Differential Equations

Gudfit

Contents

0

Introduction: Ordinary Differential Equations

Mathematical modelling of physical and engineering problems frequently leads to functional equations relating an independent variable, an unknown function, and its derivatives. When the unknown function depends on a single variable, such equations are termed ordinary differential equations.

0.1 Definitions and Classification

> **Definition 0.1.** ***Ordinary Differential Equation.***
> An **ordinary differential equation** (ODE) is a relation of the form
>
> $$F(x, y, y', \dots, y^{(n)}) = 0, \tag{1}$$
>
> where x is the independent variable, $y = y(x)$ is the unknown function, and $y^{(k)}$ denotes the k-th derivative of y with respect to x. The integer $n \geq 1$, representing the highest derivative appearing in the equation, is called the **order** of the differential equation.
>
> 定義

It is implicitly assumed that F depends non-trivially on $y^{(n)}$. Equations involving composition of the unknown function, such as $y'(x) = y(y(x))$, are functional equations but fall outside the scope of standard differential equations.

We distinguish between linear and non-linear equations based on the dependence of F on the dependent variable and its derivatives.

> **Definition 0.2.** ***Linearity.***
> The differential equation (1) is **linear** if F is a linear function of the variables $y, y', \dots, y^{(n)}$. Explicitly, an n-th order linear ODE can be written as:
>
> $$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = f(x),$$
>
> where the coefficients $a_i(x)$ and the term $f(x)$ depend only on x. If the

equation is not linear, it is termed **nonlinear**.

定義

Example 0.1. Classification of Equations. Consider the following equations:

First-order, linear ($x \neq 0$):

$$\frac{dy}{dx} + \frac{1}{x}y = x^3.$$

First-order, nonlinear (due to y^2):

$$\frac{dy}{dx} = 1 + y^2.$$

Second-order, nonlinear (due to the product yy'):

$$y'' + yy' = x.$$

Second-order, linear (independent variable t, unknown θ):

$$\frac{d^2\theta}{dt^2} + a^2\theta = 0.$$

範例

Remark.

If the unknown function depends on multiple independent variables (e.g., $u(x, y)$), the equation involves partial derivatives and is called a **partial differential equation** (PDE). For instance, the Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is a second-order linear PDE. This text focuses exclusively on ordinary differential equations.

Solutions and Integrals

Solving a differential equation entails finding a function that satisfies the relationship identically over a specific domain.

Definition 0.3. ***Solution.***
A function $y = \varphi(x)$ is a **solution** of the differential equation (1) on an interval J if φ is n-times differentiable on J and satisfies

$$F(x, \varphi(x), \varphi'(x), \ldots, \varphi^{(n)}(x)) = 0$$

for all $x \in J$.

定義

Example 0.2. Verification of Solutions. Consider the equation

$$y' = 1 + y^2.$$

The function $y = \tan x$ is a solution on the interval $(-\pi/2, \pi/2)$, since

$$\frac{d}{dx}(\tan x) = \sec^2 x = 1 + \tan^2 x.$$

More generally, $y = \tan(x - C)$ is a solution on $(-C - \pi/2, -C + \pi/2)$ for any constant C.

範例

Note

$y = C \tan x$ is not a solution for $C \neq 1$.

Integration of differential equations typically introduces arbitrary constants. The nature of these constants defines the general solution.

Definition 0.4. ***General and Particular Solutions.***
Let

$$y = \varphi(x, C_1, \ldots, C_n)$$

be a family of solutions to an n-th order differential equation, depending on n constants.

- This family is called the **general solution** if the constants $C_1, \ldots, C_n$ are independent.
- A solution obtained by assigning specific values to these constants is called a **particular solution**.

Independence of the constants requires that the Jacobian determinant of the function and its derivatives with respect to the constants is nonzero:

$$\frac{D(\varphi, \varphi', \ldots, \varphi^{(n-1)})}{D(C_1, C_2, \ldots, C_n)} = \begin{vmatrix} \frac{\partial \varphi}{\partial C_1} & \cdots & \frac{\partial \varphi}{\partial C_n} \\ \frac{\partial \varphi'}{\partial C_1} & \cdots & \frac{\partial \varphi'}{\partial C_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi^{(n-1)}}{\partial C_1} & \cdots & \frac{\partial \varphi^{(n-1)}}{\partial C_n} \end{vmatrix} \not\equiv 0,$$

where

$$\begin{aligned} \varphi &= \varphi(x, C_1, C_2, \cdots, C_n), \\ \varphi' &= \varphi'(x, C_1, C_2, \cdots, C_n), \\ \varphi^{(n-1)} &= \varphi^{(n-1)}(x, C_1, C_2, \cdots, C_n). \end{aligned}$$

定義

Example 0.3. Harmonic Oscillator. For the equation

$$\theta'' + a^2\theta = 0,$$

the function

$$\theta(t) = C_1 \sin(at) + C_2 \cos(at)$$

is the general solution, containing two independent constants. Functions such as $\theta = 3\sin(at)$ are particular solutions.

範例

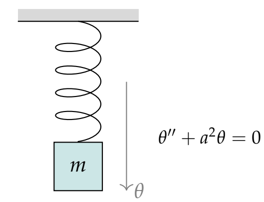

Figure 1: A simple harmonic oscillator, such as a mass on a spring, is modelled by a second-order linear ODE.

Initial Value Problems

A differential equation characterises the local, instantaneous laws governing a system. However, these laws alone are insufficient to determine a unique global state; one must also constrain the system at a specific initial point.

Consider the motion of a particle under gravity, neglecting air resistance. Let $y(t)$ denote the vertical position at time t, with the y-axis oriented upwards. We employ Newton's notation for time derivatives, where $\dot{y} \equiv \frac{dy}{dt}$ represents velocity and $\ddot{y} \equiv \frac{d^2y}{dt^2}$ represents acceleration.

By Newton's second law, the acting force is $-mg$, where g denotes the acceleration due to gravity. This yields the equation of motion:

$$m\ddot{y} = -mg \implies \ddot{y} = -g. \tag{2}$$

Integrating (2) with respect to t introduces an arbitrary constant C_1:

$$\dot{y}(t) = -gt + C_1.$$

Integrating a second time yields the position function:

$$y(t) = -\frac{1}{2}gt^2 + C_1 t + C_2. \tag{3}$$

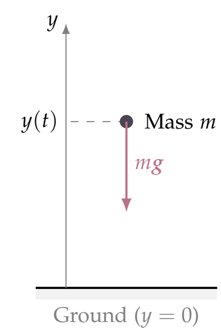

Figure 2: Free-fall motion under gravity.

The family of functions (3) represents the **general solution**. In classical literature, the process of solving a differential equation is often termed **integrating** the equation, and the resulting solution is called an **integral**. This terminology reflects the fact that solving an ODE is fundamentally an inverse operation to differentiation. The solution contains two arbitrary constants, indicating that the differential equation (2) admits infinitely many trajectories. This indeterminacy arises because (2) only expresses the instantaneous law of motion; objects released from different heights or with different initial velocities will exhibit distinct paths while satisfying the same law.

To determine a specific motion, we must consider the initial state of the system. We prescribe the **initial conditions** at time $t = 0$:

$$y(0) = y_0, \quad \dot{y}(0) = v_0, \tag{4}$$

where y_0 and v_0 are known data. Substituting these initial conditions into the general solution determines the constants: $C_2 = y_0$ and $C_1 = v_0$. The unique solution for the motion is:

$$y(t) = -\frac{1}{2}gt^2 + v_0 t + y_0. \tag{5}$$

The combination of the differential equation (2) and the initial conditions (4) constitutes an **Initial Value Problem** (IVP). We also refer to (5) as the solution to the system:

$$\begin{cases} \ddot{y} = -g, \\ y(0) = y_0, \quad \dot{y}(0) = v_0. \end{cases}$$

This formulation is also known as the **Cauchy Problem**.

Definition 0.5. ***The Cauchy Problem.***
The Initial Value Problem (or Cauchy Problem) for an n-th order differential equation asks for a solution $y(x)$ satisfying:

$$\begin{cases} y^{(n)} = F(x, y, y', \dots, y^{(n-1)}), \\ y(x_0) = y_0, \quad y'(x_0) = y'_0, \quad \dots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}, \end{cases} \tag{6}$$

where x_0 is the initial value of the independent variable, and $\{y_0, \dots, y_0^{(n-1)}\}$ are the specified initial states.

定義

The fundamental theory of ODEs concerns the existence and uniqueness of solutions to (6). We shall establish that continuity of F guarantees local existence, while stronger conditions, such as Lipschitz continuity (in the dependent variables), ensure uniqueness. These results are proven in the subsequent chapters.

Note

For a first-order IVP $y' = f(x, y)$, a typical uniqueness hypothesis is a *Lipschitz condition in y*: there is an $L > 0$ such that $|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$ on a region of interest.

0.2 *Geometric Interpretation*

A first-order equation prescribes the slope of a solution curve at each point, so it admits a geometric interpretation.

Consider the first-order differential equation resolved with respect to the derivative:

$$\frac{dy}{dx} = f(x, y), \tag{7}$$

where f is a continuous function defined on a domain $G \subseteq \mathbb{R}^2$.
If $y = \varphi(x)$ is a solution on an interval I, its graph

$$\Gamma = \{(x, \varphi(x)) \mid x \in I\}$$

is a smooth curve in the plane. We call Γ an **integral curve** of the differential equation. Since $\varphi'(x) = f(x, \varphi(x))$, the slope of the tangent at any point $P(x, y)$ on Γ equals $f(x, y)$.

Definition 0.6. ***Direction Field.***
A **line element** at a point $P(x, y) \in G$ is a short line segment passing through P with slope $k = f(x, y)$. The collection of all such line elements in G is called the **direction field** (or line element field) of the differential equation.

定義

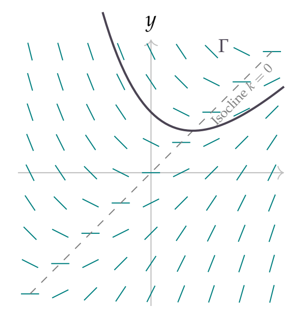

Figure 3: Direction field for $y' = x - y$. The dashed line is $y = x$ (slope 0). The curve Γ is a particular solution.

Solving (7) amounts to finding a curve whose tangent agrees with the line elements of the direction field.

Proposition 0.1. ***Tangency Condition.***
A smooth curve $y = \psi(x)$ in G is an integral curve of equation (7) if and only if, at every point P on the curve, the tangent line coincides with the line element of the direction field at P.

命題

Proof

If $y = \psi(x)$ matches the direction field, then $\psi'(x) = f(x, \psi(x))$ for all x, so it satisfies (7). Conversely, any solution has tangent slope $f(x, \psi(x))$ at each point, hence matches the line elements.

■

To assist in sketching direction fields by hand, one identifies the loci of points where the slope is constant.

Definition 0.7. ***Isoclines.***
An **isocline** of the differential equation $y' = f(x, y)$ is a curve defined by the equation

$$f(x, y) = k,$$

where k is a constant. Along this curve, the line elements of the direction field are parallel, sharing the fixed slope k.

定義

Example 0.4. Isoclines of a Linear Equation. Consider the equation $y' = x - y$. By *Isoclines*, the isoclines are the family of parallel lines

$x - y = k$, or $y = x - k$.

- On the line $y = x$ (where $k = 0$), all solution curves must have horizontal tangents.
- On the line $y = x - 1$ (where $k = 1$), solution curves have slope 1.

This structure is illustrated in *Figure 3*.

範例

Note

$y = x - 1$ is itself a solution (a line with slope 1), making it a particular integral curve.

Symmetric Form and Singularities

In the explicit form $y' = f(x, y)$, the geometric interpretation fails where the tangent is vertical (slope infinite). To treat the variables impartially and accommodate arbitrary tangents, we employ the symmetric form.
Consider the ratio:

$$\frac{dy}{dx} = -\frac{P(x, y)}{Q(x, y)}. \tag{8}$$

If $Q(x_0, y_0) \neq 0$, the slope is finite. If $Q(x_0, y_0) = 0$ but $P(x_0, y_0) \neq 0$, the slope is undefined. However, by considering x as a function of y, we obtain the inverse relation:

$$\frac{dx}{dy} = -\frac{Q(x, y)}{P(x, y)}.$$

At such points, $dx/dy = 0$, indicating a well-defined vertical tangent in the xy-plane.

Definition 0.8. ***Symmetric Form.***
The **symmetric form** of a first-order differential equation is

$$P(x, y)\, dx + Q(x, y)\, dy = 0. \tag{9}$$

A curve is an integral curve of (9) if at every point its tangent vector (dx, dy) is orthogonal to the field vector (P, Q).

定義

The direction field becomes indeterminate only when both components vanish.

Definition 0.9. ***Singular Point.***
A point $(x_0, y_0) \in G$ is a **singular point** of (9) if

$$P(x_0, y_0) = 0 \quad \text{and} \quad Q(x_0, y_0) = 0.$$

At a singular point, the direction field is undefined.

定義

Example 0.5. Radial and Circular Fields.

Radial Field ($ydx - xdy = 0$): Equivalent to $y' = y/x$. The origin is a singular point (*definition* 0.9). Isoclines are lines $y = kx$; integral curves are rays $y = Cx$ emanating from the origin (a *node*).

Circular Field ($xdx + ydy = 0$): Equivalent to $y' = -x/y$. The origin is a singular point (*definition* 0.9). Integral curves are concentric circles $x^2 + y^2 = C^2$ around the origin (a *center*).

範例

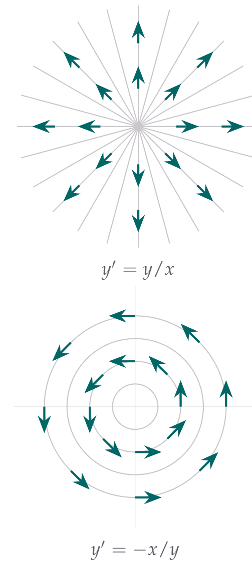

Figure 4: Integral curves near singular points. Top: A radial node (source). Bottom: A center (vortex).

Example 0.6. Magnetic Dipole Field. We model the bar magnet as two point magnetic charges of strengths $m_1 = +1$ at $(-a, 0)$ and $m_2 = -1$ at $(a, 0)$.
At a point (x, y), let r_1 and r_2 be the distances to these points. We take the field of each charge to be directed along the displacement to (x, y), so

$$\mathbf{H}_1 = \frac{m_1}{r_1^3}(x + a, y), \qquad \mathbf{H}_2 = \frac{m_2}{r_2^3}(x - a, y).$$

Thus $\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2$, and its components are:

$$U(x, y) = \frac{x + a}{r_1^3} - \frac{x - a}{r_2^3},$$
$$V(x, y) = \frac{y}{r_1^3} - \frac{y}{r_2^3},$$

where $r_1 = \sqrt{(x + a)^2 + y^2}$ and $r_2 = \sqrt{(x - a)^2 + y^2}$. The field lines are integral curves of:

$$\begin{aligned} \frac{dy}{dx} &= \frac{V(x, y)}{U(x, y)} \\ &= \frac{\{[(x - a)^2 + y^2]^{3/2} - [(x + a)^2 + y^2]^{3/2}\}y}{(x + a)[(x - a)^2 + y^2]^{3/2} - (x - a)[(x + a)^2 + y^2]^{3/2}}. \end{aligned} \tag{10}$$

The points $(\pm a, 0)$ are singularities, acting as the source and sink of the field lines (see *figure* 5).

範例

Remark.

Singular points govern the global topology of the integral curves. Later, we will study these systematically as **equilibrium points** of dynamical systems, classifying them into types such as nodes, saddles, centers, and foci.

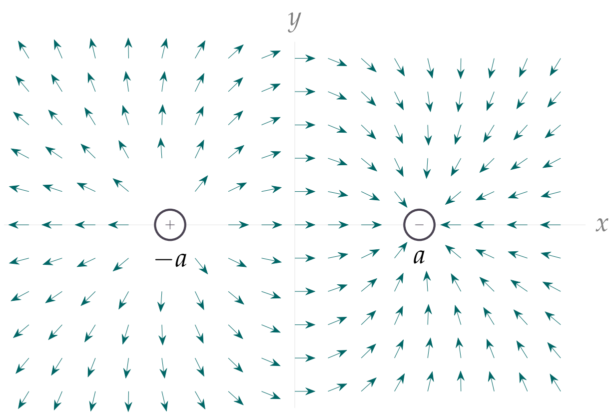

Figure 5: Direction field of the magnetic dipole.

Geometric View of the Initial Value Problem. Solving the Initial Value Problem (6) for $n = 1$:

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0,$$

is geometrically equivalent to finding a smooth curve passing through (x_0, y_0) tangent to the direction field. This perspective motivates **Euler's method**, which constructs an approximate solution as a polygonal line following the field vectors. Even without an explicit formula, the direction field reveals the stability and asymptotic behaviour of the system.

0.3 *Exercises*

In the following exercises, assume all functions and derivatives are defined on appropriate domains unless specified otherwise.

1. **Verification of Solutions.** Verify that the functions $y(x)$ given in the left column are solutions (or general solutions) to the corresponding differential equations in the right column.

   (a)
   $$y = C_1 e^{2x} + C_2 e^{-2x}, \; y'' - 4y = 0$$

   (b)
   $$y = \frac{\sin x}{x}, \; xy' + y = \cos x \quad (x \neq 0)$$

   (c)
   $$y = x\left(\int x^{-1} e^x \, dx + C\right), \; xy' - y = xe^x$$

(d)

$$y = \begin{cases} -\frac{1}{4}(x - C_1)^2, & x < C_1 \\ 0, & C_1 \leq x \leq C_2 \\ \frac{1}{4}(x - C_2)^2, & x > C_2 \end{cases}, \quad y' = \sqrt{|y|}$$

For (d): pay particular attention to the continuity of the derivative at the points $x = C_1$ and $x = C_2$.

2. **Solving Initial Value Problems.** Find the unique solution to the following Initial Value Problems.
   (a) $y''' = x$, subject to $y(0) = a_0$, $y'(0) = a_1$, $y''(0) = a_2$.
   (b) $\frac{dy}{dx} = f(x)$, subject to $y(0) = 0$, where f is a continuous function.
   (c) $\frac{dR}{dt} = -aR$, subject to $R(0) = 1$, where $a > 0$ is a constant.
   (d) $\frac{dy}{dx} = 1 + y^2$, subject to $y(x_0) = y_0$.

For (c): Separate Variables.

3. **Sketching Direction Fields.** Sketch the direction fields for the following differential equations on the domain $[-2, 2] \times [-2, 2]$. Indicate any symmetries or singular points.
   (a) $y' = \frac{xy}{|xy|}$ (Signum field; undefined when $xy = 0$)
   (b) $y' = (y - 1)^2$
   (c) $y' = x^2 + y^2$

4. **Qualitative Analysis via Isoclines.** Utilise the method of isoclines to sketch the direction field and describe the qualitative behaviour of the family of integral curves for the following equations.
   (a) $y' = 1 + xy$
   (b) $y' = x^2 - y^2$

For (b): consider the regions divided by the lines $y = x$ and $y = -x$. Where is the slope positive, negative, or zero?

5. **The Magnetic Dipole.** Based on the physical discussion of the magnetic field in *Figure 5*, defined by Equation (10):
   (a) Verify analytically that the points $(\pm a, 0)$ are singular points of the associated symmetric form (equivalently, zeros of the planar field (U, V)).
   (b) Sketch the behaviour of the field lines near the origin $(0, 0)$. What is the slope of the field line passing exactly through the origin?

6. **Curvature of Integral Curves.** The geometric interpretation of the ODE $y' = f(x, y)$ assigns a tangent direction to every point. It also implicitly assigns a *curvature* to the integral curve passing through that point.
   (a) The curvature κ of a plane curve $y(x)$ is given by

$$\kappa = \frac{|y''|}{(1 + (y')^2)^{3/2}},$$

show that for an integral curve of $y' = f(x,y)$, the curvature at any point (x,y) can be expressed solely in terms of f and its partial derivatives as:

$$\kappa(x,y) = \frac{\left|\frac{\partial f}{\partial x} + f(x,y)\frac{\partial f}{\partial y}\right|}{(1+f(x,y)^2)^{3/2}}.$$

(b) Apply this formula to the "Circular Field" equation $y' = -\frac{x}{y}$. Show that $\kappa(x,y)$ is constant along any circle $x^2 + y^2 = R^2$, and find the value of this curvature in terms of R.

7. **Locus of Inflection Points.** An integral curve has an inflection point where $y'' = 0$.

   (a) Using the Chain Rule, show that the locus of all inflection points of the solutions to $y' = f(x,y)$ is given by the algebraic curve:
$$\frac{\partial f}{\partial x} + f(x,y)\frac{\partial f}{\partial y} = 0.$$

   (b) Consider the differential equation $y' = y^2 - x$. Determine the equation of the curve of inflection points.

   (c) Sketch the direction field for $y' = y^2 - x$ together with the isocline $y^2 - x = 0$ (where slopes are horizontal) and the curve of inflection points derived in (b). How do the integral curves cross these loci?

8. **Independence of Parameters.** We defined the general solution of a second-order equation as a family $y = \varphi(x, C_1, C_2)$ where the constants are independent, meaning the Jacobian determinant is non-zero. Consider the Harmonic Oscillator solution

$$\varphi(t, C_1, C_2) = C_1 \sin(at) + C_2 \cos(at)$$

(with $a \neq 0$).

   (a) Compute the Jacobian determinant:

$$J = \det \begin{bmatrix} \frac{\partial \varphi}{\partial C_1} & \frac{\partial \varphi}{\partial C_2} \\ \frac{\partial \varphi'}{\partial C_1} & \frac{\partial \varphi'}{\partial C_2} \end{bmatrix}.$$

   (b) Verify that J is nowhere zero, thereby proving that C_1 and C_2 are independent parameters for this family.

1

Elementary Integration Methods

While the study of differential equations frequently requires numerical or qualitative analysis, the classical theory focuses on constructing explicit solutions using elementary functions and their integrals. This approach, developed largely by Newton, Leibniz, the Bernoullis, and Euler, provides the foundational tools for exact solvability. Although Liouville (1841) demonstrated that most differential equations cannot be solved by elementary integration, the methods in this chapter remain essential for handling the structured equations that arise in physical and geometric problems.

1.1 *Exact Equations*

We consider the first-order differential equation in the symmetric form introduced in *definition* 0.8:

$$P(x, y)\, dx + Q(x, y)\, dy = 0. \tag{1.1}$$

Definition 1.1. ***Exact Equation.***
The differential equation (1.1) is called an **exact equation** if there exists a differentiable function $\Phi(x, y)$ such that

$$d\Phi(x, y) = P(x, y)\, dx + Q(x, y)\, dy. \tag{1.2}$$

This condition implies that the vector field (P, Q) is the gradient of a potential function Φ:

$$\frac{\partial \Phi}{\partial x} = P(x, y), \quad \frac{\partial \Phi}{\partial y} = Q(x, y).$$

定義

If (1.1) is exact, it reduces to the total differential equation $d\Phi(x, y) = 0$. Consequently, the general solution is given implicitly by the level sets of the potential function:

$$\Phi(x, y) = C, \tag{1.3}$$

where C is an arbitrary constant. We refer to (1.3) as the **general integral** of the equation.

Example 1.1. Simple Exact Equation. Consider the equation

$$2xy^3\,dx + 3x^2y^2\,dy = 0.$$

By inspection, the left-hand side is the total differential of $\Phi(x,y) = x^2y^3$, since

$$\frac{\partial}{\partial x}(x^2y^3) = 2xy^3 \quad \text{and} \quad \frac{\partial}{\partial y}(x^2y^3) = 3x^2y^2.$$

Thus, the equation is exact, and its general integral is $x^2y^3 = C$.

範例

Criterion for Exactness

To determine solvability without relying on inspection, we require a systematic criterion for exactness. This is provided by the compatibility condition for the existence of a potential.

Theorem 1.1. ***Exactness Criterion.***
Let $P(x,y)$ and $Q(x,y)$ be functions defined on a simply connected rectangular region $R = (\alpha,\beta)\times(\gamma,\delta)$ (simply connected here means every closed loop in R can be continuously shrunk to a point while staying in R). Assume that P and Q are continuous and possess continuous first-order partial derivatives $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ on R.
The differential equation $P\,dx + Q\,dy = 0$ is exact if and only if

$$\frac{\partial P}{\partial y} \equiv \frac{\partial Q}{\partial x} \tag{1.4}$$

identically on R. When this condition holds, the general integral is given by

$$\int_{x_0}^{x} P(t,y)\,dt + \int_{y_0}^{y} Q(x_0,t)\,dt = C, \tag{1.5}$$

where (x_0,y_0) is any fixed point in R.

定理

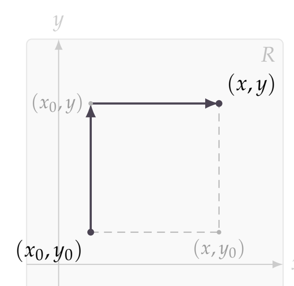

Figure 1.1: Integration paths in the rectangular domain R. The solid path corresponds to the derivation of formula (1.5) (vertical then horizontal), while the dashed path represents an alternative route, yielding the same potential due to exactness.

Necessity.
Suppose the equation is exact, so there is Φ with $\Phi_x = P$ and $\Phi_y = Q$. Fix $(x,y)\in R$, and for small h,k define

$$\Delta(h,k) = \Phi(x+h,y+k) - \Phi(x+h,y) - \Phi(x,y+k) + \Phi(x,y).$$

Computing Δ in the x-direction gives

$$\Delta(h,k) = \int_x^{x+h} \big(P(s,y+k) - P(s,y)\big)\,ds.$$

Computing Δ in the y-direction gives

$$\Delta(h,k) = \int_y^{y+k} \big(Q(x+h,t) - Q(x,t)\big)\,dt.$$

Hence

$$\frac{\Delta(h,k)}{hk} = \frac{1}{h}\int_x^{x+h} \frac{P(s,y+k)-P(s,y)}{k}\,ds = \frac{1}{k}\int_y^{y+k} \frac{Q(x+h,t)-Q(x,t)}{h}\,dt.$$

Using continuity of P_y and Q_x, letting $(h,k) \to (0,0)$ yields

$$\frac{\partial P}{\partial y}(x,y) = \frac{\partial Q}{\partial x}(x,y).$$

Therefore $\partial P/\partial y \equiv \partial Q/\partial x$ on R.

証明終

Sufficiency.

Assume (1.4) holds. We construct Φ explicitly. Define a candidate function by integrating P with respect to x:

$$\Phi(x,y) = \int_{x_0}^{x} P(t,y)\,dt + \psi(y), \tag{1.6}$$

where $\psi(y)$ is an arbitrary function of y. Differentiating (1.6) with respect to y:

$$\frac{\partial \Phi}{\partial y} = \frac{\partial}{\partial y}\int_{x_0}^{x} P(t,y)\,dt + \psi'(y) = \int_{x_0}^{x} \frac{\partial P}{\partial y}(t,y)\,dt + \psi'(y).$$

Using the condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$:

$$\frac{\partial \Phi}{\partial y} = \int_{x_0}^{x} \frac{\partial Q}{\partial x}(t,y)\,dt + \psi'(y) = \Big[Q(t,y)\Big]_{x_0}^{x} + \psi'(y) = Q(x,y) - Q(x_0,y) + \psi'(y).$$

For Φ to satisfy $\frac{\partial \Phi}{\partial y} = Q(x,y)$, we require:

$$Q(x,y) - Q(x_0,y) + \psi'(y) = Q(x,y) \implies \psi'(y) = Q(x_0,y).$$

Integrating $\psi'(y)$ gives $\psi(y) = \int_{y_0}^{y} Q(x_0,t)\,dt$. Substituting this back into (1.6) yields the potential function:

$$\Phi(x,y) = \int_{x_0}^{x} P(t,y)\,dt + \int_{y_0}^{y} Q(x_0,t)\,dt.$$

This function satisfies $d\Phi = P\,dx + Q\,dy$, proving the equation is exact.

証明終

Remark.

The formula (1.5) represents a line integral of the field (P, Q) along a rectilinear path from (x_0, y_0) to (x, y). Alternative paths yield the same potential (up to a constant) due to the condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, which ensures path independence in a simply connected domain.

Example 1.2. Solving via Integration. Solve the differential equation

$$(2x \sin y + 3x^2 y)\, dx + (x^3 + x^2 \cos y + y^2)\, dy = 0.$$

Let $P = 2x \sin y + 3x^2 y$ and $Q = x^3 + x^2 \cos y + y^2$. We check the exactness condition (*theorem* 1.1):

$$\frac{\partial P}{\partial y} = 2x \cos y + 3x^2, \quad \frac{\partial Q}{\partial x} = 3x^2 + 2x \cos y.$$

Since the partial derivatives are identical, the equation is exact. We seek Φ such that $\Phi_x = P$. Integrating P with respect to x:

$$\Phi(x, y) = \int (2x \sin y + 3x^2 y)\, dx = x^2 \sin y + x^3 y + \psi(y).$$

To determine $\psi(y)$, we differentiate with respect to y and equate to Q:

$$\frac{\partial \Phi}{\partial y} = x^2 \cos y + x^3 + \psi'(y) = x^3 + x^2 \cos y + y^2.$$

Cancelling common terms, we find $\psi'(y) = y^2$, which implies $\psi(y) = \frac{1}{3} y^3$. The general integral is therefore

$$x^2 \sin y + x^3 y + \frac{1}{3} y^3 = C.$$

範例

Note

Experienced practitioners often solve exact equations by **grouping terms** to recognise standard differentials, avoiding explicit partial integration. For the previous example:

$$\begin{aligned} &(2x \sin y\, dx + x^2 \cos y\, dy) + (3x^2 y\, dx + x^3\, dy) + y^2\, dy \\ &= d(x^2 \sin y) + d(x^3 y) + d\left(\tfrac{1}{3} y^3\right) = 0. \end{aligned}$$

This method is efficient but relies on inspection.

Topology and Multivalued Potentials

The sufficiency in *theorem* 1.1 relies on the domain R being simply connected (e.g., a rectangle). If the domain contains "holes", the con-

dition
$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$
is necessary locally but does not guarantee a single-valued potential globally.

Example 1.3. The Vortex Field. Consider the equation on $\mathbb{R}^2 \setminus \{(0,0)\}$:
$$\frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 0.$$
Here $P = \frac{-y}{x^2+y^2}$ and $Q = \frac{x}{x^2+y^2}$. Checking the derivatives:
$$\frac{\partial P}{\partial y} = \frac{y^2-x^2}{(x^2+y^2)^2}, \quad \frac{\partial Q}{\partial x} = \frac{y^2-x^2}{(x^2+y^2)^2}.$$
The condition holds everywhere except at the origin. However, the integral around the unit circle is non-zero (2π). The potential function is the angle $\theta = \arctan(y/x)$, which is **multivalued**. While we can write the solution as $\arctan(y/x) = C$, one must be careful with the domain of definition branches.

範例

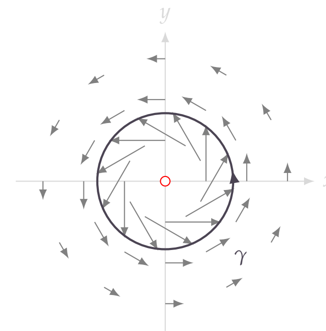

Figure 1.2: The vortex field $\mathbf{F} = \frac{-y\mathbf{i}+x\mathbf{j}}{x^2+y^2}$. Vectors are tangent to circles with magnitude decreasing as $1/r$. The line integral along the unit circle γ is 2π, illustrating the failure of exactness due to the hole at the origin.

1.2 *Separable Equations*

The method of separation of variables is the simplest and most commonly used technique for solving first-order differential equations.

Definition 1.2. ***Separable Equation.***
A first-order differential equation in symmetric form
$$P(x,y)\,dx + Q(x,y)\,dy = 0$$
is called a **separable equation** if the coefficients can be factored as products of single-variable functions:
$$P(x,y) = X(x)Y_1(y), \quad Q(x,y) = X_1(x)Y(y).$$
Substituting these into the equation yields
$$X(x)Y_1(y)\,dx + X_1(x)Y(y)\,dy = 0. \tag{1.7}$$

定義

The strategy is to regroup terms so that each differential is multiplied only by a function of its own variable.

General Solution Procedure

If $X_1(x)Y_1(y) \neq 0$, we may divide (1.7) by this factor to separate the variables:

$$\frac{X(x)}{X_1(x)}\,dx + \frac{Y(y)}{Y_1(y)}\,dy = 0. \tag{1.8}$$

Equation (1.8) is now an **exact equation** (*theorem* 1.1), as the first term depends only on x and the second only on y. The general integral is obtained by direct integration:

$$\int \frac{X(x)}{X_1(x)}\,dx + \int \frac{Y(y)}{Y_1(y)}\,dy = C. \tag{1.9}$$

> **Proposition 1.1.** ***Loss of Solutions.***
> The division by $X_1(x)Y_1(y)$ assumes this product is non-zero. The zeros of the divisors correspond to additional special solutions (integral curves of the symmetric equation) that may be lost in the general integral (1.9).
>
> · If $X_1(a) = 0$, then $x = a$ is an integral curve of the symmetric equation (1.7) (a vertical line in the xy-plane).
>
> · If $Y_1(b) = 0$, then $y = b$ is an integral curve of (1.7) (a horizontal line).
>
> A complete solution set consists of the general integral (1.9) together with these additional special solutions.
>
> 命題

For the proof we work with the separable equation in symmetric form (1.7).

> *Vertical solutions.*
>
> Fix a such that $X_1(a) = 0$ and consider the curve $x \equiv a$. Along this curve we have $dx = 0$, so the equation reduces to
>
> $$X_1(a)Y(y)\,dy = 0,$$
>
> which holds identically since $X_1(a) = 0$. Hence $x = a$ is an integral curve.
>
> 証明終

> *Horizontal solutions.*
>
> Fix b such that $Y_1(b) = 0$ and consider the curve $y \equiv b$. Along this curve we have $dy = 0$, so the equation reduces to
>
> $$X(x)Y_1(b)\,dx = 0,$$
>
> which holds identically since $Y_1(b) = 0$. Hence $y = b$ is an integral curve.
>
> Finally, the separation step divides (1.7) by $X_1(x)Y_1(y)$ and there-

fore applies only where $X_1(x)Y_1(y) \neq 0$. The integral curves $x = a$ and $y = b$ lie entirely in the set where this product vanishes, and so they may be excluded by the division and fail to appear in (1.9). Adding them back yields the full solution set.

証明終

Example 1.4. Separation of Variables. Solve the equation

$$(x^2+1)(y^2-1)\,dx + xy\,dy = 0. \tag{1.10}$$

範例

Solution

We separate the variables by dividing by $x(y^2-1)$ (assuming $x \neq 0, y \neq \pm 1$):

$$\frac{x^2+1}{x}\,dx + \frac{y}{y^2-1}\,dy = 0.$$

Integrating term by term:

$$\int \left(x + \frac{1}{x}\right) dx + \frac{1}{2}\int \frac{2y}{y^2-1}\,dy = C_1$$

$$\frac{1}{2}x^2 + \ln|x| + \frac{1}{2}\ln|y^2-1| = C_1.$$

Multiplying by 2 and exponentiating:

$$e^{x^2}x^2|y^2-1| = e^{2C_1}.$$

Letting $C = \pm e^{2C_1}$, we write the solution explicitly for y^2:

$$y^2 = 1 + \frac{C}{x^2}e^{-x^2}. \tag{1.11}$$

Special Solutions: The division excluded $x = 0$ and $y = \pm 1$.

$y = \pm 1$ satisfies the original equation (1.10) identically. These solutions are recovered from (1.11) by setting $C = 0$.

$x = 0$ satisfies the equation (vertical line) but is singular in the form (1.11).

The direction field and integral curves for this system are shown in *figure* 1.3.

■

Note

$(0,1)$ and $(0,-1)$ are singular points where the direction field is undefined.

Separable equations often reveal subtle issues regarding the uniqueness of solutions at points where the uniqueness theorem's hypothe-

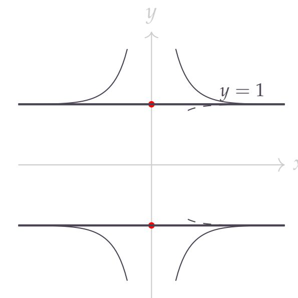

Figure 1.3: Integral curves for $(x^2+1)(y^2-1)dx + xydy = 0$. The lines $y = \pm 1$ are solutions. The points $(0, \pm 1)$ are singularities.

ses fail.

Example 1.5. Non-Uniqueness. Solve

$$y' = \frac{3}{2}y^{1/3}.$$

範例

Solution

Separating variables for $y \neq 0$:

$$y^{-1/3}\,dy = \frac{3}{2}\,dx \implies \frac{3}{2}y^{2/3} = \frac{3}{2}x + C_1.$$

Simplifying yields the family of semi-cubic parabolas:

$$y^2 = (x+C)^3, \quad x \geq -C. \tag{1.12}$$

Additionally, $y = 0$ is a solution.

Observing the integral curves, we find a distinct difference between this example and the previous one. In eq. (1.10), solutions are locally unique away from the singular points. However, here, for any point $(x^*, 0)$ on the x-axis, infinitely many integral curves pass through it. A solution can travel along the line $y = 0$ and branch off onto a parabola $(x+C)^3$ at any point.

Thus, the solution satisfying an initial condition $y(x_0) = y_0$ is locally unique if $y_0 \neq 0$, but locally non-unique if $y_0 = 0$. The theoretical explanation for this phenomenon is reserved for the general discussion on existence and uniqueness in the later chapters.

■

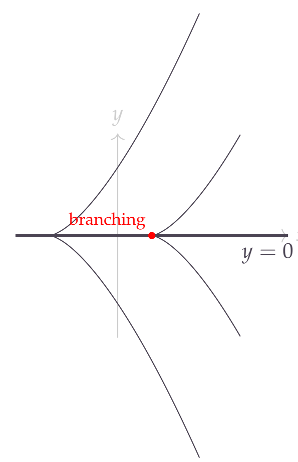

Figure 1.4: Non-uniqueness for $y' = \frac{3}{2}y^{1/3}$. Solutions can stay on the singular solution $y = 0$ and branch off at any point x_0 (e.g., red dot) onto a cubic parabola $y^2 = (x - x_0)^3$.

Application: Terminal Velocity

We consider an object of mass m falling in air with resistance proportional to the square of velocity. Let the x-axis point vertically downwards. By Newton's second law:

$$m\ddot{x} = mg - k\dot{x}^2,$$

where $k > 0$ is the damping coefficient. Letting $v = \dot{x}$, the equation becomes:

$$\frac{dv}{dt} = g - \frac{k}{m}v^2 \quad (v > 0). \tag{1.13}$$

This is a separable equation. When the right-hand side is not zero, we separate variables:

$$\frac{dv}{g - \frac{k}{m}v^2} = dt.$$

Integration yields the general solution (for $v \neq A$):

$$v(t) = A\frac{Ce^{2at}+1}{Ce^{2at}-1} \quad (t \geq 0), \tag{1.14}$$

where $a = \sqrt{\frac{kg}{m}}$, $A = \sqrt{\frac{mg}{k}}$, and C is an arbitrary constant determined by the initial velocity v_0.

> *Note*
>
> The separation step excludes the constant solution $v(t) \equiv A$, which also satisfies (1.13) (and is consistent with the restriction $v > 0$).

- If $0 < v_0 < A$, then $v(t)$ approaches A from below.
- If $v_0 > A$, then $v(t)$ approaches A from above.

The value A represents the limiting **terminal velocity**.

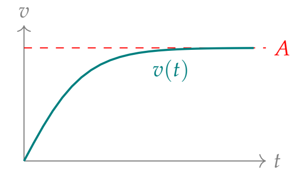

Figure 1.5: Velocity approaching terminal velocity A.

> *Remark (Sky Diving).*
>
> This model applies to skydiving. A skydiver has a small damping coefficient k_1 before opening the parachute (falling rapidly) and a large coefficient k_2 afterwards (falling slowly). The problem of skydiving is to determine the time T to open the parachute such that the landing velocity approximates $\sqrt{mg/k_2}$ within a fixed height H_0.

1.3 *Linear First-Order Equations*

We now turn our attention to the most important class of differential equations: those in which the unknown function and its derivative appear linearly.

> **Definition 1.3.** ***Linear Differential Equation.***
> A first-order differential equation is **linear** if it can be written in the form
>
> $$\frac{dy}{dx} + p(x)y = q(x), \tag{1.15}$$
>
> where the functions $p(x)$ and $q(x)$ are continuous on an interval $I = (a, b)$.
>
> - If $q(x) \equiv 0$, the equation is termed **homogeneous**.
> - If $q(x) \not\equiv 0$, the equation is **non-homogeneous**.
>
> 定義

The Homogeneous Case

Consider the homogeneous equation

$$\frac{dy}{dx} + p(x)y = 0. \tag{1.16}$$

Writing this in the symmetric form $dy + p(x)y\,dx = 0$, we observe that it is a separable equation. Assuming $y \neq 0$, we divide by y to obtain

$$\frac{dy}{y} = -p(x)\,dx.$$

Integration yields $\ln|y| = -\int p(x)\,dx + c$, which simplifies to

$$y(x) = Ce^{-\int p(x)\,dx}, \tag{1.17}$$

where C is a non-zero constant. However, we observe that $y \equiv 0$ is also a solution to (1.16). By allowing $C = 0$ in (1.17), we encompass this trivial solution. Thus, (1.17) represents the general solution for any arbitrary constant C.

The Non-Homogeneous Case: Integrating Factors

To solve the non-homogeneous equation (1.15), we rewrite it in symmetric differential form:

$$dy + [p(x)y - q(x)]\,dx = 0. \tag{1.18}$$

In general, this is not an exact equation (satisfying *theorem* 1.1). However, we may seek an **integrating factor** $\mu(x)$ such that multiplying (1.15) by $\mu(x)$ renders the left-hand side the derivative of a product. Multiplying (1.15) by a non-zero function $\mu(x)$:

$$\mu(x)\frac{dy}{dx} + \mu(x)p(x)y = \mu(x)q(x).$$

We require the left-hand side to be identically $\frac{d}{dx}[\mu(x)y] = \mu(x)\frac{dy}{dx} + \mu'(x)y$. Comparing terms, we must satisfy

$$\mu'(x) = \mu(x)p(x).$$

This is a separable homogeneous equation for μ. Choosing the particular solution where the integration constant is 1:

$$\mu(x) = e^{\int p(x)\,dx}. \tag{1.19}$$

Since the exponential function is never zero, this $\mu(x)$ is a valid integrating factor. The differential equation becomes

$$\frac{d}{dx}\left[e^{\int p(x)\,dx}y\right] = q(x)e^{\int p(x)\,dx}.$$

Integrating both sides with respect to x yields

$$e^{\int p(x)\,dx}y = \int q(x)e^{\int p(x)\,dx}\,dx + C.$$

Solving for y, we obtain the general solution.

Theorem 1.2. ***General Solution of Linear Equations.***
The general solution to the first-order linear differential equation $y' + p(x)y = q(x)$ is given by

$$y(x) = e^{-\int p(x)\,dx}\left(C + \int q(x)e^{\int p(x)\,dx}\,dx\right), \qquad (1.20)$$

where C is an arbitrary constant.

定理

Note

This technique is known as the **method of integrating factors**. A second approach, the *method of variation of constants*, appears in the exercises below (see the dedicated variation-of-constants problem).

Example 1.6. A Polynomial Perturbation. Solve the differential equation

$$\frac{dy}{dx} + \frac{1}{x}y = x^3 \quad (x \neq 0).$$

Here $p(x) = 1/x$. The integrating factor is

$$\mu(x) = e^{\int \frac{1}{x}\,dx} = e^{\ln|x|} = |x|.$$

Since the sign of μ does not affect the equation, we take $\mu(x) = x$. Multiplying the ODE by x:

$$x\frac{dy}{dx} + y = x^4 \implies \frac{d}{dx}(xy) = x^4.$$

Integrating gives $xy = \frac{1}{5}x^5 + C$, or

$$y = \frac{1}{5}x^4 + \frac{C}{x}.$$

This solution is valid on $(-\infty, 0)$ or $(0, \infty)$.

範例

The Initial Value Problem

For the Initial Value Problem (IVP)

$$y' + p(x)y = q(x), \quad y(x_0) = y_0, \qquad (1.21)$$

it is advantageous to express the indefinite integrals in (1.20) as definite integrals with a variable upper limit.

Corollary 1.1. *Cauchy Formula for Linear Equations.* The unique solution to the IVP (1.21) is

$$y(x) = y_0 e^{-\int_{x_0}^{x} p(t)\,dt} + \int_{x_0}^{x} q(s) e^{-\int_{s}^{x} p(t)\,dt}\,ds. \tag{1.22}$$

推論

Proof

Multiply the differential equation by the integrating factor

$$\mu(t) = e^{\int_{x_0}^{t} p(\tau)\,d\tau}.$$

Then

$$\frac{d}{dt}\left(y(t) e^{\int_{x_0}^{t} p(\tau)\,d\tau}\right) = q(t) e^{\int_{x_0}^{t} p(\tau)\,d\tau}.$$

Integrating from $t = x_0$ to $t = x$:

$$\left[y(t) e^{\int_{x_0}^{t} p(\tau)\,d\tau}\right]_{x_0}^{x} = \int_{x_0}^{x} q(s) e^{\int_{x_0}^{s} p(\tau)\,d\tau}\,ds.$$

Substituting $y(x_0) = y_0$ and rearranging yields the result. ∎

Structure of the Solution Space

The linearity of the operator $L[y] = y' + p(x)y$ imposes a rigid structure on the space of solutions. We summarise these properties below; proofs follow from the explicit formulae and the properties of the integral.

Proposition 1.2. ***Properties of Linear Solutions.***
Let $p(x)$ and $q(x)$ be continuous on I.

1. **Non-vanishing Homogeneous Solutions:** A solution to the homogeneous equation $y' + p(x)y = 0$ is either identically zero or never zero on I.
2. **Global Existence:** Solutions to linear equations exist on the entire interval I where p and q are continuous. Unlike non-linear equations (e.g., $y' = y^2$), they do not blow up in finite time.
3. **Superposition Principle:**
   - If y_1, y_2 solve the homogeneous equation, so does $c_1 y_1 + c_2 y_2$.
   - If y_h solves the homogeneous equation and y_p solves the non-homogeneous equation, then $y_h + y_p$ solves the non-homogeneous equation.
   - If y_1, y_2 solve the non-homogeneous equation, their difference $y_1 - y_2$ solves the homogeneous equation.
4. **General Solution Structure:** The general solution of the non-homogeneous

equation is the sum of the general solution of the homogeneous equation and any particular solution of the non-homogeneous equation.

5. **Uniqueness:** The solution to the Initial Value Problem (1.21) is unique.

命題

Property 1.

Let $y = \varphi(x)$ be a solution to $y' + p(x)y = 0$. By (1.17), every solution has the form

$$\varphi(x) = C\exp\left(-\int p(x)\,dx\right).$$

If $\varphi \not\equiv 0$, then $C \neq 0$. Since the exponential factor never vanishes, $\varphi(x) \neq 0$ for all $x \in I$. Hence a homogeneous solution is either identically zero or nowhere zero on I.

証明終

Property 5 (Uniqueness).

Suppose $y = \varphi_1(x)$ and $y = \varphi_2(x)$ solve (1.21). By linearity, their difference $\psi(x) = \varphi_1(x) - \varphi_2(x)$ satisfies the homogeneous equation with initial condition $\psi(x_0) = 0$. By (1.17), $\psi(x) = C\exp(-\int p(x)\,dx)$. Evaluating at $x = x_0$ gives $C = 0$, so $\psi \equiv 0$. Therefore $\varphi_1 \equiv \varphi_2$.

証明終

Example 1.7. Periodic Solutions. Consider the equation

$$\frac{dy}{dx} + ay = f(x), \tag{1.23}$$

where $a > 0$ is a constant and $f(x)$ is a continuous 2π-periodic function. We seek a 2π-periodic solution.

範例

Using (1.22) with $x_0 = 0$, the general solution is

$$y(x) = Ce^{-ax} + \int_0^x e^{-a(x-s)} f(s)\,ds.$$

For $y(x)$ to be 2π-periodic, we must have $y(x+2\pi) \equiv y(x)$. It suffices to enforce the condition at a single point, say $x = 0$, provided we invoke uniqueness.

Claim 1.1. $y(2\pi) = y(0)$ implies $y(x+2\pi) = y(x)$ for all x.

主張

Proof

Let $u(x) = y(x+2\pi) - y(x)$. Since $f(x+2\pi) = f(x)$, differentiation shows

$$u' + au = (y(x+2\pi)' + ay(x+2\pi)) - (y'(x) + ay(x)) = f(x) - f(x) = 0.$$

Thus $u(x)$ satisfies the homogeneous equation. If $y(2\pi) = y(0)$, then $u(0) = 0$. By Property 1, $u(x) \equiv 0$.

■

Solution

To find the required C, we set $y(2\pi) = y(0) = C$:

$$Ce^{-2\pi a} + \int_0^{2\pi} e^{-a(2\pi - s)} f(s)\, ds = C.$$

Solving for C:

$$C(1 - e^{-2\pi a}) = e^{-2\pi a} \int_0^{2\pi} e^{as} f(s)\, ds \implies C = \frac{1}{e^{2\pi a} - 1} \int_0^{2\pi} e^{as} f(s)\, ds.$$

Substituting C back into the general solution and simplifying using the periodicity of f yields the periodic solution:

$$y(x) = \frac{1}{e^{2a\pi} - 1} \int_x^{x+2\pi} e^{-a(x-s)} f(s) ds. \tag{1.24}$$

■

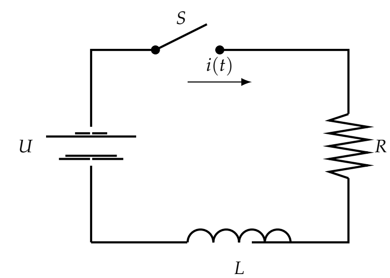

Figure 1.6: An RL series circuit with inductance L, resistance R, and voltage source U.

Example 1.8. RL Series Circuit. Consider a series circuit containing a resistor R, an inductor L, and a constant voltage source U (see *figure* 1.6). By Kirchhoff's voltage law, the sum of the voltage drops across the inductor ($L\frac{di}{dt}$) and the resistor (Ri) equals the supplied voltage:

$$L\frac{di}{dt} + Ri = U. \tag{1.25}$$

This is a linear first-order equation for the current $i(t)$. The homogeneous equation $Li' + Ri = 0$ has the general solution

$$i_h(t) = Ce^{-(R/L)t}.$$

A particular solution to the non-homogeneous equation is the constant current $i_p = U/R$. By the Superposition Principle (Property 4), the general solution is

$$i(t) = \frac{U}{R} + Ce^{-\frac{R}{L}t}.$$

If the switch is closed at $t = 0$ such that $i(0) = 0$, we find $0 = \frac{U}{R} + C$, so $C = -U/R$. The current evolves as:

$$i(t) = \frac{U}{R}\left(1 - e^{-\frac{R}{L}t}\right).$$

範例

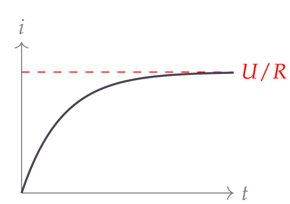

Figure 1.7: Current response in an RL circuit.

1.4 *Exercises*

1. **Exact Equations.** Determine whether the following differential equations are exact. For those that are exact, determine the general solution.

   (a) $(3x^2 - 1)\, dx + (2x + 1)\, dy = 0.$

   (b) $(x + 2y)\, dx + (2x - y)\, dy = 0.$

   (c) $(ax + by)\, dx + (bx + cy)\, dy = 0 \quad (a, b,$ and c are constants$)$.

   (d) $(ax - by)\, dx + (bx - cy)\, dy = 0 \quad (b \neq 0).$

   (e) $(t^2 + 1) \cos u\, du + 2t \sin u\, dt = 0.$

   (f) $(ye^x + 2e^x + y^2)\, dx + (e^x + 2xy)\, dy = 0.$

   (g) $\left(\frac{y}{x} + x^2\right) dx + (\ln x - 2y)\, dy = 0.$

   (h) $(ax^2 + by^2)\, dx + cxy\, dy = 0 \quad (a, b,$ and c are constants$)$.

   (i) $\frac{2s-1}{t}\, ds + \frac{s-s^2}{t^2}\, dt = 0.$

   (j) $xf(x^2 + y^2)\, dx + yf(x^2 + y^2)\, dy = 0$, where f is a continuously differentiable function.

2. **Separable Equations and Domains.** Solve the following differential equations. For each equation, indicate the region in the xy-plane where the differential equation is meaningful.

   (a) $\frac{dy}{dx} = \frac{x^2}{y}$

   (b) $\frac{dy}{dx} = \frac{x^2}{y(1+x^3)}$

   (c) $\frac{dy}{dx} + y^2 \sin x = 0$

   (d) $\frac{dy}{dx} = 1 + x + y^2 + xy^2$

   (e) $\frac{dy}{dx} = (\cos x \cos 2y)^2$

   (f) $x\frac{dy}{dx} = \sqrt{1 - y^2}$

   (g) $\frac{dy}{dx} = \frac{x - e^{-x}}{y + e^x}$

3. **Initial Value Problems.** Find the unique solutions to the following initial value problems:

   (a) $\sin 2x\, dx + \cos 3y\, dy = 0, \quad y\left(\frac{\pi}{2}\right) = \frac{\pi}{3}$

   (b) $x\, dx + ye^{-x}\, dy = 0, \quad y(0) = 1$

   (c) $\frac{dr}{d\theta} = r, \quad r(0) = 2$

   (d) $\frac{dy}{dx} = \frac{\ln |x|}{1+y^2}, \quad y(1) = 0$

   (e) $\sqrt{1 + x^2}\, \frac{dy}{dx} = xy^3, \quad y(0) = 1$

4. **Families of Integral Curves.** Solve the following differential equa-

tions and sketch the corresponding family of integral curves.

(a) $\frac{dy}{dx} = \cos x$

(b) $\frac{dy}{dx} = ay \quad (a \neq 0 \text{ is a constant})$

(c) $\frac{dy}{dx} = 1 - y^2$

(d) $\frac{dy}{dx} = y^n \quad \text{for } n \in \left\{\frac{1}{3}, 1, 2\right\}$

5. **The Pursuit Problem.** Suppose Person A starts from the origin on the xy-plane and moves along the positive x-axis with constant speed. Simultaneously, Person B starts from the point $(0, b)$, moves with the same constant speed as A, and tracks A; that is, B's velocity vector always points towards A. Derive the differential equation governing B's path and find the smooth trajectory of B's motion.

6. ⋆ **Osgood's Uniqueness Criterion.** Consider the differential equation

$$\frac{dy}{dx} = f(y), \tag{1.26}$$

where $f(y)$ is continuous in some neighbourhood of $y = a$ (specifically, the interval $[a - \epsilon, a + \epsilon]$), and $f(y) = 0$ if and only if $y = a$. Prove that at each point on the line $y = a$, the solution of equation (1.26) is locally unique if and only if the integral

$$\int_a^{a+\epsilon} \frac{dy}{f(y)} = +\infty \quad \text{(i.e., the integral diverges).}$$

7. **Application of Uniqueness.** Using the result of the previous exercise (without explicitly solving the equations), sketch the family of integral curves for the following differential equations and discuss the uniqueness of solutions at $y = 0$:

(a) $\frac{dy}{dx} = \sqrt{|y|}$

(b) $\frac{dy}{dx} = \begin{cases} y \ln |y|, & y \neq 0, \\ 0, & y = 0. \end{cases}$

8. **Solving Linear Equations.** Find the general solutions to the following differential equations. If an initial condition is provided, find the unique particular solution.

(a) $\frac{dy}{dx} + 2y = xe^{-x}$

(b) $\frac{dy}{dx} + y \tan x = \sin 2x$

(c) $x\frac{dy}{dx} + 2y = \sin x, \quad y(\pi) = \frac{1}{\pi}$

(d) $\frac{dy}{dx} - \frac{1}{1-x^2} y = 1 + x, \quad y(0) = 1$

9. **Reduction to Linear Form.** By choosing an appropriate substitution (change of variables), transform the following non-linear

differential equations into linear differential equations and solve them.

(a) $\frac{dy}{dx} = \frac{x^2+y^2}{2y}$

(b) $\frac{dy}{dx} = \frac{y}{x+y^2}$

(c) $3xy^2\frac{dy}{dx} + y^3 + x^3 = 0$

(d) $\frac{dy}{dx} = \frac{1}{\cos y} + x \tan y$

For (b): Consider x as the dependent variable of y.

10. Differential Inequalities (Gronwall's Lemma type). Let $y = \varphi(x)$ be a differentiable function satisfying the differential inequality

$$y' + a(x)y \leq 0 \quad (x \geq 0).$$

Prove that:

$$\varphi(x) \leq \varphi(0)e^{-\int_0^x a(s)\,ds} \quad (x \geq 0).$$

11. Method of Variation of Constants. In the text, we derived the general solution to the linear equation using an integrating factor. An alternative method, due to Lagrange, is the *variation of constants*. Consider the non-homogeneous linear equation:

$$\frac{dy}{dx} + p(x)y = q(x).$$

Assume the solution has the form of the homogeneous solution, but replace the constant C with an undetermined function $C(x)$:

$$y(x) = C(x)e^{-\int p(x)\,dx}.$$

Substitute this form into the original equation to determine $C(x)$, and show that this yields the same general solution formula derived in eq. (1.20).

12. Periodic Solutions. Consider the linear equation with periodic coefficients:

$$\frac{dy}{dx} + p(x)y = q(x),$$

where $p(x)$ and $q(x)$ are continuous functions with period $\omega > 0$.

(a) Prove that if $q(x) \equiv 0$, then any non-zero solution has period ω if and only if the average value of $p(x)$ over one period is zero:

$$\bar{p} = \frac{1}{\omega}\int_0^\omega p(x)\,dx = 0.$$

(b) Prove that if $q(x)$ is not identically zero, then the equation has a unique ω-periodic solution if and only if $\bar{p} \neq 0$. Find the expression for this solution.

13. **Bounded Solutions.** Suppose $f(x)$ is a continuous function that is bounded on the interval $(-\infty, +\infty)$. Prove that the equation

$$y' + y = f(x)$$

possesses exactly one solution that is bounded on $(-\infty, +\infty)$. Find this bounded solution explicitly. Furthermore, prove that if $f(x)$ is also periodic with period ω, then this unique bounded solution is also periodic with period ω.

14. **Linear Operators on Function Spaces.** Let

$$H^0 = \{f(x) \mid f \text{ is a continuous function with period } 2\pi\}.$$

The set H^0 forms a linear space over the real numbers. For any $f \in H^0$, define the norm:

$$\|f\| = \max_{0 \le x \le 2\pi} |f(x)|.$$

It is a known result in analysis that H^0 is a complete space (a Banach space) under this norm. Using the solution formula derived for periodic solutions, we define a transformation φ mapping f to the solution y. Prove that φ is a bounded linear operator from H^0 to H^0. Specifically, prove:

(a) **Linearity:** For any constants C_1, C_2 and any $f_1, f_2 \in H^0$:

$$\varphi(C_1 f_1 + C_2 f_2) = C_1 \varphi(f_1) + C_2 \varphi(f_2).$$

(b) **Boundedness:** For any $f \in H^0$:

$$\|\varphi(f)\| \le k\|f\|,$$

where $k > 0$ is a constant independent of f.

2

Elementary Transformation Methods

Having established methods for exact, separable, and linear first-order equations, we now expand our capability to solve differential equations by introducing **elementary transformations**. Many equations that appear intractable in their original form can be reduced to one of the standard types discussed in the previous chapter through a judicious change of variables.

2.1 Canonical Substitutions

We begin by examining equations where the structure of the function suggests a natural substitution.

Example 2.1. Affine Argument. Consider the differential equation

$$\frac{dy}{dx} = f(x+y). \tag{2.1}$$

The composition depends solely on the sum $x + y$. This suggests the substitution $u = x + y$. Differentiating with respect to x yields

$$\frac{du}{dx} = 1 + \frac{dy}{dx}.$$

Substituting this into (2.1), we obtain

$$\frac{du}{dx} - 1 = f(u) \implies \frac{du}{dx} = 1 + f(u).$$

This is a separable equation for $u(x)$.

範例

Example 2.2. Non-linear Reductions. Consider the equation

$$\frac{dy}{dx} = \frac{xy^2 + \sin x}{2y}. \tag{2.2}$$

Rearranging terms to isolate the derivative of a power of y, we write

$$2y\frac{dy}{dx} = xy^2 + \sin x.$$

Recognising that $2yy' = \frac{d}{dx}(y^2)$, we introduce the transformation $v = y^2$. The equation becomes

$$\frac{dv}{dx} = xv + \sin x \implies \frac{dv}{dx} - xv = \sin x.$$

This is a first-order linear differential equation for v, which may be solved using the integrating factor method derived in eq. (1.20).

範例

Homogeneous Equations

A particularly important class of equations arises when the direction field is invariant under uniform scaling of the coordinates.

Definition 2.1. ***Homogeneous Function.***
A function $F(x, y)$ is said to be **homogeneous of degree** m if, for all $t > 0$,

$$F(tx, ty) = t^m F(x, y).$$

定義

Definition 2.2. ***Homogeneous Differential Equation.***
The differential equation

$$P(x, y)\, dx + Q(x, y)\, dy = 0 \tag{2.3}$$

is called a **homogeneous equation** if both coefficients $P(x, y)$ and $Q(x, y)$ are homogeneous functions of the same degree m.

定義

Note

This definition is distinct from the concept of a *homogeneous linear equation* ($y' + p(x)y = 0$). In the present context, "homogeneous" refers to the scaling symmetry of the variables x and y.

An equivalent characterisation is that the derivative can be expressed solely as a function of the ratio y/x. If P and Q are homogeneous of degree m, we may write

$$\frac{dy}{dx} = -\frac{P(x, y)}{Q(x, y)} = -\frac{x^m P(1, y/x)}{x^m Q(1, y/x)} = \phi\left(\frac{y}{x}\right).$$

Method of Solution

To solve a homogeneous equation, we introduce the scaling variable $u = y/x$.

Proposition 2.1. ***Reduction of Homogeneous Equations.***
The substitution $y = ux$ transforms a homogeneous differential equation into a separable equation.

命题

Proof

Let $y = ux$. Differentiating yields $dy = u\,dx + x\,du$. Substituting into (2.3) and utilising the homogeneity property $P(x, ux) = x^m P(1, u)$ and $Q(x, ux) = x^m Q(1, u)$:

$$P(x, ux)\,dx + Q(x, ux)(u\,dx + x\,du) = 0$$
$$x^m P(1, u)\,dx + x^m Q(1, u)(u\,dx + x\,du) = 0.$$

Assuming $x \neq 0$, we divide by x^m and group terms by differential:

$$[P(1, u) + uQ(1, u)]\,dx + xQ(1, u)\,du = 0.$$

Separating variables yields

$$\frac{dx}{x} + \frac{Q(1, u)}{P(1, u) + uQ(1, u)}\,du = 0. \tag{2.4}$$

■

Note

The line $x = 0$ may be a solution to the original equation but is excluded by the transformation. It must be checked separately.

Example 2.3. Logarithmic Spirals. Solve the differential equation

$$\frac{dy}{dx} = \frac{x + y}{x - y}.$$

The function on the right is homogeneous of degree 0. We set $y = ux$, so that $y' = u + xu'$. The equation becomes

$$u + x\frac{du}{dx} = \frac{x + ux}{x - ux} = \frac{1 + u}{1 - u}.$$

Rearranging to separate variables:

$$x\frac{du}{dx} = \frac{1 + u}{1 - u} - u = \frac{1 + u - u + u^2}{1 - u} = \frac{1 + u^2}{1 - u}.$$

Separating variables:

$$\frac{1 - u}{1 + u^2}\,du = \frac{dx}{x}.$$

We integrate the left side by splitting the numerator:

$$\int \frac{1}{1 + u^2}\,du - \frac{1}{2}\int \frac{2u}{1 + u^2}\,du = \arctan u - \frac{1}{2}\ln(1 + u^2).$$

Equating to the integral of the right side:

$$\arctan u - \ln \sqrt{1+u^2} = \ln|x| - \ln C \quad (C > 0).$$

Rearranging gives

$$\ln(Ce^{\arctan u}) = \ln\left(|x|\sqrt{1+u^2}\right).$$

Exponentiating and substituting $u = y/x$:

$$\sqrt{x^2+y^2} = Ce^{\arctan(y/x)}.$$

In polar coordinates ($x = r\cos\theta, y = r\sin\theta$), this simplifies to

$$r = Ce^{\theta}.$$

This solution represents a family of **logarithmic spirals** focusing on the origin O.

範例

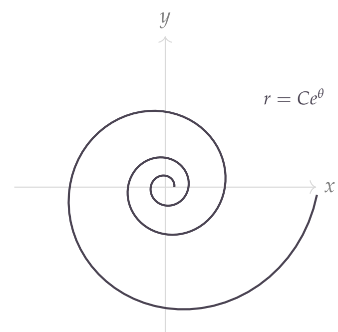

Figure 2.1: The logarithmic spiral, the integral curve of the homogeneous equation $y' = (x+y)/(x-y)$.

Linear Fractional Equations

A natural generalisation of the homogeneous equation involves ratios of linear functions with constant terms. Consider the equation

$$\frac{dy}{dx} = f\left(\frac{ax+by+c}{mx+ny+l}\right), \tag{2.5}$$

where a, b, c, m, n, l are constants. If $c = l = 0$, the argument is homogeneous of degree 0, and the method of the previous section applies. If not, the behaviour depends on the determinant of the coefficients.

Case I: Intersecting Lines ($\Delta \neq 0$)

Suppose the determinant $\Delta = an - bm \neq 0$. The lines

$$L_1 : ax+by+c = 0, \quad L_2 : mx+ny+l = 0$$

intersect at a unique point (α, β). We may satisfy the system

$$\begin{cases} a\alpha + b\beta + c = 0, \\ m\alpha + n\beta + l = 0. \end{cases} \tag{2.6}$$

We introduce a translation of coordinates to move the origin to (α, β):

$$x = \xi + \alpha, \quad y = \eta + \beta.$$

Since $dx = d\xi$ and $dy = d\eta$, substituting these into (2.5) eliminates the constant terms:

$$\frac{d\eta}{d\xi} = f\left(\frac{a(\xi+\alpha)+b(\eta+\beta)+c}{m(\xi+\alpha)+n(\eta+\beta)+l}\right)$$
$$= f\left(\frac{a\xi+b\eta+(a\alpha+b\beta+c)}{m\xi+n\eta+(m\alpha+n\beta+l)}\right)$$
$$= f\left(\frac{a\xi+b\eta}{m\xi+n\eta}\right).$$

This transformed equation is homogeneous in ξ and η. The substitution $\eta = u\xi$ renders it separable.

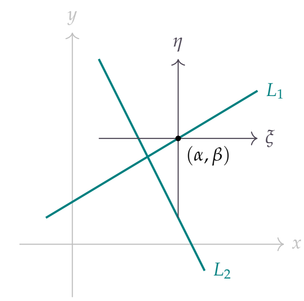

Figure 2.2: Translation of coordinates to the intersection of the lines L_1 and L_2 reduces the equation to homogeneous form.

Case II: Parallel Lines ($\Delta = 0$)

If $\Delta = an - bm = 0$, the lines are parallel. Consequently, the coefficients are proportional:

$$\frac{m}{a} = \frac{n}{b} = \lambda.$$

The denominator in (2.5) becomes

$$mx + ny + l = \lambda(ax+by) + l.$$

The differential equation simplifies to a function of a single linear combination:

$$\frac{dy}{dx} = f\left(\frac{ax+by+c}{\lambda(ax+by)+l}\right).$$

We employ the substitution $v = ax + by$. Differentiating with respect to x:

$$\frac{dv}{dx} = a + b\frac{dy}{dx}.$$

Substituting for y':

$$\frac{dv}{dx} = a + bf\left(\frac{v+c}{\lambda v + l}\right).$$

This is a separable equation for $v(x)$.

2.2 *Bernoulli Equations*

We now consider a class of non-linear equations that can be linearised through a simple substitution.

Definition 2.3. ***Bernoulli Equation.***
A first-order differential equation of the form

$$\frac{dy}{dx} + p(x)y = q(x)y^n, \qquad (2.7)$$

where $n \in \mathbb{R}$ is a constant, is called a **Bernoulli equation**.

定義

If $n = 0$ or $n = 1$, the equation is linear. We assume $n \neq 0, 1$. The non-linearity arises from the y^n term. To eliminate it, we divide (2.7) by y^n:

$$y^{-n}\frac{dy}{dx} + p(x)y^{1-n} = q(x).$$

Recognising that $\frac{d}{dx}(y^{1-n}) = (1-n)y^{-n}y'$, we introduce the substitution:

$$z = y^{1-n}. \tag{2.8}$$

Differentiating z with respect to x yields $z' = (1-n)y^{-n}y'$. Substituting this into the modified equation:

$$\frac{1}{1-n}\frac{dz}{dx} + p(x)z = q(x).$$

Multiplying by $(1-n)$ reduces the system to a standard linear form:

$$\frac{dz}{dx} + (1-n)p(x)z = (1-n)q(x). \tag{2.9}$$

This linear equation may be solved for $z(x)$ using an integrating factor, after which $y(x)$ is recovered via $y = z^{1/(1-n)}$.

Example 2.4. Logistic Growth. The logistic equation describing population growth under limited resources is a Bernoulli equation ($n = 2$):

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \implies \frac{dN}{dt} - rN = -\frac{r}{K}N^2.$$

Here $p(t) = -r$ and $q(t) = -r/K$. We substitute $z = N^{1-2} = N^{-1}$. Then:

$$\frac{dz}{dt} - (-1)(-r)z = (-1)\left(-\frac{r}{K}\right) \implies \frac{dz}{dt} + rz = \frac{r}{K}.$$

This linear equation is easily solved to yield the sigmoid growth curve.

範例

2.3 *Riccati Equations*

The Riccati equation represents the simplest non-linear differential equation that does not generally admit a solution by elementary quadratures (i.e., by a finite combination of elementary operations and indefinite integrals of known functions).

Definition 2.4. ***Riccati Equation.***
A **Riccati equation** is a first-order ordinary differential equation quadratic

in the unknown function:

$$\frac{dy}{dx} = p(x)y^2 + q(x)y + r(x), \tag{2.10}$$

where p, q, r are continuous functions on an interval I, and $p(x) \not\equiv 0$.

定義

Named after Jacopo Riccati (1676–1754), this equation plays a significant role in various fields, from the theory of Bessel functions to modern control theory (e.g., the matrix Riccati equation). Unlike the Bernoulli equation, there is no general method to solve (2.10) analytically. However, if partial information about the solution space is available, the equation can be reduced.

Theorem 2.1. ***Reduction of the Riccati Equation.***
If a particular solution $y = \varphi_1(x)$ of the Riccati equation (2.10) is known, the general solution can be obtained by elementary integration.

定理

Proof

We express the general solution as a perturbation of the known solution:

$$y(x) = u(x) + \varphi_1(x).$$

Substituting this into (2.10):

$$\frac{du}{dx} + \varphi_1'(x) = p(x)(u + \varphi_1)^2 + q(x)(u + \varphi_1) + r(x).$$

Expanding the quadratic term:

$$\frac{du}{dx} + \varphi_1' = p(x)(u^2 + 2u\varphi_1 + \varphi_1^2) + q(x)u + q(x)\varphi_1 + r(x).$$

Since φ_1 satisfies the original equation, $\varphi_1' = p\varphi_1^2 + q\varphi_1 + r$. Subtracting this identity from the equation leaves:

$$\frac{du}{dx} = p(x)u^2 + 2p(x)\varphi_1 u + q(x)u.$$

Rearranging terms yields a Bernoulli equation for u (with $n = 2$):

$$\frac{du}{dx} - [2p(x)\varphi_1(x) + q(x)]u = p(x)u^2.$$

This can be solved using the substitution $z = u^{-1}$, reducing the problem to a linear equation.

■

While the general Riccati equation is not integrable, specific forms allow for systematic reduction.

Theorem 2.2. ***Bernoulli's Solvability Condition (1725).***
Consider the special Riccati equation:

$$\frac{dy}{dx} + ay^2 = bx^m, \tag{2.11}$$

where $a, b \neq 0$ and m are constants. This equation can be transformed into a separable equation (and thus solved by elementary integration) if m takes any of the values:

$$m = 0, \quad m = -2, \quad \text{or} \quad m = \frac{-4k}{2k \pm 1} \quad (k \in \mathbb{Z}^+). \tag{2.12}$$

定理

Without loss of generality, we set $a = 1$ by scaling x if necessary. We consider $y' + y^2 = bx^m$.

Cases

Case $m = 0$***:*** The equation is separable: $y' = b - y^2$.

Case $m = -2$***:*** We introduce the transformation $z = xy$. Then $y = z/x$ and $y' = (z'x - z)/x^2$. Substituting into the equation:

$$\frac{z'x - z}{x^2} + \frac{z^2}{x^2} = \frac{b}{x^2} \implies xz' - z + z^2 = b.$$

Rearranging gives the separable equation

$$z' = \frac{b + z - z^2}{x}.$$

証明終

Recursive Reduction:

For

$$m = \frac{-4k}{2k+1},$$

we employ a sequence of transformations to reduce the index k.
First, apply the change of variables:

$$x = \xi^{\frac{1}{m+1}}, \quad y = \frac{b}{m+1}\eta^{-1}. \tag{2.13}$$

The differential transforms as:

$$\frac{dy}{dx} = \frac{dy/d\eta}{dx/d\xi}\frac{d\eta}{d\xi} = \frac{-\frac{b}{m+1}\eta^{-2}}{\frac{1}{m+1}\xi^{\frac{-m}{m+1}}}\frac{d\eta}{d\xi} = -b\eta^{-2}\xi^{\frac{m}{m+1}}\frac{d\eta}{d\xi}.$$

Substituting into $y' + y^2 = bx^m$:

$$-b\eta^{-2}\xi^{\frac{m}{m+1}}\frac{d\eta}{d\xi} + \frac{b^2}{(m+1)^2}\eta^{-2} = b\xi^{\frac{m}{m+1}}.$$

Multiplying by $-\frac{1}{b}\eta^2\xi^{-\frac{m}{m+1}}$ simplifies this to:

$$\frac{d\eta}{d\xi} - \frac{b}{(m+1)^2}\xi^{-\frac{m}{m+1}} = -\eta^2 \implies \frac{d\eta}{d\xi} + \eta^2 = \frac{b}{(m+1)^2}\xi^n, \quad (2.14)$$

where the new exponent is $n = -\frac{m}{m+1}$. If $m = \frac{-4k}{2k+1}$, a direct calculation shows:

$$n = -\frac{\frac{-4k}{2k+1}}{1+\frac{-4k}{2k+1}} = -\frac{-4k}{(2k+1)-4k} = \frac{4k}{2k-1} = \frac{-4k}{1-2k}.$$

This matches the form of the exponent in the second branch of (2.12), but with a sign change in the denominator. Next, we apply a second transformation:

$$\xi = \frac{1}{t}, \quad \eta = t - zt^2.$$

This transforms (2.14) into a new Riccati equation for $z(t)$:

$$\frac{dz}{dt} + z^2 = \frac{b}{(m+1)^2}t^l, \quad (2.15)$$

where the exponent l satisfies the recurrence relation

$$l(k) = \frac{-4(k-1)}{2(k-1)+1}.$$

Thus, one iteration of these transformations reduces the index from k to $k - 1$. Repeating this process k times reduces the exponent to $m = 0$, which is separable. The case $m = \frac{-4k}{2k-1}$ is solved similarly by entering the recursion at the intermediate stage (2.14).

証明終

Remark (Liouville's Theorem on Integrability).

The condition (2.12) was proven to be sufficient by Daniel Bernoulli in 1725. In 1841, Joseph Liouville proved that this condition is also **necessary**. That is, the Riccati equation $y' + y^2 = bx^m$ cannot be solved by elementary integration for any other values of m. This result demonstrated that even simple-looking equations (e.g., $y' = x^2 + y^2$) typically generate non-elementary functions. Consequently, the focus of differential equation theory shifted from finding explicit formulae to investigating existence, uniqueness, and qualitative properties (such as stability and periodicity).

2.4 *Exercises*

1. **Solving by Transformations.** Solve the following differential equations by identifying and applying the appropriate substitution (homogeneous, linear fractional, or Bernoulli).

   (a) $y' = \frac{2y-x}{2x-y}$

   (b) $y' = \frac{2y-x+5}{2x-y-4}$

   (c) $y' = \frac{x+2y+1}{2x+4y-1}$

   (d) $y' = x^3y^3 - xy$

2. **Advanced Substitutions.** Use appropriate transformations to solve the following equations.

   (a) $y' = \cos(x-y)$

   (b) $(3uv + v^2)\, du + (u^2 + uv)\, dv = 0$

   (c) $(x^2 + y^2 + 3)\frac{dy}{dx} = 2x\left(2y - \frac{x^2}{y}\right)$

   (d) $\frac{dy}{dx} = \frac{2x^3+3xy^2-7x}{3x^2y+2y^3-8y}$

3. **Riccati Equations.** Solve the following Riccati equations.

   (a) $y' = -\frac{y^2}{4x^2}$

   (b) $x^2y' = x^2y^2 + xy + 1$

4. **Reduction to Riccati Form.** Show that the general second-order linear homogeneous differential equation

   $$y'' + p(x)y' + q(x)y = 0$$

   can be transformed into a Riccati equation for a new variable $u(x)$. Explicitly state the transformation used.

5. **Geometric Properties.** Find the equation of a curve such that the angle between the tangent at any point on the curve and the radius vector to that point is always 45°. Express the result in polar coordinates.

In polar form $r = r(\theta)$, if ψ is the angle between the tangent and radius vector, then $\tan\psi = \frac{r}{dr/d\theta}$.

6. **The Parabolic Reflector.** The law of reflection says that the angle of incidence equals the angle of reflection, both measured with respect to the normal line at the point of reflection. Determine the shape of a reflector (assumed to be a surface of revolution) such that light rays emitted from a point source at the origin are reflected into a beam of parallel rays. Formulate the governing differential equation for the profile curve and solve it.

3
The Method of Integrating Factors

In the preceding chapters, we established that the general solution to an exact differential equation $P(x,y)\,dx + Q(x,y)\,dy = 0$ is found by direct integration. Furthermore, we observed that linear equations $y' + p(x)y = q(x)$ and separable equations can be rendered exact by multiplication with a suitable function. We now generalise this approach to arbitrary first-order equations.

3.1 Existence and Definition

Consider the general first-order differential equation in symmetric form:

$$P(x,y)\,dx + Q(x,y)\,dy = 0. \tag{3.1}$$

If the compatibility condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ is not satisfied, the equation is not exact. We seek a non-zero, differentiable function $\mu(x,y)$ such that the modified equation

$$\mu(x,y)P(x,y)\,dx + \mu(x,y)Q(x,y)\,dy = 0 \tag{3.2}$$

is exact. Such a function μ is termed an **integrating factor**.
The condition for (3.2) to be exact is

$$\frac{\partial}{\partial y}[\mu(x,y)P(x,y)] = \frac{\partial}{\partial x}[\mu(x,y)Q(x,y)].$$

Expanding these derivatives using the product rule yields a first-order partial differential equation for μ:

$$P\frac{\partial \mu}{\partial y} - Q\frac{\partial \mu}{\partial x} = \mu\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right). \tag{3.3}$$

While local solutions to (3.3) follow under standard regularity hypotheses (for instance via the method of characteristics, to be discussed later), finding a general solution is typically as difficult as solving the original ordinary differential equation. However, in specific instances, we may assume a simplified structure for μ (such as dependence on a single variable) to reduce (3.3) to a tractable ordinary differential equation.

Single-Variable Integrating Factors

We first investigate whether an integrating factor exists that depends solely on the variable x.

Theorem 3.1. ***Integrating Factor of One Variable.***
The differential equation (3.1) admits an integrating factor $\mu = \mu(x)$ depending only on x if and only if the expression

$$G(x) = \frac{1}{Q(x,y)} \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right) \tag{3.4}$$

is a function of x alone. In this case, the integrating factor is given by

$$\mu(x) = \exp\left(\int G(x)\, dx \right).$$

定理

Proof

Assume $\mu = \mu(x)$. Then $\frac{\partial \mu}{\partial y} = 0$ and $\frac{\partial \mu}{\partial x} = \frac{d\mu}{dx}$. Substituting these into the PDE (3.3):

$$-Q(x,y)\frac{d\mu}{dx} = \mu(x) \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right).$$

Rearranging terms separates the variables μ and the coefficients:

$$\frac{1}{\mu}\frac{d\mu}{dx} = \frac{1}{Q} \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right).$$

Since the left-hand side depends only on x, a solution exists if and only if the right-hand side is independent of y. Denoting this function by $G(x)$, we integrate to find $\ln|\mu| = \int G(x)\, dx$, yielding the stated formula.

■

By symmetry, we deduce the criterion for an integrating factor depending only on y.

Theorem 3.2. ***Integrating Factor in y.***
The equation admits an integrating factor $\mu = \mu(y)$ if and only if

$$H(y) = \frac{1}{P(x,y)} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \tag{3.5}$$

is independent of x. The integrating factor is then $\mu(y) = \exp\left(\int H(y)\, dy\right)$.

定理

Example 3.1. Polynomial Coefficients. Solve the equation

$$(3x^3 + y)\,dx + (2x^2y - x)\,dy = 0. \tag{3.6}$$

Let $P = 3x^3 + y$ and $Q = 2x^2y - x$. We check exactness by comparing the mixed partial derivatives:

$$\frac{\partial P}{\partial y} = 1, \quad \frac{\partial Q}{\partial x} = 4xy - 1.$$

The equation is not exact since $1 \neq 4xy - 1$. We test the condition for $\mu(x)$:

$$\frac{1}{Q}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = \frac{1-(4xy-1)}{2x^2y - x} = \frac{2-4xy}{-x(1-2xy)} = \frac{2(1-2xy)}{-x(1-2xy)} = -\frac{2}{x}.$$

This depends solely on x. Thus, an integrating factor is

$$\mu(x) = \exp\left(\int -\frac{2}{x}\,dx\right) = e^{-2\ln|x|} = \frac{1}{x^2}.$$

Multiplying (3.6) by x^{-2} (for $x \neq 0$):

$$\left(3x + \frac{y}{x^2}\right)\,dx + \left(2y - \frac{1}{x}\right)\,dy = 0.$$

This is now exact. By inspection or integration, the general integral is

$$\frac{3}{2}x^2 + y^2 - \frac{y}{x} = C.$$

We must also include the solution $x = 0$, which was lost during the division.

範例

The Method of Grouping

Often, the terms of a differential equation can be partitioned into groups, each of which admits an obvious integrating factor. We may exploit this structure by finding a multiplier that reconciles these individual factors.

Theorem 3.3. ***Generalised Integrating Factors.***
Let $\mu(x, y)$ be an integrating factor for the equation $P\,dx + Q\,dy = 0$, such that $\mu P\,dx + \mu Q\,dy = d\Phi(x, y)$. Then, for any differentiable function g, the product

$$\tilde{\mu} = \mu(x, y)g(\Phi(x, y))$$

is also an integrating factor.

定理

Proof

Multiplying the original equation by $\tilde{\mu}$ yields

$$g(\Phi)(\mu P\,dx + \mu Q\,dy) = g(\Phi)\,d\Phi = d\left(\int g(u)\,du\right)\Big|_{u=\Phi},$$

which is a total differential.

■

This theorem suggests a strategy: partition the equation into two parts,

$$(P_1\,dx + Q_1\,dy) + (P_2\,dx + Q_2\,dy) = 0.$$

Suppose μ_1 is an integrating factor for the first group (yielding potential Φ_1) and μ_2 for the second (yielding Φ_2). We seek functions g_1, g_2 such that the combined integrating factors match:

$$\mu = \mu_1 g_1(\Phi_1) = \mu_2 g_2(\Phi_2).$$

If such functions exist, μ serves as an integrating factor for the entire equation.

Example 3.2. Grouping Terms. Solve the differential equation

$$(x^3y - 2y^2)\,dx + x^4\,dy = 0.$$

We rearrange the terms to identify standard differentials:

$$(x^3y\,dx + x^4\,dy) - 2y^2\,dx = 0. \tag{3.7}$$

Group 1: $x^3y\,dx + x^4\,dy = x^3(y\,dx + x\,dy) = x^3\,d(xy)$. The factor x^{-3} reduces this to $d(xy)$. Thus, the general integrating factor for this group is

$$\mu_1 = \frac{1}{x^3}g_1(xy).$$

Group 2: $-2y^2\,dx$. The factor y^{-2} reduces this to $-2\,dx = d(-2x)$. The general integrating factor is

$$\mu_2 = \frac{1}{y^2}g_2(x).$$

We require $\mu_1 = \mu_2$, so

$$\frac{1}{x^3}g_1(xy) = \frac{1}{y^2}g_2(x).$$

To satisfy this, we choose $g_1(u) = u^{-2}$ and $g_2(x) = x^{-5}$. Then:

$$\mu = \frac{1}{x^3}\frac{1}{(xy)^2} = \frac{1}{x^5y^2}, \quad \text{and} \quad \frac{1}{y^2}\frac{1}{x^5} = \frac{1}{x^5y^2}.$$

Multiplying (3.7) by $\mu = x^{-5}y^{-2}$:

$$\frac{1}{(xy)^2}\,d(xy) - \frac{2}{x^5}\,dx = 0.$$

Integrating directly:

$$-\frac{1}{xy} - 2\left(\frac{x^{-4}}{-4}\right) = C \implies -\frac{1}{xy} + \frac{1}{2x^4} = C.$$

Rearranging for y yields the general solution:

$$y = \frac{2x^3}{1 - 2Cx^4}.$$

The special solutions $x = 0$ and $y = 0$, excluded by the integrating factor, satisfy the original equation.

範例

Integrating Factors for Homogeneous Equations

Recall that a homogeneous equation $P\,dx + Q\,dy = 0$ is invariant under scaling. While the substitution $y = ux$ is the standard method of solution, the integrating factor method offers a direct alternative.

Proposition 3.1. ***Homogeneous Integrating Factor.***
If $P(x, y)$ and $Q(x, y)$ are homogeneous functions of the same degree, then

$$\mu(x, y) = \frac{1}{xP(x, y) + yQ(x, y)} \tag{3.8}$$

is an integrating factor, provided $xP + yQ \not\equiv 0$.

命題

Example 3.3. The Logarithmic Spiral Revisited. Consider the homogeneous equation from the previous chapter:

$$(x + y)\,dx - (x - y)\,dy = 0.$$

Here $P = x + y$ and $Q = y - x$. We compute the denominator for the integrating factor:

$$xP + yQ = x(x + y) + y(y - x) = x^2 + xy + y^2 - xy = x^2 + y^2.$$

The integrating factor is $\mu = (x^2 + y^2)^{-1}$. Multiplying the equation by μ:

$$\frac{x\,dx + y\,dy}{x^2 + y^2} + \frac{y\,dx - x\,dy}{x^2 + y^2} = 0.$$

We recognise standard differentials:

$$\frac{x\,dx + y\,dy}{x^2 + y^2} = \frac{1}{2}d(\ln(x^2 + y^2)),$$

$$\frac{y\,dx - x\,dy}{x^2 + y^2} = -d\left(\arctan\frac{y}{x}\right).$$

Integrating yields:

$$\frac{1}{2}\ln(x^2+y^2)-\arctan\frac{y}{x}=C'.$$

Exponentiating recovers the polar form $r = Ce^{\theta}$, consistent with the result derived via substitution in *Figure 2.1*.

範例

3.2 Application Examples

We conclude this chapter by examining specific applications of first-order differential equations in geometry and the natural sciences. These examples illustrate how the theoretical framework developed thus far enables the modelling of continuous systems.

Isogonal and Orthogonal Trajectories

A classical geometric problem involves finding a family of curves that intersects a given family at a prescribed angle.
Let a one-parameter family of smooth curves in the xy-plane be given by the implicit equation:

$$\Phi(x,y,C)=0. \tag{3.9}$$

We seek a second family of curves, $\Psi(x,y,K) = 0$, such that every curve of the second family intersects every curve of the first family at a constant angle $\alpha \in (-\pi/2, \pi/2]$. Such a family is called a family of **isogonal trajectories**. If $\alpha = \pi/2$, they are termed **orthogonal trajectories**.
To determine the equation of the trajectories, we first find the differential equation satisfied by the original family. Differentiating (3.9) with respect to x and eliminating the parameter C yields an equation of the form:

$$\frac{dy}{dx}=f(x,y). \tag{3.10}$$

Here, $f(x,y)$ represents the slope $\tan\theta$ of the tangent to the curve $\Phi = 0$ at the point (x,y). Let ϕ be the inclination of the tangent to the required trajectory. The condition of intersection at angle α implies $\phi = \theta + \alpha$. Using the addition formula for tangents:

$$y'_{\text{traj}}=\tan\phi=\tan(\theta+\alpha)=\frac{\tan\theta+\tan\alpha}{1-\tan\theta\tan\alpha}.$$

Substituting $\tan\theta = f(x,y)$, the differential equation for the isogonal trajectories is:

$$\frac{dy}{dx}=\frac{f(x,y)+\tan\alpha}{1-f(x,y)\tan\alpha}. \tag{3.11}$$

Orthogonal Trajectories ($\alpha = \pi/2$). In the limit as $\alpha \to \pi/2$, the slope relationship becomes $y'_{\text{traj}} = -1/f(x,y)$. Thus, the orthogonal trajectories satisfy:

$$\frac{dy}{dx} = -\frac{1}{f(x,y)}. \tag{3.12}$$

Example 3.4. Electric Field Lines and Equipotentials. Consider the family of parabolas $y = Cx^2$, representing, for instance, a planar electric field. Differentiating with respect to x gives $y' = 2Cx$. Eliminating $C = y/x^2$ yields the differential equation of the field lines:

$$\frac{dy}{dx} = 2\left(\frac{y}{x^2}\right)x = \frac{2y}{x}.$$

By (3.12), the differential equation for the orthogonal trajectories is:

$$\frac{dy}{dx} = -\frac{x}{2y}.$$

Separating variables yields $2y\,dy = -x\,dx$, or $x\,dx + 2y\,dy = 0$. Integration gives the general solution:

$$\frac{1}{2}x^2 + y^2 = K' \implies x^2 + 2y^2 = K, \quad (K > 0).$$

The orthogonal trajectories are a family of ellipses centred at the origin (see figure 3.1). In electrostatics, if the parabolas represent field lines, these ellipses represent the equipotential curves.

範例

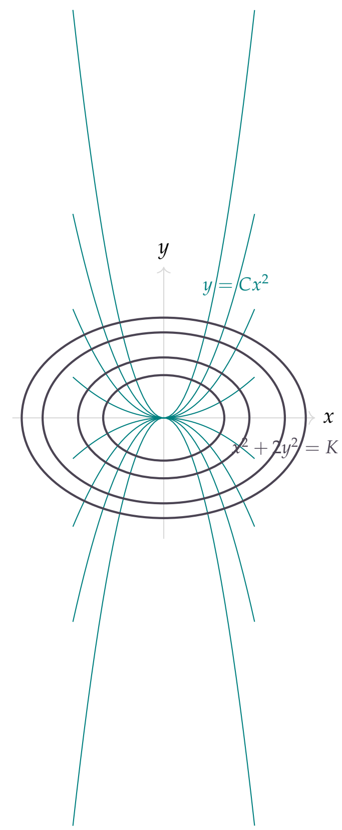

Figure 3.1: Orthogonal trajectories: parabolic field lines and elliptical equipotentials.

Population Dynamics

Differential equations provide a natural language for modelling the growth of populations, where the rate of change is dependent on the current state. Let $N(t)$ denote the population size at time t, assumed to be a continuously differentiable function. The relative growth rate is defined as:

$$r(N) = \frac{1}{N}\frac{dN}{dt}.$$

The Malthusian Model. The simplest assumption, proposed by Thomas Malthus (1798), is that the growth rate is a positive constant k. This yields the initial value problem:

$$\frac{dN}{dt} = kN, \quad N(t_0) = N_0.$$

The solution is exponential growth, $N(t) = N_0 e^{k(t-t_0)}$. While accurate for short intervals with unlimited resources, this model fails as $t \to \infty$ due to environmental constraints.

The Logistic Model. A more realistic model incorporates the limitations of resources. We assume the growth rate r decreases linearly with population size:

$$r(N) = a - bN,$$

where $a, b > 0$ are the **life coefficients**. This leads to the **logistic equation** (Verhulst, 1838):

$$\frac{dN}{dt} = N(a - bN). \tag{3.13}$$

This is a separable equation (and also a Bernoulli equation). Separating variables for $N \neq 0, a/b$:

$$\frac{dN}{N(a - bN)} = dt.$$

Using the partial fraction decomposition $\frac{1}{N(a-bN)} = \frac{1}{aN} + \frac{b}{a(a-bN)}$, we integrate:

$$\frac{1}{a}\left(\int \frac{dN}{N} + \int \frac{b\,dN}{a - bN}\right) = \int dt \implies \ln|N| - \ln|a - bN| = a(t - t_0) + C_1.$$

Exponentiating and solving for N (with initial condition $N(t_0) = N_0$) yields the logistic function:

$$N(t) = \frac{aN_0 e^{a(t-t_0)}}{a - bN_0 + bN_0 e^{a(t-t_0)}}. \tag{3.14}$$

As $t \to \infty$, $N(t) \to a/b$, which represents the carrying capacity of the environment.

The Lotka-Volterra Equations

We now consider a system of two coupled differential equations modelling the interaction between two species: a **prey** population $y(t)$ and a **predator** population $x(t)$.

> *Note*
>
> Following the convention of the source material for this section, we denote the predator by x and the prey by y. Standard modern texts often reverse this, so care must be taken with notation.

Prey (y): In the absence of predators, the prey grows Malthusianly at a rate $\mu > 0$. The presence of predators increases the death rate proportional to encounters, modelled by $-\delta x$.

Predator (x): In the absence of prey, the predator population decays at a rate $-\lambda < 0$. The consumption of prey contributes to growth proportional to encounters, modelled by $+\sigma y$.

This leads to the **Lotka-Volterra equations** (1925):

$$\begin{aligned} \frac{dx}{dt} &= x(-\lambda + \sigma y), \\ \frac{dy}{dt} &= y(\mu - \delta x), \end{aligned} \tag{3.15}$$

where $\lambda, \sigma, \mu, \delta$ are positive constants.

Phase Plane Analysis. To analyse the system without solving for t explicitly, we determine the relationship between x and y by eliminating dt:

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{y(\mu - \delta x)}{x(-\lambda + \sigma y)}.$$

This is a separable equation:

$$\frac{-\lambda + \sigma y}{y}\, dy = \frac{\mu - \delta x}{x}\, dx \implies \left(-\frac{\lambda}{y} + \sigma\right) dy = \left(\frac{\mu}{x} - \delta\right) dx.$$

Integration yields the conserved quantity (first integral):

$$H(x, y) \equiv \delta x - \mu \ln x + \sigma y - \lambda \ln y = C. \tag{3.16}$$

The function $H(x, y)$ is convex (i.e., its Hessian is positive definite) for $x, y > 0$, possessing a unique global minimum at the equilibrium point

$$(x_*, y_*) = \left(\frac{\mu}{\delta}, \frac{\lambda}{\sigma}\right).$$

Consequently, the level sets $H(x, y) = C$ form a family of closed curves surrounding the equilibrium (x_*, y_*). This implies that the populations $x(t)$ and $y(t)$ are periodic functions of time.
The dynamics exhibit a phase lag: an increase in prey supports an increase in predators, which subsequently causes the prey to collapse, leading to a predator collapse, allowing prey recovery.

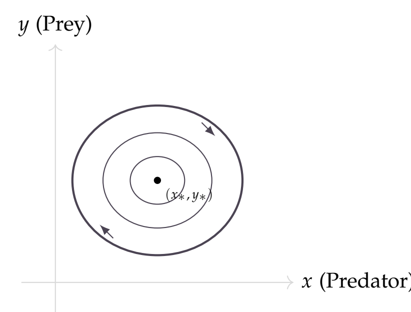

Figure 3.2: Phase portrait of the Lotka-Volterra system. Solutions are closed orbits, indicating periodic population cycles.

Volterra's Principle. While the amplitudes of oscillation depend on initial conditions, the *average* populations over one period T are independent of the orbit. Integrating $\frac{1}{y}\frac{dy}{dt} = \mu - \delta x$ over a period T:

$$\int_0^T \frac{d}{dt}(\ln y)\, dt = \int_0^T (\mu - \delta x)\, dt.$$

The left side vanishes because $y(T) = y(0)$. Thus $\mu T = \delta \int_0^T x\, dt$. The average predator population is:

$$\bar{x} = \frac{1}{T}\int_0^T x(t)\, dt = \frac{\mu}{\delta}.$$

Similarly, using the equation for x, the average prey population is $\bar{y} = \lambda/\sigma$.

Consider now the effect of constant rate harvesting (e.g., fishing) with coefficient $\varepsilon > 0$, removing both species proportional to their populations. The system becomes:

$$\frac{dx}{dt} = x(-(\lambda+\varepsilon)+\sigma y), \quad \frac{dy}{dt} = y((\mu-\varepsilon)-\delta x).$$

Provided $\mu > \varepsilon$, the form of the equations is unchanged, but the constants shift. The new averages are:

$$\bar{x}_{\text{new}} = \frac{\mu-\varepsilon}{\delta} < \bar{x}, \quad \bar{y}_{\text{new}} = \frac{\lambda+\varepsilon}{\sigma} > \bar{y}.$$

Conclusion: Moderate harvesting reduces the average predator population and *increases* the average prey population. This counter-intuitive result, known as **Volterra's Principle**, explained the observations of D'Ancona (1926) regarding the surge in predatory fish in the Adriatic Sea during World War I, when fishing activity (ε) was curtailed.

3.3 *Exercises*

1. **General Solutions.** Solve the following differential equations by finding an appropriate integrating factor.
   (a) $(3x^2y+2xy+y^3)\,dx+(x^2+y^2)\,dy=0$
   (b) $y\,dx+(2xy-e^{-2y})\,dy=0$
   (c) $\left(3x+\frac{6}{y}\right)dx+\left(\frac{x^2}{y}+\frac{3y}{x}\right)dy=0$
   (d) $y\,dx-(x^2+y^2+x)\,dy=0$
   (e) $2xy^3\,dx+(x^2y^2-1)\,dy=0$
   (f) $y(1+xy)\,dx-x\,dy=0$
   (g) $y^3\,dx+2(x^2-xy^2)\,dy=0$
   (h) $e^x\,dx+(e^x\cos y+2y\cos y)\,dy=0$
2. **Criteria for Specific Integrating Factors.**
   (a) Prove that a necessary and sufficient condition for the differential equation

   $$P(x,y)\,dx+Q(x,y)\,dy=0$$

   to admit an integrating factor of the form $\mu=\mu(\varphi(x,y))$ is that the expression

   $$\frac{\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}}{Q\frac{\partial \varphi}{\partial x}-P\frac{\partial \varphi}{\partial y}}$$

   depends only on $\varphi(x,y)$. Explicitly state the formula for such an integrating factor.

(b) Apply this result to derive the necessary and sufficient conditions for the existence of an integrating factor of the following forms:

(1). $\mu = \mu(x \pm y)$
(2). $\mu = \mu(x^2 + y^2)$
(3). $\mu = \mu(xy)$
(4). $\mu = \mu\left(\frac{y}{x}\right)$
(5). $\mu = \mu(x^a y^b)$

3. **Uniqueness of the Integrating Factor Structure.** Let $\mu(x, y)$ be an integrating factor for $P\,dx + Q\,dy = 0$, yielding the potential $\Phi(x, y)$. The forward direction is already established in *theorem 3.3*. Prove the converse: if μ_1 is any other integrating factor, then μ_1 must be expressible in the form $\mu_1 = \mu g(\Phi)$ for some function g.

4. **General Integral from Ratio of Factors.** Suppose the functions P, Q, μ_1, and μ_2 are continuously differentiable, and that μ_1 and μ_2 are two integrating factors of the differential equation $P\,dx + Q\,dy = 0$. Prove that if the ratio μ_1/μ_2 is not constant, then
$$\frac{\mu_1(x, y)}{\mu_2(x, y)} = C$$
is a general integral of the equation.

5. **Orthogonal Trajectories.** Find the equation of the family of curves that intersects each of the following families orthogonally (at an angle of $\pi/2$).

(a) $x^2 + y^2 = Cx$ (Family of circles tangent to the y-axis at the origin)
(b) $xy = C$ (Rectangular hyperbolas)
(c) $y^2 = Cx^3$ (Semicubic parabolas)
(d) $x^2 + C^2y^2 = 1$ (Family of ellipses with variable semi-minor axis)

6. **Isogonal Trajectories.** Find the family of curves that intersects each of the following families at a constant angle of $\pi/4$.

(a) $x - 2y = C$ (Family of parallel lines)
(b) $xy = C$ (Rectangular hyperbolas)
(c) $y = x \ln(Cx)$
(d) $y^2 = 4Cx$ (Family of parabolas)

7. **A Specific Trajectory.** Consider the family of hyperbolas given by $x^2 - y^2 = C$. A particle P moves in the xy-plane such that its trajectory intersects every hyperbola of the family at a constant angle of $\pi/6$. If the particle starts at the point $P_0(0, 1)$, determine the equation of its path.

8. **Population Models.** Consider a population $N(t)$ governed by the Gompertz equation:

$$\frac{dN}{dt} = -rN \ln\left(\frac{N}{K}\right),$$

where $r > 0$ is the growth rate and $K > 0$ is the carrying capacity.

   (a) Solve the differential equation subject to the initial condition $N(0) = N_0$.
   (b) Show that $\lim_{t\to\infty} N(t) = K$.
   (c) Determine the time t at which the growth rate dN/dt is maximised.

9. **Predator-Prey Dynamics.** In the Lotka-Volterra system (3.15), consider the specific case where the prey has unlimited resources (logistic term is absent) but the predator is harvested at a rate proportional to its population. The equations are:

$$\frac{dx}{dt} = -x + xy - hx, \quad \frac{dy}{dt} = y - xy.$$

   (a) Find the non-trivial equilibrium point (x_*, y_*) in terms of the harvest rate h.

   (b) How does increasing the harvest rate h affect the equilibrium population of the prey? Interpret this result biologically.

10. **Pursuit Curves.** A rabbit runs along the y-axis with constant speed v. A dog chases the rabbit with constant speed $w > v$. At time $t = 0$, the rabbit is at the origin and the dog is at the point $(L, 0)$. The dog always runs directly towards the rabbit.

   (a) Let (x, y) be the position of the dog. Show that the tangent to the dog's path passes through the rabbit's position $(0, vt)$, and derive the differential equation:

$$x\frac{d^2y}{dx^2} = \frac{v}{w}\sqrt{1 + \left(\frac{dy}{dx}\right)^2}.$$

   (b) Solve this second-order equation to find the path $y(x)$ of the dog.

   Let $p = dy/dx$

   (c) Determine the distance the dog runs before catching the rabbit.

4
Existence and Uniqueness Theory

The fundamental problem of ordinary differential equations concerns the conditions under which the initial value problem admits a solution and whether that solution is unique. In this chapter, we establish the classical existence and uniqueness theorem using the method of successive approximations, attributed to Émile Picard. We also explore weaker conditions for uniqueness and the limitations of the iterative method.

4.1 *The Picard-Lindelöf Theorem*

We consider the initial value problem

$$\frac{dy}{dx} = f(x,y), \quad y(x_0) = y_0. \tag{4.1}$$

To ensure the convergence of iterative approximations, we require a regularity condition on f slightly stronger than continuity.

> **Definition 4.1.** ***Lipschitz Condition.***
> Let $f(x,y)$ be defined on a region $D \subseteq \mathbb{R}^2$. The function f is said to satisfy a **Lipschitz condition** with respect to y in D if there exists a constant $L > 0$ such that
>
> $$|f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2| \tag{4.2}$$
>
> for all $(x,y_1), (x,y_2) \in D$. The constant L is called the Lipschitz constant.
>
> 定義

> **Proposition 4.1.** ***Differentiability implies Lipschitz.***
> If $f(x,y)$ admits a continuous partial derivative $\frac{\partial f}{\partial y}$ on a bounded, closed, convex region D (that is, for any two points in D, the entire line segment joining them lies in D), then f satisfies the Lipschitz condition on D.
>
> 命題

Proof

Since D is closed and bounded and $\frac{\partial f}{\partial y}$ is continuous, there exists $K > 0$ such that $\left|\frac{\partial f}{\partial y}\right| \leq K$ on D. For any fixed x, applying the Mean Value Theorem to f as a function of y yields

$$f(x, y_1) - f(x, y_2) = \frac{\partial f}{\partial y}(x, \xi)(y_1 - y_2),$$

where ξ lies between y_1 and y_2. Taking absolute values, we obtain the Lipschitz condition with $L = K$.

■

Note

The converse is false. For example, $f(x, y) = |y|$ satisfies the Lipschitz condition with $L = 1$ but is not differentiable at $y = 0$.

Picard's Existence and Uniqueness Theorem

We now state and prove the central result of this chapter.

Theorem 4.1. ***Picard's Existence and Uniqueness Theorem.*** Consider the initial value problem (4.1). Let $f(x, y)$ be continuous on the rectangular region

$$R = \{(x, y) : |x - x_0| \leq a, \ |y - y_0| \leq b\}$$

and satisfy the Lipschitz condition with respect to y on R. Let $M > \max_{(x,y) \in R} |f(x, y)|$. Then the problem (4.1) possesses a unique solution $y(x)$ defined on the interval

$$I = [x_0 - h, x_0 + h], \quad \text{where } h = \min\left\{a, \frac{b}{M}\right\}.$$

定理

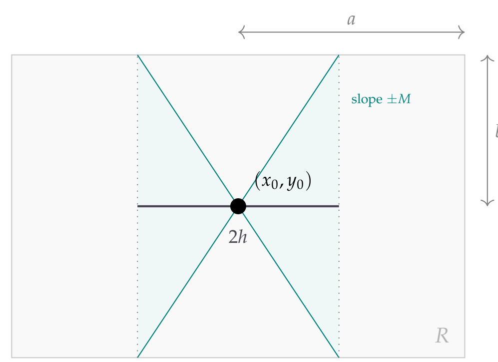

Figure 4.1: The solution is confined within the region bounded by lines of slope $\pm M$. The interval width h is determined by where these lines intersect the boundary of R.

Proof

We proceed in four steps: establishing an equivalent integral equation, constructing the Picard iterates, proving uniform convergence, and demonstrating uniqueness.

Equivalence to an Integral Equation. A continuous function $y(x)$ on I is a solution to (4.1) if and only if it satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt. \tag{4.3}$$

If $y(x)$ solves the IVP, integration of $y'(t) = f(t, y(t))$ from x_0 to x yields (4.3). Conversely, if $y(x)$ is a continuous solution of (4.3), the right-hand side is continuously differentiable. Differentiating with respect to x recovers the differential equation, and setting $x = x_0$ yields $y(x_0) = y_0$.

Construction of the Picard Sequence. We define the sequence of functions $\{y_n(x)\}$ iteratively:

$$y_0(x) = y_0, \quad y_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, y_n(t))\, dt, \quad n \geq 0. \tag{4.4}$$

We must first show that this sequence is well-defined, i.e., that the graph of each $y_n(x)$ remains within the rectangle R for $x \in I$.

- *Base case:* $(x, y_0) \in R$ for all $x \in I$.
- *Inductive step:* Assume $|y_n(x) - y_0| \leq b$ for $x \in I$. Then $f(t, y_n(t))$ is defined. From (4.4):

  $$|y_{n+1}(x) - y_0| = \left| \int_{x_0}^{x} f(t, y_n(t))\, dt \right| \leq \left| \int_{x_0}^{x} M\, dt \right| = M|x - x_0| \leq Mh.$$

  Since $h \leq b/M$, we have $|y_{n+1}(x) - y_0| \leq b$.

Thus, by induction, $y_n(x)$ is defined and continuous on I for all n.

Uniform Convergence. We prove that $\{y_n(x)\}$ converges uniformly on I by examining the series

$$y_0(x) + \sum_{n=0}^{\infty} [y_{n+1}(x) - y_n(x)]. \tag{4.5}$$

We assert the bound

$$|y_{n+1}(x) - y_n(x)| \leq \frac{ML^n |x - x_0|^{n+1}}{(n+1)!}. \tag{4.6}$$

For $n = 0$, $|y_1 - y_0| \leq M|x - x_0|$, which matches the formula. Assume the bound holds for k. Then

$$\begin{aligned}
|y_{k+2}(x) - y_{k+1}(x)| &= \left| \int_{x_0}^{x} [f(t, y_{k+1}(t)) - f(t, y_k(t))]\, dt \right| \\
&\leq \left| \int_{x_0}^{x} L|y_{k+1}(t) - y_k(t)|\, dt \right| \\
&\leq L \left| \int_{x_0}^{x} \frac{ML^k |t - x_0|^{k+1}}{(k+1)!}\, dt \right| \\
&= \frac{ML^{k+1} |x - x_0|^{k+2}}{(k+2)!}.
\end{aligned}$$

This bound is uniform on I. Let $a_n = \frac{ML^n h^{n+1}}{(n+1)!}$. For any $m > n$,

$$|y_m(x) - y_n(x)| \leq \sum_{k=n}^{m-1} |y_{k+1}(x) - y_k(x)| \leq \sum_{k=n}^{\infty} a_k.$$

The series $\sum a_k$ converges (it is a tail of an exponential-type series), so the sequence $\{y_n\}$ is uniformly Cauchy. By the completeness of continuous functions, it converges uniformly to a continuous limit $\varphi(x)$. Passing the limit $n \to \infty$ in (4.4), the uniform convergence allows us to exchange limit and integral:

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.$$

Thus $\varphi(x)$ is a solution.

Uniqueness. Suppose $u(x)$ and $v(x)$ are two solutions on a subinterval $J = [x_0 - d, x_0 + d] \subseteq I$. Then

$$u(x) - v(x) = \int_{x_0}^{x} [f(t, u(t)) - f(t, v(t))]\, dt.$$

Using the Lipschitz condition,

$$|u(x) - v(x)| \leq L \left| \int_{x_0}^{x} |u(t) - v(t)|\, dt \right|.$$

Since $|u - v|$ is bounded on J by some constant K, iterating this inequality yields

$$|u(x) - v(x)| \leq K \frac{(L|x - x_0|)^n}{n!}$$

for any $n \geq 1$. As $n \to \infty$, the right-hand side vanishes. Hence $u(x) \equiv v(x)$ on J.

■

Example 4.1. Riccati Equation. Consider

$$\frac{dy}{dx} = x^2 + y^2.$$

範例

Solution

The function $f(x, y) = x^2 + y^2$ is continuously differentiable everywhere. By *theorem* 4.1, a unique solution exists through any point (x_0, y_0) in the plane, at least locally. Note that elementary integration cannot solve this equation.

■

The Lipschitz condition is sufficient for uniqueness, but not necessary for existence.

Note

If $f(x, y)$ is merely continuous but does not satisfy a Lipschitz condition, the **Peano Existence Theorem** guarantees at least one solution, though it may not be unique.

We introduce a weaker condition due to Osgood that still preserves uniqueness.

Definition 4.2. ***Osgood Condition.***
A continuous function $f(x, y)$ satisfies the **Osgood condition** with respect to y if

$$|f(x, y_1) - f(x, y_2)| \leq F(|y_1 - y_2|),$$

where $F(r) > 0$ is continuous for $r > 0$ and satisfies the integral condition

$$\int_0^{r_1} \frac{dr}{F(r)} = +\infty \tag{4.7}$$

for some constant $r_1 > 0$.

定義

Note
The Lipschitz condition corresponds to $F(r) = Lr$, for which $\int \frac{dr}{Lr}$ diverges at 0. Thus, Osgood generalizes Lipschitz.

Theorem 4.2. ***Osgood Uniqueness Theorem.***
If $f(x, y)$ satisfies the Osgood condition in a region G, then through any point in G the initial value problem for $y' = f(x, y)$ admits at most one solution (i.e., uniqueness holds whenever a solution exists).

定理

Proof
Suppose, for the sake of contradiction, that two distinct solutions $y_1(x)$ and $y_2(x)$ pass through (x_0, y_0). Assume there exists $x_1 > x_0$ such that $y_1(x_1) > y_2(x_1)$. Let

$$\bar{x} = \sup\{x \in [x_0, x_1] : y_1(x) = y_2(x)\}.$$

Clearly $x_0 \leq \bar{x} < x_1$, and for $x \in (\bar{x}, x_1]$, the difference $r(x) = y_1(x) - y_2(x)$ is positive. Differentiating $r(x)$:

$$r'(x) = y_1'(x) - y_2'(x) = f(x, y_1) - f(x, y_2) \leq |f(x, y_1) - f(x, y_2)|.$$

Applying the Osgood condition:

$$r'(x) \leq F(r(x)).$$

Separating variables and integrating from $\bar{x}$ to x_1:

$$\int_{\bar{x}}^{x_1} \frac{r'(x)}{F(r(x))}\, dx \leq \int_{\bar{x}}^{x_1} dx \implies \int_0^{r(x_1)} \frac{dr}{F(r)} \leq x_1 - \bar{x}.$$

The right-hand side is finite, but by hypothesis (4.7), the integral on the left diverges to $+\infty$. This contradiction implies that no such splitting point $\bar{x}$ exists; hence $y_1 \equiv y_2$.

∎

Finally, we demonstrate that without the Lipschitz condition, the Picard iteration may fail to converge entirely, even if the solution is unique.

Example 4.2. Müller's Counterexample. Consider the IVP $y' = F(x, y)$ with $y(0) = 0$, where F is defined on $[0, 1] \times \mathbb{R}$ by:

$$F(x,y) = \begin{cases} 0 & x = 0, \\ 2x & 0 < x \leq 1,\, y < 0, \\ 2x - \frac{4y}{x} & 0 < x \leq 1,\, 0 \leq y \leq x^2, \\ -2x & 0 < x \leq 1,\, y > x^2. \end{cases}$$

This function is continuous but fails the Lipschitz condition near $x = 0$. The Picard iterates starting from $y_0(x) = 0$ are:

$$y_1(x) = \int_0^x F(t, 0)\, dt = \int_0^x 2t\, dt = x^2.$$
$$y_2(x) = \int_0^x F(t, t^2)\, dt = \int_0^x (2t - 4t)\, dt = -x^2.$$
$$y_3(x) = \int_0^x F(t, -t^2)\, dt = \int_0^x 2t\, dt = x^2.$$

By induction, $y_n(x) = (-1)^{n+1} x^2$. This sequence oscillates and does not converge. However, the actual unique solution is $y(x) = \frac{1}{3}x^2$.

範例

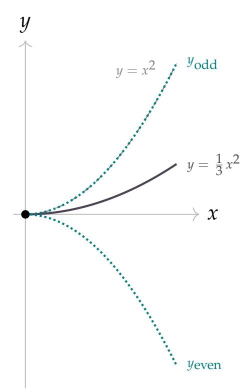

Figure 4.2: The failure of Picard iteration in Müller's example. The approximations oscillate between x^2 and $-x^2$, failing to converge to the solution $x^2/3$.

4.2 *The Peano Existence Theorem*

While the Picard-Lindelöf theorem guarantees both existence and uniqueness, its reliance on the Lipschitz condition is restrictive. If we assume only that $f(x, y)$ is continuous, the method of successive approximations may fail to converge (as seen in *figure* 4.2). To establish existence under continuity alone, we employ a constructive approach dating back to Euler: the method of polygonal approximations. This leads to the **Peano Existence Theorem**, which asserts existence but not uniqueness.

Consider the initial value problem

$$(E): \quad y' = f(x, y), \quad y(x_0) = y_0,$$

where f is continuous on the rectangle $R = \{(x, y) : |x - x_0| \leq a, |y - y_0| \leq b\}$. Let $M = \max_{(x,y)\in R} |f(x, y)|$. If $M = 0$, then $f \equiv 0$

on R, so $y(x) \equiv y_0$ is a solution on $[x_0 - a, x_0 + a]$. Hence the only nontrivial case is $M > 0$, for which we set

$$h = \min\{a, b/M\},$$

and restrict attention to $I = [x_0 - h, x_0 + h]$.
We construct an approximate solution by discretising the interval. For a positive integer n, let $h_n = h/n$. We define the nodes $x_k = x_0 + kh_n$ for $k = 0, \pm 1, \ldots, \pm n$. The values y_k approximating $y(x_k)$ are generated recursively. For the right-hand interval $[x_0, x_0 + h]$:

$$y_0 = y(x_0), \quad y_{k+1} = y_k + h_n f(x_k, y_k), \quad k = 0, \ldots, n-1. \tag{4.8}$$

We connect these points linearly to form the **Euler polygon** $\varphi_n(x)$. Explicitly, for $x \in [x_k, x_{k+1}]$:

$$\varphi_n(x) = y_k + f(x_k, y_k)(x - x_k). \tag{4.9}$$

The construction for the left-hand interval is symmetric.
Geometrically, $\varphi_n(x)$ is a polygonal curve starting at (x_0, y_0) whose slope on each sub-interval matches the direction field at the beginning of that sub-interval. Since $|f(x, y)| \leq M$, the slope of each segment lies in $[-M, M]$. Consequently, the graph of $\varphi_n(x)$ is confined within the angular region

$$\Delta_h = \{(x, y) : |x - x_0| \leq h, \; |y - y_0| \leq M|x - x_0|\} \subseteq R.$$

This ensures that the recursive step (4.8) is always well-defined.
To show that the sequence of functions $\{\varphi_n(x)\}$ converges to a solution, we require a criterion for compactness in the space of continuous functions $C(I)$.

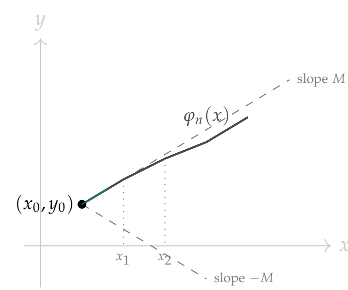

Figure 4.3: Construction of an Euler polygon. The slope of the segment on $[x_k, x_{k+1}]$ is determined by the field at (x_k, y_k).

> **Definition 4.3.** ***Equicontinuity.***
> A family of functions $\mathcal{F}$ defined on an interval I is **equicontinuous** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $f \in \mathcal{F}$ and all $x_1, x_2 \in I$:
> $$|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \varepsilon.$$
> 定義

Essentially, δ depends only on ε, not on the choice of function $f \in \mathcal{F}$.

> **Theorem 4.3.** ***The Arzelà-Ascoli Theorem.***
> Let $\{f_n\}$ be a sequence of functions on a closed bounded interval I. If the sequence is:
> 1. **Uniformly bounded:** There exists K such that $|f_n(x)| \leq K$ for all n, x.
> 2. **Equicontinuous:** As defined above.
>
> Then $\{f_n\}$ contains a subsequence $\{f_{n_k}\}$ that converges uniformly on I to a continuous function.

定理

We now apply these tools to prove the main existence result.

Theorem 4.4. ***Peano Existence Theorem.***
Let $f(x,y)$ be continuous on the region R. Then the initial value problem (E) admits at least one solution $y(x)$ defined on the interval $I = [x_0 - h, x_0 + h]$.

定理

Proof

We prove this in two steps.

Compactness of the Euler Sequence. Consider the sequence of Euler polygons $\{\varphi_n(x)\}$.

- **Uniform Boundedness:** By construction, the graph of each φ_n lies within $\Delta_h \subseteq R$. Thus $|\varphi_n(x) - y_0| \leq b$ for all n and all $x \in I$.
- **Equicontinuity:** For any $x_1, x_2 \in I$, the difference quotient is bounded by the maximum slope:

$$|\varphi_n(x_1) - \varphi_n(x_2)| \leq \sup_{\xi} |\varphi_n'(\xi)||x_1 - x_2| \leq M|x_1 - x_2|.$$

  Since the bound M on $|f|$ is independent of n, the family is equicontinuous.

By the *The Arzelà-Ascoli Theorem*, there exists a subsequence $\{\varphi_{n_k}\}$ converging uniformly on I to a continuous function $y(x)$.

The Integral Equation. We must verify that the limit $y(x)$ satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt.$$

We rewrite the definition of the Euler polygon φ_n in integral form with an error term. For $x \in [x_0, x_0 + h]$, let $x(t) = x_k$ whenever $t \in [x_k, x_{k+1})$. Then the derivative is $\varphi_n'(t) = f(x(t), \varphi_n(x(t)))$ almost everywhere (specifically, everywhere except at the nodes x_k). Integrating yields:

$$\varphi_n(x) = y_0 + \int_{x_0}^{x} f(x(t), \varphi_n(x(t)))\, dt. \tag{4.10}$$

Let $\Delta_n(t) = f(x(t), \varphi_n(x(t))) - f(t, \varphi_n(t))$. The equation becomes

$$\varphi_n(x) = y_0 + \int_{x_0}^{x} f(t, \varphi_n(t))\, dt + \int_{x_0}^{x} \Delta_n(t)\, dt.$$

As $n \to \infty$, the step size $h_n \to 0$, so $|x(t) - t| \to 0$. Since φ_n is Lipschitz with constant M, $|\varphi_n(x(t)) - \varphi_n(t)| \leq Mh_n \to 0$.

Because f is continuous on a compact set R, it is uniformly continuous. Therefore, $\Delta_n(t) \to 0$ uniformly. Passing to the limit $n_k \to \infty$ in (4.10):

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt.$$

Thus $y(x)$ is a solution to the initial value problem.

■

Note

Unlike the Picard sequence, the Euler sequence (or a subsequence thereof) converges to a solution even without the Lipschitz condition. However, if the solution is not unique, different subsequences may converge to different solutions.

*Remark (**The Lavrentieff Phenomenon.**).*

While Peano's theorem guarantees existence, the structure of the solution set can be complex. In 1925, Lavrentieff constructed a continuous function $f(x, y)$ such that for every point in a region, there exist at least two distinct integral curves passing through it. Conversely, if f is not continuous (or not even defined) at the initial point, a classical solution may fail to exist. For example, the IVP $y' = \frac{1}{y}$, $y(0) = 0$ has no classical solution because the right-hand side is undefined at $y = 0$. For $y' = \text{sgn}(y)$ (where $\text{sgn}(y) = 1$ if $y > 0$, -1 if $y < 0$, and 0 if $y = 0$), one must specify the convention at 0: with this choice, the solution $y \equiv 0$ is classical.

4.3 *Exercises*

1. **Uniqueness Analysis.** Using the properties of the right-hand side function (Lipschitz or Osgood conditions), discuss the uniqueness of the solution for the following initial value problems with $y(0) = 0$:
   (a) $\frac{dy}{dx} = |y|^a$, where $a > 0$ is a constant.
   (b) $\frac{dy}{dx} = \begin{cases} \dfrac{y}{1 + |\ln|y||}, & y \neq 0, \\ 0, & y = 0. \end{cases}$

2. **Picard Iteration.** Construct the Picard sequence $\{y_n(x)\}$ for the initial value problem:

$$\frac{dy}{dx} = x + y + 1, \quad y(0) = 0.$$

   Compute the limit of this sequence to find the exact solution.

3. ⋆ **One-Sided Uniqueness.** Suppose the continuous function $f(x, y)$ is decreasing with respect to y. Prove that the solution of the initial value problem $y' = f(x, y), y(x_0) = y_0$ is unique for $x > x_0$. Does uniqueness hold for $x < x_0$? Provide a proof or a counterexample.

4. **Compactness Criteria.**
   (a) Let $\{\varphi_n\}$ be the Euler polygons for the IVP $y' = f(x, y),\ y(x_0) = y_0$ on a finite interval I, with f continuous on a compact rectangle R. Prove: (i) $\{\varphi_n\}$ is uniformly bounded and equicontinuous on I; (ii) every uniformly convergent subsequence has a limit y satisfying the integral equation

   $$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,dt;$$

   (iii) conclude that the IVP has at least one solution on I.

   (b) Provide a counterexample to show that this conclusion generally fails if I is an infinite interval (e.g., $[0, \infty)$).

5. **Picard vs. Peano.** We established that the Picard sequence satisfies the conditions of the Arzelà-Ascoli theorem (uniform boundedness and equicontinuity). Can the Picard sequence alone be used to prove Peano's Existence Theorem (i.e., existence without the Lipschitz condition)? Explain your reasoning.

6. **Tonelli Approximations.** Consider the integral equation equivalent to the initial value problem:

   $$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,dt.$$

   We construct a sequence of approximate solutions $y_n(x)$ on $I = [x_0, x_0 + h]$ known as the **Tonelli sequence**. Let $d_n = h/n$ and define the function recursively on sub-intervals $[x_0 + kd_n, x_0 + (k + 1)d_n]$:

   $$y_n(x) = \begin{cases} y_0, & x_0 \le x \le x_0 + d_n, \\ y_0 + \int_{x_0}^{x-d_n} f(t, y_n(t))\,dt, & x_0 + d_n < x \le x_0 + h. \end{cases}$$

   Use the Arzelà-Ascoli theorem to prove that a subsequence of these functions converges to a solution of the integral equation, thereby providing an alternative proof of the Peano Existence Theorem.

7. ⋆ **Non-Convergence of Euler Polygons.** This exercise demonstrates that while a subsequence of Euler polygons converges to *a* solution (by Peano's theorem), the full sequence need not con-

verge if the solution is non-unique. Let

$$\alpha(x) = \int_0^x e^{-1/t^2}\, dt$$

for $x > 0$, with $\alpha(0) = 0$. Define a continuous function $f^*(x,y)$ on the strip $G : 0 \leq x \leq 1$, $y \in \mathbb{R}$ such that $f^*(0,0) = 0$ and:

$$f^*(x,y) = \begin{cases} x, & y > \alpha(x), \\ x\cos(\pi/x), & y = 0, \\ -x, & y < -\alpha(x). \end{cases}$$

For $0 < |y| < \alpha(x)$, assume $f^*(x,y)$ is defined by linear interpolation in y to ensure continuity on the strip. Consider the problem $y' = f^*(x,y)$, $y(0) = 0$. Let $\varphi_n^*(x)$ be the Euler polygon with step size $1/n$. Show that for $x \in [1/n, 1]$:

(a) If n is even, $\varphi_n^*(x) \geq \alpha(x)$.

(b) If n is odd, $\varphi_n^*(x) \leq -\alpha(x)$.

Conclude that the sequence $\{\varphi_n^*(x)\}$ does not converge as $n \to \infty$.

5
Extension of Solutions

The existence theorems established in the preceding chapter — specifically those of Picard and Peano — are local in nature. They guarantee a solution $y(x)$ only within a sufficiently small neighbourhood $[x_0 - h, x_0 + h]$ of the initial point. However, in physical applications, one typically requires knowledge of the solution's behaviour over a larger domain, or indeed for all $x \geq 0$.

In this chapter, we investigate the conditions under which a local solution can be extended to a larger interval. We establish that a solution can be continued until it either approaches the boundary of the domain of definition or becomes unbounded.

5.1 The Extension Theorem

Consider the first-order differential equation

$$\frac{dy}{dx} = f(x, y), \tag{5.1}$$

where f is continuous on an open region $G \subseteq \mathbb{R}^2$. By *Peano Existence Theorem*, for any $P_0(x_0, y_0) \in G$, there exists at least one integral curve passing through P_0 defined on a small interval. We seek to extend this interval to its maximal possible range.

By the *maximal interval of existence* of a solution, we mean an interval on which the solution is defined and outside which no continuation of the same integral curve is possible.

Theorem 5.1. *Extension Principle.*
Let $f(x, y)$ be continuous in an open region G. Let $y = \varphi(x)$ be a solution of (5.1) passing through a point $P_0 \in G$. Then the solution can be extended in both directions until the integral curve $(x, \varphi(x))$ approaches the boundary of G.
More precisely, if G_1 is any bounded closed region such that $P_0 \in G_1 \subset G$, the integral curve must eventually exit G_1.

定理

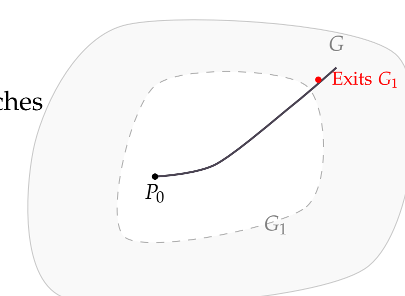

Figure 5.1: The Extension Theorem. A solution starting at P_0 cannot remain trapped inside a compact subset $G_1 \subset G$; it must reach the boundary of G_1.

Proof

Let J denote the maximal interval of existence for the solution

$$\Gamma : y = \varphi(x).$$

We focus on the extension to the right of x_0; the argument for the left is symmetric. Let $J^+ = J \cap [x_0, \infty)$ be the right-maximal interval. If $J^+ = [x_0, \infty)$, the solution extends indefinitely, satisfying the theorem (as it leaves any bounded set in the x-direction). Suppose J^+ is bounded. We distinguish two cases:

Case 1: ***J^+ is a closed interval*** $[x_0, x_1]$**.** Let $y_1 = \varphi(x_1)$. Since the interval is closed, the point $P_1(x_1, y_1)$ must lie within G. Since G is open, there exists a rectangular neighbourhood $R_1 \subset G$ centred at P_1:

$$R_1 = \{(x, y) : |x - x_1| \leq a_1, |y - y_1| \leq b_1\}.$$

By the local existence theorem (*Peano Existence Theorem*), there exists a solution $\psi(x)$ defined on $[x_1, x_1 + h]$ satisfying $\psi(x_1) = y_1$, for some $h > 0$. We may patch φ and ψ together:

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x) & x_0 \leq x \leq x_1, \\ \psi(x) & x_1 \leq x \leq x_1 + h. \end{cases}$$

This function is continuously differentiable and satisfies the ODE on $[x_0, x_1 + h]$, contradicting the assumption that x_1 was the rightmost endpoint. Thus, J^+ cannot be a closed interval.

Case 2: ***J^+ is a half-open interval*** $[x_0, x_1)$**.** We must show that the curve leaves any compact set $G_1 \subset G$. Suppose, for the sake of contradiction, that the integral curve remains entirely within a bounded closed region $G_1 \subset G$ for all $x \in [x_0, x_1)$.

Since f is continuous on the compact set G_1, it is bounded; say $|f(x, y)| \leq K$. Then for any $x \in [x_0, x_1)$:

$$|\varphi'(x)| = |f(x, \varphi(x))| \leq K.$$

By the Mean Value Theorem, for any $\tau, \sigma \in [x_0, x_1)$:

$$|\varphi(\tau) - \varphi(\sigma)| \leq K|\tau - \sigma|.$$

This Lipschitz property implies that $\varphi(x)$ is uniformly continuous on $[x_0, x_1)$. By the Cauchy criterion, the limit $\lim_{x \to x_1^-} \varphi(x)$ exists. Let this limit be y_1.

We define the extended function:

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x) & x_0 \leq x < x_1, \\ y_1 & x = x_1. \end{cases}$$

Integrating the differential equation:

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.$$

Taking the limit as $x \to x_1$, and using the continuity of f, we find that $\tilde{\varphi}$ satisfies the integral equation on $[x_0, x_1]$. Thus, $\tilde{\varphi}$ is a solution on the closed interval $[x_0, x_1]$. This reduces to Case 1, where we proved such a solution can be extended further to $[x_0, x_1 + h]$. This contradicts the maximality of the interval $[x_0, x_1)$.

Therefore, the solution cannot be trapped in any compact subset G_1 over a finite interval. It must eventually exit G_1.

■

Combining this result with the uniqueness theorem yields a stronger corollary.

Corollary 5.1. *Global Uniqueness and Extension.* Let $f(x, y)$ be continuous in G and satisfy a local Lipschitz condition with respect to y (i.e., around each point of G, there is a neighbourhood where f is Lipschitz in y). Then for any $P_0 \in G$, there exists a unique integral curve passing through P_0, and this curve extends to the boundary of G.

推論

Proof

Fix $P_0 = (x_0, y_0) \in G$.

Local existence and local uniqueness at P_0**.** By local Lipschitz at P_0, there exists a rectangle $Q \subset G$ centered at P_0 where f is Lipschitz in y. Since f is continuous on Q, *Picard's Existence and Uniqueness Theorem* applies on a smaller closed rectangle inside Q, yielding a (possibly small-interval) solution through P_0, and that solution is unique on that interval.

Extension. By the *Extension Principle*, this local solution extends (to the right and left) until its graph approaches the boundary of G.

Uniqueness of the whole integral curve. Let φ and ψ be two solutions through P_0, with domains I_φ and I_ψ. Consider $I = I_\varphi \cap I_\psi$, and define

$$E := \{x \in I : \varphi(x) = \psi(x)\}.$$

Clearly $x_0 \in E$, so $E \neq \varnothing$. Continuity of φ, ψ implies E is closed in I. Now take any $\xi \in E$, and set $y_\xi = \varphi(\xi) = \psi(\xi)$. By local Lipschitz at (ξ, y_ξ), there is a neighborhood where *Picard's Existence and Uniqueness Theorem* gives uniqueness for the IVP $y' = f(x, y),\ y(\xi) = y_\xi$. Hence $\varphi = \psi$ on some interval around ξ, so E is open in I.

Thus E is nonempty, open, and closed in the interval I; therefore $E = I$. So φ and ψ coincide on their common domain, i.e., the integral curve through P_0 is unique.

Hence there is a unique integral curve through P_0, and by the extension theorem it extends to the boundary of G.

■

Remark.

If G is the entire plane $\mathbb{R}^2$, "extending to the boundary" means that as x approaches the endpoints of the maximal interval, either $x \to \pm\infty$ or $|y| \to \infty$. It does *not* imply that the interval of existence is $(-\infty, \infty)$.

Finite Escape Time

A common misconception is that if $f(x, y)$ is smooth and defined everywhere (e.g., on $\mathbb{R}^2$), solutions must exist for all x. This is false for non-linear equations.

Example 5.1. Finite Interval of Existence. Consider the differential equation

$$\frac{dy}{dx} = x^2 + y^2. \tag{5.2}$$

The function $f(x, y) = x^2 + y^2$ is continuous and locally Lipschitz on the entire plane. By the Extension Principle, the unique solution through any point (x_0, y_0) extends to the boundary of $\mathbb{R}^2$ (infinity). We show that it reaches infinity in finite time.

範例

Solution

Let $y(x)$ be the solution with $y(x_0) = y_0$. Let $[x_0, \beta)$ be the maximal right interval. If $\beta \leq 0$, the interval is already finite. If $\beta > 0$, choose $x_1 \in (x_0, \beta)$ such that $x_1 > 0$. For $x \in [x_1, \beta)$, we have

$$y'(x) = x^2 + y^2 \geq x_1^2 + y^2.$$

Separating variables to bound the derivative:

$$\frac{y'(x)}{x_1^2 + y^2} \geq 1.$$

Integrate from x_1 to x:

$$\int_{x_1}^{x} \frac{y'(t)}{x_1^2 + y^2(t)}\, dt \geq \int_{x_1}^{x} dt \implies \frac{1}{x_1}\left[\arctan\frac{y(x)}{x_1} - \arctan\frac{y(x_1)}{x_1}\right] \geq x - x_1.$$

Rearranging for the length of the interval:

$$x - x_1 \leq \frac{1}{x_1}\left(\frac{\pi}{2} - \arctan\frac{y(x_1)}{x_1}\right) < \frac{\pi}{x_1}.$$

Thus, x cannot grow arbitrarily large; it is bounded by $x_1 + \pi/x_1$. This proves β is finite. The solution $y(x)$ tends to $+\infty$ as $x \to \beta$. For the branch $\beta \leq 0$, finiteness is immediate; and since β is a finite endpoint of a maximal interval, the Extension Principle implies $|y(x)| \to \infty$ as $x \to \beta^-$. Because $y'(x) = x^2 + y^2 > 0$, y is increasing, hence the only possibility is $y(x) \to +\infty$.

■

Qualitative Extension Analysis

Often we can determine global existence by examining the direction field, even without an explicit solution.

Example 5.2. Global Existence via Isoclines. Prove that the solution to the initial value problem

$$\frac{dy}{dx} = (x - y)e^{xy}, \quad y(x_0) = y_0 \tag{5.3}$$

extending to the right exists on $[x_0, \infty)$.

範例

Solution

By the corollary, a unique solution exists and extends to the boundary of $\mathbb{R}^2$. We must show it does not blow up (i.e., $|y| \to \infty$) at a finite x. The line $y = x$ is an isocline where $y' = 0$.

- Above the line ($y > x$), $x - y < 0$, so $y' < 0$ (decreasing).
- Below the line ($y < x$), $x - y > 0$, so $y' > 0$ (increasing).

Case $y_0 > x_0$**:** The solution starts above the isocline. Since $y' < 0$, the curve descends, so it cannot blow up to $+\infty$. If it stayed strictly above $y = x$ for all $x \geq x_0$, then $y(x) > x$, so in particular $y(x)$ cannot tend to $-\infty$ at any finite x. Hence no finite-time blow-up is possible while the solution remains above $y = x$; the only alternative is that it intersects $y = x$ and then enters the region $y < x$.

Case $y_0 \leq x_0$**:** The solution is below (or on) the isocline. Set $w(x) = y(x) - x$. Then

$$w'(x) = y'(x) - 1 = (x - y)e^{xy} - 1 = -w(x)e^{xy} - 1.$$

If $w(\xi) = 0$ at some ξ, then $w'(\xi) = -1 < 0$. Hence w cannot cross from negative to positive values at such a point. Therefore a solution starting with $w(x_0) \leq 0$ cannot cross from below $y = x$ to above $y = x$.

In the region $y < x$, we have $y'(x) > 0$. Also, since $y(x)$ is trapped below $y = x$, it cannot go to $+\infty$ for finite x. Can it go to $-\infty$? No, because y is increasing ($y' > 0$). Thus, for $x \geq x_0$, the solution is confined to a region where $|y|$ does not tend to infinity in finite time. The maximal interval is $[x_0, \infty)$.

■

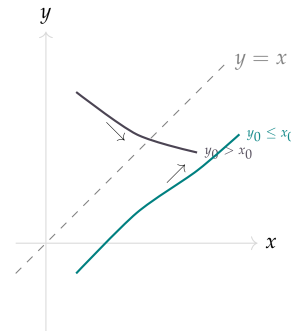

Figure 5.2: Solutions for $y' = (x - y)e^{xy}$. Curves above $y = x$ descend; curves below ascend but cannot cross $y = x$.

Criteria for Global Existence

If we cannot analyze the direction field geometrically, we may rely on growth bounds. Linearity in y ensures global existence.

Theorem 5.2. ***Linear Growth Bound.***
Let $f(x, y)$ be continuous on the strip

$$S = \{(x, y) : \alpha < x < \beta, -\infty < y < \infty\}.$$

Suppose there exist continuous non-negative functions $A(x)$ and $B(x)$ on (α, β) such that

$$|f(x, y)| \leq A(x)|y| + B(x) \tag{5.4}$$

for all $(x, y) \in S$. Then every solution exists on the entire interval (α, β).

定理

Proof

Let $y(x)$ be a solution with maximal right interval $[x_0, \beta_0)$ where $\beta_0 < \beta$. We proceed by contradiction.
Assume $\beta_0 < \beta$. Select x_1, x_2 such that $x_0 < x_1 < \beta_0 < x_2 < \beta$, with $x_2 - x_1$ sufficiently small. Specifically, let A_0, B_0 be the maxima of $A(x), B(x)$ on $[x_0, x_2]$. We choose x_1 close enough to β_0 such that $a_1 = x_2 - x_1$ satisfies

$$\begin{cases} a_1 < \dfrac{1}{4A_0}, & A_0 > 0, \\ a_1 > 0 \text{ arbitrary}, & A_0 = 0. \end{cases}$$

Let $y(x_1) = y_1$. We define a rectangle R_1 centered at (x_1, y_1) with width a_1 and height b_1 (where b_1 is large). On this rectangle,

$$|f(x, y)| \leq A_0(|y_1| + b_1) + B_0.$$

Let $M_1 = A_0(|y_1| + b_1) + B_0 + 1$. The local existence theorem guarantees a solution on an interval of radius

$$h_1 = \min\left(a_1, \frac{b_1}{M_1}\right).$$

We examine the ratio $\frac{b_1}{M_1}$ for large b_1:

$$\lim_{b_1 \to \infty} \frac{b_1}{M_1} = \begin{cases} \dfrac{1}{A_0}, & A_0 > 0, \\ \infty, & A_0 = 0. \end{cases}$$

Hence, in both cases, by taking b_1 sufficiently large we obtain $\frac{b_1}{M_1} > a_1$. Thus $h_1 = a_1$. This implies the solution exists on $[x_1, x_1 + a_1] = [x_1, x_2]$. Since $x_2 > \beta_0$, this contradicts the assumption that β_0 is the maximal endpoint. Therefore, $\beta_0 = \beta$. By symmetry, the left endpoint is α.

■

Note

This theorem explains why linear differential equations $y' + p(x)y = q(x)$, always have global solutions on the interval where p and q are continuous (since $|p(x)y - q(x)| \leq |p(x)||y| + |q(x)|$).

5.2 *Comparison Theorems*

While the Extension Principle guarantees that solutions continue until the boundary of the domain, it provides little insight into the quantitative behaviour of the solution, such as its magnitude or specific interval of existence. To estimate these properties, we compare the solutions of a difficult differential equation with those of a simpler one.

Theorem 5.3. ***First Comparison Theorem.***
Let $f(x, y)$ and $F(x, y)$ be continuous functions on a domain G satisfying the strict inequality

$$f(x, y) < F(x, y) \quad \text{for all } (x, y) \in G. \tag{5.5}$$

Let $y = \varphi(x)$ and $y = \Phi(x)$ be solutions on an interval (a, b) to the initial value problems

$$\varphi' = f(x, \varphi), \quad \Phi' = F(x, \Phi),$$

sharing the same initial condition $\varphi(x_0) = \Phi(x_0) = y_0$, where $(x_0, y_0) \in G$. Then:

$$\varphi(x) < \Phi(x) \quad \text{for } x \in (x_0, b), \tag{5.1}$$

$$\varphi(x) > \Phi(x) \quad \text{for } x \in (a, x_0). \tag{5.2}$$

定理

Proof

Consider the difference function $\psi(x) = \Phi(x) - \varphi(x)$. At the initial point, $\psi(x_0) = 0$. The derivative at x_0 is

$$\psi'(x_0) = \Phi'(x_0) - \varphi'(x_0) = F(x_0, y_0) - f(x_0, y_0) > 0.$$

By continuity, $\psi'(x) > 0$ in a small neighbourhood of x_0. Thus, $\psi(x)$ is increasing at x_0, implying $\psi(x) > 0$ for $x > x_0$ sufficiently close to x_0.

Suppose the inequality (5.1) fails at some point in (x_0, b). Let β be the first point greater than x_0 where the solutions meet:

$$\beta = \inf\{x \in (x_0, b) : \Phi(x) = \varphi(x)\}.$$

By continuity, $\psi(\beta) = 0$, and $\psi(x) > 0$ for $x \in (x_0, \beta)$. Hence for every $x \in (x_0, \beta)$,

$$\frac{\psi(\beta) - \psi(x)}{\beta - x} = \frac{-\psi(x)}{\beta - x} < 0.$$

Passing to the limit $x \to \beta^-$ gives $\psi'(\beta) \leq 0$. However, at the intersection point (β, γ) where $\gamma = \varphi(\beta) = \Phi(\beta)$, the differential equations require:

$$\psi'(\beta) = F(\beta, \gamma) - f(\beta, \gamma) > 0.$$

This contradiction ($\psi'(\beta) \leq 0$ vs $\psi'(\beta) > 0$) implies no such intersection point β exists. Thus $\varphi(x) < \Phi(x)$ for all $x \in (x_0, b)$. The proof for $x < x_0$ is analogous.

■

Remark.

Geometric meaning: at a fixed abscissa x, an integral curve whose slope is everywhere smaller cannot cross from below to above an integral curve whose slope is everywhere larger.

When the function $f(x, y)$ is continuous but fails the Lipschitz condition (e.g., $y' = \sqrt{|y|}$), uniqueness is not guaranteed. In such cases, solutions through (x_0, y_0) need not be unique, and extremal solutions provide sharp upper and lower envelopes.

Definition 5.1. ***Maximal and Minimal Solutions.***
Let S be the set of all solutions $y(x)$ to the initial value problem $y' = f(x, y), y(x_0) = y_0$ defined on an interval I.

- A solution $y_M(x)$ is the **maximal solution** if $y(x) \leq y_M(x)$ for all $y \in S$ and all $x \in I$.
- A solution $y_m(x)$ is the **minimal solution** if $y_m(x) \leq y(x)$ for all $y \in S$ and all $x \in I$.

定義

Definition 5.2. ***One-Sided Extremal Solutions.***
Let $I = (a, b)$ with $x_0 \in I$, and let S be the set of solutions of $y' = f(x, y),\ y(x_0) = y_0$ on I.

- A solution u is **right-maximal and left-minimal** if, for every $y \in S$,

$$y(x) \leq u(x)\ (x \in [x_0, b)), \qquad y(x) \geq u(x)\ (x \in (a, x_0]).$$

- A solution v is **right-minimal and left-maximal** if, for every $y \in S$,

$$y(x) \geq v(x)\ (x \in [x_0, b)), \qquad y(x) \leq v(x)\ (x \in (a, x_0]).$$

定義

The existence of these boundary solutions is guaranteed by approximating the equation from above and below.

Theorem 5.4. ***Existence of Extremal Solutions.***
Let $f(x, y)$ be continuous on a rectangle

$$R: \ |x - x_0| \leq a, \quad |y - y_0| \leq b.$$

Set

$$M = \max_{(x,y) \in R} |f(x, y)|, \qquad h = \begin{cases} \min\left\{a, \dfrac{b}{M}\right\}, & M > 0, \\ a, & M = 0. \end{cases}$$

Then there exists σ with $0 < \sigma < h$ such that on

$$I = [x_0 - \sigma, x_0 + \sigma],$$

the initial value problem $(E) : y' = f(x, y),\ y(x_0) = y_0$ admits both a maximal solution and a minimal solution.

定理

Proof
Choose σ with

$$0 < \sigma < \min\left\{a, \frac{b}{M+1}\right\} \leq h.$$

For $k \in \mathbb{N}$, consider the perturbed IVPs

$$\frac{dy}{dx} = f(x, y) + \frac{1}{k}, \quad y(x_0) = y_0. \tag{5.6}$$

By *Peano Existence Theorem*, each IVP has a solution φ_k on I. On I,

$$|\varphi_k'(x)| = |f(x, \varphi_k(x)) + 1/k| \leq M + 1.$$

Hence

$$|\varphi_k(x) - y_0| \leq (M+1)|x - x_0| \leq (M+1)\sigma < b,$$

so all graphs stay inside R and are uniformly bounded. Also,

$$|\varphi_k(x_1) - \varphi_k(x_2)| \leq (M+1)|x_1 - x_2|,$$

so $\{\varphi_k\}$ is equicontinuous on I. By *The Arzelà-Ascoli Theorem*, a subsequence converges uniformly on I to a continuous function y_M. Writing (5.6) in integral form and passing to the limit gives

$$y_M(x) = y_0 + \int_{x_0}^{x} f(t, y_M(t))\, dt,$$

so y_M solves the original IVP on I. To show maximality, let $y(x)$ be any solution of the original problem. Since $f(x,y) < f(x,y) + 1/k$, the *First Comparison Theorem* implies

$$y(x) < \varphi_k(x) \ (x_0 < x \leq x_0 + \sigma), \qquad y(x) > \varphi_k(x) \ (x_0 - \sigma \leq x < x_0).$$

Letting $k \to \infty$ yields

$$y(x) \leq y_M(x) \ (x_0 \leq x \leq x_0 + \sigma), \qquad y(x) \geq y_M(x) \ (x_0 - \sigma \leq x \leq x_0).$$

Thus y_M is right-maximal and left-minimal. Replacing $+1/k$ by $-1/k$ yields a limit solution y_m that is left-maximal and right-minimal. Define

$$\overline{y}(x) = \begin{cases} y_M(x), & x \in [x_0, x_0 + \sigma], \\ y_m(x), & x \in [x_0 - \sigma, x_0], \end{cases} \qquad \underline{y}(x) = \begin{cases} y_m(x), & x \in [x_0, x_0 + \sigma], \\ y_M(x), & x \in [x_0 - \sigma, x_0]. \end{cases}$$

Since $y_M(x_0) = y_m(x_0) = y_0$, both $\overline{y}$ and $\underline{y}$ are continuous at x_0. Moreover, since $y'_M(x_0) = f(x_0, y_0) = y'_m(x_0)$, the one-sided derivatives match, ensuring $\overline{y}$ and $\underline{y}$ are differentiable at x_0 and solve (E) on the entire interval I. For any solution y on $I = [x_0 - \sigma, x_0 + \sigma]$:

$$\underline{y}(x) \leq y(x) \leq \overline{y}(x) \qquad (x \in I).$$

Hence $\overline{y}$ is the maximal solution and $\underline{y}$ is the minimal solution on all of I.

■

Remark.

As in the extension theorem, local maximal/minimal solutions can be continued up to the boundary of G along their maximal intervals. Moreover, the IVP is unique if and only if its maximal and minimal solutions coincide.

This framework allows us to relax the strict inequality in *theorem* 5.3.

Theorem 5.5. ***Second Comparison Theorem.***
Let $f(x,y)$ and $F(x,y)$ be continuous on G such that $f(x,y) \leq F(x,y)$

for all $(x,y) \in G$. Let (a,b) be an interval with $x_0 \in (a,b)$. Assume φ and Φ solve on (a,b)

$$\varphi' = f(x,\varphi), \quad \Phi' = F(x,\Phi), \quad \varphi(x_0) = \Phi(x_0) = y_0.$$

Assume φ is right-minimal/left-maximal for $y' = f(x,y)$, and Φ is right-maximal/left-minimal for $y' = F(x,y)$. Then

$$\varphi(x) \leq \Phi(x) \ (x_0 \leq x < b), \qquad \varphi(x) \geq \Phi(x) \ (a < x \leq x_0).$$

定理

Proof

Fix a compact subinterval $J = [\ell, r] \subset (a,b)$ with $\ell < x_0 < r$. For each $k \in \mathbb{N}$, consider

$$\Phi_k' = F(x,\Phi_k) + \frac{1}{k}, \qquad \Phi_k(x_0) = y_0.$$

Using the same perturbation argument as in *Existence of Extremal Solutions* (applied on finite intervals, then continued along J), one obtains a right-maximal/left-minimal solution Φ_k of this perturbed IVP on J; moreover, a subsequence converges uniformly on J to the right-maximal/left-minimal solution of

$$\Phi' = F(x,\Phi), \qquad \Phi(x_0) = y_0,$$

which is exactly the given Φ.
Now $f \leq F < F + 1/k$, so *First Comparison Theorem* yields

$$\varphi(x) < \Phi_k(x) \ (x_0 < x \leq r), \qquad \varphi(x) > \Phi_k(x) \ (\ell \leq x < x_0).$$

Passing to the convergent subsequence limit on J gives

$$\varphi(x) \leq \Phi(x) \ (x_0 \leq x \leq r), \qquad \varphi(x) \geq \Phi(x) \ (\ell \leq x \leq x_0).$$

Since $J = [\ell, r] \subset (a,b)$ is arbitrary, we conclude

$$\varphi(x) \leq \Phi(x) \ (x_0 \leq x < b), \qquad \varphi(x) \geq \Phi(x) \ (a < x \leq x_0).$$

■

Comparison theorems are powerful tools for determining the long-term behaviour of solutions and estimating their intervals of existence.

Example 5.3. Asymptotic Decay. Consider the differential equation

$$\frac{dy}{dx} = \sin(xy). \tag{5.7}$$

Because $f(x,y) = \sin(xy)$ is continuous and satisfies $|f(x,y)| \leq 1$, theorem 5.2 implies every solution exists for all $x \in \mathbb{R}$. Since $\partial f/\partial y$ is continuous, f is locally Lipschitz in y, so local uniqueness holds by *Picard's Existence and Uniqueness Theorem*. Fix $y(0) = y_0 > 0$. By symmetry of (5.7), it is enough to study $x \to +\infty$ in the first quadrant.

Let

$$H_k: \; xy = k\pi, \qquad G_k: \; (k-1)\pi < xy < k\pi \quad (k \in \mathbb{N}).$$

Then

$$0 < y' < 1 \text{ in } G_{2n-1}, \qquad -1 < y' < 0 \text{ in } G_{2n}.$$

Construct a polygonal line $A : y = u(x)$ starting from $P_0(0, y_0)$ recursively. For $n \geq 0$:

1. From P_{2n}, draw a segment with slope 1 through G_{2n+1} until it intersects H_{2n+1} at P_{2n+1}.
2. From P_{2n+1}, draw a segment with slope 0 through G_{2n+2} until it intersects H_{2n+2} at P_{2n+2}.

Apply *First Comparison Theorem* strip-by-strip: if Γ is below A at one vertex, then it stays below on the next strip (since $u' \geq \sup f$ in each region). By induction, $\Gamma : y = y(x)$ satisfies $y(x) \leq u(x)$ for all $x \geq 0$.

Write vertices as $P_k = (x_k, y_k)$ with $x_k y_k = k\pi$. For odd $n = 2k-1$, one has $y_{n+1} = y_n$ (slope 0 step), $y_{n+2} > y_n$ (slope 1 step), and

$$x_{n+1} - x_n = \frac{\pi}{y_n}, \qquad x_{n+2} - x_{n+1} < \frac{\pi}{y_n},$$

hence $x_{n+2} - x_{n+1} < x_{n+1} - x_n$, so

$$\frac{y_{n+2} - y_n}{x_{n+2} - x_n} < \frac{1}{2}.$$

Thus, for every odd n, the secant slope of $P_n P_{n+2}$ is $< 1/2$. Let L_1 be the line through P_1 with slope $1/2$. Iterating the previous inequality over odd indices gives that all later vertices of A lie below L_1, hence A itself lies below L_1. Since L_1 has slope $1/2 < 1$, there exists x^* such that $L_1(x) < x$ for $x > x^*$. Therefore, for $x > x^*$,

$$y(x) \leq u(x) \leq L_1(x) < x.$$

At $x = 0$, $y(0) = y_0 > 0 = x$. By continuity, Γ intersects $L : y = x$ at some $(\bar{x}, \bar{x})$.

Now compare Γ with $z(x) = x$, the solution of $z' = 1$, $z(\bar{x}) = \bar{x}$. Since $\sin(xy) \leq 1$, and local uniqueness holds for both equations (hence maximal/minimal solutions coincide; see the remark after theorem 5.4), theorem 5.5 gives

$$y(x) \leq x \qquad (x \geq \bar{x}).$$

Choose the smallest integer m with $\bar{x}^2 < (2m - \frac{1}{2})\pi$, and set

$$H: \ xy = (2m - \tfrac{1}{2})\pi.$$

At any point of H, the ODE slope is $y' = \sin((2m - \frac{1}{2})\pi) = -1$, while the hyperbola slope is

$$K_H = -\frac{y}{x} = -\frac{(2m - \frac{1}{2})\pi}{x^2}.$$

If Γ crossed H upward (from below to above) at some $x_1 > \bar{x}$, we would need $y'(x_1) \geq K_H(x_1)$. Also, from (5.7) and *theorem* 5.5, $y(x) \leq x$ for $x \geq \bar{x}$. Hence on H we have

$$K_H(x_1) = -\frac{y(x_1)}{x_1} \geq -1.$$

If equality held, then $y(x_1) = x_1$ and therefore $y'(x_1) = \sin(x_1^2) = -1 < 1 = z'(x_1)$, so immediately to the right of x_1 we would still have $y(x) < x$, and no upward crossing of H can occur there. Thus at an upward crossing point we must have $y(x_1) < x_1$, so $K_H(x_1) > -1$. But on H, $y'(x_1) = \sin((2m - \frac{1}{2})\pi) = -1 < K_H(x_1)$, contradicting the crossing criterion. Hence for $x > \bar{x}$,

$$y(x) < \frac{(2m - \frac{1}{2})\pi}{x}.$$

Furthermore, since $y(0) = y_0 > 0$ and $y \equiv 0$ is a solution (as $f(x, 0) = 0$), uniqueness implies $y(x)$ never vanishes. Thus $y(x) > 0$ for all x. By squeezing, $\lim_{x\to\infty} y(x) = 0$.

範例

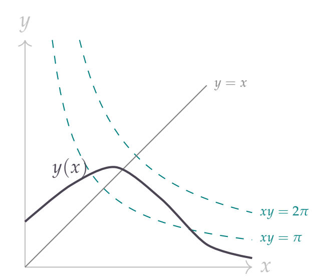

Figure 5.3: Solution of $y' = \sin(xy)$. The solution oscillates but is eventually trapped by the decaying hyperbolas.

Example 5.4. Estimation of Existence Interval. Let $[0, \beta^*)$ be the maximal right interval for the solution to

$$\frac{dy}{dx} = x^2 + (y+1)^2, \quad y(0) = 0. \tag{5.8}$$

We prove

$$\frac{\pi}{4} < \beta^* < 1.$$

For $x \in [0, 1]$,

$$(y+1)^2 \leq x^2 + (y+1)^2 \leq 1 + (y+1)^2.$$

Compare with

$$W' = (W+1)^2, \ W(0) = 0, \qquad Z' = 1 + (Z+1)^2, \ Z(0) = 0.$$

Explicitly,

$$W(x) = \frac{1}{1-x} - 1 \quad \text{(blow-up at } x = 1\text{)},$$

$$\arctan(Z+1) - \frac{\pi}{4} = x \quad (\text{blow-up at } x = \pi/4).$$

By *theorem 5.5*,

$$W(x) \le y(x) \le Z(x) \quad (0 \le x < \min\{\beta^*, \pi/4, 1\}),$$

so

$$\frac{\pi}{4} \le \beta^* \le 1.$$

To prove strict $\beta^* < 1$, pick $\xi \in (0, \beta^*)$ and set $\eta = y(\xi)$. Compare with

$$Y' = (Y+1)^2, \qquad Y(\xi) = \eta.$$

Its blow-up time is

$$C(\xi) = \xi + \frac{1}{\eta+1}.$$

Since $y'(\xi) = \xi^2 + (\eta+1)^2$, we get

$$C'(\xi) = 1 - \frac{y'(\xi)}{(\eta+1)^2} = 1 - \frac{\xi^2 + (\eta+1)^2}{(\eta+1)^2} < 0.$$

Also $C(0) = 1$, hence $C(\xi) < 1$ for $\xi > 0$. Comparison gives $\beta^* \le C(\xi) < 1$, so $\beta^* < 1$. To prove strict $\beta^* > \pi/4$, choose $\lambda \in (0,1)$ and compare with

$$Y_\lambda' = \lambda^2 + (Y_\lambda + 1)^2, \qquad Y_\lambda(0) = 0.$$

This has blow-up time

$$C(\lambda) = \frac{1}{\lambda}\left(\frac{\pi}{2} - \arctan\frac{1}{\lambda}\right).$$

Since $C(1) = \pi/4$ and

$$C'(\lambda) = -\frac{1}{\lambda^2}\left(\frac{\pi}{2} - \arctan\frac{1}{\lambda}\right) + \frac{1}{\lambda(1+\lambda^2)},$$

$$C'(1) = -\frac{\pi}{4} + \frac{1}{2} < 0,$$

we can choose $\lambda \in (\pi/4, 1)$ sufficiently close to 1 so that $C(\lambda) > \pi/4$. Then on $[0, \pi/4]$, $x^2 \le \lambda^2$, hence

$$x^2 + (y+1)^2 \le \lambda^2 + (y+1)^2.$$

By *theorem 5.5*, $y(x) \le Y_\lambda(x)$ on $[0, \min\{\beta^*, \pi/4\}]$. Also $C(\lambda) > \frac{\pi}{4}$, so Y_λ does not blow up on $[0, \pi/4]$, and therefore y does not blow up on $[0, \pi/4]$. Thus $\beta^* > \pi/4$. Therefore

$$\frac{\pi}{4} < \beta^* < 1.$$

範例

5.3 Exercises

1. **Existence Intervals.** Discuss the intervals of existence for solutions to the following differential equations:

   (a) $\frac{dy}{dx} = \frac{1}{x^2+y^2}$

   (b) $\frac{dy}{dx} = y(y-1)$

   (c) $\frac{dy}{dx} = y\sin(xy)$

   (d) $\frac{dy}{dx} = 1 + x^2 + y^2$

2. **Extension and Compactness.** Consider the differential equation in symmetric form $x\,dx + y\,dy = 0$ defined on the domain

$$G = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 > 0\}.$$

   The unit circle $x^2 + y^2 = 1$ is an integral curve contained entirely within G, yet it is a closed curve and does not "extend to the boundary" in the sense of becoming unbounded. Does this contradict the Extension Principle (*theorem* 5.1)? Explain your reasoning carefully.

3. ⋆ **One-Sided Unboundedness.** Let (a,b) be the maximal interval of existence for the solution to the initial value problem:

$$\frac{dy}{dx} = (y^2 - 2y - 3)e^{(x+y)^2}, \quad y(x_0) = y_0.$$

   Prove that at least one of the conditions $a = -\infty$ or $b = +\infty$ must hold.

4. ⋆ **Global Extension for Small Initial Data.** Consider the initial value problem

$$\frac{dy}{dx} = (x^2 - y^2)f(x,y), \quad y(x_0) = y_0,$$

   where $f(x,y)$ is continuous on $\mathbb{R}^2$ and satisfies $f(x,y) > 0$ for $y \neq 0$. Prove that for each fixed $x_0 < 0$, there exists $\delta = \delta(x_0, f) > 0$ such that if $|y_0| < \delta$, then the solution extends to $(-\infty, +\infty)$.

5. **Filling the Gap Between Extremal Solutions.** Let the initial value problem (E), and the interval $I = [x_0 - \sigma, x_0 + \sigma]$ be as in *Existence of Extremal Solutions*. Let $y = W(x)$ and $y = Z(x)$ be the minimal and maximal solutions of (E) on I, respectively. Prove that the region between these solutions is filled by other solutions. Specifically, show that for any point (x_1, y_1) such that

$$x_1 \in I \quad \text{and} \quad W(x_1) \leq y_1 \leq Z(x_1),$$

   there exists at least one solution $y = u(x)$ defined on I satisfying $u(x_1) = y_1$.

6

Implicit Equations and Singular Solutions

We now extend our geometric and analytic framework to **first-order implicit differential equations**, where the derivative is intertwined non-linearly with the variables. Generally, the families of integral curves resolving such equations possess a general solution parameterised by an arbitrary constant. However, the non-linear algebraic structure of the implicit equation often permits the existence of **singular solutions** — curves along which the uniqueness of the initial value problem is geometrically destroyed.

6.1 *First-Order Implicit Differential Equations*

We consider equations of the form

$$F\left(x, y, \frac{dy}{dx}\right) = 0. \tag{6.1}$$

To simplify notation, it is standard practice to denote the derivative by $p = \frac{dy}{dx}$. Thus, (6.1) is written as $F(x, y, p) = 0$.

The Method of Differentiation

When (6.1) cannot be solved explicitly for p, it may instead be solvable for y (or x). Suppose we can isolate y such that

$$y = f(x, p). \tag{6.2}$$

Assuming f is continuously differentiable with respect to both arguments, we differentiate (6.2) with respect to x, recalling that $y' = p$:

$$p = \frac{\partial f}{\partial x}(x, p) + \frac{\partial f}{\partial p}(x, p)\frac{dp}{dx}.$$

Rearranging this yields a first-order differential equation in the variables x and p:

$$\left[\frac{\partial f}{\partial x}(x, p) - p\right] + \frac{\partial f}{\partial p}(x, p)\frac{dp}{dx} = 0. \tag{6.3}$$

If we can determine the general solution to (6.3), say $p = u(x, C)$ for an arbitrary constant C, substituting this parameterisation back into (6.2) furnishes the general solution of the original equation:

$$y = f(x, u(x, C)).$$

Alternatively, if (6.3) is more naturally solved for x as a function of p, yielding $x = v(p, C)$, the general solution is obtained parametrically:

$$\begin{cases} x = v(p, C) \\ y = f(v(p, C), p) \end{cases} \tag{6.4}$$

where p operates as the parameter. Crucially, if (6.3) admits a specific solution $p = w(x)$ free of an arbitrary constant, the corresponding curve $y = f(x, w(x))$ may represent a singular solution.
A paramount example of the differentiation method is the resolution of a specific class of equations introduced by Alexis Clairaut.

Theorem 6.1. ***Clairaut's Equation.***
A differential equation of the form

$$y = xp + f(p), \quad \left(p = \frac{dy}{dx}\right) \tag{6.5}$$

where f is a continuously differentiable function such that $f''(p) \neq 0$, possesses a general solution consisting of a family of straight lines:

$$y = Cx + f(C),$$

and a singular solution given parametrically by

$$x = -f'(p), \quad y = -pf'(p) + f(p).$$

This singular solution is the envelope of the general family of lines.

定理

Proof
Applying the differentiation method to (6.5), we differentiate with respect to x:

$$p = p + x\frac{dp}{dx} + f'(p)\frac{dp}{dx}.$$

Cancelling p from both sides and factoring yields:

$$\left[x + f'(p)\right]\frac{dp}{dx} = 0. \tag{6.6}$$

This equation is satisfied if either factor vanishes.

Case 1: $\frac{dp}{dx} = 0$. Integration gives $p = C$, a constant. Substituting this into the original equation (6.5) yields the general solution:

$$y = Cx + f(C).$$

Geometrically, this is a one-parameter family of straight lines.

Case 2: $x + f'(p) = 0$. This implies $x = -f'(p)$. Substituting this into (6.5) yields $y = -pf'(p) + f(p)$. Since $f''(p) \neq 0$, the mapping $p \mapsto -f'(p)$ is locally invertible, meaning p can be viewed as a smooth function of x, say $p = w(x)$. The solution curve is thus $y = xw(x) + f(w(x))$.

To verify that this second case represents an envelope, we examine the derivative of the singular curve. By the chain rule:

$$\frac{dy}{dx} = w(x) + xw'(x) + f'(w(x))w'(x) = w(x) + w'(x)\left[x + f'(w(x))\right].$$

Since $x = -f'(w(x))$ on this curve, the bracketed term vanishes, leaving $\frac{dy}{dx} = w(x)$. Thus, at any point (x_0, y_0) on the singular curve, its tangent line has slope $C_0 = w(x_0)$ and passes through (x_0, y_0), matching precisely the line $y = C_0 x + f(C_0)$ from the general solution. The singular solution is everywhere tangent to a member of the general solution family, thereby forming its envelope.

■

Example 6.1. Parabolic Envelope. Solve the differential equation

$$y = xp - \frac{1}{4}p^2.$$

範例

Solution

This is Clairaut's equation (6.5) with $f(p) = -\frac{1}{4}p^2$. By *theorem* 6.1, the general solution is immediately given by the family of lines:

$$y = Cx - \frac{1}{4}C^2.$$

To find the singular solution, we set $x + f'(p) = 0$. Since $f'(p) = -\frac{1}{2}p$, this requires $x - \frac{1}{2}p = 0$, or $p = 2x$. Substituting $p = 2x$ into the original differential equation yields:

$$y = x(2x) - \frac{1}{4}(2x)^2 = 2x^2 - x^2 = x^2.$$

The singular solution is the parabola $y = x^2$. As illustrated in *figure* 6.1, the family of lines defined by the general solution envelops this parabola. Because $w'(x) = 2 \neq 0$, the parabolic solution is not a member of the general linear family.

■

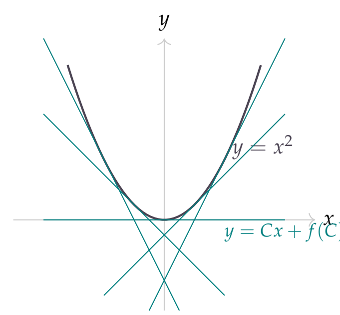

Figure 6.1: The general solutions of $y = xp - \frac{1}{4}p^2$ are lines, which collectively envelop the singular solution $y = x^2$.

The method of differentiation extends beyond Clairaut's equation to more general forms where y appears linearly or can be algebraically isolated.

Example 6.2. Non-linear Implicit Equation. Determine the general and singular solutions of

$$x(y')^2 - 2yy' + 9x = 0. \tag{6.7}$$

範例

Solution

Substituting $p = y'$ and isolating y, we divide by $2p$ (assuming $p \neq 0$):

$$y = \frac{xp}{2} + \frac{9x}{2p}. \tag{6.8}$$

Differentiating with respect to x, we apply the product and quotient rules:

$$p = \frac{p}{2} + \frac{x}{2}\frac{dp}{dx} + \frac{9}{2p} - \frac{9x}{2p^2}\frac{dp}{dx}.$$

Grouping the terms without derivatives on the left:

$$\frac{p}{2} - \frac{9}{2p} = \frac{x}{2}\left(1 - \frac{9}{p^2}\right)\frac{dp}{dx}.$$

Factoring $\frac{1}{p}$ from the left side gives $\frac{1}{2p}(p^2 - 9)$. Similarly, factoring the right side yields $\frac{x}{2p^2}(p^2 - 9)\frac{dp}{dx}$. Bringing all terms to one side, we obtain:

$$\left(\frac{p^2-9}{2p^2}\right)\left(p - x\frac{dp}{dx}\right) = 0. \tag{6.9}$$

This equation is satisfied if either of the two distinct conditions holds.

Case 1: $p - x\frac{dp}{dx} = 0$. Separating variables yields $\frac{dp}{p} = \frac{dx}{x}$, which integrates to $p = Cx$. Substituting this back into (6.8) furnishes the general solution:

$$y = \frac{x(Cx)}{2} + \frac{9x}{2Cx} = \frac{C}{2}x^2 + \frac{9}{2C}.$$

This defines a family of parabolas parameterised by $C \neq 0$.

Case 2: $p^2 - 9 = 0$. This forces $p = 3$ or $p = -3$. Since these are constants, $dp/dx = 0$, and substituting these constant values directly into (6.8) yields two special solutions:

$$y = \frac{3x}{2} + \frac{9x}{6} = 3x, \quad \text{and} \quad y = -\frac{3x}{2} - \frac{9x}{6} = -3x.$$

The straight lines $y = \pm 3x$ are special solutions. Restricted to $x \neq 0$, each branch is singular. As seen in *figure* 6.2, the general solution

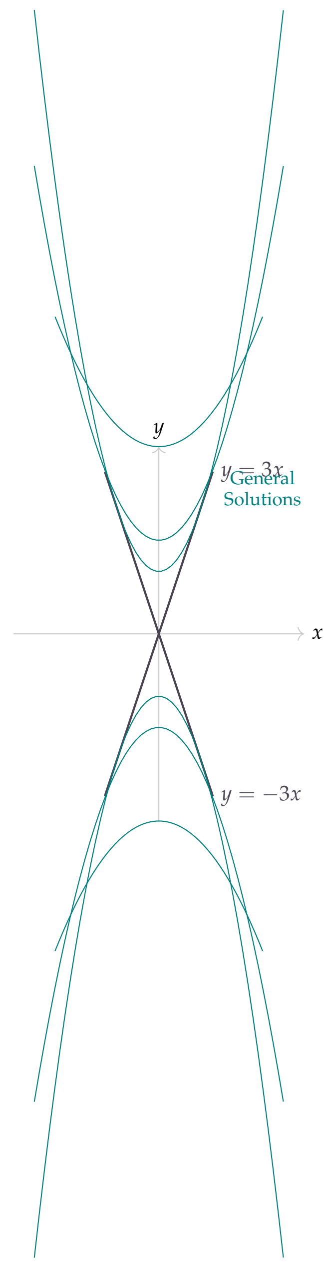

Figure 6.2: The general solution family of parabolas and their linear envelope branches (singular for $x \neq 0$).

constitutes a family of parabolas that does not include these lines for any real value of C. Moreover, every point on these lines with $x \neq 0$ acts as a point of tangency to exactly one of the parabolas. They form the bounding envelope of the system away from the origin.

■

6.2 *The Parametric Method*

When a first-order implicit differential equation cannot be easily resolved for y or x, we may instead parameterise the variables. Geometrically, the equation $F(x, y, p) = 0$ defines a surface in the three-dimensional space with coordinates (x, y, p). A solution is a curve on this surface that additionally satisfies the fundamental constraint $dy = p\,dx$.

Equations Lacking the Independent Variable

Consider the autonomous implicit equation where x is explicitly absent:

$$F(y, p) = 0. \tag{6.10}$$

This equation typically represents a family of curves in the yp-plane. Suppose we can parameterise such a curve by a real variable t, yielding:

$$y = g(t), \quad p = h(t).$$

Assuming g and h are continuous, g is differentiable, and $h(t) \neq 0$, we substitute this parameterisation into the differential constraint $dx = \frac{1}{p}\,dy$. Using the chain rule $dy = g'(t)\,dt$, we find:

$$dx = \frac{g'(t)}{h(t)}\,dt.$$

Direct integration furnishes x as a function of t:

$$x = \int \frac{g'(t)}{h(t)}\,dt + C.$$

Thus, the general solution to (6.10) is given parametrically by

$$\begin{cases} x = \int \frac{g'(t)}{h(t)}\,dt + C \\ y = g(t). \end{cases} \tag{6.11}$$

Example 6.3. Circular Parameterisation. Solve the differential equation

$$y^2 + \left(\frac{dy}{dx}\right)^2 = 1. \tag{6.12}$$

範例

Solution

The relation $y^2 + p^2 = 1$ describes a circle in the yp-plane. A natural parameterisation is:

$$y = \cos t, \quad p = \sin t, \quad (-\infty < t < \infty).$$

Applying the parametric relation $dx = \frac{1}{p}\,dy$, we compute the differential $dy = -\sin t\,dt$. Thus:

$$dx = \frac{-\sin t\,dt}{\sin t} = -dt.$$

Integrating yields $x = -t + C$. The general solution in parametric form is therefore $x = -t + C$ and $y = \cos t$. Eliminating the parameter $t = C - x$, we recover the explicit general solution:

$$y = \cos(C - x) = \cos(x - C).$$

We must also inspect the boundary of the parameterisation where $p = \sin t = 0$. This corresponds to $y = \pm 1$. Substituting $y = \pm 1$ back into (6.12), we see that $(\pm 1)^2 + 0^2 = 1$, which is satisfied identically. Hence, $y = 1$ and $y = -1$ are valid solutions. As illustrated in *figure* 6.3, these constant solutions form the upper and lower envelopes of the family of translated cosine waves.

■

Note

Setting $y = 0$ implies $p = \pm 1$, which corresponds to points where the integral curves cross the x-axis. However, $y = 0$ is not a solution because its derivative is identically zero, failing the condition $0^2 + 0^2 = 1$.

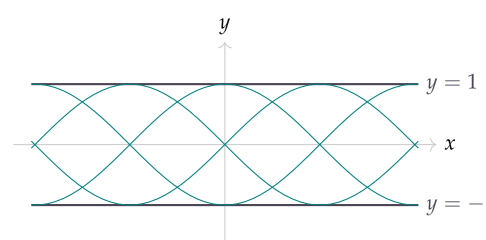

Figure 6.3: The integral curves $y = \cos(x - C)$ bounded by the singular envelopes $y = \pm 1$.

The General Parametric Approach

The technique naturally generalises to the full implicit equation $F(x, y, p) = 0$. Suppose the surface defined by $F = 0$ admits a global parameterisation via two parameters u and v:

$$x = f(u, v), \quad y = g(u, v), \quad p = h(u, v).$$

We enforce the differential constraint $dy = p\,dx$. Expanding the total differentials $dy = g_u\,du + g_v\,dv$ and $dx = f_u\,du + f_v\,dv$, where subscripts denote partial derivatives, we obtain:

$$g_u\,du + g_v\,dv = h(u,v)\,(f_u\,du + f_v\,dv)\,.$$

Gathering the coefficients of du and dv yields a first-order explicit differential equation in terms of the parameters:

$$[g_u - h(u,v)f_u]\,du + [g_v - h(u,v)f_v]\,dv = 0. \tag{6.13}$$

If (6.13) possesses a general solution of the form $v = Q(u,C)$, substituting this back into the original parameterisation resolves the full system:

$$x = f(u, Q(u,C)), \quad y = g(u, Q(u,C)).$$

Furthermore, any particular solution $v = S(u)$ lacking the arbitrary constant corresponds to an additional special solution $x = f(u, S(u)), y = g(u, S(u))$. Such a curve is singular only if it satisfies the tangency criterion of *definition* 6.1.

Example 6.4. Parameterisation by Variables. Find the general and additional special solutions to

$$\left(\frac{dy}{dx}\right)^2 + y - x = 0. \tag{6.14}$$

範例

Solution

This equation can be expressed as $y = x - p^2$. We may directly use x and p as our parameters by setting $u = x$ and $v = p$. The parameterisation of the surface is:

$$x = u, \quad p = v, \quad y = u - v^2.$$

We now enforce $dy = p\,dx$. Differentiating our parametric equations yields $dx = du$ and $dy = du - 2v\,dv$. Substituting these into the constraint gives:

$$du - 2v\,dv = v\,du.$$

Rearranging to separate variables:

$$(1 - v)\,du - 2v\,dv = 0 \implies du = \frac{2v}{1-v}\,dv.$$

We rewrite the right-hand side to facilitate integration:

$$du = \left(-2 + \frac{2}{1-v}\right)dv.$$

Integrating both sides produces the general solution for u in terms of v:

$$u = -2v - 2\ln|v-1| + C = -2v - \ln(v-1)^2 + C.$$

Returning to the parameterisation $x = u$ and $y = u - v^2$, the general solution to (6.14) is given parametrically by:

$$\begin{cases} x = C - 2v - \ln(v-1)^2 \\ y = C - 2v - \ln(v-1)^2 - v^2 \end{cases} \tag{6.15}$$

where v acts as the parameter.
To identify additional special solutions, we examine the separation step where we divided by $1 - v$. Setting $1 - v = 0$ yields $v = 1$. This is a particular solution to $(1 - v)\,du - 2v\,dv = 0$ that cannot be obtained from the general family for any finite C. Substituting $v = 1$ into our initial parameterisation gives $p = 1$ and $y = u - 1$. Since $x = u$, this defines the straight line:

$$y = x - 1.$$

Substituting this back into the original equation (6.14) verifies it is indeed a solution: $(1)^2 + (x - 1) - x = 0$. Thus, $y = x - 1$ is a special solution not contained in the general family (6.15).

■

6.3 *Singular Solutions and the p-Discriminant*

In the preceding sections, we observed that certain particular solutions to first-order implicit differential equations—such as the envelope of Clairaut's equation—possess distinct geometric properties. Specifically, they violate the uniqueness of solutions for initial value problems. We now formalise this phenomenon.

Definition 6.1. ***Singular Solution.***
Let a first-order differential equation be given by

$$F\left(x, y, \frac{dy}{dx}\right) = 0. \tag{6.16}$$

A solution $\Gamma : y = \varphi(x)$ defined on an interval J is termed a **singular solution** if, for every point $Q \in \Gamma$, any neighbourhood of Q contains a distinct solution of (6.16) that is tangent to Γ at Q.

定義

Geometrically, a singular solution is the envelope of a family of integral curves. Examples we have already encountered include the branches $y = \pm 3x$ for equation (6.7) and the unit bounds $y = \pm 1$ for equation (6.12). For the lines $y = \pm 3x$, the singularity statement is made on the intervals $J = (0, \infty)$ and $J = (-\infty, 0)$ separately (equivalently, $x \neq 0$), not on an interval containing $x = 0$.

We now seek an analytic criterion to systematically locate such solutions without first requiring the general solution.

Theorem 6.2. ***The p-Discriminant (Necessary Condition).***
Let the function $F(x, y, p)$ be continuously differentiable on a domain G (equivalently, F_x, F_y, and F_p are continuous on G). If the curve $y = \varphi(x)$ is a singular solution of $F(x, y, p) = 0$ such that $(x, \varphi(x), \varphi'(x)) \in G$ for all $x \in J$, then $\varphi(x)$ necessarily satisfies the simultaneous equations:

$$F(x, y, p) = 0, \quad \frac{\partial F}{\partial p}(x, y, p) = 0, \quad \left(p = \frac{dy}{dx}\right). \tag{6.17}$$

定理

Proof

Because $y = \varphi(x)$ is a solution to the differential equation, it trivially satisfies the first condition $F(x, y, p) = 0$. We must prove it satisfies the second.

Assume, for the sake of contradiction, that there exists a point $x_0 \in J$ where $F_p(x_0, y_0, p_0) \neq 0$, with $y_0 = \varphi(x_0)$ and $p_0 = \varphi'(x_0)$. Since $F(x_0, y_0, p_0) = 0$ and the partial derivatives are continuous, the Implicit Function Theorem guarantees that the equation $F(x, y, p) = 0$ can be uniquely solved for p in a neighbourhood of (x_0, y_0). That is, there exists a unique function $f(x, y)$ such that

$$\frac{dy}{dx} = f(x, y) \tag{6.18}$$

with $f(x_0, y_0) = p_0$. Thus, any solution of the implicit equation passing through (x_0, y_0) with slope p_0 must also satisfy the explicit equation (6.18).

Furthermore, in this neighbourhood, the partial derivative of f with respect to y is given by

$$\frac{\partial f}{\partial y}(x, y) = -\frac{F_y(x, y, f(x, y))}{F_p(x, y, f(x, y))}.$$

Because $F_p \neq 0$, f_y is continuous, implying that $f(x, y)$ is locally Lipschitz in y. By Picard's Existence and Uniqueness Theorem, the initial value problem for (6.18) with $y(x_0) = y_0$ admits exactly one solution. Consequently, $y = \varphi(x)$ is the unique integral curve passing through (x_0, y_0) with slope p_0.

This precludes the existence of any other solution to (6.16) tangent to $\varphi(x)$ at x_0, directly contradicting the premise that $y = \varphi(x)$ is a singular solution. The assumption $F_p \neq 0$ must therefore be false, completing the proof.

■

Remark.

Eliminating the parameter p from the system (6.17) yields an algebraic relation $\Delta(x, y) = 0$, known as the **p-discriminant curve**. To borrow a perspective established in this section of the geometry notes, the p-discriminant determines the geometric locus where the uniqueness of the tangent direction breaks down. However, the curve $\Delta(x, y) = 0$ is a set of candidates; it is not guaranteed to be a singular solution, nor even a solution at all.

To demonstrate the limitations of the necessary condition, we examine two cases where the p-discriminant fails to yield a singular solution.

Example 6.5. Non-Solution Discriminant. Find the p-discriminant curve of the equation

$$\left(\frac{dy}{dx}\right)^2 + y - x = 0. \tag{6.19}$$

範例

Solution

Let $F(x, y, p) = p^2 + y - x$. We compute $F_p = 2p$. Setting $F_p = 0$ requires $p = 0$. Substituting $p = 0$ into $F = 0$ yields the p-discriminant curve:

$$y = x.$$

However, if $y = x$, then its derivative is $p = 1$. Substituting these back into the original equation (6.19) gives $(1)^2 + x - x = 1 \neq 0$. Thus, the curve $y = x$ is not a solution to the differential equation.

■

Example 6.6. Non-Singular Solution. Analyse the p-discriminant of the equation

$$\left(\frac{dy}{dx}\right)^2 - y^2 = 0. \tag{6.20}$$

範例

Solution

Let $F(x, y, p) = p^2 - y^2$. Setting $F_p = 2p = 0$ gives $p = 0$. Substituting this into $F = 0$ yields $y^2 = 0$, so the p-discriminant curve is $y = 0$.

Since the derivative of $y = 0$ is $p = 0$, it identically satisfies (6.20), meaning it is a valid solution. To determine if it is singular, we find the general solution. Factoring the differential equation gives $y' = y$ or $y' = -y$. The general solutions are:

$$y = Ce^x \quad \text{and} \quad y = Ke^{-x}.$$

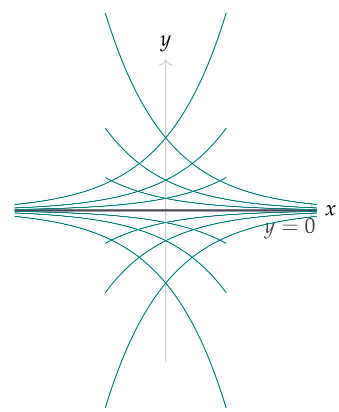

Figure 6.4: Integral curves for $(y')^2 - y^2 = 0$. The trivial solution $y = 0$ is found via the p-discriminant, but it acts as an asymptote rather than an envelope.

For any point on $y = 0$, a tangent solution would require $y = 0$ and $y' = 0$ simultaneously. For the general families, $Ce^x = 0$ only if $C = 0$, which merely returns the trivial solution $y = 0$ itself. No distinct solutions touch $y = 0$ tangentially. Therefore, uniqueness holds along the x-axis, and $y = 0$ is an ordinary, not singular, solution (see *figure* 6.4).

■

Because verifying tangency against the general solution is computationally burdensome (and often impossible if the general solution is unknown), a sufficient algebraic criterion is highly desirable.

Theorem 6.3. ***Sufficient Conditions for Singularity.***
Let $F(x, y, p)$ be twice continuously differentiable on a domain G. Suppose the curve $y = \psi(x)$, obtained by eliminating p from the p-discriminant system (6.17), is a valid solution to the differential equation $F(x, y, p) = 0$.
If the following strict conditions hold along the curve for all $x \in J$:

$$F_p(x, \psi, \psi') = 0, \quad F_y(x, \psi, \psi') \neq 0, \quad F_{pp}(x, \psi, \psi') \neq 0, \tag{6.21}$$

then $y = \psi(x)$ is a singular solution.

定理

Note

The main chapter uses *theorem* 6.3 as an algebraic criterion; a full proof is deferred to the appendix at the end of this chapter for optional reading.

Example 6.7. Applying the Sufficient Condition. Locate the singular solutions of the equation

$$[(y-1)p]^2 = ye^{xy}, \quad \left(p = \frac{dy}{dx}\right). \tag{6.22}$$

範例

Solution

We define $F(x, y, p) = (y-1)^2p^2 - ye^{xy}$. The p-discriminant requires $F_p = 0$:

$$F_p(x, y, p) = 2p(y-1)^2 = 0.$$

This implies either $p = 0$ or $y = 1$. If $y = 1$, substituting into $F = 0$ gives $0 - 1 \cdot e^x = 0$, which has no real solution. If $p = 0$, substituting into $F = 0$ requires $-ye^{xy} = 0$, which strictly forces $y = 0$.
We test the candidate $y = 0$. Its derivative is $p = 0$. Substituting $y = 0, p = 0$ into (6.22) yields $0 = 0$, confirming it is a valid solution. We

now evaluate the necessary higher derivatives along $y = 0, p = 0$:

$$F_y(x, y, p) = 2p^2(y-1) - e^{xy} - xye^{xy} \implies F_y(x, 0, 0) = -1 \neq 0,$$
$$F_{pp}(x, y, p) = 2(y-1)^2 \implies F_{pp}(x, 0, 0) = 2 \neq 0.$$

Because $F_y \neq 0$ and $F_{pp} \neq 0$ identically along the curve, *theorem* 6.3 applies. Thus, $y = 0$ is the unique singular solution to the differential equation.

■

6.4 *Envelopes and the* C*-Discriminant*

To fully elucidate the relationship between general solutions and singular solutions, we borrow the concept of an envelope from differential geometry. As established in *theorem* 6.2, the p-discriminant identifies where the differential equation itself possesses multiple tangent directions. Conversely, the C-discriminant operates on the general solution, identifying the geometric envelope of the integral curves.

Consider a one-parameter family of curves in the plane, given implicitly by:

$$K(C): \quad V(x, y, C) = 0, \tag{6.23}$$

where V is a continuously differentiable function of its three arguments on some domain. For instance, $x^2 + y^2 = C$ describes a family of concentric circles, whilst $y - (x - C)^2 = 1$ describes a family of horizontally translated parabolas.

Definition 6.2. ***Envelope of a Family of Curves.***
Let Γ be a continuously differentiable curve. If, at every point $q \in \Gamma$, there exists a curve $K(C^*)$ from the family (6.23) that passes through q and is tangent to Γ at q, and $K(C^*)$ is not locally identical to Γ in any neighbourhood of q, then Γ is termed an **envelope** of the family.

定義

Note

This definition departs slightly from the strictest classical differential geometry formulation, which typically demands that *every* branch of the curve family contributes to the envelope. Our relaxed definition permits families that contain intersecting sub-families (as seen in example 6.9), which is essential for analysing the complete general integrals of non-linear differential equations.

Theorem 6.4. ***Envelopes are Singular Solutions.***
Suppose the differential equation $F(x, y, y') = 0$ admits a general integral $V(x, y, C) = 0$. If this family of integral curves possesses an en-

velope $\Gamma : y = \varphi(x)$ defined on an interval J, then Γ is a singular solution of the differential equation.

定理

Proof

By the geometric definition of an envelope, we must first verify that Γ is a valid solution to the differential equation.
Let (x_0, y_0) be an arbitrary point on Γ, such that $y_0 = \varphi(x_0)$. By *definition* 6.2, there exists a member of the general integral family, say $y = u(x, C_0)$, that is tangent to Γ at (x_0, y_0). Tangency implies the functions share both a coordinate and a first derivative at this point:

$$\varphi(x_0) = u(x_0, C_0), \quad \text{and} \quad \varphi'(x_0) = u_x(x_0, C_0).$$

Because $y = u(x, C_0)$ is, by definition, a solution to the differential equation, it satisfies:

$$F(x_0, u(x_0, C_0), u_x(x_0, C_0)) = 0.$$

Substituting the tangency conditions yields $F(x_0, \varphi(x_0), \varphi'(x_0)) = 0$. As $x_0 \in J$ was chosen arbitrarily, the envelope $y = \varphi(x)$ satisfies the differential equation everywhere on J.
Finally, because the envelope is everywhere tangent to a distinct member of the general family, uniqueness of the initial value problem fails at every point on Γ. Thus, Γ is a singular solution.

■

Consequently, the search for a singular solution reduces to the purely algebraic problem of finding the envelope of the general integral. We now establish the necessary conditions for a curve to be an envelope.

Theorem 6.5. ***The C-Discriminant (Necessary Condition).***
If Γ is an envelope of the family $V(x, y, C) = 0$ and can be smoothly parameterised by the family parameter C, then Γ necessarily satisfies the simultaneous equations:

$$V(x, y, C) = 0, \quad \frac{\partial V}{\partial C}(x, y, C) = 0. \tag{6.24}$$

The algebraic relation $\Omega(x, y) = 0$ obtained by eliminating C is called the C-**discriminant curve**.

定理

Proof

By hypothesis, the envelope Γ can be parameterised smoothly by the parameter C of the curves it touches:

$$\Gamma : \quad x = f(C), \quad y = g(C), \quad (C \in I).$$

Because the point $(f(C), g(C))$ lies on the curve $K(C)$, it satisfies the family equation identically for all C:

$$V(f(C), g(C), C) = 0. \tag{6.25}$$

Differentiating this identity with respect to C using the chain rule yields:

$$V_x f'(C) + V_y g'(C) + V_C = 0, \tag{6.26}$$

where the partial derivatives of V are evaluated at $(f(C), g(C), C)$. We now exploit the tangency condition. The tangent vector to the envelope Γ at C is $\mathbf{t}_\Gamma = (f'(C), g'(C))$. The normal vector to the curve $V(x, y, C) = 0$ at this same point is $\mathbf{n}_K = (V_x, V_y)$. Because the curves are tangent, the tangent of Γ must be orthogonal to the normal of $K(C)$. Therefore, their dot product vanishes:

$$\mathbf{n}_K \cdot \mathbf{t}_\Gamma = V_x f'(C) + V_y g'(C) = 0.$$

Substituting this orthogonality relation into (6.26) leaves strictly $V_C(f(C), g(C), C) = 0$. Since (6.25) also holds, the theorem is proved.

■

As with the p-discriminant, satisfying the C-discriminant is merely a necessary condition. The locus $\Omega(x, y) = 0$ may contain envelopes, but also loci of nodes, cusps, or isolated points of the family. A sufficient condition guarantees the extraction of the true envelope.

Theorem 6.6. ***Sufficient Conditions for an Envelope.***
Suppose the system $V = 0, V_C = 0$ defines a continuously differentiable curve $\Lambda : x = \varphi(C), y = \psi(C)$ that is not a member of the family $V(x, y, C) = 0$. If the non-degeneracy conditions

$$(\varphi'(C), \psi'(C)) \neq (0, 0), \quad \text{and} \quad (V_x, V_y) \neq (0, 0) \tag{6.27}$$

hold along Λ, then Λ is an envelope of the family.

定理

Proof

Pick an arbitrary point $q(C) = (\varphi(C), \psi(C))$ on Λ. By definition, we have $V(\varphi, \psi, C) = 0$ and $V_C(\varphi, \psi, C) = 0$.
Because the spatial gradient (V_x, V_y) does not vanish, the Implicit Function Theorem ensures that $V(x, y, C) = 0$ locally defines a smooth curve Γ_C near $q(C)$, with a well-defined normal vector $\mathbf{n}_K = (V_x, V_y)$. The tangent vector to the candidate envelope Λ is $\mathbf{t}_\Lambda = (\varphi'(C), \psi'(C))$, which is non-zero by the first non-degeneracy condition.

Differentiating the identity $V(\varphi(C), \psi(C), C) = 0$ entirely with respect to C gives:

$$V_x \varphi'(C) + V_y \psi'(C) + V_C = 0.$$

Since Λ lies on the C-discriminant locus, $V_C = 0$. The equation reduces to $V_x \varphi'(C) + V_y \psi'(C) = 0$, or $\mathbf{n}_K \cdot \mathbf{t}_\Lambda = 0$. This proves that the tangent vector of Λ is orthogonal to the normal vector of Γ_C. Thus, Λ is tangent to the curve Γ_C at $q(C)$. Because Λ is explicitly assumed not to be a member of the family itself, it fulfils *definition* 6.2 and constitutes an envelope.

■

Example 6.8. Nodal Loci vs. Envelopes. Determine the singular solutions of the differential equation

$$(y-1)^2 \left(\frac{dy}{dx}\right)^2 = \frac{4}{9}y. \tag{6.28}$$

範例

Solution

By separating variables and integrating, one obtains the general integral:

$$V(x, y, C) = (x-C)^2 - y(y-3)^2 = 0, \quad (C \in \mathbb{R}). \tag{6.29}$$

To find the envelope, we construct the C-discriminant system. We compute V_C:

$$V_C(x, y, C) = -2(x-C) = 0.$$

This implies $x = C$. Substituting $x = C$ back into $V = 0$ yields:

$$-y(y-3)^2 = 0.$$

This algebraic relation provides two distinct candidate curves: the line $\Lambda_1 : y = 0$ and the line $\Lambda_2 : y = 3$. We can parameterise both natively by C:

$$\Lambda_1 : x = C, y = 0 \quad \text{and} \quad \Lambda_2 : x = C, y = 3.$$

We apply the sufficient conditions (*theorem* 6.6). For both curves, the tangent vector is $(x'(C), y'(C)) = (1, 0) \neq (0, 0)$. The gradient of V is:

$$V_x = 2(x-C), \quad V_y = -(y-3)^2 - 2y(y-3).$$

Testing Λ_1 **(**$y = 0$**):** At $x = C, y = 0$, the gradient is $(V_x, V_y) = (0, -9) \neq (0, 0)$. The non-degeneracy conditions are met. Thus, $y = 0$ is an envelope, and by *theorem* 6.4, it is a singular solution. (Direct substitution into (6.28) trivially confirms $y = 0$ is a solution).

Testing Λ_2 ***($y = 3$):*** At $x = C, y = 3$, the gradient is $(V_x, V_y) = (0, 0)$. The non-degeneracy condition fails. Consequently, *theorem* 6.6 is silent. However, we can test $y = 3$ in the original differential equation (6.28). Its derivative is $y' = 0$. Substituting yields $(2)^2(0)^2 = \frac{4}{9}(3)$, or $0 = \frac{4}{3}$, which is a contradiction. Thus, $y = 3$ is not a solution, and cannot be a singular solution.

Geometrically, the curves $(x - C)^2 = y(y-3)^2$ exhibit self-intersections (nodes) along the line $y = 3$. The C-discriminant extracts both the geometric envelope ($y = 0$) and the locus of nodes ($y = 3$).

■

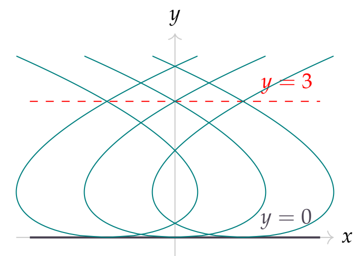

Figure 6.5: The family $(x - C)^2 = y(y-3)^2$. The C-discriminant identifies the envelope $y = 0$ and the nodal locus $y = 3$. Only $y = 0$ is a singular solution.

Example 6.9. Intersecting Sub-families. Find the singular solutions to the equation

$$\left(\frac{dy}{dx}\right)^4 - \left(\frac{dy}{dx}\right)^3 - y^2\frac{dy}{dx} + y^2 = 0. \tag{6.30}$$

範例

Solution

Let $p = y'$. The equation factors neatly by grouping:

$$p^3(p-1) - y^2(p-1) = 0 \implies (p^3 - y^2)(p-1) = 0.$$

This splits the differential equation into two distinct branches: $p = 1$ and $p^3 = y^2$. Integrating $y' = 1$ gives the family of lines $y = x - C_1$. Integrating $y^{-2/3}dy = dx$ gives $3y^{1/3} = x - C_2$, or $y = \frac{1}{27}(x - C_2)^3$. We can unify these into a single general integral by taking $C_1 = C_2 = C$:

$$V(x, y, C) = \left[y - \frac{1}{27}(x-C)^3\right][y - (x - C)] = 0. \tag{6.31}$$

We apply the C-discriminant. Setting $V_C = 0$ requires differentiating the product:

$$V_C = \frac{1}{9}(x-C)^2[y - (x-C)] + \left[y - \frac{1}{27}(x-C)^3\right](1) = 0.$$

We must solve the system $V = 0$ and $V_C = 0$. From $V = 0$, either the cubic term or the linear term is zero. If $y - (x - C) = 0$, substituting into $V_C = 0$ leaves $y - \frac{1}{27}(x-C)^3 = 0$. For both brackets to vanish simultaneously, $(x - C) = \frac{1}{27}(x-C)^3$, so

$$x - C = 0,\ \pm 3\sqrt{3},$$

and therefore

$$y = 0,\ \pm 3\sqrt{3}.$$

If $y - \frac{1}{27}(x-C)^3 = 0$, substituting into $V_C = 0$ leaves:

$$\frac{1}{9}(x-C)^2\,[y-(x-C)] = 0.$$

This gives either $x = C$ (hence $y = 0$), or $y = x - C$, which again yields $y = 0, \pm 3\sqrt{3}$ from the previous compatibility condition. Thus, the C-discriminant locus contains the three horizontal lines

$$y = 0, \quad y = 3\sqrt{3}, \quad y = -3\sqrt{3}.$$

We now test them in (6.30). For a constant solution $y = a$, we have $y' = 0$, and substitution gives $a^2 = 0$. Hence only $a = 0$ is admissible: $y = \pm 3\sqrt{3}$ are discriminant artifacts (intersection loci of sub-families), not solutions. Therefore, the only singular-solution candidate from the C-discriminant that survives the equation check is $y = 0$. Taking the derivative gives $y' = 0$, and substituting into (6.30) yields $0 - 0 - 0 + 0 = 0$. Hence, $y = 0$ is a valid solution. It acts as the envelope for the family of cubics $y = \frac{1}{27}(x-C)^3$, while the linear family $y = x - C$ simply passes through it transversely. By our relaxed *definition* 6.2, $y = 0$ touches a curve of the family tangentially at every point, destroying uniqueness. It is therefore a singular solution.

■

6.5 *Exercises*

1. **Differentiation Method.** Solve the following differential equations:

   (a) $2y = p^2 + 4px + 2x^2, \quad p = \dfrac{dy}{dx}.$

   (b) $y = px\ln x + (xp)^2.$

   (c) $2xp = 2\tan y + p^3\cos^2 y.$

2. **Parametric Method.** Use the parametric method to solve the following differential equations:

   (a) $2y^2 + 5\left(\dfrac{dy}{dx}\right)^2 = 4.$

   (b) $x^2 - 3\left(\dfrac{dy}{dx}\right)^2 = 1.$

   (c) $\left(\dfrac{dy}{dx}\right)^2 + y - x^2 = 0.$

   (d) $x^3 + \left(\dfrac{dy}{dx}\right)^3 = 4x\dfrac{dy}{dx}.$

3. **The p-Discriminant.** Use the p-discriminant to find singular solutions of the following differential equations:

   (a) $y = x\dfrac{dy}{dx} + \left(\dfrac{dy}{dx}\right)^2$.

   (b) $y = 2x\dfrac{dy}{dx} + \left(\dfrac{dy}{dx}\right)^2$.

   (c) $(y-1)^2\left(\dfrac{dy}{dx}\right)^2 = \dfrac{4}{9}y$.

4. **Necessity of the Non-Degeneracy Conditions.** Give an example to show that the two non-vanishing conditions in eq. (6.21) are indispensable.

5. **Necessity of $F_p = 0$.** Study the following example to show that the vanishing condition $F_p = 0$ in eq. (6.21) is indispensable:

$$y = 2x + y' - \frac{1}{3}(y')^3.$$

   Let the continuous function $E(y)$ satisfy

$$E(0) = 0, \qquad E(y) \neq 0 \text{ for } 0 < y \leq 1.$$

   Prove that $y = 0$ is a singular solution of

$$\frac{dy}{dx} = E(y)$$

   if and only if the integral

$$\int_0^1 \frac{dy}{E(y)}$$

   converges. (A similar integral criterion appeared in the starred Osgood uniqueness exercise in Chapter 1; there it concerns uniqueness, while here it concerns singularity.)

6. **Clairaut Revisited.** Find the general solution of Clairaut's equation and determine its envelope.

7. **Inverse Construction.** Construct a first-order differential equation whose singular solution is $y = \sin x$.

6.6 ⋆ *Proof of the Sufficient Condition for Singularity*

In *theorem* 6.3, we asserted without proof that a solution $y = \psi(x)$ derived from the p-discriminant is a singular solution of the differential equation

$$F(x, y, p) = 0, \quad \left(p = \frac{dy}{dx}\right) \tag{6.32}$$

provided the following strict non-degeneracy conditions hold along the curve for all $x \in J$:

$$F_p = 0, \quad F_y \neq 0, \quad F_{pp} \neq 0, \tag{6.33}$$

where all partial derivatives are evaluated at $(x, \psi(x), \psi'(x))$.
This appendix provides a rigorous proof of this existence theorem. The argument relies on a perturbation method: we shift the origin to the candidate curve and construct an intersecting family of distinct integral curves using the Implicit Function Theorem and Peano's Existence Theorem.

Proof

By hypothesis, $y = \psi(x)$ is a valid solution to (6.32), meaning:

$$F(x, \psi(x), \psi'(x)) = 0 \quad \text{for all } x \in J. \tag{6.34}$$

Additionally, because $y = \psi(x)$ is derived from the p-discriminant, it satisfies:

$$F_p(x, \psi(x), \psi'(x)) = 0 \quad \text{for all } x \in J. \tag{6.35}$$

Step 1: Shifting the Origin. We introduce a change of variables to centre our analysis on the candidate curve. Let $u(x)$ represent the deviation from $\psi(x)$:

$$y(x) = \psi(x) + u(x).$$

The derivative shifts correspondingly: $p(x) = \psi'(x) + q(x)$, where $q = \frac{du}{dx}$. Substituting these into the original differential equation defines a new function $H(x, u, q)$:

$$H(x, u, q) = F(x, \psi(x) + u, \psi'(x) + q) = 0. \tag{6.36}$$

Because F is assumed to be twice continuously differentiable, H inherits this regularity in a neighbourhood of $(x, 0, 0)$.

Under this transformation, the candidate solution $y = \psi(x)$ corresponds precisely to the trivial solution $u \equiv 0$, $q \equiv 0$. Equations (6.34) and (6.35) translate to:

$$H(x, 0, 0) = 0, \quad \text{and} \quad H_q(x, 0, 0) = 0 \quad \text{for all } x \in J. \tag{6.37}$$

Because these identities hold identically for all x, their total derivatives with respect to x must also vanish. Differentiating $H(x, 0, 0) = 0$ gives $H_x(x, 0, 0) = 0$. Differentiating $H_q(x, 0, 0) = 0$ gives $H_{qx}(x, 0, 0) = 0$. Thus:

$$H_x = H_{xx} = H_{xq} = 0 \quad \text{at } (x, 0, 0). \tag{6.38}$$

The strict non-degeneracy conditions (6.33) translate directly to:

$$H_u(x, 0, 0) \neq 0, \quad \text{and} \quad H_{qq}(x, 0, 0) \neq 0 \quad \text{for all } x \in J. \tag{6.39}$$

Step 2: The Implicit Function Theorem. Fix an arbitrary point $x_0 \in J$. Because $H(x_0, 0, 0) = 0$ and $H_u(x_0, 0, 0) \neq 0$, the Implicit Function Theorem guarantees the existence of a neighbourhood V of $(x_0, 0)$ in the xq-plane, and a unique, twice continuously differentiable function $u = \phi(x, q)$ defined on V, such that:

$$H(x, \phi(x, q), q) \equiv 0, \quad \text{and} \quad \phi(x_0, 0) = 0. \tag{6.40}$$

We determine the partial derivatives of ϕ at $(x, 0)$ by differentiating the identity $H(x, \phi(x, q), q) = 0$.

- Differentiating with respect to x: $H_x + H_u\phi_x = 0$. Since $H_x(x, 0, 0) = 0$ and $H_u \neq 0$, we conclude $\phi_x(x, 0) = 0$. Differentiating again confirms $\phi_{xx}(x, 0) = 0$.
- Differentiating with respect to q: $H_u\phi_q + H_q = 0$. Since $H_q(x, 0, 0) = 0$, we conclude $\phi_q(x, 0) = 0$. Differentiating with respect to x confirms $\phi_{qx}(x, 0) = 0$.
- Differentiating twice with respect to q: $H_u\phi_{qq} + H_{uu}(\phi_q)^2 + 2H_{uq}\phi_q + H_{qq} = 0$. Evaluating at $q = 0$ where $\phi_q = 0$ leaves $H_u\phi_{qq} + H_{qq} = 0$. Thus:

$$\phi_{qq}(x, 0) = -\frac{H_{qq}(x, 0, 0)}{H_u(x, 0, 0)} \neq 0. \tag{6.41}$$

Step 3: Constructing the Differential Equation for q. We seek a non-trivial solution $u(x)$ by treating q as an unknown function of x. Recalling $q = \frac{du}{dx}$, we apply the chain rule to $u = \phi(x, q)$:

$$q = \phi_x(x, q) + \phi_q(x, q)\frac{dq}{dx}.$$

Rearranging this yields a first-order explicit differential equation for q:

$$\frac{dq}{dx} = \frac{q - \phi_x(x, q)}{\phi_q(x, q)} \equiv h(x, q). \tag{6.42}$$

We must evaluate the behaviour of $h(x, q)$ as $q \to 0$. Because $\phi_x(x, 0) = 0$ and $\phi_q(x, 0) = 0$, the function h presents an indeterminate form of type $0/0$ at $q = 0$. We apply L'Hôpital's Rule with respect to q:

$$\lim_{q \to 0} h(x, q) = \lim_{q \to 0} \frac{1 - \phi_{xq}(x, q)}{\phi_{qq}(x, q)} = \frac{1 - 0}{\phi_{qq}(x, 0)} = \frac{1}{\phi_{qq}(x, 0)} \equiv \tilde{h}(x).$$

Because $\phi_{qq}(x, 0) \neq 0$ by (6.41), $\tilde{h}(x)$ is well-defined, continuous, and non-zero. Specifically, $\tilde{h}(x_0) \neq 0$. We define a continuous extension $s(x, q)$:

$$s(x, q) = \begin{cases} h(x, q) & \text{if } q \neq 0, \\ \tilde{h}(x) & \text{if } q = 0. \end{cases}$$

Step 4: Integration and Tangency. Consider the initial value problem for $q(x)$:

$$\frac{dq}{dx} = s(x, q), \quad q(x_0) = 0. \tag{6.43}$$

Because $s(x, q)$ is continuous in a neighbourhood of $(x_0, 0)$, Peano's Existence Theorem (*theorem* 4.4) guarantees at least one solution $q = r(x)$ on some small interval $\tilde{J} \subset J$ around x_0.

We construct the corresponding deviation function $u = \xi(x) = \phi(x, r(x))$. By the definition of ϕ in (6.40), this function satisfies $H(x, \xi(x), \xi'(x)) = 0$ on $\tilde{J}$. Translating back to our original variables, the curve

$$\tilde{y}(x) = \psi(x) + \xi(x)$$

is a valid solution to the original differential equation (6.32).

We now prove that $\tilde{y}(x)$ is tangent to the candidate curve $\psi(x)$ at x_0, but is geometrically distinct from it. At x_0, we have $r(x_0) = 0$. Therefore:

$$\xi(x_0) = \phi(x_0, 0) = 0.$$

The derivative is $\xi'(x) = r(x)$, so at x_0:

$$\xi'(x_0) = r(x_0) = 0.$$

This proves that $\tilde{y}(x_0) = \psi(x_0)$ and $\tilde{y}'(x_0) = \psi'(x_0)$; the two solutions are tangent at x_0. To demonstrate they are distinct, we examine the second derivative. By (6.43), $r'(x_0) = s(x_0, 0) = \tilde{h}(x_0)$. Since $\tilde{h}(x) \neq 0$, we have $r'(x_0) \neq 0$. The second derivative of ξ at x_0 is:

$$\xi''(x_0) = r'(x_0) \neq 0.$$

Because $\xi''(x_0) \neq 0$, the deviation $\xi(x)$ cannot be identically zero in any neighbourhood of x_0. Thus, $\tilde{y}(x)$ is a distinct integral curve that touches $\psi(x)$ tangentially at x_0.

Since $x_0 \in J$ was chosen arbitrarily, every point on $y = \psi(x)$ possesses a distinct, tangent integral curve. By *definition* 6.1, $y = \psi(x)$ is a singular solution.

■

7 Higher-Order Differential Equations

The analysis of dynamic systems frequently generates differential equations containing a function and its derivatives up to order n. The integer n, termed the **order** of the differential equation, serves as a natural measure of the system's geometric and analytic complexity. A primary strategic goal when faced with an n-th order equation is **order reduction**. If an n-th order equation can be systematically reduced to an $(n-1)$-th order equation, we advance the problem closer to a first-order system, which may be analysed using the existence and uniqueness frameworks previously established. In this chapter, we exploit symmetries—specifically, the absence of the independent variable—to achieve this reduction. This algebraic reduction naturally motivates the geometric study of the **phase plane**, allowing us to extract qualitative behaviour even when the resulting integrals are analytically intractable.

7.1 Order Reduction for Autonomous Equations

A differential equation that does not explicitly depend on the independent variable (typically time, t, or spatial coordinate, x) is termed **autonomous**. Such systems arise universally in conservative physical models where the governing laws are invariant under time translations.

Proposition 7.1. ***Reduction of Autonomous Equations.***
Let

$$F\left(y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}\right) = 0$$

be an n-th order autonomous differential equation. Setting $z = \frac{dy}{dx}$ gives

$$\frac{d^2y}{dx^2} = z\frac{dz}{dy}, \quad \frac{d^3y}{dx^3} = z^2\frac{d^2z}{dy^2} + z\left(\frac{dz}{dy}\right)^2, \quad \dots, \quad \frac{d^ny}{dx^n} = \varphi\left(z, \frac{dz}{dy}, \dots, \frac{d^{n-1}z}{dy^{n-1}}\right).$$

Hence the equation reduces to an $(n-1)$-th order equation with independent variable y and unknown $z(y)$:

$$F_1\left(y, z, \frac{dz}{dy}, \ldots, \frac{d^{n-1}z}{dy^{n-1}}\right) = 0.$$

命題

Proof

By the chain rule,

$$\frac{d^2y}{dx^2} = \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = z\frac{dz}{dy}.$$

Differentiating once more,

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(z\frac{dz}{dy}\right) = z\frac{d}{dy}\left(z\frac{dz}{dy}\right) = z^2\frac{d^2z}{dy^2} + z\left(\frac{dz}{dy}\right)^2.$$

Higher derivatives have the stated form by iteration, so substituting removes x and yields the reduced equation.

■

Note

Rigorously, treating z as a function of y is a local step on intervals where $y'(x) \neq 0$. Constant branches with $y' \equiv 0$ are handled separately as equilibrium solutions.

Second-Order Systems and Phase Plane Analysis

The most prominent application of *proposition* 7.1 is the second-order autonomous equation, which dictates the motion of Newtonian particles in conservative force fields:

$$\frac{d^2x}{dt^2} = f(x). \tag{7.1}$$

Applying the reduction $v = \frac{dx}{dt}$, we obtain the first-order equation

$$v\frac{dv}{dx} = f(x). \tag{7.2}$$

This equation is separable. Integrating both sides with respect to x yields a **first integral** of the motion:

$$\frac{1}{2}v^2 = F(x) - \frac{1}{2}C_1 \implies v^2 = 2F(x) - C_1, \tag{7.3}$$

where $F(x)$ is an antiderivative of $f(x)$ and C_1 is an arbitrary constant of integration.

Solving (7.3) for v provides a separable first-order differential equation for $x(t)$:

$$\frac{dx}{dt} = \pm\sqrt{2F(x) - C_1}.$$

Integrating this yields the general implicit solution

$$\int \frac{dx}{\pm\sqrt{2F(x) - C_1}} = t + C_2. \tag{7.4}$$

In practice, evaluating the integral in (7.4) and inverting it to find $x(t)$ explicitly is often impossible; for instance, if $F(x)$ is a cubic polynomial, the integral becomes a non-elementary elliptic integral. However, complete explicit integration is rarely necessary to understand the system's dynamics.

Instead, we examine the relationship between displacement x and velocity v directly using (7.3). For a fixed constant C_1, the equation $v^2 = 2F(x) - C_1$ defines a curve (or set of curves) Γ_{C_1} in the xv-plane. The xv-plane is called the **phase plane**, and the family of curves representing all possible states of the system is the **phase portrait**. A trajectory that separates regions with qualitatively different motion is called a **separatrix**.

Example 7.1. The Harmonic Oscillator. Consider the linear restoring force $f(x) = -x$, corresponding to the differential equation

$$\frac{d^2x}{dt^2} + x = 0.$$

範例

Solution

Applying the first integral (7.3) with $F(x) = -\frac{1}{2}x^2$, we have:

$$v^2 + x^2 = -C_1.$$

For real solutions, we must have $C_1 \leq 0$. Setting $C_1 = -C^2$ for $C \geq 0$, the trajectories Γ_C are concentric circles $x^2 + v^2 = C^2$. The origin $(0, 0)$ represents the stationary trivial solution. Because $v = \frac{dx}{dt}$, x must increase when $v > 0$. Thus, the trajectories are traversed clockwise. The closed nature of the curves indicates that the motion is strictly periodic (see *figure* 7.1, left).

■

Example 7.2. The Repulsive Oscillator. Consider the repulsive force $f(x) = x$, corresponding to

$$\frac{d^2x}{dt^2} - x = 0.$$

範例

Solution

Here, $F(x) = \frac{1}{2}x^2$, leading to the first integral:

$$v^2 - x^2 = -C_1.$$

This describes a family of hyperbolas. The motion is not periodic. If $C_1 < 0$, then $v^2 - x^2 > 0$, trajectories cross $x = 0$ with nonzero velocity, and then diverge. If $C_1 > 0$, the curves satisfy $x^2 - v^2 = C_1$, so trajectories remain in either $x > 0$ or $x < 0$ and never cross $x = 0$. The critical case $C_1 = 0$ gives the separatrices $v = \pm x$, which approach the origin asymptotically in one time direction. The origin remains a stationary point, but it is unstable (see *figure 7.1*, right).

■

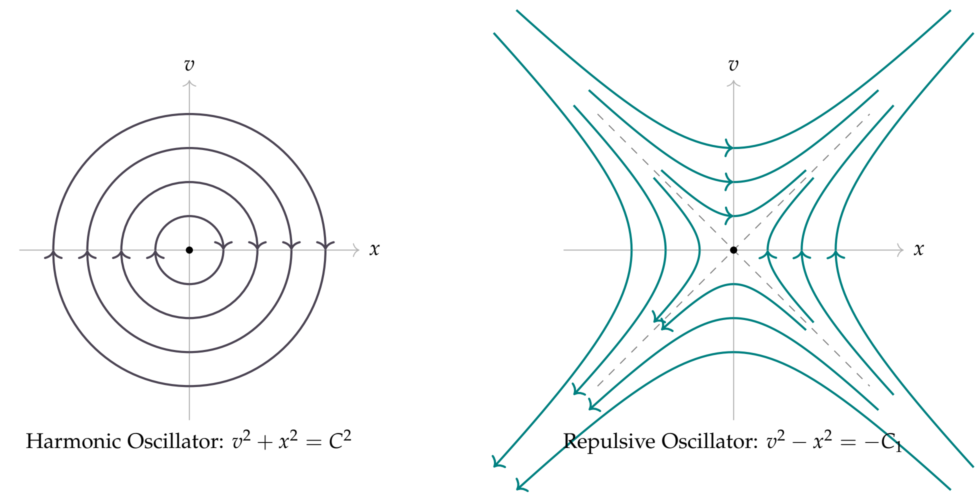

Figure 7.1: Phase portraits for linear autonomous systems. Arrows indicate the direction of forward time flow ($v > 0 \implies x$ increases).

Geometric Construction of Phase Trajectories

When $f(x)$ is highly non-linear, analytic classification of the phase curves becomes difficult. We instead employ a geometric construction.

Consider the auxiliary plane with coordinates (x, u). Plot the curve $u = 2F(x)$. For a specified integration constant C_1, the physically permissible region where $v^2 \geq 0$ corresponds strictly to the domain where

$$2F(x) - C_1 \geq 0.$$

Draw the horizontal line $u = C_1$. The section of the curve $u = 2F(x)$ that lies above this horizontal line is denoted $\Delta^+(C_1)$. For every point $(x, u) \in \Delta^+(C_1)$, the corresponding velocities in the phase plane are mapping symmetrically as

$$v = \pm\sqrt{u - C_1}.$$

This procedure translates the geometric "depth" of the curve $u = 2F(x)$ above C_1 into the vertical amplitude of the phase trajectory Γ_{C_1}. If $\Delta^+(C_1)$ forms a bounded interval, the resulting trajectory is a closed curve (a periodic orbit). As C_1 increases to the local maximum of $2F(x)$, the closed trajectory shrinks to a single point, identifying a stable equilibrium.

The Non-Linear Pendulum

To witness the full power of phase plane analysis paired with exact integration, we examine the classical simple pendulum.
Suppose a mass m is suspended by an inextensible, massless thread of length l from a fixed point. It swings freely in a vertical plane under gravity. Let $x(t)$ be the directed angle between the thread and the downward vertical. The tangential acceleration along the circular path is $l\frac{d^2x}{dt^2}$. Newton's second law yields the equation of motion:

$$m\left(l\frac{d^2x}{dt^2}\right) = -mg\sin x \implies \frac{d^2x}{dt^2} + a^2\sin x = 0, \tag{7.5}$$

where $a = \sqrt{g/l} > 0$.
We multiply the equation by $\frac{dx}{dt}$ to deduce the first integral:

$$\frac{dx}{dt}\frac{d^2x}{dt^2} + a^2\sin x\frac{dx}{dt} = 0 \implies \frac{1}{2}v^2 - a^2\cos x = -\frac{1}{2}C_1.$$

Rearranging, we map the trajectories in the phase plane:

$$v = \pm\sqrt{2a^2\cos x - C_1}. \tag{7.6}$$

Linearization and Isochronism

Before tackling the exact non-linear integral, we consider small amplitude vibrations ($|x| \ll 1$). We invoke the Taylor expansion to first order, $\sin x \approx x$, linearising the differential equation to:

$$\frac{d^2x}{dt^2} + a^2x = 0. \tag{7.7}$$

This is the harmonic oscillator. Its first integral is $v^2 + a^2x^2 = C^2$, and separation of variables yields the general solution:

$$x(t) = A\sin(at + D),$$

where the amplitude A and phase D are arbitrary constants. The time required for one complete swing is precisely $\frac{2\pi}{a}$. Remarkably, this period is entirely independent of the amplitude A. This phenomenon, known as **isochronism**, forms the theoretical basis for pendulum clocks.
However, isochronism is an artifact of the linear approximation. As amplitude increases, the truncation $\sin x \approx x$ fails, and the true period diverges from the linear prediction.

The Exact Non-Linear Period

We return to the non-linear first integral (7.6) to extract the exact period. Suppose the pendulum is released from rest at a maximum angular displacement (amplitude) A, where $0 < A < \pi$. At this peak, $x(t_1) = A$ and $v(t_1) = 0$.
Substituting these conditions into (7.6) determines the constant C_1:

$$0 = 2a^2 \cos A - C_1 \implies C_1 = 2a^2 \cos A.$$

The phase trajectory for this specific oscillation is therefore governed by:

$$\frac{dx}{dt} = \pm a\sqrt{2(\cos x - \cos A)}.$$

The time taken to travel from the lowest point $x = 0$ to the maximum amplitude $x = A$ constitutes exactly one quarter of the full period $T(A)$. Separating variables and integrating gives:

$$\frac{1}{4}T(A) = \int_0^A \frac{dx}{a\sqrt{2(\cos x - \cos A)}}.$$

To evaluate this, we introduce the substitution $x = Au$, which implies $dx = A\,du$. The bounds transform from $[0, A]$ to $[0, 1]$:

$$T(A) = \frac{4A}{a\sqrt{2}} \int_0^1 \frac{du}{\sqrt{\cos(Au) - \cos A}} = \frac{2\sqrt{2}A}{a} \int_0^1 \frac{du}{\sqrt{\cos(Au) - \cos A}}. \tag{7.8}$$

This formula encapsulates the exact period as a non-elementary integral. Using rigorous analytic limits on (7.8), one can prove the boundary behaviours:

$$\lim_{A \to 0} T(A) = \frac{2\pi}{a},$$
$$\lim_{A \to \pi} T(A) = \infty.$$

The first limit rigorously recovers the linear isochronism for infinitesimal vibrations. The second limit proves that as the release angle approaches the vertical upright position (π), the pendulum requires infinite time to complete a cycle, resting asymptotically at the unstable equilibrium.

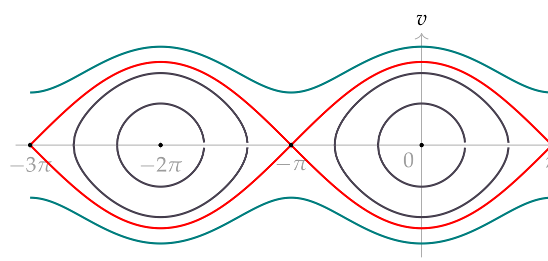

Figure 7.2: Global pendulum phase portrait in normalized units ($a = 1$) over $x \in [-3\pi, 3\pi]$. Closed curves (accent) represent libration, open curves (teal) represent rotation, and the separatrices (red) divide the two regimes.

Global Phase Portrait

Plotting the relation $v = \pm\sqrt{2a^2 \cos x - C_1}$ across all possible values of C_1 yields the global phase portrait of the pendulum (*figure 7.2*), which is strictly periodic in x with period 2π.
We can classify the dynamics by the value of C_1:

1. **Libration (Oscillation):** When $-2a^2 < C_1 < 2a^2$, trajectories are closed loops around stable equilibria $(2k\pi, 0)$, $k \in \mathbb{Z}$.

2. **Precession (Rotation):** When $C_1 < -2a^2$, the radicand is strictly positive for all x. Trajectories are open curves with fixed sign of v, corresponding to continuous rotation.
3. **Separatrix:** When $C_1 = -2a^2$, the curves
$$v = \pm 2a \left|\cos\left(\frac{x}{2}\right)\right|$$
(equivalently, as a set, $v = \pm 2a\cos(x/2)$) connect consecutive unstable equilibria $((2k-1)\pi, 0)$ and $((2k+1)\pi, 0)$, separating libration from rotation.
4. **Stable Equilibria:** When $C_1 = 2a^2$, the only real phase points are $(2k\pi, 0)$, $k \in \mathbb{Z}$.
5. **No Real States:** When $C_1 > 2a^2$, no real trajectory exists because $2a^2 \cos x - C_1 < 0$ for every x.

Example 7.3. The Catenary Equation. Determine the analytic curve assumed by an ideal, perfectly flexible, and inextensible uniform heavy chain suspended between two fixed points P_1 and P_2 under the influence of uniform gravity.

範例

Solution

Let the points lie in the xy-plane with the y-axis directed vertically upwards. We model the chain as a curve $y = y(x)$. Let γ denote the linear weight density of the chain. Consider an arbitrary arc segment between x and $x + \Delta x$. The length of this segment is Δs. The segment is in static equilibrium under three forces: the tension $\mathbf{T}(x)$ at the left endpoint pulling downwards and leftwards, the tension $\mathbf{T}(x + \Delta x)$ at the right endpoint pulling upwards and rightwards, and the gravitational force $\mathbf{W} = (0, -\gamma \Delta s)$ acting strictly downwards.

Resolving the forces into horizontal (H) and vertical (V) components, equilibrium dictates:
$$H(x + \Delta x) - H(x) = 0, \quad V(x + \Delta x) - V(x) - \gamma \Delta s = 0.$$
The first condition implies the horizontal tension is a global constant, $H(x) \equiv H_0$. Dividing the second condition by Δx and taking the limit as $\Delta x \to 0$ yields the differential relation:
$$V'(x) = \gamma \frac{ds}{dx}. \tag{7.9}$$
Because the chain is perfectly flexible, the tension vector $\mathbf{T}$ must be everywhere tangent to the curve. Thus, the slope of the curve

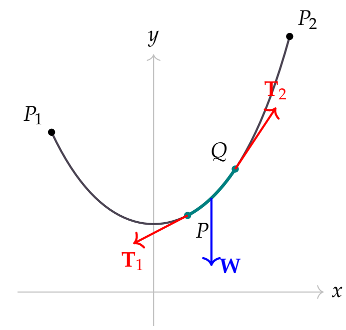

Figure 7.3: A segment PQ of the catenary is in static equilibrium under the tensions $\mathbf{T}_1, \mathbf{T}_2$ at its endpoints and its own weight $\mathbf{W}$.

dictates the ratio of the force components:

$$y'(x) = \frac{V(x)}{H_0} \implies V(x) = H_0 y'(x).$$

Differentiating this and substituting into (7.9), we obtain the governing equation. Noting that the differential arc length is $\frac{ds}{dx} = \sqrt{1 + (y')^2}$, we arrive at the second-order autonomous differential equation:

$$y'' = a\sqrt{1 + (y')^2}, \tag{7.10}$$

where $a = \frac{\gamma}{H_0} > 0$ is a constant.
We reduce the order of (7.10) using the substitution $z = y'$. The equation becomes a separable first-order equation in z:

$$z' = a\sqrt{1 + z^2} \implies \frac{dz}{\sqrt{1 + z^2}} = a\,dx.$$

Integrating both sides yields $\operatorname{arsinh}(z) = a(x + C_1)$, or $z = \sinh(a(x + C_1))$. Replacing z with y' and integrating with respect to x provides the general solution:

$$y(x) = \frac{1}{a}\cosh(a(x + C_1)) + C_2. \tag{7.11}$$

The shape of the suspended chain is therefore a hyperbolic cosine curve, commonly termed a **catenary**.
The constants C_1, C_2 and the physical parameter a are completely determined by the coordinates of the suspension points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ and the total length L of the chain. These manifest as boundary conditions rather than initial conditions. Explicitly, the arc length integral restricts a:

$$\begin{aligned} L &= \int_{x_1}^{x_2} \sqrt{1 + (y')^2}\,dx = \int_{x_1}^{x_2} \cosh(a(x + C_1))\,dx \\ &= \frac{1}{a}\left[\sinh(a(x_2 + C_1)) - \sinh(a(x_1 + C_1))\right]. \end{aligned}$$

Combining this with the altitude difference $y_2 - y_1 = \frac{1}{a}\left[\cosh(a(x_2 + C_1)) - \cosh(a(x_1 + C_1))\right]$ and applying hyperbolic trigonometric identities isolates the parameter a as the unique positive root of a transcendental equation dependent solely on $L, x_2 - x_1$, and $y_2 - y_1$.

■

The reduction of order technique extends naturally to higher-dimensional autonomous systems. We illustrate this through the classical derivation of Keplerian orbits from Newtonian gravity, arguably the most historically significant application of differential equations.

Example 7.4. The Two-Body Problem. Determine the trajectory of a planetary body of mass m_e orbiting a central star of mass m_s under the influence of mutual gravitation.

範例

Solution

Let the star be fixed at the origin O of an inertial reference frame (assuming $m_s \gg m_e$). The position of the planet is given by the vector $\mathbf{r}(t) = (x(t), y(t), z(t))$. By Newton's law of universal gravitation, the attractive force is directed towards the origin with magnitude proportional to the inverse square of the distance $r = |\mathbf{r}|$. Newton's second law yields the equation of motion:

$$m_e\ddot{\mathbf{r}} = -G\frac{m_s m_e}{r^2}\frac{\mathbf{r}}{r} \implies \ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r}, \tag{7.12}$$

where $\mu = Gm_s$. Expanding this into Cartesian components produces a coupled system of three second-order autonomous differential equations (equivalently, a 6-dimensional first-order autonomous system).

We first demonstrate that the motion is strictly planar. Taking the cross product of the position vector $\mathbf{r}$ with the acceleration $\ddot{\mathbf{r}}$ gives:

$$\mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{r} \times \left(-\frac{\mu}{r^3}\mathbf{r}\right) = \mathbf{0}.$$

By the product rule, $\frac{d}{dt}(\mathbf{r} \times \dot{\mathbf{r}}) = \dot{\mathbf{r}} \times \dot{\mathbf{r}} + \mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{0}$. Integrating this yields a constant vector $\mathbf{h}$:

$$\mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h}. \tag{7.13}$$

This first integral represents the conservation of specific angular momentum. Because $\mathbf{r} \cdot \mathbf{h} = \mathbf{r} \cdot (\mathbf{r} \times \dot{\mathbf{r}}) = 0$, the position vector $\mathbf{r}(t)$ is perpetually orthogonal to the constant vector $\mathbf{h}$. Thus, the orbit is confined to a fixed plane. We may align our coordinate axes such that this is the xy-plane ($z \equiv 0$), reducing the problem to a 4-dimensional first-order system in $\mathbb{R}^4$.

We extract a second first integral by taking the dot product of the equation of motion with the velocity vector $\dot{\mathbf{r}}$:

$$\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} = -\frac{\mu}{r^3}(\mathbf{r} \cdot \dot{\mathbf{r}}).$$

Noting that $\frac{d}{dt}(\frac{1}{2}|\dot{\mathbf{r}}|^2) = \dot{\mathbf{r}} \cdot \ddot{\mathbf{r}}$ and $\mathbf{r} \cdot \dot{\mathbf{r}} = r\dot{r}$, the equation becomes:

$$\frac{d}{dt}\left(\frac{1}{2}v^2\right) = -\frac{\mu}{r^2}\dot{r} = \frac{d}{dt}\left(\frac{\mu}{r}\right).$$

Integrating with respect to time yields the conservation of specific energy E:

$$\frac{1}{2}v^2 - \frac{\mu}{r} = E. \tag{7.14}$$

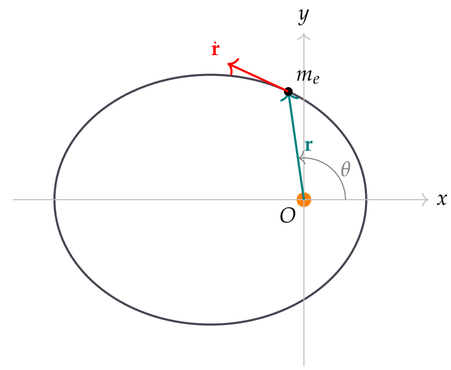

Figure 7.4: The elliptical orbit ($e < 1$) of a planetary body. The strict periodicity of the solution establishes analytic bounds that forbid the planet from spiralling into the central star.

To determine the spatial geometry of the orbit independently of time, we transition to polar coordinates (r, θ). The velocity squared is $v^2 = \dot{r}^2 + (r\dot{\theta})^2$. The conserved planar angular momentum is $h = r^2\dot{\theta}$ (hence $|h|$ is the conserved magnitude). For non-radial trajectories ($h \neq 0$), substituting $\dot{\theta} = h/r^2$ into the energy equation (7.14) gives:

$$\dot{r}^2 + \frac{h^2}{r^2} - \frac{2\mu}{r} = 2E. \tag{7.15}$$

We eliminate time (still assuming $h \neq 0$ and $r > 0$) by applying the chain rule $\dot{r} = \frac{dr}{d\theta}\dot{\theta} = \frac{dr}{d\theta}\frac{h}{r^2}$. Substituting this into (7.15) provides a first-order differential equation for the orbital path $r(\theta)$:

$$\left(\frac{h}{r^2}\frac{dr}{d\theta}\right)^2 = 2E + \frac{2\mu}{r} - \frac{h^2}{r^2}.$$

When $h = 0$, the motion is purely radial ($\dot{\theta} \equiv 0$), so the representation $r = r(\theta)$ is not used and the radial equation is treated separately. Isolating the derivative and completing the square under the radical yields:

$$\frac{h}{r^2}\frac{dr}{d\theta} = \pm\sqrt{2E + \left(\frac{\mu}{h}\right)^2 - \left(\frac{h}{r} - \frac{\mu}{h}\right)^2}.$$

This equation is separable. We divide by the radical and integrate with respect to θ:

$$\int \frac{\frac{h}{r^2}\,dr}{\sqrt{2E + \left(\frac{\mu}{h}\right)^2 - \left(\frac{h}{r} - \frac{\mu}{h}\right)^2}} = \pm\int d\theta.$$

Using the standard integral $\int \frac{-du}{\sqrt{c^2 - u^2}} = \arccos(u/c)$ where $u = \frac{h}{r} - \frac{\mu}{h}$, we obtain:

$$\arccos\left(\frac{\frac{h}{r} - \frac{\mu}{h}}{\sqrt{2E + \left(\frac{\mu}{h}\right)^2}}\right) = \theta - \theta_0.$$

Inverting the arccosine and algebraically solving for r produces the classical polar equation for a conic section:

$$r(\theta) = \frac{p}{1 + e\cos(\theta - \theta_0)}, \tag{7.16}$$

where the semi-latus rectum p and eccentricity e are purely functions of the integration constants (the conserved physical quantities):

$$p = \frac{h^2}{\mu}, \quad e = \sqrt{1 + \frac{2Eh^2}{\mu^2}}.$$

From analytic geometry, (7.16) dictates that the orbit is an ellipse if $e < 1$ (which requires negative total energy $E < 0$), a parabola if $e = 1$ ($E = 0$), and a hyperbola if $e > 1$ ($E > 0$).

■

7.2 Systems of First-Order Differential Equations

The analytic framework developed for a single first-order differential equation extends naturally to systems of equations. This generalisation is structurally crucial, because any higher-order differential equation (or coupled system of higher-order equations) can be systematically reduced to an equivalent first-order system operating in a higher-dimensional linear space.

Reduction to Standard Form

Consider the general n-th order differential equation

$$\frac{d^n y}{dx^n} = F\left(x, y, \frac{dy}{dx}, \dots, \frac{d^{n-1}y}{dx^{n-1}}\right). \tag{7.17}$$

We introduce n new variables corresponding to the unknown function and its derivatives up to order $n - 1$:

$$y_1 = y, \quad y_2 = \frac{dy}{dx}, \quad \dots, \quad y_n = \frac{d^{n-1}y}{dx^{n-1}}.$$

Differentiating these definitions yields an equivalent system of n first-order differential equations:

$$\begin{cases} \frac{dy_1}{dx} = y_2, \\ \frac{dy_2}{dx} = y_3, \\ \vdots \\ \frac{dy_n}{dx} = F(x, y_1, y_2, \dots, y_n). \end{cases} \tag{7.18}$$

The equivalence is absolute: any solution $y = \varphi(x)$ to (7.17) generates a solution tuple $(\varphi, \varphi', \dots, \varphi^{(n-1)})$ to (7.18), and conversely, the first component $y_1(x)$ of any solution to (7.18) satisfies (7.17).
This reduction procedure applies equally to coupled systems of higher-order equations. If a system comprises multiple unknown functions, the dimension of the resulting first-order system equals the sum of the highest derivative orders appearing for each respective function.

Example 7.5. Coupled Higher-Order System. Reduce the following system to standard first-order form:

$$\begin{cases} \frac{d^2u}{dx^2} = F\left(x, u, \frac{du}{dx}, v, w, \frac{dw}{dx}, \frac{d^2w}{dx^2}\right), \\ \frac{dv}{dx} = G\left(x, u, \frac{du}{dx}, v, w, \frac{dw}{dx}, \frac{d^2w}{dx^2}\right), \\ \frac{d^3w}{dx^3} = H\left(x, u, \frac{du}{dx}, v, w, \frac{dw}{dx}, \frac{d^2w}{dx^2}\right). \end{cases}$$

範例

Solution

The sum of the highest derivative orders is $2 + 1 + 3 = 6$. We define the six state variables:

$$y_1 = u, \quad y_2 = \frac{du}{dx}, \quad y_3 = v, \quad y_4 = w, \quad y_5 = \frac{dw}{dx}, \quad y_6 = \frac{d^2w}{dx^2}.$$

The equivalent 6-dimensional first-order system is:

$$\begin{cases} y_1' = y_2, \\ y_2' = F(x, y_1, \ldots, y_6), \\ y_3' = G(x, y_1, \ldots, y_6), \\ y_4' = y_5, \\ y_5' = y_6, \\ y_6' = H(x, y_1, \ldots, y_6). \end{cases}$$

■

Vector Formulation and Normed Spaces

To abstract the theory, we express systems of the form (7.18) using vector notation. Let the n-dimensional column vector be $\mathbf{y} = (y_1, y_2, \ldots, y_n)^\top \in \mathbb{R}^n$, and map the right-hand side functions into a vector field $\mathbf{f}(x, \mathbf{y}) = (f_1(x, \mathbf{y}), \ldots, f_n(x, \mathbf{y}))^\top$. The standard system becomes:

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}). \tag{7.19}$$

The initial value problem appends the condition $\mathbf{y}(x_0) = \mathbf{y}_0$, where $\mathbf{y}_0 \in \mathbb{R}^n$.

To evaluate convergence and limits in $\mathbb{R}^n$, we equip the space with a vector norm $|\cdot|$. Standard choices include:

- Euclidean norm: $|\mathbf{y}|_2 = \sqrt{y_1^2 + \cdots + y_n^2}$,
- Manhattan norm: $|\mathbf{y}|_1 = |y_1| + \cdots + |y_n|$,
- Maximum norm: $|\mathbf{y}|_\infty = \max\{|y_1|, \ldots, |y_n|\}$.

Because all norms on a finite-dimensional vector space are topologically equivalent, the specific choice of norm does not affect the convergence of function sequences or the validity of existence theorems. We assume the standard properties: $|\mathbf{y}| \geq 0$ (with equality if and only if $\mathbf{y} = \mathbf{0}$) and the triangle inequality $|\mathbf{y} + \mathbf{z}| \leq |\mathbf{y}| + |\mathbf{z}|$.

> **Definition 7.1.** ***Lipschitz Condition in $\mathbb{R}^n$.***
> A vector field $\mathbf{f}(x, \mathbf{y})$ satisfies a **Lipschitz condition** with respect to $\mathbf{y}$ on a domain $D \subset \mathbb{R}^{n+1}$ if there exists a constant $L > 0$ such that
>
> $$|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(x, \mathbf{z})| \leq L|\mathbf{y} - \mathbf{z}|$$
>
> for all $(x, \mathbf{y}), (x, \mathbf{z}) \in D$.
>
> 定義

With the topological machinery of $\mathbb{R}^n$ established, the foundational existence and uniqueness theorems translate seamlessly from the scalar case.

> **Theorem 7.1.** ***Picard and Peano Theorems for Systems.***
> Consider the initial value problem $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ with $\mathbf{y}(x_0) = \mathbf{y}_0$. Let $\mathbf{f}$ be continuous on the domain
>
> $$R = \{(x, \mathbf{y}) : |x - x_0| \leq a, \ |\mathbf{y} - \mathbf{y}_0| \leq b\}.$$
>
> 1. **Peano (Existence):** The system admits at least one solution defined on an interval $|x - x_0| \leq h$ for some $h > 0$.
> 2. **Picard (Uniqueness):** If $\mathbf{f}$ additionally satisfies a Lipschitz condition with respect to $\mathbf{y}$ on R, the solution is unique.
>
> 定理

> *Proof*
>
> The proofs are structurally identical to those of the scalar *Peano Existence Theorem* and *Picard's Existence and Uniqueness Theorem*. Absolute values are uniformly replaced by the vector norm $|\cdot|$, and integration of the vector-valued function $\mathbf{f}$ is performed component-wise. The Arzelà-Ascoli theorem (*theorem* 4.3) and the completeness of continuous functions extend natively to $\mathbb{R}^n$. The construction of the Picard iterates
>
> $$\mathbf{y}_{n+1}(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(t, \mathbf{y}_n(t))\,dt$$
>
> proceeds unaltered.
>
> ■

Linear Systems of Differential Equations

A fundamental sub-class of (7.19) arises when the vector field $\mathbf{f}$ is strictly affine with respect to the state variables $\mathbf{y}$.

Definition 7.2. ***Linear System.***
A differential equation system is **linear** if it can be expressed in the form

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y} + \mathbf{e}(x), \tag{7.20}$$

where $A(x) = (a_{ij}(x))_{n\times n}$ is an $n \times n$ matrix-valued function and $\mathbf{e}(x)$ is an n-dimensional column vector-valued function.

定義

In components, (7.20) is

$$\frac{dy_k}{dx} = \sum_{i=1}^{n} a_{ki}(x)y_i + e_k(x), \quad k = 1, \dots, n. \tag{7.21}$$

If one works with row vectors, the same system can be written as

$$\frac{d\mathbf{y}}{dx} = \mathbf{y}\,\widetilde{A}(x) + \widetilde{\mathbf{e}}(x), \tag{7.22}$$

with $\widetilde{A}(x) = (a_{ik}(x))_{n\times n}$ and $\widetilde{\mathbf{e}}(x) = \mathbf{e}(x)^\top$ (transpose indexing convention). In this text we use column vectors:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} e_1(x) \\ e_2(x) \\ \vdots \\ e_n(x) \end{bmatrix},$$

so the standard form remains (7.20).

Proposition 7.2. ***Linear Growth Bound in $\mathbb{R}^n$.***
Let $\mathbf{f}(x, \mathbf{y})$ be continuous on the strip

$$S = \{(x, \mathbf{y}) : \alpha < x < \beta,\ \mathbf{y} \in \mathbb{R}^n\}.$$

Assume there exist continuous non-negative functions $A(x), B(x)$ on (α, β) such that

$$|\mathbf{f}(x, \mathbf{y})| \le A(x)|\mathbf{y}| + B(x)$$

for all $(x, \mathbf{y}) \in S$. Then every solution of $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ exists on the entire interval (α, β).

命題

Proof

The contradiction argument in *theorem* 5.2 carries over verbatim after replacing scalar absolute value by a vector norm and using local existence from *theorem* 7.1. Hence finite-time blow-up cannot occur inside (α, β).

■

Proposition 7.3. ***Global Existence for Linear Systems.***
If the matrix $A(x)$ and the vector $\mathbf{e}(x)$ are continuous on an open interval (a, b), then for any $x_0 \in (a, b)$ and any $\mathbf{y}_0 \in \mathbb{R}^n$, the initial value problem for (7.20) admits a unique solution defined on the entire interval (a, b).

命題

Proof

Fix a vector norm $|\cdot|$ and let $\|\cdot\|$ denote its induced operator norm:

$$\|M\| = \sup_{|\mathbf{y}|=1} |M\mathbf{y}|.$$

By continuity, on any compact sub-interval $J \subset (a, b)$, the functions $\|A(x)\|$ and $|\mathbf{e}(x)|$ are bounded by constants M and K. Therefore

$$|A(x)\mathbf{y} + \mathbf{e}(x)| \leq \|A(x)\||\mathbf{y}| + |\mathbf{e}(x)| \leq M|\mathbf{y}| + K.$$

Applying *proposition* 7.2, solutions cannot blow up in finite time inside (a, b). Together with local existence and uniqueness from *theorem* 7.1, the solution extends uniquely to the full interval (a, b).

■

Example 7.6. A 2×2 Linear System. Solve the initial value problem

$$\begin{cases} y_1' = y_1 + 2y_2 + e^x, \\ y_2' = -y_2 + x, \\ y_1(0) = 1, \quad y_2(0) = 0. \end{cases}$$

Write it in both column-vector and row-vector matrix forms.

範例

Solution

In column form,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} e^x \\ x \end{bmatrix},$$

so the system is $\mathbf{y}' = A\mathbf{y} + \mathbf{e}(x)$. In row form, with $\widetilde{A} = A^\top$,

$$\mathbf{y}' = \mathbf{y}\widetilde{A} + \widetilde{\mathbf{e}}(x), \qquad \widetilde{\mathbf{e}}(x) = \mathbf{e}(x)^\top, \qquad \widetilde{A} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}.$$

We solve directly from the triangular structure. The second equation is independent:

$$y_2' + y_2 = x \implies y_2 = x - 1 + Ce^{-x}.$$

Using $y_2(0) = 0$ gives $C = 1$, hence $y_2(x) = x - 1 + e^{-x}$. Substitute into the first equation:

$$y_1' - y_1 = 2y_2 + e^x = 2x - 2 + 2e^{-x} + e^x.$$

Multiplying by e^{-x},

$$\frac{d}{dx}(e^{-x}y_1) = (2x-2)e^{-x} + 2e^{-2x} + 1.$$

Integrating:

$$e^{-x}y_1 = -2xe^{-x} - e^{-2x} + x + C_1.$$

Therefore $y_1(x) = e^x(x + C_1) - 2x - e^{-x}$. Using $y_1(0) = 1$ gives $C_1 = 2$. Thus the unique solution is

$$\begin{aligned} y_1(x) &= e^x(x+2) - 2x - e^{-x}, \\ y_2(x) &= x - 1 + e^{-x}. \end{aligned}$$

Since A and $\mathbf{e}$ are continuous on $\mathbb{R}$, *proposition* 7.3 also guarantees this solution exists uniquely for all $x \in \mathbb{R}$.

■

7.3 *Exercises*

1. **Estimating g from a Pendulum.** Use the linearized pendulum equation (7.7) to estimate the gravitational acceleration g at your location. Describe the measurement procedure (choice of length, period measurement, and averaging), derive the estimation formula, and discuss the dominant sources of error.

2. **Cubic Pendulum Approximation.** Starting from the non-linear pendulum model (7.5), replace $\sin x$ by the cubic approximation

$$\sin x \approx x - \frac{1}{6}x^3$$

to obtain

$$\frac{d^2x}{dt^2} + a^2\left(x - \frac{1}{6}x^3\right) = 0.$$

In the oscillatory regime ($0 < A < \sqrt{6}$), derive the quarter-period integral for oscillations of amplitude A and use it to prove that the period $T(A)$ depends on A (non-isochrony). Then analyze the corresponding energy level curves near the origin, and explain why this cubic truncation should not be used to infer the global precession (full-rotation) regime of the exact pendulum.

3. **Degenerate Catenary Length.** In the catenary boundary-value problem governed by (7.10) with general profile (7.11), analyze the

limiting case

$$L = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Explain how this case should be handled geometrically and analytically.

4. **Collision and Circular Orbits.** For the two-body equation of motion (7.12), use the first integrals (7.13) and (7.14) to choose integration constants so that:

   (a) the trajectory is a straight line approaching O (collision; radial motion),

   (b) the trajectory is a circle.

   For part (b), relate your constants to the conic form (7.16). For part (a), explain separately why the representation $r = r(\theta)$ is not used.

5. **Standard First-Order Systems.** Rewrite each of the following as a standard first-order system:

   (a) the pendulum equation (7.5),

   (b) the catenary equation (7.10),

   (c) the two-body motion equation (7.12).

6. **Peano Without Picard in $\mathbb{R}^2$.** Consider the IVP

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}), \qquad \mathbf{f}(y_1, y_2) = \left(\sqrt{|y_1|}, -y_2\right), \qquad \mathbf{y}(0) = (0, 0).$$

   (a) Show that $\mathbf{f}$ is continuous but not locally Lipschitz near $(0, 0)$.

   (b) Construct infinitely many distinct solutions.

   (c) Identify the maximal and minimal first components among your solutions, and compare with the scalar non-uniqueness picture from earlier chapters.

7. **A Linear System with a Singular Point.** Consider

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{e}(x), \qquad A(x) = \begin{bmatrix} \frac{1}{x} & 1 \\ -1 & \frac{2}{x} \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} \sin x \\ x^{-2} \end{bmatrix}.$$

   (a) For $x_0 > 0$, use *proposition* 7.3 to justify a unique solution on $(0, \infty)$.

   (b) For $x_0 < 0$, justify a unique solution on $(-\infty, 0)$.

   (c) Explain precisely why this theorem does not give a solution interval that crosses $x = 0$.

8

Continuous Dependence on Initial Conditions and Parameters

Mathematical models of dynamic physical systems inherently rely on empirical measurements to define their initial states and governing parameters. Because any physical measurement possesses a finite boundary of error, a deterministic differential equation model is physically meaningful if and only if small variations in these parameters and initial conditions induce only bounded, strictly controlled deviations in the resulting integral curves. This geometric robustness is termed **continuous dependence**.

In this chapter, we extend the Picard analytic framework established in *theorem 7.1* to prove continuous dependence on both parameters and initial data under the same continuity and Lipschitz hypotheses that guarantee uniqueness.

8.1 Reduction of Initial Conditions to Parameters

Consider a general first-order initial value problem for an n-dimensional system containing an m-dimensional parameter vector $\boldsymbol{\lambda} \in K \subset \mathbb{R}^m$:

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda}), \quad \mathbf{y}(x_0) = \mathbf{y}_0. \tag{8.1}$$

A prototypical example is the linear pendulum equation $\frac{d^2z}{dt^2} + a^2 z = 0$, where the parameter $a = \sqrt{g/l}$ encapsulates the physical constants of gravity and length.

Example 8.1. Harmonic Oscillator Parameters. Solve the initial value problem for the linear pendulum $\frac{d^2z}{dt^2} + a^2 z = 0$ subject to $z(t_0) = z_0$ and $z'(t_0) = v_0$, and observe its parameter dependence.

範例

Solution

The characteristic equation yields purely imaginary roots $\pm ia$. The general solution and its derivative are:

$$z(t) = A\cos(a(t-t_0)) + B\sin(a(t-t_0)),$$
$$z'(t) = -Aa\sin(a(t-t_0)) + Ba\cos(a(t-t_0)).$$

Applying the initial conditions at $t = t_0$ strictly forces $A = z_0$ and $B = \frac{v_0}{a}$. The unique solution is:

$$z(t) = z_0\cos(a(t-t_0)) + \frac{v_0}{a}\sin(a(t-t_0)).$$

The solution $z(t)$ is explicitly a continuous (and continuously differentiable) function of the initial data (t_0, z_0, v_0) and the physical parameter a for all $a \neq 0$.

■

To formalise this continuous dependence for non-linear systems, we first unify the concepts of initial values and parameters. We apply a linear translation to the coordinates in (8.1):

$$t = x - x_0, \quad \mathbf{u} = \mathbf{y} - \mathbf{y}_0.$$

The system transforms into

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(t + x_0, \mathbf{u} + \mathbf{y}_0, \boldsymbol{\lambda}), \quad \mathbf{u}(0) = \mathbf{0}. \tag{8.2}$$

Crucially, the initial conditions $(x_0, \mathbf{y}_0)$ now operate exclusively as parameters within the vector field, whilst the new initial condition is trivially fixed at the origin. Consequently, it suffices to study the dependence of the solution on a generic parameter vector $\boldsymbol{\lambda}$ for the normalised initial value problem $\mathbf{y}(0) = \mathbf{0}$.

Local Continuous Dependence

We construct the analytic proof of continuous dependence by demonstrating that the Picard iterates converge uniformly not merely in the spatial variable x, but jointly in $(x, \boldsymbol{\lambda})$.

Theorem 8.1. ***Local Continuous Dependence on Parameters.***
Let $\mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda})$ be a continuous n-dimensional vector field on the closed domain

$$G: \quad |x| \leq a, \quad |\mathbf{y}| \leq b, \quad |\boldsymbol{\lambda} - \boldsymbol{\lambda}_0| \leq c,$$

satisfying a Lipschitz condition with respect to $\mathbf{y}$ with constant $L > 0$:

$$|\mathbf{f}(x, \mathbf{y}_1, \boldsymbol{\lambda}) - \mathbf{f}(x, \mathbf{y}_2, \boldsymbol{\lambda})| \leq L|\mathbf{y}_1 - \mathbf{y}_2|.$$

If $M > 0$ bounds $|\mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda})|$ on G, and we define $h = \min\left\{a, \frac{b}{M}\right\}$, then the unique solution $\mathbf{y} = \boldsymbol{\varphi}(x, \boldsymbol{\lambda})$ to the problem

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda}), \quad \mathbf{y}(0) = \mathbf{0}$$

is a continuous function of both x and $\boldsymbol{\lambda}$ on the cylinder $D : |x| \leq h, |\boldsymbol{\lambda} - \boldsymbol{\lambda}_0| \leq c$.

定理

Proof

The problem is entirely equivalent to the integral equation:

$$\mathbf{y} = \int_0^x \mathbf{f}(t, \mathbf{y}(t), \boldsymbol{\lambda})\, dt. \tag{8.3}$$

We define the Picard sequence of successive approximations $\boldsymbol{\varphi}_0(x, \boldsymbol{\lambda}) = \mathbf{0}$, and

$$\boldsymbol{\varphi}_{k+1}(x, \boldsymbol{\lambda}) = \int_0^x \mathbf{f}(t, \boldsymbol{\varphi}_k(t, \boldsymbol{\lambda}), \boldsymbol{\lambda})\, dt, \quad (k \geq 0).$$

Because $\mathbf{f}$ is continuous on G, we may prove by induction that each iterate $\boldsymbol{\varphi}_k(x, \boldsymbol{\lambda})$ is continuous jointly in $(x, \boldsymbol{\lambda})$ on D, and that its spatial image remains bounded within $|\boldsymbol{\varphi}_k| \leq M|x| \leq b$.
We now establish the contraction bound. For $k = 0$, we have

$$|\boldsymbol{\varphi}_1(x, \boldsymbol{\lambda}) - \boldsymbol{\varphi}_0(x, \boldsymbol{\lambda})| \leq \left| \int_0^x |\mathbf{f}(t, \mathbf{0}, \boldsymbol{\lambda})|\, dt \right| \leq M|x|.$$

Applying the Lipschitz condition, the inductive step assumes the bound for $k - 1$ and yields:

$$\begin{aligned} |\boldsymbol{\varphi}_{k+1}(x, \boldsymbol{\lambda}) - \boldsymbol{\varphi}_k(x, \boldsymbol{\lambda})| &\leq \left| \int_0^x |\mathbf{f}(t, \boldsymbol{\varphi}_k, \boldsymbol{\lambda}) - \mathbf{f}(t, \boldsymbol{\varphi}_{k-1}, \boldsymbol{\lambda})|\, dt \right| \\ &\leq L \left| \int_0^x |\boldsymbol{\varphi}_k(t, \boldsymbol{\lambda}) - \boldsymbol{\varphi}_{k-1}(t, \boldsymbol{\lambda})|\, dt \right| \\ &\leq L \left| \int_0^x \frac{M}{L} \frac{(L|t|)^k}{k!}\, dt \right| = \frac{M}{L} \frac{(L|x|)^{k+1}}{(k+1)!}. \end{aligned}$$

This sequence of differences is majorised by the terms of the strictly convergent exponential series $\frac{M}{L} e^{Lh}$. By the Weierstrass M-test, the telescopic sum

$$\boldsymbol{\varphi}(x, \boldsymbol{\lambda}) = \sum_{k=0}^{\infty} [\boldsymbol{\varphi}_{k+1}(x, \boldsymbol{\lambda}) - \boldsymbol{\varphi}_k(x, \boldsymbol{\lambda})]$$

converges uniformly on the domain D. The uniform limit of continuous functions is strictly continuous. Hence, the solution limit $\boldsymbol{\varphi}(x, \boldsymbol{\lambda})$ is continuous jointly in x and $\boldsymbol{\lambda}$.

∎

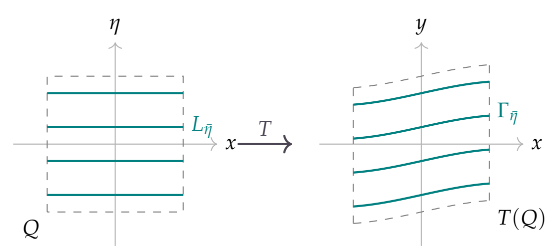

Figure 8.1: The topological transformation T locally straightens the non-linear vector field. The constant coordinate lines $L_{\bar{\eta}}$ in the mapping space map bijectively to the integral curves $\Gamma_{\bar{\eta}}$.

Corollary 8.1. ***Local Straightening of Integral Curves.***
Let $\mathbf{f}(x, \mathbf{y})$ satisfy a local Lipschitz condition near $(x_0, \mathbf{y}_0)$. The mapping

$$T : (x, \boldsymbol{\eta}) \mapsto (x, \boldsymbol{\varphi}(x, \boldsymbol{\eta}))$$

where $\boldsymbol{\varphi}(x, \boldsymbol{\eta})$ is the solution to $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ through $(x_0, \boldsymbol{\eta})$, is a topological homeomorphism from a rectangular neighbourhood Q of $(x_0, \mathbf{y}_0)$ onto its image $T(Q)$.

推論

Proof

By the reduction argument (8.2) and *theorem* 8.1, $\boldsymbol{\varphi}(x, \boldsymbol{\eta})$ is continuous jointly in $(x, \boldsymbol{\eta})$, so T is continuous on Q (chosen as a closed rectangle around $(x_0, \mathbf{y}_0)$). By Picard's uniqueness theorem (*theorem* 7.1), distinct initial conditions $\boldsymbol{\eta}_1 \neq \boldsymbol{\eta}_2$ generate distinct integral curves, hence T is injective and therefore bijective onto $T(Q)$.
It remains to show continuity of T^{-1}. Let $(x_\nu, \mathbf{y}_\nu) = T(x_\nu, \boldsymbol{\eta}_\nu) \to (x_*, \mathbf{y}_*)$ in $T(Q)$. Because Q is compact, $\{(x_\nu, \boldsymbol{\eta}_\nu)\}$ has a convergent subsequence $(x_{\nu_j}, \boldsymbol{\eta}_{\nu_j}) \to (\bar{x}, \bar{\boldsymbol{\eta}}) \in Q$. Continuity of T gives

$$T(\bar{x}, \bar{\boldsymbol{\eta}}) = \lim_{j \to \infty} T(x_{\nu_j}, \boldsymbol{\eta}_{\nu_j}) = (x_*, \mathbf{y}_*).$$

By injectivity, $(\bar{x}, \bar{\boldsymbol{\eta}})$ is the unique preimage of $(x_*, \mathbf{y}_*)$, so every convergent subsequence has the same limit. Hence $(x_\nu, \boldsymbol{\eta}_\nu) \to (\bar{x}, \bar{\boldsymbol{\eta}}) = T^{-1}(x_*, \mathbf{y}_*)$, proving T^{-1} is continuous. Therefore T is a homeomorphism.

■

As illustrated in *figure* 8.1, this topological equivalence allows us to view the local family of non-linear integral curves near a point where $\mathbf{f}$ is continuous and locally Lipschitz in $\mathbf{y}$ as a parallel family of straight lines.

Global Continuous Dependence

The local boundaries of continuous dependence can be systematically extended along any existing valid integral curve. If a solution is known to exist over a large interval, small perturbations to the initial condition will not destroy the solution before the interval concludes.

Theorem 8.2. ***Global Continuous Dependence.***
Let $\mathbf{f}(x, \mathbf{y})$ be continuous on an open domain G and satisfy a local Lipschitz condition with respect to $\mathbf{y}$. Suppose $\mathbf{y} = \boldsymbol{\xi}(x)$ is a specific solution of the differential equation $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ defined on an interval J.
For any compact sub-interval $[a, b] \subset J$, there exists a constant $\delta >$

0 such that for any initial conditions $(x_0, \mathbf{y}_0)$ satisfying

$$a \le x_0 \le b, \quad |\mathbf{y}_0 - \boldsymbol{\xi}(x_0)| \le \delta,$$

the initial value problem

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}), \quad \mathbf{y}(x_0) = \mathbf{y}_0$$

possesses a unique solution $\mathbf{y} = \boldsymbol{\varphi}(x; x_0, \mathbf{y}_0)$ defined across the entire interval $[a, b]$. Furthermore, this solution is jointly continuous with respect to $(x, x_0, \mathbf{y}_0)$ on the closed domain:

$$D_\delta : \quad x \in [a, b], \quad x_0 \in [a, b], \quad |\mathbf{y}_0 - \boldsymbol{\xi}(x_0)| \le \delta.$$

定理

Proof

The curve segment $\Gamma = \{(x, \mathbf{y}) : \mathbf{y} = \boldsymbol{\xi}(x),\ a \le x \le b\}$ is a compact subset of the open domain G. By the finite sub-cover property, there exists a uniform radius $\sigma > 0$ such that the closed tubular neighbourhood

$$E_\sigma = \{(x, \mathbf{y}) : a \le x \le b,\ |\mathbf{y} - \boldsymbol{\xi}(x)| \le \sigma\}$$

is strictly contained within G. Since $\mathbf{f}$ is locally Lipschitz in $\mathbf{y}$ on G, each point of E_σ has a neighbourhood carrying a Lipschitz constant in $\mathbf{y}$. A finite sub-cover of E_σ then yields a single constant L valid on all of E_σ.

We construct a modified Picard sequence parametrised by the variable initial conditions. The crucial structural adaptation is selecting an initial approximation that inherently incorporates the base solution $\boldsymbol{\xi}(x)$:

$$\boldsymbol{\varphi}_0(x; x_0, \mathbf{y}_0) = \mathbf{y}_0 + \boldsymbol{\xi}(x) - \boldsymbol{\xi}(x_0). \tag{8.4}$$

The iteration proceeds normally:

$$\boldsymbol{\varphi}_{k+1}(x; x_0, \mathbf{y}_0) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(t, \boldsymbol{\varphi}_k(t; x_0, \mathbf{y}_0))\, dt. \tag{8.5}$$

We must guarantee that the iterates never exit the tubular neighbourhood E_σ. We select the restrictive initial bound:

$$\delta = \frac{1}{2} e^{-L(b-a)} \sigma. \tag{8.6}$$

Since $e^{-L(b-a)} < 1$, we trivially have $\delta < \sigma$. We prove by induction that for any $(x, x_0, \mathbf{y}_0) \in D_\delta$, the distance from the base curve remains bounded by σ, and the sequential differences converge.

For the base case $k = 0$, (8.4) implies:

$$|\boldsymbol{\varphi}_0(x; x_0, \mathbf{y}_0) - \boldsymbol{\xi}(x)| = |\mathbf{y}_0 - \boldsymbol{\xi}(x_0)| \le \delta < \sigma.$$

Because $\boldsymbol{\xi}(x)$ is an exact solution, it satisfies its own integral equation $\boldsymbol{\xi}(x) = \boldsymbol{\xi}(x_0) + \int_{x_0}^{x} \mathbf{f}(t, \boldsymbol{\xi}(t))\, dt$. Applying this to the difference between $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_0$:

$$\begin{aligned} |\boldsymbol{\varphi}_1(x) - \boldsymbol{\varphi}_0(x)| &= \left| \int_{x_0}^{x} [\mathbf{f}(t, \boldsymbol{\varphi}_0(t)) - \mathbf{f}(t, \boldsymbol{\xi}(t))]\, dt \right| \\ &\le \left| \int_{x_0}^{x} L|\boldsymbol{\varphi}_0(t) - \boldsymbol{\xi}(t)|\, dt \right| \le L|x - x_0||\mathbf{y}_0 - \boldsymbol{\xi}(x_0)|. \end{aligned}$$

Assume inductively that for all $i < k$:

$$|\boldsymbol{\varphi}_{i+1}(x) - \boldsymbol{\varphi}_i(x)| \le \frac{(L|x - x_0|)^{i+1}}{(i+1)!} |\mathbf{y}_0 - \boldsymbol{\xi}(x_0)|.$$

By telescoping the sum of differences up to k, we bound the absolute deviation of the k-th iterate from the base solution:

$$\begin{aligned} |\boldsymbol{\varphi}_k(x) - \boldsymbol{\xi}(x)| &\le \sum_{i=0}^{k-1} |\boldsymbol{\varphi}_{i+1}(x) - \boldsymbol{\varphi}_i(x)| + |\boldsymbol{\varphi}_0(x) - \boldsymbol{\xi}(x)| \\ &\le \sum_{i=0}^{k} \frac{(L|x - x_0|)^i}{i!} |\mathbf{y}_0 - \boldsymbol{\xi}(x_0)| \\ &\le e^{L|x - x_0|} \delta \le e^{L(b-a)} \left(\frac{1}{2} e^{-L(b-a)} \sigma \right) = \frac{\sigma}{2} < \sigma. \end{aligned}$$

This strictly confirms that $\boldsymbol{\varphi}_k(x)$ never exits E_σ, allowing the continued application of the Lipschitz constant L. The standard inductive step for the sequence difference follows immediately.
Because the iterates are uniformly bounded and the bounding series converges, the sequence $\boldsymbol{\varphi}_k(x; x_0, \mathbf{y}_0)$ converges uniformly on D_δ. Consequently, the limit function $\boldsymbol{\varphi}(x; x_0, \mathbf{y}_0)$ is a continuous solution to the initial value problem defined everywhere on the compact interval $[a, b]$.

■

Remark.

theorem 8.2 geometrically extends the "local straightening" of *corollary* 8.1. Rather than being constrained to a microscopic square near one point where $\mathbf{f}$ is continuous and locally Lipschitz in $\mathbf{y}$, the family of integral curves can be topologically straightened within a long, thin tubular domain surrounding the reference trajectory segment.

8.2 *Smooth Dependence and Variational Equations*

While *theorem* 8.1 and *theorem* 8.2 guarantee that the integral curves of a differential equation deform continuously under parameter perturbations, modern dynamical systems theory often requires linearising this deformation. This necessitates that the solution map is not merely continuous, but continuously differentiable (C^1) with respect to its initial data and parameters.

As established previously, arbitrary initial conditions can be absorbed into the system's parameter vector $\boldsymbol{\lambda} \in \mathbb{R}^m$ via a simple coordinate translation. We therefore restrict our analytic proof to the normalised initial value problem parametrised by $\boldsymbol{\lambda}$:

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda}), \quad \mathbf{y}(0) = \mathbf{0}. \tag{8.7}$$

Theorem 8.3. ***Smooth Dependence on Parameters.***
Let $\mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda})$ be an n-dimensional vector field defined on the compact domain

$$G: \quad |x| \leq a, \quad |\mathbf{y}| \leq b, \quad |\boldsymbol{\lambda} - \boldsymbol{\lambda}_0| \leq c.$$

If $\mathbf{f}$ is continuous and possesses continuous partial derivatives with respect to $\mathbf{y}$ (the Jacobian $J_{\mathbf{y}}\mathbf{f}$) and $\boldsymbol{\lambda}$ (the Jacobian $J_{\boldsymbol{\lambda}}\mathbf{f}$) uniformly on G, then the unique solution $\mathbf{y} = \boldsymbol{\varphi}(x, \boldsymbol{\lambda})$ of (8.7) is continuously differentiable with respect to $\boldsymbol{\lambda}$ on the restricted cylinder $D : |x| \leq h$, $|\boldsymbol{\lambda} - \boldsymbol{\lambda}_0| \leq c$, where h is defined as in *theorem* 8.1.

定理

Proof

We transform the initial value problem into the equivalent integral equation

$$\boldsymbol{\varphi}(x, \boldsymbol{\lambda}) = \int_0^x \mathbf{f}(t, \boldsymbol{\varphi}(t, \boldsymbol{\lambda}), \boldsymbol{\lambda})\, dt, \tag{8.8}$$

and consider the standard Picard sequence $\boldsymbol{\varphi}_0(x, \boldsymbol{\lambda}) = \mathbf{0}$, with

$$\boldsymbol{\varphi}_{k+1}(x, \boldsymbol{\lambda}) = \int_0^x \mathbf{f}(t, \boldsymbol{\varphi}_k(t, \boldsymbol{\lambda}), \boldsymbol{\lambda})\, dt.$$

Because $J_{\mathbf{y}}\mathbf{f}$ is continuous on the compact set G, $\mathbf{f}$ globally satisfies a Lipschitz condition on G. By *theorem* 8.1, the sequence $\{\boldsymbol{\varphi}_k\}$ converges uniformly on D to the unique continuous solution $\boldsymbol{\varphi}(x, \boldsymbol{\lambda})$.

To prove continuous differentiability, we differentiate the iteration formally with respect to the parameter vector $\boldsymbol{\lambda}$. Applying the chain rule, we define the sequence of Jacobian matrices $U_k(x, \boldsymbol{\lambda}) = \frac{\partial \boldsymbol{\varphi}_k}{\partial \boldsymbol{\lambda}}$:

$$U_{k+1}(x, \boldsymbol{\lambda}) = \int_0^x \left[J_{\mathbf{y}}\mathbf{f}(t, \boldsymbol{\varphi}_k, \boldsymbol{\lambda}) U_k(t, \boldsymbol{\lambda}) + J_{\boldsymbol{\lambda}}\mathbf{f}(t, \boldsymbol{\varphi}_k, \boldsymbol{\lambda}) \right] dt, \tag{8.9}$$

where $U_0 \equiv \mathbf{0}$. Because the partial derivatives of $\mathbf{f}$ are continuous on G, they are bounded by some constant $\alpha > 0$. By induction, one verifies that $|U_k| \leq \beta \leq e^{\alpha h}$ for all $(x, \boldsymbol{\lambda}) \in D$, meaning the derivative sequence is uniformly bounded.

To establish uniform convergence, we examine the Cauchy differences $V_{k,s} = |U_{k+s} - U_k|$ for any integer $s \geq 1$. Writing

$$A_k(t, \boldsymbol{\lambda}) = J_{\mathbf{y}}\mathbf{f}(t, \boldsymbol{\varphi}_k(t, \boldsymbol{\lambda}), \boldsymbol{\lambda}), \quad B_k(t, \boldsymbol{\lambda}) = J_{\boldsymbol{\lambda}}\mathbf{f}(t, \boldsymbol{\varphi}_k(t, \boldsymbol{\lambda}), \boldsymbol{\lambda}),$$

equation (8.9) gives

$$V_{k+1,s}(x, \boldsymbol{\lambda}) \leq \left| \int_0^x (A_{k+s}U_{k+s} - A_k U_k + B_{k+s} - B_k)\, dt \right|.$$

Add and subtract $A_k U_{k+s}$ inside the integrand:

$$V_{k+1,s}(x, \boldsymbol{\lambda}) \leq \alpha \int_{x_-}^{x_+} V_{k,s}(t, \boldsymbol{\lambda})\, dt + \int_{x_-}^{x_+} (\beta |A_{k+s} - A_k| + |B_{k+s} - B_k|)\, dt.$$

where $x_- = \min\{0, x\}$ and $x_+ = \max\{0, x\}$. Define

$$W_{k,s}(r, \boldsymbol{\lambda}) = \sup_{|t| \leq r} V_{k,s}(t, \boldsymbol{\lambda}), \qquad 0 \leq r \leq h.$$

Since $\boldsymbol{\varphi}_k \to \boldsymbol{\varphi}$ uniformly on D, and $J_{\mathbf{y}}\mathbf{f}, J_{\boldsymbol{\lambda}}\mathbf{f}$ are uniformly continuous on G, there exists $\varepsilon_k \downarrow 0$ such that, uniformly in s,

$$\int_{x_-}^{x_+} (\beta |A_{k+s} - A_k| + |B_{k+s} - B_k|)\, dt \leq \varepsilon_k |x|.$$

Hence

$$W_{k+1,s}(r, \boldsymbol{\lambda}) \leq \alpha \int_0^r W_{k,s}(\rho, \boldsymbol{\lambda})\, d\rho + \varepsilon_k r, \qquad (0 \leq r \leq h).$$

Iterating this inequality yields, for each integer $m \geq 1$,

$$W_{k+m,s}(r, \boldsymbol{\lambda}) \leq \varepsilon_k r \sum_{j=0}^{m-1} \frac{(\alpha r)^j}{j!} + 2\beta \frac{(\alpha r)^m}{m!},$$

because $V_{k,s} \leq |U_{k+s}| + |U_k| \leq 2\beta$. On $0 \leq r \leq h$, the exponential series is bounded by $e^{\alpha h}$ and the factorial term tends to 0 as $m \to \infty$. Therefore $\{U_k\}$ is uniformly Cauchy on D.
Thus, $\{U_k(x, \boldsymbol{\lambda})\}$ is a uniformly Cauchy sequence on D. It converges to a continuous limit matrix $U(x, \boldsymbol{\lambda})$. By the classical theorem on the exchange of limits and derivatives for uniformly convergent sequences, $U(x, \boldsymbol{\lambda}) = \frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{\lambda}}$. Furthermore, the spatial derivative $\frac{\partial \boldsymbol{\varphi}}{\partial x} = \mathbf{f}(x, \boldsymbol{\varphi}, \boldsymbol{\lambda})$ is intrinsically continuous. Hence, $\boldsymbol{\varphi}(x, \boldsymbol{\lambda})$ is continuously differentiable on D.

■

The Variational Equations

The continuous differentiability established in *theorem* 8.3 permits the explicit calculation of the sensitivity of a solution to its initial conditions and parameters without requiring a closed-form general solution.

Let $\mathbf{y} = \boldsymbol{\varphi}(x; x_0, \mathbf{y}_0, \boldsymbol{\lambda})$ denote the solution to the general initial value problem

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}, \boldsymbol{\lambda}), \quad \mathbf{y}(x_0) = \mathbf{y}_0. \tag{8.10}$$

The solution identically satisfies the integral equation:

$$\boldsymbol{\varphi}(x; x_0, \mathbf{y}_0, \boldsymbol{\lambda}) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(t, \boldsymbol{\varphi}(t; x_0, \mathbf{y}_0, \boldsymbol{\lambda}), \boldsymbol{\lambda})\, dt.$$

Assuming $\mathbf{f}$ is continuously differentiable in all arguments, we may differentiate this identity with respect to the initial state $\mathbf{y}_0$, the initial time x_0, and the parameter $\boldsymbol{\lambda}$. We denote the Jacobian of $\mathbf{f}$ with respect to the state variables evaluated strictly along the base trajectory as

$$A(x) = J_{\mathbf{y}}\mathbf{f}(x, \boldsymbol{\varphi}(x; x_0, \mathbf{y}_0, \boldsymbol{\lambda}), \boldsymbol{\lambda}).$$

Differentiating the integral identity directly yields three linear initial value problems along the same base trajectory. These are termed the **variational equations** of the system:

Sensitivity to Initial State: Let $\Phi(x) = \frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{y}_0}$ be the $n \times n$ matrix governing the propagation of spatial perturbations. Differentiating the identity with respect to $\mathbf{y}_0$ yields:

$$\frac{d\Phi}{dx} = A(x)\Phi(x), \quad \Phi(x_0) = I_n, \tag{8.11}$$

where I_n is the $n \times n$ identity matrix.

Sensitivity to Initial Time: Let $\mathbf{z}(x) = \frac{\partial \boldsymbol{\varphi}}{\partial x_0}$ be the n-dimensional vector governing the response to a delay in the starting time. Differentiating the integral with respect to its lower bound x_0 yields:

$$\frac{d\mathbf{z}}{dx} = A(x)\mathbf{z}(x), \quad \mathbf{z}(x_0) = -\mathbf{f}(x_0, \mathbf{y}_0, \boldsymbol{\lambda}). \tag{8.12}$$

Sensitivity to Parameters: Let $\Psi(x) = \frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{\lambda}}$ be the $n \times m$ matrix governing the structural response to parameter changes. Differentiating the identity with respect to $\boldsymbol{\lambda}$ yields:

$$\frac{d\Psi}{dx} = A(x)\Psi(x) + J_{\boldsymbol{\lambda}}\mathbf{f}(x, \boldsymbol{\varphi}(x), \boldsymbol{\lambda}), \quad \Psi(x_0) = \mathbf{0}. \tag{8.13}$$

The profound geometric implication is that while the underlying differential equation (8.10) may be highly non-linear, the propagation of perturbations is strictly governed by linear differential equations.

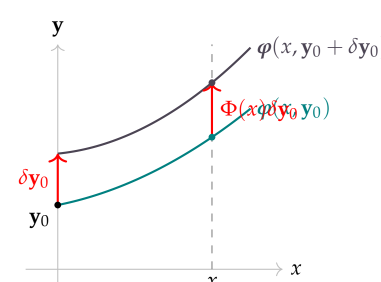

Figure 8.2: Geometric interpretation of the sensitivity matrix $\Phi(x)$. An initial perturbation $\delta\mathbf{y}_0$ is transported along the flow to a spatial displacement $\Phi(x)\delta\mathbf{y}_0$ at time x.

Example 8.2. Sensitivity of a Linear Equation. Without directly computing the general solution, determine the initial state and time sensitivities $\frac{\partial \varphi}{\partial y_0}$ and $\frac{\partial \varphi}{\partial x_0}$ for the scalar linear initial value problem

$$\frac{dy}{dx} + p(x)y = q(x), \quad y(x_0) = y_0.$$

範例

Solution

We rewrite the equation in standard form $y' = -p(x)y + q(x)$. The spatial derivative is strictly $A(x) = \frac{\partial f}{\partial y} = -p(x)$.

By (8.11), the sensitivity to the initial state $z(x) = \frac{\partial \varphi}{\partial y_0}$ satisfies the homogeneous linear equation:

$$\frac{dz}{dx} = -p(x)z, \quad z(x_0) = 1.$$

Separating variables and integrating directly provides the fundamental solution:

$$\frac{\partial \varphi}{\partial y_0} = \exp\left(-\int_{x_0}^{x} p(t)\,dt\right).$$

Similarly, by (8.12), the sensitivity to the initial time $w(x) = \frac{\partial \varphi}{\partial x_0}$ satisfies:

$$\frac{dw}{dx} = -p(x)w, \quad w(x_0) = p(x_0)y_0 - q(x_0).$$

Because this is the exact same homogeneous equation multiplied by a different initial constant, linearity dictates:

$$\frac{\partial \varphi}{\partial x_0} = (p(x_0)y_0 - q(x_0)) \exp\left(-\int_{x_0}^{x} p(t)\,dt\right).$$

■

For non-linear systems, the matrix $A(x)$ inevitably depends on the explicit trajectory $\boldsymbol{\varphi}(x)$, meaning the variational equations cannot generally be solved without first solving the base equation. However, if a trivial or known base solution exists, the variational equations permit isolated analysis of perturbations surrounding it.

Example 8.3. Pointwise Evaluation of Sensitivities. Let $y = \varphi(x; x_0, y_0, \lambda)$ be the unique solution to the non-linear initial value problem

$$\frac{dy}{dx} = \sin(\lambda xy), \quad y(x_0) = y_0. \tag{8.14}$$

Determine the partial derivatives $\frac{\partial \varphi}{\partial x_0}$, $\frac{\partial \varphi}{\partial y_0}$, and $\frac{\partial \varphi}{\partial \lambda}$ evaluated at the trivial origin state $(x_0, y_0, \lambda) = (0, 0, 0)$.

範例

Solution

We define the vector field $f(x, y, \lambda) = \sin(\lambda xy)$. Its relevant partial derivatives are:

$$\frac{\partial f}{\partial y} = \lambda x \cos(\lambda xy), \quad \frac{\partial f}{\partial \lambda} = xy \cos(\lambda xy).$$

We evaluate the sensitivities specifically around the trivial configuration. When $x_0 = 0, y_0 = 0$, the initial value problem (8.14) admits the identically zero solution $\varphi(x; 0, 0, \lambda) \equiv 0$ for all real x.
Substituting $\varphi \equiv 0$ into our partial derivatives gives the simplified coefficients for the variational equations:

$$A(x) = \lambda x \cos(0) = \lambda x, \quad B(x) = \frac{\partial f}{\partial \lambda}(x, 0, \lambda) = 0.$$

We construct the three variational problems using (8.11), (8.12), and (8.13):

For $z_1 = \frac{\partial \varphi}{\partial x_0}$, the initial condition is $z_1(0) = -f(0, 0, \lambda) = 0$. The initial value problem is $z_1' = \lambda x z_1$ with $z_1(0) = 0$, which dictates $z_1(x) \equiv 0$. Thus,

$$\left.\frac{\partial \varphi}{\partial x_0}\right|_{(0,0,0)} = 0.$$

For $z_2 = \frac{\partial \varphi}{\partial y_0}$, the problem is $z_2' = \lambda x z_2$ with $z_2(0) = 1$. Integrating yields $z_2(x) = \exp(\frac{1}{2}\lambda x^2)$. At $\lambda = 0$, this returns

$$\left.\frac{\partial \varphi}{\partial y_0}\right|_{(0,0,0)} = e^0 = 1.$$

For $z_3 = \frac{\partial \varphi}{\partial \lambda}$, the problem is $z_3' = \lambda x z_3 + 0$ with $z_3(0) = 0$. The unique solution is strictly zero. Hence,

$$\left.\frac{\partial \varphi}{\partial \lambda}\right|_{(0,0,0)} = 0.$$

■

8.3 Exercises

1. **Explicit Straightening Map for a Linear Equation.** Consider the scalar linear equation

$$y' + p(x)y = q(x), \qquad y(0) = \eta,$$

where p, q are continuous on $|x| \leq h$. Let $\varphi(x, \eta)$ denote the unique solution.

(a) Compute $\varphi(x, \eta)$ explicitly.

(b) Write the map $T : (x, \eta) \mapsto (x, \varphi(x, \eta))$ and an explicit formula for T^{-1}.

(c) Verify directly that T is a homeomorphism from a rectangle in the (x, η)-plane onto its image.

2. **Continuous Dependence from Uniqueness (Local Region).** Suppose $f(x, y)$ is continuous in the region

$$R: \quad |x - x_0| \leq a, \qquad |y - y_0| \leq b,$$

and the integral curve of

$$\frac{dy}{dx} = f(x, y)$$

passing through any interior point of R is unique. Prove that solutions depend continuously on initial values in a local region around (x_0, y_0).

3. **Failure Without Uniqueness.** Give an example showing that if the differential equation does not satisfy uniqueness, then its family of integral curves cannot be viewed locally as a family of parallel straight lines.

4. **Parameter Sensitivity for a Linear Family.** For the parameter-dependent linear initial value problem

$$y' + p(x)y = \lambda q(x), \qquad y(x_0) = y_0,$$

with continuous p, q, let $z(x) = \dfrac{\partial \varphi}{\partial \lambda}(x; x_0, y_0, \lambda)$.

(a) Derive the linear equation and initial condition satisfied by z.

(b) Solve that equation explicitly.

(c) Differentiate the Cauchy formula for y directly with respect to λ and verify agreement.

5. **Relation Between $\partial\varphi/\partial x_0$ and $\partial\varphi/\partial y_0$.** Let $\Phi(x) = \dfrac{\partial \varphi}{\partial y_0}(x; x_0, y_0, \lambda)$ and $\mathbf{z}(x) = \dfrac{\partial \varphi}{\partial x_0}(x; x_0, y_0, \lambda)$. Using (8.11) and (8.12), prove

$$\mathbf{z}(x) = -\Phi(x)\, \mathbf{f}(x_0, y_0, \lambda).$$

Then write the corresponding scalar formula.

6. **Positivity of Initial-Value Sensitivity.** Suppose the scalar function $y = y(x, \eta)$ is the solution of $\frac{dy}{dx} = \sin(xy), y(0) = \eta$. Prove that $\frac{\partial y}{\partial \eta}(x, \eta) > 0$ for all x and η.

9
General Theory of Linear Systems

Mathematical models of dynamic systems frequently generate non-linear differential equations. A ubiquitous analytical strategy is to linearise these equations around an equilibrium state or a known reference trajectory, transforming the non-linear problem into a locally equivalent linear system (as previously observed with the pendulum approximation). The robust structural properties of linear differential equations provide both a fundamental toolkit for explicit computation and the theoretical scaffolding for advanced stability analysis. This chapter establishes the general algebraic and analytic theory governing linear differential equation systems.

9.1 Structure of the Solution Space

Recall from eq. (7.20) that an n-dimensional linear differential equation system on an interval (a, b) takes the standard vector form

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y} + \mathbf{e}(x),$$

where $A(x)$ is an $n \times n$ matrix-valued function and $\mathbf{e}(x)$ is an n-dimensional vector-valued function, both continuous on (a, b).
If the forcing term $\mathbf{e}(x)$ is not identically zero, the system is termed **non-homogeneous**. If $\mathbf{e}(x) \equiv \mathbf{0}$, we obtain the corresponding **homogeneous** linear differential equation system:

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y}. \tag{9.1}$$

Our immediate objective is to characterise the algebraic structure of the set of all valid solutions to (9.1), denoted by S. We establish that S natively inherits the structure of a finite-dimensional vector space.

> **Lemma 9.1.** ***Superposition Principle.***
> Let $\mathbf{y}_1(x)$ and $\mathbf{y}_2(x)$ be solutions to the homogeneous linear differential equation system (9.1) on (a, b). Then for any real scalars C_1, C_2, the

linear combination

$$\mathbf{y}(x) = C_1\mathbf{y}_1(x) + C_2\mathbf{y}_2(x)$$

is also a solution on (a, b).

引理

Proof

By the linearity of the derivative operator and matrix multiplication,

$$\begin{aligned}\frac{d}{dx}(C_1\mathbf{y}_1 + C_2\mathbf{y}_2) &= C_1\frac{d\mathbf{y}_1}{dx} + C_2\frac{d\mathbf{y}_2}{dx} \\ &= C_1 A(x)\mathbf{y}_1 + C_2 A(x)\mathbf{y}_2 \\ &= A(x)\left(C_1\mathbf{y}_1 + C_2\mathbf{y}_2\right).\end{aligned}$$

Thus, the linear combination identically satisfies (9.1).

■

lemma 9.1 confirms that the solution set S is a linear subspace of the vector space of continuously differentiable functions from (a, b) to $\mathbb{R}^n$.

Theorem 9.1. ***Dimension of the Solution Space.***
Let the matrix $A(x)$ be continuous on the interval (a, b). The linear space S of solutions to the homogeneous system (9.1) is exactly n-dimensional.

定理

Proof

Fix a base point $x_0 \in (a, b)$. By the global existence and uniqueness theorem for linear systems (*proposition* 7.3), for every constant vector $\mathbf{y}_0 \in \mathbb{R}^n$, there exists a unique global solution $\mathbf{y}(x) \in S$ satisfying the initial condition $\mathbf{y}(x_0) = \mathbf{y}_0$.
This uniqueness guarantees that the evaluation mapping

$$\begin{aligned}&H : \mathbb{R}^n \to S \\ &\mathbf{y}_0 \mapsto \mathbf{y}(x) \quad \text{such that } \mathbf{y}(x_0) = \mathbf{y}_0\end{aligned}$$

is a well-defined bijection. Specifically:

1. **Surjectivity:** Any solution $\mathbf{y}(x) \in S$ trivially possesses a well-defined initial state $\mathbf{y}(x_0) \in \mathbb{R}^n$, which serves as its pre-image.
2. **Injectivity:** By uniqueness in *proposition* 7.3, if $\mathbf{y}_1(x_0) \neq \mathbf{y}_2(x_0)$, the corresponding global solutions must remain distinct as functions on (a, b), hence $\mathbf{y}_1(x) \neq \mathbf{y}_2(x)$.

Furthermore, *lemma* 9.1 ensures H preserves linear combinations:

$$H(C_1\mathbf{y}_{0,1} + C_2\mathbf{y}_{0,2}) = C_1 H(\mathbf{y}_{0,1}) + C_2 H(\mathbf{y}_{0,2}).$$

Therefore, H is a linear isomorphism. Two isomorphic vector spaces must share the same dimension, concluding that $\dim(S) = \dim(\mathbb{R}^n) = n$.

■

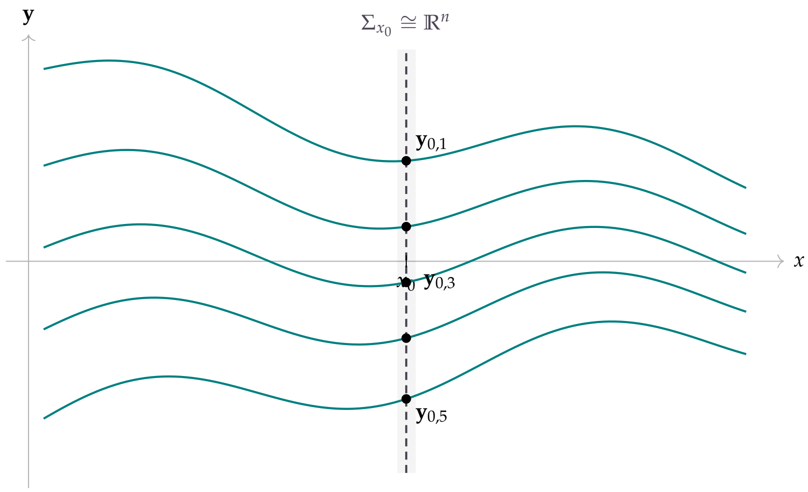

Figure 9.1: Geometric interpretation of the isomorphism $H^{-1}\colon S \to \mathbb{R}^n$. Each solution curve intersects the hyperplane $\Sigma_{x_0} = \{x = x_0\}$ in exactly one point. Evaluation at x_0 bijectively identifies the n-dimensional solution space S with $\mathbb{R}^n$.

Definition 9.1. ***Fundamental Set of Solutions.***
A set of n linearly independent solutions $\boldsymbol{\varphi}_1(x), \boldsymbol{\varphi}_2(x), \ldots, \boldsymbol{\varphi}_n(x)$ to the homogeneous system (9.1) is called a **fundamental set of solutions**.

定義

Because these n solutions form a basis for the n-dimensional space S, the general solution to (9.1) is given by their arbitrary linear combination:

$$\mathbf{y}(x) = \sum_{i=1}^{n} C_i \boldsymbol{\varphi}_i(x), \tag{9.2}$$

where $C_1, \ldots, C_n$ are arbitrary real constants.

9.2 *The Wronskian Determinant and Liouville's Formula*

To pragmatically determine whether a given set of n solutions is linearly independent, we assemble them into a single matrix.
Let $\mathbf{y}_1(x), \ldots, \mathbf{y}_n(x)$ be n solutions to (9.1). We define the $n \times n$ **solution matrix** $Y(x)$ by aligning the solution vectors as its columns:

$$Y(x) = \begin{bmatrix} \mathbf{y}_1(x) & \mathbf{y}_2(x) & \cdots & \mathbf{y}_n(x) \end{bmatrix} = \begin{bmatrix} y_{11}(x) & y_{12}(x) & \cdots & y_{1n}(x) \\ y_{21}(x) & y_{22}(x) & \cdots & y_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1}(x) & y_{n2}(x) & \cdots & y_{nn}(x) \end{bmatrix}.$$

Because each column inherently satisfies the differential equation, matrix multiplication dictates that $Y(x)$ itself is a matrix-valued solution to the system:

$$\frac{dY(x)}{dx} = A(x)Y(x).$$

Definition 9.2. ***Wronskian Determinant*****.**
The **Wronskian determinant** $W(x)$ of a set of n solutions is the determinant of their corresponding solution matrix:

$$W(x) = \det Y(x).$$

定義

The propagation of this determinant along the interval (a, b) is governed by an elegant scalar differential equation dependent strictly on the trace of the coefficient matrix $A(x)$.

Lemma 9.2. ***Liouville's Formula*****.**
Let $W(x)$ be the Wronskian determinant of a solution matrix $Y(x)$ to (9.1). Then $W(x)$ explicitly satisfies the relation

$$W(x) = W(x_0) \exp\left(\int_{x_0}^{x} \mathrm{tr}[A(t)]\, dt\right), \tag{9.3}$$

where $x_0 \in (a, b)$ and $\mathrm{tr}[A(t)] = \sum_{j=1}^{n} a_{jj}(t)$ is the trace of $A(t)$.

引理

Proof

By the standard expansion rules for differentiating a determinant, the derivative $\frac{dW}{dx}$ equals the sum of n distinct determinants. In the i-th determinant of the sum, the i-th row is differentiated, while all other rows remain identical to $Y(x)$:

$$\frac{dW}{dx} = \sum_{i=1}^{n} \begin{vmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{i1} & y'_{i2} & \cdots & y'_{in} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{vmatrix}.$$

Because $Y(x)$ satisfies $Y' = AY$, the derivative of the i-th row can be expanded using the differential equation system $y'_{ij} = \sum_{k=1}^{n} a_{ik} y_{kj}$:

$$\frac{dW}{dx} = \sum_{i=1}^{n} \begin{vmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{ik} y_{k1} & \sum_{k=1}^{n} a_{ik} y_{k2} & \cdots & \sum_{k=1}^{n} a_{ik} y_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{vmatrix}.$$

By the multilinearity of determinants, we can split the i-th row of the i-th determinant into a sum over k. If $k \neq i$, the resulting matrix has two proportional rows (the k-th row and the i-th row are linearly dependent), thus yielding a determinant of zero. The only non-zero contribution occurs when $k = i$, isolating the diagonal coefficient $a_{ii}(x)$:

$$\frac{dW}{dx} = \sum_{i=1}^{n} a_{ii}(x) \begin{vmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{i1} & \cdots & y_{in} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nn} \end{vmatrix} = \left(\sum_{i=1}^{n} a_{ii}(x) \right) W(x) = \mathrm{tr}[A(x)] W(x).$$

This is a separable first-order scalar differential equation $W' = \mathrm{tr}[A(x)]W$. Integrating from x_0 to x directly yields (9.3).

■

Because the exponential function is strictly positive, *lemma* 9.2 dictates an extreme geometric dichotomy: $W(x)$ is either identically zero everywhere on (a, b), or it is non-zero everywhere.

Theorem 9.2. ***Wronskian Criterion for Linear Independence.***
A set of n solutions to the homogeneous linear system (9.1) is linearly independent on (a, b) if and only if its Wronskian determinant is nowhere zero:

$$W(x) \neq 0 \quad \text{for all } x \in (a, b).$$

定理

Proof

By Liouville's Formula (*lemma* 9.2), $W(x) \neq 0$ globally on (a, b) is strictly equivalent to the local condition $W(x_0) \neq 0$ at a single evaluation point x_0.
From classical linear algebra, the determinant $W(x_0) = \det[\mathbf{y}_1(x_0) \ldots \mathbf{y}_n(x_0)]$ is non-zero if and only if the initial state vectors $\mathbf{y}_1(x_0), \ldots, \mathbf{y}_n(x_0)$ are linearly independent in $\mathbb{R}^n$.
As established in the proof of *theorem* 9.1, the evaluation mapping

$H^{-1} : \mathbf{y}(x) \mapsto \mathbf{y}(x_0)$ is a linear isomorphism. Because isomorphisms strictly preserve linear independence, the initial vectors in $\mathbb{R}^n$ are linearly independent if and only if the continuous solution functions $\mathbf{y}_1(x), \ldots, \mathbf{y}_n(x)$ are linearly independent in S.

■

Corollary 9.1. ***Wronskian Linear Dependence.***
A set of n solutions to (9.1) is linearly dependent if and only if $W(x) \equiv 0$ on (a, b).

推論

Remark.

The constraint that the vectors must be *solutions* to the same linear differential equation is absolute. For arbitrary sets of smooth functions, a vanishing Wronskian does not necessarily imply linear dependence. The equivalence is unique to the solution spaces of linear systems.

Example 9.1. Wronskian Verification for a 2 × 2 System. Verify that the general solution to the coupled system

$$\frac{d}{dx}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos^2 x & \frac{1}{2}\sin(2x) - 1 \\ \frac{1}{2}\sin(2x) + 1 & \sin^2 x \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{9.4}$$

is generated by the basis vectors

$$\boldsymbol{\varphi}_1(x) = \begin{bmatrix} e^x \cos x \\ e^x \sin x \end{bmatrix}, \quad \boldsymbol{\varphi}_2(x) = \begin{bmatrix} -\sin x \\ \cos x \end{bmatrix}. \tag{9.5}$$

範例

Solution

Assuming direct substitution verifies that $\boldsymbol{\varphi}_1(x)$ and $\boldsymbol{\varphi}_2(x)$ identically satisfy (9.4) over $\mathbb{R}$, we evaluate their linear independence using the Wronskian at a computationally convenient point, such as $x = 0$.
The initial states are:

$$\boldsymbol{\varphi}_1(0) = \begin{bmatrix} e^0 \cos 0 \\ e^0 \sin 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\varphi}_2(0) = \begin{bmatrix} -\sin 0 \\ \cos 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Constructing the solution matrix at the origin gives the identity matrix, whose determinant is trivial:

$$W(0) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \neq 0.$$

By *theorem* 9.2, the solutions are linearly independent globally, thus forming a fundamental set. The general solution is strictly their linear combination:

$$\begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix} = C_1 \begin{bmatrix} e^x \cos x \\ e^x \sin x \end{bmatrix} + C_2 \begin{bmatrix} -\sin x \\ \cos x \end{bmatrix}.$$

■

9.3 *Fundamental Solution Matrices*

The linear algebra mechanics of S can be seamlessly abstracted into matrix algebra by defining global fundamental matrices.

Definition 9.3. ***Fundamental Solution Matrix.***
A solution matrix $\Phi(x)$ is termed a **fundamental solution matrix** if its columns constitute a fundamental set of solutions.

定義

Because the columns form a basis, $\det \Phi(x) \neq 0$, implying that $\Phi(x)$ is an invertible matrix everywhere on (a, b). Furthermore, the vector general solution (9.2) can be compactly written as a matrix-vector product:

$$\mathbf{y}(x) = \Phi(x)\mathbf{c}, \tag{9.6}$$

where $\mathbf{c} = (C_1, C_2, \ldots, C_n)^\top$ is an arbitrary constant column vector.

Proposition 9.1. ***Transformations of Fundamental Matrices.***
Let $\Phi(x)$ be a fundamental solution matrix of (9.1).

1. For any non-singular constant $n \times n$ matrix C, the matrix $\Psi(x) = \Phi(x)C$ is also a fundamental solution matrix.
2. Conversely, if $\Psi(x)$ is any other fundamental solution matrix of the same system, there exists a unique non-singular constant matrix C such that $\Psi(x) = \Phi(x)C$.

命題

Proof

For (1), we check that $\Psi(x)$ satisfies the differential equation:

$$\frac{d\Psi}{dx} = \frac{d}{dx}(\Phi(x)C) = \left(\frac{d\Phi}{dx}\right)C = (A(x)\Phi(x))C = A(x)(\Phi(x)C) = A(x)\Psi(x).$$

Furthermore, $\det \Psi(x) = \det \Phi(x) \det C$. Since $\det \Phi(x) \neq 0$ and $\det C \neq 0$, it follows that $\det \Psi(x) \neq 0$, thus establishing $\Psi(x)$ as a fundamental solution matrix.
For (2), define

$$C(x) = \Phi^{-1}(x)\Psi(x).$$

Then

$$C'(x) = \Phi^{-1}(x)\Psi'(x) - \Phi^{-1}(x)\Phi'(x)\Phi^{-1}(x)\Psi(x) = \Phi^{-1}(x)A(x)\Psi(x) - \Phi^{-1}(x)A(x)\Phi(x)\Phi^{-1}(x)\Psi(x) = 0.$$

Hence $C(x)$ is constant, so $\Psi(x) = \Phi(x)C$ for a unique constant matrix C. Because $\det \Psi(x) \neq 0$ and $\det \Phi(x) \neq 0$, we get

$$\det C = \det\big(\Phi^{-1}(x)\Psi(x)\big) \neq 0,$$

so C is non-singular.

■

9.4 *Non-Homogeneous Linear Systems*

With the algebraic structure of the homogeneous system fully characterised, we now construct the general solution for the non-homogeneous linear differential equation system:

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y} + \mathbf{e}(x). \tag{9.7}$$

The underlying geometry mirrors that of scalar linear equations: the complete solution space of (9.7) is an affine space, generated by translating the linear subspace S of the homogeneous system (9.1) by a single particular solution.

Lemma 9.3. ***Affine Structure of the Solution Space.***
Let $\Phi(x)$ be a fundamental solution matrix of the corresponding homogeneous system $\mathbf{y}' = A(x)\mathbf{y}$. If $\boldsymbol{\varphi}^*(x)$ is any particular solution to the non-homogeneous system (9.7), then every solution $\mathbf{y} = \boldsymbol{\varphi}(x)$ of (9.7) can be uniquely expressed as

$$\boldsymbol{\varphi}(x) = \Phi(x)\mathbf{c} + \boldsymbol{\varphi}^*(x),$$

where $\mathbf{c} \in \mathbb{R}^n$ is a constant column vector.

引理

Proof

Consider the difference between the arbitrary solution and the particular solution: $\mathbf{u}(x) = \boldsymbol{\varphi}(x) - \boldsymbol{\varphi}^*(x)$. By the linearity of differentiation,

$$\frac{d\mathbf{u}}{dx} = \frac{d\boldsymbol{\varphi}}{dx} - \frac{d\boldsymbol{\varphi}^*}{dx} = [A(x)\boldsymbol{\varphi}(x) + \mathbf{e}(x)] - [A(x)\boldsymbol{\varphi}^*(x) + \mathbf{e}(x)] = A(x)\mathbf{u}(x).$$

Thus, $\mathbf{u}(x)$ strictly satisfies the homogeneous equation (9.1). By *theorem* 9.1 and *definition* 9.3, every homogeneous solution lies in the column space of the fundamental matrix, meaning there exists a unique constant vector $\mathbf{c}$ such that $\mathbf{u}(x) = \Phi(x)\mathbf{c}$. Rearranging yields the stated result.

■

Variation of Parameters

To explicitly construct the particular solution $\boldsymbol{\varphi}^*(x)$, we generalise the method of integrating factors from the scalar case. We assume the particular solution takes the structural form of the homogeneous general solution, but allow the parameter vector $\mathbf{c}$ to vary dynamically with x:

$$\boldsymbol{\varphi}^*(x) = \Phi(x)\mathbf{c}(x). \tag{9.8}$$

Differentiating (9.8) via the product rule yields:

$$\frac{d\boldsymbol{\varphi}^*}{dx} = \Phi'(x)\mathbf{c}(x) + \Phi(x)\mathbf{c}'(x).$$

Substituting this and the ansatz (9.8) into the non-homogeneous system (9.7) provides:

$$\Phi'(x)\mathbf{c}(x) + \Phi(x)\mathbf{c}'(x) = A(x)\Phi(x)\mathbf{c}(x) + \mathbf{e}(x).$$

Because $\Phi(x)$ is a solution matrix for the homogeneous system, it inherently satisfies $\Phi'(x) = A(x)\Phi(x)$. The terms involving $\mathbf{c}(x)$ strictly cancel, isolating the derivative of the parameter vector:

$$\Phi(x)\mathbf{c}'(x) = \mathbf{e}(x). \tag{9.9}$$

By *theorem* 9.2, the fundamental matrix $\Phi(x)$ possesses a non-zero Wronskian determinant everywhere on the interval (a, b), ensuring $\Phi^{-1}(x)$ exists globally. Multiplying (9.9) by the inverse matrix isolates $\mathbf{c}'(x)$:

$$\mathbf{c}'(x) = \Phi^{-1}(x)\mathbf{e}(x).$$

Direct integration from a base point $x_0 \in (a, b)$ yields the time-varying parameter vector:

$$\mathbf{c}(x) = \int_{x_0}^{x} \Phi^{-1}(s)\mathbf{e}(s)\, ds.$$

Substituting this back into our initial ansatz (9.8) secures the particular solution:

$$\boldsymbol{\varphi}^*(x) = \Phi(x) \int_{x_0}^{x} \Phi^{-1}(s)\mathbf{e}(s)\, ds. \tag{9.10}$$

We synthesise this derivation with *lemma* 9.3 to establish the master theorem for non-homogeneous linear systems.

> **Theorem 9.3.** ***General Solution of Non-Homogeneous Systems.***
> Let $\Phi(x)$ be a fundamental solution matrix of $\mathbf{y}' = A(x)\mathbf{y}$. The general solution of the non-homogeneous system $\mathbf{y}' = A(x)\mathbf{y} + \mathbf{e}(x)$ on (a, b) is
>
> $$\mathbf{y}(x) = \Phi(x)\mathbf{c} + \Phi(x) \int_{x_0}^{x} \Phi^{-1}(s)\mathbf{e}(s)\, ds, \tag{9.11}$$
>
> where $\mathbf{c} \in \mathbb{R}^n$ is an arbitrary constant vector.
> Furthermore, the unique solution satisfying the initial condition $\mathbf{y}(x_0) =$

$\mathbf{y}_0$ is strictly given by:

$$\mathbf{y}(x) = \Phi(x)\Phi^{-1}(x_0)\mathbf{y}_0 + \Phi(x)\int_{x_0}^{x} \Phi^{-1}(s)\mathbf{e}(s)\, ds. \tag{9.12}$$

定理

Remark.

While (9.11) structurally parallels the general solution for a first-order scalar equation, there is a severe operational distinction. In the scalar case, the fundamental factor is the simple exponential $\exp(\int A(x)\, dx)$, so the solution is always reducible to quadrature, even when no elementary closed form exists. For $n \geq 2$, the matrix $\Phi(x)$ generally cannot be expressed as a closed-form integral of $A(x)$ unless $A(x)$ commutes with its own integral. Consequently, (9.12) functions primarily as a theoretical representation framework rather than a universal computational algorithm, except in cases where $\Phi(x)$ is analytically deducible.

Example 9.2. Variation of Parameters on a 2×2 System. Solve the initial value problem:

$$\begin{cases} \frac{d}{dx}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos^2 x & \frac{1}{2}\sin(2x) - 1 \\ \frac{1}{2}\sin(2x) + 1 & \sin^2 x \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \cos x \\ \sin x \end{bmatrix}, \\ \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{cases} \tag{9.13}$$

範例

Solution

From the analysis of the homogeneous counterpart (9.4), we previously established the fundamental solution matrix:

$$\Phi(x) = \begin{bmatrix} e^x \cos x & -\sin x \\ e^x \sin x & \cos x \end{bmatrix}.$$

The determinant is $\det \Phi(x) = e^x \cos^2 x + e^x \sin^2 x = e^x$. The inverse matrix is therefore:

$$\Phi^{-1}(x) = e^{-x} \begin{bmatrix} \cos x & \sin x \\ -e^x \sin x & e^x \cos x \end{bmatrix} = \begin{bmatrix} e^{-x}\cos x & e^{-x}\sin x \\ -\sin x & \cos x \end{bmatrix}.$$

We compute the integrand of the particular solution (9.10) by multiplying the inverse fundamental matrix by the forcing vector $\mathbf{e}(s) = (\cos s, \sin s)^\top$:

$$\Phi^{-1}(s)\mathbf{e}(s) = \begin{bmatrix} e^{-s}\cos s & e^{-s}\sin s \\ -\sin s & \cos s \end{bmatrix} \begin{bmatrix} \cos s \\ \sin s \end{bmatrix} = \begin{bmatrix} e^{-s}\cos^2 s + e^{-s}\sin^2 s \\ -\sin s \cos s + \sin s \cos s \end{bmatrix} = \begin{bmatrix} e^{-s} \\ 0 \end{bmatrix}.$$

Integrating this vector from the initial time $x_0 = 0$ to x yields the dynamic parameter vector:

$$\mathbf{c}(x) = \int_0^x \begin{bmatrix} e^{-s} \\ 0 \end{bmatrix} ds = \begin{bmatrix} 1 - e^{-x} \\ 0 \end{bmatrix}.$$

The particular solution is generated by mapping this back through $\Phi(x)$:

$$\boldsymbol{\varphi}^*(x) = \Phi(x)\mathbf{c}(x) = \begin{bmatrix} e^x \cos x & -\sin x \\ e^x \sin x & \cos x \end{bmatrix} \begin{bmatrix} 1 - e^{-x} \\ 0 \end{bmatrix} = \begin{bmatrix} (e^x - 1)\cos x \\ (e^x - 1)\sin x \end{bmatrix}.$$

Finally, we evaluate the homogeneous component satisfying the initial condition. Since $\Phi(0) = I_2$, we have $\Phi^{-1}(0)\mathbf{y}_0 = \mathbf{y}_0 = (0, 1)^\top$. The free response is:

$$\Phi(x) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin x \\ \cos x \end{bmatrix}.$$

By *theorem* 9.3, the unique solution is the sum of the homogeneous and particular components:

$$\begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix} = \begin{bmatrix} (e^x - 1)\cos x - \sin x \\ (e^x - 1)\sin x + \cos x \end{bmatrix}.$$

■

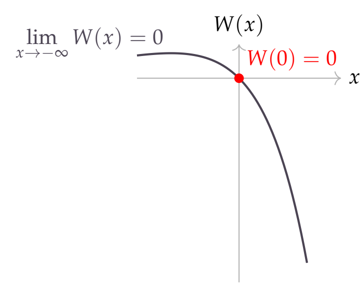

Figure 9.2: The Wronskian $W(x) = -xe^x$ evaluated for the system in Example 3. The point $x = 0$ is a regular zero of W; the singularity occurs in the coefficient matrix $A(x)$.

Example 9.3. A Singular Coefficient Matrix. Determine a fundamental solution matrix and its corresponding Wronskian for the homogeneous differential equation system

$$\frac{d}{dx} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{x} & 1 + \frac{1}{x} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \tag{9.14}$$

defined over the domain $D = (-\infty, 0) \cup (0, \infty)$.

範例

Solution

Expanding the system into components, we have:

$$y_1' = y_2, \quad y_2' = -\frac{1}{x} y_1 + \left(1 + \frac{1}{x}\right) y_2.$$

Substituting y_2 and its derivative into the second equation isolates y_1 into a second-order linear scalar equation:

$$y_1'' = -\frac{1}{x} y_1 + \left(1 + \frac{1}{x}\right) y_1' \implies x y_1'' - (x+1) y_1' + y_1 = 0.$$

By inspection, $y_1(x) = e^x$ is a solution, as $xe^x - (x+1)e^x + e^x = 0$.

This forces $y_2 = (e^x)' = e^x$, yielding the first solution vector $\boldsymbol{\varphi}_1 = (e^x, e^x)^\top$.
A second independent solution is a linear polynomial $y_1(x) = x + 1$, which gives $y_1' = 1$ and $y_1'' = 0$. Verifying this: $0 - (x+1)(1) + (x+1) = 0$. This forces $y_2 = 1$, yielding the second vector $\boldsymbol{\varphi}_2 = (x+1, 1)^\top$.
The resulting fundamental solution matrix is:

$$\Phi(x) = \begin{bmatrix} e^x & x+1 \\ e^x & 1 \end{bmatrix}.$$

We evaluate the Wronskian determinant:

$$W(x) = \det \Phi(x) = e^x(1) - e^x(x+1) = -xe^x.$$

By *theorem* 9.2, $\Phi(x)$ is a fundamental matrix precisely where $W(x) \neq 0$, which is strictly true for all $x \in D$. The general solution over either contiguous interval $(-\infty, 0)$ or $(0, \infty)$ is therefore:

$$\begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix} = C_1 \begin{bmatrix} e^x \\ e^x \end{bmatrix} + C_2 \begin{bmatrix} x+1 \\ 1 \end{bmatrix}.$$

We critically observe that at $x = 0$, the Wronskian evaluates to $W(0) = 0$ (see *figure* 9.2). Does this violate *theorem* 9.2, which states that the Wronskian of linearly independent solutions is nowhere zero?
It does not. *theorem* 9.2 natively inherits the hypothesis of *proposition* 7.3, requiring the coefficient matrix $A(x)$ to be continuous on the interval (a, b). In system (9.14), the elements $-\frac{1}{x}$ and $1 + \frac{1}{x}$ introduce a hard singularity at $x = 0$, severing the real line into two distinct intervals of continuity. The linear independence of solutions is topologically bounded by the domain of the differential equation itself; the behavior exactly at the singularity is undefined by the classical theory.

■

9.5 *Exercises*

1. **General Solutions for Homogeneous Systems.** Find the general solution of the homogeneous linear differential equation system

$$\frac{d\mathbf{y}}{dt} = A(t)\mathbf{y}$$

for each coefficient matrix:

$$\text{(a)} \quad A(t) = \begin{bmatrix} \frac{1}{t} & 0 \\ 0 & \frac{1}{t} \end{bmatrix}, \quad t \neq 0.$$

Solve part (a) separately on each connected interval $(-\infty, 0)$ and $(0, \infty)$. If you regard the full domain as $(-\infty, 0) \cup (0, \infty)$, state explicitly how the integration constants may differ across the two components.

$$\text{(b)} \quad A(t) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

2. **Initial Value Problems for Non-Homogeneous Systems.** Solve the following initial value problems:

Both systems are triangular; solve one scalar equation first, then substitute into the other.

$$\text{(a)} \quad \begin{cases} \frac{dx}{dt} = 1 - \frac{2}{t}x, \\ \frac{dy}{dt} = x + y - 1 + \frac{2}{t}x, \quad t > 0, \\ x(1) = \frac{1}{3}, \quad y(1) = -\frac{1}{3}. \end{cases}$$

$$\text{(b)} \quad \begin{cases} \frac{dx}{dt} = \frac{2t}{1+t^2}x, \\ \frac{dy}{dt} = -\frac{1}{t}y + x + t, \quad t > 0, \\ x(1) = 0, \quad y(1) = \frac{4}{3}. \end{cases}$$

3. **A Vanishing Wronskian Outside the Hypotheses.** Prove that the vector-function set

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x^2 \\ 0 \\ 0 \end{bmatrix}$$

is linearly independent on any interval (a, b). Then compute its Wronskian and verify that it vanishes identically. Explain why this does not contradict theorem 9.2.

4. **Reconstructing the Coefficient Matrix from a Fundamental Matrix.** Let $\Phi(x)$ be an invertible C^1 matrix on an interval I.

   (a) Show that if Φ is a fundamental solution matrix of $\mathbf{y}' = A(x)\mathbf{y}$, then necessarily

   $$A(x) = \Phi'(x)\Phi^{-1}(x).$$

   (b) Deduce the uniqueness statement: if

   $$\mathbf{y}' = A(x)\mathbf{y} \qquad \text{and} \qquad \mathbf{y}' = B(x)\mathbf{y}$$

   have the same fundamental solution matrix $\Phi(x)$ on I, then $A(x) \equiv B(x)$ on I.

   (c) Apply part (a) to compute the coefficient matrices generated by

   $$\Phi_1(x) = \begin{bmatrix} e^x & 0 \\ 0 & e^{-x} \end{bmatrix}, \qquad \Phi_2(x) = \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}.$$

5. **Equivalent Integral Equation for a Non-Linear Perturbation.** Let $\Phi(x)$ be a fundamental solution matrix of the homogeneous system (9.1), and let $\mathbf{f}(x,\mathbf{y})$ be continuous on

$$E = \{(x,\mathbf{y}) : a < x < b,\ \mathbf{y} \in \mathbb{R}^n\}.$$

Prove that solving the initial value problem

$$\begin{cases} \frac{d\mathbf{y}}{dx} = A(x)\mathbf{y} + \mathbf{f}(x,\mathbf{y}), \\ \mathbf{y}(x_0) = \mathbf{y}_0, \end{cases}$$

where $x_0 \in (a,b)$, is equivalent to solving the vector integral equation

$$\mathbf{y}(x) = \Phi(x)\Phi^{-1}(x_0)\mathbf{y}_0 + \int_{x_0}^{x} \Phi(x)\Phi^{-1}(s)\mathbf{f}(s,\mathbf{y}(s))\,ds.$$

6. **Affine Geometry of Non-Homogeneous Solution Families.** Assume $\mathbf{e}(x) \not\equiv \mathbf{0}$ on (a,b) in (9.7), and let $\boldsymbol{\varphi}^*(x)$ be a fixed particular solution.

   (a) Let $\boldsymbol{\psi}^*(x)$ be any other particular solution of (9.7). Prove that

   $$\boldsymbol{\varphi}^*(x) + S = \boldsymbol{\psi}^*(x) + S,$$

   i.e., the affine solution set does not depend on which particular solution is chosen.

   (b) Show that for any $n+1$ solutions $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_n$ of (9.7), the following are equivalent:

   $$\{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_n\} \text{ is linearly independent} \iff \{\mathbf{y}_1 - \mathbf{y}_0, \dots, \mathbf{y}_n - \mathbf{y}_0\} \text{ is linearly independent in } S.$$

   (c) Construct explicitly $n+1$ linearly independent solutions of (9.7) from one fundamental solution matrix and one particular solution, and deduce that the system has at most $n+1$ linearly independent solutions.

10

Linear Systems with Constant Coefficients

In the preceding chapter, we established that the fundamental solution matrix $\Phi(x)$ completely characterises the geometric structure of the solution space for the homogeneous linear differential system (9.1). However, because $\Phi(x)$ generally only possesses theoretical existence for an arbitrary time-dependent matrix $A(x)$, the general variation of parameters formula (*theorem* 9.3) serves primarily as an abstract representation rather than a direct computational tool. We now restrict our analysis to systems governed by a constant matrix, where the fundamental matrix can be explicitly constructed.
A **linear differential system with constant coefficients** takes the form

$$\frac{d\mathbf{y}}{dx} = A\mathbf{y} + \mathbf{e}(x), \tag{10.1}$$

where $A \in \mathbb{R}^{n\times n}$ is an $n \times n$ constant matrix, and $\mathbf{e}(x)$ is a vector-valued function continuous on an interval (a, b).
By *theorem* 9.3, the core requirement for solving (10.1) is finding a fundamental solution matrix for the corresponding homogeneous system:

$$\frac{d\mathbf{y}}{dx} = A\mathbf{y}. \tag{10.2}$$

When $n = 1$, the matrix A reduces to a single real scalar a, and (10.2) simplifies to the scalar equation $\frac{dy}{dx} = ay$. The general solution to this scalar equation is generated by the strictly positive exponential function e^{ax}. To extend this analytic structure to $\mathbb{R}^n$, we must rigorously define the exponential of a square matrix.

10.1 Definition and Properties of the Matrix Exponential Function

Let $M_n(\mathbb{R})$ denote the vector space of all $n \times n$ real matrices, which possesses a natural dimension of n^2. To define convergence for sequences and series of matrices, we introduce a matrix norm.

Definition 10.1. ***The Entrywise ℓ^1-Norm on $M_n(\mathbb{R})$.***
For any matrix $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$, we define its norm as the absolute sum of all its entries:

$$\|A\| = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|. \tag{10.3}$$

This is the entrywise ℓ^1-norm (not the induced operator 1-norm).
定義

Proposition 10.1. ***Analytic Properties of the Matrix Norm.***
The mapping $\|\cdot\|\colon M_n(\mathbb{R}) \to [0,\infty)$ satisfies the standard properties of a norm and is sub-multiplicative. For any $A, B \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$, we have:

1. $\|A\| \geq 0$, with strict equality if and only if $A = O$ (the zero matrix).
2. $\|A + B\| \leq \|A\| + \|B\|$.
3. $\|cA\| = |c|\,\|A\|$.
4. $\|AB\| \leq \|A\|\|B\|$.

命題

Proof

Properties (1), (2), and (3) are immediate consequences of the scalar absolute value. To prove the sub-multiplicativity (4), we expand the components of the matrix product $C = AB$:

$$\begin{aligned}\|AB\| = \sum_{i=1}^{n}\sum_{j=1}^{n} |c_{ij}| &= \sum_{i=1}^{n}\sum_{j=1}^{n} \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \\ &\leq \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} |a_{ik}||b_{kj}| \\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} \left(|a_{ik}| \sum_{j=1}^{n} |b_{kj}|\right).\end{aligned}$$

Because the inner sum $\sum_{j=1}^{n} |b_{kj}|$ represents the absolute sum of the k-th row of B, it is bounded above by the total norm $\|B\|$. Substituting this uniform bound yields:

$$\|AB\| \leq \sum_{i=1}^{n}\sum_{k=1}^{n} |a_{ik}|\|B\| = \|A\|\|B\|.$$

■

An immediate inductive consequence of *proposition* 10.1 is the power bound $\|A^k\| \leq \|A\|^k$ for all integers $k \geq 1$. (We adopt the convention $A^0 = I_n$, the $n \times n$ identity matrix, where this inequality generally fails, but this boundary case does not impact sequential convergence). Because $M_n(\mathbb{R})$ is finite-dimensional, it is topologically complete; any Cauchy sequence of matrices converges strictly to a limit matrix within $M_n(\mathbb{R})$.

Proposition 10.2. ***Absolute Convergence of the Matrix Exponential.***
For any matrix $A \in M_n(\mathbb{R})$, the infinite matrix power series

$$I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{k!}A^k + \ldots$$

converges absolutely in $M_n(\mathbb{R})$.

命題

Proof

Taking the norm of the general term and applying the sub-multiplicative power bound gives:

$$\left\| \frac{1}{k!}A^k \right\| = \frac{1}{k!}\|A^k\| \leq \frac{1}{k!}\|A\|^k.$$

The sum of these upper bounds over all $k \geq 0$ is exactly the Taylor series for the scalar exponential $e^{\|A\|}$, which converges strictly for all real numbers $\|A\|$. By the Weierstrass M-test (the completeness of $M_n(\mathbb{R})$), the sequence of partial sums of the matrix series is absolutely Cauchy, and thus converges to a well-defined matrix.

■

Definition 10.2. ***Matrix Exponential.***
For any matrix $A \in M_n(\mathbb{R})$, the **matrix exponential** e^A (alternatively denoted $\exp(A)$) is defined as the unique limit of the absolutely convergent power series:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{10.4}$$

定義

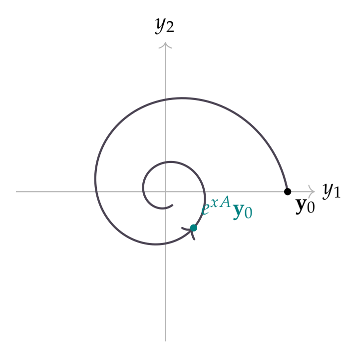

Figure 10.1: The matrix exponential evaluated at a real parameter x produces a linear operator e^{xA} that maps the initial state $\mathbf{y}_0$ continuously along the flow of the differential equation.

When A is a 1×1 matrix, (10.4) recovers the classical scalar exponential series. In higher dimensions, the matrix exponential preserves the fundamental algebraic structure of the scalar exponential, subject only to the constraints of matrix commutativity.

Proposition 10.3. ***Algebraic Properties of the Matrix Exponential.***
The matrix exponential function identically satisfies the following structural properties:

1. If matrices A and B commute (i.e., $AB = BA$), then

$$e^{A+B} = e^A e^B.$$

2. For any matrix A, the exponential e^A is everywhere invertible, and its exact inverse is
$$(e^A)^{-1} = e^{-A}.$$

3. If P is a non-singular $n \times n$ matrix, then for any matrix A,
$$e^{PAP^{-1}} = Pe^AP^{-1}.$$

命题

Proof

1. Because A and B commute, the binomial expansion strictly applies to arbitrary powers:
$$(A+B)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j}.$$
Because the series for e^A and e^B converge absolutely (*proposition* 10.2), we may compute their Cauchy product by uniformly rearranging the terms:
$$\begin{aligned} e^A e^B = \left(\sum_{i=0}^{\infty} \frac{A^i}{i!}\right)\left(\sum_{j=0}^{\infty} \frac{B^j}{j!}\right) &= \sum_{k=0}^{\infty}\sum_{j=0}^{k}\left(\frac{A^j}{j!}\right)\left(\frac{B^{k-j}}{(k-j)!}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j} \\ &= \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = e^{A+B}. \end{aligned}$$

2. A matrix A inherently commutes with its own negation $-A$, as $A(-A) = (-A)A = -A^2$. Applying Property 1 yields:
$$e^A e^{-A} = e^{A-A} = e^O = I_n.$$
Thus e^{-A} acts as both the left and right inverse of e^A, proving e^A is non-singular.

3. We observe inductively that $(PAP^{-1})^k = PA^kP^{-1}$ for any integer $k \geq 0$. Because matrix multiplication by a fixed matrix is a continuous linear operator on $M_n(\mathbb{R})$, we may factor P and P^{-1} out of the convergent infinite series:
$$\begin{aligned} e^{PAP^{-1}} = \sum_{k=0}^{\infty} \frac{(PAP^{-1})^k}{k!} &= \sum_{k=0}^{\infty} \frac{PA^kP^{-1}}{k!} \\ &= P\left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)P^{-1} = Pe^AP^{-1}. \end{aligned}$$

■

10.2 The Fundamental Solution Matrix via Matrix Exponentials

With the matrix exponential rigorously established, we construct the definitive fundamental solution matrix for any homogeneous linear differential system equipped with constant coefficients.

Theorem 10.1. ***The Matrix Exponential as a Fundamental Solution.*** The matrix-valued function $\Phi(x) = e^{xA}$ is a fundamental solution matrix for the constant-coefficient homogeneous system $\mathbf{y}' = A\mathbf{y}$. Furthermore, it satisfies the standard initial condition $\Phi(0) = I_n$.

定理

Proof

On any bounded interval $[-R, R]$, the infinite matrix power series

$$\Phi(x) = e^{xA} = I_n + xA + \frac{x^2}{2!}A^2 + \cdots + \frac{x^k}{k!}A^k + \ldots$$

converges uniformly. This uniform convergence rigorously justifies term-by-term differentiation with respect to the scalar independent variable x. Differentiating the series yields:

$$\begin{aligned}\frac{d}{dx}\Phi(x) &= \frac{d}{dx}I_n + \frac{d}{dx}(xA) + \frac{d}{dx}\left(\frac{x^2}{2!}A^2\right) + \cdots + \frac{d}{dx}\left(\frac{x^k}{k!}A^k\right) + \ldots \\ &= O + A + xA^2 + \cdots + \frac{x^{k-1}}{(k-1)!}A^k + \ldots \\ &= A\left(I_n + xA + \cdots + \frac{x^{k-1}}{(k-1)!}A^{k-1} + \ldots\right) \\ &= Ae^{xA} = A\Phi(x).\end{aligned}$$

Thus, $\Phi(x)$ identically satisfies the matrix differential equation. Evaluating the series at $x = 0$ trivially annihilates all higher-order terms, leaving $\Phi(0) = I_n$. Because its Wronskian determinant at the origin is $\det(I_n) = 1 \neq 0$, *theorem* 9.2 confirms that $\Phi(x)$ is a globally valid fundamental solution matrix.

■

The profound utility of *theorem* 10.1 becomes apparent when applied to the variation of parameters formula (*theorem* 9.3). Because the fundamental matrix is an exponential, the inverse matrix $\Phi^{-1}(s)$ is simply e^{-sA} (*proposition* 10.3). The matrix multiplication simplifies algebraically.

Corollary 10.1. ***Variation of Parameters for Constant Coefficients.*** The general solution to the constant-coefficient non-homogeneous linear differential system $\mathbf{y}' = A\mathbf{y} + \mathbf{e}(x)$ on an interval (a, b) is strictly given by:

$$\mathbf{y}(x) = e^{xA}\mathbf{c} + \int_{x_0}^{x} e^{(x-s)A}\mathbf{e}(s)\, ds, \tag{10.5}$$

where $\mathbf{c} \in \mathbb{R}^n$ is an arbitrary constant vector.
The unique solution satisfying the initial value problem $\mathbf{y}(x_0) = \mathbf{y}_0$ is:

$$\mathbf{y}(x) = e^{(x-x_0)A}\mathbf{y}_0 + \int_{x_0}^{x} e^{(x-s)A}\mathbf{e}(s)\, ds. \tag{10.6}$$

推論

Proof

Substituting $\Phi(x) = e^{xA}$ into the general integral formula (9.11) yields the integrand $\Phi(x)\Phi^{-1}(s)\mathbf{e}(s) = e^{xA}e^{-sA}\mathbf{e}(s)$. Because the matrices xA and $-sA$ are scalar multiples of the same matrix A, they strictly commute. By property (1) of *proposition* 10.3, we may consolidate the exponentials:

$$e^{xA}e^{-sA} = e^{(x-s)A}.$$

The representations (10.5) and (10.6) follow immediately.

■

The analytic problem is now entirely reduced to an algebraic one: expressing the infinite series defining e^{xA} in a finite, computable form using elementary functions. We demonstrate this reduction through two structural archetypes.

Example 10.1. Exponential of a Diagonal Matrix. Compute the matrix exponential function e^{xA} for the diagonal matrix

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}.$$

範例

Solution

Matrix multiplication dictates that any power A^k of a diagonal matrix simply exponentiates the individual elements along the principal diagonal:

$$A^k = \begin{bmatrix} a_1^k & 0 & \cdots & 0 \\ 0 & a_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^k \end{bmatrix}.$$

Substituting this into the definition of the matrix exponential (10.4) and operating component-wise yields:

$$e^{xA} = I_n + x \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix} + \frac{x^2}{2!} \begin{bmatrix} a_1^2 & & \\ & \ddots & \\ & & a_n^2 \end{bmatrix} + \dots$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{(a_1 x)^k}{k!} & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{(a_2 x)^k}{k!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{\infty} \frac{(a_n x)^k}{k!} \end{bmatrix} = \begin{bmatrix} e^{a_1 x} & 0 & \cdots & 0 \\ 0 & e^{a_2 x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{a_n x} \end{bmatrix}.$$

The differential system completely uncouples into n independent scalar equations.

■

Example 10.2. Exponential via Nilpotent Decomposition. Find the exact matrix exponential function e^{xA} for the non-diagonal matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

範例

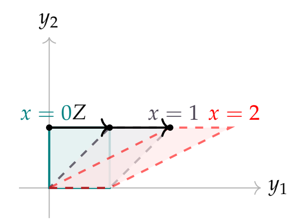

Figure 10.2: The geometric action of the nilpotent flow e^{xZ}. The matrix induces a continuous horizontal shear directly proportional to x, leaving the y_1-axis invariant.

Solution

The matrix A cannot be diagonalised. We instead decompose it into the sum of a scalar matrix and a strictly triangular matrix:

$$A = I_2 + Z, \quad \text{where} \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The identity matrix I_2 naturally commutes with Z. By *proposition* 10.3, we may factor the exponential mapping:

$$e^{xA} = e^{x(I_2+Z)} = e^{xI_2} e^{xZ}. \tag{10.7}$$

Applying the established result of example 10.1 to the identity matrix provides:

$$e^{xI_2} = \begin{bmatrix} e^x & 0 \\ 0 & e^x \end{bmatrix} = e^x I_2. \tag{10.8}$$

The matrix Z is **nilpotent** of index 2; explicit calculation shows $Z^2 = O$, which identically forces $Z^k = O$ for all integers $k \geq 2$. Consequently, the infinite power series for e^{xZ} strictly truncates after the linear term:

$$e^{xZ} = I_2 + xZ + \frac{x^2}{2!} O + \dots = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}. \tag{10.9}$$

This matrix geometrically represents a continuous shear transformation (*figure* 10.2).
Substituting (10.8) and (10.9) into (10.7) determines the complete exponential:

$$e^{xA} = (e^x I_2)\begin{bmatrix}1 & x\\0 & 1\end{bmatrix} = \begin{bmatrix}e^x & xe^x\\0 & e^x\end{bmatrix}.$$

The infinite series has thus been exactly compressed into a finite sum of elementary functions.

■

The decomposition methodology exhibited in example 10.2 possesses universal analytic significance. Over $\mathbb{C}$, any square matrix A can be transformed via a non-singular matrix P into its Jordan Canonical Form $J = P^{-1}AP$. (For real matrices, this may be carried out in the complexified space, or equivalently via the real Jordan form.) Each independent Jordan block fundamentally constitutes the sum of a scalar matrix λI and a nilpotent shift matrix Z. By strategically invoking property (3) of *proposition* 10.3, $e^{xA} = Pe^{xJ}P^{-1}$, we reduce the matrix exponential to polynomial-exponential blocks.

10.3 *Computation of the Fundamental Matrix via the Jordan Canonical Form*

The methodology exhibited in example 10.2 scales algebraically to matrices of arbitrary dimension. Over $\mathbb{C}$, while a general $n \times n$ matrix A may not be strictly diagonalisable, it is always similar to a block-diagonal **Jordan canonical form** J. Specifically, there exists an invertible transition matrix $P \in \mathbb{C}^{n\times n}$ (whose columns are generalised eigenvectors of A) such that

$$A = PJP^{-1}, \quad \text{where} \quad J = \begin{bmatrix}J_1 & 0 & \cdots & 0\\ 0 & J_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & J_m\end{bmatrix}. \tag{10.10}$$

When all eigenvalues are real, P and J may be taken in $\mathbb{R}^{n\times n}$.
Each Jordan block J_i is an $n_i \times n_i$ matrix corresponding to an eigenvalue λ_i of A (where $\sum_{i=1}^{m} n_i = n$). Crucially, every Jordan block can be linearly decomposed into the exact sum of a scalar matrix and a

standard nilpotent shift matrix:

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} = \lambda_i I_{n_i} + Z_{n_i},$$

where I_{n_i} is the identity matrix of dimension n_i, and Z_{n_i} is the nilpotent matrix carrying ones strictly on its first superdiagonal and zeros elsewhere. By definition, Z_{n_i} is nilpotent of index n_i, meaning $(Z_{n_i})^{n_i} = O$.

Because the scalar matrix $\lambda_i I_{n_i}$ natively commutes with any matrix of the same dimension, we may apply property (1) of *proposition* 10.3 to factor the exponential of the block:

$$e^{xJ_i} = e^{x(\lambda_i I_{n_i} + Z_{n_i})} = e^{\lambda_i x I_{n_i}} e^{xZ_{n_i}} = e^{\lambda_i x} e^{xZ_{n_i}}.$$

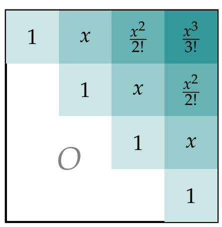

Figure 10.3: The upper-triangular band structure of the nilpotent exponential factor e^{xZ_4}. The truncated Taylor series perfectly fills the superdiagonals until reaching the matrix boundary.

The infinite power series for the nilpotent exponential strictly truncates after n_i terms. Calculating the successive powers of Z_{n_i} reveals that each power shifts the band of ones one position higher, until it annihilates off the top-right corner. The finite sum evaluates explicitly to:

$$e^{xZ_{n_i}} = I_{n_i} + xZ_{n_i} + \frac{x^2}{2!} Z_{n_i}^2 + \cdots + \frac{x^{n_i-1}}{(n_i-1)!} Z_{n_i}^{n_i-1} = \begin{bmatrix} 1 & x & \frac{x^2}{2!} & \cdots & \frac{x^{n_i-1}}{(n_i-1)!} \\ 0 & 1 & x & \cdots & \frac{x^{n_i-2}}{(n_i-2)!} \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & x \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Multiplying this upper-triangular matrix by the scalar exponential $e^{\lambda_i x}$ yields the exact elementary representation of e^{xJ_i}.

Because powers of block-diagonal matrices remain strictly block-diagonal, the exponential of the entire Jordan canonical form J is simply the block-diagonal assembly of the block exponentials:

$$e^{xJ} = \begin{bmatrix} e^{xJ_1} & 0 & \cdots & 0 \\ 0 & e^{xJ_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{xJ_m} \end{bmatrix}. \tag{10.11}$$

Finally, we apply property (3) of *proposition* 10.3 to map the solution space back into the original coordinates of the system:

$$e^{xA} = e^{x(PJP^{-1})} = Pe^{xJ}P^{-1}. \tag{10.12}$$

Formula (10.12) completely resolves the analytic problem of expressing e^{xA} using a finite sum of polynomial-exponential functions. However, computing the explicit matrix e^{xA} requires calculating the inverse matrix P^{-1}.

In the real-spectrum case (where P is real), we can systematically optimise this computation. By *proposition* 9.1, post-multiplying any fundamental solution matrix by a non-singular constant matrix generates another valid fundamental solution matrix. Because P is non-singular by definition, the matrix

$$\Psi(x) = e^{xA}P = (Pe^{xJ}P^{-1})P = Pe^{xJ} \tag{10.13}$$

is also a fundamental solution matrix for the homogeneous system (10.2).

> *Remark.*
>
> Utilising $\Psi(x) = Pe^{xJ}$ instead of e^{xA} strictly bypasses the algebraic inversion of P and eliminates one matrix multiplication. Nevertheless, calculating the generalised eigenvectors required to construct P and transforming A into J remains an operationally heavy spectral process. For large-dimensional systems, numerical instability in computing repeated eigenvalues often dictates the use of alternative algebraic methods, such as Putzer's algorithm, which evaluate the matrix exponential directly from the characteristic polynomial without invoking eigenvectors.

10.4 *Spectral Construction of Fundamental Matrices*

While the theoretical reduction of e^{xA} via the Jordan canonical form confirms that the fundamental solution matrix is built from exponential, polynomial, and (in the real form) trigonometric factors, calculating the transition matrix P and its inverse is computationally intensive. By directly exploiting the spectral properties of the coefficient matrix A—specifically its eigenvalues and eigenvectors—we can construct the columns of the fundamental solution matrix without global matrix inversion.

Diagonalisable Matrices

We first consider the case where A possesses a complete set of linearly independent eigenvectors. This inherently includes matrices with n distinct real eigenvalues, but also encompasses diagonalisable matrices with repeated eigenvalues.

Lemma 10.1. ***Eigenvector Solutions.***
The homogeneous linear differential system $\mathbf{y}' = A\mathbf{y}$ possesses a non-trivial solution of the form $\mathbf{y}(x) = e^{\lambda x}\mathbf{v}$ if and only if λ is an eigenvalue of A and $\mathbf{v}$ is an associated non-zero eigenvector.

引理

Proof

Direct substitution of the ansatz $\mathbf{y}(x) = e^{\lambda x}\mathbf{v}$ into the differential equation yields:

$$\frac{d}{dx}\left(e^{\lambda x}\mathbf{v}\right) = \lambda e^{\lambda x}\mathbf{v}, \quad \text{and} \quad A\left(e^{\lambda x}\mathbf{v}\right) = e^{\lambda x}A\mathbf{v}.$$

Equating these expressions and dividing by the non-zero scalar $e^{\lambda x}$ produces the algebraic condition $\lambda\mathbf{v} = A\mathbf{v}$, which is identically the eigenvalue equation $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$.

■

Theorem 10.2. ***Fundamental Matrix for Diagonalisable Systems.***
Let A be an $n \times n$ real constant matrix possessing n linearly independent *real* eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, corresponding to real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. The matrix-valued function

$$\Phi(x) = \begin{bmatrix} e^{\lambda_1 x}\mathbf{v}_1 & e^{\lambda_2 x}\mathbf{v}_2 & \cdots & e^{\lambda_n x}\mathbf{v}_n \end{bmatrix} \tag{10.14}$$

is a fundamental solution matrix for the system $\mathbf{y}' = A\mathbf{y}$.

定理

Proof

By *lemma* 10.1, each column of $\Phi(x)$ is a valid solution to the differential system. To establish that $\Phi(x)$ is a fundamental matrix, we evaluate its Wronskian determinant at the origin:

$$\det\Phi(0) = \det\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}.$$

Because the eigenvectors are strictly linearly independent by hypothesis, the columns of $\Phi(0)$ form a basis for $\mathbb{R}^n$, ensuring $\det\Phi(0) \neq 0$. *theorem* 9.2 immediately confirms that $\Phi(x)$ is a fundamental solution matrix everywhere on $\mathbb{R}$.

■

Complex Conjugate Eigenvalues

If the real matrix A possesses a complex eigenvalue $\lambda = \alpha + i\beta$, the corresponding eigenvector $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ is also generally complex. The real-eigenvector construction above no longer directly yields a real basis. We therefore pass temporarily to the complexified system and then extract real solutions using the reality of A.
If $\mathbf{y}(x) = \mathbf{p}(x) + i\mathbf{q}(x)$ is a complex solution to $\mathbf{y}' = A\mathbf{y}$ (with real A), then

$$\mathbf{p}'(x) + i\mathbf{q}'(x) = A\mathbf{p}(x) + iA\mathbf{q}(x).$$

Matching real and imaginary parts yields

$$\mathbf{p}'(x) = A\mathbf{p}(x), \quad \mathbf{q}'(x) = A\mathbf{q}(x),$$

so $\mathbf{p}$ and $\mathbf{q}$ are real solutions. Equivalently,

$$\mathbf{p}(x) = \frac{1}{2}(\mathbf{y} + \overline{\mathbf{y}}), \quad \mathbf{q}(x) = \frac{1}{2i}(\mathbf{y} - \overline{\mathbf{y}}).$$

For an eigen-solution, expanding $e^{(\alpha+i\beta)x}(\mathbf{u} + i\mathbf{w})$ via Euler's formula produces two linearly independent real solutions:

$$\mathbf{y}_{\mathrm{Re}}(x) = e^{\alpha x}(\cos(\beta x)\mathbf{u} - \sin(\beta x)\mathbf{w}), \tag{10.1}$$
$$\mathbf{y}_{\mathrm{Im}}(x) = e^{\alpha x}(\sin(\beta x)\mathbf{u} + \cos(\beta x)\mathbf{w}). \tag{10.2}$$

Example 10.3. Real Fundamental Matrix from Complex Eigenvalues. Construct the general real solution to the 2×2 differential system

$$\frac{d\mathbf{y}}{dx} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \mathbf{y}. \tag{10.15}$$

範例

Solution

The characteristic polynomial is $\det(A - \lambda I_2) = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2$. The roots are the complex conjugate pair $\lambda_{1,2} = 1 \pm i$.
We isolate the eigenvector $\mathbf{v}_1$ corresponding to $\lambda_1 = 1 + i$:

$$(A - (1+i)I_2)\mathbf{v}_1 = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Choosing $z_1 = 1$ forces $z_2 = i$, yielding the complex eigenvector $\mathbf{v}_1 = (1, i)^\top$. The associated complex solution is:

$$\mathbf{y}(x) = e^{(1+i)x} \begin{bmatrix} 1 \\ i \end{bmatrix} = e^x(\cos x + i \sin x) \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right).$$

Separating this into real and imaginary vector components provides

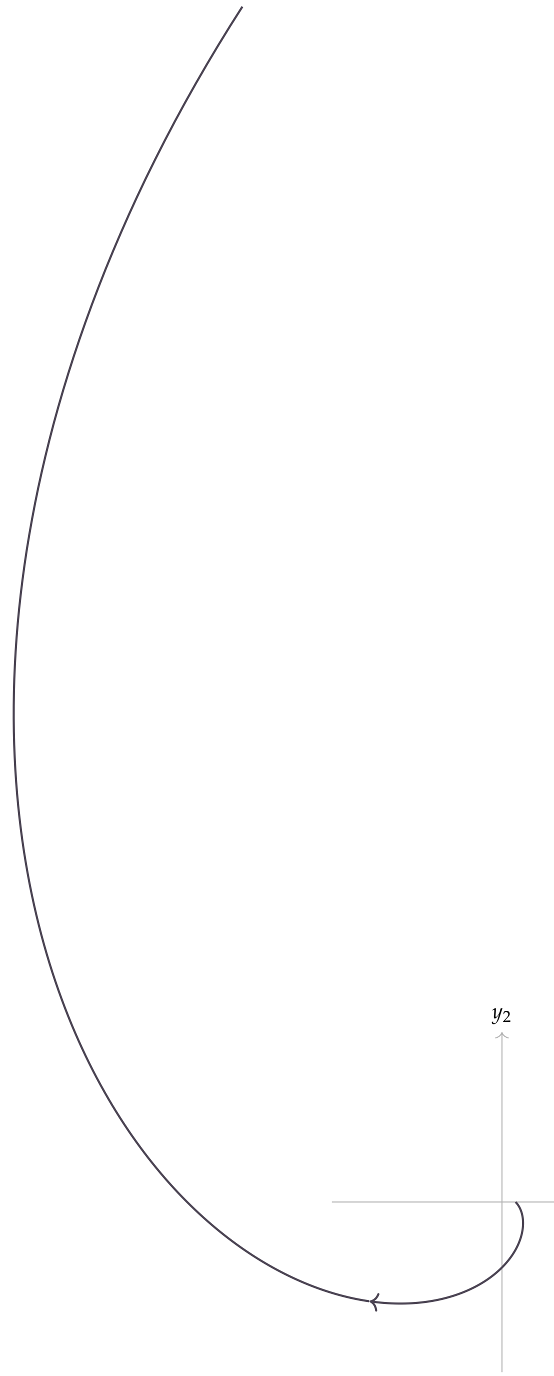

Figure 10.4: Phase portrait for the system in example 10.3. The complex eigenvalues generate an expanding spiral trajectory, combining exponential growth (e^x) with periodic rotation.

two linearly independent real solutions:

$$\mathbf{y}_{\text{Re}}(x) = e^x \left(\cos x \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \sin x \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = e^x \begin{bmatrix} \cos x \\ -\sin x \end{bmatrix},$$

$$\mathbf{y}_{\text{Im}}(x) = e^x \left(\sin x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \cos x \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = e^x \begin{bmatrix} \sin x \\ \cos x \end{bmatrix}.$$

The real general solution is strictly their arbitrary linear combination:

$$\mathbf{y}(x) = C_1 e^x \begin{bmatrix} \cos x \\ -\sin x \end{bmatrix} + C_2 e^x \begin{bmatrix} \sin x \\ \cos x \end{bmatrix}.$$

■

Defective Matrices and Generalised Eigenvectors

When the algebraic multiplicity of an eigenvalue strictly exceeds its geometric multiplicity (the dimension of its eigenspace), the matrix A is termed **defective**. It lacks a complete basis of eigenvectors, rendering *theorem* 10.2 insufficient to span the solution space.
To recover the missing dimensions, we invoke the polynomial structure identified in the Jordan canonical form (*figure* 10.3).

Lemma 10.2. ***Solutions from Generalised Eigenvectors.***
Let λ be an eigenvalue of A with algebraic multiplicity m. Suppose there exists a non-zero vector $\mathbf{r}_0$ satisfying the generalised eigenvector condition $(A - \lambda I_n)^m \mathbf{r}_0 = \mathbf{0}$.
If we recursively define the vector sequence

$$\mathbf{r}_k = (A - \lambda I_n)\mathbf{r}_{k-1}, \quad \text{for } k = 1, 2, \ldots, m-1, \tag{10.16}$$

then the vector polynomial function

$$\mathbf{y}(x) = e^{\lambda x} \left(\mathbf{r}_0 + x\mathbf{r}_1 + \frac{x^2}{2!}\mathbf{r}_2 + \cdots + \frac{x^{m-1}}{(m-1)!}\mathbf{r}_{m-1} \right) \tag{10.17}$$

is a strict solution to the differential system $\mathbf{y}' = A\mathbf{y}$.

引理

Proof

We differentiate the ansatz (10.17) directly:

$$\mathbf{y}'(x) = \lambda e^{\lambda x} \left(\mathbf{r}_0 + x\mathbf{r}_1 + \cdots + \frac{x^{m-1}}{(m-1)!}\mathbf{r}_{m-1} \right) + e^{\lambda x} \left(\mathbf{r}_1 + x\mathbf{r}_2 + \cdots + \frac{x^{m-2}}{(m-2)!}\mathbf{r}_{m-1} \right).$$

Multiplying the ansatz by A and applying the recurrence relation

$A\mathbf{r}_k = \lambda \mathbf{r}_k + \mathbf{r}_{k+1}$ gives:

$$A\mathbf{y}(x) = e^{\lambda x}\left(A\mathbf{r}_0 + xA\mathbf{r}_1 + \cdots + \frac{x^{m-1}}{(m-1)!}A\mathbf{r}_{m-1}\right)$$
$$= e^{\lambda x}\left[(\lambda\mathbf{r}_0 + \mathbf{r}_1) + x(\lambda\mathbf{r}_1 + \mathbf{r}_2) + \cdots + \frac{x^{m-1}}{(m-1)!}(\lambda\mathbf{r}_{m-1} + \mathbf{r}_m)\right].$$

By the termination condition of the general eigenspace, $\mathbf{r}_m = (A - \lambda I_n)^m \mathbf{r}_0 = \mathbf{0}$. The highest-order term truncates correctly. Matching coefficients of powers of x confirms that $\mathbf{y}'(x) \equiv A\mathbf{y}(x)$.

■

Theorem 10.3. ***Fundamental Matrix via Generalised Eigenvectors.*** Let $A \in \mathbb{C}^{n\times n}$ have distinct eigenvalues $\lambda_1, \ldots, \lambda_s$ of algebraic multiplicities $m_1, \ldots, m_s$ (where $\sum m_i = n$).
For each eigenvalue λ_i, choose vectors $\mathbf{r}_{j,0}^{(i)} \in \ker((A - \lambda_i I_n)^{m_i})$ such that the associated chains generated by (10.16) form a basis of the generalised eigenspace $\ker((A - \lambda_i I_n)^{m_i})$. Generating the polynomial solutions via *lemma* 10.2 for all these chains completely populates an $n \times n$ fundamental solution matrix $\Phi(x)$ for $\mathbf{y}' = A\mathbf{y}$ on $\mathbb{C}^n$.
If A is real and the relevant eigenvalues are real, this fundamental matrix is real-valued on $\mathbb{R}^n$; non-real eigenvalues are paired through the complex-conjugate reduction above.

定理

Proof

From linear algebra, the direct sum of the generalised eigenspaces exactly spans $\mathbb{C}^n$. Constructing $\Phi(x)$ according to (10.17) forces the evaluation at the origin to be $\Phi(0) = \begin{bmatrix}\mathbf{r}_{1,0}^{(1)} & \cdots & \mathbf{r}_{m_s,0}^{(s)}\end{bmatrix}$. Because these initial vectors form a complete basis for $\mathbb{C}^n$, $\det\Phi(0) \neq 0$, satisfying the same Wronskian invertibility criterion.

■

Example 10.4. Generalised Eigenvectors in a 3×3 System. Determine the general solution for the system:

$$\frac{d\mathbf{y}}{dx} = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{bmatrix}\mathbf{y}. \tag{10.18}$$

範例

Solution

We compute the characteristic polynomial:

$$\det(A-\lambda I_3) = \begin{vmatrix} 3-\lambda & 1 & 0 \\ -4 & -1-\lambda & 0 \\ 4 & -8 & -2-\lambda \end{vmatrix} = -(\lambda+2)(\lambda-1)^2.$$

The eigenvalues are $\lambda_1 = -2$ (algebraic multiplicity $m_1 = 1$) and $\lambda_2 = 1$ ($m_2 = 2$).
For the simple eigenvalue $\lambda_1 = -2$, we locate the standard eigenvector:

$$A-(-2)I_3 = \begin{bmatrix} 5 & 1 & 0 \\ -4 & 1 & 0 \\ 4 & -8 & 0 \end{bmatrix} \implies \mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

This generates the first basis solution $\mathbf{y}_1(x) = e^{-2x}\mathbf{v}_1$.
For the repeated eigenvalue $\lambda_2 = 1$, we must locate a basis for the generalised eigenspace $\ker((A-I_3)^2)$:

$$(A-I_3)^2 = \begin{bmatrix} 2 & 1 & 0 \\ -4 & -2 & 0 \\ 4 & -8 & -3 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 28 & 44 & 9 \end{bmatrix}.$$

We select two linearly independent vectors annihilated by this squared matrix:

$$\mathbf{r}_{1,0} = \begin{bmatrix} 11 \\ -7 \\ 0 \end{bmatrix}, \quad \mathbf{r}_{2,0} = \begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix}.$$

Applying the recurrence relation (10.16) calculates their first-order shifts:

$$\mathbf{r}_{1,1} = (A-I_3)\mathbf{r}_{1,0} = \begin{bmatrix} 2 & 1 & 0 \\ -4 & -2 & 0 \\ 4 & -8 & -3 \end{bmatrix} \begin{bmatrix} 11 \\ -7 \\ 0 \end{bmatrix} = \begin{bmatrix} 15 \\ -30 \\ 100 \end{bmatrix}.$$

$$\mathbf{r}_{2,1} = (A-I_3)\mathbf{r}_{2,0} = \begin{bmatrix} 2 & 1 & 0 \\ -4 & -2 & 0 \\ 4 & -8 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

By *lemma* 10.2, these vector chains construct the remaining two independent solutions:

$$\mathbf{y}_2(x) = e^x(\mathbf{r}_{1,0} + x\mathbf{r}_{1,1}) = e^x \begin{bmatrix} 11+15x \\ -7-30x \\ 100x \end{bmatrix},$$

$$\mathbf{y}_3(x) = e^x(\mathbf{r}_{2,0} + x\mathbf{r}_{2,1}) = e^x \begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix}.$$

The general solution is the arbitrary linear combination of $\mathbf{y}_1, \mathbf{y}_2$, and $\mathbf{y}_3$:

$$\mathbf{y}(x) = C_1 e^{-2x} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + C_2 e^{x} \begin{bmatrix} 11 + 15x \\ -7 - 30x \\ 100x \end{bmatrix} + C_3 e^{x} \begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix}.$$

■

10.5 *Exercises*

1. **General Solutions for Homogeneous Constant-Coefficient Systems.** Find the general solution of the constant-coefficient homogeneous linear differential system (10.2), where the matrix A is respectively:

(1)

$$A = \begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}.$$

(2)

$$A = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}.$$

(3)

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & -4 \end{bmatrix}.$$

(4)

$$A = \begin{bmatrix} 1 & 2 & -\frac{2}{3} \\ 0 & 2 & \frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{4}{3} \end{bmatrix}.$$

(5)

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

2. **General Solutions for Non-Homogeneous Constant-Coefficient Systems.** Find the general solution of the constant-coefficient non-homogeneous linear differential system (10.1), where:

(1)

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

(2)

$$A = \begin{bmatrix} 0 & -n^2 \\ -n^2 & 0 \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} \cos nx \\ \sin nx \end{bmatrix}.$$

Assume $n \neq 0$.

(3)

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} 0 \\ 2e^x \end{bmatrix}.$$

(4)

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -1 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} 2-x \\ 0 \\ 1-x \end{bmatrix}.$$

(5)

$$A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \mathbf{e}(x) = \begin{bmatrix} x^2 \\ 2x \\ x \end{bmatrix}.$$

3. **Initial Value Problems for Non-Homogeneous Systems.** Find the solution of (10.1) satisfying the initial condition $\mathbf{y}(0) = \boldsymbol{\eta}$, where:

(1)

$$A = \begin{bmatrix} -5 & -1 \\ 1 & -3 \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} e^x \\ e^{2x} \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

(2)

$$A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} 3x \\ 4 \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

(3)

$$A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} \sin x \\ -2\cos x \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(4)

$$A = \begin{bmatrix} 16 & 14 & 38 \\ -9 & -7 & -18 \\ -4 & -4 & -11 \end{bmatrix}, \quad \mathbf{e}(x) = \begin{bmatrix} -2e^{-x} \\ -3e^{-x} \\ 2e^{-x} \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

4. **A Two-Dimensional Spiral-Scaling System.** Solve the differential system

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

where a and b are real constants with $a \neq 0$ and $b \neq 0$ (thus excluding the pure-rotation special case from Exercise 1(2)).

5. **Asymptotic Vanishing Criterion via Spectrum.** Prove: every solution of the constant-coefficient homogeneous linear differential system (10.2) tends to zero as $x \to +\infty$ if and only if all eigenvalues of its coefficient matrix A have strictly negative real parts.

11
Higher-Order Linear Differential Equations

We now specialise the general order reduction procedure established in eq. (7.17) to the linear domain. This allows the structural theorems of linear systems to be applied directly to scalar equations of arbitrary order.

Let $a_1(x), a_2(x), \ldots, a_n(x)$ and $f(x)$ be continuous functions on an interval (a, b). The general n-th order linear differential equation is

$$y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_{n-1}(x)y' + a_n(x)y = f(x). \tag{11.1}$$

When $f(x)$ is not identically zero, (11.1) is a non-homogeneous linear differential equation. The corresponding homogeneous equation is

$$y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_{n-1}(x)y' + a_n(x)y = 0. \tag{11.2}$$

As established in (7.18), we introduce n state variables corresponding to the unknown function and its derivatives:

$$y_1 = y, \quad y_2 = y', \quad \ldots, \quad y_n = y^{(n-1)}. \tag{11.3}$$

Differentiating these definitions and isolating $y^{(n)}$ from (11.1) translates the scalar equation into an equivalent n-dimensional first-order linear system:

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y} + \mathbf{f}(x), \tag{11.4}$$

where the state vector $\mathbf{y}$ and the forcing vector $\mathbf{f}(x)$ are

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(x) \end{bmatrix},$$

and the coefficient matrix $A(x)$ is the companion matrix

$$A(x) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n(x) & -a_{n-1}(x) & -a_{n-2}(x) & \cdots & -a_1(x) \end{bmatrix}. \quad (11.5)$$

The homogeneous scalar equation (11.2) correspondingly transforms into the homogeneous system

$$\frac{d\mathbf{y}}{dx} = A(x)\mathbf{y}. \quad (11.6)$$

Because the system formulation (11.4) is strictly linear, all analytic and algebraic results derived for first-order linear systems natively govern the n-th order scalar equation. The most immediate consequence is the absolute global existence of solutions.

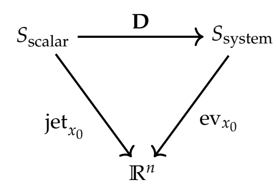

Figure 11.1: For the homogeneous equation, the strict linear isomorphisms connect the scalar solution space S_{scalar}, the companion system solution space S_{system}, and $\mathbb{R}^n$. Here $\mathbf{D}y = (y, y', \ldots, y^{(n-1)})^\top$, $\text{jet}_{x_0}(y) = (y(x_0), y'(x_0), \ldots, y^{(n-1)}(x_0))^\top$, and $\text{ev}_{x_0}(\mathbf{y}) = \mathbf{y}(x_0)$.

Theorem 11.1. ***Global Existence and Uniqueness for Higher-Order Equations.***
Let the coefficients $a_1(x), \ldots, a_n(x)$ and the forcing term $f(x)$ be continuous on (a, b). For any base point $x_0 \in (a, b)$ and any prescribed initial constants $y_0, y_0', \ldots, y_0^{(n-1)} \in \mathbb{R}$, the initial value problem for the differential equation (11.1) satisfying

$$y(x_0) = y_0, \quad y'(x_0) = y_0', \quad \ldots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}$$

admits a unique solution $y = \varphi(x)$ defined on the entire interval (a, b).

定理

Proof

By hypothesis, the elements of the companion matrix $A(x)$ and the vector $\mathbf{f}(x)$ consist exclusively of constants, $-a_k(x)$, and $f(x)$. Since these functions are continuous on (a, b), $A(x)$ and $\mathbf{f}(x)$ are globally continuous on (a, b).
Applying *proposition* 7.3 to the equivalent system (11.4), there exists a unique global vector solution $\mathbf{y}(x)$ satisfying the initial condition $\mathbf{y}(x_0) = (y_0, y_0', \ldots, y_0^{(n-1)})^\top$. Extracting the first component $y_1(x)$ yields the unique continuous solution to (11.1) that preserves the required initial derivatives.

■

Corollary 11.1. ***Dimension of the Scalar Solution Space.***
The set of all solutions to the homogeneous n-th order linear differential equation (11.2) forms a vector space of dimension n.

推論

Proof

By *theorem* 9.1, the solution space of the equivalent homogeneous system (11.6) is exactly n-dimensional. The differential mapping $y \mapsto (y, y', \ldots, y^{(n-1)})^\top$ is a linear isomorphism between the scalar solutions of (11.2) and the vector solutions of (11.6). Isomorphisms strictly preserve dimension.

■

This systemic reduction also dictates that the linear independence of n scalar solutions $y_1(x), \ldots, y_n(x)$ is completely evaluated by the Wronskian determinant (*definition* 9.2) of their corresponding state vectors. Formally, the standard scalar Wronskian

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{vmatrix}$$

is structurally identical to the determinant of the fundamental solution matrix for the companion system. Consequently, the independence criteria developed in *theorem* 9.2 strictly govern higher-order scalar equations without requiring separate proofs.

11.1 General Theory of Higher-Order Linear Differential Equations

Assume that the n scalar functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ are solutions to the homogeneous linear differential equation (11.2). As established in the preceding section, the linear independence of these scalar functions on an interval (a, b) is strictly equivalent to the linear independence of their corresponding state vectors in $\mathbb{R}^n$.
Consequently, we may directly inherit the algebraic structure of the solution space from *theorem* 9.1. A set of n linearly independent solutions forms a **fundamental set of solutions**, and the general solution to (11.2) is uniquely expressed as their linear combination:

$$y(x) = C_1\varphi_1(x) + C_2\varphi_2(x) + \cdots + C_n\varphi_n(x),$$

where $C_1, \ldots, C_n$ are arbitrary real constants.
The Wronskian determinant (*definition* 9.2) of these scalar solutions takes the explicit form

$$W(x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \cdots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \cdots & \varphi_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \cdots & \varphi_n^{(n-1)}(x) \end{vmatrix}. \tag{11.7}$$

Denote by $\mathcal{W}(x)$ the matrix appearing in (11.7), so $W(x) = \det \mathcal{W}(x)$. By *theorem* 9.2, the solutions constitute a fundamental set if and only if $W(x)$ is nowhere zero on (a, b).

Remark.

By inspecting the companion matrix $A(x)$ defined in (11.5), we observe that its trace is strictly the negative of the second leading coefficient of the differential equation: $\operatorname{tr}[A(x)] = -a_1(x)$. Substituting this into Liouville's Formula (*lemma* 9.2) yields a highly tractable scalar identity:

$$W(x) = W(x_0) \exp\left(-\int_{x_0}^{x} a_1(s)\, ds \right). \tag{11.8}$$

For second-order equations ($n = 2$), (11.8) geometrically tracks the expanding or contracting area of the solution span in the phase plane (*figure* 11.2). Analytically, it permits the complete construction of the general solution from a single known non-zero solution.

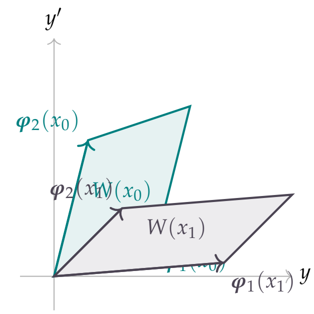

Figure 11.2: Geometric interpretation of the Wronskian for a second-order equation. The determinant $W(x)$ measures the oriented area of the parallelogram spanned by the state vectors in the phase plane. Liouville's formula dictates the precise volumetric evolution of this area under the flow.

Example 11.1. Reduction of Order for Second-Order Equations. Let $y = \varphi(x)$ be a known, non-zero solution to the second-order homogeneous linear differential equation

$$y'' + p(x)y' + q(x)y = 0, \tag{11.9}$$

where $p(x)$ and $q(x)$ are continuous on (a, b). Determine its general solution.

範例

Solution

For simplicity, we restrict our analysis to sub-intervals where $\varphi(x) \neq 0$. Let $y(x)$ be an arbitrary linearly independent solution to (11.9). We construct their Wronskian:

$$W(x) = \begin{vmatrix} \varphi(x) & y(x) \\ \varphi'(x) & y'(x) \end{vmatrix} = \varphi(x)y'(x) - \varphi'(x)y(x).$$

By the scalar Liouville formula (11.8), the Wronskian also satisfies:

$$W(x) = C \exp\left(-\int_{x_0}^{x} p(s)\, ds \right),$$

for some non-zero constant C. Equating the two expressions and dividing both sides by the integrating factor $\varphi^2(x)$ isolates the exact derivative of the ratio y/φ:

$$\frac{\varphi(x)y'(x) - \varphi'(x)y(x)}{\varphi^2(x)} = \frac{d}{dx}\left(\frac{y(x)}{\varphi(x)} \right) = \frac{C}{\varphi^2(x)} \exp\left(-\int_{x_0}^{x} p(s)\, ds \right).$$

Integrating this relation with respect to x and multiplying by $\varphi(x)$ recovers the second solution. Absorbing the base constants yields the general solution:

$$y(x) = \varphi(x)\left[C_1 + C_2\int_{x_0}^{x}\frac{1}{\varphi^2(s)}\exp\left(-\int_{x_0}^{s}p(t)\,dt\right)ds\right].$$

■

We now leverage the general variation of parameters formula (*theorem* 9.3) to explicitly construct particular solutions for higher-order non-homogeneous equations.

Theorem 11.2. ***Variation of Parameters for Higher-Order Equations.***

Let $\varphi_1(x), \dots, \varphi_n(x)$ be a fundamental set of solutions to the homogeneous equation (11.2) on (a, b), with Wronskian $W(x)$. The general solution to the non-homogeneous equation (11.1) is

$$y(x) = \sum_{k=1}^{n} C_k\varphi_k(x) + \varphi^*(x), \tag{11.10}$$

where the particular solution $\varphi^*(x)$ is strictly given by the integral formula:

$$\varphi^*(x) = \sum_{k=1}^{n}\varphi_k(x)\int_{x_0}^{x}\frac{W_k(s)}{W(s)}f(s)\,ds. \tag{11.11}$$

Here, $W_k(x)$ denotes the algebraic cofactor corresponding to the (n, k)-th entry of the Wronskian matrix $\mathcal{W}(x)$. Equivalently, $W_k(x)$ is the determinant obtained by replacing the k-th column of $\mathcal{W}(x)$ with the basis vector $(0, \dots, 0, 1)^\top$.

定理

Proof

We apply *theorem* 9.3 to the equivalent differential system (11.4). The fundamental matrix $\Phi(x)$ is precisely $\mathcal{W}(x)$, and its determinant is the scalar Wronskian $W(x)$. The general vector solution is

$$\mathbf{y}(x) = \Phi(x)\mathbf{c} + \Phi(x)\int_{x_0}^{x}\Phi^{-1}(s)\mathbf{f}(s)\,ds.$$

The scalar solution $y(x)$ corresponds strictly to the first component of $\mathbf{y}(x)$. The homogeneous component $\Phi(x)\mathbf{c}$ trivially projects to the linear combination $\sum C_k\varphi_k(x)$. We must verify that the first component of the particular integral evaluates to (11.11).
By Cramer's rule, the vector $\mathbf{u}(s) = \Phi^{-1}(s)\mathbf{f}(s)$ is the unique solution to the linear system $\Phi(s)\mathbf{u}(s) = \mathbf{f}(s)$. Because the forcing vector $\mathbf{f}(s)$ consists of zeros in its first $n-1$ entries and $f(s)$ in the n-th entry, the k-th component of $\mathbf{u}(s)$ is exactly:

$$u_k(s) = \frac{W_k(s)}{W(s)}f(s).$$

The integral of the vector $\mathbf{u}(s)$ is performed component-wise. Finally, multiplying by the fundamental matrix $\Phi(x)$ to extract the first row yields the linear combination of the top-row elements $\varphi_k(x)$ with the integrated coefficients:

$$\varphi^*(x) = \sum_{k=1}^{n} \varphi_k(x) \int_{x_0}^{x} u_k(s)\, ds = \sum_{k=1}^{n} \varphi_k(x) \int_{x_0}^{x} \frac{W_k(s)}{W(s)} f(s)\, ds.$$

■

Example 11.2. Variation of Parameters for Second-Order Equations. Derive the explicit variation of parameters formula for the second-order non-homogeneous linear differential equation

$$y'' + p(x)y' + q(x)y = f(x), \tag{11.12}$$

given a fundamental set of solutions $\varphi_1(x)$ and $\varphi_2(x)$ to the corresponding homogeneous equation.

範例

Solution

We execute the classical derivation, which locally assumes a dynamic parameter ansatz $y(x) = C_1(x)\varphi_1(x) + C_2(x)\varphi_2(x)$. Differentiating this ansatz yields:

$$y'(x) = \left[C_1(x)\varphi_1'(x) + C_2(x)\varphi_2'(x)\right] + \left[C_1'(x)\varphi_1(x) + C_2'(x)\varphi_2(x)\right].$$

To prevent the introduction of second derivatives of the unknown parameters, we impose the strict algebraic constraint that the second bracket identically vanishes:

$$C_1'(x)\varphi_1(x) + C_2'(x)\varphi_2(x) = 0. \tag{11.13}$$

This reduces the first derivative to $y'(x) = C_1(x)\varphi_1'(x) + C_2(x)\varphi_2'(x)$. Differentiating once more produces:

$$y''(x) = C_1(x)\varphi_1''(x) + C_2(x)\varphi_2''(x) + C_1'(x)\varphi_1'(x) + C_2'(x)\varphi_2'(x).$$

Substituting y, y', and y'' into the non-homogeneous equation (11.12) and regrouping by $C_1(x)$ and $C_2(x)$ yields:

$$\begin{aligned} &C_1(x)\left[\varphi_1''(x) + p(x)\varphi_1'(x) + q(x)\varphi_1(x)\right] \\ &+ C_2(x)\left[\varphi_2''(x) + p(x)\varphi_2'(x) + q(x)\varphi_2(x)\right] \\ &+ C_1'(x)\varphi_1'(x) + C_2'(x)\varphi_2'(x) = f(x). \end{aligned}$$

Because φ_1 and φ_2 are exact homogeneous solutions, the bracketed terms identically vanish. We are left with the second constraint:

$$C_1'(x)\varphi_1'(x) + C_2'(x)\varphi_2'(x) = f(x). \tag{11.14}$$

Equations (11.13) and (11.14) constitute a 2×2 linear algebraic system for the derivatives $C_1'(x)$ and $C_2'(x)$. The determinant of this system is precisely the Wronskian $W(x) = \varphi_1(x)\varphi_2'(x) - \varphi_2(x)\varphi_1'(x)$. Applying Cramer's rule yields:

$$C_1'(x) = \frac{-\varphi_2(x)f(x)}{W(x)}, \quad C_2'(x) = \frac{\varphi_1(x)f(x)}{W(x)}.$$

Direct integration from a base point x_0 determines the parameters. Substituting these back into the original ansatz provides the exact particular solution:

$$\varphi^*(x) = \varphi_1(x)\int_{x_0}^{x} \frac{-\varphi_2(s)f(s)}{W(s)}\,ds + \varphi_2(x)\int_{x_0}^{x} \frac{\varphi_1(s)f(s)}{W(s)}\,ds.$$

Factoring the integrand over a common integral generates the consolidated formula for the general solution:

$$y(x) = C_1\varphi_1(x) + C_2\varphi_2(x) + \int_{x_0}^{x} \frac{\varphi_1(s)\varphi_2(x) - \varphi_2(s)\varphi_1(x)}{\varphi_1(s)\varphi_2'(s) - \varphi_2(s)\varphi_1'(s)} f(s)\,ds. \tag{11.15}$$

This explicitly confirms the abstract matrix projection derived in *theorem* 11.2 for the $n = 2$ case.

■

11.2 *Linear Differential Equations with Constant Coefficients*

We now restrict our attention to the n-th order linear differential equation where the coefficient functions are strictly constant:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = f(x), \tag{11.16}$$

and its corresponding homogeneous equation:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0. \tag{11.17}$$

Here, $a_1, a_2, \ldots, a_n \in \mathbb{R}$, and the forcing term $f(x)$ is a real-valued continuous function on (a, b).
Because the differential operator is governed by constant coefficients, the companion matrix A defined in (11.5) is a constant matrix in $\mathbb{R}^{n\times n}$. The problem of determining a fundamental set of solutions for (11.17) reduces entirely to calculating the matrix exponential e^{xA} and extracting the top row of the resulting fundamental solution matrix, as established in *theorem* 10.1.

The Characteristic Equation and Spectral Solutions

To diagonalise the companion system $\mathbf{y}' = A\mathbf{y}$, we must determine the eigenvalues of the companion matrix A. The characteristic determinant is:

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda + a_1 \end{vmatrix}. \tag{11.18}$$

Expanding this determinant along the bottom row (or applying induction) yields the characteristic polynomial of A:

$$P(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0. \tag{11.19}$$

Equation (11.19) is termed the **characteristic equation** of the differential equation (11.17). Notice that it can be constructed trivially by applying the algebraic substitution $y^{(k)} \mapsto \lambda^k$ into the differential equation.

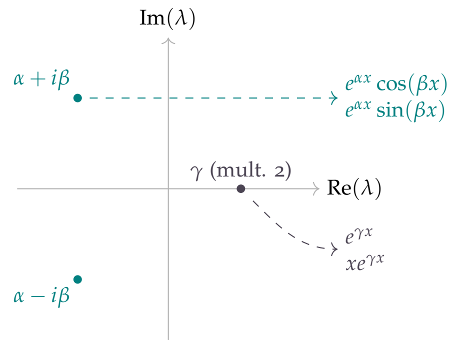

Figure 11.3: The mapping from the spectral roots of the characteristic equation (11.19) to the real fundamental basis functions. Complex roots strictly appear in conjugate pairs, spanning oscillatory solutions.

Theorem 11.3. ***Fundamental Solutions for Constant Coefficients.*** Let the characteristic equation (11.19) possess s distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_s$ in $\mathbb{C}$, with algebraic multiplicities $m_1, m_2, \ldots, m_s$ respectively (where $\sum m_k = n$). The collection of n functions

$$\bigcup_{k=1}^{s} \left\{ e^{\lambda_k x},\ xe^{\lambda_k x},\ \ldots,\ x^{m_k-1}e^{\lambda_k x} \right\} \tag{11.20}$$

constitutes a fundamental set of solutions for the homogeneous equation (11.17) in the complex solution space.

定理

Proof

Let $D = \frac{d}{dx}$ and define the operator

$$L = D^n + a_1 D^{n-1} + \cdots + a_n,$$

so (11.17) is exactly $L[y] = 0$. Over $\mathbb{C}$, the characteristic polynomial factorisation gives

$$P(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = \prod_{k=1}^{s}(\lambda - \lambda_k)^{m_k},$$

and therefore

$$L = \prod_{k=1}^{s}(D - \lambda_k)^{m_k}.$$

For fixed k and $j = 0, 1, \ldots, m_k - 1$, set $y_{k,j}(x) = x^j e^{\lambda_k x}$. Using

$$(D - \lambda_k)\big(e^{\lambda_k x}u(x)\big) = e^{\lambda_k x}u'(x),$$

repeated application yields

$$(D - \lambda_k)^{m_k} y_{k,j} = e^{\lambda_k x}\frac{d^{m_k}}{dx^{m_k}}(x^j) = 0 \quad (m_k > j).$$

Hence each $y_{k,j}$ satisfies $L[y_{k,j}] = 0$, so every function in (11.20) is a solution.

To prove linear independence, assume

$$\sum_{k=1}^{s} \sum_{j=0}^{m_k-1} c_{k,j} x^j e^{\lambda_k x} = 0.$$

Fix $r \in \{1, \dots, s\}$ and apply

$$Q_r(D) = \prod_{\substack{k=1 \\ k \neq r}}^{s} (D - \lambda_k)^{m_k}.$$

Every term with index $k \neq r$ is annihilated, so

$$Q_r(D) \left(\sum_{j=0}^{m_r-1} c_{r,j} x^j e^{\lambda_r x} \right) = 0.$$

On the space $e^{\lambda_r x} \mathcal{P}_{m_r-1}$ (where $\mathcal{P}_{m_r-1}$ denotes polynomials of degree at most $m_r - 1$), each factor $(D - \lambda_k)$ acts as $D + (\lambda_r - \lambda_k)$ on polynomials. Since $\lambda_r - \lambda_k \neq 0$, each factor is invertible, hence $Q_r(D)$ is invertible on this finite-dimensional space. Therefore

$$\sum_{j=0}^{m_r-1} c_{r,j} x^j e^{\lambda_r x} = 0,$$

forcing $c_{r,j} = 0$ for all j. Because r is arbitrary, every coefficient vanishes. Thus the n functions in (11.20) are linearly independent, and therefore form a complex fundamental set.

■

Because $a_1, \dots, a_n$ are real, any non-real root appears with its conjugate. If $\lambda = \alpha \pm i\beta$ is such a pair, the corresponding complex solutions $x^j e^{(\alpha \pm i\beta)x}$ can be replaced by the real pair

$$x^j e^{\alpha x} \cos(\beta x) \quad \text{and} \quad x^j e^{\alpha x} \sin(\beta x),$$

for each $j = 0, 1, \dots, m-1$ (where m is the multiplicity of $\alpha + i\beta$). This replacement yields a real fundamental set (*figure* 11.3).

Example 11.3. Higher-Order Equation with Repeated and Complex Roots. Find the general solution to the homogeneous equation

$$y^{(4)} - 4y''' + 13y'' - 36y' + 36y = 0.$$

範例

Solution

We immediately write the characteristic equation:

$$P(\lambda) = \lambda^4 - 4\lambda^3 + 13\lambda^2 - 36\lambda + 36 = 0.$$

By inspection or the rational root theorem, $\lambda = 2$ is a root. Factoring out $(\lambda - 2)$ yields:

$$P(\lambda) = (\lambda - 2)(\lambda^3 - 2\lambda^2 + 9\lambda - 18) = (\lambda - 2)^2(\lambda^2 + 9) = 0.$$

The roots are $\lambda_1 = 2$ (with multiplicity $m_1 = 2$) and the purely imaginary conjugate pair $\lambda_{2,3} = \pm 3i$ (multiplicity $m_2 = 1$).
By *theorem* 11.3, the repeated real root generates the basis functions e^{2x} and xe^{2x}. The complex roots yield $\cos(3x)$ and $\sin(3x)$. The general real solution is their linear combination:

$$y(x) = (C_1 + C_2 x)e^{2x} + C_3 \cos(3x) + C_4 \sin(3x).$$

■

The Method of Undetermined Coefficients

While the variation of parameters formula (*theorem* 11.2) universally integrates any continuous forcing function $f(x)$, it often demands the evaluation of non-elementary integrals. When $f(x)$ takes the specific form of polynomials multiplying exponentials and sinusoids, we may construct a particular solution $\varphi^*(x)$ algebraically by predicting its structural form.
Let the differential equation be written as

$$L[y] = P(D)y = f(x), \quad D = \frac{d}{dx},$$

where

$$P(D) = D^n + a_1 D^{n-1} + \cdots + a_n$$

is the constant-coefficient differential operator associated with (11.19).
We classify the forcing function into two primary structural archetypes:

Case 1: Exponential-Polynomial Forcing Suppose $f(x) = Q_m(x)e^{\mu x}$, where $Q_m(x)$ is a polynomial of degree m.

- If μ is *not* a root of the characteristic equation $P(\lambda) = 0$, the particular solution takes the form:

  $$\varphi^*(x) = R_m(x)e^{\mu x},$$

  where $R_m(x)$ is a general polynomial of the same degree m.

- If μ *is* a root of $P(\lambda) = 0$ with algebraic multiplicity k, the solution space experiences resonance, and we must elevate the polynomial degree by the multiplicity:

$$\varphi^*(x) = x^k R_m(x) e^{\mu x}.$$

Case 2: Sinusoidal Forcing Suppose $f(x) = [A_m(x)\cos(\beta x) + B_l(x)\sin(\beta x)]e^{\alpha x}$.

- We define the complex frequency $\mu = \alpha + i\beta$.
- Let $M = \max\{m, l\}$. If μ is a root of $P(\lambda) = 0$ with multiplicity k (where $k = 0$ indicates it is not a root), the particular solution takes the form:

  $$\varphi^*(x) = x^k [C_M(x)\cos(\beta x) + D_M(x)\sin(\beta x)]e^{\alpha x},$$

  where $C_M(x)$ and $D_M(x)$ are general polynomials of degree M.

The undetermined coefficients of these general polynomials are resolved by directly substituting the ansatz $\varphi^*(x)$ back into the differential equation and equating identical functional components.

Example 11.4. Undetermined Coefficients with Resonance. Solve the non-homogeneous equation

$$y'' + 6y' + 9y = 12e^{-3x}. \tag{11.21}$$

範例

Solution

The characteristic equation is $\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2 = 0$. The homogeneous root is $\lambda = -3$ with multiplicity $k = 2$. The complementary solution is therefore $y_c(x) = (C_1 + C_2 x)e^{-3x}$.
The forcing function is $f(x) = 12e^{-3x}$. Here, the forcing exponent is $\mu = -3$ and the polynomial is a constant $Q_0(x) = 12$ (degree $m = 0$). Because $\mu = -3$ identically matches the homogeneous root with multiplicity $k = 2$, we are in a state of double resonance. The predicted particular solution must be multiplied by x^2:

$$\varphi^*(x) = x^2(A)e^{-3x} = Ax^2e^{-3x}.$$

We differentiate the ansatz to substitute it into (11.21):

$$(\varphi^*)' = A(2x - 3x^2)e^{-3x},$$
$$(\varphi^*)'' = A(2 - 12x + 9x^2)e^{-3x}.$$

Substituting these into the left-hand side of the differential equation:

$$A(2 - 12x + 9x^2)e^{-3x} + 6A(2x - 3x^2)e^{-3x} + 9Ax^2e^{-3x} = 12e^{-3x}.$$

Factoring out Ae^{-3x} and gathering the polynomial terms:

$$A[(9 - 18 + 9)x^2 + (-12 + 12)x + 2] = 12 \implies 2A = 12 \implies A = 6.$$

The particular solution is exactly $\varphi^*(x) = 6x^2e^{-3x}$. Combining this with the complementary function yields the general solution:

$$y(x) = (C_1 + C_2x + 6x^2)e^{-3x}.$$

■

Algebraic Elimination of Coupled Systems

The operational equivalence between n-th order scalar equations and n-dimensional linear systems is fully bidirectional. For systems of differential equations with constant coefficients, we can frequently exploit algebraic elimination to decouple the system, reducing it to a single higher-order scalar equation for one of the state variables.

Example 11.5. Decoupling a 2×2 System via Elimination. Determine the general solution for the coupled autonomous system:

$$\begin{cases} \frac{dx}{dt} = 4x - 2y, \\ \frac{dy}{dt} = 5x + 2y. \end{cases} \tag{11.22}$$

範例

Solution

We aim to eliminate $y(t)$ and construct a second-order scalar differential equation strictly in terms of $x(t)$. From the first equation, we algebraically isolate y:

$$2y = 4x - \frac{dx}{dt}. \tag{11.23}$$

Differentiating this relation with respect to t provides the first derivative of y:

$$2\frac{dy}{dt} = 4\frac{dx}{dt} - \frac{d^2x}{dt^2}. \tag{11.24}$$

We double the second equation of the original system (11.22) to align the coefficients: $2\frac{dy}{dt} = 10x + 2(2y)$. Substituting (11.23) and (11.24) into this scaled equation yields an equation exclusively in x:

$$4\frac{dx}{dt} - \frac{d^2x}{dt^2} = 10x + 2\left(4x - \frac{dx}{dt}\right).$$

Expanding and rearranging terms isolates the standard form of the homogeneous scalar equation:

$$\frac{d^2x}{dt^2} - 6\frac{dx}{dt} + 18x = 0.$$

The characteristic equation is $\lambda^2 - 6\lambda + 18 = 0$. Completing the square, $(\lambda - 3)^2 + 9 = 0$, yielding the complex conjugate roots $\lambda =$

$3 \pm 3i$. By *theorem* 11.3, the general solution for $x(t)$ is:

$$x(t) = e^{3t}(C_1 \cos(3t) + C_2 \sin(3t)). \tag{11.25}$$

To recover $y(t)$, we simply differentiate (11.25) and substitute the result back into the algebraic isolation (11.23), circumventing any further integration:

$$x'(t) = e^{3t}[(3C_1 + 3C_2)\cos(3t) + (-3C_1 + 3C_2)\sin(3t)].$$

Substituting $x(t)$ and $x'(t)$ into $2y = 4x - x'$ and dividing by 2 exactly determines the corresponding state variable:

$$y(t) = \frac{1}{2}e^{3t}[(C_1 - 3C_2)\cos(3t) + (3C_1 + C_2)\sin(3t)].$$

The arbitrary constants C_1 and C_2 govern both components, perfectly reflecting the two-dimensional nature of the original system's solution space.

■

11.3 *Exercises*

1. **Companion Reduction for a Third-Order Equation.** Consider the non-homogeneous equation

$$y''' + a_1(x)y'' + a_2(x)y' + a_3(x)y = f(x),$$

where all coefficients and the forcing term are continuous on (a, b).

(a) Introduce $y_1 = y$, $y_2 = y'$, $y_3 = y''$, and write the equivalent first-order system in matrix form $\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$.

(b) Prove directly that the mapping

$$y \mapsto (y, y', y'')^\top$$

sends scalar solutions to system solutions, and that the first component of any system solution solves the original scalar equation.

(c) Deduce that for any $x_0 \in (a, b)$ and any prescribed $y(x_0), y'(x_0), y''(x_0)$, the corresponding initial value problem has a unique solution on (a, b).

2. **Direct Derivation of the Scalar Liouville Formula.** Prove (11.8) directly for the second-order equation (11.9) by differentiating the Wronskian

$$W(x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{vmatrix}$$

and showing that $W'(x) = -p(x)W(x)$.

3. **A Second-Order Equation with Zero First-Derivative Term.** Consider the differential equation

$$y'' + q(x)y = 0.$$

   (a) Let $y = \varphi(x)$ and $y = \psi(x)$ be any two solutions. Prove that the Wronskian determinant of φ and ψ is constant.
   (b) Given that the equation has a particular solution $y = e^x$, find the general solution and determine $q(x)$.

4. **Zeros of Non-Trivial Solutions.** Consider the differential equation (11.9), where $p(x)$ and $q(x)$ are continuous on an interval $I = (a, b)$.

   (a) Let $y = \varphi(x)$ be a non-zero solution on I. Prove that every zero of φ is simple; that is, if $\varphi(x_0) = 0$ for some $x_0 \in I$, then $\varphi'(x_0) \neq 0$. Deduce that on every finite subinterval of I, the zeros of φ are isolated and hence finite in number.
   (b) Generalise this uniqueness phenomenon to the higher-order equation (11.2): prove that if a solution y satisfies

$$y(x_0) = y'(x_0) = \cdots = y^{(n-1)}(x_0) = 0$$

   at some point $x_0 \in (a, b)$, then $y \equiv 0$ on (a, b).

5. **Resonance Orders for Undetermined Coefficients.** Consider

$$(D-1)^2(D^2+4)\,y = f(x),$$

where $D = \frac{d}{dx}$. For each forcing term below:

$$f_1(x) = e^x, \quad f_2(x) = xe^x, \quad f_3(x) = \cos(2x), \quad f_4(x) = \sin(2x), \quad f_5(x) = e^x\cos(2x),$$

determine the resonance multiplicity k and write the minimal undetermined-coefficients ansatz for a particular solution.

6. **Variation of Constants for Higher-Order Equations.** Using the ansatz

$$y(x) = \sum_{k=1}^{n} C_k(x)\varphi_k(x),$$

derive *theorem* 11.2 without invoking *theorem* 9.3.

7. **Euler Equations.** Consider the Euler equation

$$x^n y^{(n)} + a_1 x^{n-1} y^{(n-1)} + \cdots + a_{n-1} x y' + a_n y = 0, \qquad x > 0,$$

where $a_1, a_2, \ldots, a_n$ are constants. Find a change of variables that converts this equation into a homogeneous linear differential equation with constant coefficients, and then describe the resulting general solution procedure.

8. **Damped Spring Vibrations.** Solve the damped spring vibration equation

$$m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = 0,$$

where m, c, and k are positive constants. Explain the physical meaning of the corresponding solutions in the three cases $\Delta = c^2 - 4mk > 0$, $\Delta = 0$, and $\Delta < 0$.

9. **Forced Vibrations Without Damping.** Solve the forced vibration equation

$$m\frac{d^2x}{dt^2} + kx = p\cos(\omega t),$$

where m, k, p, and ω are positive constants. Explain the physical meaning of the solutions in the two cases $\omega \neq \omega_0$ and $\omega = \omega_0$, where $\omega_0 = \sqrt{\frac{k}{m}}$ is the natural frequency of the oscillator.

10. **Constant-Coefficient and Euler-Type Problems.** Solve the following differential equations:

    (a) $y'' + y' - 2y = 2x$, with $y(0) = 0$ and $y'(0) = 1$.
    (b) $2y'' - 4y' - 6y = 3e^{2x}$.
    (c) $y'' + 2y' = 3 + 4\sin(2x)$.
    (d) $y''' + 3y' - 4y = 0$.
    (e) $y''' - 2y'' - 3y' + 10y = 0$.
    (f) $y''' - 3ay'' + 3a^2y' - a^3y = 0$.
    (g) $y^{(4)} - 4y''' + 8y'' - 8y' + 3y = 0$.
    (h) $y^{(5)} + 2y'' + y' = 0$.
    (i) $y^{(4)} + 2y'' + y = \sin x$, with $y(0) = 1$, $y'(0) = -2$, $y''(0) = 3$, and $y'''(0) = 0$.
    (j) $y^{(4)} + y = 2e^x$, with $y(0) = y'(0) = y''(0) = y'''(0) = 1$.
    (k) $y'' - 2y' + 2y = 4e^x\cos x$.
    (l) $y'' - 5y' + 6y = (12x - 7)e^{-x}$.
    (m) $x^2y'' + 5xy' + 13y = 0$, for $x > 0$.
    (n) $(2x+1)^2y'' - 4(2x+1)y' + 8y = 0$, solved separately on $(-\infty, -\frac{1}{2})$ and $(-\frac{1}{2}, \infty)$.