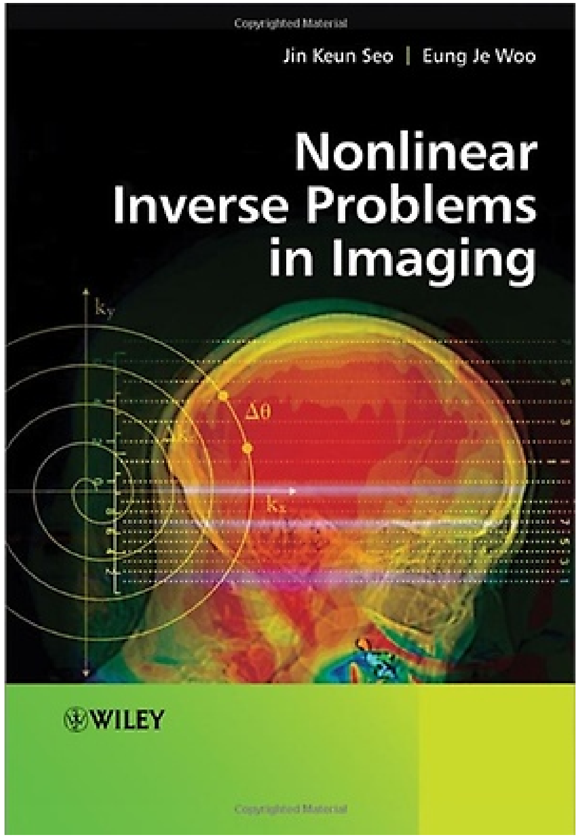
Partial Differential Equations
for Computational Science & Engineering



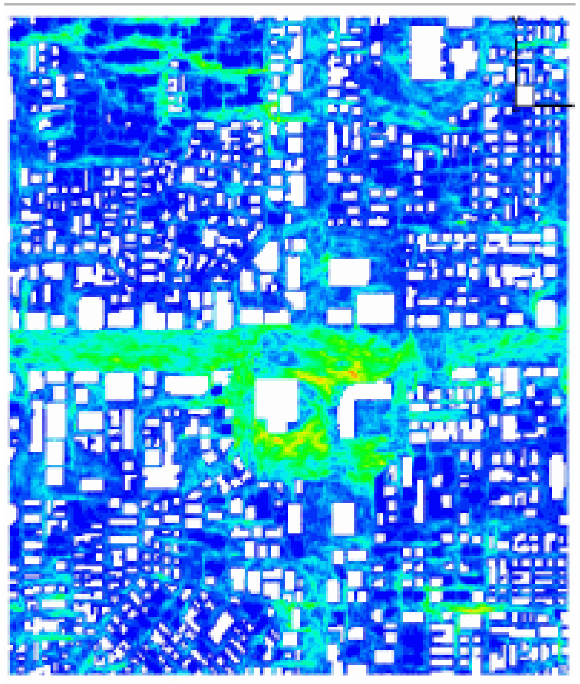
Lecture 1. Introduction of PDEs

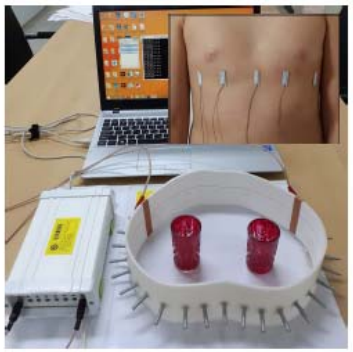
Homepage: https://www.deepmediview.com/blank-15

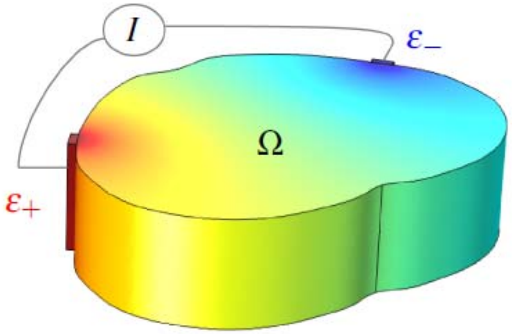
Jin Keun Seo (CSE@Yonsei university)

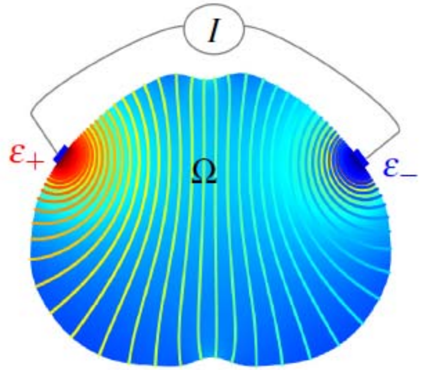
Partial Differential Equations (PDEs)

- **Mathematical techniques in science and engineering have evolved to expand our ability to visualize various physical phenomena of interest and their characteristics in detail.**
- **A large variety of natural phenomena occurring in real-life applications from fluid flows to biology and medical imaging fields are described by means of PDEs.**
- **Finding solutions with practical significance and value requires an in-depth understanding of the underlying physical phenomena with data acquisition systems as well as the implementation details of algorithms.**



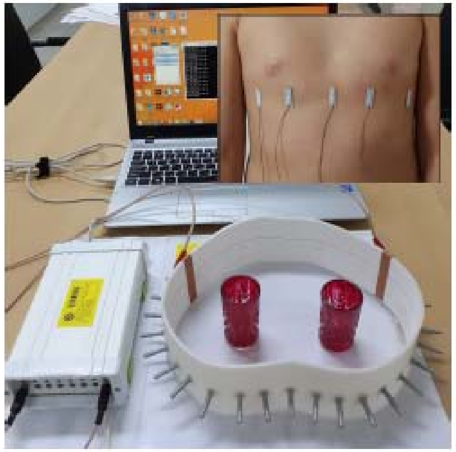


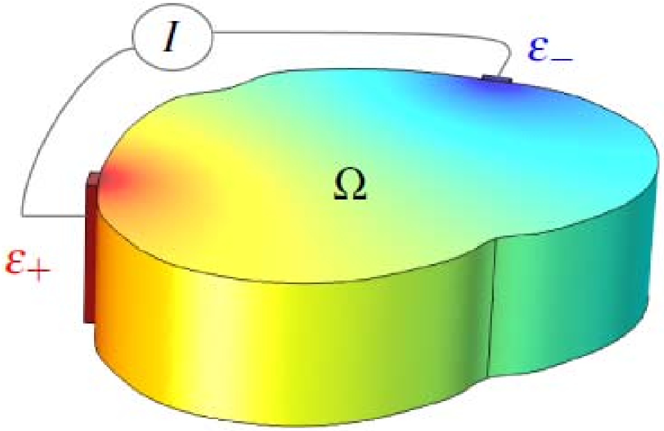


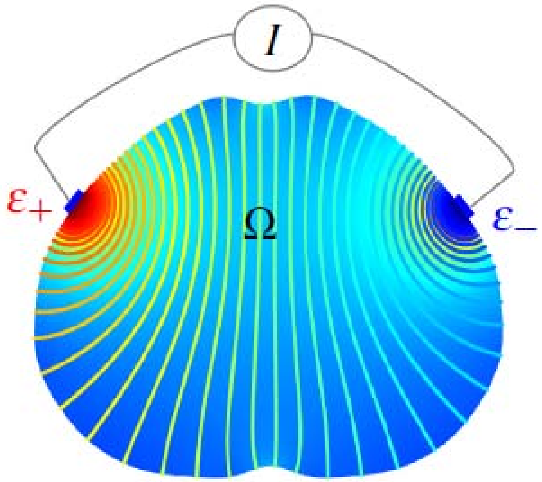


Laplace Equation

Laplace equation $\Delta u = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) u = 0$ where $u = u(x, y, z)$ is the potential at position (x, y, z).

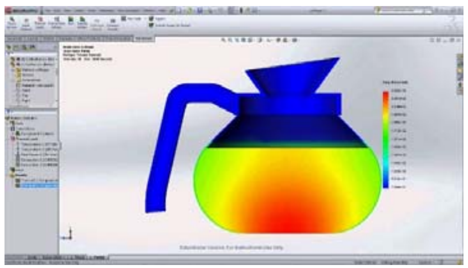


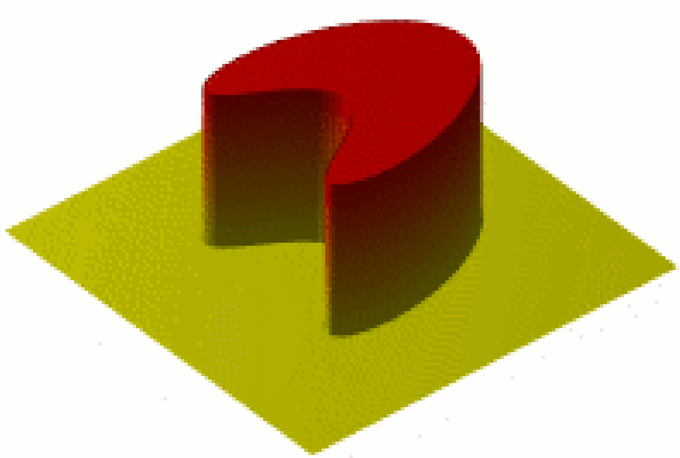


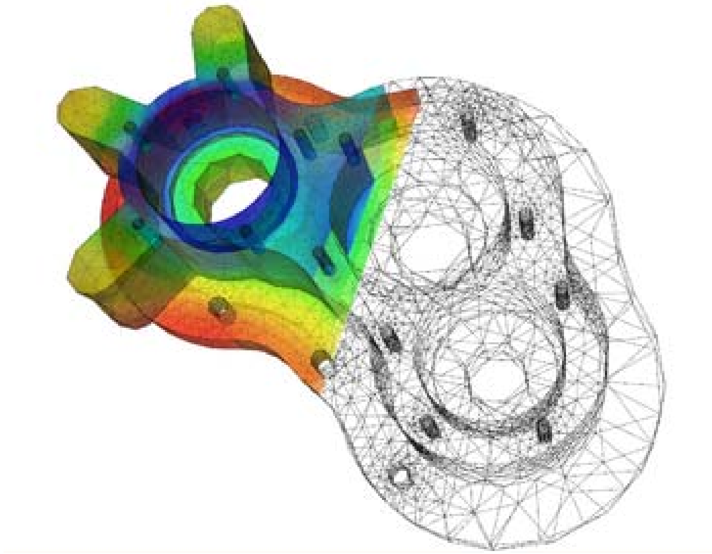


Heat equation

Heat equation $\frac{\partial}{\partial t}u - \Delta u = 0$ **where** $u = u(x, y, z, t)$ **is the temperature at position** (x, y, z) **and time** t.





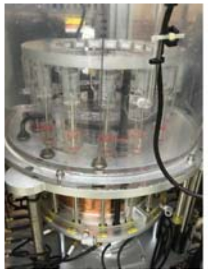


Animated plot of the evolution of the temperature in a square metal plate as predicted by the heat equation. The height and redness indicate the temperature at each point.

Image credit: Wikipedia

Helmholtz equation

Helmholtz equation $\boldsymbol{\Delta u + k^2 u = 0}$ where $u = u(x, y, z)$ is the pressure at position (x, y, z).



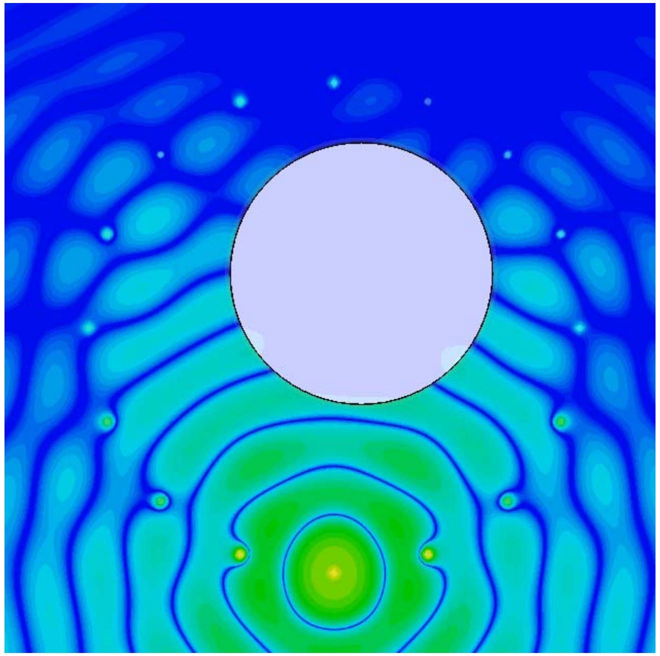
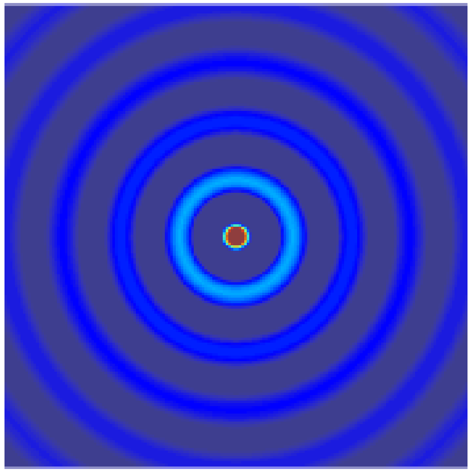


Image credit: Samsung HS60

Wave equation

Wave equation $\frac{\partial^2}{\partial t^2}u - \Delta u = 0$ where $u = u(x, y, z, t)$ is the displacement at position (x, y, z) and time t.



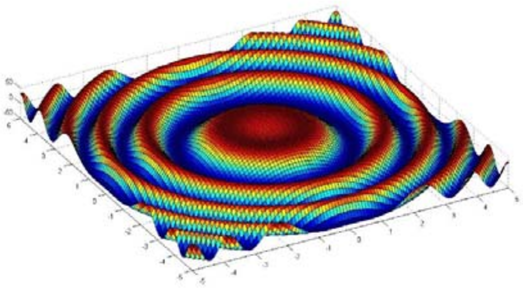
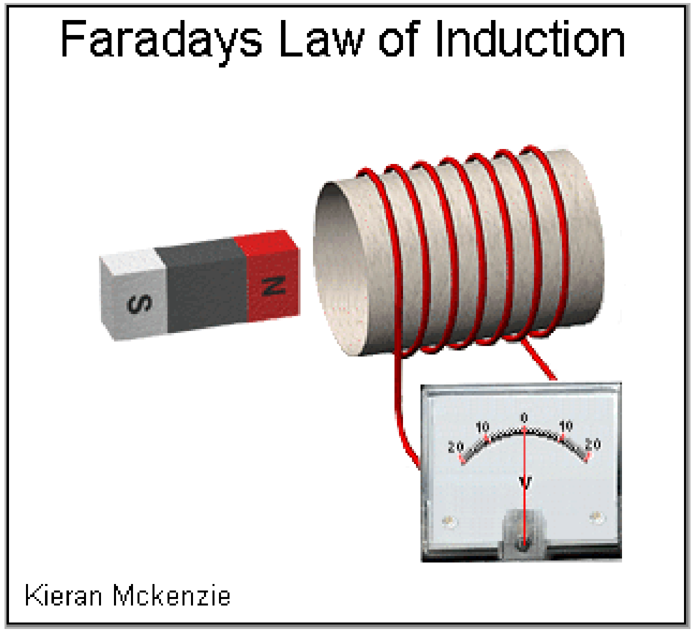
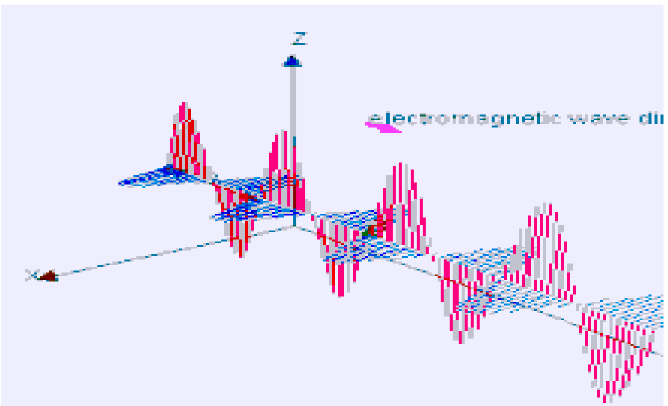


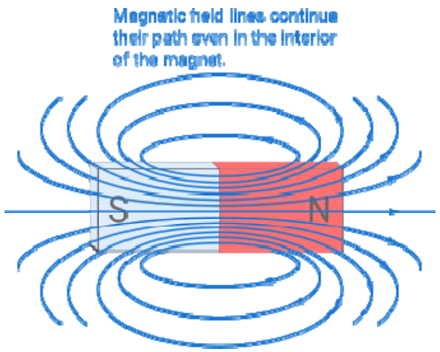
Image credit: Wikipedia

Maxwell’s equations in electromagnetism

- Faraday’s law of induction: $\boldsymbol{\nabla \times E = -\frac{\partial}{\partial t} B}$
- Ampere’s circuital law: $\boldsymbol{\nabla \times H = J + \frac{\partial}{\partial t} D}$
- Gauss’s Law for magnetism: $\boldsymbol{\nabla \cdot B = 0}$
- Gauss’s Law : $\boldsymbol{\nabla \cdot E = \rho / \epsilon}$







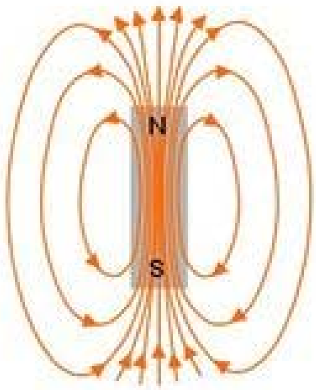
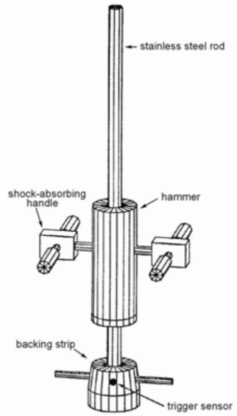


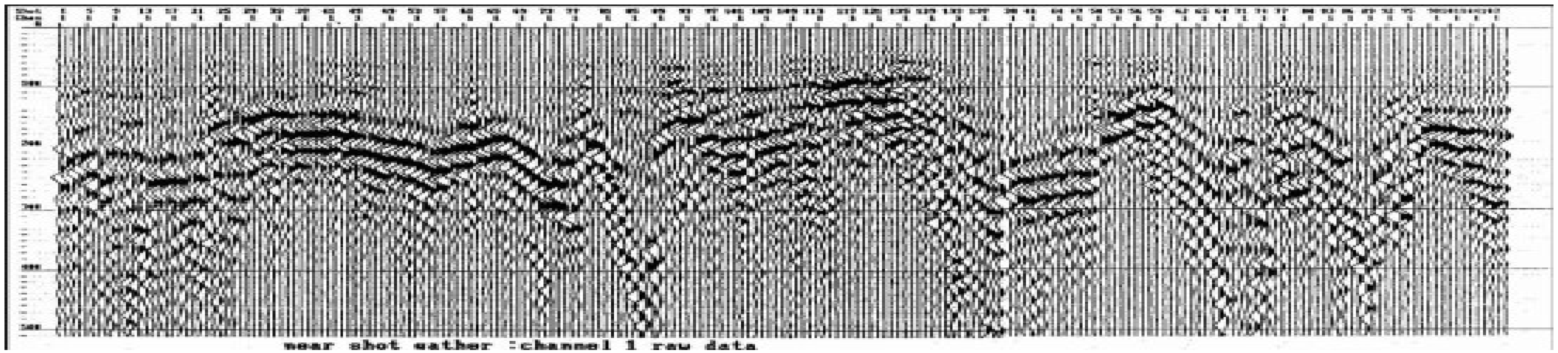
Image credit: Wikipedia

Elastic equation

Linear elasticity equation $\frac{\partial^2}{\partial t^2}\boldsymbol{u} = \mu\Delta\boldsymbol{u} + (\mu+\lambda)\nabla\nabla\cdot\boldsymbol{u}$

where $u = u(x,y,z,t)$ is the displacement at position (x,y,z) and time t.





Navier-Stokes Equation in Fluid Mechanics

Inertia per volume

Divergence of stress

$$\rho\left(\frac{\partial \boldsymbol{v}}{\partial t}+\boldsymbol{v}\cdot\nabla\boldsymbol{v}\right) = -\nabla p + \mu\nabla^2\boldsymbol{v} + \boldsymbol{f}$$

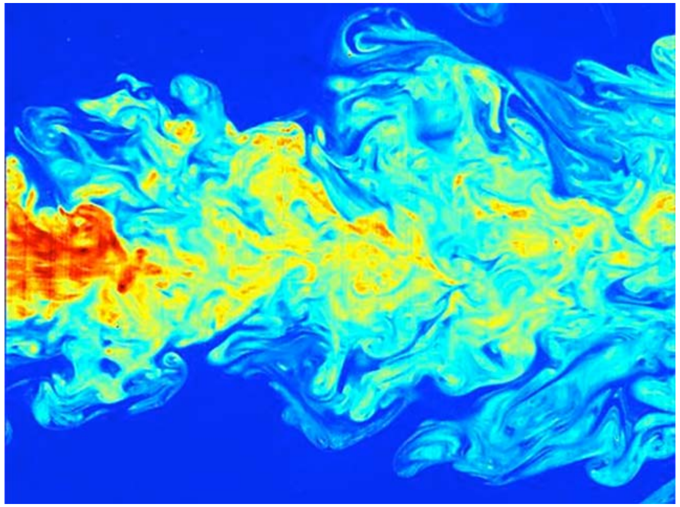
Eulerian acceleration

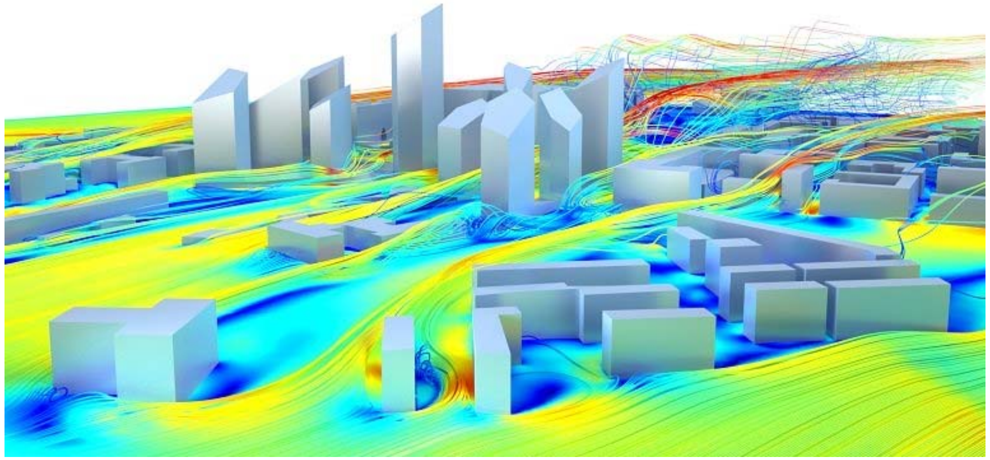
advection

Pressure gradient

Viscosity

Other body force





Hadamard (born in 1865) believed that mathematical models must satisfy three properties: existence, uniqueness, and stability.

Linear system

$Ax = b$ is well-posed if the following three conditions hold:

- [Existence] for each b, there exist at least one possible solution of $Ax = b$;
- [Uniqueness] for each b, $Ax = b$ has a unique solution;
- [Stability] the solution is stable under perturbation of b.

Well-posedness

Constructing a mathematical model, which transforms physical phenomena into a collection of mathematical expressions and data, we should consider the following three properties:

- **Existence:** at least one solution exists. For example, the problem $u''(x) = 1$ in $[0, 1]$ has at least one possible solution of $u = \frac{x^2}{2}$, whereas the problem $|u''(x)|^2 = -1$ has no solution.
- **Uniqueness:** only one solution exists. For example, the boundary value problem

$$u''(x) = u(x) \quad (\, 0 < \forall x < 1), \;\; u(0) = 1, \;\; u'(0) = e$$

  has the unique solution of $u(x) = e^x$, whereas the boundary value problem

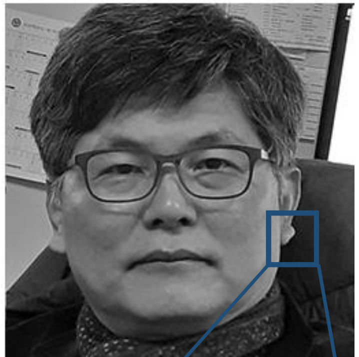
$$\left(\frac{u'(x)}{|u'(x)|}\right)' = 0 \quad (\, 0 < \forall x < 1), \;\; u(0) = 0, \;\; u(1) = 1$$

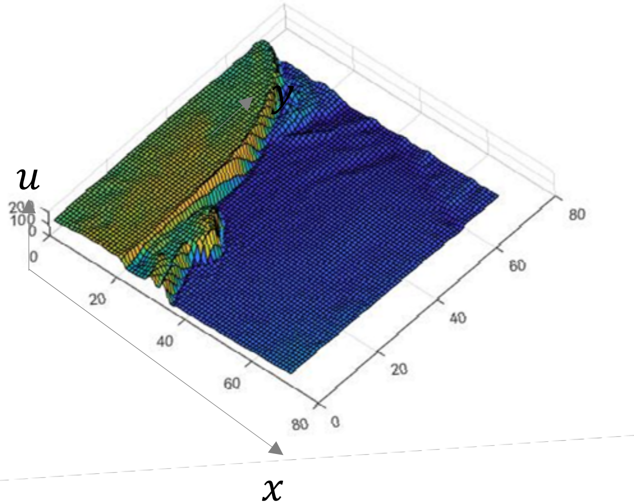
  has infinitely many solutions of $u(x) = x, x^2, x^3, \cdots$.
- **Continuity or stability:** a solution depends continuously on the data.

Understanding PDEs using 2D images as examples

- **A function $u(x, y)$ of two variables (x, y) can be represented as a grayscale image.**
- **The grayscale image is represented as a matrix (u_{ij}), with each element corresponding to one image pixel.**

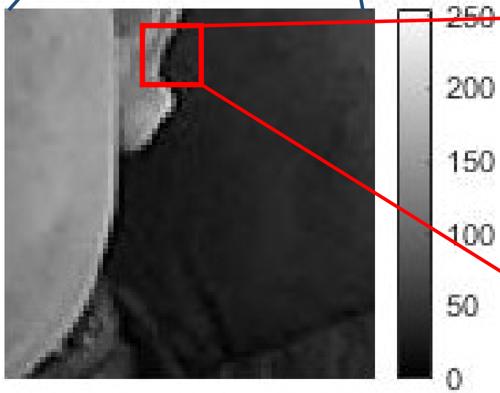
$$\begin{bmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mn} \end{bmatrix} =$$





$$u(x, y) \approx u_{ij}$$

white-255
black-0



108	148	168	105	105	170	133	25	23	41
114	155	152	92	127	169	82	12	31	36
122	161	135	87	147	145	34	21	46	37
149	148	112	112	157	104	17	41	35	38
154	147	101	128	139	64	19	37	38	35
152	131	99	146	106	21	28	35	42	35
142	110	117	147	63	6	39	37	44	39
126	105	142	122	24	18	45	39	41	41
111	113	155	84	5	32	42	37	37	39
110	130	152	50	12	36	36	36	39	38

Each pixel is assigned a value of grayscale level between 0 and 255

Poisson's equation

This image $u(x,y) \approx u_{ij}$ can be viewed as a solution of Poisson's equation

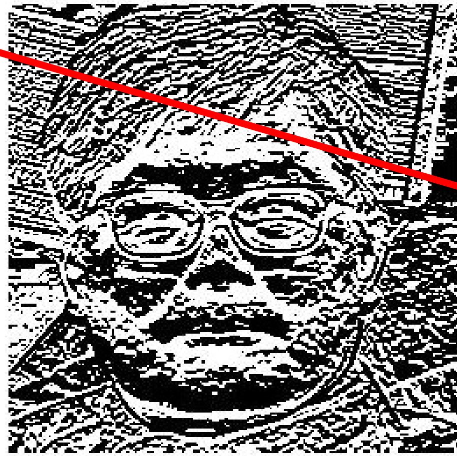
$$\nabla^2 u = \rho \quad in \quad \Omega$$

The sparse data ρ contains almost the full information of the image u .

u

$\frac{\partial u}{\partial x}$

$\frac{\partial u}{\partial y}$



$\rho = \nabla^2 u$

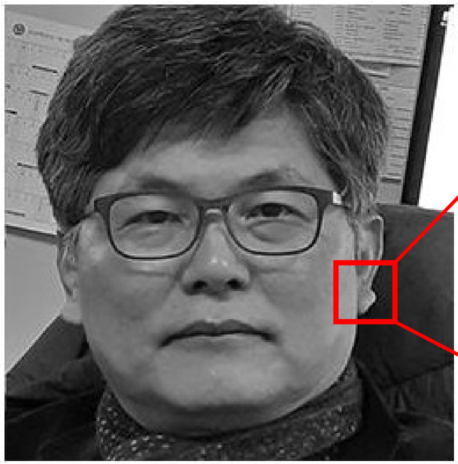
Wave equation

This image $u(x, y) \approx u_{ij}$ can be viewed as a solution of the wave equation

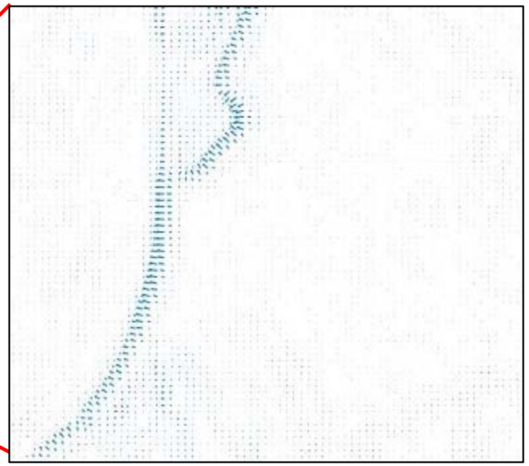
$$2\frac{\partial u}{\partial x} + 3\frac{\partial u}{\partial y} = \phi \quad \text{in} \quad \Omega$$

with the initial condition $u(x, 0) = f_1(x)$, $u(0, y) = f_2(y)$ for $0 < x, y < L$. Here, ϕ is plotted at the right side of figure and the initial data is the boundary intensity of the image u on the left.

u



∇u



ϕ

Heat equation

These images shows the solution $w(x,y,t)$ of heat equation with the initial condition $w(x,y,0) = u(x,y)$.

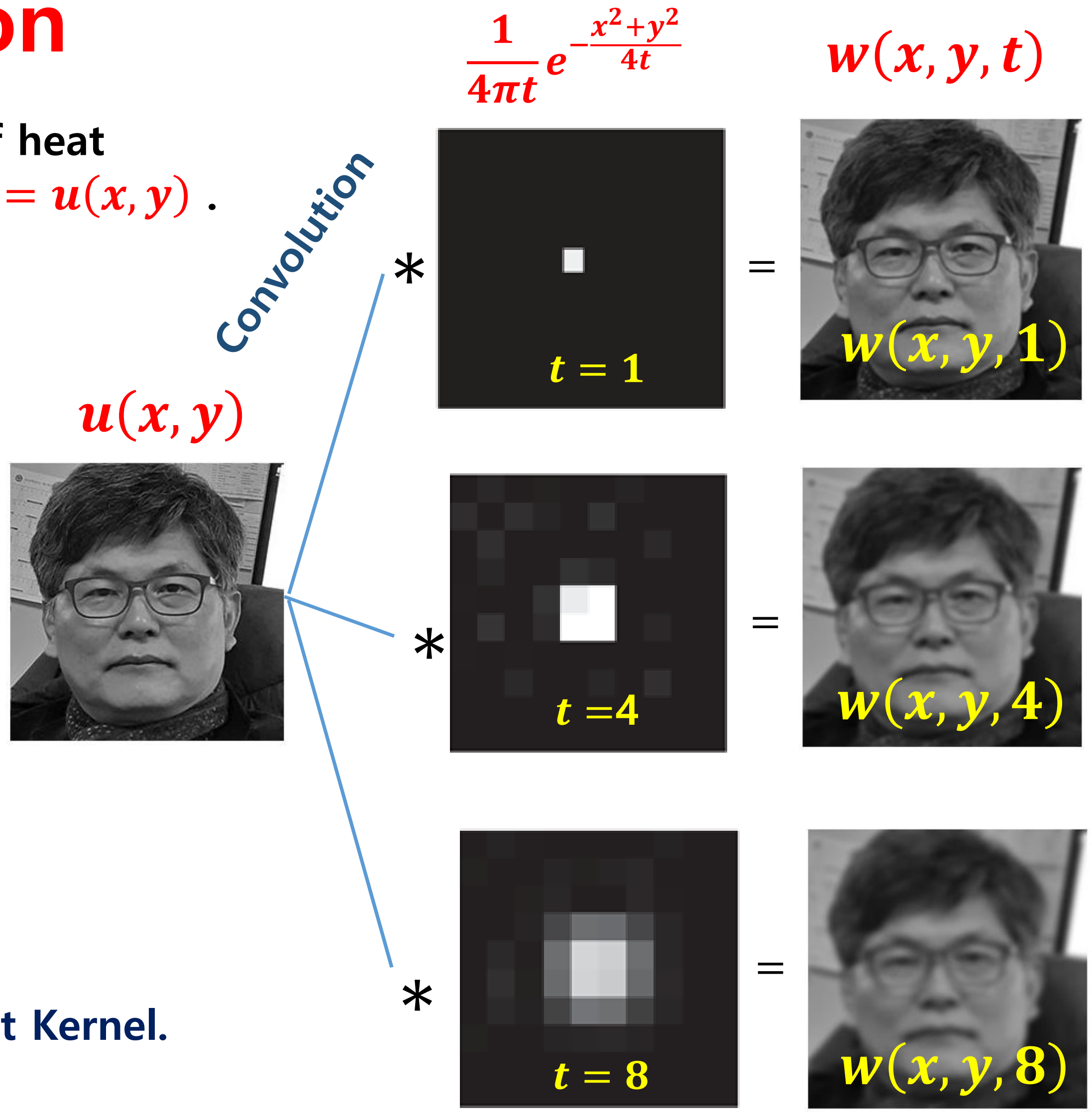
The convolution $\mathbf{w(x,y,t) = G_t * u(x,y)}$ satisfies the heat equation

$$\left[\frac{\partial}{\partial t} - \nabla^2\right] w(x,y,t) = 0$$

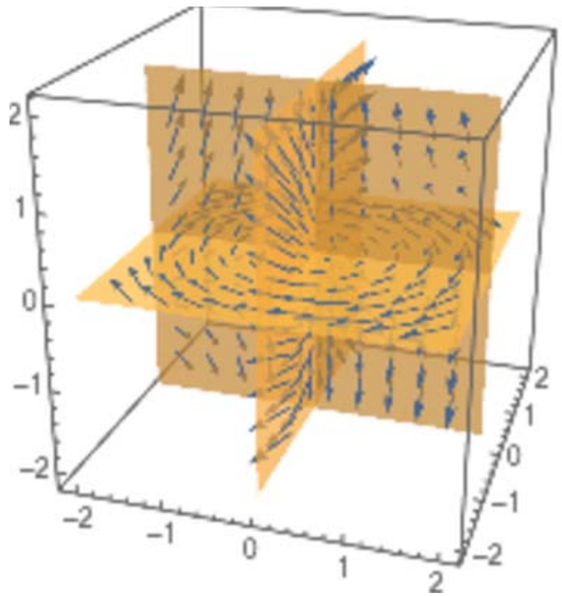
with the initial condition

$$w(x,y,0) = \lim_{t \to 0^+} w(x,y,t) = u(x,y) .$$

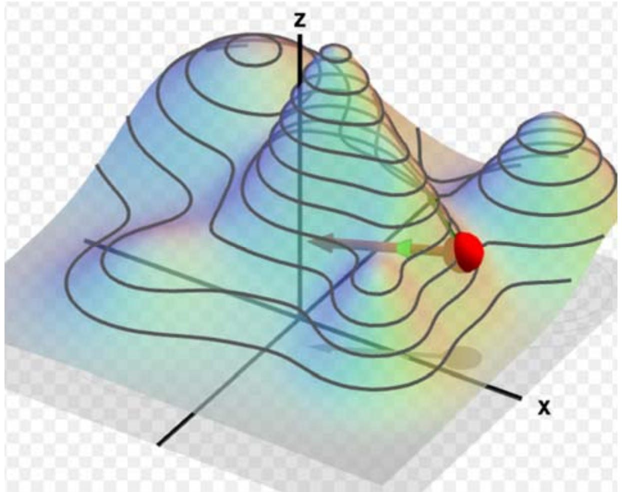
Here, $G_t(x,y) \coloneqq \frac{1}{4\pi t} e^{-\frac{x^2+y^2}{4t}}$ is the Gaussian Heat Kernel.



Review: Vector Analysis



- A scalar valued function $u(\mathbf{r})$ defined in $3-$dimensional Euclidean space $\mathbb{R}^3$ usually represent a numerical quantity depending on the location point $\mathbf{r} = (x, y, z)$ such as temperature, pressure, voltages, altitude, etc.
- $C(\mathbb{R}^n)$: the set of all functions that are continuous in $\mathbb{R}^n$.
- $C^1(\mathbb{R}^n)$: the set of all functions in $C(\mathbb{R}^n)$ with continuous derivative in $\mathbb{R}^n$.
- $C^2(\mathbb{R}^n)$: the set of all functions in $C(\mathbb{R}^n)$ with continuously twice differentiable in $\mathbb{R}^n$.



Vector Calculus: Gradient of $f \in C^2(\mathbb{R}^n)$

- Directional derivative of f at $\mathbf{x}$ in the direction $\mathbf{d}$ represents a rate of an increase in f at $\mathbf{x}$ in the direction of $\mathbf{d}$:

$$\partial_{\mathbf{d}} f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{d}) - f(\mathbf{x})}{h}.$$

- The gradient of f is a vector-valued function which points to the direction of maximum increase of f:

$$\nabla f(\mathbf{x}) = \partial_{\mathbf{d}^*} f(\mathbf{x})\mathbf{d}^* \qquad \text{where} \qquad \partial_{\mathbf{d}^*} f(\mathbf{x}) = \sup_{|\mathbf{d}|=1} \partial_{\mathbf{d}} f(\mathbf{r}).$$

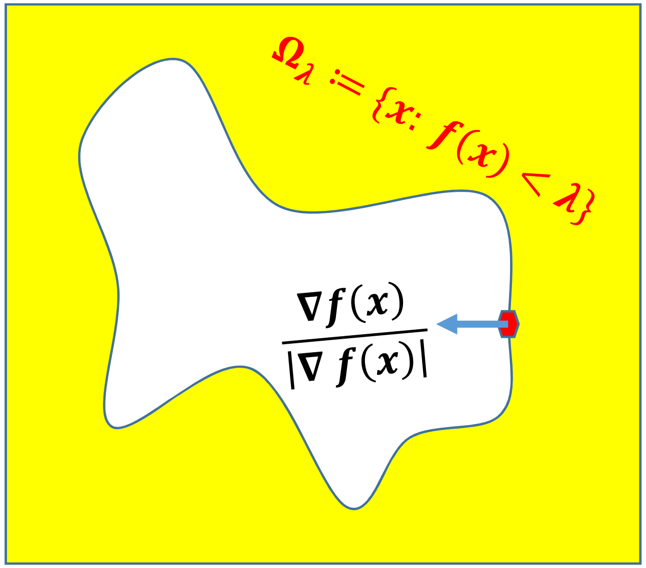
- The vector $\nabla f(\mathbf{x}_0)$ is perpendicular to the level set $\mathcal{L}_\lambda = \{\mathbf{y} \in \mathbb{R}^n : f(\mathbf{y}) = \lambda\}$ because there is no increase of f along the level set $\mathcal{L}_\lambda$.
- If $\Omega_\lambda := \{\mathbf{y} \in \mathbb{R}^n : f(\mathbf{y}) < \lambda\}$, the unit outward normal vector $\mathbf{n}(\mathbf{x}_0)$ at $\mathbf{x}_0$ on the boundary $\partial\Omega_\lambda$ is

$$\mathbf{n}(\mathbf{x}_0) = \frac{\nabla f(\mathbf{x}_0)}{|\nabla f(\mathbf{x}_0)|},$$

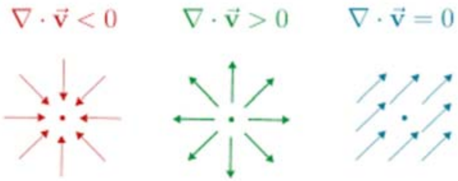
which points to the steepest ascending direction.

- The curvature along the level set $\mathcal{L}_\lambda$ is given by

$$\kappa(\mathbf{x}) := \nabla \cdot \frac{\nabla f(\mathbf{x})}{|\nabla f(\mathbf{x})|}, \quad \mathbf{x} \in \mathcal{L}_\lambda.$$



Vector Calculus: Divergence of $\mathbf{F} = (F_1, F_2, F_3) \in [C^1(\mathbb{R}^3)]^3$



- The divergence of $\mathbf{F}(\mathbf{r})$ at a point $\mathbf{r}$, written by div$\mathbf{F}$, is the net outward flux of $\mathbf{F}$ per unit volume of a ball centered at $\mathbf{r}$ as the ball shrinks to zero:

$$\mathrm{div}\mathbf{F}(\mathbf{r}) := \lim_{r\to 0} \frac{3}{4\pi r^3} \int_{\partial B_r(\mathbf{r})} \mathbf{F}(\mathbf{r}') \cdot d\mathbf{S}_{\mathbf{r}'}$$

  where $d\mathbf{S}$ is the surface element, $B_r(\mathbf{r})$ is the ball with radius r and center $\mathbf{r}$ and ∂B is the boundary of B which is a sphere.

- **Divergence theorem**. The volume integral of the divergence of a C^1-vector field $\mathbf{F} = (F_1, F_2, F_3)$ equals the total outward flux of the vector $\mathbf{F}$ through the boundary of Ω:

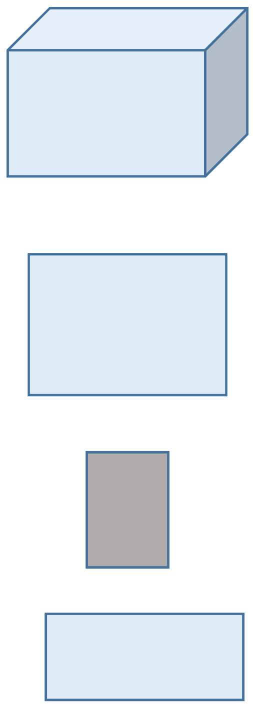
$$\int_\Omega \mathrm{div}\mathbf{F}(\mathbf{y}) d\mathbf{y} = \int_{\partial\Omega} \mathbf{F}(\mathbf{y}) \cdot d\mathbf{S}.$$

- Prove that div$\mathbf{F} = \partial_1 F_1 + \partial_2 F_2 + \partial_3 F_3$ and the divergence theorem.

Divergence theorem. The volume integral of the divergence of a C^1-vector field $\mathbf{F} = (F_1, F_2, F_3)$ equals the total outward flux of the vector $\mathbf{F}$ through the boundary of Ω:

$$\int_{\Omega} \mathrm{div}\mathbf{F}(\mathbf{y})d\mathbf{y} = \int_{\partial\Omega} \mathbf{F}(\mathbf{y}) \cdot d\mathbf{S}.$$

Proof for the special case of $\Omega = \{ (x, y, z): 0 < x < \boldsymbol{a},\ 0 < y < \boldsymbol{b}, 0 < y < \boldsymbol{c}\}$ (cuboid)



$$\iiint \partial_x F_1 + \partial_y F_2 + \partial_z F_3$$

$$= \iint F_1(\boldsymbol{a}, y, z) - F_1(\mathbf{0}, y, z) dydz$$

$$+ \iint F_2(x, \boldsymbol{b}, z) - F_2(x, \mathbf{0}, z)\, dxdz$$

$$+ \iint F_3(x, y, \boldsymbol{c}) - F_3(x, y, \mathbf{0}) dydz$$

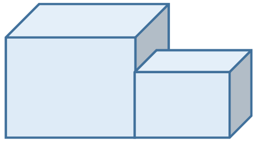
Divergence theorem. The volume integral of the divergence of a C^1-vector field $\mathbf{F} = (F_1, F_2, F_3)$ equals the total outward flux of the vector $\mathbf{F}$ through the boundary of Ω:

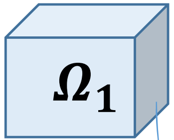
$$\int_\Omega \mathrm{div}\mathbf{F}(\mathbf{y})d\mathbf{y} = \int_{\partial\Omega} \mathbf{F}(\mathbf{y}) \cdot d\mathbf{S}.$$

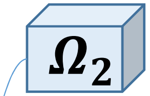
Proof for the case $\boldsymbol{\Omega} = \boldsymbol{\Omega_1} \cup \boldsymbol{\Omega_2}$ (union of two cuboids).

In general, any volume can be approximated by a union of many cuboids.

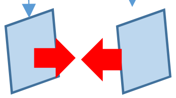
$$\iiint \nabla \cdot F \; = \; \iiint \nabla \cdot F + \iiint \nabla \cdot F$$





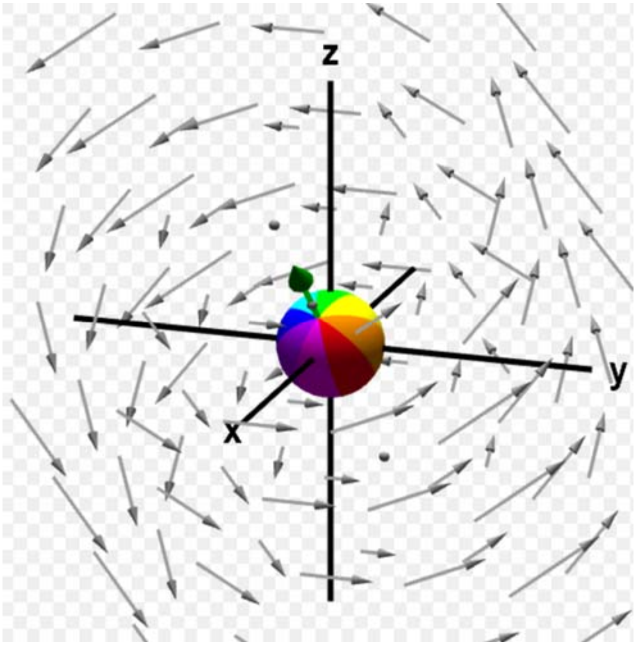


$$= \iint_{\partial\Omega_1} \boldsymbol{F} \cdot d\boldsymbol{S} + \iint_{\partial\Omega_2} \boldsymbol{F} \cdot d\boldsymbol{S} = \iint_{\partial\Omega} \boldsymbol{F} \cdot d\boldsymbol{S}$$



At the common interface, the outward normal vectors of the two volumes are opposite and therefore the interface cancel each other out.

Curl of $\mathbf{F} = (F_1, F_2, F_3) \in [C^1(\mathbb{R}^3)]^3$



- The circulation of a vector field $\mathbf{F} = (F_1, F_2, F_3)$ around a closed path C in $\mathbb{R}^3$ is defined as a scalar line integral of the vector $\mathbf{F}$ over the path C:

$$\oint_C \mathbf{F} \cdot d\mathbf{l} = \oint_C F_1 dx_1 + F_2 dx_2 + F_3 dx_3.$$

- The curl of a vector field $\mathbf{F}$, denoted by curl$\mathbf{F}$, is a vector whose magnitude is the maximum net circulation of $\mathbf{F}$ per unit area as the area shrinks to zero and whose direction is the normal direction of the area when the area is oriented to make the net circulation maximum.

- **Stokes's theorem** Let C_{area} be an open smooth surface with its boundary as a smooth contour C. The surface integral of the curl of a C^1-vector field $\mathbf{F}$ over the surface C_{area} is equal to the closed line integral of the vector $\mathbf{F}$ along the contour C:

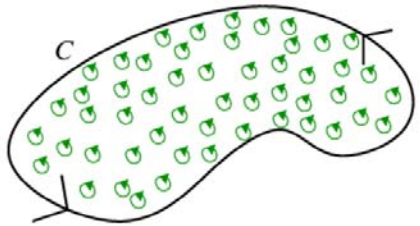
$$\int_{C_{area}} \nabla \times \mathbf{F}(\mathbf{y}) \cdot dS = \oint_C \mathbf{F}(\mathbf{y}) \cdot d\mathbf{l}$$

where

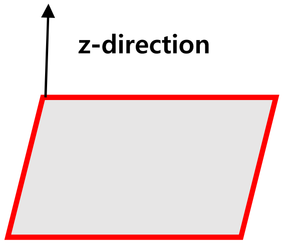
$$\mathrm{curl}\mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{vmatrix}$$

Stokes's theorem Let C_{area} be an open smooth surface with its boundary as a smooth contour C. The surface integral of the curl of a C^1-vector field $\mathbf{F}$ over the surface C_{area} is equal to the closed line integral of the vector $\mathbf{F}$ along the contour C:

$$\int_{C_{area}} \nabla \times \mathbf{F}(\mathbf{y}) \cdot dS = \oint_C \mathbf{F}(\mathbf{y}) \cdot d\mathbf{l}$$



Proof for the special case of $C_{area} = \{\,(x, y): 0 < x < \boldsymbol{a},\ 0 < y < \boldsymbol{b}\}$



$$\iint \nabla \times F \cdot dS = \iint \nabla \times F \cdot n\, dxdy = \int_0^b \int_0^a \partial_x F_2 - \partial_y F_1 dxdy$$

$$\boxed{\int_0^b F_2(a, y) - F_2(0, y) dy}$$

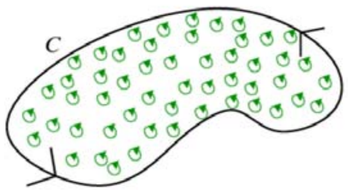
$$\boxed{\int_0^a -F_1(x, b) + F_1(x, 0) dy}$$

$$\int F \cdot d\ell$$

$$\int F \cdot d\ell$$

Stokes's theorem Let C_{area} be an open smooth surface with its boundary as a smooth contour C. The surface integral of the curl of a C^1-vector field $\mathbf{F}$ over the surface C_{area} is equal to the closed line integral of the vector $\mathbf{F}$ along the contour C:

$$\int_{C_{area}} \nabla \times \mathbf{F}(\mathbf{y}) \cdot dS = \oint_C \mathbf{F}(\mathbf{y}) \cdot d\mathbf{l}$$

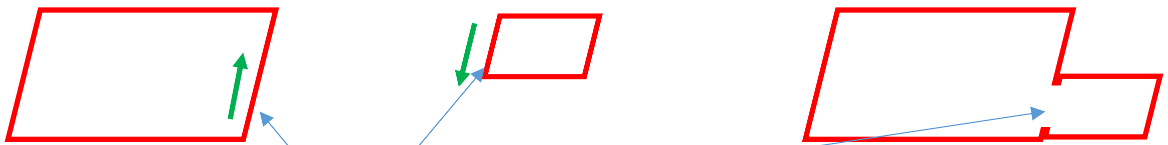


In general, any surface can be approximated by a union of rectangular patches.

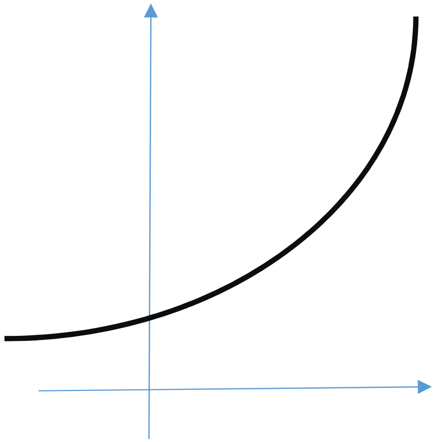
Proof for the case of $C_{area} = D_1 \cup D_2$

$$\iint \boldsymbol{\nabla} \times \boldsymbol{F} \cdot \boldsymbol{dS} \;=\; \iint \boldsymbol{\nabla} \times \boldsymbol{F} \cdot \boldsymbol{dS} + \iint \boldsymbol{\nabla} \times \boldsymbol{F} \cdot \boldsymbol{dS}$$

$$= \int \boldsymbol{F} \cdot \boldsymbol{d\ell} \;+\; \int \boldsymbol{F} \cdot \boldsymbol{d\ell} \;=\; \int \boldsymbol{F} \cdot \boldsymbol{d\ell}$$



The two line integrals are evaluated around counter-clockwise contours. At the common interface, the directions of contours are opposite and therefore they cancel each other out.



Taylor's Expansion

Taylor's expansion for $f \in C^{m+1}(\mathbb{R})$ about x is

$$f(x+h) = f(x) + f'(x)h + \cdots + \frac{f^{(m)}(x)}{m!}h^m + O(|h|^{m+1})$$

where $O(|h|^{m+1})$ is the remainder term containing $(m+1)$th order of h:

$$R_m(x,h) = \int_x^{x+h} \frac{(x-y)^m}{m!} f^{m+1}(y)dy = O(|h|^{m+1}).$$

This expansion leads to numerical differential formulae of f in various ways.

- $f'(x) = \frac{f(x+h)-f(x)}{h} + O(h)$ (forward difference).
- $f'(x) = \frac{f(x)-f(x-h)}{h} + O(h)$ (backward difference).
- $f'(x) = \frac{f(x+h)-f(x-h)}{2h} + O(h^2)$ (centered difference).
- $f'(x) = \frac{1}{12h}[f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)] + O(h^4)$.

Newton-Raphson method to find a root of $f(x) = 0$

- The Newton-Raphson method to find a root of $\boldsymbol{f(x) = 0}$ can be explained from the first order Taylor's approximation

$$\boldsymbol{f(x+h) = f(x) + hf'(x) + O(h^2)}$$

  ignoring the term $\boldsymbol{O(h^2)}$, which is negligible when $\boldsymbol{h}$ is small.
- The method is based on the approximation

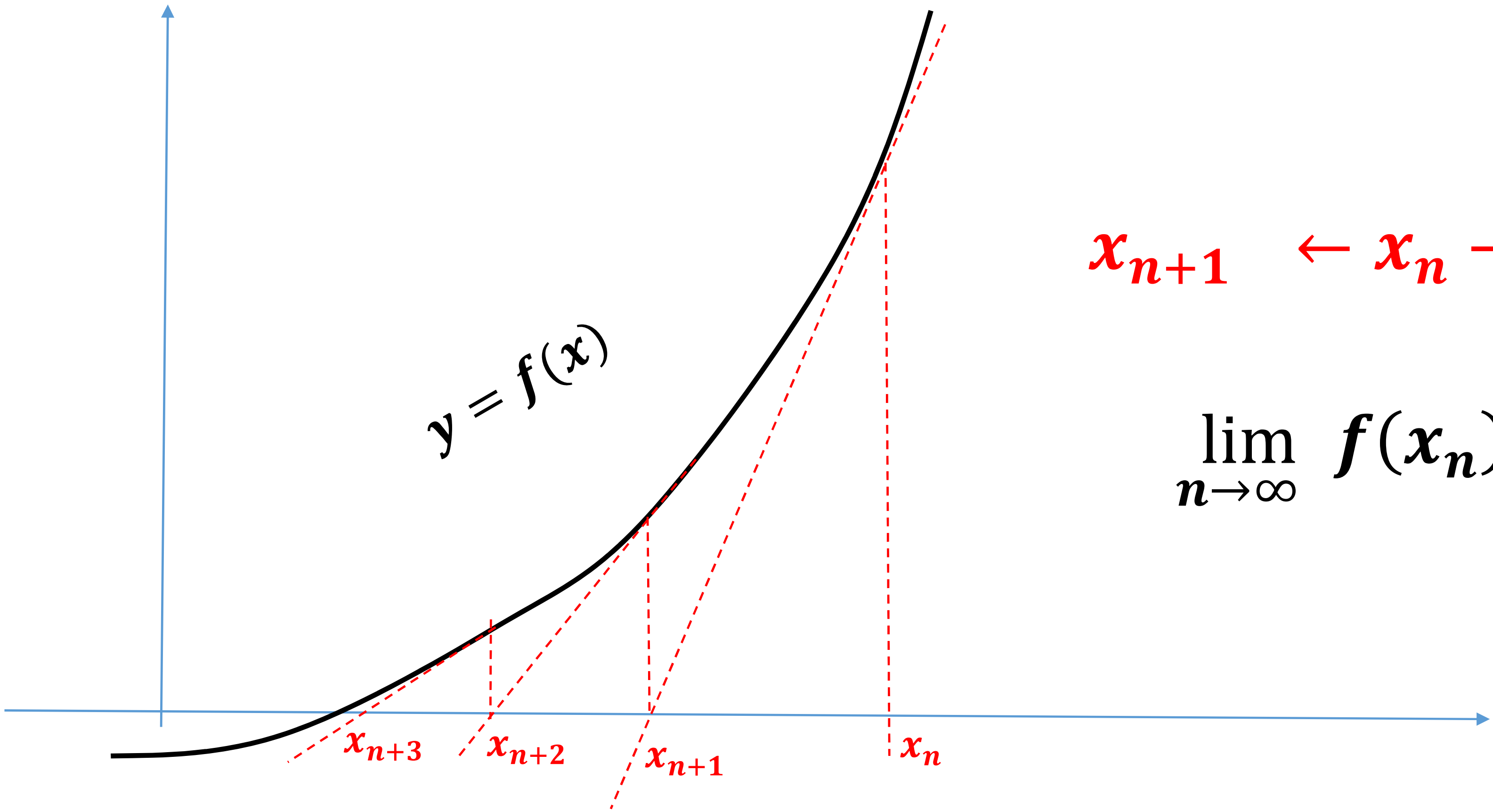
$$\boldsymbol{0 \longleftarrow f(x_{n+1}) \approx f(x_n) + h_n f'(x_n), \quad (h_n := x_{n+1} - x_n).}$$

- It starts with an initial guess $\boldsymbol{x_0}$ and generates a sequence $\boldsymbol{\{x_n\}}$ by the formula

$$\boldsymbol{x_{n+1} \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}.}$$

- Need to check the convergence issue.

Newton's Iteration Method for finding roots.



$$x_{n+1} \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\lim_{n\to\infty} f(x_n) = 0$$

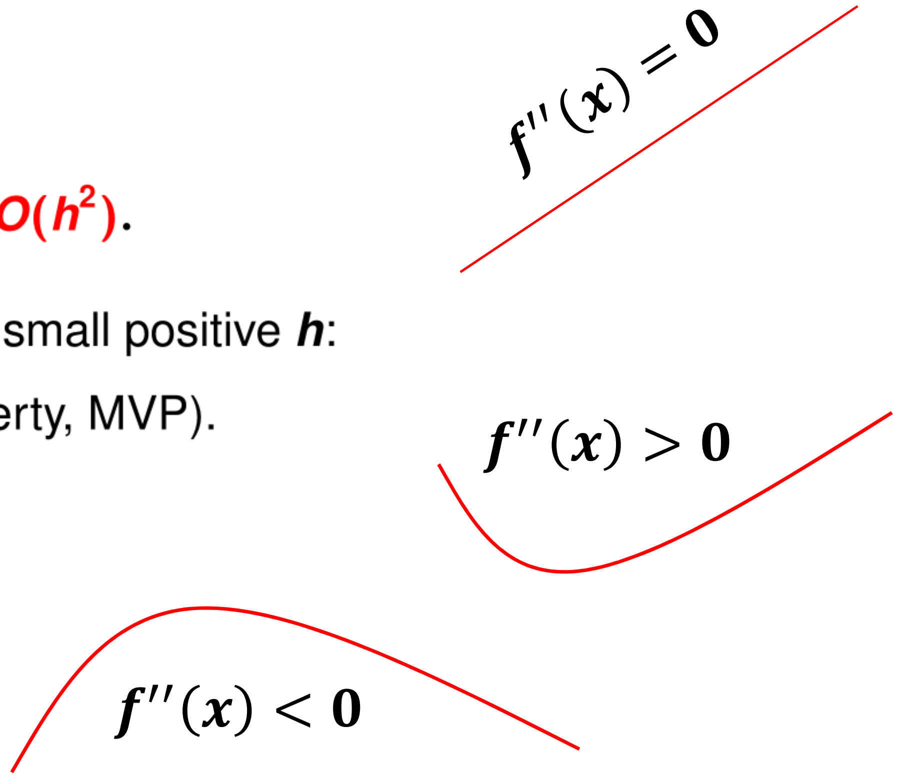
Sub-Mean, Mean & Super-Mean Value Properties

By Taylor's expansion, we approximate $f''(x)$ by

$$f''(x) = \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} + O(h^2).$$

The sign of $f''(x)$ gives local information of f for a sufficiently small positive h:

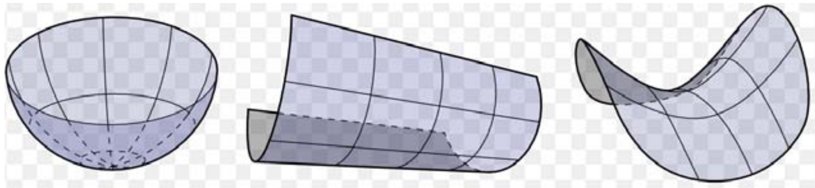
- $f''(x) = 0 \Rightarrow f(x) \approx \frac{f(x+h)+f(x-h)}{2}$ (mean value property, MVP).
- $f''(x) > 0 \Rightarrow f(x) \lesssim \frac{f(x+h)+f(x-h)}{2}$ (sub-MVP).
- $f''(x) < 0 \Rightarrow f(x) \gtrsim \frac{f(x+h)+f(x-h)}{2}$ (super-MVP).



Taylor's approximation for $f \in C^3(\mathbb{R}^n)$

$$f(\mathbf{x}+\mathbf{h}) \quad = \quad f(\mathbf{x}) + \nabla f(\mathbf{x})\cdot\mathbf{h} + \frac{1}{2}(D^2 f(\mathbf{x})\mathbf{h})\cdot\mathbf{h} + O(|\mathbf{h}|^3)$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, $\mathbf{h} = (h_1, \cdots, h_n)$ and $D^2 f(\mathbf{x})$ is the Hessian matrix:



$$D^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ & \ddots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}.$$

Local Maxima & Minima

- If f has a local maximum (minimum) at $\mathbf{x}_0$, then the Hessian matrix $D^2 f(\mathbf{x}_0)$ is negative (positive) semi-definite.
- If $D^2 f(\mathbf{x}_0)$ is negative (positive) definite and $\nabla f(\mathbf{x}_0) = 0$, then f has a local maximum (minimum) at x_0.

Proof of Taylor's approximation for $f \in C^3(\mathbb{R}^n)$

$$f(\mathbf{x}+\mathbf{h}) \quad = \quad f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{h} + \frac{1}{2}(D^2 f(\mathbf{x})\mathbf{h}) \cdot \mathbf{h} + +O(|\mathbf{h}|^3)$$

Proof.

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \textstyle\int_0^1 \frac{d}{dt} f(\mathbf{x}+t\mathbf{h})dt = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}+t\mathbf{h})h_i dt$$

$$= \textstyle\sum_{i=1}^n \int_0^1 \frac{\partial f}{\partial x_i}(\mathbf{x}+t\mathbf{h})h_i \frac{d(t-1)}{dt} dt \qquad \left(\text{Why? } \frac{d(t-1)}{dt} = 1\right)$$

$$= \textstyle\sum_{i=1}^n \left[\frac{\partial f}{\partial x_i}(\mathbf{x})h_i - \int_0^1 \frac{d}{dt}\left(\frac{\partial f}{\partial x_i}(\mathbf{x}+t\mathbf{h})h_i\right)(t-1)\, dt\right]$$

$$= \underbrace{\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x})h_i}_{\nabla f(\mathbf{x})\cdot\mathbf{h}} + \underbrace{\sum_{i,j=1}^n \int_0^1 (1-t)\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}+t\mathbf{h})h_i h_j\right)\, dt}_{R_1(\mathbf{h},\mathbf{x})}$$

Integration by part yields

$$R_1(\mathbf{h},\mathbf{x}) = \textstyle\sum_{i,j=1}^n \int_0^1 \underbrace{\frac{d}{dt}\left(-\frac{(t-1)^2}{2!}\right)}_{(1-t)} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}+t\mathbf{h})h_i h_j\right)\, dt$$

$$= \underbrace{\frac{1}{2!}\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})h_i h_j}_{(D^f(\mathbf{x})\mathbf{h})\cdot\mathbf{h}} + \underbrace{\sum_{i,j,k=1}^n \int_0^1 \frac{(t-1)^2}{2!}\left(\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}+t\mathbf{h})h_i h_j h_k\right)\, dt}_{R_2(\mathbf{h},\mathbf{x})=O(|\mathbf{h}|^3)}$$

Vector space $\mathbb{R}^n$ & Inner Product

- For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define inner product and norm:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x(j)y(j), \quad \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

- The distance (or metric) between $\mathbf{x}$ and $\mathbf{y}$ is defined by $\|\mathbf{x} - \mathbf{y}\|$, and hence $\|\mathbf{x} - \mathbf{y}\| = 0$ implies $\mathbf{x} = \mathbf{y}$.
- If $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, $\mathbf{x}$ and $\mathbf{y}$ are said to be orthogonal.
- $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ is said to be an orthonormal basis of $\mathbb{R}^n$ if

$$\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\} \qquad \& \qquad \langle \mathbf{e}_j, \mathbf{e}_i \rangle = \delta_{ij} \text{ for all } i, j$$

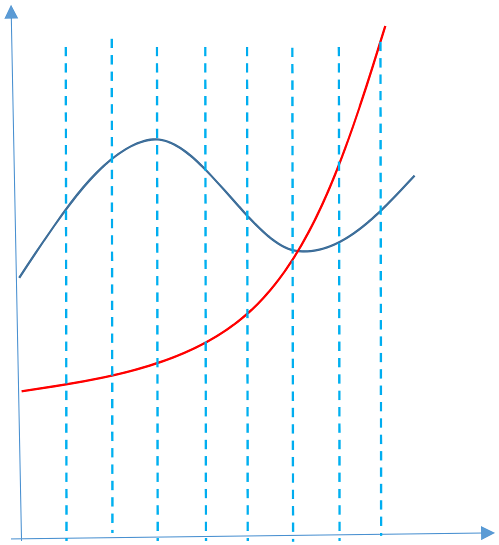
- If $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ is an orthonormal basis, then every $\mathbf{x} \in \mathbb{R}^n$ can be represented uniquely by

$$\mathbf{x} = \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j.$$

- The distance between $\mathbf{x}$ and $V_m = span\{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$ is

$$\text{dist}(\mathbf{x}, V_m) = \|\mathbf{x} - \sum_{j=1}^{m} \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j\| = \sqrt{\sum_{j=m}^{n} \langle \mathbf{x}, \mathbf{e}_j \rangle^2}$$

Vector space $C[0, 2\pi]$ & Inner Product



- Let $C[0, 2\pi]$ be the space of all continuous functions $f : [0, 2\pi] \to \mathbb{C}$.
- For $f, g \in V$, we define **the inner product**

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\overline{g(x)}\, dx$$

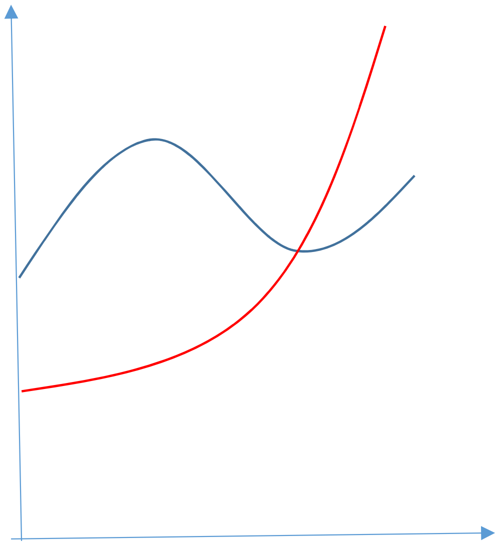
  where $\overline{g(x)}$ denotes the complex conjugate of $g(x)$.
- The above inner product can be approximated by

$$\langle f, g \rangle \approx \sum_{j=1}^{n} f(x_j)\bar{g}(x_j)\, \Delta x = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix} \cdot \begin{pmatrix} \overline{g(x_1)} \\ \vdots \\ \overline{g(x_n)} \end{pmatrix} \Delta x.$$

  where we divide the interval $[0, 2\pi]$ into n subintervals with endpoints $x_0 = 0 < x_1 < \cdots < x_n = 2\pi$ and equal width $\Delta x = \frac{2\pi}{n}$.
- Two functions f and g are said to be **orthogonal** if

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\overline{g(x)} dx = 0.$$

Hilbert space: Complete vector space equipped with an inner product and norm



- **L^2-space**. Let I be the interval $[0,1]$. We denote the set of all square integrable complex functions by $L^2(I)$, that is,

$$L^2(I) = \{ f \mid \int_I |f(x)|^2 \, dx < \infty \}.$$

- **Inner product**. For $f, g \in L^2(I)$, we define the inner product as

$$\langle f, g \rangle = \int_I f(x)\overline{g(x)} \, dx$$

where $\overline{g(x)}$ denotes the complex conjugate of $g(x)$.

- **Norm**. For $f, g \in V$, the distance between f and g can be defined by

$$\|f - g\| = \sqrt{\int_0^1 |f(x) - g(x)|^2 dx}.$$

The vector space $L^2(I)$ with the above inner product is a Hilbert space and retains features of Euclidean space.

Vector space

The followings are examples of vector space.

- $\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) \ : \ x_1, \cdots, x_n \in \mathbb{R}\}$: n-dimensional Euclidean space.
- $\mathbb{C}^n = \{(x_1, x_2, \cdots, x_n) \ : \ x_1, \cdots, x_n \in \mathbb{C}\}$.
- $C([a,b])$: the set of all complex valued functions that are continuous on the interval $[a,b]$.
- $C^1([a,b]) := \{f \in C([a,b]) \ : \ f' \in C[a,b]\}$: the set of all functions in $C([a,b])$ with continuous derivative on the interval $[a,b]$.
- $L^2((a,b)) := \{f : (a,b) \to \mathbb{R} \ : \int_a^b |f(x)|^2 dx \ < \ \infty\}$, the set of all square integrable functions on the open interval (a,b).
- $H^1((a,b)) := \{f \in L^2(a,b) : \ \int_a^b |f'(x)|^2 dx \ < \ \infty\}$.