Fourier Transform Theory

Fourier Transform:

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx}dx, \quad f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(k)e^{ikx}dk$$

Key Properties:

- Linearity: $\mathcal{F}\{af + bg\} = a\hat{f} + b\hat{g}$
- Time Shift: $\mathcal{F}\{f(x-a)\} = e^{-ika}\hat{f}(k)$
- Frequency Shift: $\mathcal{F}\{e^{iax}f(x)\} = \hat{f}(k-a)$
- Scaling: $\mathcal{F}\{f(ax)\} = \frac{1}{|a|}\hat{f}(k/a)$
- Derivatives: $\mathcal{F}\{f^{(n)}(x)\} = (ik)^n\hat{f}(k)$
- Multiplication by x^n: $\mathcal{F}\{x^n f(x)\} = i^n\hat{f}^{(n)}(k)$

Convolution Theorem:

$$(f * g)(x) = \int f(x-y)g(y)dy, \quad \mathcal{F}\{f * g\} = \hat{f}(k)\hat{g}(k)$$

Parseval Identity:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk$$

Delta Function Identities:

$$\mathcal{F}\{1\} = 2\pi\delta(k), \quad \mathcal{F}\{\delta(x-a)\} = e^{-ika}$$

Poisson Summation Formula:

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{k\in\mathbb{Z}} \hat{f}(2\pi k)$$

Standard Transform Pairs:

$$\mathcal{F}\{e^{-ax^2}\} = \sqrt{\frac{\pi}{a}}e^{-k^2/4a}$$
$$\mathcal{F}\{\text{rect}(x)\} = \text{sinc}(k/2)$$
$$\mathcal{F}\{\text{sinc}(x)\} = \text{rect}(k/2\pi)$$
$$\mathcal{F}\{H(x)\} = \pi\delta(k) + \frac{1}{ik} \quad \text{(principal value)}$$
$$\mathcal{F}\{e^{-|x|}\} = \frac{2}{1+k^2}$$

Example: Heat Equation via Fourier Transform

$$u_t = Du_{xx}, \quad u(x,0) = f(x)$$

Take FT in x: $\hat{u}_t = -Dk^2\hat{u}$

$$\Rightarrow \hat{u}(k,t) = \hat{f}(k)e^{-Dk^2t} \Rightarrow u(x,t) = \frac{1}{2\pi}\int \hat{f}(k)e^{-Dk^2t}e^{ikx}dk$$

Or via convolution:

$$u(x,t) = \int f(\xi)\frac{1}{\sqrt{4\pi Dt}}e^{-\frac{(x-\xi)^2}{4Dt}}d\xi$$

PDE Examples and Techniques

Poisson's Equation in $\mathbb{R}^2$ and $\mathbb{R}^3$

Equation: $\nabla^2 u = -f(x)$ on $\mathbb{R}^n$

- In $\mathbb{R}^2$, Green's function:

$$G(x,\xi) = -\frac{1}{2\pi}\log|x-\xi|$$
$$u(x) = \int_{\mathbb{R}^2} -\frac{1}{2\pi}\log|x-\xi|\, f(\xi)\, d^2\xi$$

- In $\mathbb{R}^3$, Green's function:

$$G(x,\xi) = \frac{1}{4\pi|x-\xi|}$$
$$u(x) = \int_{\mathbb{R}^3} \frac{1}{4\pi|x-\xi|} f(\xi)\, d^3\xi$$

Separation of Variables: Heat Equation

Problem:

$$u_t = \alpha^2 u_{xx}, \quad x \in [0,L], \quad u(0,t) = u(L,t) = 0, \quad u(x,0) = f(x)$$

Ansatz: $u(x,t) = X(x)T(t)$

$$\frac{T'}{\alpha^2 T} = \frac{X''}{X} = -\lambda \Rightarrow X_n = \sin\left(\frac{n\pi x}{L}\right), \quad T_n = e^{-\alpha^2\left(\frac{n\pi}{L}\right)^2 t}$$

General solution:

$$u(x,t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right)e^{-\alpha^2\left(\frac{n\pi}{L}\right)^2 t}$$

$$A_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi x}{L}\right)dx$$

Fourier Transform Method for PDEs

Example: Solve $u_t = Du_{xx}$, $u(x,0) = f(x)$ on $\mathbb{R}$

- Take FT in x: $\hat{u}_t = -Dk^2\hat{u}$
- Solve ODE: $\hat{u}(k,t) = \hat{f}(k)e^{-Dk^2t}$
- Inverse FT:

$$u(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(k)e^{-Dk^2t}e^{ikx}dk$$

- If $\hat{f}(k)$ known (e.g. Gaussian), result is explicit.

Green's Functions

ODE Case: Constructing $G(x,\xi)$

For a linear second-order ODE:

$$L[y] = \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = f(x)$$

We want: $y(x) = \int G(x,\xi)f(\xi)d\xi$

Step-by-step:

1. Solve the homogeneous equation $Ly = 0$ to find $y_1(x)$ and $y_2(x)$
2. Compute Wronskian: $W = y_1y_2' - y_1'y_2$

3. Construct:

$$G(x,\xi) = \begin{cases} \frac{y_1(x)y_2(\xi)}{p(\xi)W(\xi)}, & x < \xi \\ \frac{y_1(\xi)y_2(x)}{p(\xi)W(\xi)}, & x > \xi \end{cases}$$

Properties:

- $G(x,\xi)$ is continuous in x at $x = \xi$
- Discontinuity in $\partial_x G$ at $x = \xi$ satisfies:

$$\left.\frac{\partial G}{\partial x}\right|_{x=\xi^+} - \left.\frac{\partial G}{\partial x}\right|_{x=\xi^-} = \frac{1}{p(\xi)}$$

PDE Case: Whole Space

Poisson Equation: $\nabla^2 u = -f$ in $\mathbb{R}^n$

- In $\mathbb{R}^2$:

$$G(x,\xi) = -\frac{1}{2\pi}\log|x-\xi|$$

- In $\mathbb{R}^3$:

$$G(x,\xi) = \frac{1}{4\pi|x-\xi|}$$

Then:

$$u(x) = \int_{\mathbb{R}^n} G(x,\xi)f(\xi)d\xi$$

Green's Function in Bounded Domains

$$Lu = f \text{ in } \Omega, \quad u|_{\partial\Omega} = 0 \Rightarrow u(x) = \int_\Omega G(x,\xi)f(\xi)d\xi$$

If nonhomogeneous BCs, boundary terms arise:

$$u(x) = \int_\Omega Gf + \int_{\partial\Omega}\left[G\frac{\partial u}{\partial n} - u\frac{\partial G}{\partial n}\right]dS$$

Green's Identities

Identity I (Integration by parts):

$$\int_\Omega \left(u\nabla^2 v + \nabla u\cdot\nabla v\right)d\Omega = \int_{\partial\Omega} u\frac{\partial v}{\partial n}dS$$

Identity II (Symmetric form):

$$\int_\Omega (u\nabla^2 v - v\nabla^2 u)d\Omega = \int_{\partial\Omega}\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right)dS$$

Jump and Symmetry Conditions

For $LG = \delta(x-\xi)$:

- $G(x,\xi)$ is symmetric: $G(x,\xi) = G(\xi,x)$
- $G(x,\xi)$ is continuous at $x = \xi$
- Derivative jump:

$$\left.\frac{\partial G}{\partial n_x}\right|_{x=\xi^+} - \left.\frac{\partial G}{\partial n_x}\right|_{x=\xi^-} = -1$$

Sturm–Liouville Theory

General Form

A Sturm–Liouville problem is:

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + \left[\lambda w(x) - q(x)\right]y = 0 \quad \text{on } [a,b]$$

with suitable boundary conditions, often:

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0, \quad \beta_1 y(b) + \beta_2 y'(b) = 0$$

Self-Adjointness

The Sturm–Liouville operator is self-adjoint with respect to the inner product:

$$\langle f,g\rangle = \int_a^b f(x)g(x)w(x)dx$$

This implies:

- Real eigenvalues: $\lambda \in \mathbb{R}$
- Orthogonal eigenfunctions: $\langle y_m, y_n\rangle = 0$ if $m \neq n$
- Completeness: eigenfunctions form a basis for suitable function space

Integration Factor Trick

To write a second-order ODE in self-adjoint form:

Given:

$$y'' + r(x)y' + s(x)y = 0$$

Multiply by $\mu(x)$ such that:

$$\mu(x)y'' + \mu(x)r(x)y' = \frac{d}{dx}\left(\mu(x)y'\right) \Rightarrow \mu(x) = e^{\int r(x)dx}$$

This gives self-adjoint form:

$$\frac{d}{dx}\left(\mu(x)y'\right) + \mu(x)s(x)y = 0$$

Eigenfunction Expansion

If $\{y_n\}$ is the set of eigenfunctions:

$$f(x) = \sum_{n=1}^\infty a_n y_n(x), \quad a_n = \frac{\int_a^b f(x)y_n(x)w(x)dx}{\int_a^b y_n^2(x)w(x)dx}$$

Example: Classic Dirichlet Problem

$$y'' + \lambda y = 0, \quad y(0) = y(L) = 0 \Rightarrow y_n(x) = \sin\left(\frac{n\pi x}{L}\right), \quad \lambda_n = \left(\frac{n\pi}{L}\right)^2$$

These form an orthogonal basis on $[0,L]$ with weight $w(x) = 1$:

$$\int_0^L \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)dx = 0 \quad \text{for } n \neq m$$

Method of Characteristics

Purpose

Used to solve first-order PDEs of the form:

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u)$$

Idea: Along special curves called *characteristics*, the PDE becomes an ODE.

Characteristic System

Let $x(s), y(s), u(s)$ describe the characteristic curve. Then:

$$\frac{dx}{ds} = a(x, y, u), \quad \frac{dy}{ds} = b(x, y, u), \quad \frac{du}{ds} = c(x, y, u)$$

Solve these ODEs with appropriate initial conditions.

Linear Case

If a, b, c are independent of u:

$$a(x, y)u_x + b(x, y)u_y = c(x, y) \Rightarrow \text{Solve } \frac{dx}{a} = \frac{dy}{b} = \frac{du}{c}$$

Quasilinear Case

If a, b, c depend on u, characteristic curves must be found together with u:

$$\frac{dx}{ds} = a(u), \quad \frac{dy}{ds} = b(u), \quad \frac{du}{ds} = c(u)$$

Initial condition: $u(x, 0) = f(x)$ gives a curve in (x, y, u) space to start integrating from.

Well-Posedness Condition

Let initial curve be $\gamma(s) = (x(s), y(s))$. The problem is well-posed if:

$$(a, b) \cdot (\gamma'(s)) \neq 0$$

That is, the characteristic direction is not tangent to the initial curve.

Shock Formation (Quasilinear Case)

In $u_t + uu_x = 0$, characteristics:

$$\frac{dx}{dt} = u, \quad \frac{du}{dt} = 0 \Rightarrow u = f(x_0) \Rightarrow x = x_0 + f(x_0)t$$

Characteristics intersect (shock forms) when:

$$\frac{dx}{dx_0} = 1 + f'(x_0)t = 0 \Rightarrow t_s = -\frac{1}{f'(x_0)} \quad \text{if } f'(x_0) < 0$$

Example: Linear Transport

Solve:

$$u_t + 2u_x = 0, \quad u(x, 0) = \phi(x)$$

Characteristic ODEs:

$$\frac{dx}{dt} = 2 \Rightarrow x = 2t + x_0, \quad u = \phi(x_0) \Rightarrow u(x, t) = \phi(x - 2t)$$

Bessel Functions and Fourier–Bessel Series

Origin: Laplace in Polar Coordinates

Consider Laplace's equation in 2D polar coordinates:

$$\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0$$

Try separation of variables: $u(r, \theta) = R(r)\Theta(\theta)$. Substitute and divide:

$$\frac{r^2 R'' + rR'}{R} + \frac{\Theta''}{\Theta} = 0 \Rightarrow \text{Each term } = -n^2$$

Angular part:

$$\Theta(\theta) = A\cos(n\theta) + B\sin(n\theta)$$

Radial part gives Bessel's equation:

$$r^2 R'' + rR' + (r^2 - n^2)R = 0 \Rightarrow R(r) = J_n(r),\, Y_n(r)$$

Bessel's Differential Equation

$$x^2 y'' + xy' + (x^2 - n^2)y = 0$$

Solutions:

- $J_n(x)$ — Bessel function of the first kind (finite at $x = 0$)
- $Y_n(x)$ — Bessel function of the second kind (singular at $x = 0$)

Zeros and Boundary Conditions

The zeros $\alpha_{n,m}$ of $J_n(x)$ are used in boundary conditions like $u(R, \theta) = 0$.

Orthogonality of J_n

$$\int_0^R r J_n\left(\frac{\alpha_{n,m} r}{R}\right) J_n\left(\frac{\alpha_{n,k} r}{R}\right) dr = 0 \quad \text{for } m \neq k$$

Fourier–Bessel Series

If $f(r)$ is defined on $[0, R]$, expand:

$$f(r) = \sum_{m=1}^{\infty} A_m J_n\left(\frac{\alpha_{n,m} r}{R}\right)$$

$$A_m = \frac{\int_0^R r f(r) J_n\left(\frac{\alpha_{n,m} r}{R}\right) dr}{\int_0^R r\left[J_n\left(\frac{\alpha_{n,m} r}{R}\right)\right]^2 dr}$$

Example: Circular Membrane

Solution:

$$u(r, \theta, t) = \sum_{n,m} A_{n,m} J_n\left(\frac{\alpha_{n,m} r}{R}\right)\left(B_{n,m}\cos(n\theta) + C_{n,m}\sin(n\theta)\right) e^{-\lambda_{n,m}}$$

Where:

$$\lambda_{n,m} = \left(\frac{\alpha_{n,m}}{R}\right)^2$$

Legendre Polynomials and Spherical Harmonics

Laplace Equation in Spherical Coordinates

$$\nabla^2 u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial u}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial u}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 u}{\partial \phi^2}$$

Assume $u(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$ and separate variables.

Angular Separation

$\Phi(\phi)$:

$$\frac{\partial^2 \Phi}{\partial \phi^2} + m^2 \Phi = 0 \Rightarrow \Phi = e^{im\phi}$$

$\Theta(\theta)$ gives:

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta \frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right) + \left[\ell(\ell+1) - \frac{m^2}{\sin^2\theta}\right]\Theta = 0 \Rightarrow \Theta = P_\ell^m(\cos\theta)$$

Legendre's Equation

For $m = 0$, standard Legendre equation:

$$(1-x^2)\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}y}{\mathrm{d}x} + \ell(\ell+1)y = 0$$

Solutions: $P_\ell(x)$ — Legendre polynomials.

Associated Legendre Functions

For general m:

$$P_\ell^m(x) = (1-x^2)^{|m|/2}\frac{\mathrm{d}^{|m|}}{\mathrm{d}x^{|m|}}P_\ell(x)$$

Orthogonality Relations

$$\int_{-1}^{1} P_\ell(x)P_k(x)dx = 0 \quad \text{if } \ell \neq k$$

$$\int_0^\pi \int_0^{2\pi} Y_\ell^m(\theta,\phi)\overline{Y_k^n(\theta,\phi)}\sin\theta\, d\phi\, d\theta = \delta_{\ell k}\delta_{mn}$$

Spherical Harmonics

$$Y_\ell^m(\theta,\phi) = N_{\ell m}P_\ell^m(\cos\theta)e^{im\phi} \quad \text{with } N_{\ell m} = \sqrt{\frac{(2\ell+1)}{4\pi}\cdot\frac{(\ell-m)!}{(\ell+m)!}}$$

Used in solving Laplace and Helmholtz equations in spherical domains.

Application: Expanding $f(\theta,\phi)$

If f is defined on the sphere:

$$f(\theta,\phi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}Y_\ell^m(\theta,\phi)$$

$$a_{\ell m} = \int f(\theta,\phi)\overline{Y_\ell^m}(\theta,\phi)\sin\theta\, d\theta\, d\phi$$

Green's Function via Method of Images

Problem Setup: Exterior of a Sphere

Let $\Omega = \mathbb{R}^3 \setminus B(0,R)$ — the region outside a sphere of radius R.

Goal: Find Green's function $G(\vec{x},\vec{\xi})$ for Poisson's equation:

$$\nabla^2 G = -\delta(\vec{x}-\vec{\xi}), \quad \vec{x}\in\Omega$$

with boundary condition: $G = 0$ on $|\vec{x}| = R$.

Assume source point $\vec{\xi}$ is outside the sphere, i.e., $|\vec{\xi}| > R$.

Image Method Idea

Place an image charge at:

$$\vec{\xi}^* = \frac{R^2}{|\vec{\xi}|^2}\vec{\xi}$$

with strength:

$$q^* = -\frac{R}{|\vec{\xi}|}$$

Then define:

$$G(\vec{x},\vec{\xi}) = \frac{1}{4\pi|\vec{x}-\vec{\xi}|} - \frac{R}{|\vec{\xi}|}\cdot\frac{1}{4\pi|\vec{x}-\vec{\xi}^*|}$$

Properties of G

- G satisfies $\nabla^2 G = -\delta(\vec{x}-\vec{\xi})$ in Ω
- $G = 0$ on $|\vec{x}| = R$
- G is symmetric: $G(\vec{x},\vec{\xi}) = G(\vec{\xi},\vec{x})$

Use in Solving Poisson's Equation

Given $\nabla^2 u = -f$ in Ω, with $u|_{\partial\Omega} = 0$, the solution is:

$$u(\vec{x}) = \int_\Omega G(\vec{x},\vec{\xi})f(\vec{\xi})\, d^3\xi$$

Remarks

The method of images only works in highly symmetric domains (e.g. spheres, half-spaces) where image charges can be explicitly constructed to satisfy boundary conditions.

Euler–Lagrange Equation

Given a functional of the form

$$F[y] = \int_a^b f(x, y, y')\, dx,$$

a necessary condition for $y(x)$ to extremise F is the **Euler–Lagrange equation**:

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0.$$

First Integral: Eliminating y

Special Case

If the integrand $f(x, y, y')$ does **not** depend explicitly on y, then:

$$\frac{\partial f}{\partial y} = 0 \quad \Rightarrow \quad \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0.$$

This implies:

$$\frac{\partial f}{\partial y'} = \text{constant}.$$

First Integral: Eliminating x

Special Case

If the integrand $f(x, y, y')$ does **not** depend explicitly on x, then:

$$\frac{\partial f}{\partial x} = 0.$$

Define the quantity:

$$H = f - y'\frac{\partial f}{\partial y'}.$$

$$\frac{dH}{dx} = 0 \quad \Rightarrow \quad f - y'\frac{\partial f}{\partial y'} = \text{constant}.$$

Interpretation

This is analogous to conservation of the Hamiltonian in mechanics — a conserved quantity associated with translational symmetry in x.

Example: Brachistochrone Problem

$$f = \frac{\sqrt{1 + (y')^2}}{\sqrt{-y}}.$$

This has no explicit x-dependence, so

$$f - y'\frac{\partial f}{\partial y'} = \text{constant} \quad \Rightarrow \quad \frac{1}{\sqrt{1 + (y')^2}} = k\sqrt{-y}.$$

Solve for $y(x)$ using a cycloidal parametrisation.

Euler–Lagrange Equation with Constraints

Setup

Suppose we wish to extremise

$$F[y] = \int_a^b f(x, y, y')\, dx$$

subject to a constraint

$$G[y] = \int_a^b g(x, y, y')\, dx = k.$$

Lagrange Multiplier Method

Define a new functional:

$$\Phi[y; \lambda] = F[y] - \lambda G[y] = \int_a^b (f - \lambda g)\, dx.$$

Then apply the usual Euler–Lagrange procedure to the integrand $f - \lambda g$:

$$\frac{d}{dx}\left(\frac{\partial (f - \lambda g)}{\partial y'}\right) - \frac{\partial (f - \lambda g)}{\partial y} = 0.$$

Example: Dido's Problem

Maximise the area under a curve with a fixed arc length.

$$\text{Objective: } A[y] = \int_a^b y(x)\, dx, \quad \text{Constraint: } L[y] = \int_a^b \sqrt{1 + (y')^2}\, dx$$

Define:

$$h = y - \lambda\sqrt{1 + (y')^2}.$$

Use the first integral form (no x-dependence) on h to find:

$$\frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) - \frac{\partial h}{\partial y} = 0 \quad \Rightarrow \quad y - \lambda\sqrt{1 + (y')^2} = \text{constant}.$$

Multiple Dependent Variables

Setup

Let $\vec{y}(x) = (y_1(x), y_2(x), \ldots, y_n(x))$, and define the functional:

$$F[\vec{y}] = \int_a^b f(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx.$$

Euler–Lagrange System

The necessary conditions for $\vec{y}(x)$ to extremise F are:

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) - \frac{\partial f}{\partial y_i} = 0 \quad \text{for all } i = 1, \ldots, n.$$

First Integrals

- If $\partial f/\partial y_j = 0$, then:

$$\frac{\partial f}{\partial y_j'} = \text{constant}.$$

- If f has no x-dependence:

$$f - \sum_{i=1}^{n} y_i'\frac{\partial f}{\partial y_i'} = \text{constant}.$$

Multiple Independent Variables

Setup

Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function defined on a region $\mathcal{D} \subseteq \mathbb{R}^n$. The functional is:

$$F[\phi] = \int_{\mathcal{D}} f(x_1, \ldots, x_n, \phi, \nabla\phi)\, d^n x.$$

Euler–Lagrange PDE

The generalised Euler–Lagrange equation is:

$$\frac{\partial f}{\partial \phi} - \nabla \cdot \left(\frac{\partial f}{\partial(\nabla\phi)} \right) = 0.$$

In index notation (summation implied):

$$\frac{\partial f}{\partial \phi} - \partial_i \left(\frac{\partial f}{\partial(\partial_i \phi)} \right) = 0.$$

Derivation Sketch

- Perturb $\phi \to \phi + \varepsilon\eta$ with η vanishing on $\partial\mathcal{D}$.
- Use divergence theorem to move derivatives off η.
- Apply fundamental lemma $\Rightarrow$ PDE above.

Example: Laplace's Equation

Minimise potential energy:

$$F[\phi] = \int_{\mathcal{D}} \frac{1}{2}(\phi_x^2 + \phi_y^2)\, dx\, dy.$$

Then:

$$\frac{\partial f}{\partial \phi} = 0, \quad \frac{\partial f}{\partial \phi_x} = \phi_x, \quad \frac{\partial f}{\partial \phi_y} = \phi_y,$$

$$\Rightarrow \quad \phi_{xx} + \phi_{yy} = 0.$$

Higher Derivatives

Setup

$$F[y] = \int_a^b f(x, y, y', y'', \ldots, y^{(n)})\, dx.$$

Euler–Lagrange Equation (General Form)

The generalised form is:

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dx^k} \left(\frac{\partial f}{\partial y^{(k)}} \right) = 0.$$

Boundary Conditions

The variation $\eta(x)$ must satisfy:

$$\eta^{(k)}(a) = \eta^{(k)}(b) = 0 \quad \text{for } k = 0, 1, \ldots, n-1,$$

Example: Elastic Beam

Minimise bending energy:

$$F[y] = \int_0^1 (y'')^2\, dx, \quad \text{subject to } y(0) = y'(0) = 0,\ y(1) = 0,\ y'(1) = 1$$

$$\frac{\partial f}{\partial y''} = 2y'', \quad \frac{d^2}{dx^2}(2y'') = 0 \quad \Rightarrow \quad y^{(4)} = 0.$$

First Integral for n = 2

Special Case

If the integrand $f(x, y', y'')$ does **not** depend explicitly on y, then:

$$\frac{\partial f}{\partial y} = 0 \quad \Rightarrow \quad \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) - \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''} \right) = 0.$$

$$\frac{\partial f}{\partial y'} - \frac{d}{dx}\left(\frac{\partial f}{\partial y''} \right) = \text{constant}.$$

Example

Extremise:

$$F[y] = \int_0^1 (y'')^2\, dx$$

$$y(0) = y'(0) = 0, \quad y(1) = 0, \quad y'(1) = 1.$$

Since $f = (y'')^2$ and does not depend on y or y', we get:

$$\frac{d^2}{dx^2}(2y'') = 0 \quad \Rightarrow \quad y^{(4)} = 0.$$

$$y(x) = x^3 - x^2.$$

Principle of Least Action

Setup

In classical mechanics, the trajectory of a particle is found by extremising the **action functional**:

$$S[x] = \int_{t_1}^{t_2} L(x(t), \dot{x}(t), t)\, dt,$$

where $L = T - V$ is the Lagrangian: kinetic energy minus potential energy.

Euler–Lagrange Equation

The action $S[x]$ is extremised when:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} = 0.$$

This is the Euler–Lagrange equation in time-dependent form, yielding the equations of motion.

Example: Newton's Second Law

Let
$$T = \frac{1}{2}m\dot{x}^2, \quad V = V(x), \quad L = T - V.$$
Then:
$$\frac{d}{dt}(m\dot{x}) = -\frac{dV}{dx} \quad \Rightarrow \quad m\ddot{x} = -\nabla V.$$
This recovers Newton's law: **force equals mass times acceleration**.

Lagrange Multipliers

Finite-Dimensional Case

Single Constraint

To extremise $f(x, y)$ subject to $g(x, y) = c$, the condition is:
$$\nabla f = \lambda \nabla g.$$

Interpretation: At an extremum, the level set of f is tangent to the constraint surface $g = c$.

Multiple Constraints

If $g_1(x, y) = c_1$ and $g_2(x, y) = c_2$, then:
$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2.$$

Variational Case (Functional Constraints)

Single Constraint

To extremise
$$F[y] = \int_a^b f(x, y, y')\,dx \quad \text{subject to} \quad \int_a^b g(x, y, y')\,dx = c,$$
define the augmented functional:
$$H[y] = \int_a^b (f - \lambda g)\,dx.$$
Apply the Euler–Lagrange equation:
$$\frac{d}{dx}\left(\frac{\partial(f - \lambda g)}{\partial y'}\right) - \frac{\partial(f - \lambda g)}{\partial y} = 0.$$

Example: Dido's Problem

- Objective: maximise $A[y] = \int_a^b y(x)\,dx$
- Constraint: $L[y] = \int_a^b \sqrt{1 + (y')^2}dx = L$
- Lagrangian:
$$h = y - \lambda\sqrt{1 + (y')^2}$$
- Apply first-integral form (no x-dependence):
$$h - y'\frac{\partial h}{\partial y'} = \text{const} \Rightarrow y - \lambda\frac{1}{\sqrt{1 + (y')^2}} = \text{const}.$$

Legendre Transform

Definition

Let $f(x)$ be a convex, differentiable function. The **Legendre transform** of f is the function $g(p)$ defined by:
$$g(p) = \sup_x (px - f(x)).$$

If f is strictly convex and differentiable, then the supremum occurs where:
$$p = f'(x),$$
and the transform becomes:
$$g(p) = px - f(x), \quad \text{with } x = (f')^{-1}(p).$$

Inverse

The Legendre transform is **involutive**:
$$f(x) = \sup_p (px - g(p)).$$

Setup

Given a Lagrangian $L(q, \dot{q})$, define the **conjugate momentum**:
$$p = \frac{\partial L}{\partial \dot{q}}.$$

Hamiltonian

The **Hamiltonian** is defined as the Legendre transform of L with respect to $\dot{q}$:
$$H(q, p) = p\dot{q} - L(q, \dot{q}).$$
This expression must be rewritten in terms of (q, p) by solving $\dot{q}$ as a function of p.

Geometric Interpretation

- The Legendre transform replaces the variable x with the slope $p = f'(x)$ of the tangent line to the graph of f.
- The value $g(p)$ gives the vertical intercept of that tangent line:
$$g(p) = px - f(x), \quad \text{with } p = f'(x).$$
- The transform captures the geometry of a convex function in terms of its tangents, encoding all the information in terms of slope rather than position.
- **Involutive property:** This viewpoint makes clear why the transform is symmetric:
$$f(x) = \sup_p(px - g(p)).$$

Duality Identity

If $f(x)$ and $g(p)$ are Legendre transforms of each other, then:
$$f(x) + g(p) = px, \quad \text{where } p = f'(x) \text{ and } x = g'(p).$$

Interpretation: This expresses a dual pairing between the variable x and its conjugate p. The total quantity px is split into two contributions.

Young's Inequality

Statement

Let $a, b \geq 0$, and let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

with equality if and only if $a^p = b^q$.

Derivation (via Convexity)

Let $f(x) = \frac{x^p}{p}$. Then f is convex on $[0, \infty)$ because:

$$f''(x) = (p-1)x^{p-2} \geq 0 \quad \text{for } x > 0.$$

Let $a, b \geq 0$, and define:

$$f(a) + f^*(b) \geq ab,$$

where f^* is the **Legendre transform** of f. For $f(x) = \frac{x^p}{p}$, we have:

$$f^*(y) = \frac{y^q}{q}, \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1.$$

Therefore:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Equality Condition

Equality holds if and only if the slopes match, i.e.:

$$a^{p-1} = b^{q-1} \quad \Leftrightarrow \quad a^p = b^q.$$

Convexity and Optimisation

Convex Functions

A function $f : \mathbb{R} \to \mathbb{R}$ is **convex** if for all x, y and $\lambda \in [0, 1]$:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

If the inequality is strict for $x \neq y$, then f is **strictly convex**.

Second Derivative Test

If f is twice differentiable:

- $f''(x) \geq 0$ for all x implies f is convex.
- $f''(x) > 0$ for all x implies f is strictly convex.

Convex Optimisation (Single-variable)

Global Minima

If f is convex on an interval, then:

- Any **local minimum is a global minimum**.
- If f is **strictly convex**, the global minimum is **unique**.

First-order Condition

If f is differentiable, then:

$$f'(x^*) = 0 \quad \Rightarrow \quad x^* \text{ is a local extremum.}$$

If f is convex, this is a global minimum.

Geometric Interpretation

For convex f, the graph lies above all its tangents:

$$f(y) \geq f(x) + f'(x)(y - x).$$

This inequality characterises convexity and forms the basis of duality theory.

Convexity in Variational Calculus

Let $F[y] = \int_a^b f(x, y, y')\, dx$ be a functional.

Second Variation

To assess whether F is minimised at y, consider the second variation:

$$\delta^2 F[y](\eta) = \int_a^b \left(\eta^2 \frac{\partial^2 f}{\partial y^2} + 2\eta\eta' \frac{\partial^2 f}{\partial y \partial y'} + (\eta')^2 \frac{\partial^2 f}{\partial y'^2} \right) dx.$$

Sufficient Condition for a Minimum

If $\delta^2 F[y](\eta) > 0$ for all admissible $\eta \neq 0$, then F is **strictly convex** and y is a **strict local minimiser**.

Second Variation and Minimisation Criteria

Second Variation

Let $F[y] = \int_a^b f(x, y, y')\, dx$ be a functional, and suppose y is a critical point (i.e. satisfies the Euler–Lagrange equation).

The **second variation** is defined by:

$$\delta^2 F[y](\eta) = \int_a^b \left(\eta^2 \frac{\partial^2 f}{\partial y^2} + 2\eta\eta' \frac{\partial^2 f}{\partial y \partial y'} + (\eta')^2 \frac{\partial^2 f}{\partial y'^2} \right) dx.$$

Interpretation

This arises from expanding $F[y + \varepsilon\eta]$ to second order in ε:

$$F[y + \varepsilon\eta] = F[y] + \varepsilon \delta F[y](\eta) + \frac{\varepsilon^2}{2} \delta^2 F[y](\eta) + \mathcal{O}(\varepsilon^3).$$

Sufficient Condition for a Minimum

Suppose y satisfies the Euler–Lagrange equation for the functional

$$F[y] = \int_a^b f(x, y, y')\, dx,$$

and consider perturbations $y + \varepsilon\eta$ with $\eta(a) = \eta(b) = 0$.

If the second variation satisfies:

$$\delta^2 F[y](\eta) > 0 \quad \text{for all admissible } \eta \neq 0,$$

then y is a **strict local minimiser** of F.

Analogy

This is analogous to the second derivative test in one-variable calculus:

$$f'(x) = 0 \quad \text{and} \quad f''(x) > 0 \quad \Rightarrow \quad \text{local minimum.}$$

Convexity and Functional Minimisation

If the integrand $f(x, y, y')$ is convex in (y, y') for each x, then the functional

$$F[y] = \int_a^b f(x, y, y')\, dx$$

is convex on the space of admissible functions.

Implication

If y satisfies the Euler–Lagrange equation and f is convex in (y, y'), then y is a **global minimiser** of F.

Strict Convexity

If f is strictly convex in (y, y'), then any solution y is a **strict global minimiser**.

Example: Second Variation for $F[y] = \int (y')^2\, dx$

Let

$$F[y] = \int_a^b (y')^2\, dx.$$

First Variation

The Euler–Lagrange equation gives:

$$\frac{d}{dx}(2y') = 0 \quad \Rightarrow \quad y'' = 0.$$

Solutions are straight lines: $y(x) = ax + b$.

Second Variation

Perturb y by η with $\eta(a) = \eta(b) = 0$:

$$\delta^2 F[y](\eta) = \int_a^b 2(\eta')^2\, dx > 0 \quad \text{for all } \eta \neq 0.$$

Hamiltonian Formulation

From Lagrangian to Hamiltonian

Given a Lagrangian $L(q, \dot{q})$, define the **conjugate momentum**:

$$p = \frac{\partial L}{\partial \dot{q}}.$$

Assuming this relation can be inverted to write $\dot{q}$ in terms of p, the **Hamiltonian** is defined via a Legendre transform:

$$H(q, p) = p\dot{q} - L(q, \dot{q}).$$

Hamilton's Equations

The dynamics are governed by the system:

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}.$$

Fermat's Principle

Light travels between two points along the path that minimises travel time.

Time Functional

In a medium with variable speed $v(x, y)$ (or refractive index $n(x, y) = 1/v$):

$$T[y] = \int \frac{\sqrt{1 + (y')^2}}{v(x, y)}\, dx = \int n(x, y)\sqrt{1 + (y')^2}\, dx.$$

Euler–Lagrange Equation

Apply the Euler–Lagrange equation to:

$$f(x, y, y') = n(x, y)\sqrt{1 + (y')^2}.$$

Special Case: Snell's Law

If n jumps across a boundary (e.g. piecewise constant), minimising $T[y]$ leads to:

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2},$$

i.e., Snell's Law.

Surface Area Minimisation (Soap Film)

Consider a surface of revolution about the x-axis generated by a curve $y(x)$.

Surface Area Functional

$$A[y] = \int_a^b 2\pi y\sqrt{1 + (y')^2}\, dx.$$

Euler–Lagrange Equation

For

$$f(y, y') = 2\pi y\sqrt{1 + (y')^2},$$

the Euler–Lagrange equation gives the minimal surface of revolution:

$$\frac{d}{dx}\left(\frac{y'}{\sqrt{1 + (y')^2}}\right) = \frac{1}{y} \cdot \frac{1}{\sqrt{1 + (y')^2}}.$$

Solution: The minimising surface is a **catenoid**:

$$y(x) = a\cosh\left(\frac{x - x_0}{a}\right).$$

Bernoulli's Equation

Steady Flow (Inviscid, Incompressible, Irrotational)

For a steady, inviscid, incompressible, and irrotational flow:

$$\frac{1}{2}\rho u^2 + p + \rho\Phi = \text{constant along a streamline.}$$

Interpretation: Kinetic + pressure + potential energy per unit volume is conserved.

Special Cases

- $\Phi = gz$ for gravity potential $\Rightarrow$ classic Bernoulli:

$$\frac{1}{2}u^2 + \frac{p}{\rho} + gz = \text{const.}$$

- Applies globally only for irrotational flow. Otherwise, only along streamlines.

Unsteady Bernoulli Equation

Assume unsteady, irrotational, incompressible, inviscid flow. The velocity field is a potential flow:

$$\vec{u} = \nabla\phi.$$

From Euler's Equation:

$$\frac{\partial \vec{u}}{\partial t} + \nabla\left(\frac{1}{2}u^2 + \frac{p}{\rho} + \Phi\right) = 0.$$

Integrate in space:

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}u^2 + \frac{p}{\rho} + \Phi = f(t),$$

where $f(t)$ is a function of time only.

Standard form:

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}u^2 + \frac{p}{\rho} + \Phi = \text{const (in space).}$$

Derivation (Steady Case)

$$(\vec{u}\cdot\nabla)\vec{u} = -\frac{1}{\rho}\nabla p - \nabla\Phi.$$

Use vector identity:

$$(\vec{u}\cdot\nabla)\vec{u} = \nabla\left(\frac{1}{2}u^2\right) - \vec{u}\times(\nabla\times\vec{u}).$$

If flow is irrotational: $\nabla\times\vec{u} = 0$, so:

$$\nabla\left(\frac{1}{2}u^2 + \frac{p}{\rho} + \Phi\right) = 0.$$

Integrate

Streamfunction and Velocity Potential

Streamfunction ψ (2D Incompressible Flow)

Defined such that:

$$u = \frac{\partial \psi}{\partial y}, \qquad v = -\frac{\partial \psi}{\partial x}.$$

Properties:

- Automatically satisfies incompressibility: $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0$.
- Lines of constant ψ are streamlines.

Velocity Potential ϕ (Irrotational Flow)

Defined such that:

$$\vec{u} = \nabla\phi \quad \text{(i.e., } u = \frac{\partial \phi}{\partial x},\ v = \frac{\partial \phi}{\partial y}\text{).}$$

Properties:

- Flow is irrotational: $\boldsymbol{\nabla}\times\vec{u} = 0$.
- ϕ satisfies Laplace's equation in incompressible flow: $\nabla^2\phi = 0$.

Poiseuille and Couette Flow

Poiseuille Flow (Pressure-Driven Flow)

Description: Steady, fully-developed, incompressible viscous flow between two stationary parallel plates (or inside a circular pipe), driven by a pressure gradient.

Assumptions:

- Unidirectional flow: $\vec{u} = u(y)\,\hat{\mathbf{i}}$
- $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial z} = 0$
- $\frac{\partial p}{\partial y} = 0$, $\frac{\partial p}{\partial z} = 0$

s

$$u(y) = \frac{1}{2\mu}\frac{\mathrm{d}p}{\mathrm{d}x}$$

Boundary Conditions:

- No-slip: $u(0) = 0$, $u(H) = 0$

Profile: Parabolic velocity profile with maximum at channel center.

Couette Flow (Shear-Driven Flow)

Description: Steady viscous flow between two parallel plates, with the bottom plate stationary and the top plate moving at velocity U.

Assumptions:

- Unidirectional flow: $\vec{u} = u(y)\,\hat{\mathbf{i}}$
- No pressure gradient: $\frac{\partial p}{\partial x} = 0$

$$u(y) = Ay + B$$

Boundary Conditions:

- $u(0) = 0$ (stationary plate)
- $u(H) = U$ (moving plate)

Solution:

$$u(y) = \frac{U}{H}y$$

Profile: Linear velocity profile.

Vorticity Equation and 2D Incompressible Flow

Vorticity Transport Equation

For an **incompressible, inviscid flow**, taking the curl of the Euler equation yields:

$$\frac{D\boldsymbol{\omega}}{Dt} = (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{u}.$$

In **2D incompressible flow**, the right-hand side vanishes (since $\boldsymbol{\omega}$ is perpendicular to the plane of motion), so:

$$\frac{D\omega}{Dt} = 0.$$

Implications in 2D Incompressible Flow

- $\omega =$ constant along trajectories.
- If the initial flow is **irrotational** ($\omega = 0$), it remains irrotational.
- If initially **vortical**, the distribution of vorticity is merely **advected** by the flow.

Relation to Streamfunction

For 2D incompressible flows, define a **streamfunction** ψ such that:

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial x}.$$

Then the vorticity becomes:

$$\omega = -\nabla^2 \psi.$$

Surface Wave Theory

Boundary Conditions

- **Bottom Boundary Condition (Impenetrability)**:

  $$w = 0 \quad \text{at } z = -h,$$

  where w is the vertical velocity and h is the fluid depth.
- **Kinematic Boundary Condition (Free Surface)**:

  $$\frac{\partial \eta}{\partial t} + u\frac{\partial \eta}{\partial x} = w \quad \text{at } z = \eta(x,t),$$

  where $\eta(x,t)$ is the surface elevation. For linear theory, this is often approximated as:

  $$\frac{\partial \eta}{\partial t} = w \quad \text{at } z = 0.$$

- **Dynamic Boundary Condition (Free Surface Pressure)**:

  $$p = p_{\text{atm}} \quad \text{at } z = \eta(x,t).$$

  In linearised theory:

  $$p = 0 \quad \text{at } z = 0,$$

  assuming constant atmospheric pressure and neglecting surface tension.

Dispersion Relation

For inviscid, incompressible, irrotational flow with a free surface under gravity, the dispersion relation is:

$$\omega^2 = gk\tanh(kh),$$

Potential Flow Solutions

Flow Around a Cylinder

Velocity potential:

$$\phi = U\left(r + \frac{a^2}{r}\right)\cos\theta$$

Streamfunction:

$$\psi = U\left(r - \frac{a^2}{r}\right)\sin\theta$$

Velocity components:

$$u_r = U\left(1 - \frac{a^2}{r^2}\right)\cos\theta, \quad u_\theta = -U\left(1 + \frac{a^2}{r^2}\right)\sin\theta$$

Pressure (via Bernoulli):

$$\frac{p}{\rho} + \frac{1}{2}|\vec{u}|^2 = \text{const}$$

Flow Around a Cylinder with Circulation

Velocity potential:

$$\phi = U\left(r + \frac{a^2}{r}\right)\cos\theta + \frac{\Gamma}{2\pi}\theta$$

Streamfunction:

$$\psi = U\left(r - \frac{a^2}{r}\right)\sin\theta - \frac{\Gamma}{2\pi}\ln r$$

Lift (Kutta–Joukowski theorem):

$$L = \rho U \Gamma$$

Stagnation point shift: Circulation displaces the stagnation points off the horizontal axis.

Flow Around a Sphere

Velocity potential:

$$\phi = U\left(r + \frac{a^3}{2r^2}\right)\cos\theta$$

Velocity field:

$$u_r = U\left(1 - \frac{a^3}{r^3}\right)\cos\theta, \quad u_\theta = -U\left(1 + \frac{a^3}{2r^3}\right)\sin\theta$$

Result: Zero drag (D'Alembert's paradox), symmetric pressure distribution.

Added Mass

Definition: Effective mass of fluid accelerated along with the body.

Sphere moving in fluid: Added mass is:

$$m_{\text{added}} = \frac{1}{2}\rho\left(\frac{4}{3}\pi a^3\right)$$

Cylinder (2D) moving in fluid:

$$m_{\text{added}} = \pi\rho a^2$$

Kelvin's Circulation Theorem

Statement

In an inviscid, barotropic fluid with conservative body forces, the circulation around a material (fluid-following) loop is conserved:

$$\frac{D\Gamma}{Dt} = 0, \quad \text{where } \Gamma = \oint_{\mathcal{C}(t)} \vec{u} \cdot d\vec{x}.$$

Rotating Fluids and Shallow/Deep Water

Rotating Fluids and the Coriolis Force

In a frame rotating with angular velocity $\boldsymbol{\Omega}$, the Navier–Stokes equation becomes:

$$\rho\frac{D\vec{u}}{Dt} + 2\rho\,\boldsymbol{\Omega} \times \vec{u} = -\nabla p + \rho\vec{g} + \mu\nabla^2\vec{u},$$

where $2\boldsymbol{\Omega} \times \vec{u}$ is the **Coriolis force**.

- In large-scale geophysical flows, this term becomes dominant.
- Leads to phenomena such as **geostrophic balance**:

$$-\frac{1}{\rho}\nabla p = 2\boldsymbol{\Omega} \times \vec{u}.$$

Shallow Water Equations

Assume horizontal length scales $\gg$ vertical scale. Let $h(x,t)$ be fluid height and $\vec{u}(x,t)$ horizontal velocity.

- **Mass conservation:**

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\vec{u}) = 0$$

- **Momentum conservation:**

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla)\vec{u} = -g\nabla h$$

Linearised Shallow Water Equations

Linearise about rest state with small perturbations. Then:

$$\frac{\partial \eta}{\partial t} + H\nabla \cdot \vec{u} = 0, \qquad \frac{\partial \vec{u}}{\partial t} = -g\nabla\eta$$

Leads to the wave equation:

$$\frac{\partial^2 \eta}{\partial t^2} = c^2\nabla^2\eta, \quad \text{with } c = \sqrt{gH}.$$

Deep Water Waves

For waves where water depth $h \to \infty$, the dispersion relation becomes:

$$\omega^2 = gk$$

with:

$$\lambda = \frac{2\pi}{k}, \qquad c = \frac{\omega}{k} = \sqrt{\frac{g}{k}}, \qquad c_g = \frac{d\omega}{dk} = \frac{1}{2}c.$$

Key features:

- Phase speed decreases with increasing k (shorter wavelengths move slower).
- Group velocity is half the phase speed.

Boundary Conditions in Fluid Dynamics

1. Free Surface

Let the free surface be given by $z = \eta(x, y, t)$.

Kinematic Boundary Condition

No fluid crosses the surface:

$$\frac{D\eta}{Dt} = \frac{\partial \eta}{\partial t} + u\frac{\partial \eta}{\partial x} + v\frac{\partial \eta}{\partial y} = w \quad \text{on } z = \eta(x, y, t)$$

Linearised:

$$\frac{\partial \eta}{\partial t} = w \quad \text{at } z = 0$$

Dynamic Boundary Condition

Balance of pressure at the surface:

$$p = p_{\text{atm}} \quad \text{on } z = \eta(x, y, t)$$

2. Rigid Boundary

Let the rigid boundary be at $z = h(x, y)$ (or simply $z = 0$).

Kinematic Boundary Condition

No flow through the wall:

$$\vec{u} \cdot \hat{n} = 0 \quad \text{(normal component vanishes)}$$

If the wall is flat at $z = 0$, then:

$$w = 0 \quad \text{on } z = 0$$

More generally: if the surface is $z = h(x, y)$,

$$w = u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} \quad \text{on } z = h(x, y)$$

Dynamic Boundary Condition

Not generally required — the wall can exert any normal stress.

3. Interface Between Two Fluids

Let two fluids (densities $\rho_1,\ \rho_2$) meet at a surface $z = \eta(x, y, t)$.

Kinematic Condition

The interface must move with the fluid:

$$\left.\frac{D\eta}{Dt}\right|_{\text{fluid 1}} = \left.\frac{D\eta}{Dt}\right|_{\text{fluid 2}} = w \quad \text{on interface}$$

Dynamic Condition

Normal stress must be continuous **No surface tension:**

$$p_1 = p_2 \quad \text{across the interface}$$

Tangential Stress Condition (Viscous case)

If viscosity is included, continuity of tangential stress is also required:

$$\left[\mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right]_{\text{across interface}} = 0 \quad \text{and similarly for } v$$

Conservation of Momentum (Integral Form)

Let $\mathcal{V}(t)$ be a control volume with boundary $\partial\mathcal{V}(t)$, velocity field $\vec{u}$, pressure p, and external body force per unit mass $\vec{f}$ (e.g., gravity).

Statement

$$\frac{d}{dt}\int_{\mathcal{V}(t)} \rho\vec{u}\, dV + \int_{\partial\mathcal{V}(t)} \rho\vec{u}(\vec{u}\cdot\hat{n})\, dS = \int_{\partial\mathcal{V}(t)} \boldsymbol{\sigma}\cdot\hat{n}\, dS + \int_{\mathcal{V}(t)} \rho\vec{f}\, dV$$

Terms

- $\int_{\mathcal{V}} \rho\vec{u}\, dV$: total momentum in the control volume.
- $\int_{\partial\mathcal{V}} \rho\vec{u}(\vec{u}\cdot\hat{n})\, dS$: momentum flux across the boundary.
- $\int_{\partial\mathcal{V}} \boldsymbol{\sigma}\cdot\hat{n}\, dS$: surface forces (normal + viscous stresses).
- $\int_{\mathcal{V}} \rho\vec{f}\, dV$: body forces (e.g. gravity).

Simplification (Inviscid Flow)

If the fluid is inviscid, then the stress tensor reduces to $-p\boldsymbol{I}$, so:

$$\boldsymbol{\sigma}\cdot\hat{n} = -p\hat{n} \quad \Rightarrow \quad \int_{\partial\mathcal{V}} \boldsymbol{\sigma}\cdot\hat{n}\, dS = -\int_{\partial\mathcal{V}} p\hat{n}\, dS$$

Interfacial Wave Dispersion Relation

Consider two immiscible, incompressible fluid layers of densities ρ_1 (upper) and ρ_2 (lower), with $\rho_2 > \rho_1$. The interface lies at $z = 0$ in the undisturbed state. We assume irrotational motion and linearise the governing equations.

Velocity Potentials

Let the velocity potentials be:

$$\phi_1(x,z,t) = A_1 e^{kz} e^{i(kx-\omega t)} \quad \text{for } z < 0 \quad \text{(upper fluid)},$$

$$\phi_2(x,z,t) = A_2 e^{-kz} e^{i(kx-\omega t)} \quad \text{for } z > 0 \quad \text{(lower fluid)}.$$

Boundary Conditions at the Interface

- **Kinematic condition (both fluids):**

$$\frac{\partial\phi_1}{\partial z} = \frac{\partial\eta}{\partial t}, \quad \frac{\partial\phi_2}{\partial z} = \frac{\partial\eta}{\partial t} \quad \text{at } z = 0$$

- **Dynamic condition (pressure continuity):**

$$\rho_1\left(\frac{\partial\phi_1}{\partial t} + g\eta\right) = \rho_2\left(\frac{\partial\phi_2}{\partial t} + g\eta\right) \quad \text{at } z = 0$$

Dispersion Relation

Combining the conditions yields the dispersion relation for interfacial gravity waves:

$$\omega^2 = gk\frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}.$$

Lift on a Cylinder (Pressure Integration)

Setup

Consider steady, incompressible, irrotational flow with circulation Γ past a cylinder of radius a. The velocity field in polar coordinates (r, θ) is:

$$v_r = U\left(1 - \frac{a^2}{r^2}\right)\cos\theta, \quad v_\theta = -U\left(1 + \frac{a^2}{r^2}\right)\sin\theta + \frac{\Gamma}{2\pi r}.$$

On the surface $r = a$:

$$v_r = 0, \quad v_\theta = -2U\sin\theta + \frac{\Gamma}{2\pi a}.$$

Pressure from Bernoulli

Apply Bernoulli's equation (assuming constant pressure at infinity):

$$p = p_\infty + \frac{1}{2}\rho(U^2 - v^2),$$

where $v^2 = v_\theta^2$ on the surface.

Lift Force

The lift is the vertical component of pressure force:

$$L = -\int_0^{2\pi} p(\theta)\cdot a\sin\theta\, d\theta.$$

Substitute v_θ and compute:

$$v_\theta = -2U\sin\theta + \frac{\Gamma}{2\pi a}, \quad v^2 = \left(-2U\sin\theta + \frac{\Gamma}{2\pi a}\right)^2.$$

Only the cross term in v^2 contributes to the integral:

$$L = \rho U\int_0^{2\pi}\left(2U\sin\theta\cdot\frac{\Gamma}{2\pi a}\right)a\sin\theta\, d\theta = \rho U\Gamma.$$

Continuity Equation in Quantum Mechanics

The continuity equation expresses the local conservation of probability:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$$

where:
- $\rho = |\psi|^2$ is the **probability density**
- $\mathbf{j}$ is the **probability current density**

Derivation from the Schrödinger Equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \psi + V\psi$$

Multiply by ψ^* and subtract the complex conjugate of the equation multiplied by ψ:

$$\psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t} = \frac{\hbar}{2im}\nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*)$$

$$\frac{\partial}{\partial t}|\psi|^2 + \nabla \cdot \left(\frac{\hbar}{2im}(\psi^* \nabla \psi - \psi \nabla \psi^*) \right) = 0$$

Thus, the probability current density is:

$$\mathbf{j} = \frac{\hbar}{2im}(\psi^* \nabla \psi - \psi \nabla \psi^*)$$

Integral Form

The integral form over a region Ω with boundary $\partial\Omega$:

$$\frac{d}{dt}\int_\Omega |\psi|^2 \, dV = -\int_{\partial\Omega} \mathbf{j} \cdot \hat{\mathbf{n}} \, dS$$

The continuity equation ensures that probability is conserved. The total probability inside a region changes only due to the flux of probability current across the boundary.

Angular Momentum Operator

The **orbital angular momentum operator** is defined as:

$$\vec{L} = \vec{r} \times \vec{p} \quad \Rightarrow \quad L_i = \epsilon_{ijk} x_j (-i\hbar \partial_k) = -i\hbar \epsilon_{ijk} x_j \partial_k.$$

Cartesian Components

$$L_x = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right), \quad L_y = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right),$$

$$L_z = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right).$$

Commutation Relations

$$[L_i, L_j] = i\hbar \epsilon_{ijk} L_k.$$

$$[L^2, L_i] = 0 \quad \text{for all } i.$$

$$[L_i, x_j] = i\hbar \epsilon_{ijk} x_k, \quad [L_i, p_j] = i\hbar \epsilon_{ijk} p_k.$$

Eigenfunctions in Spherical Coordinates

For central potentials, the Schrödinger equation reduces to an equation involving L^2.
- The eigenfunctions of L^2 and L_z are the spherical harmonics:

$$L^2 Y_{\ell m} = \hbar^2 \ell(\ell+1) Y_{\ell m}, \quad L_z Y_{\ell m} = \hbar m Y_{\ell m}.$$

- $\ell \in \mathbb{Z}_{\geq 0}$, and $m = -\ell, -\ell+1, \ldots, \ell$.

Heisenberg Uncertainty Principle

Statement

For any pair of observables A and B, represented by Hermitian operators $\hat{A}$, $\hat{B}$, the uncertainty relation is:

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i}\langle [\hat{A}, \hat{B}] \rangle \right)^2,$$

where $\sigma_A^2 = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle$ is the variance of A, and similarly for σ_B^2.

Cauchy–Schwarz Derivation

Let $\psi \in \mathcal{H}$, and define:

$$\Delta\hat{A} = \hat{A} - \langle \hat{A} \rangle, \quad \Delta\hat{B} = \hat{B} - \langle \hat{B} \rangle.$$

Apply the Cauchy–Schwarz inequality:

$$|\langle \Delta\hat{A}\psi, \Delta\hat{B}\psi \rangle|^2 \leq \langle \Delta\hat{A}\psi, \Delta\hat{A}\psi \rangle \cdot \langle \Delta\hat{B}\psi, \Delta\hat{B}\psi \rangle.$$

This becomes:

$$|\langle \psi | \Delta\hat{A}\Delta\hat{B} | \psi \rangle|^2 \leq \sigma_A^2 \sigma_B^2.$$

Write:

$$\langle \Delta\hat{A}\Delta\hat{B} \rangle = \frac{1}{2}\langle \{\Delta\hat{A}, \Delta\hat{B}\} \rangle + \frac{1}{2}\langle [\Delta\hat{A}, \Delta\hat{B}] \rangle.$$

Use:

$$|\langle \Delta\hat{A}\Delta\hat{B} \rangle|^2 \geq \left(\frac{1}{2i}\langle [\hat{A}, \hat{B}] \rangle \right)^2.$$

This gives the uncertainty relation.

Application to Position and Momentum

For $\hat{x}$ and $\hat{p}$, we have:

$$[\hat{x}, \hat{p}] = i\hbar, \quad \Rightarrow \quad \sigma_x \sigma_p \geq \frac{\hbar}{2}.$$

General Observables

For any pair of Hermitian operators $\hat{A}, \hat{B}$, this yields a constraint on the product of variances:

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i}\langle [\hat{A}, \hat{B}] \rangle \right)^2.$$

Harmonic Oscillator Example

Ground state wavefunction:

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}.$$

Then:

$$\sigma_x^2 = \frac{\hbar}{2m\omega}, \quad \sigma_p^2 = \frac{\hbar m\omega}{2} \quad \Rightarrow \quad \sigma_x\sigma_p = \frac{\hbar}{2}.$$

This saturates the inequality — the ground state of the harmonic oscillator is a **minimum uncertainty state**.

Interpretation

- The lower bound of the uncertainty relation is reached only when the state is a Gaussian wavepacket. - Equality occurs when the commutator and anticommutator terms are aligned in phase.

Hydrogen Atom

Time-Independent Schrödinger Equation

The potential for the hydrogen atom is:

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}.$$

In spherical coordinates, the time-independent Schrödinger equation is:

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi + V(r)\psi = E\psi.$$

Separation of Variables

Assume $\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$, where $Y(\theta,\phi)$ are spherical harmonics satisfying:

$$L^2Y = \hbar^2\ell(\ell+1)Y.$$

Radial Equation Derivation

The radial equation becomes:

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{\ell(\ell+1)}{r^2}R\right] - \frac{e^2}{4\pi\varepsilon_0 r}R = ER.$$

Let $u(r) = rR(r)$, giving the simpler form:

$$-\frac{\hbar^2}{2\mu}\frac{d^2u}{dr^2} + \left[\frac{\hbar^2\ell(\ell+1)}{2\mu r^2} - \frac{e^2}{4\pi\varepsilon_0 r}\right]u = Eu.$$

Energy Levels

Solutions for $u(r)$ lead to the quantised energy levels:

$$E_n = -\frac{\mu e^4}{32\pi^2\varepsilon_0^2\hbar^2 n^2} = -\frac{13.6\,\text{eV}}{n^2}, \quad n = 1, 2, 3, \ldots$$

Quantum Numbers and Allowed Values

- $n \in \mathbb{Z}^+$: principal quantum number. - $\ell = 0, 1, \ldots, n-1$: angular momentum quantum number. - $m = -\ell, -\ell+1, \ldots, \ell$: magnetic quantum number.

Inequality: $|m| \le \ell$.

Wavefunction Interpretation

- The total wavefunction is:

$$\psi_{n\ell m}(r,\theta,\phi) = R_{n\ell}(r)Y_{\ell m}(\theta,\phi).$$

- $R_{n\ell}(r)$ determines the radial distribution; nodes increase with $n-\ell-1$. - $Y_{\ell m}$ determines angular dependence; complex-valued in general. - The probability density is:

$$|\psi(r,\theta,\phi)|^2 = |R(r)|^2|Y_{\ell m}(\theta,\phi)|^2.$$

- The radial probability density is:

$$P(r)\,dr = |R(r)|^2 r^2\,dr = |u(r)|^2\,dr.$$

Degeneracy

The degeneracy of the n-th energy level is:

$$g(n) = \sum_{\ell=0}^{n-1}(2\ell+1) = n^2.$$

Asymptotic Behaviour of $u(r)$

For large r: $u(r) \sim e^{-\alpha r}$ for some $\alpha > 0$. For small r: $u(r) \sim r^{\ell+1}$ to cancel the $1/r^2$ centrifugal term.

Hermitian Operators and Expectation Values

Definition

An operator A is **Hermitian** if:

$$\int \psi_1^*(A\psi_2)\,dx = \int (A\psi_1)^*\psi_2\,dx \quad \text{for all } \psi_1, \psi_2.$$

This implies $\langle\psi, A\psi\rangle \in \mathbb{R}$, i.e., expectation values are real.

Expectation Value

Given a normalized state ψ, the expectation value of observable A is:

$$\langle A\rangle = \int \psi^* A\psi\,dx.$$

If A is Hermitian, then $\langle A\rangle \in \mathbb{R}$.

Eigenfunctions and Spectra

If A is Hermitian:

- Eigenvalues are real.
- Eigenfunctions corresponding to distinct eigenvalues are orthogonal.

Hermiticity of Common Operators

- Position operator: $\hat{x}$ is Hermitian.
- Momentum operator: $\hat{p} = -i\hbar\partial_x$ is Hermitian with suitable boundary conditions.
- Hamiltonian $\hat{H}$ is Hermitian so that energy is real and probability is conserved.

Commutators and Hermiticity

For two Hermitian operators A, B, the commutator $[A, B]$ is anti-Hermitian:

$$[A, B]^\dagger = -[A, B].$$

So $i[A, B]$ is Hermitian.

Ehrenfest's Theorem

Derivation from Time-Dependent Schrödinger Equation (TDSE)

Let $\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle$ be the expectation value of an operator $\hat{A}$ in the state $\psi(x, t)$ satisfying the TDSE:

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi.$$

Then the time derivative of the expectation value is:

$$\frac{d}{dt}\langle A \rangle = \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle + \frac{1}{i\hbar}\langle [\hat{A}, \hat{H}] \rangle.$$

If $\hat{A}$ has no explicit time dependence, this simplifies to:

$$\frac{d}{dt}\langle A \rangle = \frac{1}{i\hbar}\langle [\hat{A}, \hat{H}] \rangle.$$

Applications to Position and Momentum

Take the standard Hamiltonian:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}).$$

Using the commutation relation $[\hat{x}, \hat{p}] = i\hbar$, we get:

For position:

$$\frac{d}{dt}\langle \hat{x} \rangle = \frac{1}{i\hbar}\langle [\hat{x}, \hat{H}] \rangle = \frac{1}{m}\langle \hat{p} \rangle.$$

For momentum:

$$\frac{d}{dt}\langle \hat{p} \rangle = \frac{1}{i\hbar}\langle [\hat{p}, \hat{H}] \rangle = -\left\langle \frac{dV}{dx} \right\rangle.$$

Classical Correspondence

These resemble Newton's laws:

$$\frac{d}{dt}\langle \hat{x} \rangle = \frac{\langle \hat{p} \rangle}{m}, \qquad \frac{d}{dt}\langle \hat{p} \rangle = \langle F \rangle.$$

Interpretation: The expectation values of quantum observables obey classical equations of motion — this is a form of the quantum-to-classical correspondence principle.

1D Potential Square Well

Infinite Square Well

Potential

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ \infty, & \text{otherwise.} \end{cases}$$

The wavefunction $\psi(x)$ must vanish at $x = 0$ and $x = a$, since the potential is infinite outside.

In the region $0 < x < a$, the time-independent Schrödinger equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi,$$

with general solution:

$$\psi(x) = A\sin(kx) + B\cos(kx), \quad k = \frac{\sqrt{2mE}}{\hbar}.$$

$$\psi(0) = 0 \Rightarrow B = 0, \quad \psi(a) = 0 \Rightarrow \sin(ka) = 0 \Rightarrow k = \frac{n\pi}{a}.$$

Energy Levels

$$E_n = \frac{\hbar^2\pi^2 n^2}{2ma^2}, \quad n = 1, 2, 3, \ldots$$

Normalised Wavefunctions

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right), \quad 0 < x < a.$$

Quantum Harmonic Oscillator and Ladder Operators

Hamiltonian

The quantum harmonic oscillator has Hamiltonian:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2.$$

Ladder Operators

Define the annihilation and creation operators:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right), \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i}{m\omega}\hat{p}\right).$$

Commutation Relation

$$[\hat{a}, \hat{a}^\dagger] = 1.$$

Hamiltonian in Terms of Ladder Operators

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right).$$

Energy Eigenstates and Eigenvalues

Define the number operator:

$$\hat{N} = \hat{a}^\dagger \hat{a}, \quad \text{with } \hat{N}|n\rangle = n|n\rangle.$$

Then the eigenstates and energies are:

$$\hat{H}|n\rangle = \hbar\omega \left(n + \frac{1}{2}\right)|n\rangle, \quad n = 0, 1, 2, \ldots$$

Action of Ladder Operators

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

Ground State Wavefunction

The ground state satisfies $\hat{a}|0\rangle = 0$, leading to:

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}}.$$

Excited States

Higher wavefunctions can be obtained by applying $\hat{a}^\dagger$ repeatedly:

$$\psi_n(x) = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n \psi_0(x).$$

They can also be expressed in terms of Hermite polynomials:

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-\frac{m\omega x^2}{2\hbar}}.$$

Time-Dependent Perturbation Theory

General Setup

Suppose the Hamiltonian is of the form:

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{V}(t),$$

where $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{V}(t)$ is a small time-dependent perturbation.

Assume:

$$\hat{H}_0|n\rangle = E_n|n\rangle.$$

First-Order Transition Amplitude

The first-order probability amplitude to transition from state $|i\rangle$ to $|f\rangle$ is:

$$c_f^{(1)}(t) = \frac{1}{i\hbar}\int_0^t \langle f|\hat{V}(t')|i\rangle e^{i\omega_{fi}t'}\, dt', \quad \text{where } \omega_{fi} = \frac{E_f - E_i}{\hbar}.$$

Transition Probability

$$P_{i\to f}(t) = |c_f^{(1)}(t)|^2.$$

Fermi's Golden Rule

If $\hat{V}(t) = \hat{V}e^{-i\omega t} + \text{c.c.}$, and there's a continuum of final states:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}|\langle f|\hat{V}|i\rangle|^2 \rho(E_f),$$

where $\rho(E_f)$ is the density of final states.

Delta Function Potential

Potential Definition

$$V(x) = -\alpha\delta(x), \quad \alpha > 0.$$

Time-Independent Schrödinger Equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} - \alpha\delta(x)\psi(x) = E\psi(x).$$

Bound State Solution

Seek solution $\psi(x) = Ae^{-\kappa|x|}$ with $\kappa > 0$. Matching discontinuity in derivative:

$$\psi'(0^+) - \psi'(0^-) = -\frac{2m\alpha}{\hbar^2}\psi(0).$$

Gives:

$$\kappa = \frac{m\alpha}{\hbar^2}, \quad E = -\frac{m\alpha^2}{2\hbar^2}.$$

Normalised Wavefunction

$$\psi(x) = \sqrt{\kappa}e^{-\kappa|x|}.$$

Boundary Conditions

- $\psi(x)$ is continuous at $x = 0$.
- Discontinuity in derivative given by delta potential:

$$\psi'(0^+) - \psi'(0^-) = -\frac{2m\alpha}{\hbar^2}\psi(0).$$

Two-Level Systems and Time Evolution

General State

Let $|\psi(t)\rangle = c_1(t)|1\rangle + c_2(t)|2\rangle$, where:

$$\hat{H}|1\rangle = E_1|1\rangle, \quad \hat{H}|2\rangle = E_2|2\rangle.$$

Time Evolution

If no perturbation:

$$|\psi(t)\rangle = c_1(0)e^{-iE_1t/\hbar}|1\rangle + c_2(0)e^{-iE_2t/\hbar}|2\rangle.$$

Revival Time

Define $\Delta E = E_2 - E_1$. Then the **revival time** is:

$$T = \frac{2\pi\hbar}{\Delta E}.$$

At time T, the system returns to its initial state up to a global phase.

Oscillations in Probability

$$|c_1(t)|^2 = |c_1(0)|^2, \quad |c_2(t)|^2 = |c_2(0)|^2,$$

but relative phase oscillates, giving interference in observables.

With Coupling (e.g. Rabi oscillations)

If a time-dependent coupling exists between $|1\rangle$ and $|2\rangle$, governed by $\hat{V}(t) = \hbar\Omega\cos(\omega t)\hat{\sigma}_x$, transition probabilities oscillate sinusoidally (Rabi formula).

Spectral Theorem for Real Self-Adjoint Operators

Definition (Self-Adjoint Operator): Let V be a finite-dimensional inner product space over $\mathbb{R}$. A linear map $\alpha \in \mathcal{L}(V)$ is called *self-adjoint* if

$$\langle \alpha(v), w \rangle = \langle v, \alpha(w) \rangle \quad \text{for all } v, w \in V.$$

Lemma: Let $\alpha \in \mathcal{L}(V)$ be self-adjoint. Then all eigenvalues of α are real.

Proof sketch: Let $v \in V$ be an eigenvector with eigenvalue λ, so $\alpha(v) = \lambda v$. Then

$$\lambda \langle v, v \rangle = \langle \alpha(v), v \rangle = \langle v, \alpha(v) \rangle = \overline{\lambda} \langle v, v \rangle,$$

so $\lambda = \overline{\lambda} \in \mathbb{R}$.

Lemma: Eigenvectors of a self-adjoint operator corresponding to distinct eigenvalues are orthogonal.

Proof sketch: Let $\alpha(v) = \lambda v$ and $\alpha(w) = \mu w$ with $\lambda \neq \mu$. Then

$$\lambda \langle v, w \rangle = \langle \alpha(v), w \rangle = \langle v, \alpha(w) \rangle = \mu \langle v, w \rangle,$$

so $(\lambda - \mu)\langle v, w \rangle = 0$ implies $\langle v, w \rangle = 0$.

Spectral Theorem (Real Case): Let V be a finite-dimensional inner product space over $\mathbb{R}$. Then every self-adjoint linear operator $\alpha \in \mathcal{L}(V)$ is diagonalizable, and there exists an orthonormal basis of V consisting of eigenvectors of α.

Equivalently: If A is a real symmetric matrix, then A is orthogonally diagonalizable: there exists $Q \in \mathrm{O}(n)$ such that $Q^T A Q$ is diagonal.

Proof: We proceed by induction on $\dim V$.

Base case: $\dim V = 1$ is trivial.

Inductive step: Assume the result for dimension $n-1$. Since α is self-adjoint, it has a real eigenvalue λ with eigenvector $v \neq 0$. Let $U = \langle v \rangle^{\perp}$.

Claim: U is α-invariant.

Let $u \in U$, so $\langle u, v \rangle = 0$. Then

$$\langle \alpha(u), v \rangle = \langle u, \alpha(v) \rangle = \langle u, \lambda v \rangle = \lambda \langle u, v \rangle = 0,$$

so $\alpha(u) \in U$.

Thus, $\alpha|_U$ is self-adjoint on a space of dimension $n-1$. By the inductive hypothesis, U has an orthonormal basis of eigenvectors of α. Together with $v/\|v\|$, this gives an orthonormal basis of eigenvectors for V.

■

Sylvester's Law of Inertia

Theorem (Sylvester's Law of Inertia): Let V be a finite-dimensional real inner product space, and let $b : V \times V \to \mathbb{R}$ be a symmetric bilinear form. Then there exists a basis of V in which the matrix of b is diagonal with entries in $\{-1, 0, 1\}$, and the number of each of these entries (the signature) is independent of the choice of diagonalizing basis.

In other words, any real symmetric bilinear form is congruent to a diagonal matrix with only $+1$, -1, and 0 entries, and the number of each is an invariant of the form.

Definition: The *signature* of a symmetric bilinear form is the triple (n_+, n_-, n_0) where:

- n_+ = number of +1s (positive index),
- n_- = number of -1s (negative index),
- n_0 = number of 0s (nullity).

Proof Sketch:

- Choose a basis and represent b by a real symmetric matrix A.
- Use orthogonal change of basis (Gram–Schmidt and congruence) to bring A into diagonal form.
- Diagonal entries must be real (since A is symmetric).
- Using congruence (not similarity), the diagonal matrix has entries in $\{-1, 0, 1\}$ after scaling.
- Sylvester's Law says the counts of each type (+1, -1, 0) are invariant under congruence transformations.

Example: Let

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

This matrix is real symmetric, so it represents a symmetric bilinear form on $\mathbb{R}^2$.

Compute eigenvalues: $2 \pm 1 = 3, 1$ (both positive), so it is positive definite. Thus, its signature is $(2, 0, 0)$.

Now consider:

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Signature is clearly $(1, 1, 1)$ — Sylvester's Law asserts that no congruence transformation can change this signature.

Gram–Schmidt Orthogonalization

Let V be a finite-dimensional inner product space over $\mathbb{R}$. Suppose $\{v_1, v_2, \ldots, v_n\}$ is a basis for V. The *Gram–Schmidt process* produces an orthonormal basis $\{e_1, e_2, \ldots, e_n\}$ such that $\mathrm{span}(e_1, \ldots, e_k) = \mathrm{span}(v_1, \ldots, v_k)$ for all k.

Algorithm:

Define $u_1 = v_1$, and then recursively

$$u_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, u_j \rangle}{\langle u_j, u_j \rangle} u_j, \quad \text{for } k = 2, \ldots, n.$$

Then define

$$e_k = \frac{u_k}{\|u_k\|}, \quad \text{for } k = 1, \ldots, n.$$

Output: $\{e_1, \ldots, e_n\}$ is an orthonormal basis of V.

Properties:

- Each e_k is orthogonal to $e_1, \ldots, e_{k-1}$.
- $\langle e_i, e_j \rangle = \delta_{ij}$.
- The process is numerically unstable in floating-point arithmetic; modified Gram–Schmidt can be used in practice.

Orthogonal Complements and Direct Sum Decomposition

Let V be a finite-dimensional inner product space over $\mathbb{R}$, and let $U \subseteq V$ be a subspace.

Definition (Orthogonal Complement): The *orthogonal complement* of U is the subspace

$$U^\perp := \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in U\}.$$

Proposition:

$$V = U \oplus U^\perp,$$

that is, every vector $v \in V$ can be uniquely written as $v = u + w$ with $u \in U$ and $w \in U^\perp$.

Proof Sketch:

- Choose a basis $\{u_1, \ldots, u_k\}$ for U.
- Extend it to a basis $\{v_1, \ldots, v_n\}$ for V.
- Apply Gram–Schmidt to obtain an orthonormal basis $\{e_1, \ldots, e_n\}$ of V, where $\{e_1, \ldots, e_k\}$ spans U.
- Then $\{e_{k+1}, \ldots, e_n\}$ spans $U^\perp$.
- Every $v \in V$ can then be written as $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$ with components in U and $U^\perp$.

Corollary:

$$\dim V = \dim U + \dim U^\perp.$$

Orthogonal Projection: The map $\pi_U : V \to U$ defined by

$$\pi_U(v) = \sum_{i=1}^k \langle v, e_i\rangle e_i \quad \text{where } \{e_1, \ldots, e_k\} \text{ is an orthonormal basis of } U$$

is the orthogonal projection onto U along $U^\perp$.

Characterization:

$$v - \pi_U(v) \in U^\perp, \quad \text{and } \pi_U(v) \in U.$$

Hence $v = \pi_U(v) + (v - \pi_U(v))$ is the unique decomposition of v into U and $U^\perp$ components.

Orthogonal Projections

Let V be a finite-dimensional inner product space over $\mathbb{R}$, and let $U \subseteq V$ be a subspace with orthonormal basis $\{e_1, \ldots, e_k\}$.

Definition (Orthogonal Projection): The *orthogonal projection* of $v \in V$ onto U is

$$\pi_U(v) := \sum_{i=1}^k \langle v, e_i\rangle e_i.$$

Properties:

- $\pi_U(v) \in U$, and $v - \pi_U(v) \in U^\perp$.
- π_U is linear: $\pi_U(\lambda v + \mu w) = \lambda\pi_U(v) + \mu\pi_U(w)$.
- π_U is idempotent: $\pi_U(\pi_U(v)) = \pi_U(v)$.
- π_U is self-adjoint: $\langle \pi_U(v), w\rangle = \langle v, \pi_U(w)\rangle$.

Matrix Form: Let E be the $n\times k$ matrix whose columns are the orthonormal vectors $e_1, \ldots, e_k$. Then the projection matrix is

$$P = EE^T,$$

and for $v \in \mathbb{R}^n$ viewed as a column vector,

$$\pi_U(v) = Pv.$$

Application (Least Squares): Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, the least squares solution to $Ax \approx b$ is the x minimizing $\|Ax - b\|^2$, given by solving

$$A^T Ax = A^T b,$$

which corresponds to projecting b orthogonally onto $\mathrm{Im}(A)$.

Simultaneous Diagonalisation

Theorem (Simultaneous Diagonalisation): Let V be a finite-dimensional inner product space over $\mathbb{R}$. Suppose $\alpha, \beta \in \mathcal{L}(V)$ are self-adjoint and commute: $\alpha\beta = \beta\alpha$. Then α and β are simultaneously diagonalizable: there exists an orthonormal basis of V consisting of vectors that are eigenvectors for both α and β.

More generally: If $\{\alpha_i\}_{i=1}^k$ is a commuting family of self-adjoint operators on V, then there exists an orthonormal basis of V consisting of simultaneous eigenvectors for all α_i.

Proof Sketch:

- Since each α_i is self-adjoint, it is diagonalizable with an orthonormal basis of eigenvectors.
- Because the operators commute, the eigenspaces of one are invariant under the others.
- Proceed inductively: diagonalize α_1 to get decomposition into orthogonal eigenspaces.
- Restrict each α_i to these eigenspaces; since they commute and are self-adjoint, repeat the process.
- The final basis simultaneously diagonalizes all α_i.

Application: Let A, B be real symmetric $n \times n$ matrices such that $AB = BA$. Then there exists an orthogonal matrix Q such that

$$Q^T AQ = D_A, \quad Q^T BQ = D_B$$

with both D_A and D_B diagonal.

Remark: Commuting alone does not imply simultaneous diagonalisation unless the matrices are also diagonalizable — which self-adjointness guarantees in the real inner product case.

Matrix Similarity, Commutators, and Centralisers

Definition (Similarity): Two matrices $A, B \in \mathrm{Mat}_n(F)$ are *similar* if there exists an invertible matrix $P \in \mathrm{GL}_n(F)$ such that

$$B = P^{-1}AP.$$

Similarity preserves many algebraic properties: determinant, trace, characteristic polynomial, eigenvalues, minimal polynomial, and rank.

Definition (Commutator): Given $A, B \in \mathrm{Mat}_n(F)$, their *commutator* is

$$[A, B] := AB - BA.$$

We say A and B *commute* if $[A, B] = 0$.

Observation: Similarity is an equivalence relation. If $A \sim B$, then A and B represent the same linear operator with respect to different bases.

Definition (Centraliser): Let $A \in \mathrm{Mat}_n(F)$. The set

$$\{B \in \mathrm{Mat}_n(F) : AB = BA\}$$

is a subspace of $\mathrm{Mat}_n(F)$ called the *centraliser* of A (though not named this in our course). It consists of all matrices that commute with A.

Examples:

- If $A = \lambda I_n$ for some scalar λ, then $AB = BA$ for all B.
- If A is diagonal, then B commutes with A if and only if B is diagonal (in general, if A has distinct eigenvalues).

Application (Simultaneous Diagonalisation): If A is diagonalizable and B commutes with A, then B preserves the eigenspaces of A. This often enables simultaneous diagonalisation.

Exercise: Show that the set of all B commuting with a given A forms a vector space, and compute its dimension in a few concrete examples.

Fitting's Lemma and Image–Kernel Decomposition

Fitting's Lemma: Let V be a finite-dimensional vector space and let $\alpha \in \mathcal{L}(V)$. Then there exists an integer $m \geq 0$ such that

$$V = \ker(\alpha^m) \oplus \mathrm{Im}(\alpha^m).$$

Details:

- The sequence $\ker(\alpha^k)$ is increasing, and $\mathrm{Im}(\alpha^k)$ is decreasing.
- Since V is finite-dimensional, both sequences stabilize: there exists m such that
$$\ker(\alpha^m) = \ker(\alpha^{m+1}), \quad \mathrm{Im}(\alpha^m) = \mathrm{Im}(\alpha^{m+1}).$$
- Then $V = \ker(\alpha^m) \oplus \mathrm{Im}(\alpha^m)$.

Interpretation:

- $\ker(\alpha^m)$ is the *generalized null space* of α.
- $\mathrm{Im}(\alpha^m)$ is the *stable image*.
- The restriction $\alpha|_{\mathrm{Im}(\alpha^m)}$ is injective.
- The restriction $\alpha|_{\ker(\alpha^m)}$ is nilpotent.

Remarks:

- The decomposition is not α-invariant in general.
- In the case where α is a linear operator on V and its minimal polynomial splits into relatively prime factors, Fitting's Lemma is used to decompose V into primary components.

Example: Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $A^2 = 0$, so

$$\ker(A^2) = \mathbb{R}^2, \quad \mathrm{Im}(A^2) = \{0\}.$$

So $V = \ker(A^2) \oplus \mathrm{Im}(A^2)$ trivially.

Addendum: Connection to Primary Decomposition and Jordan Form

Primary Decomposition Theorem: Let $\alpha \in \mathcal{L}(V)$ have minimal polynomial

$$m_\alpha(t) = p_1(t)^{r_1} \cdots p_k(t)^{r_k},$$

where the p_i are distinct monic irreducible polynomials over F. Then

$$V = \bigoplus_{i=1}^k V_i, \quad \text{where } V_i := \ker(p_i(\alpha)^{r_i}).$$

Each V_i is α-invariant, and $\alpha|_{V_i}$ has minimal polynomial $p_i^{r_i}$.

Use of Fitting's Lemma: Within each V_i, we apply Fitting's Lemma to write

$$V_i = \ker(\alpha|_{V_i}^m) \oplus \mathrm{Im}(\alpha|_{V_i}^m),$$

giving a decomposition into a nilpotent part and an invertible (or semi-simple) part.

Application to Jordan Form: For α with minimal polynomial splitting into linear factors over F (e.g. over $\mathbb{C}$), the primary decomposition groups generalized eigenspaces:

$$V = \bigoplus_\lambda \ker((\alpha - \lambda I)^r).$$

Fitting's Lemma provides the basis for constructing Jordan blocks: each $\ker((\alpha - \lambda I)^r)$ is where the nilpotent action lives, layered by powers of $(\alpha - \lambda I)$.

Summary: Fitting's Lemma gives the internal structure of each generalized eigenspace, forming the backbone of the Jordan canonical form construction.

Trace Identity: $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

tr(AB) = tr(BA)

Theorem: Let $A \in \mathrm{Mat}_{n\times m}(F)$ and $B \in \mathrm{Mat}_{m\times n}(F)$. Then

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

Proof Sketch: Write out the trace:

$$\mathrm{tr}(AB) = \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{j=1}^m A_{ij}B_{ji}.$$

$$\mathrm{tr}(BA) = \sum_{j=1}^m (BA)_{jj} = \sum_{j=1}^m \sum_{i=1}^n B_{ji}A_{ij}.$$

The two sums are equal by rearranging the order of summation.

Remarks:

- This holds even when AB and BA are not the same size (e.g. AB is $n \times n$, BA is $m \times m$).
- In particular, if $A, B \in \mathrm{Mat}_n(F)$ then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
- Generalization: For any $k \in \mathbb{N}$, $\mathrm{tr}(A_1A_2 \cdots A_k)$ is invariant under cyclic permutations.

Commutators and the Kernel of Trace

Fact: The set of all commutators $[A, B] := AB - BA$ in $\mathrm{Mat}_n(F)$ spans the kernel of the trace map:

$$\ker(\mathrm{tr}) = \mathrm{span}\{AB - BA : A, B \in \mathrm{Mat}_n(F)\}.$$

Proof Sketch:

- For any $A, B \in \mathrm{Mat}_n(F)$, we have $\mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0$.
- So every commutator lies in $\ker(\mathrm{tr})$.
- Conversely, $\ker(\mathrm{tr})$ is a hyperplane in $\mathrm{Mat}_n(F)$ (codimension 1), and it can be shown that the span of commutators is already $n^2 - 1$-dimensional, hence equals the kernel.

Interpretation: The trace map $\mathrm{tr} : \mathrm{Mat}_n(F) \to F$ is surjective with 1-dimensional image, and the space of trace-zero matrices is exactly the space generated by commutators.

Example:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad AB - BA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This matrix has trace zero and lies in the span of commutators.

Uniqueness of the Trace Function

Theorem: Let $f : \mathrm{Mat}_n(F) \to F$ be a linear map such that

$$f(AB) = f(BA) \quad \text{for all } A, B \in \mathrm{Mat}_n(F).$$

Then f is a scalar multiple of the trace: there exists $c \in F$ such that

$$f(A) = c \cdot \mathrm{tr}(A) \quad \text{for all } A.$$

Proof: Let $E_{ij} \in \mathrm{Mat}_n(F)$ denote the standard matrix units: $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$.

1. For $i \neq j$, compute:

$$f(E_{ij}) = f(E_{ii}E_{ij}) = f(E_{ij}E_{ii}) = f(E_{ij}),$$

and

$$f(E_{ij}) = f(E_{ik}E_{kj}) = f(E_{kj}E_{ik}) = 0,$$

for $k \neq i, j$ — and more generally, we can construct A, B so that $AB = E_{ij}$ but $BA = 0$ if $i \neq j$, forcing $f(E_{ij}) = 0$.

So all off-diagonal entries vanish under f:

$$f(E_{ij}) = 0 \quad \text{for } i \neq j.$$

2. For diagonal entries E_{ii} and E_{jj}:

$$f(E_{ii}) = f(E_{ij}E_{ji}) = f(E_{ji}E_{ij}) = f(E_{jj}),$$

so all $f(E_{ii})$ are equal. Let $c := f(E_{11})$.

Then for any $A = \sum_{i,j} a_{ij}E_{ij}$, we have:

$$f(A) = \sum_{i=1}^{n} a_{ii} f(E_{ii}) = c \sum_{i=1}^{n} a_{ii} = c \cdot \mathrm{tr}(A).$$

■

Corollary: The trace map is the unique linear map $f : \mathrm{Mat}_n(F) \to F$ satisfying

$$f(AB) = f(BA) \quad \text{and} \quad f(I) = n.$$

Proof: From the above, $f = c \cdot \mathrm{tr}$, and $f(I) = c \cdot \mathrm{tr}(I) = cn$. So to force $f(I) = n$, we must have $c = 1$, hence $f = \mathrm{tr}$.

Useful Problem-Solving Identities

Polarisation Trick (Hermitian/Symmetric Forms): Let ψ be a symmetric bilinear form on a complex vector space. Then

$$\psi(u, v) = \frac{1}{n} \sum_{k=1}^{n} \zeta^k \, \psi(u + \zeta^k v, u + \zeta^k v), \quad \text{where } \zeta = e^{2\pi i/n},\ n \geq 2.$$

Averaging isolates the cross-term via orthogonality of roots of unity.

Cyclic Trace Identity:

$$\mathrm{tr}(A_1 A_2 \cdots A_k) = \mathrm{tr}(A_k A_1 \cdots A_{k-1}).$$

Useful when trace appears in a product — allows cyclic rearrangement.

Rank–Nullity Theorem: For any $\alpha \in \mathcal{L}(V)$:

$$\dim V = \dim \ker \alpha + \dim \mathrm{Im}\, \alpha.$$

Diagonalisation by Spectral Theorem: For real symmetric A, there exists $Q \in \mathrm{O}(n)$ such that:

$$Q^T A Q = D \quad \text{(diagonal)}.$$

Projection Formula: Let $\{e_1, \ldots, e_k\}$ be an orthonormal basis for subspace U. Then the orthogonal projection onto U is:

$$\pi_U(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i.$$

Minimal Polynomial Identity: If $m_\alpha(t)$ is the minimal polynomial of $\alpha \in \mathcal{L}(V)$, then:

$$m_\alpha(\alpha) = 0, \quad \text{and } m_\alpha \text{ is the monic polynomial of least degree with th}$$

Schur Decomposition (over $\mathbb{C}$): Any $A \in \mathrm{Mat}_n(\mathbb{C})$ is unitarily triangularizable:

$$A = UTU^*, \quad \text{with } U \in U(n),\ T \text{ upper triangular}.$$

Commutator Trace Identity:

$$\mathrm{tr}([A, B]) = \mathrm{tr}(AB - BA) = 0.$$

Double Commutator Identity (Lie-type trick):

$$[A, [A, B]] = A^2 B - 2ABA + BA^2.$$

Useful in induction or polynomial identity manipulation.

Gauss's Lemma and Eisenstein's Criterion (General Form)

Let R be a unique factorisation domain (UFD), with field of fractions F. Define the *content* of $f \in R[x]$ as $\operatorname{cont}(f) = \gcd$ of its coefficients. Say f is *primitive* if $\operatorname{cont}(f) = 1$.

Gauss I: $\operatorname{cont}(fg) = \operatorname{cont}(f) \cdot \operatorname{cont}(g)$

Proof: Let $f = c \cdot f'$, $g = d \cdot g'$ where f', g' are primitive and $c = \operatorname{cont}(f)$, $d = \operatorname{cont}(g)$. Then

$$fg = cd \cdot f'g'.$$

We prove $f'g'$ is primitive.

Suppose p is an irreducible in R dividing all coefficients of $f'g'$. Then reduce modulo (p) to get $\overline{f'} \cdot \overline{g'} = 0$ in $(R/(p))[x]$. But since $R/(p)$ is an integral domain (as R is a UFD), this implies $\overline{f'} = 0$ or $\overline{g'} = 0$, i.e., all coefficients of f' or g' are divisible by p — contradicting primitivity.

Hence $f'g'$ is primitive, and $\operatorname{cont}(fg) = cd = \operatorname{cont}(f)\operatorname{cont}(g)$.

Gauss II: Primitive Irreducible in $R[x] \Rightarrow$ Irreducible in $F[x]$

Proof: Let $f \in R[x]$ be primitive and irreducible in $R[x]$. Suppose $f = gh$ in $F[x]$. Then write $g = a^{-1}g'$, $h = b^{-1}h'$ with $g', h' \in R[x]$, and $a, b \in R \setminus \{0\}$.

Then
$$f = \frac{1}{ab} g'h' \Rightarrow abf = g'h'.$$

Take contents: $\operatorname{cont}(abf) = ab$ (since f primitive), and $\operatorname{cont}(g') \cdot \operatorname{cont}(h') = ab$ by Gauss I. Set

$$g'' := \operatorname{cont}(g')^{-1}g', \quad h'' := \operatorname{cont}(h')^{-1}h'$$

so g'', h'' are primitive, and

$$f = g''h'' \in R[x].$$

Then f is factored into non-unit elements of $R[x]$, contradicting irreducibility.

Gauss III: Primitive $\Rightarrow$ Irreducible in $F[x] \Leftrightarrow$ Irreducible in $R[x]$

Proof: $(\Rightarrow)$ is Gauss II.

$(\Leftarrow)$: Let f be primitive and irreducible in $F[x]$. Suppose $f = gh$ in $R[x]$. Then f is reducible in $F[x]$, contradiction.

Eisenstein's Criterion

Let $f(x) = \sum_{i=0}^{n} a_i x^i \in R[x]$, and let $p \in R$ be an irreducible element such that:
- $p \mid a_0, \dots, a_{n-1}$,
- $p \nmid a_n$,
- $p^2 \nmid a_0$.

Then f is irreducible in $F[x]$, where F is the field of fractions of R.

Proof: Suppose $f = gh$ in $R[x]$ with $\deg g, \deg h < \deg f$.

Let $g = \sum b_i x^i$, $h = \sum c_j x^j$ with degrees r, s, $r + s = n$. Then

$$a_0 = b_0 c_0.$$

But $p \mid a_0$ implies $p \mid b_0$ or $p \mid c_0$.

WLOG $p \mid b_0$. Let k be the least index such that $p \nmid b_k$. Such k exists since $p \nmid a_n$.

Consider the coefficient $a_k = \sum_{i+j=k} b_i c_j$. All b_i with $i < k$ satisfy $p \mid b_i$, and b_k is the first with $p \nmid b_k$. Since $p \mid a_k$ by assumption, the sum

$$a_k = b_k c_0 + (\text{terms divisible by } p)$$

implies $p \mid b_k c_0 \Rightarrow p \mid c_0$ (since $p \nmid b_k$), so $p \mid b_0$ and $p \mid c_0$.

Hence $p^2 \mid a_0 = b_0 c_0$, contradicting the assumption that $p^2 \nmid a_0$.

Therefore, f is irreducible in $R[x]$ and hence in $F[x]$ by Gauss II.

Sylow's Theorems

Let G be a finite group, and let $|G| = p^n m$, where p is prime and $p \nmid m$.

First Sylow Theorem

Statement: G has a subgroup of order p^k for every $0 \le k \le n$. In particular, there exists a subgroup of order p^n (a Sylow p-subgroup).

Proof (Sketch for maximal $k = n$): Act on the set X of subsets of G of size p^n by left multiplication. Then count the number of such subsets and show that some stabiliser must have order divisible by p^n. Alternatively, induct on $|G|$ using Cauchy's theorem and normalisers.

Second Sylow Theorem

Statement: Any two Sylow p-subgroups of G are conjugate. Moreover, every p-subgroup is contained in a Sylow p-subgroup.

Proof: Let P be a Sylow p-subgroup and let Q be any p-subgroup. Consider the action of Q on the left coset space G/P by left multiplication. Then $|G/P| = m$ and $p \nmid m$, so the number of fixed points is congruent to $|G/P| \mod p$.

By the orbit-counting lemma, some fixed point exists — i.e., some gP is fixed by Q, which implies $Q \le gPg^{-1}$.

Third Sylow Theorem

Statement: Let n_p be the number of Sylow p-subgroups of G. Then:

$$n_p \equiv 1 \pmod{p}, \quad n_p \mid m.$$

Proof: Let G act on the set $\mathcal{S}$ of Sylow p-subgroups by conjugation. Then the orbit of any Sylow p-subgroup P under this action has size equal to $|G : N_G(P)|$.

But $P \le N_G(P)$, and $|P| = p^n$, so $p \nmid |G : N_G(P)|$. Thus each orbit has size not divisible by p.

Now apply orbit counting or consider the conjugation action directly: by counting fixed points, we obtain $n_p \equiv 1 \mod p$.

Smith Normal Form and Classification of Abelian Groups

Let R be a PID (typically $R = \mathbb{Z}$), and let M be a finitely generated R-module.

Smith Normal Form (SNF)

Theorem: Let A be an $m \times n$ matrix with entries in a PID R. Then there exist invertible matrices $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$ such that

$$PAQ = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_r \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

with $d_1 \mid d_2 \mid \cdots \mid d_r$ in R.

Proof Outline: This is done by performing elementary row and column operations to diagonalise A, clearing lower entries via division (possible in PIDs), using the Euclidean algorithm to enforce divisibility conditions.

Interpretation: If A is the matrix of a homomorphism $R^n \to R^m$, then the cokernel

$$\operatorname{coker} A \cong \bigoplus_{i=1}^{r} R/(d_i R) \oplus R^{m-r}.$$

Structure Theorem for Finitely Generated Abelian Groups

Let G be a finitely generated abelian group. Then:

Theorem (Invariant Factor Form): There exists an isomorphism

$$G \cong \mathbb{Z}^r \oplus \mathbb{Z}/d_1 \oplus \cdots \oplus \mathbb{Z}/d_k,$$

with $d_1 \mid d_2 \mid \cdots \mid d_k$ and $d_i \geq 2$.

Theorem (Elementary Divisor Form): Alternatively,

$$G \cong \mathbb{Z}^r \oplus \mathbb{Z}/p_1^{e_1} \oplus \cdots \oplus \mathbb{Z}/p_n^{e_n},$$

with p_i prime and exponents $e_i \geq 1$.

Proof Idea: Regard G as a module over $\mathbb{Z}$, present it as $\mathbb{Z}^n/\operatorname{im}(A)$ for some integer matrix A, and apply Smith Normal Form to A. The invariant factors d_i are the diagonal entries in SNF.

Corollaries

- Every finite abelian group is isomorphic to a finite direct sum of cyclic groups of prime power order.
- The invariant factors are unique up to isomorphism of G.
- The torsion subgroup of G is isomorphic to the torsion part $\bigoplus \mathbb{Z}/d_i$.

Structure Theorem for Finitely Generated Modules over a PID

Let R be a principal ideal domain (in our case, typically a Euclidean domain), and let M be a finitely generated R-module.

Statement (Invariant Factor Form)

There exists an isomorphism

$$M \cong R^r \oplus R/(d_1) \oplus R/(d_2) \oplus \cdots \oplus R/(d_k),$$

where $d_i \in R$ are nonzero and satisfy $d_1 \mid d_2 \mid \cdots \mid d_k$. The d_i are called the *invariant factors* of M, and r is the rank of the free part of M.

Alternate Form (Elementary Divisors)

Equivalently, M decomposes as

$$M \cong R^r \oplus \bigoplus_{i=1}^{n} R/(p_i^{e_i}),$$

where each p_i is irreducible (typically prime) in R.

Proof Sketch (Euclidean Domain Version)

1. Present M as a quotient: Let M be generated by n elements, so there exists a surjective map:

$$\phi : R^n \twoheadrightarrow M.$$

Then $M \cong R^n/\ker\phi$.

2. Represent ϕ by an $m \times n$ matrix A with entries in R, corresponding to a presentation of M.

3. Use elementary row and column operations (invertible over R) to bring A into Smith Normal Form:

$$PAQ = \operatorname{diag}(d_1, d_2, \ldots, d_k, 0, \ldots, 0), \quad \text{with } d_1 \mid d_2 \mid \cdots \mid d_k.$$

4. Then $M \cong R/(d_1) \oplus \cdots \oplus R/(d_k) \oplus R^{n-k}$.

5. The free rank $r := n - k$ is uniquely determined as $\dim_R(M \otimes_R F)$ for $F = \operatorname{Frac}(R)$.

Remarks

- The d_i are uniquely determined up to associates and satisfy $d_1 \mid d_2 \mid \cdots \mid d_k$. - The decomposition reflects torsion and free parts:

$$M_{\text{tor}} = \bigoplus R/(d_i), \quad M_{\text{free}} = R^r.$$

- This generalises the classification of finitely generated abelian groups when $R = \mathbb{Z}$.

Hilbert's Basis Theorem

Theorem: Let R be a Noetherian ring. Then the polynomial ring $R[x]$ is also Noetherian.

Proof: Let $I \subseteq R[x]$ be an ideal. We aim to show that I is finitely generated.

For each $n \geq 0$, define

$$I_n := \{a \in R \mid \text{there exists } f(x) \in I$$

$$\text{with } \deg f = n \text{ and leading coefficient } a\}.$$

That is, I_n consists of all possible leading coefficients of degree-n polynomials in I.

Each I_n is an ideal of R. Since R is Noetherian, the ascending chain

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots$$

stabilises. That is, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $I_n = I_N$.

Now, for each $0 \leq n \leq N$, choose finitely many polynomials $f_{n,1}, \ldots, f_{n,r_n} \in I$ of degree n such that their leading coefficients generate I_n.

Let S be the finite set of all such polynomials across all $n \leq N$. We claim that S generates I.

Let $f \in I$ be arbitrary. We induct on the degree $d = \deg f$.

If $d > N$, then the leading coefficient a of f lies in $I_d = I_N$. So $a = \sum_i r_i a_i$ where each a_i is the leading coefficient of some $f_{N,j}$.

Then define

$$g := \sum_i r_i x^{d-N} f_{N,j_i} \in \langle S \rangle.$$

Note that g and f have the same degree and leading coefficient, so $f - g$ has degree $< d$ and lies in I. By induction, $f - g \in \langle S \rangle$, so $f \in \langle S \rangle$.

If $d \leq N$, the same argument applies using $f_{d,j}$.

Therefore, I is generated by S and $R[x]$ is Noetherian.

Noetherian Rings and Modules

Definition: A ring R is *Noetherian* if every ascending chain of ideals

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

stabilises; i.e., there exists n such that $I_k = I_n$ for all $k \geq n$.

Equivalently, R is Noetherian if every ideal of R is finitely generated.

Noetherian Modules

Let M be an R-module. Then M is called *Noetherian* if every submodule is finitely generated. This is equivalent to every ascending chain of submodules of M stabilising.

Examples:

- $\mathbb{Z}$ is Noetherian, since every ideal is of the form (n).
- A field F is trivially Noetherian.
- $\mathbb{Z}[x]$ is Noetherian by Hilbert's Basis Theorem.

Standard Lemmas and Properties

Lemma (Submodule Lemma): If M is a Noetherian R-module and $N \leq M$, then N is Noetherian.

Lemma (Quotient Lemma): If M is Noetherian and $N \leq M$, then M/N is Noetherian.

Lemma (Extension Lemma): Let $0 \to N \to M \to P \to 0$ be a short exact sequence of R-modules. Then:

$$M \text{ Noetherian} \iff N, P \text{ Noetherian}.$$

Corollary: If R is Noetherian and M is a finitely generated R-module, then M is Noetherian.

Proof Sketch: Let $M = Rm_1 + \cdots + Rm_n$ be generated by n elements. Define a surjective map:

$$\phi : R^n \to M, \quad (r_1, \ldots, r_n) \mapsto \sum r_i m_i.$$

Then $M \cong R^n / \ker \phi$, and since R^n is Noetherian and quotients of Noetherian modules are Noetherian, M is Noetherian.

Important Consequences

- Every ideal in a Noetherian ring is finitely generated.
- Every submodule of a finitely generated module over a Noetherian ring is finitely generated.
- Hilbert's Basis Theorem: If R is Noetherian, then $R[x]$ is Noetherian.
- Every finitely generated algebra over a Noetherian ring is Noetherian.

Theorems on PIDs and UFDs

Definitions

- A *principal ideal domain* (PID) is an integral domain in which every ideal is of the form (a) for some $a \in R$.
- A *unique factorisation domain* (UFD) is an integral domain in which every nonzero non-unit can be written as a product of irreducibles, uniquely up to unit and order.

Theorems on PIDs

- Every PID is Noetherian.
- Every PID is a UFD.
- Every finitely generated torsion-free module over a PID is free.
- Every submodule of a free module over a PID is free (in the finitely generated case).
- Let R be a PID, and M a finitely generated R-module. Then:

$$M \cong R^r \oplus R/(d_1) \oplus \cdots \oplus R/(d_k), \quad \text{with } d_1 \mid \cdots \mid d_k.$$

Theorems on UFDs

- Every PID is a UFD, but not every UFD is a PID (e.g. $k[x, y]$).
- In a UFD, irreducibles are primes (i.e. $p \mid ab \Rightarrow p \mid a$ or $p \mid b$).

- In a UFD, GCDs exist and can be expressed as linear combinations in special cases (e.g. Euclidean domains).
- If R is a UFD, then so is $R[x]$.
- Gauss's Lemma holds: if R is a UFD, then $R[x]$ is a UFD.
- In a UFD, any polynomial in $R[x]$ is reducible in $R[x]$ if and only if it is reducible in $\mathrm{Frac}(R)[x]$, provided it is primitive.

Consequences and Comparisons

- Euclidean $\Rightarrow$ PID $\Rightarrow$ UFD $\Rightarrow$ integral domain.
- $\mathbb{Z}$, $k[x]$ are PIDs; $k[x, y]$ is a UFD but not a PID.
- The structure theorem for finitely generated modules applies over PIDs (but fails in general UFDs).
- Gauss's Lemma and Eisenstein's Criterion require UFD assumptions (or PID for cleaner argument).

Rational Canonical Form via the Structure Theorem

Let V be a finite-dimensional vector space over a field k, and let $T : V \to V$ be a linear map.

Idea: Treat V as a module over $k[x]$ via:

$$f(x) \cdot v := f(T)(v), \quad \text{for } f(x) \in k[x],\ v \in V.$$

Then V becomes a finitely generated $k[x]$-module. Since $k[x]$ is a PID, the structure theorem applies.

Structure Theorem Application

There exists an isomorphism of $k[x]$-modules:

$$V \cong \bigoplus_{i=1}^{r} k[x]/(f_i(x)), \quad \text{where } f_i \mid f_{i+1}.$$

The $f_i(x)$ are the *invariant factors* of T, determined uniquely up to associates. The largest invariant factor is the minimal polynomial of T.

Matrix Form

With respect to a suitable basis, the matrix of T is block-diagonal:

$$\mathrm{RCF}(T) = \begin{pmatrix} C(f_1) & & \\ & \ddots & \\ & & C(f_r) \end{pmatrix},$$

where $C(f)$ is the companion matrix of $f(x)$.

Companion matrix: If $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, then

$$C(f) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Uniqueness and Minimal Polynomial

- The invariant factors $f_i(x)$ are uniquely determined by T.
- The product $\prod f_i(x) = \chi_T(x)$ is the characteristic polynomial of T.
- The largest invariant factor equals the minimal polynomial of T.

Summary

- The Rational Canonical Form is unique up to similarity over k.
- It provides a complete invariant for similarity classes over arbitrary fields.
- The module-theoretic view avoids eigenvalues and diagonalisation.

Core Definitions (GRM)

- A **group** $(G, \cdot)$ is a set with an associative binary operation, an identity element e, and inverses: $\forall g \in G, \exists g^{-1} \in G$ such that $gg^{-1} = e$.
- A **ring** $(R, +, \cdot)$ is a set with two operations: $(R, +)$ is an abelian group, multiplication is associative, and distributive over addition.
- A ring is a **domain** if it is commutative with $1 \neq 0$ and has no zero divisors.
- A **field** is a commutative ring in which every nonzero element has a multiplicative inverse.
- An **ideal** $I \subseteq R$ is a subset such that I is an additive subgroup and $r \in R,\ a \in I \Rightarrow ra \in I$.
- A **module** over a ring R is an abelian group M with a scalar multiplication $R \times M \to M$ satisfying:

$$r(m+n) = rm+rn, \quad (r+s)m = rm+sm, \quad (rs)m = r(sm), \quad 1_R m$$

- A module is **Noetherian** if every submodule is finitely generated (equivalently, satisfies the ascending chain condition).
- A **principal ideal domain** (PID) is an integral domain in which every ideal is of the form (a) for some $a \in R$.
- A **unique factorisation domain** (UFD) is a domain where every nonzero non-unit factors into irreducibles, uniquely up to unit and order.
- A **finitely generated module** is a module M with a finite generating set: $\exists m_1, \ldots, m_n \in M$ such that $M = \sum Rm_i$.
- A linear operator $T : V \to V$ is **diagonalisable** if V has a basis of eigenvectors of T.
- The **minimal polynomial** of T is the monic polynomial $m(x)$ of least degree such that $m(T) = 0$.

Contraction Mapping Theorem (Banach Fixed Point Theorem)

Theorem: Let (X, d) be a complete metric space, and let $f : X \to X$ be a contraction; that is, there exists $0 < c < 1$ such that

$$d(f(x), f(y)) \leq c\, d(x, y), \quad \forall x, y \in X.$$

Then:

1. f has a unique fixed point $x^* \in X$, i.e. $f(x^*) = x^*$.
2. For any $x_0 \in X$, the sequence defined by $x_{n+1} = f(x_n)$ converges to x^*.
3. Moreover, the convergence is geometric: $d(x_n, x^*) \leq \frac{c^n}{1-c} d(x_1, x_0)$.

Proof:

Let $x_0 \in X$ be arbitrary and define the sequence $x_n := f(x_{n-1})$ for $n \geq 1$.

Step 1: Show (x_n) is Cauchy.

We have:

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq c\, d(x_n, x_{n-1}).$$

By induction:

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0).$$

Then for $m > n$,

$$d(x_m, x_n) \leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq d(x_1, x_0) \sum_{k=n}^{m-1} c^k \leq \frac{c^n}{1-c} d(x_1, x_0).$$

Hence (x_n) is Cauchy, and since X is complete, $x_n \to x^*$ for some $x^* \in X$.

Step 2: $f(x^*) = x^*$.

Since f is continuous (as it is Lipschitz), we have:

$$f(x^*) = f(\lim x_n) = \lim f(x_n) = \lim x_{n+1} = x^*.$$

Step 3: Uniqueness.

If $f(y) = y$ and $f(z) = z$, then:

$$d(y, z) = d(f(y), f(z)) \leq c\, d(y, z),$$

implying $(1 - c)d(y, z) \leq 0 \Rightarrow y = z$.

Application: Iteration for Solving Equations

Given a recurrence or functional equation of the form $x = f(x)$, one can apply the Banach Fixed Point Theorem to prove: - Existence and uniqueness of a solution - Convergence of the iteration $x_{n+1} = f(x_n)$ - Geometric rate of convergence

Example: Solve $x = \cos x$.

Let $f(x) = \cos x$ on $[0, 1]$ with the usual metric. Then f is a contraction:

$$|f'(x)| = |\sin x| \leq \sin 1 < 1.$$

Thus, iteration converges to the unique fixed point in $[0, 1]$.

Inverse Function Theorem

Definitions:

- A function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable at** $x \in U$ if there exists a linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$\lim_{h \to 0} \frac{\|f(x + h) - f(x) - Df(x)(h)\|}{\|h\|} = 0.$$

- The matrix of $Df(x)$ in standard coordinates is the **Jacobian matrix**:

$$J_f(x) = \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{1 \leq i \leq m, 1 \leq j \leq n}.$$

- A function is of class C^k if it is k times continuously differentiable.

Theorem (Inverse Function Theorem): Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 function on an open set U, and let $a \in U$.

Suppose $Df(a)$ is invertible (i.e. $\det Df(a) \neq 0$). Then:

1. There exists an open neighbourhood V of a and an open neighbourhood W of $f(a)$ such that

$$f : V \to W \text{ is a bijection, and } f^{-1} : W \to V \text{ is } C^1.$$

2. For all $y \in W$, the derivative of the inverse is given by:

$$D(f^{-1})(y) = Df(f^{-1}(y))^{-1}.$$

Proof Outline:

Let $T = Df(a)$, which is invertible. Define $g(x) := f(x) - T(x - a)$. Then $g(a) = f(a)$ and $Dg(a) = 0$.

We rewrite $f(x)$ near a as:

$$f(x) = f(a) + T(x-a) + R(x), \quad \text{with } \frac{\|R(x)\|}{\|x - a\|} \to 0 \text{ as } x \to a.$$

Then define the map:

$$\Phi(x) := x - T^{-1}(f(x) - f(a)).$$

This map has a fixed point at a, and one shows that Φ is a contraction near a. By the Contraction Mapping Theorem, the fixed point is unique and depends continuously on the image value — constructing f^{-1} and proving differentiability.

Example Application:

Let $f(x, y) = (x + y + \sin(xy), x - y)$.

Compute $Df(0, 0)$:

$$Df(0, 0) = \begin{pmatrix} 1 + 0 & 1 + 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \det = -2 \neq 0.$$

So f is locally invertible near $(0, 0)$ and f^{-1} is differentiable.

Addendum: Implicit Function Theorem

Theorem (Implicit Function Theorem):

Let $F : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a C^1 function, and suppose $F(a,b) = 0$ for some $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$.

If the Jacobian matrix

$$\left(\frac{\partial F_i}{\partial y_j}(a,b) \right)_{1 \le i,j \le m}$$

is invertible, then there exist open neighbourhoods:

$$U \subseteq \mathbb{R}^n \text{ around } a, \quad V \subseteq \mathbb{R}^m \text{ around } b,$$

and a unique C^1 function $g : U \to V$ such that:

$$F(x, g(x)) = 0 \quad \text{for all } x \in U.$$

Moreover, for all $x \in U$, we have:

$$Dg(x) = -\left(\frac{\partial F}{\partial y}(x, g(x)) \right)^{-1} \cdot \frac{\partial F}{\partial x}(x, g(x)).$$

Interpretation: If $F(x,y) = 0$ implicitly defines y in terms of x, then under the above conditions, this can be solved locally as $y = g(x)$ with g differentiable.

Example: Let $F(x,y) = x^2 + y^2 - 1$. Then $F(0,1) = 0$, and

$$\frac{\partial F}{\partial y}(0,1) = 2y = 2 \neq 0.$$

So there exists a differentiable function $y = g(x)$ near $x = 0$ such that $x^2 + g(x)^2 = 1$ — i.e., one branch of the unit circle.

Uniform Convergence and Continuity / Integrability / Differentiability

Definition (Uniform Convergence): Let $f_n : X \to \mathbb{R}$ be a sequence of functions. We say $f_n \to f$ *uniformly* on X if:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } \forall n \ge N, \forall x \in X,\ |f_n(x) - f(x)| < \varepsilon.$$

Theorems:

- **Continuity Preserved:** If each f_n is continuous on a metric space X, and $f_n \to f$ uniformly, then f is continuous.
- **Integrability Preserved:** If $f_n \in L^1[a,b]$, and $f_n \to f$ uniformly, then:

$$f \in L^1[a,b], \quad \text{and} \int_a^b f_n \to \int_a^b f.$$

- **Differentiability Not Preserved:** Even if each f_n is differentiable, and $f_n \to f$ uniformly, the limit may not be differentiable. However, if $f_n' \to g$ uniformly and $f_n(x_0) \to f(x_0)$, then $f_n \to f$ uniformly and $f' = g$.
- **Boundedness Preserved:** Uniform limits of bounded functions are bounded.

Function Spaces

Let $C_0(\mathbb{R}^d)$ denote the space of continuous functions $f : \mathbb{R}^d \to \mathbb{R}$ that vanish at infinity:

$$\forall \varepsilon > 0, \exists K \subset \mathbb{R}^d \text{ compact such that } |f(x)| < \varepsilon \text{ for all } x \notin K.$$

Theorem: $C_0(\mathbb{R}^d)$, with the sup norm $\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$, is a complete metric space.

Proof Sketch: Let $f_n \in C_0(\mathbb{R}^d)$ be a Cauchy sequence in $\|\cdot\|_\infty$. Then $f_n \to f$ uniformly, and hence $f \in C_b(\mathbb{R}^d)$. One checks that $f \in C_0(\mathbb{R}^d)$ using the ϵ–K definition above, since uniform convergence preserves vanishing at infinity.

Connectedness and Path Connectedness

Definition (Connectedness): A topological space X is *connected* if there do not exist disjoint non-empty open sets $U, V \subseteq X$ such that:

$$X = U \cup V.$$

Definition (Path Connectedness): A topological space X is *path connected* if for all $x, y \in X$, there exists a continuous map:

$$\gamma : [0,1] \to X \quad \text{with } \gamma(0) = x,\ \gamma(1) = y.$$

Theorems and Proofs

Theorem: The continuous image of a connected space is connected.

Proof: Let $f : X \to Y$ be continuous and X connected. Suppose $f(X) = U \cup V$, where $U, V \subseteq Y$ are disjoint non-empty open. Then $f^{-1}(U), f^{-1}(V)$ are open, disjoint, cover X, and non-empty $\Rightarrow$ contradiction. So $f(X)$ is connected.

Theorem: Path connected $\Rightarrow$ connected.

Proof: Let X be path connected. Suppose $X = U \cup V$, disjoint open non-empty. Pick $x \in U, y \in V$, and let $\gamma : [0,1] \to X$ be a path from x to y. Then $\gamma^{-1}(U), \gamma^{-1}(V)$ are disjoint open in $[0,1]$ and cover it, with $0 \in \gamma^{-1}(U)$, $1 \in \gamma^{-1}(V) \Rightarrow$ contradicts connectedness of $[0,1]$.

Theorem: The closure of a connected set is connected.

Proof: Let $A \subseteq X$ connected, and suppose $\overline{A} = U \cup V$ with U, V disjoint non-empty open in $\overline{A}$. Then $U \cap A, V \cap A$ are disjoint open in A, cover A, and non-empty $\Rightarrow$ contradiction.

Theorem (Finite Intersection Criterion): If $\{A_i\}_{i \in I}$ is a collection of connected subspaces with non-empty pairwise intersections and $X = \bigcup_{i \in I} A_i$, then X is connected.

Proof: If each A_i is connected and intersects a fixed A_0, their union is connected by induction using the fact that the union of two connected sets with non-empty intersection is connected.

Characterisation Theorem (Three-Way Equivalence)

Let X be a topological space. The following are equivalent:

1. X is connected.
2. Every continuous map $f : X \to \{0,1\}$ is constant.
3. Every continuous map $f : X \to \mathbb{Z}$ is constant.

Proof:

(i) ⇒ (ii): Suppose $f : X \to \{0,1\}$ is continuous. Then $f^{-1}(0)$, $f^{-1}(1)$ are open in X, disjoint, and cover X. If both are non-empty, this is a disconnection of X, contradicting connectedness. So f is constant.

(ii) ⇒ (iii): Let $f : X \to \mathbb{Z}$ be continuous. Then for each $n \in \mathbb{Z}$, the set $f^{-1}(n)$ is open, since $\mathbb{Z}$ has the discrete topology.

Since $X = \bigsqcup_{n\in\mathbb{Z}} f^{-1}(n)$, and the disjoint union of open sets is open, only one of these can be non-empty by (ii). Hence f is constant.

(iii) ⇒ (i): Suppose X is not connected. Then there exist disjoint non-empty open sets $U, V \subseteq X$ such that $X = U \cup V$. Define:

$$f(x) := \begin{cases} 0 & x \in U \\ 1 & x \in V \end{cases}$$

Then $f : X \to \mathbb{Z}$ is continuous (as $\mathbb{Z}$ is discrete), but not constant ⇒ contradiction.

Therefore, X must be connected.

Compactness in Topological Spaces

Definition (Compactness): A topological space X is *compact* if every open cover has a finite subcover:

$$\forall \{U_\alpha\}_{\alpha\in A} \text{ open with } X = \bigcup_\alpha U_\alpha, \quad \exists \alpha_1, \dots, \alpha_n$$

$$\text{such that } X = \bigcup_{i=1}^{n} U_{\alpha_i}.$$

Theorem: A closed subset of a compact space is compact.

Sketch Proof: Let $A \subseteq X$ be closed and X compact. Given an open cover of A, extend it with $X \setminus A$ to cover X. Extract a finite subcover — those not covering $X \setminus A$ already cover A.

Definition (Hausdorff): A space X is *Hausdorff* if for all $x \neq y$, there exist disjoint open sets U, V with $x \in U, y \in V$.

Theorem: If X is compact and Hausdorff, then every continuous bijection $f : X \to Y$ is a homeomorphism.

Sketch Proof: Show that f is closed: Let $A \subseteq X$ be closed. Then A is compact, and so $f(A)$ is compact in Y. If Y is Hausdorff, compact sets are closed ⇒ $f(A)$ is closed ⇒ f is closed ⇒ inverse is continuous.

Definition (Homeomorphism): A map $f : X \to Y$ is a homeomorphism if it is a bijective continuous map with continuous inverse.

Theorem: The quotient of a compact space is compact.

Sketch Proof: Let $\pi : X \to X/\sim$ be the quotient map. Given an open cover of $X/\sim$, pull back to get an open cover of X, which has a finite subcover. Push this forward to get a finite subcover of $X/\sim$.

Remark: Compactness is preserved under quotients. Hausdorffness is not: the quotient of a Hausdorff space need not be Hausdorff unless equivalence classes are closed.

Example (Non-Hausdorff Quotient): Identify all points of $[0,1]$ to a point. Quotient is compact but not Hausdorff.

Differentiability in Multivariable Calculus

Definition (Differentiability at a Point): Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a function, and let $a \in U$. We say that f is *differentiable at a* if there exists a linear map $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|}{\|h\|} = 0.$$

The matrix of $Df(a)$ in standard bases is the *Jacobian matrix*:

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1\le i\le m, 1\le j\le n}.$$

Note: Differentiability implies continuity. Partial derivatives existing does not imply differentiability unless they are continuous (i.e. $f \in C^1$).

Completeness of Metric Spaces

Definition (Complete Metric Space): A metric space (X,d) is *complete* if every Cauchy sequence converges to a limit in X.

Definition (Cauchy Sequence): A sequence $(x_n) \subseteq X$ is Cauchy if:

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } \forall m, n \ge N,\ d(x_n, x_m) < \varepsilon.$$

Examples:

- $\mathbb{R}^n$ with the Euclidean metric is complete.
- $C([a,b])$ with the sup norm is complete.
- Any closed subset of a complete metric space is complete.

Definitions

Continuity (Topological): A function $f : X \to Y$ between topological spaces is *continuous* if:

$$\forall V \subseteq Y \text{ open, } f^{-1}(V) \subseteq X \text{ is open.}$$

Differentiability (Euclidean): Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. f is *differentiable at $a \in U$* if there exists a linear map $Df(a)$ such that:

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|}{\|h\|} = 0.$$

Quotient Topology: Let X be a topological space, and $\sim$ an equivalence relation on X. The *quotient topology* on

$X/\sim$ is the finest topology such that the projection map $\pi : X \to X/\sim$ is continuous. Explicitly:

$$U \subseteq X/\sim \text{ is open } \iff \pi^{-1}(U) \text{ is open in } X.$$

Product Topology: Let $\{X_i\}_{i\in I}$ be a family of topological spaces. The *product topology* on $\prod_{i\in I} X_i$ is the coarsest topology such that all projections $\pi_j : \prod X_i \to X_j$ are continuous. A subbasis is given by products $\prod U_i$, where $U_i \subseteq X_i$ is open and $U_i = X_i$ for all but finitely many i.

Gluing Lemma

Theorem (Gluing Lemma): Let $X = A \cup B$ with $A, B \subseteq X$ closed (or open). Let $f : X \to Y$ be a function such that:

- $f|_A : A \to Y$ is continuous
- $f|_B : B \to Y$ is continuous
- $f|_{A\cap B}$ agrees on the overlap

Then $f : X \to Y$ is continuous.

Proof: Let $U \subseteq Y$ be open. Then:

$$f^{-1}(U) = (f|_A)^{-1}(U) \cup (f|_B)^{-1}(U)$$

Each preimage is open in A, B respectively, hence:

$$(f|_A)^{-1}(U) = A \cap V_1, \quad (f|_B)^{-1}(U) = B \cap V_2$$

for some open $V_1, V_2 \subseteq X$. So:

$$f^{-1}(U) = (A \cap V_1) \cup (B \cap V_2)$$

is open in X. Therefore, f is continuous.

Continuity: Sequential and Topological Criteria

Sequential Characterisation of Continuity: Let $f : X \to Y$ be a function between topological spaces.

Theorem: If X is first countable, then:

$$f \text{ is continuous at } x \in X \iff \forall (x_n) \to x,\ f(x_n) \to f(x).$$

Proof Sketch: "$\Rightarrow$": Follows from openness of inverse images. "$\Leftarrow$": Assume inverse image of open $V \subseteq Y$ is not open in X. Then there exists a sequence $x_n \to x \in f^{-1}(V)$ with $f(x_n) \notin V$, contradicting $f(x_n) \to f(x) \in V$.

Closure Characterisation of Continuity:

Theorem: $f : X \to Y$ is continuous iff for every $A \subseteq X$,

$$f(\overline{A}) \subseteq \overline{f(A)}.$$

Proof Sketch: Let $x \in \overline{A}$. Then every open neighbourhood of x intersects A, so every neighbourhood of $f(x)$ intersects $f(A)$, implying $f(x) \in \overline{f(A)}$.

Continuity in Function Spaces

Let (Y, d) be a metric space and $C(X, Y)$ the set of continuous functions from $X \to Y$, endowed with the sup norm:

$$\|f - g\|_\infty = \sup_{x\in X} d(f(x), g(x)).$$

Theorem: If (Y, d) is complete and X is compact, then $(C(X, Y), \|\cdot\|_\infty)$ is complete.

Proof Sketch: Let $(f_n) \subset C(X, Y)$ be a Cauchy sequence. Then for each x, $f_n(x)$ is Cauchy in Y, so converges to $f(x) \in Y$. Define $f : X \to Y$, show uniform convergence, and use uniform limit of continuous functions is continuous $\Rightarrow f \in C(X, Y)$.

Continuity in Product Spaces:

Theorem: A map $f : Z \to \prod_\alpha X_\alpha$ is continuous iff each composition $\pi_\alpha \circ f : Z \to X_\alpha$ is continuous.

Proof Sketch: Follows by subbasis definition of product topology: basic open sets are preimages under projections $\Rightarrow$ continuity of f is equivalent to continuity of each coordinate function.

Topological vs Non-Topological Properties

Topological Property: A property of a space X that is preserved under homeomorphism. If $X \cong Y$ (i.e., there exists a homeomorphism), then X has the property $\iff$ Y does.

Examples of Topological Properties:

- Connectedness
- Compactness
- Hausdorffness
- Second countability
- Local connectedness / local compactness
- Continuity of maps

Non-Topological Properties: Depend on more than the topology — e.g., algebraic, geometric, metric.

Examples of Non-Topological Properties:

- Metrizability (unless specified otherwise)
- Boundedness (not preserved under homeomorphism)
- Total disconnectedness (in some settings)
- Smoothness / differentiability
- Distance and angles

Remark: Topological properties are defined in terms of open sets, closures, and continuous functions. Non-topological properties usually require extra structure (metric, vector space, etc.).

Weierstrass M-Test and Power Series

Weierstrass M-Test: Let $f_n : X \to \mathbb{R}$ be functions with:

$$|f_n(x)| \le M_n \quad \forall x \in X, \quad \text{and} \sum M_n < \infty.$$

Then $\sum f_n(x)$ converges uniformly and absolutely on X, and the sum is continuous if all f_n are.

Local Uniform Convergence of Power Series: Let $f(x) = \sum_{n=0}^\infty a_n x^n$ with radius of convergence $R > 0$. Then:

$$\sum a_n x^n \text{ converges uniformly on every compact subset of } (-R, R).$$

$\Rightarrow$ Power series define continuous functions; term-by-term differentiation/integration valid on compact subsets within $(-R, R)$.