Lecture Notes of MTH201

Erfang Ma

December 13, 2024

Contents

Chapter 1

Scalar and Vector Fields

1.1 Scalars and Vectors

A scalar is a quantity that has magnitude only and can be specified by a real number with a unit. Temperature, potential, pressure, density (defined as mass per unit volume) are examples of scalars. A vector is a quantity that has both a magnitude and a direction. Force, velocity, displacement are examples of vectors.

A scalar is often denoted by a letter in italic e.g., a, b, w. A vector is often denoted by a letter in bold face italic e.g., $\boldsymbol{E}, \boldsymbol{F}$ in print and by a letter with an arrow above it e.g. $\overrightarrow{E}, \overrightarrow{F}$.

The position vector of a point P in a (Cartesian, Spherical or Cylindrical) coordinate system with origin O is the vector $\overrightarrow{OP}$. If P has Cartesian coordinates (x, y, z), then

$$\overrightarrow{OP} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}} \tag{1.1}$$

Here $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$ is the unit vector in the positive direction of x, y and z axis of the Cartesian coordinate system, respectively. If P has Spherical coordinates (r, ϕ, θ), then

$$\overrightarrow{OP} = r\hat{\boldsymbol{r}}(P) + \phi\hat{\boldsymbol{\phi}}(P) + \theta\hat{\boldsymbol{\theta}}(P). \tag{1.2}$$

Here $\hat{\boldsymbol{r}}(P)$, $\hat{\boldsymbol{\phi}}(P)$ and $\hat{\boldsymbol{\theta}}(P)$ are unit vectors at P and in the increasing directions of r, ϕ and θ coordinates, respectively. If P has Cylindrical coordinates (ρ, ϕ, z), then

$$\overrightarrow{OP} = \rho\hat{\boldsymbol{\rho}}(P) + \phi\hat{\boldsymbol{\phi}}(P) + z\hat{\boldsymbol{z}} \tag{1.3}$$

Here $\hat{\boldsymbol{\rho}}(P)$ and $\hat{\boldsymbol{\phi}}(P)$ are unit vectors at P and in the increasing directions of ρ and ϕ coordinates, respectively.

Infinitesimal Displacement Vector is a vector that represents an extremely small change in position. It is

$$\mathrm{d}\boldsymbol{s} = \mathrm{d}x\,\hat{\boldsymbol{x}} + \mathrm{d}y\,\hat{\boldsymbol{y}} + \mathrm{d}z\,\hat{\boldsymbol{z}}. \tag{1.4}$$

in Cartesian coordinates,

$$\mathrm{d}\boldsymbol{s} = \mathrm{d}r\,\hat{\boldsymbol{r}} + r\sin\theta\mathrm{d}\phi\,\hat{\boldsymbol{\phi}} + r\mathrm{d}\theta\,\hat{\boldsymbol{\theta}} \tag{1.5}$$

in Spherical coordinates and

$$\mathrm{d}\boldsymbol{s} = \mathrm{d}\rho\,\hat{\boldsymbol{\rho}} + \rho\mathrm{d}\phi\,\hat{\boldsymbol{\phi}} + \mathrm{d}z\,\hat{\boldsymbol{z}} \tag{1.6}$$

in cylindrical coordinates.

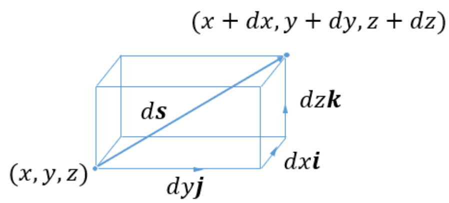

Figure 1.1: Decomposition of infinitesimal displacement vector in Cartesian coordinates

1.2 Scalar and Vector Fields

A scalar field is a distribution of scalars each of which is associated with a point in space.

Example 1.1. *The distribution of temperature u in a room Ω is an example of scalar field. Assume*

$$u(P) = 2x + 3y + z \qquad \text{in } \Omega. \tag{1.7}$$

where (x, y, z) are the Cartesian coordinates of P.

Then the temperature u at a point $P \in \Omega$ with Cartesian coordinates $(x = 1, y = 2, z = 3)$ is

$$u(P) = 2 \cdot 1 + 3 \cdot 2 + 1 \cdot 3 = 11 \tag{1.8}$$

The temperature u at another point $P' \in \Omega$ with Cartesian coordinates $(x = 10, y = 20, z = 30)$ is

$$u(P) = 2 \cdot 10 + 3 \cdot 20 + 1 \cdot 30 = 110 \tag{1.9}$$

Example 1.2. *The distribution of electrical potential u in a region is also an example of scalar field. According to electrostatics, the electrical potential u due to a point charge q fixed at the origin O of a coordinate system could be*

$$u(P) = \frac{q}{4\pi\epsilon_0 |\overrightarrow{OP}|} \qquad P \neq O \tag{1.10}$$

where ϵ_0 is some physical constant.

Therefore, the potential u at a point P with Cartesian coordinates $(x = 3, y = 2, z = 1)$ is

$$u(P) = \frac{q}{4\pi\epsilon_0\sqrt{3^2 + 2^2 + 1^2}} = \frac{q}{4\pi\epsilon_0\sqrt{14}}. \tag{1.11}$$

The potential at another point P' with Spherical coordinates $(r = 2, \theta = \pi/3, \phi = \pi/5)$ is

$$u(P) = \frac{q}{4\pi\epsilon_0 r} = \frac{q}{8\pi\epsilon_0}. \tag{1.12}$$

A vector field is a distribution of vectors each of which is associated with a point in space.

Example 1.3. *The electric field $\boldsymbol{E}$ is a typical example of vector field. According to electrostatics, the electric field $\boldsymbol{E}$ induced by a point charge q fixed at the origin O of a coordinate systems is*

$$\boldsymbol{E}(P) = \frac{q}{4\pi\epsilon_0} \frac{\overrightarrow{OP}}{|\overrightarrow{OP}|^3} \qquad P \neq O \tag{1.13}$$

Here ϵ_0 is the same physical constant as in Example (1.2).

The electric field vector $\boldsymbol{E}$ at a point P with Cartesian coordinates $(x = 1, y = 2, z = 3)$ is

$$\boldsymbol{E}(P) = \frac{q}{4\pi\epsilon_0} \frac{1 \cdot \hat{\boldsymbol{x}} + 2 \cdot \hat{\boldsymbol{y}} + 3 \cdot \hat{\boldsymbol{z}}}{\sqrt{1^2 + 2^2 + 3^2}^3} \tag{1.14}$$

The electric field vector $\boldsymbol{E}$ at another point P' with Spherical coordinates $(r = 2, \theta = \pi/3, \phi = \pi/5)$ is

$$\boldsymbol{E}(P) = \frac{q}{4\pi\epsilon_0} \frac{r\hat{\boldsymbol{r}}}{r^3} = \frac{q}{4\pi\epsilon_0} \frac{\hat{\boldsymbol{r}}}{r^2} = \frac{q}{4\pi\epsilon_0} \frac{\hat{\boldsymbol{r}}}{2^2} = \frac{q}{16\pi\epsilon_0} \hat{\boldsymbol{r}} \tag{1.15}$$

By evaluating $\boldsymbol{E}$ similarly at other points, we could obtain a sketch of this vector field $(q > 0)$ as shown in Figure 1.2. Here each electric field vector is placed at its corresponding point. The length of all the vectors are appropriately scaled so that they do not overlap with each other.

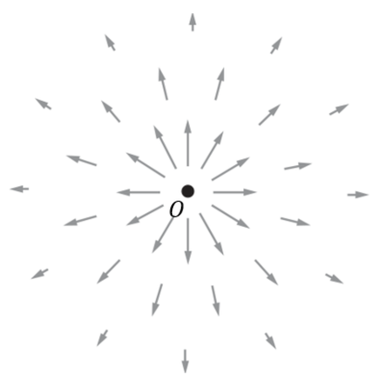

Figure 1.2: A sketch of vector field defined in (1.13) with $q > 0$. This sketch consists of electric field vectors at various points around the origin O at the center. The tail of each electric field vector is placed at the corresponding point.

Example 1.4. *Magnetic field $\boldsymbol{B}$ is also an example of vector field. Assume an infinitely long straight wire lying along the z-axis carries an electric current I flowing in the $\hat{\boldsymbol{z}}$ direction. According to Ampere's law in Magnetostatics, the magnetic field due to this current is*

$$\boldsymbol{B}(P) = \frac{\mu_0 I}{2\pi\rho} \hat{\boldsymbol{\phi}}(P) \qquad \phi > 0 \tag{1.16}$$

where ρ and ϕ are the cylindrical coordinates of P.

For example, the magnetic field vector $\boldsymbol{B}$ at a point P with cylindrical coordinates $(\rho = 3, \phi = \pi/3, z = 13)$ is

$$\boldsymbol{B}(P) = \frac{\mu_0 I}{6\pi}\hat{\boldsymbol{\phi}}(P) \tag{1.17}$$

We discuss scalar and vector fields in the first four chapters. We study these fields by considering their integrals, 'derivatives' and the relations between integrals and derivatives.

Exercises

1.1. *(scalar or vector) Bellow are a list of some important quantities in electromagnetism. Are they scalar or vector quantities ?*

electric charge, the electric force between two charges, energy associated with the electric field, electrical conductivity, capacitance, current density

1.2. *(conversion between different coordinates) Assume O is the origin of the coordinate system. P is a point with Cartesian coordinates $(x = 2, y = 1, z = 3)$. What are the Spherical and Cylindrical coordinates of P ? What is the vector $\overrightarrow{OP}$ in different coordinate? Assume Q is another point with Spherical coordinates $(r = 3, \theta = 0, \phi = \pi)$. What is the Cartesian coordinates of Q ? Evaluate the following two products*

$$\overrightarrow{OP} \cdot \overrightarrow{OQ} \quad \textit{and} \quad \overrightarrow{OP} \times \overrightarrow{OQ} \tag{1.18}$$

1.3. *(identities of vector algebra) Fill in the blanks, so that the following equations are true for any vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$*

$$\boldsymbol{a} + \boldsymbol{b} = \underline{\hspace{2cm}} + \boldsymbol{a} \tag{1.19a}$$
$$\boldsymbol{a} \cdot \boldsymbol{b} = \underline{\hspace{2cm}} \cdot \boldsymbol{a} \tag{1.19b}$$
$$\boldsymbol{a} \times \boldsymbol{b} = \underline{\hspace{2cm}} \times \boldsymbol{a} \tag{1.19c}$$
$$\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \underline{\hspace{2cm}} + \underline{\hspace{2cm}} \cdot \boldsymbol{c} \tag{1.19d}$$
$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \underline{\hspace{2cm}} \times \boldsymbol{b} + \boldsymbol{a} \times \underline{\hspace{2cm}} \tag{1.19e}$$
$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\underline{\hspace{2cm}}(\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot \underline{\hspace{2cm}} \tag{1.19f}$$
$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \underline{\hspace{2cm}})\boldsymbol{b} + \underline{\hspace{2cm}}\boldsymbol{c} \tag{1.19g}$$

1.4. *(evaluation of scalar fields) Consider the potential field in Example 1.2. Compare the potential at a point with Cartesian coordinates* $(x = 1, y = 0, z = 2)$*, to another point with spherical coordinates* $(r = 2, \theta = \pi, \phi = \pi/2)$*. Which one is bigger ?*

1.5. *(evaluation of vector fields) Consider the electric field in Example 1.13. Find three points* P_1, P_2 *and* P_3 *such that*

$$\boldsymbol{E}(P_1) = -\boldsymbol{E}(P_2) = 2\boldsymbol{E}(P_3) \tag{1.20}$$

1.6. *(Sketch of vector fields) A vector field* $\boldsymbol{E}$ *defined as*

$$\boldsymbol{E}(P) = f(x, y, z)\hat{\boldsymbol{x}} + g(x, y, z)\hat{\boldsymbol{y}} + h(x, y, z)\hat{\boldsymbol{z}} \tag{1.21}$$

for any point P *with Coordinates* (x, y, z) *is also often written as*

$$\boldsymbol{E}(x, y, z) = f(x, y, z)\hat{\boldsymbol{x}} + g(x, y, z)\hat{\boldsymbol{y}} + h(x, y, z)\hat{\boldsymbol{z}}. \tag{1.22}$$

Sketch the following three vector fields

$$\begin{aligned}
\boldsymbol{F}(x, y, z) &= \hat{\boldsymbol{y}} \\
\boldsymbol{F}(x, y, z) &= x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} \\
\boldsymbol{F}(x, y, z) &= \frac{1}{\sqrt{x^2 + y^2}}(x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}})
\end{aligned}$$

1.7. *(Sketch of vector fields) Match vector fields a-d that are defined on the xy-plane, with graphs A-D in Figure 1.3.*

a. $\boldsymbol{F}(x, y) = x^2\hat{\boldsymbol{y}}$ *b.* $\boldsymbol{F}(x, y) = (x - y)\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}}$

c. $\boldsymbol{F}(x, y) = 2x\hat{\boldsymbol{x}} - y\hat{\boldsymbol{y}}$ *d.* $\boldsymbol{F}(x, y) = y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}}$

1.8. *Find a vector field* $\mathbf{E}$ *that is defined on the xy-plane with the following properties, respectively.*

1. $\mathbf{E}$ *is everywhere normal to the line* $y = x$ *in the xy-plane.*
2. $\mathbf{E}$ *is everywhere normal to the line* $x = 2$ *in the xy-plane.*
3. *At all points except* $(0, 0)$*,* $\boldsymbol{E}$ *has unit magnitude and points away from the origin along radial lines.*

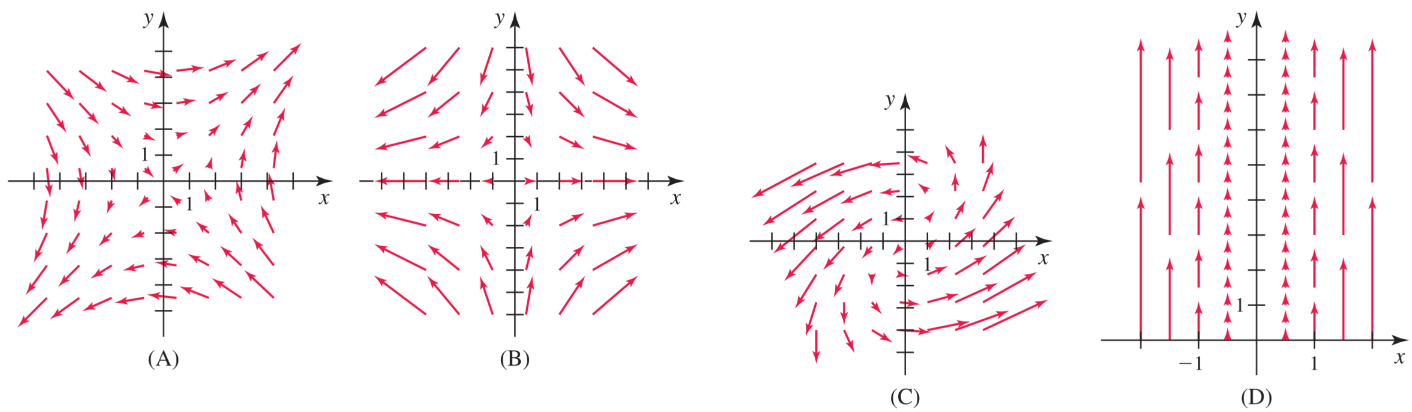

Figure 1.3

Chapter 2

Line Integral and Surface Integral

2.1 Line Integral

Assume f is a real-valued function and defined on a curve Γ in space. The line integral of f along the curve Γ is

$$\int_\Gamma f \, \mathrm{d}s \overset{\text{def}}{=} \lim_{\Delta s_i \to 0} \sum_i f(P_i)\Delta s_i \tag{2.1}$$

Here curve Γ is divided into many parts. P_i is a point on the i-th part of the curve with length Δs_i (Figure 2.1) and $f(P_i)$ is the value of f at the point $P_i, i = 1, 2, ...$ As $\Delta s_i \to 0$, the curve Γ is divided into more and more parts and the length of each part approaches zero.

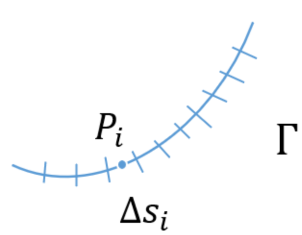

Figure 2.1: A curve Γ is divided into many parts. The i-th part is of length Δs_i and contains a point P_i.

If Γ is a closed path, the line integral along Γ is denoted by

$$\oint_\Gamma f \, \mathrm{d}s$$

Example 2.1. *Assume function f takes a constant value 2 everywhere on a curve Γ. In other words,*

$$f(P) = 2 \qquad \text{for any } P \in \Gamma \tag{2.2}$$

Then the line integral of f along Γ is

$$\int_\Gamma f \, \mathrm{d}\Gamma = \lim_{\Delta s_i \to 0} 2\Delta s_i = 2 \lim_{\Delta s_i \to 0} \Delta s_i = 2|\Gamma| \tag{2.3}$$

with $|\Gamma|$ the arc length of Γ.

Example 2.2. *Γ is a straight line segment between the point with Cartesian coordinate $(x = 0, y = 0, z = 0)$ (i.e. the origin) and the point with Cartesian coordinate $(x = 1, y = 0, z = 0)$. Function f is defined on Γ such that*

$$f(P) \stackrel{\text{def}}{=} x + y. \tag{2.4}$$

for any $P \in \Gamma$ with Cartesian coordinates (x, y, z).

In this case, the line integral of f along Γ is

$$\int_\Gamma f \, \mathrm{d}s = \int_0^1 x \, \mathrm{d}x = \frac{1}{2} \tag{2.5}$$

Example 2.3. *Assume Γ is the half circle from the origin O to the point $(x = 1, y = 0, z = 0)$ as shown in Figure. The value of line integral of f defined in (2.4) along Γ could be estimated as follows.*

For any point $P \in \Gamma$, its Cartesian coordinates x, y satisfy

$$0 \le x \le 1, \quad 0 \le y \le 1. \tag{2.6}$$

So

$$0 \le f(P) = x + y \le 2. \tag{2.7}$$

Then

$$\int_\Gamma f \, \mathrm{d}s = \lim_{\Delta s_i \to 0} f(P_i)\Delta s_i \le \lim_{\Delta s_i \to 0} 2 \cdot \Delta s_i = 2|\Gamma| \tag{2.8}$$

$$\int_\Gamma f \, \mathrm{d}s = \lim_{\Delta s_i \to 0} f(P_i)\Delta s_i \ge \lim_{\Delta s_i \to 0} 0 \cdot \Delta s_i = 0 \tag{2.9}$$

In short,

$$0 \leq \int_\Gamma f \, \mathrm{d}s \leq 2|\Gamma| \tag{2.10}$$

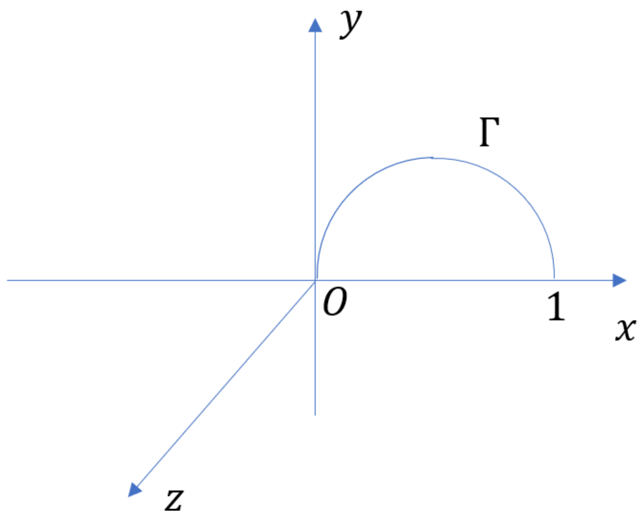

Figure 2.2: The half circle Γ on the xy-plane. It is centered at $(x = 1/2, y = 0, z = 0)$ and has radius $1/2$.

2.2 Surface Integral

Surface integral is defined similarly as the line integral on the above. Assume f is real-valued function and defined on a surface S in space. The surface integral of f over the surface S is

$$\int_S f \, \mathrm{d}S \stackrel{\text{def}}{=} \lim_{\Delta S_i \to 0} \sum_i f(P_i) \Delta S_i. \tag{2.11}$$

Here the surface S is divided into many parts. P_i is a point on the i-th part of the surface with area Δs_i, and $f(P_i)$ is the value of f at the point $P_i, i = 1, 2, \ldots.$ As $\Delta S_i \to 0$, the surface S is divided into more and more parts, and the area of each part approaches zero.

If S is a closed surface, then the surface integral over S is denoted by

$$\oint_S f \, \mathrm{d}S.$$

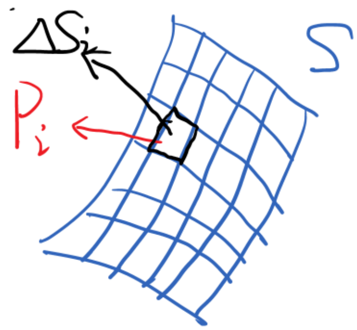

Figure 2.3: The surface is divided into many parts. The i-th part has area ΔS_i and contains a point $P_i, i = 1, 2, \ldots$.

Example 2.4. *Assume function f takes a constant value 2 everywhere on a surface S. In other words,*

$$f(P) = 2 \qquad \text{for any } P \in S \tag{2.12}$$

Then the surface integral of f over S is

$$\int_S f \, \mathrm{d}S = \lim_{\Delta S_i \to 0} 2\Delta S_i = 2 \lim_{\Delta S_i \to 0} \Delta S_i = 2|S| \tag{2.13}$$

with $|S|$ denoting the area of S.

Example 2.5. *S is a surface on the xy-plane and function f is defined on S such that*

$$f(P) = x + y \tag{2.14}$$

for any $P \in S$ with Cartesian coordinates (x, y, z).

In this case, the surface integral of f over S is

$$\int_S f \, \mathrm{d}S = \int_S f(x, y) \mathrm{d}x\mathrm{d}y = \int_S (x + y) \, \mathrm{d}x\mathrm{d}y \tag{2.15}$$

i.e. a double integral of function $f(x, y) = x + y$ over the region S on the xy-plane.

Example 2.6. *Assume surface S is defined be*

$$\{(x, y, z) | x^2 + y^2 + z^2 = 3, x \geq 0, y \geq 0\}. \tag{2.16}$$

Function f is defined on S such that

$$f(P) \overset{\text{def}}{=} x + y \tag{2.17}$$

for any $P \in S$ with Cartesian coordinates (x, y, z).

The value of surface integral of f over S could be estimated as follows. For any point $P \in S$, its Cartesian coordinates satisfy

$$0 \leq x \leq 3, \quad 0 \leq y \leq 3. \tag{2.18}$$

So

$$0 \leq f(P) = x + y \leq 3 + 3 = 6 \tag{2.19}$$

Then

$$\int_S f \,\mathrm{d}S = \lim_{\Delta S_i \to 0} f(P_i)\Delta S_i \leq \lim_{\Delta S_i \to 0} 6\Delta S_i = 6|S| \tag{2.20}$$

$$\int_S f \,\mathrm{d}S = \lim_{\Delta S_i \to 0} f(P_i)\Delta S_i \geq \lim_{\Delta S_i \to 0} 0\Delta S_i = 0 \tag{2.21}$$

In short,

$$0 \leq \int_S f \,\mathrm{d}S \leq 6|S| \tag{2.22}$$

2.3 Line Integral of Vector Field

Definition. Assume $\boldsymbol{F}$ is a vector field defined on a curve Γ. $\boldsymbol{T}$ is defined on Γ such that $\boldsymbol{T}(P)$ is a unit vector and tangent to Γ at P for any $P \in \Gamma$. The line integral of $\boldsymbol{F}$ along Γ in the direction of $\boldsymbol{T}$ is

$$\int_\Gamma \boldsymbol{F} \cdot \boldsymbol{T} \,\mathrm{d}s \tag{2.23}$$

i.e. the line integral of f along Γ with

$$f(P) \overset{\text{def}}{=} \boldsymbol{F}(P) \cdot \boldsymbol{T}(P) \qquad \text{for any } P \in \Gamma. \tag{2.24}$$

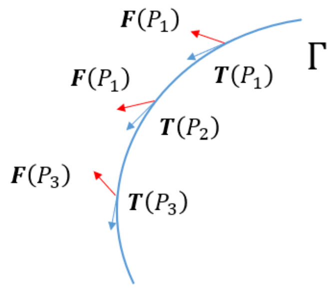

Figure 2.4: The unit tangent vector $\boldsymbol{T}$ (in blue) and the vector $\boldsymbol{F}$ (in red) at three typical points P_1, P_2 and P_3 on the curve Γ. $\boldsymbol{T}(P_i)$ should be tangent to Γ at P_i while $\boldsymbol{F}(P_i)$ does not have to.

Physical Meaning. If $\boldsymbol{F}$ is the force field, then (2.23) means the work done by $\boldsymbol{F}$ on a particle moving along the curve Γ in the direction of $\boldsymbol{T}$. If the curve Γ is closed, then (2.23) is often written as

$$\oint_\Gamma \boldsymbol{F} \cdot \boldsymbol{T}\,\mathrm{d}s \tag{2.25}$$

and means the **circulation** of the vector field $\boldsymbol{F}$ along Γ in the direction of $\boldsymbol{T}$.

Computation With $\mathrm{d}\boldsymbol{s} = \boldsymbol{T}\,\mathrm{d}s$, we have

$$\int_\Gamma \boldsymbol{F} \cdot \boldsymbol{T}\,\mathrm{d}s = \int_\Gamma \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} \tag{2.26}$$

If $\boldsymbol{F} = F_x\hat{\boldsymbol{x}} + F_y\hat{\boldsymbol{y}} + F_z\hat{\boldsymbol{z}}$, then

$$\begin{aligned}\int_\Gamma \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} &= \int_\Gamma (F_x\hat{\boldsymbol{x}} + F_y\hat{\boldsymbol{y}} + F_z\hat{\boldsymbol{z}}) \cdot (\mathrm{d}x\,\hat{\boldsymbol{x}} + \mathrm{d}y\,\hat{\boldsymbol{y}} + \mathrm{d}z\,\hat{\boldsymbol{z}}) \\ &= \int_\Gamma (F_x\,\mathrm{d}x + F_y\,\mathrm{d}y + F_z\,\mathrm{d}z) \\ &\equiv \int_\Gamma F_x\,\mathrm{d}x + F_y\,\mathrm{d}y + F_z\,\mathrm{d}z. \end{aligned} \tag{2.27}$$

Here

$$\int_\Gamma F_x \, \mathrm{d}x = \lim_{\Delta s_i \to 0} \sum_i F_x(P_i)\Delta x_i \tag{2.28}$$

$$\int_\Gamma F_y \, \mathrm{d}y = \lim_{\Delta s_i \to 0} \sum_i F_y(P_i)\Delta y_i \tag{2.29}$$

$$\int_\Gamma F_z \, \mathrm{d}z = \lim_{\Delta s_i \to 0} \sum_i F_z(P_i)\Delta z_i \tag{2.30}$$

If $\boldsymbol{F} = F_r\hat{\boldsymbol{r}} + F_\phi\hat{\boldsymbol{\phi}} + F_\theta\hat{\boldsymbol{\theta}}$, then

$$\begin{aligned}\int_\Gamma \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{s} &= \int_\Gamma (F_r\hat{\boldsymbol{r}} + F_\phi\hat{\boldsymbol{\phi}} + F_\theta\hat{\boldsymbol{\theta}}) \cdot (\mathrm{d}r\,\hat{\boldsymbol{r}} + r\sin\theta \mathrm{d}\phi\,\hat{\boldsymbol{\phi}} + r\mathrm{d}\theta\,\hat{\boldsymbol{\theta}}) \\ &= \int_\Gamma F_r\,\mathrm{d}r + F_\phi r\sin\theta\,\mathrm{d}\phi + F_\theta r\,\mathrm{d}\theta \end{aligned} \tag{2.31}$$

If $\boldsymbol{F} = F_\rho\hat{\boldsymbol{\rho}} + F_\phi\hat{\boldsymbol{\phi}} + F_z\hat{\boldsymbol{z}}$, then

$$\begin{aligned}\int_\Gamma \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{s} &= \int_\Gamma (F_\rho\hat{\boldsymbol{\rho}} + F_\phi\hat{\boldsymbol{\phi}} + F_z\hat{\boldsymbol{z}}) \cdot (\mathrm{d}\rho\,\hat{\boldsymbol{\rho}} + \rho\mathrm{d}\phi\,\hat{\boldsymbol{\phi}} + \mathrm{d}z\,\hat{\boldsymbol{z}}) \\ &= \int_\Gamma F_\rho\,\mathrm{d}\rho + F_\phi\rho\,\mathrm{d}\phi + F_\theta\,\mathrm{d}z. \end{aligned} \tag{2.32}$$

Example 2.7. Γ *is a straight line segment between the point* P_1 *with* $(x = 0, y = 0, z = 0)$ *and the point* P_2 *with* $(x = 1, y = 0, z = 0)$*.* $\boldsymbol{F}$ *is a vector field defined on* Γ *such that*

$$\boldsymbol{F}(P) = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} \tag{2.33}$$

for any $P \in \Gamma$ *and has Cartesian coordinates* (x, y, z)*.*

If

$$\boldsymbol{T}(P) = \hat{\boldsymbol{x}} \quad \textit{for any } P \in \Gamma, \tag{2.34}$$

then the line integral of $\boldsymbol{F}$ *along* Γ *in the direction of* $\boldsymbol{T}$*, i.e. from* P_1 *to* P_2 *is*

$$\int_\Gamma \boldsymbol{F}\cdot\boldsymbol{T}\,\mathrm{d}s = \int_\Gamma (x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}})\cdot\hat{\boldsymbol{x}}\,\mathrm{d}s = \int_\Gamma x\mathrm{d}s = \int_0^1 x\mathrm{d}x = 1/2 \tag{2.35}$$

If

$$\boldsymbol{T}(P) = -\hat{\boldsymbol{x}} \quad \textit{for any } P \in \Gamma, \tag{2.36}$$

then the line integral of $\boldsymbol{F}$ along Γ in the direction of $\boldsymbol{T}$, i.e. from P_2 to P_1 is

$$\int_\Gamma \boldsymbol{F}\cdot\boldsymbol{T}\,\mathrm{d}s = \int_\Gamma (x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}})\cdot(-\hat{\boldsymbol{x}})\,\mathrm{d}s = -\int_\Gamma x\mathrm{d}s = -\int_0^1 x\mathrm{d}x = -1/2 \quad (2.37)$$

Example 2.8. *Let $\boldsymbol{F}$ be a force field such that*

$$\boldsymbol{F}(P) = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + y\hat{\boldsymbol{z}} \qquad \textit{for any } P \quad (2.38)$$

In this field, a particle moves in a straight line Γ from the point with Cartesian coordinates $(x=-1, y=-2, z=0)$ to the point with Cartesian coordinates $(x=3, y=1, z=1)$.

The work done by $\boldsymbol{F}$ on this particle is (2.23) where $\boldsymbol{T}(P)$ is a unit vector and in direction of $\overrightarrow{P_1P_2}$.

We use (2.27) to evaluate this line integral. In this case,

$$F_x = x,\ \ F_y = y,\ \ F_z = y \quad (2.39)$$

and

$$\int_\Gamma F_x\,\mathrm{d}x = \int_\Gamma x\,\mathrm{d}x = \int_{-1}^3 x\,\mathrm{d}x = \frac{3^2-(-1)^2}{2} = 8/2 \quad (2.40)$$

$$\int_\Gamma F_y\,\mathrm{d}y = \int_\Gamma y\,\mathrm{d}y = \int_{-2}^1 y\,\mathrm{d}y = \frac{1^2-(-2)^2}{2} = -3/2 \quad (2.41)$$

For any $P\in\Gamma$, its y and z coordinates are related as

$$y = 3z-2 \quad (2.42)$$

So

$$\int_\Gamma F_z\,\mathrm{d}z = \int_\Gamma y\,\mathrm{d}z = \int_0^1 (3z-2)\,\mathrm{d}z = \frac{3}{2} - 2 = -1/2 \quad (2.43)$$

Overall, the line integral is

$$\begin{aligned}\int_\Gamma \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s} &= \int_\Gamma F_x\,\mathrm{d}x + F_y\,\mathrm{d}y + F_z\,\mathrm{d}z \\ &= \int_\Gamma F_x\,\mathrm{d}x + \int_\Gamma F_y\,\mathrm{d}y + \int_\Gamma F_z\,\mathrm{d}z \\ &= 8/2 - 3/2 - 1/2 = 2 \end{aligned} \quad (2.44)$$

Example 2.9. *The electrical field* $\boldsymbol{E}$ *in Example* (1.3) *exerts a force*

$$\boldsymbol{F} = q^* \boldsymbol{E} \tag{2.45}$$

on another charge q^* *placed within this field. The force* $\boldsymbol{F}$ *does work*

$$\int_\Gamma \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s \tag{2.46}$$

on the charge q^* *that moves along a curve* Γ *in the direction of* $\boldsymbol{T}$.

Assume Γ *is a circle around the origin and the charge* q^* *moves in the circle once. In this case,*

$$\boldsymbol{E}(P) \cdot \boldsymbol{T}(P) = 0 \tag{2.47}$$

for any $P \in \Gamma$. *Then the work done by the force* $\boldsymbol{F}$ *is*

$$\oint_\Gamma \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s = \oint_\Gamma (q^* \boldsymbol{E}) \cdot \boldsymbol{T} \, \mathrm{d}s = q^* \oint_\Gamma \boldsymbol{E} \cdot \boldsymbol{T} \, \mathrm{d}s = 0 \tag{2.48}$$

In other words, the force does not do any work.

Assume Γ *is a curve between the point* P_1 *with spherical coordinates* $(r = 1, \phi = \pi/2, \theta = \pi/3)$ *and the point with spherical coordinates* $(r = 3, \phi = \pi/3, \theta = \pi/2)$. *In this case, the work done by the force is*

$$\begin{aligned}
\int_\Gamma \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s &= q^* \int_\Gamma \boldsymbol{E} \cdot \boldsymbol{T} \, \mathrm{d}s = q^* \int_\Gamma \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} \\
&= q^* \int_\Gamma \frac{q}{4\pi\epsilon_0} \frac{\hat{\boldsymbol{r}}}{r^2} \cdot (\mathrm{d}r \, \hat{\boldsymbol{r}} + r \sin\theta \, \mathrm{d}\phi \, \hat{\boldsymbol{\phi}} + r \mathrm{d}\theta \, \hat{\boldsymbol{\theta}}) \\
&= \frac{q^* q}{4\pi\epsilon_0} \int_\Gamma \frac{1}{r^2} \, \mathrm{d}r = \frac{q^* q}{4\pi\epsilon_0} \int_1^3 \frac{1}{r^2} \, \mathrm{d}r = \frac{q^* q}{4\pi\epsilon_0} \cdot \frac{2}{3} = \frac{q^* q}{6\pi\epsilon_0}
\end{aligned} \tag{2.49}$$

Note that the value of line integral does NOT depend on the part of the curve Γ *between its two end points* P_1 *and* P_2. *Actually, this is true for any two points (except for the one at the origin, verify this). Because of this, the line integral in this case is called* **path-independent**.

2.4 Surface Integral of Vector Field

Definition Assume $\boldsymbol{F}$ is a vector field defined on a surface S. The $\boldsymbol{n}$ is defined on S such that $\boldsymbol{n}(P)$ is a unit vector and normal (i.e. perpendicular) to S at P for any $P \in S$. The surface integral of $\boldsymbol{F}$ through the surface S in the direction of $\boldsymbol{n}$ is

$$\int_S \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{2.50}$$

i.e. the surface integral of f over S with

$$f(P) \stackrel{\text{def}}{=} \boldsymbol{F}(P) \cdot \boldsymbol{n}(P) \qquad \text{for any } P \in \Gamma \tag{2.51}$$

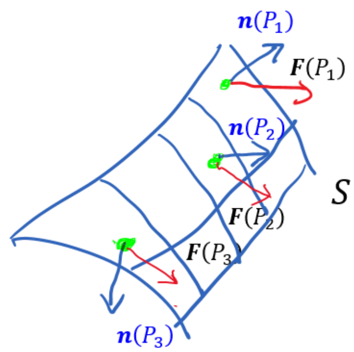

Figure 2.5: The unit normal vector $\boldsymbol{n}$ (in blue) and the vector $\boldsymbol{F}$ (in red) at three typical points P_1, P_2 and P_3 on the surface Γ. $\boldsymbol{n}(P_i)$ should be normal to Γ at P_i while $\boldsymbol{F}(P_i)$ does not have to.

Physical Meaning. The (2.50) means the **flux** of vector field $\boldsymbol{F}$ through the surface S in the direction of $\boldsymbol{n}$. When S is a closed surface, the (2.50) is often written as

$$\oint_S \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S. \tag{2.52}$$

Unless specified otherwise, $\boldsymbol{n}$ is often also assumed to point outward, away from the interior of the surface. In this case, (2.52) means the **net flux** of $\boldsymbol{F}$ out of the closed surface S.

Computation The surface integral (2.50) could be evaluated exactly only in some special cases, and evaluated approximately most of the time based on its definition.

Example 2.10. *Assume S is a surface on the xy-plane. $\boldsymbol{n}$ is defined on S and*

$$\boldsymbol{n}(P) = \hat{\boldsymbol{z}} \tag{2.53}$$

for any $P \in S$. $\boldsymbol{F}$ is vector field defined on S and

$$\boldsymbol{F}(P) = y\hat{\boldsymbol{x}} + z\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}} \tag{2.54}$$

for any $P \in S$ with Cartesian coordinates (x, y, z). The surface integral of $\boldsymbol{F}$ through S in the direction of $\boldsymbol{n}$ is

$$\int_S \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = \int_S (y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}}) \cdot \hat{\boldsymbol{z}} \, \mathrm{d}S = \int_S x \ \mathrm{d}S$$

In other words, the surface integral in this case is actually a double integral of function $f(x, y) = x$ over the region S on the xy-plane.

Example 2.11. *Let S be a sphere of radius R and centered at the origin in the electric field $\boldsymbol{E}$ in Example (1.3). The net flux of $\boldsymbol{E}$ out of S is*

$$\oint_S \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{2.55}$$

with

$$\boldsymbol{n}(P) = \frac{\overrightarrow{OP}}{|\overrightarrow{OP}|} \qquad \textit{for any } P \in S \tag{2.56}$$

Therefore,

$$\begin{aligned}
\oint_S \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S &= \oint_S \frac{q}{4\pi\epsilon_0} \frac{\overrightarrow{OP}}{|\overrightarrow{OP}|^3} \cdot \frac{\overrightarrow{OP}}{|\overrightarrow{OP}|} \, \mathrm{d}S \\
&= \oint_S \frac{q}{4\pi\epsilon_0 |\overrightarrow{OP}|^2} \, \mathrm{d}S = \oint_S \frac{q}{4\pi\epsilon_0 R^2} \, \mathrm{d}S \\
&= \frac{q}{4\pi\epsilon_0 R^2} \oint_S \mathrm{d}S = \frac{q}{4\pi\epsilon_0 R^2} \cdot 4\pi R^2 = q/\epsilon_0
\end{aligned} \tag{2.57}$$

Example 2.12. *Let S be the surface consisting of points whose Cartesian coordinates satisfy*

$$x = 1, -2 \leq y \leq 2, -2 \leq z \leq 2. \tag{2.58}$$

The flux of electric field $\boldsymbol{E}$ *through* S *in the positive direction of* x*-axis is*

$$\int_S \boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}S = \int_S \boldsymbol{E}\cdot\hat{\boldsymbol{x}}\,\mathrm{d}S = \int_S \frac{q}{4\pi\epsilon_0}\frac{\overrightarrow{OP}}{|\overrightarrow{OP}|^3}\cdot\hat{\boldsymbol{x}}\,\mathrm{d}S \tag{2.59}$$

Because

$$\overrightarrow{OP}\cdot\hat{\boldsymbol{x}} > 0 \tag{2.60}$$

for any $P \in S$*,*

$$\int_S \boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}S = \int_S \frac{q}{4\pi\epsilon_0}\frac{\overrightarrow{OP}}{|\overrightarrow{OP}|^3}\cdot\hat{\boldsymbol{x}}\,\mathrm{d}S > 0 \tag{2.61}$$

when $q > 0$*.*

Exercises

2.1. *(The* $\boldsymbol{T}$ *) Assume a particle moves along a curve* Γ *in a force field* $\boldsymbol{F}$ *as shown in Figure. If the line integral*

$$\int_\Gamma \boldsymbol{F}\cdot\boldsymbol{T}\,\mathrm{d}s \tag{2.62}$$

represents the work done by $\boldsymbol{F}$ *on this particle, then what is* $\boldsymbol{T}$ *and indicates its direction in the Figure.*

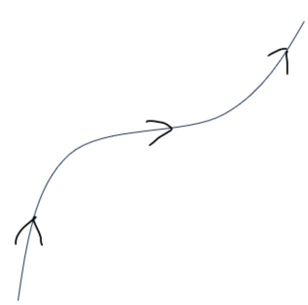

Figure 2.6: A directed curve. The arrow on the curve indicates the direction of the movement of a particle moving along the curve.

2.2. *Assume* Γ *is a circle on the plane* $z = 1$*. It is of radius 3 and centered at the point with Cartesian coordinates* $(x = 0, y = 0, z = 1)$*. What is the line integral of the magnetic field* $\boldsymbol{B}$ *in Example* (1.4) *along* Γ *in the direction* $\boldsymbol{T}$ *as indicated in Figure ?*

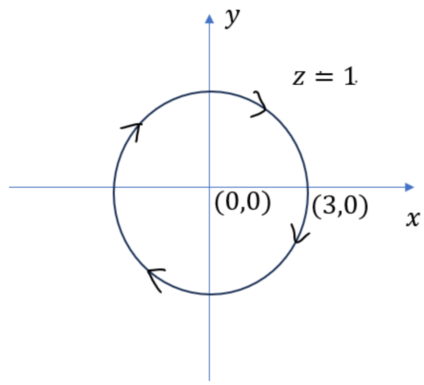

Figure 2.7: A circle on the plane of equation $z = 1$. The circle is centered at $(x = 0, y = 0, z = 1)$ and has radius 3. The arrows on the circle are clockwise

2.3. *Assume* Γ *is a straight line segment between the point with Cartesian coordinates* $(-2, 1, 1)$ *to the point with Cartesian coordinates* $(2, 1, 1)$*.* $\boldsymbol{B}$ *is the magnetic field in Example* (1.4)*.* $\boldsymbol{T}$ *is defined on* Γ *such that*

$$\boldsymbol{T}(P) = \hat{\boldsymbol{x}} \quad \textit{for any } P \in \Gamma \tag{2.63}$$

Consider the line integral of $\boldsymbol{B}$ *along* Γ *and in the direction of* $\boldsymbol{T}$*. Is the value of this line integral positive or negative ? Why ?*

2.4. *Assume point* P_1 *has Cartesian coordinates* $(x = 1, y = 0, z = 0)$*,* P_2 *has Cartesian coordinates* $(x = 0, y = 2, z = 0)$ *and* P_3 *has Cartesian coordinates* $(x = 0, y = 0, z = 3)$*. Let* $\boldsymbol{F}$ *be a force field such that*

$$\boldsymbol{F}(P) = y\hat{\boldsymbol{x}} + z\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}} \qquad \textit{for any } P \tag{2.64}$$

In this field, a particle moves along a straight line from P_1 *to* P_2*, then along a straight line from* P_2 *to the point* P_3*, finally along a straight line from* P_3 *to* P_1*. What is the total work done by* $\boldsymbol{F}$ *on this particle ?*

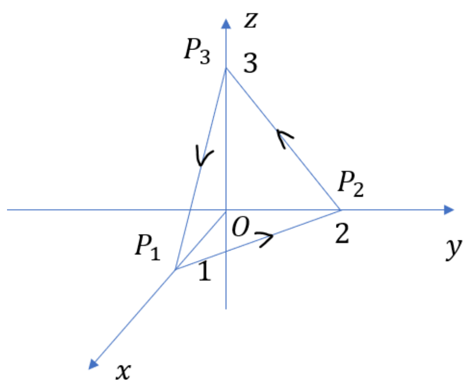

Figure 2.8: A closed and directed path Γ that consists of straight line segment from $P_1(1, 0, 0)$ to $P_2(0, 2, 0)$, line segment from P_2 to $P_3(0, 0, 3)$, and line segment from P_3 to P_1.

2.5. *In the electric field* $\boldsymbol{E}$ *in Example* (1.3)*, a point charge* q^* *moves along a straight line from the point with Cartesian coordinates* $(x = 1, y = 1, z = 1)$ *to another point with Cartesian coordinates* $(x = 2, y = 3, z = 4$*. Find the work done by the electric force exerted by* $\boldsymbol{E}$ *on the charge* q^**.*

2.6. *Let* S *be a surface on the plane of equation* $y = 2$ *and* $\boldsymbol{J}$ *is current density.*

(i) Assume

$$\boldsymbol{J}(P) = \hat{\boldsymbol{y}} \qquad \text{for any } P \in S. \tag{2.65}$$

If the surface integral

$$\int_S \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{2.66}$$

represents the current flowing through S *from the side* $y < 2$ *to the side* $y > 2$*, what is the* $\boldsymbol{n}$ *in* (2.66) *? If* (2.66) *represents the current flowing through* S *from the side* $y > 2$ *to the side* $y < 2$*, what is the* $\boldsymbol{n}$ *in* (2.66) *?*

(ii) Assume

$$\boldsymbol{J}(P) = \hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} \qquad \text{for any } P \in S. \tag{2.67}$$

What is the $\boldsymbol{n}$ *in* (2.66) *if* (2.66) *represents the current flowing through* S *from the side* $y < 2$ *to the side* $y > 2$ *?*

2.7. *Assume electric current with current density* $\boldsymbol{J}$ *flows through a surface* S *as shown in Figure 2.9. Indicate in the Figure the direction of* $\boldsymbol{n}$ *at several points on* S *if*

$$\int_S \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{2.68}$$

represent the current flowing through S *from the side Left to the side Right as shown in the Figure.*

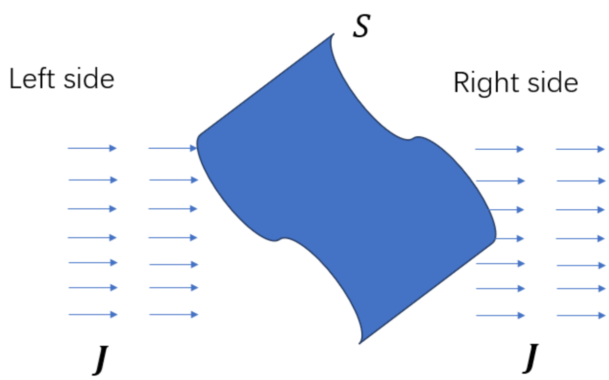

Figure 2.9: Current with density $\boldsymbol{J}$ flows through a surface S.

2.8. *Let* $\boldsymbol{E}$ *be a vector field such that*

$$\boldsymbol{E}(P) = y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}} \tag{2.69}$$

for any P *with Cartesian coordinates* (x, y, z)*. Find the flux of* $\boldsymbol{E}$ *into the tetrahedron* $\overline{OP_1P_2P_3}$ *in Figure 2.8, through each of its four faces, respectively.*

2.9. *Assume stationary charge with charge density* σ *is distributed over an infinite flat plane* $z = 0$*. According to electrostatics, it results in an electric field* $\boldsymbol{E}$ *that is*

$$\boldsymbol{E}(P) = \begin{cases} \sigma\hat{\boldsymbol{z}}/(2\epsilon_0) & z > 0 \\ -\sigma\hat{\boldsymbol{z}}/(2\epsilon_0) & z < 0 \end{cases} \tag{2.70}$$

for any P *with Cartesian coordinates* (x, y, z)*.*

(i) Find a surface and a direction so that the flux of $\boldsymbol{E}$ *through this surface and in some direction is* 10.

(ii) Let S *be the half sphere that consists of all the points whose Cartesian coordinates satisfy*

$$x^2 + y^2 + (z-1)^2 = 1, \quad z \geq 1. \tag{2.71}$$

Estimate the flux of $\boldsymbol{E}$ *through* S *upward (i.e. from the side that contains the point* $(x = 0, y = 0, z = 1)$ *to the side that contain* $(x = 0, y = 0, z = 5)$*).*

Chapter 3

Derivatives of Fields and Fundamental Theorems

3.1 Gradient and Gradient Theorem

Definition The gradient of a scalar field u is a vector field $\operatorname{grad} u$. Its vector at point P, denoted by $\operatorname{grad} u(P)$, satisfies

$$\operatorname{grad} u(P) \cdot \boldsymbol{T} = \text{derivative of } u \text{ at } P \text{ and in the direction of } \boldsymbol{T} \tag{3.1}$$

for any unit vector $\boldsymbol{T}$. Alternatively, $\operatorname{grad} u(P)$ is a vector such that

$$\mathrm{d}u \equiv u(P + \mathrm{d}\boldsymbol{s}) - u(P) = \operatorname{grad} u(P) \cdot \mathrm{d}\boldsymbol{s} \tag{3.2}$$

for any infinitesimal displacement $\mathrm{d}\boldsymbol{s}$. Here $u(P + \mathrm{d}\boldsymbol{s})$ is the value of u at the point whose position vector is the vector sum of $\overrightarrow{OP}$ and $\mathrm{d}\boldsymbol{s}$.

Informally, (3.1) or (3.2) says that

$$u(P) - u(P_0) \approx \operatorname{grad} u(P_0) \cdot \overrightarrow{P_0P} \qquad P \text{ is close to } P_0 \tag{3.3}$$

Geometric Meaning $\operatorname{grad} u(P)$ points in the direction of the greatest rate of increase of u at P. The greatest rate of increase is the magnitude of $\operatorname{grad} u(P)$.

Computation Let $\boldsymbol{T} = \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ in (3.1), respectively. Then

$$\operatorname{grad} u(P)\cdot\hat{\boldsymbol{x}} = \left.\frac{\partial u}{\partial x}\right|_P, \quad \operatorname{grad} u(P)\cdot\hat{\boldsymbol{y}} = \left.\frac{\partial u}{\partial y}\right|_P, \quad \operatorname{grad} u(P)\cdot\hat{\boldsymbol{z}} = \left.\frac{\partial u}{\partial z}\right|_P, \tag{3.4}$$

which implies

$$\text{grad}u(P) = \left.\frac{\partial u}{\partial x}\right|_P \hat{\boldsymbol{x}} + \left.\frac{\partial u}{\partial y}\right|_P \hat{\boldsymbol{y}} + \left.\frac{\partial u}{\partial z}\right|_P \hat{\boldsymbol{z}}. \tag{3.5}$$

Gradient Theorem The line integral of $\text{grad}\, u$ along a directed curve from point P_1 to P_2 is $u(P_2) - u(P_1)$, i.e.

$$\int_{P_1}^{P_2} \text{grad}\, u \cdot \text{d}\boldsymbol{s} = u(P_2) - u(P_1). \tag{3.6}$$

Note that (3.6) holds for any directed curve from P_1 to P_2. The line integral of gradient is path-**independent**.

Informal justification.

$$\begin{aligned}\int_{P_{begin}}^{P_{end}} \text{grad}\, u \cdot \text{d}\boldsymbol{s} \approx \sum_i \text{grad}\, u(P_i) \cdot \Delta \boldsymbol{s}_i \approx \sum_i u(P_{i+1}) - u(P_i) \\ = u(P_{end}) - u(P_{begin})\end{aligned} \tag{3.7}$$

Example 3.1. *Find the gradient of the scalar field in* (1.10) *in Example* (1.2). *How is its gradient related to the electric field in* (1.13) *in Example* (1.3) *? Finally, use its gradient to determine the direction of the largest rate of decrease of potential at the point* $(x = 1, y = 2, z = 3)$.

Solution. *In terms of Cartesian coordinates* x, y *and* z, *the* u *in* (1.10) *is*

$$u(x, y, z) = \frac{q}{4\pi\epsilon_0} \frac{1}{\sqrt{x^2 + y^2 + z^2}}. \tag{3.8}$$

So

$$\frac{\partial u}{\partial x} = \frac{q}{4\pi\epsilon_0} \cdot \left(-\frac{1}{2}\right) \cdot \frac{2x}{\sqrt{x^2 + y^2 + z^2}^3} = -\frac{q}{4\pi\epsilon_0} \frac{x}{\sqrt{x^2 + y^2 + z^2}^3} \tag{3.9}$$

and similarly

$$\frac{\partial u}{\partial y} = -\frac{q}{4\pi\epsilon_0} \frac{y}{\sqrt{x^2 + y^2 + z^2}^3} \tag{3.10}$$

$$\frac{\partial u}{\partial z} = -\frac{q}{4\pi\epsilon_0} \frac{z}{\sqrt{x^2 + y^2 + z^2}^3} \tag{3.11}$$

According to (3.5),

$$\textit{grad}\, u(P) = -\frac{q}{4\pi\epsilon_0}\frac{x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}}{\sqrt{x^2+y^2+z^2}^3} \tag{3.12}$$

provided that P *has Cartesian coordinates* x, y *and* z*. In other words,*

$$\textit{grad}\, u(P) = -\frac{q}{4\pi\epsilon_0}\frac{\overrightarrow{OP}}{|OP|^3}. \tag{3.13}$$

Comparing (3.13) *with the electric field* $\boldsymbol{E}$ *in* (1.13)*, we have*

$$-\textit{grad}\, u = \boldsymbol{E} \tag{3.14}$$

According to (3.13)*, the gradient of* u *at the point* $(x = 1, y = 2, z = 3)$ *points in the opposite direction of the vector*

$$\hat{\boldsymbol{x}} + 2\hat{\boldsymbol{y}} + 3\hat{\boldsymbol{z}}. \tag{3.15}$$

Therefore, the vector (3.15) *points in the direction of largest rate of* **decrease** *of* u *at this point.*

Example 3.2. *In electrostatics, the electric potential* u *is defined in terms of electric field* $\boldsymbol{E}$ *as follows*

$$u(P) \stackrel{\text{def}}{=} -\int_{P_0}^{P} \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{s} \qquad \textit{for any point } P. \tag{3.16}$$

That is to say, the potential u *at a point* P *is minus the line integral of* $\boldsymbol{E}$ *along a directed curve from a fixed point* P_0 *to* P*. We can choose any curve from* P_0 *to* P *in evaluating the line integral. This is because the line integral of this field is path-independent as shown in Example* (2.9)*. The fixed point* P_0 *is called the reference point and* $u(P_0) = 0$ *(Why ?). From* (3.16)*, we have*

$$\boldsymbol{E} = -\textit{grad}\, u \tag{3.17}$$

i.e. the electric field is the gradient of $-u$*.*

Proof. According to the Gradient Theorem,

$$u(P) = u(P) - u(P_0) = \int_{P_0}^{P} \mathrm{grad}\, u \cdot \mathrm{d}\boldsymbol{s} \tag{3.18}$$

Comparing (3.16) with (3.18), we have

$$\int_{P_0}^{P} (\mathrm{grad}\, u + \boldsymbol{E}) \cdot \mathrm{d}\boldsymbol{s} = 0 \tag{3.19}$$

Since (3.19) holds for any curve, we must have

$$\mathrm{grad}\, u + \boldsymbol{E} = 0 \quad \text{or} \quad \boldsymbol{E} = -\mathrm{grad}\, u \tag{3.20}$$

everywhere □

Example 3.3. *Assume*

$$\boldsymbol{F}(P) = yz\hat{\boldsymbol{x}} + zx\hat{\boldsymbol{y}} + xy\hat{\boldsymbol{z}} \tag{3.21}$$

for any point P with Cartesian coordinates x, y, z. Find the line integral of $\boldsymbol{F}$ along a directed curve $\Gamma : A \to B \to C \to D \to E$ as shown in Figure 3.1. The points A, B, C, D, E are all the vertices of a cube and their Cartesian coordinates are

$$A = (x = \ 1, y = -1, z = 1) \tag{3.22}$$
$$B = (x = 1, y = 1, z = 1) \tag{3.23}$$
$$C = (x = -1, y = 1, z = 1) \tag{3.24}$$
$$D = (x = -1, y = 1, z = -1) \tag{3.25}$$
$$E = (x = -1, y = -1, z = -1) \tag{3.26}$$

Solution. *Because*

$$\boldsymbol{F} = \mathit{grad}\, u \tag{3.27}$$

with

$$u(P) = xyz \text{ for any } P \text{ with Cartesian coordinates } x, y, z \tag{3.28}$$

Therefore, the line integral of $\boldsymbol{F}$ is path-independent and

$$\begin{aligned}\int_{\Gamma} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} &= \int_{A}^{E} \mathit{grad}\, u \cdot \mathrm{d}\boldsymbol{s} = u(E) - u(A) \\ &= (-1) \cdot (-1) \cdot (-1) - 1 \cdot (-1) \cdot 1 = 0 \end{aligned} \tag{3.29}$$

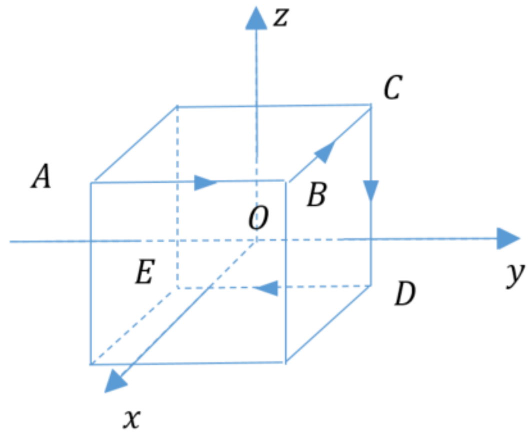

Figure 3.1: A path from point A to E. It consists of straight line segment AB, followed by BC, CD and DE. The A, B, C, D, E are the vertices of a cube and have Cartesian coordinates $(x = 1, y = -1, z = 1), (x = 1, y = 1, z = 1), (x = -1, y = 1, z = 1), (x = -1, y = 1, z = -1)$ and $(x = -1, y = -1, z = -1)$, respectively.

3.2 Divergence and Divergence Theorem

Definition The divergence of a vector field $\boldsymbol{F}$ is a scalar field $\operatorname{div} \boldsymbol{F}$. Its scalar at point P is

$$\operatorname{div} \boldsymbol{F}(P) \stackrel{\text{def}}{=} \lim_{\Omega \to P} \frac{1}{|\Omega|} \oint_{\partial\Omega} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{3.30}$$

Here Ω is a small region that contains P, has volume $|\Omega|$ and boundary surface $\partial\Omega$. $\boldsymbol{n}$ is the outward unit normal vector to $\partial\Omega$. (i.e. For any $P \in \partial\Omega$, $\boldsymbol{n}(P)$ is a unit vector perpendicular to $\partial\Omega$ at P and pointing away from the interior of Ω.) As Ω shrinks to P, both its volume $|\Omega|$ and the area of surface $\partial\Omega$ approaches zero. This makes it possible for the limit of the ratio in (3.30) to exist.

Informally,

$$\operatorname{div} \boldsymbol{F}(P) \approx \frac{1}{|\Omega_i|} \oint_{\partial\Omega_i} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{3.31}$$

where Ω_i is a small region around P.

Meaning $\operatorname{div} \boldsymbol{F}(P)$ is approximately the net flux of $\boldsymbol{F}$ out of a small region around P per unit volume. If $\operatorname{div} \boldsymbol{F}(P) < 0$, then its absolute value is approximatly the net flux of $\boldsymbol{F}$ into the region per unit area.

Computation Assume vector field $\boldsymbol{F}$ is

$$\boldsymbol{F}(P) = F_1(x,y,z)\hat{\boldsymbol{x}} + F_2(x,y,z)\hat{\boldsymbol{y}} + F_3(x,y,z)\hat{\boldsymbol{z}} \tag{3.32}$$

for any point P with Cartesian coordinates x, y, z. Here functions F_1, F_2 and F_3 all have continuous partial derivatives. Then the divergence of $\boldsymbol{F}$ at point P is

$$\operatorname{div} \boldsymbol{F}(P) = \left.\frac{\partial F_1}{\partial x}\right|_P + \left.\frac{\partial F_2}{\partial y}\right|_P + \left.\frac{\partial F_3}{\partial z}\right|_P \tag{3.33}$$

Example 3.4. *Find the divergence of the vector field $\boldsymbol{E}$ defined in* (1.3) *in Example* (1.3)*.*

Solution. *In terms of Cartesian coordinates,*

$$\boldsymbol{E}(x,y,z) = E_1\hat{\boldsymbol{x}} + E_2\hat{\boldsymbol{y}} + E_3\hat{\boldsymbol{z}} \tag{3.34}$$

with

$$E_1 = \frac{q}{4\pi\epsilon_0}\frac{x}{r^3}, \quad E_2 = \frac{q}{4\pi\epsilon_0}\frac{y}{r^3}, \quad E_3 = \frac{q}{4\pi\epsilon_0}\frac{z}{r^3} \tag{3.35}$$

and

$$r = \sqrt{x^2+y^2+z^2}. \tag{3.36}$$

Since

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r}, \tag{3.37}$$

$$\frac{\partial E_1}{\partial x} = \frac{q}{4\pi\epsilon_0}\left(\frac{r^3 - x\cdot 3r^2\cdot(x/r)}{r^6}\right) \tag{3.38}$$

$$\frac{\partial E_2}{\partial y} = \frac{q}{4\pi\epsilon_0}\left(\frac{r^3 - y\cdot 3r^2\cdot(y/r)}{r^6}\right) \tag{3.39}$$

$$\frac{\partial E_3}{\partial z} = \frac{q}{4\pi\epsilon_0}\left(\frac{r^3 - y\cdot 3r^2\cdot(y/r)}{r^6}\right) \tag{3.40}$$

Therefore,

$$div\, \boldsymbol{E} = \frac{q}{4\pi\epsilon_0}\left(\frac{3r^3 - 3r(x^2+y^2+z^2)}{r^6}\right) = 0 \tag{3.41}$$

That is to say, the divergence of the electric field (1.13) *is zero everywhere except for the origin.*

Example 3.5. *The divergence of the following four vector fields*

$$\boldsymbol{F}(x,y,z) = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} \qquad \textit{for any } x,y,z \tag{3.42}$$
$$\boldsymbol{G}(x,y,z) = 3x\hat{\boldsymbol{x}} + 3y\hat{\boldsymbol{y}} \qquad \textit{for any } x,y,z \tag{3.43}$$
$$\boldsymbol{H}(x,y,z) = -x\hat{\boldsymbol{x}} - y\hat{\boldsymbol{y}} \qquad \textit{for any } x,y,z \tag{3.44}$$
$$\boldsymbol{K}(x,y,z) = 8\hat{\boldsymbol{x}} + 8\hat{\boldsymbol{y}} \qquad \textit{for any } x,y,z \tag{3.45}$$

at the origin O $(x=0, y=0, z=0)$ is 2, 6, -2 and 0, respectively. Figure 3.2 shows these fields around the origin on the xy-plane. In (a) and (b), the vector fields indicate that something is spreading out from the center (i.e. the origin) of the figure. These patterns correspond to a positive value for the divergence $\boldsymbol{F}$ and $\boldsymbol{G}$ at the origin. In (b), the degree of spreading out is apparently larger than the degree of spreading out in (a). This corresponds to that the positive value of divergence of $\boldsymbol{G}$ is larger than the positive value of divergence of $\boldsymbol{F}$, at the origin. In (c), the vector field indicates that something is converging towards the origin. This is associated with a negative value of the divergence of vector field $\boldsymbol{H}$ there. Finally, (d) shows a constant vector field. Nothing is spreading out or converging. This is associated with a zero value of divergence of the vector field.

Divergence Theorem Assume $\boldsymbol{F}$ is continuously differentiable (i.e. its Cartesian components are all continuousely differentable) everywhere in a region Ω. Then

$$\int_\Omega \operatorname{div} \boldsymbol{F} \,\mathrm{d}\Omega = \oint_{\partial\Omega} \boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S \tag{3.46}$$

Here $\boldsymbol{n}$ is outward unit normal vector to the boundary surface $\partial\Omega$ of Ω

Informal Justification.

$$\int_\Omega \operatorname{div} \boldsymbol{F}\,\mathrm{d}\Omega \approx \sum_i \operatorname{div} \boldsymbol{F}(P_i)\Delta|\Omega_i| \approx \sum_i \oint_{\partial\Omega_i} \boldsymbol{F}\cdot\boldsymbol{n}_i\,\mathrm{d}S$$
$$= \oint_{\partial\Omega} \boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S \tag{3.47}$$

Example 3.6. *Gauss's law says:*

If $\boldsymbol{E}$ is the electric field due to some stationary charges, then the net flux of $\boldsymbol{E}$ through any closed surface S is proportional to the electric charge enclosed by S.

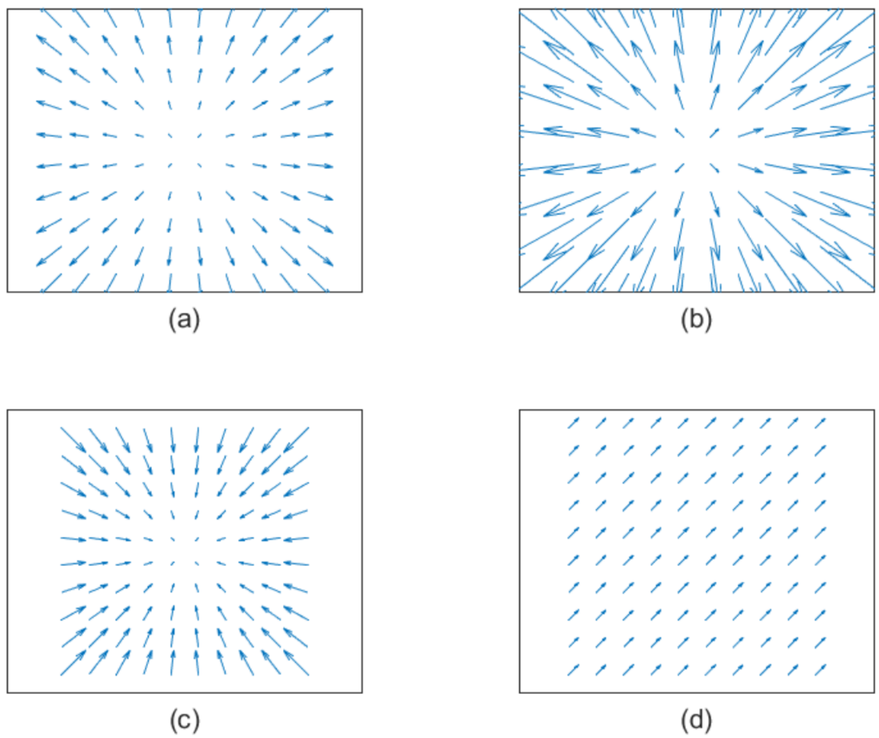

Figure 3.2: The four vector fields around the origin O that is at the center of each figure: $(a) \leftrightarrow \boldsymbol{F}$, $(b) \leftrightarrow \boldsymbol{G}$, $(c) \leftrightarrow \boldsymbol{H}$ and $(d) \leftrightarrow \boldsymbol{K}$

Justify this law.

Solution.
First, assume the electric field $\boldsymbol{E}$ *is due to a single fixed charge. We set up a Coordinate system with origin placed at the fixed charge as in Example* (1.3). *Within the surface* S*, we draw a sphere* S^* *that is centered at the origin as shown in Figure 3.3. According to* **Example 2.11**,

$$\int_{S^*} \boldsymbol{E} \cdot \boldsymbol{n}^* \, \mathrm{d}S = q/\epsilon_0. \tag{3.48}$$

Here $\boldsymbol{n}^*$ *is the outward unit normal vector to* S^*.

Consider the region Ω *between* S *and* S^*. *It has been shown in* **Example 3.4** *that the divergence of* $\boldsymbol{E}$ *at any point within* Ω *is 0. Therefore*

$$\oint_{\partial\Omega} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = \int_{\Omega} div \, \boldsymbol{F} \, \mathrm{d}\Omega = \int_{\Omega} 0 \, \mathrm{d}\Omega = 0 \tag{3.49}$$

Here $\boldsymbol{n}$ *is the outward unit normal vector to the boundary* $\partial\Omega$ *of* Ω.

Since Ω *is a region between* S *and the sphere* S^*, $\partial\Omega = S \cup S^*$ *and*

$$\oint_{\partial\Omega} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = \int_{S^*} \boldsymbol{E} \cdot (-\boldsymbol{n}^*) \, \mathrm{d}S + \int_S \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{3.50}$$

Combing (3.48), (3.49) *and* (3.50), *we have*

$$\int_S \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S = \int_{S^*} \boldsymbol{E} \cdot \boldsymbol{n}^* \, \mathrm{d}S = q/\epsilon_0. \tag{3.51}$$

Second, assume $\boldsymbol{E}$ *is due to several fixed point charges* $q_1, q_2, \ldots$. *According to the principle of superposition,*

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots \tag{3.52}$$

where $\boldsymbol{E}_i$ *is the electric field due to the fixed point charge* $q_i, i = 1, 2, \ldots$. *Therefore,*

$$\begin{aligned}
\oint_S \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S &= \oint_S (\boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots) \cdot \boldsymbol{n} \, \mathrm{d}S \\
&= \oint_S \boldsymbol{E}_1 \cdot \boldsymbol{n} \, \mathrm{d}S + \oint_S \boldsymbol{E}_2 \cdot \boldsymbol{n} \, \mathrm{d}S + \cdots \\
&= q_1/\epsilon_0 + q_2/\epsilon_0 + \cdots \\
&= (q_1 + q_2 + \cdots)/\epsilon_0 \qquad (3.53)
\end{aligned}$$

Therefore, the net flux of $\boldsymbol{E}$ *is proportional to the total charge enclosed.*

Example 3.7. *(Gauss's Law in differential equations) Consider the electric field* $\boldsymbol{E}$ *due to the electric charge that is continuously distributed with charge density* ρ *over a region* Ω'. *According to Gauss's law, the net flux of* $\boldsymbol{E}$ *out of a closed surface* $\partial\Omega$ *within* Ω' *is*

$$\oint_{\partial\Omega} \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S = \frac{1}{\epsilon_0} \int_\Omega \rho \, \mathrm{d}\Omega \tag{3.54}$$

Here Ω *is a region bounded by* $\partial\Omega$. *Dividing both sides of* (3.54) *by the volume* $|\Omega|$ *of* Ω, *we obtain*

$$\frac{1}{|\Omega|} \oint_{\partial\Omega} \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S = \frac{1}{\epsilon_0} \frac{1}{|\Omega|} \int_\Omega \rho \, \mathrm{d}\Omega \tag{3.55}$$

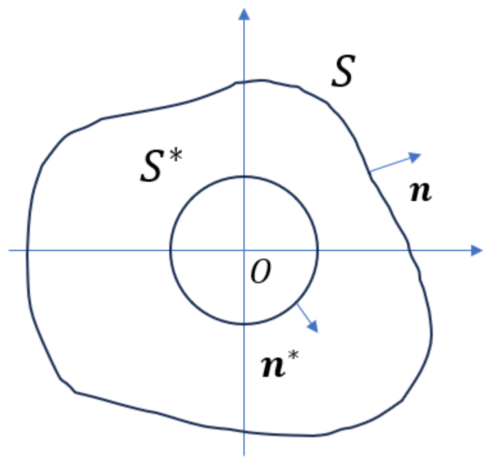

Figure 3.3: Two closed surfaces S and S^* enclosing the origin. S^* is a sphere centered at the origin; while S is the other surface that encloses S^* and has irregular shape.

As the region Ω shrinks to a point P, we have

$$\lim_{\Omega \to P} \frac{1}{|\Omega|} \oint_{\partial\Omega} \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}S = \frac{1}{\epsilon_0} \lim_{\Omega \to P} \frac{1}{|\Omega|} \int_{\Omega} \rho \, \mathrm{d}\Omega \tag{3.56}$$

or equivalently

$$div \, \boldsymbol{E}(P) = \rho(P)/\epsilon_0 \tag{3.57}$$

That is to say, the divergence of electric field at a point P is proportional to the charge density there.

3.3 Curl and Curl Theorem

Definition The curl of a vector field $\boldsymbol{F}$ is another vector field $\operatorname{curl} \boldsymbol{F}$. Its vector at point P, i.e. $\operatorname{curl} \boldsymbol{F}(P)$ satisfies

$$\operatorname{curl} \boldsymbol{F}(P) \cdot \boldsymbol{n} = \lim_{S \text{ shrinks to} P} \frac{1}{|S|} \oint_{\partial S} \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s \tag{3.58}$$

for any unit vector $\boldsymbol{n}$. Here S is small surface that contains P, has area $|S|$ and boundary curve ∂S. $\boldsymbol{T}$ is the unit tangent vector to ∂S and is specified so that $\boldsymbol{n}$ and $\boldsymbol{T}$ follow the right hand rule as shown in Figure 3.4.

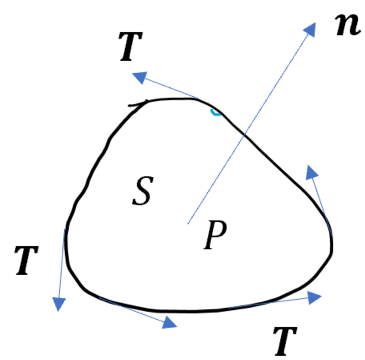

Figure 3.4: The unit tangent vector $\boldsymbol{T}$ at various places on the boundary curve of the surface S. These vectors $\boldsymbol{T}$ and the unit normal vector n to S follow the right-hand rule.

Informally,

$$\operatorname{curl} \boldsymbol{F}(P) \cdot \boldsymbol{n}(P) \approx \frac{1}{|S|} \oint_{\partial S} \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s \tag{3.59}$$

for any unit vector $\boldsymbol{n}(P)$. Here S is small surface containing P such that $\boldsymbol{n}(P)$ is perpendicular to S at P.

Meaning The curl of a vector field at a point is such a vector, that its project in a direction is the circulation per unit area of this vector field at this point and in this direction.

Comptuation Assume vector field $\boldsymbol{F}$ is

$$\boldsymbol{F}(P) = F_1(x, y, z)\hat{\boldsymbol{x}} + F_2(x, y, z)\hat{\boldsymbol{y}} + F_3(x, y, z)\hat{\boldsymbol{z}} \tag{3.60}$$

for any point P with Cartesian coordinates x, y, z. Here function F_1, F_2 and F_3 all have continuous partial derivatives. Then the curl of $\boldsymbol{F}$ at point P is

$$\operatorname{curl} \boldsymbol{F}(P) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\boldsymbol{x}} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\hat{\boldsymbol{y}} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{\boldsymbol{z}}. \tag{3.61}$$

Example 3.8. *Find the curl of the vecor field* $\boldsymbol{E}$ *in* (1.13) *in example* (1.3).

Solution. *Recall that*

$$\boldsymbol{E}(x, y, z) = E_1\hat{\boldsymbol{x}} + E_2\hat{\boldsymbol{y}} + E_3\hat{\boldsymbol{z}} \tag{3.62}$$

with

$$E_1 = \frac{q}{4\pi\epsilon_0}\frac{x}{r^3}, \quad E_2 = \frac{q}{4\pi\epsilon_0}\frac{y}{r^3}, \quad E_3 = \frac{q}{4\pi\epsilon_0}\frac{z}{r^3} \tag{3.63}$$

and

$$r = \sqrt{x^2 + y^2 + z^2}. \tag{3.64}$$

Because of (3.37)*,*

$$\frac{\partial E_1}{\partial y} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{x}{r^4}\cdot\frac{y}{r}, \qquad \frac{\partial E_1}{\partial z} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{x}{r^4}\cdot\frac{z}{r} \tag{3.65}$$

$$\frac{\partial E_2}{\partial x} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{y}{r^4}\cdot\frac{x}{r}, \qquad \frac{\partial E_2}{\partial z} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{y}{r^4}\cdot\frac{z}{r} \tag{3.66}$$

$$\frac{\partial E_3}{\partial x} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{z}{r^4}\cdot\frac{x}{r}, \qquad \frac{\partial E_3}{\partial y} = \frac{q}{4\pi\epsilon_0}(-3)\cdot\frac{z}{r^4}\cdot\frac{y}{r} \tag{3.67}$$

So

$$\begin{aligned} curl\,\boldsymbol{E}(x,y,z) &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\boldsymbol{x}} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\hat{\boldsymbol{y}} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{\boldsymbol{z}} \\ &= 0\hat{\boldsymbol{x}} + 0\hat{\boldsymbol{y}} + 0\hat{\boldsymbol{z}} = \boldsymbol{0} \end{aligned} \tag{3.68}$$

whenever $(x, y, z) \neq (0, 0, 0)$*.*

Example 3.9. *The curl of the following four vector fields*

$$\boldsymbol{F}(x,y,z) = -y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}} \qquad \textit{for any } x, y, z \tag{3.69}$$

$$\boldsymbol{G}(x,y,z) = -3y\hat{\boldsymbol{x}} + 3x\hat{\boldsymbol{y}} \qquad \textit{for any } x, y, z \tag{3.70}$$

$$\boldsymbol{H}(x,y,z) = y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}} \qquad \textit{for any } x, y, z \tag{3.71}$$

$$\boldsymbol{K}(x,y,z) = 8\hat{\boldsymbol{x}} + 8\hat{\boldsymbol{y}} \qquad \textit{for any } x, y, z \tag{3.72}$$

at the origin O $(x = 0, y = 0, z = 0)$ *is*

$$curl\,\boldsymbol{F}(O) = 2\hat{\boldsymbol{z}} \qquad curl\,\boldsymbol{F}(O) = 6\hat{\boldsymbol{z}} \tag{3.73}$$

$$curl\,\boldsymbol{H}(O) = -2\hat{\boldsymbol{z}} \qquad curl\,\boldsymbol{K}(O) = \boldsymbol{0} \tag{3.74}$$

Figure 3.5 shows these four fields around the origin that is at the center of the figure, and in the xy-plane. In (a), (b) and (c), vectors around the origin are rotating around the z-axis. This correspond to that the curl of these fields at the origin are all vectors in the direction of the z-axis. In (a) and (b), vectors are rotating anti-clockwisely and its rotation axis points in *the positive direction of z-axis (which is determined by the right-hand rule). This corresponds to a positive component of $\hat{\boldsymbol{z}}$ in the curl vector. The rotating vectors in (b) has larger magnitude than the ones in (a). This corresponds to that the postitive component of $\hat{\boldsymbol{z}}$ in the curl vector of (b) is larger. In (c), vectors are rotating clockwisely and its rotational axis points in the negative direction of z-axis. This corresponds to a negative component of $\hat{\boldsymbol{z}}$. In (d), vectors are not rotating at all in the xy-plane. This corresponds that the curl vector has no component of $\hat{\boldsymbol{z}}$.*

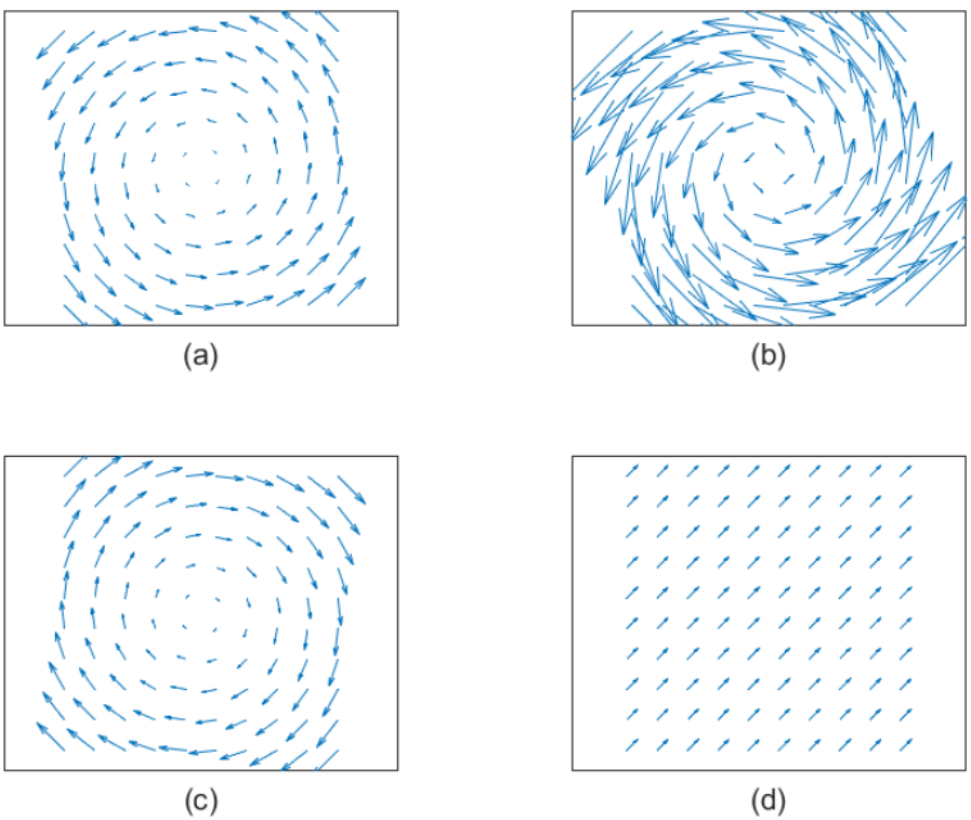

Figure 3.5: The vector field $\boldsymbol{F}$ in (a), $\boldsymbol{G}$ in (b), $\boldsymbol{H}$ in (c) and $\boldsymbol{K}$ in (d).

Example 3.10. *A vector field whose curl is a zero vector everywhere is called* **curl free**. *Any vector field that is a gradient of some scalar field must be curl free.*

Proof. Assume $\boldsymbol{F}$ is a gradient, i.e. $\boldsymbol{F} = \nabla u$ for some scalar field u. Then the line integral of $\boldsymbol{F}$ along any closed path Γ is

$$\oint_\Gamma \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} = \oint_\Gamma \nabla u \cdot \mathrm{d}\boldsymbol{s} = 0 \tag{3.75}$$

Since the line integral along any closed path is zero,

$$\operatorname{curl} \boldsymbol{F} = \boldsymbol{0} \qquad \text{everywhere,} \tag{3.76}$$

according to the definition of curl (3.58). □

Curl Theorem (or Stokes' Theorem) Assume $\boldsymbol{F}$ is continuously differentiable everywhere in a region that contains the surface S. Then

$$\int_S \operatorname{curl} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = \oint_{\partial S} \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s \tag{3.77}$$

Here ∂S is the boundary of S and is a closed curve. The unit normal vector $\boldsymbol{n}$ and unit tangent vector $\boldsymbol{T}$ follow the right hand rule as shown in Figure 3.4.

Informal Justification.

$$\begin{aligned}\int_S \operatorname{curl} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S \approx \sum_i \operatorname{curl} \boldsymbol{F}(P_i) \cdot \boldsymbol{n}_i(P) \Delta |S_i| \approx \sum_i \oint_{\partial S_i} \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s \\ = \oint_{\partial S} \boldsymbol{F} \cdot \boldsymbol{T} \, \mathrm{d}s\end{aligned} \tag{3.78}$$

Example 3.11. *Steady electric current with current density $\boldsymbol{J}$ produce a magnetic field $\boldsymbol{B}$. According to Ampere's law,*

$$\oint_{\partial S} \boldsymbol{B} \cdot \boldsymbol{T} \, \mathrm{d}s = \mu_0 \int_S \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{3.79}$$

for any surface S. Here μ_0 is some physical constants. The unit tangent vector $\boldsymbol{T}$ and the unit normal vector $\boldsymbol{n}$ follow the right hand rule as shown in Figure 3.4.

From (3.79), *we have*

$$\frac{1}{|S|} \oint_{\partial S} \boldsymbol{B} \cdot \boldsymbol{T} \, \mathrm{d}s = \frac{1}{|S|} \mu_0 \int_S \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}S \tag{3.80}$$

and

$$\lim_{S \text{ shrinks to } P} \frac{1}{|S|} \oint_{\partial S} \boldsymbol{B} \cdot \boldsymbol{T} \, \mathrm{d}s = \lim_{S \text{ shrinks to } P} \frac{1}{|S|} \mu_0 \int_S \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}S$$
$$= \mu_0 \boldsymbol{J}(P) \cdot \boldsymbol{n}(P) \tag{3.81}$$

Since (3.81) *holds for any unit vector* $\boldsymbol{n}(P)$*, we have*

$$curl\, \boldsymbol{B}(P) = \mu_0 \boldsymbol{J}(P) \tag{3.82}$$

after comparing (3.81) *and* (3.58)*. The* (3.82) *holds for any point* P*, so we simply write*

$$curl\, \boldsymbol{B} = \mu_0 \boldsymbol{J}. \tag{3.83}$$

This implies that the curl of magnetic field is proportional to the current density.

Example 3.12. *The surface integral of a vector field* $\boldsymbol{F}$ *is called surface independent if*

$$\int_{S_1} \boldsymbol{F} \cdot \boldsymbol{n}_1 \mathrm{d}S = \int_{S_2} \boldsymbol{F} \cdot \boldsymbol{n}_2 \mathrm{d}S. \tag{3.84}$$

Here S_1 *and* S_2 *are any two surfaces that share the same boundary curve, i.e.* $\partial S_1 = \partial S_2$*.* $\boldsymbol{n}_1$ *and* $\boldsymbol{n}_2$ *are the unit normal vector to* S_1 *and* S_2*, respectively. They both point to the same side of* S_1 *and* S_2 *as shown in Figure 3.6.*

If $\boldsymbol{F}$ *is the curl of another vector field, i.e.* $\boldsymbol{F} = \nabla \times \boldsymbol{G}$ *for some vector field* $\boldsymbol{G}$*, then the surface integral of* $\boldsymbol{F}$ *is surface independent.*

Proof.

$$\int_{S_1} \boldsymbol{F} \cdot \boldsymbol{n}_1 \mathrm{d}S = \int_{S_1} (\nabla \times \boldsymbol{G}) \cdot \boldsymbol{n}_1 \mathrm{d}S = \int_{\partial S_1} \boldsymbol{G} \cdot \mathrm{d}\boldsymbol{s}$$
$$= \int_{\partial S_2} \boldsymbol{G} \cdot \mathrm{d}\boldsymbol{s} = \int_{S_2} (\nabla \times \boldsymbol{G}) \cdot \boldsymbol{n}_2 \mathrm{d}S = \int_{S_2} \boldsymbol{F} \cdot \boldsymbol{n}_2 \mathrm{d}S \tag{3.85}$$

□

Example 3.13. *A vector field whose divergence is zero everywhere is called* **divergence free***. A vector field that is the curl of another vector field must be divergence free.*

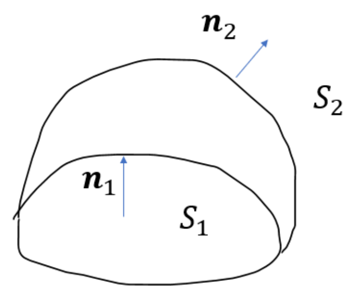

Figure 3.6: Two surfaces S_1 and S_2 that share the same boundary curve. The unit vector $\boldsymbol{n}_1$ (normal to S_1) and $\boldsymbol{n}_2$ (normal to S_2) are pointing to the same side of S_1 and S_2.

Proof. Assume $\boldsymbol{F} = \nabla \times \boldsymbol{G}$ for some $\boldsymbol{G}$. For any closed surface S,

$$\oint_S \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = \int_{S_1} \boldsymbol{F} \cdot \boldsymbol{n}_1 \, \mathrm{d}S + \int_{S_2} \boldsymbol{F} \cdot (-\boldsymbol{n}_2) \, \mathrm{d}S \tag{3.86}$$

Here S_1 and S_2 are the two parts of S. $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are unit normal vector to S_1 and S_2, respectively as shown in Figure 3.7.

Because the surface integral of $\boldsymbol{F}$ is surface independent and S_1, S_2 share the same boundary curve, we have

$$\int_{S_1} \boldsymbol{F} \cdot \boldsymbol{n}_1 \, \mathrm{d}S = \int_{S_2} \boldsymbol{F} \cdot \boldsymbol{n}_2 \, \mathrm{d}S. \tag{3.87}$$

From (3.86) and (3.87), we have

$$\oint_S \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}S = 0. \tag{3.88}$$

That is to say, the surface integral of $\boldsymbol{F}$ over any closed surface is zero. According to the definition of divergence (3.30), the divergence of $\boldsymbol{F}$ must be zero everywhere. □

3.4 Fundamental Theorems

Gradient theorem, Divergence theorem and Curl theorem are the three fundamental theorems of vector calculus. They can be considered as generalization

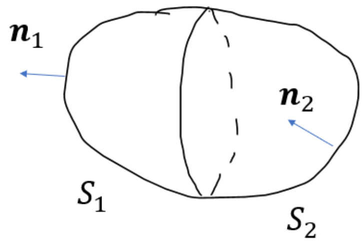

Figure 3.7: Two surfaces S_1 and S_2 that share the same boundary curve. The unit vector $\boldsymbol{n}_1$ (normal to S_1) and $\boldsymbol{n}_2$ (normal to S_2) are pointing to the same side of S_1 and S_2.

of Newton-Leibniz theorem in higher-dimension. All of them relate the integral of the 'derivative' of a field over a 'region' to the integral of the field along the boundary of this 'region'.

3.5 Notation ∇

The gradient of the scalar field, the divergence and the curl of the vector field are the three 'derivatives' of fields in vector calculus. They are often also denoted with symbol ∇ (called del or nabla) as follows.

$$\nabla u \equiv \operatorname{grad} u \tag{3.89a}$$

$$\nabla \cdot \boldsymbol{F} \equiv \operatorname{div} \boldsymbol{F} \tag{3.89b}$$

$$\nabla \times \boldsymbol{F} \equiv \operatorname{curl} \boldsymbol{F} \tag{3.89c}$$

To make sense of (3.89), we consider ∇ as the vector and differential operator

$$\nabla \equiv \hat{\boldsymbol{x}} \frac{\partial}{\partial x} + \hat{\boldsymbol{y}} \frac{\partial}{\partial y} + \hat{\boldsymbol{z}} \frac{\partial}{\partial z} \tag{3.90}$$

that acts as follows

$$\nabla u = \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right) u = \hat{\boldsymbol{x}}\frac{\partial u}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial u}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial u}{\partial z} \tag{3.91}$$

$$\begin{aligned} \nabla \cdot \boldsymbol{F} &= \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right) \cdot (F_1\hat{\boldsymbol{x}} + F_2\hat{\boldsymbol{y}} + F_3\hat{\boldsymbol{z}}) \\ &= \frac{\partial F_1}{\partial x}\hat{\boldsymbol{x}} + \frac{\partial F_2}{\partial y}\hat{\boldsymbol{y}} + \frac{\partial F_3}{\partial z}\hat{\boldsymbol{z}} \end{aligned} \tag{3.92}$$

$$\begin{aligned} \nabla \times \boldsymbol{F} &= \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right) \times (F_1\hat{\boldsymbol{x}} + F_2\hat{\boldsymbol{y}} + F_3\hat{\boldsymbol{z}}) \\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\boldsymbol{x}} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\hat{\boldsymbol{y}} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{\boldsymbol{z}}. \end{aligned} \tag{3.93}$$

Like ordinary derivatives, the calculation of these 'derivatives' of fields is also faciliated by a number of rules. Assume u and v are scalar fields, $\boldsymbol{F}$ and $\boldsymbol{G}$ are vector fields. Then we have

(Addition Rules)

$$\nabla(u+v) = \nabla u + \nabla v \tag{3.94}$$

$$\nabla \cdot (\boldsymbol{F} + \boldsymbol{G}) = \nabla \cdot \boldsymbol{F} + \nabla \cdot \boldsymbol{G} \tag{3.95}$$

$$\nabla \times (\boldsymbol{F} + \boldsymbol{G}) = \nabla \times \boldsymbol{F} + \nabla \times \boldsymbol{G} \tag{3.96}$$

(Scalar k multiplication)

$$\nabla(ku) = k\nabla u, \quad \nabla \cdot (k\boldsymbol{F}) = k(\nabla \cdot F), \quad \nabla \times (k\boldsymbol{F}) = k(\nabla \times \boldsymbol{F}) \tag{3.97}$$

(Product rules I: applying ∇ on uv and $\boldsymbol{F} \cdot \boldsymbol{G}$)

$$\nabla(uv) = u\nabla v + v\nabla u \tag{3.98}$$

$$\nabla(\boldsymbol{F} \cdot \boldsymbol{G}) = \text{not useful in MTH201} \tag{3.99}$$

(Product rules II: applying $\nabla\cdot$ on $u\boldsymbol{F}$ and $\boldsymbol{F} \times \boldsymbol{G}$)

$$\nabla \cdot (u\boldsymbol{F}) = u\nabla \cdot \boldsymbol{F} + \boldsymbol{F} \cdot \nabla u \tag{3.100}$$

$$\nabla \cdot (\boldsymbol{F} \times \boldsymbol{G}) = \text{not useful in MTH201} \tag{3.101}$$

(Product rules III: applying $\nabla\times$ on $u\boldsymbol{F}$ and $\boldsymbol{F} \times \boldsymbol{G}$)

$$\nabla \times (u\boldsymbol{F}) = u\nabla \times \boldsymbol{F} + \nabla u \times \boldsymbol{F} \tag{3.102}$$

$$\nabla \times (\boldsymbol{F} \times \boldsymbol{G}) = \text{not useful in MTH201} \tag{3.103}$$

Furthermore, by applying ∇ operator twice, we could get the following five kinds of 'second derivatives':

$$u \xrightarrow{\nabla} \nabla u \xrightarrow{\nabla \cdot} \nabla \cdot (\nabla u) \equiv \nabla^2 u \tag{3.104a}$$

$$u \xrightarrow{\nabla} \nabla u \xrightarrow{\nabla \times} \nabla \times (\nabla u) = \mathbf{0} \tag{3.104b}$$

$$\boldsymbol{F} \xrightarrow{\nabla \cdot} \nabla \cdot \boldsymbol{F} \xrightarrow{\nabla} \text{(not useful in MTH201)} \tag{3.104c}$$

$$\boldsymbol{F} \xrightarrow{\nabla \times} \nabla \times \boldsymbol{F} \xrightarrow{\nabla \cdot} \nabla \cdot (\nabla \times \boldsymbol{F}) = 0 \tag{3.104d}$$

$$\boldsymbol{F} \xrightarrow{\nabla \times} \nabla \times \boldsymbol{F} \xrightarrow{\nabla \times} \nabla \times (\nabla \times \boldsymbol{F}) = \nabla(\nabla \cdot \boldsymbol{F}) - \nabla^2 \boldsymbol{F} \tag{3.104e}$$

provided that both u and $\boldsymbol{F}$ can be differentiated many times.

The $\nabla^2 u$ in (3.104a) is called Laplacian u and will be discussed in the Chapter of Laplace and Poisson equation later on.

Exercises

3.1. *Assume electric charge is continuously distributed along the z-axis with linear charge density λ. The resulting potential ditribution u is*

$$u(P) = -\frac{\lambda}{2\pi\epsilon_0} \ln \rho \tag{3.105}$$

when P has Cylindrical coordinats (ρ, ϕ, θ) and $\rho \neq 0$. Use the relation

$$\boldsymbol{E} = -\nabla u \tag{3.106}$$

to find the corresponding electric field $\boldsymbol{E}$ due to this charge distribution.

3.2. *By taking the unit tangent vector $\boldsymbol{T} = \hat{\boldsymbol{r}}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}$, respectively in (3.1), we could get the gradient in Spherical coordinates (r, θ, ϕ):*

$$\nabla u = \frac{\partial u}{\partial r}\hat{\boldsymbol{r}} + \frac{1}{r}\frac{\partial u}{\partial \theta}\hat{\boldsymbol{\theta}} + \frac{1}{r \sin\theta}\frac{\partial u}{\partial \phi}\hat{\boldsymbol{\phi}} \tag{3.107}$$

Give details about this process.

3.3. *If scalar field u and v are different, then is it possible for their gradients to be the same ? If so, give an example. If not, prove it.*

3.4. *Assume the gradient of a scalar field is a zero vector field, i.e. (its vector is a zero vector everywhere). What is this scalar field ?*

3.5. *Verify the identity* (3.104b), *i.e.*

$$\nabla \times \nabla u = \mathbf{0} \tag{3.108}$$

by using the Cartesian representation of the gradient (3.5) *and curl* (3.61), *provided that u has continuous partial derivatives.*

3.6. *Assume stationary electric charge is continuously distributed over a region Ω. The resulting electric field $\boldsymbol{E}$ is*

$$\boldsymbol{E}(P) = kr^3\hat{\boldsymbol{r}} \tag{3.109}$$

for any point $P \in \Omega$ and has Spherical coordinates (r, ϕ, θ). Here k is a constant. Find the charge density of this continuous charge distribution.

3.7. *Find a nonzero vector field that is divergence free everywhere. Find another vector field whose divergence is 1 everywhere.*

3.8. *A student found the net flux of electric field $\boldsymbol{E}$ through a sphere S in Example* (2.11) *in the folowing way.*

According to Example (3.4),

$$div\, \boldsymbol{E} = 0 \qquad \textit{within the sphere} \tag{3.110}$$

Therefore,

$$\int_\Omega div\, \boldsymbol{E}\, \mathrm{d}\Omega = \int_\Omega 0\, \mathrm{d}\Omega = 0 \tag{3.111}$$

where Ω is the region bounded by the sphere.

Finally, according to the divergence theorem,

$$\oint_S \boldsymbol{E} \cdot \boldsymbol{n}\, \mathrm{d}S = \int_\Omega div\, \boldsymbol{E}\, \mathrm{d}\Omega = 0 \tag{3.112}$$

However, this regulst contradicts with the result in Example (2.11), *where the net flux was shown to be q/ϵ_0. What is wrong ?*

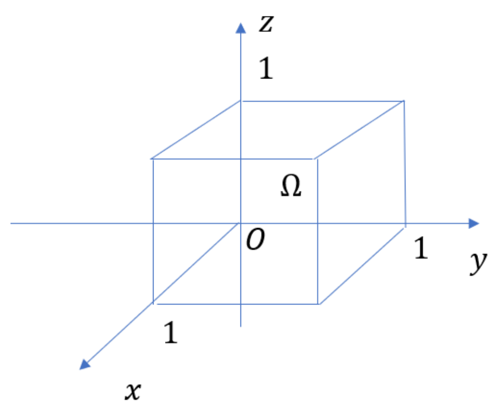

Figure 3.8: A cube of side length 1 sits on the xy-plane. One of its corder is at the origin O of the coordinates system

3.9. *Assume*

$$\boldsymbol{F}(P) = y^2\hat{\boldsymbol{x}} + (2xy + z^2)\hat{\boldsymbol{y}} + (2yz)\hat{\boldsymbol{z}} \tag{3.113}$$

for any point P with Cartesian coordinates (x, y, z). Ω is a unit cube. One of its corner is at the origin of the coordinate system as shown in Figure 3.8. Check the divergence theorem in this case. In other words, verify

$$\int_{\Omega} div\,\boldsymbol{F}\,\mathrm{d}\Omega = \oint_{\partial\Omega} \boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S \tag{3.114}$$

3.10. *Consider the vector field $\boldsymbol{F}$ and the cube Ω in exercise (3.9). Assume we cut Ω into two parts: Ω_1 and Ω_2, with a plane parallel to the xz-plane. Compare*

$$\oint_{\partial\Omega} \boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S \tag{3.115}$$

with

$$\oint_{\partial\Omega_1} \boldsymbol{F}\cdot\boldsymbol{n}_1\,\mathrm{d}S + \oint_{\partial\Omega_2} \boldsymbol{F}\cdot\boldsymbol{n}_2\,\mathrm{d}S \tag{3.116}$$

Which one is bigger ? In (3.116), $\partial\Omega_i$ is the boundary surface of $\Omega_i, i = 1, 2$. $\boldsymbol{n}_i$ is the outward unit normal vector to $\partial\Omega_i, i = 1, 2$.

3.11. *The line integral of some vector fields are path-independent while others are not. Evaluting line integral that is path independent is much easier. Therefore, it is worthwhile to check if the line integral is path independent, before evaluating it. How to check this ?*

3.12. *Justify the identity* (3.104d), *i.e.*

$$\nabla \cdot (\nabla \times \boldsymbol{F}) = 0 \tag{3.117}$$

by using the Cartesian representation of the divergence (3.33) *and curl* (3.61).

3.13. *Find a nonzero vector field that is curl free. Find another vector field whose curl is* $\hat{\boldsymbol{x}}$ *everywhere.*

3.14. *If a vector field is a constant vector field i.e. the vector of this field is the same everywhere, then this vector field is both divergence free and curl free. In additon to constant vector field, is it possible for some other vector fields to be both divergence free and curl free ? If so, give one example. If not, prove it.*

3.15. *Assume*

$$\boldsymbol{F}(P) = (2xz + 3y^2)\hat{\boldsymbol{y}} + (4yz)\hat{\boldsymbol{z}} \tag{3.118}$$

for any point P *with Cartesian coordinates* (x, y, z). Γ *is a closed path along the boundary of a square on the yz-plane as shown in Figure 3.9. Check the curl theorem (stokes' theorem) in this case by verifying*

$$\oint_{\Gamma} \boldsymbol{F} \cdot \boldsymbol{T} \,\mathrm{d}s = \int_S curl\, \boldsymbol{F} \cdot \boldsymbol{n} \,\mathrm{d}S \tag{3.119}$$

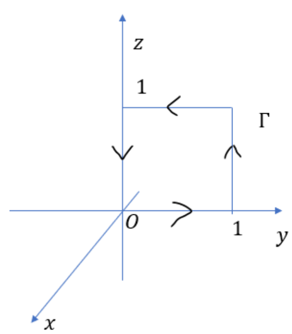

Figure 3.9: The closed path Γ is the boundary of a square on the yz-plane. The square is of length 1 on each side. Its bottom left cornder is at the origin of the coordinate system. In the line integral, the unit tangent vector $\boldsymbol{T}$ is chosen so that Γ is traced out anti-clockwisely.

3.16. *Divide the unit square in Exercise* (3.15) *into two parts as in Figure 3.10*

The one part has directed boundary $\Gamma_1 : O \to E \to D \to C \to O$*. The other part has directed boundary* $\Gamma_2 : E \to A \to B \to D \to E$*.*

Consider the line integral of $\boldsymbol{F}$ *in* (3.118) *along the directed* Γ_1 *and* Γ_2*. Compare their sum, i.e.*

$$\oint_{\Gamma_1} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} + \oint_{\Gamma_2} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} \tag{3.120}$$

with the line integral in Exercise (3.15)*. Which one is bigger ?*

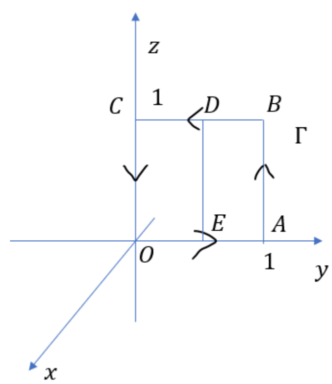

Figure 3.10: The unit square is divided into two parts. The one part has directed boundary $O \to E \to D \to C \to O$. The other part has directed boundary $E \to A \to B \to D \to E$.

3.17. *Figure 3.11 shows four vector fields in the xy-plane. The vector of all four vector fields are independent of the z-coordinate of the point. Two of the vector fields are known to be divergence free. Two of the vector fields are known to be curl free. Do you know which two fields are divergence free ? Which two fields are curl free ? (Hint: use the geometric meaning of divergence and curl)*

3.18. *Use the divergence theorem to find the surface integral of the vector field* (2.69) *through the face* $\overline{P_1P_2P_3}$ *in Exercise 2.8*

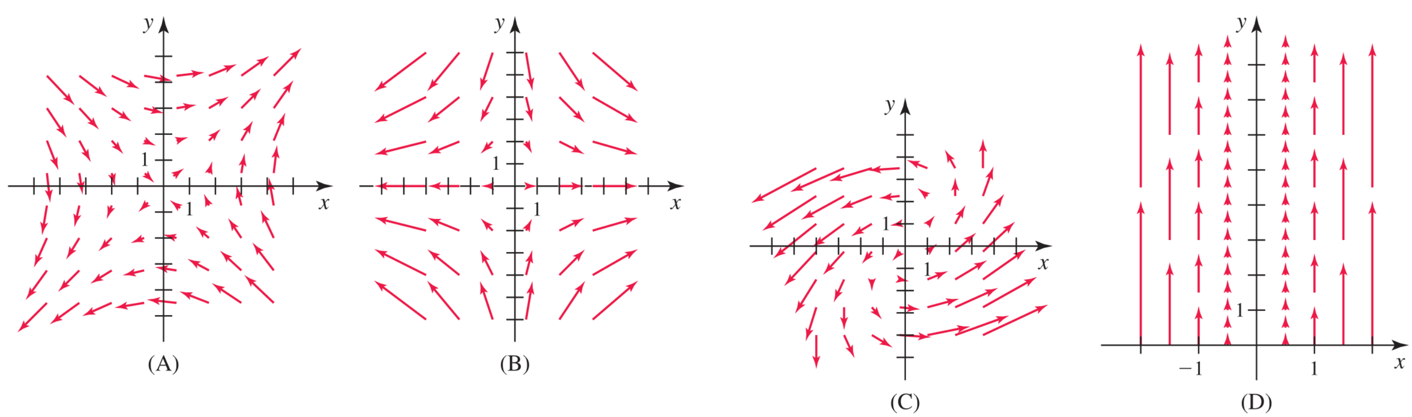

Figure 3.11

Chapter 4

Introduction to PDE

A **partial differential equation (PDE)** is an equation about a (real-valued) function. This equation has to contain the partial derivatives of this function.

Example 4.1.

$$u_t = 0, \qquad 0 < x < 1, t > 0 \tag{4.1}$$

is a PDE about the function

$$u = u(x, t), \qquad 0 \le x \le 1, t \ge 0. \tag{4.2}$$

In (4.1),

$$u_t \equiv \frac{\partial u}{\partial t} \tag{4.3}$$

is the partial derivative of u *with respect to* t*.*

The PDE (4.1) *specifies the value of* u_t *at any* (x, t) *with* $0 < x < 1, t > 0$*. They are all zeros.*

(A physical interpretation) Assume a thin rod lies on the x*-axis from* $x = 0$ *to* $x = 1$*. Its temperature at section* x *and at time* t *is* $u(x, t), 0 \le x \le 1, t \ge 0$*. Then the PDE* (4.1) *means that the temperature at any place of the rod does not change after time* $t = 0$*.*

Example 4.2. *The equation*

$$u_t(x = 1, t = 1) = 0 \tag{4.4}$$

also contains the partial derivative of u*. But it only specifies* u_t *at a single* (x, t)*. Because of this, it is usually NOT considered as a PDE. Normally,* *the PDE is required to be about the partial derivatives of the function at all possible points.*

A **solution** of PDE is a **function** that satisfies this PDE.

Example 4.3. *A solution of PDE* (4.1) *in Example* (4.1) *is the function*

$$u(x,t) = x \qquad 0 \le x \le 1, t \ge 0 \tag{4.5}$$

Another solution of this PDE is

$$u(x,t) = x^2 + 2024 \qquad 0 \le x \le 1, t \ge 0 \tag{4.6}$$

There are a lot more solutions.

The **general solution** of a PDE is the collection of all the solutions of this PDE.

Example 4.4. *The general solution of the PDE* (4.1) *in Example* (4.1) *is*

$$u(x,t) = \phi(x) \qquad 0 \le x \le 1, t \ge 0 \tag{4.7}$$

where $\phi = \phi(x)$ represents any function of x that is defined on the interval $[0,1]$.

Proof. Let

$$v(t) \overset{\text{def}}{=} u(x_0, t) \qquad t \ge 0 \tag{4.8}$$

where x_0 is an arbitrary constant in $(-\infty, \infty)$. Then (4.1) implies

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 0 \qquad t > 0 \tag{4.9}$$

whose solution is

$$v(t) = v(0) \qquad t \ge 0 \tag{4.10}$$

Given (4.8), the (4.10) means

$$u(x_0, t) = u(x_0, 0) \qquad t \ge 0. \tag{4.11}$$

Since x_0 is an arbitray constant in $(-\infty, \infty)$, the (4.11) means

$$u(x,t) = u(x,0) \qquad -\infty < x < \infty, t \ge 0 \tag{4.12}$$

The (4.12) is the solution of PDE (4.1). The (4.7) is (4.12) with

$$u(x,0) = \phi(x) \qquad -\infty < x < \infty \tag{4.13}$$

□

The **order** of a PDE about function u is the **highest** order of the partial derivatives of u in the PDE.

Example 4.5. *Bellow are four more PDEs about the function* $u = u(x,t), 0 \leq x \leq 1, t \geq 0$.

$$u_t = u_x \qquad 0 < x < 1, t > 0 \tag{4.14}$$
$$u_t = u_{xx} \qquad 0 < x < 1, t > 0 \tag{4.15}$$
$$u_{xx} = 1 \qquad 0 < x < 1, t > 0 \tag{4.16}$$
$$u_{tt} = u_{xx} \qquad 0 < x < 1, t > 0 \tag{4.17}$$

The (4.14) *is a 1st order PDE, while the rest are all 2nd order PDEs.*

Exercises

4.1. *Here are five more PDEs about the function* $u = u(x,t), 0 \leq x \leq 1, t \geq 0$ *in Example* (4.1). *Find a solution and then the general solution of each of these PDEs.*

$$u_x = 0 \qquad 0 < x < 1, t > 0 \tag{4.18}$$
$$u_t = 1 \qquad 0 < x < 1, t > 0 \tag{4.19}$$
$$u_t = x \qquad 0 < x < 1, t > 0 \tag{4.20}$$
$$u_t = t \qquad 0 < x < 1, t > 0 \tag{4.21}$$
$$u_t = u \qquad 0 < x < 1, t > 0 \tag{4.22}$$

If $u = u(x,t)$ *is about the temperature of the rod as in Example 4.1, what does the PDE* (4.18) *mean ?*

4.2. *Here are five PDEs*

$$u_x = 0 \qquad 0 < x, y, z < 1, t > 0 \tag{4.23}$$
$$u_t = 1 \qquad 0 < x, y, z < 1, t > 0 \tag{4.24}$$
$$u_t = x \qquad 0 < x, y, z < 1, t > 0 \tag{4.25}$$
$$u_t = t \qquad 0 < x, y, z < 1, t > 0 \tag{4.26}$$
$$u_t = u \qquad 0 < x, y, z < 1, t > 0 \tag{4.27}$$

about the function

$$u = u(x, y, z, t) \qquad 0 \leq x, y, z \leq 1, t \geq 0. \tag{4.28}$$

Find a solution and then the general solution of each of all these PDEs.

Chapter 5

1D Advection Equation

A simple and important 1st order PDE about the function

$$u = u(x,t), \qquad -\infty < x < \infty, t \geq 0 \tag{5.1}$$

is

$$u_t = cu_x, \qquad -\infty < x < \infty, t > 0. \tag{5.2}$$

where c is a constant.

The (5.2) is called the 1D advection equation because it could describe the one-dimensional advection process.

5.1 Physical Interpretation

Assume some chemical moves along an infinitely long tube that lies on the x-axis (Figure 5.1). Let $u(x,t)$ and $\phi(x,t)$ be its mass density and flux to the right, at cross-section x and at time t, respectively. We have

$$u_t = -\phi_x \qquad -\infty < x < \infty, t > 0 \tag{5.3}$$

provided that the amount of chemical is conserved everywhere in the tube at any time.

If

$$\phi(x,t) = -cu(x,t), \qquad -\infty < x < \infty, t > 0, \tag{5.4}$$

with some constant c, then (5.3) and (5.4) give (5.2).

The relation (5.4) characterizes the advection process. It means that the chemical moves in bulk and to the left at a speed c when $c > 0$.

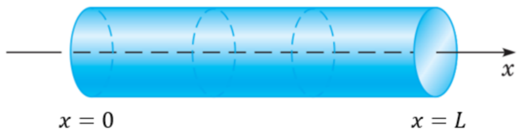

Figure 5.1: A portion of the tube in the shape of right circular cylinder. Its axis is on the x-axis.

5.2 IVP

To determine u that satisfies (5.2), we need to know what u is initially. For example, if u is initially distributed as

$$u(x, t = 0) = e^{-x^2} \qquad -\infty < x < \infty, \tag{5.5}$$

then there must be a unique function $u = u(x, t)$ that satisfy

$$u_t = cu_x, \qquad -\infty < x < \infty, t > 0 \tag{5.6a}$$

$$u(x, 0) = e^{-x^2} \qquad -\infty < x < \infty \tag{5.6b}$$

simultaneously. Finding a function $u = u(x, t)$ that satisfies (5.6) is called an **initial value problem (IVP)**.

Solution. *(of IVP* (5.6)*)*
To determine the function $u = u(x, t)$ is to find its value at any point (x^*, t^*) with $t^* \geq 0$. We find $u(x^*, t^*)$ by relating the point (x^*, t^*) to some point on the line $t = 0$ (in the xt-plane) at which u is known.

According to the chain rule,

$$\frac{\mathrm{d}u(x(t), t)}{\mathrm{d}t} = u_x \frac{\mathrm{d}x}{\mathrm{d}t} + u_t. \tag{5.7}$$

Therefore,

$$\frac{\mathrm{d}u(x(t), t)}{\mathrm{d}t} = 0 \tag{5.8}$$

on each solution curve of the equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -c \qquad -\infty < x < \infty, t > 0. \tag{5.9}$$

That is to say, $u(x(t),t)$ does not change its value, as point $(x(t),t)$ moves along any curve that is determined by the equation (5.9).

Curve of the equation

$$x - x^* = -c(t - t^*) \tag{5.10}$$

is determined by (5.9) *and contains* (x^*, t^*). *So u is constant on this curve and we have*

$$u(x = x^*, x = t^*) = u(x = x^* + ct^*, t = 0) = e^{-(x^*+ct^*)^2} \tag{5.11}$$

Overall, the solution of IVP (5.6) *is*

$$u(x,t) = e^{-(x+ct)^2}, \qquad -\infty < x < \infty, t \geq 0. \tag{5.12}$$

Exercises

5.1. *We justify the relation* (5.3) *under the condition that the total amount of chemical is conserved at any time and at any section of the tube. Fill in the blanks to complete the proof.*

Proof. The conservation of chemical means that the increase in the amount of the chemical in each portion of the tube is always equal to the amount of chemical that flows into this portion, minus the amount that flows out of this portion.

Assume the tube is a right cylinder and its cross-section has area A. Consider a short portion of the tube between $x = x_0$ and $x = x_0 + \Delta x, \Delta x > 0$. In terms of ϕ, the increase of chemical in this portion during a short time period from $t = t_0$ to $t = t_0 + \Delta t, \Delta t > 0$ is approximately

$$(\phi(x_0, t_0) - \underline{\hspace{2em}})A\Delta t \tag{5.13}$$

In terms of u, the increase is approximately

$$(u(x_0, t_0 + \Delta t) - u(x_0, t_0))A\underline{\hspace{2em}} \tag{5.14}$$

Since both (5.13) and (5.14) represent the same amount, they should be equal to each other. So we have

$$(\phi(x_0, t_0) - \phi(x_0 + \Delta x, t_0))A\Delta t = (u(x_0, t_0 + \Delta t) - u(x_0, t_0))A\Delta x \tag{5.15}$$

or equivalently

$$\frac{\phi(x_0,t_0)-\phi(x_0+\Delta x,t_0)}{\Delta x}=\frac{u(x_0,t_0+\Delta t)-u(x_0,t_0)}{\underline{\qquad\qquad}} \tag{5.16}$$

As $\Delta t\to 0, \Delta x\to 0$, (5.16) becomes

$$-\phi_x(x_0,t_0)=u_t(x_0,t_0) \tag{5.17}$$

Since (5.17) holds for any x_0 and t_0, we obtain (5.3). $\square$

5.2. *Find some functions that satisfy the PDE* (5.2)*.*

5.3. *(Linearity) Assume u and $\hat{u}$ are both solutions of* (5.2)*. Prove that $c_1u+c_2\hat{u}$ is also a solution of* (5.2)*, where c_1 and c_2 are some arbitrary constants.*

5.4. *(General Solution) Prove that function*

$$u(x,t)\stackrel{\text{def}}{=}f(x+ct) \qquad -\infty<x<\infty, t\ge 0 \tag{5.18}$$

with f an arbitrary differentiable function, satisfies (5.2)*.*

5.5. *Function* (5.18) *represents some disturbance moving to the left at speed c when $c>0$. To understand this, assume*

$$f(x)=e^{-x^2} \qquad -\infty<x<\infty \tag{5.19}$$

and $c=2$. Then sketch the graph of $u(x,t)$ at time $t=0,1,10$ (i.e. $u(x,t=0), u(x,t=1), u(x,t=10)$) on the same ux-plane.

5.6. *Consider the following initial value problem of another 1st order PDE*

$$u_t=-2tu_x, \qquad -\infty<x<\infty, t>0 \tag{5.20a}$$
$$u(x,0)=x^2, \qquad -\infty<x<\infty \tag{5.20b}$$

(i) What is a possible physical interpretation of equation (5.20a) *?*

(ii) Find the unique solution of this IVP.

5.7. *If the c in* (5.4) *depends on x and t, then* (5.4) *becomes*

$$\phi(x,t)=-c(x,t)u(x,t), \qquad -\infty<x<\infty, t>0. \tag{5.21}$$

Given (5.3) *and* (5.21)*, what will the PDE about $u=u(x,t)$ be in this case?*

Chapter 6

1D Diffusion Equation

A simple and important 2nd order PDE about a function

$$u = u(x,t) \qquad 0 \le x \le L, t \ge 0 \tag{6.1}$$

is

$$u_t = cu_{xx}, \qquad 0 < x < L, t > 0. \tag{6.2}$$

Here c and L are both positive constants.

The (6.2) is often called the 1D diffusion equation because it could describe the one-dimensional diffusion process.

6.1 Physical Interpretation

Consider again the chemical that is transported along a tube as in section (5.1). If the tube is finite and

$$\phi(x,t) = -cu_x(x,t), \qquad 0 < x < L, t > 0 \tag{6.3}$$

with c a positive constant, then (5.3) and (6.3) give (6.2).

The relation (6.3) characterizes the diffusion process. It implies that the chemical always moves in the direction from the place with high density of chemical to the place of low density.

A typical example of diffusion process is the heat conduction that follows the Fourier's law. Assume a rod of length L lies on the x-axis and its temperature at cross-section x and at time t is $u(x,t), 0 \le x \le L, t \ge 0$. It can be shown that the function $u = u(x,t)$ satisfies (6.2) under the following three conditions

(i) conservation of heat energy

$$q_t = -\phi_x \qquad 0 < x < L, t > 0 \tag{6.4}$$

Here $\phi(x,t)$ and $q(x,t)$ are the flux of heat energy to the right, and the density of heat energy at cross-section x and at time t, respectively.

(ii) diffusion of heat energy

$$\phi = -kq_x \qquad 0 < x < L, t > 0 \tag{6.5}$$

with $k > 0$ the thermal diffusivity.

(iii) the change in temperature being proportional to the change in heat energy

$$\mathrm{d}q \propto \mathrm{d}u \tag{6.6}$$

Because of this, the 1D diffusion equation (6.2) is also often called 1D heat equation.

6.2 IBVP

To determine u that satisfies (6.2), we need to know what u is initially and the behavior of u at the two ends at all the time.

For example, if u is initially distributed as

$$u(x, t = 0) = x(L - x) \qquad 0 \leq x \leq L, \tag{6.7}$$

and u is always zero at both ends, then there must be a unique function $u = u(x,t)$ satisfying all the following conditions

$$u_t = cu_{xx} \qquad 0 < x < L, t > 0 \tag{6.8a}$$

$$u(x, t = 0) = x(L - x) \qquad 0 \leq x \leq L \tag{6.8b}$$

$$u(x = 0, t) = 0 \qquad t \geq 0 \tag{6.8c}$$

$$u(x = L, t) = 0 \qquad t \geq 0 \tag{6.8d}$$

simultaneously.

To find a function $u = u(x,t)$ that satisfies (6.8), is called an **initial boundary value problem (IBVP)**.

6.3 Behavior of Solution

We discuss the behavior of the solution of the (IBVP of) diffusion equation, without knowing the explicit solution formula.

Smoothing Property. The equation (6.8a) means that at any $x_0 \in (0, L)$ and at any time $t > 0$

$$u_t(x_0, t) \approx \frac{2c}{(\Delta x)^2} \left[\frac{1}{2}(u(x_0 + \Delta x, t) + u(x_0 - \Delta x, t)) - u(x_0, t) \right] \tag{6.9}$$

when $|\Delta x|$ is small. This implies that at any time t

- u at x_0 will **increase** when its value is **smaller** than the average of its two neighbours, i.e. $\frac{1}{2}(u(x_0, t) + u(x_0 + \Delta x, t))$
- u at x_0 will **decrease** when its value is **greater** than the average of its two neighbours
- u at x_0 will **not change** when its value is **equal to** the average of its two neighbours

These have the effect of smoothing out the profile of u against x over the time. For example, Figure 6.1 shows a profile of u against x at a particular time, and how u would change at two different locations at that time.

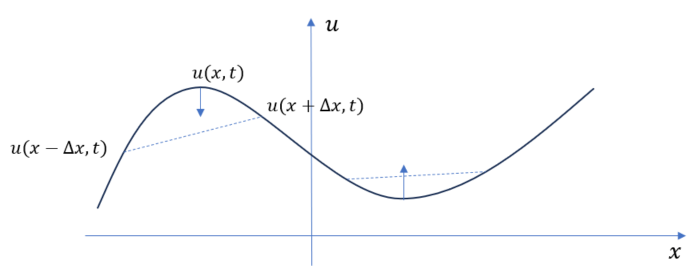

Figure 6.1: A profile of u against x at a particular time t

Example 6.1. *Consider the IBVP* (6.8) *of the heat equation. Its solution at time* $t = 0, t_1, t_2, t_3$ *with* $0 < t_1 < t_2 < t_3$ *are sketched in Figure 6.2. In the long run, the temperature of the rod will be zero everywhere. Mathematically, this means*

$$\lim_{t\to\infty} u(x, t) = 0 \qquad 0 \le x \le L \tag{6.10}$$

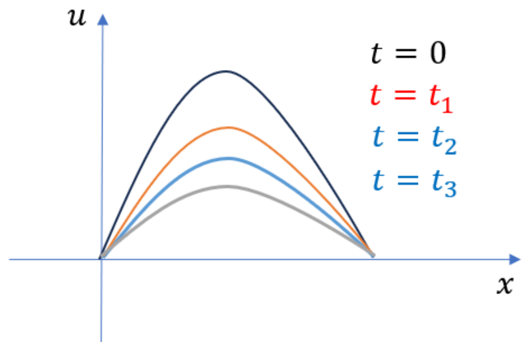

Figure 6.2: Profiles of u against x at time $t = 0, t_1, t_2, t_3$ with $0 < t_1 < t_2 < t_3$

Exercises

6.1. *In heat conduction along the rod, show that the temperature function* $u = u(x, t)$ *satisfies the heat equation* (6.2) *under the condition* (6.4), (6.5) *and* (6.6).

6.2. *Find the solution of the following IBVP of the heat equation.*

$$\begin{aligned} u_t &= 2u_{xx} && 0 < x < 3, t > 0 && \text{(6.11a)} \\ u(x, t = 0) &= 2x + 1 && 0 \leq x \leq 3 && \text{(6.11b)} \\ u(x = 0, t) &= 1 && t \geq 0 && \text{(6.11c)} \\ u(x = 3, t) &= 7 && t \geq 0 && \text{(6.11d)} \end{aligned}$$

6.3. *(Linearity) Assume* u *and* $\hat{u}$ *are both solutions of diffusion equation* (6.2). *Show that* $c_1 u + c_2 \hat{u}$ *is also a solution of* (6.2), *where* c_1 *and* c_2 *are some arbitrary constants.*

6.4. *We look for functions that satisfy the heat equation* (6.2) *in the following form*

$$u(x, t) = e^{at} \cos(bx) \qquad 0 < x < L, t > 0 \tag{6.12}$$

where a, b *are constants. What numbers should* a, b *be so that* (6.12) *is a solution of* (6.2).

6.5. *Here is another IBVP of the diffusion equation.*

$$\begin{aligned}
u_t &= cu_{xx} && 0 < x < L, t > 0 && (6.13a)\\
u(x,t=0) &= x && 0 \le x \le L && (6.13b)\\
u(x=0,t) &= 0 && t \ge 0 && (6.13c)\\
u(x=L,t) &= 0 && t \ge 0 && (6.13d)
\end{aligned}$$

It has no solution because some condition contradicts with some other condition. This makes it impossible to find a function that satisfy all these conditions in (6.13) *simultaneously. Can you find out which conditions are contradicting with each other ?*

6.6. *Consider the following IBVP of the heat equation. Could you sketch its solution at several different times as in Example* (6.1)*.*

$$\begin{aligned}
u_t &= 2u_{xx} && 0 < x < 1, t > 0 && (6.14a)\\
u(x,t=0) &= \sin(\pi x) && 0 \le x \le 1 && (6.14b)\\
u_x(x=0,t) &= 0 && t > 0 && (6.14c)\\
u_x(x=1,t) &= 0 && t > 0 && (6.14d)
\end{aligned}$$

6.7. *Here is a variant of the basic heat equation*

$$u_t = cu_{xx} + h(x) \qquad 0 < x < L, t > 0 \qquad (6.15)$$

where $h(x)$ is a nonzero term. In the context of heat conduction along the rod, $h(x)$ account for the heat transfer through the side surface of the rod, and its value is proportional to the heat following into (if $h(x) > 0$) the rod from its lateral side at location x, per unit time.

Now consider the following IBVP of an example of (6.15)*.*

$$\begin{aligned}
u_t &= \frac{1}{4}u_{xx} + 1 && 0 < x < 1, t > 0 && (6.16a)\\
u(x,t=0) &= x && 0 \le x \le 1 && (6.16b)\\
u(x=0,t) &= 0 && t \ge 0 && (6.16c)\\
u(x=1,t) &= 1 && t \ge 0 && (6.16d)
\end{aligned}$$

Could you sketch its solution at several different times as in the exercise above. What is the long-term behavior of its solution, i.e. the limit $\lim_{t\to\infty} u(x,t)$ *?*

Chapter 7

1D Laplace/Poisson Equation

The 1D Poisson equation about a function $u = u(x), 0 \leq x \leq L$ is

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = f(x) \qquad 0 < x < L. \tag{7.1}$$

Here $f(x)$ is known for any $x \in (0, L)$. If $f(x) \equiv 0, 0 < x < L$, then (7.1) is called the 1D Laplace equation.

The (7.1) is an ordinary differential equation (ODE), we see it as a special case of PDE that models some equilibrium phenomena.

7.1 Physical Interpretation

Consider the heat conduction along a rod that is subject to external heating. According to the last exercise of Chapter 6, the temperature $u = u(x, t)$ should satisfy

$$u_t = cu_{xx} + h(x) \qquad 0 < x < L, t > 0 \tag{7.2}$$

where $h(x), 0 \leq x \leq L$ accounts for external heating. If the distribution of temperature at some time $t = t_0$ satisfies

$$cu_{xx}(x, t_0) + h(x) = 0 \qquad 0 < x < L, \tag{7.3}$$

then

$$u_t(x,t)|_{t=t_0} = 0 \qquad 0 < x < L. \tag{7.4}$$

The (7.4) means that at time $t = t_0$, u does not change at any x. Then after time $t = t_0$, the distribution of temperature will remain the same. We say that the temperature distribution along the rod reaches equilibrium or is stationary.

The stationary distribution of temperature is characterized by (7.3) which is essentially an ODE (7.1) with

$$u(x) = u(x, t_0) \qquad 0 \leq x \leq L \tag{7.5}$$

and

$$f(x) = -h(x)/c \qquad 0 < x < L \tag{7.6}$$

7.2 BVP

To determine the function u that satisfies (7.1), we also need to know its behavior at the two ends of the interval $[0, L]$.

Example 7.1. *Assume* $L = 1$ *and* $f(x) = x, 0 < x < 1$. *If*

$$u(x = 0) = 1 \quad \textit{and} \quad u(x = 1) = 0, \tag{7.7}$$

then there must be a unique function $u = u(x)$ *that satisfies*

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = x \qquad 0 < x < 1. \tag{7.8a}$$

$$u(x = 0) = 1 \tag{7.8b}$$

$$u(x = 1) = 0 \tag{7.8c}$$

simultaneously.

To find a function $u = u(x)$ *that satisfies all the conditions in* (7.8) *is called a* **boundary value problem (BVP)**

Exercises

7.1. *Find two different functions that are the solutions of the Laplace equation*

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = 0 \qquad 0 < x < L \tag{7.9}$$

7.2. *Find all the solutions (i.e. the general solution) of the Laplace equation* (7.9).

7.3. *Among the solutions of* (7.9)*, find the one that also satisfies the following two conditions*

$$u(x = 0) = 1 \quad \textit{and} \quad u(x = 1) = 0. \tag{7.10}$$

In other words, find the solution of the following BVP of 1D Laplace equation (7.9)

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = x \qquad 0 < x < 1. \tag{7.11a}$$
$$u(x = 0) = 1 \tag{7.11b}$$
$$u(x = 1) = 0 \tag{7.11c}$$

7.4. *Find the solution of BVP* (7.8)*.*

7.5. *Here is another BVP of the Poisson equation*

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = x \qquad 0 < x < 1. \tag{7.12a}$$
$$u_x(x = 0) = 1 \tag{7.12b}$$
$$u_x(x = 1) = 0 \tag{7.12c}$$

Find all of its solutions.

7.6. *The ODE*

$$\frac{1}{x^2}\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}\right) = 5 \qquad 0 < x < 3 \tag{7.13}$$

about function

$$u = u(x), \quad 0 \le x \le 3 \tag{7.14}$$

is sometimes called an example of generalized Poisson equation.

1. *Find all of its solutions.*
2. *Among all of its solutions, find those that also satisfy the following condition*

$$\lim_{x\to 0^+} u(x) \ \textit{is finite.} \tag{7.15}$$

Chapter 8

1D Wave equation

Another important 2nd order PDE about the function $u = u(x,t)$ is

$$u_{tt} = v^2 u_{xx} \qquad -\infty < x < \infty, t > 0. \tag{8.1}$$

Here v is a positive constant.

The (8.1) is called the 1D wave equation because it is satisfied by the functions representing one-dimensional (travelling) waves, i.e.

$$u(x,t) = f(x - vt) \quad \text{or} \quad g(x + vt) \qquad -\infty < x < \infty, t \geq 0 \tag{8.2}$$

Here f and g could be any differentiable functions.

Proof. If

$$u(x,t) = f(x - vt) \qquad -\infty < x < \infty, t \geq 0, \tag{8.3}$$

then

$$u_x = f'(x - vt) \qquad u_{xx} = f''(x - vt) \tag{8.4}$$

and

$$u_t = -vf'(x - vt) \qquad u_{tt} = (-v)^2 f''(x - vt) \tag{8.5}$$

So we have (8.1).

If

$$u(x,t) = g(x + vt) \qquad -\infty < x < \infty, t \geq 0, \tag{8.6}$$

then

$$u_x = g'(x + vt) \qquad u_{xx} = g''(x + vt) \tag{8.7}$$

and

$$u_t = vg'(x + v) \qquad u_{tt} = v^2 g''(x + vt) \tag{8.8}$$

So we again have (8.1). □

Example 8.1. *If*

$$u(x,t) = f(x - vt) \qquad -\infty < x < \infty, t \geq 0 \tag{8.9}$$

with

$$f(x) = A\sin(kx) \qquad -\infty < x < \infty, \tag{8.10}$$

then

$$u(x,t) = A\sin k(x - vt), \qquad -\infty < x < \infty, t \geq 0 \tag{8.11}$$

Here A and k are some positive constants.

The function (8.11) represents a one-dimensional sinusoidal (or harmonic) wave travelling at speed v and in the positive x-direction. The constant A is the amplitude of the wave and k is called the propagation number.

Sinusoidal waves are important, because almost any wave could be represented as a superposition of these waves (i.e. Fourier series representation).

8.1 Physical Interpretation

The wave equation (8.1) could be used to model the voltage and current in a long transmission line as shown in Figure 8.1. Consider a section of this transmission line located at x, and assign currents and voltages as shown in Figure 8.2.

According to Kirchhoff's voltage and current law, we have

$$V(x,t) = L\frac{\partial I(x,t)}{\partial t} + V(x + \Delta x, t) \tag{8.12a}$$

$$I(x,t) = C\frac{\partial V(x + \Delta x, t)}{\partial t} + I(x + \Delta x, t) \tag{8.12b}$$

With linear approximation from Calculus

$$V(x + \Delta x, t) \approx V(x,t) + \frac{\partial V(x,t)}{\partial x}\Delta x \tag{8.13a}$$

$$I(x + \Delta x, t) \approx I(x,t) + \frac{\partial I(x,t)}{\partial x}\Delta x, \tag{8.13b}$$

the equations (8.12) then become

$$\frac{\partial V(x,t)}{\partial x}\Delta x \approx -L\frac{\partial I(x,t)}{\partial t} \tag{8.14a}$$

$$\frac{\partial I(x,t)}{\partial x}\Delta x \approx -C\frac{\partial V(x,t)}{\partial t} - C\frac{\partial^2 V(x,t)}{\partial t \partial x}\Delta x \tag{8.14b}$$

or equivalently

$$\frac{\partial V(x,t)}{\partial x} \approx -\frac{L}{\Delta x}\frac{\partial I(x,t)}{\partial t} \tag{8.15a}$$

$$\frac{\partial I(x,t)}{\partial x} \approx -\frac{C}{\Delta x}\frac{\partial V(x,t)}{\partial t} - C\frac{\partial^2 V(x,t)}{\partial t \partial x} \tag{8.15b}$$

If $L/\Delta x$ approaches constant $\tilde{L}$ and $C/\Delta x$ approaches constant $\tilde{C}$ as $\Delta x \to 0$, then (8.15) becomes

$$\frac{\partial V(x,t)}{\partial x} = -\tilde{L}\frac{\partial I(x,t)}{\partial t} \tag{8.16a}$$

$$\frac{\partial I(x,t)}{\partial x} = -\tilde{C}\frac{\partial V(x,t)}{\partial t} \tag{8.16b}$$

as $\Delta \to 0$.

Differentiating both sides of (8.16) with respect to x or t, we have

$$\frac{\partial^2 V(x,t)}{\partial t^2} = \frac{1}{\tilde{L}\tilde{C}}\frac{\partial^2 V(x,t)}{\partial x^2} \tag{8.17a}$$

$$\frac{\partial^2 I(x,t)}{\partial t^2} = \frac{1}{\tilde{L}\tilde{C}}\frac{\partial^2 I(x,t)}{\partial x^2} \tag{8.17b}$$

If the transmission line is infinitely along, then (8.17) contains two examples of wave equations (8.1) with $u = V$ or I and $v^2 = 1/(\tilde{L}\tilde{C})$.

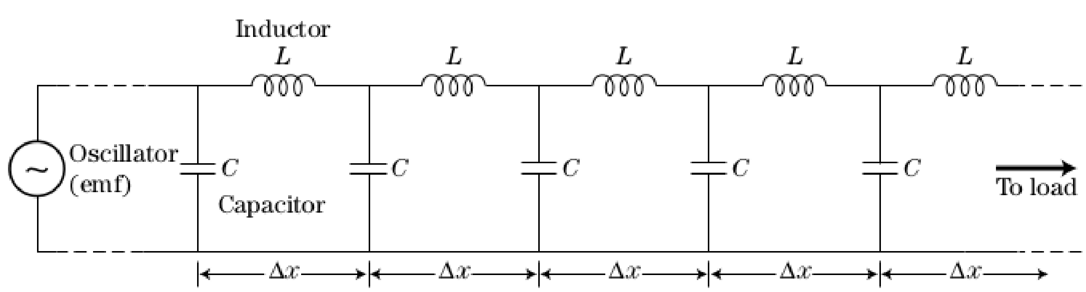

Figure 8.1: LC transmission line (Ref: Fundamental of wave phenomena, Hirose, Ch.9)

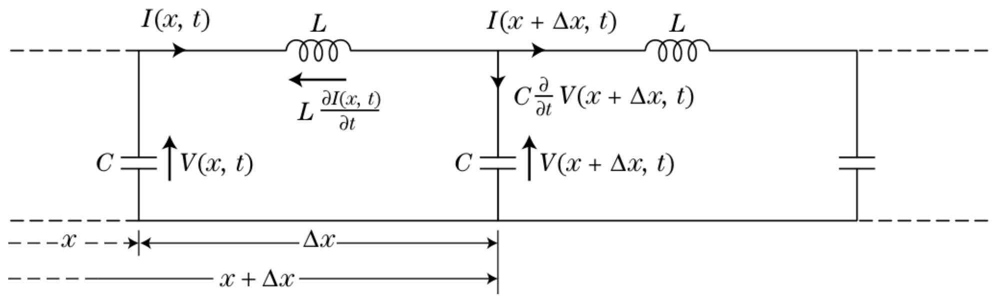

Figure 8.2: Voltages and currents at one section of an LC transmission line (Ref: Fundamental of wave phenomena, Hirose, Ch.9)

8.2 IVP

To determine u that satisfies (8.1), we also need to know both u and u_t in the beginning.

Example 8.2. *If we know*

$$u(x, t = 0) = e^{-x^2} \qquad -\infty < x < \infty \tag{8.18}$$
$$u_t(x, t = 0) = 0 \qquad -\infty < x < \infty \tag{8.19}$$

then there must be a unique function $u = u(x, t)$ that satisfies all the following three conditions

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.20a}$$
$$u(x, 0) = e^{-x^2} \qquad -\infty < x < \infty \tag{8.20b}$$
$$u_t(x, 0) = 0 \qquad -\infty < x < \infty \tag{8.20c}$$

simultaneously.

To find a function $u = u(x, t)$ that satisfies (8.20) *is called an initial value problem (IVP) of the wave equation* (8.1)*.*

The solution of IVP (8.20) *is*

$$u(x, t) = \frac{1}{2}e^{-(x-vt)^2} + \frac{1}{2}e^{-(x+vt)^2} \qquad -\infty < x < \infty \tag{8.21}$$

Proof. First, we let

$$u(x, t) = f(x - vt) + g(x + vt) \qquad -\infty < x < \infty, t \geq 0 \tag{8.22}$$

for some function f and g. According to the discussion in the beginning of this chapter, the function (8.22) satisfies the condition (8.20a).

Next, we choose appropriate functions f and g so that the function (8.22) also meet the other two conditions in (8.20).

Given (8.22), the condition (8.20b) becomes

$$f(x) + g(x) = e^{-x^2} \qquad -\infty < x < \infty, \tag{8.23}$$

the condition (8.20c) becomes

$$-vf'(x) + vg'(x) = 0 \qquad -\infty < x < \infty, \tag{8.24}$$

or equivalently

$$f'(x) = g'(x) \qquad -\infty < x < \infty \tag{8.25}$$

The (8.25) means

$$g(x) = f(x) + C \qquad -\infty < x < \infty \tag{8.26}$$

From (8.23) and (8.26), we know that

$$f(x) = \frac{1}{2}e^{-x^2/2} - \frac{C}{2} \qquad -\infty < x < \infty \tag{8.27a}$$

$$g(x) = \frac{1}{2}e^{-x^2/2} + \frac{C}{2} \qquad -\infty < x < \infty \tag{8.27b}$$

Given (8.27), the (8.22) becomes (8.21). □

Interpretation *The solution function* (8.21) *could be interpreted as a superposition of two waves* $\frac{1}{2}e^{-(x-vt)^2}$ *and* $\frac{1}{2}e^{-(x+vt)^2}$ *(See an animation of an example of this in the module page of LMO). As a function of* x, t, $\frac{1}{2}e^{-(x-vt)^2}$ *represents a wave of profile* $\frac{1}{2}e^{-x^2}$ *that travels in the positive* x*-direction with speed* v*; while* $\frac{1}{2}e^{-(x+vt)^2}$ *represents a wave of the same profile that travels in the negative* x*-direction with the same speed* v*.*

Example 8.3. *Solve the following IVP of the 1D wave equation* (8.1)

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.28a}$$

$$u(x,0) = 0 \qquad -\infty < x < \infty \tag{8.28b}$$

$$u_t(x,0) = \frac{1}{1+x^2} \qquad -\infty < x < \infty \tag{8.28c}$$

In other words, find the function $u = u(x,t)$ *that satisfies all the conditions in* (8.28) *simultaneously.*

Solution. *As in the solution of last example, we first assume u is in the form* (8.22) *for some function f and g. Then we choose appropriate f and g so that u also meet the other two conditions.*

Given (8.22), *condition* (8.28b) *becomes*

$$f(x) + g(x) = 0 \qquad -\infty < x < \infty \tag{8.29}$$

and condition (8.28c) *becomes*

$$-vf'(x) + vg'(x) = -\frac{1}{1+x^2} \qquad -\infty < x < \infty \tag{8.30}$$

From (8.29) *and* (8.30), *we have*

$$f(x) = -g(x) \qquad -\infty < x < \infty \tag{8.31}$$

$$f'(x) = \frac{1}{2v\,(1+x^2)} \qquad -\infty < x < \infty \tag{8.32}$$

which means

$$f(x) = -g(x) = \frac{1}{2v}\arctan x + f(0) \qquad -\infty < x < \infty \tag{8.33}$$

Therefore,

$$\begin{aligned} u(x,t) &= f(x-vt) + g(x+vt) \\ &= \frac{1}{2v}\left(\arctan(x-vt) - \arctan(x+vt)\right) \qquad -\infty < x < \infty, t \geq 0 \end{aligned} \tag{8.34}$$

Example 8.4. *Here is another IVP of the 1D wave equation* (8.1)

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.35a}$$

$$u(x,0) = e^{-x^2} \qquad -\infty < x < \infty \tag{8.35b}$$

$$u_t(x,0) = \frac{1}{1+x^2} \qquad -\infty < x < \infty \tag{8.35c}$$

Its solution is the sum of the solution of the IVP (8.20) *and IVP* (8.28), *i.e.*

$$u(x,t) = \frac{1}{2}e^{-(x-vt)^2} + \frac{1}{2}e^{-(x+vt)^2} + \frac{1}{2v}\left(\arctan(x-vt) - \arctan(x+vt)\right) \qquad -\infty < x < \infty, t \geq 0 \tag{8.36}$$

8.3 IBVP

Sometimes the wave equation is defined only on a finite interval, e.g.

$$u_{tt} = v^2 u_{xx} \qquad 0 < x < L, t > 0. \tag{8.37}$$

Here L is a positive constant.

To determine u that satisfies (8.37), we also need to know the behavior of u at the two ends of the interval, in addition to the initial values of u and u_t.

Example 8.5. *If we know*

$$u(0,t) = u(L,t) = 0 \qquad t \geq 0 \tag{8.38}$$

then there must be a unique function $u = u(x,t)$ that satisfies all the following five conditions

$$\begin{aligned}
u_{tt} &= v^2 u_{xx}, && 0 < x < L, t > 0 && (8.39a)\\
u(x,0) &= \sin\frac{\pi x}{L} && 0 \leq x \leq L && (8.39b)\\
u_t(x,0) &= 0 && 0 < x < L && (8.39c)\\
u(0,t) &= 0 && t \geq 0 && (8.39d)\\
u(L,t) &= 0 && t \geq 0 && (8.39e)
\end{aligned}$$

simultaneously.

To find a function $u = u(x,t)$ that satisfies all the conditions in (8.39) *is called an initial boundary value problem (IBVP) of the wave equation* (8.37)*.*

The solution of IBVP (8.39) *is*

$$u(x,t) = \sin\frac{\pi x}{L}\cos\frac{v\pi t}{L} \qquad 0 \leq x \leq L, t \geq 0 \tag{8.40}$$

Check that this function indeed satisfies all the conditions in (8.39)*.*

Interpretation. *The solution function* (8.40) *represents a standing wave. (See an animation of an example of this in the module page of LMO). Assume all the points that are initially on the interval $[0, L]$ on the x-axis move in a transverse direction. Let $u(x,t), 0 \leq x \leq L, t \geq 0$ be the displacement at time t of the point that is initially at x on the x-axis (Figure 8.3). The displacement is positive if the point moves above x-axis, and negative if it*

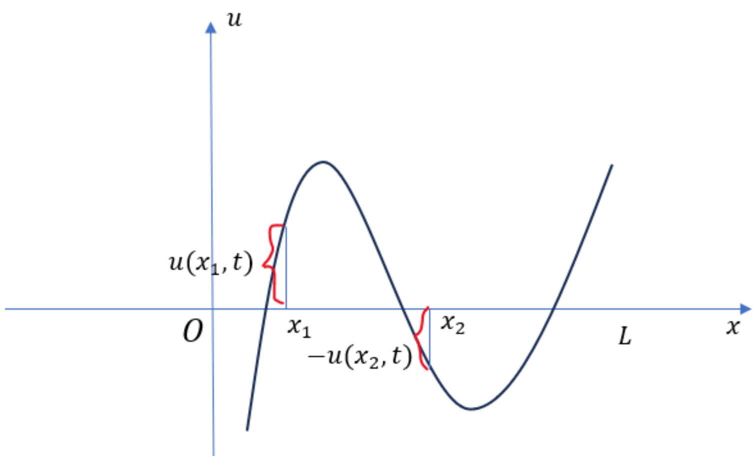

Figure 8.3: A snapshot of a collection of moving points at a particular time t. Each point moves in the u-direction. For the point that is initially at x_i on the x-axis, $u(x_i, t)$ is its displacement from its initial position on the x-axis. The displacement takes a positive value if the point moves above the x-axis, and a negative value if it is bellow.

moves below. Then the solution function (8.40) *means each point does a simple harmonic motion with the same frequency and is in phase with each other.*

According to trigonometric identity

$$\sin\alpha\cos\beta = \frac{1}{2}\left(\sin(\alpha+\beta) + \sin(\alpha-\beta)\right) \tag{8.41}$$

the standing wave (8.40) *could also be written as*

$$\begin{aligned} u(x,t) &= \frac{1}{2}\left(\sin\left(\frac{\pi x}{L} + \frac{v\pi t}{L}\right) + \sin\left(\frac{\pi x}{L} - \frac{v\pi t}{L}\right)\right) \\ &= \frac{1}{2}\sin\left(\frac{\pi}{L}(x+vt)\right) + \frac{1}{2}\sin\left(\frac{\pi}{L}(x-vt)\right) \end{aligned} \tag{8.42}$$

This indicates that the standing wave is a superposition of two travelling waves. The two waves have the same shape at $t = 0$ (i.e. $\sin(\pi x/L)$), the same speed (i.e. v), but travel in the opposite directions.

Example 8.6. *Here is another IBVP of the wave equation* (8.1).

$$u_{tt} = v^2 u_{xx}, \qquad 0 < x < L, t > 0 \tag{8.43a}$$

$$u(x,0) = \sin\frac{k\pi x}{L} \qquad 0 \le x \le L \tag{8.43b}$$

$$u_t(x,0) = 0 \qquad 0 < x < L \tag{8.43c}$$

$$u(0,t) = 0 \qquad t \ge 0 \tag{8.43d}$$

$$u(L,t) = 0 \qquad t \ge 0 \tag{8.43e}$$

where k is some positive integer.

Check that its solution is

$$u(x,t) = \sin\frac{k\pi x}{L}\cos\frac{kv\pi t}{L} \qquad 0 \le x \le L, t \ge 0 \tag{8.44}$$

Example 8.7. *The IBVP* (8.43) *is the IBVP*

$$u_{tt} = v^2 u_{xx}, \qquad 0 < x < L, t > 0 \tag{8.45a}$$

$$u(x,0) = f(x) \qquad 0 \le x \le L \tag{8.45b}$$

$$u_t(x,0) = 0 \qquad 0 < x < L \tag{8.45c}$$

$$u(0,t) = 0 \qquad t \ge 0 \tag{8.45d}$$

$$u(L,t) = 0 \qquad t \ge 0 \tag{8.45e}$$

when

$$f(x) = \sin\frac{k\pi x}{L}, \qquad k = 1, 2, ... \tag{8.46}$$

When

$$f(x) = 3\sin\frac{2\pi x}{L} + 5\sin\frac{3\pi x}{L} \qquad -\infty < x < \infty, \tag{8.47}$$

the solution of IBVP (8.45) *is*

$$u(x,t) = 3\sin\frac{2\pi x}{L}\cos\frac{2v\pi t}{L} + 5\sin\frac{3\pi x}{L}\cos\frac{3v\pi t}{L} \quad 0 \le x \le L, t \ge 0 \tag{8.48}$$

When

$$f(x) = \sum_{k=1}^{\infty} c_k \sin\frac{k\pi x}{L} \qquad -\infty < x < \infty \tag{8.49}$$

where $c_1, c_2, ...$ are constants, the solution of IBVP (8.45) *is*

$$u(x,t) = \sum_{k=1}^{\infty} c_k \sin\frac{k\pi x}{L}\cos\frac{kv\pi t}{L} \quad 0 \le x \le L, t \ge 0 \tag{8.50}$$

Exercises

8.1. *In addition to* (8.11)*, Sinusoid waves are sometimes also represented in following ways*

$$u(x,t) = A\sin\frac{2\pi}{\lambda}(x - vt) \tag{8.51a}$$

$$u(x,t) = A\sin 2\pi\left(\frac{x}{\lambda} \pm \frac{t}{\tau}\right) \tag{8.51b}$$

$$u(x,t) = A\sin 2\pi(\kappa x \pm \nu t) \tag{8.51c}$$

$$u(x,t) = A\sin(kx \pm \omega t) \tag{8.51d}$$

$$u(x,t) = A\sin 2\pi\nu\left(\frac{x}{v} \pm t\right) \tag{8.51e}$$

Except for x, t and u, all the letters in the above expression are constants. What are the requirement on these constants so that each of them satisfies the 1D wave equation (8.1).

8.2. *(Linearity) Assume u and $\hat{u}$ are both solutions of the wave equation* (8.1)*. Prove that $c_1 u + c_2\hat{u}$ is also a solution of* (8.1)*, where c_1 and c_2 are some arbitrary constants.*

8.3. *Consider the following IVP of 1D wave equation* (8.1)

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.52a}$$

$$u(x,0) = f(x) \qquad -\infty < x < \infty \tag{8.52b}$$

$$u_t(x,0) = 0 \qquad -\infty < x < \infty \tag{8.52c}$$

1. *Find its solution (in terms of f). (Hint: when $f(x) = e^{-x^2}$, this IVP becomes the IVP* (8.20)*).*

2. *Sketch the graph of its solution $u = u(x,t)$ at time $t = 0, 1, 10$, respectively, provided that*

$$f(x) = \begin{cases} 1 & -3 < x < 3 \\ 0 & \text{otherwise} \end{cases} \tag{8.53}$$

   and $v = 2$.

8.4. *Find the solution of the IVP*

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.54a}$$
$$u(x,0) = 0 \qquad -\infty < x < \infty \tag{8.54b}$$
$$u_t(x,0) = g(x) \qquad -\infty < x < \infty \tag{8.54c}$$

(Hint: when $g(x) = 1/(1+x^2)$*, this IVP becomes the IVP* (8.28)*).*

8.5. *Find the solution of the IVP*

$$u_{tt} = v^2 u_{xx}, \qquad -\infty < x < \infty, t > 0 \tag{8.55a}$$
$$u(x,0) = f(x) \qquad -\infty < x < \infty \tag{8.55b}$$
$$u_t(x,0) = g(x) \qquad -\infty < x < \infty \tag{8.55c}$$

8.6. *(This exercise is OPTIONAL)*
We could also solve the IVP of 1D wave equation (8.1) *by solving two IVPs of advection equations*

Take the IVP (8.20) *for example. Let*

$$w(x,t) = u_t(x,t) - vu_x(x,t) \qquad -\infty < x < \infty, t \geq 0 \tag{8.56}$$

Then

$$\begin{aligned} w(x,0) &= u_t(x,0) - vu_x(x,0) \\ &= 0 - v(-2x)e^{-x^2} = 2vxe^{-x^2} \end{aligned} \tag{8.57}$$

and (8.20a) *becomes*

$$w_t + vw_x = 0 \qquad -\infty < x < \infty, t > 0. \tag{8.58}$$

1. *Find the* w *that satisfies both* (8.57) *and* (8.58)*. (an IVP of 1D advection about* w*).*
2. *After finding the* w *in step (1), we find the* u *that satisfies both* (8.56) *and* (8.20b)*. (an IVP of 1st order PDE about* u*)*

8.7. *Find the solutions of the following four IBVPs of the 1D wave equation* (8.1)*, respectively*

$$u_{tt} = v^2 u_{xx}, \qquad 0 < x < L, t > 0 \tag{8.59a}$$
$$u(x,0) = 0 \qquad 0 \leq x \leq L \tag{8.59b}$$
$$u_t(x,0) = \sin\frac{\pi x}{L} \qquad 0 < x < L \tag{8.59c}$$
$$u(0,t) = 0 \qquad t \geq 0 \tag{8.59d}$$
$$u(L,t) = 0 \qquad t \geq 0 \tag{8.59e}$$

$$
\begin{aligned}
u_{tt} &= v^2 u_{xx}, && 0 < x < L, t > 0 && (8.60a)\\
u(x,0) &= 0 && 0 \le x \le L && (8.60b)\\
u_t(x,0) &= \sin\frac{k\pi x}{L} && 0 < x < L, k = 1, 2, ... && (8.60c)\\
u(0,t) &= 0 && t \ge 0 && (8.60d)\\
u(L,t) &= 0 && t \ge 0 && (8.60e)
\end{aligned}
$$

$$
\begin{aligned}
u_{tt} &= v^2 u_{xx}, && 0 < x < L, t > 0 && (8.61a)\\
u(x,0) &= 0 && 0 \le x \le L && (8.61b)\\
u_t(x,0) &= \sum_{k=1}^{\infty} c_k \sin\frac{k\pi x}{L} && 0 < x < L && (8.61c)\\
u(0,t) &= 0 && t \ge 0 && (8.61d)\\
u(L,t) &= 0 && t \ge 0 && (8.61e)
\end{aligned}
$$

$$
\begin{aligned}
u_{tt} &= v^2 u_{xx}, && 0 < x < L, t > 0 && (8.62a)\\
u(x,0) &= \sum_{k=1}^{\infty} d_k \sin\frac{k\pi x}{L} && 0 \le x \le L && (8.62b)\\
u_t(x,0) &= \sum_{k=1}^{\infty} c_k \sin\frac{k\pi x}{L} && 0 < x < L && (8.62c)\\
u(0,t) &= 0 && t \ge 0 && (8.62d)\\
u(L,t) &= 0 && t \ge 0 && (8.62e)
\end{aligned}
$$

8.8. *Sketch the solution of IBVP* (8.59) *at times* $t = 0, 1, 2, 3$ *respectively, provided that* $v = 2$ *and* $L = 3$.

8.9. *Could you find an IBVP of the 1D wave equation* (8.1) *so that its unique solution is*

$$
u(x,t) = \cos\frac{\pi x}{L}\cos\frac{v\pi t}{L} \qquad 0 \le x \le L, t \ge 0 \text{ ?} \tag{8.63}
$$

How about an IBVP whose unique solution is

$$u(x,t) = \cos\frac{\pi x}{L}\sin\frac{v\pi t}{L} \qquad 0 \le x \le L, t \ge 0 \text{ ?} \tag{8.64}$$

8.10. *Consider another IBVP of 1D wave equation* (8.1) *that is little different from the IBVP* (8.39). *In* (8.39), *u is required to be zero at the two ends for all the time. Here, however, u is required to take some nonzero values at the two ends. What is the solution of this IBVP ?*

$$u_{tt} = v^2 u_{xx}, \qquad 0 < x < L, t > 0 \tag{8.65a}$$

$$u(x,0) = \sin\frac{\pi x}{L} + \frac{x}{L} + 1 \qquad 0 \le x \le L \tag{8.65b}$$

$$u_t(x,0) = 0 \qquad 0 < x < L \tag{8.65c}$$

$$u(0,t) = 1 \qquad t \ge 0 \tag{8.65d}$$

$$u(L,t) = 2 \qquad t \ge 0 \tag{8.65e}$$

Chapter 9

Fourier Series

9.1 Motivation

In solving the IBVP (8.45) in Example (8.7), we find that if function $f = f(x)$ could be represented as a linear combination of functions

$$\sin\frac{\pi x}{L}, \sin\frac{2\pi x}{L}, \sin\frac{3\pi x}{L}, \cdots, \tag{9.1}$$

over the interval $[0, L]$, then the solution of this IBVP could be found easily.

A natural question to ask is: Could any function $f = f(x)$ be represented as a linear combination of functions from (9.1), over the interval $[0, L]$?

9.2 Fourier Series

We start with approximating or representing function $f = f(x)$ over the interval $(-L, L)$, with standard functions

$$f_1(x) \overset{\text{def}}{=} \cos\frac{0\cdot x}{L} = 1, \quad f_2(x) \overset{\text{def}}{=} \sin\frac{\pi x}{L}, \quad f_3(x) \overset{\text{def}}{=} \cos\frac{\pi x}{L}, \quad f_4(x) \overset{\text{def}}{=} \sin\frac{2\pi x}{L},$$
$$\cdots, f_{2k+1}(x) \overset{\text{def}}{=} \cos\frac{k\pi x}{L}, \quad f_{2k+2} \overset{\text{def}}{=} \sin\frac{(k+1)\pi x}{L}, \cdots \tag{9.2}$$

that satisfy

$$\int_{-L}^{L} f_m(x) f_n(x)\,\mathrm{d}x = 0 \qquad \text{(Exercise 1)} \tag{9.3}$$

where m, n are distinct positive integers.

First, we choose some functions $f_{k_1}, f_{k_2}, ..., f_{k_m}$ from (9.2), where $k_1, k_2, ..., k_m$ are distinct positive integers. We want to find constants $c_{k_1}, c_{k_2}, ..., c_{k_m}$ so that the linear combination of these functions, i.e.

$$c_{k_1} f_{k_1} + c_{k_2} f_{k_2} + \cdots + c_{k_m} f_{k_m} \tag{9.4}$$

is as close to the function f over the interval $(-L, L)$ as possible. We use

$$\int_{-L}^{L} \left(f(x) - \sum_{i=1}^{m} c_{k_i} f_{k_i}(x) \right)^2 \mathrm{d}x \tag{9.5}$$

to measure the closeness. We find that the approximation error (9.5) attains its minimum value when

$$c_{k_i} = \frac{1}{\int_{-L}^{L} f_{k_i}^2(x)\,\mathrm{d}x} \int_{-L}^{L} f(x) f_{k_i}(x)\,\mathrm{d}x, \quad i = 1, 2, ..., m \tag{9.6}$$

and the minimum value is

$$\int_{-L}^{L} f^2(x)\,\mathrm{d}x - \sum_{i=1}^{m} \frac{1}{\int_{-L}^{L} f_{k_i}^2(x)\,\mathrm{d}x} \left(\int_{-L}^{L} f(x) f_{k_i}(x)\,\mathrm{d}x \right)^2 \tag{9.7}$$

(Exercise 2)

Example 9.1. *Approximate function $f(x) \stackrel{\text{def}}{=} x$ over the interval $(-\pi, \pi)$, using some standard functions as described above.*

Solution. *In this case, $L = \pi$. So we choose functions from*

$$1, \quad \sin x, \ \cos x, \quad \cos(2x), \quad \sin(2x), \cdots, \cos(kx), \quad \sin(kx), \cdots \tag{9.8}$$

for the approximation.

Suppose we choose functions

$$1, \quad \sin x, \quad \cos x. \tag{9.9}$$

and want to find constants c_1, c_2, c_3, so that the linear combination

$$c_1 + c_2 \sin x + c_3 \cos x \tag{9.10}$$

is closest to function $f(x) = x$ *over the interval* $(-\pi, \pi)$. *The closeness is measured with*

$$\int_{-\pi}^{\pi} (x - (c_1 + c_2 \sin x + c_3 \cos x))^2 \, \mathrm{d}x \tag{9.11}$$

According to the above discussion, the approximation error (9.11) *is smallest when*

$$c_1 = \frac{1}{\int_{-\pi}^{\pi} 1^2 \, \mathrm{d}x} \int_{-\pi}^{\pi} f(x) \cdot 1 \, \mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} x \mathrm{d}x = 0 \tag{9.12}$$

$$c_2 = \frac{1}{\int_{-\pi}^{\pi} \sin^2 x \, \mathrm{d}x} \int_{-\pi}^{\pi} f(x) \sin x \, \mathrm{d}x = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin x \, \mathrm{d}x = 2 \tag{9.13}$$

$$c_3 = \frac{1}{\int_{-\pi}^{\pi} \cos^2 x \, \mathrm{d}x} \int_{-\pi}^{\pi} f(x) \cos x \, \mathrm{d}x = \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos x \, \mathrm{d}x = 0 \tag{9.14}$$

In this case, the linear combination is

$$c_1 + c_2 \sin x + c_3 \cos x = 2 \sin x \tag{9.15}$$

and the approximation error (9.11) *is*

$$\begin{aligned} &\int_{-\pi}^{\pi} (x - (c_1 x + c_2 \sin x + c_3 \cos x))^2 \, \mathrm{d}x \\ &= \int_{-\pi}^{\pi} (x - 2 \sin x)^2 \, \mathrm{d}x = \int_{-\pi}^{\pi} x^2 \, \mathrm{d}x - \frac{1}{\pi} \left(\int_{-\pi}^{\pi} x \sin x \, \mathrm{d}x \right)^2 \\ &= \frac{2\pi^3}{3} - 2^2 \pi \end{aligned} \tag{9.16}$$

Second, we use all the functions from (9.2) for the approximation or representation. As

$$(k_1, k_2, ..., k_m) \rightarrow (1, 2, 3, 4, ...) \tag{9.17}$$

it can be shown that the approximation error (9.7) approches zero provided that

$$\int_{-L}^{L} f^2(x) \, \mathrm{d}x \quad \text{is finite.} \tag{9.18}$$

The linear combination (9.4) with optimal coefficients (9.6) becomes an infinite series

$$c_1 + c_2 \sin\frac{\pi x}{L} + c_3 \cos\frac{\pi x}{L} + c_4 \sin\frac{2\pi x}{L} + c_{2k+1} \cos\frac{k\pi x}{L} + c_{2k+2} \sin\frac{(k+1)\pi x}{L}, \cdots \tag{9.19}$$

with

$$c_k = \frac{1}{\int_{-L}^{L} f_k^2(x)\,\mathrm{d}x} \int_{-L}^{L} f(x) f_k(x)\,\mathrm{d}x, \quad k = 1, 2, ..., \tag{9.20}$$

or

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right) \tag{9.21}$$

with

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\frac{n\pi x}{L}\,\mathrm{d}x \quad (n = 0, 1, 2, ...,) \tag{9.22a}$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\frac{n\pi x}{L}\,\mathrm{d}x \quad (n = 1, 2, ...) \tag{9.22b}$$

We call this series the **Fourier series** of function $f = f(x)$ over the interval $(-L, L)$, and write

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right) \qquad (-L < x < L) \tag{9.23}$$

Remarks

- The triagonometric functions in the Fourier series depend on the interval over which function is to be represented. In Example (9.1), the interval is $(-\pi, \pi)$. So we use triagnometric functions in (9.8) to construct the Fourier series. To represent a function over the interval $(-1, 1)$ (i.e. $L = 1$), we need to use functions from

$$1,\ \sin(\pi x),\ \cos(\pi x),\ \sin(2\pi x),\ \cos(2\pi x) \cdots, \cos(k\pi x),\ \sin(k\pi x), \cdots \tag{9.24}$$

- The Fourier series (9.21) with coefficients (9.22) only represents function $f = f(x)$ over the interval $(-L, L)$, and may fail to do so outside of this interval.

- Since each function in (9.2) is periodic with period $2L$, the Fourier series (9.21) with coefficients (9.22) also represents the periodic extension $\tilde{f}$ of f over the entire interval $(-\infty, \infty)$. Hence it is also called the Fourier series of the periodic function $\tilde{f}$.

- If we change the value of function $f = f(x)$ at only finite many points, then its Fourier series remains the same.

- The Fourier series (9.21) with coefficients (9.22) represents $f = f(x)$ over the interval $(-L, L)$, in the sense

$$\lim_{N\to\infty} \int_{-L}^{L} (S_N(x) - f(x))^2 \,\mathrm{d}x = 0 \tag{9.25}$$

with

$$S_N(x) \overset{\text{def}}{=} \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right) \tag{9.26}$$

. This does not implies that

$$\lim_{N\to\infty} S_N(x) = f(x) \qquad \text{for any } x \in (-L, L) \tag{9.27}$$

Because of this, we do not use $' =' $ in (9.23). Instead, we use $' \sim' $ to mean that the Fourier series written on the right **corresponds** to the function $f(x)$ on the left.

Example 9.2. *Find the Fourier series of function $f(x) = x$ over the interval $(-\pi, \pi)$.*

Solution. *Since $L = \pi$,*

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos(nx)\,\mathrm{d}x = 0, \quad n = 0, 1, 2, ... \tag{9.28}$$

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\sin(nx)\,\mathrm{d}x \\ &= \frac{2}{\pi}\int_0^{\pi} x\mathrm{d}\left(-\frac{1}{n}\cos(nx)\right) = \frac{2}{\pi}\left(\left[-\frac{x}{n}\cos(nx)\right]_0^{\pi} + \int_0^{\pi}\frac{1}{n}\cos(nx)\mathrm{d}x\right) \\ &= -\frac{2}{n}\cos(n\pi) = \frac{2}{n}(-1)^{n+1}, \ n = 1, 2, ... \end{aligned} \tag{9.29}$$

Therefore the Fourier series is

$$\frac{a_0}{2} + \left(\sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right)$$
$$= \sum_{n=1}^{\infty} \frac{2}{n}(-1)^{n+1} \sin(nx) \tag{9.30}$$

In summary, we write

$$x \sim \sum_{n=1}^{\infty} \frac{2}{n}(-1)^{n+1} \sin(nx) \qquad (-\pi < x < \pi) \tag{9.31}$$

or

$$x \sim 2\sin x - \sin(2x) + \frac{2}{3}\sin(3x) + \cdots \qquad (-\pi < x < \pi) \tag{9.32}$$

9.3 Fourier sine series

The Fourier sine series of function $f = f(x)$ over the interval $(0, L)$ is the Fourier series of the odd extension of f_o of f over the interval $(-L, L)$. It is the series (9.21) with coefficients

$$a_n = \frac{1}{L}\int_{-L}^{L} f_o(x) \cos \frac{n\pi x}{L}\,\mathrm{d}x = 0, \quad n = 0, 1, 2, ... \tag{9.33}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f_o(x) \sin \frac{n\pi x}{L}\,\mathrm{d}x = \frac{2}{L}\int_{0}^{L} f_o(x) \sin \frac{n\pi x}{L}\,\mathrm{d}x$$
$$= \frac{2}{L}\int_{0}^{L} f(x) \sin \frac{n\pi x}{L}\,\mathrm{d}x, \quad n = 1, 2, 3, ... \tag{9.34}$$

or the series

$$\sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L} \tag{9.35}$$

with coefficients (9.34). This series represents f_o over the interval $(-L, L)$, hence also represents f over the interval $(0, L)$.

In summary, we write

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L} \qquad (0 < x < L) \tag{9.36}$$

with

$$b_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x, \quad n = 1, 2, 3, ... \tag{9.37}$$

Example 9.3. *Find the Fourier sine series of function $f(x) = x$ over the interval $(0, \pi)$.*

Solution. *An odd extension of f is*

$$f_o(x) = x, \qquad -\pi < x < \pi \tag{9.38}$$

The series is the Fourier series of f_o over the interval $(-\pi, \pi)$, i.e. (9.30), according to Example (9.2).

In summary, we write

$$x \sim \sum_{n=1}^{\infty} \frac{2}{n}(-1)^{n+1}\sin(nx) \qquad (0 < x < \pi) \tag{9.39}$$

or

$$x \sim 2\sin x - \sin(2x) + \frac{2}{3}\sin(3x) + \cdots \qquad (0 < x < \pi) \tag{9.40}$$

9.4 Fourier cosine series

The Fourier cosine series of function $f = f(x)$ over the interval $(0, L)$ is the Fourier series of the even extension f_e of f over the interval $(-L, L)$. It is the series (9.21) with coefficients

$$\begin{aligned} a_n &= \frac{1}{L}\int_{-L}^{L} f_e(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x = \frac{2}{L}\int_0^L f_e(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x \\ &= \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x, \quad n = 0, 1, 2, ... \qquad (9.41) \\ b_n &= \frac{1}{L}\int_{-L}^{L} f_e(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x = 0, \quad n = 1, 2, ... \qquad (9.42) \end{aligned}$$

or the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{L} \tag{9.43}$$

with coefficients (9.41). This series represents f_e over the interval $(-L, L)$, hence also represents f over the interval $(0, L)$.

We write

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{L} \qquad (0 < x < \pi) \tag{9.44}$$

with

$$a_n = \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,\mathrm{d}x, \quad n = 0, 1, 2, ... \tag{9.45}$$

Example 9.4. *Find the Fourier cosine series of function $f(x) = x$ over the interval $(0, \pi)$.*

Solution. *In this case $L = \pi$, so it is the series*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) \tag{9.46}$$

with coefficients

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x\mathrm{d}x = \pi \tag{9.47}$$

$$\begin{aligned} a_n &= \frac{2}{\pi}\int_0^{\pi} x\cos(nx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\mathrm{d}\left(\frac{1}{n}\sin(nx)\right) \\ &= \frac{2}{\pi}\left(\left[\frac{x}{n}\sin(nx)\right]_0^{\pi} - \int_0^{\pi}\frac{1}{n}\sin(nx)\,\mathrm{d}x\right) \\ &= \frac{2}{n^2\pi}[\cos(nx)]_0^{\pi} = \frac{2}{n^2\pi}\left((-1)^n - 1\right), \quad n = 1, 2, ... \end{aligned} \tag{9.48}$$

In summary, we write

$$x \sim \frac{\pi}{2} + \sum_{n=1}^{\infty}\frac{2}{n^2\pi}\left((-1)^n - 1\right)\cos(nx) \qquad (0 < x < \pi) \tag{9.49}$$

or

$$x \sim \frac{\pi}{2} - \frac{4}{\pi}\cos x - \frac{4}{9\pi}\cos(3x) - \frac{4}{25\pi}\cos(5x) - \cdots \quad (0 < x < \pi) \tag{9.50}$$

Exercises

9.1. *Let* $f_1(x), f_2(x), \ldots$ *be functions defined in* (9.2). *Prove that*

$$\int_{-L}^{L} f_m(x) f_n(x)\,\mathrm{d}x = \begin{cases} 0 & m \neq n \\ L & m = n \neq 0 \\ 2L & m = n = 0 \end{cases} \tag{9.51}$$

Here m *and* n *are nonnegative integers.*

9.2. *Assume* $f_1, f_2, \ldots$ *are functions defined in* (9.2). $k_1, k_2, \ldots, k_m$ *are distinct positive integers. Prove that the approximation error*

$$E(c_1, c_2, \ldots, c_m) \stackrel{\text{def}}{=} \int_{-L}^{L} \left(f(x) - \sum_{i=1}^{m} c_{k_i} f_{k_i}(x) \right)^2 \mathrm{d}x \tag{9.52}$$

attains its minimum value when constants

$$c_{k_i} = \frac{1}{F_i} \int_{-L}^{L} f(x) f_{k_i}(x)\,\mathrm{d}x, \quad i = 1, 2, \ldots, m, \tag{9.53}$$

and the minimum value is

$$\int_{-L}^{L} f^2(x)\,\mathrm{d}x - \sum_{i=1}^{m} \frac{1}{F_i} \left(\int_{-L}^{L} f(x) f_{k_i}(x)\,\mathrm{d}x \right)^2 \tag{9.54}$$

where

$$F_i = \int_{-L}^{L} f_{k_i}^2(x)\,\mathrm{d}x, \qquad i = 1, 2, \ldots, m \tag{9.55}$$

Proof. According to (9.51),

$$\begin{aligned} \int_{-L}^{L} \left(\sum_{i=1}^{m} c_{k_i} f_{k_i}(x) \right)^2 \mathrm{d}x &= \int_{-L}^{L} \left(\sum_{i=1}^{m} c_{k_i}^2 f_{k_i}^2(x) + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{m} c_{k_i} c_{k_j} f_{k_i}(x) f_{k_j}(x) \right) \mathrm{d}x \\ &= \sum_{i=1}^{m} c_{k_i}^2 F_i + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{m} c_{k_i} c_{k_j} \int_{-L}^{L} f_{k_i}(x) f_{k_j}(x)\,\mathrm{d}x \\ &= \sum_{i=1}^{m} c_{k_i}^2 F_i \end{aligned} \tag{9.56}$$

Therefore the approximation error (9.52) is

$$\begin{aligned}
E(c_1, c_2, ..., c_m) &= \int_{-L}^{L} \left(f^2(x) - 2f(x) \sum_{i=1}^{m} c_{k_i} f_{k_i}(x) + \left(\sum_{i=1}^{m} c_{k_i} f_{k_i} \right)^2 \right) \mathrm{d}x \\
&= \int_{-L}^{L} f^2(x) \,\mathrm{d}x - \sum_{i=1}^{m} 2c_{k_i} \int_{-L}^{L} f(x) f_{k_i}(x) \,\mathrm{d}x + \sum_{i=1}^{m} c_{k_i}^2 F_i \\
&= \int_{0}^{L} f^2(x) \,\mathrm{d}x + \sum_{i=1}^{m} \frac{1}{F_i} (c_{k_i} - A_i)^2 - \sum_{i=1}^{m} A_i^2 F_i \qquad (9.57)
\end{aligned}$$

where

$$A_i = \frac{1}{F_i} \int_{-L}^{L} f(x) f_{k_i}(x) \,\mathrm{d}x \qquad (9.58)$$

Therefore the approximation error E attains its minimum value when

$$c_{k_i} = A_i = \frac{1}{F_i} \int_{-L}^{L} f(x) f_{k_i}(x) \,\mathrm{d}x, \qquad i = 1, 2, ..., m \qquad (9.59)$$

and the minimum value is

$$\int_{-L}^{L} f^2(x) \,\mathrm{d}x - \sum_{i=1}^{m} A_i^2 F_i = \int_{-L}^{L} f^2(x) \,\mathrm{d}x - \sum_{i=1}^{m} \frac{1}{F_i} \left(\int_{-L}^{L} f(x) f_{k_i}(x) \,\mathrm{d}x \right)^2 \qquad (9.60)$$

□

9.3. *Find the Fourier series of constant function $f(x) = 1$ over the interval $(-3, 3)$.*

9.4. *Find the Fourier series of function*

$$f(x) = \begin{cases} 1 & -2 < x < 2 \\ 0 & -\pi < x \leq -2, \ \textit{or}\ 2 \leq x < \pi \end{cases} \qquad (9.61)$$

over the interval $(-\pi, \pi)$.

9.5. *Find the Fourier sine series of constant function $f(x) = 1$ over the interval $(0, \pi)$.*

9.6. *Find the Fourier cosine series of constant function* $f(x) = 1$ *over the interval* $(0, \pi)$.

9.7. *Prove that the fourier series of function* $f = f(x)$ *over the interval* $(-L, L)$ *could also be represented as*

$$\sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L} \tag{9.62}$$

with coefficients

$$c_n = \frac{1}{2L} \int_{-L}^{L} f(x) e^{-in\pi x/L} \, \mathrm{d}x \qquad n = 0, \pm 1, \pm 2, ... \tag{9.63}$$

Here i *is a complex number and* $i^2 = -1$.

9.8. *Find the solution of the following IBVP of the 1D wave equation*

$$u_{tt} = 3^2 u_{xx} \qquad 0 < x < 1, t > 0 \tag{9.64a}$$
$$u(x, 0) = x(1 - x) \qquad 0 \leq x \leq 1 \tag{9.64b}$$
$$u_t(x, 0) = 0 \qquad 0 \leq x \leq 1 \tag{9.64c}$$
$$u(0, t) = 0 \qquad t \geq 0 \tag{9.64d}$$
$$u(1, t) = 0 \qquad t \geq 0 \tag{9.64e}$$

Chapter 10

3D PDEs

By 3D PDEs, we mean those PDEs that are associated with an unknown function that has three independent spatial variables i.e. x, y, z. Most of the PDEs in previous chapters are associated with an unknown function that has only one independent spatial variable x. These PDEs are therefore called 1D PDEs.

10.1 Advection Equation

Advection equation is associated with advection process. The advection process refers to the transport of a property of a fluid, such as some chemical, by the bulk motion of the fluid itself.

Suppose the chemical is transported in space with **concentration** u and flux density $\boldsymbol{\phi}$. $u(P, t)$ is the concentration of the pollutant at point P and at time t. $\boldsymbol{\phi}(P, t)$ is a vector that points in the direction along which the chemical is transported at point P and at time t. The magnitude of this vector is the amount (measured in mass) of chemical moving in the direction of this vector, per unit area and unit time.

If the chemical is neither generated nor absorbed at any time and location during the transportation, then conservation of chemical yields (Exercise 1)

$$u_t = -\nabla \cdot \boldsymbol{\phi} \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.1}$$

If the chemical moves in the direction of a constant unit vector $\boldsymbol{d}$, with speed v, then we have

$$\boldsymbol{\phi} = vu\boldsymbol{d} \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.2}$$

Combing equation (10.1) with equation (10.2), we obtain

$$u_t = -v\nabla u \cdot \boldsymbol{d} \qquad \text{in } \mathbb{R}^3, t > 0. \tag{10.3}$$

The PDE (10.3) is called the **3D advection equation**. This is a 1st order PDE and associated with an unknown function

$$u = u(x, y, z, t) \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.4}$$

This PDE says that at any point $P \in \mathbb{R}^3$ and at any time $t > 0$, u_t is always equal to, minus v times the dot product of $\boldsymbol{d}$ and the gradient of u at the same point and same time.

An initial value problem (IVP) of the advection equation (10.3) is

$$u_t = -v\nabla u \cdot \boldsymbol{d} \qquad \text{in } \mathbb{R}^3, t > 0. \tag{10.5a}$$

$$u|_{t=0} = f \qquad \text{in } \mathbb{R}^3 \tag{10.5b}$$

The condition (10.5b) means that: at time $t = 0$, the value of u at any point P is equal to the value of function f at the same point. Given that function f is already known, the goal of this IVP (10.5) is to find a function (10.4) that satisfies both the advection equation (10.5a) and the initial condition (10.5b).

The physical interpretation of this problem is as follows: we aim to determine the concentration of the chemical at any time and location, provided that we know

- the initial distribution of the chemical concentration (i.e. the initial condition (10.5b)
- how the concentration of the chemical changes over time (i.e. the differential equation (10.5a)

Solution. *(of the IVP* (10.5)*)*
According to the transportation rule (10.2)*, the chemical at any point P at time $t = 0$ is transported to the point P^* with*

$$\overrightarrow{OP^*} = \overrightarrow{OP} + vt^*\boldsymbol{d} \tag{10.6}$$

by the time $t = t^* > 0$. Therefore,

$$u(P^*, t = t^*) = u(P, t = 0) = f(P) \tag{10.7}$$

In other words, the value of u at any point P^* and any time $t^* > 0$ is

$$u(P^*, t = t^*) = f(P) \tag{10.8}$$

with P satisfying (10.6). This means that the function

$$u(x, y, z, t) = f(x - vd_1t, y - vd_2t, z - vd_3t) \qquad (x, y, z) \in \mathbb{R}^3, t \geq 0 \tag{10.9}$$

satisfies all the conditions in the IVP (10.5) and hence is its solution. Here d_1, d_2, d_3 are Cartesian components of $\boldsymbol{d}$, i.e. $\boldsymbol{d} = (d_1, d_2, d_3)$.

Example 10.1. *The PDE*

$$u_t = -(2u_x - 3u_y + 4u_z) \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.10}$$

with an unknown function

$$u(x, y, z, t), \quad -\infty < x, y, z < \infty, t \geq 0 \tag{10.11}$$

is an example of 3D advection (10.3) with

$$\boldsymbol{d} = \frac{1}{\sqrt{2^2 + 3^2 + 4^2}}(2\hat{\boldsymbol{x}} - 3\hat{\boldsymbol{y}} + 4\hat{\boldsymbol{z}}) \tag{10.12}$$

and $v = \sqrt{2^2 + 3^2 + 4^2} = \sqrt{29}$.

Therefore, if PDE (10.10) models the advection or transportation of something in three-dimensional space, then the transportation occurs in the direction of the vector $\boldsymbol{d}$ in (10.12), and has speed $\sqrt{2^2 + 3^2 + 4^2} = \sqrt{29}$.

10.2 Diffusion Equation

The diffusion equation is associated with diffusion phenomena. Like advection, diffusion is also a type of transportation. Assume something e.g. heat energy is transported in space with concentration u and flux density $\boldsymbol{\phi}$ as defined in Section 10.1. The concentration u and flux density $\boldsymbol{\phi}$ are also related by conservation law (10.1), provided that heat energy is neither generated nor absorbed at any place and at any time.

In diffusion, however, heat flux goes from regions of high concentration to the regions of lower concentration, at a rate (per unit area and unit time) proportional to the concentration gradient. Mathematically, this means

$$\boldsymbol{\phi} = -c\nabla u \qquad \text{in} \mathbb{R}^3, t > 0 \tag{10.13}$$

with c a positive constant. Combing (10.1) with (10.13), we get

$$u_t = c\nabla^2 u \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.14}$$

This PDE is called the **3D diffusion equation**. It is a 2nd order PDE and also associated with an unknown function u in (10.4). This PDE says that at any point $P \in \mathbb{R}^3$ and at any time t, u_t is always equal to c times the Laplacian of u at the same point and same time.

In practice, the diffusion equation is often defined over a finite region Ω. An IBVP for this kind of diffusion equation is

$$u_t = c\nabla^2 u \qquad \text{in } \Omega, t > 0 \tag{10.15a}$$
$$u|_{t=0} = g \qquad \text{on } \Omega \tag{10.15b}$$
$$u = f \qquad \text{on } \partial\Omega, t > 0 \tag{10.15c}$$

Here $\partial\Omega$ is the boundary of the region Ω. The (10.15b) is called the initial condition. It says that: at time $t = 0$, the value of u at any point $P \in \Omega$, is equal to the value of function g at the same point. The (10.15c) is called the boundary condition. It says that: at any time t, the value of u at any point P on the boundary, is equal to the value of f at the same point and the same time. Given that function f and g are already known, the goal of IBVP (10.15) is to find a function

$$u = (x, y, z, t) \qquad (x, y, z) \in \Omega, t \geq 0 \tag{10.16}$$

that satisfies all the conditions in (10.15).

The physical interpretation of the IBVP (10.15) is as follows: we aim to determine the temperature at any location in the region Ω at any time, provided that we know

- how the temperature changes over time (i.e. the differential equation (10.15a))

- the initial temperature distribution over the region Ω (i.e. the initial condition (10.15b))

- the temperature on the boundary of the region at all times (i.e. the boundary condition (10.15c))

10.3 Laplace/Poisson Equation

The Laplace and Poisson equations are associated with equilibrium phenomena or time-independent problems. They could be interpreted as characterizing the equilibrium temperature distribution as in Chapter 7. Here we consider its application in electrostatics.

Assume a region Ω is occupied by stationary charges with charge density ρ. The resulting electric field $\boldsymbol{E}$ has to satisfy

$$\nabla \cdot \boldsymbol{E} = \rho/\epsilon_0 \qquad \text{in } \Omega \tag{10.17}$$

according to Gauss's law. Here ϵ_0 is some physical constant. In terms of electrical potential

$$\boldsymbol{E} = -\nabla u \qquad \text{in } \Omega \tag{10.18}$$

the relation (10.17) becomes

$$\nabla^2 u = f \qquad \text{in } \Omega \tag{10.19}$$

with $f = -\rho/\epsilon_0$ everywhere in Ω.

The PDE (10.19) specifies the Laplacian of u at every point in Ω. It is a 2nd order PDE that is associated with an unknown function

$$u = u(x, y, z) \qquad \text{on } \Omega. \tag{10.20}$$

The PDE (10.19) is called the **3D Poisson equation** if function f is not identically equal to 0. If f is identically equal to zero, then the PDE (10.19) becomes

$$\nabla^2 u = 0 \qquad \text{in } \Omega \tag{10.21}$$

and is called the **3D Laplace Equation**.

A BVP (boundary value problem) for the Poisson equation is

$$\nabla^2 u = f \qquad \text{in } \Omega \tag{10.22a}$$

$$u = g \qquad \text{on } \partial\Omega \tag{10.22b}$$

Given that functions f and g are already known, the goal of this BVP is to find a function

$$u = u(x, y, z) \qquad \text{on } \Omega \tag{10.23}$$

that satisfies all the conditions in (10.22).

It can be shown that there is at most one solution of the BVP (10.22), due to the property of Laplacian operator.

Example 10.2. *Assume stationary charges are uniformly distributed throughout a ball, denoted as Ω. This ball has a radius of 4 units and is positioned at the origin O of a coordinate system. Your task is to compute the resulting electric field (symbolized by the vector $\boldsymbol{E}$) and the electrical potential (denoted as u).*

Proof. According to the discussion above, the electrical potential u satisfies the Poisson equation

$$\nabla^2 u = -\rho/\epsilon_0 \quad \text{in } \Omega \tag{10.24}$$

Here ρ is the charge density in Ω and ϵ_0 is a physical constant. Since ρ is constant throughout Ω and Ω is a ball centered at the origin, the potential u at any point P must be independent of the spherical coordinates θ and ϕ of P. Because of this, the Laplacian in spherical coordinates (formula (10.25) is available in the Final exam paper)

$$\nabla^2 u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2} \tag{10.25}$$

is simplified to be

$$\nabla^2 u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right). \tag{10.26}$$

Then the Poisson equation (10.24) becomes

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}u}{\mathrm{d}r}\right) = -\frac{\rho}{\epsilon_0} \qquad 0 < r < 4. \tag{10.27}$$

or equivalently

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}u}{\mathrm{d}r}\right) = -\frac{\rho r^2}{\epsilon_0} \qquad 0 < r < 4. \tag{10.28}$$

The ODE (10.28) means

$$r^2\frac{\mathrm{d}u}{\mathrm{d}r} = -\int\frac{\rho r^2}{\epsilon_0}\,\mathrm{d}r = -\frac{\rho r^3}{3\epsilon_0} + C_1 \quad 0 < r < 4 \tag{10.29}$$

where C_1 is some constant. So

$$\frac{\mathrm{d}u}{\mathrm{d}r} = -\frac{\rho r}{3\epsilon_0} + \frac{C_1}{r^2} \quad 0 < r < 4 \tag{10.30}$$

which implies that

$$u(r) = \int \left(-\frac{\rho r}{3\epsilon_0} + \frac{C_1}{r^2} \right) \mathrm{d}r = -\frac{\rho r^2}{6\epsilon_0} - \frac{C_1}{r} + C_2 \qquad 0 \le r \le 4 \tag{10.31}$$

Here C_2 is also some constant.

We claim that $C_1 = 0$ in (10.31). Otherwise, the potential u at the center of the ball would become infinite. When $C_1 = 0$, the expression (10.31) becomes

$$u(r) = -\frac{\rho r^2}{6\epsilon_0} + C_2 \qquad 0 \le r \le 4 \tag{10.32}$$

This formula gives the potential distribution throughout the region. It says that the potential at a point P is $-r^2/(6\epsilon_0) + C_2$, if the distance from P to the origin O is r.

From (10.32), we know

$$\boldsymbol{E}(r) = -\nabla u = -\frac{\mathrm{d}u}{\mathrm{d}r}\hat{\boldsymbol{r}} = \frac{r\rho}{3\epsilon_0}\hat{\boldsymbol{r}} \qquad 0 \le r < 4 \tag{10.33}$$

The formula (10.33) gives the electric field that is generated by the stationary charge distribution. It says that the electric field vector at a point P is in the direction of $\overrightarrow{OP}$, and has magnitude $|\overrightarrow{OP}|\rho/(3\epsilon_0)$. □

10.4 Wave equation

The wave equation is associated with wave phenomena. It occurs for example in describing the electromagnetic waves. In a vacuum region Ω, the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ are related by the following famous Maxwell's equations:

$$\nabla \cdot \boldsymbol{E} = 0 \qquad \text{in } \Omega \tag{10.34a}$$

$$\nabla \cdot \boldsymbol{B} = 0 \qquad \text{in } \Omega \tag{10.34b}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad \text{in } \Omega \tag{10.34c}$$

$$\nabla \times \boldsymbol{B} = \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} \qquad \text{in } \Omega \tag{10.34d}$$

Here c is the speed of light. The equations (10.34) imply

$$\frac{\partial^2 E_i}{\partial t^2} = c^2 \nabla^2 E_i \qquad \text{in } \Omega,\ i = 1, 2, 3 \tag{10.35a}$$

$$\frac{\partial^2 B_i}{\partial t^2} = c^2 \nabla^2 B_i \qquad \text{in } \Omega,\ i = 1, 2, 3 \tag{10.35b}$$

Here $E_i, i = 1, 2, 3$ are the Cartesian components of $\boldsymbol{E}$, i.e. $\boldsymbol{E} = (E_1, E_2, E_3)$. $B_i, i = 1, 2, 3$ are the Cartesian components of $\boldsymbol{B}$, i.e. $\boldsymbol{B} = (B_1, B_2, B_3)$.

Proof. The curl of the left side of (10.34c) is

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} = -\nabla^2 \boldsymbol{E} \tag{10.36}$$

The first identity in (10.36) is actually the identity (3.104e) in Chapter 3. Here

$$\nabla^2 \boldsymbol{E} = \left(\nabla^2 E_1, \nabla^2 E_2, \nabla^2 E_3\right) \tag{10.37}$$

by definition. The second identity in (10.36) is because of (10.34a).

The curl of the right side of (10.34c) is

$$\nabla \times \left(-\frac{\partial \boldsymbol{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla \times \boldsymbol{B}) = -\frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} \tag{10.38}$$

Here

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \left(\frac{\partial^2 E_1}{\partial t^2}, \frac{\partial^2 E_2}{\partial t^2}, \frac{\partial^2 E_3}{\partial t^2}\right) \tag{10.39}$$

by definition.

The curl of the two sides of (10.34c) should be the same. So we obtain (10.35) based on (10.36) and (10.38).

Similarly, we could obtain identity (10.35b) by taking the curl on both sides of (10.34d). $\square$

The equations (10.35) mean that each Cartesian components of $\boldsymbol{E}$ and $\boldsymbol{B}$ satisfies the following equation

$$u_{tt} = v^2 \nabla^2 u, \qquad \text{in } \Omega, t > 0, \tag{10.40}$$

with $v = c$. In other words, if $u = E_i$ or $B_i, i = 1, 2, 3$, then identity (10.40) holds true.

The PDE (10.40) is called 3D wave equation. It is a 2nd order PDE and associated with an unknown function $u = u(x, y, z, t)$. This PDE says that at any point $P \in \Omega$ and at any time $t > 0$, u_{tt} is always equal to v^2 times the Laplacian of u at the same point and same time.

Example 10.3. *A pair of electric field* $\boldsymbol{E}$ *and magnetic field* $\boldsymbol{B}$ *that satisfy the Maxwell's equations* (10.34) *is an example of electromagnetic wave propagating in a vacuum* Ω. *Assume the electric field vector at any point* $P \in \Omega$ *and any time* $t \geq 0$ *is*

$$\boldsymbol{E}(P,t) = \hat{\boldsymbol{z}} E_0 \sin(ky - vt) \tag{10.41}$$

Here y *is the* y*-coordinate of* P. *The* E_0, v *and* k *are constants.*

- *What are the requirements on the constants* E_0, k *and* v, *so that the* $\boldsymbol{E}$ *in* (10.41) *satisfies the 3D wave equation* (10.35a) *?*
- *Assume* $E_0 = 2, k = 1$ *and* $v = c$. *Then sketch the electric field* $\boldsymbol{E}$ *at time* $t = 0$ *and* $t = 1$, *respectively.*
- *Find a magnetic field* $\boldsymbol{B}$, *so that this* $\boldsymbol{B}$ *and the* $\boldsymbol{E}$ *in* (10.41) *satisfy the Maxwell's equations* (10.34).

Solution.
(i) We require

$$\frac{v}{k} = \pm c \tag{10.42}$$

in order for the $\boldsymbol{E}$ *in* (10.41) *to satisfy the 3D wave equation* (10.35a).

(ii). Figure 10.1 shows a sketch of a portion of the vector field $\boldsymbol{E}$ *in* (10.41) *when* $E_0 = 2, k = 1$ *and* $v = c$.

(iii). From (10.34c), *we know*

$$\frac{\partial \boldsymbol{B}}{\partial t} = -\nabla \times \boldsymbol{E} = -E_0 k \cos(ky - vt)\hat{\boldsymbol{x}} \tag{10.43}$$

So the magnetic field vector at point P *and time* t *is*

$$\boldsymbol{B}(P,t) = B_1 \hat{\boldsymbol{x}} \tag{10.44}$$

with

$$B_1 = -\int E_0 k \cos(ky - vt)\,\mathrm{d}t = \frac{E_0 k}{v}\sin(ky - vt) + \phi(y) \tag{10.45}$$

Here $\phi(y)$ *is some function of* y. *Then*

$$\nabla \times \boldsymbol{B} = -\left(\frac{E_0 k^2}{v}\cos(ky - vt) + \phi'(y)\right)\hat{\boldsymbol{z}} \tag{10.46}$$

Comparing (10.46) *with*

$$\frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} = -\frac{E_0 v}{c^2}\cos(ky - vt)\hat{\boldsymbol{z}} = -\frac{E_0 k^2}{v}\cos(ky - vt)\hat{\boldsymbol{z}} \tag{10.47}$$

we find that the $\boldsymbol{E}$ *in* (10.41) *and the* $\boldsymbol{B}$ *in* (10.46) *satisfy the Ampere's law* (10.34d), *provided that* $\phi(y) \equiv C$ *(or* $\phi'(y) = 0$ *for all possible values of* y*).*

Apparently both the $\boldsymbol{E}$ *in* (10.41) *and the* $\boldsymbol{B}$ *in* (10.44) *are divergence free in* Ω*. (Check this yourself)*

Overall, the magnetic field defined as

$$\boldsymbol{B}(P,t) = \left(\frac{E_0 k}{v}\sin(ky - vt) + C\right)\hat{\boldsymbol{x}} \quad \textit{for any } P \in \Omega \textit{ and time } t \geq 0 \tag{10.48}$$

is the required magnetic field.

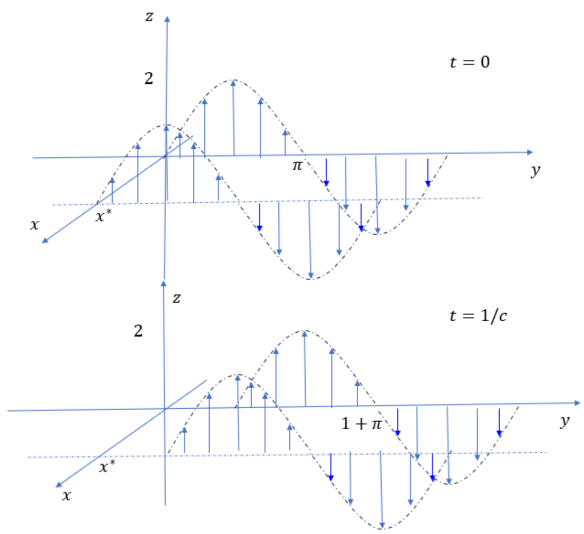

Figure 10.1: A sketch of a portion of the vector fields $\boldsymbol{E}$ in (10.41) when $E_0 = 2, k = 1, v = c$.

Exercises

10.1. *Assume chemical is transported in space with concentration* u *and flux density* $\boldsymbol{\phi}$ *as defined in Section 10.1. Prove that its concentration* u *and flux* *density* $\boldsymbol{\phi}$ *are related by*

$$u_t = -\nabla \cdot \boldsymbol{\phi} \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.49}$$

provided that the chemical is neither generated nor absorbed at any location and at any time.

10.2. *The PDEs discussed in this chapter can be considered as generalization of their 1D versions in the previous chapters. In some scenarios,*

- *the 3D advection equation* (10.3) *can be simplified to the 1D advection equation* (5.2)*,*
- *the 3D diffusion equation* (10.14) *can be simplified to the 1D diffusion equation* (6.2)*,*
- *the 3D Poisson equation* (10.19) *can be simplified to the 1D Poisson equation* (7.1)*,*
- *the 3D Wave equation* (10.40) *can be simplified to the 1D wave equation* (8.1)*.*

Can you give more details on the process of these simplifications ?

10.3. *Consider the PDE*

$$u_t = u_x + u_y + u_z \qquad \text{in } \mathbb{R}^3, t > 0 \tag{10.50}$$

that is associated with an unknown function

$$u = u(x, y, z, t) \qquad -\infty < x, y, z < \infty, t \geq 0. \tag{10.51}$$

Give a physical interpretation of this PDE. You could start by writing the PDE in the form (10.3) *with some suitable constant number* v *and constant vector* $\boldsymbol{d}$*.*

10.4. *(Faraday Cage) Let us consider a perfect conductor that has been shaped to enclose an empty region, denoted as* Ω*. The conductor is uniformly charged, with voltage of 3 units, and its centered at the origin of a coordinate system. Can you determine the voltage at the very center of this spherical region* Ω *? Moreover, would your answer change if the region* Ω *was a cube or had a different shape ? Please explain with your reasoning.*

10.5. *Consider an infinitely long annular cylinder, which has stationary electric charge uniformly distributed over it. The inner and outer radii of the annular cylinder are 1 and 2 units, respectively. The voltages on the inner and outer surface of the annular cylinder are set to 0 and 2 units, respectively. Based on this information, can you determine the distribution of electric potential throughout this annular cylinder ?*

10.6. *In addition to the pair of* $(\boldsymbol{E}, \boldsymbol{B})$ *in Example* (10.3)*, we look for other examples of electromagnetic waves. We start by considering the following questions*

- *Assume both* $\boldsymbol{E}$ *and* $\boldsymbol{B}$ *are constant vector fields. Both are time-independent. Then could the pair of* $(\boldsymbol{E}, \boldsymbol{B})$ *possibly satisfy the Maxwell's equation* (10.34) *?*

- *Assume one of* $\boldsymbol{E}$ *and* $\boldsymbol{B}$ *is time-independent, the other is time-dependent. Then, could the pair of* $(\boldsymbol{E}, \boldsymbol{B})$ *possibly satisfy the Maxwell's equation* (10.34) *?*

- *Find a suitable* $\boldsymbol{B}$ *such that the pair* $(\boldsymbol{E}, \boldsymbol{B})$ *satisfy the Maxwell's equation* (10.34)*, when* $\boldsymbol{E}$ *is defined as*

$$\boldsymbol{E}(P,t) = \hat{\boldsymbol{z}} E_0 \sin(kx - vt) \tag{10.52}$$

or

$$\boldsymbol{E}(P,t) = \hat{\boldsymbol{z}} E_0 \sin(kz - vt) \tag{10.53}$$

or

$$\boldsymbol{E}(P,t) = \hat{\boldsymbol{z}} E_0 \cos(ky - vt) \tag{10.54}$$

for any point $P \in \Omega$ *and time* $t \geq 0$*.*

Appendix A

Extension of a function

An **odd extension** of a function f defined on $(0, L)$ is a function f_o defined on $(-L, L)$ such that

$$f_o(x) = \begin{cases} f(x) & 0 < x < L \\ -f(-x) & -L < x < 0 \end{cases} \tag{A.1}$$

An **even extension** of a function f defined on $(0, L)$ is a function f_e defined on $(-L, L)$ such that

$$f_e(x) = \begin{cases} f(x) & 0 < x < L \\ f(-x) & -L < x < 0 \end{cases} \tag{A.2}$$

A **periodic extension** of a function f defined on (a, b) is the function $\tilde{f}$ defined everwhere such that

$$\tilde{f}(x) = \begin{cases} f(x) & a < x < b \\ f(x) & a + k(b-a) < x < b + k(b-a), k = 1, 2, ... \end{cases} \tag{A.3}$$

A **periodic odd extension** of a function f defined on $(0, L)$ is the periodic extension $\tilde{f}_o$ of the odd extension f_o of f.
A **periodic even extension** of a function f defined on $(0, L)$ is the periodic extension $\tilde{f}_e$ of the even extension f_e of f.