LECTURE "HIGHER MATHEMATICS"

PART I

Vector Calculus

1 Spatial vectors

A vector $\mathbf{a}$ in physical space is an arrow, pointing from some point P to another point Q. The **magnitude** (length) $|\mathbf{a}| \geq 0$ of $\mathbf{a}$ is defined as the linear distance between P and Q. Generally, a vector $\mathbf{a}$ is fixed by its magnitude and its direction. As an arrow is shifted in space, without changing magnitude and direction, it still describes the same vector.

1.1 Vector algebra

As $\mathbf{a}$ points from P to Q and $\mathbf{b}$ from Q to R, **vector addition** is defined by

$$\mathbf{c} \;=\; \mathbf{a} \;+\; \mathbf{b}, \tag{1}$$

where $\mathbf{c}$ is the vector pointing from P to R.
Multiplication of $\mathbf{a}$ with a real number $\lambda \in \mathbb{R}$ shall, by definition, yield another vector

$$\mathbf{a}' \;=\; \lambda\mathbf{a} \tag{2}$$

with length $|\mathbf{a}'| = |\lambda||\mathbf{a}|$ and direction parallel ($\lambda > 0$) ore opposite ($\lambda < 0$) to $\mathbf{a}$.
For $\lambda = -1$, we also write

$$\lambda\mathbf{a} \;\equiv\; (-1)\mathbf{a} \;=\; -\mathbf{a}, \qquad \mathbf{a} \;+\; (-1)\mathbf{b} \;=\; \mathbf{a} \;-\mathbf{b}. \tag{3}$$

We choose three mutually orthogonal directions "1", "2", and "3" (which shall form a right-handed system). Then, a typical vector is written as

$$\mathbf{a} \;=\; \begin{pmatrix} 2\,\mathrm{m} \\ -5\,\mathrm{m} \\ 4\,\mathrm{m} \end{pmatrix} \;=\; \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \;=\; \underbrace{a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{=\,\mathbf{a}_1} \;+\; \underbrace{a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{=\,\mathbf{a}_2} \;+\; \underbrace{a_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{=\,\mathbf{a}_3}, \tag{4}$$

where $\mathbf{a}_1$ is the vector with length $a_1 = 2\,\mathrm{m}$ pointing in direction 1,
$\mathbf{a}_2$ is the vector with length $-a_2 = 5\,\mathrm{m}$ pointing opposite to direction 2,
$\mathbf{a}_3$ is the vector with length $a_3 = 4\,\mathrm{m}$ pointing in direction 3.
The dimensionless objects

$$\mathbf{u}_1 \;=\; \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 \;=\; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{5}$$

with $|\mathbf{u}_n| = 1$, are called **unit vectors**,

$$\mathbf{a} \;=\; a_1\mathbf{u}_1 \;+\; a_2\mathbf{u}_2 \;+\; a_3\mathbf{u}_3. \tag{6}$$

The **numbers** a_n are called the **coordinates** of the vector $\mathbf{a}$. Theorem of Pythagoras:

$$|\mathbf{a}| \;=\; \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{7}$$

The **vectors** $\mathbf{a}_n$ are called the **components** of the vector $\mathbf{a}$.

Vectors may have **units other than length**. While the a_n are carrying the physical units of the vectorial quantity $\mathbf{a}$, the $\mathbf{u}_n$ are "dimensionless". E.g., a **force vector** is

$$\begin{aligned}
\mathbf{F} &= F_1\,\mathbf{u}_1 \;+\; F_2\,\mathbf{u}_2 \;+\; F_3\,\mathbf{u}_3 \\
&= (2.13\,\mathrm{N})\,\mathbf{u}_1 \quad + \quad (-1.95\,\mathrm{N})\,\mathbf{u}_2 \quad + \quad (4.79\,\mathrm{N})\,\mathbf{u}_3 \\
&= (2.13\,\mathrm{N})\begin{pmatrix}1\\0\\0\end{pmatrix} + (-1.95\,\mathrm{N})\begin{pmatrix}0\\1\\0\end{pmatrix} + (4.79\,\mathrm{N})\begin{pmatrix}0\\0\\1\end{pmatrix} \;=\; \begin{pmatrix}2.13\,\mathrm{N}\\-1.95\,\mathrm{N}\\4.79\,\mathrm{N}\end{pmatrix}. \qquad (8)
\end{aligned}$$

Its **magnitude** is $|\mathbf{F}| \;=\; \sqrt{F_1^2+F_2^2+F_3^2} \;\approx\; 5.59\,\mathrm{N}$.

1.2 Scalar (dot) and vector (cross) products

The **scalar product** (or **dot product**) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$\mathbf{a}\cdot\mathbf{b} \;:=\; |\mathbf{a}||\mathbf{b}|\cos\gamma \qquad \text{(a number !)}, \qquad (9)$$

where γ is the **angle** between $\mathbf{a}$ and $\mathbf{b}$.
Consequently, $\mathbf{b}\cdot\mathbf{a} = \mathbf{a}\cdot\mathbf{b}$ and, with some $\lambda\in\mathbb{R}$, $(\lambda\mathbf{a})\cdot\mathbf{b} = \lambda(\mathbf{a}\cdot\mathbf{b})$.
Furthermore, $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}$ (see problem 1.1), and thus

$$\begin{aligned}
\mathbf{a}\cdot\mathbf{b} &\equiv (a_1\mathbf{u}_1+a_2\mathbf{u}_2+a_3\mathbf{u}_3)\cdot(b_1\mathbf{u}_1+b_2\mathbf{u}_2+b_3\mathbf{u}_3) \\
&\equiv \Big(\sum_{i=1}^{3} a_i\mathbf{u}_i\Big)\cdot\Big(\sum_{j=1}^{3} b_j\mathbf{u}_j\Big) \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3} a_ib_j\,\underbrace{(\mathbf{u}_i\cdot\mathbf{u}_j)}_{\delta_{ij}} \;=\; \sum_{i=1}^{3} a_ib_i \;\equiv\; a_1b_1+a_2b_2+a_3b_3. \qquad (10)
\end{aligned}$$

Therefore, in terms of the coordinates,

$$\mathbf{a}\cdot\mathbf{b} \;\equiv\; \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\cdot\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix} \;=\; a_1b_1+a_2b_2+a_3b_3. \qquad (11)$$

Let $\mathbf{u}_\perp$ be the unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, such that $\{\mathbf{a},\mathbf{b},\mathbf{u}_\perp\}$ is right-handed. The **vector product** (or **cross product**) of $\mathbf{a}$ with $\mathbf{b}$ is defined as

$$\mathbf{a}\times\mathbf{b} \;:=\; \Big(|\mathbf{a}||\mathbf{b}|\sin\gamma\Big)\mathbf{u}_\perp \qquad \text{(a vector !)}. \qquad (12)$$

Since $\mathbf{b}\times\mathbf{a} = -\mathbf{a}\times\mathbf{b}$ and $(\lambda\mathbf{a})\times\mathbf{b} = \lambda(\mathbf{a}\times\mathbf{b})$, etc. (exercises), we now have

$$\mathbf{a}\times\mathbf{b} \;\equiv\; ... \;=\; \big(a_2b_3-a_3b_2\big)\,\mathbf{u}_1 \;+\; \big(a_3b_1-a_1b_3\big)\,\mathbf{u}_2 \;+\; \big(a_1b_2-a_2b_1\big)\,\mathbf{u}_3. \qquad (13)$$

Therefore, in terms of the coordinates,

$$\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\times\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix} \;=\; \begin{pmatrix}a_2b_3-a_3b_2\\a_3b_1-a_1b_3\\a_1b_2-a_2b_1\end{pmatrix}. \qquad (14)$$

2 Curves in space: Vector functions of one argument

2.1 Orbit of a moving particle as an example

Selecting one point O in space as origin, any other point P is fixed by its **position vector**

$$\mathbf{r} \;=\; \vec{OP} \;=\; x\mathbf{u}_1 + y\mathbf{u}_2 + z\mathbf{u}_3. \tag{15}$$

The position vector of a **moving particle** has time-dependet coordinates x, y, and z,

$$\mathbf{r}(t) \;=\; \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{16}$$

This is a **vector function**. As t grows, the tip of the vector $\mathbf{r}(t)$ "draws" the **orbit** of the particle which is a **curve in space**.

Example 1a: Motion on a helical (helix shaped) orbit is described by the function

$$\mathbf{r}(t) = \begin{pmatrix} R\cos(\omega t) \\ R\sin(\omega t) \\ ut \end{pmatrix}. \tag{17}$$

2.2 Derivative of vector functions: velocity vector

The **derivative** of a vector function $\mathbf{r}(t)$ is another vector function $\dot{\mathbf{r}}(t)$, defined as

$$\dot{\mathbf{r}}(t) \;\equiv\; \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r}(t) \;:=\; \lim_{\Delta t\to 0}\frac{\mathbf{r}(t+\Delta t)-\mathbf{r}(t)}{\Delta t} \tag{18}$$

$$\equiv\; \lim_{\Delta t\to 0}\frac{1}{\Delta t}\begin{pmatrix} x(t+\Delta t)-x(t) \\ y(t+\Delta t)-y(t) \\ z(t+\Delta t)-z(t) \end{pmatrix} \;\equiv\; \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix}. \tag{19}$$

The vectors $\mathbf{r}(t)$ and $\mathbf{r}(t+\Delta t)$ (blue) and their difference (green) are shown in Fig. 1.

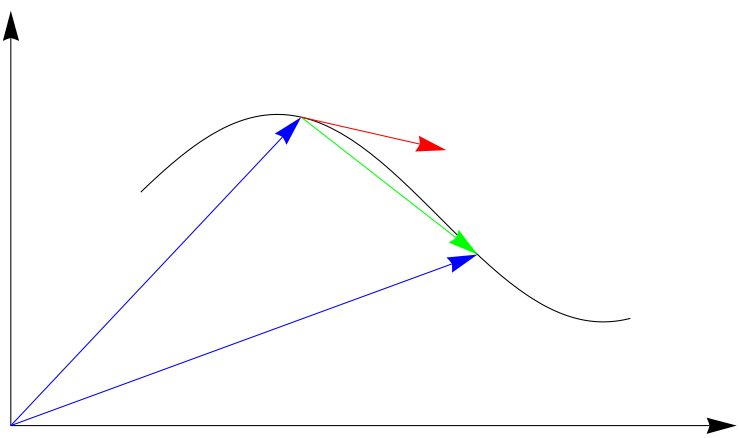


Figure 1: .

At any given time t, the **magnitude** of the vector $\dot{\mathbf{r}}(t)$ is

$$\begin{aligned} |\dot{\mathbf{r}}(t)| &= \Big| \lim_{\Delta t\to 0} \frac{\mathbf{r}(t+\Delta t)-\mathbf{r}(t)}{\Delta t} \Big| \\ &= \lim_{\Delta t\to 0} \frac{|\mathbf{r}(t+\Delta t)-\mathbf{r}(t)|}{\Delta t}. \end{aligned} \tag{20}$$

For sufficiently small Δt (when orbital curvature becomes negligible), the numerator $|\mathbf{r}(t+\Delta t)-\mathbf{r}(t)|$ approaches the distance Δs on the orbit, travelled by the particle between times t and $t+\Delta t$. Consequently, the limit (20) is the temporary **speed** $v(t)$ of the particle,

$$|\dot{\mathbf{r}}(t)| \;\equiv\; \sqrt{\dot{x}(t)^2+\dot{y}(t)^2+\dot{z}(t)^2} \;=\; \frac{\Delta s}{\Delta t} \;=\; v(t). \tag{21}$$

As is obvious from the figure, the **direction** of the vector $\dot{\mathbf{r}}(t)$, indicated in red color, is **tangential** to the orbit at $\mathbf{r}(t)$. Therefore, the vector

$$\mathbf{v}(t) \;\equiv\; \dot{\mathbf{r}}(t) \tag{22}$$

is called the **velocity (vector)** of the particle.

Example 1b: In Example 1a, we obtain

$$\dot{\mathbf{r}}(t) = \begin{pmatrix} -R\omega\sin(\omega t) \\ +R\omega\cos(\omega t) \\ u \end{pmatrix}, \qquad v(t) \;=\; \sqrt{(R\omega)^2+u^2} \;\equiv\; v. \tag{23}$$

2.3 Parametrization of a curve

Instead of t, any monotonic function $p(t)$ may serve as a parameter p of a curve,

$$\mathbf{r}(p) \;=\; \begin{pmatrix} x(p) \\ y(p) \\ z(p) \end{pmatrix}. \tag{24}$$

A typical **example** is the **arc length** $p(t)=s(t)$ along the curve,

$$s(t) \;=\; \int_0^t \mathrm{d}t' \,\big|\dot{\mathbf{r}}(t')\big| \;\equiv\; \int_0^t \mathrm{d}t' \sqrt{\dot{x}(t')^2+\dot{y}(t')^2+\dot{z}(t')^2}. \tag{25}$$

Example 1c: For the curve of Example 1a, we obtain

$$s(t) \;=\; \int_0^t \mathrm{d}t' \sqrt{(R\omega)^2+u^2} \;\equiv\; vt \qquad \Rightarrow \qquad \mathbf{r}(s) = \begin{pmatrix} R\cos(\frac{\omega}{v}s) \\ R\sin(\frac{\omega}{v}s) \\ \frac{u}{v}s \end{pmatrix}. \tag{26}$$

Exercise: Writing $\frac{\mathrm{d}\mathbf{r}(p)}{\mathrm{d}p} = \dot{\mathbf{r}}(p)$, we find **chain** and **product rules** for vector functions

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{r}\big(p(t)\big) \;=\; \big[\dot{\mathbf{r}}(p)\big]\Big|_{p=p(t)} \dot{p}(t), \tag{27}$$

$$\frac{\mathrm{d}}{\mathrm{d}p}\Big[\mathbf{r}_1(p)\cdot\mathbf{r}_2(p)\Big] \;=\; \dot{\mathbf{r}}_1(p)\cdot\mathbf{r}_2(p) \;+\; \mathbf{r}_1(p)\cdot\dot{\mathbf{r}}_2(p). \tag{28}$$

2.4 Unit tangent vector

Since the vector $\dot{\mathbf{r}}(p)$ is tangential to the orbit, the ("normalized") vector

$$\boldsymbol{\mathcal{T}}(p) \;=\; \frac{\dot{\mathbf{r}}(p)}{|\dot{\mathbf{r}}(p)|} \;=\; \begin{pmatrix} \mathcal{T}_1(p) \\ \mathcal{T}_2(p) \\ \mathcal{T}_3(p) \end{pmatrix} \tag{29}$$

is the **unit tangent vector** of the orbit.
As a consequence, $\boldsymbol{\mathcal{T}}(p)\cdot\boldsymbol{\mathcal{T}}(p) = 1$ and the **product rule** yields

$$\begin{aligned} 0 \;&=\; \frac{\mathrm{d}}{\mathrm{d}p}\Big[\boldsymbol{\mathcal{T}}(p)\cdot\boldsymbol{\mathcal{T}}(p)\Big] \\ &=\; 2\,\boldsymbol{\mathcal{T}}(p)\cdot\frac{\mathrm{d}\boldsymbol{\mathcal{T}}(p)}{\mathrm{d}p} \qquad \Rightarrow \qquad \frac{\mathrm{d}\boldsymbol{\mathcal{T}}(p)}{\mathrm{d}p} \;\perp\; \boldsymbol{\mathcal{T}}(p). \end{aligned} \tag{30}$$

Consequently, the vectors $\boldsymbol{\mathcal{T}}(p)$ and $\frac{\mathrm{d}\boldsymbol{\mathcal{T}}(p)}{\mathrm{d}p}$ are **orthogonal** to each other.

Remark 1: When $p = s$ is the arc length along the curve (with $|\dot{\mathbf{r}}(s)| = 1$), we write

$$\frac{\mathrm{d}\boldsymbol{\mathcal{T}}(s)}{\mathrm{d}s} \;=:\; \frac{1}{\rho(s)}\boldsymbol{\mathcal{N}}(s) \qquad \begin{cases} |\boldsymbol{\mathcal{N}}(s)| = 1, \\ \rho(s) \geq 0. \end{cases} \tag{31}$$

- $\boldsymbol{\mathcal{N}}(s)$, the **unit normal vector**, points toward the center of that circle in space which best approximates the curve at the point $\mathbf{r}(s)$.
- $\rho(s)$ is the radius of that circle ("**radius of curvature**").

Remark 2: Choosing again the time t as parameter, with $v(t) = \frac{\mathrm{d}s(t)}{\mathrm{d}t} \equiv \dot{s}(t)$, we have

$$\mathbf{v}(t) \;\equiv\; \dot{\mathbf{r}}(t) \;=\; v(t)\,\boldsymbol{\mathcal{T}}\big(s(t)\big). \tag{32}$$

Taking here the derivative $\frac{\mathrm{d}}{\mathrm{d}t}$ (for the second time), we obtain another vector function of time t which is called the vector $\mathbf{a}(t)$ of the particle's **acceleration**,

$$\begin{aligned} \mathbf{a}(t) \;\equiv\; \ddot{\mathbf{r}}(t) \;&=\; \dot{v}(t)\,\boldsymbol{\mathcal{T}}\big(s(t)\big) \;+\; v(t)\,\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{T}}\big(s(t)\big) \\ &=\; \dot{v}(t)\,\boldsymbol{\mathcal{T}}\big(s(t)\big) \;+\; v(t)\,\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\mathcal{T}}(s)\Big|_{s=s(t)}\underbrace{\dot{s}(t)}_{v(t)} \\ &=\; \ddot{s}(t)\,\boldsymbol{\mathcal{T}}\big(s(t)\big) \;+\; \frac{\dot{s}(t)^2}{\rho\big(s(t)\big)}\boldsymbol{\mathcal{N}}\big(s(t)\big). \end{aligned} \tag{33}$$

We see that $\mathbf{a}(t)$ has a component $\ddot{s}(t)\,\boldsymbol{\mathcal{T}}\big(s(t)\big)$ **tangential** to the orbit of the moving particle, but, in addition, also has a component $\frac{\dot{s}(t)^2}{\rho(s(t))}\boldsymbol{\mathcal{N}}\big(s(t)\big)$ **normal** to the orbit. Consequently, while $|\mathbf{v}(t)| = |\dot{s}(t)|$, we have

$$|\mathbf{a}(t)| \;=\; \sqrt{\ddot{s}(t)^2 + \left[\frac{\dot{s}(t)^2}{\rho\big(s(t)\big)}\right]^2} \;\geq\; |\ddot{s}(t)|. \tag{34}$$

3 Scalar fields

A **scalar field** is a function f that assigns a number

$$f(\mathbf{r}) \equiv f(x, y, z) \tag{35}$$

to each point $\mathbf{r} = (x, y, z)$ in space.
Examples: $T(x, y, z)$ can be the air temperature (in K) in a room or the temperature distribution in a solid, $\rho(x, y, z)$ the density (in mg/m^3) of CO_2 in air (e.g., above a city). [In addition, these values are typically changing with time t, $T = T(x, y, z, t)$, etc.]

3.1 Graphical representations

3.1.1 2D fields: 3D plot

An example for a 2D scalar field $f(x, y)$ is the temperature $T(x, y)$ on a planar metal surface (xy-plane). As a model, we consider the simple analytical function

$$T(x, y) = \frac{y}{1 + x^2}. \tag{36}$$

Such a function with only **two** variables can be illustrated by a graph in 3D space: For each point (x, y) on the xy-plane consider the number $z = T(x, y)$ as the third coordinate of a **point** (x, y, z) **in 3D space** above (or, for $z < 0$, below) the xy-plane.

x	0	1	2	3	0	1	2	3
y	1	1	1	1	2	2	2	2
z	1.0	0.5	0.2	0.1	2.0	1.0	0.4	0.2

For a few selected points (x, y) in the xy-plane (black dots), the resulting points (x, y, z) are plotted in Fig. 2 (red dots). The union of all these image points forms a curved surface in 3D space, indicated by blue curves in Fig. 2, called the **3D plot** of f.

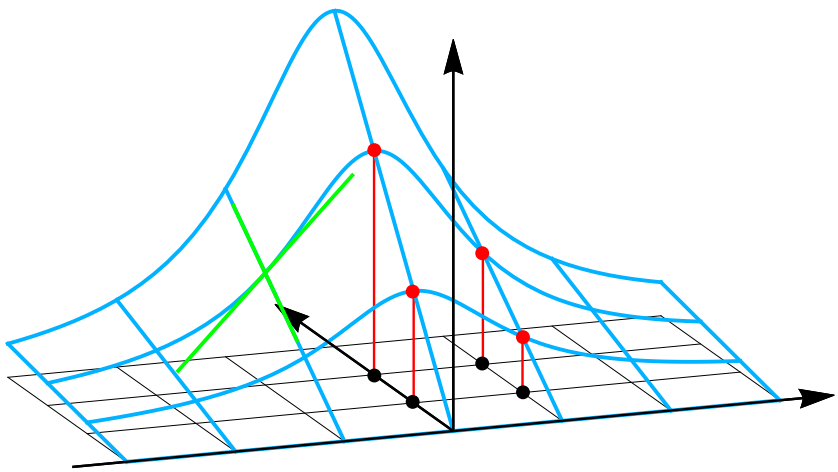


Figure 2: 3D plot of the function $T(x, y) = \frac{y}{1+x^2}$ (blue solid curves).

3.1.2 2D fields: Contour plot

Alternatively, we can plot curves of constant temperature in the xy-plane: Resolving

$$T(x,y) \equiv \frac{y}{1+x^2} = z \tag{37}$$

for y (we could resolve for x as well), we find

$$\begin{aligned} y &= z(1+x^2) \\ &\equiv y_z(x). \end{aligned} \tag{38}$$

For each fixed value of z, this equation describes a curve in the xy-plane, joining points with constant temperature $T(x,y) = z$. For $z \in \{0.5, 1.0, 1.5, ...\}$, these "isothermals" on the xy-plane are shown as blue curves in Fig. 3.

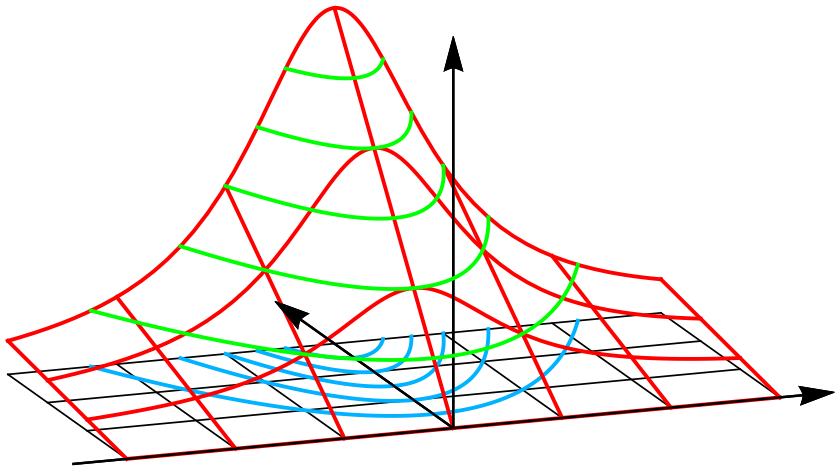


Figure 3: Contour plot of the function $T(x,y) = \frac{y}{1+x^2}$ (blue solid curves).

Such a set of curves in the xy-plane is called a **contour plot** of the scalar field $T(x,y)$.

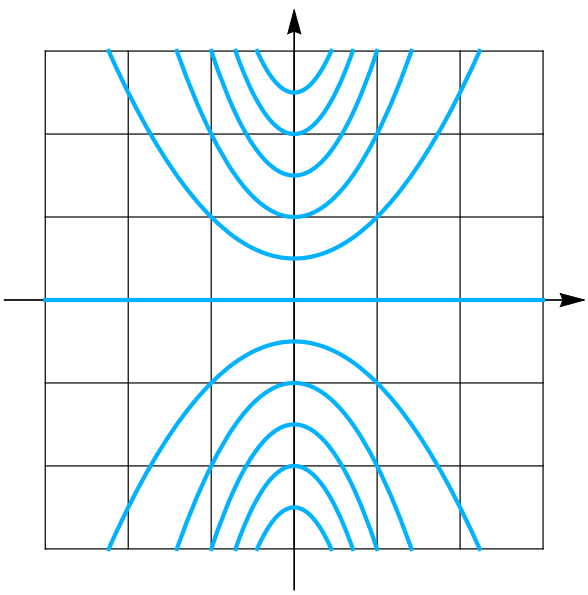


Figure 4: Contour plot of the function $T(x,y) = \frac{y}{1+x^2}$ in the xy-plane (blue solid curves).

3.2 Partial derivatives

For a given scalar field $f(x, y, z)$, the three functions (scalar fields as well!)

$$\begin{aligned} f_x(x,y,z) &\equiv \frac{\partial}{\partial x} f(x,y,z) := \lim_{h\to 0} \frac{f(x+h,y,z) - f(x,y,z)}{h}, \\ f_y(x,y,z) &\equiv \frac{\partial}{\partial y} f(x,y,z), \\ f_z(x,y,z) &\equiv \frac{\partial}{\partial z} f(x,y,z) \end{aligned} \tag{39}$$

are called its **partial derivatives**. They are evaluated in the same way as the derivative of a function with only **one** variable, treating all the other variables simply as **parameters**.

Example: The partial derivatives of $f(x, y, z) = x^2 y \cos(z^2)$ are

$$\frac{\partial f}{\partial x} = 2xy\cos(z^2), \qquad \frac{\partial f}{\partial y} = x^2\cos(z^2), \qquad \frac{\partial f}{\partial x} = -2x^2yz\sin(z^2). \tag{40}$$

For a 2D scalar field $f(x, y)$, the partial derivatives are visible in the 3D plot of Fig. 2:
• $f_x(x_0, y_0)$ is the slope of the tangent (green) to the curve $z = f(x, y_0)$ in the xz-plane,
• $f_y(x_0, y_0)$ is the slope of the tangent (green) to the curve $z = f(x_0, y)$ in the yz-plane,
both tangents drawn at the point (x_0, y_0, z_0), where $z_0 = f(x_0, y_0)$.

3.3 Gradient

To find an alternative way of illustrating, applicable also to a 3D field $f(x, y, z)$, we take the quantities $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ as the coordinates of a column vector

$$\mathbf{G}_f(x,y) = \begin{pmatrix} \frac{\partial}{\partial x} f(x,y) \\ \frac{\partial}{\partial y} f(x,y) \end{pmatrix} \equiv \begin{pmatrix} f_x(x,y) \\ f_y(x,y) \end{pmatrix}. \tag{41}$$

This vector $\mathbf{G}_f(x, y)$, actually a vector field, is called the **gradient** of the scalar field f at the point $\mathbf{r} = (x, y)$. Introducing the **nabla operator** ∇,

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}, \tag{42}$$

the gradient is usually written as

$$\mathbf{G}_f(x,y) = \nabla f(x,y) \qquad \Leftrightarrow \qquad \mathbf{G}_f(\mathbf{r}) = \nabla f(\mathbf{r}). \tag{43}$$

At the point (x_0, y_0) on the planar metal surface (xy plane), the vector

$$\mathbf{G}_T(x_0,y_0) = \nabla T(x,y)\Big|_{x=x_0, y=y_0} \tag{44}$$

points into that direction where the temperature is increasing at the strongest rate. Consequently, it is exactly perpendicular to the corresponding curve in the contour plot.

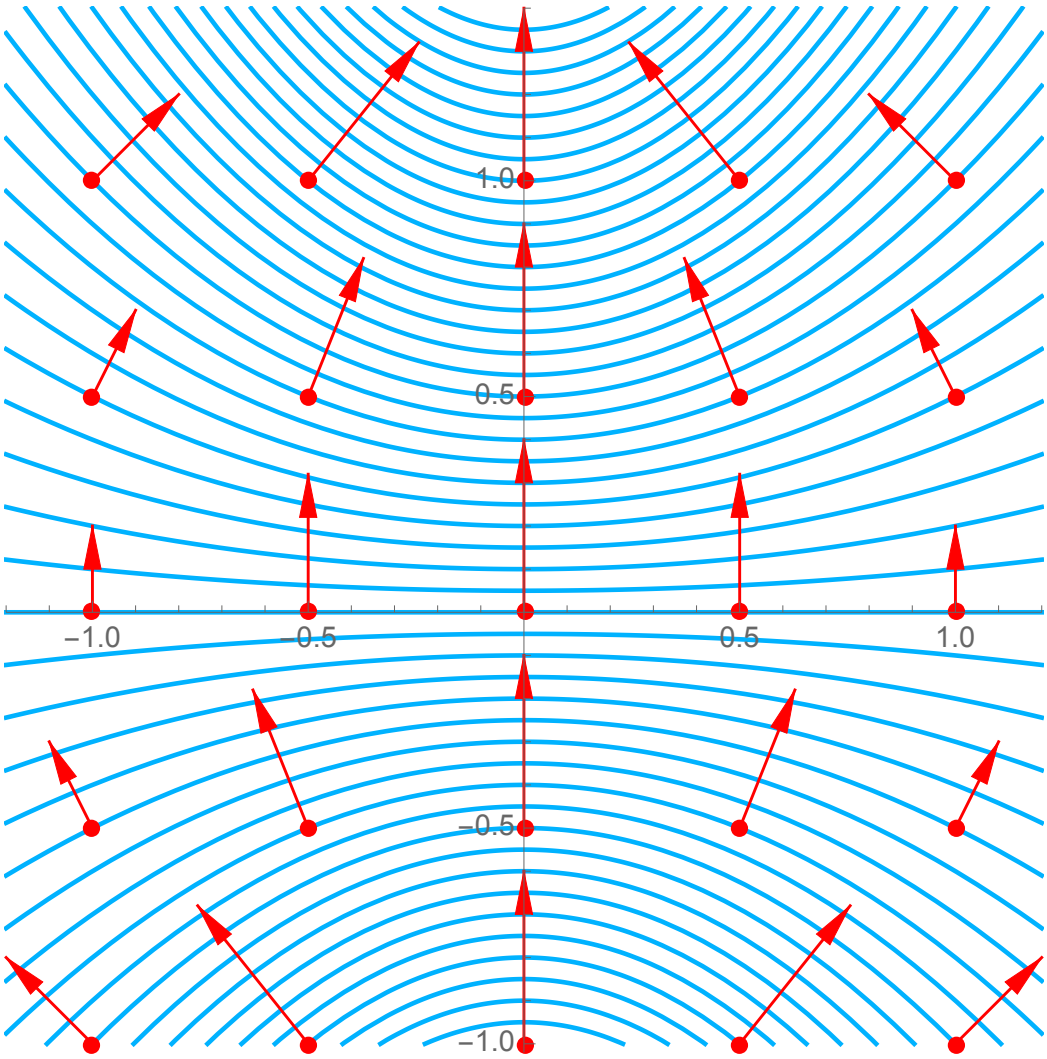


Figure 5: Gradient field of the function $T(x, y) = \frac{y}{1+x^2}$ (red arrows).

A 3D scalar field $f(x, y, z)$ such as the air temperature $T(x, y, z)$, cannot be represented graphically (since there is no such thing as a "4D plot"!). In this case, the gradient is a 3D column vector,

$$\mathbf{G}_f(x, y, z) \;:=\; \nabla f(x, y, z) \;\equiv\; \begin{pmatrix} f_x(x, y, z) \\ f_y(x, y, z) \\ f_z(x, y, z) \end{pmatrix}, \qquad \nabla \;:=\; \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}. \tag{45}$$

Still, now in 3D space, this vector points into the direction where the function $f(x, y, z)$, e.g., the air temperature $T(x, y, z)$, increases at the strongest rate.

3.4 Multiple integrals

3.4.1 Double or surface integrals

On a water surface (xy-plane), let $\rho(x,y)$ be the mass density (in μg/cm^2) of an oil film, representing a 2D scalar field. For simplicity, we assume the function

$$\rho(x,y) \;=\; \left\{ \begin{array}{ll} \rho_0\Big(1 - \dfrac{x^2+y^2}{R^2}\Big) & (x^2+y^2 \le R^2) \\ 0 & (x^2+y^2 \ge R^2) \end{array} \right\}, \quad \rho_0 = 2.7\,\frac{\mu\mathrm{g}}{\mathrm{cm}^2}, \quad R = 0.50\,\mathrm{m}. \tag{46}$$

From the function $\rho(x,y)$, we wish to compute the amount m_Ω of oil (in μg) distributed within a given piece Ω of water surface, e.g., the right triangle (red in Fig. 6)

$$\Omega \;=\; \left\{ (x,y) \in \mathbb{R}^2 \,\middle|\, 0 \le x \le a, \quad 0 \le y \le \frac{bx}{a} \right\} \qquad (a,b>0, \quad a^2+b^2 < R^2). \tag{47}$$

For an approximate solution to this problem, we introduce small step sizes Δx and Δy, generating a mesh of discrete points $\mathbf{r}_{ij}$ in the xy-plane (black dots in Fig. 6),

$$\mathbf{r}_{ij} \;=\; (x_i, y_j) \;=\; (i\cdot\Delta x,\; j\cdot\Delta y) \qquad (i,j \in \mathbb{Z}). \tag{48}$$

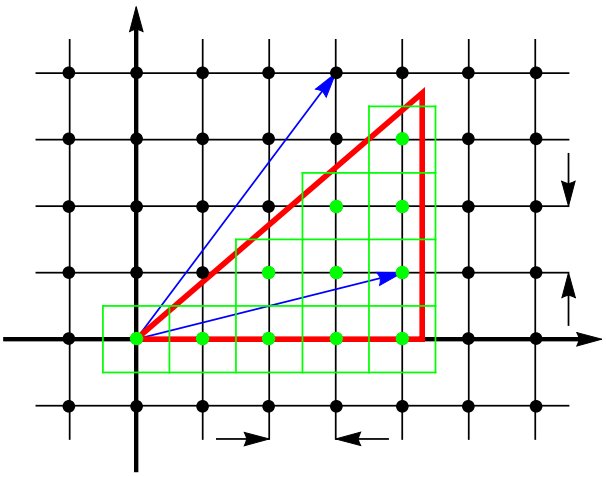


Figure 6: The right triangle Ω (red) and the position vectors $\mathbf{r}_{41}$ and $\mathbf{r}_{34}$ (blue).

For sufficiently small Δx and Δy, the amount Δm_{ij} of oil within the small rectangle with side lengths Δx and Δy, centered at the point $\mathbf{r}_{ij}$, is approximately given by

$$\Delta m_{ij} \;\approx\; \Delta x \cdot \Delta y \cdot \rho(x_i, y_j). \tag{49}$$

Considering all points $\mathbf{r}_{ij} \in \Omega$ (green in Fig. 6) and taking the sum, we have

$$m_\Omega \;\approx\; \sum_{\mathbf{r}_{ij}\in\Omega} \Delta m_{ij} \;=\; \sum_{i=i_{\min}}^{i_{\max}} \Delta x \left[\sum_{j=j_{\min}(i)}^{j_{\max}(i)} \Delta y \cdot \rho(x_i, y_j) \right]. \tag{50}$$

For the situation in Fig. 6, e.g., the outer sum has the fixed limits $i_{\min} = 0$ and $i_{\max} = 4$, while the limits $j_{\min}$ and $j_{\max}$ of the inner sum are depending on the current value of i:

$$j_{\min}(i) = 0 \;(\forall i), \qquad j_{\max}(0) = j_{\max}(1) = 0, \quad j_{\max}(2) = 1, \quad j_{\max}(3) = 2, \quad j_{\max}(4) = 3.$$

At this point, it is easy to see[1] that the limit $\Delta x, \Delta y \to 0$ yields the **double integral**

$$m_\Omega \;=\; \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_{y_{\min}(x)}^{y_{\max}(x)} \mathrm{d}y\, \rho(x,y), \tag{58}$$

where $x_{\min}$ and $x_{\max}$ are the extreme values of the x-coordinates of all points (x,y) in Ω, while $y_{\min}(x)$ and $y_{\max}(x)$ are, for a given value of x, the ones of the y-coordinates,

$$m_\Omega \;=\; \int_0^a \mathrm{d}x \int_0^{bx/a} \mathrm{d}y\, \rho(x,y). \tag{59}$$

[1]Fixed by the boundary of the domain $\Omega \in \mathbb{R}^2$, the numbers $i_{\min}$ and $i_{\max}$ depend on Δx and, for each i, the numbers $j_{\min}(i)$ and $j_{\max}(i)$ depend on Δy, with the limits

$$\lim_{\Delta x\to 0}\left[i_{\min}\cdot\Delta x\right] \;=\; x_{\min}, \qquad \lim_{\Delta x\to 0}\left[i_{\max}\cdot\Delta x\right] \;=\; x_{\max}\,; \tag{51}$$

$$\lim_{\Delta y\to 0}\left[j_{\min}(i)\cdot\Delta y\right] \;=\; y_{\min}(x_i), \qquad \lim_{\Delta y\to 0}\left[j_{\max}(i)\cdot\Delta y\right] \;=\; y_{\max}(x_i), \tag{52}$$

where $x_{\min}/x_{\max}$ are the smallest/largest x-coordinates of any $\mathbf{r} \in \Omega$, while $y_{\min}(x)/y_{\max}(x)$ are the smallest/largest y-coordinates of any $\mathbf{r} \in \Omega$ with given x-coordinate x.
E.g., for the triangle Ω in Fig. 6/Eq. (47), we have

$$x_{\min} = 0, \quad x_{\max} = a, \qquad y_{\min}(x) \equiv 0, \quad y_{\max}(x) = \frac{bx}{a}. \tag{53}$$

The approximation (50) becomes exact in the limit $\Delta x, \Delta y \to 0$,

$$m_\Omega \;=\; \lim_{\Delta x\to 0} \sum_{i=i_{\min}}^{i_{\max}} \Delta x \left[\lim_{\Delta y\to 0} \sum_{j=j_{\min}(i)}^{j_{\max}(i)} \Delta y\, \rho(x_i, y_j)\right]. \tag{54}$$

Obviously, for any fixed value of x_i, the limit [...] in squared brackets is the ordinary integral of the function $f(y) \equiv \rho(x_i, y)$,

$$\lim_{\Delta y\to 0} \sum_{j=j_{\min}(i)}^{j_{\max}(i)} \Delta y\, f(y_j) \;=\; \int_{y_{\min}(x_i)}^{y_{\max}(x_i)} \mathrm{d}y\, f(y). \tag{55}$$

It can be evaluated in the familiar way,

$$\begin{aligned}
\int_{y_{\min}(x_i)}^{y_{\max}(x_i)} \mathrm{d}y\, \rho(x_i, y) \;&=\; \rho_0 \int_0^{bx_i/a} \mathrm{d}y\, (A - By^2) \qquad \left(A = 1 - \frac{x_i^2}{R^2}, \quad B = \frac{1}{R^2}\right) \\
&=\; \rho_0 \left[Ay - \frac{B}{3}\, y^3\right]_{y=0}^{y=bx_i/a} \\
&=\; \rho_0 \left[\frac{b}{a}\, x_i \;-\; \frac{b}{aR^2}\left(1 + \frac{b^2}{3a^2}\right) x_i^3\right] \;\equiv\; F(x_i).
\end{aligned} \tag{56}$$

In a second step, the limit $\Delta x \to 0$ in Eq. (54) yields the ordinary integral of $F(x)$,

$$\begin{aligned}
m_\Omega \;&=\; \lim_{\Delta x\to 0} \sum_{i=i_{\min}}^{i_{\max}} \Delta x\, F(x_i) \\
&=\; \int_{x_{\min}}^{x_{\max}} \mathrm{d}x\, F(x) \\
&=\; \int_0^a \mathrm{d}x\, F(x) \;=\; \rho_0 \left[\frac{ab}{2} - \frac{ab}{12\, R^2}(3a^2 + b^2)\right] \;=\; \rho_0 \frac{ab}{2}\left[1 - \frac{3a^2 + b^2}{6\, R^2}\right].
\end{aligned} \tag{57}$$

These two subsequent integrations are usually written as the double integral of Eq. (59).

In evaluating a double integral, we first perform the "inner" (here: y) integration (note that the upper limit bx/a of the y integral depends on the variable x of the "outer" integral!), treating the other variable x as a **constant parameter**,

$$\begin{aligned}
m_\Omega &= \int_0^a \mathrm{d}x \int_0^{bx/a} \mathrm{d}y\, \rho_0\Big(1 - \frac{x^2+y^2}{R^2}\Big) \\
&= \int_0^a \mathrm{d}x \int_0^{bx/a} \mathrm{d}y\, \frac{\rho_0}{R^2}\Big(R^2 - (x^2+y^2)\Big) \\
&= \frac{\rho_0}{R^2}\int_0^a \mathrm{d}x \int_0^{bx/a} \mathrm{d}y\, \left(R^2 - x^2 - y^2\right) \\
&= \frac{\rho_0}{R^2}\int_0^a \mathrm{d}x \left[(R^2-x^2)y - \frac{y^3}{3}\right]_{y=0}^{y=bx/a} \\
&= \frac{\rho_0}{R^2}\int_0^a \mathrm{d}x \left[\frac{bR^2}{a}x - \Big(\frac{b}{a} + \frac{b^3}{3a^3}\Big)x^3\right].
\end{aligned} \tag{60}$$

Remaining is an ordinary integral over **one** variable x,

$$m_\Omega = \frac{\rho_0}{R^2}\left[\frac{bR^2}{a}\frac{x^2}{2} - \Big(\frac{b}{a} + \frac{b^3}{3a^3}\Big)\frac{x^4}{4}\right]_0^a = \rho_0\frac{ab}{2}\left[1 - \frac{3a^2+b^2}{6\,R^2}\right]. \tag{61}$$

Remark: It is also possible, to take the y integral as the outer one. Then, however, the integration limits must be re-defined,

$$\begin{aligned}
m_\Omega &= \frac{\rho_0}{R^2}\int_{y_{\min}}^{y_{\max}} \mathrm{d}y \int_{x_{\min}(y)}^{x_{\max}(y)} \mathrm{d}x\, \left(R^2 - y^2 - x^2\right) \\
&= \frac{\rho_0}{R^2}\int_0^b \mathrm{d}y \int_{ay/b}^{a} \mathrm{d}x\, \left(R^2 - y^2 - x^2\right) \\
&= \frac{\rho_0}{R^2}\int_0^b \mathrm{d}y \left[(R^2-y^2)x - \frac{x^3}{3}\right]_{x=ay/b}^{x=a} \\
&= \frac{\rho_0}{R^2}\int_0^b \mathrm{d}y \Big[(R^2-y^2)\Big(a - \frac{a}{b}y\Big) - \frac{1}{3}\Big(a^3 - \frac{a^3}{b^3}y^3\Big)\Big] \\
&= \frac{\rho_0}{R^2}\int_0^b \mathrm{d}y \Big[\Big(aR^2 - \frac{a^3}{3}\Big) - \frac{aR^2}{b}y - ay^2 + \Big(\frac{a}{b} + \frac{a^3}{3b^3}\Big)y^3\Big]
\end{aligned} \tag{62}$$

Again, we are left with an ordinary integral over **one** variable y,

$$\begin{aligned}
m_\Omega &= \frac{\rho_0}{R^2}\Big[\Big(aR^2 - \frac{a^3}{3}\Big)b - \frac{abR^2}{2} - \frac{ab^3}{3} + \Big(\frac{ab^3}{4} + \frac{a^3b}{12}\Big)\Big] \\
&= \rho_0\Big[\frac{ab}{2} - \frac{ab}{12\,R^2}(3a^2+b^2)\Big].
\end{aligned} \tag{63}$$

The result, of course, cannot depend on the particular way chosen for evaluating the integrals. It is completely fixed by the domain $\Omega \subseteq \mathbb{R}^2$ and by the integrand (scalar field) $\rho(x,y) \equiv \rho(\mathbf{r})$. Therefore, we use the notation

$$m_\Omega = \int_\Omega \mathrm{d}^2r\, \rho(\mathbf{r}), \tag{64}$$

and call this the **double integral** of the scalar field $\rho(\mathbf{r})$ over the domain $\Omega \subseteq \mathbb{R}^2$.

Example: We consider the **circle** Ω with radius a and center at $\mathbf{r} = \mathbf{0}$,

$$\Omega \;=\; \Big\{(x,y) \in \mathbb{R}^2 \,\Big|\, x^2 + y^2 \le a^2\Big\}. \tag{65}$$

In this case, $x_{\min}, x_{\max} = \mp a$ and $y_{\min}(x), y_{\max}(x) = \mp\sqrt{a^2 - x^2}$,

$$\begin{aligned}
\int_\Omega \mathrm{d}^2r\, \rho(\mathbf{r}) \;&=\; \frac{\rho_0}{R^2}\int_{-a}^{a} \mathrm{d}x \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \mathrm{d}y\, \left(R^2 - x^2 - y^2\right) \\
&=\; \frac{\rho_0}{R^2}\int_{-a}^{a} \mathrm{d}x \left[(R^2 - x^2)y - \frac{y^3}{3}\right]_{y=-\sqrt{a^2-x^2}}^{y=\sqrt{a^2-x^2}} \\
&=\; \frac{\rho_0}{R^2}\int_{-a}^{a} \mathrm{d}x \left[(R^2 - x^2)2\sqrt{a^2-x^2} - \frac{2}{3}(a^2-x^2)^{3/2}\right] \\
&=\; \frac{\rho_0}{R^2}\int_{-a}^{a} \mathrm{d}x \left[2\Big(R^2 - \frac{a^2}{3}\Big) - \frac{4}{3}x^2\right]\sqrt{a^2-x^2}.
\end{aligned} \tag{66}$$

From standard integral tables (e.g.: Bronstein, integrals nos. 157 and 159), we extract

$$\int \mathrm{d}x\, \sqrt{a^2-x^2} \;=\; \frac{1}{2}\Big(x\sqrt{a^2-x^2} \,+\, a^2 \arcsin\frac{x}{a}\Big), \tag{67}$$

$$\int \mathrm{d}x\, x^2\sqrt{a^2-x^2} \;=\; \frac{a^2}{8}\Big(x\sqrt{a^2-x^2} \,+\, a^2 \arcsin\frac{x}{a}\Big) \,-\, \frac{x}{4}\,(a^2-x^2)^{3/2}, \tag{68}$$

to find

$$\int_\Omega \mathrm{d}^2r\, \rho(\mathbf{r}) \;=\; \rho_0\Big(\pi a^2 \,-\, \frac{\pi a^4}{2R^2}\Big). \tag{69}$$

This derivation is simplified considerably by using polar coordinates, see Eq. (94) below.

3.4.2 Triple or volume integrals

Integrals of scalar fields $f(\mathbf{r})$ over a 3D domain Ω are called **triple** or **volume integrals**,

$$\int_\Omega \mathrm{d}^3r\, f(\mathbf{r}) \;=\; \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_{y_{\min}(x)}^{y_{\max}(x)} \mathrm{d}y \int_{z_{\min}(x,y)}^{z_{\max}(x,y)} \mathrm{d}z\, f(x,y,z). \tag{70}$$

Again, the order of these three integrations may be chosen differently, when the limits of each integral are re-defined properly.

Example: Let Ω be a pyramid with height H and quadratic base with side a,

$$\Omega \;=\; \Big\{(x,y,z) \,\Big|\, 0 \le z \le H, \quad 0 \le |x|, |y| \le \frac{a}{2}\Big(1 - \frac{z}{H}\Big)\Big\}. \tag{71}$$

In this case, we choose the z integral as the outer one to obtain

$$\begin{aligned}
\int_\Omega \mathrm{d}^3r\, f(\mathbf{r}) \;&=\; \int_{z_{\min}}^{z_{\max}} \mathrm{d}z \int_{x_{\min}(z)}^{x_{\max}(z)} \mathrm{d}x \int_{y_{\min}(z)}^{y_{\max}(z)} \mathrm{d}y\, f(x,y,z) \\
&=\; \int_0^H \mathrm{d}z \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}x \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}y\, f(x,y,z).
\end{aligned} \tag{72}$$

For the simple function $f(x,y,z) = az$, this integral will be evaluated in Eq. (78) below.

3.4.3 Average value of a function

A solid rod, lying on the x-axis (with its endpoints at $x = a$ and at $x = b$), may have a non-uniform (x-dependent) temperature distribution, given by a function $T(x)$.

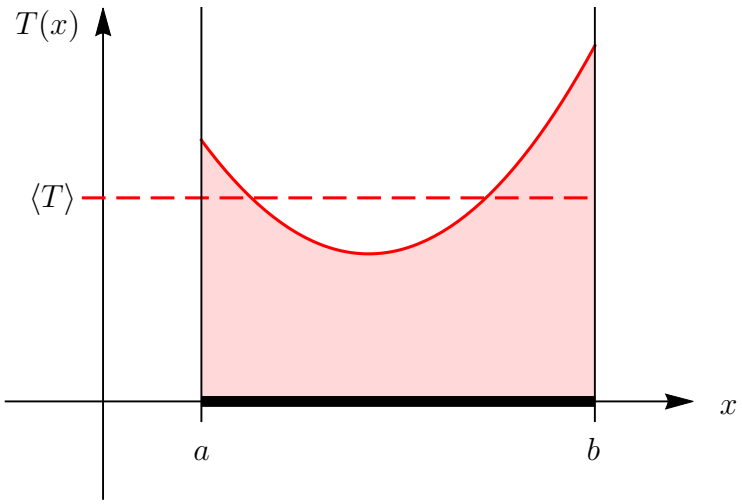


Figure 7: The rod on the x-axis (black), its non-uniform temperature distribution (red curve), and its average temperature (dashed red line).

From the figure, it is obvious that the **average temperature** of the rod is given by

$$\langle T \rangle \;=\; \frac{1}{L}\int_a^b \mathrm{d}x\, T(x), \tag{73}$$

where $L = b - a$ is the length of the rod.

Instead of the rod, covering the 1D interval $\Omega = [a, b] \subset \mathbb{R}$ with **length** $L_\Omega = b - a$, we may also consider a thin metal sheet, covering a 2D surface $\Omega \subset \mathbb{R}^2$ with **area** A_Ω, or a piece of solid material, covering a 3D region $\Omega \subset \mathbb{R}^3$ with **volume** V_Ω.

Area A_Ω and length L_Ω can each be called a "volume" V_Ω in D-dimensional space (with $D = 2$ and $D = 1$, respectively). In this spirit, we generally define:

Def.: The **average value** of a scalar field $T(\mathbf{r})$ in a D-dimensional domain $\Omega \subset \mathbb{R}^D$ is

$$\Big\langle T(\mathbf{r}) \Big\rangle_{\mathbf{r}\in\Omega} \;\equiv\; \frac{1}{V_\Omega}\int_\Omega \mathrm{d}^D r\, T(\mathbf{r}), \tag{74}$$

where V_Ω is the D-dimensional "volume" of Ω (with $D = 3$, $D = 2$ or $D = 1$).

Example 1: Consider a metal sheet covering the "upper half " Ω of a disk,

$$\Omega \;=\; \Big\{(x|y) \in \mathbb{R}^2 \;\Big|\; x^2 + y^2 \le a^2,\; y \ge 0\Big\} \qquad (V_\Omega \equiv A_\Omega = \frac{\pi}{2}a^2). \tag{75}$$

Given the function $T(x, y) = T_0 \cdot \frac{y}{a}$, the average temperature on the sheet is

$$\begin{aligned}\Big\langle T(\mathbf{r}) \Big\rangle_{\mathbf{r}\in\Omega} &= \frac{1}{\frac{\pi}{2}a^2}\int_{-a}^{a}\mathrm{d}x\int_0^{\sqrt{a^2-x^2}}\mathrm{d}y\, T_0\cdot\frac{y}{a} \\ &= \frac{2}{\pi a^2}\frac{T_0}{a}\int_{-a}^{a}\mathrm{d}x\left[\frac{y^2}{2}\right]_0^{\sqrt{a^2-x^2}} \;=\; \frac{4}{3\pi}T_0.\end{aligned} \tag{76}$$

Example 2: Inside the pyramid Ω of Eq. (71) the temperature shall increase linearly from 0°C at the floor ($z = 0$) to some maximum value T_0 at the tip ($z = H$),

$$T(\mathbf{r}) \;=\; \frac{T_0}{H}\, z \;\equiv\; T(x, y, z). \tag{77}$$

Given the volume $V_\Omega = \frac{1}{3}a^2 H$ of this pyramid, the **average temperature** inside is

$$\begin{aligned}
\Big\langle T(\mathbf{r})\Big\rangle_{\mathbf{r}\in\Omega} &\equiv \frac{1}{V_\Omega}\int_\Omega \mathrm{d}^3r\; T(\mathbf{r}) \\
&= \frac{3}{a^2 H}\int_0^H \mathrm{d}z \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}x \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}y\, \frac{T_0}{H}\, z \\
&= \frac{3T_0}{a^2H^2}\int_0^H \mathrm{d}z \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}x \Big[zy\Big]_{y=-\frac{a}{2}(1-\frac{z}{H})}^{y=\frac{a}{2}(1-\frac{z}{H})} \\
&= \frac{3T_0}{a^2H^2}\int_0^H \mathrm{d}z \int_{-\frac{a}{2}(1-\frac{z}{H})}^{\frac{a}{2}(1-\frac{z}{H})} \mathrm{d}x\; za\Big(1-\frac{z}{H}\Big) \\
&= \frac{3T_0}{a^2H^2}\int_0^H \mathrm{d}z \Big[za\Big(1-\frac{z}{H}\Big)x\Big]_{x=-\frac{a}{2}(1-\frac{z}{H})}^{x=\frac{a}{2}(1-\frac{z}{H})} \\
&= \frac{3T_0}{a^2H^2}\int_0^H \mathrm{d}z\; za^2\Big(1-\frac{z}{H}\Big)^2 \\
&= \frac{3T_0}{H^2}\int_0^H \mathrm{d}z\,\Big(\frac{z^3}{H^2}-\frac{2z^2}{H}+z\Big) \;=\; \frac{T_0}{4}.
\end{aligned} \tag{78}$$

Remark A: Computing a volume (in 3D or in 2D).
For the trivial function $T(\mathbf{r}) \equiv 1$, we have $\langle T(\mathbf{r})\rangle_\Omega = 1$, implying the formula

$$V_\Omega \;=\; \int_\Omega \mathrm{d}^D r\, 1. \tag{79}$$

Example 3: Find the volume V_Ω of the tetrahedron Ω, bounded by the four planes $x = 0$, $y = 0$, $z = 0$ and $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ (with $a, b, c > 0$; SKETCH!),

$$\begin{aligned}
V_\Omega \;=\; \int_\Omega \mathrm{d}^3r\, 1 \;&=\; \int_0^a \mathrm{d}x \int_0^{b(1-\frac{x}{a})} \mathrm{d}y \int_0^{c(1-\frac{x}{a}-\frac{y}{b})} \mathrm{d}z\, 1 \\
&= c\int_0^a \mathrm{d}x \int_0^{b(1-\frac{x}{a})} \mathrm{d}y\,\Big(1-\frac{x}{a}-\frac{y}{b}\Big) \\
&= c\int_0^a \mathrm{d}x\,\Big[\Big(1-\frac{x}{a}\Big)y-\frac{y^2}{2b}\Big]_{y=0}^{y=b(1-\frac{x}{a})} \\
&= bc\int_0^a \mathrm{d}x\,\Big[\Big(1-\frac{x}{a}\Big)^2-\frac{1}{2}\Big(1-\frac{x}{a}\Big)^2\Big] \\
&= \frac{bc}{2}\int_0^a \mathrm{d}x\,\Big(1-\frac{x}{a}\Big)^2 \;=\; -\frac{bc}{2}\int_1^0 a\mathrm{d}u\, u^2 \;=\; \frac{abc}{6}.
\end{aligned} \tag{80}$$

Remark B: Double integral as a 3D volume

Let Σ be a domain in the xy-plane, fixed by $x_{\min}$, $x_{\max}$, $y_{\min}(x)$, and $y_{\max}(x)$. Then, the double integral of a 2D scalar field $f(\mathbf{r}) = f(x, y)$ is

$$\begin{aligned} \int_{\Sigma} \mathrm{d}^2 r\, f(\mathbf{r}) &= \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_{y_{\min}(x)}^{y_{\max}(x)} \mathrm{d}y\; f(x, y) \\ &\equiv \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_{y_{\min}(x)}^{y_{\max}(x)} \mathrm{d}y \int_{0}^{f(x,y)} \mathrm{d}z\, 1, \end{aligned} \tag{81}$$

where we have utilized the trivial identity $f(x, y) = \int_0^{f(x,y)} \mathrm{d}z\, 1$.
Consequently, provided that $f(\mathbf{r}) \geq 0$ for $\mathbf{r} \in \Sigma$, the double integral $\int_{\Sigma} \mathrm{d}^2 r\, f(\mathbf{r})$ equals the **volume** of the 3D region Ω bounded by Σ and the (curved) 3D-plot of f,

$$\int_{\Sigma} \mathrm{d}^2 r\, f(\mathbf{r}) \;=\; \int_{\Omega} \mathrm{d}^3 r\, 1. \tag{82}$$

Example 4: The tetrahedral volume from example 2 can also be obtained as the double integral of the 2D scalarfield $f(x, y) = c(1 - \frac{x}{a} - \frac{y}{b}) \equiv z$ over $\Sigma = \{(x, y) \in \mathbb{R}^2 \,|\, ...\}$,

$$V_{\Omega} \;=\; \int_{\Sigma} \mathrm{d}^2 r\, f(\mathbf{r}) \;=\; c \int_0^a \mathrm{d}x \int_0^{b(1-x/a)} \mathrm{d}y \left(1 - \frac{x}{a} - \frac{y}{b}\right) \;=\; ... \;=\; \frac{abc}{6}. \tag{83}$$

4 Curvilinear coordinates

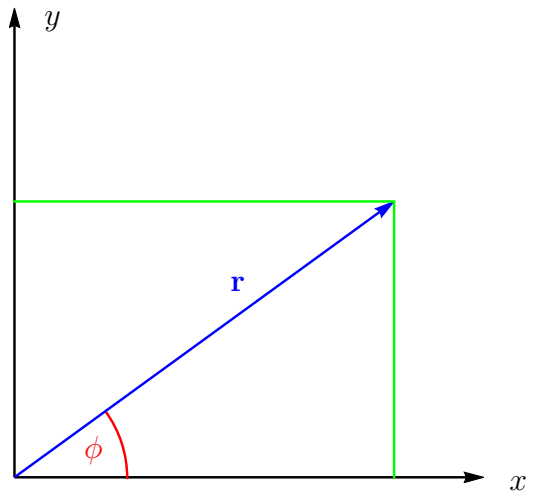
4.1 Planar polar coordinates

The position vector $\mathbf{r}$ of a point P in the xy-plane with **cartesian coordinates** (x, y),

$$\mathbf{r} \;=\; x\mathbf{u}_1 + y\mathbf{u}_2 \;=\; \begin{pmatrix} x \\ y \end{pmatrix}, \tag{84}$$

is fixed as well by its **planar polar coordinates** (r, ϕ):

$r = \lvert\mathbf{r}\rvert$	$0 \leq r < +\infty$	**Length** of $\mathbf{r}$	Distance of P from the origin O.
ϕ	$0 \leq \phi < 2\pi$	**Azimuth angle** of $\mathbf{r}$	Angle from the x-axis to $\mathbf{r}$ (ccw)



4.1.1 Coordinate lines

We express the "old" coordinates (x, y) in terms of the "new" ones (r, ϕ),

$$\mathbf{r}(r, \phi) \equiv \begin{pmatrix} x(r, \phi) \\ y(r, \phi) \end{pmatrix} = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}. \tag{85}$$

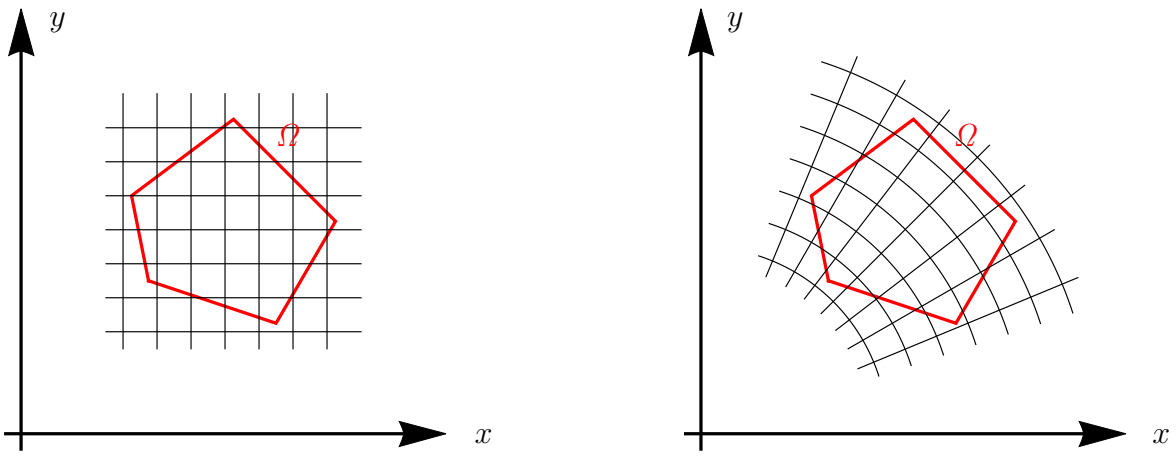
This is a vector function of **two** variables (r and ϕ).

- For fixed $\phi = \phi_0$ and r growing from $r = 0$ to $r = \infty$, the point $\mathbf{r}(r, \phi_0)$ in the xy-plane moves on a **ray** to infinity, starting at the origin.
- For fixed $r = r_0$ and ϕ growing from $\phi = 0$ to $\phi = 2\pi$, the point $\mathbf{r}(r_0, \phi)$ moves ccw on a **circle** of radius r_0 centered at the origin.

These rays and circles are the **coordinate lines** of planar polar coordinates.

4.1.2 Double integrals

As in section 3.4.1, let $\rho(x, y)$ be the 2D mass density of a thin oil film on a water surface (xy-plane $\mathbb{R}^2$). To compute the amount m_Ω of oil distributed within a given piece Ω of water surface (red pentagon in the figures below), we introduce a grid of coordinate lines.



Cartesian coordinate lines (left part of the figure) cut the xy-plane into small rectangles ("tiles") with side lengths Δx, Δy, and with the uniform surface area (per tile)

$$\Delta A \;=\; \Delta x \cdot \Delta y. \tag{86}$$

By proper choice of the coordinate lines, these cartesian tiles have their midpoints at

$$\mathbf{r}_{ij} \;=\; \begin{pmatrix} x_i \\ y_j \end{pmatrix}, \tag{87}$$

where $x_i = i \cdot \Delta x$ and $y_j = j \cdot \Delta y$ (with $i, j \in \mathbb{Z}$). Consequently, the amount m_{ij} of oil distributed within the tile at $\mathbf{r}_{ij}$ is approximately given by $\rho(\mathbf{r}_{ij}) \cdot \Delta A$,

$$m_{ij} \;\approx\; \rho(\mathbf{r}_{ij}) \cdot \Delta x \cdot \Delta y, \tag{88}$$

which becomes exact in the limit $\Delta x, \Delta y \to 0$.
Polar coordinate lines (right part of the figure) cut the xy-plane into curved tiles which, spanning a small angle $\Delta\phi$, are sections of concentric rings with thickness Δr. For small $\Delta\phi$, such a curved tile, with outer and inner radii $r \pm \frac{1}{2}\Delta r$, is approximately rectangular, with side lengths Δr and $r\Delta\phi$. Consequently, the surface area per tile is r-dependent,

$$\Delta\widetilde{A}(r) \;\approx\; r \cdot \Delta r \cdot \Delta\phi. \tag{89}$$

In terms of $x(r, \phi)$ and $y(r, \phi)$ from Eq. (85), these tiles have their midpoints at

$$\mathbf{s}_{\alpha\beta} \;=\; \begin{pmatrix} x(r_\alpha, \phi_\beta) \\ y(r_\alpha, \phi_\beta) \end{pmatrix}, \tag{90}$$

where $r_\alpha = \alpha \cdot \Delta r$ and $\phi_\beta = \beta \cdot \Delta\phi$ (with $\alpha, \beta = 1, 2, 3, ...$). Consequently, the amount $\widetilde{m}_{\alpha\beta}$ of oil distributed within the tile at $\mathbf{s}_{\alpha\beta}$ is approximately given by $\rho(\mathbf{s}_{\alpha\beta}) \cdot \Delta\widetilde{A}(r_\alpha)$,

$$\widetilde{m}_{\alpha\beta} \;\approx\; \rho(\mathbf{s}_{\alpha\beta})\, r_\alpha \cdot \Delta r \cdot \Delta\phi, \tag{91}$$

which becomes exact in the limit $\Delta r, \Delta\phi \to 0$.

Eq. (88) for **cartesian** coordinates implies [cf. Eq. (50) in section 3.4.1]

$$m_\Omega \;\approx\; \sum_{\mathbf{r}_{ij}\in\Omega} m_{ij} \;\approx\; \sum_{i=i_{\min}}^{i_{\max}} \Delta x \sum_{j=j_{\min}(i)}^{j_{\max}(i)} \Delta y \cdot \rho(x_i, y_j)$$
$$\to \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_{y_{\min}(x)}^{y_{\max}(x)} \mathrm{d}y\, \rho(x,y). \tag{92}$$

Similarly, writing $\rho(\mathbf{s}_{\alpha\beta}) \equiv \rho\big(x(r_\alpha,\phi_\beta), y(r_\alpha,\phi_\beta)\big) = \tilde\rho(r_\alpha,\phi_\beta)$, Eq. (91) yields

$$m_\Omega \;\approx\; \sum_{\mathbf{s}_{\alpha\beta}\in\Omega} \widetilde{m}_{\alpha\beta} \;=\; \sum_{\alpha=\alpha_{\min}}^{\alpha_{\max}} \Delta r \sum_{\beta=\beta_{\min}(\alpha)}^{\beta_{\max}(\alpha)} \Delta\phi \cdot r_\alpha \cdot \tilde\rho(r_\alpha,\phi_\beta)$$
$$\to \int_{r_{\min}}^{r_{\max}} \mathrm{d}r \int_{\phi_{\min}(r)}^{\phi_{\max}(r)} \mathrm{d}\phi\; r\, \tilde\rho(r,\phi). \tag{93}$$

The extra factor $J(r,\phi) = r$ in the integrand is the **Jacobian** of planar **polar** coordinates.

Example 1: For the disk Ω with radius a of Eq. (65), Eq. (66) now becomes

$$\begin{aligned}
m_\Omega \;=\; \int_\Omega \mathrm{d}^2r\, \rho(\mathbf{r}) \;&=\; \int_{-a}^{a} \mathrm{d}x \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \mathrm{d}y\, \rho(x,y), \qquad \rho(x,y) = \frac{\rho_0}{R^2}\Big[R^2 - (x^2+y^2)\Big] \\
&=\; \int_0^a \mathrm{d}r \int_0^{2\pi} \mathrm{d}\phi\, r\, \tilde\rho(r,\phi) \\
&=\; \frac{\rho_0}{R^2} \int_0^a \mathrm{d}r \int_0^{2\pi} \mathrm{d}\phi\, r\, (R^2 - r^2) \\
&=\; \frac{\rho_0}{R^2} \int_0^a \mathrm{d}r \Big[(R^2 r - r^3)\,\phi\Big]_{\phi=0}^{\phi=2\pi} \\
&=\; \frac{\rho_0}{R^2} \int_0^a \mathrm{d}r\, 2\pi (R^2 r - r^3) \\
&=\; \frac{\rho_0}{R^2}\, 2\pi \left[R^2 \frac{r^2}{2} - \frac{r^4}{4}\right]_{r=0}^{r=a} \;=\; \rho_0\Big(\pi a^2 \;-\; \frac{\pi a^4}{2R^2}\Big).
\end{aligned} \tag{94}$$

Example 2: If Ω, more generally, is a **section of a circular ring**, we have

$$\int_\Omega \mathrm{d}^2r\, f(\mathbf{r}) \;=\; \int_{r_1}^{r_2} \mathrm{d}r \int_{\phi_1}^{\phi_2} \mathrm{d}\phi\, r\, \tilde f(r,\phi). \tag{95}$$

Remark: For general curvilinear coordinates (v,w), the surface area $\Delta\widetilde{A}$ of the curved tile depends on both v and w. The corresponding **Jacobian** $J(v,w)$ is defined by

$$J(v,w) \;=\; \lim_{\Delta v,\Delta w\to 0} \frac{\Delta\widetilde{A}(v,w)}{\Delta v\cdot\Delta w}. \tag{96}$$

For small but **finite** Δv and Δw, we then have $\Delta\widetilde{A}(v,w) \;\approx\; J(v,w)\cdot\Delta v\cdot\Delta w$ and

$$m_\Omega \;=\; \int_{v_{\min}}^{v_{\max}} \mathrm{d}v \int_{w_{\min}(v)}^{w_{\max}(v)} \mathrm{d}w\; J(v,w)\, \tilde\rho(r,\phi). \tag{97}$$

Generally, the Jacobian is the determinant of the **functional matrix**,

$$J(v,w) \;=\; \det\begin{pmatrix} \partial x(v,w)/\partial v & \partial x(v,w)/\partial w \\ \partial y(v,w)/\partial v & \partial y(v,w)/\partial w \end{pmatrix}. \tag{98}$$

4.1.3 Scale factors

Consider the partial derivatives of the two-argument vector function $\mathbf{r}(r,\phi)$,

$$\begin{aligned}
\mathbf{h}_r(r,\phi) &= \frac{\partial \mathbf{r}(r,\phi)}{\partial r} = \frac{\partial}{\partial r}\begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix} = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \\
\mathbf{h}_\phi(r,\phi) &= \frac{\partial \mathbf{r}(r,\phi)}{\partial \phi} = \frac{\partial}{\partial \phi}\begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix} = \begin{pmatrix} -r\sin\phi \\ +r\cos\phi \end{pmatrix}.
\end{aligned} \tag{99}$$

At each point $\mathbf{r}(r,\phi)$, these are **tangential vectors** to the two coordinate lines. The magnitudes

$$h_r(r,\phi) \equiv |\mathbf{h}_r(r,\phi)| = 1, \qquad h_\phi(r,\phi) \equiv |\mathbf{h}_\phi(r,\phi)| = r, \tag{100}$$

are called **scale factors**. For orthogonal coordinates, their product yields the **Jacobian**,

$$J(r,\phi) = |\mathbf{h}_r(r,\phi)|\,|\mathbf{h}_\phi(r,\phi)| = r. \tag{101}$$

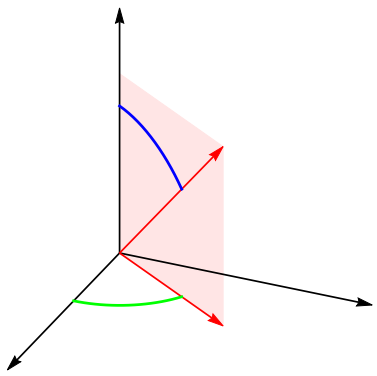
4.2 Spherical (polar) coordinates

The position vector $\mathbf{r}$ of a point P in xyz-space with **cartesian coordinates** (x,y,z),

$$\mathbf{r} = x\mathbf{u}_1 + y\mathbf{u}_2 + z\mathbf{u}_3 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{102}$$

is fixed as well by its **spherical (polar) coordinates** (r,θ,ϕ):

r	$0 \le r < \infty$	**Length** $\lvert\mathbf{r}\rvert$ of $\mathbf{r}$	Distance of P from the origin O.
θ	$0 \le \theta \le \pi$	**Polar angle** of $\mathbf{r}$	Angle between $\mathbf{r}$ and the z-axis (polar axis) $\mathbf{u}_3$
ϕ	$0 \le \phi < 2\pi$	**Azimuth angle** of $\mathbf{r}$	Angle in the xy-plane from the x-axis $\mathbf{u}_1$ (ccw) to the projection $\mathbf{r}_{xy}$ of $\mathbf{r}$ onto the xy-plane



4.2.1 Coordinate lines and surfaces

We express the "old" coordinates (x,y,z) in terms of the "new" ones (r,θ,ϕ),

$$\mathbf{r}(r,\theta,\phi) \equiv \begin{pmatrix} x(r,\theta,\phi) \\ y(r,\theta,\phi) \\ z(r,\theta,\phi) \end{pmatrix} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}. \tag{103}$$

This is a vector function of **three** arguments (r, θ, and ϕ).

Keeping **two** arguments fixed, we are generating the **coordinate lines**:

- The point $\mathbf{r}(r, \theta_0, \phi_0)$, with growing r, moves on a **ray**, starting at $\mathbf{r} = \mathbf{0}$.
- The point $\mathbf{r}(r_0, \theta, \phi_0)$, with growing θ, moves on a vertical **semi circle**.
- The point $\mathbf{r}(r_0, \theta_0, \phi)$, with growing ϕ, moves on a horizontal **circle**.

Keeping only **one** argument fixed, we are generating the **coordinate surfaces**:

- The point $\mathbf{r}(r_0, \theta, \phi)$ moves on a **spherical surface** with radius r_0.
- The point $\mathbf{r}(r, \theta_0, \phi)$ moves on a **conical surface** with opening angle θ_0.
- The point $\mathbf{r}(r, \theta, \phi_0)$ moves on a **semi plane**, bounded by the z-axis.

Except for points on the z-axis, the functions $x(r, \theta, \phi)$, $y(...)$, $z(...)$ can be **inverted**:
• Generally (including points on the z-axis), we have

$$r(x, y, z) \;=\; \sqrt{x^2 + y^2 + z^2}. \tag{104}$$

• For $\mathbf{r} \neq \mathbf{0} \;\Leftrightarrow\; r \neq 0$, we have

$$\begin{aligned} \theta(x, y, z) \;&=\; \arccos \frac{z}{r} \\ &\equiv\; \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}}, \end{aligned} \tag{105}$$

while θ is not defined for $r = 0$.
• Eventually, for $x^2 + y^2 \neq 0$, we have

$$\phi(x, y) \;=\; \begin{cases} \arctan \frac{y}{x} & (x > 0, y \geq 0), \\ \pi/2 & (x = 0, y > 0), \\ \pi + \arctan \frac{y}{x} & (x < 0), \\ 3\pi/2 & (x = 0, y < 0), \\ 2\pi + \arctan \frac{y}{x} & (x > 0, y < 0), \end{cases} \tag{106}$$

while ϕ is not defined for points on the z-axis (with $x^2 + y^2 = 0$).

4.2.2 Triple integrals

A sphere (radius R, center at the origin) has the **electric dipole** charge density

$$\rho(\mathbf{r}) \;\equiv\; \rho(x, y, z) \;=\; a \cdot z \qquad (a > 0). \tag{107}$$

(What is the physical dimension of the constant $a > 0$?)
The amount q_Ω of charge on the upper half sphere Ω ($z \geq 0$),

$$\Omega \;=\; \Big\{ (x, y, z) \;\Big|\; x^2 + y^2 + z^2 \leq R^2, \quad z \geq 0 \Big\}, \tag{108}$$

is in **cartesian coordinates** given by the triple integral

$$\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r}) \;=\; \int_{-R}^{R}\mathrm{d}x\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}\mathrm{d}y\int_0^{\sqrt{R^2-x^2-y^2}}\mathrm{d}z\,\rho(x,y,z). \tag{109}$$

In terms of **spherical polar coordinates** (r,θ,ϕ), writing

$$\tilde\rho(r,\theta,\phi) \;=\; \rho\big(r\sin\theta\cos\phi,\; r\sin\theta\sin\phi,\; r\cos\theta\big), \tag{110}$$

the limits of this integral are simplified considerably,

$$\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r}) \;=\; \int_0^R\mathrm{d}r\int_0^{\pi/2}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,J(r,\theta,\phi)\,\tilde\rho(r,\theta,\phi), \tag{111}$$

with the **Jacobian of spherical polar coordinates**,

$$J(r,\theta,\phi) \;=\; r^2\sin\theta. \tag{112}$$

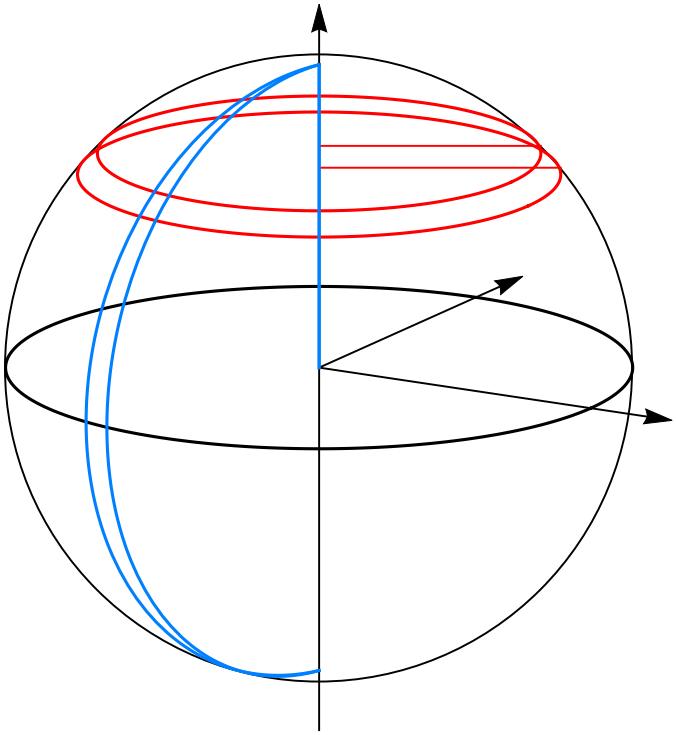
This extra factor arises, since

$$\Delta V_0 \;=\; \Delta r\cdot\Delta\theta\cdot\Delta\phi\cdot J(r_0,\theta_0,\phi_0) \;=\; \Delta r\cdot(r_0\Delta\theta)\cdot(r_0\sin\theta_0\Delta\phi) \tag{113}$$

is the volume of the small almost-cuboid

$$\Delta\Omega \;=\; \left\{\mathbf{r}(r,\theta,\phi)\;\middle|\;\begin{array}{ccccc} r_0-\frac12\Delta r & \le & r & \le & r_0+\frac12\Delta r,\\ \theta_0-\frac12\Delta\theta & \le & \theta & \le & \theta_0+\frac12\Delta\theta,\\ \phi_0-\frac12\Delta\phi & \le & \phi & \le & \phi_0+\frac12\Delta\phi\end{array}\right\} \tag{114}$$

surrounding the point $\mathbf{r}_0=\mathbf{r}(r_0,\theta_0,\phi_0)$.



Now, using $\tilde\rho(r,\theta,\phi)=a\cdot r\cos\theta$, we can evaluate the integral,

$$\begin{aligned}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r}) &= \int_0^R\mathrm{d}r\int_0^{\pi/2}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,r^2\sin\theta\,(a\cdot r\cos\theta)\\ &= 2\pi a\int_0^R\mathrm{d}r\int_0^{\pi/2}\mathrm{d}\theta\,r^3\sin\theta\,\cos\theta\\ &= 2\pi a\int_0^R\mathrm{d}r\left[\frac{r^3}{2}\sin^2\theta\right]_0^{\pi/2} \;=\; \pi a\left[\frac{r^4}{4}\right]_0^R \;=\; \frac{\pi}{4}aR^4.\end{aligned} \tag{115}$$

4.2.3 Scale factors

Consider the partial derivatives of the three-argument vector function $\mathbf{r}(r,\theta,\phi)$,

$$\begin{aligned}
\mathbf{h}_r(r,\theta,\phi) &= \frac{\partial \mathbf{r}(r,\theta,\phi)}{\partial r} = \frac{\partial}{\partial r}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \\
\mathbf{h}_\theta(r,\theta,\phi) &= \frac{\partial \mathbf{r}(r,\theta,\phi)}{\partial \theta} = \frac{\partial}{\partial \theta}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} r\cos\theta\cos\phi \\ r\cos\theta\sin\phi \\ -r\sin\theta \end{pmatrix}, \\
\mathbf{h}_\phi(r,\theta,\phi) &= \frac{\partial \mathbf{r}(r,\theta,\phi)}{\partial \phi} = \frac{\partial}{\partial \phi}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} -r\sin\theta\sin\phi \\ r\sin\theta\cos\phi \\ 0 \end{pmatrix},
\end{aligned} \tag{116}$$

At each point $\mathbf{r}(r,\phi)$, these are **tangential vectors** to the three coordinate lines.
The magnitudes $h_u(r,\theta,\phi) \equiv |\mathbf{h}_u(r,\theta,\phi)|$

$$h_r(r,\theta,\phi) = 1, \qquad h_\theta(r,\theta,\phi) = r, \qquad h_\phi(r,\theta,\phi) = r\sin\theta, \tag{117}$$

are called **scale factors**. Again, their product yields the **Jacobian**,

$$J(r,\theta,\phi) = |\mathbf{h}_r|\,|\mathbf{h}_\theta|\,|\mathbf{h}_\phi| = r^2\sin\theta. \tag{118}$$

4.3 General (orthogonal) curvilinear coordinates

Generalizing (r,θ,ϕ), we now consider **any** set $\alpha = (u,v,w)$ of three coordinates which uniquely specify the cartesian coordinates (x,y,z) of each point $\mathbf{r}$ in space

$$\mathbf{r}(u,v,w) \equiv \begin{pmatrix} x(u,v,w) \\ y(u,v,w) \\ z(u,v,w) \end{pmatrix} = \mathbf{r}(\alpha), \tag{119}$$

with **some** functions $x(u,v,w)$, $y(u,v,w)$, and $z(u,v,w)$.

4.3.1 Scale factors

Tangential vectors to the three coo.-lines through $\mathbf{r}(\alpha)$ are given by

$$\mathbf{h}_1(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial u}, \qquad \mathbf{h}_2(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial v}, \qquad \mathbf{h}_3(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial w}. \tag{120}$$

Def.: The coordinates $\alpha = (u,v,w)$ are called **orthogonal**, when

$$\mathbf{h}_i(\alpha)\cdot\mathbf{h}_j(\alpha) = 0 \quad (i \neq j), \qquad \text{at each point } \mathbf{r}(\alpha). \tag{121}$$

The set $\{\mathbf{h}_1(\alpha), \mathbf{h}_2(\alpha), \mathbf{h}_3(\alpha)\}$ is called the **local basis** at $\mathbf{r}(\alpha)$.
The magnitudes $h_i(\alpha) := |\mathbf{h}_i(\alpha)|$ are called the **scale factors** at $\mathbf{r}(\alpha)$.

Recall: For spherical coordinates $(u_1, u_2, u_3) = (r, \theta, \phi)$, the scale factors are

$$h_r \;=\; 1, \qquad\qquad h_\theta \;=\; r, \qquad\qquad h_\phi \;=\; r\sin\theta. \tag{122}$$

4.3.2 Multiple integrals and Jacobian

For a given function (scalar field) $f(\mathbf{r}) = f(x, y, z)$, we wish to evaluate the triple integral

$$\int_\Omega \mathrm{d}^3r\, f(\mathbf{r}) \tag{123}$$

over some volume region Ω.
Suppose we can find proper curvilinear coo's (u, v, w) such that

$$\Omega \;=\; \Big\{\mathbf{r}(u,v,w)\;\Big|\; a_1 \le u \le a_2, \quad b_1 \le v \le b_2, \quad c_1 \le w \le c_2\Big\}, \tag{124}$$

with some **constants** $a_{1,2}$, $b_{1,2}$, $c_{1,2}$. Then, we have

$$\int_\Omega \mathrm{d}^3r\, f(\mathbf{r}) \;=\; \int_{a_1}^{a_2} \mathrm{d}u \int_{b_1}^{b_2} \mathrm{d}v \int_{c_1}^{c_2} \mathrm{d}w\; J(u,v,w)\; \tilde{f}(u,v,w), \tag{125}$$

where

$$\tilde{f}(\alpha) = f\Big(x(\alpha),\; y(\alpha),\; z(\alpha)\Big) \tag{126}$$

and the **Jacobian** is generally given by

$$J(\alpha) \;=\; \big[\mathbf{h}_1(\alpha) \times \mathbf{h}_2(\alpha)\big] \cdot \mathbf{h}_3(\alpha) \;\equiv\; \det\begin{pmatrix} \partial x/\partial u & \partial x/\partial v & \partial x/\partial w \\ \partial y/\partial u & \partial y/\partial v & \partial y/\partial w \\ \partial z/\partial u & \partial z/\partial v & \partial z/\partial w \end{pmatrix}. \tag{127}$$

When the coo's (u, v, w) are **orthogonal**, this expression is simplified to

$$J(\alpha) \;=\; h_1(\alpha)\, h_2(\alpha)\, h_3(\alpha). \tag{128}$$

Ideas for a proof: With $\mathbf{r}(\alpha_0) = \mathbf{r}(u_0, v_0, w_0)$, we have for sufficiently small Δu

$$\begin{aligned} \Delta u \cdot h_1(\alpha_0) \;&\equiv\; \Delta u \cdot \lim_{\delta u \to 0} \frac{1}{\delta u}\Big|\mathbf{r}(u_0 + \delta u, v_0, w_0) - \mathbf{r}(u_0, v_0, w_0)\Big| \\ &\approx\; \Delta s_u, \end{aligned} \tag{129}$$

with the **arc length** Δs_u on the u-line between $\mathbf{r}(\alpha_0) \equiv \mathbf{r}(u_0, v_0, w_0)$ and $\mathbf{r}(u_0 + \Delta u, v_0, w_0)$.
Consequently, for sufficiently small Δu, Δv, and Δw, the product

$$\Delta u \Delta v \Delta w\, h_1(\alpha_0) h_2(\alpha_0) h_3(\alpha_0) \;\approx\; \Delta s_u \Delta s_v \Delta s_w \;\approx\; \Delta V \tag{130}$$

in case of orthogonal coo.'s becomes the volume of the small region (almost-cuboid)

$$\Delta\Omega \;=\; \Big\{\mathbf{r}(u,v,w)\;\Big|\; |u - u_0| \le \tfrac{1}{2}\Delta u, \quad |v - v_0| \le \tfrac{1}{2}\Delta v, \quad |w - w_0| \le \tfrac{1}{2}\Delta w\Big\} \tag{131}$$

surrounding the point $\mathbf{r}(\alpha_0)$, cf. Eq. (114) and red marking in the figure there.

For later purposes: For sufficiently small $\Delta u, \Delta v > 0$, the product

$$\Delta u \Delta v\, \big[\mathbf{h}_1(\alpha_0) \times \mathbf{h}_2(\alpha_0)\big] \;\approx\; \Delta\mathbf{A} \tag{132}$$

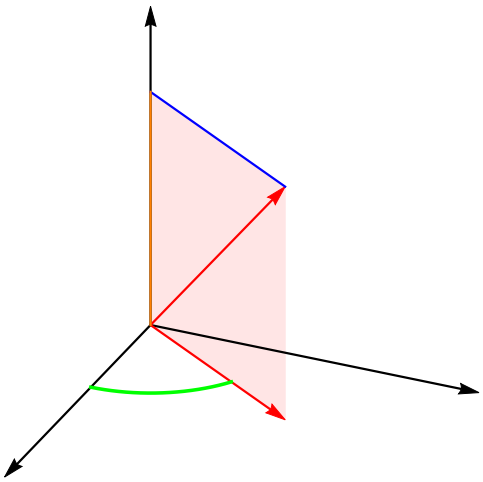
is that **vector**, which is orthogonal to the coordinate surface $w = w_0$ at the point $\mathbf{r}(\alpha_0)$ and whose magnitute $|\Delta\mathbf{A}|$ equals the area of the piece of surface, given by

$$\Delta\Sigma \;=\; \Big\{\mathbf{r}(u,v,w_0)\;\Big|\; |u - u_0| \le \tfrac{1}{2}\Delta u, \quad |v - v_0| \le \tfrac{1}{2}\Delta v\Big\}. \tag{133}$$

4.3.3 Further example: Cylindrical (polar) coordinates

These are $\alpha = (s, \phi, z)$, defined as

$s = \lvert\mathbf{r}_{xy}\rvert$	$0 \leq s < \infty$		Distance of P from the z-axis.
ϕ	$0 \leq \phi < 2\pi$	**Azimuth angle**	(the same as for spherical polars)
z	$-\infty < z < \infty$		the cartesian z coordinate



Again, we express the "old" coordinates (x, y, z) in terms of the "new" ones (s, ϕ, z),

$$\mathbf{r}(s, \phi, z) \equiv \begin{pmatrix} x(s, \phi, z) \\ y(s, \phi, z) \\ z(s, \phi, z) \end{pmatrix} = \begin{pmatrix} s\cos\phi \\ s\sin\phi \\ z \end{pmatrix}. \tag{134}$$

This is a vector function of **three** arguments (s, ϕ, and z).

- Keeping **two** arguments fixed, we are generating the **coordinate lines**:
  s-lines: Horizontal **rays**, ϕ-lines: Horizontal **circles**, z-lines: Vertical **lines**.
- Keeping only **one** argument fixed, we are generating the **coordinate surfaces**:
  Vertical cylindrical surfaces, vertical semi planes, and hoizontal planes.

In terms of the **tangential vectors**

$$\mathbf{h}_s(\alpha) = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix}, \qquad \mathbf{h}_\phi(\alpha) = \begin{pmatrix} -s\sin\phi \\ +s\cos\phi \\ 0 \end{pmatrix}, \qquad \mathbf{h}_z(\alpha) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{135}$$

(which are mutually **orthogonal**!), the **Jacobian** is

$$J(s, \phi, z) \;\equiv\; |\mathbf{h}_s(\alpha)||\mathbf{h}_\phi(\alpha)||\mathbf{h}_z(\alpha)| \;=\; s. \tag{136}$$

5 Vector fields

A **vector field** assigns to each point $\mathbf{r}$ in space a **vector**

$$\mathbf{v}(\mathbf{r}) \;\equiv\; \mathbf{v}(x,y,z) \;=\; \begin{pmatrix} v_1(x,y,z) \\ v_2(x,y,z) \\ v_3(x,y,z) \end{pmatrix}. \tag{137}$$

$\mathbf{v}(\mathbf{r})$ may be, at windy weather for example, the **velocity of flowing air** at position $\mathbf{r}$. Another example is the **electrostatic field** $\mathbf{E}(\mathbf{r})$ due to a charge distribution.

We also consider 2D vector fields (on the xy-plane),

$$\mathbf{v}(\mathbf{r}) \;\equiv\; \mathbf{v}(x,y) \;=\; \begin{pmatrix} v_1(x,y) \\ v_2(x,y) \end{pmatrix}. \tag{138}$$

5.1 Gradient fields als a special case

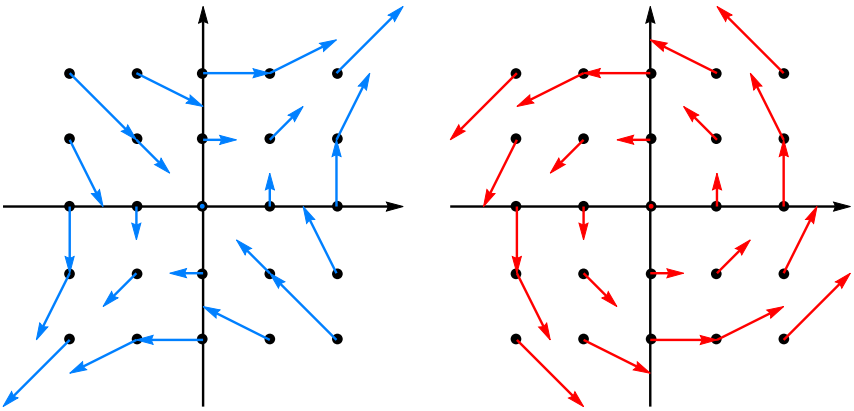
A particularly important example is the **gradient of a scalar field** $f(\mathbf{r}) \equiv f(x,y,z)$,

$$\mathbf{G}_f(\mathbf{r}) \;=\; \nabla f(\mathbf{r}) \;\equiv\; \begin{pmatrix} f_x(x,y,z) \\ f_y(x,y,z) \\ f_z(x,y,z) \end{pmatrix}. \tag{139}$$

Example: Consider the 2D vector field (blue in the figure below, NEW 2019 !!!)

$$\mathbf{p}(\mathbf{r}) \;\equiv\; \begin{pmatrix} p_1(x,y) \\ p_2(x,y) \end{pmatrix} \;=\; \begin{pmatrix} y \\ x \end{pmatrix}. \tag{140}$$

Obviously, $\mathbf{p}(\mathbf{r}) = \nabla f(\mathbf{r})$, with $f(x,y) = xy$.



In contrast, the quite similar 2D vector field (red in the figure)

$$\mathbf{q}(\mathbf{r}) \;\equiv\; \begin{pmatrix} q_1(x,y) \\ q_2(x,y) \end{pmatrix} \;=\; \begin{pmatrix} -y \\ x \end{pmatrix} \tag{141}$$

can obviously **not** be the gradient of any continuous 2D scalar field $f(\mathbf{r}) = f(x,y)$: The value $f(\mathbf{r})$ would increase constantly as the point $\mathbf{r}$ is moving on a circle centered at the origin $\mathbf{r} = \mathbf{0}$. After one full turn, this would conflict with the continuity of f. $\square$

• When a 2D vector field $\mathbf{v}(\mathbf{r})$ shall be gradient of some scalar field $f(\mathbf{r})$,

$$\mathbf{v}(\mathbf{r}) \;\equiv\; \begin{pmatrix} v_1(x,y) \\ v_2(x,y) \end{pmatrix} \;=\; \nabla f(\mathbf{r}) \;\equiv\; \begin{pmatrix} \frac{\partial}{\partial x} f(x,y) \\ \frac{\partial}{\partial y} f(x,y) \end{pmatrix}, \tag{142}$$

then its components v_1 and v_2 must satisfy the **necessary condition**

$$\frac{\partial}{\partial y}\, v_1(x,y) \;=\; \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\, f(x,y)\right] \;\equiv\; \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\, f(x,y)\right] \;=\; \frac{\partial}{\partial x}\, v_2(x,y). \tag{143}$$

Example: For the above fields $\mathbf{p}(\mathbf{r})$ and $\mathbf{q}(\mathbf{r})$, we have

$$\frac{\partial}{\partial y}\, p_1(x,y) \;=\; 1 \;=\; \frac{\partial}{\partial x}\, p_2(x,y), \qquad \frac{\partial}{\partial y}\, q_1(x,y) \;=\; -1 \;\neq\; 1 \;=\; \frac{\partial}{\partial x}\, q_2(x,y), \tag{144}$$

reveiling immediately that $\mathbf{p}(\mathbf{r})$ is a gradient field, while $\mathbf{q}(\mathbf{r})$ is not.

• For a 3D vector field, Eq. (137), we obtain **three necessary conditions**,

$$\frac{\partial v_2}{\partial x} = \frac{\partial v_1}{\partial y}, \qquad \frac{\partial v_3}{\partial y} = \frac{\partial v_2}{\partial z}, \qquad \frac{\partial v_1}{\partial z} = \frac{\partial v_3}{\partial x}. \tag{145}$$

Using the nabla operator ∇ and writing $(x,y,z) = (x_1,x_2,x_3)$, $\frac{\partial}{\partial x_n} = \partial_n$, these conditions are summarized to a single **vector condition**,

$$\nabla \times \mathbf{v}(\mathbf{r}) \;\equiv\; \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \;\equiv\; \begin{pmatrix} \partial_2 v_3 - \partial_3 v_2 \\ \partial_3 v_1 - \partial_1 v_3 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix} \;=\; \mathbf{0}. \tag{146}$$

The **vector field** $\nabla \times \mathbf{v}(\mathbf{r})$ is called the **curl** of the vector field $\mathbf{v}(\mathbf{r})$.

Remark 1: A 2D vector field is a 3D one with a zero third component,

$$\begin{pmatrix} v_1(x,y) \\ v_2(x,y) \end{pmatrix} \to \begin{pmatrix} v_1(x,y) \\ v_2(x,y) \\ 0 \end{pmatrix}, \quad \nabla \times \mathbf{v}(\mathbf{r}) = \begin{pmatrix} 0 \\ 0 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix}. \tag{147}$$

Remark 2: In section 6.4, we shall see that, under mild assumptions, the condition $\nabla \times \mathbf{v}(\mathbf{r}) \equiv \mathbf{0}$ is even **sufficient** for $\mathbf{v}(\mathbf{r})$ to be a gradient field.

5.2 Line integrals

In a uniform electric filed $\mathbf{E}$ (plate capacitor), a point charge q (mass m) experiences the force $\mathbf{F} = q\mathbf{E}$. As q is moving without friction on a **straight** track from position $\mathbf{r}_\mathrm{A}$ to position $\mathbf{r}_\mathrm{B}$, its velocities $v_\mathrm{A,B}$ at these positions are related by

$$\frac{m}{2}\, v_\mathrm{B}^2 \;=\; \frac{m}{2}\, v_\mathrm{A}^2 \;+\; W_\mathrm{AB}, \tag{148}$$

where W_AB is the **work** done by the force $\mathbf{F}$,

$$W_\mathrm{AB} \;=\; \Big(|\mathbf{F}|\cos\gamma\Big)|\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}| \;\equiv\; \mathbf{F}\cdot\big[\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}\big]. \tag{149}$$

Note: $|\mathbf{F}|\cos\gamma$ is the component of $\mathbf{F}$ in direction of the vector $\Delta\mathbf{r} = \mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}$.

More generally, let $\mathbf{E}(\mathbf{r})$ be a **non-uniform vector field**. Then, both magnitude and direction of the force vector $\mathbf{F} = q\mathbf{E}(\mathbf{r}) = \mathbf{F}(\mathbf{r})$ will change as q is moving. Furthermore, the track may have the shape of some **curve** Γ, given by the **vector function**

$$\mathbf{r}(u) \;\equiv\; \begin{pmatrix} x(u) \\ y(u) \\ z(u) \end{pmatrix}, \qquad \Gamma \;=\; \Big\{\mathbf{r}(u) \;\Big|\; u_{\rm A} \leq u \leq u_{\rm B}\Big\}. \tag{150}$$

We partition the interval $[u_{\rm A}, u_{\rm B}]$ into N pieces with equal lengths $\Delta u = \frac{u_{\rm B}-u_{\rm A}}{N}$,

$$u_n = u_{\rm A} + n\Delta u \qquad (n = 0, 1, 2, ..., N). \tag{151}$$

Consequently, as q is moving on Γ from $\mathbf{r}_{\rm A} = \mathbf{r}(u_{\rm A})$ to $\mathbf{r}_{\rm B} = \mathbf{r}(u_{\rm B})$, the work done by the force field $\mathbf{F}(\mathbf{r})$ is

$$W_{\rm AB} \;=\; \lim_{N\to\infty} \sum_{n=1}^{N} \mathbf{F}\big(\mathbf{r}(u_n)\big) \cdot \big[\mathbf{r}(u_n) - \mathbf{r}(u_{n-1})\big]. \tag{152}$$

As $\Delta u \to 0$ ($N \to \infty$), we obtain

$$\mathbf{r}(u_n) - \mathbf{r}(u_{n-1}) \;\equiv\; \begin{pmatrix} x(u_n) - x(u_{n-1}) \\ y(u_n) - y(u_{n-1}) \\ z(u_n) - z(u_{n-1}) \end{pmatrix} \;\to\; \begin{pmatrix} \dot{x}(u_n) \\ \dot{y}(u_n) \\ \dot{z}(u_n) \end{pmatrix} \Delta u \;=\; \dot{\mathbf{r}}(u_n)\,\Delta u, \tag{153}$$

where the derivatives of $x(u)$, etc. are denoted as $\dot{x}(u)$. Eventually,

$$W_{\rm AB} \;=\; \lim_{N\to\infty} \sum_{n=1}^{N} \Big[\mathbf{F}\big(\mathbf{r}(u_n)\big) \cdot \dot{\mathbf{r}}(u_n)\Big] \Delta u \;=\; \int_{u_{\rm A}}^{u_{\rm B}} {\rm d}u \,\Big[\underbrace{\dot{\mathbf{r}}(u) \cdot \mathbf{F}\big(\mathbf{r}(u)\big)}_{f(u)}\Big]. \tag{154}$$

This is an **ordinary integral** over one real variable u.

Def.: The **line integral** of the vector field $\mathbf{F}(\mathbf{r})$ along the curve Γ is defined as

$$\int_{\Gamma} {\rm d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;:=\; \int_{u_{\rm A}}^{u_{\rm B}} {\rm d}u\, \dot{\mathbf{r}}(u) \cdot \mathbf{F}\big(\mathbf{r}(u)\big) \qquad (u_A < u_B). \tag{155}$$

While its **absolute value** is independent of the parametrization $\mathbf{r}(u)$ of Γ, its **sign** depends on which one of the two endpoints of Γ is chosen as the "starting point A".

Remark: When Γ is a **closed loop**, with $\mathbf{r}_B = \mathbf{r}_A$, the integral symbol may be decorated with a circle,

$$\int_{\Gamma} {\rm d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;\equiv\; \oint_{\Gamma} {\rm d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}). \tag{156}$$

Example: We consider the line integral of the vector field

$$\mathbf{F}(\mathbf{r}) \;=\; \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \tag{157}$$

(with constants $a, b, c, d \in \mathbb{R}$) along a circle Γ of radius R in the xy-plane,

$$\Gamma \;=\; \Big\{\mathbf{r}(\phi) \;\Big|\; 0 \leq \phi \leq 2\pi\Big\}, \qquad \mathbf{r}(\phi) = \begin{pmatrix} x(\phi) \\ y(\phi) \end{pmatrix} = \begin{pmatrix} R\cos\phi \\ R\sin\phi \end{pmatrix}. \tag{158}$$

In this case (closed loop!), we have

$$\dot{\mathbf{r}}(u) = \begin{pmatrix} -R\sin\phi \\ R\cos\phi \end{pmatrix}. \tag{159}$$

Consequently, with $u_A \equiv \phi_\mathrm{A} = 0$ and $u_B \equiv \phi_\mathrm{B} = 2\pi$, Eq. (155) yields

$$\begin{aligned}
\oint_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}) &= \int_{\phi_\mathrm{A}}^{\phi_\mathrm{B}} \mathrm{d}\phi\,\dot{\mathbf{r}}(\phi)\cdot\mathbf{F}\big(\mathbf{r}(\phi)\big) \\
&= \int_0^{2\pi} \mathrm{d}\phi \begin{pmatrix} -R\sin\phi \\ R\cos\phi \end{pmatrix} \cdot \begin{pmatrix} aR\cos\phi + bR\sin\phi \\ cR\cos\phi + dR\sin\phi \end{pmatrix} \\
&= R^2 \int_0^{2\pi} \mathrm{d}\phi \Big\{ (d-a)\sin\phi\cos\phi - b\sin^2\phi + c\cos^2\phi \Big\} \\
&= R^2 \left\{ (d-a)\left[\frac{\sin^2\phi}{2}\right]_0^{2\pi} - b\left[\frac{\phi - \sin^2\phi\cos\phi}{2}\right]_0^{2\pi} + c\left[\frac{\phi + \sin^2\phi\cos\phi}{2}\right]_0^{2\pi} \right\} \\
&= R^2\pi(c-b).
\end{aligned} \tag{160}$$

Remarkably, this result does not depend on the values of a and d.

Choosing $a = d = 0$, $c = 1$ and $b = \pm 1$, respectively, the vector fields $\mathbf{p}(\mathbf{r}) = \binom{y}{x}$ and $\mathbf{q}(\mathbf{r}) = \binom{-y}{x}$ of section 5.1 are recovered [where $\mathbf{p}(\mathbf{r}) = \nabla xy$ is a gradient], implying that

$$\oint_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{p}(\mathbf{r}) \;=\; 0, \qquad \oint_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{q}(\mathbf{r}) \;=\; 2R^2\pi. \tag{161}$$

The first one of these two special results has a simple reason:

Theorem: For the gradient $\mathbf{G}_f(\mathbf{r})$ of a scalar field, $\mathbf{G}_f(\mathbf{r}) = \nabla f(\mathbf{r})$, the line integral is

$$\int_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{G}_f(\mathbf{r}) \;\equiv\; \int_\Gamma \mathrm{d}\mathbf{r}\cdot\nabla f(\mathbf{r}) \;=\; f(\mathbf{r}_B) - f(\mathbf{r}_A), \tag{162}$$

where $\mathbf{r}_\mathrm{A} = \mathbf{r}(u_A)$ is the starting and $\mathbf{r}_\mathrm{B} = \mathbf{r}(u_B)$ the end point of Γ.
Proof: The chain rule for partial derivatives, $\frac{\mathrm{d}}{\mathrm{d}u} f(\mathbf{r}(u)) = \dot{\mathbf{r}}(u)\cdot\mathbf{G}_f(\mathbf{r}(u))$, implies

$$\begin{aligned}
\int_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{G}_f(\mathbf{r}) &\equiv \int_{u_\mathrm{A}}^{u_\mathrm{B}} \mathrm{d}u\,\dot{\mathbf{r}}(u)\cdot\mathbf{G}_f\big(\mathbf{r}(u)\big) \\
&= \int_{u_\mathrm{A}}^{u_\mathrm{B}} \mathrm{d}u\,\frac{\mathrm{d}}{\mathrm{d}u} f\big(\mathbf{r}(u)\big) \;=\; f\big(\mathbf{r}(u_\mathrm{B})\big) - f\big(\mathbf{r}(u_\mathrm{A})\big).
\end{aligned} \tag{163}$$

In other words: The line integral of a **gradient field** only depends on the end points of the curve, but not on its shape. Consequently, for a **closed loop** Γ (with $\mathbf{r}_\mathrm{A} = \mathbf{r}_\mathrm{B}$),

$$\oint_\Gamma \mathrm{d}\mathbf{r}\cdot\nabla f(\mathbf{r}) \;=\; 0. \tag{164}$$

Remark: When the **arc length** s is used as the curve parameter, then $\dot{\mathbf{r}}(s) = \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{r}(s)$ is the **unit** vector $\boldsymbol{\mathcal{T}}(\mathbf{r})$ tangential to Γ at the point $\mathbf{r} = \mathbf{r}(s) \in \Gamma$. Therefore,

$$\int_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}) \;\equiv\; \int_{s_\mathrm{A}}^{s_\mathrm{B}} \mathrm{d}s\,\dot{\mathbf{r}}(s)\cdot\mathbf{F}\big(\mathbf{r}(s)\big) \;=\; L_\Gamma \left\langle \boldsymbol{\mathcal{T}}(\mathbf{r})\cdot\mathbf{F}(\mathbf{r}) \right\rangle_{\mathbf{r}\in\Gamma} \tag{165}$$

or, in words: The line integral equals the average value of the tangential component $\boldsymbol{\mathcal{T}}\cdot\mathbf{F}$ of the vector field $\mathbf{F}(\mathbf{r})$ on the curve Γ times the length $L_\Gamma = s_B - s_A$ of Γ.

5.3 Flux (or surface) integrals

We consider **flowing air** with uniform and time-independent density $\rho(\mathbf{r}) = \rho_0$, e.g.,

$$\rho_0 \;\approx\; \frac{0.7 \cdot 28\,\mathrm{kg} + 0.3 \cdot 32\,\mathrm{kg}}{22\,400\,\ell} \;=\; \frac{29.2\,\mathrm{kg}}{22.4\,\mathrm{m}^3} \;=\; 1.30\,\frac{\mathrm{kg}}{\mathrm{m}^3}. \tag{166}$$

The flow velocity shall also be uniform, $\mathbf{v}(\mathbf{r}) = \mathbf{v}_0$, e.g. with $|\mathbf{v}_0| = 2.00\,\frac{\mathrm{m}}{\mathrm{s}}$. The vector

$$\mathbf{J} \;=\; \rho_0 \mathbf{v}_0 \;=\; 2.60\,\mathbf{u}_J\,\frac{\mathrm{kg}}{\mathrm{m}^2\mathrm{s}} \qquad \Big(\mathbf{u}_J := \frac{\mathbf{v}_0}{|\mathbf{v}_0|}\Big) \tag{167}$$

is called **current density** or **flux**. Its magnitude $J = |\mathbf{J}|$ gives the amount of air (in kg) flowing per unit time (1 s) through a unit area (1 m^2) which has $\mathbf{J}$ as a normal vector.

Consequently, the amount of air flowing per unit time through a planar piece of surface $\varSigma$, with area A and unit normal vector $\boldsymbol{\mathcal{N}}$, is given by the scalar product

$$\dot{M} \;=\; \big(A\boldsymbol{\mathcal{N}}\big) \cdot \mathbf{J} \;=\; AJ\,\cos\gamma, \tag{168}$$

where γ is the angle between the vectors $\mathbf{J}$ and $\mathbf{n}$. Note that $\dot{M}$ changes sign when the vector $\boldsymbol{\mathcal{N}}$ is replaced with $-\boldsymbol{\mathcal{N}}$. The two possible choices $\pm\boldsymbol{\mathcal{N}}$ are called the **orientations** of $\varSigma$. If its orientation is specified, $\varSigma$ is called **orientated**.

We wish to generalize this result to a flow with **non-uniform** current density, represented by a vector field $\mathbf{J}(\mathbf{r})$, through a **curved** (and orientated[2]) piece of surface

$$\varSigma \;=\; \Big\{\mathbf{r}(u,v,w_0)\;\Big|\; a_1 \le u \le a_2,\; b_1 \le v \le b_2\Big\}, \tag{169}$$

represented by a proper set $(u,v,w) \equiv \alpha$ of curvilinear coordinates.

Let $\boldsymbol{\mathcal{N}}(u,v)$ be the unit normal vector to $\varSigma$ at the point $\mathbf{r}(u,v,w_0) \in \varSigma$ and let

$$\delta A \;=\; \delta u\,\delta v\,A(u,v) \tag{170}$$

be the area of the small piece $\delta\varSigma$ of $\varSigma$, consisting of the points $\mathbf{r}(u',v',w_0)$ with

$$u \le u' \le u + \delta u, \qquad v \le v' \le v + \delta v. \tag{171}$$

Then, the wanted generalization of Eq. (168) is obviously given by the double integral

$$\dot{M} \;=\; \int_{a_1}^{a_2} \mathrm{d}u \int_{b_1}^{b_2} \mathrm{d}v\,A(u,v)\,\boldsymbol{\mathcal{N}}(u,v) \cdot \mathbf{J}\big(\mathbf{r}(u,v,w_0)\big). \tag{172}$$

Generally, it can be shown (see section 4.3) that

$$A(u,v)\,\boldsymbol{\mathcal{N}}(u,v) \;=\; \pm\mathbf{h}_u(u,v,w_0) \times \mathbf{h}_v(u,v,w_0) \tag{173}$$

(with a sign depending on the orientation of $\varSigma$) and we obtain the compact formula

$$\dot{M} \;=\; \pm\int_{a_1}^{a_2} \mathrm{d}u \int_{b_1}^{b_2} \mathrm{d}v\,\big[\mathbf{h}_u(\alpha) \times \mathbf{h}_v(\alpha)\big] \cdot \mathbf{J}\big(\mathbf{r}(\alpha)\big)\Big|_{w=w_0} \;\equiv\; \int_{\varSigma} \mathrm{d}\mathbf{A} \cdot \mathbf{J}(\mathbf{r}). \tag{174}$$

- This is called the **flux (or surface) integral** of the vector field $\mathbf{J}(\mathbf{r})$ through $\varSigma$.
- $\mathrm{d}\mathbf{A}$, the **surface vector element** of $\varSigma$, represents the product $A(u,v)\,\boldsymbol{\mathcal{N}}(u,v)\,\mathrm{d}u\,\mathrm{d}v$.

[2] There are curved surfaces $\varSigma$ that cannot be orientated, e.g., a Moebius tape!

Remark: An alternative notation for flux integrals is

$$\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;\equiv\; \int_\Sigma \mathrm{d}^2r\,\boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\mathbf{F}(\mathbf{r}), \tag{175}$$

where $\boldsymbol{\mathcal{N}}(\mathbf{r})$ is the unit vector normal to Σ at the point $\mathbf{r}\in\Sigma$. In fact, one can show that

$$\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;=\; A_\Sigma\,\big\langle\boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\mathbf{F}(\mathbf{r})\big\rangle_{\mathbf{r}\in\Sigma}. \tag{176}$$

This is the average of the normal component $\boldsymbol{\mathcal{N}}\cdot\mathbf{F}$ of $\mathbf{F}(\mathbf{r})$ on Σ times the area A_Σ of Σ.

Example: Choosing spherical polar coordinates,

$$\mathbf{r}(r,\theta,\phi) \;=\; \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} \tag{177}$$

we compute the flux of the following vector field $\mathbf{F}(\mathbf{r})$ **out of** the following surface Σ,

$$\mathbf{F}(\mathbf{r}) \;=\; \begin{pmatrix} Ax+By+Cz \\ Lx+My+Nz \\ Ux+Vy+Wz \end{pmatrix}, \qquad \Sigma \;=\; \Big\{\mathbf{r}(R,\theta,\phi)\;\Big|\;0\le\theta\le\beta,\quad 0\le\phi\le 2\pi\Big\} \tag{178}$$

(Σ is the region "north" from the "latitude" $\theta=\beta$ on a "planet" with radius R),

$$\begin{aligned}
\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;&=\; \int_0^\beta \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,\big[\mathbf{h}_\theta(\alpha)\times\mathbf{h}_\phi(\alpha)\big]\cdot\mathbf{F}\big(\mathbf{r}(r,\theta,\phi)\big)\Big|_{r=R} \\
=\; \int_0^\beta \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,&\left[\begin{pmatrix} R\cos\theta\cos\phi \\ R\cos\theta\sin\phi \\ -R\sin\theta \end{pmatrix}\times\begin{pmatrix} -R\sin\theta\sin\phi \\ +R\sin\theta\cos\phi \\ 0 \end{pmatrix}\right]\cdot\mathbf{F}\big(\mathbf{r}(R,\theta,\phi)\big) \\
=\; \int_0^\beta \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,&\begin{pmatrix} R^2\sin^2\theta\cos\phi \\ R^2\sin^2\theta\sin\phi \\ R^2\cos\theta\sin\theta \end{pmatrix}\cdot\begin{pmatrix} R\sin\theta(A\cos\phi+B\sin\phi)+RC\cos\theta \\ R\sin\theta(L\cos\phi+M\sin\phi)+RN\cos\theta \\ R\sin\theta(U\cos\phi+V\sin\phi)+RW\cos\theta \end{pmatrix}.
\end{aligned}$$

It is easy to see that the integral $\int_0^{2\pi}\mathrm{d}\phi...$ yields zero for the B,C,L,N,U, and V terms. Only the A,M and W terms do contribute, and we are left with

$$\begin{aligned}
... \;&=\; R^3\int_0^\beta \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\Big[A\sin^3\theta\cos^2\phi \;+\; M\sin^3\theta\sin^2\phi \;+\; W\cos^2\theta\sin\theta\Big] \\
&=\; R^3\left[\pi(A+M)\int_0^\beta \mathrm{d}\theta\sin^3\theta \;+\; 2\pi W\int_0^\beta \mathrm{d}\theta\cos^2\theta\sin\theta\right]
\end{aligned} \tag{179}$$

Writing $\sin^3\theta=\sin\theta(1-\cos^2\theta)$, we obtain

$$\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;=\; \pi R^3(A+M)\left[-\cos\theta+\frac{1}{3}\cos^3\theta\right]_0^\beta + 2\pi R^3 W\left[-\frac{1}{3}\cos^3\theta\right]_0^\beta. \tag{180}$$

In the case $A=M=W$ (and $B,C,L,N,U,V=0$), this result can be found even without integration: Since the "current density" $\mathbf{F}(\mathbf{r})=W\mathbf{r}$ is **perpendicular** to Σ at any point $\mathbf{r}\in\Sigma$, with **constant magnitude** $|\mathbf{F}|=WR$ for $\mathbf{r}\in\Sigma$, the flux (integral) is simply the product of the area A_Σ of Σ with this magnitude WR,

$$\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;=\; A_\Sigma WR. \tag{181}$$

In the case $\beta=\pi$, Σ is the full spherical surface, and Eq. (180) is simplified to

$$\oint_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;=\; \frac{4\pi}{3}R^3\,(A+M+W). \tag{182}$$

6 Integral theorems

6.1 Integral types in vector calculus, revisited

Consider a set of **orthogonal curvilinear coordinates** $\alpha = (u, v, w)$, given by

$$\mathbf{r}(u, v, w) \;\equiv\; \mathbf{r}(\alpha) \;=\; \begin{pmatrix} x(u, v, w) \\ y(u, v, w) \\ z(u, v, w) \end{pmatrix}. \tag{183}$$

Tangent vectors to the coordinate curves through $\mathbf{r}(\alpha)$ are given by

$$\mathbf{h}_u(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial u}, \qquad \mathbf{h}_v(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial v}, \qquad \mathbf{h}_w(\alpha) = \frac{\partial \mathbf{r}(\alpha)}{\partial w}. \tag{184}$$

Their magnitudes $h_u(\alpha) = |\mathbf{h}_u(\alpha)|$, etc. are called **scale factors**.

In terms of the coordinates $\alpha = (u, v, w)$, a **volume region** Ω, a piece of **surface** Σ, and a **curve** Γ can be defined by

$$\begin{aligned} \Omega &= \Big\{ \mathbf{r}(\alpha) \,\Big|\, u_1 \le u \le u_2,\; v_1 \le v \le v_2,\; w_1 \le w \le w_2 \Big\}, \\ \Sigma &= \Big\{ \mathbf{r}(\alpha) \,\Big|\, u_1 \le u \le u_2,\; v_1 \le v \le v_2,\; w = w_0 \Big\}, \\ \Gamma &= \Big\{ \mathbf{r}(\alpha) \,\Big|\, u_1 \le u \le u_2,\; v = v_0,\; w = w_0 \Big\}. \end{aligned} \tag{185}$$

We now consider a **scalar field** $f(\mathbf{r})$ and a **vector field** $\mathbf{F}(\mathbf{r})$,

$$f(\mathbf{r}) \;=\; f(x, y, z), \qquad \mathbf{F}(\mathbf{r}) \;=\; \begin{pmatrix} F_1(x, y, z) \\ F_2(x, y, z) \\ F_3(x, y, z) \end{pmatrix}. \tag{186}$$

Replacing $\mathbf{r}$ with $\mathbf{r} = \mathbf{r}(\alpha)$, we obtain functions of $\alpha = (u, v, w)$,

$$f\big(\mathbf{r}(\alpha)\big) = \tilde{f}(u, v, w), \qquad \mathbf{F}\big(\mathbf{r}(\alpha)\big) = \begin{pmatrix} \widetilde{F}_1(u, v, w) \\ \widetilde{F}_1(u, v, w) \\ \widetilde{F}_1(u, v, w) \end{pmatrix}. \tag{187}$$

The **volume integral** of the scalar field $f(\mathbf{r})$ over Ω is evaluated as

$$\int_\Omega \mathrm{d}^3 r\, f(\mathbf{r}) \;=\; \int_{u_1}^{u_2} \mathrm{d}u \int_{v_1}^{v_2} \mathrm{d}v \int_{w_1}^{w_2} \mathrm{d}w \, \big[h_u(\alpha) h_v(\alpha) h_w(\alpha)\big]\, f\big(\mathbf{r}(\alpha)\big), \tag{188}$$

where $\big[h_u(\alpha) h_v(\alpha) h_w(\alpha)\big] \equiv J(u, v, w)$ is the **Jacobian** of the coordinates $\alpha = (u, v, w)$.
The **flux integral** of the vector field $\mathbf{F}(\mathbf{r})$ through Σ is evaluated as

$$\int_\Sigma \mathrm{d}\mathbf{A} \cdot \mathbf{F}(\mathbf{r}) \;=\; \int_{u_1}^{u_2} \mathrm{d}u \int_{v_1}^{v_2} \mathrm{d}v \, \big[\mathbf{h}_u(\alpha) \times \mathbf{h}_v(\alpha)\big] \cdot \mathbf{F}\big(\mathbf{r}(\alpha)\big)\Big|_{w=w_0}, \tag{189}$$

where the **orientation** of Σ is fixed by the direction of $\big[\mathbf{h}_u(\alpha) \times \mathbf{h}_v(\alpha)\big]$.
The **line integral** of the vector field $\mathbf{F}(\mathbf{r})$ along Γ is evaluated as

$$\int_\Gamma \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;=\; \int_{u_1}^{u_2} \mathrm{d}u \, \mathbf{h}_u(\alpha) \cdot \mathbf{F}\big(\mathbf{r}(\alpha)\big)\Big|_{v=v_0,\; w=w_0}. \tag{190}$$

It is instructive to notice that, in terms of the volume V_Ω of Ω, the surface area A_Σ of Σ, and the length L_Γ of Γ, we may write

$$\int_\Omega \mathrm{d}^3r\, f(\mathbf{r}) \;=\; V_\Omega \left\langle f(\mathbf{r})\right\rangle_{\mathbf{r}\in\Omega}, \tag{191}$$

$$\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) \;=\; A_\Sigma \left\langle \boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\mathbf{F}(\mathbf{r})\right\rangle_{\mathbf{r}\in\Sigma}, \tag{192}$$

$$\int_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}) \;=\; L_\Gamma \left\langle \boldsymbol{\mathcal{T}}(\mathbf{r})\cdot\mathbf{F}(\mathbf{r})\right\rangle_{\mathbf{r}\in\Gamma}, \tag{193}$$

where $\boldsymbol{\mathcal{N}}(\mathbf{r})$ is the unit normal vector at $\mathbf{r}\in\Sigma$ and $\boldsymbol{\mathcal{T}}(\mathbf{r})$ the unit tangent vector at $\mathbf{r}\in\Gamma$. In practice, the quantities within $\langle ...\rangle$ are often **constant** (independent of $\mathbf{r}$).

6.2 Divergence of a vector field and Gauss' theorem

Def.: The **divergence** of a vector field $\mathbf{F}(\mathbf{r})$ is the scalar field $D_\mathbf{F}(\mathbf{r})$, given by

$$D_\mathbf{F}(\mathbf{r}_0) \;=\; \lim_{a\to 0}\frac{1}{a^3}\oint_{\partial\Omega_a}\mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}), \tag{194}$$

where Ω_a is a cube, centered at $\mathbf{r}=\mathbf{r}_0$, with edges of length a in x-, y-, and z-directions.
Remark 1: The limit $D_\mathbf{F}(\mathbf{r}_0)$ does not change when the cube is rotated.
Remark 2: In usual notation, $D_\mathbf{F}(\mathbf{r}) \;\equiv\; \mathrm{div}\,\mathbf{F}(\mathbf{r})$.

Corollary (Gauss' theorem): For any finite volume region Ω, we have

$$\int_\Omega \mathrm{d}^3r\, D_\mathbf{F}(\mathbf{r}) \;=\; \oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}). \tag{195}$$

Lemma: In cartesian coordinates, $\mathbf{F}(\mathbf{r}) = \begin{pmatrix} F_1(x,y,z)\\ F_2(x,y,z)\\ F_3(x,y,z)\end{pmatrix}$, the divergence is given by

$$D_\mathbf{F}(\mathbf{r}) \;=\; \frac{\partial F_1(\mathbf{r})}{\partial x}+\frac{\partial F_2(\mathbf{r})}{\partial y}+\frac{\partial F_3(\mathbf{r})}{\partial z} \;\equiv\; \nabla\cdot\mathbf{F}(\mathbf{r}). \tag{196}$$

Proof: Exercises!

Example: Consider flowing CO_2 gas in air. At any position $\mathbf{r}$ and at any time t, let be:
- $\rho(\mathbf{r},t)$ the mass density (in $\frac{\mathrm{g}}{\mathrm{m}^3}$) of CO_2,
- $\mathbf{J}(\mathbf{r},t)$ the current density (in $\frac{\mathrm{g}}{\mathrm{m}^2\mathrm{s}}$) of the CO_2 flow,
- $s(\mathbf{r},t)$ the source density (in $\frac{\mathrm{g}}{\mathrm{m}^3\mathrm{s}}$) of chemical CO_2 production.

Then, for any volume region Ω (with surface $\partial\Omega$), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t) \;=\; \int_\Omega \mathrm{d}^3r\, s(\mathbf{r},t) \;-\; \oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t). \tag{197}$$

For the special case of **stationary flow**, $f(\mathbf{r},t)=f(\mathbf{r})$ where $f\in\{\rho,\mathbf{J},s\}$, we conclude

$$\int_\Omega \mathrm{d}^3r\, s(\mathbf{r}) \;=\; \oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r}). \tag{198}$$

Since this equation holds for any volume region $\Omega\subseteq\mathbb{R}^3$, Gauss' theorem implies

$$s(\mathbf{r}) \;=\; D_\mathbf{J}(\mathbf{r}) \;\equiv\; \nabla\cdot\mathbf{J}(\mathbf{r}). \tag{199}$$

Therefore, the divergence $D_\mathbf{F}(\mathbf{r})$ of a vector field $\mathbf{F}(\mathbf{r})$ is often called its **source density**.

Alternative version of subsection 6.2

Consider flowing CO_2 gas in air. At any position $\mathbf{r}$ and at any time t, let be:
• $\rho(\mathbf{r},t)$ the mass density (in $\frac{\mathrm{g}}{\mathrm{m}^3}$) of CO_2,
• $\mathbf{J}(\mathbf{r},t)$ the current density (in $\frac{\mathrm{g}}{\mathrm{m}^2\mathrm{s}}$) of the CO_2 flow,
• $s(\mathbf{r},t)$ the source density (in $\frac{\mathrm{g}}{\mathrm{m}^3\mathrm{s}}$) of chemical CO_2 production.
Then, for any volume region Ω with (closed) surface $\partial\Omega$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t) \;=\; \int_\Omega \mathrm{d}^3r\,s(\mathbf{r},t) \;-\; \oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t). \tag{200}$$

For the special case of **stationary flow**, $f(\mathbf{r},t)=f(\mathbf{r})$ where $f\in\{\rho,\mathbf{J},s\}$, we obtain

$$\int_\Omega \mathrm{d}^3r\,s(\mathbf{r}) \;=\; \oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r}). \tag{201}$$

Consequently, the scalar field $s(\mathbf{r})$ is, in terms of the vector field $\mathbf{J}(\mathbf{r})$, given by

$$\begin{aligned} s(\mathbf{r}_0) \;\equiv\; \lim_{V_\Omega\to 0}\big\langle s(\mathbf{r})\big\rangle_{\mathbf{r}\in\Omega} \;&\equiv\; \lim_{V_\Omega\to 0}\frac{1}{V_\Omega}\int_\Omega \mathrm{d}^3r\,s(\mathbf{r}) \\ &=\; \lim_{V_\Omega\to 0}\frac{1}{V_\Omega}\oint_{\partial\Omega_a} \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r}), \end{aligned} \tag{202}$$

where V_Ω is the volume of Ω (with $\mathbf{r}_0\in\Omega$; SKETCH!). More precisely, we define

Def.: Let Ω_a be the cube with center $\mathbf{r}_0$ and edges of length a in x-, y-, and z-directions. The **divergence** of a vector field $\mathbf{F}(\mathbf{r})$ is the scalar field $D_{\mathbf{F}}(\mathbf{r})$, given by

$$D_{\mathbf{F}}(\mathbf{r}_0) \;=\; \lim_{a\to 0}\frac{1}{a^3}\oint_{\partial\Omega_a} \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}). \tag{203}$$

Remark 1: The number $D_{\mathbf{F}}(\mathbf{r}_0)$ is not affected by a rotation of the x-, y-, and z-axes.
Remark 2: In usual notation, $D_{\mathbf{F}}(\mathbf{r}) \;\equiv\; \mathrm{div}\,\mathbf{F}(\mathbf{r})$.

Lemma: For a differentiable vector field $\mathbf{F}(\mathbf{r})$, the divergence is given by

$$D_{\mathbf{F}}(\mathbf{r}) \;=\; \frac{\partial F_1(\mathbf{r})}{\partial x}+\frac{\partial F_2(\mathbf{r})}{\partial y}+\frac{\partial F_3(\mathbf{r})}{\partial z} \;\equiv\; \nabla\cdot\mathbf{F}(\mathbf{r}). \tag{204}$$

Proof: Exercises!

Consequently, during stationary flow, source and current densities are related by

$$s(\mathbf{r}) \;=\; D_{\mathbf{J}}(\mathbf{r}) \;\equiv\; \nabla\cdot\mathbf{J}(\mathbf{r}). \tag{205}$$

Therefore, the divergence $D_{\mathbf{F}}(\mathbf{r})$ of a vector field $\mathbf{F}(\mathbf{r})$ is often called its **source density**. Generalizing Eq. (201), one can show:

Theorem (Gauss): For a continuously differentiable vector field $\mathbf{F}(\mathbf{r})$, we have

$$\int_\Omega \mathrm{d}^3r\big[\nabla\cdot\mathbf{F}(\mathbf{r})\big] \;=\; \oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}), \tag{206}$$

for any compact volume region $\Omega\subset\mathbb{R}^3$ with piecewise smooth boundary $\partial\Omega$.
(Of course, the field $\mathbf{F}$ must be defined on a neighborhood of Ω.)

6.3 Curl of a vector field and Stokes' theorem

Let Σ be a piece of surface, e.g. a planar **disk** or a curved **saddle**. By $\partial\Sigma$, we denote the **closed curve** bordering Σ (rim or edge of Σ). The line integral

$$\oint_{\partial\Sigma} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \tag{207}$$

is called the **circulation** of the vector field $\mathbf{F}(\mathbf{r})$ around Σ. (To fix the sign of this integral, the orientation of Σ must be specified.)

While the **flux** of $\mathbf{F}(\mathbf{r})$ out of some volume region Ω is a measure for the tendency of the vectors $\mathbf{F}(\mathbf{r})$ to "diverge away" from a point $\mathbf{r}_0 \in \Omega$, the **circulation** of $\mathbf{F}(\mathbf{r})$ around a piece Σ of surface addresses the tendency of the vectors $\mathbf{F}(\mathbf{r})$ to "rotate around" $\mathbf{r}_0 \in \Sigma$.

For illustration, let $\mathbf{F}(\mathbf{r})$ be the velocity field $\mathbf{v}(\mathbf{r})$ of a solid, rotating about the origin,

$$\mathbf{v}(\mathbf{r}) \;=\; \boldsymbol{\omega} \times \mathbf{r}. \tag{208}$$

$\mathbf{v}(\mathbf{r})$ is the velocity vector of a point with position vector $\mathbf{r}$ on the solid. The vector $\boldsymbol{\omega}$ has the direction of the axis of rotation (right-hand rule). $|\boldsymbol{\omega}|$ is the angular velocity. Furthermore, let Σ be a disk centered at $\mathbf{r}_0$ with radius R and unit normal vector $\boldsymbol{\mathcal{N}}_\Sigma$.
• In the simplest case, when $\boldsymbol{\mathcal{N}}_\Sigma \parallel \boldsymbol{\omega}$ and $\mathbf{r}_0$ lies on the axis, we obviously have

$$\oint_{\partial\Sigma} \mathrm{d}\mathbf{r} \cdot \mathbf{v}(\mathbf{r}) \;=\; 2\pi R \left|\mathbf{v}(\mathbf{r})\right|_{\mathbf{r}\in\partial\Sigma} \;=\; 2\pi R\, |\boldsymbol{\omega}| R \;=\; \pi R^2\, 2|\boldsymbol{\omega}|. \tag{209}$$

• In the general case, we find (Exercise on Worksheet 6)

$$\oint_{\partial\Sigma} \mathrm{d}\mathbf{r} \cdot \mathbf{v}(\mathbf{r}) \;=\; \pi R^2\, (2\boldsymbol{\omega} \cdot \boldsymbol{\mathcal{N}}_\Sigma). \tag{210}$$

In words: The circulation of the velocity field $\mathbf{v}(\mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}$ around a disk Σ with unit normal vector $\boldsymbol{\mathcal{N}}_\Sigma$ equals the area $A_\Sigma = \pi R^2$ of the disk times the scalar product $2\boldsymbol{\omega} \cdot \boldsymbol{\mathcal{N}}_\Sigma$.

Theorem: For a vector field $\mathbf{F}(\mathbf{r})$, differentiable at $\mathbf{r} = \mathbf{r}_0$, there is a vector $\boldsymbol{\omega}_{\mathbf{F}}(\mathbf{r}_0)$ with

$$2\boldsymbol{\omega}_{\mathbf{F}}(\mathbf{r}_0) \cdot \boldsymbol{\mathcal{N}}_\Sigma \;=\; \lim_{R\to 0} \frac{1}{A_R} \oint_{\partial\Sigma_R} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}), \tag{211}$$

where Σ_R is the disk centered at $\mathbf{r}_0$ with unit normal vector $\boldsymbol{\mathcal{N}}_\Sigma$, radius R (and area $A_\Sigma = \pi R^2$). In cartesian coordinates, the vector $\boldsymbol{\omega}_{\mathbf{F}}(\mathbf{r})$ is given by

$$2\boldsymbol{\omega}_{\mathbf{F}}(\mathbf{r}) \;=\; \nabla \times \mathbf{F}(\mathbf{r}) \;\equiv\; \begin{pmatrix} \partial_2 F_3 - \partial_3 F_2 \\ \partial_3 F_1 - \partial_1 F_3 \\ \partial_1 F_2 - \partial_2 F_1 \end{pmatrix} \;\equiv\; \begin{pmatrix} \frac{\partial F_3(x,y,z)}{\partial y} - \frac{\partial F_2(x,y,z)}{\partial z} \\ \frac{\partial F_1(x,y,z)}{\partial z} - \frac{\partial F_3(x,y,z)}{\partial x} \\ \frac{\partial F_2(x,y,z)}{\partial x} - \frac{\partial F_1(x,y,z)}{\partial y} \end{pmatrix}. \tag{212}$$

[The vector field $\nabla \times \mathbf{F}(\mathbf{r})$ is called the **curl**, rot $\mathbf{F}(\mathbf{r})$, of the vector field $\mathbf{F}(\mathbf{r})$.]
Proof: See section **??**.

The "curl-analogy" to Gauss' divergence theorem reads:

Stokes' theorem: Let Σ be some piece of oriented surface with boundary ("rim") $\partial\Sigma$ and $\mathbf{F}(\mathbf{r})$ a vector field, defined and differentiable for all $\mathbf{r} \in \Sigma$. Then,

$$\oint_{\partial\Sigma} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;=\; \int_\Sigma \mathrm{d}\mathbf{A} \cdot \big[\nabla \times \mathbf{F}(\mathbf{r})\big]. \tag{213}$$

Alternative version of subsection 6.3

In section 6.2 we have considered the **flux** Φ of a vector field $\mathbf{F}(\mathbf{r})$,

$$\Phi \;=\; \oint_{\partial\Omega_a} \mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}), \tag{214}$$

out of the surface of a small **cube** Ω_a with side length a and center $\mathbf{r}_0$.
• The number Φ measures the tendency of $\mathbf{F}(\mathbf{r})$ to "diverge" away from the point $\mathbf{r}_0$.
In contrast, we now consider the **circulation** Θ of $\mathbf{F}(\mathbf{r})$ (a curve integral),

$$\Theta \;=\; \oint_{\partial\Sigma_a} \mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}), \tag{215}$$

along the rim of a small **planar square** Σ_a with side length a and center $\mathbf{r}_0$.
• The number Θ measures the tendency of $\mathbf{F}(\mathbf{r})$ to "rotate" around $\mathbf{r}_0$ in the plane of Σ_a.

Example: For a given vector $\boldsymbol{\omega}$, the vector field

$$\mathbf{v}(\mathbf{r}) \;=\; \boldsymbol{\omega}\times\mathbf{r} \tag{216}$$

gives the velocity of a point (with instantaneous position $\mathbf{r}$) on a solid, rotating with angular velocity $\boldsymbol{\omega}$ around an axis through the origin.
Let $\mathbf{a}_1$ and $\mathbf{a}_2$ be two vectors with $\mathbf{a}_1\cdot\mathbf{a}_2 = 0$ and $|\mathbf{a}_1| = |\mathbf{a}_2| = a$. Then,

$$\mathbf{r}_{1,2,3,4} \;=\; \mathbf{r}_0 \pm \frac{\mathbf{a}_1}{2} \pm \frac{\mathbf{a}_2}{2}. \tag{217}$$

are the corners of a square Σ_a with center $\mathbf{r}_0$, side length a, unit normal vector $\boldsymbol{\mathcal{N}}_\Sigma = \frac{\mathbf{a}_1\times\mathbf{a}_2}{a^2}$.
In the exercises, we shall show

$$\frac{\Theta}{a^2} \;=\; \frac{1}{a^2}\oint_{\partial\Sigma_a} \mathrm{d}\mathbf{r}\cdot\mathbf{v}(\mathbf{r}) \;=\; \boldsymbol{\mathcal{N}}_\Sigma\cdot 2\boldsymbol{\omega}. \tag{218}$$

Def.: Let Σ_a be a square with center $\mathbf{r}_0$, side length a, and unit normal vector $\boldsymbol{\mathcal{N}}_\Sigma$.
The **curl** of a vector field $\mathbf{F}(\mathbf{r})$ is the vector field $\mathbf{C}_\mathbf{F}(\mathbf{r})$ with the property

$$\boldsymbol{\mathcal{N}}_\Sigma\cdot\mathbf{C}_\mathbf{F}(\mathbf{r}_0) \;=\; \lim_{a\to 0}\frac{1}{a^2}\oint_{\partial\Sigma_a} \mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}). \tag{219}$$

Remark 1: The number $\boldsymbol{\mathcal{N}}_\Sigma\cdot\mathbf{C}_\mathbf{F}(\mathbf{r}_0)$ is not affected by a rotation of the x-, y-, and z-axes.
Remark 2: In usual notation, $\mathbf{C}_\mathbf{F}(\mathbf{r}) \;\equiv\; \mathrm{rot}\,\mathbf{F}(\mathbf{r})$.

Lemma: For a differentiable vector field $\mathbf{F}(\mathbf{r})$, the curl is given by

$$\mathbf{C}_\mathbf{F}(\mathbf{r}) \;=\; \begin{pmatrix} \frac{\partial F_3(\mathbf{r})}{\partial y} - \frac{\partial F_2(\mathbf{r})}{\partial z} \\ \frac{\partial F_1(\mathbf{r})}{\partial z} - \frac{\partial F_3(\mathbf{r})}{\partial x} \\ \frac{\partial F_2(\mathbf{r})}{\partial x} - \frac{\partial F_1(\mathbf{r})}{\partial y} \end{pmatrix} \;\equiv\; \begin{pmatrix} \partial_2 F_3 - \partial_3 F_2 \\ \partial_3 F_1 - \partial_1 F_3 \\ \partial_1 F_2 - \partial_2 F_1 \end{pmatrix} \;\equiv\; \nabla\times\mathbf{F}(\mathbf{r}). \tag{220}$$

Proof: Exercises!

Example: For the velocity field of the rotating solid, we find in fact

$$\nabla\times\mathbf{v}(\mathbf{r}) \;\equiv\; \nabla\times\big(\boldsymbol{\omega}\times\mathbf{r}\big) \;=\; \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \times \begin{pmatrix} \omega_2 z - \omega_3 y \\ \omega_3 x - \omega_1 z \\ \omega_1 y - \omega_2 x \end{pmatrix} \;=\; 2\boldsymbol{\omega}. \tag{221}$$

6.4 Curl and independence of path

Def.: A volume region Ω is called **simply connected**, when every closed loop Γ in Ω can be shrunk continuously within Ω to a point $\mathbf{r}_0 \in \Omega$.
(Then, there is a piece of surface Σ with: $\Gamma = \partial\Sigma$ **and** $\Sigma \subset \Omega$.)

Example: $\Omega = \{(x, y, z) \,|\, x^2 + y^2 \geq R^2\} \subset \mathbb{R}^3$ is **not** simply connected when $R > 0$.

Theorem: Let $\mathbf{F}(\mathbf{r})$ be a differentiable vector field, defined on a simply connected region $\Omega \subseteq \mathbb{R}^3$. Then, the following three statements are equivalent:

(a) There is a scalar field $f(\mathbf{r})$ with the property $\mathbf{F}(\mathbf{r}) = \nabla f(\mathbf{r})$ for all $\mathbf{r} \in \Omega$.

(b) $\nabla \times \mathbf{F}(\mathbf{r}) = \mathbf{0}$ for all $\mathbf{r} \in \Omega$.

(c) For a curve $\Gamma \in \Omega$, the value of the line integral $\int_\Gamma \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r})$ only depends on the initial and final points $\mathbf{r}_A$ and $\mathbf{r}_B$, respectively, of Γ, but **not on its path**.

Proof: • From section 5.1, we already know that (a) $\Rightarrow$ (b).
[We also know that (a) $\Rightarrow$ (c), since $\int_\Gamma \mathrm{d}\mathbf{r} \cdot \nabla f(\mathbf{r}) \equiv f(\mathbf{r}_B) - f(\mathbf{r}_A)$, see section 5.2.]
• The conclusion (b) $\Rightarrow$ (c) follows immediately from Stokes' Theorem: Given two curves Γ_1 and Γ_2 which have the same initial point $\mathbf{r}_A$ and the same final point $\mathbf{r}_B$, let $\overline{\Gamma}_2$ follow the same path as Γ_2, but in opposite direction (from $\mathbf{r}_B$ to $\mathbf{r}_A$). Then, the composition $\overline{\Gamma}_2 \circ \Gamma_1$ makes a closed loop $\Gamma = \partial\Sigma$ in Ω (starting at $\mathbf{r}_A$, following Γ_1 up to $\mathbf{r}_B$ and returning along the path of Γ_2 to $\mathbf{r}_A$). Therefore, since $\nabla \times \mathbf{F}(\mathbf{r}) = \mathbf{0}$ for $\mathbf{r} \in \Sigma \subset \Omega$,

$$\int_{\Gamma_1} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) - \int_{\Gamma_2} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;\equiv\; \oint_{\partial\Sigma} \mathrm{d}\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \;=\; \int_\Sigma \mathrm{d}\mathbf{A} \cdot \big[\nabla \times \mathbf{F}(\mathbf{r})\big] \;=\; 0. \tag{222}$$

• To prove the conclusion (c) $\Rightarrow$ (a) we define for any $\mathbf{r} \in \Omega$ the scalar field

$$f(\mathbf{r}) \;:=\; \int_{\Gamma(\mathbf{r}_0, \mathbf{r})} \mathrm{d}\mathbf{r}' \cdot \mathbf{F}(\mathbf{r}'), \tag{223}$$

where $\Gamma(\mathbf{r}_0, \mathbf{r})$ is **any** curve in Ω running from some fixed point $\mathbf{r}_0 \in \Omega$ to $\mathbf{r}$. For $h \geq 0$, let $\Gamma(\mathbf{r}_0, \mathbf{r}, h)$ be the extension of $\Gamma(\mathbf{r}_0, \mathbf{r})$ by a straight line Γ_h, running in x-direction from $\mathbf{r}$ to the new endpoint $\mathbf{r} + h\mathbf{u}_x$ (where $\mathbf{u}_x$ is the unit vector in x-direction). Then,

$$\begin{aligned} \frac{\partial}{\partial x} f(x, y, z) \;&\equiv\; \lim_{h \to 0} \frac{1}{h} \Big[f(x + h, y, z) \;-\; f(x, y, z) \Big] \\ &=\; \lim_{h \to 0} \frac{1}{h} \left[\int_{\Gamma(\mathbf{r}_0, \mathbf{r}, h)} \mathrm{d}\mathbf{r}' \cdot \mathbf{F}(\mathbf{r}') - \int_{\Gamma(\mathbf{r}_0, \mathbf{r})} \mathrm{d}\mathbf{r}' \cdot \mathbf{F}(\mathbf{r}') \right] \;=\; \lim_{h \to 0} \frac{1}{h} \int_{\Gamma_h} \mathrm{d}\mathbf{r}' \cdot \mathbf{F}(\mathbf{r}'). \end{aligned} \tag{224}$$

Using the parametrisation

$$\Gamma_h \;=\; \Big\{ \mathbf{r}'(s) \,\Big|\, 0 \leq s \leq h \Big\}, \qquad \mathbf{r}'(s) = \mathbf{r} + s\mathbf{u}_x, \qquad \frac{\mathrm{d}\mathbf{r}'(s)}{\mathrm{d}s} = \mathbf{u}_x, \tag{225}$$

we have $\int_{\Gamma_h} \mathrm{d}\mathbf{r}' \cdot \mathbf{F}(\mathbf{r}') \equiv \int_0^h \mathrm{d}s \, \frac{\mathrm{d}\mathbf{r}'(s)}{\mathrm{d}s} \cdot \mathbf{F}(\mathbf{r}'(s)) = \int_0^h \mathrm{d}s \, F_1(x + s, y, z)$ and, eventually,

$$\frac{\partial}{\partial x} f(x, y, z) \;=\; \lim_{h \to 0} \frac{1}{h} \int_0^h \mathrm{d}s \, F_1(x + s, y, z) \;=\; F_1(x, y, z), \quad \text{q.e.d.} \tag{226}$$

7 Application of vector calculus: Electrodynamics

7.1 Electric charge q

We consider electric charge as a continuum with density $\rho(\mathbf{r},t)$ (in $\frac{\mathrm{C}}{\mathrm{m}^3}$) and current density $\mathbf{J}(\mathbf{r},t)$ (in $\frac{\mathrm{C}}{\mathrm{m}^2\,\mathrm{s}}$). In such a description, the case of a **point charge** can be included, when we introduce Dirac's δ-function $\delta(\mathbf{r}-\mathbf{r}_0)=\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)$: A moving point charge q, with time-dependent position vector $\mathbf{r}_0(t)$, has the charge density

$$\rho(\mathbf{r},t) \;=\; q\,\delta\big(\mathbf{r}-\mathbf{r}_0(t)\big). \tag{227}$$

Similarly, a thin wire (covering the z-axis) with electric current I has the current density

$$\mathbf{J}(\mathbf{r},t) \;=\; I\,\delta(x)\delta(y)\,\mathbf{e}_z. \tag{228}$$

Unlike a chemical compound (such as CO_2 in our former examples), electric charge is a **conserved quantity**: It can (chemically) neither be created nor destroyed. Consequently, the source density of charge is $s(\mathbf{r},t)=0$ and the amount $\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t)$ of charge within any volume region Ω can change only by a flow of charge across the surface $\partial\Omega$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t) \;=\; -\oint_{\partial\Omega}\mathrm{d}\mathbf{S}\cdot\mathbf{J}(\mathbf{r},t). \tag{229}$$

Since this equation holds for any $\Omega\subset\mathbb{R}^3$, Gauss' theorem $\oint_{\partial\Omega}\mathrm{d}\mathbf{S}\cdot\mathbf{J}=\int_\Omega \mathrm{d}^3r\,\nabla\cdot\mathbf{J}$ implies

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) \;+\; \nabla\cdot\mathbf{J}(\mathbf{r},t) \;=\; 0. \tag{230}$$

This differential relation between the scalar field $\rho(\mathbf{r},t)$ and the vector field $\mathbf{J}(\mathbf{r},t)$ is called **continuity equation**. It is the mathematical expression for a **conservation law**.

7.2 Electromagnetic fields E and B

In space, let $\mathbf{F}(\mathbf{r},\mathbf{v},t)$ be the total force acting on a test (point) charge q at the time t when it has position $\mathbf{r}$ and is moving with velocity $\mathbf{v}$. If $\mathbf{F}$ is purely an electromagnetic force, then there exist (time dependent) vector fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$, called **electric** and **magnetic field**, respectively, such that

$$\mathbf{F}(\mathbf{r},\mathbf{v},t) \;=\; q\Big[\mathbf{E}(\mathbf{r},t) \;+\; \mathbf{v}\times\mathbf{B}(\mathbf{r},t)\Big]. \tag{231}$$

Consequently, the fields have the physical dimensions[3]

$$[\mathbf{E}] \;=\; 1\frac{\mathrm{N}}{\mathrm{C}}, \qquad [\mathbf{B}] \;=\; 1\frac{\mathrm{N\,s}}{\mathrm{C\,m}} \;\equiv\; 1\mathrm{T}\;\;(1\text{ Tesla}). \tag{232}$$

Due to Eq. (231), the vectors $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$, at any position-time pair $(\mathbf{r},t)$, can be measured by observing the forces acting on various moving test charges q.

[3] Giving $\mathbf{B}$ the same dimension N/C as $\mathbf{E}$, one may define alternatively

$$\mathbf{F} \;=\; q\,\mathbf{E}(\mathbf{r},t) \;+\; q\,\frac{\mathbf{v}}{c}\times\mathbf{B}(\mathbf{r},t),$$

where c is the speed of light.

7.3 Stationary fields $\mathbf{E}(\mathbf{r})$, $\mathbf{B}(\mathbf{r})$

In situations when the fields $\mathbf{E}$ and $\mathbf{B}$ are time-independent ("stationary"), they seem to be "generated" by (stationary) charges $\rho(\mathbf{r})$ and, respectively, (stationary) currents $\mathbf{J}(\mathbf{r})$. This relationship between the fields and their "sources" is governed by the constants

$$\epsilon_0 \;=\; 8.854 \cdot 10^{-12}\,\frac{\mathrm{A\,s}}{\mathrm{V\,m}}, \qquad\qquad \mu_0 \;=\; 4\pi \cdot 10^{-7}\,\frac{\mathrm{V\,s}}{\mathrm{A\,m}}. \tag{233}$$

7.3.1 Coulomb's law

A time-independent charge distribution $\rho(\mathbf{r})$ gives rise to a stationary electric field,

$$\mathbf{E}(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\rho(\mathbf{r}')\,\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}, \tag{234}$$

seen by a (test) charge at position $\mathbf{r}$. In this context, the "generating" charge $\rho(\mathbf{r})$ is also called "source" charge. In the special case of a point source, $\rho(\mathbf{r}) = Q\delta(\mathbf{r}-\mathbf{r}_0)$, we find

$$\mathbf{E}(\mathbf{r}) \;=\; \frac{Q}{4\pi\epsilon_0}\,\frac{\mathbf{r}-\mathbf{r}_0}{|\mathbf{r}-\mathbf{r}_0|^3} \qquad \Rightarrow \qquad |\mathbf{E}(\mathbf{r})| \;=\; \frac{|Q|}{4\pi\epsilon_0}\,\frac{1}{|\mathbf{r}-\mathbf{r}_0|^2}. \tag{235}$$

Usually, Eq. (235) for a point source is called **Coulomb's law**. Then, Eq. (234) results from the **superposition principle**.

7.3.2 Biot-Savart law

A time-independent current distribution $\mathbf{J}(\mathbf{r})$ gives rise to a stationary magnetic field,

$$\mathbf{B}(\mathbf{r}) \;=\; \frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\,\mathbf{J}(\mathbf{r}')\times\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}, \tag{236}$$

seen by a (moving) test charge at position $\mathbf{r}$. In the special case, when the source current is an infinitely long straight wire, carrying the current I, $\mathbf{J}(\mathbf{r}) = I\,\delta(x)\delta(y)\,\mathbf{e}_z$, we find

$$\begin{aligned}
\mathbf{B}(\mathbf{r}) \;&=\; \frac{\mu_0}{4\pi}\,I\int_{-\infty}^{\infty}\mathrm{d}z'\,\frac{\begin{pmatrix}0\\0\\1\end{pmatrix}\times\begin{pmatrix}x\\y\\z-z'\end{pmatrix}}{\Big[x^2\,+\,y^2\,+\,(z-z')^2\Big]^{3/2}} \\
&=\; \frac{\mu_0}{4\pi}\,I\begin{pmatrix}-y\\x\\0\end{pmatrix}\int_{-\infty}^{\infty}\mathrm{d}z'\,\Big[x^2\,+\,y^2\,+\,(z-z')^2\Big]^{-3/2} \\
&=\; \frac{\mu_0}{2\pi}\,I\begin{pmatrix}-y\\x\\0\end{pmatrix}\frac{1}{x^2\,+\,y^2},
\end{aligned} \tag{237}$$

where we have used $\int_{-\infty}^{\infty}\mathrm{d}u\,[a^2+u^2]^{-3/2} \;=\; \frac{u}{a^2}[a^2+u^2]^{-1/2}$.

$$|\mathbf{B}(\mathbf{r})| \;=\; \frac{\mu_0\,I}{2\pi}\,\frac{1}{|\mathbf{r}_\perp|}. \tag{238}$$

7.3.3 Some mathematical identities (problems from Worksheet 7/2019)

Problem 4: The scalar field $\phi(\mathbf{r}) = \frac{1}{|\mathbf{r}-\mathbf{r}_0|}$.

Use the explicit representation of $|\mathbf{r}-\mathbf{r}_0|$ in cartesian coordinates,

$$|\mathbf{r}-\mathbf{r}_0| \;=\; \Big[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2\Big]^{1/2} \;\equiv\; f(x,y,z),$$

to prove the identities

$$\nabla\,\frac{1}{|\mathbf{r}-\mathbf{r}_0|} \;=\; -\frac{\mathbf{r}-\mathbf{r}_0}{|\mathbf{r}-\mathbf{r}_0|^3}, \qquad\qquad \nabla\times\frac{\mathbf{a}}{|\mathbf{r}-\mathbf{r}_0|} \;=\; \frac{\mathbf{a}\times(\mathbf{r}-\mathbf{r}_0)}{|\mathbf{r}-\mathbf{r}_0|^3}, \tag{239}$$

where the differential operator ∇ is understood to act on the variables $\mathbf{r}=(x,y,z)$, while $\mathbf{r}_0=(x_0,y_0,z_0)$ and $\mathbf{a}=(a_1,a_2,a_3)$ are constant vectors.

Problem 5: Laplacian (operator) ∇^2

(a) For a differentiable scalar field $g(\mathbf{r})=g(x,y,z)$, show that

$$\nabla\cdot\big[\nabla g(\mathbf{r})\big] \;=\; \nabla^2 g(\mathbf{r}),$$

where (in cartesian coordinates) the **Laplacian** ∇^2 reads $\nabla^2 \;=\; \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$.

(b) Evaluate the Laplacian $\nabla^2 g(\mathbf{r})$ for the following scalar fields (where $A,B,C,D,E,F,R,\,\alpha,\,\beta$ are constants).

$$\begin{aligned}
g_1(\mathbf{r}) &= Ax^2+By^2-(A+B)z^2 \;+\; Dx+Ey+Fz \;+\; C,\\
g_2(\mathbf{r}) &= A\,\sin(\alpha x)\,\sin(\beta y)\,\mathrm{e}^{-\gamma z}, \qquad\qquad \gamma=\sqrt{\alpha^2+\beta^2},\\
g_3(\mathbf{r}) &= \frac{A}{\sqrt{x^2+y^2+z^2}} \;\equiv\; A(x^2+y^2+z^2)^{-1/2} \qquad (\mathbf{r}\neq\mathbf{0}),\\
g_4(\mathbf{r}) &= (Ax+By+Cz)\,(x^2+y^2+z^2)^{-3/2} \qquad (\mathbf{r}\neq\mathbf{0}),\\
g_5(\mathbf{r}) &= \frac{Q}{4\pi\epsilon_0}\begin{cases}\frac{3R^2-r^2}{2R^3} & (r\le R),\\ \frac{1}{r} & (r>R).\end{cases}
\end{aligned}$$

(c) Use the result of part (b) for $g_5(\mathbf{r})$, to verify the identity

$$\nabla^2\,\frac{1}{|\mathbf{r}-\mathbf{r}_0|} \;=\; -4\pi\,\delta(\mathbf{r}-\mathbf{r}_0). \tag{240}$$

Problem 6: Second derivatives and integration by parts

(a) For a differentiable scalar field $g(\mathbf{r})$ and a differentiable vector field $\mathbf{F}(\mathbf{r})$, show that

$$\nabla\cdot\big[\nabla g(\mathbf{r})\big] \;=\; \nabla^2 g(\mathbf{r}) \qquad \text{(already shown in problem 5(a))}, \tag{241}$$

$$\nabla\times\big[\nabla g(\mathbf{r})\big] \;=\; \mathbf{0}, \tag{242}$$

$$\nabla\cdot\big[\nabla\times\mathbf{F}(\mathbf{r})\big] \;=\; 0, \tag{243}$$

$$\nabla\times\big[\nabla\times\mathbf{F}(\mathbf{r})\big] \;=\; \nabla\big[\nabla\cdot\mathbf{F}(\mathbf{r})\big]-\nabla^2\mathbf{F}(\mathbf{r}). \tag{244}$$

(b) Integrating by parts, show that

$$\int \mathrm{d}^3r\,\mathbf{F}(\mathbf{r})\cdot\big[\nabla g(\mathbf{r})\big] \;=\; -\int \mathrm{d}^3r\,\big[\nabla\cdot\mathbf{F}(\mathbf{r})\big]\,g(\mathbf{r}), \tag{245}$$

provided that the product $\mathbf{P}(\mathbf{r})=g(\mathbf{r})\,\mathbf{F}(\mathbf{r})$ goes to zero as $|\mathbf{r}|\to\infty$.

7.3.4 Scalar and vector potentials ϕ and $\mathbf{A}$

Defining the so-called (electrostatic) **scalar potential**

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{246}$$

and the so-called (electromagnetic) **vector potential**

$$\mathbf{A}(\mathbf{r}) \;=\; \frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\,\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{247}$$

and applying the identities (239), the integrals (234) and (236) show that

$$\mathbf{E}(\mathbf{r}) \;=\; -\nabla\phi(\mathbf{r}), \qquad\qquad \mathbf{B}(\mathbf{r}) \;=\; \nabla\times\mathbf{A}(\mathbf{r}). \tag{248}$$

Maxwell's equations for **stationary** fields,

$$\text{(I)} \qquad \nabla\cdot\mathbf{E}(\mathbf{r}) \;=\; \frac{1}{\epsilon_0}\,\rho(\mathbf{r}), \qquad\qquad \text{(II)} \qquad \nabla\cdot\mathbf{B}(\mathbf{r}) \;=\; 0, \tag{249}$$

$$\text{(III)} \qquad \nabla\times\mathbf{B}(\mathbf{r}) \;=\; \mu_0\,\mathbf{J}(\mathbf{r}), \qquad\qquad \text{(IV)} \qquad \nabla\times\mathbf{E}(\mathbf{r}) \;=\; \mathbf{0}, \tag{250}$$

can be derived from the representations (248) as follows:

- (II) and (IV) are immediate consequences of (248) due to Eqs. (242) and (243).
- (I) is obtained from (248) as follows:

$$\nabla\cdot\mathbf{E}(\mathbf{r}) \;\equiv\; -\nabla^2\phi(\mathbf{r}) \;=\; -\frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\rho(\mathbf{r}')\,\nabla^2\frac{1}{|\mathbf{r}-\mathbf{r}'|} \;=\; \frac{1}{\epsilon_0}\,\rho(\mathbf{r}), \tag{251}$$

  where we have used the famous identity Eq. (240) in the last step.
- To derive Equation (III) from (248), we first notice that

$$\nabla\times\mathbf{B}(\mathbf{r}) \;\equiv\; \nabla\times\big[\nabla\times\mathbf{A}(\mathbf{r})\big] \;=\; \nabla\big[\nabla\cdot\mathbf{A}(\mathbf{r})\big] - \nabla^2\mathbf{A}(\mathbf{r}), \tag{252}$$

  due to the identity (244). Here, we have

$$\nabla\cdot\mathbf{A}(\mathbf{r}) \;=\; \frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\,\mathbf{J}(\mathbf{r}')\cdot\nabla_{\mathbf{r}}\frac{1}{|\mathbf{r}-\mathbf{r}'|} \;\equiv\; -\frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\,\mathbf{J}(\mathbf{r}')\cdot\nabla_{\mathbf{r}'}\frac{1}{|\mathbf{r}-\mathbf{r}'|}, \tag{253}$$

  and integration by parts, see Eq. (245), yields

$$\nabla\cdot\mathbf{A}(\mathbf{r}) \;=\; +\frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\,\big[\nabla_{\mathbf{r}'}\cdot\mathbf{J}(\mathbf{r}')\big]\,\frac{1}{|\mathbf{r}-\mathbf{r}'|} \;=\; 0, \tag{254}$$

  where we have used the continuity equation $\nabla\cdot\mathbf{J}(\mathbf{r}) = -\frac{\partial\rho(\mathbf{r})}{\partial t} = 0$. Consequently,

$$\nabla\times\mathbf{B}(\mathbf{r}) \;=\; -\nabla^2\mathbf{A}(\mathbf{r}) \;=\; -\frac{\mu_0}{4\pi}\int \mathrm{d}^3r'\;\mathbf{J}(\mathbf{r}')\,\nabla^2\frac{1}{|\mathbf{r}-\mathbf{r}'|} \;=\; \mu_0\,\mathbf{J}(\mathbf{r}), \tag{255}$$

  where we have again used the identity Eq. (240) in the last step.

It seems that the divergence of $\mathbf{E}$ is the charge density ρ and the curls of $\mathbf{B}$ are the currents $\mathbf{J}$, while $\mathbf{E}$ has no curls and $\mathbf{B}$ has zero divergence. In the following section 7.4, however, we shall switch to **time-dependent** fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$, reveiling that $\frac{\partial\mathbf{B}}{\partial t}$ produces curls of $\mathbf{E}$ and $\frac{\partial\mathbf{E}}{\partial t}$ produces (additional) curls of $\mathbf{B}$.

7.4 Maxwell's equations

7.4.1 Differential form

Generally, along with the electric densities $\rho(\mathbf{r},t)$ and $\mathbf{J}(\mathbf{r},t)$ of charge and current, the time-dependent fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ satisfy **Maxwell's equations**,

$$\nabla\cdot\mathbf{E}(\mathbf{r},t) \;=\; \frac{1}{\epsilon_0}\,\rho(\mathbf{r},t), \qquad \nabla\cdot\mathbf{B}(\mathbf{r},t) \;=\; 0, \tag{256}$$

$$\nabla\times\mathbf{B}(\mathbf{r},t) \;=\; \mu_0\,\mathbf{J}(\mathbf{r},t) \;+\; \epsilon_0\mu_0\frac{\partial\mathbf{E}(\mathbf{r},t)}{\partial t}, \qquad \nabla\times\mathbf{E}(\mathbf{r},t) \;=\; -\frac{\partial\mathbf{B}(\mathbf{r},t)}{\partial t}. \tag{257}$$

Remark: The extra term $\epsilon_0\mu_0\frac{\partial\mathbf{E}(\mathbf{r},t)}{\partial t}$ makes ρ and $\mathbf{J}$ satisfy the continuity equation,

$$\frac{1}{\mu_0}\nabla\cdot\Big[\nabla\times\mathbf{B}(\mathbf{r},t)\Big] \;\equiv\; \nabla\cdot\mathbf{J}(\mathbf{r},t) \;+\; \frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;=\; 0. \tag{258}$$

The first one of Eqs. (256) is called **Gauss' Law**.
Eqs. (257) are called **Ampere's law** and **Faraday's law**, respectively.

7.4.2 Integral form

Applying Gauss' theorem $\oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{F}(\mathbf{r}) = \int_\Omega \mathrm{d}^3r[\nabla\cdot\mathbf{F}(\mathbf{r})]$ to Maxwell's equations (256), we obtain the integral relations

$$\begin{aligned}
&\text{(I)} \quad \oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{E}(\mathbf{r},t) \;=\; \frac{1}{\epsilon_0}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t) \;\equiv\; \frac{Q_\Omega(t)}{\epsilon_0},\\
&\text{(II)} \quad \oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{B}(\mathbf{r},t) \;=\; 0.
\end{aligned} \tag{259}$$

In words, (I): *The flux of the electric field out of the surface $\partial\Omega$ of a volume region Ω always equals ($\frac{1}{\epsilon_0}$ times) the amount $Q_\Omega(t)$ of charge contained within Ω at the same time.*
(II): *There is no magnetic flux out of closed surfaces. The field $\mathbf{B}(\mathbf{r},t)$ has no sources.*

Similarly, applying Stokes' theorem $\oint_{\partial\Sigma}\mathrm{d}\mathbf{r}\cdot\mathbf{F}(\mathbf{r}) = \int_\Sigma \mathrm{d}\mathbf{A}\cdot[\nabla\times\mathbf{F}(\mathbf{r})]$ to Maxwell's equations (257), we obtain the integral relations

$$\text{(III)} \quad \oint_{\partial\Sigma}\mathrm{d}\mathbf{r}\cdot\mathbf{B}(\mathbf{r},t) \;=\; \mu_0\int_\Sigma \mathrm{d}\mathbf{A}\cdot\left[\mathbf{J}(\mathbf{r},t)+\epsilon_0\frac{\partial\mathbf{E}(\mathbf{r},t)}{\partial t}\right], \tag{260}$$

$$\text{(IV)} \quad \oint_{\partial\Sigma}\mathrm{d}\mathbf{r}\cdot\mathbf{E}(\mathbf{r},t) \;=\; -\frac{\mathrm{d}}{\mathrm{d}t}\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{B}(\mathbf{r},t). \tag{261}$$

(IV) is **Faraday's induction law**: *The circulation of the electric field around the rim $\partial\Sigma$ of any piece Σ of surface, called the* **induction voltage** *$U_{\mathrm{ind}}(t)$ at the time t, always equals minus the rate of change of the magnetic flux $\Phi(t)$ through Σ at that time,*

$$U_{\mathrm{ind}}(t) \;=\; -\frac{\mathrm{d}}{\mathrm{d}t}\,\Phi(t). \tag{262}$$

In the absence of a time-dependent electric field, Eq. (III) becomes

$$\oint_{\partial\Sigma}\mathrm{d}\mathbf{r}\cdot\mathbf{B}(\mathbf{r},t) \;=\; \mu_0\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t) \;\equiv\; \mu_0\,I_\Sigma(t). \tag{263}$$

In words: *The circulation of the magnetic field around the rim $\partial\Sigma$ at the time t equals (μ_0 times) the total electric current $I_\Sigma(t)$ flowing through Σ at that time (***Ampere's law***).*

7.4.3 Derivation of the laws of Coulomb and Biot-Savart

A uniformly charged sphere with charge Q and finite radius R_0 has the charge density

$$\rho(r,\theta,\phi) \;=\; \left\{ \begin{array}{ll} \rho_0 & (r \le R_0) \\ 0 & (r > R_0) \end{array} \right\}, \qquad \rho_0 = \frac{Q}{\frac{4\pi}{3} R_0^3}. \tag{264}$$

Due to the spherical symmetry of ρ, the electric field has the form

$$\mathbf{E}(\mathbf{r}) \;=\; E(r)\,\frac{\mathbf{r}}{r} \;\equiv\; E(r)\,\mathbf{u}_r, \tag{265}$$

with a yet unknown r-dependent magnitude $E(r)$. Let Ω be a concentric sphere with radius R. Then, Maxwell's equation $\oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{E}(\mathbf{r}) = \frac{1}{\epsilon_0}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r})$ has the LHS

$$\oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{E}(\mathbf{r}) \;\equiv\; A_{\partial\Omega}\Big\langle \boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\mathbf{E}(\mathbf{r})\Big\rangle_{\mathbf{r}\in\partial\Omega} \;=\; 4\pi R^2\,E(R), \tag{266}$$

where we have utilised that $A_{\partial\Omega} = 4\pi R^2$ and $\boldsymbol{\mathcal{N}}(\mathbf{r}) = \mathbf{u}_r$. On the RHS, we have

$$\frac{1}{\epsilon_0}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r}) \;=\; \frac{1}{\epsilon_0}\left\{ \begin{array}{ll} V_\Omega\rho_0 \equiv Q\,\frac{R^3}{R_0^3} & (R \le R_0), \\ Q & (R \ge R_0), \end{array} \right. \tag{267}$$

since $V_\Omega = \frac{4\pi}{3}R^3$. Comparing the last two results (and setting $R = r$), we obtain

$$E(r) \;=\; \frac{1}{4\pi\epsilon_0}\,Q\left\{ \begin{array}{ll} \frac{r}{R_0^3} & (r \le R_0), \\ \frac{1}{r^2} & (r \ge R_0). \end{array} \right. \tag{268}$$

Coulomb's law (235) for a point charge Q is recovered in the limit $R_0 \to 0$.

A uniform current I inside an infinitely long cylinder with radius R_0 has the density

$$\mathbf{J}(s,\phi,z) \;=\; \left\{ \begin{array}{ll} J_0\,\mathbf{u}_z & (s \le R_0) \\ 0 & (s > R_0) \end{array} \right\}, \qquad J_0 = \frac{I}{\pi R_0^2}. \tag{269}$$

Due to the cylindrical symmetry of $\mathbf{J}$, the magnetic field has the form

$$\mathbf{B}(\mathbf{r}) \;=\; B(s)\,\frac{\mathbf{u}_z\times\mathbf{r}}{|\mathbf{u}_z\times\mathbf{r}|} \;\equiv\; B(s)\,\mathbf{u}_\phi, \tag{270}$$

with a yet unknown s-dependent magnitude $B(s)$. Let Σ be a disk with radius R, centered on the z-axis and normal to it. Then, Maxwell's equation $\oint_{\partial\Sigma} \mathrm{d}\mathbf{r}\cdot\mathbf{B}(\mathbf{r}) = \mu_0\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r})$ has the LHS

$$\oint_{\partial\Sigma} \mathrm{d}\mathbf{r}\cdot\mathbf{B}(\mathbf{r}) \;\equiv\; L_{\partial\Sigma}\Big\langle \boldsymbol{\mathcal{T}}(\mathbf{r})\cdot\mathbf{B}(\mathbf{r})\Big\rangle_{\mathbf{r}\in\partial\Sigma} \;=\; 2\pi R\,B(R), \tag{271}$$

where we have utilised that $L_{\partial\Sigma} = 2\pi R$ and $\boldsymbol{\mathcal{T}}(\mathbf{r}) = \mathbf{u}_\phi$. On the RHS, we have

$$\mu_0\int_\Sigma \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r}) \;=\; \mu_0\left\{ \begin{array}{ll} A_\Sigma J_0 \equiv I\,\frac{R^2}{R_0^2} & (R \le R_0), \\ I & (R \ge R_0), \end{array} \right. \tag{272}$$

since $A_\Sigma = \pi R^2$. Comparing the last two results (and setting $R = s$), we obtain

$$B(s) \;=\; \frac{\mu_0}{2\pi}\,I\left\{ \begin{array}{ll} \frac{s}{R_0^2} & (s \le R_0), \\ \frac{1}{s} & (s \ge R_0). \end{array} \right. \tag{273}$$

Eq. (237) is recovered in the limit $R_0 \to 0$.

7.4.4 Electromagnetic waves

The laws of Coulomb/Biot-Savart are suggesting that the fields $\mathbf{E}$ and $\mathbf{B}$ are "generated" by source charges Q and currents I, respectively. To demonstrate that this is not the full truth, we now consider Maxwell's equations in "vacuum", $\rho(\mathbf{r},t) = 0$ and $\mathbf{J}(\mathbf{r},t) = \mathbf{0}$,

$$\text{(I)} \quad \nabla \cdot \mathbf{E}(\mathbf{r},t) = 0, \qquad \text{(II)} \quad \nabla \cdot \mathbf{B}(\mathbf{r},t) = 0, \tag{274}$$

$$\text{(III)} \quad \nabla \times \mathbf{B}(\mathbf{r},t) = \epsilon_0\mu_0 \frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t}, \qquad \text{(IV)} \quad \nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}. \tag{275}$$

Taking the curl on both sides of (III), we obtain on the LHS

$$\nabla \times \big[\nabla \times \mathbf{B}(\mathbf{r},t)\big] \equiv \nabla\big[\underbrace{\nabla \cdot \mathbf{B}(\mathbf{r},t)}_{0}\big] - \nabla^2 \mathbf{B}(\mathbf{r},t), \tag{276}$$

where we have used an identity (exercises!) and (II). Consequently, (III) becomes

$$-\nabla^2 \mathbf{B}(\mathbf{r},t) = \epsilon_0\mu_0 \frac{\partial}{\partial t}\big[\nabla \times \mathbf{E}(\mathbf{r},t)\big], \tag{277}$$

Using now (IV) for $\nabla \times \mathbf{E}$, we eventually find

$$\nabla^2 \mathbf{B}(\mathbf{r},t) = \frac{1}{c^2}\frac{\partial^2 \mathbf{B}(\mathbf{r},t)}{\partial t^2}, \qquad c = \frac{1}{\sqrt{\epsilon_0\mu_0}} \equiv 2.998 \cdot 10^8\,\frac{\mathrm{m}}{\mathrm{s}}. \tag{278}$$

This is the **wave equation** for the components B_1, B_2, B_3 of $\mathbf{B}$,

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) B_i(x,y,z,t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} B_i(x,y,z,t) \qquad (i = 1,2,3), \tag{279}$$

a **partial differential equation (PDE)** for the unknown function $B_i(\mathbf{r},t)$.

Starting with (IV), a similar reasoning leads to the equation

$$\nabla^2 \mathbf{E}(\mathbf{r},t) = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}(\mathbf{r},t)}{\partial t^2}. \tag{280}$$

We see: There are solutions $\mathbf{E}(\mathbf{r},t) \neq \mathbf{0}$ and $\mathbf{B}(\mathbf{r},t) \neq \mathbf{0}$, even when no charges Q and no currents I are around.

7.5 Electrostatic potential and Poisson's equation

When the magnetic field is stationary, $\frac{\partial}{\partial t}\mathbf{B}(\mathbf{r},t) = \mathbf{0}$, Faraday's law implies

$$\nabla \times \mathbf{E}(\mathbf{r},t) \;=\; \mathbf{0}. \tag{281}$$

Consequently, due to the theorem of section 6.4, there is a scalar field $\phi(\mathbf{r},t)$ such that

$$\mathbf{E}(\mathbf{r},t) \;=\; -\nabla\phi(\mathbf{r},t). \tag{282}$$

In such a situation, the vector field $\mathbf{E}(\mathbf{r},t)$ is fixed by a (simpler) scalar field $\phi(\mathbf{r},t)$, called the **electrostatic potential**.[4]

Combining Eq. (282) with Gauss' law, we see that $\phi(\mathbf{r},t)$ satisfies **Poisson's equation**,

$$\nabla^2\phi(\mathbf{r},t) \;=\; -\frac{1}{\epsilon_0}\,\rho(\mathbf{r},t). \tag{283}$$

For any given $\rho(\mathbf{r},t)$, this is a **partial differential equation (PDE)** for the unknown function $\phi(\mathbf{r},t)$. Once Eq. (283) is solved for $\phi(\mathbf{r},t)$, $\mathbf{E}(\mathbf{r},t)$ is obtained from Eq. (282).

Remark: For time-independent $\rho = \rho(\mathbf{r})$, a solution to Poisson's equation is given by

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{284}$$

since (as we shall prove in the exercises)[5]

$$\begin{aligned}\nabla^2\phi(\mathbf{r}) \;&\equiv\; \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\rho(\mathbf{r}')\,\nabla^2\,\frac{1}{|\mathbf{r}-\mathbf{r}'|}\\ &=\; \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3r'\,\rho(\mathbf{r}')\,\big[-4\pi\,\delta(\mathbf{r}-\mathbf{r}')\big] \;\equiv\; -\frac{1}{\epsilon_0}\,\rho(\mathbf{r}).\end{aligned} \tag{285}$$

However, Eq. (284) is not the only solution. In particular, it is useless in cases when the charge distribution $\rho(\mathbf{r},t)$ is not known from the beginning. As an example, take a point charge Q at a distance a in front of a grounded metal wall: The potential $\phi(\mathbf{r})$, and thus also the field $\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$, in front of the wall is not only due to Q itself, but also due to an unknown charge distribution on the metal surface, **influenced** by Q.

[4]For a time-independent field $\mathbf{E}(\mathbf{r})$, the potential $\phi(\mathbf{r})$ can be chosen to be time-independent as well. Then, the work done by $\mathbf{E}(\mathbf{r})$ on a charge q with mass m, moving on some curve $\Gamma \subset \Omega$, is given by

$$W_\Gamma \;\equiv\; q\int_\Gamma \mathrm{d}\mathbf{r}\cdot\mathbf{E}(\mathbf{r}) \;=\; -q\Big[\phi(\mathbf{r}_B)-\phi(\mathbf{r}_A)\Big].$$

This integral only depends on the starting point $\mathbf{r}_A$ and the endpoint $\mathbf{r}_B$ of Γ. With the corresponding speeds v_A and v_B of the charge, energy conservation then implies

$$\frac{m}{2}v_B^2 \;=\; \frac{m}{2}v_A^2 + W_\Gamma \qquad\Leftrightarrow\qquad \frac{m}{2}v_A^2 \;+\; q\phi(\mathbf{r}_A) \;=\; \frac{m}{2}v_B^2 \;+\; q\phi(\mathbf{r}_B).$$

Consequently, as a function of its position $\mathbf{r}$, the value $q\phi(\mathbf{r})$ is the electrostatic **potential energy** of the charge q due to the electric field $\mathbf{E}(\mathbf{r})$.

[5]Dirac's delta function $\delta(\mathbf{r})$ in 3D is the limit $R \to 0$ of the finite function

$$\Delta_R(\mathbf{r}) \;=\; \begin{cases} \frac{1}{\frac{4\pi}{3}R^3} & (|\mathbf{r}| \le R),\\ 0 & (|\mathbf{r}| > R).\end{cases}$$

PART II

Partial Differential Equations (PDEs)

8 Differential equations (DEs)

8.1 Partial versus ordinary

With a given charge density $\rho(\mathbf{r})$, e.g.: $\rho(\mathbf{r}) = Q\frac{2a^{5/2}}{\pi^{3/2}}\, z^2\, \mathrm{e}^{-a(x^2+y^2+z^2)}$, Poisson's equation

$$\nabla^2\phi(\mathbf{r}) \;\equiv\; \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi(x,y,z) \;=\; -\frac{1}{\epsilon_0}\rho(x,y,z) \tag{286}$$

for the unknown wanted function $\phi(\mathbf{r})$ is a differential equation (DE), since derivatives of ϕ enter. More precisely, it is a partial DE **(PDE)**, since (partial) derivatives with respect to **more** variables **than one** are involved.

In contrast, an ordinary DE **(ODE)** is for a function $f(x)$ with only one variable, e.g.,

$$f''(x) \;+\; f'(x) \;=\; 3x(2+x). \tag{287}$$

This ODE is of **second** order, since the highest derivative involved is $f''(x)$. Consequently (see below), its infinitely many solutions form a **two-parameter** set of functions,

$$f(x) \;=\; x^3 + A\mathrm{e}^{-x} + B \qquad (A, B \in \mathbb{R}). \tag{288}$$

The set of solutions to a PDE, in contrast, is more involved. E.g., Laplace's equation,

$$\nabla^2\phi(\mathbf{r}) \;\equiv\; \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi(x,y,z) \;=\; 0, \tag{289}$$

which is a special case of Poisson's (with $\rho \equiv 0$), has very different types of solutions. We list a few of them (with free parameters $A, B, C, D, E, F, \alpha, \beta \in \mathbb{R}$):

$$\begin{aligned}
\phi_1(\mathbf{r}) &= Ax^2 + By^2 - (A+B)z^2 \;+\; Dx + Ey + Fz \;+\; C, \\
\phi_2(\mathbf{r}) &= A\,\sin(\alpha x)\,\sin(\beta y)\,\mathrm{e}^{-\gamma z}, \qquad\qquad \gamma = \sqrt{\alpha^2+\beta^2}, \\
\phi_3(\mathbf{r}) &= \frac{A}{\sqrt{x^2+y^2+z^2}} \;\equiv\; A(x^2+y^2+z^2)^{-1/2} \qquad (\mathbf{r} \neq \mathbf{0}), \\
\phi_4(\mathbf{r}) &= (Ax + By + Cz)\,(x^2+y^2+z^2)^{-3/2} \qquad\qquad (\mathbf{r} \neq \mathbf{0}).
\end{aligned} \tag{290}$$

8.2 Some facts about ODEs

The ODE Eq. (287) is an **inhomogeneous linear ODE** with **constant coefficients**. In general, this important type of ODE, when of n-th order, has the form

$$f^{(n)}(x) \;+\; a_{n-1}f^{(n-1)}(x) \;+\; \ldots \;+\; a_1 f'(x) \;+\; a_0 f(x) \;=\; b(x). \tag{291}$$

The a_k are given numbers and the **inhomogeneity** $b(x)$ is a given function.
In Eq. (287), e.g., $a_2 = a_1 = 1$, $a_0 = 0$, and $b(x) = 3x(2+x)$.
In the **general solution** of Eq. (291),

$$f(x) \;=\; f^{\mathrm{inh}}_{\mathrm{spec}}(x) \;+\; f^{\mathrm{hom}}_{\mathrm{gen}}(x), \tag{292}$$

• $f^{\rm inh}_{\rm spec}(x)$ is **one particular** solution (found by an educated guess) and
• $f^{\rm hom}_{\rm gen}(x)$ is the **general solution** of the related **homogeneous** ODE with $b(x) = 0$.
Generally, $f^{\rm hom}_{\rm gen}(x)$ is a linear combination of n linearly independent functions. The ansatz

$$f^{\rm hom}(x) \;=\; {\rm e}^{\lambda x} \tag{293}$$

transforms the homogeneous ODE into an **algebraic equation** for λ,

$$\lambda^n \;+\; a_{n-1}\lambda^{n-1} \;+\; ... \;+\; a_1\lambda \;+\; a_0 \;=\; 0. \tag{294}$$

In typical cases, its roots $\lambda_1, ..., \lambda_k$ yield the wanted n independent functions.

Example 1: A third-order inh. lin. ODE with constant coefficients is

$$f'''(x) \;-\; 4f''(x) \;+\; 13f'(x) \;+\; 50f(x) \;=\; 50x^2 + 26x - 8. \tag{295}$$

By educated guess, we find

$$f^{\rm inh}_{\rm spec}(x) \;=\; x^2. \tag{296}$$

The ansatz $f^{\rm hom}(x) \;=\; {\rm e}^{\lambda x}$ for the homogeneous equation yields the algebraic equation

$$\lambda^3 - 4\lambda^2 + 13\lambda + 50 \;\equiv\; (\lambda + 2)(\lambda^2 - 6\lambda + 25) \;=\; 0, \tag{297}$$

which has three different roots,

$$\lambda_1 \;=\; -2, \qquad \lambda_{2,3} \;=\; 3 \pm 4\,{\rm i}\,. \tag{298}$$

The corresponding complex-valued solutions of the homogeneous ODE are

$$f_1(x) \;=\; {\rm e}^{-2x}, \qquad f_{2,3}(x) \;=\; {\rm e}^{(3\pm 4\,{\rm i})x} \;\equiv\; {\rm e}^{3x}\big[\cos 4x \pm {\rm i}\,\sin 4x\big]. \tag{299}$$

In each of them, both the real and the imaginary parts provide a real-valued solution. Since three linearly independent solutions of that kind are provided, we conclude

$$f^{\rm hom}_{\rm gen}(x) \;=\; A\,{\rm e}^{-2x} \;+\; B\,{\rm e}^{3x}\cos 4x \;+\; C\,{\rm e}^{3x}\sin 4x \qquad (A, B, C \in \mathbb{R}). \tag{300}$$

Consequently, any solution of the original inhomogeneous ODE has the form

$$\begin{aligned} f_(x) \;&=\; x^2 \;+\; A\,{\rm e}^{-2x} \;+\; B\,{\rm e}^{3x}\cos 4x \;+\; C\,{\rm e}^{3x}\sin 4x. && (301)\\ &=\; x^2 \;+\; A\,{\rm e}^{-2x} \;+\; \big(B\,\cos 4x \;+\; C\,\sin 4x\big)\,{\rm e}^{3x}. && (302) \end{aligned}$$

Example 2: In section 10 we shall face the following ODEs,

$$\begin{aligned} X''(x) \;&=\; -\alpha^2 X(x),\\ Y''(y) \;&=\; -\beta^2 Y(y),\\ Z''(z) \;&=\; (\alpha^2 + \beta^2)Z(z). \end{aligned} \tag{303}$$

These are homogeneous 2nd-order ($n = 2$) ODEs of the type of Eq. (291), with $b(x) = 0$. In the case $\alpha, \beta > 0$, the exponential ansatz yields the following general solutions

$$\begin{aligned} X(x) \;&=\; A_1\cos(\alpha x) \;+\; A_2\sin(\alpha x),\\ Y(y) \;&=\; B_1\cos(\beta y) \;+\; B_2\sin(\beta y),\\ Z(z) \;&=\; C_1{\rm e}^{\gamma z} \;+\; C_2{\rm e}^{-\gamma z}, \qquad \gamma = \sqrt{\alpha^2 + \beta^2}. \end{aligned} \tag{304}$$

Remark: A completely different method **("separable variables")**, applicable to many **linear or non-linear first-order ODEs**, will be treated in section 11.4.

9 Arising of PDEs in physics

Poisson's equation and the **wave equation** were our first examples of PDEs. We shall now derive several further important PDEs in physics.

9.1 Diffusion

9.1.1 Densities of mass and current

We consider a gas (or liquid, etc.) of N identical particles ("molecules") with masses m. For $i \in \{1, ..., N\}$, $\mathbf{r}_i(t)$ is the position and $\mathbf{v}_i(t)$ the velocity of particle i at time t.

Let $\Omega_a(\mathbf{r})$ be a small sphere with radius a (and volume $\frac{4\pi}{3}a^3$), centered at $\mathbf{r}$. The **mass density** $\rho(\mathbf{r}, t)$ of the gas at position $\mathbf{r}$ and time t is defined as

$$\rho(\mathbf{r}, t) \;=\; \frac{N_a(\mathbf{r}, t)\, m}{\frac{4\pi}{3}a^3}, \tag{305}$$

where $N_a(\mathbf{r}, t)$ is the number of particles with positions $\mathbf{r}_i(t)$ inside $\Omega_a(\mathbf{r})$ at the time t,

$$N_a(\mathbf{r}, t) \;=\; \sum_{\mathbf{r}_i(t) \in \Omega_a(\mathbf{r})} 1. \tag{306}$$

$\rho(\mathbf{r}, t) \equiv \rho(x, y, z, t)$ is a **smooth function** of x, y, z, t, when the radius a, smaller than the resolution L of position measurement, is considerably larger than the typical distance ℓ between molecules,

$$L \;\gg\; a \;\gg\; \ell. \tag{307}$$

The **average velocity** of a molecule with position vector $\mathbf{r}$ at time t is defined as

$$\mathbf{v}(\mathbf{r}, t) \;=\; \frac{1}{N_a(\mathbf{r}, t)} \sum_{\mathbf{r}_i(t) \in \Omega_a(\mathbf{r})} \mathbf{v}_i(t). \tag{308}$$

Defining the vector field of the **current density** of the gas as the product

$$\mathbf{J}(\mathbf{r}, t) \;=\; \rho(\mathbf{r}, t)\, \mathbf{v}(\mathbf{r}, t), \tag{309}$$

the amount $I_\Sigma(t)$ of gas (in kg/s) passing at the time t per unit time through a piece Σ of surface is given by the flux integral

$$I_\Sigma(t) \;=\; \int_\Sigma \mathrm{d}\mathbf{A} \cdot \mathbf{J}(\mathbf{r}, t) \;\equiv\; \int_\Sigma \mathrm{d}^2 r\, \mathbf{n}(\mathbf{r}) \cdot \mathbf{J}(\mathbf{r}, t). \tag{310}$$

9.1.2 Continuity equation

We now consider some arbitrary finite volume region Ω (with surface $\partial\Omega$). If the mass of CO_2 created chemically per unit time within a small volume d^3r around the position $\mathbf{r}$ at time t is given by $s(\mathbf{r},t)\,\mathrm{d}^3r$, the rate of change of its mass within Ω is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{d}^3r\,\rho(\mathbf{r},t) \;=\; \int_\Omega \mathrm{d}^3r\,s(\mathbf{r},t) \;-\; \oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t). \tag{311}$$

Here, the RHS represents the amount of CO_2 created chemically within Ω minus the amount leaving Ω by flow through the surface $\partial\Omega$. [Note that $\oint_{\partial\Omega}\mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t)$ is defined as the flux of $\mathbf{J}$ *out* of Ω.] By Gauß's law, we have

$$\oint_{\partial\Omega} \mathrm{d}\mathbf{A}\cdot\mathbf{J}(\mathbf{r},t) \;=\; \int_\Omega \mathrm{d}^3r\,\nabla\cdot\mathbf{J}(\mathbf{r},t). \tag{312}$$

Therefore, since Eq. (311) must hold for any region Ω, we conclude

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) \;+\; \nabla\cdot\mathbf{J}(\mathbf{r},t) \;=\; s(\mathbf{r},t). \tag{313}$$

When the gas is a **conserved quantity**, $s(\mathbf{r},t)=0$, we have the **continuity equation**

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) \;+\; \nabla\cdot\mathbf{J}(\mathbf{r},t) \;=\; 0. \tag{314}$$

9.1.3 Diffusion and convection

The velocity field is typically due to both **diffusion** and **convection**,

$$\mathbf{v}(\mathbf{r},t) \;=\; \mathbf{v}_{\mathrm{diff}}(\mathbf{r},t) \;+\; \mathbf{v}_{\mathrm{conv}}(\mathbf{r},t). \tag{315}$$

For example, if the "gas" is CO_2 in air, $\mathbf{v}_{\mathrm{conv}}(\mathbf{r},t)$ is the **wind velocity** of the carrier medium (air). In contrast, $\mathbf{v}_{\mathrm{diff}}(\mathbf{r},t)$ is the **average velocity** of the CO_2 molecules in the absence of wind, due to **diffusion** between the air molecules (N_2, O_2, Ar).

Diffusion law: The origin of diffusion is the density gradient,

$$\mathbf{J}_{\mathrm{diff}}(\mathbf{r},t) \;\equiv\; \rho(\mathbf{r},t)\,\mathbf{v}_{\mathrm{diff}}(\mathbf{r},t) \;=\; -D\,\nabla\rho(\mathbf{r},t). \tag{316}$$

The **diffusion constant** D has the physical dimension $1\ \frac{\mathrm{m}^2}{\mathrm{s}}$.[6]

The total flux of the gas is

$$\mathbf{J}(\mathbf{r},t) \;=\; -D\,\nabla\rho(\mathbf{r},t) \;+\; \rho(\mathbf{r},t)\mathbf{v}_{\mathrm{conv}}(\mathbf{r},t). \tag{317}$$

[6] Boltzmann's transport theory yields the expression

$$D \;=\; \frac{t_{\mathrm{c}}k_{\mathrm{B}}T}{m},$$

where the relaxation time t_{c} is close to the average collision time between molecules. Consequently, this "constant" can actually be a function $D(\mathbf{r},t)$.

9.1.4 Diffusion equation

For simplicity, we assume $\mathbf{v}_{\rm conv}(\mathbf{r},t) = 0$,

$$\mathbf{J}(\mathbf{r},t) \;=\; -D\,\nabla\rho(\mathbf{r},t). \tag{318}$$

Taking here the divergence $\nabla\cdot$... and using Eq. (313), we obtain the **diffusion equation**,

$$\frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;=\; D\,\nabla^2\rho(\mathbf{r},t) \;+\; s(\mathbf{r},t). \tag{319}$$

For a given value of the constant D and a given function $s(\mathbf{r},t)$, this is a PDE for the unknown wanted function $\rho(\mathbf{r},t) \equiv \rho(x,y,z,t)$ of four variables x, y, z, and t,

$$\frac{\partial}{\partial t}\,\rho(x,y,z,t) \;=\; D\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\rho(x,y,z,t) \;+\; s(x,y,z,t). \tag{320}$$

9.1.5 Example for a solution

$\rho(\mathbf{r},t)$ shall be the time-dependent density distribution of **ink in water**. In this case, the diffusion constant has the empirical value

$$D \;\approx\; 2.5\cdot 10^{-4}\,\frac{\rm mm^2}{\rm s} \qquad \text{[Eur. J. Phys. } \mathbf{25}\text{, 331 (2004)]}. \tag{321}$$

At the time $t = 0$, a small spherical droplet of ink is placed at the position $\mathbf{r} = \mathbf{0}$ deep inside the water. No more ink is added to this system at any time $t > 0$,

$$s(\mathbf{r},t) \;=\; 0 \qquad (\text{for } t > 0). \tag{322}$$

Due to diffusion, the ink, originally concentrated around the point $\mathbf{r} = \mathbf{0}$, will slowly spread out into the surrounding water. This effect is qualitatively described by the function

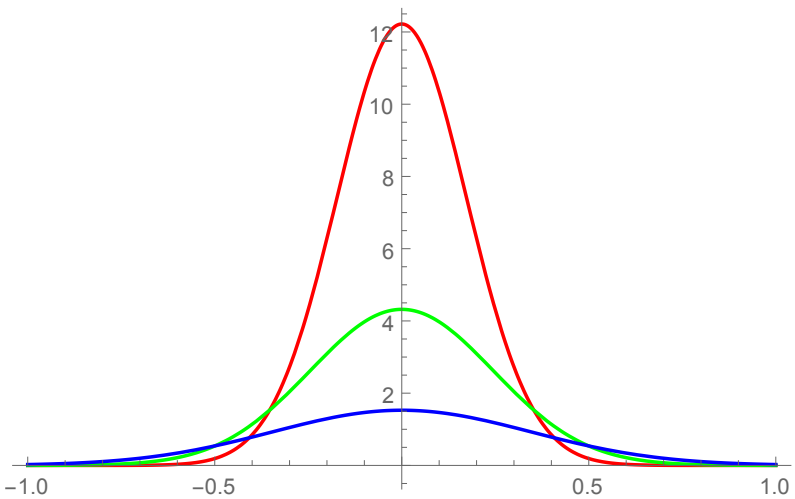
$$\begin{aligned}\rho(\mathbf{r},t) \;=\; \rho_{\rm G}(\mathbf{r},t) \;&\equiv\; \frac{M}{(4\pi Dt)^{3/2}}\,{\rm e}^{-\mathbf{r}^2/4Dt}\\ &\equiv\; \frac{M}{(4\pi Dt)^{3/2}}\,{\rm e}^{-(x^2+y^2+z^2)/4Dt} \;=\; \rho_{\rm G}(x,y,z,t) \qquad (t>0),\end{aligned} \tag{323}$$

where M is a constant with the dimension of a mass (kg). We shall see below that $\rho_{\rm G}(x,y,z,t)$ is in fact an exact solution of the diffusion equation (320). Note that the exponent $\frac{\mathbf{r}^2}{4Dt}$ is dimensionless, while the prefactor has the correct dimension $1\frac{\rm kg}{\rm m^3}$ of a mass density. The function $\rho_{\rm G}(x,0,0,t)$ is plotted versus x in the following figure for the three different time values $t = 60$ s (red), $t = 120$ s (green), and $t = 240$ s (blue).

Since no ink is added after $t = 0$, the total mass in the density distribution $\rho_{\rm G}(\mathbf{r},t)$ should be a constant, not depending on t. In fact, elementary triple integration yields

$$\int {\rm d}^3r\,\rho_{\rm G}(\mathbf{r},t) \;\equiv\; \frac{M}{(4\pi Dt)^{3/2}}\int_{-\infty}^{\infty}{\rm d}x\int_{-\infty}^{\infty}{\rm d}y\int_{-\infty}^{\infty}{\rm d}z\,{\rm e}^{-(x^2+y^2+z^2)/4Dt} \;=\; M, \tag{324}$$

where we have used $\int_0^\infty {\rm d}u\,u^2{\rm e}^{-u^2} = \frac{1}{4}\sqrt{\pi}$. After all, $\rho_{\rm G}(\mathbf{r},t)$ is a reasonable candidate for the unknown true time-dependent density distribution of the ink.



To see that $\rho_{\rm G}(\mathbf{r},t)$ is in fact a solution of Eq. (320), we apply the product rule to find

$$\begin{aligned}\frac{\partial}{\partial t}\rho_{\rm G}(x,y,z,t) &= -\frac{3}{2}\frac{\rho_{\rm G}(x,y,z,t)}{t} - \Big[-\frac{(x^2+y^2+z^2)}{4Dt^2}\rho_{\rm G}(x,y,z,t)\Big] \\ &= \Big(-\frac{3}{2t}+\frac{(x^2+y^2+z^2)}{4Dt^2}\Big)\rho_{\rm G}(x,y,z,t). \end{aligned} \tag{325}$$

On the other hand, we have

$$\begin{aligned} D\frac{\partial^2}{\partial x^2}\rho_{\rm G}(x,y,z,t) &= D\frac{\partial}{\partial x}\Big(\frac{-2x}{4Dt}\rho_{\rm G}(x,y,z,t)\Big) \\ &= D\Big(-\frac{2}{4Dt}+\frac{(-2x)^2}{(4Dt)^2}\Big)\rho_{\rm G}(x,y,z,t) \\ &= \Big(-\frac{1}{2t}+\frac{x^2}{4Dt^2}\Big)\rho_{\rm G}(x,y,z,t). \end{aligned} \tag{326}$$

Consequently, adding the corresponding expressions for $D\frac{\partial^2\rho_{\rm G}}{\partial y^2}$ and $D\frac{\partial^2\rho_{\rm G}}{\partial z^2}$, we obtain

$$D\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\rho_{\rm G}(x,y,z,t) = \Big(-\frac{3}{2t}+\frac{(x^2+y^2+z^2)}{4Dt^2}\Big)\rho_{\rm G}(x,y,z,t), \tag{327}$$

which is the same as the result (325). Therefore, $\rho_{\rm G}(\mathbf{r},t)$ is an exact solution of the PDE (320) in the special case $s(\mathbf{r},t)=0$.

Remark 1: The integral (324) is more appropriately evaluated in spherical coordinates,

$$\int {\rm d}^3r\,\rho_{\rm G}(\mathbf{r},t) = \frac{M}{(4\pi Dt)^{3/2}}\int_0^\infty {\rm d}r\,(4\pi r^2){\rm e}^{-r^2/4Dt} = \frac{M}{\pi^{3/2}}4\pi\int_0^\infty {\rm d}u\,u^2{\rm e}^{-u^2} = M. \tag{328}$$

Remark 2: The result (327) can be obtained in vector notation: With the gradient

$$\nabla\rho_{\rm G}(\mathbf{r},t) = -\frac{\mathbf{r}}{2Dt}\rho_{\rm G}(\mathbf{r},t), \tag{329}$$

we have

$$\begin{aligned} D\,\nabla^2\rho_{\rm G}(\mathbf{r},t) \equiv D\,\nabla\cdot\big[\nabla\rho_{\rm G}(\mathbf{r},t)\big] &= D\,\nabla\cdot\Big[-\frac{\mathbf{r}}{2Dt}\rho_{\rm G}(\mathbf{r},t)\Big] \\ &= -\frac{1}{2t}\Big[(\nabla\cdot\mathbf{r})\rho_{\rm G}(\mathbf{r},t) + \mathbf{r}\cdot\nabla\rho_{\rm G}(\mathbf{r},t)\Big] \\ &= -\frac{1}{2t}\Big[3\,\rho_{\rm G}(\mathbf{r},t) - \frac{\mathbf{r}^2}{2Dt}\rho_{\rm G}(\mathbf{r},t)\Big] \\ &= \Big(-\frac{3}{2t}+\frac{\mathbf{r}^2}{4Dt^2}\Big)\rho_{\rm G}(\mathbf{r},t). \end{aligned} \tag{330}$$

9.2 Heat flow

9.2.1 Specific heat

In a **uniform medium** (liquid water, solid silicon, etc.), we treat **heat (energy)** as a continuous quantity with generally non-uniform density $\rho(\mathbf{r},t)$ (in kJ/m^3), satisfying

$$\rho(\mathbf{r},t) \;=\; c\,\mu\,\big[T(\mathbf{r},t) - T_0\big]. \tag{331}$$

Here, $T(\mathbf{r},t)$ is the **temperature distribution** in the medium, $\mu = \frac{\Delta m}{\Delta V}$ is its uniform mass density (in kg/m^3), and c is its **specific heat** (in kJ/kgK). $T_0 \approx 300$ K is a constant reference temperature.

Remark: Multiplying with ΔV yields the familiar relation $Q = c\,m\,\Delta T$.

If there are no heat sources or sinks in the medium, density $\rho(\mathbf{r},t)$ and current density $\mathbf{J}(\mathbf{r},t)$ of heat satisfy the continuity equation (314),

$$\frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;+\; \nabla\cdot\mathbf{J}(\mathbf{r},t) \;=\; 0. \tag{332}$$

On the other hand, Eq. (331) implies

$$\frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;=\; c\,\mu\,\frac{\partial}{\partial t}\,T(\mathbf{r},t). \tag{333}$$

9.2.2 Heat conduction

Law of heat conduction: $\mathbf{J}(\mathbf{r},t)$ is proportional to the gradient of $T(\mathbf{r},t)$,

$$\mathbf{J}(\mathbf{r},t) \;=\; -\lambda\,\nabla\,T(\mathbf{r},t). \tag{334}$$

This law is quite reminiscent of the diffusion law (316).

Combining Eqs. (332) and (333), we may use Eq. (334) to eliminate $\mathbf{J}$,

$$c\,\mu\,\frac{\partial}{\partial t}\,T(\mathbf{r},t) \;=\; -\nabla\cdot\mathbf{J}(\mathbf{r},t) \;=\; \lambda\nabla\cdot\big[\nabla\,T(\mathbf{r},t)\big] \;\equiv\; \lambda\,\nabla^2 T(\mathbf{r},t). \tag{335}$$

This is a PDE for the temperature distribution $T(\mathbf{r},t)$ in the medium,

$$\frac{\partial}{\partial t}\,T(\mathbf{r},t) \;=\; \frac{\lambda}{c\,\mu}\,\nabla^2 T(\mathbf{r},t). \tag{336}$$

9.2.3 Laplace's equation

A **stationary** temperature distribution, with $\frac{\partial}{\partial t}\,T(\mathbf{r},t) = 0$, obeys **Laplace's equation**,

$$\nabla^2 T(\mathbf{r}) \;=\; 0. \tag{337}$$

9.3 Waves

9.3.1 Young's modulus of a spring

By **Hooke's law**, the restoring force of a spring with spring constant $k > 0$ is

$$F(u) \;=\; -ku, \tag{338}$$

where u is the displacement ($u > 0$: stretched spring, $u < 0$: compressed spring). In any case, F has the opposite sign (direction) of u.

As the spring is cut into two equal pieces, the resulting two smaller springs have spring constant $k' = 2k$ each. Generally, a spring with equilibrium length L has the constant

$$k \;=\; \frac{Y}{L}, \tag{339}$$

where **Young's modulus** Y (in N) is a material constant, independent of L.

9.3.2 Continuum limit of a linear chain

A linear chain is an array of (infinitely) many identical point masses m on the x-axis which at equilibrium occupy the equidistant sites

$$x_\alpha \;=\; \alpha \cdot \Delta x \qquad (\alpha \in \mathbb{Z}). \tag{340}$$

Joined by identical springs (each with spring constant k and equilibrium length $L = \Delta x$), the point masses may perform longitudinal oscillations in x-direction.

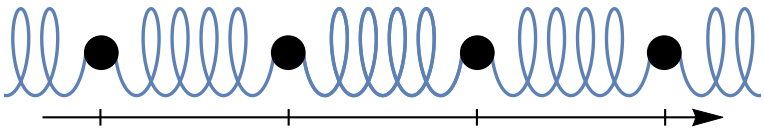


Figure 8: A section of a linear chain, with four masses (black dots) at their equilibrium positions $x_{\alpha-1}$, x_α, $x_{\alpha+1}$, and $x_{\alpha+2}$, respectively, on the x-axis.

Let $u_\alpha(t)$ be the displacement of mass α from its equilibrium position x_α at time t. Then, its equation of motion reads

$$m\,\ddot{u}_\alpha(t) \;=\; k\big[u_{\alpha+1}(t) - u_\alpha(t)\big] \;-\; k\big[u_\alpha(t) - u_{\alpha-1}(t)\big]. \tag{341}$$

Writing $u_\alpha(t) = u(x_\alpha, t)$, with a function $u(x,t)$ of x and t, this equation reads

$$m\,\frac{\partial^2 u(x_\alpha,t)}{\partial t^2} \;=\; k\,\big[u(x_\alpha + \Delta x, t) \;-\; u(x_\alpha, t)\big] \;-\; k\,\big[u(x_\alpha, t) \;-\; u(x_\alpha - \Delta x, t)\big]. \tag{342}$$

We divide by Δx, and write

$$\frac{m}{\Delta x}\,\frac{\partial^2 u(x_\alpha,t)}{\partial t^2} \;=\; k\,\Delta x \cdot \frac{\dfrac{u(x_\alpha + \Delta x, t) \;-\; u(x_\alpha, t)}{\Delta x} \;-\; \dfrac{u(x_\alpha, t) \;-\; u(x_\alpha - \Delta x, t)}{\Delta x}}{\Delta x}. \tag{343}$$

Here, $\frac{m}{\Delta x} = \mu$ is the mass per unit length (in $\frac{\mathrm{kg}}{\mathrm{m}}$) of the chain and $k\Delta x = Y$ is Young's modulus of the springs.

In the limit $\Delta x \to 0$, when the chain becomes a **continuous rod**, we obtain

$$\begin{aligned} \mu \frac{\partial^2 u(x,t)}{\partial t^2} &= Y \lim_{\Delta x \to 0} \frac{\frac{\partial u(x,t)}{\partial x} - \frac{\partial u(x-\Delta x,t)}{\partial x}}{\Delta x} \\ &\equiv Y \frac{\partial^2 u(x,t)}{\partial x^2}. \end{aligned} \tag{344}$$

This is a PDE for the function $u(x,t)$. Describing **longitudinal vibrations** of the rod, $u(x,t)$ is the displacement u at time t of a point with equilibrium position x on the rod.

9.3.3 The wave equation

The latter PDE is the one-dimensional **wave equation**,

$$\frac{1}{c^2} \frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2}. \tag{345}$$

The constant c, here $c = \sqrt{Y/\mu}$, in general turns out to be the **velocity** in x-direction of **waves** that are described by the solution $u(x,t)$ (see section 10.3.1).

The same equation holds for the **transversal vibration** in y-direction of a **violin string** on the x-axis, when Y is replaced with the **string tension** F. When a 1D **sound wave** is propagating in a thin rigid straight pipe (along the x-axis), the pressure $P(x,t)$ of the medium in the pipe satisfies the wave equation

$$\frac{\partial^2 P(x,t)}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 P(x,t)}{\partial t^2} = 0, \qquad c = \frac{1}{\sqrt{\kappa \rho}}, \tag{346}$$

where ρ is the mass density and κ the **adiabatic compressibility** of the medium,

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_S. \tag{347}$$

For perfect gases, with $\mathrm{d}Q = \frac{f+2}{2} P \mathrm{d}V + \frac{f}{2} V \mathrm{d}P$, we find: $\kappa = \frac{f}{f+2} \frac{1}{P}$.

System	$u(x,t)$	c
Vibrating **rod**	**Longitudinal displacement** at time t in x-direction of a point on the rod which at rest would be at position x.	$\sqrt{\frac{Y}{\mu}}$
Vibrating **string**	**Transversal displacement** at time t in y-direction of a point on the string with x-coordinate x.	$\sqrt{\frac{F}{\mu}}$
Sound in a thin straight pipe	**Air pressure** $u(x,t) = P(x,t)$ at time t at position x in the pipe	$\frac{1}{\sqrt{\kappa\rho}}$

Eq. (345) is a special case of the wave equation in three dimensions,

$$\frac{1}{c^2} \frac{\partial^2 u(x,y,z,t)}{\partial t^2} = \frac{\partial^2 u(x,y,z,t)}{\partial x^2} + \frac{\partial^2 u(x,y,z,t)}{\partial y^2} + \frac{\partial^2 u(x,y,z,t)}{\partial z^2} \tag{348}$$

or, in short-hand notation,

$$\frac{1}{c^2} \frac{\partial^2 u(\mathbf{r},t)}{\partial t^2} = \nabla^2 u(\mathbf{r},t). \tag{349}$$

9.4 Classification of PDEs

A second-order linear PDE for a function $u(x,y)$ with two variables has the general form

$$A\,u_{xx} + B\,u_{xy} + C\,u_{yy} + D\,u_x + E\,u_y + F\,u + G \;=\; 0, \tag{350}$$

in the short-hand notation $u_{xx} = \frac{\partial^2}{\partial x^2}\,u(x,y)$, $u_{xy} = \frac{\partial^2}{\partial x \partial y}\,u(x,y)$, $u_x = \frac{\partial}{\partial x}\,u(x,y)$, etc. The coefficients $A, B, ..., G$ are given functions of (x,y) or, typically, constant numbers.

Examples:

Name of the PDE	Form of the PDE	non-zero coeffs	PDE type
1D wave equation	$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\,u(x,t) \;=\; \frac{\partial^2}{\partial x^2}\,u(x,t)$	$A=-1,\ C=\frac{1}{c^2}.$	hyperbolic
1D heat conduction	$\frac{\partial}{\partial t}\,u(x,t) \;=\; \frac{\lambda}{c\mu}\frac{\partial^2}{\partial x^2}\,u(x,t)$	$A=-\frac{\lambda}{c\mu},\ E=1$	parabolic
2D Laplace equation	$\frac{\partial^2}{\partial x^2}\,u(x,y) + \frac{\partial^2}{\partial y^2}\,u(x,y) \;=\; 0$	$A=1,\ C=1$	elliptic

Def.: The PDE Eq. (350) is called[7]

$$\left.\begin{array}{c}\textbf{hyperbolic}\\ \textbf{parabolic}\\ \textbf{elliptic}\end{array}\right\}\ \text{when}\ \left\{\begin{array}{l} B^2-4AC \;>\; 0,\\ B^2-4AC \;=\; 0,\\ B^2-4AC \;<\; 0.\end{array}\right.$$

Theorem: The PDE Eq. (350) has n **real characteristics** (RCs), where $n=2$ in the hyperbolic case, $n=1$ in the parabolic case, and $n=0$ in the elliptic case.

[7]Geometrical reason: The solutions $(x,y)\in\mathbb{R}^2$ to the two-variable quadratic equation

$$ax^2 + b\,xy + c\,y^2 \;=\; 1$$

form a hyperbola ($b^2-4ac>0$), a parabola ($b^2-4ac=0$) or an ellipse ($b^2-4ac<0$).
To see this, we rotate in the case $b\neq 0$ the axes in the xy-plane by a proper angle ϕ,

$$\begin{aligned} x &= \xi\cos\phi - \eta\sin\phi,\\ y &= \xi\sin\phi + \eta\cos\phi,\end{aligned}$$

to obtain the simpler equation $\frac{\xi^2}{\alpha}+\frac{\eta^2}{\beta}=1$. The required angle ϕ only needs to satisfy

$$\cos^2\phi \;=\; \frac{1}{2}\left[1 \pm \sqrt{\frac{(c-a)^2}{b^2+(c-a)^2}}\,\right].$$

10 Analytical methods for solving PDEs

10.1 Diffusion equation: Green's function

Theorem: Let $\rho(\mathbf{r}, t = 0) = \rho_0(\mathbf{r})$ be the **initial density distribution** (at the time $t = 0$) of a diffusing substance (e.g., ink in water). Then, provided that there are no sources or sinks, the future distribution (at times $t > 0$) is given by the volume integral

$$\rho(\mathbf{r}, t) \;=\; \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \, G(\mathbf{r}' - \mathbf{r}, t) \qquad \text{(a CONVOLUTION integral !)}, \tag{351}$$

with the so-called **Green's function** of the diffusion equation (cf. section 9.1.5)

$$G(\mathbf{r}' - \mathbf{r}, t) \;=\; \frac{\mathrm{e}^{-(\mathbf{r}'-\mathbf{r})^2/4Dt}}{(4\pi D t)^{3/2}} \;\equiv\; \frac{1}{M} \, \rho_{\mathrm{G}}(\mathbf{r}' - \mathbf{r}, t). \tag{352}$$

Proof: • Since $G(\mathbf{r}, t) = \frac{\rho_{\mathrm{G}}(\mathbf{r},t)}{M}$ satisfies the diffusion equation, see section 9.1.5, we have

$$\begin{aligned} \frac{\partial}{\partial t} G(\mathbf{r}' - \mathbf{r}, t) \;&\equiv\; \left[\frac{\partial}{\partial t} G(\mathbf{s}, t)\right]\Big|_{\mathbf{s}=\mathbf{r}'-\mathbf{r}} \\ &=\; \Big[D \, \nabla_{\mathbf{s}}^2 G(\mathbf{s}, t)\Big]\Big|_{\mathbf{s}=\mathbf{r}'-\mathbf{r}} \;=\; D \, \nabla_{\mathbf{r}}^2 G(\mathbf{r}' - \mathbf{r}, t). \end{aligned} \tag{353}$$

Consequently, we have in Eq. (351)

$$\begin{aligned} \frac{\partial}{\partial t} \rho(\mathbf{r}, t) \;&\equiv\; \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \, \frac{\partial}{\partial t} G(\mathbf{r}' - \mathbf{r}, t) \\ &=\; \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \; D \, \nabla_{\mathbf{r}}^2 G(\mathbf{r}' - \mathbf{r}, t) \\ &=\; D \, \nabla_{\mathbf{r}}^2 \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \, G(\mathbf{r}' - \mathbf{r}, t) \;\equiv\; D \, \nabla^2 \rho(\mathbf{r}, t), \end{aligned} \tag{354}$$

and $\rho(\mathbf{r}, t)$ is in fact a solution of the diffusion equation.
• We still have to show that $\rho(\mathbf{r}, t) \to \rho_0(\mathbf{r})$ as $t \to 0$:
Since $\int \mathrm{d}^3 r \, \rho_{\mathrm{G}}(\mathbf{r}, t) = M$, see Eq. (324) in section 9.1.5, we have at any time $t > 0$

$$\int \mathrm{d}^3 r' \, G(\mathbf{r}' - \mathbf{r}, t) \;=\; 1. \tag{355}$$

On the other hand, for any $\mathbf{r}' \neq \mathbf{r}$, we have $G(\mathbf{r}' - \mathbf{r}, t) \to 0$ as $t \to 0$, see Eq. (352). Consequently, as a function of $\mathbf{r}'$, $G(\mathbf{r}' - \mathbf{r}, t)$ becomes infinitely **peaked** around $\mathbf{r}' = \mathbf{r}$,

$$\lim_{t \to 0-0} \rho(\mathbf{r}, t) \;\equiv\; \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \, \Big[\underbrace{\lim_{t \to 0-0} G(\mathbf{r}' - \mathbf{r}, t)}_{\delta(\mathbf{r}'-\mathbf{r})} \Big] \;=\; \rho_0(\mathbf{r}). \tag{356}$$

Example: A large water container is separated by a wall (yz-plane) into two halfs. The left half ($x < 0$) contains ink with uniform density ρ_1, the right half ($x > 0$) is ink-free,

$$\rho_0(\mathbf{r}) \;\equiv\; \rho_0(x, y, z) \;=\; \begin{cases} \rho_1 & (x < 0), \\ 0 & (x > 0). \end{cases} \tag{357}$$

As the wall is removed at $t = 0$, the ink distribution $\rho(\mathbf{r}, t)$ for any time $t > 0$ is

$$\rho(\mathbf{r}, t) \;\equiv\; \int \mathrm{d}^3 r' \, \rho_0(\mathbf{r}') \, G(\mathbf{r}' - \mathbf{r}, t) \;=\; \rho_1 \int_{-\infty}^{0} \mathrm{d}x' \int_{-\infty}^{\infty} \mathrm{d}y' \int_{-\infty}^{\infty} \mathrm{d}z' \, G(\mathbf{r}' - \mathbf{r}, t). \tag{358}$$

With Eq. (352), we explicitly have

$$\rho(\mathbf{r}, t) \;=\; \frac{\rho_1}{(4\pi D t)^{3/2}} \int_{-\infty}^{0} \mathrm{d}x' \, \mathrm{e}^{-(x'-x)^2/4Dt} \int_{-\infty}^{\infty} \mathrm{d}y' \, \mathrm{e}^{-(y'-y)^2/4Dt} \int_{-\infty}^{\infty} \mathrm{d}z' \, \mathrm{e}^{-(z'-z)^2/4Dt}. \tag{359}$$

Substituting $\frac{x'-x}{\sqrt{4Dt}} = u$, with $\mathrm{d}x' = (4Dt)^{1/2} \mathrm{d}u$, as well as $\frac{y'-y}{\sqrt{4Dt}} = v$, $\frac{z'-z}{\sqrt{4Dt}} = w$, we get

$$\begin{aligned} \rho(\mathbf{r}, t) \;&=\; \frac{\rho_1}{(4\pi D t)^{3/2}} (4Dt)^{3/2} \int_{-\infty}^{-x/\sqrt{4Dt}} \mathrm{d}u \, \mathrm{e}^{-u^2} \underbrace{\int_{-\infty}^{\infty} \mathrm{d}v \, \mathrm{e}^{-v^2}}_{\sqrt{\pi}} \underbrace{\int_{-\infty}^{\infty} \mathrm{d}w \, \mathrm{e}^{-w^2}}_{\sqrt{\pi}} \\ &=\; \frac{\rho_1}{\sqrt{\pi}} \int_{-\infty}^{-x/\sqrt{4Dt}} \mathrm{d}u \, \mathrm{e}^{-u^2} \;\equiv\; \rho(x, t). \end{aligned} \tag{360}$$

In terms of the **error function** $\mathrm{erf}(s) \;=\; \frac{2}{\sqrt{\pi}} \int_0^s \mathrm{d}u \, \mathrm{e}^{-u^2}$, we obtain

$$\begin{aligned} \rho(x, t) \;&=\; \frac{\rho_1}{\sqrt{\pi}} \left\{ \int_{-\infty}^{0} + \int_{0}^{-x/\sqrt{4Dt}} \right\} \mathrm{d}u \, \mathrm{e}^{-u^2} \\ &=\; \frac{\rho_1}{\sqrt{\pi}} \left\{ \frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{2} \mathrm{erf}\Big(- \frac{x}{\sqrt{4Dt}} \Big) \right\} \\ &=\; \frac{\rho_1}{2} \left\{ 1 + \mathrm{erf}\Big(- \frac{x}{\sqrt{4Dt}} \Big) \right\}. \end{aligned} \tag{361}$$

• The function $f(x) = \frac{1}{2}\{1 + \mathrm{erf}(-\frac{x}{a})\}$ is monotonically decreasing from $f(-\infty) = 1$ to $f(+\infty) = 0$, with $f(0) = \frac{1}{2}$ and $f'(0) = -\frac{1}{a\sqrt{\pi}}$, see Fig. 9.

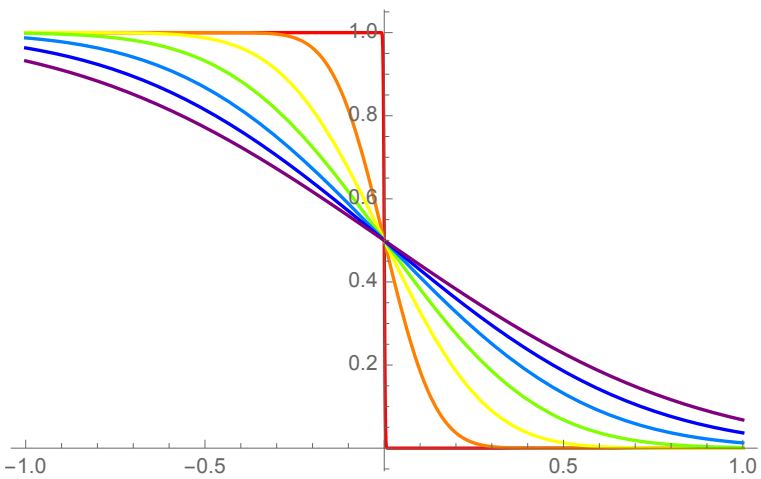


Figure 9: The function $\rho(x, t)$ of Eq. (361) in units of ρ_1, plotted versus x (in mm), for different times t (in s): $t = 0.01$ (red), 25 (orange) 100, 225, 400, 625 and 900 (violet), using the diffusion constant $D = 2.5 \cdot 10^{-4} \, \frac{\mathrm{mm}^2}{\mathrm{s}}$ for **ink in water**, Eq. (321).

• Notice that $\rho(x, t)$ of Eq. (361) is a solution of the **1D diffusion equation** $\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2}$.

10.2 Heat flow: Separation ansatz

A rectangular metal plate in the xy-plane, with corners $(x, y) = (0, 0)$, $(a, 0)$, $(0, b)$, (a, b), has its top edge $y = b$ held at some temperature $T = T_0$ (red in the figure), while the other three edges (blue in the figure) are held at $T = 0$ (in °C).

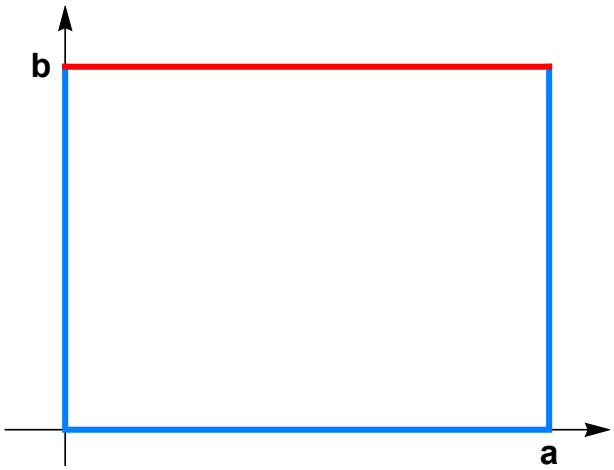


Figure 10: The boundary conditions of Eqs. (363) and (364).

When the plate, away from its edges, has no heat sources or sinks, its temperature distribution $T(x, y)$ at steady-state equilibrium is that solution to **Laplace's equation**,

$$\nabla^2 T(x, y) \;\equiv\; \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) T(x, y) \;=\; 0, \tag{362}$$

which satisfies the corresponding **boundary conditions** (BCs)

$$T(0, y) \;=\; T(a, y) \;=\; T(x, 0) \;=\; 0 \qquad (0 \le x \le a,\; 0 \le y \le b), \tag{363}$$

$$T(x, b) \;=\; t(x) \qquad (0 \le x \le a), \tag{364}$$

where $t(x)$ may be any **given function**. In our present example, $t(x) = T_0$ (a constant).

10.2.1 Factorizing solutions

As a starting point, we look for solutions of the factorized form (**separation ansatz**)

$$T_{\text{sep}}(x, y) \;=\; X(x)Y(y). \tag{365}$$

Typically, such a solution can satisfy only the homogeneous BCs, Eq. (363), but not the inhomogeneous one, Eq. (364), with $t(x) \neq 0$, see section 10.2.2.

With the ansatz (365), our PDE Eq. (362) reads

$$\nabla^2 T_{\text{sep}}(x, y) \;\equiv\; X''(x)Y(y) \;+\; X(x)Y''(y) \;=\; 0.$$

Dividing both sides by $X(x)Y(y)$, we obtain

$$\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} \;=\; 0. \tag{366}$$

Since the variables $x \in [0, a]$ and $y \in [0, b]$ may vary independently of each other, Eq. (366) can be true only when each summand on the LHS is a **constant**, $\mp k^2$ say (with $k \in \mathbb{R}$),

$$X''(x) \;=\; -k^2 X(x), \qquad Y''(y) \;=\; k^2 Y(y). \tag{367}$$

(Choosing opposite signs, we could not satisfy the first two of the four conditions below!) The general solutions of these ODEs (for any value of the yet unknown k) are

$$X(x) \;=\; A\cos(kx) \,+\, B\sin(kx), \qquad Y(y) \;=\; C\mathrm{e}^{ky} \,+\, D\mathrm{e}^{-ky} \qquad (A, B, C, D \in \mathbb{R}).$$

- The <u>condition $T_{\text{sep}}(0, y) = 0$</u> requires that $X(0) = 0$ or, equivalently, $A = 0$.
- Then, the <u>condition $T_{\text{sep}}(a, y) = 0$</u> requires that $X(a) = 0$ or $ka = n\pi$, with $n \in \mathbb{Z}$,

$$k \;=\; n\,\frac{\pi}{a} \;\equiv\; k_n \qquad (n \in \mathbb{Z}).$$

- The <u>condition $T_{\text{sep}}(x, 0) = 0$</u> requires that $Y(0) = 0$ or $D = -C$, and we have

$$\begin{aligned} T_{\text{sep}}(x, y) \;&=\; B\sin(k_n x)\, C\big[\mathrm{e}^{k_n y} - \mathrm{e}^{-k_n y}\big] \\ &\equiv\; B'\sin(k_n x)\,\sinh(k_n y) \qquad (n \in \mathbb{N}), \end{aligned} \tag{368}$$

where $\sinh(u) = \frac{1}{2}(\mathrm{e}^u - \mathrm{e}^{-u})$ and thus $B' = 2BC$; the integer $n \in \mathbb{N}$ is not fixed. For $n = 5$ and $B' = 1$, this function is plotted in Fig. 11.

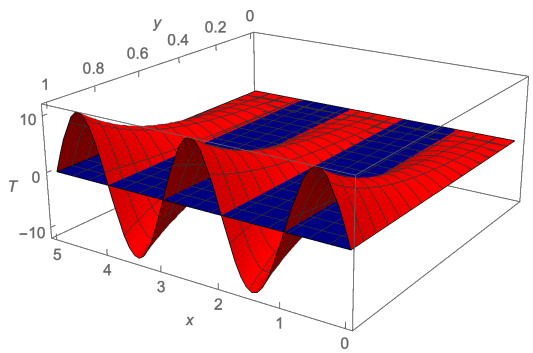


Figure 11: The function $T_{\text{sep}}(x, y) = \sin(k_5 x)\,\sinh(k_5 y)$ for $a = 5$ and $0 \le y \le 1$.

10.2.2 Superposition

$T_{\text{sep}}(x, y)$ cannot satisfy the inh BC $T(x, b) = t(x)$, with a given function $t(x)$, since

$$T_{\text{sep}}(x, b) \;\equiv\; B'\sin(k_n x)\,\sinh(k_n b) \tag{369}$$

is either zero (for $n = 0$) or (for $n \neq 0$) is a **sine function** of $x \in [0, a]$, see Fig. 11. However, any **superposition** of these factorising functions,

$$T(x, y) \;=\; \sum_{n=1}^{\infty} B'_n \sin(k_n x)\sinh(k_n y), \tag{370}$$

with arbitrary values of B'_1, B'_2, B'_3, ..., still satisfies the homogeneous conditions (363). Choosing, e.g., $(B'_1, B'_2, B'_3) = (15,\ 3,\ 2)$ and $B'_n = 0$ for $n \ge 4$, we obtain the plot:

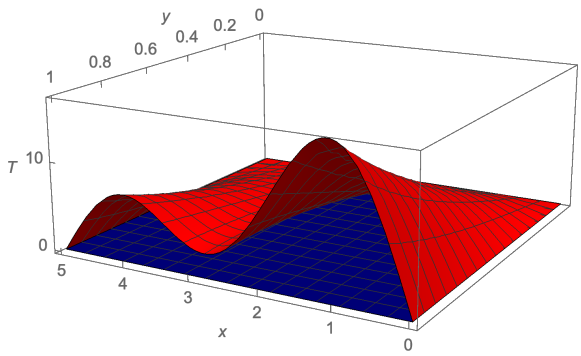


Figure 12: $T(x,y) = \sum_{n=1}^{3} B'_n \sin(k_n x)\sinh(k_n y)$ for $(B'_1, B'_2, B'_3) = (15,\ 3,\ 2)$.

Fig. 12 is indicating that, by a proper choice of the B'_n, we may succeed to satisfy also the inhomogeneous BC $T(x,b) = t(x)$, for any given function $t(x)$,

$$T(x,b) \ \equiv\ \sum_{n=1}^{\infty} \underbrace{B'_n \sinh(k_n b)}_{\beta_n} \sin(k_n x) \ =\ t(x). \tag{371}$$

Writing $x = \frac{a}{\pi} z$ (so that $k_n x = nz$) and $t(\frac{a}{\pi} z) = f(z)$, this reads

$$\underbrace{\sum_{n=1}^{\infty} \beta_n \sin(nz)}_{f_S(z)} \ =\ f(z) \qquad (0 < z < \pi). \tag{372}$$

The LHS is the **Fourier series** $f_{\rm S}(z)$ of a function $f(z)$ for $z \in [0,\pi]$.
In the case $f(z) = T_0 = \text{const}$, the coefficients β_n are given by Eq. (386) of section 10.2.4,

$$(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, ...) \ =\ \frac{4T_0}{\pi} \times \left(1, 0, \tfrac{1}{3}, 0, \tfrac{1}{5}, 0, \tfrac{1}{7}, ...\right), \tag{373}$$

and the exact solution of our original problem, Eqs. (362) with the BCs (363) and (364), is given by Eq. (370) with $B'_n = \frac{\beta_n}{\sinh(k_n b)}$,

$$T(x,y) \ =\ \frac{4T_0}{\pi} \sum_{n=1,3,5,...} \frac{1}{n\ \sinh(k_n b)} \sin(k_n x)\sinh(k_n y) \qquad \left(k_n \ =\ n\,\frac{\pi}{a}\right). \tag{374}$$

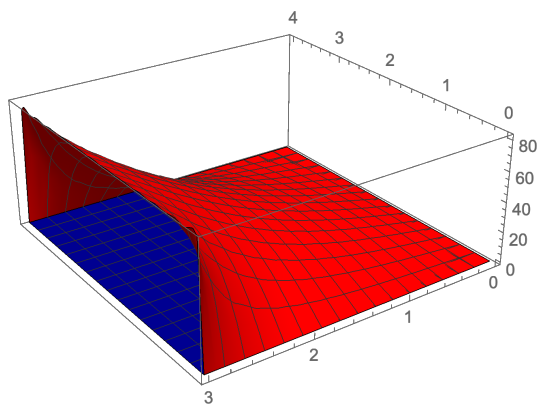


Figure 13: $T(x,y)$ of Eq. (374), truncated at $n = 199$, using $(T_0, a, b) = (80, 4, 3)$.

10.2.3 More general boundary conditions

Now, each one of the four edges of the rectangle shall have a different temperature,

$$T(a,y) \;=\; T_1, \qquad T(x,b) \;=\; T_2, \qquad T(0,y) \;=\; T_3, \qquad T(x,0) \;=\; T_4. \tag{375}$$

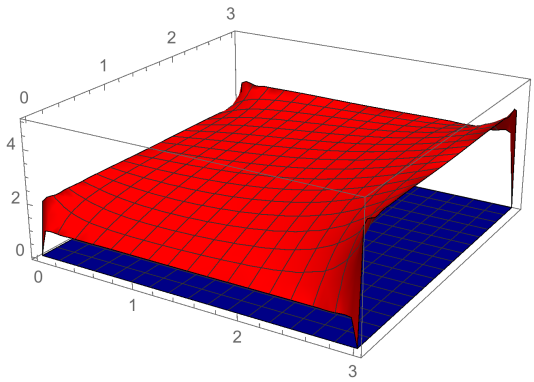


Figure 14: $T(x,y)$ for $a=b=3$ and $(T_1,T_2,T_3,T_4)=(1,2,3,4)$.
In this case, $T(1,1)=3.02287$, $T(1,2)=T(2,1)=\frac{5}{2}$ and $T(2,2)=1.97713$.

Then, in terms of the functions

$$t_U(x,y) \;=\; \frac{4}{\pi}\sum_{n=1,3,5,\ldots}\frac{1}{n\,\sinh(n\pi\frac{b}{a})}\,\sin\left(n\pi\frac{x}{a}\right)\,\sinh\left(n\pi\frac{y}{a}\right), \tag{376}$$

$$t_R(x,y) \;=\; \frac{4}{\pi}\sum_{n=1,3,5,\ldots}\frac{1}{n\,\sinh(n\pi\frac{a}{b})}\,\sinh\left(n\pi\frac{x}{b}\right)\,\sin\left(n\pi\frac{y}{b}\right), \tag{377}$$

the solution is obviously given by the linear combination

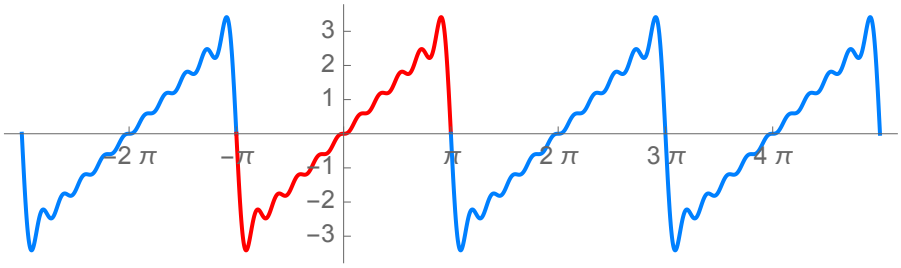
$$T(x,y) \;=\; T_1\cdot t_R(x,y) \;+\; T_2\cdot t_U(x,y) \;+\; T_3\cdot t_R(a-x,y) \;+\; T_4\cdot t_U(x,b-y). \tag{378}$$

10.2.4 Fourier analysis

Consider (for $z \in \mathbb{R}$) the function

$$\begin{aligned} f_{\rm S}^{(N)}(z) &= 2\left[\sin z \;-\; \frac{1}{2}\sin 2z \;+\; \frac{1}{3}\sin 3z \;-+\ldots\;+\;(-1)^{N+1}\frac{1}{N}\sin(Nz)\right] \\ &\equiv 2\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}\sin(nz), \end{aligned} \tag{379}$$

plotted here versus $z \in \mathbb{R}$ for the case $N = 10$:



Clearly, this function
- has the **periodicity** interval $[-\pi, \pi]$.
- is an **odd function**, $f_{\rm S}^{(N)}(-z) = -f_{\rm S}^{(N)}(z)$.
- for $z \in [-\pi, \pi]$ seems to approach the function $f(z) = z$, as $N \to \infty$.

Theorem: For $z \in [-\pi, \pi]$, let $f(z)$ be **any given continuous** function.
Then, there are numbers $\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots$ (called the Fourier coefficients of f) with the property that for each $z \in \,]-\pi, \pi[\,$, the value $f(z)$ is given by the limit

$$f(z) \;=\; \frac{\alpha_0}{2} \;+\; \lim_{N\to\infty}\sum_{n=1}^{N}\Big[\alpha_n\cos(nz) \;+\; \beta_n\sin(nz)\Big] \qquad (-\pi < z < \pi). \tag{380}$$

The right-hand side of Eq. (380) is called a **Fourier series**, in short-hand notation

$$f_{\rm S}(z) \;=\; \frac{\alpha_0}{2} \;+\; \sum_{n=1}^{\infty}\Big[\alpha_n\cos(nz) \;+\; \beta_n\sin(nz)\Big]. \tag{381}$$

Example 1: In the case $f(z) = z$, the coefficients are

$$\begin{aligned} \alpha_n &= 0 \qquad\qquad (n = 0, 1, 2, 3, \ldots), \\ \beta_n &= (-1)^{n+1}\,\frac{2}{n} \qquad (n = 1, 2, 3, \ldots). \end{aligned} \tag{382}$$

Remark 1: To prove this theorem, the set $\{1, \cos z, \sin z, \cos 2z, \sin 2z, \ldots\}$ must be shown to be a **"complete set" of functions** on the interval $-\pi < z < \pi$.

Corollary: Given $f(z)$, the Fourier coefficients are easily obtained as integrals,

$$\begin{aligned} \alpha_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} {\rm d}z\, f(z)\,\cos(nz) \qquad (n = 0, 1, 2, 3, \ldots), \\ \beta_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} {\rm d}z\, f(z)\,\sin(nz) \qquad (n = 1, 2, 3, \ldots). \end{aligned} \tag{383}$$

Proof: Due to the theorem, we have $f(z) = f_S(z)$ for $-\pi < z < \pi$. Consequently, e.g,

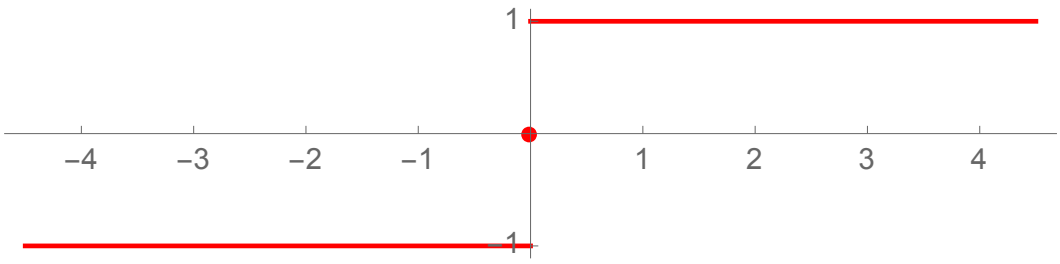
$$\begin{aligned}
\int_{-\pi}^{\pi} \mathrm{d}z\, f(z)\, \cos(mz) &= \int_{-\pi}^{\pi} \mathrm{d}z\, f_S(z)\, \cos(mz) \\
&= \int_{-\pi}^{\pi} \mathrm{d}z \left\{ \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} \Big[\alpha_n \cos(nz) + \beta_n \sin(nz) \Big] \right\} \cos(mz) \\
= \frac{\alpha_0}{2} \underbrace{\int_{-\pi}^{\pi} \mathrm{d}z\, \cos(mz)}_{0} &+ \sum_{n=1}^{\infty} \alpha_n \underbrace{\int_{-\pi}^{\pi} \mathrm{d}z\, \cos(nz) \cos(mz)}_{\pi\, \delta_{n,m}} + \sum_{n=1}^{\infty} \beta_n \underbrace{\int_{-\pi}^{\pi} \mathrm{d}z\, \sin(nz) \cos(mz)}_{0} \\
&= \alpha_m \pi.
\end{aligned}$$

Remark 2: The theorem can be generalized to **discontinuous functions** with a finite number ℓ of discontinuities $z_1, ..., z_\ell \in [-\pi, \pi]$:

Example 2: The **sign function** $f(z) = \mathrm{sgn}(z)$,

$$\mathrm{sgn}(z) = \begin{cases} -1 & (z < 0), \\ 0 & (z = 0), \\ 1 & (z > 0), \end{cases} \tag{384}$$

has one single ($\ell = 1$) discontinuity $z_1 = 0$ within $[-\pi, \pi]$.



It has the Fourier series

$$\begin{aligned}
f_S(z) &= \frac{4}{\pi} \cdot \Big\{ \sin z + \tfrac{1}{3} \sin 3z + \tfrac{1}{5} \sin 5z + ... \Big\} \\
&= \frac{4}{\pi} \cdot \sum_{n=1,3,5,...}^{\infty} \frac{1}{n} \sin(nz).
\end{aligned} \tag{385}$$

In this case, $\alpha_n = 0$ (for all $n = 0, 1, 2, 3, ...$) and

$$\beta_n = \begin{cases} \frac{4}{\pi} \cdot \frac{1}{n} & (n = 1, 3, 5, ...), \\ 0 & (n = 2, 4, 6, ...). \end{cases} \tag{386}$$

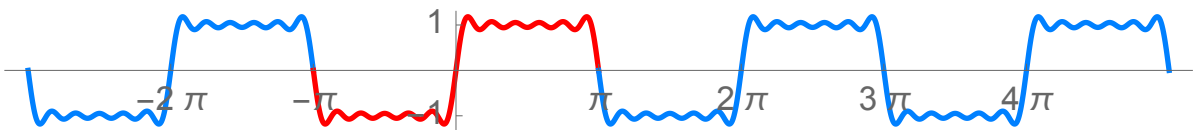


Figure 15: The truncated Fourier series $f_S(z) = \frac{4}{\pi} \cdot \sum_{n=1,3,5,..}^{N} \frac{1}{n} \sin(nz)$ with $N = 11$.

10.3 Wave equation in 1D

The transversal displacement $y = u(x,t)$ (in y-direction) of a violin string (which at rest covers the x-axis) satisfies the 1D wave equation (345),

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}u(x,t) \;=\; \frac{\partial^2}{\partial x^2}u(x,t), \qquad c^2 = \frac{F}{\mu}. \tag{387}$$

F (in N) is the string tension and μ (in $\frac{\mathrm{g}}{\mathrm{m}}$) is the mass density of the string.

10.3.1 D'Alembert solution

The general solution of Eq. (387) is given by

$$u(x,t) \;=\; f(x-ct) + g(x+ct), \tag{388}$$

where $f(x)$ and $g(x)$ are arbitrary functions, with derivatives $f'(x)$, $f''(x)$ and $g'(x)$, $g''(x)$, respectively. To see this, consider

$$\begin{aligned} \frac{\partial u(x,t)}{\partial t} &= (-c)\,f'(x-ct) \;+\; c\,g'(x+ct), \\ \Rightarrow \quad \frac{\partial^2 u(x,t)}{\partial t^2} &= (-c)^2 f''(x-ct) \;+\; c^2 g''(x+ct) \;\equiv\; c^2\,\frac{\partial^2 u(x,t)}{\partial x^2}. \end{aligned} \tag{389}$$

Plotted as a function of x for different (fixed) values of t, the term $f(x-ct)$ describes a **rigid profile**, moving at constant **velocity** c in positive x-direction along the x-axis. Similarly, $g(x+ct)$ describes another profile, moving in negative x-direction, cf. Fig. 16. Generally, such a motion of a rigid profile is called a **wave**.

In particular, the constant $c = \sqrt{F/\mu}$ is the **speed of waves** on the string. E.g.,

$$c \;=\; \sqrt{\frac{100\,\mathrm{N}}{0.010\,\frac{\mathrm{kg}}{\mathrm{m}}}} \;=\; \sqrt{10\;000\,\frac{\mathrm{m}^2}{\mathrm{s}^2}} \;=\; 100\,\frac{\mathrm{m}}{\mathrm{s}}. \tag{390}$$

10.3.2 Initial-value conditions

A solution $u(x,t)$ of Eq. (387) is fixed uniquely by **initial value conditions** (ICs),

$$u(x,0) \;=\; y(x), \qquad\qquad u_t(x,0) \;=\; v(x), \tag{391}$$

where $u_t(x,t) \equiv \frac{\partial u(x,t)}{\partial t}$. Here, $y(x)$ and $v(x)$ are **two given functions**.

Using Eq. (388) for $u(x,t)$, these ICs read

$$f(x) + g(x) \;=\; y(x), \qquad c\big[g'(x) - f'(x)\big] \;=\; v(x). \tag{392}$$

Defining $V(x) = \int_0^x v(x')\mathrm{d}x' + C$, we obtain $f(x)$ and $g(x)$ in terms of $y(x)$ and $v(x)$,

$$f(x) \;=\; \frac{1}{2}\left[y(x) - \frac{V(x)}{c}\right] + \frac{C}{2c}, \qquad g(x) \;=\; \frac{1}{2}\left[y(x) + \frac{V(x)}{c}\right] - \frac{C}{2c}. \tag{393}$$

The resulting solution (388) does not depend on C,

$$\begin{aligned} u(x,t) &= \frac{1}{2}\Big[y(x-ct)+y(x+ct)\Big] + \frac{1}{2c}\Big[V(x+ct)-V(x-ct)\Big] \\ &\equiv \frac{1}{2}\Big[y(x-ct)+y(x+ct)\Big] + \frac{1}{2c}\int_{x-ct}^{x+ct} \mathrm{d}x'\, v(x'). \end{aligned} \tag{394}$$

An example with $v(x)=0$: $u(x,t)=\frac{1}{2}[y(x-ct)+y(x+ct)]$, is illustrated in Fig. 16.

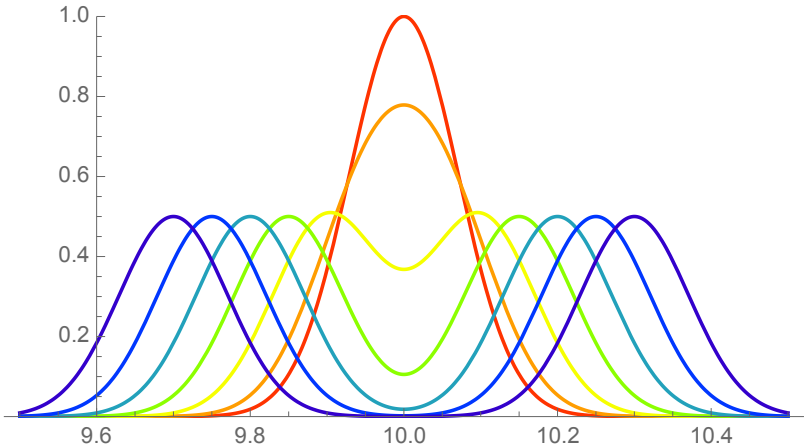


Figure 16: The solution $u(x,t)$ to the ICs $y(x) = \mathrm{e}^{-(x-10)^2/0.1^2}$ and $v(x) = 0$ for an infinitely long string (choosing the value $c = 1$), plotted versus x for a few **small** times: $t = 0$ (red), $t = 0.05$ (orange), $t = 0.1$ (yellow), $t = 0.15$ (green), ..., $t = 0.3$ (dark-blue).

10.3.3 Boundary conditions

A violin string, fixed at $x = 0$ and at $x = \ell$, satisfies the **boundary conditions** (BCs)

$$u(0,t) \;=\; 0, \qquad u(\ell,t) \;=\; 0 \qquad \text{(for all } t\text{)}. \tag{395}$$

More generally, the BCs can have the form

$$u(0,t) \;=\; a(t), \qquad u(\ell,t) \;=\; b(t) \qquad \text{(for all } t\text{)}, \tag{396}$$

with two given functions $a(t)$ and $b(t)$, imposing on the ICs the constraints

$$y(0) = a(0), \qquad y(\ell) = b(0), \qquad v(0) = \dot{a}(0), \qquad v(\ell) = \dot{b}(0). \tag{397}$$

In the presence of BCs, the application of d'Alembert's solution is **not quite clear**, since then the functions $y(x)$ and $v(x)$ in Eq. (394) are known for $0 \le x \le \ell$ only, while the argument $|x \pm ct|$ can grow beyond any limit as t increases. E.g., in the case $\ell = 20$, the behavior of the wave in Fig. 16 becomes wrong as soon as $t \approx 9.7$ when both partial waves will reach the boundaries at $x = 0$ and at $x = 20$, respectively.

To find the proper continuation of $y(x)$ onto all of $\mathbb{R}$ [we choose $v(x) = 0$ for simplicity], we shall now solve the wave equation directly.

10.3.4 Factorizing solutions

As in section 10.2.1, we first attempt to satisfy all **homogeneous** conditions [which are: the BCs (395) plus the second IC of Eq. (391), here $v(x) = 0$] by a separation ansatz

$$u(x,t) \;=\; X(x)T(t). \tag{398}$$

In this case, the homogeneous conditions read

$$X(0) \;=\; 0, \qquad X(\ell) \;=\; 0, \qquad \dot{T}(0) \;=\; 0. \tag{399}$$

Similar steps as in section 10.2.1 (see problem 9.7) yield

$$X(x)T(t) \;=\; U\sin(k_n x)\cos(k_n\, ct), \qquad k_n \;=\; n\,\frac{\pi}{\ell} \quad (n = 1, 2, 3, ...). \tag{400}$$

All homogeneous conditions are still satisfied by any arbitrary linear combination

$$u(x,t) \;=\; \sum_{n=1}^{\infty} U_n \sin(k_n x)\cos(k_n\, ct). \tag{401}$$

The addition theorem $\sin\phi\cos\chi = \frac{1}{2}[\sin(\phi-\chi) + \sin(\phi+\chi)]$ yields

$$\begin{aligned} u(x,t) \;&=\; \sum_{n=1}^{\infty} \frac{U_n}{2}\Big[\sin\big(k_n(x-ct)\big) \;+\; \sin\big(k_n(x+ct)\big)\Big] \\ &\equiv\; \frac{1}{2}\Big[F(x-ct) \;+\; F(x+ct)\Big], \end{aligned} \tag{402}$$

where $F(x)$ is an odd function with period $P = 2\ell$, given for all $x \in \mathbb{R}$ by the series

$$F(x) \;=\; \sum_{n=1}^{\infty} U_n \sin(k_n x) \qquad (x \in \mathbb{R}). \tag{403}$$

Now, the first (**inhomogeneous**) IC of Eq. (391) implies that $F(x) \equiv y(x)$ for $0 < x < \ell$. Consequently, $F(x)$ is the wanted continuation of $y(x)$,

$$F(x) \;=\; \left\{ \begin{array}{cc} -y(|x|) & (-\ell \le x \le 0) \\ y(x) & (0 \le x \le \ell) \end{array} \right\}, \qquad F(x \pm 2\ell) \;=\; F(x) \quad (x \in \mathbb{R}). \tag{404}$$

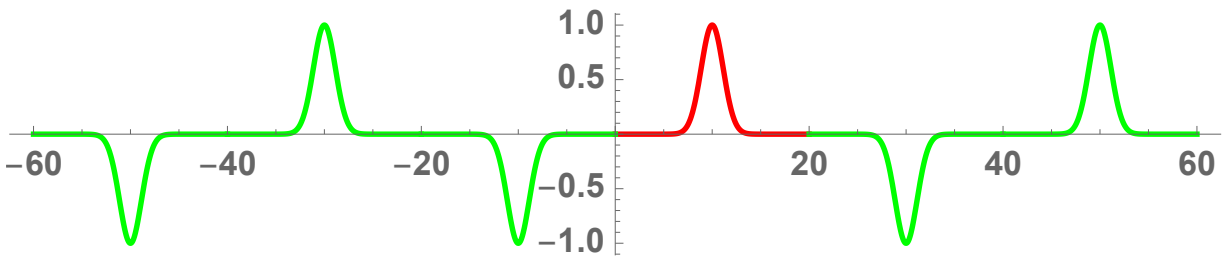


Figure 17: The periodic continuation $F(x)$ to $x \in \mathbb{R}$ (green) of the IC $y(x) = b\mathrm{e}^{-(x-10)^2/a^2}$ for $0 < x < 20$ (red). The parameters are $a = 1.6$ and $b = 1.0$.

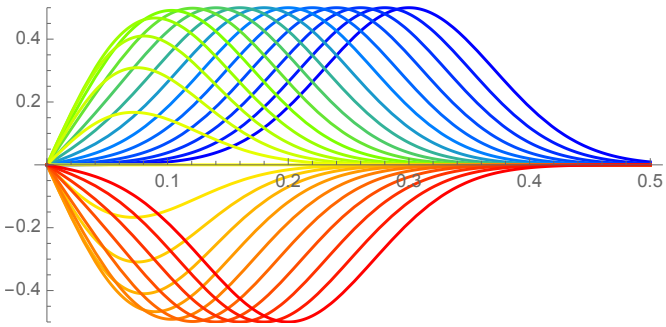


Figure 18: The solution $u(x,t)$ to the ICs $y(x) = \mathrm{e}^{-(x-10)^2/0.1^2}$ and $v(x) = 0$ for the string fixed at $x = 0$ and at $x = 20$ (choosing the value $c = 1$), plotted versus x for the times: $t = 9.7,\ 9.72,\ 9.74$ (blue), ..., $t = 10.16,\ 10.18,\ 10.2$ (red).

10.3.5 Fourier analysis

Fixed by Eq. (403), the coefficients U_n are not needed for constructing the function $F(x)$. However, they yield the **spectrum** of the sound produced by the vibrating string.

For a Fourier analysis of $F(x)$, we consider one period, $-\ell < x < \ell$. Due to Eqs. (383), with $f(z) = F(\frac{\ell}{\pi}z)$, $x_0 = 0$, $a = \ell$, $-\pi < z < \pi$, the coefficients are

$$\begin{aligned} U_n \;\equiv\; \beta_n \;&=\; \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}z\, F\Big(\frac{\ell}{\pi}z\Big)\,\sin(nz) \\ &=\; \frac{2}{\pi}\int_{0}^{\pi} \mathrm{d}z\, y\Big(\frac{\ell}{\pi}z\Big)\,\sin(nz) \;\equiv\; \frac{2}{\ell}\int_{0}^{\ell} \mathrm{d}x\, y(x)\,\sin(k_n x). \end{aligned} \tag{405}$$

and $\alpha_n = 0$ for all n.

As an example, with the BCs (395), we consider a string that was picked at the time $t = 0$,

$$y(x) \;=\; \left\{ \begin{array}{ll} \frac{b}{a}\cdot x & (0 \le x \le a), \\ b\cdot\frac{\ell - x}{\ell - a} & (a \le x \le \ell). \end{array} \right\}, \qquad v(x) \;=\; 0. \tag{406}$$

In this case, we have (writing $k_n = k$),

$$\begin{aligned} U_n \;&=\; \frac{2}{\ell}\left\{\frac{b}{a}\int_0^a \mathrm{d}x\,\sin(kx)\,x \;+\; \frac{b}{\ell - a}\int_a^\ell \mathrm{d}x\,\sin(kx)\,(\ell - x)\right\} \\ &=\; \frac{2}{\ell}\left\{\frac{b}{a}\Big[\frac{\sin(kx)}{k^2} - \frac{x\cos(kx)}{k}\Big]_0^a \;+ \right. \\ &\qquad \left. + \frac{b\ell}{\ell - a}\Big[-\frac{\cos(kx)}{k}\Big]_a^\ell \;-\; \frac{b}{\ell - a}\Big[\frac{\sin(kx)}{k^2} - \frac{x\cos(kx)}{k}\Big]_a^\ell\right\} \\ &=\; \frac{2b}{a(\ell - a)}\,\frac{\sin(ka)}{k^2} \\ &=\; \frac{2b}{a(\ell - a)}\,\frac{\ell^2}{\pi^2}\,\frac{\sin\big(n\frac{\pi}{\ell}a\big)}{n^2}. \end{aligned} \tag{407}$$

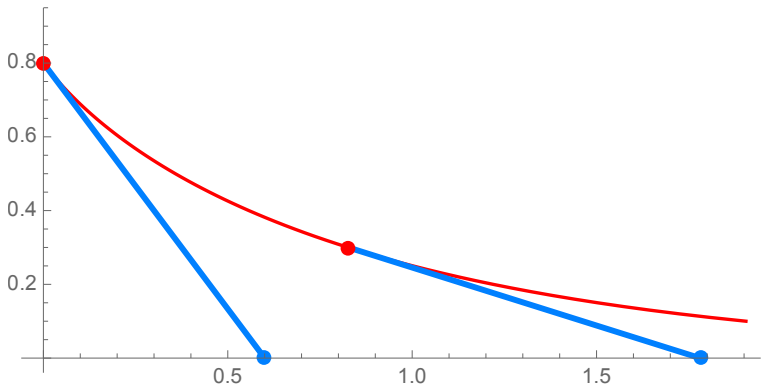
PART III

Numerical Solution of Differential Equations

11 Ordinary equations (ODE's)

11.1 The trailing curve as an example

On a horizontal table, one end A of a stretched thread (length a, **blue**) is being pulled along the x-axis. What is the curve $y = y(x) > 0$ drawn by a piece of **red chalk** dragged at the other end B of the thread, when B was initially at $(x|y) = (0|b)$ (with $b < a$)?



For any point $B(x_\mathrm{B}|y(x_\mathrm{B}))$ on that curve, the pulling end A of the thread on the x-axis has an x-coordinate $x_\mathrm{A} > x_\mathrm{B}$, fixed by Pythagoras' theorem,

$$y(x_\mathrm{B})^2 + (x_\mathrm{A} - x_\mathrm{B})^2 = a^2 \qquad \Rightarrow \qquad x_\mathrm{A} - x_\mathrm{B} = \sqrt{a^2 - y(x_\mathrm{B})^2}\,. \tag{408}$$

In particular, since the pulling thread lies **tangential** to the curve in B, it has the slope

$$y'(x_\mathrm{B}) = -\frac{y(x_\mathrm{B})}{x_\mathrm{A} - x_\mathrm{B}}, \tag{409}$$

where $y'(x)$ is the derivative of the wanted function $y(x)$. Using Eq. (408) for $x_\mathrm{A} - x_\mathrm{B}$ and replacing x_B with x, we obtain a **non-linear ODE** for $y(x)$,

$$y'(x) = -\frac{y(x)}{\sqrt{a^2 - y(x)^2}}\,. \tag{410}$$

Being non-linear, this ODE cannot be solved by the strategy discussed in section 8.2. Before we shall solve it exactly in section 11.4 (using a different strategy), we here want to use it for demonstrating the performance of **approximate numerical** techniques.

11.2 Finite difference methods (FDMs)

11.2.1 Euler's method

Instead of treating x as a **continuous** variable, we choose a finite **step size** $h > 0$ and introduce $N + 1$ equidistant **discrete mesh points**

$$x_n = nh \qquad (n = 0, 1, 2, ..., N). \tag{411}$$

The definition of the derivative, $y'(x) \equiv \lim_{h\to 0} \frac{y(x+h)-y(x)}{h}$, for **finite** but sufficiently small $h > 0$ yields an approximation for a **finite difference**,

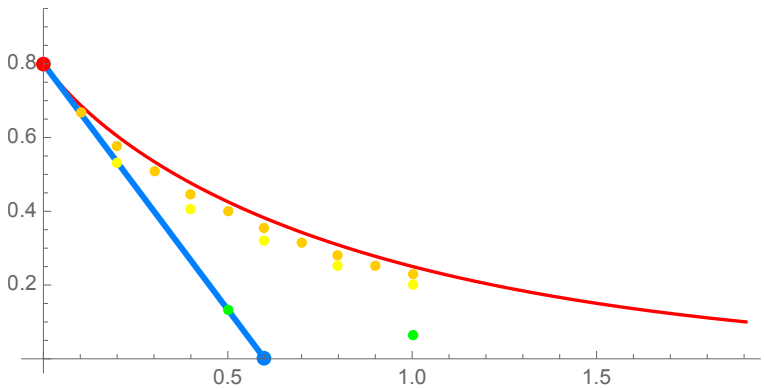
$$y(\underbrace{x_n + h}_{x_{n+1}}) \;-\; y(x_n) \;\approx\; h \cdot y'(x_n). \tag{412}$$

Using for $y'(x_n)$ the ODE Eq. (410), we obtain a **recursion formula**,

$$y(x_{n+1}) \;\approx\; y(x_n) \;-\; h \cdot \frac{y(x_n)}{\sqrt{a^2 - y(x_n)^2}}. \tag{413}$$

Starting with the initial value $y(x_0) = b$, it successively yields approximate values for $y_h(x_1)$, $y_h(x_2)$, etc. This FDM is known as **Euler's method**.

x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$y_{h=0.5}$	0.8					0.133					0.066
$y_{h=0.2}$	0.8		0.533		0.407		0.318		0.251		0.199
$y_{h=0.1}$	0.8	0.667	0.577	0.507	0.448	0.398	0.354	0.316	0.283	0.254	0.227
y_{exact}	0.8	0.689	0.604	0.535	0.476	0.426	0.382	0.343	0.309	0.278	0.251



Here is a **MATLAB code**, using the values $a = 1$, $b = 0.8$, $h = 0.01$ and $N = 100$:

```
a=1;
b=0.8;
dx=0.01;
N=100;
X=zeros(1,N+1);          % zero matrix 'X' with 1 row and N+1 columns
Y=zeros(1,N+1);          % zero matrix 'Y' with 1 row and N+1 columns
x=0;
y=b;
Y(1)=y;
for (n=2:N+1)
    yp=-y/sqrt(a.^2-y.^2);
    y=y+yp*dx;
    x=x+dx;
    [x,y]                % print the pair (x,y) on the screen
    X(n)=x;
    Y(n)=y;
end
plot(X,Y)
axis equal
grid on
```

11.2.2 Runge-Kutta methods

11.3 The MATLAB tool "ode23"

11.3.1 First-order ODEs

"ode23" **numerically** solves **first-order ODEs** (or, more generally, systems of several first-order ODEs). Generally, such a (typically **non-linear**) ODE has the form

$$y'(x) \;=\; f\big(x, y(x)\big), \tag{414}$$

with a given function $f(x, y)$ of two variables.

• **Example 1:** A **fantasy ODE** and the corresponding function f are

$$y'(x) \;=\; -2x\, y(x)^2, \qquad\qquad f(x, y) \;=\; -2xy^2. \tag{415}$$

In passing, we note that the **exact** solution is (check!)

$$y(x) \;=\; \frac{1}{a + x^2} \qquad (a \in \mathbb{R}). \tag{416}$$

To apply "ode23" to such an ODE, we first define the function $f(x, y)$ in a separate m-file. For the above example, this file (introducing the name "fantasy" for f) could read:

```
% fantasy.m
% Function for the fantasy ODE
function out=fantasy(x,y)
out=-2*x*y.^2
```

• **Example 2:** For the ODE (410) of the **trailing curve**, $f(x, y)$ does not depend on x,

$$y'(x) \;=\; -\frac{y(x)}{\sqrt{a^2 - y(x)^2}}\,, \qquad\qquad f(x, y) \;=\; -\frac{y}{\sqrt{a^2 - y^2}}. \tag{417}$$

In this case, the separate m-file (with the name "trailing" for f) could read:

```
% trailing.m
% Function for the trailing-curve ODE
function out=trailing(x,y)
a=1.;
out=-y/sqrt(a.^2-y.^2)
```

After having defined the ODE this way, we can call "ode23" by the command:

```
[X Y]=ode23('trailing', [0 3], 0.8);
```

The first argument `'trailing'` calls the function $f(x, y)$ that defines the ODE, the second one `[0 3]` fixes the x-interval $x_0\,(= 0) \le x \le x_{\rm f}\,(= 3)$ for which the solution is wanted. The third argument `0.8` gives the initial value $y(x_0)$, here $y(0) = b(= 0.8)$.

The output `[X Y]` gives two vectors `X` and `Y`, containing, respectively, x- and y-values of the solution $y(x)$ found for the ODE. This result can be plotted by

```
plot(X,Y)
axis equal
grid on
```

11.3.2 Higher-order ODEs

For completeness, we apply "ode23" to a system of two first-order ODEs:
The exact equation of motion for a mathematical pendulum is a **second-order** ODE,

$$\ddot{\phi}(t) \;=\; -\frac{g}{\ell}\,\sin\phi(t). \tag{418}$$

Here, $g = 9.81\frac{\mathrm{m}}{\mathrm{s}^2}$ (gravitational acceleration on the surface of Earth) and ℓ is the length of the pendulum. $\phi(t)$ is, as a function of time t, the **angular displacement** from the pendulum's equilibrium position ($\phi = 0$).

Writing $\phi(t) = y_1(t)$ and $\dot{\phi}(t) = y_2(t)$, we obtain a **system of two first-order** ODEs,

$$\begin{aligned} \dot{y}_1(t) &= y_2(t), \\ \dot{y}_2(t) &= -\frac{g}{\ell}\,\sin\big(y_1(t)\big). \end{aligned} \tag{419}$$

Combining the two solutions to a **vector**, $\binom{y_1(t)}{y_2(t)} = \mathbf{y}(t)$, we formally have

$$\dot{\mathbf{y}}(t) \;=\; \mathbf{f}\big(t, \mathbf{y}(t)\big), \tag{420}$$

where the vector function $\mathbf{f}(t, \mathbf{y})$ here does not depend on t, but only on $\mathbf{y} = (y_1, y_2)$,

$$\mathbf{f}(t, \mathbf{y}) \;\equiv\; \begin{pmatrix} f_1(t, \mathbf{y}) \\ f_2(t, \mathbf{y}) \end{pmatrix} \;=\; \begin{pmatrix} y_2 \\ -\frac{g}{\ell}\,\sin y_1 \end{pmatrix}. \tag{421}$$

The small-amplitude ($\phi_{\max} \ll \pi$) oscillation period $T_{\mathrm{sa}} = 2\pi\sqrt{\ell/g}$ is exactly 1 s when we choose $\ell = 24.85\,\mathrm{cm}$ or $g/\ell = 39.48\,\mathrm{s}^{-2}$. Then, the **vector function** $\mathbf{f}(t, \mathbf{y})$ is defined by

```
% pend.m
function out=pend(t,y)
out=zeros(2,1);
out(1)=y(2);
out(2)=-39.48*sin(y(1));
```

Notice that the input `y` is a vector, with two components `y(1)` and `y(2)`. So is the output.

Now, calling "ode23" using the time interval $0 \le t \le 3$ s and the initial conditions $y_1(0) \equiv \phi(0) = 0.15 \ll \pi$ (**small amplitude**) and $y_2(0) \equiv \dot{\phi}(0) = 0$,

```
[T Y]=ode23('pend', [0 3], [0.15 0]);
```

we obtain an almost perfect sine curve with period ≈ 1 s and amplitude 0.15:

```
plot(X,Y(:,1))
grid on
```

In contrast, the initial conditions $y_1(0) \equiv \phi(0) = 3.13 \approx \pi$ (**almost maximum possible amplitude**) and $y_2(0) \equiv \dot{\phi}(0) = 0$,

```
[T Y]=ode23('pend', [0 3], [3.13 0]);
```

yield a curve totally different from a sine, and a period T considerably larger than 1 s.

Entertaining example: The Lorenz attractor.

11.4 Exact trailing curve

Our first-order ODE Eq. (410) is a **separable-variables** equation: It can be written as

$$\frac{\mathrm{d}y}{\mathrm{d}x} \;=\; -\frac{y}{\sqrt{a^2-y^2}} \qquad \Rightarrow \qquad -\frac{\sqrt{a^2-y^2}}{y}\,\mathrm{d}y \;=\; \mathrm{d}x. \tag{422}$$

Now, only y's appear on the LHS while the RHS has x's only !

A simpler examlple for this kind of ODE is $y'(x) = \frac{5y(x)}{x}$. In this case we write

$$\frac{\mathrm{d}y}{\mathrm{d}x} \;=\; \frac{5y(x)}{x} \qquad \Rightarrow \qquad \frac{1}{y}\,\mathrm{d}y \;=\; \frac{5}{x}\,\mathrm{d}x. \tag{423}$$

Formally integrating on the left side over y and on the right side over x yields

$$\ln y \;+\; C_1 \;=\; 5\ln x \;+\; C_2. \tag{424}$$

There is only one independent constant, $C_2 - C_1 = C$.
Eventually, resolving for $y = y(x)$ yields the general solution

$$y(x) \;=\; \mathrm{e}^{5\ln x + C} \;=\; a\,x^5, \qquad a = \mathrm{e}^C. \tag{425}$$

By checking, we confirm that this is in fact a solution to $y'(x) = \frac{5y(x)}{x}$, for all $a \in \mathbb{R}$.

Now, we apply the same integration technique on both sides of Eq. (422), to obtain[8]

$$a\ln\frac{a+\sqrt{a^2-y^2}}{y} \;-\; \sqrt{a^2-y^2} \;+\; C_1 \;=\; x \;+\; C_2. \tag{426}$$

We note in passing: $\ln(u+\sqrt{u^2-1}\,) \equiv \operatorname{arcosh}(u)$, where $u = \frac{a}{y} \geq 1$.
For plotting the (invertible!) function $y = y(x)$, it is sufficient to resolve for x,

$$x(y) \;=\; a\ln\frac{a+\sqrt{a^2-y^2}}{y} \;-\; \sqrt{a^2-y^2} \;-\; C \qquad (C = C_2 - C_1). \tag{427}$$

To satisfy the starting condition $y(0) = b$ or, equivalently, $x(b) = 0$, we choose

$$C \;=\; a\ln\frac{a+\sqrt{a^2-b^2}}{b} \;-\; \sqrt{a^2-b^2}. \tag{428}$$

[8] We check Eq. (426) by diffferentiating it:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}y}\left\{a\ln\frac{a+\sqrt{a^2-y^2}}{y} - \sqrt{a^2-y^2}\right\} \;&=\; a\,\frac{y}{a+\sqrt{\cdots}}\cdot\frac{y\cdot\frac{-2y}{2\sqrt{\cdots}} \;-\; 1\cdot(a+\sqrt{\cdots})}{y^2} - \frac{-2y}{2\sqrt{\cdots}} \\
&=\; a\,\frac{y(a-\sqrt{\cdots})}{y^2}\cdot\frac{-1}{\sqrt{\cdots}} \;-\; \frac{a}{y} \;+\; \frac{y}{\sqrt{\cdots}} \\
&=\; -\frac{a^2}{y\sqrt{\cdots}} \;+\; \frac{a}{y} \;-\; \frac{a}{y} \;+\; \frac{y^2}{y\sqrt{\cdots}} \;=\; -\frac{\sqrt{a^2-y^2}}{y}.
\end{aligned}$$

12 Partial differential equations (PDE's)

We here consider, as an example, the PDE

$$-p\,\nabla^2 u(\mathbf{r}) \;=\; s(\mathbf{r}) \tag{429}$$

for the **wanted function** $u(\mathbf{r})$ (a scalar field) in $n = 2$ or $n = 3$ dimensions, $\mathbf{r} \in \mathbb{R}^n$. $p > 0$ is a given number (**parameter**) and $s(\mathbf{r})$ is a given function (**source**).

Eq. (429) represents the time-independent (stationary) case of any one of the PDEs

$$\frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;=\; D\,\nabla^2\rho(\mathbf{r},t) \;+\; s(\mathbf{r},t) \qquad \text{(diffusion equation)}, \tag{430}$$

$$c\mu\,\frac{\partial}{\partial t}\,T(\mathbf{r},t) \;=\; \lambda\,\nabla^2 T(\mathbf{r},t) \;+\; s(\mathbf{r},t) \qquad \text{(heat conduction equation)}, \tag{431}$$

$$\nabla^2\,\Phi(\mathbf{r},t) \;=\; -\frac{1}{\epsilon_0}\rho(\mathbf{r},t) \qquad \text{(Poisson's equation)}, \tag{432}$$

when parameter p, source $s(\mathbf{r})$, and solution $u(\mathbf{r})$ are interpreted properly:

PDE	p	$u(\mathbf{r})$	$s(\mathbf{r})$
Diffusion	D	$\rho(\mathbf{r})$	$s(\mathbf{r})$
Heat conduction	λ	$T(\mathbf{r})$	$s(\mathbf{r})$
Poisson	ϵ_0	$\Phi(\mathbf{r})$	$\rho(\mathbf{r})$

In the **diffusion equation** ($p = D$), $u = \rho$ is the **mass density** of the diffusing substance, in the **heat conduction eqn.** ($p = \lambda$), $u = T$ is the **temperature distribution**, while in **Poisson's eqn.** ($p = \epsilon_0$), where the source is given by the **electric charge density**, $s = \rho$, the wanted function is the **electrostatic potential** $u = \Phi$.

Time-dependent problems (including waves) will be considered later.

12.1 Boundary conditions

Typically, a solution $u(\mathbf{r})$ of Eq. (429) is wanted for $\mathbf{r} \in \Omega$, where $\Omega \subset \mathbb{R}^n$ is a **given region** in 3D or 2D space. In addition, on the **surface** $\partial\Omega$ of Ω (in the case $n = 2$, $\partial\Omega$ is the **rim** of Ω), $u(\mathbf{r})$ shall satisfy the **boundary conditions**

$$\text{(D)} \qquad u(\mathbf{r}) \;=\; f(\mathbf{r}) \qquad (\text{for all } \mathbf{r} \in \partial\Omega_1), \tag{433}$$

$$\text{(N)} \qquad \boldsymbol{\mathcal{N}}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) \;=\; g(\mathbf{r}) \qquad (\text{for all } \mathbf{r} \in \partial\Omega_2). \tag{434}$$

Here, $\partial\Omega$ is divided into two pieces $\partial\Omega_1$ and $\partial\Omega_2$ with zero overlap, $\partial\Omega_1 \cap \partial\Omega_2 = \emptyset$.
$\boldsymbol{\mathcal{N}}(\mathbf{r})$ is the unit normal vector at $\mathbf{r} \in \partial\Omega$, pointing outward of Ω.
$f(\mathbf{r})$ and $g(\mathbf{r})$ are given functions for $\mathbf{r} \in \partial\Omega_1$ and $\mathbf{r} \in \partial\Omega_2$, respectively.

Interpretation: In the case of **Poisson's equation**, the **Dirichlet** BC (D) fixes the electrostatic potential $\Phi(\mathbf{r}) = f(\mathbf{r})$ on the part $\partial\Omega_1$ of the surface, while the **Neumann** BC (N) fixes the surface charge density $\sigma(\mathbf{r})$ on $\partial\Omega_2$.

Theorem: A solution $u(\mathbf{r})$ of Eq. (429) is fixed **uniquely** for all $\mathbf{r} \in \Omega$ by the conditions (433) and (434).

• For the proof, we need **Green's first identity**:
Consider two scalar fields $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$.
Applying Gauß's theorem to the vectorfield $\mathbf{F}(\mathbf{r}) = \phi(\mathbf{r})\nabla\psi(\mathbf{r})$ yields
Green's first identity,

$$\oint_{\partial\Omega} \mathrm{d}\mathbf{S} \cdot \big[\phi\nabla\psi\big] \;=\; \int_{\Omega} \mathrm{d}^3r \,\Big[\phi\nabla^2\psi \,+\, (\nabla\phi)\cdot(\nabla\psi)\Big]. \tag{435}$$

Interchanging here ϕ and ψ and subtracting the resulting equation yields
Green's second identity,

$$\oint_{\partial\Omega} \mathrm{d}\mathbf{S} \cdot \big[\phi\nabla\psi \,-\, \psi\nabla\phi\big] \;=\; \int_{\Omega} \mathrm{d}^3r \,\Big[\phi\nabla^2\psi \,-\, \psi\nabla^2\phi\Big]. \tag{436}$$

Proof: Let $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$ be two solutions satisfying (433) and (434).
Their difference $w(\mathbf{r})$,

$$w(\mathbf{r}) \;=\; u_1(\mathbf{r}) - u_2(\mathbf{r}), \tag{437}$$

satisfies (433) and (434) with $f = g = 0$. Consequently,

$$\begin{aligned}
\oint_{\partial\Omega} \mathrm{d}\mathbf{S} \cdot \Big[w(\mathbf{r})\,\nabla w(\mathbf{r})\Big] \;&\equiv\; \oint_{\partial\Omega} \mathrm{d}^{n-1}r\; w(\mathbf{r})\,\Big[\boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\nabla w(\mathbf{r})\Big] \\
&=\; \int_{\partial\Omega_1} \mathrm{d}^{n-1}r\;\; 0\;\Big[\boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\nabla w(\mathbf{r})\Big] \;+\; \int_{\partial\Omega_2} \mathrm{d}^{n-1}r\; w(\mathbf{r})\;\; 0 \\
&=\; 0.
\end{aligned} \tag{438}$$

Moreover, for $\mathbf{r} \in \Omega$, $w(\mathbf{r})$ satisfies Laplace's Equation,

$$\nabla^2 w(\mathbf{r}) \;=\; 0 \qquad (\text{for all } \mathbf{r} \in \Omega). \tag{439}$$

Consequently, writing Green's first identity with $\phi(\mathbf{r}) = \psi(\mathbf{r}) = w(\mathbf{r})$, we obtain

$$0 \;\equiv\; \oint_{\partial\Omega} \mathrm{d}\mathbf{S} \cdot \Big[w(\mathbf{r})\,\nabla w(\mathbf{r})\Big] \;=\; \int_{\Omega} \mathrm{d}^n r\,|\nabla w(\mathbf{r})|^2. \tag{440}$$

Since $w(\mathbf{r})$ is a continuous function, the RHS then implies that

$$|\nabla w(\mathbf{r})|^2 = 0 \qquad \Leftrightarrow \qquad u_1(\mathbf{r}) - u_2(\mathbf{r}) = C \qquad (\text{for all } \mathbf{r} \in \Omega), \tag{441}$$

with a constant C. Since $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$ both satisfy the first BC (433), we have $C = 0$, provided that $\partial\Omega_1 \neq \emptyset$.

12.2 Finite difference methods (FDM)

We here consider the PDE, Eq. (429), in **two dimensions** ($n = 2$) and set $p = 1$,

$$\nabla^2 u(\mathbf{r}) \;\equiv\; \frac{\partial^2 u(x,y)}{\partial x^2} \;+\; \frac{\partial^2 u(x,y)}{\partial y^2} \;=\; -s(x,y). \tag{442}$$

Given some region Ω in the xy-plane and a function $f(\mathbf{r})$ defined on the rim $\partial\Omega$ of Ω, we want the solution $u(x,y)$ that is fixed by the Dirichlet BC

$$u(\mathbf{r}) \;=\; f(\mathbf{r}), \qquad \mathbf{r} \in \partial\Omega. \tag{443}$$

12.2.1 Transformation into a FD equation

Using a small **step size** $h > 0$, we introduce discrete **mesh points** in the xy-plane,

$$\mathbf{r}_{nm} \;\equiv\; (x_n, y_m) \;=\; (nh,\; mh) \qquad (n, m \in \mathbb{Z}). \tag{444}$$

Writing $\frac{\partial u(x,y)}{\partial x} = u_x(x,y)$, the definition of derivatives implies

$$\begin{aligned}
\frac{\partial^2 u(x,y)}{\partial x^2} &= \lim_{h\to 0} \frac{u_x(x,y) - u_x(x-h,y)}{h} \\
&= \lim_{h\to 0} \frac{\frac{u(x+h,y)-u(x,y)}{h} - \frac{u(x,y)-u(x-h,y)}{h}}{h} \\
&= \lim_{h\to 0} \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2}.
\end{aligned} \tag{445}$$

Writing $\frac{\partial^2 u(x,y)}{\partial x^2} = u_{xx}(x,y)$, we obtain for sufficiently small $h > 0$ the approximation

$$\begin{aligned}
u_{xx}(x_n, y_m) &\approx \frac{u(x_{n+1}, y_m) - 2u(x_n, y_m) + u(x_{n-1}, y_m)}{h^2} \\
&\equiv \frac{u_{n+1,m} - 2u_{n,m} + u_{n-1,m}}{h^2},
\end{aligned} \tag{446}$$

where we have introduced the notation $u(x_n, y_m) = u_{n,m}$ in the second step. With the corresponding approximation for $u_{yy}(x_n, y_m)$, the **FD version of the PDE** reads

$$\frac{u_{n+1,m} - 2u_{n,m} + u_{n-1,m}}{h^2} \;+\; \frac{u_{n,m+1} - 2u_{n,m} + u_{n,m-1}}{h^2} \;=\; -s_{n,m}. \tag{447}$$

12.2.2 Iteration method

Resolving Eq. (447) for $u_{n,m}$, we obtain a set of **iteration equations**,

$$u_{n,m} \;=\; \frac{1}{4}\Big[u_{n+1,m} + u_{n-1,m} + u_{n,m+1} + u_{n,m-1}\Big] \;+\; \frac{h^2}{4}\, s_{n,m}\;. \tag{448}$$

Using on the RHS approximate **trial values** $u_{k,\ell} \equiv u(x_k, y_\ell)$ that cover the region Ω, the LHS yields new values. Provided that these "new" values are a better approximation to the wanted solution, we may continue, using them as the new "old" ones, etc.

Remark: In the case $s = 0$ (as in the following example), each "new" value u_{nm} is simply the **average** of the **four neighboring** "old" values!

Example 1: Revisiting the **heat-flow example** with $s(x,y) = 0$ (Laplace's equation) of section 10.2, Eq. (362) there, we here consider the special case $b = a$ of the square

$$\Omega \;=\; \Big\{(x,y)\;\Big|\;0 \le x,y \le a\Big\}. \tag{449}$$

Writing $T(x,y) = u(x,y)$, the BCs (363) and (364) now read

$$u(0,y) \;=\; u(x,0) \;=\; u(a,y) \;=\; 0, \qquad\qquad u(x,a) \;=\; T_0. \tag{450}$$

Choosing the step size $h = \frac{a}{3}$, we have as few as $N = 4$ mesh points $\mathbf{r}_{n,m} = (x_n, y_m)$ within Ω (and twelve mesh points on the rim $\partial\Omega$), with the coordinates (see the figure)

$$\begin{aligned} \mathbf{r}_1 \;\equiv\; \mathbf{r}_{12} \;&=\; (\tfrac{a}{3}, \tfrac{2a}{3}), \qquad & \mathbf{r}_2 \;\equiv\; \mathbf{r}_{22} \;&=\; (\tfrac{2a}{3}, \tfrac{2a}{3}), \\ \mathbf{r}_3 \;\equiv\; \mathbf{r}_{11} \;&=\; (\tfrac{a}{3}, \tfrac{a}{3}), \qquad & \mathbf{r}_4 \;\equiv\; \mathbf{r}_{21} \;&=\; (\tfrac{2a}{3}, \tfrac{a}{3}). \end{aligned} \tag{451}$$

Here, we have re-labeled the points $\mathbf{r}_{nm}$ within Ω as $\mathbf{r}_\nu$, with a **single index** $\nu \in \{1, ..., N\}$.

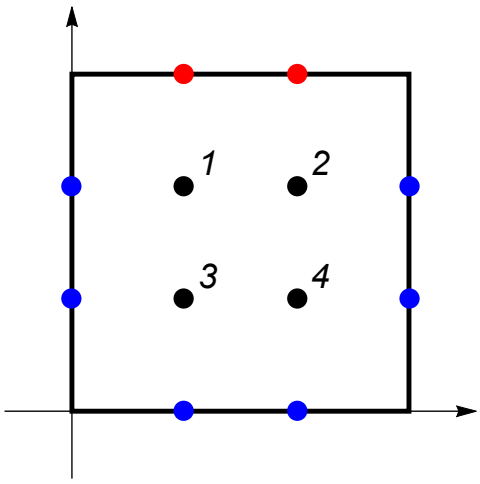


Figure 19: The region Ω, Eq. (449), in the xy-plane, with four mesh points inside (black). On the boundary $\partial\Omega$, the blue/red dots correspond to the first/second BC of Eq. (450).

Writing $u(\mathbf{r}_\nu) = u_\nu$, the four relevant **iteration equations** (448) read

$$\begin{aligned} u_1 \;&=\; \tfrac{1}{4}(u_2 \,+\, u_3 \,+\, T_0), \\ u_2 \;&=\; \tfrac{1}{4}(u_1 \,+\, u_4 \,+\, T_0), \\ u_3 \;&=\; \tfrac{1}{4}(u_1 \,+\, u_4), \\ u_4 \;&=\; \tfrac{1}{4}(u_2 \,+\, u_3). \end{aligned} \tag{452}$$

We write them in matrix notation, $u = Au + c$, with

$$A \;=\; \frac{1}{4}\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \qquad c \;=\; \frac{T_0}{4}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{453}$$

Choosing $T_0 = 80$ and starting with the **trial values** $u^{(0)} = (40, 40, 20, 20)^T$, the first two iterations $u^{(k+1)} = Au^{(k)} + c$ (with $k = 0$ and $k = 1$) successively yield

$$u^{(1)} = \begin{pmatrix} 15 \\ 15 \\ 15 \\ 15 \end{pmatrix} + \begin{pmatrix} 20 \\ 20 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 35 \\ 35 \\ 15 \\ 15 \end{pmatrix}, \quad u^{(2)} = \begin{pmatrix} 12.5 \\ 12.5 \\ 12.5 \\ 12.5 \end{pmatrix} + \begin{pmatrix} 20 \\ 20 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 32.5 \\ 32.5 \\ 12.5 \\ 12.5 \end{pmatrix}. \tag{454}$$

In this **simple example**, we can guess that this iteration procedure converges toward

$$u \;=\; \lim_{k\to\infty} u^{(k)} \;=\; \begin{pmatrix} 30 \\ 30 \\ 10 \\ 10 \end{pmatrix}, \tag{455}$$

which is the obvious solution of the vector equation $u = Au + c$. For the **general case**, the convergence behavior of this iteration strategy is studied in section 12.2.3 below.

The exact solution $u(x,y) = T(x,y)$ is given by Eq. (374) of section 10.2, with $b = a$,

$$T(x,y) \;=\; \frac{4T_0}{\pi} \sum_{n=1,3,5,\ldots} \frac{1}{n\,\sinh(n\pi)} \sin\left(\frac{n\pi}{a}\,x\right) \sinh\left(\frac{n\pi}{a}\,y\right). \tag{456}$$

Numerical summation of the terms $n = 1, 3, 5, ..., 199$ yields the values

$$\begin{aligned} u_1 \;\equiv\; T(\tfrac{a}{3}, \tfrac{2a}{3}) \;&=\; 30.45747 \;=\; u_2, \\ u_3 \;\equiv\; T(\tfrac{a}{3}, \tfrac{a}{3}) \;&=\; 9.54253 \;=\; u_4. \end{aligned} \tag{457}$$

Our present result (455) is **not exact**, since we have used a finite (and very rough) step size $h = \frac{a}{3}$. Slightly refining the step size to $h = \frac{a}{4}$ would more than double the number N of mesh points within Ω, see Fig. 20.

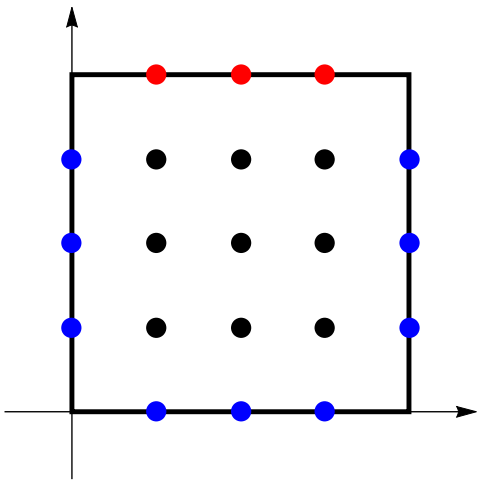


Figure 20: The region Ω, Eq. (449), in the xy-plane, with nine mesh points inside (black), cf. Fig. 19, corresponding to the slightly refined step size $h = \frac{a}{4}$.

12.2.3 Theory

Using in Eq. (448), $u = Au + c$, a trial vector $u^{(0)}$, we obtain to 1st and 2nd orders

$$u^{(1)} \;=\; Au^{(0)} + c, \tag{458}$$

$$\begin{aligned} u^{(2)} \;&=\; A\big(Au^{(0)} + c\big) + c \\ &\equiv\; A^2u^{(0)} + A^1c + A^0c \end{aligned} \tag{459}$$

(where $A^0 = 1$ denotes the unit matrix). Generally, to kth order, this expression becomes

$$u^{(k)} \;=\; A^k u^{(0)} + \sum_{j=0}^{k-1} A^j c. \tag{460}$$

As the $N \times N$-matrix A is **symmetric**, $A_{\mu\nu} = A_{\nu\mu}$,

$$A_{\mu\nu} \;=\; \begin{cases} \frac{1}{4} & \text{(when site } \mu \text{ is neighboring site } \nu\text{)}, \\ 0 & \text{(otherwise)}, \end{cases} \tag{461}$$

there is a **basis set** $\{v_1, ..., v_N\}$ of N (pairwise orthogonal) **eigenvectors** v_ν of A,

$$Av_\nu \;=\; \lambda_\nu v_\nu \qquad (\nu = 1, ..., N), \tag{462}$$

with **real-valued** eigenvalues $\lambda_\nu \in \mathbb{R}$. Consequently, we may expand

$$u^{(0)} \;=\; \sum_{\nu=1}^{N} \alpha_\nu v_\nu, \qquad c \;=\; \sum_{\nu=1}^{N} \gamma_\nu v_\nu, \tag{463}$$

to obtain in Eq. (460)

$$\begin{aligned} u^{(k)} \;&=\; A^k \left(\sum_{\nu=1}^{N} \alpha_\nu v_\nu \right) \;+\; \sum_{j=0}^{k-1} A^j \left(\sum_{\nu=1}^{N} \gamma_\nu v_\nu \right) \\ &=\; \sum_{\nu=1}^{N} \left(\alpha_\nu \lambda_\nu^k \;+\; \gamma_\nu \sum_{j=0}^{k-1} \lambda_\nu^j \right) v_\nu. \end{aligned} \tag{464}$$

Since $|\lambda_\nu| < 1$ for $\nu = 1, ..., N$ (see Theorem below), this expression converges for $k \to \infty$, and the solution of Eq. (448) is given by the limit

$$u \;=\; \lim_{k\to\infty} u^{(k)} \;=\; \sum_{\nu=1}^{N} \frac{\gamma_\nu}{1 - \lambda_\nu} v_\nu. \tag{465}$$

Theorem: Each eigenvalue λ of the matrix A has the property

$$|\lambda| \;<\; 1. \tag{466}$$

Proof: Notice that the sum S_ν of the elements in the ν-th row of A,

$$S_\nu \;=\; \sum_{\rho=1}^{N} A_{\nu\rho}, \tag{467}$$

has for all ν the property $|S_\nu| \leq 1$. In particular, $|S_\mu| < 1$ for at least one $\mu \in \{1, ..., N\}$. (This is so, since a point $\mathbf{r}_{nm}$ inside Ω cannot have more than four next neighbors inside Ω. Moreover, all the points next to the boundary $\partial\Omega$ have less than four such neighbors.) If there is an eigenvector $v \equiv (x_1, ..., x_N)$ with $|x_1| = |x_2| = ... = |x_N| = 1$, we have

$$|\lambda| \;=\; |\lambda x_\mu| \;=\; |(Av)_\mu| \;=\; \left| \sum_{\rho=1}^{N} A_{\mu\rho} x_\rho \right| \;\leq\; \sum_{\rho=1}^{N} A_{\mu\rho} |x_\rho| \;=\; S_\mu \;<\; 1, \tag{468}$$

where we have utilized, that $A_{\nu\rho} \geq 0$ for all ν and ρ. Any other eigenvector $v \equiv (x_1, ..., x_N)$ can be normalized in such a way that $x_\nu = 1$ for some ν, $|x_\rho| \leq 1$ for $\rho \neq \nu$, and $|x_\rho| < 1$ for at least one $\rho \neq \nu$. Then,

$$|\lambda| \;=\; |\lambda x_\nu| \;=\; |(Av)_\nu| \;=\; \left| \sum_{\rho=1}^{N} A_{\nu\rho} x_\rho \right| \;\leq\; \sum_{\rho=1}^{N} A_{\nu\rho} |x_\rho| \;<\; S_\nu \;\leq\; 1. \tag{469}$$

Example 2: We now consider a problem with $s(x,y) = 1$,

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} = -1. \tag{470}$$

Aluminium, e.g., has the heat conductivity $\lambda = 0.23\,\frac{\mathrm{kJ}}{\mathrm{s\cdot m\cdot K}}$. Choosing a uniform heating $s = 0.23\,\frac{\mathrm{kJ}}{\mathrm{s\cdot m^3}}$, the heat flow equation $c\mu\frac{\partial T}{\partial t} = \lambda\nabla^2 T + s$ in the stationary case reads

$$\nabla^2 T(\mathbf{r}) = -1\,\frac{\mathrm{K}}{\mathrm{m}^2}. \tag{471}$$

For comparison with the FEM treatment in section 12.3.4 below, we choose Dirichlet BCs, $u(\mathbf{r}) = 0$ for $\mathbf{r} \in \partial\Omega$, with respect to the region Ω shown in Fig. 21.
In this case, the analytical method (**separation ansatz**) of section 10.2 does **not apply**!

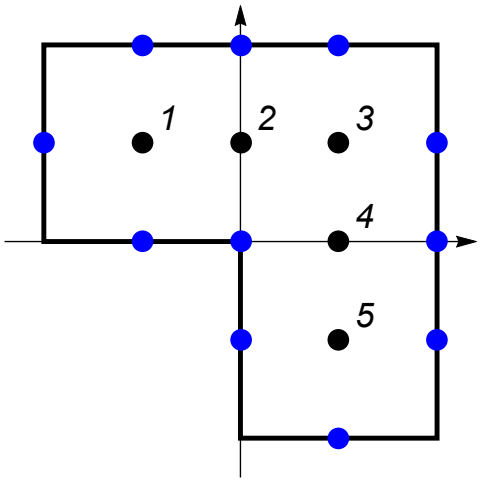


Figure 21: The region Ω in the xy-plane, with $\mathbf{r}_3 = (\frac{1}{2}|\frac{1}{2})$ (cf. Fig. 23 in section 12.3.4).

Now, since $s(x,y) \neq 0$, the step size $h = \frac{1}{2}$ enters the relevant **iteration equations** (448),

$$\begin{aligned} u_1 &= \tfrac{1}{4}(u_2 + h^2), \\ u_2 &= \tfrac{1}{4}(u_1 + u_3 + h^2), \\ u_3 &= \tfrac{1}{4}(u_2 + u_4 + h^2), \\ u_4 &= \tfrac{1}{4}(u_3 + u_5 + h^2), \\ u_5 &= \tfrac{1}{4}(u_4 + h^2). \end{aligned} \tag{472}$$

In matrix notation, these equations are summarized to read $u = Au + c$, with

$$A = \frac{1}{4}\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad c = \frac{h^2}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{473}$$

The eigenvectors $v_1, v_2, ..., v_5$ of the symmetric matrix A represent an ON basis of $\mathbb{R}^5$,

$$\frac{1}{2\sqrt{3}}\begin{pmatrix} 1 \\ -\sqrt{3} \\ 2 \\ -\sqrt{3} \\ 1 \end{pmatrix}, \quad \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 \\ \sqrt{3} \\ 2 \\ \sqrt{3} \\ 1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}. \tag{474}$$

Consequently, the representation $c = \sum_{\nu=1}^{5} \gamma_\nu v_\nu$ has the coefficients $\gamma_\nu = v_\nu \cdot c$,

$$\gamma_{1,2} \;=\; \frac{h^2}{4\sqrt{3}}\big(2 \mp \sqrt{3}\,\big), \qquad \gamma_{3,4} \;=\; 0, \qquad \gamma_5 \;=\; \frac{h^2}{4\sqrt{3}}. \tag{475}$$

With the corresponding eigenvalues,

$$\lambda_{1,2} = \mp\frac{1}{4}\sqrt{3}, \qquad \lambda_{3,4} = \mp\frac{1}{4}, \qquad \lambda_5 = 0, \tag{476}$$

Eq. (465) eventually yields the converged iterative solution

$$u \;=\; \frac{h^2(2-\sqrt{3})}{4\sqrt{3}+3}\,v_1 \;+\; \frac{h^2(2+\sqrt{3})}{4\sqrt{3}-3}\,v_2 \;+\; \frac{h^2}{4\sqrt{3}}\,v_5 \;=\; \frac{h^2}{52}\begin{pmatrix} 19 \\ 24 \\ 25 \\ 24 \\ 19 \end{pmatrix} \;=\; \begin{pmatrix} 0.091346 \\ 0.115385 \\ 0.120192 \\ 0.115385 \\ 0.091346 \end{pmatrix}, \tag{477}$$

where $h = \frac{1}{2}$ was used in the last step.
The first 8 iterations: Starting with the trial vector

$$u^{(0)} \;=\; \begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{pmatrix}, \tag{478}$$

we numerically find for $u^{(k+1)} = Au^{(k)} + c$ the (rounded) results

$$u^{(1)} = \begin{pmatrix} 0.31 \\ 0.56 \\ 0.56 \\ 0.56 \\ 0.31 \end{pmatrix}, \quad u^{(2)} = \begin{pmatrix} 0.20 \\ 0.28 \\ 0.34 \\ 0.28 \\ 0.20 \end{pmatrix}, \quad u^{(3)} = \begin{pmatrix} 0.13 \\ 0.20 \\ 0.20 \\ 0.20 \\ 0.13 \end{pmatrix}, \quad u^{(4)} = \begin{pmatrix} 0.11 \\ 0.15 \\ 0.16 \\ 0.15 \\ 0.11 \end{pmatrix},$$

$$u^{(5)} = \begin{pmatrix} 0.099 \\ 0.131 \\ 0.136 \\ 0.131 \\ 0.099 \end{pmatrix}, \quad u^{(6)} = \begin{pmatrix} 0.095 \\ 0.121 \\ 0.128 \\ 0.121 \\ 0.095 \end{pmatrix}, \quad u^{(7)} = \begin{pmatrix} 0.093 \\ 0.118 \\ 0.123 \\ 0.118 \\ 0.093 \end{pmatrix}, \quad u^{(8)} = \begin{pmatrix} 0.092 \\ 0.116 \\ 0.122 \\ 0.116 \\ 0.092 \end{pmatrix}.$$

12.3 Finite elements method (FEM)

Again, we consider for the unknown function $u(\mathbf{r})$ the PDE

$$\nabla^2 u(\mathbf{r}) \;=\; -s(\mathbf{r}) \qquad (\mathbf{r} \in \Omega), \tag{479}$$

where $s(\mathbf{r})$ is a given source function.
Again, $\partial\Omega_1$ and $\partial\Omega_2 = \partial\Omega \backslash \partial\Omega_1$ shall be two disjoint pieces of the boundary $\partial\Omega$ of Ω.

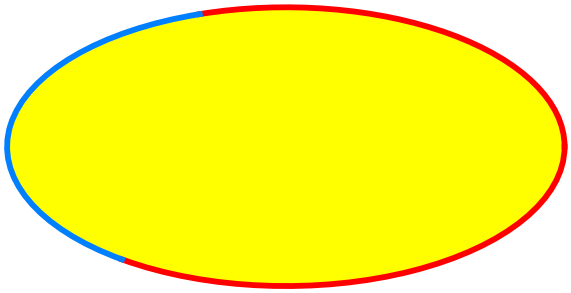


Figure 22: Two disjoint pieces $\partial\Omega_1$ (blue) and $\partial\Omega_2$ (red) of the boundary $\partial\Omega$ of a region Ω (yellow).

12.3.1 Weak formulation of a PDE

For a given region $\Omega \subset \mathbb{R}^n$, let $\mathcal{B}$ be the set of all functions $v : \Omega \to \mathbb{R},\ \mathbf{r} \mapsto v(\mathbf{r})$ with "reasonable" properties (continuity, two-fold partial differentiability, etc.).

Theorem: The following two statements are equivalent:

(I) The function $u \in \mathcal{B}$ is a solution of $\quad \nabla^2 u = -s \quad$ that satisfies the Neumann BC

$$\text{(N)} \qquad \boldsymbol{\mathcal{N}}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) \;=\; g(\mathbf{r}) \qquad \forall\, \mathbf{r} \in \partial\Omega_2. \tag{480}$$

(II) The function $u \in \mathcal{B}$ satisfies the equation

$$\int_\Omega \mathrm{d}^n r \left\{ w(\mathbf{r}) s(\mathbf{r}) \;-\; \big[\nabla w(\mathbf{r})\big] \cdot \big[\nabla u(\mathbf{r})\big] \right\} \;-\; \int_{\partial\Omega_2} \mathrm{d}^{n-1} r\; w(\mathbf{r}) g(\mathbf{r}) \;=\; 0, \tag{481}$$

for **all** functions $w \in \mathcal{B}_1^{\mathrm{hom}}$, where

$$\mathcal{B}_1^{\mathrm{hom}} \;=\; \left\{ w \in \mathcal{B} \;\middle|\; w(\mathbf{r}) = 0 \qquad \forall\, \mathbf{r} \in \partial\Omega_1 \right\}. \tag{482}$$

Corollary: The **unique** solution $u(\mathbf{r})$ of $\quad \nabla^2 u = -s \quad$ satisfying both the BCs (N) and

$$\text{(D)} \qquad u(\mathbf{r}) \;=\; f(\mathbf{r}) \qquad \forall\, \mathbf{r} \in \partial\Omega_1 \tag{483}$$

is that function $u \in \mathcal{B}$ that (i) satisfies Eq. (481) for **all** $w \in \mathcal{B}_1^{\mathrm{hom}}$ and (ii) satisfies (D).
[*This is called the* **weak formulation** *of the PDE* $\quad \nabla^2 u = -s \quad$ *with* (D) *plus* (N).]

Proof of the theorem: Statement (I) is equivalent to saying that $u(\mathbf{r})$ has the property

$$0 \;=\; \int_\Omega \mathrm{d}^n r\; w(\mathbf{r})\Big[\underbrace{s(\mathbf{r}) + \nabla^2 u(\mathbf{r})}_{0}\Big] \;+\; \int_{\partial\Omega_2} \mathrm{d}^{n-1} r\; w(\mathbf{r})\Big[\underbrace{\boldsymbol{\mathcal{N}}(\mathbf{r})\cdot\nabla u(\mathbf{r}) \;-\; g(\mathbf{r})}_{0}\Big], \tag{484}$$

for all $w \in \mathcal{B}_1^{\mathrm{hom}}$. Since $\nabla\cdot(w\nabla u) = w\nabla^2 u + (\nabla w)\cdot(\nabla u)$, we may write this as

$$0 \;=\; \int_\Omega \mathrm{d}^n r\,\Big[ws \;+\; \nabla\cdot(w\nabla u) \;-\; (\nabla w)\cdot(\nabla u)\Big] \;+\; \int_{\partial\Omega_2} \mathrm{d}^{n-1} r\; w\Big[\boldsymbol{\mathcal{N}}\cdot\nabla u \;-\; g\Big]. \tag{485}$$

Now, Gauss' theorem yields

$$\begin{aligned} \int_\Omega \mathrm{d}^n r\,\Big[\nabla\cdot(w\nabla u)\Big] \;&=\; \oint_{\partial\Omega} \mathrm{d}^{n-1} r\; \boldsymbol{\mathcal{N}}\cdot(w\nabla u) \\ &\equiv\; \oint_{\partial\Omega} \mathrm{d}^{n-1} r\; w\Big[\boldsymbol{\mathcal{N}}\cdot\nabla u\Big]. \end{aligned} \tag{486}$$

Since $w \in \mathcal{B}_1^{\mathrm{hom}}$, we may replace $\partial\Omega$ with $\partial\Omega_2$ here, and Eq. (485) becomes Eq. (481). Consequently, statements (I) and (II) are equivalent.

12.3.2 System of linear equations

Confining the set $\mathcal{B}_1^{\mathrm{hom}}$ properly (e.g., by excluding all functions w that are not square integrable), we may select an infinite series $\{w_j\}_{j\in\mathbb{N}}$ of linearly independent functions $w_j \in \mathcal{B}_1^{\mathrm{hom}}$ such that each function $w \in \mathcal{B}_1^{\mathrm{hom}}$ has a representation

$$w(\mathbf{r}) \;=\; \sum_{j=1}^{\infty} \xi_j\, w_j(\mathbf{r}), \tag{487}$$

where the ξ_j are constant ($\mathbf{r}$-independent) coefficients, fixed uniquely by $w(\mathbf{r})$.

Let $u_{\mathrm{D}}(\mathbf{r})$ be **some** function satisfying (D) and let $u(\mathbf{r})$ be the **unique solution** of the PDE (479) that also satisfies (N). Since $u - u_{\mathrm{D}} \in \mathcal{B}_1^{\mathrm{hom}}$, we have

$$u(\mathbf{r}) \;=\; u_{\mathrm{D}}(\mathbf{r}) \;+\; \sum_{j=1}^{\infty} w_j(\mathbf{r})\,\xi_j. \tag{488}$$

Given the functions $u_{\mathrm{D}}(\mathbf{r})$ and $w_1(\mathbf{r})$, $w_2(\mathbf{r})$, ..., we wish to find the coefficients ξ_1, ξ_2, ..., that yield via Eq. (488) the unknown wanted function $u(\mathbf{r})$.

Choosing in Eq. (481) $w = w_i$ and using expression (488) for u, we obtain

$$\sum_{j=1}^{\infty} A_{ij}\,\xi_j \;=\; \eta_i, \tag{489}$$

with the constants

$$\begin{aligned} A_{ij} \;&=\; \int_\Omega \mathrm{d}^n r\,\big[\nabla w_i(\mathbf{r})\big]\cdot\big[\nabla w_j(\mathbf{r})\big], \\ \eta_i \;&=\; \int_\Omega \mathrm{d}^n r\,\Big\{s(\mathbf{r})w_i(\mathbf{r}) - \big[\nabla u_{\mathrm{D}}(\mathbf{r})\big]\cdot\big[\nabla w_i(\mathbf{r})\big]\Big\} \;-\; \int_{\partial\Omega_2} \mathrm{d}^{n-1} r\; g(\mathbf{r})w_i(\mathbf{r}). \end{aligned} \tag{490}$$

These can be evaluated for any given choice of the functions $u_{\mathrm{D}}(\mathbf{r})$ and $w_1(\mathbf{r})$, $w_2(\mathbf{r})$, ... Then, Eq. (489) provides an infinite system of linear equations for the wanted coefficients $\xi_1, \xi_2, \ldots$. In matrix form, this system reads

$$A\xi \;=\; \eta. \tag{491}$$

12.3.3 Summary (3D notation)

• **Given:** A volume region $\Omega \subset \mathbb{R}^3$ with surface

$$\partial\Omega \;=\; \partial\Omega_1 \cup \partial\Omega_2 \qquad \Big(\partial\Omega_1 \cap \partial\Omega_2 \;=\; \emptyset\Big) \tag{492}$$

and three functions

$$s:\; \Omega \;\to\; \mathbb{R}, \qquad \mathbf{r} \;\mapsto\; s(\mathbf{r}), \tag{493}$$
$$f:\; \partial\Omega_1 \;\to\; \mathbb{R}, \qquad \mathbf{r} \;\mapsto\; f(\mathbf{r}), \tag{494}$$
$$g:\; \partial\Omega_2 \;\to\; \mathbb{R}, \qquad \mathbf{r} \;\mapsto\; g(\mathbf{r}). \tag{495}$$

• **Wanted:** The unique function $u:\; \Omega \;\to\; \mathbb{R},\; \mathbf{r} \;\mapsto\; u(\mathbf{r})$ that satisfies:

$$\text{(PDE)} \qquad \nabla^2 u(\mathbf{r}) \;=\; -s(\mathbf{r}) \qquad (\forall\, \mathbf{r} \;\in\; \Omega), \tag{496}$$
$$\text{(D)} \qquad u(\mathbf{r}) \;=\; f(\mathbf{r}) \qquad (\forall\, \mathbf{r} \;\in\; \partial\Omega_1), \tag{497}$$
$$\text{(N)} \qquad \boldsymbol{\mathcal{N}}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) \;=\; g(\mathbf{r}) \qquad (\forall\, \mathbf{r} \;\in\; \partial\Omega_2). \tag{498}$$

Here, $\boldsymbol{\mathcal{N}}(\mathbf{r})$ is the unit vector normal to $\partial\Omega_2$ (and pointing outward) at $\mathbf{r} \in \partial\Omega_2$.

• **Solution:**

$$u(\mathbf{r}) \;=\; u_D(\mathbf{r}) \;+\; \sum_{j=1}^{\infty} \xi_j\, w_j(\mathbf{r}). \tag{499}$$

Here:

(a) The function $u_D(\mathbf{r})$ is **some** function $u_D:\; \Omega \to \mathbb{R}$ satisfying (D),

$$u_D(\mathbf{r}) = f(\mathbf{r}) \qquad (\forall\, \mathbf{r} \;\in\; \partial\Omega_1), \tag{500}$$

(b) $\{w_1, w_2, w_3, ...\}$ is a basis set of the infinite-dimensional function space[9]

$$\mathcal{B}_1^{\text{hom}} \;=\; \Big\{\, w : \Omega \to \mathbb{R} \;\Big|\; w(\mathbf{r}) = 0 \quad (\forall\, \mathbf{r} \;\in\; \partial\Omega_1) \,\Big\}, \tag{502}$$

(c) the coefficients ξ_j are fixed by the linear system of equations

$$A\xi \;=\; \eta, \tag{503}$$

where the elements A_{ij} (of the matrix A) and η_i (of the column η) are fixed by: the functions s, f (via u_D), g, and $\{w_1, w_2, w_3, ...\}$ via Eqs. (490),

$$A_{ij} \;=\; \int_\Omega \mathrm{d}^n r\, \big[\nabla w_i(\mathbf{r})\big] \cdot \big[\nabla w_j(\mathbf{r})\big],$$
$$\eta_i \;=\; \int_\Omega \mathrm{d}^n r \Big\{ s(\mathbf{r}) w_i(\mathbf{r}) - \big[\nabla u_{\mathrm{D}}(\mathbf{r})\big] \cdot \big[\nabla w_i(\mathbf{r})\big] \Big\} \;-\; \int_{\partial\Omega_2} \mathrm{d}^{n-1} r\; g(\mathbf{r}) w_i(\mathbf{r}).$$

[9] This means: For any function $w \in \mathcal{B}_1^{\text{hom}}$ there is a unique sequence $(\xi_1, \xi_2, \xi_3, ...)$ of coefficients ξ_i such that

$$w(\mathbf{r}) \;=\; \sum_{i=1}^{\infty} \xi_i\, w_i(\mathbf{r}). \tag{501}$$

12.3.4 Finite elements

In the FEM, the vectorspace $\mathcal{B}_1^{\rm hom}$ is approximated by a finite-dimensional one, selecting a finite number of basis functions $w_1(\mathbf{r}), ..., w_N(\mathbf{r})$.

For illustration, we consider the 2D problem of Example 2 in section 12.2.3: Ω is given by Fig. 21/23, with $\partial\Omega_1 = \partial\Omega$ and $\partial\Omega_2 = \emptyset$, $s(\mathbf{r}) \equiv 1$ and $f(\mathbf{r}) \equiv 0$,

$$\nabla^2 u(\mathbf{r}) = -1 \quad (\mathbf{r} \in \Omega), \tag{504}$$

$$u(\mathbf{r}) = 0 \quad (\mathbf{r} \in \partial\Omega). \tag{505}$$

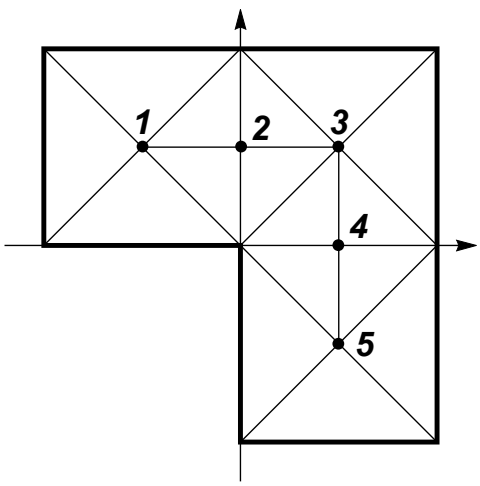


Figure 23: The region Ω in the xy-plane, Fig. 21, divided up into 16 triangles.

For the FEM, we divide Ω up into 16 triangles (**finite elements**), as shown in Fig. 23. Five vertices $\mathbf{r}_1, ..., \mathbf{r}_5$ of these triangles are lying **inside** Ω (dots in Fig. 23),

$$\mathbf{r}_1 = (-\tfrac{1}{2}|\tfrac{1}{2}), \qquad \mathbf{r}_2 = (0|\tfrac{1}{2}), \qquad \mathbf{r}_3 = (\tfrac{1}{2}|\tfrac{1}{2}), \qquad \mathbf{r}_4 = (\tfrac{1}{2}|0), \qquad \mathbf{r}_5 = (\tfrac{1}{2}|-\tfrac{1}{2}), \tag{506}$$

the remaining eight vertices (with negative subscripts) $\mathbf{r}_{-1}, ..., \mathbf{r}_{-8}$ are on the boundary $\partial\Omega$ (not marked in Fig. 23). For each vertex $\mathbf{r}_i$, a function $w_i(\mathbf{r})$ is introduced as that **continuous** function of $\mathbf{r} \in \Omega$, which (i) within each one of the 16 triangles is a **linear function** $w(x, y) = ax + by + c$ of x and y and (ii) on the 13 vertices assumes the values

$$w_i(\mathbf{r}_j) = \delta_{ij} \equiv \begin{cases} 1 & (j = i), \\ 0 & (j \neq i). \end{cases} \tag{507}$$

In other words: $w_i(\mathbf{r})$ assumes its maximum value 1 at $\mathbf{r} = \mathbf{r}_i$, falls linearly from 1 to 0 within each triangle sharing the vertex $\mathbf{r}_i$, and is zero within all the remaining triangles.[10] The functions $w_i(\mathbf{r})$ with $i < 0$ are needed only for the approximation

$$u_{\rm D}(\mathbf{r}) \approx U_{\rm D}(\mathbf{r}) = \sum_{i=-1}^{-8} f(\mathbf{r}_i) w_i(\mathbf{r}) \tag{508}$$

[10] Writing $\mathbf{r}_i = (x_i, y_i)$ and $X = x - x_i$, $Y = y - y_i$, we explicitly have for $i = 1, 3, 5$

$$w_i(x, y) = \begin{cases} 1 - 2X & \big(X \in [0, \tfrac{1}{2}],\ |Y| \leq X\big), \\ 1 - 2Y & \big(Y \in [0, \tfrac{1}{2}],\ |X| \leq Y\big), \\ 1 + 2X & \big(X \in [-\tfrac{1}{2}, 0],\ |Y| \leq -X\big), \\ 1 + 2Y & \big(Y \in [-\tfrac{1}{2}, 0],\ |X| \leq -Y\big), \end{cases}$$

to the function $u_{\rm D}(\mathbf{r})$ in Eq. (488), while the complete $u(\mathbf{r})$ there is approximated by

$$U(\mathbf{r}) \;=\; U_{\rm D}(\mathbf{r}) \;+\; \sum_{i=1}^{5} \xi_i \, w_i(\mathbf{r}). \tag{509}$$

We see that $\mathcal{B}_1^{\rm hom}$ is approximated here by a **5-dimensional** vectorspace. Due to Eq. (507), we also see that $\xi_i = U(\mathbf{r}_i)$ for $i > 0$. Moreover, since $f(\mathbf{r}) = 0$ here, we have

$$U_{\rm D}(\mathbf{r}) \;=\; 0. \tag{510}$$

Now, with $\mathbf{r} = (x, y)$, $u_{\rm D}(\mathbf{r}) \approx U_{\rm D}(\mathbf{r}) = 0$, $s(\mathbf{r}) = 1$, and $\partial\Omega_2 = \emptyset$, Eqs. (490) read

$$\eta_i \;=\; \int_\Omega {\rm d}^2 r \; w_i(\mathbf{r}), \qquad A_{ij} \;=\; \int_\Omega {\rm d}^2 r \, \big[\nabla w_i(\mathbf{r})\big] \cdot \big[\nabla w_j(\mathbf{r})\big]. \tag{511}$$

The support $\Sigma_i = \{\mathbf{r} \in \Omega \mid w_i(\mathbf{r}) \neq 0\}$ of $w_i(\mathbf{r})$ is a square with area $S_i = 1$ for $i = 1, 3, 5$ and $S_i = \frac{1}{2}$ for $i = 2, 4$. Consequently, since $\big\langle w_i(\mathbf{r}) \big\rangle_{\mathbf{r} \in S_i} = \frac{1}{3}$, we have

$$\eta_i \;\equiv\; S_i \, \big\langle w_i(\mathbf{r}) \big\rangle_{\mathbf{r} \in S_i} \;=\; \frac{S_i}{3} \;=\; \begin{cases} \frac{1}{3} & (i = 1, 3, 5), \\ \frac{1}{6} & (i = 2, 4). \end{cases} \tag{512}$$

For $\mathbf{r} \in \Sigma_i$, it is easy to verify that $|\nabla w_i(\mathbf{r})|^2 = 4$ for $i = 1, 3, 5$, while $|\nabla w_i(\mathbf{r})|^2 = 8$ for $i = 2, 4$. Consequently, we have

$$A_{ii} \;\equiv\; S_i \, |\nabla w_i(\mathbf{r})|^2 \;=\; 4 \qquad (i = 1, ..., 5). \tag{513}$$

In the overlap region $\Sigma_1 \cap \Sigma_2$ between Σ_1 and Σ_2, we have $\nabla w_1(\mathbf{r}) = -\binom{2}{0}$, while $\nabla w_2(\mathbf{r}) = \binom{+2}{-2}$ for $y > \frac{1}{2}$ and $\nabla w_2(\mathbf{r}) = \binom{2}{2}$ for $y < \frac{1}{2}$. Consequently,

$$A_{12} \;=\; \int_{\Sigma_1 \cap \Sigma_2} {\rm d}^2 r \begin{pmatrix} -2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ \pm 2 \end{pmatrix} \;=\; \frac{1}{4} \cdot (-4) \;=\; -1, \tag{514}$$

where $\frac{1}{4}$ is the area of the triangle $\Sigma_1 \cap \Sigma_2$. Similarly, we find $A_{ij} = -1$ for $|i - j| = 1$ and $A_{ij} = 0$ for all other matrix elements A_{ij} with $i \neq j$. In summary,

$$A = \begin{pmatrix} 4 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 4 \end{pmatrix}, \qquad \eta = \frac{1}{6} \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}. \tag{515}$$

The solution ξ of the corresponding system $A\xi = \eta$ of 5 linear equations is

$$\xi \;=\; \frac{1}{156} \begin{pmatrix} 17 \\ 16 \\ 21 \\ 16 \\ 17 \end{pmatrix} \;=\; \begin{pmatrix} 0.1090 \\ 0.1026 \\ 0.1346 \\ 0.1026 \\ 0.1090 \end{pmatrix}. \tag{516}$$

An accurate computation yields $u(\frac{1}{2}, \frac{1}{2}) = 0.1310$ [Sewell, p. 224], to be compared with the present value $U(\mathbf{r}_3) \equiv \xi_3 = 0.1346$.

---

and for $i = 2, 4$

$$w_i(x, y) \;=\; \begin{cases} 1 - 2X - 2Y & \big(X \in [0, \frac{1}{2}],\; Y \in [0, \frac{1}{2} - X]\big), \\ 1 + 2X - 2Y & \big(X \in [-\frac{1}{2}, 0],\; Y \in [0, \frac{1}{2} + X]\big), \\ 1 + 2X + 2Y & \big(X \in [-\frac{1}{2}, 0],\; Y \in [-\frac{1}{2} - X, 0]\big), \\ 1 - 2X + 2Y & \big(X \in [0, \frac{1}{2}],\; Y \in [-\frac{1}{2} + X, 0]\big), \end{cases}$$

while $w_i(x, y) = 0$ for all other values of X, Y.