Recap 18.03

Tristan Ozuch

Contents

1 Introduction to differential equations

1.1 Modeling

1. Identify relevant quantities (known/unknown) to study and assign symbols.

2. Identify the variable(s) these quantities depend upon (often time t).
3. Write equations describing how the functions change (rates of change) and any laws relating the variables.

1.2 Introduction to differential equations

Differential equations. A *differential equation (DE)* is an equation relating an unknown function and some of its derivatives, e.g.

$$y' = 3y.$$

Notation for derivatives (with respect to t).

$$y' = \dot{y} = \frac{dy}{dt}, \quad y'' = \ddot{y} = \frac{d^2y}{dt^2}, \quad y^{(n)} = \frac{d^ny}{dt^n}.$$

If y has units of meters and t of seconds, then

$$y' \text{ has units m/s}, \qquad y'' \text{ has units m/s}^2.$$

General and particular solutions. For the DE

$$y' = 3y,$$

a *particular solution* is $t \mapsto -17e^{3t}$ the *general solution* is the 1-parameter family

$$y(t) = ce^{3t}, \qquad c \in \mathbb{R}.$$

An *initial value problem (IVP)* specifies an initial condition, e.g.

$$y(0) = 6 \quad \Rightarrow \quad y(t) = 6e^{3t}.$$

1.3 Classification of differential equations

ODE vs PDE.

- *Ordinary differential equation (ODE):* derivatives of a function of one variable, e.g.

$$\ddot{y} = -9y \quad \text{(solve for } y(t)\text{)}.$$

- *Partial differential equation (PDE):* partial derivatives of a multivariable function, e.g.

$$\frac{\partial u}{\partial t} = 9\frac{\partial^2 u}{\partial x^2} \quad \text{(solve for } u(x,t)\text{)}.$$

Order. The *order* of a DE is the highest derivative that appears. For an ODE in $y(t)$, if $y^{(n)}$ appears and no higher derivative does, the order is n.

2 First order ODEs and introduction to linear ODEs

2.1 Separation of variables

Separable first-order ODE.

$y' = f(t, y)$ is *separable* if it can be written as $y' = g(t)\,h(y)$.

Rewrite as

$$\frac{dy}{dt} = g(t)\,h(y) \quad \Longleftrightarrow \quad \frac{1}{h(y)}\,dy = g(t)\,dt.$$

(be careful not to divide by zero!).

Integrate:

$$\int \frac{1}{h(y)}\,dy = \int g(t)\,dt \quad \Longrightarrow \quad F(y) = G(t) + C.$$

(If possible) solve for $y(t)$ to obtain the general solution.

2.2 Linear ODEs

General linear ODE (order n).

$$p_n(t)y^{(n)} + p_{n-1}(t)y^{(n-1)} + \cdots + p_1(t)y' + p_0(t)y = q(t),$$

where $p_n, \ldots, p_0, q$ depend only on t.

Homogeneous vs. inhomogeneous.

$$\text{Homogeneous: } q(t) \equiv 0, \qquad \text{Inhomogeneous: } q(t) \not\equiv 0.$$

The function $t \mapsto y(t) = 0$ is solution to all homogeneous equations, but it is never a solution to an inhomogeneous equation. If $t \mapsto y(t)$ is a solution to a homogeneous equation, then so is $t \mapsto c\,y(t)$ for any $c \in \mathbb{R}$.

Standard linear form. Divide by the leading coefficient $p_n(t) \neq 0$ (be careful not to divide by zero!):

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y = b(t).$$

First-order linear ODE (standard form).

$$y' + p(t)\,y = q(t).$$

Nonlinear ODE examples. Nonlinear if $y, y', \ldots$ appear with powers/products or inside nonlinear functions, e.g.

$$y' = y^2, \quad y' = \cos(y + t), \quad y'y'' - ty = 0.$$

2.3 Solving a first-order linear ODE

Consider

$$y' + p(t)\,y = q(t), \qquad P(t) = \int p(t)\,dt.$$

Homogeneous case ($q \equiv 0$).

Theorem 2.1.

$$y' + p(t)\,y = 0 \quad \Longrightarrow \quad y_h(t) = c\,e^{-P(t)}, \text{ for some } c \in \mathbb{R}.$$

Inhomogeneous case: variation of parameters. Seek $y(t) = u(t)\,y_h(t)$ with $y_h(t) = e^{-P(t)} \neq 0$. Plugging into $y' + p(t)y = q(t)$ gives

$$u'(t)\,y_h(t) = q(t) \quad \Longrightarrow \quad u'(t) = q(t)\,e^{P(t)},$$

$$u(t) = \int q(t)\,e^{P(t)}\,dt.$$

hence a *particular solution is given by*

$$y_p(t) = e^{-P(t)} \int q(t)\,e^{P(t)}\,dt.$$

(this is a formula you can use directly, but the method will apply to more complicated equations).

General solution.

Theorem 2.2.

$$y(t) = y_p(t) + y_h(t) = e^{-P(t)}\left(C + \int q(t)\,e^{P(t)}\,dt\right).$$

2.4 Linear combinations and superposition

Linear combination of functions. For functions $f_1, \dots, f_k$,

$$c_1 f_1(t) + \cdots + c_k f_k(t) \quad (c_1, \dots, c_k \in \mathbb{R})$$

is a *linear combination*.

Superposition principle (homogeneous). Let L be a *linear* differential operator (e.g. $L[y] = y' + p(t)y$). If $L[y_1] = 0$ and $L[y_2] = 0$, then

$$L[c_1 y_1 + c_2 y_2] = 0 \quad \text{for all } c_1, c_2.$$

General / particular solution rule (inhomogeneous). If

$$L[y] = q(t), \qquad L[y_h] = 0, \qquad L[y_p] = q(t),$$

then

$$y(t) = y_p(t) + y_h(t)$$

is the general solution.

2.5 Strategy for solving first-order ODEs

1. First, find the *general* solution.
    - Try separation of variables (if the ODE is separable).
    - If the ODE is linear:
        - Homogeneous: $y' + p(t)y = 0 \Rightarrow$ separation of variables/solve directly.
        - Inhomogeneous: $y' + p(t)y = q(t)$:

    - ∗ Guess a particular solution y_p (and check), then use $y = y_p + y_h$, or
    - ∗ Use the above formula for y_p (which might be complicated), or
    - ∗ Use the variation of parameters method (if the equation is higher order) to find some y_p.

2. Finally, apply the initial condition(s) to determine the constant(s) in the homogeneous solution y_h.

2.6 Existence and uniqueness for linear ODEs

Theorem 2.3. Let $p_{n-1}(t), \dots, p_0(t), q(t)$ be continuous on an open interval $I \subset \mathbb{R}$, and consider

$$y^{(n)} + p_{n-1}(t)y^{(n-1)} + \cdots + p_1(t)y' + p_0(t)y = q(t).$$

Given $a \in I$ and numbers $b_0, \dots, b_{n-1}$, there is exactly one solution $y(t)$ on I such that the following *initial conditions* are satisfied

$$y(a) = b_0, \quad y'(a) = b_1, \quad \dots, \quad y^{(n-1)}(a) = b_{n-1}.$$

3 Complex numbers and complex exponential

3.1 Basic properties

Complex numbers. $z = a + ib, \quad a, b \in \mathbb{R}, \quad i^2 = -1. \qquad \Re z = a, \quad \Im z = b.$

Conjugate and modulus.

$$\overline{z} = a - ib, \qquad z\,\overline{z} = |z|^2, \qquad |z| = \sqrt{a^2 + b^2}.$$

$$|zw| = |z|\,|w|, \qquad \overline{z + w} = \overline{z} + \overline{w}, \quad \overline{zw} = \overline{z}\,\overline{w}.$$

3.2 Roots of polynomials

Complex polynomial of degree *n*. $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0, \quad a_j \in \mathbb{C},\ a_n \neq 0.$

Fundamental Theorem of Algebra. Every polynomial of degree $n \geq 1$ with complex coefficients has at least one root in $\mathbb{C}$, and exactly n roots counted with multiplicity.

Quadratic equations. To solve $az^2 + bz + c = 0, \qquad a \neq 0$, define the discriminant $\Delta = b^2 - 4ac$. Then

$$z = \frac{-b \pm \sqrt{\Delta}}{2a},$$

where $\sqrt{\Delta}$ is the complex square root. If $a, b, c \in \mathbb{R}$ and $\Delta < 0$, the two roots are a complex conjugate pair.

3.3 Complex exponential

Definition. $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad z \in \mathbb{C}.$

Euler's formula. For $t \in \mathbb{R}$, $e^{it} = \cos t + i \sin t$. That is $\cos t = \Re(e^{it})$ and $\sin t = \Im(e^{it})$.

General complex exponential. For $z = x + iy$,

$$e^{x+iy} = e^x(\cos y + i \sin y).$$

Cosine and sine via exponentials. $\cos t = \frac{e^{it} + e^{-it}}{2}, \qquad \sin t = \frac{e^{it} - e^{-it}}{2i}.$

3.4 Polar form of a complex number

Polar form. For $z \neq 0$,

$$z = r(\cos\theta + i \sin\theta) = re^{i\theta}, \quad r = |z| \geq 0,\ \theta \in \mathbb{R}.$$

$$\theta \text{ determined up to } 2\pi k,\ k \in \mathbb{Z}.$$

Multiplication. If $z_1 = r_1 e^{i\theta_1},\ z_2 = r_2 e^{i\theta_2}, z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$

Roots of complex numbers. Solve $w^n = z \neq 0$. Write $z = re^{i\theta}$ with $r > 0$. Then the n distinct n^{th} roots are

$$w_k = r^{1/n} e^{i(\theta + 2k\pi)/n}, \quad k = 0, \ldots, n - 1.$$

Roots of unity. Solutions of $z^n = 1$ are called "n^{th} roots of unity": $z_k = e^{2\pi i k/n}, \quad k = 0, \dots, n-1$.

Complex solutions to real ODEs. Let L be a linear differential operator with *real* coefficients (e.g. $L[y] = y'' + p(t)y' + q(t)y$ with real p, q). If $y : \mathbb{R} \to \mathbb{C}$ satisfies $L[y] = 0$, then:

$$L[\overline{y}] = 0, \qquad L[\Re y] = 0, \qquad L[\Im y] = 0.$$

Conversely, if y_1, y_2 are real solutions of $L[y] = 0$, then

$$y = y_1 + i y_2$$

is a complex solution. If y and $\overline{y}$ are linearly independent complex solutions, then $\Re y$ and $\Im y$ form a real basis of the solution space.

In particular, if $y(t) = e^{zt}$ with $z = a + ib$ solves a linear ODE with *real* coefficients, then so do

$$t \mapsto e^{at}\cos(bt) \qquad \text{and} \qquad t \mapsto e^{at}\sin(bt).$$

4 Linear ODEs with constant coefficients: second order homogeneous case

Equation of motion. $m\ddot{x} = -kx - b\dot{x} + F_{\text{ext}}(t) \iff m\ddot{x} + b\dot{x} + kx = F_{\text{ext}}(t).$

Homogeneous damped harmonic oscillator: $m\ddot{x} + b\dot{x} + kx = 0.$

4.1 Superposition for second order homogeneous ODEs

General second order homogeneous linear ODE (constant coefficients).

$$a_2\ddot{y} + a_1\dot{y} + a_0 y = 0, \qquad a_2 \neq 0.$$

Solution space. Let S be the set of all solutions $y : \mathbb{R} \to \mathbb{R}$. Then:

- If $y_1, y_2 \in S$ and $c_1, c_2 \in \mathbb{R}$, then $c_1 y_1 + c_2 y_2 \in S$.
- S is a 2-dimensional vector space.

Span, linear independence, basis.

- $\text{Span}(y_1, y_2) = \{c_1 y_1 + c_2 y_2 : c_1, c_2 \in \mathbb{R}\}$.
- y_1, y_2 *linearly independent* if neither is a scalar multiple of the other.
- If y_1, y_2 are linearly independent solutions, then every solution is

  $$y(t) = c_1 y_1(t) + c_2 y_2(t),$$

  so $\{y_1, y_2\}$ is a basis of S.

Example: harmonic oscillator. $\ddot{x} + x = 0.$
Independent solutions: $\cos t,\ \sin t$. General solution:

$$x(t) = a\cos t + b\sin t.$$

4.2 General solution via characteristic polynomial

Characteristic equation. For

$$a_2\ddot{y} + a_1\dot{y} + a_0 y = 0,$$

look for $y(t) = e^{rt}$. This leads to

$$a_2 r^2 + a_1 r + a_0 = 0$$

(the *characteristic equation*).

Distinct real roots. If $r_1 \neq r_2$ are real roots, then

$$y(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}.$$

Repeated real root. If $r_1 = r_2 = r$ (double root), then

$$y(t) = (c_1 + c_2 t)e^{rt}.$$

Complex conjugate roots. If roots are $r_{1,2} = \alpha \pm i\beta$ with $\beta > 0$, then a complex basis is $e^{(\alpha+i\beta)t}, e^{(\alpha-i\beta)t}$. A real basis is

$$e^{\alpha t}\cos(\beta t), \quad e^{\alpha t}\sin(\beta t),$$

so

$$y(t) = e^{\alpha t}\big(c_1\cos(\beta t) + c_2\sin(\beta t)\big).$$

4.3 Sinusoidal functions

General sinusoid.

$$y(t) = A\cos(\omega t + \varphi), \quad A \geq 0,\ \omega > 0.$$

Parameters.

- Amplitude: A.
- Angular frequency: ω (rad/s).
- Period: $T = \dfrac{2\pi}{\omega}$.
- Frequency: $f = \dfrac{1}{T} = \dfrac{\omega}{2\pi}$.
- Phase (phase shift): φ (radians).
- Time lag: write $A\cos(\omega t - \delta) = A\cos\big(\omega(t - t_0)\big)$ with

$$t_0 = \frac{\delta}{\omega}.$$

Forms of a sinusoid.

- **Amplitude-phase form:**

$$y(t) = A\cos(\omega t - \delta).$$

- **Linear combination form:**

$$y(t) = c_1\cos(\omega t) + c_2\sin(\omega t),$$

  with

$$A = \sqrt{c_1^2 + c_2^2}, \quad c_1 = A\cos\delta, \quad c_2 = A\sin\delta.$$

- **Complex form:**

$$y(t) = \Re\big(Ce^{i\omega t}\big), \quad C \in \mathbb{C}.$$

Beats (optional). Sum of two close frequencies:

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\Big(\frac{\omega_1 - \omega_2}{2}t\Big)\cos\Big(\frac{\omega_1 + \omega_2}{2}t\Big).$$

4.4 Harmonic oscillators and damped frequency

Damped oscillator.

$$m\ddot{x} + b\dot{x} + kx = 0, \quad m > 0,\ k > 0,\ b \geq 0.$$

Characteristic equation:

$$mr^2 + br + k = 0.$$

Natural and damped frequencies.

$$\omega_0 = \sqrt{\frac{k}{m}} \quad \text{(natural frequency).}$$

If $0 < b < 2\sqrt{mk}$ (underdamped), roots are

$$r = -\frac{b}{2m} \pm i\omega_d, \quad \omega_d = \sqrt{\omega_0^2 - \left(\frac{b}{2m}\right)^2},$$

and

$$x(t) = e^{-\frac{b}{2m}t}\big(A\cos(\omega_d t - \delta)\big).$$

Cases.

- **Undamped:** $b = 0$.
$$\ddot{x} + \omega_0^2 x = 0, \quad x(t) = A\cos(\omega_0 t - \delta).$$
- **Underdamped:** $0 < b < 2\sqrt{mk}$: oscillatory, exponentially decaying.
- **Critically damped:** $b = 2\sqrt{mk}$ (double root $r = -\omega_0$):
$$x(t) = (c_1 + c_2 t)e^{-\omega_0 t},$$
fastest non-oscillatory return to equilibrium.
- **Overdamped:** $b > 2\sqrt{mk}$: two distinct real negative roots r_1, r_2,
$$x(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t},$$
slow non-oscillatory return.

5 Linear ODEs with constant coefficients: general case

5.1 Vector spaces of solutions

Span. Given functions $f_1, \dots, f_n$, their span is

$$\mathrm{Span}(f_1, \dots, f_n) := \{c_1 f_1 + \cdots + c_n f_n \mid c_1, \dots, c_n \in \mathbb{R} \text{ (or } \mathbb{C})\}.$$

Linear independence. $f_1, \dots, f_n$ are linearly independent if

$$c_1 f_1 + \cdots + c_n f_n = 0 \quad \Rightarrow \quad c_1 = \cdots = c_n = 0.$$

Basis. A list $(f_1, \dots, f_n)$ is a basis of a vector space S if

$$S = \mathrm{Span}(f_1, \dots, f_n) \quad \text{and} \quad f_1, \dots, f_n \text{ are linearly independent.}$$

Then every $f \in S$ can be written uniquely as $f = c_1 f_1 + \cdots + c_n f_n$ (coordinates).

Dimension. The dimension of a vector space is the number of elements in any basis.

Dimension theorem for homogeneous linear ODEs. For an nth order homogeneous linear ODE

$$a_n(t) y^{(n)} + \cdots + a_1(t) y' + a_0(t) y = 0,$$

the solution space is an n-dimensional vector space. Hence the general solution depends on n parameters.

5.2 Solving a homogeneous linear ODE with constant coefficients

Characteristic polynomial. For

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0, \quad a_n \neq 0,$$

with constant coefficients, define

$$p(r) := a_n r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0.$$

The numbers r with $p(r) = 0$ are the *characteristic roots.*

Real roots. If r is a real root of p of multiplicity m, then

$$e^{rt},\ te^{rt},\ t^2 e^{rt}, \dots, t^{m-1} e^{rt}$$

are m linearly independent solutions.

Complex roots, real coefficients. If coefficients are real and $r = \alpha + i\beta$ ($\beta \neq 0$) is a root of multiplicity m, then $\bar{r} = \alpha - i\beta$ is also a root of multiplicity m, and the associated real solutions can be chosen as

$$t^k e^{\alpha t} \cos(\beta t), \quad t^k e^{\alpha t} \sin(\beta t), \qquad k = 0, \dots, m-1.$$

General solution. If the distinct roots of p are $r_1, \dots, r_k$ with multiplicities $m_1, \dots, m_k$ (so $\sum m_j = n$), then a basis of solutions is obtained by taking, for each r_j,

$$e^{r_j t},\ te^{r_j t}, \dots, t^{m_j - 1} e^{r_j t},$$

and the general solution is their linear combination.

5.3 Operator notation for linear ODEs

Differential operator. Let $D := \frac{d}{dt}$. This lets us rewrite

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = q(t),$$

as

$$p(D)y = q(t), \quad p(D) := a_n D^n + \cdots + a_1 D + a_0.$$

Linearity. For functions y_1, y_2 and scalars c_1, c_2,

$$p(D)(c_1 y_1 + c_2 y_2) = c_1 p(D) y_1 + c_2 p(D) y_2.$$

Action on exponentials. For any number r,

$$D^k e^{rt} = r^k e^{rt}, \qquad p(D) e^{rt} = p(r)\, e^{rt}.$$

This connects the differential equation to the characteristic polynomial.

5.4 Inhomogeneous linear ODEs

General structure. For $p(D)y = f(t)$, let y_h be the general solution of $p(D)y = 0$ and y_p any particular solution of $p(D)y = f(t)$. Then

$$y = y_h + y_p$$

is the general solution.

Exponential Response Formula (ERF). For $p(D)y = e^{rt}$ with $p(r) \neq 0$, a particular solution is

$$y_p(t) = \frac{1}{p(r)} e^{rt}.$$

For a linear combination $\sum_j A_j e^{r_j t}$ with each $p(r_j) \neq 0$,

$$y_p(t) = \sum_j \frac{A_j}{p(r_j)} e^{r_j t}.$$

Generalized ERF (resonant case). If r is a root of p of multiplicity m, then a particular solution of $p(D)y = e^{rt}$ is

$$y_p(t) = \frac{1}{p^{(m)}(r)}\, t^m e^{rt}, \qquad \textbf{no factor } m! \textbf{ at the denominator}$$

where $p^{(m)}$ is the mth derivative of p. Similarly, if factorizes as $p(x) = (x - r)^m q(x)$ with $q(r) \neq 0$, then we have:

$$y_p(t) = \frac{1}{m!\, q(r)}\, t^m e^{rt}, \qquad \textbf{factor } m! \textbf{ at the denominator}$$

Complex replacement for sinusoidal forcing. For real p and ω,

$$p(D)x = \cos(\omega t).$$

Method:

1. Solve the complex ODE
$$p(D)z = e^{i\omega t}$$
(using ERF or generalized ERF).
2. Take $x_p(t) := \Re z_p(t)$; then x_p is a particular solution of the original ODE.

More generally, for input $q(t) = A_{\text{in}} \cos(\omega t - \phi_{\text{in}}) = \Re\left(c_{\text{in}} e^{i\omega t}\right)$, $\quad c_{\text{in}} = A_{\text{in}} e^{-i\phi_{\text{in}}}$, solve

$$p(D)z = c_{\text{in}} e^{i\omega t}$$

to get

$$z_p(t) = \frac{c_{\text{in}}}{p(i\omega)} e^{i\omega t} = c_{\text{out}} e^{i\omega t}.$$

Complex gain, gain, phase lag. Define the complex gain

$$G(\omega) := \frac{1}{p(i\omega)}, \qquad c_{\text{out}} = G(\omega)\, c_{\text{in}}.$$

Write $G(\omega) = |G(\omega)| e^{-i\theta(\omega)}$. Then the steady-state output is

$$x_{\text{ss}}(t) = A_{\text{out}} \cos(\omega t - \phi_{\text{out}}),$$

with

$$A_{\text{out}} = |G(\omega)| A_{\text{in}}, \qquad \phi_{\text{out}} = \phi_{\text{in}} + \theta(\omega),$$

and

$$|G(\omega)| = \frac{1}{|p(i\omega)|}, \qquad \theta(\omega) = -\arg G(\omega) = \arg p(i\omega).$$

5.5 Stability

Transient and steady-state parts. For $p(D)x = f(t)$ with constant coefficients,

$$x(t) = x_{\text{tr}}(t) + x_{\text{ss}}(t),$$

where x_{tr} solves $p(D)x = 0$ (transient) and x_{ss} is a particular solution determined by f (steady-state).

Stability via characteristic roots. Consider the homogeneous equation

$$p(D)x = 0,$$

with characteristic roots $r_1, \ldots, r_n$.

- If $\Re(r_j) < 0$ for all j, then every solution tends to 0 as $t \to +\infty$ (asymptotically stable).
- If some r_j has $\Re(r_j) > 0$, then some solutions grow without bound (unstable).

Second-order test. For $mx'' + bx' + kx = 0, \quad m > 0$, all solutions are asymptotically stable if both roots have negative real part, i.e.

$$b > 0, \quad k > 0.$$

5.6 Resonance

Frequency response. For a sinusoidal input $A_{\text{in}} \cos(\omega t - \phi_{\text{in}})$, the steady-state output has amplitude

$$A_{\text{out}}(\omega) = |G(\omega)| A_{\text{in}}, \qquad G(\omega) = \frac{1}{p(i\omega)},$$

and phase lag $\theta(\omega)$ as above. The function $\omega \mapsto |G(\omega)|$ is the gain (magnitude) of the frequency response.

Near resonance (undamped case). Example: $x'' + \omega_0^2 x = \cos(\omega t)$ has

$$p(r) = r^2 + \omega_0^2, \quad G(\omega) = \frac{1}{\omega_0^2 - \omega^2}, \quad |G(\omega)| = \frac{1}{|\omega_0^2 - \omega^2|}.$$

As $\omega \to \omega_0$, $|G(\omega)| \to \infty$, so the steady-state amplitude becomes arbitrarily large.

Pure resonance (no damping, exact frequency). If $p(i\omega_0) = 0$ (e.g. $x'' + \omega_0^2 x = \cos(\omega_0 t)$), ERF does not apply. Generalized ERF gives a particular solution of the form

$$x_p(t) = Ct \sin(\omega_0 t) \quad \text{or} \quad Ct \cos(\omega_0 t),$$

whose amplitude grows linearly in t (unbounded, non-periodic response).

Resonance with damping. For $x'' + bx' + \omega_0^2 x = \cos(\omega t), \quad b > 0$,

$$p(r) = r^2 + br + \omega_0^2, \quad G(\omega) = \frac{1}{(\omega_0^2 - \omega^2) + ib\omega}.$$

Then

$$|G(\omega)| = \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + (b\omega)^2}},$$

which has a finite peak (large but bounded amplitude) near a frequency close to ω_0.

5.7 RLC circuits

Series RLC equation. Let $Q(t)$ be the charge on the capacitor in a series RLC circuit with inductance L, resistance R, capacitance C, driven by a voltage $V(t)$. Kirchhoff's law gives

$$LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = V(t).$$

Analogy with mass-spring-dashpot. Compare

$$mx'' + bx' + kx = F(t) \quad \leftrightarrow \quad LQ'' + RQ' + \frac{1}{C}Q = V(t),$$

with the identifications

$$x \leftrightarrow Q, \quad m \leftrightarrow L, \quad b \leftrightarrow R, \quad k \leftrightarrow \frac{1}{C}, \quad F(t) \leftrightarrow V(t).$$

All results on natural frequency, damping, stability, resonance, gain, and phase carry over via this dictionary.

6 Systems of linear ODEs with constant coefficients

6.1 Motivation: a two-loop circuit

See Bjorn Poonen's notes.

6.2 Linear algebra for 2×2 systems

Column vector: $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

Matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}, \qquad A\mathbf{x} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \in \mathbb{R}^2.$$

Identity:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad AI = IA = A.$$

Trace and determinant:

$$\operatorname{tr} A = a + d, \qquad \det A = ad - bc.$$

Characteristic polynomial and eigenvalues:

$$p_A(\lambda) = \det(\lambda I - A) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A.$$

Eigenvalues λ_1, λ_2 are the roots of $p_A(\lambda) = 0$:

$$\lambda_{1,2} = \frac{\operatorname{tr} A \pm \sqrt{(\operatorname{tr} A)^2 - 4 \det A}}{2}, \quad \lambda_1 + \lambda_2 = \operatorname{tr} A, \quad \lambda_1 \lambda_2 = \det A.$$

Eigenvector for eigenvalue λ: solve the system

$$\mathbf{v} \neq 0, \quad A\mathbf{v} = \lambda \mathbf{v} \iff (A - \lambda I)\mathbf{v} = 0.$$

6.3 Basics of systems

General first-order system in two unknowns:

$$\begin{cases} x' = f(t, x, y), \\ y' = g(t, x, y), \end{cases} \qquad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \ \mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}, \ F(t, \mathbf{x}) = \begin{pmatrix} f(t, x, y) \\ g(t, x, y) \end{pmatrix}.$$

Vector form: $\mathbf{x}' = F(t, \mathbf{x})$.

Linear systems. Linear (in (x, y)) system:

$$\mathbf{x}' = A(t)\mathbf{x} + \mathbf{b}(t),$$

where $A(t)$ is a 2×2 matrix, $\mathbf{b}(t)$ a vector.

- *Homogeneous*: $\mathbf{b}(t) \equiv 0$, so $\mathbf{x}' = A(t)\mathbf{x}$.
- *Constant coefficients*: $A(t) \equiv A$ constant.

Dimension of the solution space. For a homogeneous linear system of 2 first-order ODEs, the solution space is 2-dimensional: any two linearly independent solutions $\mathbf{x}_1, \mathbf{x}_2$ give the general solution

$$\mathbf{x}(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t).$$

Converting between a 2nd-order ODE and a 2×2 system. Given

$$x'' + ax' + bx = 0,$$

let $x_1 = x$, $x_2 = x'$:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Conversely, from a 2×2 system $\mathbf{x}' = A\mathbf{x}$ (with A constant), one can eliminate a variable to obtain a scalar 2nd-order ODE.

6.4 Homogeneous 2×2 systems with constant coefficients

Consider

$$\mathbf{x}' = A\mathbf{x}, \quad A \text{ constant } 2 \times 2.$$

Exponential solutions. Guess $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$:

$$\lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}A\mathbf{v} \implies (A - \lambda I)\mathbf{v} = 0, \quad \det(\lambda I - A) = 0.$$

So eigenvalues of A give possible λ, and eigenvectors give the corresponding $\mathbf{v}$.

Distinct real eigenvalues. If $\lambda_1 \neq \lambda_2$ are real eigenvalues with independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2$, then

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2.$$

Complex eigenvalues. If eigenvalues are $\lambda_{1,2} = a \pm bi$ ($b \neq 0$), and $\mathbf{w}$ is a complex eigenvector for $\lambda = a + bi$, then

$$\mathbf{z}(t) = e^{(a+bi)t}\mathbf{w}$$

is a complex solution. Real solutions:

$$\mathbf{x}_1(t) = \Re\mathbf{z}(t), \quad \mathbf{x}_2(t) = \Im\mathbf{z}(t),$$

and the general real solution is

$$\mathbf{x}(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t).$$

Repeated real eigenvalue. If λ is a real eigenvalue with algebraic multiplicity 2:

- If $A = \lambda I$ (2D eigenspace), any two independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ give

  $$\mathbf{x}(t) = e^{\lambda t}\big(c_1\mathbf{v}_1 + c_2\mathbf{v}_2\big).$$

- If there is only one eigenvector $\mathbf{v}$, choose $\mathbf{w}$ such that $(A - \lambda I)\mathbf{w} = \mathbf{v}$. Then

  $$\mathbf{x}_1(t) = e^{\lambda t}\mathbf{v}, \quad \mathbf{x}_2(t) = e^{\lambda t}(t\mathbf{v} + \mathbf{w}),$$

  and the general solution is $c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t)$.

6.5 Phase portraits

For $\mathbf{x}' = A\mathbf{x}$, the origin is an equilibrium. Each solution gives a trajectory in the (x, y)-plane (phase plane). Eigenvectors give *eigenlines* (invariant lines through the origin).

Classification by eigenvalues (qualitative picture):

- **Real, distinct:**
  - $\det A < 0$ (eigenvalues of opposite signs): saddle.
  - $\det A > 0$, $\operatorname{tr} A > 0$ (both > 0): repelling node.
  - $\det A > 0$, $\operatorname{tr} A < 0$ (both < 0): attracting node.
  - $\det A = 0$, $\operatorname{tr} A \neq 0$: comb.
- **Complex:** eigenvalues $a \pm bi$, $b \neq 0$.
  - $a = 0$: center (closed ellipses).
  - $a > 0$: repelling spiral.
  - $a < 0$: attracting spiral.
- **Repeated real:** λ with multiplicity 2.
  - $\lambda > 0$: repelling degenerate/star node.
  - $\lambda < 0$: attracting degenerate/star node.

Direction of rotation (for complex eigenvalues) is determined by the direction of $\mathbf{x}' = A\mathbf{x}$ at a convenient point, e.g. at $(1, 0)$.

Trace–determinant plane Characteristic polynomial:

$$p_A(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A.$$

Discriminant:

$$\Delta = (\operatorname{tr} A)^2 - 4 \det A.$$

- $\Delta > 0$ (below the parabola $\det = \frac{1}{4}(\operatorname{tr} A)^2$): real, distinct eigenvalues.
- $\Delta < 0$ (above the parabola): complex conjugate eigenvalues.
- $\Delta = 0$: repeated real eigenvalue.

Summary (in terms of $\operatorname{tr} A$ and $\det A$):

- $\det < 0$: saddle.
- $\det > 0$, $\Delta > 0$, $\operatorname{tr} > 0$: repelling node.
- $\det > 0$, $\Delta > 0$, $\operatorname{tr} < 0$: attracting node.
- $\det > 0$, $\Delta < 0$, $\operatorname{tr} > 0$: repelling spiral.
- $\det > 0$, $\Delta < 0$, $\operatorname{tr} < 0$: attracting spiral.
- $\det > 0$, $\operatorname{tr} = 0$: center.
- $\det = 0$, $\operatorname{tr} \neq 0$: comb.

6.6 Stability

Stability. For $\mathbf{x}' = A\mathbf{x}$:

- *Stable* $\iff$ all trajectories tend to 0 as $t \to +\infty$.
- *Unstable* $\iff$ some trajectory is unbounded as $t \to +\infty$.
- *Semistable* (neutrally stable): all trajectories are bounded, but not all tend to 0 (e.g. a center).

Tests:

$$\text{stable} \iff \Re\lambda < 0 \text{ for all eigenvalues } \lambda \iff p_A(\lambda) \text{ has positive coefficients} \iff \operatorname{tr} A < 0,\ \det A > 0.$$

Borderline semistable 2×2 cases (all $\Re\lambda \le 0$, some $\Re\lambda = 0$):

- $\operatorname{tr} A < 0$, $\det A = 0$: distinct real eigenvalues, semistable.
- $\operatorname{tr} A = 0$, $\det A > 0$: non-real eigenvalues, center, semistable.
- $\operatorname{tr} A = \det A = 0$, $A = 0$: all solutions constant, semistable.
- $\operatorname{tr} A = \det A = 0$, $A \neq 0$: unstable.

Analyzing nonlinear systems

Consider an autonomous system

$$\begin{cases} x' = f(x, y), \\ y' = g(x, y). \end{cases}$$

Nullclines.

- The *x-nullcline* is the set
$$f(x, y) = 0.$$
  On this set, $x' = 0$: the vector field is vertical.
- The *y-nullcline* is the set
$$g(x, y) = 0.$$
  On this set, $y' = 0$: the vector field is horizontal.

Away from the nullclines, the signs of f and g determine the direction of motion:

$$x' > 0 \Rightarrow \text{motion to the right}, \qquad x' < 0 \Rightarrow \text{motion to the left},$$

$$y' > 0 \Rightarrow \text{motion upward}, \qquad y' < 0 \Rightarrow \text{motion downward}.$$

Fixed points. A *fixed point* (or *equilibrium*, *stationary point*) is a point (x_0, y_0) such that

$$f(x_0, y_0) = 0, \qquad g(x_0, y_0) = 0.$$

Hence fixed points are the intersections of the two nullclines.

Qualitative phase portrait. To sketch the phase portrait:

1. Find the nullclines $f = 0$ and $g = 0$.
2. Find their intersections (fixed points).
3. In each region cut out by the nullclines, determine the signs of $x' = f(x, y)$ and $y' = g(x, y)$.
4. Draw the local direction field: $(+, +)$, $(+, -)$, $(-, +)$, $(-, -)$.
5. Analyze each fixed point by linearization (stability, phase portrait type, eigendirections).

Linearization at a fixed point. At a fixed point (x_0, y_0), define the Jacobian matrix

$$J(x_0, y_0) = \begin{pmatrix} \partial_x f(x_0, y_0) & \partial_y f(x_0, y_0) \\ \partial_x g(x_0, y_0) & \partial_y g(x_0, y_0) \end{pmatrix}.$$

Near (x_0, y_0), denoting $A = J(x_0, y_0)$, the system is approximated by the linear system

$$Y' = AY, \qquad \text{where} \qquad Y = \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}.$$

Stability. A fixed point (x_0, y_0) is:

- *stable* if solutions starting sufficiently close stay close for all $t \geq 0$;
- *(asymptotically) stable* if it is stable and, moreover,

$$(x(t), y(t)) \to (x_0, y_0) \qquad (t \to +\infty);$$

- *unstable* if it is not stable.

Stability test from the linearization. Let

$$A = J(x_0, y_0).$$

Generic cases:

- $\det A < 0$: saddle $\Rightarrow$ unstable.
- $\det A > 0,\ \operatorname{tr} A < 0$: (asymptotically) stable
  - $(\operatorname{tr} A)^2 - 4 \det A > 0$: attracting node;
  - $(\operatorname{tr} A)^2 - 4 \det A < 0$: attracting spiral.
- $\det A > 0,\ \operatorname{tr} A > 0$: unstable
  - $(\operatorname{tr} A)^2 - 4 \det A > 0$: repelling node;
  - $(\operatorname{tr} A)^2 - 4 \det A < 0$: repelling spiral;

Equivalent criterion (that extends to higher dimensions):

- all eigenvalues of A have negative real part $\Rightarrow$ (asymptotically) stable;
- at least one eigenvalue has positive real part $\Rightarrow$ unstable.

Inconclusive cases. If $A = J(x_0, y_0)$ has an eigenvalue with zero real part (e.g. double zero eigenvalue or purely imaginary eigenvalues), then linearization is *inconclusive*: higher-order terms must be studied.

Typical procedure at a fixed point.

$$\text{fixed point} \longrightarrow \text{Jacobian } J \longrightarrow \text{eigenvalues / trace-determinant} \longrightarrow \text{local phase portrait and stability.}$$

Note: there are many different, more or less subtle notions of stability. The nomenclature is very field-dependent. In this class, we focus on the nondegenerate cases and simply use “stable” and “unstable.”

7 Linear algebra

7.1 Vectors, matrices, and linear maps

Vectors and matrices. A vector in $\mathbb{R}^n$ is

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

An $n \times n$ matrix is

$$A = (a_{ij})_{1 \le i,j \le n}.$$

Matrix-vector multiplication is

$$(AX)_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Identity matrix. The identity matrix I_n with ones on the diagonal and zeros elsewhere satisfies

$$I_n X = X, \qquad A I_n = I_n A = A.$$

Linear transformation. An $n \times n$ matrix A defines a linear map

$$T_A : \mathbb{R}^n \to \mathbb{R}^n, \qquad T_A(X) = AX.$$

Composition. If A, B are $n \times n$ matrices, then

$$A(BX) = (AB)X.$$

Thus matrix multiplication corresponds to composition of linear maps.

7.2 Trace, determinant, characteristic polynomial

Trace and determinant. For $A \in \mathbb{R}^{n\times n}$, the trace is the sum of the diagonal elements

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}, \qquad \det A \in \mathbb{R}.$$

Characteristic polynomial. The characteristic polynomial of A is

$$p_A(\lambda) = \det(\lambda I_n - A).$$

It is a polynomial of degree n. (Note: the definition of the determinant in higher dimensions is quite complicated, see 18.06).

Eigenvalues. The eigenvalues of A are the roots of p_A.

7.3 Eigenvectors and eigenspaces

Eigenvector. A nonzero vector $v \in \mathbb{R}^n$ (or $\mathbb{C}^n$) is an eigenvector of A with eigenvalue λ if

$$Av = \lambda v.$$

Equivalently,

$$(A - \lambda I_n)v = 0.$$

Eigenspace. The eigenspace associated to λ is

$$E_\lambda = \ker(A - \lambda I_n).$$

Algebraic and geometric multiplicities. For an eigenvalue λ:

- its *algebraic multiplicity* is its multiplicity as a root of p_A;
- its *geometric multiplicity* is $\dim E_\lambda$.

We always have:

$$1 \leq \dim E_\lambda \leq \text{ algebraic multiplicity of } \lambda.$$

Complete / deficient matrices. A matrix A is *complete* or *diagonalizable* if it has a basis of eigenvectors. Otherwise it is *deficient*. (Here, in 18.03, we focus on the generic and simpler complete case).

7.4 Diagonalization

Diagonalizable matrices. If A has a basis of eigenvectors $v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$, define

$$S = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}, \qquad D = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Then

$$A = SDS^{-1}.$$

Interpretation. In the coordinates given by the eigenvector basis, the action of A is just

$$Y \mapsto DY,$$

i.e. each coordinate evolves independently.

In the 2×2 **case.** If $v_1 = \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix}$, $v_2 = \begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix}$ are linearly independent eigenvectors, then with $S = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$, one has

$$A = S \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} S^{-1}.$$

8 $n \times n$ systems of linear ODEs with constant coefficients

8.1 Homogeneous systems

General form. A homogeneous linear system with constant coefficients is

$$X'(t) = AX(t),$$

where A is a constant $n \times n$ matrix and

$$X(t) \in \mathbb{R}^n.$$

Dimension of the solution space. The set of solutions is an n-dimensional vector space.

Exponential solutions from eigenvectors. If $Av = \lambda v$, then $X(t) = e^{\lambda t} v$ is a solution of $X' = AX$.

General solution when A is complete. If $v_1, \dots, v_n$ is a basis of eigenvectors with eigenvalues $\lambda_1, \dots, \lambda_n$, then

$$X(t) = c_1 e^{\lambda_1 t} v_1 + \cdots + c_n e^{\lambda_n t} v_n.$$

Real solutions from complex eigenvalues. If A is real and $w \in \mathbb{C}^n$ is an eigenvector with eigenvalue $\lambda = \alpha + i\beta$, then $Z(t) = e^{(\alpha+i\beta)t} w$ is a complex solution, and real solutions are given by

$$\Re Z(t) \qquad \text{and} \qquad \Im Z(t).$$

In the 2×2 case. This recovers exactly the familiar formulas

$$X(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2$$

or, for complex eigenvalues,

$$X(t) = c_1\, \Re\big(e^{(\alpha+i\beta)t} w\big) + c_2\, \Im\big(e^{(\alpha+i\beta)t} w\big).$$

8.2 Fundamental matrix

Definition. A *fundamental matrix* for $X' = AX$ is a matrix-valued solution

$$\Phi(t) \in \mathbb{R}^{n \times n}$$

whose columns form a basis of solutions.

General solution from a fundamental matrix. If $\Phi(t)$ is a fundamental matrix, then every solution is of the form

$$X(t) = \Phi(t)c, \qquad c \in \mathbb{R}^n \text{ constant vector.}$$

Construction from a basis of solutions. If $X_1, \dots, X_n$ is a basis of solutions, then

$$\Phi(t) = \begin{pmatrix} | & & | \\ X_1(t) & \cdots & X_n(t) \\ | & & | \end{pmatrix}$$

is a fundamental matrix.

Complete case. If A has eigenbasis $v_1, \dots, v_n$, then a fundamental matrix is

$$\Phi(t) = \begin{pmatrix} | & & | \\ e^{\lambda_1 t} v_1 & \cdots & e^{\lambda_n t} v_n \\ | & & | \end{pmatrix}.$$

8.3 Matrix exponential

Definition. For any square matrix A,

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = I_n + At + \frac{(At)^2}{2!} + \cdots .$$

Key properties.

$$\frac{d}{dt} e^{At} = A e^{At} = e^{At} A, \qquad e^{A \cdot 0} = I_n.$$

Therefore e^{At} is the unique fundamental matrix satisfying

$$\Phi(0) = I_n.$$

General solution. The solution of the initial value problem

$$\begin{cases} X' = AX, \\ X(0) = X_0 \end{cases}$$

is

$$X(t) = e^{At} X_0.$$

Diagonalizable case. If $A = SDS^{-1}$, then

$$e^{At} = S e^{Dt} S^{-1}, \qquad e^{Dt} = \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}).$$

Meaning. Knowing the eigenvectors of A reduces the computation of e^{At} to exponentiating diagonal entries.

In the 2×2 **case.** If

$$A = S \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} S^{-1},$$

then

$$e^{At} = S \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} S^{-1}.$$

8.4 Deficient case: one basic example

Jordan block of size 2**.** If

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \lambda I + N, \qquad N^2 = 0,$$

then

$$e^{At} = e^{\lambda t}(I + tN) = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Corresponding solutions. A basis of solutions is

$$e^{\lambda t}\begin{pmatrix}1\\0\end{pmatrix}, \qquad e^{\lambda t}\begin{pmatrix}t\\1\end{pmatrix}.$$

(Repeated eigenvalues may produce factors of t, just as for scalar constant-coefficient ODEs.)

8.5 Inhomogeneous systems

General form. For $F(t) \in \mathbb{R}^n$ given:

$$X'(t) = AX(t) + F(t).$$

Variation of parameters. If $\Phi(t)$ is a fundamental matrix for $X' = AX$, write

$$X(t) = \Phi(t)u(t).$$

Then

$$u'(t) = \Phi(t)^{-1}F(t),$$

so

$$X(t) = \Phi(t)\left(c + \int^t \Phi(s)^{-1}F(s)\,ds\right).$$

Using e^{At}. Since $\Phi(t) = e^{At}$ is a fundamental matrix, the general solution can be written

$$X(t) = e^{At}c + e^{At}\int^t e^{-As}F(s)\,ds.$$

General / particular solution rule.

$$X(t) = X_h(t) + X_p(t),$$

where X_h solves $X' = AX$ and X_p is any one particular solution.

8.6 Decoupling

Change of variables. If $A = SDS^{-1}$ and we set

$$X = SY,$$

then

$$X' = AX + F(t)$$

becomes

$$Y' = DY + S^{-1}F(t).$$

Decoupled system. Since D is diagonal,

$$y_i' = \lambda_i y_i + \widetilde{F}_i(t), \qquad i = 1, \dots, n,$$

so the system splits into n independent first-order linear ODEs.

8.7 Exponential forcing

Exponential Response Formula for systems. Consider

$$X' = AX + e^{rt}B,$$

where $r \in \mathbb{C}$ and $B \in \mathbb{C}^n$ is constant.

If $rI - A$ is invertible, then a particular solution is

$$X_p(t) = e^{rt}(rI - A)^{-1}B.$$

Complex replacement. For sinusoidal forcing, replace $\cos(\omega t)$, $\sin(\omega t)$ by $e^{i\omega t}$, solve the complex system, then take real or imaginary parts.

9 Fourier series

9.1 Coordinates and orthogonal bases

Orthogonal and orthonormal bases. Let V be a vector space with inner product $\langle\cdot,\cdot\rangle$. A basis (e_n) is *orthogonal* if

$$\langle e_m, e_n\rangle = 0 \qquad (m \neq n).$$

It is *orthonormal* if moreover $\langle e_n, e_n\rangle = 1$.

Coordinates in an orthogonal basis. If

$$f = \sum_n c_n e_n$$

in an orthogonal basis (e_n), then

$$c_n = \frac{\langle f, e_n\rangle}{\langle e_n, e_n\rangle}.$$

If the basis is orthonormal, this becomes

$$c_n = \langle f, e_n\rangle.$$

Inner product for functions. For real-valued 2π-periodic functions f, g, define

$$\langle f, g\rangle := \int_{-\pi}^{\pi} f(t)g(t)\,dt.$$

9.2 Fourier series for periodic functions

Fourier series. For a 2π-periodic function f, its Fourier series is

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt).$$

Orthogonality relations.

$$\begin{gathered}
\langle 1, \cos(nt)\rangle = 0, \qquad \langle 1, \sin(nt)\rangle = 0 \qquad (n \geqslant 1),\\
\langle \cos(mt), \sin(nt)\rangle = 0 \qquad (m, n \geqslant 1),\\
\langle \cos(mt), \cos(nt)\rangle = \begin{cases} 0, & m \neq n,\\ \pi, & m = n \geqslant 1,\end{cases}\\
\langle \sin(mt), \sin(nt)\rangle = \begin{cases} 0, & m \neq n,\\ \pi, & m = n \geqslant 1,\end{cases}
\end{gathered}$$

and

$$\langle 1, 1\rangle = 2\pi.$$

Fourier coefficients. Using the orthogonal-basis formula,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)\,dt \qquad (n \geqslant 0),$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)\,dt \qquad (n \geqslant 1).$$

Constant term / average value.

$$\frac{a_0}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt.$$

So the constant term is the average value of f on $(-\pi, \pi)$.

Fourier theorem. For a "reasonable" (e.g. smooth) 2π-periodic function f,

$$f(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt).$$

At points where f is continuous, the Fourier series converges to $f(t)$. At a jump discontinuity $t = \tau$, it converges to

$$\frac{f(\tau^-) + f(\tau^+)}{2}.$$

9.3 Sine and cosine Fourier series for functions on an interval

Even and odd functions.

$$f \text{ even} \iff f(-t) = f(t), \qquad f \text{ odd} \iff f(-t) = -f(t).$$

Even/odd symmetry and Fourier series. If

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt),$$

then:

- if f is even, then $b_n = 0$ for all n, so

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt);$$

- if f is odd, then $a_n = 0$ for all n (including $a_0 = 0$), so

$$f(t) = \sum_{n=1}^{\infty} b_n \sin(nt).$$

Cosine series on $(0, \pi)$**.** Given a function f on $(0, \pi)$, extend it evenly to $(-\pi, \pi)$ and then periodically with period 2π. Its cosine series is

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) \qquad (0 < t < \pi),$$

with

$$a_n = \frac{2}{\pi}\int_0^{\pi} f(t)\cos(nt)\,dt \qquad (n \geqslant 0).$$

Sine series on $(0, \pi)$**.** Given a function f on $(0, \pi)$, extend it oddly to $(-\pi, \pi)$ and then periodically with period 2π. Its sine series is

$$f(t) = \sum_{n=1}^{\infty} b_n \sin(nt) \qquad (0 < t < \pi),$$

with

$$b_n = \frac{2}{\pi}\int_0^{\pi} f(t)\sin(nt)\,dt \qquad (n \geqslant 1).$$

9.4 Solving ODEs with Fourier series

Philosophy. If the input signal is periodic, write it as a Fourier series and use linearity/superposition: solve the ODE separately for each Fourier mode, then add the responses.

Mode-by-mode solution. Consider a linear ODE with constant coefficients

$$p(D)x = f(t),$$

where f is 2π-periodic with Fourier series

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt).$$

Then a periodic particular solution can be found by solving separately for each frequency n.

Complex replacement for one mode. For the input e^{int},

$$p(D)z = e^{int}.$$

If $p(in) \neq 0$, then by ERF,

$$z_n(t) = \frac{1}{p(in)} e^{int}.$$

Hence:

$$p(D)x = \cos(nt) \quad \Longrightarrow \quad x_n(t) = \Re\left(\frac{1}{p(in)} e^{int}\right),$$

$$p(D)x = \sin(nt) \quad \Longrightarrow \quad x_n(t) = \Im\left(\frac{1}{p(in)} e^{int}\right).$$

Complex gain for the n-th mode.

$$G(n) := \frac{1}{p(in)}.$$

For the n-th Fourier mode, the output has the same frequency n, multiplied by the complex gain $G(n)$.

Superposition. If $x_n^{(c)}$ solves $p(D)x = \cos(nt)$ and $x_n^{(s)}$ solves $p(D)x = \sin(nt)$, then

$$x_p(t) = \frac{a_0}{2} \frac{1}{p(0)} + \sum_{n=1}^{\infty} a_n x_n^{(c)}(t) + \sum_{n=1}^{\infty} b_n x_n^{(s)}(t)$$

is a periodic particular solution (assuming the denominators are nonzero).

Resonance. If $p(in) = 0$ for some n, then the n-th mode is resonant:

- if there is no damping, exact resonance produces a factor t;
- if $p(in)$ is small but nonzero, the corresponding gain $\frac{1}{|p(in)|}$ is large.

9.5 Listening to Fourier series (optional)

Fundamental and overtones. A periodic sound wave can be written as a Fourier series. If the first nonconstant mode has frequency ν, then:

$$\text{fundamental frequency } = \nu, \qquad \text{overtones } = 2\nu, 3\nu, \ldots$$

The n-th Fourier mode corresponds to the frequency $n\nu$.

Pitch. Higher frequency means higher pitch.

Phase. The phases of the different sinusoidal components are not heard directly.

9.6 More on Fourier series

Fourier series of period $2L$. A $2L$-periodic function has Fourier series

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi t}{L}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi t}{L}\right),$$

with coefficients

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt \qquad (n \geq 0),$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt \qquad (n \geq 1).$$

Convergence. For a "reasonable" periodic function f:

- at points where f is continuous, the Fourier series converges to $f(t)$;
- at a jump discontinuity $t = \tau$, it converges to

$$\frac{f(\tau^-) + f(\tau^+)}{2}.$$

Periodic antiderivative. If f is periodic and has average value 0, i.e. $\frac{a_0}{2} = 0$, then f has periodic antiderivatives. If

$$f(t) = \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt),$$

then for a constant C, a periodic antiderivative is

$$F(t) = C + \sum_{n=1}^{\infty} \frac{a_n}{n}\sin(nt) - \sum_{n=1}^{\infty} \frac{b_n}{n}\cos(nt).$$

Term-by-term integration.

$$\int \cos(nt)\,dt = \frac{1}{n}\sin(nt), \qquad \int \sin(nt)\,dt = -\frac{1}{n}\cos(nt).$$

Warning. A periodic antiderivative exists only if the average value of f is 0 or equivalently $a_0 = 0$. Also, Fourier series can be integrated term by term much more safely than they can be differentiated term by term.

10 Introduction to PDEs

10.1 Boundary value problems

Initial value problem vs. boundary value problem. For an ODE on an interval:

- *initial conditions*: conditions at **one** point, e.g.
$$v(0) = a, \qquad v'(0) = b;$$
- *boundary conditions*: conditions at **different** points, e.g.
$$v(0) = a, \qquad v(\pi) = b.$$

Warning. For boundary value problems, there is in general *no* existence/uniqueness theorem like for IVPs: there may be no solution, one solution, or infinitely many solutions.

Dirichlet eigenvalue problem on $(0, \pi)$. Consider

$$\begin{cases} v''(x) = \lambda v(x), \\ v(0) = 0, \\ v(\pi) = 0. \end{cases}$$

Nonzero solutions exist only for

$$\lambda = -n^2, \qquad n = 1, 2, 3, \ldots$$

and we say that $v_n(x) = \sin(nx)$ is an associated eigenfunction.

Neumann eigenvalue problem on $(0, \pi)$. Consider

$$\begin{cases} v''(x) = \lambda v(x), \\ v'(0) = 0, \\ v'(\pi) = 0. \end{cases}$$

Nonzero solutions exist for

$$\lambda = 0, -1, -4, -9, \ldots$$

with eigenfunctions $1,\ \cos x,\ \cos 2x,\ \cos 3x, \ldots$.

Philosophy. Boundary value problems for

$$\frac{d^2}{dx^2}$$

play the role of the eigenvalue problem for a matrix:

$$Av = \lambda v \qquad \longleftrightarrow \qquad v'' = \lambda v \text{ (with boundary conditions).}$$

Homogeneous vs. inhomogeneous boundary value problems. For a linear differential operator $p(D)$, one often studies

$$p(D)y = 0 \qquad \text{and} \qquad p(D)y = f$$

together with boundary conditions. The general philosophy is the same as for linear algebra:

- first understand the homogeneous problem $p(D)y = 0$;
- then solve the inhomogeneous problem by writing

  $$y = y_p + y_h,$$

  where y_p is one particular solution of $p(D)y = f$ and y_h is the general solution of $p(D)y = 0$;
- finally impose the boundary conditions.

Typical second-order example. A basic model is

$$p_\mu(D)y = y'' + \mu^2 y.$$

The homogeneous equation

$$y'' + \mu^2 y = 0$$

describes the free modes of the system. The boundary conditions quantize the values of μ that are actually allowed.

Solving inhomogeneous boundary value problems. Given $p_\mu(D)y = f$ with boundary conditions:

1. solve the homogeneous equation $Ly = 0$;
2. find one particular solution y_p of $Ly = f$;
3. write $y = y_p + y_h$;
4. impose the boundary conditions to determine the constants in y_h if possible.

If the homogeneous boundary value problem has only the zero solution, then there is at most one solution of the inhomogeneous problem. If the homogeneous problem has nonzero solutions, then uniqueness fails and there may also be an obstruction to existence.

Integration by parts. A key identity behind many solvability conditions is the integration by parts formula

$$\int_a^b \big(y''(t)z(t) - y(t)z''(t)\big)\,dt = \big[y'(t)z(t) - y(t)z'(t)\big]_a^b.$$

More generally, for a self-adjoint operator of the form

$$Ly = -(p(t)y'(t))' + q(t)y(t),$$

one has

$$\int_a^b \big(Ly\big)z\,dt - \int_a^b y\,(Lz)\,dt = \text{boundary term}.$$

Under common boundary conditions (Dirichlet, Neumann, mixed), this boundary term often vanishes.

Fredholm alternative / compatibility condition. The Fredholm alternative says, philosophically, that there are two possibilities:

- either the homogeneous boundary value problem has only the trivial zero solution, and then one expects a unique solution for every forcing f;
- or the homogeneous problem has nontrivial solutions, and then $Ly = f$ is solvable only for forcings f satisfying compatibility conditions.

For self-adjoint problems, these compatibility conditions usually take the form

$$\boxed{\int_a^b f(t)\, v(t)\, dt = 0}$$

for every homogeneous solution v satisfying the same boundary conditions. So nontrivial homogeneous solutions are the source of both:

- failure of uniqueness;
- possible failure of existence.

Eigenfunctions and expansions. When the homogeneous problem produces a family of eigenfunctions e_n, one often expands

$$f = \sum_n c_n e_n, \qquad y = \sum_n a_n e_n.$$

The coefficients are obtained from the inner product formula

$$c_n = \frac{\langle f, e_n \rangle}{\langle e_n, e_n \rangle},$$

whenever the eigenfunctions are orthogonal. For example, on $(0, \pi)$, Dirichlet conditions naturally lead to $\sin(nt)$, while Neumann conditions naturally lead to $1, \cos(nt)$, but the main principle is more general:

10.2 Heat equation

Heat equation.

$$\partial_t u = \partial_x \partial_x u, \qquad t > 0.$$

Dirichlet boundary conditions. For a rod with ends held at 0 and π, we impose

$$u(0, t) = 0, \qquad u(\pi, t) = 0.$$

Separation of variables. Look for a solution of the form $u(x, t) = X(x)T(t)$. Then the equation becomes $X(x)T'(t) = X''(x)T(t)$ so

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = \lambda.$$

Thus

$$X''(x) = \lambda X(x), \qquad T'(t) = \lambda T(t).$$

Normal modes for Dirichlet conditions. Using the boundary-value problem above,

$$X_n(x) = \sin(nx), \qquad \lambda_n = -n^2,$$

and therefore $T_n(t) = e^{-n^2 t}$. So the normal modes are

$$u_n(x,t) = e^{-n^2 t}\sin(nx).$$

General solution with Dirichlet conditions.

$$u(x,t) = \sum_{n=1}^{\infty} b_n e^{-n^2 t}\sin(nx).$$

The coefficients b_n are determined by the sine series of f, the initial condition on $(0,\pi)$:

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} b_n \sin(nx).$$

Long-time behavior. Higher modes decay faster:

$$e^{-n^2 t} \to 0 \qquad (t \to +\infty).$$

So the solution becomes smoother and tends to the lowest surviving mode.

Energy identity. Define

$$E(t) := \frac{1}{2}\int_0^{\pi} u(x,t)^2\,dx.$$

Then

$$E'(t) = \int_0^{\pi} u\,\partial_t u\,dx = \int_0^{\pi} u\,\partial_x\partial_x u\,dx.$$

Integrating by parts gives

$$\int_0^{\pi} u\,\partial_x\partial_x u\,dx = \big[u\,\partial_x u\big]_0^{\pi} - \int_0^{\pi} (\partial_x u)^2\,dx.$$

Since $u(0,t) = u(\pi,t) = 0$, we have $\big[u\,\partial_x u\big]_0^{\pi} = 0$, hence

$$E'(t) = -\int_0^{\pi} (\partial_x u(x,t))^2\,dx \le 0.$$

So heat energy **decreases** in time.

Entropy along the heat flow. Assume $u(x,t) > 0$ solves $\partial_t u = \partial_x\partial_x u$ and that the boundary terms vanish (for instance for Neumann boundary conditions, periodic boundary conditions, or sufficient decay at infinity).

Define the entropy

$$H(t) := -\int u(x,t)\log u(x,t)\,dx.$$

Then

$$H'(t) = -\int (1+\log u)\,\partial_t u\,dx = -\int (1+\log u)\,\partial_x\partial_x u\,dx.$$

Integrating by parts, we get

$$H'(t) = \int \frac{(\partial_x u)^2}{u}\,dx \ge 0.$$

So the entropy increases along the heat flow.

10.3 Duhamel principle for the heat equation

From matrices to differential operators. For a finite-dimensional inhomogeneous linear system $X'(t) = AX(t) + F(t)$, variation of parameters gives

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}F(s)\,ds.$$

The same idea applies to linear PDEs. The matrix A is replaced by a differential operator. For the heat equation on $(0, \pi)$ with homogeneous Dirichlet boundary conditions,

$$\begin{cases} \partial_t u = \partial_x\partial_x u + F(x,t), \\ u(0,t) = 0, \\ u(\pi,t) = 0, \\ u(x,0) = u_0(x), \end{cases}$$

the operator playing the role of A is the *Dirichlet Laplacian* (i.e. the Laplacian seen as acting on functions satisfying Dirichlet conditions) $\Delta_D = \partial_x\partial_x$, where the subscript D reminds us that we impose Dirichlet boundary conditions. Thus the analogue of e^{At} is the **heat operator**

$$e^{t\Delta_D}.$$

Duhamel formula. The solution is formally

$$\boxed{u(t) = e^{t\Delta_D}u_0 + \int_0^t e^{(t-s)\Delta_D}F(s)\,ds.}$$

Here $u(t)$ denotes the function $x \mapsto u(x,t)$, and $F(s)$ denotes the function $x \mapsto F(x,s)$. This is the PDE version of variation of parameters:

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}F(s)\,ds.$$

Interpretation. The first term $e^{t\Delta_D}u_0$ is the heat flow coming from the initial temperature u_0. The second term

$$\int_0^t e^{(t-s)\Delta_D}F(s)\,ds$$

adds up the effect of the forcing. A small amount of heat inserted at time s then evolves freely for time $t-s$. This is called *Duhamel's principle*: an inhomogeneous problem is solved by superposing the homogeneous evolution of all infinitesimal inputs.

Fourier sine series diagonalize the Dirichlet Laplacian. The Dirichlet eigenfunctions on $(0, \pi)$ are $\sin(nx)$, $n = 1, 2, 3, \ldots$, and

$$\Delta_D \sin(nx) = \partial_x\partial_x \sin(nx) = -n^2 \sin(nx).$$

Thus the Laplacian is diagonal in the sine basis: on the n-th mode, Δ_D acts like multiplication by $-n^2$. Therefore the heat operator acts by

$$e^{t\Delta_D}\sin(nx) = e^{-n^2 t}\sin(nx).$$

So Fourier sine series “diagonalize” the heat operator.

Homogeneous heat flow. If $u_0(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$, then each Fourier mode evolves independently:

$$e^{t\Delta_D} u_0 = \sum_{n=1}^{\infty} b_n e^{-n^2 t} \sin(nx).$$

Inhomogeneous heat equation mode-by-mode. Now suppose

$$F(x,t) = \sum_{n=1}^{\infty} f_n(t) \sin(nx), \qquad u(x,t) = \sum_{n=1}^{\infty} a_n(t) \sin(nx).$$

Plugging into $\partial_t u = \Delta_D u + F$ gives

$$\sum_{n=1}^{\infty} a_n'(t) \sin(nx) = \sum_{n=1}^{\infty} \left(-n^2 a_n(t) + f_n(t)\right) \sin(nx).$$

Therefore each coefficient satisfies the scalar first-order ODE $a_n'(t) = -n^2 a_n(t) + f_n(t)$. Solving this ODE gives

$$a_n(t) = e^{-n^2 t} a_n(0) + \int_0^t e^{-n^2(t-s)} f_n(s)\, ds.$$

Hence

$$\boxed{u(x,t) = \sum_{n=1}^{\infty} \left[e^{-n^2 t} a_n(0) + \int_0^t e^{-n^2(t-s)} f_n(s)\, ds \right] \sin(nx).}$$

Comparison with matrices. The dictionary is

$$X(t) \in \mathbb{R}^n \quad \longleftrightarrow \quad u(t) = u(\cdot, t), \qquad\qquad A \quad \longleftrightarrow \quad \Delta_D,$$

$$e^{At} \quad \longleftrightarrow \quad e^{t\Delta_D}, \qquad\qquad F(t) \quad \longleftrightarrow \quad F(\cdot, t).$$

Thus Duhamel's formula for heat flow is exactly the operator version of variation of parameters.

Heat kernel version. Using the Dirichlet heat kernel

$$K_D(t,x,y) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2 t} \sin(nx) \sin(ny),$$

the heat operator may also be written as

$$\left(e^{t\Delta_D} g\right)(x) = \int_0^{\pi} K_D(t,x,y) g(y)\, dy.$$

Therefore Duhamel's formula becomes

$$\boxed{u(x,t) = \int_0^{\pi} K_D(t,x,y) u_0(y)\, dy + \int_0^t \int_0^{\pi} K_D(t-s,x,y) F(y,s)\, dy\, ds.}$$

The first integral evolves the initial condition. The double integral adds the effect of the forcing F, with heat inserted at position y and time s, then diffused until time t.

10.4 Laplace equation on a disk

Laplace equation. Let $D \subset \mathbb{R}^2$ be the unit disk: $D = \{(r,\theta) : 0 \le r < 1,\ 0 \le \theta < 2\pi\}$. The Laplace equation is

$$\Delta u = 0.$$

In polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$,

$$\Delta u = \partial_x\partial_x u + \partial_y\partial_y u = \partial_r\partial_r u + \frac{1}{r}\partial_r u + \frac{1}{r^2}\partial_\theta\partial_\theta u.$$

Dirichlet problem on the disk. Given boundary data $u(1,\theta) = f(\theta)$, we want to solve

$$\begin{cases} \Delta u = 0, & 0 \le r < 1, \\ u(1,\theta) = f(\theta). \end{cases}$$

Fourier expansion of the boundary data. Write

$$f(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta).$$

Then the harmonic function inside the disk is

$$\boxed{u(r,\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n a_n \cos(n\theta) + \sum_{n=1}^{\infty} r^n b_n \sin(n\theta).}$$

Why the factors r^n appear. For each angular mode $\cos(n\theta)$ or $\sin(n\theta)$, separation of variables gives radial powers

$$r^n \quad \text{and} \quad r^{-n}.$$

The term r^{-n} blows up at $r = 0$, so it is excluded if we want a bounded solution in the disk. The regular harmonic modes $r^n\cos(n\theta)$ and $r^n\sin(n\theta)$ are in fact harmonic polynomials.

10.5 Dirichlet-to-Neumann map and tomography

Dirichlet-to-Neumann map. For a domain Ω, suppose u solves

$$\begin{cases} \Delta u = 0, & \text{in } \Omega, \\ u = f, & \text{on } \partial\Omega. \end{cases}$$

The *Dirichlet-to-Neumann map* sends the boundary value f to the normal derivative of the solution at the boundary:

$$\Lambda f = \partial_\nu u\Big|_{\partial\Omega}.$$

Here ν is the outward unit normal vector.

Interpretation. The input f is a prescribed voltage on the boundary. The output $\partial_\nu u$ is the corresponding current flux through the boundary. Thus Λ records how the inside of the domain responds to boundary measurements.

Example: unit disk. For the unit disk, if $f(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta)$, then

$$u(r,\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n a_n \cos(n\theta) + \sum_{n=1}^{\infty} r^n b_n \sin(n\theta).$$

Since the outward normal derivative on the unit circle is ∂_r, we get

$$\partial_\nu u\Big|_{r=1} = \partial_r u(1,\theta) = \sum_{n=1}^{\infty} n a_n \cos(n\theta) + \sum_{n=1}^{\infty} n b_n \sin(n\theta).$$

Therefore

$$\Lambda\left(\cos(n\theta)\right) = n\cos(n\theta), \qquad \Lambda\left(\sin(n\theta)\right) = n\sin(n\theta).$$

Tomography philosophy. In *inverse problems*, one asks whether boundary measurements determine what is inside a domain. The Dirichlet-to-Neumann map is a mathematical idealization of such measurements:

$$\text{impose boundary voltage} \quad \longrightarrow \quad \text{measure boundary current.}$$

The guiding question is:

Can we recover information about the interior from Λ?

This is the basic idea behind the widely used electrical impedance tomography.

11 Optimization schemes and numerical methods

11.1 Numerical methods and time discretization

Why numerical methods? Many differential equations cannot be solved explicitly. Even when an exact formula exists, it may be too complicated to use in practice. Numerical methods replace a continuous-time trajectory $y(t)$ by a sequence of approximate values

$$y_0,\ y_1,\ y_2, \dots,$$

where y_k is meant to approximate $y(t_k)$ at discrete times

$$t_k = t_0 + kh.$$

The number $h > 0$ is called the *time step*.

Euler's method. Consider an ODE (it could be in any dimension)

$$\begin{cases} y' = f(t,y), \\ y(t_0) = y_0. \end{cases}$$

For small h, we approximate the derivative by

$$y'(t_k) \approx \frac{y_{k+1} - y_k}{h}.$$

Since the ODE says $y'(t_k) = f(t_k, y_k)$, this gives

$$y_{k+1} = y_k + hf(t_k, y_k).$$

This is *Euler's method.* It is the simplest numerical method for solving an ODE.

Euler's method for autonomous ODEs. If $y' = f(y)$, then Euler's method becomes

$$y_{k+1} = y_k + hf(y_k).$$

Thus a continuous-time flow is replaced by a discrete-time dynamical system.

More advanced methods. Euler's method is only the beginning. More accurate methods, such as Runge–Kutta methods, use several evaluations of f at each step to better approximate the true solution. Other methods, called *implicit methods*, define y_{k+1} using information at the future time t_{k+1}. For example, implicit Euler is

$$y_{k+1} = y_k + hf(t_{k+1}, y_{k+1}).$$

This is usually harder to compute because y_{k+1} appears on both sides, but implicit methods often have better stability properties.

11.2 Gradient descent as discretized gradient flow

Gradient flow. The gradient flow of a function $F : \mathbb{R}^d \to \mathbb{R}$ is

$$y'(t) = -\nabla F(y(t)).$$

Along a solution,

$$\frac{d}{dt}F(y(t)) = \nabla F(y(t)) \cdot y'(t) = -|\nabla F(y(t))|^2 \leq 0.$$

Thus F decreases along the flow. This is why gradient flows model optimization: the trajectory moves downhill.

Numerical methods and optimization. Optimization algorithms can be viewed as numerical methods applied to gradient flow equations. The gradient flow

$$y' = -\nabla F(y)$$

moves downhill along the function F. Applying Euler's method gives

$$y_{k+1} = y_k - h\nabla F(y_k),$$

which is exactly the **gradient descent** algorithm. Thus gradient descent is not an isolated trick: it is the most basic numerical method applied to the most basic optimization flow.

Why the step size matters. Consider the simplest quadratic function

$$F(x) = \frac{\lambda}{2}x^2, \qquad \lambda > 0.$$

Then the gradient flow is $x' = -\lambda x$, so

$$x(t) = e^{-\lambda t}x(0).$$

The continuous-time system is always stable and converges exponentially fast to the minimum at $x = 0$.

Gradient descent gives $x_{k+1} = x_k - h\lambda x_k = (1 - h\lambda)x_k$. Therefore

$$x_k = (1 - h\lambda)^k x_0.$$

This converges to 0 **only** when $|1 - h\lambda| < 1$.

Moral. The continuous gradient flow is stable for every $\lambda > 0$, but the discrete algorithm is stable only if the step size is small enough. This is one reason learning rates are delicate in optimization.

Implicit Euler and proximal methods. If instead we apply implicit Euler to the gradient flow

$$y' = -\nabla F(y),$$

we get

$$\frac{y_{k+1} - y_k}{h} = -\nabla F(y_{k+1}).$$

This equation says that y_{k+1} is a critical point of $y \mapsto F(y) + \frac{1}{2h}|y - y_k|^2$. Thus we can search for it through

$$y_{k+1} = \operatorname*{argmin}_y \left(F(y) + \frac{1}{2h}|y - y_k|^2 \right).$$

This is the idea behind *proximal methods*: instead of moving directly downhill, we minimize F while penalizing moves that are too far from the current point.

11.3 Hessian eigenvalues and learning rates

Near a critical point. Suppose y_0 is a critical point of F, so

$$\nabla F(y_0) = 0.$$

Near y_0, write $z = y - y_0$. The linearization of the gradient flow is

$$z' = -\operatorname{Hess} F(y_0)\, z.$$

Since the Hessian is symmetric, it has real eigenvalues and an orthonormal eigenbasis.

Diagonalizing the gradient flow. If the eigenvalues of Hess $F(y_0)$ are

$$\lambda_1, \dots, \lambda_d,$$

then in eigenvector coordinates the linearized flow is

$$z_i' = -\lambda_i z_i.$$

Hence

$$z_i(t) = e^{-\lambda_i t} z_i(0).$$

Interpretation. If $\lambda_i > 0$, the i-th direction is stable. If $\lambda_i < 0$, the i-th direction is unstable. Thus:

- local minima have positive Hessian eigenvalues and are stable;
- local maxima have negative Hessian eigenvalues and are unstable;
- saddle points have both positive and negative Hessian eigenvalues and are unstable.

Note: if some of the eigenvalues vanish, then nothing can generally be said. Some rates can still be obtained through so-called Łojasiewicz inequalities under convexity or analyticity assumptions.

Gradient descent near a minimum. Suppose y_0 is a local minimum of F, and assume for simplicity that $\nabla F(y_0) = 0$ and Hess $F(y_0)$ has positive eigenvalues. Near y_0, write $z_k = y_k - y_0$. The gradient descent algorithm is

$$y_{k+1} = y_k - h\nabla F(y_k).$$

Using the linear approximation

$$\nabla F(y_k) \approx \text{Hess}\, F(y_0)(y_k - y_0) = \text{Hess}\, F(y_0)z_k,$$

we get the linearized gradient descent iteration

$$z_{k+1} = z_k - h\,\text{Hess}\, F(y_0)z_k = \big(I - h\,\text{Hess}\, F(y_0)\big)z_k.$$

Since Hess $F(y_0)$ is symmetric, we can diagonalize it in an orthonormal basis of eigenvectors. If

$$\text{Hess}\, F(y_0)v_i = \lambda_i v_i,$$

then the component of z_k in the v_i-direction evolves independently:

$$z_{i,k+1} = (1 - h\lambda_i)z_{i,k}.$$

For convergence in that direction, we need

$$|1 - h\lambda_i| < 1.$$

To be stable in every eigendirection, this must hold for every eigenvalue λ_i. Therefore the most restrictive condition comes from the largest eigenvalue:

$$0 < h < \frac{2}{\lambda_{\max}}.$$

Interpretation. Large Hessian eigenvalues correspond to directions in which F is very curved. In those directions, a step size that is too large overshoots the minimum and can cause oscillation or divergence. Thus the largest eigenvalue of the Hessian controls the largest safe learning rate.

11.4 Heat equation as a gradient flow

Finite-dimensional gradient flow. For a function $F : \mathbb{R}^d \to \mathbb{R}$, the gradient flow is

$$x' = -\nabla F(x).$$

Infinite-dimensional analogue. The heat equation $\partial_t u = \partial_x\partial_x u$ can be interpreted as the gradient flow of the energy

$$E(u) = \frac{1}{2}\int_0^\pi |\partial_x u|^2\, dx.$$

To compute its gradient, perturb u in the direction of a test function v. Define $u_\varepsilon = u + \varepsilon v$. Then

$$E(u + \varepsilon v) = \frac{1}{2}\int_0^\pi |\partial_x u + \varepsilon\partial_x v|^2\, dx.$$

Differentiating at $\varepsilon = 0$, we get

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} E(u + \varepsilon v) = \int_0^\pi \partial_x u\, \partial_x v\, dx.$$

Integrating by parts,

$$\int_0^\pi \partial_x u \, \partial_x v \, dx = -\int_0^\pi (\partial_x \partial_x u) \, v \, dx + \left[\partial_x u \, v\right]_0^\pi.$$

If we impose Dirichlet boundary conditions, then the allowed variations satisfy $v(0) = v(\pi) = 0$, so the boundary term vanishes. Therefore

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} E(u + \varepsilon v) = -\int_0^\pi (\partial_x \partial_x u) \, v \, dx.$$

By definition of the gradient, $\nabla E(u)$ is the function satisfying

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} E(u + \varepsilon v) = \langle \nabla E(u), v \rangle = \int_0^\pi \nabla E(u) \, v \, dx$$

for every variation v. Hence the negative gradient direction is $-\nabla E(u) = \partial_x \partial_x u$. So the gradient flow equation

$$\partial_t u = -\nabla E(u)$$

is exactly the heat equation $\partial_t u = \partial_x \partial_x u$. Thus $\partial_t u = \Delta u$ is a gradient descent equation in the space of functions.

Energy decay. Along the heat flow, just like for gradient flows,

$$\frac{d}{dt}\frac{1}{2}\int_0^\pi |\partial_x u|^2 \, dx = -\int_0^\pi |\partial_t u|^2 \, dx \leq 0.$$

So the heat equation is not just a PDE: it is an optimization process. It tries to decrease the roughness of u.

Note: from the perspective of Wasserstein spaces in Optimal Transport theory, the heat flow can also be seen as a gradient flow of minus the entropy:

$$-H(u) = \int_0^\pi u \log u.$$