F})$ . In this course, any array space, e.g.,  $\mathbb{F}^n$ ,  $M_{m,n}(\mathbb{F})$ ,  $M_{i_1, \dots, i_d}(\mathbb{F})$ , will implicitly represent the linear space  $(\mathbb{F}^n, \mathbb{F})$ ,  $(M_{m,n}(\mathbb{F}), \mathbb{F})$ ,  $(M_{i_1, \dots, i_d}(\mathbb{F}), \mathbb{F})$ , respectively, under the operations of component-wise addition and multiplication unless otherwise specified.

#### Example 2.1: Matrix subspace

Consider the linear space  $M_{m,n}(\mathbb{F})$ . Then the following subset

$$\mathcal{V} = \{A \in M_{m,n}(\mathbb{F}) \mid A_{ij} = 0, \quad i, j > 1\} \quad (2.6)$$

is a linear subspace of  $M_{m,n}(\mathbb{F})$  because for any  $A, B \in \mathcal{V}$  and  $\alpha, \beta \in \mathbb{R}$ ,  $\alpha A + \beta B \in \mathcal{V}$ .

#### Example 2.2: Linear combination in $\mathbb{R}^2$

Consider the linear space  $\mathbb{R}^2$ . The vector  $(x_1, x_2) \in \mathbb{R}^2$  is a linear combination of  $(1, 0), (0, 0.1) \in \mathbb{R}^2$  because

$$(x_1, x_2) = x_1(1, 0) + 10x_2(0, 0.1).$$

### Example 2.3: Canonical basis and dimension

Consider the linear space  $\mathbb{F}^n$ . Define the set of *canonical unit vectors*  $\{e^{(i)}\}_{i=1}^n$  where  $e^{(i)} \in \mathbb{F}^n$  for  $i = 1, \dots, n$  is the vector with 1  $\in \mathbb{F}$  as its  $i$ th component and all other components are 0  $\in \mathbb{F}$ :

$$e_j^{(i)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \quad (2.7)$$

The set of canonical unit vectors is independent because

$$\sum_{k=1}^n \alpha_k e^{(k)} = (\alpha_1, \dots, \alpha_n)$$

is the zero element of  $\mathbb{F}^n$  if and only if  $\alpha_1 = \dots = \alpha_n = 0 \in \mathbb{F}$ . The canonical unit vectors span all of  $\mathbb{F}^n$ : consider any  $x \in \mathbb{F}^n$  with  $x = (x_1, \dots, x_n)$ , then

$$x = \sum_{k=1}^n x_k e^{(k)}.$$

### 2.3.3 Sequence spaces

Let  $\mathcal{S}(\mathbb{F})$  denote the space of sequences of elements in  $\mathbb{F}$  (field), i.e.,

$$\mathcal{S}(\mathbb{F}) = \{(x_1, x_2, \dots) \mid x_1, x_2, \dots \in \mathbb{F}\}.$$

For convenience we use  $\{x_n\}$  as short-hand notation for  $(x_1, x_2, \dots)$ . Addition and scalar multiplication are defined as

$$\{x_n\} + \{y_n\} = \{x_n + y_n\}, \quad \lambda\{x_n\} = \{\lambda x_n\} \quad (2.8)$$

for any  $\{x_n\}, \{y_n\} \in \mathcal{S}(\mathbb{F})$  and  $\lambda \in \mathbb{F}$ . The space  $(\mathcal{S}(\mathbb{F}), \mathbb{F})$  under the above operations is a linear space, which can be verified using the same argument used to prove the Cartesian product of linear spaces is a linear space (Section 2.3.6).

The space of all bounded sequences of elements in  $\mathbb{F}$ , denoted  $\mathcal{S}_b(\mathbb{F})$ , is a linear subspace of  $\mathcal{S}(\mathbb{F})$ , which we prove as follows. Let  $\alpha, \beta \in \mathbb{F}$  and  $\{x_n\}, \{y_n\} \in \mathcal{S}_b(\mathbb{F})$ , which implies there exists  $M_1, M_2 > 0$  such that  $|x_n| < M_1$  and  $|y_n| < M_2$  for all  $n \in \mathbb{N}$  (definition of bounded sequence). From the definition of sequence addition and scalar multiplication we have

$$\alpha\{x_n\} + \beta\{y_n\} = \{z_n\}, \quad (2.9)$$

where  $z_n = \{\alpha x_n + \beta y_n\}$ .  $\{z_n\}$  is a bounded sequence because  $z_n < M_3$ , where  $M_3 = \alpha M_1 + \beta M_2$ . Therefore  $\{z_n\} \in \mathcal{S}_b(\mathbb{F})$ , which implies  $\mathcal{S}_b(\mathbb{F})$  is a subspace of  $\mathcal{S}(\mathbb{F})$ . Similarly one can show that the space of all convergent sequences in  $\mathbb{F}$  is a subspace of  $\mathcal{S}_b(\mathbb{F})$ . In this course, it is implicitly assumed that  $\mathcal{S}(\mathbb{F})$  and  $\mathcal{S}_b(\mathbb{F})$  refer to the linear spaces  $(\mathcal{S}(\mathbb{F}), \mathbb{F})$  and  $(\mathcal{S}_b(\mathbb{F}), \mathbb{F})$  under the operations defined in (2.8).

### 2.3.4 Function spaces

Let  $\mathcal{X}$  be a nonempty set and  $(\mathcal{V}, \mathbb{F})$  a linear space. The set of all mappings from  $\mathcal{X}$  to  $\mathcal{V}$ , denoted  $\mathcal{F}_{\mathcal{X} \rightarrow \mathcal{V}}$ , is a linear space over the field  $\mathbb{F}$  under the following operations

$$(f + g)(x) = f(x) + g(x), \quad (\lambda f)(x) = \lambda f(x) \quad (2.10)$$

for  $f, g \in \mathcal{F}_{\mathcal{X} \rightarrow \mathcal{V}}$  and  $\lambda \in \mathbb{F}$ . Properties (a)-(b), (e)-(h) follow from the corresponding properties of the linear space  $\mathcal{V}$  since addition and scalar multiplication are defined pointwise, i.e., for each  $x \in \mathcal{X}$ . The additive identity (c) is the function  $F_0 : \mathcal{X} \rightarrow \mathcal{V}$  such that  $x \mapsto 0 \in \mathcal{V}$ . The additive inverse (d) of any  $f \in \mathcal{F}_{\mathcal{X} \rightarrow \mathcal{V}}$  is  $-f : \mathcal{X} \rightarrow \mathcal{V}$  such that  $x \mapsto -f(x)$ , where  $-f(x) \in \mathcal{V}$  is the additive inverse of  $f(x) \in \mathcal{V}$ .

### 2.3.5 Polynomial spaces

A particularly important function space in the study of the finite element method is the space of polynomial functions. Let  $\Omega$  be an open set of  $\mathbb{R}$  (we restrict our attention to real, scalar polynomial spaces; multidimensional polynomials are introduced later in the course) and define the space of polynomials over  $\Omega$  as the function space

$$\mathcal{P}(\Omega) := \left\{ p \in \mathcal{F}_{\Omega \rightarrow \mathbb{R}} \mid p(x) = \sum_{n=0}^{\infty} a_n x^n, x \in \Omega, a_n \in \mathbb{R}, n \in \mathbb{N}_0 \right\}$$

The space of polynomials  $\mathcal{P}(\Omega)$  is a linear space over the field  $\mathbb{R}$  with operations

$$(p+q)(x) = \sum_{k=0}^{\infty} (a_k + b_k)x^k, \quad (\lambda p)(x) = \sum_{k=0}^{\infty} (\lambda a_k)x^k \quad (2.11)$$

for  $\alpha, \beta \in \mathbb{R}$  and  $p, q \in \mathcal{P}(\Omega)$  where  $p(x) = \sum_{k=0}^{\infty} a_k x^k$  ( $a_k \in \mathbb{R}$ ) and  $q(x) = \sum_{k=0}^{\infty} b_k x^k$  ( $b_k \in \mathbb{R}$ ). Properties (a)-(b) follow from the corresponding property of real numbers. The additive identity (c) is the polynomial  $p_0 \in \mathcal{P}(\Omega)$  with coefficients  $a_k = 0$  for  $k \in \mathbb{N}_0$ . The additive inverse (d) of the polynomial  $p \in \mathcal{P}(\Omega)$  where  $p(x) = \sum_{k=0}^{\infty} a_k x^k$  is  $-p \in \mathcal{P}(\Omega)$  defined as  $(-p)(x) = \sum_{k=0}^{\infty} (-a_k)x^k$ . Properties (e)-(f) follow from the definition of scalar multiplication and associativity of real numbers. Finally, let  $\alpha, \beta \in \mathbb{R}$  and  $p, q \in \mathcal{P}(\Omega)$  where  $p(x) = \sum_{k=0}^{\infty} a_k x^k$  ( $a_k \in \mathbb{R}$ ) and  $q(x) = \sum_{k=0}^{\infty} b_k x^k$  ( $b_k \in \mathbb{R}$ ). Then property (g) follows from

$$\alpha(p+q)(x) = \alpha(p(x) + q(x)) = \alpha \left( \sum_{k=0}^{\infty} a_k x^k + \sum_{k=0}^{\infty} b_k x^k \right) = \alpha \sum_{k=0}^{\infty} a_k x^k + \alpha \sum_{k=0}^{\infty} b_k x^k = (\alpha p + \beta q)(x)$$

and (h) follows from

$$(\alpha + \beta)p(x) = (\alpha + \beta) \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} (\alpha + \beta)a_k x^k = \alpha \sum_{k=0}^{\infty} a_k x^k + \beta \sum_{k=0}^{\infty} a_k x^k = \alpha p(x) + \beta q(x).$$

The *degree* of a polynomial  $p \in \mathcal{P}(\Omega)$ , denoted  $\deg p$ , is the exponent of the highest degree monomial term with non-zero coefficient, i.e., if  $k = \deg p$  then  $a_k \neq 0$  and  $a_{k+1} = a_{k+2} = \dots = 0$ . An extremely important subspace of  $\mathcal{P}(\Omega)$  that will be used extensively in our study of the finite element method is the space of polynomials of degree at most  $k$

$$\mathcal{P}^k(\Omega) := \left\{ p \in \mathcal{P}(\Omega) \mid p(x) = \sum_{n=0}^k a_n x^n, x \in \Omega, a_0, \dots, a_k \in \mathbb{R} \right\}.$$

$\mathcal{P}^k(\Omega)$  is indeed a linear subspace of  $\mathcal{P}(\Omega)$  because addition and multiplication by a scalar (independent of  $x$ ) cannot increase the degree of the polynomial. In this course, we only consider real-valued polynomials over real domains. Therefore the notation  $\mathcal{P}(\Omega)$ ,  $\mathcal{P}^k(\Omega)$  ( $\Omega \subset \mathbb{R}^d$ ) will be used to denote the linear space  $(\mathcal{P}(\Omega), \mathbb{R})$  and  $(\mathcal{P}^k(\Omega), \mathbb{R})$  under the standard operations of addition and scalar multiplication of polynomials (2.11).

#### Example 2.4: Linear combination of polynomials

Consider the linear space  $\mathcal{P}^3(\mathbb{R})$ . The element  $v = x^2 + x^3 \in \mathcal{P}^3(\mathbb{R})$  is a linear combination of  $2 + x^2 \in \mathcal{P}^3(\mathbb{R})$  and  $1 - x^3/2 \in \mathcal{P}^3(\mathbb{R})$  because

$$v = x^2 + x^3 = 1(2 + x^2) - 2(1 - x^3/2).$$

#### Example 2.5: Canonical basis of $\mathcal{P}^k(\Omega)$ and dimension

Consider the polynomial space  $\mathcal{P}^k(\Omega)$  where  $\Omega \subset \mathbb{R}$  (open). The set of *monomials*  $\{x^n\}_{n=0}^k$  defines a basis of  $\mathcal{P}^k(\Omega)$  because monomials are linearly independent and any polynomial of order at most  $k$  can be represented as a linear combination of all monomials to degree  $k$ . Since there are  $k+1$  independent monomials of degree at most  $k$  and this set of monomials is a basis of  $\mathcal{P}^k(\Omega)$ , we have  $\dim \mathcal{P}^k(\Omega) = k+1$

(finite-dimensional). In contrast,  $\mathcal{P}(\Omega)$  is an infinite-dimensional linear space.

### 2.3.6 Cartesian product of linear spaces

Consider a linear space  $(\mathcal{V}, \mathbb{F})$  and let  $\mathcal{V}_1, \dots, \mathcal{V}_n \subset \mathcal{V}$  be linear subspaces. The Cartesian product space  $(\mathcal{W}, \mathbb{F})$  is a linear space, where  $\mathcal{W} = \mathcal{V}_1 \times \dots \times \mathcal{V}_n$  and the operations are defined component-wise

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n) \\ \lambda(x_1, \dots, x_n) &= (\lambda x_1, \dots, \lambda x_n).\end{aligned}$$

Commutativity (a) and associativity (b) of addition follow from the corresponding property of the generating space  $\mathcal{V}$ . The additive identity (c) is  $z = (0, \dots, 0)$  and the additive inverse (d) of any  $(x_1, \dots, x_n) \in \mathbb{F}^n$  is  $(-x_1, \dots, -x_n)$ . Properties (e)-(f) are directly inherited from  $\mathcal{V}$  since scalar multiplication is component-wise. Finally, let  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathcal{W}$  and  $\alpha, \beta \in \mathbb{F}$ . Then (g) is established as

$$\alpha((x_1, \dots, x_n) + (y_1, \dots, y_n)) = \alpha(x_1 + y_1, \dots, x_n + y_n) = (\alpha(x_1 + y_1), \dots, \alpha(x_n + y_n))$$

and (h) is established as

$$\begin{aligned}(\alpha + \beta)(x_1, \dots, x_n) &= ((\alpha + \beta)x_1, \dots, (\alpha + \beta)x_n) \\ &= (\alpha x_1 + \beta x_1, \dots, \alpha x_n + \beta x_n) \\ &= (\alpha x_1, \dots, \alpha x_n) + (\beta x_1, \dots, \beta x_n) \\ &= \alpha(x_1, \dots, x_n) + \beta(x_1, \dots, x_n).\end{aligned}$$

### 2.3.7 Addition of linear spaces

Consider a linear space  $(\mathcal{V}, \mathbb{F})$  and let  $\mathcal{V}_1, \mathcal{V}_2$  be linear subspaces. The sum space  $(\mathcal{W}, \mathbb{F})$  is a linear subspace of  $(\mathcal{V}, \mathbb{F})$ , where

$$\mathcal{W} = \mathcal{V}_1 + \mathcal{V}_2 := \{v_1 + v_2 \mid v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2\}, \quad (2.12)$$

under the same operations of addition and scalar multiplication defined for  $(\mathcal{V}, \mathbb{F})$ . To show this take  $x, y \in \mathcal{W}$  and  $\alpha, \beta \in \mathbb{F}$ . From the definition of  $\mathcal{W}$ , there must exist  $x_1, y_1 \in \mathcal{V}_1$  and  $x_2, y_2 \in \mathcal{V}_2$  such that

$$x = x_1 + x_2, \quad y = y_1 + y_2,$$

which implies

$$\alpha x + \beta y = \alpha(x_1 + x_2) + \beta(y_1 + y_2).$$

Since both  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are linear spaces, we have  $\alpha(x_1 + x_2) \in \mathcal{V}_1$  and  $\beta(y_1 + y_2) \in \mathcal{V}_2$ , which implies  $\alpha x + \beta y \in \mathcal{W}$ . Therefore,  $\mathcal{W}$  is a linear subspace.

### 2.3.8 Affine spaces

Consider a linear space  $(\mathcal{V}, \mathbb{F})$ . An *affine* subspace is a subset  $\mathcal{A} \subset \mathcal{V}$  that can be expressed as the sum of an *arbitrary* element  $a \in \mathcal{A}$  and a linear space  $(\mathcal{W}, \mathbb{F})$  under the operations of the original space  $(\mathcal{V}, \mathbb{F})$

$$\mathcal{A} = a + \mathcal{W} := \{a + w \mid w \in \mathcal{W}\}. \quad (2.13)$$

Affine spaces are not linear spaces because they are not closed under addition or scalar multiplication. To see this, take  $x, y \in \mathcal{A}$  then, by definition, for any  $a \in \mathcal{A}$  (in general  $a \notin \mathcal{W}$ ) there exists  $u, v \in \mathcal{W}$  such that  $x = a + u$  and  $y = a + v$ . Taking a linear combination of these elements with arbitrary scalars  $\alpha, \beta \in \mathbb{F}$  leads to

$$z := \alpha x + \beta y = \alpha(a + u) + \beta(a + v) = a + \underbrace{(\alpha + \beta - 1)a + \alpha u + \beta v}_{\notin \mathcal{W} \text{ unless } \alpha + \beta = 1 \text{ or } a \in \mathcal{W}}, \quad (2.14)$$

which proves affine subspaces are not closed under addition and therefore not linear (sub)spaces. However, they will play an important role in defining essential boundary conditions in a variational setting.

## 2.4. Normed spaces

In this section we associate a *norm* with a linear space to provide a notion of magnitude or length, which will be used to define, e.g., distance between vectors and convergence of sequences. The concept of a norm generalizes the absolute value function in  $\mathbb{R}$  or  $\mathbb{C}$ .

**Definition 2.4.1** (Norm). A function  $x \mapsto \|x\|$  from a linear space  $\mathcal{V}$  into  $\mathbb{R}$  is called a *norm* if

- (a)  $\|x\| \geq 0$  and  $\|x\| = 0 \implies x = 0$
- (b)  $\|\lambda x\| = |\lambda| \|x\|$  for every  $x \in \mathcal{V}, \lambda \in \mathbb{F}$
- (c) triangle inequality:  $\|x + y\| \leq \|x\| + \|y\|$  for every  $x, y \in \mathcal{V}$ .

### Example 2.6: Norms in $\mathbb{R}^n$

Consider the linear space  $\mathbb{R}^n$  and define the following functions:

$$\begin{aligned}\|\cdot\|_1 : \mathbb{R}^n &\rightarrow \mathbb{R}, & x = (x_1, \dots, x_n) &\mapsto |x_1| + \dots + |x_n| \\ \|\cdot\|_p : \mathbb{R}^n &\rightarrow \mathbb{R}, & x = (x_1, \dots, x_n) &\mapsto (|x_1|^p + \dots + |x_n|^p)^{1/p} \\ \|\cdot\|_\infty : \mathbb{R}^n &\rightarrow \mathbb{R}, & x = (x_1, \dots, x_n) &\mapsto \max_{i \in \{1, \dots, n\}} |x_i|.\end{aligned}$$

These functions are called the one-norm,  $p$ -norm (the most common being the  $p = 2$  norm), and infinity norm, respectively, and are valid norms according to Definition 2.4.1. For concreteness, we verify  $\|\cdot\|_1$  is a valid norm. Property (a) follows from the fact that for any  $x \in \mathbb{R}^n$ ,  $\|x\|_1$  is a summation of nonnegative numbers, i.e., must be nonnegative and can only be zero if all terms are zero (no negative numbers to cancel out positive numbers). Property (b) is established as follows: for any  $x \in \mathbb{R}^n$  and  $\lambda \in \mathbb{R}$ ,

$$\|\lambda x\|_1 = \sum_{k=1}^n |\lambda x_k| = \sum_{k=1}^n |\lambda| |x_k| = |\lambda| \sum_{k=1}^n |x_k| = |\lambda| \|x\|_1.$$

Finally, the triangle inequality follows from the triangle inequality of the absolute value: for any  $x, y \in \mathbb{R}^n$ , we have

$$\|x + y\|_1 = \sum_{k=1}^n |x_k + y_k| \leq \sum_{k=1}^n (|x_k| + |y_k|) = \sum_{k=1}^n |x_k| + \sum_{k=1}^n |y_k| = \|x\|_1 + \|y\|_1,$$

which establishes  $\|\cdot\|_1$  as a valid norm.

**Definition 2.4.2** (Normed space). A *normed space* is a linear space  $\mathcal{V}$  endowed with a norm  $\|\cdot\|$ .

It is possible to define different norms on the same linear space so both must be specified to completely define the normed space:  $(\mathcal{V}, \|\cdot\|)$ . If ambiguous, the field and/or operations of the linear space will be included in the description:  $((\mathcal{V}, \mathbb{F}), \|\cdot\|)$  or  $((\mathcal{V}, \mathbb{F}, +, \cdot), \|\cdot\|)$ . A linear subspace of a normed space is a normed space with the same norm. A *unit vector* associated with normed space  $(\mathcal{V}, \|\cdot\|)$  is any  $x \in \mathcal{V}$  such that  $\|x\| = 1$ .

**Definition 2.4.3** (Convergence in normed spaces). Let  $(\mathcal{V}, \|\cdot\|)$  be a normed space. A sequence  $\{x_n\}$  of elements of  $\mathcal{V}$  converges to  $x \in \mathcal{V}$  if for every  $\epsilon > 0$  there exists  $M > 0$  such that for every  $n \geq M$ ,  $\|x_n - x\| < \epsilon$ . In this case we say  $\{x_n\}$  converges to  $x$  and write  $\lim_{n \rightarrow \infty} x_n = x$  or  $x_n \rightarrow x$ .

Let  $x \in \mathcal{V}$ , where  $(\mathcal{V}, \|\cdot\|)$  is a normed space, and let  $r \in \mathbb{R}_{>0}$ . The *open ball*, *closed ball*, and *sphere* are defined as

$$\begin{aligned}B(x, r) &:= \{y \in \mathcal{V} \mid \|y - x\| < r\} \\ \overline{B}(x, r) &:= \{y \in \mathcal{V} \mid \|y - x\| \leq r\} \\ S(x, r) &:= \{y \in \mathcal{V} \mid \|y - x\| = r\},\end{aligned}\tag{2.15}$$

respectively. In each case,  $x \in \mathcal{V}$  is the center and  $r$  is the radius.

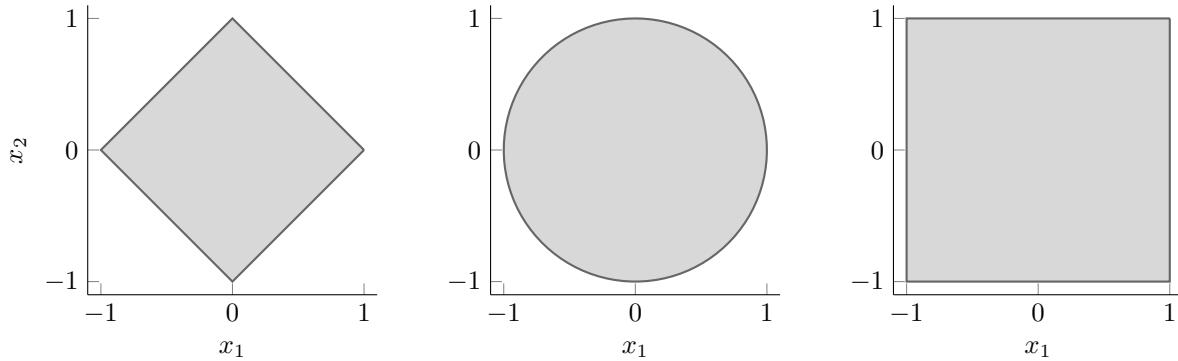


Figure 2.2: The gray region alone (without boundary (—)) defines the open unit ball  $B(0, 1)$ , the union of the gray region and boundary (—) defines the closed unit ball  $\overline{B}(0, 1)$ , and the boundary (—) alone defines the unit sphere  $S(0, 1)$  corresponding to the one-norm  $\|\cdot\|_1$  (left), two-norm  $\|\cdot\|_2$  (center), and infinity-norm  $\|\cdot\|_\infty$  (right).

**Definition 2.4.4** (Open and closed sets). A subset  $\mathcal{S}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is called *open* if for every  $x \in \mathcal{S}$  there exists  $\epsilon > 0$  such that  $B(x, \epsilon) \subset \mathcal{S}$ . A subset  $\mathcal{S}$  is called *closed* if its complement is open, i.e.,  $\mathcal{V} \setminus \mathcal{S}$  is open.

**Theorem 2.1.** *A subset  $\mathcal{S}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is closed if and only if every sequence of elements in  $\mathcal{S}$  convergent in  $\mathcal{V}$  has its limit in  $\mathcal{S}$ , i.e.,*

$$x_1, x_2, \dots \in \mathcal{S} \text{ and } x_n \implies x \in \mathcal{S}. \quad (2.16)$$

**Definition 2.4.5** (Interior). Let  $(\mathcal{V}, \|\cdot\|)$  be a normed space. The interior of  $\mathcal{V}$ , denoted  $\mathcal{V}^\circ$ , is the union of all open subsets of  $\mathcal{V}$ , or equivalently, the set of all points  $x \in \mathcal{V}$  such that  $B(x, \epsilon) \subset \mathcal{V}$  for some  $\epsilon > 0$ .

**Definition 2.4.6** (Closure). Let  $\mathcal{S}$  be a subset of a normed space  $(\mathcal{V}, \|\cdot\|)$ . The closure of  $\mathcal{S}$ , denoted  $\overline{\mathcal{S}}$ , is the intersection of all closed sets containing  $\mathcal{S}$ , i.e., the smallest closed set containing  $\mathcal{S}$ .

**Theorem 2.2.** *Let  $\mathcal{S}$  be a subset of a normed space  $(\mathcal{V}, \|\cdot\|)$ . The closure of  $\mathcal{S}$  is the set of limits of all convergent sequences of elements of  $\mathcal{S}$ .*

**Definition 2.4.7** (Boundary). The boundary  $\partial\mathcal{V}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is the (set) difference between the closure of  $\mathcal{V}$  and its interior

$$\partial\mathcal{V} := \overline{\mathcal{V}} \setminus \mathcal{V}^\circ.$$

**Definition 2.4.8** (Dense subset). A subset  $\mathcal{S}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is *dense* in  $\mathcal{V}$  if  $\overline{\mathcal{S}} = \mathcal{V}$ .

**Definition 2.4.9** (Compact set). A subset  $\mathcal{S}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is *compact* if every sequence  $\{x_n\}$  in  $\mathcal{S}$  contains a convergent subsequence whose limit belongs to  $\mathcal{S}$ .

**Definition 2.4.10** (Bounded subsets). A subset  $\mathcal{S}$  of a normed space  $(\mathcal{V}, \|\cdot\|)$  is *bounded* if  $\mathcal{S} \subset B(0, r)$  for some  $r > 0$ .

**Theorem 2.3.** *Compact sets are closed and bounded.*

**Definition 2.4.11** (Cauchy sequence). Let  $(\mathcal{V}, \|\cdot\|)$  be a normed space. A sequence  $\{x_n\}_{n=1}^\infty \subset \mathcal{V}$  is a *Cauchy sequence* if for every  $\epsilon \in \mathbb{R}_{>0}$ , there exists  $N \in \mathbb{N}$  such that for all  $m, n > N$  ( $m, n \in \mathbb{N}$ ),  $\|x_n - x_m\| < \epsilon$ .

**Definition 2.4.12** (Banach space). A normed space  $(\mathcal{V}, \|\cdot\|)$  is called *complete* if every Cauchy sequence of points in  $\mathcal{V}$  converges to a point in  $\mathcal{V}$ . A complete normed space is called a *Banach space*.

## 2.5. Inner product spaces

**Definition 2.5.1** (Inner product). A mapping  $(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$  on a linear space  $\mathcal{V}$  is called a *inner product* if it satisfies

- (i)  $(u + v, w) = (u, w) + (v, w)$  for all  $u, v, w \in \mathcal{V}$
- (ii)  $(\alpha u, v)$  for all  $u, v \in \mathcal{V}, \alpha \in \mathbb{F}$
- (iii)  $(u, v) = (v, u)$  for all  $u, v \in \mathcal{V}$
- (iv)  $(v, v) \geq 0$  for all  $v \in \mathcal{V}$
- (v)  $(v, v) = 0 \iff v = 0$

**Definition 2.5.2** (Inner product space). A linear space  $\mathcal{V}$  together with an inner product defined on it is called an *inner product space* and is denoted  $(\mathcal{V}, (\cdot, \cdot))$ .

**Theorem 2.4** (Cauchy-Schwarz Inequality). *If  $(\mathcal{V}, (\cdot, \cdot))$  is an inner product space, then for any  $u, v \in \mathcal{V}$*

$$|(u, v)| \leq \sqrt{(u, u)}\sqrt{(v, v)}. \quad (2.17)$$

*The equality holds if and only if  $u$  and  $v$  are linearly dependent.*

**Proposition 2.1.**  $\|v\| := \sqrt{(v, v)}$  defines a norm in the inner product space  $(\mathcal{V}, (\cdot, \cdot))$  and is called the norm induced by the inner product  $(\cdot, \cdot)$ .

Thus any inner product space  $(\mathcal{V}, (\cdot, \cdot))$  can be made into a normed space  $(\mathcal{V}, \|\cdot\|)$  under the norm induced by the inner product  $\|v\| := \sqrt{(v, v)}$ .

**Definition 2.5.3.** Let  $(\mathcal{V}, (\cdot, \cdot))$  be an inner product space. If the associated normed space  $(\mathcal{V}, \|\cdot\|)$  is complete, then  $(\mathcal{V}, (\cdot, \cdot))$  is called a *Hilbert space*.

**Proposition 2.2.** *Let  $(\mathcal{V}, (\cdot, \cdot))$  be a Hilbert space and  $\mathcal{S} \subset \mathcal{V}$  a linear subspace. Then  $(\mathcal{S}, (\cdot, \cdot))$  is a Hilbert space.*

## 2.6. Mappings

Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two sets and consider a *mapping*  $f : \mathcal{X} \rightarrow \mathcal{Y}$ . If  $y = f(x)$ , then  $y$  is the *image* of  $x$ . More generally, let  $\mathcal{A} \subset \mathcal{X}$  and  $\mathcal{B} \subset \mathcal{Y}$ , then  $f(\mathcal{A})$  is the *image* of the set  $\mathcal{X}$  and  $f^{-1}(\mathcal{B})$  is the *inverse image* or *preimage* of  $\mathcal{B}$ , where

$$f(\mathcal{A}) := \{f(x) \mid x \in \mathcal{A}\}, \quad f^{-1}(\mathcal{B}) := \{x \in \mathcal{X} \mid f(x) \in \mathcal{B}\}.$$

Note that the notation  $f^{-1}$  does not imply  $f$  is invertible. If  $f(\mathcal{X}) = \mathcal{Y}$ ,  $f$  maps  $\mathcal{X}$  onto  $\mathcal{Y}$  (surjective). If, for each  $y \in \mathcal{Y}$ ,  $f^{-1}(y)$  consists of at most one element of  $\mathcal{X}$ , then  $f$  is a *one-to-one* mapping of  $\mathcal{X}$  into  $\mathcal{Y}$  (injective). Mappings that are both injective and surjective are *bijective* (invertible). The *domain* of  $f$ , denoted  $\mathcal{D}(f)$ , is the source set  $\mathcal{X}$  of the mapping. The *range* of the mapping, denoted  $\mathcal{R}(f)$ , is the set of points mapped from the domain

$$\mathcal{R}(f) := f(\mathcal{D}(f)) = \{y \in \mathcal{Y} \mid f(x) = y \text{ for some } x \in \mathcal{D}(f)\}. \quad (2.18)$$

Now let  $f \in \mathcal{F}_{\mathcal{X} \rightarrow \mathcal{Y}}$  where  $\mathcal{Y}$  is a linear space. The *null space* of the mapping,  $\mathcal{N}(f) \subset \mathcal{D}(f)$ , is the set of points in the domain that map to zero, i.e.,

$$\mathcal{N}(f) := \{x \in \mathcal{D}(f) \mid f(x) = 0\}. \quad (2.19)$$

The *support* of a mapping,  $\text{supp}(f) \subset \mathcal{D}(f)$ , is the set of points in the domain that do not map to zero, i.e.,

$$\text{supp } f := \{x \in \mathcal{D}(f) \mid f(x) \neq 0\}. \quad (2.20)$$

### Example 2.7: Sine function

Consider the sine function over  $\mathbb{R}$ :  $\sin \in \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}}, x \mapsto \sin x$  (Figure 2.3). The domain is  $\mathcal{D}(\sin) = \mathbb{R}$  and

the range is  $\mathcal{R}(\sin) = [-1, 1]$ . The image of  $\mathcal{A} := [0, \pi/2]$  and pre-image of  $\mathbb{R}_{>0}$  are

$$\sin \mathcal{A} = [0, 1], \quad \sin^{-1} \mathbb{R}_{>0} = \bigcup_{k \in \mathbb{Z}} (2k\pi, (2k+1)\pi).$$

The null space of the sine function is  $\mathcal{N}(\sin) = \{k\pi \mid k \in \mathbb{Z}\}$ . The sine function maps  $\mathbb{R}$  onto  $[-1, 1]$  (surjective with respect to the range  $[-1, 1]$ ), but is not injective (null space contains more than one point).

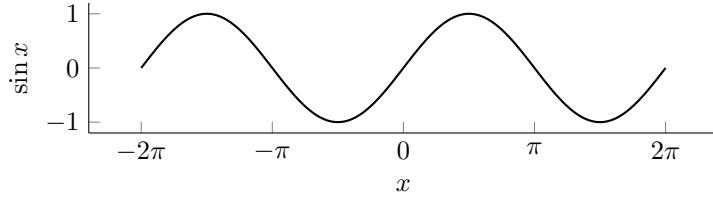


Figure 2.3: The sine function over the domain  $[-2\pi, 2\pi]$ .

### 2.6.1 Linear and bilinear mappings

**Definition 2.6.1** (Linear mapping). Let  $(\mathcal{V}, \mathbb{F})$  be a linear space with linear subspaces  $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}$ . A mapping  $T : \mathcal{V}_1 \rightarrow \mathcal{V}_2$  is a *linear mapping* if  $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$  for all  $x, y \in \mathcal{V}_1$  and scalars  $\alpha, \beta \in \mathbb{F}$ .

**Definition 2.6.2** (Bilinear mapping). Let  $(\mathcal{V}, \mathbb{F})$  be a linear space with linear subspaces  $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \subset \mathcal{V}$ . A mapping  $T : \mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \mathcal{V}_3$  is a *bilinear mapping* if  $T(\alpha x_1 + \beta y_1, z_2) = \alpha T(x_1, z_2) + \beta T(y_1, z_2)$  and  $T(z_1, \alpha x_2 + \beta y_2) = \alpha T(z_1, x_2) + \beta T(z_1, y_2)$  and for all  $x_1, y_1, z_1 \in \mathcal{V}_1, x_2, y_2, z_2 \in \mathcal{V}_2$ , and scalars  $\alpha, \beta \in \mathbb{F}$ .

Let  $(\mathcal{V}, \mathbb{F})$  be a linear space with linear subspaces  $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \subset \mathcal{V}$ . The space of all linear mappings from  $\mathcal{V}_1$  and  $\mathcal{V}_2$  is a linear subspace of  $\mathcal{F}_{\mathcal{V}_1 \rightarrow \mathcal{V}_2}$  over the field  $\mathbb{F}$  with addition and scalar multiplication defined as in (2.10):

$$(T_1 + T_2)(x) = T_1(x) + T_2(x), \quad (\lambda T)(x) = \lambda T(x)$$

for any  $x \in \mathcal{V}_1$  and  $\lambda \in \mathbb{F}$ . Similarly the space of all bilinear mappings between  $\mathcal{V}_1 \times \mathcal{V}_2$  and  $\mathcal{V}_3$  is a linear subspace of  $\mathcal{F}_{\mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \mathcal{V}_3}$  over the field  $\mathbb{F}$  with addition and scalar multiplication are defined as

$$(T_1 + T_2)(x, y) = T_1(x, y) + T_2(x, y), \quad (\lambda T)(x, y) = \lambda T(x, y)$$

for any  $x \in \mathcal{V}_1, y \in \mathcal{V}_2$ , and  $\lambda \in \mathbb{F}$ .

### 2.6.2 Continuity and boundedness

**Definition 2.6.3** (Continuity). Let  $(\mathcal{V}_1, \|\cdot\|)$  and  $(\mathcal{V}_2, \|\cdot\|)$  be normed spaces. A mapping  $F : \mathcal{V}_1 \rightarrow \mathcal{V}_2$  is *continuous* at  $x_0 \in \mathcal{V}_1$  if, for any sequence  $\{x_n\}$  of elements of  $\mathcal{V}_1$  convergent to  $x_0$ , the sequence  $\{F(x_n)\}$  converges to  $F(x_0)$ , i.e.,  $\|x_n - x_0\| \rightarrow 0$  implies  $\|F(x_n) - F(x_0)\| \rightarrow 0$ . If  $F$  is continuous at every  $x_0 \in \mathcal{V}_1$ ,  $F$  is called *continuous*.

**Definition 2.6.4** (Bounded linear mapping). A linear mapping  $T : \mathcal{V}_1 \rightarrow \mathcal{V}_2$  is *bounded* if there exists  $\alpha > 0$  such that  $\|T(x)\| \leq \alpha \|x\|$  for all  $x \in \mathcal{V}_1$ .

**Theorem 2.5.** *A linear mapping is continuous if and only if it is bounded.*

### 2.6.3 Classification: function vs. functional vs. operator

Three types of mappings, classified based on their domain and range, will be used extensively throughout the course. *Functions* are mappings between numeric or array spaces. Let  $\mathbb{F}$  be a field ( $\mathbb{R}$  or  $\mathbb{C}$ ) and

consider  $\Omega \subset \mathbb{F}^d$  (open). A mapping  $f : \Omega \rightarrow \mathbb{F}^m$  is called a *m-vector-valued* function of  $d$  variables because it maps a vector with  $d$  entries to one with  $m$  entries. In the special case where  $m = 1$ , the function is *scalar-valued*. Vector-valued functions can be written as a collection of  $m$  scalar-valued functions, i.e.,  $x \mapsto f(x) = (f_1(x), \dots, f_m(x))$  where  $f_i : \Omega \rightarrow \mathbb{F}$  for  $i = 1, \dots, m$ . A mapping  $f : \Omega \rightarrow M_{m,n}(\mathbb{F})$  is a *(m,n)-matrix-valued* function of  $d$  variables and mappings to higher dimensional arrays are defined similarly, e.g.,  $f : \Omega \rightarrow M_{m,n,k}(\mathbb{F})$  is an *array-valued* function. Matrix-valued and array-valued functions can also be written as arrays of scalar-valued functions, e.g.,  $f : \Omega \rightarrow M_{m,n}(\mathbb{F})$  can be written as

$$x \mapsto f(x) = \begin{bmatrix} f_{11}(x) & \dots & f_{1n}(x) \\ \vdots & \ddots & \vdots \\ f_{m1}(x) & \dots & f_{mn}(x) \end{bmatrix},$$

where  $f_{ij} : \Omega \rightarrow \mathbb{F}$  for  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ . In addition to functions, we will consider *functionals*, mappings between a function space and a field, and *operators*, mappings between function spaces.

### Example 2.8: Functions vs. functionals vs. operators

Any element of  $\mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}}$  is a function, e.g.,

$$\sin : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto \sin x, \quad |\cdot| : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto |x|, \quad \sqrt{\cdot} : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto \sqrt{x} \quad (2.21)$$

are functions. Any definite integral  $\int_a^b : \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}} \rightarrow \mathbb{R}$  is a functional, e.g.,

$$\int_a^b \sin x \, dx = \cos(a) - \cos(b), \quad \int_{-1}^2 |x| \, dx = 2.5, \quad \int_0^a \sqrt{x} \, dx = \frac{2}{3} a^{3/2} \quad (2.22)$$

for  $a, b \in \mathbb{R}$  are functionals. The derivative and antiderivative are operators

$$\frac{d}{dx} : \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}} \rightarrow \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}}, \quad \int_0^x : \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}} \rightarrow \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}} \quad (2.23)$$

because they map functions to other functions, e.g., the derivative operator maps  $\sin \in \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}}$  to  $\cos \in \mathcal{F}_{\mathbb{R} \rightarrow \mathbb{R}}$  and the antiderivative operator maps  $\cos$  to  $\sin$ :

$$\frac{d}{dx}(\sin) = \cos, \quad \int_0^x \cos = \sin. \quad (2.24)$$

## 2.7. Indicial notation

For convenience, brevity, and simplicity, we will make extensive use of *indicial notation* and Einstein summation convention that writes vector/tensor operations explicitly in terms of their components in an abbreviated form by eliminating summation symbols.

### 2.7.1 Range convention

When an equation involves indices, called *free indices*, with each character appearing *once*, it is understood that the equation holds over the range of that variable. Free indices must appear exactly once in every term of an equation, except when the term is numeric, in which case it is repeated for each equation.

### Example 2.9: Range convention

Let  $x \in \mathbb{R}^2$ , then

$$x_i = 1 \iff \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

Let  $a, b, c \in \mathbb{R}^3$ , then

$$a_i^2 = b_i^2 + c_i^2 \iff \begin{cases} a_1^2 = b_1^2 + c_1^2 \\ a_2^2 = b_2^2 + c_2^2 \\ a_3^2 = b_3^2 + c_3^2 \end{cases}$$

Let  $A \in M_{m,n}(\mathbb{R})$ ,  $u \in \mathbb{R}^m$ ,  $v \in \mathbb{R}^n$ , then

$$A_{ij} = u_i v_j \iff \begin{cases} A_{11} = u_1 v_1 \\ \dots \\ A_{m1} = u_m v_1 \\ A_{12} = u_1 v_2 \\ \dots \\ A_{mn} = u_m v_n \end{cases}$$

## 2.7.2 Summation convention

If an index, called a *dummy index*, appears twice in a term of an equation, summation over the range of the variable is implied (Einstein summation convention). Because a dummy index is defined as an index appearing twice in a term, they must come in *pairs*.

### Example 2.10: Summation convention

Let  $A \in M_{2,3,2}(\mathbb{R})$ ,  $B \in M_{2,2}(\mathbb{R})$ , and  $y \in \mathbb{R}^3$ , then

$$B_{ij} = \sum_{k=1}^3 A_{ikj} y_k \iff B_{ij} = A_{ikj} y_k \iff \begin{cases} B_{11} = A_{111} y_1 + A_{121} y_2 + A_{131} y_3 \\ B_{21} = A_{211} y_1 + A_{221} y_2 + A_{231} y_3 \\ B_{12} = A_{112} y_1 + A_{122} y_2 + A_{132} y_3 \\ B_{22} = A_{212} y_1 + A_{222} y_2 + A_{232} y_3 \end{cases}$$

Let  $A \in M_{m,n}(\mathbb{R})$ ,  $v \in \mathbb{R}^m$  and  $w \in \mathbb{R}^n$ , then

$$v = Aw \iff v_i = \sum_{j=1}^n A_{ij} w_j \iff A_{ij} w_j$$

## 2.7.3 Rules

Indicial notation is a collection of conventions or rules that can save considerable time and minimize mistakes if *they are meticulously followed*. Two indicial notation equations are equivalent provided they correspond to the same expanded equation. This implies: 1) the *relative index position* is important, i.e.,  $v_i = A_{ij} w_j$  is not the same as  $v_i = A_{ji} w_j$  unless  $A$  is symmetric (the latter may not even be valid unless  $m = n$ ), 2) the character used to represent indices is not important provided the relative position is maintained, i.e.,  $v_i = A_{ij} w_j$  is equivalent to  $v_r = A_{rs} w_s$ , 3) the relative ordering of terms in the equation and factors in a term is unimportant, i.e.,  $v_i = A_{ij} w_j + y_i$  is equivalent to  $v_i = y_i + w_j A_{ij}$ , and 4) the character used to represent dummy index pairs can be swapped freely provided the pair structure is maintained, i.e.,  $a = b_{jj} + c_{kk} = b_{kk} + c_{jj} = b_{jj} + c_{jj}$ . The rules are summarized as:

- 1) The same free indices must appear in every term of an equation.
- 2) A character used as a free index should never be used as a dummy index.
- 3) The same character should never be used to represent multiple dummy index pairs in a term of an equation.

- 4) An index should appear no more than twice in any term of an equation.

**Example 2.11: Vector and matrix operations**

Standard vector and matrix operations in vector and indicial notation:

- *Scalar times vector*: let  $\alpha \in \mathbb{R}$  and  $v, w \in \mathbb{R}^n$

$$w = \alpha v \iff w_i = \alpha v_i$$

- *Vector addition*: let  $a, b, c \in \mathbb{R}^n$

$$c = a + b \iff c_i = a_i + b_i$$

- *Vector contraction (dot product)*: let  $\alpha \in \mathbb{R}$  and  $a, b \in \mathbb{R}^n$

$$a \cdot b = \sum_{i=1}^n a_i b_i = a_i b_i$$

- *Two norm of vector*:

$$\|v\|_2 = \sqrt{v \cdot v} = \sqrt{\sum_{i=1}^n v_i^2} = \sqrt{v_i v_i}$$

- *Scalar times matrix*: let  $\alpha \in \mathbb{R}$  and  $A, B \in M_{m,n}(\mathbb{R})$

$$B = \alpha A \iff B_{ij} = \alpha A_{ij}$$

- *Matrix addition*: let  $A, B, C \in M_{m,n}(\mathbb{R})$

$$C = A + B \iff C_{ij} = A_{ij} + B_{ij}$$

- *Matrix transpose*: let  $A \in M_{m,n}(\mathbb{R}), B \in M_{n,m}(\mathbb{R})$

$$B = A^T \iff B_{ij} = A_{ji}$$

- *Matrix-vector contraction (matrix-vector product)*: let  $A \in M_{m,n}(\mathbb{R}), v \in \mathbb{R}^n, w \in \mathbb{R}^m$

$$w = Av \iff w_i = A_{ij}v_j$$

- *Matrix-matrix product I*: let  $A \in M_{m,l}(\mathbb{R}), B \in M_{l,n}(\mathbb{R}), C \in M_{m,n}(\mathbb{R})$

$$C = AB \iff C_{ik} = \sum_{j=1}^l A_{ij}B_{jk} = A_{ij}B_{jk}$$

- *Matrix-matrix product II*: let  $A \in M_{l,m}(\mathbb{R}), B \in M_{l,n}(\mathbb{R}), C \in M_{m,n}(\mathbb{R})$

$$C = A^T B \iff C_{ik} = \sum_{j=1}^l A_{ji}B_{jk} = A_{ji}B_{jk}$$

- *Matrix-matrix product III*: let  $A \in M_{m,l}(\mathbb{R}), B \in M_{n,l}(\mathbb{R}), C \in M_{m,n}(\mathbb{R})$

$$C = AB^T \iff C_{ik} = \sum_{j=1}^l A_{ij}B_{kj} = A_{ij}B_{kj}$$

- Matrix-matrix product IV: let  $A \in M_{l,m}(\mathbb{R})$ ,  $B \in M_{n,l}(\mathbb{R})$ ,  $C \in M_{m,n}(\mathbb{R})$

$$C = A^T B^T \iff C_{ik} = \sum_{j=1}^l A_{ji} B_{kj} = A_{ji} B_{kj}$$

- Double contraction of matrix:

$$A : B = \sum_{i=1}^m \sum_{j=1}^n A_{ij} B_{ij} = A_{ij} B_{ij}$$

- Trace of matrix:

$$\text{tr} A = \sum_{i=1}^n A_{ii} = A_{ii}$$

- Frobenius norm of matrix:

$$\|A\|_F = \sqrt{\text{tr}(AA^T)} = \sqrt{A_{ij} A_{ij}}$$

## 2.7.4 Kronecker Delta

The Kronecker delta  $\delta_{ij}^N$  is the indicial notation representation of the  $N \times N$  identity matrix, i.e.,

$$\delta_{ij}^N = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

for  $i, j = 1, \dots, N$ . We will drop the superscript when no risk of confusion. This implies its trace, is  $\delta_{kk}^N = N$ . The Kronecker delta also acts as a *replacement operator* that replaces  $a_i$  with  $a_j$  when multiplied by  $\delta_{ij}$ , i.e.,  $a_i \delta_{ij} = a_1 \delta_{1j} + a_2 \delta_{2j} + \dots + a_N \delta_{Nj} = a_j$ .

### Example 2.12: Kronecker Delta

The replacement property implies

$$A_{ij} \delta_{js} = A_{is}, \quad \delta_{jk} \delta_{km} = \delta_{jm}, \quad A_{ij} \delta_{jk} B_{kl} \delta_{lm} = A_{ik} B_{km}, \quad C_{ijkl} \delta_{ij} \delta_{kl} = C_{jikk}.$$

## 2.8. Differentiation and smoothness of functions

In this section we consider only functions between *real* array spaces; the extension to complex spaces is straightforward, but not needed for this course.

### 2.8.1 Scalar-valued functions of one variable

Consider a real-valued function over  $\Omega \subset \mathbb{R}$  (open)  $f : \Omega \rightarrow \mathbb{R}$ . The derivative of  $f$  at  $x \in \Omega$  is defined as

$$f'(x) = \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x}, \tag{2.25}$$

provided the limit exists, in which case  $f$  is said to be *differentiable at  $x$* . If  $f$  is differentiable at all  $x \in \Omega$  it is simply referred to as *differentiable*. In this case we define the derivative function  $f' : \Omega \rightarrow \mathbb{R}$ , also denoted  $\frac{df}{dx}$ , according to the limit (2.25). If the derivative function  $f'$  is continuous at all  $x \in \Omega$ , the function  $f$  is said to be *continuously differentiable*. Higher derivatives of  $f$  are defined by applying (2.25) recursively, i.e., the third derivative  $f^{(3)} : \Omega \rightarrow \mathbb{R}$  is the derivative function of the second derivative  $f'' : \Omega \rightarrow \mathbb{R}$ , which is the derivative function of  $f'(x)$ . A function  $g : \Omega \rightarrow \mathbb{R}$  is the  $k$ th *order derivative* of  $f : \Omega \rightarrow \mathbb{R}$  if it results from

recursively applying the differentiation formula  $k$  times. Any function that possesses  $k$  derivative functions is said to be  *$k$  times differentiable*. If the derivative functions are continuous it is  *$k$  times continuously differentiable*. A function is *infinitely differentiable* if  $k$  derivative functions exist for any  $k \in \mathbb{N}$ .

### 2.8.2 Array-valued functions of multiple variables

Now consider a vector-valued function over  $\Omega \subset \mathbb{R}^m$  (open),  $f : \Omega \rightarrow \mathbb{R}^n$ , where  $x \mapsto (f_1(x), \dots, f_n(x))$  and  $f_i : \Omega \rightarrow \mathbb{R}$  for  $i = 1, \dots, n$  are real-valued functions. The partial derivative of  $f_i : \Omega \rightarrow \mathbb{R}$  at  $x \in \Omega$  is

$$\frac{\partial f_i}{\partial x_j}(x) := \lim_{t \rightarrow 0} \frac{f_i(x + te^{(j)}) - f_i(x)}{t}, \quad (2.26)$$

provided the limit exists, where  $e^{(j)} \in \mathbb{R}^n$  is the canonical unit vector defined in (2.7). In the case where a partial derivative exist for all  $x \in \Omega$ , we define the partial derivative function  $\frac{\partial f_i}{\partial x_j} : \Omega \rightarrow \mathbb{R}$ . Similar to the single variable case, partial differentiation can be applied recursively to construct higher order derivative functions, including mixed derivatives, i.e., the mixed second derivative function  $\frac{\partial^2 f_k}{\partial x_i \partial x_j}$  is defined as

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j}(x) := \frac{\partial \left( \frac{\partial f_k}{\partial x_i} \right)}{\partial x_j}(x). \quad (2.27)$$

The *order* of the partial derivative is the total number of partial derivatives taken. For sufficiently smooth functions, the ordering in which partial differentiation is applied can be interchanged without modifying the resulting function.

To succinctly describe partial derivatives of vector-valued functions we rely on *multi-index* notation. An  $n$ -dimensional multi-index is an element  $\alpha \in \mathbb{N}_0^n$  with entries  $\alpha = (\alpha_1, \dots, \alpha_n)$ . The *order* or magnitude of the multi-index is given by its sum:  $|\alpha| = \sum_{i=1}^n \alpha_i$ . We will use the following multi-index notation to construct a monomial over  $\mathbb{R}^n$ : let  $x \in \mathbb{R}^n$  with components  $x = (x_1, \dots, x_n)$  and define

$$x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}.$$

It will also be convenient to use the multi-index to index into a multi-dimensional array  $A \in M_{m_1 \dots m_n}(\mathbb{R})$ : define

$$A_\alpha := A_{\alpha_1 \dots \alpha_n}.$$

Finally, the multi-index notation will be used to define a partial derivative of a multi-dimensional function:

$$(D_\alpha f)(x) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(x), \quad (2.28)$$

The order of the derivative is given by  $|\alpha|$ .

For convenience, we adopt the *comma* convention as a shorthand notation for partial derivatives. For an array of any order, all indices appearing after a comma indicate coordinates along which derivatives are taken; all indices appearing before the comma have their usual meaning as indices of the array. The number of indices appearing after the comma determines the order of the derivative. For example, we write

$$f_{i,j} \leftarrow \frac{\partial f_i}{\partial x_j}, \quad f_{k,ij} \leftarrow \frac{\partial^2 f_k}{\partial x_i \partial x_j}. \quad (2.29)$$

In the remainder of the section, we introduce a number of useful derivative functions that will be used extensively throughout the course.

**Definition 2.8.1** (Jacobian matrix). Let  $f : \Omega \rightarrow \mathbb{R}^m$  be a function from an open set  $\Omega \subset \mathbb{R}^n$ . The *Jacobian matrix function*, or Jacobian, denoted  $J : \Omega \rightarrow M_{m,n}(\mathbb{R})$  is the matrix-valued function of partial derivatives of  $f$ , i.e.,

$$J_{ij}(x) := \frac{\partial f_i}{\partial x_j}(x) = f_{i,j}(x). \quad (2.30)$$

**Definition 2.8.2** (Gradient of scalar-valued function). Let  $f : \Omega \rightarrow \mathbb{R}$  be a scalar-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *gradient* is defined as the vector-valued function of partial derivative functions  $\nabla f : \Omega \rightarrow \mathbb{R}^n$

$$[\nabla f(x)]_i = \frac{\partial f}{\partial x_i}(x) = f_{,i}(x) \quad (2.31)$$

**Definition 2.8.3** (Gradient of vector-valued function). Let  $f : \Omega \rightarrow \mathbb{R}^m$  be a vector-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *gradient* is defined as the matrix-valued function of partial derivative functions  $\nabla f : \Omega \rightarrow M_{n,m}(\mathbb{R})$

$$[\nabla f(x)]_{ij} = \frac{\partial f_j}{\partial x_i}(x) = f_{j,i}(x) = J_{ji}(x), \quad (2.32)$$

which coincides with the transpose of the Jacobian of  $f$ .

**Definition 2.8.4** (Gradient of array-valued function). Let  $f : \Omega \rightarrow M_{m_1, \dots, m_d}$  be a array-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *gradient* is defined as the array-valued function of partial derivative functions  $\nabla f : \Omega \rightarrow M_{n, m_1, \dots, m_d}(\mathbb{R})$

$$[\nabla f(x)]_{i_1 \dots i_{d+1}} = \frac{\partial f_{i_2 \dots i_{d+1}}}{\partial x_{i_1}}(x) = f_{i_2 \dots i_{d+1}, i_1}. \quad (2.33)$$

**Definition 2.8.5** (Divergence of vector-valued function). Let  $f : \Omega \rightarrow \mathbb{R}^m$  be a vector-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *divergence* is defined as the real function  $\nabla \cdot f : \Omega \rightarrow \mathbb{R}$

$$\nabla \cdot f(x) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) = f_{,i}(x). \quad (2.34)$$

**Definition 2.8.6** (Divergence of matrix-valued function). Let  $f : \Omega \rightarrow M_{m,n}(\mathbb{R}^m)$  be a matrix-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *divergence* is defined as the vector-valued function  $\nabla \cdot f : \Omega \rightarrow \mathbb{R}^m$

$$[\nabla \cdot f(x)]_i = \sum_{j=1}^n \frac{\partial f_{ij}}{\partial x_j}(x) = f_{ij,j}(x). \quad (2.35)$$

**Definition 2.8.7** (Divergence of array-valued function). Let  $f : \Omega \rightarrow M_{n_1, \dots, n_d}(\mathbb{R}^m)$  be a array-valued function from an open set  $\Omega \subset \mathbb{R}^{n_d}$ . The *divergence* is defined as the array-valued function  $\nabla \cdot f : \Omega \rightarrow M_{n_1, \dots, n_{d-1}}$

$$[\nabla \cdot f(x)]_{i_1 \dots i_{d-1}} = \sum_{j=1}^{n_d} \frac{\partial f_{i_1 \dots i_{d-1} j}}{\partial x_j}(x) = f_{i_1 \dots i_{d-1} j, j}(x) \quad (2.36)$$

**Definition 2.8.8** (Laplacian of scalar-valued function). Let  $f : \Omega \rightarrow \mathbb{R}$  be a scalar-valued function from an open set  $\Omega \subset \mathbb{R}^n$ . The *Laplacian* is defined as the divergence of the gradient  $\Delta f : \Omega \rightarrow \mathbb{R}$

$$\Delta f(x) = (\nabla \cdot \nabla f)(x) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_i}(x) = f_{,ii}(x). \quad (2.37)$$

### 2.8.3 Spaces of smooth functions

Let  $\Omega \subset \mathbb{R}^n$  be an open set and define  $\mathcal{C}(\Omega)$  as the space of continuous real-valued functions over  $\Omega$

$$\mathcal{C}(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ continuous on } \Omega\}.$$

Similarly, define  $\mathcal{C}^k(\Omega)$  ( $k \in \mathbb{N}_0$ ) as the space of real-valued functions over  $\Omega$  with continuous partial derivatives of order  $k$

$$\mathcal{C}^k(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ has continuous partial derivatives of order } k \text{ on } \Omega\}$$

and  $\mathcal{C}^\infty(\Omega)$  as the space of infinitely differentiable (partial derivatives of any order exists and are continuous) real-valued functions over  $\Omega$

$$\mathcal{C}^\infty(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ infinitely differentiable on } \Omega\}.$$

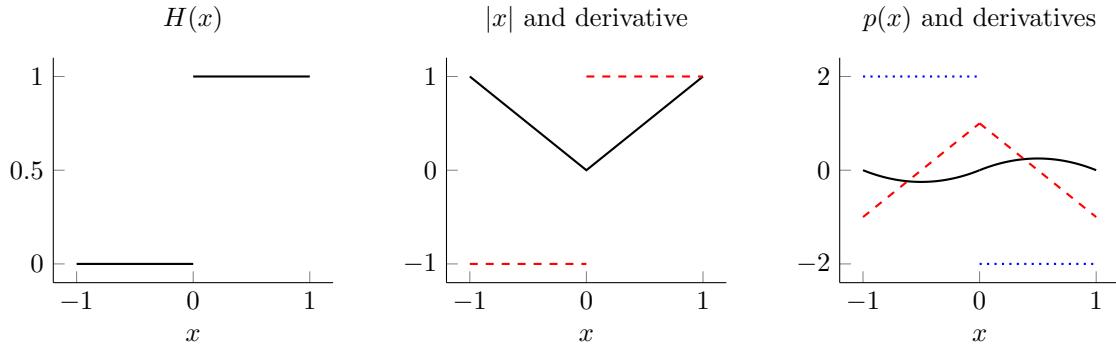


Figure 2.4: The value (—), first derivative (if exists) (---), and second derivative (if exists) (···) of the Heaviside function (left), absolute value function (center), and piecewise polynomial function defined in (2.40) (right).

An important subset of  $\mathcal{C}^\infty(\Omega)$ , denoted  $\mathcal{C}_c^\infty(\Omega)$ , is the set of compactly supported  $\mathcal{C}^\infty(\Omega)$  functions, i.e.,  $f \in \mathcal{C}_c^\infty(\Omega)$  implies  $f \in \mathcal{C}^\infty(\Omega)$  and  $\text{supp } f$  is a compact set.

By definition we have  $\mathcal{C}^0(\Omega) = \mathcal{C}(\Omega)$ . We also use  $\mathcal{C}^{-1}(\Omega)$  to denote the space of real-valued discontinuous, i.e., not continuous, functions over  $\Omega$

$$\mathcal{C}^{-1}(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ discontinuous}\}.$$

### Example 2.13: Real-valued functions

Most real-valued function with which we are familiar, e.g., polynomials, exponential and trigonometric functions, are infinitely differentiable (belong to  $\mathcal{C}^\infty(\mathbb{R})$ ). The *Heaviside function*  $H : \mathbb{R} \rightarrow \mathbb{R}$  is defined as

$$x \mapsto H(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0. \end{cases} \quad (2.38)$$

Sometimes, for convenience, the Heaviside function is defined at  $x = 0$ :  $H(0) = 0.5$ . The Heaviside function is discontinuous:  $H \in \mathcal{C}^{-1}(\mathbb{R})$  (Figure 2.4). The *absolute value function*  $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}$ , defined as

$$x \mapsto |x| := \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}, \quad (2.39)$$

is a continuous function but not differentiable at 0:  $|\cdot| \in \mathcal{C}^0(\mathbb{R})$  (Figure 2.4). Functions belonging to  $\mathcal{C}^k(\Omega)$  are trickier to construct and usually take the form of piecewise polynomial functions with special conditions where different polynomials meet. For example, the function  $p : \mathbb{R} \rightarrow \mathbb{R}$  defined as

$$x \mapsto p(x) := \begin{cases} -x^2 + x & \text{if } x \geq 0 \\ x^2 + x & \text{if } x < 0 \end{cases} \quad (2.40)$$

is continuously differentiable  $p \in \mathcal{C}^1(\Omega)$  (Figure 2.4), which can be verified by observing the function and its derivative are  $\mathcal{C}^\infty(\mathbb{R} \setminus \{0\})$  and continuous at 0 (but the second derivative is not).

To construct the corresponding function spaces for vector-valued functions, observe that any function  $f : \Omega \rightarrow \mathbb{R}^n$ , where  $\Omega$  is an open subset of  $\mathbb{R}^m$  and  $x \mapsto (f_1(x), \dots, f_n(x))$ , is a Cartesian product of real-valued functions  $f_i : \Omega \rightarrow \mathbb{R}$  for  $i = 1, \dots, n$ . This implies that function spaces for vector-valued functions can easily be constructed as Cartesian products of function spaces for scalar-valued functions, e.g.,  $[\mathcal{C}^\infty(\Omega)]^n$  is the space of infinitely differentiable  $n$ -vector-valued functions over  $\Omega$ .

## 2.9. Integral identities

### 2.9.1 Divergence theorem

The divergence theorem relates the volume integral of the divergence of a continuously differentiable function to a boundary integral of the function itself. In one-dimension, this is the well-known fundamental theorem of calculus: let  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  be continuously differentiable in the interval  $[a, b] \subset \mathbb{R}$ , i.e.,  $\phi \in C^1([a, b])$ , then

$$\int_a^b \phi'(x) dx = [\phi]_a^b = \phi(b) - \phi(a). \quad (2.41)$$

In multiple dimensions, let  $\Omega$  be an compact subset of  $\mathbb{R}^d$  with piecewise smooth boundary  $\partial\Omega$  and outward unit normal  $n : \partial\Omega \rightarrow \mathbb{R}^d$ . The divergence theorem states that the following identities hold

$$\int_{\Omega} f_{i,i} dV = \int_{\partial\Omega} f_i n_i dS, \quad (2.42)$$

where  $f : \Omega \rightarrow \mathbb{R}^n$  is a continuously differentiable, vector-valued function. It can be extended to matrix-valued functions:

$$\int_{\Omega} F_{ij,j} dV = \int_{\partial\Omega} F_{ij} n_j dS, \quad (2.43)$$

$F : \Omega \rightarrow M_{m,n}(\mathbb{R})$  is a continuously differentiable matrix-valued function.

### 2.9.2 Integration-by-parts

A particularly useful integral identity for our study of the finite element method is integration-by-parts, which follows directly from the divergence theorem, and allows one to move a derivative from one term in a product to another. Let  $\Omega$  be an compact subset of  $\mathbb{R}^d$  with piecewise smooth boundary  $\partial\Omega$  and outward unit normal  $n : \partial\Omega \rightarrow \mathbb{R}^d$ . Then consider the continuously differentiable vector-valued function  $f : \Omega \rightarrow \mathbb{R}^d$  and continuously differentiable real-valued function  $w : \Omega \rightarrow \mathbb{R}$ . The product rule of differentiation states:

$$(f_i w)_{,i} = f_{i,i} w + f_i w_{,i}. \quad (2.44)$$

Next we integrate this equation over  $\Omega$  and apply the divergence theorem to the term on the left side to obtain

$$\int_{\Omega} f_{i,i} w dV = \int_{\partial\Omega} w f_i n_i dS - \int_{\Omega} f_i w_{,i}, \quad (2.45)$$

which is written in vector notation as

$$\int_{\Omega} w \nabla \cdot f dV = \int_{\partial\Omega} w f \cdot n dS - \int_{\Omega} f \cdot \nabla w dV. \quad (2.46)$$

For the special case of  $d = 1$ , this reduces to the well-known integration-by-parts formula

$$\int_a^b w f' dx = [w f]_a^b - \int_a^b f w' dx, \quad (2.47)$$

which can easily be derived directly from the product rule  $\frac{d}{dx}(wf) = w'f + wf'$  and fundamental theorem of calculus. Finally, define  $\phi \in C^2(\Omega)$  and  $f : \Omega \rightarrow \mathbb{R}^d$ , where  $f_i := \phi_{,i}$ . In this case, the integration-by-parts formula yields

$$\int_{\Omega} w \phi_{,ii} dV = \int_{\partial\Omega} w \phi_{,i} n_i dS - \int_{\Omega} w_{,i} \phi_{,i} dV, \quad (2.48)$$

which is written in vector notation as

$$\int_{\Omega} w \Delta \phi dV = \int_{\partial\Omega} w \nabla \phi \cdot n dS - \int_{\Omega} \nabla w \cdot \nabla \phi dV. \quad (2.49)$$

## 2.10. Partial differential equations

A partial differential equation (PDE) is a relationship between an (unknown) function of several variables and its partial derivatives over a domain  $\Omega \subset \mathbb{R}^d$  with boundary  $\partial\Omega$ . If the function only depends on one variable, it is called an ordinary differential equation. Let  $u(x_1, \dots, x_d, t)$  be an unknown function (dependent variable) with independent variables  $x_1, \dots, x_d, t$ . The first  $d$  variables ( $x_1, \dots, x_d$ ) usually represent spatial variables and the last variable ( $t$ ) is usually time.

To be well-posed, a PDE must be equipped with boundary conditions (BCs) and an initial condition (IC). Problems that depend on only space  $x_1, \dots, x_d$  (not time  $t$ ) require only BCs (not an IC) and are called *boundary value problem*. On the other hand, problems that depend on only time  $t$  (not space  $x_1, \dots, x_d$ ) require only an IC (not BCs) and are called *initial value problem*. Problems that depends on both space and time are called *initial boundary value problems* and must be equipped with both BCs and an IC.

The *order* of the PDE is determined by the highest derivative in the equation. A PDE is called *linear* if it is of first degree in all its field (dependent) variables and their partial derivatives. A *system* of partial differential equations is a collection of several PDEs for several unknown functions. PDEs that do not depend on time  $t$  are called *static* or steady, otherwise they are *time-dependent* or unsteady.

### Example 2.14: Poisson equation

The Poisson equation is a second-order, linear, static PDE over a domain  $\Omega \subset \mathbb{R}^d$

$$-\Delta u = f. \quad (2.50)$$

The solution of the Poisson equation over a disk ( $d = 2$ ) subject to homogeneous essential boundary conditions is shown in Figure 2.5.

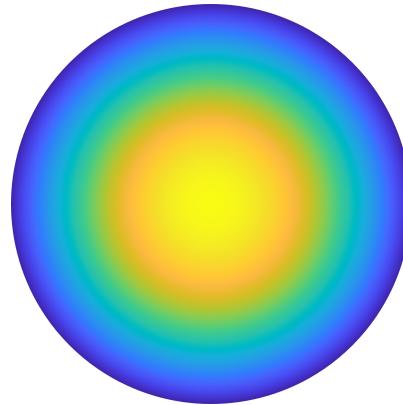


Figure 2.5: Solution of the Poisson equation over the disk.

### Example 2.15: Linear elasticity

The linear elasticity equations are a system of  $d$  second-order, linear, static PDEs over a domain  $\Omega \subset \mathbb{R}^d$  that model deformation of a linear elastic structure under infinitesimal strains

$$\sigma_{ij,j} = f_i, \quad \sigma_{ij} = C_{ijkl}\epsilon_{kl}, \quad \epsilon_{kl} = \frac{1}{2}(u_{k,l} + u_{l,k}) \quad (2.51)$$

where all indices range from  $1, \dots, d$ ,  $u : \Omega \rightarrow \mathbb{R}^d$  is the (unknown) displacement function,  $\epsilon : \Omega \rightarrow M_{d,d}(\mathbb{R})$  is the strain,  $\sigma : \Omega \rightarrow M_{d,d}(\mathbb{R})$  is the stress,  $C : \Omega \rightarrow M_{d,d,d,d}(\mathbb{R})$  is the elasticity tensor, and  $f : \Omega \rightarrow \mathbb{R}^d$  is the body load. The solution of the linear elasticity equations over the Batman domain ( $d = 2$ ) subject to some boundary conditions is shown in Figure 2.6.



Figure 2.6: Solution of the linear elasticity equations over the Batman domain.

### Example 2.16: Incompressible Navier-Stokes equations

The incompressible Navier-Stokes equations are a system of  $d+1$  second-order, nonlinear, time-dependent PDEs over a domain  $\Omega \subset \mathbb{R}^d$  that model flow of a fluid with constant density

$$-(\rho\nu v_{i,j})_j + \rho v_j v_{i,j} + P_{,i} = 0, \quad v_{j,j} = 0 \quad (2.52)$$

where  $v : \Omega \rightarrow \mathbb{R}^d$  is the (unknown) fluid velocity,  $P : \Omega \rightarrow \mathbb{R}_{>0}$  is the (unknown) fluid pressure,  $\rho \in \mathbb{R}_{>0}$  is the fluid density, and  $\nu : \Omega \rightarrow \mathbb{R}_{>0}$  is the fluid viscosity. The solution of the incompressible Navier-Stokes equations in the steady limit over the ND logo domain ( $d = 2$ ) subject to some boundary conditions is shown in Figure 2.7.

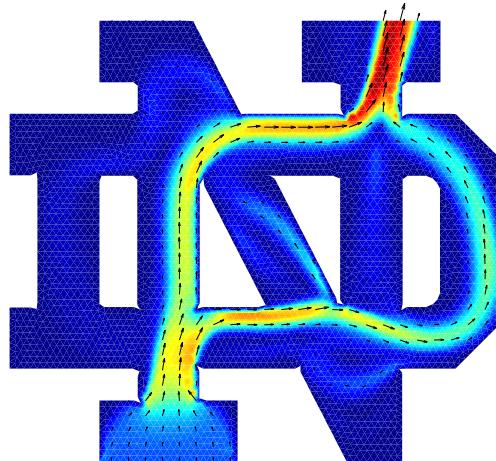


Figure 2.7: Solution of incompressible Navier-Stokes equation over the ND domain.

### Example 2.17: Compressible Euler equations

The compressible Euler equations are a system of  $d+2$  first-order, nonlinear, time-dependent PDEs over a domain  $\Omega \subset \mathbb{R}^d$  that models high-speed gas flows

$$\rho_{,t} + (\rho v_j)_{,j} = 0, \quad (\rho v_i)_{,t} + (\rho v_i v_j + P\delta_{ij})_{,j} = 0, \quad (\rho E)_{,t} + (\rho Hv_i)_{,j} = 0 \quad (2.53)$$

where  $\rho : \Omega \rightarrow \mathbb{R}_{>0}$  is the fluid density,  $v : \Omega \rightarrow \mathbb{R}$  is the fluid velocity,  $E : \Omega \rightarrow \mathbb{R}_{>0}$  is the total energy,  $H : \Omega \rightarrow \mathbb{R}_{>0}$  is the total enthalpy, and  $P : \Omega \rightarrow \mathbb{R}_{>0}$  is the pressure. The solution of the compressible Euler equations in the steady limit over a NACA0012 airfoil ( $d = 2$ ) subject to some boundary conditions is shown in Figure 2.8.

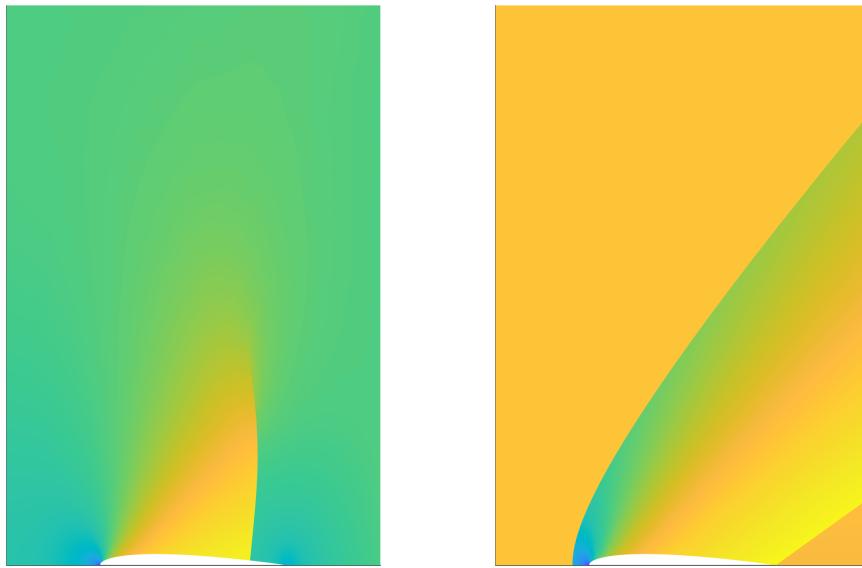


Figure 2.8: Solution (Mach) of compressible Euler equations for a transonic (*left*) and supersonic inlet flow (*right*).

## 2.11. Summary

This chapter provided a detailed account of the mathematical concepts that will be used throughout the document:

- 1) *Indicial notation* is a set of rules and conventions that can significantly reduce shorten and clarify notation involving vector/tensor algebra and calculus.
- 2) A *linear space* is the fundamental mathematical space upon which most of the development in this section is built. It is a set endowed with two operations: addition and scalar multiplication. A linear space is an abstraction of vectors in Euclidean spaces.
- 3) A *basis* is a linearly independent set of vectors that spans a linear space. The (unique) number of vectors in a basis of a linear space is the dimension of the space.
- 4) A *normed space* is a linear space equipped with a *norm*, which introduces the concept of distance and length. This provides the necessary structure to define convergence and a topology (open/closed sets).
- 5) *Mappings* are transformations between two sets. Linear and bilinear mappings possess the special property of linearity in (both) its arguments. For mappings between numerical sets, we introduced the concept of differentiation and carefully considered relevant cases: single vs. multiple variables, scalar-valued vs. vector-valued vs. matrix-valued vs. array-valued mappings.
- 6) The *divergence theorem* relates the volume integral of the divergence of a continuously differentiable function to a boundary integral of the function itself. *Integration-by-parts* is a useful identity to exchange a derivative between terms in a product.
- 7) Partial differential equations are classified by their order (highest derivative), linear vs. nonlinear, the behavior of their solutions, scalar PDE vs. systems of PDEs, and static vs. time-dependent.