Notes for *Differential Equations*[*] *(WUSTL, Math 217, Summer 2018)*

Mohammad Jabbari

August 13, 2018

Contents

0 Prologue (Optional) **4**

1 Preliminaries **7**
1.1 Calculus . . . 7
1.1.1 Real Variables . . . 7
1.1.2 Smooth Functions . . . 8
1.1.3 Differentiation of Definite Integrals . . . 8
1.1.4 Elementary Functions . . . 8
1.1.5 Binomial Integral (Optional) . . . 8
1.1.6 Dependent and Independent Variables, Implicit Equations and Locally Defined Functions . . . 9
1.1.7 Infinitesimals, Line Integrals . . . 11
1.1.8 Gamma Function . . . 12
1.2 Complex Numbers, Complex-Valued Functions . . . 13
1.3 Polynomials . . . 15
1.4 Linear Algebra . . . 16
1.4.1 Vectors, Bases . . . 16
1.4.2 Matrices . . . 16
1.4.3 Determinants . . . 19
1.4.4 Eigenvalues and Eigenvectors . . . 20
1.5 Matrix-Valued Functions, Matrix Exponentials . . . 21

*Pictures are taken from [5], [13] and [23].

2 Introduction to Differential Equations **22**
2.1 The Nature of Solutions of Ordinary Differential Equations 23
2.2 The Nature of Solutions of Partial Differential Equations 25
2.3 Using Computers 26
2.4 Separable First-Order Ordinary Differential Equations 26

3 Modeling of Dynamical Phenomena by Differential Equations **28**
3.1 Modeling by Ordinary Differential Equations 28
3.2 Modeling by Systems of Ordinary Differential Equations 33
3.3 Modeling by Partial Differential Equations 38
3.4 An Integro-differential Equation 41

4 First-Order Ordinary Differential Equations **42**
4.1 The Existence and Uniqueness of Solutions 44
4.2 Homogeneous Equations 48
4.3 Linear Equations 48
4.4 Bernoulli Equation 49
4.5 Exact Equations 50
4.6 $y' = F(ax + by + c)$ 51
4.7 $y' = F\left(\frac{ax+by+c}{\alpha x+\beta y+\gamma}\right)$ 52
4.8 When Dependent or Independent Variable Is Missing 53
4.9 Integrating Factor 54

5 Laplace Transform, Approximate Identities **56**
5.1 Basic Properties 59
5.2 Approximate Identities, Heat and Poisson Kernels (Optional) 63
5.3 Tautochrone Problem 65
5.4 Existence, Uniqueness and Inversion Formulas for Laplace Transform (Optional) . . 65

6 Linear Ordinary Differential Equations **67**
6.1 The Existence and Uniqueness of Solutions 67
6.2 Structure Theorem 68
6.3 Finding a Fundamental System of Solutions for a Second-Order Homogeneous Linear Equation Having One Solution at Hand 70
6.4 Constant Coefficient Equations 71
6.5 The Method of Variation of Parameters 74
6.6 The Operator Method of Heaviside 75
6.7 The Method of Laplace Transform, Analysis of Discontinuous and Periodic Inputs . 76

6.8 The Method of Undetermined Coefficients 83
6.9 Resonance 84
6.10 Euler-Cauchy Equation 85
6.11 Some Qualitative Analysis 85

7 Power Series Methods **86**
7.1 Linear Second-Order Ordinary Differential Equation: Nonsingular Points 89
7.2 Linear Second-Order Ordinary Differential Equation: Regular Nonsingular Points . 91
7.3 First-Order Ordinary Differential Equations 95

8 Systems of First-Order Ordinary Differential Equations I: Basics and Linear Equations **97**
8.1 The Existence and Uniqueness of Solutions 99
8.2 Linear Systems 100
8.3 The Method of Variation of Parameters 101
8.4 Linear First-Order Constant-Coefficient Systems 102
8.4.1 Elimination 102
8.4.2 Using Laplace Transform and Matrix Exponentials 103
8.4.3 Using Linear Algebra 104
8.4.4 An Explicit Fundamental System of Solutions for Equation $x' = Ax$ 106

9 Systems of Ordinary Differential Equations II: Nonlinear Equations **110**

10 First Touches on Partial Differential Equations **111**
10.1 Linear First-Order Partial Differential Equations 111
10.2 Heat Equation in Infinite Rod 111
10.3 Heat Equation in Finite Rod 112
10.3.1 Numerical Solution 113
10.3.2 Analytic Solution 114
10.4 Fourier Series 117
10.5 Wave Equation in Infinite String 118
10.6 Wave Equation in Finite String 121
10.7 Laplace Equation 121
10.8 The Convergence of Fourier Series (Optional) 123

11 Sturm-Liouville Eigenvalue Problem **125**

0 Prologue (Optional)

Many dynamic phenomena in nature and (pure or applied) sciences could be described by differential equations. Here are some witnesses:

- Human population through time $P(t)$, in many circumstances, obey either the logistic model
$$\frac{dP}{dt} = kP(M-P),$$
where k and M are positive constants.

- The temperature T of a hot object placed in an environment of temperature C decreases through time t according to
$$\frac{dT}{dt} = -k(T-C),$$
where k is a positive constant.

- The population of rabbits $x(t)$ and foxes $y(t)$, through time t, in an isolated iceland with unlimited food supply for rabbits, closely follows the following system of equations
$$\begin{cases} \frac{dx}{dt} = x(a-by) \\ \frac{dy}{dt} = -y(c-dx) \end{cases},$$
where a, b, c and d are positive constants. This is Lotka-Voterra's prey-predator equation.

- Soap film surfaces are (locally) described by the following partial differential equation of minimal surfaces:
$$\left(1+u_y^2\right)u_{xx} - 2u_x u_y u_{xy} + \left(1+u_x^2\right)u_{yy} = 0.$$

- All classical electromagnetic and optical phenomena are governed by Maxwell's equations:
$$\begin{cases} \nabla \cdot D = \rho \\ \nabla \cdot B = 0 \\ \nabla \times E = -\frac{\partial B}{\partial t} \\ \nabla \times H = J + \frac{\partial D}{\partial t} \\ \nabla \cdot J + \frac{\partial \rho}{\partial t} = 0 \end{cases},$$
where
$$\nabla := \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z},$$

is the nabla operator. They express how electric (E and D) and magnetic (H and B) field strengths and flux densities interact with themselves and with electric charge density (ρ) and electric current (J). Electrical engineers use them to design circuits, antennas, control and power devices, etc.

- Many equilibrium states in nature, for example electrostatic or magnetostatic phenomena, are described by Laplace equation
$$\Delta u = 0,$$
where
$$\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$
is the Laplacian operator.

- Many diffusion phenomena in nature, for example heat conduction in solids, is governed by the heat equation
$$u_t = \Delta u.$$

- Many oscillatory phenomena in nature, for example vibrations of elastic strings or membranes, are described by the wave equation
$$u_{tt} = \Delta u.$$

- The dynamics of viscous[1] fluids is described by Navier-Stokes equations. They are a collection of nonlinear partial differential equations, describing the evolution of fluid's velocity vector field (u, v, w) and pressure p in terms of its density ρ, viscosity ν, and external forces. They are used in the analysis and design of aircrafts, blood flow, air pollution, etc. Here is a version of them satisfied inside homogeneous isotropic incompressible fluids under no external force:
$$\begin{cases} \rho(u_t + uu_x + vu_y + wu_z) = -p_x + \nu\Delta u \\ \rho(v_t + uv_x + vv_y + wv_z) = -p_y + \nu\Delta v \\ \rho(w_t + uw_x + vw_y + ww_z) = -p_z + \nu\Delta w \\ u_x + v_y + w_z = 0 \end{cases}.$$

- The deviation $u(x, y)$ of a thin elastic plate from its equilibrium horizontal position is governed by biharmonic equation:
$$u_{xxxx} + 2u_{xxyy} + u_{yyyy} = 0.$$

[1]Viscosity accounts for friction between fluid molecules.

- Non-relativistic quantum effects of a single particle of mass m under the influence of a potential function $V(x, y, z)$ are described by the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\Psi + V\Psi,$$

  where i is the imaginary unit, $\hbar$ is the reduced Planck constant, and $\Psi(x, y, z, t)$ is systems's wave function.

- Gravitational phenomena, for example the dynamics of solar system, are governed by Einstein's field equations. They are nonlinear system of partial differential equations describing how the geometry of spacetime interacts with matter and energy.

- Atmospheric convection, more specifically the evolution of two-dimensional fluid layers uniformly warmed from below and cooled from above, is described by Lorentz system of ordinary differential equations

$$\begin{cases} \frac{dx}{dt} = \sigma(y - x) \\ \frac{dy}{dt} = x(\rho - z) - y \\ \frac{dz}{dt} = xy - \beta z \end{cases},$$

  where σ, ρ and β are positive constants, and x, y and z are specific state variables.

I guess that is already enough to convince you of the ubiquity of differential equations in science and engineering. The truth is that only very few of these equations have exact analytic general solutions; however:

- Several simplified versions of these equations have exact or asymptotic analytic solutions, and these solutions are of invaluable help to scientists and engineers in developing intuition about the dynamics of the system under study.

- Different level of qualitative analysis is possible in all cases.

- Stable numerical solutions mostly exist. However, there are chaotic systems, for example the Lorentz attractor, where numerical computations become unstable.

In these notes, after learning how to describe dynamic phenomena by differential equation in Chapter 3, we discuss basic concepts and tools to analyze these equations analytically, qualitatively and numerically. However we completely confine ourself to deterministic differential equations. You are highly encouraged to take some complementary course on stochastic differential equations. I have gathered all preliminaries in Chapter 1, and we will discuss each one in time of need during the course.

1 Preliminaries

Natural numbers are positive integers

$$\mathbb{N} = \mathbb{Z}_{>0} = \{1, 2, \cdots\}.$$

Nonnegative integers are

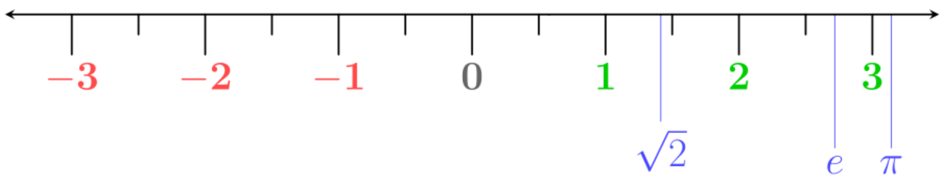
$$\mathbb{Z}_{\geqslant 0} = \{0, 1, 2, \cdots\}.$$

1.1 Calculus

1.1.1 Real Variables

Having a set S of objects, a generic element of S is called a **variable ranging over** (or **living on**) S. Variables are usually denoted by x, y, z, t, u, v, w, etc., whereas specific elements of a set are usually denoted by a, b, c, A, B, C, α, β, γ, etc. Functions are usually denoted by f, g, F, G, etc. When reading (or writing!) a mathematical statement which contains variables, one should always make sure he/she understands where each variable lives.

The most important set in this course is the set of *real numbers*, denoted by $\mathbb{R}$. Intuitively, each real number corresponds to a *point* on the *real line*. There are no *gaps* in the real line. When made abstract, this latter statement is called the **completeness axiom** for real numbers, and underlies all existence theorems in this course, and elsewhere in mathematical analysis.



The set of points in the place and more generally in the n-dimensional space are denoted by $\mathbb{R}^2$ and $\mathbb{R}^n$, respectively. Therefore

$$\mathbb{R}^n = \{(x_1, \cdots, x_n) : x_i \in \mathbb{R} \text{ for each } i\}.$$

Most important subsets of $\mathbb{R}$ are intervals, for example

$$(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}, \quad [0, 1] = \{x \in \mathbb{R} : 0 \leqslant x \leqslant 1\}.$$

The corresponding notion in higher dimensions are rectangles (in $\mathbb{R}^2$) and boxes (in $\mathbb{R}^n$). For example $(0, 1)^n$ denotes the sets of points $(x_1, \cdots, x_n) \in \mathbb{R}^n$ with each x_i strictly between 0 and 1.

1.1.2 Smooth Functions

In this course, we mostly work with **smooth** real-valued functions on open boxes.[2] Smoothness means that (ordinary or partial) derivatives of all orders exist. At special sections we also consider discontinuous, impulse, complex-valued, vector-valued or matrix-valued functions.

Example 1. Consider the function $y = x^2|x|$ defined for real variable x ranging on whole real line. It has derivatives up to order 2 everywhere, but the third derivative does not exist at $x = 0$. ■

1.1.3 Differentiation of Definite Integrals

Let $F(x,t)$ be a real-valued function of two real variables, and $u(x), v(x)$ be real-valued function of a real variable. Under mild conditions:

$$\frac{d}{dx}\int_{u(x)}^{v(x)} F(x,t)\,dt = \int_{u(x)}^{v(x)} F_x(x,t)\,dt + \frac{dv}{dx}\cdot F(x,v(x)) - \frac{du}{dx}\cdot F(x,u(x)).$$

1.1.4 Elementary Functions

Functions that are constructed by algebraic operations (namely, addition, subtraction, multiplication, division and composition) from power functions, trigonometric functions and their inverses, exponentials and logarithms are called **elementary functions**. Most integrals, for example

$$\int e^{-t^2}dt, \quad \int \frac{\sin t}{t}dt, \quad \int \frac{dt}{\log t}, \quad \int t^{\frac{1}{2}}(1+t)^{\frac{1}{3}}dt,$$

are not elementary. We will encounter many non-elementary functions in this course.

1.1.5 Binomial Integral (Optional)

The indefinite integral

$$\int x^A(B + Cx^D)^E dx,$$

where A, D and E are *rational* constants, and B and C are nonzero real constants, is called **binomial integral**. With the algebraic change of variables

$$Cx^D = Bu,$$

[2]Specially there is no hole in our boxes. The study of differential equations in the presence of holes leads to *cohomology*, which is beyond this course.

it becomes

$$B \int u^F (1+u)^G du, \tag{1}$$

where F and G are new rational constants.

Theorem 1 (Chebyshev). *The integral (1), where F and G are rational constants, and B is a nonzero real constant, is elementary if and only if at least one of F, G, or $F+G$ is integer.*

We only show the *if part*. For the other direction refer to references given in [16, Appendix I].

Case I: F *integer*. Assuming $G = \frac{G_1}{G_2}$ with integers G_1 and G_2, then change of variables

$$1 + u = v^{G_2},$$

makes (1) into an integral of a rational function, which we know is an elementary function [1, vol I, section 6.23].

Case II: G *integer*. Assuming $F = \frac{F_1}{F_2}$ with integers F_1 and F_2, then change of variables

$$u = v^{F_2},$$

makes (1) into an integral of a rational function, which is an elementary function.

Case III: $F + G$ *integer*. Writing (1) as

$$\int u^{F+G} \left(\frac{1+u}{u}\right)^G du,$$

change of variable

$$\frac{1+u}{u} = v,$$

reduces us to the first case.

1.1.6 Dependent and Independent Variables, Implicit Equations and Locally Defined Functions

Let x, y and z be three real variables, some depending on the other. This dependence is usually expressed by one or more equations. For example the equation

$$y = x^2, \tag{2}$$

determines y in terms of x. In mathematical terms, y is a *function* of x, with x the **independent variable**, and y the **dependent variable**. As another example the equation

$$z = x^2 + y^2, \tag{3}$$

expresses dependent variable z as a function of independent variables x and y. It helps to visualize these two functions as parabola and elliptic paraboloid, respectively.

As a different example, consider the equation of unit circle

$$x^2 + y^2 = 1. \tag{4}$$

Here for each x there might be one, two or three corresponding values of y. There are several ways to deal with this issue:

- Although y is not a function of x, mathematicians call y a **multi-valued "function"** of x. We will encounter other examples of multi-valued functions in this course.[3]

- It is geometrically clear that if we add the extra assumption $y \geqslant 0$ to (4), then y is uniquely determined in terms of x as a continuous function

$$x \in [-1, 1] \mapsto y = \sqrt{1 - x^2} \in \mathbb{R}.$$

  This way we turned *implicit* equation (4) into *explicit* one $y = \sqrt{1 - x^2}$. Alternatively we could add assumption $y \leqslant 0$, and get $y = -\sqrt{1 - x^2}$. In this particular example we were able to give y explicitly in terms of x, but this is not always possible or even useful in most cases. For example, although the equation $y^5 + y = x$, for each real x gives exactly one real y (why?), there is no simple way to find y explicitly in terms of x.

- It is geometrically clear that for each point $P = (x_0, y_0)$ on the unit circle, except for special points $(\pm 1, 0)$, there is a small open interval I around x_0, and unique smooth function $y = y(x)$ defined on $x \in I$, which passes through P, and satisfies (4). This function is said to be **locally-defined by equation (4) around point** P. The good news is that, even without drawing the graph of an implicit equation, or doing the tedious/impossible task of converting it into an explicit equation, there is a mathematical criterion to decide when a general implicit equation $F(x, y) = 0$ defines a local function $y = y(x)$. Here is that criterion, one of the deepest result in calculus.

Theorem 2 (Implicit Function Theorem). *Let $F(x, y)$ be a smooth real-valued function defined on an open rectangle containing (x_0, y_0). If $F_y(x_0, y_0) \neq 0$, then there is an open interval I containing x_0, and unique smooth function $y = f(x)$ defined on $x \in I$ such that*

$$f(x_0) = y_0, \quad F(x, f(x)) = 0,$$

for each $x \in I$. Furthermore, $f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}$.

[3] It is beyond this course, but let us briefly mention that Riemann had the brilliant idea of turning multi-valued "functions" into functions by doing *surgery* on their domain of definition.

Exercise 1. *Consider the curve defined by equation*

$$y^5 + y - x^3 + 3x - 1 = 0.$$

(a) Find all the points on the curve where the tangent line is vertical. Find all the points on the curve where the tangent line is horizontal.

(b) At what points on the curve you are sure that the equation locally determines y as a smooth function of x? At what points on the curve you are sure that the equation locally determines x as a smooth function of y?

(c) (Optional) *Does the equation determines y as a function of x on the whole line? Does the equation determines x as a function of y on the whole line?*

1.1.7 Infinitesimals, Line Integrals

Let real variable z be a *smooth* function of real variables x and y, notationally $z = F(x, y)$. Around a fixed point (x_0, y_0), $F(x, y)$ could be approximated by its linearizion:

$$F(x, y) \approx F(x_0, y_0) + (x - x_0)F_x(x_0, y_0) + (y - y_0)F_y(x_0, y_0).$$

Moving $F(x_0, y_0)$ to the left, we write this latter equation as

$$\boxed{dz = z_x\,dx + z_y\,dy},$$

and understand it as follows: when x and y are respectively perturbed by dx and dy, then z is perturbed by $z_x\,dx + z_y\,dy$.

Exercise 2. *Consider the following quantities:*

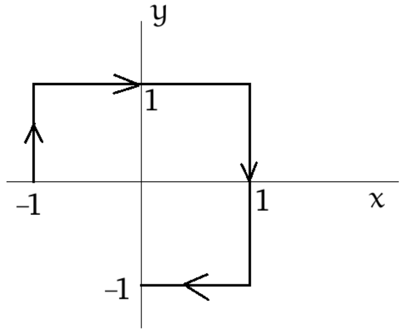
$$\sqrt[3]{1.0001}, \quad 0.001^3 \sin^3(0.001^2).$$

Clearly they approximately equal 0 *and* 1*, respectively. Without using a calculator, find a better approximation.*

Let $M(x, y)$ and $N(x, y)$ be two real-valued functions of real variables x and y, and let C be a directed curve in the plane. Then we could integrate the infinitesimal $M dx + N dy$ along C. This is done by first choose a parametrization $(x, y) = (f(t), g(t))$, $t \in [a, b]$, for C, and then setting

$$\int_C M(x, y)dx + N(x, y)dy := \int_a^b \big(M(f(t), g(t))f'(t) + N(f(t), g(t))g'(t)\big)\,dt.$$

Exercise 3. *Integrate $x dx + y dy$ along the directed curve shown below.*



1.1.8 Gamma Function

Integration by parts shows that

$$\int_0^\infty e^{-t}t^n dt = n!,$$

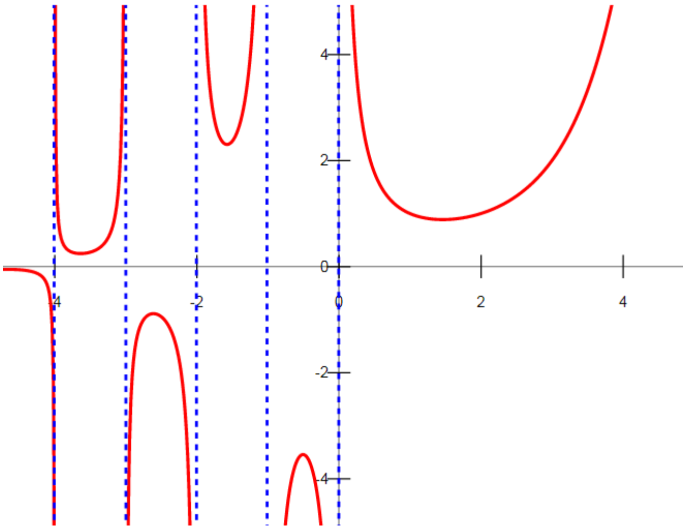
for each natural number n. This motivated mathematicians to define the factorial of a real number a by

$$a! := \int_0^\infty e^{-t}t^a dt. \tag{5}$$

This integral converges when $a > -1$. The expression on the right hand side of (5) is also denoted by $\Gamma(a+1)$, and defines the **Gamma function**

$$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt, \quad x \in (0, \infty).$$

By integration by parts $\Gamma(x+1) = x\Gamma(x)$, which enables us to define $\Gamma(x)$ for reals x except nonpositive integers. Here is its graph.



Exercise 4. *Accepting* $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, *compute* $\Gamma\left(-\frac{3}{2}\right)$.

1.2 Complex Numbers, Complex-Valued Functions

A **complex number** is a formal expression of the form $a + bi$, where a and b are real numbers, and i is a crazy object, called the **imaginary unit**, pretended to satisfy $i^2 + 1 = 0$.[4] Two complex numbers $a + bi$ and $a' + b'i$ are equal exactly when $a = a'$ and $b = b'$. We also agree that $a + 0i = a$ and $0 + bi = bi$. The set of all complex numbers is denoted by $\mathbb{C}$. This is not just a set but a *number system* in the sense that its elements are added, subtracted, multiplied and divided according to:

$$(a+bi) \pm (c+di) = a \pm c + (b \pm d)i, \quad (a+bi)(c+di) = ac - bd + (ad+bc)i,$$
$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac+bd}{c^2+d^2} + \frac{-ad+bc}{c^2+d^2}i, \quad \text{if } c+di \neq 0.$$

These four operations of complex numbers has exactly the same corresponding algebraic properties of real numbers. For example, product is commutative, associative, and distributive over addition:

$$z_1 z_2 = z_2 z_1, \quad (z_1 z_2) z_3 = z_1 (z_2 z_3), \quad z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3.$$

Exercise 5. *(a) Find a complex number z which solves equation $z^4 + 1 = 0$. (b) Find all complex numbers z which solve equation $z^4 + 1 = 0$. [Hint. For (a), find $z = x + iy$ with $z^2 = i$.]*

Fix complex number $z = a + bi$. Real numbers a and b are, respectively, called the **real part** and the **imaginary part** of z, and denoted by $\Re z$ (or $\mathrm{Re}(z)$) and $\Im z$ (or $\mathrm{Im}(z)$). The **conjugate** and **modulus** (or **absolute value**) of z are defined by

$$\bar{z} = a - bi \in \mathbb{C}, \quad |z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2} \in [0, \infty).$$

Intuitively, we think of complex number $z = a + bi$ as the *point* (a, b) in the Cartesian plane. Then, $\Re z$, $\Im z$, $\bar{z}$ and $|z|$ are, respectively, the x-abscissa, y-abscissa, the reflection of z with respect to

[4] Here comes a motivation for this pretension. The simplest equation not solvable in reals is $x^2 + 1 = 0$. If we momentarily imagine a *number* i satisfying this equation, namely $i^2 + 1 = 0$, adjoining it to reals, namely forming expressions like $a + bi$, with a and b reals, we will be able to solve all quadratic equations $Ax^2 + Bx + C = 0$, with $A \neq 0$, B and C real, even when $\Delta = B^2 - 4AC < 0$; here are the solutions:

$$x = \frac{-B \pm \sqrt{-\Delta}\, i}{2A}.$$

The miraculous fact, discovered by Laplace and Gauss, is that these numbers $a + bi$, with a and b real, are even enough to solve all higher degree equations $a_n x^n + \cdots + a_1 x + a_0 = 0$, with $a_n \neq 0, a_{n-1}, \ldots, a_0$ real, and $n \geqslant 2$ integer. Also refer to Exercise 5.

x-axis, and the distance between z and the origin. The addition of complex numbers corresponds to the addition of vectors using parallelogram law in physics. To find the geometrical interpretation of multiplication of complex numbers, consider complex numbers z and z' with polar coordinates (r, θ) and (r', θ'), respectively, namely

$$z = r\cos\theta + ir\sin\theta, \quad z' = r'\cos\theta' + ir'\sin\theta'.$$

Then the computation

$$\begin{aligned} zz' &= (r\cos\theta + ir\sin\theta)(r'\cos\theta' + ir'\sin\theta') \\ &= rr'(\cos\theta\cos\theta' - \sin\theta\sin\theta') + irr'(\cos\theta\sin\theta' + \sin\theta\cos\theta') \\ &= rr'\cos(\theta+\theta') + irr'\sin(\theta+\theta'), \end{aligned}$$

shows that the geometric interpretation of multiplying complex numbers by z with polar coordinate (r, θ) is rotation by θ around origin followed by dilation with factor r (or the other way around). Specially, multiplication by i corresponds to rotation around origin by 90° counterclockwise. We could also deduce

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \tag{6}$$

for each real θ and integer n. This motivates us to define exponential e^z of a complex number $z = x + yi$ as:

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

This is **Euler's formula**. Note that, if (r, θ) is the polar coordinates corresponding to complex number z, then Euler's formula exactly says that

$$\boxed{z = re^{i\theta}},$$

which is called the **polar representation of a complex number**.

Exercise 6. *Prove $e^{z_1+z_2} = e^{z_1}e^{z_2}$ for each two complex numbers z_1 and z_2.*

Exercise 7. *(a) Equating real and imaginary parts in (6), prove that*

$$\cos n\theta = \binom{n}{0}\cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + - \cdots,$$

$$\sin n\theta = \binom{n}{1}\cos^{n-1}\theta\sin\theta - \binom{n}{3}\cos^{n-3}\theta\sin^3\theta + - \cdots.$$

(b) Prove that:

$$\tan n\theta = \frac{\binom{n}{1}\tan\theta - \binom{n}{3}\tan^3\theta + - \cdots}{\binom{n}{0} - \binom{n}{2}\tan^2\theta + - \cdots}.$$

(c) Find a nontrivial polynomial equation with integer coefficients satisfied by $x = \cos\frac{\pi}{7}$.

For each natural number n, and each complex number $z = re^{i\theta} \neq 0$, there are exactly n distinct complex numbers which when raised to power n are equal to z. They are called **n-th roots of** z, denoted by multi-valued notation $\sqrt[n]{z}$, and computed via

$$\boxed{\sqrt[n]{re^{i\theta}} = \sqrt[n]{r}e^{i\frac{\theta+2\pi k}{n}}, \quad k = 1, \ldots, n}.$$

Exercise 8. *(a) Compute all* 8*-th roots of* 1*, and show them in the complex plane. (b) Compute all* 4*th-roots of* -1*, and show them in the complex plane. (This is Exercise 5!)*

If $f(t)$ and $g(t)$ are real-valued functions of real variable t, then $h(t) := f(t) + ig(t)$ is called a **complex-valued function** of t. Limit, continuity, differentiation, integration, etc., of $h(t)$ is defined componentwisely. For example, $h(t)$ is called **differentiable** at point $t = t_0$ if both f and g are differentiable at the the same point $t = t_0$; if so we write $h'(t) = f'(t) + ig'(t)$.

In these notes, by an **scalar**, we mean either a real or complex number. For example, a scalar-valued function is a dependent variable ranging over $\mathbb{R}$ or $\mathbb{C}$. By a **vector**, we always mean some (ordered) tuple of scalars.

Exercise 9. *(a) Show that for any complex number* c*,*

$$\frac{d}{dx}e^{cx} = ce^{cx}, \qquad \int e^{cx}dx = \frac{1}{c}e^{cx} + [\text{constant}].$$

(b) Use part (a) to compute indefinite integrals

$$\int e^{ax}\cos(bx)dx, \qquad \int e^{ax}\sin(bx)dx.$$

1.3 Polynomials

An scalar-valued function of scalar variable z of the form

$$p(z) := a_n z^n + \cdots + a_1 z + a_0, \tag{7}$$

where a_i, $i = 0, \ldots, n$, are scalars, and $a_n \neq 0$ is called a **polynomial of degree** n. Scalar z_0 is a **root** (or **zero**) of p if $p(z_0) = 0$. This happens exactly when $p(z) = (z - z_0)q(z)$ where $q(z)$ is a polynomial of degree $n - 1$. If

$$p(z) = (z - z_0)^m q(z), \quad m \in \mathbb{N}, \quad q(z) \text{ polynomial}, \quad q(z_0) \neq 0,$$

then z_0 is a root of $p(z)$ with **multiplicity** m. If $m = 1$, z_0 is called a **simple root**. If $m > 1$, $z = z_0$ is called a **multiple** (or **repeated**) **root**.

Proposition 1. *Scalar z_0 is a root of polynomial $p(z)$ with multiplicity m if and only if*

$$p(z_0) = p'(z_0) = \cdots = p^{(m-1)}(z_0) = 0, \quad p^{(m)}(z_0) \neq 0.$$

Theorem 3 (Fundamental Theorem of Algebra). *A polynomial like (7) can be factored as*

$$p(z) = a_n \prod_{1 \leqslant k \leqslant N} (z - \lambda_k)^{m_k},$$

where m_k, $k = 1, \ldots N$, are natural numbers, summing up to n.

1.4 Linear Algebra

1.4.1 Vectors, Bases

Let $v_1, \ldots, v_k$ be real n-vectors, namely elements of $\mathbb{R}^n$. They are called **linearly independent** if there are no real numbers $C_1, \ldots, C_k$, at least one nonzero, such that $C_1 v_1 + \cdots + C_k v_k = 0$. Otherwise they are **linearly dependent**. An n-vector w is a **linear combination** of $v_1, \ldots, v_k$ if there are scalars $C_1, \ldots, C_k$ such that $w = C_1 v_1 + \cdots + C_k v_k$. Here is a fundamental fact:

Theorem 4. *Consider real n-vectors $v_1, \ldots, v_k$. If two of the following properties hold then hold the third.*

(a) $k = n$; (b) $v_1, \ldots, v_k$ are independent; (c) Every n-vector can be written as a linear combination of $v_1, \ldots, v_k$.

*If all these properties hold, then the collection $\{v_1, \ldots, v_k\}$ is called a **basis** or **fundamental system** for $\mathbb{R}^n$.*

Example 2. Show that $\{(1, 1), (1, -1)\}$ is a basis for $\mathbb{R}^2$, but $\{(1, 1, 1), (-1, 2, -3), (1, 7, -3)\}$ is not a basis for $\mathbb{R}^3$. ■

1.4.2 Matrices

For us a **matrix** is a rectangular array of scalars, say

$$\begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{bmatrix}.$$

This is denoted by A, $[A_{ij}]$, or $[A_{ij}]_{m \times n}$. The pair (m, n), or $m \times n$, is the **order** of A, m the number of rows, and n the number of columns. Two matrices have the same order if they have the same number of rows, and the same number of columns. Scalar A_{ij} is the ij**-th entry** of A. Two matrices of the same order are **equal** if all their corresponding entries equal. The **transpose** of A, denoted by A^t, is the $n \times m$ matrix whose ij-th entry is ji-th entry of A, for each i and j. If $n = 1$ then A is a **column vector**, which could be identified with the corresponding point in $\mathbb{C}^m$. For typographical reasons we might sometimes denote it by $[A_{11}, \dots, A_{1m}]^t$ instead of

$$\begin{bmatrix} A_{11} \\ \vdots \\ A_{1m} \end{bmatrix}.$$

Dually, if $n = 1$ then A is a **row vector**. If $m = n = 1$ then A is identified with its single entry, hence an scalar. Rows are numbered from above, and column from left. For example, for A above,

$$[A_{i1}, A_{i2}, \dots, A_{in}],$$

is the i-th row, and

$$\begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix},$$

is the j-th column. If $v_1, \dots, v_n$ are column vectors in $\mathbb{C}^m$, then the $m \times n$ matrix whose first column is v_1, second column v_2, etc., is denoted by $[v_1, \cdots, v_n]$. Dually, if $v_1, \dots, v_m$ are row vectors in $\mathbb{C}^n$, then the $m \times n$ matrix whose first row is v_1, second row v_2, etc., is denoted by $[v_1; \cdots; v_n]$. If $m = n$ then A is a **square matrix** of order n. An square matrix $A = [A_{ij}]$ is **diagonal** if off-diagonal entries A_{ij}, $i \neq j$, are all zero. An square matrix $A = [A_{ij}]$ is **upper triangular** if entries A_{ij}, $i > j$, are all zero.

A matrix A could be multiplied by a scalar c componentwise, namely $(cA)_{ij} = cA_{ij}$. Two matrices of the same order could be added or subtracted componentwisely, namely $(A \pm B)_{ij} = A_{ij} \pm B_{ij}$. The matrix with all entries zero is called **zero matrix**, and denoted by 0. For two matrices $A = [A_{ij}]_{m\times n}$ and $B = [B_{ij}]_{p\times q}$, *only if the number of columns of* A *equals the number of rows of* B*, namely* $n = p$, the product matrix $C = AB = [C_{ij}]_{m\times q}$ is defined as follows:

$$C_{ij} = \sum_{1 \leqslant k \leqslant p} A_{ik} B_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, q.$$

In other words, the ij-th entry of AB, is the i-th row of A "multiplied" componentwise by j-th column of B.[5]

[5]There is a good reason for this strange-looking definition. Matrices are algebraic side of linear maps between vector spaces, and we define matrix multiplication in this way to correspond to the composition of linear maps.

Matrix multiplication is associative $(AB)C = A(BC)$, distributive over addition $A(B + C) = AB + AC$, but *not* commutative:

$$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} \neq \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}.$$

This might seem like a bad news at first, but it let matrices describe those operators in nature which do not commute. Quantum mechanics is full of such operators.

For each n there is a unique $n \times n$ matrix $I_n = [\delta_{ij}]$, called the **identity matrix**, and defined by

$$\delta_{ij} = \begin{cases}1, & i = j\\0, & i \neq j\end{cases},$$

such that $AI_n = I_nA = A$ for each square matrix A of order n. A square matrix A of order n is **invertible** (or **nonsingular**) if there is a square matrix B with $AB = BA = I$. Then B is called the **inverse** of A, and denoted by A^{-1}.

Exercise 10. *Prove that inverses are unique if they exist.*

Exercise 11. *Check that matrix*

$$\begin{bmatrix}a & b\\c & d\end{bmatrix},$$

is invertible if and only if $ad - bc \neq 0$. *If so then its inverse is given by*

$$\frac{1}{ad - bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}.$$

Exercise 12. *Check that the matrices of the form*

$$\begin{bmatrix}a & b\\b & -a\end{bmatrix}, \quad a, b \in \mathbb{R},$$

are added, subtracted, multiplied and divided (A/B *is defined as* AB^{-1}*) exactly as complex numbers* $a + bi$.

An application of matrices is that a linear system of m equations with n unknowns

$$y_j = a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n, \quad j = 1, \ldots, m,$$

could be concisely written as $y = Ax$ where

$$x = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}, \quad y = \begin{bmatrix}y_1\\y_2\\\vdots\\y_m\end{bmatrix}, \quad A = \begin{bmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\a_{21} & a_{22} & & a_{2n}\\\vdots & \vdots & \ddots & \vdots\\a_{m1} & a_{m2} & \ldots & a_{mn}\end{bmatrix}.$$

1.4.3 Determinants

For each natural number n, there is a unique function $A \mapsto |A|$ (also denoted by $A \mapsto \det(A)$), called **determinant**, taking $n \times n$ matrices as input and spitting a scalar, satisfying the following properties:

1. For each $j = 1, \ldots, n$, det is linear with respect to the j-th column while other columns are held fixed. This means

$$\det[v_1, \ldots, v_{j-1}, v_j + cw, v_{j+1}, \ldots, v_n] = \det[v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n] + \\ + c\det[v_1, \cdots, v_{j-1}, w, v_{j+1}, \ldots, v_n],$$

   where $v_1, \ldots, v_n, w$ are $n \times 1$ column matrices, and c is an scalar.

2. If two different columns are interchanged then det is multiplied by -1.

3. Identity matrix has determinant 1.

We have

$$\begin{vmatrix} a \end{vmatrix} = a,$$

$$\boxed{\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc},$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix}, \quad \text{etc.}$$

Exercise 13. *Show that* $\begin{vmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{vmatrix} = (\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1)$.

Theorem 5. *Let A and B be $n \times n$ matrices. Then:*

1. $|A^t| = |A|$. *Specially in the characterization of determinants above, "column" could be replaced by "row".*

2. *The determinant of an upper triangular matrix is the product of its diagonal elements.*

3. $|AB| = |BA|$.

4. *A is invertible if and only if $|A| \neq 0$.*

5. *Real (respectively, complex) n-vectors $v_1, \dots, v_n$ are a basis for $\mathbb{R}^n$ (respectively, $\mathbb{C}^n$) if and only if that $n \times n$ matrix whose j-th column, $j = 1, \dots, n$, is v_j has nonzero determinant.*

6. *Consider the linear system of equations $Ax = b$, where b is an $n \times 1$ vector. Let Δ denote the determinant of A, and for each $k = 1, \dots, n$, Δ_k be Δ where k-th column is replaced by b. Then* ***Cramer's Rule*** *says*
$$\boxed{\Delta x_k = \Delta_k, \quad \textit{for each } k}.$$

7. *The homogeneous system $Ax = 0$, has a solution besides the trivial one $x = 0$, if and only of $|A| = 0$.*

8. *For each i and j, the ij-th entry of A^{-1} is $\frac{(-1)^{i+j}}{|A|}$ times the determinant of the matrix made from A after removing i-th column and j-th row.*

1.4.4 Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix. If there is a scalar λ and a *nonzero* vector x such that $Ax = \lambda x$, then λ is called an **eigenvalue** of A with corresponding **eigenvector** x. Eigenvalues of A are exactly the roots of the **characteristic polynomial**

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda - A_{11} & -A_{12} & \dots & -A_{1n} \\ -A_{21} & \lambda - A_{22} & & -A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{n1} & -A_{n2} & \dots & \lambda - A_{nn} \end{vmatrix} \\ &= \lambda^n - (A_{11} + \cdots + A_{nn})\lambda^{n-1} + - \cdots + (-1)^n |A|. \end{aligned}$$

Proposition 2. *(a) Eigenvectors corresponding to distinct eigenvalues are independent.*

(b) If λ is an eigenvalue of A with multiplicity m (namely λ is a root of multiplicity m of characteristic polynomial of A), then there are at most m independent eigenvectors correspoding to λ.

Exercise 14. *Find eigenvalues and eigenvectors of* $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

Example 3. Consider the following matrices

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then:

Matrix A has eigenvalues $0, 0, 0$; eigenvalue 0 has three independent eigenvectors.

Matrix B has eigenvalues $0, 0, 0$; eigenvalue 0 has at most two independent eigenvectors.

Matrix C has eigenvalues $0, 0, 0$; eigenvalue 0 has at most one independent eigenvector.

Matrix D has eigenvalues $0, 0, 1$; eigenvalue 0 has at most two independent eigenvectors; eigenvalue 0 has at most one independent eigenvector.

Matrix E has eigenvalues $0, 0, 1$; eigenvalue 0 has at most one independent eigenvector; eigenvalue 0 has at most one independent eigenvector. ■

1.5 Matrix-Valued Functions, Matrix Exponentials

If all entries of a matrix A are functions of a real variable t, then A is called a **matrix-valued function** of t. Limit, continuity, infinit series, differentiation, integration, etc, of matrix-valued functions are defined componentwisely. For example, matrix-valued function $A(t) = [A_{ij}(t)]$ is called **differentiable** at point $t = t_0$ if each entry $A_{ij}(t)$ is differentiable at the same point $t = t_0$; and we set $A'(t) = \left[A'_{ij}(t)\right]$. If A and B are square matrix-valued functions of the same order, and c is a constant scalar, then

$$\frac{d}{dt}(cA) = c\frac{dA}{dt}, \quad \frac{d}{dt}(A+B) = \frac{dA}{dt} + \frac{dB}{dt}, \quad \frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}.$$

Exercise 15. *Let $A(t)$ be a square matrix-valued function. Show that if A is both invertible and differentiable at a point $t = t_0$, then A^{-1} is also differentiable at $t = t_0$, and given by*

$$\left(A^{-1}\right)' = -A^{-1}A'A^{-1}.$$

For square matrix A, it could be proved that the infinite series

$$\sum_{n\geqslant 0} \frac{(tA)^n}{n!} = I + tA + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \cdots,$$

converges for each real t. It will be denoted by e^{At}, called **matrix exponential**. It is the unique matrix-valued function $X(t)$ satisfying

$$\frac{dX}{dt} = AX, \quad X(0) = I.$$

We find a formula for it in Section 8.4.2.

Exercise 16. *Compute exponentials for matrices in Example 3.*

2 Introduction to Differential Equations

Recall our discussion in Section 1.1.6 about dependent and independt variables. The equations in all examples there were *algebraic*, in the sense that only the algebraic operations of addition, subtraction, multiplication, division and composition were used in their construction. If in addition to these algebraic operations, the equation involves (ordinary or partial) differentiations of dependent variables with respect to independent variables, then we get a **differential equation**. For example, each of the followings is a differential equation

$$\frac{dy}{dx} = 2x, \quad x + y\frac{dy}{dx} = 0, \quad 4z = \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2. \tag{8}$$

The first two are **ordinary differential equations (ODEs)** because no partial derivative appears, but the third one is a **partial differential equation (PDE)** because partial derivatives show up. Every function which satisfies a differential equation is called its **solution**. For example, functions given by (2) and (3), respectively, are solutions to the first and third equations in (8). The highest order of differentiation in a differential equation is called its **order**. For example all equations in (8) are of the first order.

The general form of an ODE is

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}\right) = 0, \tag{9}$$

where x is real independent variable, y is scalar dependent varibale, n is the order, and F a scalar-valued function of its arguments. This equation is called **linear** if F is a linear function of y and its derivatives, hence an equation of the form

$$P_0(x)y + P_1(x)y' + \dots + P_n(x)y^n + P_{n+1}(x) = 0,$$

where $P_k(x)$, $k = 0, \dots, n+1$, are scalar-valued functions of x.

The general form of a PDE is

$$F\left(x_1, \dots, x_N, y, \frac{\partial y}{\partial x_1}, \dots, \frac{\partial y}{\partial x_N}, \dots, \frac{\partial y}{\partial x_n}, \frac{\partial^2 y}{\partial x_1^2}, \dots\right) = 0, \tag{10}$$

where $x_1, \dots, x_N$ are real independent variables, y is scalar dependent variable, and F a scalar-valued function of its arguments. This equation is **linear** if F is a linear function of y and its partial derivatives.

2.1 The Nature of Solutions of Ordinary Differential Equations

The simplest differential equation is

$$\frac{dy}{dx} = F(x),$$

where F is a real-valued function of real variable x. Just by the very definition of indefinite integrals, this equation is solved as

$$y = \int F(x)dx + C.$$

Thus solving this differential equation is equivalent to find an anti-derivative for F. In many cases we are are not able to express this integral in terms of elementary function, and should accept it as the ultimate answer. (Recall examples in Section 1.1.4.)

The next simplest differential equation is

$$\frac{dy}{dx} = ky,$$

where k is a constant. It could easily be checked that $y = Ce^{kx}$, with constant C, satisfies the equation. Are there any other solutions? The following one line argument shows that the answer is NO. If y satisfies $y' = ky$, then $u := ye^{-kx}$ satisfies

$$u' = y'e^{-kx} - kye^{-kx} = 0,$$

hence u is a constant function C, hence $y = Ce^{kx}$. We say that $y = Ce^{-kx}$ is the **general solution** of $y' = ky$, in the sense that, firstly for each constant C, the function $y = Ce^{kx}$ satisfies the differential equation, and secondly, for each function $y = f(x)$ satisfying our equation, there exists a constant C such that $f(x) = Ce^{kx}$. Another way to put it is that, for any real number C, the **initial value problem**

$$\frac{dy}{dx} = ky, \quad y(0) = C,$$

has a unique solution given by $y = Ce^{kx}$. We generally expect that the general solution of a first-order ODE, contains only one constant C. We prove this for *linear* equations in Proposition 3, however each of the following *nonlinear* equations

$$\left(\frac{dy}{dx}\right)^2 + 1 = 0 \quad \text{or} \quad \left(\frac{dy}{dx}\right)^2 + y^2 = 0,$$

shows that this expectation does not hold in general.

As another example, let us consider ODE

$$\frac{d^2y}{dx^2} = y.$$

A little inspection shows that $y = e^x$ and $y = e^{-x}$ both satisfy our equation. Hence, by linearity, $y = C_1 e^x + C_2 e^{-x}$, where C_1 and C_2 are constants, also satisfies the equation. Are there any other solutions? The following one line argument shows that the answer is NO. If y satisfies $y'' = y$, then

$$u := e^{-x}(y + y'), \quad v := e^x(y - y'),$$

satisfy

$$u' = -e^{-x}(y + y') + e^{-x}(y' + y'') = 0, \quad v' = e^x(y - y') + e^x(y' - y'') = 0,$$

hence u and v are constant function, say C_1 and C_2, respectively, namely

$$e^{-x}(y + y') = C_1, \quad e^x(y - y') = C_2,$$

hence

$$y + y' = C_1 e^x, \quad y - y' = C_2 e^{-x},$$

hence $y = C_1 e^x + C_2 e^{-x}$. We won! We say that $y = C_1 e^x + C_2 e^{-x}$ is the general solution to $y'' = y$.

As these examples show it is reasonable to expect that the general solution of a n-th order ODE contains n constants. This will be proved for linear equations in Theorem 9.

Example 4. Let us now start with a family of functions containing general constant and try to find a differential equation satisfied by all of them. As our first example, we start with the family

$$y = Cx - \frac{C^2}{4},$$

where C is a general constant. C could be called the **parameter** of the family. Each specific value of parameter C gives a member of the family. Since C is a constant

$$0 = \frac{d}{dx}\left(-\frac{C^2}{4}\right) = \frac{d}{dx}(y - Cx) = y' - C,$$

hence $C = y'$. Putting this into our original equation gives

$$y = xy' - \frac{1}{4}(y')^2.$$

As another example, consider the family

$$y = Ae^{2x} + Be^{-3x},$$

where A and B are general constants. Since this family contains two constants we expect a second-order ODE, so we compute

$$y' = 2Ae^{2x} - 3Be^{-3x}, \quad y'' = 4Ae^{2x} + 9Be^{-3x}.$$

From the two equations $y = \cdots$ and $y' = \cdots$, we could find A and B, and then plug them into $y'' = \cdots$. We leave the computations to the reader. The answer will be $y'' + y' - 6y = 0$. Here is a shorter way. Trying to eliminate A form $y = \cdots$ and $y' = \cdots$, we have

$$u = y' - 2y = -3Be^{-3x} - 2Be^{-3x} = -5Be^{-3x},$$

but since we know u satisfies $u' = -3u$, we have

$$y'' - 2y' = -3(y' - 2y) = -3y' + 6y,$$

or equivalently $y'' + y' - 6y = 0$. ■

Exercise 17. *Let $y = f(x)$ be a smooth solution of the differential equation*

$$xy'' + x(y')^2 = 1 - e^{-x},$$

for all real x. If f has an extremum at a point $x = a$, show that this extremum is a minimum.

Exercise 18 (Optional. Bug Family Reunion.)**.** *Four bugs sit at the corners of a square table of side a. At the same instant they all begin to walk with the same speed, each moving steadily toward the bug on its right. If a polar coordinate system is established on the table, with the origin at the center and the polar axis along a diagonal, find the path of the bug that starts on the polar axis and the total distance it walks before all bugs meet at the center. This is taken from [23, p. 44]. [Answer. $r = \frac{a}{\sqrt{2}}e^{-\theta}$. Total distance is a.]*

2.2 The Nature of Solutions of Partial Differential Equations

Describing general solutions of PDEs, even linear ones, are much more harder than ODEs. One of the rare cases, besides first-order linear PDEs, where general solution can easily be described is one-dimensional *wave equation*

$$u_{tt} = u_{xx}. \tag{11}$$

One can easily check that, for any two second-differentiable real-valued functions $f(z)$ and $g(z)$ of real variable z, the function

$$u(x,t) = f(x+t) + g(x-t),$$

satisfies (11).

Exercise 19. *Check that all the following functions satisfy the Laplace equation $u_{xx} + u_{yy} = 0$:*

$$1, \quad x, \quad y, \quad xy, \quad x^2 - y^2, \quad x^3 - 3xy^2, \quad x^4 - 6x^2y^2 + y^4, \quad x^3y - xy^3, \quad \ldots,$$

$$e^{\pm\lambda y}\cos(\lambda x), \quad e^{\pm\lambda y}\sin(\lambda x), \quad \log(x^2+y^2), \quad \arctan\frac{y}{x},$$
$$(x^2+y^2)^{\pm\frac{n}{2}}\cos\left(n\arctan\frac{y}{x}\right), \quad (x^2+y^2)^{\pm\frac{n}{2}}\sin\left(n\arctan\frac{y}{x}\right),$$

where λ is a constant, and n is a natural number.

Exercise 20. *Check that all the following functions satisfy the heat equation $u_t = u_{xx}$:*

$$1, \quad x, \quad x^2+2t, \quad x^3+6xt, \quad t^{-\frac{1}{2}}e^{-\frac{x^2}{4t}}, \quad e^{\lambda^2 t\pm\lambda x}, \quad e^{-\lambda^2 t}\cos(\lambda x), \quad e^{-\lambda^2 t}\sin(\lambda x),$$

where λ is a constant.

2.3 Using Computers

Each three most famous numerical and symbolic computing softwates MATLAB, MAPLE and MATHEMATICA has numerous commands to deal with differential equations. One could simply use online website `http://www.wolframalpha.com` to solve simple equations. For example typing

$$\text{“solve } y' = x + y^2\text{”},$$

in its command line yields

$$y = \frac{x^{\frac{3}{2}}\left(-c_1 J_{-\frac{4}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right) + c_1 J_{\frac{2}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right) - 2J_{-\frac{2}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right)\right) - c_1 J_{-\frac{1}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right)}{2x\left(c_1 J_{-\frac{1}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right) + J_{\frac{1}{3}}\left(\frac{2x^{\frac{3}{2}}}{3}\right)\right)}, \tag{12}$$

where J_p is Bessel function of first kind introduced in page 94.

2.4 Separable First-Order Ordinary Differential Equations

Elementary functions and their integrals are not enough to express solutions of most ODEs even as simple as

$$y' = x + y^2,$$

mentioned at page 26. A fundamental special class of first-order ODEs with explicit solution are **separable** equations:

$$\boxed{\frac{dy}{dx} = G(x)H(y)},$$

where x and y are real variables, and G and H are real-valued functions of x and y respectively. This equation can readily be solved by the following procedure:

$$\frac{dy}{H(y)} = G(x)dx \Longrightarrow \int\frac{dy}{H(y)} = \int G(x)dx + C.$$

This shows the miracle of using Leibniz's symbol $\frac{dy}{dx}$ for the derivative of y with respect to x.

Example 5. The differential equation

$$\frac{dy}{dx} = 2xy^2 - x^3y^2,$$

might not look separable at first sight, but really is, because after factoring common y^2 expression on the right hand side, we have

$$\frac{dy}{dx} = y^2(2x - x^3), \tag{13}$$

or equivalently

$$\int \frac{dy}{y^2} = \int (2x - x^3)\,dx. \tag{14}$$

After integration

$$\frac{-1}{y} = x^2 - \frac{x^4}{4} + C,$$

or equivalently

$$y = \frac{4}{x^4 - 4x^2 - 4C}.$$

Replacing constant $-4C$ by another constant K, we have

$$y = \frac{4}{x^4 - 4x^2 + K},$$

as our solution.

But some solution is missing! The constantly zero function satisfies our differential equation but is not among our family of solutions. How did we miss it? We missed it when we did division by y^2 getting from (13) to (14). ■

Exercise 21. *Solve* $\frac{dy}{dx} = e^{2x-3y-1}$.

3 Modeling of Dynamical Phenomena by Differential Equations

In engineering, biology, physics, chemistry, economics, politics, etc., people translate dynamic phenomena into the language of differential equations. We illustrate this process in this chapter by several examples.

3.1 Modeling by Ordinary Differential Equations

Example 6. Let $P(t)$ denote human population of a country at time t. For simplicity of our model, we assume an isolated country so that there is no immigration or emigration. A simple way to model the dynamics of P is to assume that people are born (respectively, die) with rate BP (respectively, DP), where where B and D are constants. Therefore in short time interval $[t, t+dt]$, the change in population is

$$dP = P(t+dt) - P(t) = BP(t)dt - DP(t)dt,$$

hence P satisfies ODE

$$\frac{dP}{dt} = kP,$$

where $k = B - D$ is a constant. This is called the **simple growth** (or **decay**) **model**. This is a separable first-order ODE, hence solved by method of Section 2.4 as

$$P(t) = P_0 e^{kt},$$

where $P_0 = P(0)$ is the initial population.

Here comes a more sophiscated model. Either because of limited food supply or cultural reasons, birth parameter B above is not a constant, but decreases as population grows. If we model this by a simple linear model $B = B_1 - B_2 P$, with B_1 and B_2 positive constants, then infinitesimal variation of population is

$$dP = P(t+dt) - P(t) = (B_1 - B_2 P(t))P(t)dt - DP(t)dt,$$

hence P satisfies ODE

$$\frac{dP}{dt} = (B_1 - D - B_2 P)P = \frac{1}{B_2}\left(\frac{B_1 - D}{B_2} - P\right)P,$$

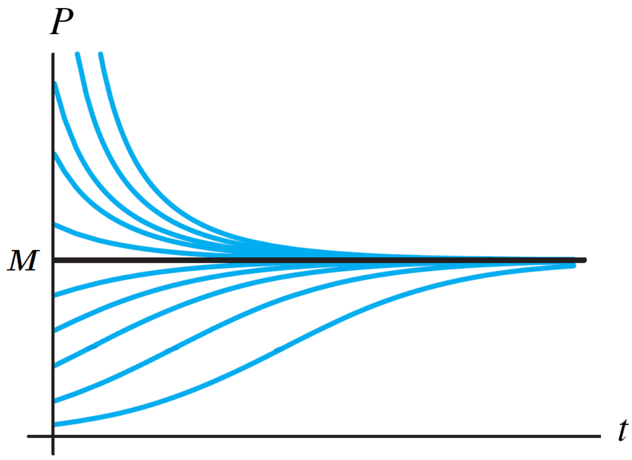
or more conveniently

$$\frac{dP}{dt} = kP(M - P),$$

with positive constants k and M. This is called the **logistic model**. This is again separable, hence solved as

$$P(t) = \frac{P_0 M}{P_0 + (M - P_0)e^{-kMt}}, \tag{15}$$

where again $P_0 = P(0)$ is the initial population. Some of these solutions, for different P_0, are depicted in the following figure.



Both the analytic solution (15) and the figure show that, with whatever initial population P_0 you start with, $P(t)$ approaches constant level M as $t \to \infty$. Constant M, called **carrying capacity**, is the maximum population that the environment can sustain on a long-term basis. ■

Example 7. It is reasonable to assume that the rate at which a hot body cools down is proportional to the difference in temperature between it and its surroundings; this is usually called **Newton's law of cooling**. If t denotes time, T the temperature of a hot bject, a the environment temperature, then the law above translates into

$$\frac{dT}{dt} = -k(T - a),$$

where k is a positive constant. The minus sign accounts for temperature decrease. This is again separable. ■

Exercise 22. *An object is heated to* $110°C$*, and then placed in an open area of temperature* $10°C$*. After* 1 *hour its temperature is* $60°C$*. How much additional time is required for it to cool to* $35°C$*?*

Example 8. Suppose that a small body of mass m kilograms is released at low height h meters above earth surface. The motion is affected by the earth's gravity with constant gravitational acceleration g (approximately 10 meters per second square), and air resistance force proportional, with constant k, to body's velocity. If $y(t)$ denotes the height of the body at time t then by Newtons's second law of motion

$$m\ddot{y} = -mg - k\dot{y},$$

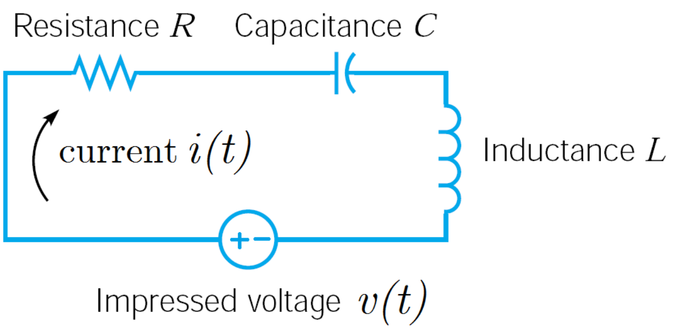
where dot stands for time derivative. To find $y(t)$, we need to solve the initial value problem

$$m\ddot{y} + k\dot{y} + mg = 0, \quad y(0) = h, \quad \dot{y}(0) = 0.$$

You can solve it using methods of Chapter 6. ■

Example 9. There are two important electrical and mechanical systems modeled by the same kind of differential equations.

RLC circuit. The following figure shows a series RLC circuit. We assume voltage $v(t)$ is known, and try to describe current $i(t)$.



Maxwell equations could be used to show that[6] as you pass along current i (measured in amperes) through a resistance R (measured in ohms), an inductance L (measure in henries), or a capacitance C (measured in farads), the voltage (measured in volts) drops, respectively, by

$$Ri, \quad L\frac{di}{dt}, \quad v_C(0) + \frac{1}{C}\int_0^t i(\tau)d\tau,$$

where $v_C(0)$ is the voltage on the capacitor at time $t = 0$. Resistors dissipate energy, but inductances and capacitors save electromagnetic energy. Quantitatively, resistor R with current i dissipates Ri^2 joules per second in the form of heat, and inductance L (resp. capacitor C) with current i (resp. voltage v) has already saved $\frac{1}{2}Li^2$ (resp. $\frac{1}{2}Cv^2$) joules of magnetic (resp. electric) energy.

According to basic information above, if we pave the RLC circuit once along current $i(t)$, we get

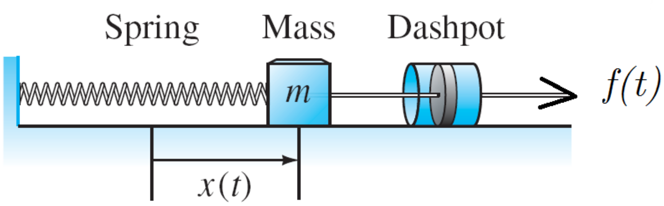
$$v = Ri + v_C(0) + \frac{1}{C}\int_0^t i(\tau)d\tau + Li',$$

[6] [22, p. 170-172].

and after differentiation

$$Li'' + Ri' + \frac{1}{C}i = v'.$$

Mass-Spring-Dashpot system. The following figure shows a mass-spring-dashpot system. We neglect friction. We assume external force $f(t)$ is known, and try to describe $x(t)$, the displacement of mass from equilibrium motionless position.



Let us recall that a stretched or compressed spring exerts restorative force proportional to the displacement from its equilibrium position. Dashpot or damper is a mechanical device resisting motion by producing force proportional to the speed but acting in the opposite direction. By Newton's second law

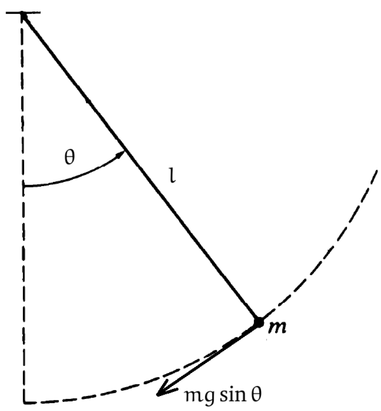
$$mx'' = f - kx - cx',$$

or

$$mx'' + cx' + kx = f,$$

where k and c are respectively spring and damper's constants. You can solve it using methods of Chapter 6. ■

Example 10. The following figure Let us model the motion of a pendulum consisting of a mass m at the end of a rod of length l and negligible mass. Let θ denote the deviation angle form the horizontal position.



By applying Newton's second law of motion at the direction tangent to circle of rotation

$$mg\sin\theta = -ml\frac{d^2\theta}{dt^2},$$

hence the evolution of θ is governed by the ODE

$$\frac{d^2\theta}{dt^2} = -k\sin\theta, \tag{16}$$

where $k = \frac{g}{l}$ is a constant. This is nonlinear equation, but for small deviations θ, could be approximated by linear equation

$$\frac{d^2\theta}{dt^2} = -k\theta.$$

Later in Chapter 6, you will see that solutions to this latter linear equation are

$$\theta = A\cos kt + B\sin kt,$$

with constants A and B. This shows that $\theta(t)$ repeats itself every

$$T = \frac{2\pi}{k} = 2\pi\sqrt{\frac{l}{g}},$$

seconds. ■

Exercise 23 (Optional)**.** *If you are interested in mechanics, try to derive (16) by either of the following ways: (a) Newton's second law for rotation* $\tau = I\alpha$*. (b) Conservation of energy. (c) Euler-Lagrange equation* $\frac{\partial L}{\partial q} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}$, $L = T - V$.

Example 11. Look at Figure 1. Each part shows two families of curves, each member of the one family intersects all the members of the other family orthogonally. These are called **orthogonal families**, and appear at numerous places is science. For example, in electrostatics, field lines and iso-potential curves, and in thermodynamics, heat lines and isothermal curves and lines of heat flow, are orthogonal families.

Let us discuss this notion with an example. The equation

$$x^2 + y^2 = R^2, \tag{17}$$

with R constant, expresses a circle of radius R in the plane centered at origin. When the *parameter* R varies, we get a family of concentric circles. Let us try to find another family of smooth curves orthogonal to this family at any point of intersection. Remember that two lines in the plane with slopes m_1 and m_2 intersect orthogonally exactly when $m_2 m_1 = -1$. This suggest a way to find the

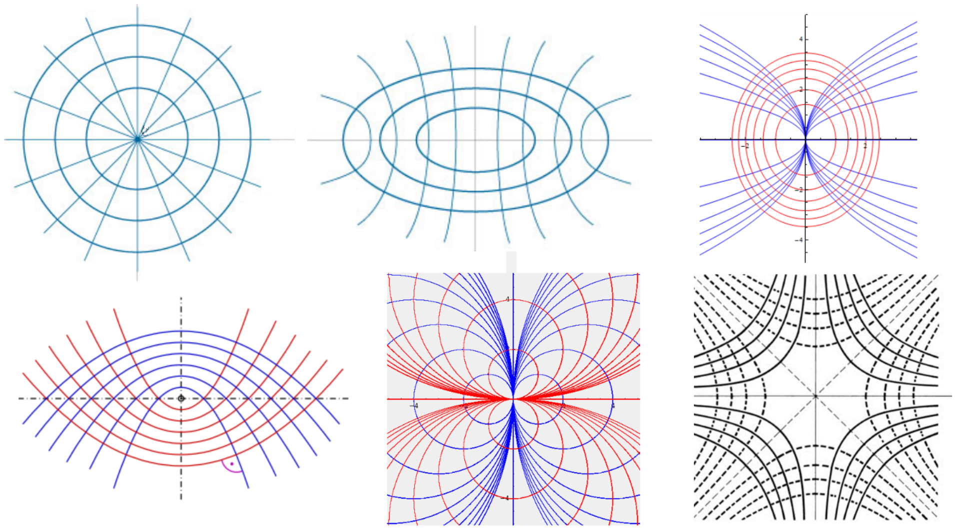


Figure 1: Some orthogonal families.

orthogonal family: *first find the differential equation describing the original family; then replacing* y' *by* $\frac{-1}{y'}$ *gives the differential equation of the orthogonal family; solving this new equation gives the* ***orthogonal family***. In our case, assuming that (17) gives y locally in terms of x, differentiating both sides with respect to x gives $2x + 2yy' = 0$ or equivalently

$$x + yy' = 0,$$

which is the ODE describing family of circles centered at origin. Therefore $y' = \frac{y}{x}$ or equivalently

$$xy' = y,$$

is the ODE describing orthogonal family. The solution to this separable equation is $y = Cx$, namely lines passing through origin. ∎

Exercise 24. *Let* $u(x, y)$ *and* $v(x, y)$ *be two smooth real-valued functions of real variables* x *and* y. *Prove that one-parameter families* $u = C$ *and* $v = K$, *with* C *and* K *constant parameters, are orthogonal if and only if* $u_x v_x + u_y v_y = 0$. *[Hint: use implicit function theorem.]*

3.2 Modeling by Systems of Ordinary Differential Equations

Example 12. Consider two points P and Q moving in xy plane. At time $t = 0$, P is at origin, and then starts to move along y-axis with constant speed of unity. At time $t = 0$, Q is at $(1, 0)$, and then *pursues* P with constant speed of a times unity, hence the direction of motion of Q is always toward Q. What is the curve of pursuit?

Let $(x(t), y(t))$ denote the position of Q. Speed condition translates into

$$a = \frac{ds}{dt} = \frac{\sqrt{dx^2 + dy^2}}{dt} = \sqrt{\dot{x}^2 + \dot{y}^2},$$

where ds is infinitesimal arc-length, and pursue condition says

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

Since $\dot{x} > 0$, pursuit curve is discribed by the system

$$\begin{cases} \frac{dx}{dt} = \frac{-ax}{\sqrt{x^2+(y-t)^2}} \\ \frac{dy}{dt} = \frac{a(t-y)}{\sqrt{x^2+(y-t)^2}} \end{cases}.$$

These are studied in Chapter 9. ■

Exercise 25 (Tractrix). *Consider two points* P *and* Q *moving in* xy *plane. At time* $t = 0$*,* P *is at origin, and then starts to move along* y*-axis with constant speed of unity. At time* $t = 0$*,* Q *is at* $(1, 0)$*, and then pursues* P *in a way that the distance between them is always* 1*. Find the equation of the the curve of pursuit.*

Example 13. This is copied from [23, p.434-436]. There is a constant struggle for survival among different species living in the same environment. One kind of animal survives by eating another; a second, by developing methods of evasion to avoid being eaten, and so on. As a simple example of this universal conflict between the predator and its prey, let us imagine an island inhabited by foxes and rabbits. The foxes eat rabbits, and the rabbits eat clover. We assume that there is so much clover that the rabbits always have an ample supply of food. When rabbits are abundant, foxes flourish their population grows, but after a while when foxes become too numerous and eat too many rabbits, they enter a period of famine and their population begins to decline. As foxes decrease, rabbits become relatively safe and their population starts to increase again. This triggers a new increase in the fox population, and as time goes on we see an endlessly repeated cycle of interrelated increases and decreases in the populations of the two species.

Let $x(t)$ and $y(t)$ denote the number of rabbits and foxes, respectively. If there where no foxes $\frac{dx}{dt} = ax$, where a is a positive constant. It is natural to assume that the number of encounters per unit time between rabbits and foxes is jointly proportional to x and y. If we further assume that a certain proportion of these encounters result in a rabbit being eaten, then we have

$$\frac{dx}{dt} = ax - bxy,$$

where b is another positive constant. With the same argument

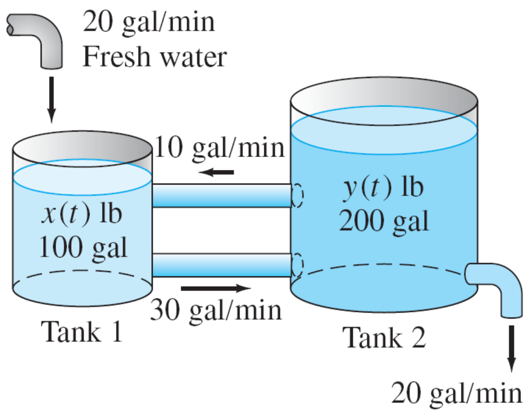
$$\frac{dy}{dt} = -cy + dxy,$$

where c and d are positive constants. Therefore we have Lotka-Voterra's prey-predator system of ODEs

$$\begin{cases} \frac{dx}{dt} = ax - bxy \\ \frac{dy}{dt} = -cy + dxy \end{cases},$$

studied in Chapter 9. ■

Example 14. Consider two brine tanks connected shown in the following figure. Let us assume that at time t, tank 1 contains $x(t)$ pounds of salt in 100 gal of brine, and tank 2 contains $y(t)$ pounds of salt in 200 gal of brine. The brine in each tank is kept uniform by stirring all the time. Also brine is pumped from each tank to the other at the rates indicated in the figure. In addition, fresh water flows into tank 1 at 20 gal per min, and the brine in tank 2 flows out at 20 gal per min.



To derive the law governing evolution of $x(t)$ and $y(t)$, let us consider infinitesimal time interval $[t, t + dt]$. During this interval, the change in salt in tank 1 is

$$\begin{aligned} dx &= x(t+dt) - x(t) = dt \text{ min} \times 10\frac{\text{gal}}{\text{min}} \times \frac{y(t)}{200} \frac{\text{lb}}{\text{gal}} - dt \text{ min} \times 30\frac{\text{gal}}{\text{min}} \times \frac{x(t)}{100} \frac{\text{lb}}{\text{gal}} \\ &= (0.05y - 0.3x)dt, \end{aligned}$$

hence

$$\frac{dx}{dt} = -0.3x + 0.05y.$$

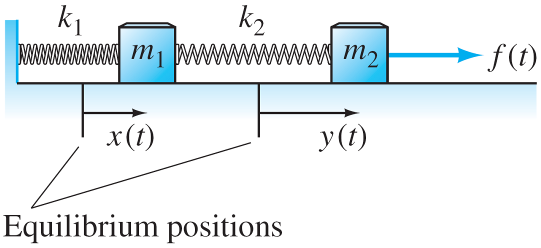
Similarly, one gets $\frac{dy}{dt} = 0.3x - 0.15y$. Therefore the dynamics is governed by the ODE system

$$\begin{cases} \frac{dx}{dt} = -0.3x + 0.05y \\ \frac{dy}{dt} = 0.3x - 0.15y \end{cases}.$$

You can solve it using methods of Chapter 8. ■

Example 15. Consider the mass-spring mechanical system shown in the following figure. Neglect friction.

Variable $f(t)$ shows the external force exerted on m_2. Variable $x_1(t)$ denote the displacement of m_1 from its equilibrium position when there is no external force and motion. Similarly for $x_2(t)$.

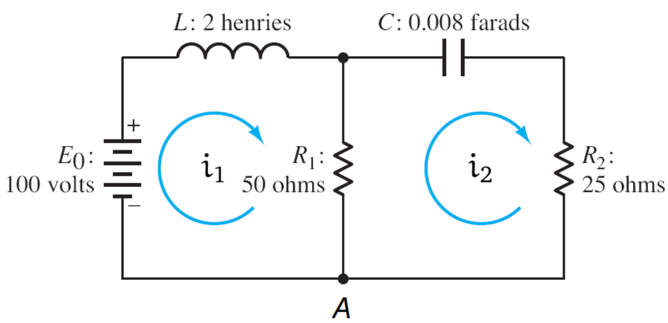


Applying Newton's second law of motion

$$\begin{cases} m_1 \frac{d^2x}{dt^2} = -k_1 x_1 + k_2(y - x) \\ m_2 \frac{d^2y}{dt^2} = -k_2(y - x) + f(t) \end{cases}.$$

You can solve it using methods of Chapter 8. ■

Example 16. Let us analyze the electrical network shown in the following figure.



Paving the left loop along current i_1 starting form point A, the change in potential is

$$0 = 100 - 2\frac{di_1}{dt} - 50(i_1 - i_2) = 0,$$

hence

$$\frac{di_1}{dt} = -25i_1 + 25i_2 + 50.$$

Paving the right loop along current i_1 starting form point A, the change in potential is

$$0 = -50(i_2 - i_1) - \frac{1}{0.008}\int i_2 dt - 25i_2,$$

and after differentiation

$$\frac{di_2}{dt} = \frac{2}{3}\frac{di_1}{dt} - \frac{5}{3}i_2,$$

and after inserting $\frac{di_1}{dt}$ from our previous equation

$$\frac{di_2}{dt} = \frac{-50}{3}i_1 + \frac{45}{3}i_2 + \frac{100}{3}.$$

Therefore the dynamics is governed by the ODE system

$$\begin{cases} \frac{di_1}{dt} = -25i_1 + 25i_2 + 50 \\ \frac{di_2}{dt} = \frac{-50}{3}i_1 + \frac{45}{3}i_2 + \frac{100}{3} \end{cases}.$$

You can solve it using methods of Chapter 8. ■

Example 17. Let us model the dynamical system consisting of N isolated bodies attracting themselves according to Newton's law of gravitation. For each $i = 1, \ldots, N$, let the i-th body has mass m_i, and is situated at the point $(x_i, y_i, z_i) \in \mathbb{R}^3$. for each i, the force

$$F_{ij} = \sum_{j\neq i} \frac{Gm_i m_j}{\left((x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2\right)^{\frac{3}{2}}} \left((x_j - x_i)\hat{x} + (y_j - y_i)\hat{y} + (z_j - z_i)\hat{x}\right),$$

affecting on i-th body gives it acceleration

$$\frac{d^2x_i}{dt^2}\hat{x} + \frac{d^2y_i}{dt^2}\hat{y} + \frac{d^2z_i}{dt^2}\hat{z} = \frac{F_{ij}}{m_i}.$$

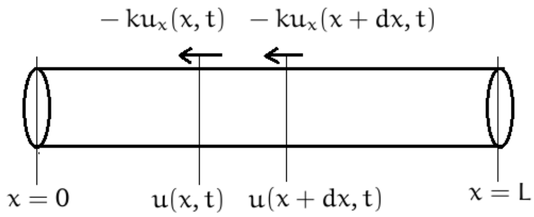
Therefore system dynamics is governed by the system

$$\begin{cases} \frac{d^2x_i}{dt^2} = \sum_{j\neq i} \frac{Gm_j(x_j - x_i)}{\left((x_j-x_i)^2+(y_j-y_i)^2+(z_j-z_i)^2\right)^{\frac{3}{2}}}, & i = 1, \ldots, N, \\ \frac{d^2y_i}{dt^2} = \sum_{j\neq i} \frac{Gm_j(y_j - y_i)}{\left((x_j-x_i)^2+(y_j-y_i)^2+(z_j-z_i)^2\right)^{\frac{3}{2}}}, & i = 1, \ldots, N, \\ \frac{d^2z_i}{dt^2} = \sum_{j\neq i} \frac{Gm_j(z_j - z_i)}{\left((x_j-x_i)^2+(y_j-y_i)^2+(z_j-z_i)^2\right)^{\frac{3}{2}}}, & i = 1, \ldots, N, \end{cases},$$

of 3N nonlinear second-order ordinary differential equations, called the N**-body problem**. Only for $N = 2$ analytic solutions been found. When $N > 2$ numerical solutions become unstable. ■

3.3 Modeling by Partial Differential Equations

Example 18. Let us model the dynamics of temperature distribution on a thin cylindrical metal rod of length L shown in the following figure.



We assume lateral surface is completely isolated, change of temperature is just because of heat conduction in x direction. We also know the temperature at boundary points $x=0$ and $x=L$. Let $u(x,t)$ denote the temperature at time t in place x. We are going to use two physical laws:

- Heat flows from hot regions to cold regions, and the amount of heat which passes per unit time per unit area of a small flat surface in space is proportional to the rate of change of temperature with respect to distance in a direction perpendicular to that surface. In mathematical terms $q=-k\nabla u$, where q is heat per unit area per unit time, k is a constant, and $\nabla u = u_x\hat{x}+u_y\hat{y}+u_z\hat{z}$ is the gradient of temperature. This reduces to $q=-ku_x$ in our one-dimensional problem.

- If the temperature of a body changes by heat exchange with environment, then this change is proportional to the amount of heat exchanged. In mathematical terms, $q=mc\Delta u$, where m is the body mass, c a constant, and Δu the change in temperature.

Returning back to our problem, we consider infinitesimal changes dx and dt in x and t, respectively. That part of the rod between x and $x+dx$, which has mass proportional to dx, during time interval $[t,t+dt]$, gets proportionally

$$(-u_x(x,t)+u_x(x+dx,t))\,dt,$$

units of heat, and this makes temperature change

$$u(x,t+dt)-u(x,t).$$

Therefore, after scaling units, we have

$$(-u_x(x,t)+u_x(x+dx,t))\,dt = (u(x,t+dt)-u_{x,t}(x,t))\,dx,$$

or equivalently:

$$\boxed{u_t = u_{xx}}.$$

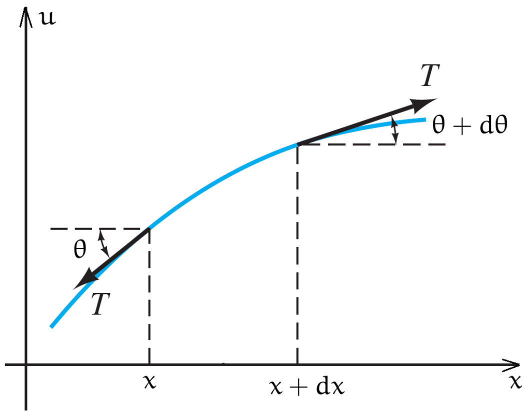
This is called one-dimensional **heat equation**.

Let us consider a specific problem. We assume $L = \pi$, the boundary points $x = 0$ and $x = \pi$ are both kept at constant temperatures 0 and 0, and that the initial temperature distribution is $f(x)$. The problem is to find the solution of

$$\begin{cases} \text{PDE}: & u_t = u_{xx}, \quad 0 < t < \infty,\ 0 < x < \pi \\ \text{BCs}: & u(0,t) = u(\pi,t) = 0, \quad 0 \leqslant t < \infty\,, \\ \text{IC}: & u(x,0) = 1, \quad 0 < x < \pi \end{cases}$$

where abbreviations BCs and IC above, respectively, stand for bondary conditions and initial condition. This is solved in Example 10.3.2. Also notice Exercise . ■

Example 19. Let us model low-amplitude vibrations of a flexible string of length L shown in the following figure.



Let $u(x,t)$ denote the displacement at time t of the point with coordinate x of the string. Experiments show that the tension force T is tangent to the string and proportional to $\sqrt{1+u_x^2}$. Consider that part of the string between x and $x+dx$, which has mass approximately proportional to dx. The force exerted on this portion along u-axis is approximately

$$-T\sin(\theta) + T\sin(\theta + d\theta) \approx T\left(\tan(\theta+d\theta) - \tan(\theta)\right) = T\left(u_x(x+dx,t) - u_x(x,t)\right),$$

hence by Newton's second law

$$\sqrt{1+u_x^2}\left(u_x(x+dx,t) - u_x(x,t)\right) = u_{tt}(x,t)dx,$$

or equivalently

$$u_{tt} = \sqrt{1+u_x^2}\,u_{xx}.$$

Assuming $|u_x|$ small, we get a linearized version

$$\boxed{u_{tt} = u_{xx}},$$

called one-dimensional **wave equation**. Here we modeled *transversal* vibrations. One could show *longitudinal* vibrations, for example sound waves, obey the same equation [24, section 2.1]. ■

Example 20. Consider a small grain suspended on the surface layer of water. Because of collision with water molecules, it follows a random path called **Brownian motion**. Let water surface is distributed in region R in xy-plane. We want to find the time it takes for the particle to reach the boundary of D. This quantity for the particle in position (x, y) at time t is denoted by $u(x, y, t)$. For simplicity we assume the particle only could move in horizontal and vertical paces. The particle at point (x, y) must have reached this point from one of its four neighbors

$$(x - dx, y), \quad (x + dx, y), \quad (x, y - dy), \quad (x, y + dy),$$

with equal probability $\frac{1}{4}$ for each neighbor. Therefore

$$u(x, y, t + dt) = dt + \frac{1}{4}\left(u(x - dx, y, t) + u(x + dx, y, t) + u(x, y - dy, t) + u(x, y + dy, t)\right).$$

Using Taylor expansion of the form

$$f(z + dz) = f(z) + f'(z)dz, \quad f(z + dz) = f(z) + f'(z)dz + \frac{1}{2}f''(z)dz^2,$$

we have

$$\begin{aligned} u + u_t dt = dt &+ \frac{1}{4}\left(u - u_x dx + \frac{1}{2}u_{xx}dx^2\right) + \frac{1}{4}\left(u + u_x dx + \frac{1}{2}u_{xx}dx^2\right) \\ &+ \frac{1}{4}\left(u - u_y dy + \frac{1}{2}u_{yy}dy^2\right) + \frac{1}{4}\left(u + u_y dy + \frac{1}{2}u_{yy}dy^2\right), \end{aligned}$$

which simplifies to

$$(u_t - 1)dt = \frac{1}{4}\left(u_{xx}dx^2 + u_{yy}dy^2\right).$$

If we assume

$$dt = Cdx^2 = Cdy^2,$$

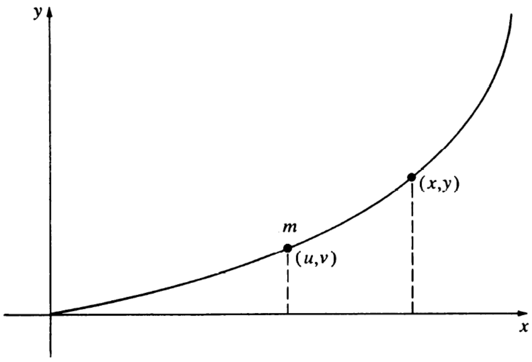
for some constant C, then our equation becomes

$$u_t = 1 + k(u_{xx} + u_{yy}),$$

for some positive constant k. This is two-dimensional **heat equation**. ■

3.4 An Integro-differential Equation

Example 21. Does there exist a curve in the plane for which the time taken by an object sliding on it, under gravity and without friction, to its lowest point is independent of the starting point? To mathematically model this problem, consider the following figure, where C is our desired curve, and the bead is released from the point (x, y), and slides to the lowest point $(0, 0)$. As bead reaches



the point (u, v) on the curve, its kinetic energy is

$$\frac{1}{2}m\left(\frac{dL}{dt}\right) = mg(y - v),$$

where L is the length of the curve between (x, y) and (u, v). Thus

$$dt = \frac{dL}{\sqrt{2g(y-v)}} = \frac{\sqrt{1 + \left(\frac{du}{dv}\right)^2}}{\sqrt{2g(y-v)}}dv.$$

Therefore the time t taken to reach $(0, 0)$ for (x, y) is

$$t = \int_0^y \frac{\sqrt{1 + \left(\frac{du}{dv}\right)^2}}{\sqrt{2g(y-v)}}dv.$$

This shows that out desired curve $u = u(v)$ satisfies

$$u(0) = 0, \quad \int_0^y \frac{\sqrt{1 + \left(\frac{du}{dv}\right)^2}}{\sqrt{y-v}}dv = C,$$

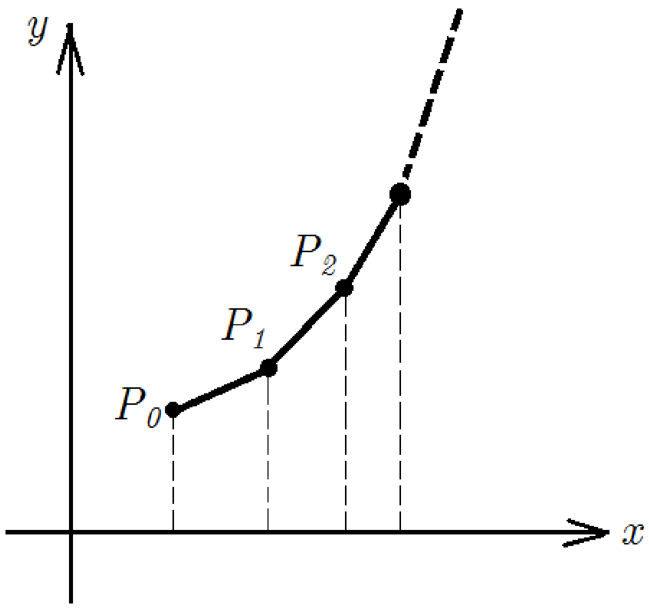
for each y, and a constant C. We solve this equation in Section 5.3. ∎

4 First-Order Ordinary Differential Equations

The general form of a first-order differential equation is

$$\boxed{\frac{dy}{dx} = F(x, y)}, \tag{18}$$

where x and y are real variables, and F is a real-valued function of x and y. There is simple *geometric* way to think about this equation. Since (18) gives the evolution law of y under infinitesimal change of x, fixing any point $P_0 = (x_0, y_0)$ in plane, one could generally find a *unique* curve $y = f(x)$ passing through P_0 and satisfying the equation. More geometrically, at each point (x, y) on the curve, the slope $f'(x) = \frac{dy}{dx}$ of the tangent line equals the prescribed value $F(x, y)$. This curve is called the **integral curve** of (18) passing through P_0. Here is a step-by-step procedure to approximately construct this curve, visualized in the following figure. Fixing a small positive number h, on interval $[x_0, x_0 + h]$, the integral curve is approximated by the tangent line $y = y_0 + F(x_0, y_0)(x - x_0)$, until it reaches the endpoint $P_1 = (x_1, y_1)$, and then similarly on interval $[x_1, , x_1 + h]$, the integral curve is approximated by $y = y_1 + F(x_1, y_1)(x - x_1)$, etc.



A byproduct of this construction is **Euler's method** to numerically solve the initial value problem

$$\frac{dy}{dx} = F(x, y), \quad y(a) = b.$$

To find a table of values of the solution $y = f(x)$ on $[a, b]$, choosing large natural number N, and setting the pace $h = (b-a)/N$, and the nodes $x_n = a+nh$, $n = 0, 1, \dots$, we compute the sequence

$$y_0 = b, \quad y_{n+1} = y_n + hF(x_n, y_n), \quad n = 0, 1, \dots .$$

Then y_n approximates $f(x_n)$ for each n.

Example 22. Let us apply Euler's method to the simple initial value problem

$$\frac{dy}{dx} = ky, \quad y(0) = C,$$

where k and C are constants To find an approximate value for $y(x)$, we choose large integer N, and set our pace $h := \frac{x}{N}$, and nodes $x_n = nh$, $n = 0, \dots, N$. Then

$$y_0 = C, \quad y_{n+1} = y_n + hky_n = (1 + hk)y_n, \quad n = 0, \dots, N-1,$$

hence

$$y(x) = y(x_N) \approx y_N = (1 + hk)^N y_0 = (1 + hk)^N C = \left(1 + \frac{xk}{N}\right)^N C$$

By the formula

$$\lim_{N\to\infty} \left(1 + \frac{a}{N}\right)^N = e^a,$$

we deduce $y(x) \approx Ce^{kx}$. ∎

To visualize 18 in one shot, through each point of a representative collection of points (x, y) in the plane, we draw a short line segment with slope $F(x, y)$. This gives a **slope field** of the equation. The following figure shows the slope field and some integral curves of the logic equation $\frac{dy}{dx} = y(1-y)$. You could use online calculator `http://www.bluffton.edu/homepages/facstaff/nesterd/java/slopefields.html` to plot slope fields.

Exercise 26. *Fix parameter $\epsilon > 0$. Let $t \geqslant 0$ be time variable measured in seconds. Separately, consider each of the following equations:*

$$\frac{dx_1}{dt} = 1 + x_1, \quad \frac{dx_2}{dt} = x_2^2, \quad \frac{dx_3}{dt} = x_3^\epsilon, \quad \frac{dx_4}{dt} = 1 + x_4^2, \quad \frac{dx_5}{dt} = 1 + x_5^{\frac{5}{2}},$$

$$\frac{dx_6}{dt} = 1 + x_6^{\frac{1}{10}}, \quad \frac{dx_7}{dt} = 1 + x_7^{\frac{3}{2}}, \quad \frac{dx_8}{dt} = 1 + x_8^{1+\epsilon}, \quad \frac{dx_9}{dt} = \epsilon(x_9 - 2)^2,$$

together with initial condition $x_i(0) = 1$, for each $i = 1, \dots, 9$.

(a) Show that $x_1(t)$ eternally exists, namely $x_1(t)$ is a smooth function of t ranging on whole $(0, \infty)$.

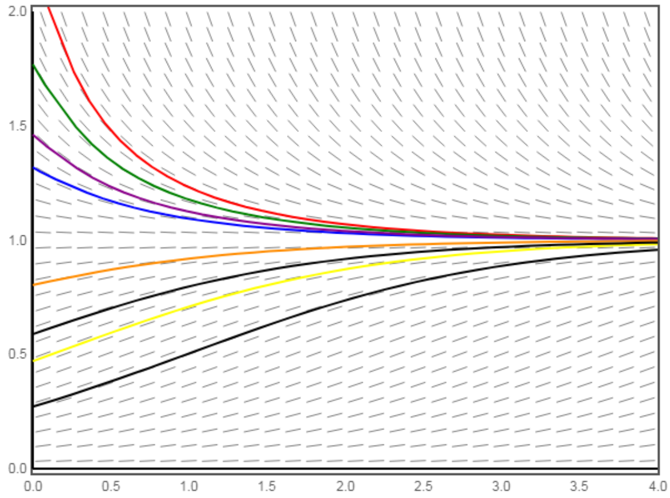


Figure 2: Slope field and some integral curves of the logic equation $\frac{dy}{dx} = y(1-y)$.

(b) Show that $x_2(t)$ *blows up in some finite time* T*, namely there is* $T > 0$ *such that* $x_2(T^-) = \infty$*. Find* T*. How does* T *change as* A *increases?*

(c) For what values of parameter ϵ*, does* $x_3(t)$ *blow blows up in finite time?*

(d) Show that $x_4(t)$ *blows up in less than* π *seconds.*

(e) Show that $x_5(t)$ *blows up in some finite time* T*. Can you say for sure that* $T < \pi$*? [Hint. Compare with (d).]*

(f) Show that $x_6(t)$ *eternally exists. [Hint. Compare with (a).]*

(g) Show that $x_7(t)$ *blows up in some finite time* T*.*

(h)(Optional) Is it true that for each value of parameter ϵ*,* $x_8(t)$ *blows up in some finite time?*

(i) Show that $x_9(t)$ *eternally exists, but in contrast with functions in parts (a) and (f), it remains bounded, namely there is some* $B > 0$ *such that* $|x_9(t)| < B$ *for all* $t \geqslant 0$*.*

4.1 The Existence and Uniqueness of Solutions

Theorem 6 (Picard-Lindelöf-Peano)**.** *Let* $F(x, y)$ *be a real-valued function of real variables* x *and* y*, defined on some open rectangle* U *containing the point* (x_0, y_0)*. If* F *is continuous on* U*, then the initial value problem*

$$\frac{dy}{dx} = F(x, y), \quad y(x_0) = y_0, \tag{19}$$

has a local solution around $x = x_0$*, namely there is a function* $y(x)$ *defined on some open interval* I *containing* a *such that*

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

for each $x \in I$. If both F and F_y are continuous on U, then (19) has exactly one local solution around $x = x_0$, namely any two solutions match on some interval around $x = x_0$.

For the proof refer to [17, p. 8-11], [4, chapter 6], [23, chapter 13], or [1, vol. II, Theorem 7.19].

Example 23. Consider the initial value problem

$$\frac{dy}{dx} = y^{\frac{1}{3}}, \quad y(0) = 0. \tag{20}$$

This is a separable equation, hence we write it as

$$y^{\frac{-1}{3}} dy = dx, \tag{21}$$

which after integration gives

$$y = \pm\left(\frac{2}{3}x - C\right)^{\frac{3}{2}}.$$

The initial condition forces C to vanish, hence $y = \pm\left(\frac{2}{3}x\right)^{\frac{3}{2}}$. This analysis shows that both functions

$$y_1 = \begin{cases} \left(\frac{2}{3}x\right)^{\frac{3}{2}}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}, \quad y_2 = -y_1,$$

satisfy (20) on whole real line. Note that in writing (21), we did a division by $y^{\frac{1}{3}}$, and so missed the trivial solution $y_3 \equiv 0$ of (20). Also for each constant $C \geqslant 0$, the functions

$$y_4 = \begin{cases} \left(\frac{2}{3}(x - C)\right)^{\frac{3}{2}}, & x \geqslant C \\ 0, & x < C \end{cases}, \quad y_5 = -y_4,$$

also satisfy (20). ∎

Exercise 27. *Fix real numbers a and b, and open interval I around a. Consider the initial value problem*

$$y(a) = b, \quad x\frac{dy}{dx} = 2y,$$

for each $x \in I$. Show that:

(a) If $a = b = 0$, the problem has infinitely many solutions.

(b) If $a = 0$ and $b \neq 0$, the problem has no solution.

(c) If $a \neq 0$, the problem has a unique solution if and only if I does not contain 0.

[Hint. It really helps to draw the family $y = Cx^2$ you get from separation of variables, and the trivial solution $y \equiv 0$ missed in this family.]

Exercise 28. *Let $y = y(x)$ be the local solution of initial value problem*

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

(a) Show that $1 \leqslant y' \leqslant x$ for all $x \geqslant 0$.

(b) Show that $x \leqslant y$ for all $x \geqslant 0$. Show that $y(x)$ does not blow up in finite $x \geqslant 0$, hence defines a unique function on whole $[0, \infty)$.

(c) Show that $\frac{x}{y^2} \to 0$ when $x \to \infty$.

(d) Show that y' approaches a constant when $x \to \infty$. Find that constant.

(e) Show that $\frac{y}{x}$ approaches a constant when $x \to \infty$. Find that constant.

[Hint. For each part use the previous one.]

Example 24. This example is from [1, vol. I, p. 342]. Let us try to investigate nonlinear ODE

$$y = xy' - \frac{1}{4}(y')^2. \tag{22}$$

A little inspection shows that $y = x^2$ is a special solution. Thus it might be useful to change dependent variable y to $u = y - x^2$.[7] Our new equation is then $u = -\frac{1}{4}(u')^2$, or equivalently $u' = \pm 2\sqrt{-u}$, which is separable, hence solved in the following way:

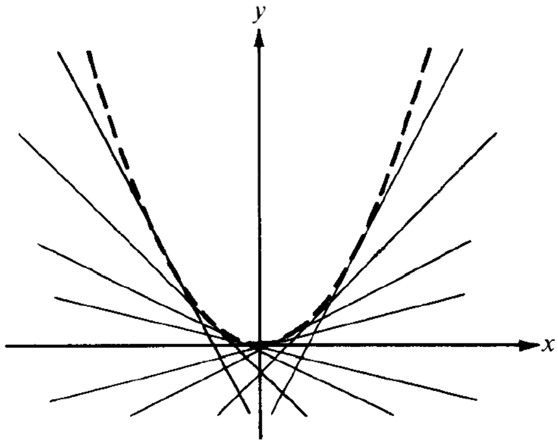
$$\int \pm 2 dx = \int (-u)^{-\frac{1}{2}} du \Longrightarrow C \pm 2x = 2(-u)^{\frac{1}{2}} \Longrightarrow u = -\left(\frac{C}{2} \pm x\right)^2,$$

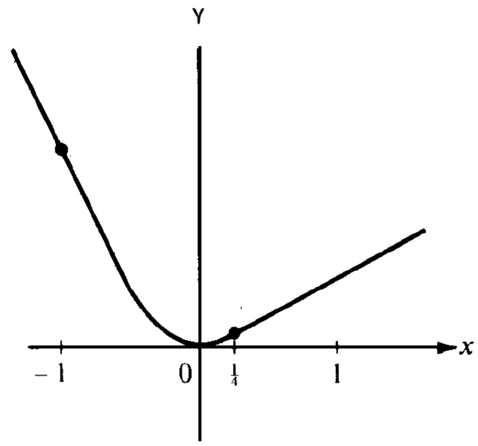
hence

$$y = x^2 - \left(\frac{C}{2} \pm x\right)^2 = -\frac{C^2}{4} \mp Cx,$$

or after digesting $\pm$ into C, we get $y = Cx - \frac{C^2}{4}$. So far we have found two solutions for (22), namely the special $y = x^2$, and the one-parameter family $y = Cx - \frac{C^2}{4}$. The left part of the following figure depicts them. One sees that the special solution, at each of its points, is tangent to one member of the one-parameter family of solutions. This is called an **envelope**.

[7] As a general suggestion, if you know a particular solution y_p to a differential equation, it might be helpful to change dependent variable y to $u = y/y_p$ or $u = y - y_p$.





Are there any other solutions? Yes! Can you see them? One can find infinitely many other solutions by piecing together members of the family with portions of the *envelope*, for example

$$y = \begin{cases} -2x-1, & x \leqslant -1 \\ x^2, & -1 \leqslant x \leqslant \frac{1}{4} \\ \frac{1}{2}x - \frac{1}{16}, & x > \frac{1}{4} \end{cases},$$

which is shown in the right part of the same figure. ∎

Exercise 29 (Optional). *Prove that the envelope of the one parameter family of curves* $F(x, y, C) = 0$ *could be found by eliminating* C *from the two equations* $F = 0$ *and* $\frac{\partial F}{\partial C} = 0$.

Exercise 30. *Consider the equation* $y^2 + (y')^2 = 1$. *Clearly constant functions* $y = 1$ *and* $y = -1$, *and also* $y = \sin(x + C)$, C *constant, are solutions. Manufacture infinitely many other solutions.*

Exercise 31. *Consider the equation*

$$y' = (\cot x)\left(\sqrt{y} - y\right).$$

(a) Solve this equation.
(b) Consider the one-parameter family of solutions

$$(1 - \sqrt{y})^2 \sin x = C.$$

Draw several members of this family, say $C = 1$, $C = -1$, $C = 0.001$, $C = 2$, $C = -2$. *(I used online plotter* `https://www.desmos.com/calculator`*.) Try to see solutions* $y \equiv 0$ *and* $y \equiv 1$ *as degenerate members of this family. Are there other solutions?*

For the rest of this chapter, we study several specal classes of first-order ODEs which can be solved explicitly.

4.2 Homogeneous Equations

The differential equation (18), if F is **homogeneous**, namely

$$F(ax, ay) = F(x, y),$$

for each x, y and a, becomes separable after changing the dependent variable u to $u = \frac{y}{x}$. Here is the process. Replacing y by xu in equation (18), we get

$$u + x\frac{du}{dx} = F(x, xu) = F(1, u),$$

which simplies into the separable equation

$$\frac{du}{dx} = \frac{F(1,u) - u}{x},$$

hence solved as

$$\int \frac{du}{F(1,u) - u} = \int \frac{dx}{x} + C.$$

Exercise 32. *Find the orthogonal trajectories of the family of all circles passing through the origin with their centers on the x-axis.*

4.3 Linear Equations

A first-order differential equation of the form

$$\boxed{\frac{dy}{dx} + P(x)y = R(x)}, \tag{23}$$

where P and R are real-valued functions of real variable x is called **linear**. Here comes a clever trick to solve it. Multiplying both sides by a function $\mu(x)$, called an **integrating factor**, to be determined later, our new equation is

$$\mu y' + \mu P y = \mu R.$$

If μ has the property that $\mu P y = \mu' y$, then our new equation could be written as $(\mu y)' = \mu R$, which is simply solved as

$$\mu y = \int \mu R dx.$$

It only remains to find μ with desired property $\mu P y = \mu' y$, or even, $\mu P = \mu'$. This is a *separable* equation, hence solved as

$$\int \frac{d\mu}{\mu} = \int P dx \Longrightarrow \log|\mu| = \int P dx \Longrightarrow \mu = \pm C e^{\int P dx}.$$

Our analysis proved the following result.

Proposition 3. *The general solution of equation (23) is*

$$y = \mu^{-1} \int \mu R dx,$$

where $\mu = e^{\int P dx}$.

Example 25. Let us solve the initial value problem

$$\frac{dy}{dx} - ay = R(x), \quad y(x_0) = b,$$

where a and b are real constants, and $R(x)$ a real-valued function. One integrating factor is $\mu = e^{\int -a dx} = e^{-ax}$, hence the general solution to $y' + ay = R(x)$ is

$$y = e^{ax} \int e^{-ax} R(x) dx = e^{ax} \left(\int_{x_0}^{x} e^{-a\xi} R(\xi) d\xi + C \right).$$

Enforcing the initial value $y(x_0)$, we get

$$b = Ce^{ax_0} \implies C = be^{-ax_0},$$

thus finally

$$\boxed{y = be^{a(x-x_0)} + \int_{x_0}^{x} e^{a(x-\xi)} R(\xi) d\xi}. \tag{24}$$

A formula worthy of remembrance. This clearly shows the effect of initial condition b and the input $R(x)$ on the output $y(x)$. We see a generalization of this formula later in Chapter 8. ■

4.4 Bernoulli Equation

A first-order differential equation of the form

$$\boxed{y' = G(x)y + H(x)y^{\alpha}},$$

where x and y are real variables, and α is a real number, is called a **Bernoulli equation**. It can be converted to a linear equation by the following trick. We change the dependent variable y to u related by $y = u^{\beta}$ where β is a real number to be determined later. Our new equation is

$$\beta u^{\beta-1} u' = Gu^{\beta} + Hu^{\alpha\beta},$$

or equivalently, after division by $u^{\beta-1}$

$$\beta u' = Gu + Hu^{\alpha\beta-\beta+1}.$$

If we choose $\beta = \frac{1}{1-\alpha}$ then this latter equation becomes linear, hence solvable by method of Section 4.3.

Exercise 33. *Solve* $y' = y - y^3$ *using separability, and Bernoulli. Make sure that the answers of two methods coincide.*

4.5 Exact Equations

There is another general way, besides the notation $y' = F(x, y)$ and drawing slope fields, to represent and gain intuition about first-order ODEs, which we now discuss. Thinking of $\frac{dy}{dx}$ as the ratio of two infinitesimals, we rewrite our equation $\frac{dy}{dx} = F(x, y)$ into

$$\boxed{M(x, y)dx + N(x, y)dy = 0}, \tag{25}$$

where $M \equiv F(x, y)$ and $N \equiv -1$. Forgetting about F, sometimes, a first-order ODE is given directly by (25), where M and N are smooth functions of x and y. The technical advantage we gain by this representation is that the expression $M dx + N dy$ could be integrated along curves in the plane. It is called a *differential* 1-*form* in higher mathematics.

Equation (25) is called **exact** if $M_y = N_x$. If this happens then the Green's theorem [1, Vol II, 11.10] shows that, fixing some point (x_0, y_0) in the plane, the line integral

$$V(x, y) = \int_{(x_0,y_0)}^{(x,y)} M dx + N dy, \tag{26}$$

does not depend on the path connecting (x_0, y_0) to (x, y), and (therefore) satisfies $V_x = M$ and $V_y = N$. Equation (25), is then equivalent to

$$0 = V_x dx + V_y dy = dV,$$

hence solved as $V(x, y) = C$, with arbitrary constant C. This is an implicit equation between x and y, which in case of need, could be analyzed by implicit function theorem to give y locally in terms of x. Function V is called the **potential of the form** $M dx + N dy$. Potential could be found by computing (26) for a simple curve (say the straight line) connecting (x_0, y_0) to (x, y), or by the process explained in following example.[8]

Example 26. Consider the differential equation

$$\frac{dy}{dx} = -\frac{e^y}{xe^y + 2y}. \tag{27}$$

This is not separable, linear, homogeneous or Bernoulli; however if we rewrite it as

$$e^y dx + (xe^y + 2y) dy = 0,$$

[8]Here is the analogy to mechanics. If a force field $F = (M, N)$ satisfy $M_y = N_y$, then the work done by it on closed curves is zero. Such forces are called conservative, and can be more easily analyzed by their potential functions V defined through $F = \nabla V$. Then $V = C$ gives equipotential curves.

it turns out to be exact, because

$$\frac{\partial}{\partial y}(e^y) = e^y = \frac{\partial}{\partial x}(xe^y + 2y).$$

We want to find function $V(x, y)$ with

$$V_x = e^y, \quad V_y = xe^y + 2y.$$

Integrating the first equation with respect to x (so y is assumed fixed momentarily), gives

$$V = e^y x + f(y),$$

where $f(y)$ is a function of y. Putting this latter form of V into the second equation $V_y = xe^y + 2y$ gives

$$e^y x + f'(y) = xe^y + 2y,$$

hence $f'(y) = 2y$, so $f(y) = y^2 + C$, where C is a constant. The whole analysis shows that $e^y x + y^2 = C$ is the general solution to our problem. ■

Exercise 34. *Rewrite equation (27) as* $\frac{dx}{dy} = \cdots$*, and solve it without your knowledge of exact equations.*

Exercise 35 (Optional). *Find potential* V *in Example 26 via formula (26).*

4.6 $y' = F(ax + by + c)$

The first-order equation of the form

$$\frac{dy}{dx} = F(ax + by + c),$$

where F is a real-valued function of a real variable, and a, b and c are constants, can be reduced to a separable equation. If $b = 0$ then it is already separable. If $b \neq 0$, changing dependent variable y to $u = ax + by + c$, new equation is

$$\frac{d}{dx}\left(\frac{u - ax - c}{b}\right) = F(u),$$

or equivalently $\frac{du}{dx} = a + bF(u)$, which is separable.

Exercise 36. *Redo Exercise 21 using the method in this section.*

4.7 $y' = F\left(\frac{ax+by+c}{\alpha x+\beta y+\gamma}\right)$

Consider the first-order equation of the form

$$\frac{dy}{dx} = F\left(\frac{ax+by+c}{\alpha x+\beta y+\gamma}\right),$$

where F is a real-valued function of a real variable, and a, b, c, α, β and γ are constants. If $a\beta - b\alpha = 0$ then our equation is of the form $y' = G(ax+by+c)$ or $y' = G(\alpha x+\beta y+\gamma)$, which are treated in Section 4.6. If $a\beta - b\alpha \neq 0$, finding the unique solution (x_0, y_0) of the system

$$\begin{cases} ax+by+c = 0, \\ \alpha x+\beta y+\gamma = 0, \end{cases}$$

and accordingly changing independent variable from x to $X = x - x_0$, and dependent variable from y to $Y = y - y_0$, we get the equivalent equation

$$\frac{dY}{dX} = F\left(\frac{aX+bY}{\alpha X+\beta Y}\right),$$

which is homogeneous, and solvable by method of Section 4.2.

Example 27. Consider the differential equation

$$\frac{dy}{dx} = \frac{2x+3y-1}{4(x+1)}. \tag{28}$$

Two lines $2x+3y-1 = 0$ and $x+1 = 0$ intersect at $(-1, 1)$, hence we introduce new independent variable $X = x+1$, and dependent variable $Y = y-1$. Then $dy = dY$ and $dx = dX$, hence

$$\frac{dY}{dX} = \frac{2X+3Y}{4X},$$

which is homogeneous. Setting $Y = Xu$, we have

$$u + X\frac{du}{dX} = \frac{2+3u}{4},$$

or equivalently

$$\frac{du}{dX} = \frac{2-u}{4X},$$

which is separable, so solved as

$$\int \frac{4du}{2-u} = \int \frac{dX}{X} \Longrightarrow \log|2-u|^{-4} = \log|X| + C \Longrightarrow X(2-u)^4 = K.$$

Finally setting $X = x+1$ and $u = \frac{y-1}{x+1}$ gives

$$(2x-y+3)^4 = K(x+1)^3,$$

as our general solution. ■

4.8 When Dependent or Independent Variable Is Missing

A second-order differential equation has the general form

$$y'' = F(x, y, y'), \tag{29}$$

where F is a function of three variables x, y and y'. In special cases, where F misses variable y or x, namely is of the form $F(x, y')$ or $F(y, y')$, respectively, then the differential equation (29), could be reduced to a first-order equation. This is shown by the following two examples.

Example 28. Consider differential equation

$$x^2 y'' = (y')^2.$$

Introducing new dependent variable $u = y'$ instead of y, our equation becomes $x^2 u' = u^2$, which is first-order and separable, hence solved as

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

where K is a constant. ■

Example 29. Consider differential equation

$$y'' + (y')^2 = 2e^{-y}.$$

Again introducing new dependent variable $u = y'$, since

$$y'' = \frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = \frac{du}{dy}u,$$

our equation becomes

$$\frac{du}{dy}u + u^2 = 2e^{-y}.$$

which is a first-order and Bernoulli. ■

Exercise 37. *Solve the following equations: (a)* $y'' = 2yy'$. *(b)* $y'' = 2(y')^2 \tan y$.

Exercise 38. *In classical mechanics, the position* $x(t)$ *of a particle restricted to move on one direction* x*, under the influence of a conservative force* $F(x)$*, satisfies differential equation* $m\frac{d^2x}{dt^2} = F(x)$*. Assume that* $F = -kx$ *where* k *is a positive constant.*

(a) Solve the equation of motion.

(b) Show that $x(t)$ *is a periodic function, and find the smallest period.*

Exercise 39. *Show that if a small body is projected away from the earth in a direction perpendicular to the earths surface with an initial velocity of more than* 11.1 *kilometers per second, then it will never return back to earth. [Hint. Neglect air resistance, but take into account the variation of the earth's gravitational field with distance. Find earth's radius, and gravitational constant online.]*

4.9 Integrating Factor

Consider the first-order differential equation (25). It could be proved[9] that there is always a function $\mu(x, y)$ such that the equation

$$M\mu dx + N\mu dy = 0,$$

is exact, hence solvable by method of Section 4.5. Any such function μ is called an **integrating factor**, and finding it an specific problem under study is an art.

Example 30. Consider the differential equation

$$\frac{dy}{dx} = -\frac{2y + 3xy^2}{3x + 4x^2y}.$$

This is not separable, linear, homogeneous or Bernoulli. To check exactness we write it in infinitesimal form

$$\left(2y + 3xy^2\right) dx + \left(3x + 4x^2y\right) dy = 0. \tag{30}$$

It is not exact because in

$$\frac{\partial}{\partial y}\left(2y + 3xy^2\right) = 2 + 6xy \quad \text{and} \quad \frac{\partial}{\partial x}\left(3x + 4x^2y\right) = 3 + 8xy,$$

although both are of the form $a+bxy$, but the coefficients do not match. Some vague feeling might tell you that an integrating factor of the form $\mu = x^\alpha y^\beta$, with constants α and β to be determined later, might fix this matching problem. Let us check. Exactness of

$$x^\alpha y^\beta \left(2y + 3xy^2\right) dx + x^\alpha y^\beta \left(3x + 4x^2y\right) dy = 0$$

is equivalent to

$$\frac{\partial}{\partial y}\left(2x^\alpha y^{\beta+1} + 3x^{\alpha+1}y^{\beta+2}\right) = \frac{\partial}{\partial x}\left(3x^{\alpha+1}y^\beta + 4x^{\alpha+2}y^{\beta+1}\right),$$

which, after taking derivatives and canceling $x^\alpha y^\beta$, is equivalent to

$$2(1+\beta) + 3xy(2+\beta) = 3(1+\alpha) + 4xy(2+\alpha),$$

which is satisfied for $\alpha = 1$ and $\beta = 2$. We found integrating factor $\mu = xy^2$. Multiplying (30) by μ, makes it exact

$$\left(2xy^3 + 3x^2y^4\right) dx + \left(3x^2y^2 + 4x^3y^3\right) dy = 0, \tag{31}$$

[9] [3, Problem 2.46] or [24, vol. I, p. 83].

with no matching problem

$$\frac{\partial}{\partial y}\left(2xy^3+3x^2y^4\right)=6xy^2+12x^2y^3=\frac{\partial}{\partial x}\left(3x^2y^2+4x^3y^3\right).$$

Exact equation (31) can be solved by the method of Example 26 to give the general solution

$$x^2y^3+x^3y^4=C,$$

with constant C. ■

Exercise 40. *Solve equation*

$$(3x^2y+y^2)dx+(2x^3+3xy)dy=0,$$

knowing that it has an integrating factor of the form $\mu=P(y)$*, where* P *is a scalar-valued function of a scalar variable. [Answer:* $x^3y^2+xy^3=C$*.]*

Exercise 41. *Solve equation*

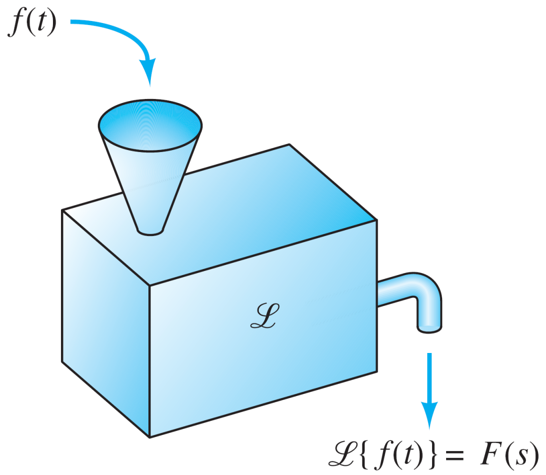
$$y\,dx+\left(x-2x^2y^3\right)dy=0,$$

knowing that it has an integrating factor of the form $\mu=P(xy)$*, where* P *is a scalar-valued function of a scalar variable. [Answer:* $x^{-1}y^{-1}+y^2=C$*.]*

5 Laplace Transform, Approximate Identities

Laplace tranform is a powerful tool in the analysis and design of linear systems specially those expressed by differential and integral equations [21, chapter 9]. We define and discuss basic properties of this transform in this chapter, and will see numerous applications in later ones.

Like differentiation or integration, Laplace transform $\mathcal{L}$ is a machine (operator, function, etc.) swallowing functions of *time* variable t as input, and spitting functions of *frequency* variable s as output, as visualized in the following figure.



There are different versions of Laplace transform, but we mostly use

$$\boxed{\mathcal{L}(f(t)) = \int_0^\infty f(t)e^{-st}dt}. \tag{32}$$

To feel comfortable with this definition, let us start computing the Laplace transform of some simple functions say t and $\cos(at)$, where a is a constant.

$$\begin{aligned}
\mathcal{L}(t) &= \int_0^\infty te^{-st}dt = \int_0^\infty td\left(-s^{-1}e^{-st}\right) \\
&= \left[-ts^{-1}e^{-st}\right]_0^\infty + \int_0^\infty s^{-1}e^{-st}dt && \text{integration by parts} \\
&= 0 + \left[-s^{-2}e^{-st}\right]_0^\infty = s^{-2}. && \text{assuming } \Re(s) > 0
\end{aligned}$$

Similarly, by doing integration by parts twice, we have

$$\mathcal{L}(\cos at) = \int_0^\infty e^{-st} d\left(a^{-1}\sin(at)\right) = \left[e^{-st}a^{-1}\sin(at)\right]_0^\infty + \int_0^\infty e^{-st} s a^{-1} \sin(at) dt =$$
$$= -sa^{-2}\int_0^\infty e^{-st} d(\cos(at)) = -sa^{-2}\left(\left[e^{-st}\cos(at)\right]_0^\infty + s\int_0^\infty e^{-st}\cos(at)dt\right) =$$
$$= -sa^{-2}\left(-1 + s\mathcal{L}(\cos at)\right),$$

assuming $\Re(s) > 0$. Therefore

$$\mathcal{L}(\cos at) = \frac{s}{s^2+a^2}, \quad \Re(s) > 0.$$

Exercise 42. *Prove that* $\mathcal{L}(\sin at) = \frac{a}{s^2+a^2}$ *assuming* $\Re(s) > 0$.

With similar computations we obtain Table 1, which lists the most encountered Laplace pairs, together with the range of s where the transform exists.

$f(t)$	$F(s)$	Region of Convergence
1	$\frac{1}{s}$	$\Re s > 0$
t^a, $a > -1$	$\frac{a!}{s^{a+1}}$	$\Re s > 0$
e^{at}	$\frac{1}{s-a}$	$\Re(s-a) > 0$
$\cos(at)$	$\frac{s}{s^2+a^2}$	$\Re s > 0$
$\sin(at)$	$\frac{a}{s^2+a^2}$	$\Re s > 0$
$\frac{1}{2a}t\sin(at)$	$\frac{s}{(s^2+a^2)^2}$	$\Re s > 0$
$\frac{1}{2a^3}(\sin(at) - at\cos(at))$	$\frac{1}{(s^2+a^2)^2}$	$\Re s > 0$
$\delta(t)$	1	$s \in \mathbb{C}$

Table 1: Most important Laplace pairs.

The last row of this table needs explanation. $\delta(t)$ is the **Dirac delta** (or **unit impulse**) **function**, characterized by two properties below:

$$\boxed{\int_{-\infty}^{\infty} \delta(t)dt = 0 \quad \text{and} \quad \delta(t) = 0 \text{ for } t \neq 0}.$$

No Riemann (or even Lebesgue) integrable function has these properties. Intuitively, it could be thought as the limit of rectangular pulses

$$\delta_\epsilon(t) = \begin{cases} \epsilon^{-1}, & 0 < t < \epsilon \\ 0, & t < 0 \text{ or } t > \epsilon \end{cases},$$

as $\epsilon \to 0+$. As a physical example, the linear charge density (measured in coulombs per meter) of a single point unit charge situated at the origin of x-axis is described by $\delta(x)$, because firstly there is no charge density outside origin, and secondly the total charge $\int_{-\infty}^{\infty} \delta(x)dx$ should equal one coulombs.

Alternatively $\delta(t)$ could be thought of as the *formal* derivative[10] of the **unit step function**

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases} .$$

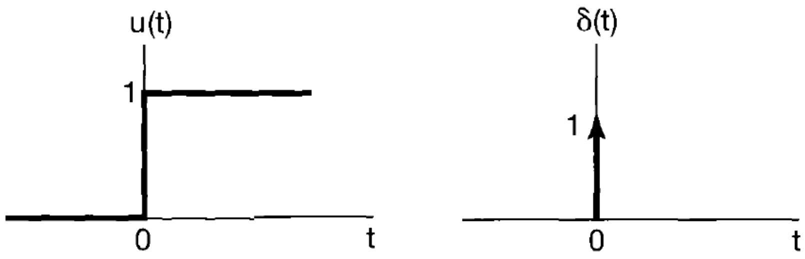


Figure 3: Unit step and unit impulse functions [21, chapter 1].

More generally, if a function $f(t)$ is differentiable everywhere except at finitely many discontinuity points $t_1, \dots, t_n$, and that at point t_k, $k = 1, \dots, n$, the functions jumps j_k units, then the summand

$$\sum_{1 \leqslant k \leqslant n} j_k \delta(t - t_k),$$

appears in $f'(t)$. For example, if $f(t) = |t|$, then

$$f'(t) = \begin{cases} -1, & x < 0 \\ 1, & x > 0 \end{cases}, \qquad f''(t) = 2\delta(t).$$

Another characteristic property of delta function is its **sifting** (or **sampling**) **property**:

$$\boxed{f(t)\delta(t-a) \equiv f(a)\delta(t-a)},$$

where a is a real constant, and f is an arbitrary smooth function. The integral version of sampling property is

$$\boxed{f(t) = \int_{-\infty}^{\infty} f(\tau)\delta(t-\tau)d\tau},$$

which expresses a function as linear combinations (superpositions) of impulses. Specially $\mathcal{L}(\delta(t)) = 1$ and $\mathcal{L}(\delta(t-a)) = e^{-as}$.

[10] Technical terminology is *distributional* or *weak* derivative.

Exercise 43 (Optional). *(a) Justify*

$$\delta(-t) = \delta(t), \quad \delta(2t) = \frac{1}{2}\delta(t), \quad t\delta(t) = 0.$$

(b) Justify

$$\delta\left(t^2 - 1\right) = \frac{1}{2}\delta(t+1) + \frac{1}{2}\delta(t-1).$$

(Hint: for part (b), use the definition $\delta(t) = \lim_{\epsilon \to 0+} \delta_\epsilon(t)$*, together with approximation* $\sqrt{1 \pm \epsilon} = 1 \pm \epsilon/2$*.)*

5.1 Basic Properties

To make Laplace transform into a powerful tool in solving differential equations, besides a table of Laplace transform of important functions (Table 1), we need to understand how Laplace transform interacts with summation, differentiation, integration etc. Table 2 gather general properties of Laplace transform.

For the Laplace transform to exist (namely the improper integral in the definition (32) to converge), satisfy the properties in Table 2, and to be one-to-one (namely $\mathcal{L}(f) = \mathcal{L}(g)$ imply $f = g$), we need to restrict attention to some special class of *admissible* functions. One admissible class usually used in elementary differential equations textbooks is the class of *piecewise continuous functions of exponential order*. Piecewise continuity means on each bounded interval the function has at most finitely many discontinuities, and all of them of jumped type (namely left and right limit exists.). Function $f(t)$ is called of exponential order if there exist real numbers a and M such that

$$|f(t)| \leqslant Me^{at},$$

for all t. This a is called an exponential order of f. For example every bounded function is of exponential order 0; however $f(t) = \exp\left(t^2\right)$ is not of exponential order. An immediate consequence of (5.1) is that $F(s) = \mathcal{L}(f(t))$ exists for $\Re s > a$, and satisfies the estimation

$$|F(s)| \leqslant \frac{M}{\Re s - a},$$

for $\Re s > a$. Specially $F(s) \to 0$ for $\Re s \to \infty$. Most functions appearing in applied science and engineering are admissible in this sense together with all their derivatives, so we could freely use Table 2, and assume that $\mathcal{L}$ is one-to-one. These technical issues are addressed in optional Section 5.4, or you could refer to [13, chapter 7].

We start studying Table 2.

$f(t)$ $g(t)$	$F(s)$ $G(s)$	
$af(t)+bg(t)$ $f(at),\ a>0$	$aF(s)+bG(s)$ $\frac{1}{a}F\left(\frac{s}{a}\right)$	linearity time scale $\leftrightarrow$ frequency scale
$f'(t)$ $f''(t)$ $f^{(n)}(t)$ $(-t)^n f(t)$	$sF(s)-f(0)$ $s^2F(s)-sf(0)-f'(0)$ $s^nF(s)-s^{n-1}f(0)-\cdots-f^{(n-1)}(0)$ $F^{(n)}(s)$	$\frac{d}{dt}\leftrightarrow s$ $t\leftrightarrow\frac{d}{ds}$
$\int_0^t f(\tau)d\tau$ $\frac{f(t)}{t}$	$\frac{F(s)}{s}$ $\int_s^\infty F(\sigma)d\sigma$	$\int\leftrightarrow\frac{1}{s}$ $\frac{1}{t}\leftrightarrow\int$
$f(t)e^{at}$ $f(t-a)u(t-a),\ a>0$	$F(s-a)$ $e^{-as}F(s)$	modulation $\leftrightarrow$ shift shift $\leftrightarrow$ modulation
$\int_0^t f(\tau)g(t-\tau)d\tau$	$F(s)G(s)$	convolution $\leftrightarrow$ multiplication
$f(t)$ with period T	$\frac{1}{1-e^{-Ts}}\int_0^T e^{-st}f(t)dt$	
If $f(\infty)$ exists then $f(\infty)=\lim_{s\to 0} sF(s)$.		Final Value Theorem

Table 2: Basic properties of the Laplace transform.

1. The verification of the following properties are straightforward, and left as exercise for the reader.

$$\mathcal{L}(af+bg)=a\mathcal{L}(f)+b\mathcal{L}(g), \quad \mathcal{L}\left(f(at)\right)=\frac{1}{a}F\left(\frac{s}{a}\right),$$

$$\mathcal{L}\left(f(t)e^{at}\right)=F(s-a), \quad \mathcal{L}\left(f(t-a)u(t-a)\right)=e^{-as}F(s).$$

2. A magic of Laplace transform is that it makes the analytic operations of differentiation and integration, respectively, into the simpler algebraic operations of multiplication and division:

$$\boxed{\mathcal{L}(f'(t))=sF(s)-f(0)}, \quad \mathcal{L}\left(\int_0^t f(\tau)d\tau\right)=\frac{F(s)}{s}. \tag{33}$$

Let us prove the first property.

$$\begin{aligned}\mathcal{L}(f'(t)) &= \int_0^\infty e^{-st}f'(t)dt=\int_0^\infty e^{-st}d\left(f(t)\right)\\ &= \left[e^{-st}f(t)\right]_0^\infty+\int_0^\infty se^{-st}f(t)dt, && \text{integration by parts}\\ &= -f(0)+sF(s). && f \text{ is of exponential order}\end{aligned}$$

This proves the first identity. Replacing $f(t)$ by $\int_0^t f(\tau)d\tau$ in this identity proves the second identity. On the other hand sending $s \to 0$ in the identity we just proved, we obtain

$$-f(0) + \lim_{s\to 0} sF(s) = \lim_{s\to 0}\int_0^\infty e^{-st}f'(t)dt = \int_0^\infty f'(t)dt = f(\infty) - f(0),$$

hence

$$\boxed{f(\infty) = \lim_{s\to 0} sF(s)},$$

provided that $f(\infty)$ exists. This useful result, every engineer should have in mind, is called the **Final Value Theorem**.

3. Also note that

$$F^{(n)}(s) = \frac{d^n}{ds^n}\int_0^\infty f(t)e^{-st}dt = \int_0^\infty f(t)\frac{d^n}{ds^n}\left(e^{-st}\right)dt = \int_0^\infty f(t)(-t)^n e^{-st}dt = \mathcal{L}\left((-t)^n f(t)\right).$$

4. To prove

$$\mathcal{L}\left(\frac{f(t)}{t}\right) = \int_s^\infty F(\sigma)d\sigma,$$

setting

$$g(t) := \frac{f(t)}{t}, \quad G(s) := \mathcal{L}(g(t)),$$

using the property just proved, we have

$$F(s) = \mathcal{L}(f(t)) = \mathcal{L}(tg(t)) = -\frac{dG(s)}{ds},$$

hence after integration

$$G(s) = -\int_{s_0}^s F(\sigma)d\sigma + C,$$

where s_0 is an arbitrary real number, and $C = G(s_0)$. Since f is assumed to be of exponential order, $G(\infty) = 0$, and after choosing $s_0 = \infty$, we have

$$G(s) = -\int_{s_0}^s F(\sigma)d\sigma + \int_{s_0}^\infty F(\sigma)d\sigma = \int_s^\infty F(\sigma)d\sigma,$$

which is what we wanted.

5. For two functions $f(t)$ and $g(t)$ defined on $t \geqslant 0$, their **convolution**, is defined by

$$\boxed{(f * g)(t) = \int_0^t f(\tau)g(t-\tau)d\tau}.$$

A black magic of Laplace transform is that it makes convolution into multiplication:

$$\boxed{\mathcal{L}(f * g) = (\mathcal{L}f)(\mathcal{L}g)}.$$

Its proof is a simple application of Fubini's theorem [23, p. 399].

Exercise 44. *Compute the following convolutions in two ways: definition, and the convolution property.*

$$t^2 * t^3, \quad e^{at} * \cos bt, \quad \cos at * \sin bt.$$

Exercise 45. *(a) Find the Laplace inverses of* $\frac{s}{(s^2+a^2)^2}$ *and* $\frac{1}{(s^2+a^2)^2}$*, where* a *is a constant. (b) Find the Laplace inverses of* $\frac{1}{(s^2+1)^3}$ *and* $\frac{s}{(s^2+1)^3}$*. [Hint: use convolution property.]*

Exercise 46. *Show that the second identity in (33) is a special case of convolution property.*

Example 31. (a) Consider rectangular pulses $A(u(t-a)-u(t-b))$ and $A'(u(t-a')-u(t-b'))$, where we are assuming $D := b - a \geqslant D' := b' - a'$. Show that their convolution is the triangular pulse of Figure 4.

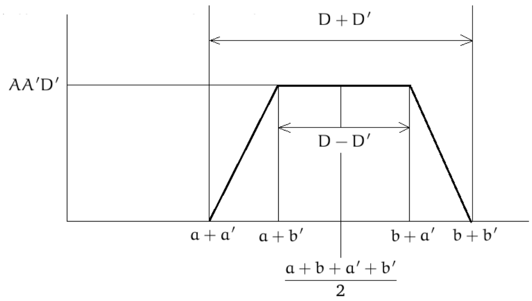


Figure 4: Convolution of two rectangular pulses.

■

Exercise 47. *(a) Find a function which violates the assertion of the Final Value Theorem.*

(b) Suppose $f(t)$ *is an admissible function containing no impulses or its derivatives at* $t = 0$*. Justify the* ***Initial Value Theorem****:*

$$f(0+) = \lim_{s\to\infty} sF(s).$$

Exercise 48. *Compute: (a)* $\int_0^\infty \frac{\sin x}{x}dx$*. (b)* $\int_0^\infty \frac{e^{-x}-e^{-2x}}{x}dx$*.*

Exercise 49 (Optional)**.** *Find a closed formula for the function*

$$f(x) := \int_0^\infty \frac{\cos(xu)}{1+u^2}du,$$

defined for real variable $x \geqslant 0$, following either of the following suggestions.

(a) Apply Laplace transform.

(b) Justify $f''(x) = f(x)$. This gives $f(x) = C_1 e^x + C_2 e^{-x}$. Show that $C_1 = 0$ and $C_2 = \frac{\pi}{2}$.

Exercise 50 (Optional)**.** *Consider the unit rectangular pulse $f(t) = u(t) - u(t-1)$. For each natural number n, let $f_n(t)$ be the function created by convolving f with itself n times.*

(a) Prove that

$$f_n(t) = \sum_{0 \leqslant k \leqslant n} (-1)^k \binom{n}{k} \frac{(t-k)^{n-1}}{(n-1)!} u(t-k).$$

(Hint: work on s domain, and use convolution property.)

(b) Prove that $f_n(t) = 0$ for $t > n$. (Hint: work on t domain, and apply induction.)

(c) For several small n, draw $f_n(t)$, and compare it with

$$g_n(t) = \frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{(t-\mu_n)^2}{2\sigma_n^2}},$$

where $\mu_n = \frac{n}{2}$ and $\sigma_n^2 = \frac{n}{12}$.

5.2 Approximate Identities, Heat and Poisson Kernels (Optional)

In the definition of the unit impulse $\delta(t)$ in (32), observe that $\delta_\epsilon(t)$ is a 1-parameter family of functions all with unit area, which concentrate more and more around origin $t = 0$ as the parameter ϵ approaches zero. Any other family with these two properties is called an **approximate identity**, and could be used to define $\delta(t)$ via formula (32). Here is a physical explanation behind this amazing fact that $\delta(t)$ does not depend on details of the family $\delta_\epsilon(t)$ except the family being an approximate identity [21, p. 36]:

> "... any real physical system has some inertia associated with it and thus does not respond instantaneously to inputs. Consequently, if a pulse of sufficiently short duration is applied to such a system, the system response will not be noticeably influenced by the pulse's duration or by the details of the shape of the pulse, for that matter. Instead, the primary characteristic of the pulse that will matter is the net, integrated effect of the pulse—i.e., its area.[11] For systems that respond much more quickly than others, the pulse will have to be of much shorter duration before the details of the pulse shape or its duration no longer matter. Nevertheless, for any physical system, we can always find a pulse that is "short enough." The unit impulse then is an idealization of this concept—the pulse that is short enough for *any* system."

[11] Here, by the area of a pulse $p(t)$, the authors mean $\int_{\infty}^{\infty} p(t)dt$, not $\int_{\infty}^{\infty} |p(t)|dt$.

For example each of the following families is an approximate identity:

$$\begin{cases} \frac{1}{2\epsilon}, & |t| < \epsilon \\ 0, & |t| > \epsilon \end{cases}, \qquad \begin{cases} \frac{1}{\epsilon}, & |t| < \epsilon \\ -\frac{1}{2\epsilon}, & \epsilon < |t| < 2\epsilon \\ 0, & |t| > 2\epsilon \end{cases}, \qquad \begin{cases} \frac{1}{\epsilon} - \frac{|t|}{\epsilon^2}, & |t| < \epsilon \\ 0, & |t| > \epsilon \end{cases},$$

$$\frac{\sin\left(\epsilon^{-1}t\right)}{\pi t}, \qquad \frac{1}{\epsilon\sqrt{2\pi}}e^{-\frac{t^2}{2\epsilon^2}}, \qquad \frac{1}{\pi}\frac{\epsilon}{t^2+\epsilon^2}.$$

It really helps to visualize these families That the area of each of the members in the last three families is 1 can be deduced by substitution in the following definite integrals:

$$1 = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin x}{x}dx = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2}dx = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{x^2+1}dx.$$

All these examples are special cases of a general recipe: *if $\varphi(x)$ is a function with unit area, then $\varphi_\epsilon(x) := \frac{1}{\epsilon}\varphi\left(\frac{x}{\epsilon}\right)$ is an approximate identity.*

Exercise 51. *Use formula $\delta(t) = \lim_{\omega\to\infty} \frac{\sin(\omega t)}{\pi t}$ to justify*

$$\boxed{\int_{-\infty}^{\infty} e^{it\omega}d\omega = 2\pi\delta(t)}.$$

This is an important identity in Fourier analysis.

Exercise 52. *Use formula $\delta(\xi) = \lim_{t\to 0+} \frac{1}{\sqrt{4\pi}}t^{-\frac{1}{2}}e^{-\frac{\xi^2}{4t}}$ to justify that*

$$\boxed{u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi}}t^{-\frac{1}{2}}e^{-\frac{(x-\xi)^2}{4t}}f(\xi)d\xi}, \tag{34}$$

solves the heat propagation in a long rod:

$$\begin{cases} \text{PDE}: & u_t = u_{xx}, \quad 0 < t < \infty, \ -\infty < x < \infty \\ \text{IC}: & u(x,0) = f(x), \quad -\infty < x < \infty \end{cases}.$$

*The function $p_t(x) = \frac{1}{\sqrt{4\pi}}t^{-\frac{1}{2}}e^{-\frac{x^2}{4t}}$ is called **heat kernel**.*

Exercise 53. *Use formula $\delta(\xi) = \lim_{y\to 0+} \frac{1}{\pi}\frac{\epsilon}{\xi^2+\epsilon^2}$ to justify that*

$$\boxed{u(x,y) = \int_{-\infty}^{\infty} \frac{1}{\pi}\frac{y}{(x-\xi)^2+y^2}f(\xi)d\xi},$$

solves the Dirichlet problem on upper-half plane:

$$\begin{cases} \text{PDE}: & u_{xx} + u_{yy} = 0, \quad -\infty < x < \infty, \quad 0 < y < \infty \\ \text{BC}: & u(x,0) = f(x), \quad -\infty < x < \infty \end{cases}.$$

*The function $p_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}$ is called **Poisson kernel**.*

5.3 Tautochrone Problem

As an application of the convolution property of Laplace transform, we solve the tautochrone problem introduces in Example 21. Recall that our desired curve in uv-plane was the one passing through origin, and satisfying

$$\int_0^y \frac{f(v)}{\sqrt{y-v}} dv = C, \quad \text{for any } y,$$

where C is a constant and $f(v) = \sqrt{1 + \left(\frac{du}{dv}\right)^2}$. The left hand side of this equation is a convolution, so applying Laplace transform gives $F(s)s^{-\frac{1}{2}} = Cs^{-1}$, with refreshed constant C. This shows that

$$\sqrt{1 + \left(\frac{dx}{dy}\right)^2} = f(y) = \frac{C}{\sqrt{y}},$$

with refreshed constant. This is a separable equation, hence solved as

$$x = \int \sqrt{\frac{C^2 - y}{y}} dy = \int 2C^2 \cos^2 \theta d\theta = C^2 \left(\theta + \frac{\sin 2\theta}{2}\right) + A,$$

where we have set $y = C^2 \sin^2 \theta$. Constant A vanishes because the curves passes through origin. Setting $2\theta = \varphi$, and $C^2/2 = B$, we get the parametric description of our desired curve:

$$x = B\left(\varphi + \sin \varphi\right), \quad u = B(1 - \cos \varphi).$$

This is a cycloid! Refer https://en.wikipedia.org/wiki/Tautochrone_curve for an animation of this problem.

5.4 Existence, Uniqueness and Inversion Formulas for Laplace Transform (Optional)

Let $f(t)$ be a real-valued function defined for $t \geqslant 0$. f is called **piecewise continuous** if on each bounded interval it has at most finitely many discontinuities, and all of them are of jumped type (namely left and right limit exists.) f is called **of exponential order** if there exists real number a such that $|f(t)| \lesssim e^{at}$ for all t. This a is called an exponential order of f. f is called **admissible** if it is both piecewise continuous and of exponential order. An admissible function has Laplace transform for $\Re s$ large enough s, because the first property make the function locally integrable, and the second one makes its Laplace transform improper integral convergent.

We gather the fundamental statements about existence and uniqueness of Laplace transform in the following theorem.

Theorem 7. *Let $f(t)$ and $g(t)$ be real-valued functions defined on $t \geqslant 0$.*

(a) If $F(s_0)$ exists then $F(s)$, $\Re s > \Re s_0$, also exists. [12, p. 16]

(b) If f is admissible, then $|F(s)| \lesssim 1/(\Re s - a)$ for $\Re s > a$, where a is an exponential order of f. Specially $F(s) \to 0$ as $\Re s \to \infty$. Furthermore, $F(s)$ is analytic on $\Re s > a$. [7, p. 188]

(c) If f and g are admissible with equal Laplace transforms on some half-plane $\Re s > a$ (or even on some real half-line $s > a$), then $f(t) = g(t)$ are equal except possibly at their points of discontinuity. Even at discontinuity point $t = t_0$, we have that

$$\frac{f(t_0-) + f(t_0+)}{2} = \frac{g(t_0-) + g(t_0+)}{2}.$$

[7, p. 201]

(d) If $F(s) = G(s)$ on some $\Re s > a$, then $f(t) = g(t)$ almost everywhere. [12, p. 21]

(e) Let

$$\int_0^\infty e^{-C_0 t}|f(t)|dt < \infty,$$

for real C_0. If f is of bounded variation on some neighborhood of $t > 0$, then

$$\frac{f(t+) + f(t-)}{2} = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{C-iT}^{C+iT} e^{st}F(s)ds,$$

for each real $C > C_0$. If f is of bounded variation on some neighborhood to the right of $t = 0$, then

$$\frac{f(t+)}{2} = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{C-iT}^{C+iT} e^{st}F(s)ds,$$

for each real $C > C_0$. This is called a ***complex inversion formula for Laplace transform****. [12, p. 157]*

(f) If $f(t)$ is continuous, and of exponential growth

$$f(t) = \lim_{n\to\infty} \frac{(-1)^n}{n!} \left(\frac{n}{t}\right)^{n+1} F^{(n)}\left(\frac{n}{t}\right), \quad t > 0.$$

This is called ***Post's inversion formula for Laplace transform****.*

6 Linear Ordinary Differential Equations

The general form of a linear ordinary differential equation of order n is

$$\boxed{\frac{d^n y}{dx^n} + P_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + P_1(x)\frac{dy}{dx} + P_0(x)y = R(x)}, \tag{35}$$

where all $P_i(x)$ and $R(x)$ are real-valued functions of real-variable x ranging on some open (connected) interval I. Equation (35) is **homogeneous** if $R \equiv 0$. The **corresponding homogeneous equation** to (35) is

$$\frac{d^n y}{dx^n} + P_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + P_1(x)\frac{dy}{dx} + P_0(x)y = 0. \tag{36}$$

Many general features of these equations already reveals itself in the special $n = 2$ case, namely an equation of the form

$$\boxed{y'' + P(x)y' + Q(x)y = R(x)}, \tag{37}$$

with corresponding homogeneous equation

$$y'' + P(x)y' + Q(x)y = 0, \tag{38}$$

and just because of notational simplicity, it is on these two latter equations that we mostly concentrate in this chapter.

6.1 The Existence and Uniqueness of Solutions

Here is the fundamental existence and uniqueness result for linear equations.

Theorem 8. *In equation (35), let all P_i and R be continuous scalar-valued functions of real variable x ranging on open interval I. For any $x_0 \in I$, and any scalars A_i, $i = 0, \ldots, n-1$, there is a unique function $y(x)$ satisfying the initial value problem*

$$y^{(n)} + P_{n-1}(x)y^{(n-1)} + \cdots + P_1(x)y' + P_0(x)y = R(x), \quad y^{(i)}(x_0) = A_i, \quad i = 0, \ldots, n-1,$$

on whole $x \in I$.

This is proved in [23, chapter 13], [4, chapter 6], [1, vol II, Theorem 6.3].

Example 32. Recall our discussion about equation

$$\frac{d^2 y}{dx^2} = y, \tag{39}$$

in Section 2.1. There we proved that $y = C_1 e^x + C_2 e^{-x}$, with C_1 and C_2 constants, is the general solution. Here is another argument for this using the *uniqueness* part of Theorem 8. Let u be an arbitrary solution of (39), and set

$$A := u(0), \quad B := u'(0), \quad C_1 := \frac{A+B}{2}, \quad C_2 := \frac{A-B}{2}, \quad v(x) = C_1 e^x + C_2 e^{-x}.$$

Now the function u and v are equal, because both satisfy the initial value problem

$$y'' + y = 0, \quad y(0) = A, \quad y'(0) = B.$$

■

6.2 Structure Theorem

Example 32 revealed some power of Theorem 8. Arguing in the same lines we will obtain the following theorem which completely determines the structure of solutions of (37).

Theorem 9. *Let y_1 and y_2 be two solutions of equation (38).*

(a) If y_1 and y_2 are ***(linearly) independent***[12]*, then*

$$\boxed{y = C_1 y_1 + C_2 y_2},$$

with contants C_1 and C_2, gives the general solution of (38). The collection $\{y_1, y_2\}$ is called a ***basis*** *or* ***fundamental system*** *for equation (38).*

(b) y_1 and y_2 are independent if and only if their ***Wronkian***

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix},$$

is nowhere-zero, if and only if $W(y_1, y_2)$ is nonzero in at least one point.

(c) If y_1 and y_2 are independent, and u is a (particular) solution of (37), then

$$\boxed{y = C_1 y_1 + C_2 y_2 + u},$$

is the general solution of (37).

Proof. (b) Since

$$W' = y_1 y_2'' - y_1'' y_2 = y_1(-Py_2' - Qy_2) - (-Py_1' - Qy_1)y_2 = -PW,$$

[12] Namely there exists no constants C_1 and C_2, at least one nonzero, such that $C_1 y_1 + C_2 y_2 \equiv 0$.

we have $W = Ce^{-\int P dx}$. This proves that W either vanishes nowhere or vanishes everywhere. Also note that under any of the dependency conditions $y_2 = ky_1$ or $y_1 = ky_2$, with k a constant, Wronskian vanishes. Conversely, assume that Wronskian vanishes at some point $x = x_0$, and we need to show that y_1 and y_2 are dependent. Putting aside the trivial case $y_1 \equiv 0$, we could assume $y_1(x_0) \neq 0$ for some x_0. Since the function $u := y_2(x_0)y_1(x) - y_1(x_0)y_2(x)$ and $v \equiv 0$ both satisfy the initial value problem

$$y'' + Py' + Qy = 0, \quad y(x_0) = y'(x_0) = 0,$$

by uniqueness part of Theorem 8, $u \equiv 0$, and this witnesses the dependency of y_1 and y_2.

(a) Let y be an arbitrary solution of (38). By the previous part, one can find x_0 such that $W(y_1, y_2)(x_0) \neq 0$. Thus the linear system

$$\begin{bmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} y(x_0) \\ y'(x_0) \end{bmatrix},$$

has solution for C_1 and C_2. Now $C_1y_1 + C_2y_2$ and y both satisfy one common initial value problem, hence equal.

(c) Trivial. ■

Remark 1 (Optional). If you are familiar with *bump functions*, you could easily construct two linearly independent smooth functions y_1 and y_2 on the whole real line whose Wronskian vanishes everywhere. Another example of this phenomena is the two functions $y_1 = x^3$ and $y_2 = x^2|x|$. This shows that the assertion of Theorem 9.b may not hold for arbitrary smooth functions, but only for those functions satisfying a common linear differential equation. ■

Fundamental Theorems 8 and 9 has corresponding analogues for n-th order equation (35). In particular, if $y_1, \dots y_n$ are solutions to homogeneous equation (36), then they are (linearly) independent[13] if and only if their Wronkian

$$W(y_1, , \dots, y_n) = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix},$$

is nonzero at a point (or equivalently, nowhere zero).

[13] Namely there exists no constants $C_1, \dots, C_n$, at least one nonzero, such that $C_1y_1 + \cdots + C_ny_n \equiv 0$.

6.3 Finding a Fundamental System of Solutions for a Second-Order Homogeneous Linear Equation Having One Solution at Hand

If one knows a nowhere-zero solution y_1 of the homogeneous equation (38), then one could find another solution, independent of y_1, by the following trick due to Liouville. We try to find function $u(x)$ such that $y = y_1 u$ also satisfies equation (38), namely

$$0 = Q\,(y_1 u) + P(y_1' u + y_1 u') + (y_1'' u + 2y_1' u' + y_1 u'') =$$
$$= (Qy_1 + Py_1' + y_1'')u + Py_1 u' + 2y_1' u' + y_1 u'' = (Py_1 + 2y_1')u' + y_1 u''.$$

Therefore we should have
$$\frac{du'}{u'} = -\frac{Py_1 + 2y_1'}{y_1} dx,$$
which after integration becomes
$$\log u' = -\int P dx - 2\log y_1,$$
hence
$$u = \int y_1^{-2} e^{-\int P dx}.$$

Retracing back, our analysis shows that
$$y_2 = y_1 \int y_1^{-2} e^{-\int P dx} dx,$$
is another solution to equation (38). To check independency of y_1 and y_2, we compute their Wronskian
$$W(y_1, y_2) = y_1(y_1' u + y_1 u') - y_1' y_1 u = y_1^2 u' = e^{-\int P dx},$$
which never vanishes. We gather what we have proved in the following proposition.

Proposition 4. *If y_1 is a nowhere-zero solution of the homogeneous equation (38), then another independent solution is given by*
$$\boxed{y_2 = y_1 \int y_1^{-2} e^{-\int P dx} dx}.$$

Exercise 54. *(a) Find the general solution of $xy'' - (x+2)y' + 2y = 0$ knowing that $y = e^x$ is one particular solution. (b) Find the general solution of $(1-x^2)y'' - 2xy' + 2y = 0$ knowing that $y = x$ is one particular solution.*

6.4 Constant Coefficient Equations

Theorem 9 tells us to find the general solution of a homogeneous linear n-th order ODE, we need to find a fundamental system of solutions, namely a collection of n independent solutions. One of the rare cases where we can do this explicitly is the constant-coefficient equation, namely an equation of the form

$$\boxed{y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0}, \tag{40}$$

where $a_0, \ldots, a_{n-1}$ are constants.

Let us start by solving

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 6y = 0.$$

Based on examples in Section 2.1, let us see if exponential function $y = e^{rx}$, where r is a constant, satisfies our equation. The computation

$$\left(\frac{d^2}{dx^2} + \frac{d}{dx} - 6\right) e^{rx} = (r^2 + r - 6)e^{rx} = (r-2)(r+3)e^{rx},$$

shows that e^{2x} and e^{-3x} satisfy our differential equation. The Wronskian

$$\begin{vmatrix} e^{2x} & e^{-3x} \\ 2e^{2x} & -3e^{-3x} \end{vmatrix} = -5e^{-x},$$

is nonzero, hence the collection $\{e^{2x}, e^{-3x}\}$ is a fundamental system, so $y = C_1e^{3x} + C_2xe^{3x}$ is the general solution.

Let us now work on

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 0.$$

Again we test exponential functions of the form $y = e^{rx}$, where r is a constant. The computation

$$\left(\frac{d^2}{dx^2} + 6\frac{d}{dx} + 9\right) e^{rx} = (r^2 + 6r + 9)e^{rx} = (r+3)^2e^{rx},$$

shows that e^{-3x} satisfy our differential equation. But to find the general solution we another solution independent of e^{3x}. How to find that? Proposition 4 gives the other one as

$$e^{3x}\int e^{6x}e^{-6x}\,dx = e^{3x}x.$$

The Wronskian

$$\begin{vmatrix} e^{-3x} & xe^{-3x} \\ -3e^{-3x} & (1-3x)e^{-3x} \end{vmatrix} = e^{-6x},$$

is nonzero, hence the collection $\{e^{-3x}, xe^{-3x}\}$ is a fundamental system, so $y = C_1e^{-3x} + C_2xe^{-3x}$ is the general solution.

Proposition 5. *In constant-coefficient equation (40), assume that the* ***characteristic polynomial***

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

factors as

$$p(r) = \prod_{k=1}^{N} (r - r_k)^{m_k}.$$

Then

$$\boxed{\{e^{r_k x}, e^{r_k x}x, \dots, e^{r_k x}x^{m_k - 1} : k = 1, \dots, N\}},$$

is a fundamental system of solutions.

Remark 2. Here comes an alternative clever *perturbative* argument to discover xe^{rx}, when r is a characteristic root of multiplicity 2, of (40). Let us think of multiple root r and r as limit degenerate case of two simple roots r and $r + h$ where h is very small. Since both e^{rx} and

$$e^{(r+h)x} = e^{rx}e^{hx} \approx e^{rx}(1 + hx),$$

are solutions of ODE, so is their difference divided by scalar h, which is $e^{rx}x$ ■

Example 33. For equation

$$y^{(5)} - 2y^{(4)} + y^{(3)} = 0,$$

since the characteristic polynomial is

$$r^5 - 2r^4 + r^3 = r^3(r - 1)^2,$$

the collection

$$\{1, x, x^3, e^x, xe^x\},$$

is a fundamental system. ■

Example 34. For equation

$$y'' + y = 0,$$

since the characteristic polynomial is

$$r^2 + 1 = (r - i)(r + i),$$

we get fundamental solutions

$$e^{ix} = \cos x + i \sin x, \quad e^{-ix} = \cos x - i \sin x,$$

so the general solution is

$$y = C_1(\cos x + i\sin x) + C_2(\cos x - i\sin x).$$

Everything is OK, but since our differential equation was given in real coefficients, it might be more suitable to express the solution in terms of real-valued functions. That is easy if we notice that, because of linearity of the equation

$$\frac{1}{2}\left(e^{ix} + e^{-ix}\right) = \cos x, \quad \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) = \sin x,$$

are also solutions, and they independent because their Wronskian

$$\begin{vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{vmatrix} = 1,$$

is nonzero. Therefore

$$y = C_1 \cos x + C_2 \sin x,$$

is also the general solution. ■

Exercise 55. *Find the general solution of* $y^{(5)} + y = 0$.

Exercise 56. *Consider the differential equation*

$$y'' + ay' + by = 0,$$

where a *and* b *are real constants. Prove that all solutions* $y(x)$ *of this equation approaches* 0 *as* $x \to \infty$ *if and only if both* a *and* b *are positive.*

Exercise 57. *(a) Find a constant coefficient linear ODE of least order such that* $y = x^3 - 2x^2 + 1$ *is one of its solutions.*

(b) Find a constant coefficient linear ODE of least order such that $y = x\sin x$ *is one of its solutions.*

(c) Find a constant coefficient linear ODE of least order such that $y = (x^2+1)\sin x$ *is one of its solutions.*

(d) Can you find a first-order linear homogeneous ODE of the form $y' + p(x)y = 0$ *with continuous* p *such that* $y = x\sin x$ *is one of its solutions on interval* $(-1, 1)$*?*

(e) Can you find a second-order linear homogeneous ODE of the form $y'' + p(x)y' + q(x)y = 0$ *with continuous* p *and* q *such that* $y = x\sin x$ *is one of its solutions on interval* $(-1, 1)$*?*

(f) Can you find a third-order linear homogeneous ODE of the form $y''' + p(x)y'' + q(x)y' + r(x)y = 0$ *with continuous* p*,* q *and* r *such that* $y = x\sin x$ *is one of its solutions on interval* $(-1, 1)$*?*

[Answer and Hint. (d, e) No! Use uniqueness part of Theorem 8. (f) Yes! Find it by inspection.]

6.5 The Method of Variation of Parameters

Here comes a beautiful idea of Lagrange to find a particular solution of inhomogeneous equation (37) having the general solution

$$y = C_1 y_1 + C_2 y_2,$$

of the corresponding homogeneous equation (38). We try to find function $u_1(x)$ and $u_2(x)$ such that

$$y = u_1 y_1 + u_2 y_2,$$

satisfies (37). This means

$$\begin{aligned} R = Q(u_1y_1 + u_2y_2) + P(u_1y_1' + u_2y_2') + (u_1y_1'' + u_2y_2'') \\ + P(u_1'y_1 + u_2'y_2) + 2(u_1'y_1' + u_2'y_2') + (u_1''y_1 + u_2''y_2) \\ = P(u_1'y_1 + u_2'y_2) + (u_1'y_1' + u_2'y_2') + (u_1'y_1 + u_2'y_2)', \end{aligned}$$

which is true if

$$\begin{cases} u_1'y_1 + u_2'y_2 = 0 \\ u_1'y_1' + u_2'y_2' = R \end{cases},$$

which has solution

$$\begin{cases} u_1' = \frac{-Ry_2}{y_1y_2' - y_1'y_2} \\ u_2' = \frac{Ry_1}{y_1y_2' - y_1'y_2} \end{cases}.$$

We have proved the first part of the following theorem; the second parts is proved similarly.

Proposition 6. *(a) If $\{y_1, y_2\}$ is a fundamental system for solutions for the homogeneous equation (38), then a particular solution to the inhomogeneous equation (37) is given by*

$$\boxed{y_p = y_1 \int \frac{-y_2}{W} R dx + y_2 \int \frac{y_1}{W} R dx},$$

where W is the Wronskian of y_1 and y_2.

(b) If $\{y_1, \ldots, y_n\}$ is a fundamental system for solutions for the homogeneous equation (35), then a particular solution to the inhomogeneous equation (36) is given by $y_p = u_1y_1 + \cdots + u_ny_n$ where $u_1, \ldots, u_n$ are solutions of the linear equation

$$\begin{bmatrix} y_1 & \cdots & y_n \\ y_1' & \cdots & y_n' \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix} \begin{bmatrix} u_1' \\ \vdots \\ u_n' \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ R \end{bmatrix}.$$

Example 35. Let us use this method to find a particular solution of the differential equation

$$y'' - 2y' + y = \frac{e^x}{1+x^2}.$$

The corresponding homogeneous differential equaltion is $y'' - 2y' + y = 0$, with charactersitic polynomial $r^2 - 2r + 1 = (r-1)^2$, which has root $r = 1$ with multiplicity 2. Thus $y_1 = e^x$ and $y_2 = xe^x$ are two independent solutions of the homogeneous equation. Their Wronskian is

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} e^x & xe^x \\ e^x & e^x(x+1) \end{vmatrix} = e^{2x}(x+1) - e^{2x}x = e^{2x}.$$

A particular solution to the inhomogeneous equation is

$$y = e^x \left(\int \frac{-xe^x}{e^{2x}} \frac{e^x}{1+x^2}\, dx \right) + xe^x \left(\int \frac{e^x}{e^{2x}} \frac{e^x}{1+x^2}\, dx \right) = e^x \left(\int \frac{-x}{1+x^2}\, dx \right) + xe^x \left(\int \frac{1}{1+x^2}\, dx \right)$$
$$= -\frac{1}{2} e^x \log\left(1+x^2\right) + xe^x \tan^{-1} x.$$

In Examples 36 and 39, we solve this equation with two other methods. ■

6.6 The Operator Method of Heaviside

We give an alternative quick way to solve constant coefficient linear equation

$$\boxed{y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_0 y = R(x)}, \tag{41}$$

where $R(x)$ is a scalar-valued function of real variable x. Let us think of differentiation as an *operator* D acting on functions, hence for any function f we denote $f'(x)$ by Df. We could compose D with itself k times, k a positive integer, and get $D^k = \frac{d}{dx^k}$. We think of D^0 as inert (identity) operator 1 (also denoted by I or $\mathbb{1}$), which leaves its input inact. D^{-1} (also denoted by $\frac{1}{D}$) is the reverse operation of integration:

$$\left(D^{-1} f\right)(x) = \int f(x) dx.$$

Also for operators acting on the space of functions and spitting out functions, we could add, subtract and compose them with each other in the natural way. With all this in mind, the equation (40) could be interpreted as

$$\left(D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0\right) y = R,$$

or even $p(D)y = R$, where

$$p(r) = r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0, \tag{42}$$

is the characteristic polynomial.

Proposition 7. *Let the characteristic polynomial (42) of equation (41) be factored as*

$$p(r) = \prod_{k=1}^{n} (r - r_k),$$

where r_k are complex numbers. Then

$$\boxed{y = \left(\prod_{k=1}^{n} \frac{1}{D - r_k}\right) R = \left(\prod_{k=1}^{n} e^{r_k x} \frac{1}{D} e^{-r_k x}\right) R = e^{r_n x} \frac{1}{D} e^{-r_n x} \cdots e^{r_1 x} \frac{1}{D} e^{-r_1 x} R},$$

is the general solution of (41).

Example 36. Let us apply this method to Example 35. A particular solution is

$$\begin{aligned}
y_p &= e^x \frac{1}{D} e^{-x} e^x \frac{1}{D} e^{-x} \frac{e^x}{1+x^2} = e^x \frac{1}{D} \frac{1}{D} \frac{1}{1+x^2} \\
&= e^x \frac{1}{D} \int \frac{dx}{1+x^2} = e^x \int \tan^{-1} x dx \\
&= e^x \left(x \tan^{-1} x - \int \frac{x dx}{1+x^2} \right), && \text{integration by parts} \\
&= e^x \left(x \tan^{-1} x - \frac{1}{2} \log\left(1+x^2\right) \right),
\end{aligned}$$

which is the same result. ■

6.7 The Method of Laplace Transform, Analysis of Discontinuous and Periodic Inputs

Laplace transform gives another way to solve the constant coefficient differential equation (41), specially when $R(x)$ has discontinuities, which is a common situation in engineering where $R(x)$ contains a pulse, mathematically modeled by Heaviside step or Dirac delta functions. We show this by several examples.

Example 37. Recall Example 25. We assume $x_0 = 0$. Applying Laplace transform, we get the *algebraic* equation

$$sY - b + aY = \mathcal{L}(R),$$

where $Y = \mathcal{L}y$. Therefore

$$Y = \frac{b + \mathcal{L}(R)}{s + a},$$

hence

$$y = \mathcal{L}^{-1}\left(\frac{b}{s+a}\right) + \mathcal{L}^{-1}\left(\frac{1}{s+a}\mathcal{L}(R)\right) = be^{-ax} + \int_0^x e^{-a(x-\xi)}R(\xi)d\xi.$$

■

Example 38. To solve

$$y'' + 4y = 9x\sin x, \quad y(0) = 1, \quad y'(0) = 4,$$

applying Laplace transform, we get

$$s^2Y - s - 4 + 4Y = -\frac{d}{ds}\mathcal{L}\left(9\sin x\right) = -\frac{d}{ds}\frac{9}{s^2+1} = \frac{18s}{(s^2+1)^2},$$

hence

$$Y = \frac{s+4}{s^2+4} + \frac{18s}{(s^2+4)(s^2+1)^2} = \frac{s}{s^2+4} + \frac{4}{s^2+4} + \frac{2s}{s^2+4} + \frac{-2s}{s^2+1} + \frac{6s}{(s^2+1)^2},$$

which we have done partial fraction decomposition in the last line. Therefore

$$y = \mathcal{L}^{-1}Y = 3\cos(2x) + 2\sin(2x) - 2\cos x + \frac{1}{2}x\sin x.$$

■

Example 39. Let us use Laplace transform to give another solution to the differential equation in Example 35. Applying Laplace transform, we get

$$(s^2 - 2s + 1)Y - sA + 2A - B = \mathcal{L}\left(\frac{e^x}{1+x^2}\right),$$

where $Y = \mathcal{L}(y)$, $A = y(0)$, and $B = y'(0)$. Therefore

$$Y = \frac{As + B - 2A}{(s-1)^2} + \frac{1}{(s-1)^2}\mathcal{L}\left(\frac{e^x}{1+x^2}\right) = \frac{C_1}{s-1} + \frac{C_2}{(s-1)^2} + \frac{1}{(s-1)^2}\mathcal{L}\left(\frac{e^x}{1+x^2}\right).$$

Therefore

$$y = \mathcal{L}^{-1}(Y) = C_1e^x + C_2e^x x + xe^x * \frac{e^x}{1+x^2} = C_1e^x + C_2e^x x + \int_0^x (x-\xi)e^{x-\xi}\frac{e^\xi}{1+\xi^2}d\xi =$$
$$= C_1e^x + C_2e^x x + e^x\int_0^x \frac{x-\xi}{1+\xi^2}d\xi = C_1e^x + C_2e^x x + xe^x\tan^{-1}x - \frac{1}{2}e^x\log\left(1+x^2\right).$$

Notice that we did not need to find Laplace transform of $\frac{e^x}{1+x^2}$. ■

More generally we have:

Proposition 8. *A particular solution of equation (41), or more precisely, the unique solution to initial value problem*

$$\begin{cases} y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = R(x) \\ y(0) = y'(0) = \cdots = y^{(n-1)}(0) = 0 \end{cases}, \tag{43}$$

is given by convolution integral

$$\boxed{y(x) = \int_0^x R(\xi)h(x-\xi)d\xi, \quad h(x) = \mathcal{L}^{-1}\left(\frac{1}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}\right)}.$$

Note that $h(x)$ above is $y(x)$ for input $R(x) := \delta(x)$. That is why $h(x)$ is called the **unit impulse response of IVP (43)**. Proposition 8 is one in a collection of results under the name of **Duhamel's** (or **Superposition**) **Principle**.

Exercise 58. *Show that the unique solution to initial value problem*

$$y'' + y = R(x), \quad y(0) = y'(0)$$

is given by convolution integral $y(x) = \int_0^x R(\xi)\sin(x-\xi)d\xi$.

Example 40. Consider initial value problem

$$\begin{cases} y'' - 2y' + y = \delta(x-2) \\ y(0) = 0, \quad y'(0) = 0 \end{cases}, \tag{44}$$

applying Laplace transform, we get

$$s^2Y - 2sY + Y = e^{-2s},$$

hence

$$Y = \frac{e^{-2s}}{(s-1)^2}.$$

Therefore

$$y = \mathcal{L}^{-1}Y = (x-2)e^{x-2}u(x-2). \tag{45}$$

Let us check that the solution just found satisfies the original initial value problem. By product rule for differentiation, and the identities

$$\frac{d}{dx}u(x-a) = \delta(x-a), \quad f(x)\delta(x-a) = f(a)\delta(x-a),$$

valid for each constant a, and each function $f(t)$, we compute

$$y' = \left((x-2)e^{x-2}\right)' u(x-2) + (x-2)e^{x-2}\delta(x-2) = e^{x-2}(x-1)u(x-2) + 0,$$

$$y'' = \left((x-1)e^{x-2}\right)' u(x-2) + (x-1)e^{x-2}\delta(x-2) = e^{x-2}xu(x-2) + \delta(x-2),$$

which you can readily check satisfies the initial value problem we started with.

Now let us solve problem (44) directly without using Laplace transform. Our problem splits into two subproblems:

$$\begin{cases} y'' - 2y' + y = 0, & 0 < x < 2 \\ y(0) = 0, \quad y'(0) = 0 \end{cases}, \qquad \begin{cases} y'' - 2y' + y = 0, & 2 < x < \infty \\ y(2+) = ?, \quad y'(2+) = ? \end{cases},$$

where we also need to find the value of $y(x)$ and $y'(x)$ for $x = 2+$. The first subproblem has unique solution $y(x) = 0$ for $0 \leqslant x < 2$. To solve the first subproblem, we should first compute $y(2+)$ and $y'(2+)$. Specially $y(2-) = y'(2-) = 0$. Firstly, note that $y(x)$ does not jump at $x = 2$, because otherwise $u(x-2)$ appears in $y(x)$, hence $\delta(x-2)$ in $y'(x)$, hence $\delta'(x-2)$ in $y''(x)$; however our main differential equation in (44) witnesses that this is not the case. Therefore

$$y(2+) = y(2-) + 0 = 0.$$

Secondly, integrating our main differential equation in (44) from $x = 2-$ to $x = 2+$ gives

$$\left(y'(2+) - y'(2-)\right) - 2\left(y(2+) - y(2-)\right) + \int_{2-}^{2+} y(x)dx = 1,$$

or equivalently

$$y'(2+) - y'(2-) - 2 \times 0 + 0 = 1,$$

hence

$$y'(2+) = y'(2-) + 1 = 1.$$

The general solution to second subproblem is $y = C_1e^x + C_2e^x x$, and enforcing $y(2+) = 0$, $y'(2+) = 1$ gives $y = (x-2)e^{x-2}$. Putting all these together, we retrieve our previous solution (45). ■

Example 41. This example comes from [13, p. 477], after a correction. Consider the circuit in Figure 5, with $E = 90$ volts, $R = 110$ ohms, $L = 1$ henries, and $C = 0.001$ farads. We assume that there is no energy in the inductance and capacitance. Only on time interval $[0, 1]$ switch is open, hence the voltage applied across the RLC part is modeled by

$$v(t) = E(u(t) - u(t-1)).$$

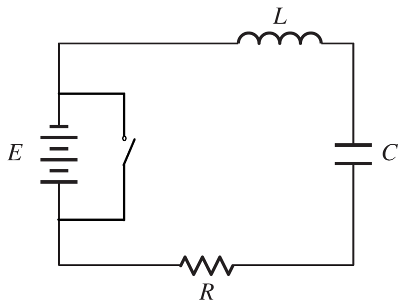


Figure 5: Circuit analyzed in Example 41. Switch is open only on time interval $[0, 1]$.

To find the variation of resulting current $i(t)$ in the circuit, we note that $i(t)$ satisfies initial value integro-differential equation

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int_0^t i(\tau)d\tau = v(t), \quad i(0) = 0. \tag{46}$$

Note that $i(0) = 0$ is because there is no initial energy in inductance. Also since there is no initial energy in capacitor, initial voltage $v_C(0)$ is zero, and the voltage drop across capacitance is

$$v_C(0) + \frac{1}{C}\int_0^t i(\tau)d\tau = \frac{1}{C}\int_0^t i(\tau)d\tau.$$

Applying Laplace transform to (46), assuming $\mathcal{L}i = I$, we have

$$LsI + RI + \frac{I}{Cs} = E\left(\frac{1}{s} - \frac{e^{-s}}{s}\right),$$

hence

$$\begin{aligned} I &= \frac{90\,(1-e^{-s})}{s^2 + 110s + 1000} = \frac{90\,(1-e^{-s})}{(s+10)(s+100)} = (1-e^{-s})\left(\frac{1}{s+10} - \frac{1}{s+100}\right) \\ &= \left(\frac{1}{s+10} - \frac{1}{s+100}\right) - e^{-s}\left(\frac{1}{s+10} - \frac{1}{s+100}\right). \end{aligned}$$

Therefore

$$\begin{aligned} i(t) &= \mathcal{L}^{-1}(I(s)) = \left(e^{-10t} - e^{-100t}\right)u(t) - \left(e^{-10(t-1)} - e^{-100(t-1)}\right)u(t-1) \\ &= \begin{cases} 0, & t < 0 \\ e^{-10t} - e^{-100t}, & 0 < t < 1 \\ e^{-10t} - e^{-10(t-1)} - e^{-100t} + e^{-100(t-1)}, & t > 1 \end{cases}. \end{aligned}$$

■

Example 42. This example comes from [13, p. 480]. Consider the initial value problem

$$\frac{d^2x}{dt^2} + 4\frac{dx}{dt} + 20x = f(t), \quad x'(0) = x(0) = 0,$$

where $f(t)$ is the 2π-periodic square wave given by

$$\begin{cases} 20, & 0 \leqslant t < \pi \\ -20 & \pi \leqslant t < 2\pi \end{cases},$$

on time interval $[0, 2\pi]$. Applying Laplace transform, we get

$$\begin{aligned} s^2X + 4sX + 20 &= \frac{\int_0^{2\pi} f(t)e^{-st}dt}{1 - e^{-2\pi s}} \\ &= \frac{\int_0^{\pi} 20e^{-st}dt - \int_{\pi}^{2\pi} 20e^{-st}dt}{1 - e^{-2\pi s}} = \frac{20}{s}\frac{(1 - e^{-\pi s})^2}{1 - e^{-2\pi s}} \\ &= \frac{20}{s}\frac{1 - e^{-\pi s}}{1 + e^{-\pi s}}, \end{aligned}$$

hence

$$\begin{aligned} X &= \frac{20}{s(s^2 + 4s + 20)}\frac{1 - e^{-\pi s}}{1 + e^{-\pi s}} \\ &= \frac{20}{s((s+2)^2 + 4^2)}(1 - e^{-\pi s})\left(1 - e^{-\pi s} + e^{-2\pi s} - + \cdots\right), \qquad \text{geometric series} \\ &= \frac{20}{s((s+2)^2 + 4^2)}\left(1 - 2e^{-\pi s} + 2e^{-2\pi s} - + \cdots\right). \end{aligned}$$

Therefore

$$x(t) = f(t) - 2f(t-\pi)u(t-\pi) + 2f(t-2\pi)u(t-2\pi) - + \cdots,$$

where

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}\frac{20}{s((s+2)^2 + 4^2)} = \mathcal{L}^{-1}\left(\frac{1}{s} + \frac{-(s+2) - 2}{(s+2)^2 + 4^2}\right) \\ &= 1 - \cos(4t)e^{-2t} - \frac{1}{2}\sin(4t)e^{-2t}. \end{aligned}$$

To find a closed formula for $x(t)$ in each interval $[n\pi, (n+1)\pi]$, $n = 0, 1, 2, \ldots$, setting

$$g := \cos(4t)e^{-2t} + \frac{1}{2}\sin(4t)e^{-2t}, \quad \alpha := e^{2\pi},$$

we have

$$\begin{aligned}
x(t) &= (1-g) - 2(1-\alpha g) + 2(1-\alpha^2 g) - + \cdots + 2(-1)^n(1-\alpha^n g)\\
&= (1-2+2-+\cdots+2(-1)^n) - \left(1 - 2\alpha + 2\alpha^2 + - \cdots + 2(-1)^n\alpha^n\right) g\\
&= (-1)^n - \left(1 - 2\alpha\frac{1-(-\alpha)^n}{1+\alpha}\right) g = (-1)^n + \frac{\alpha-1}{\alpha+1}g - \frac{2\alpha}{\alpha+1}(-\alpha)^n g\\
&= \left(\frac{\alpha-1}{\alpha+1}g\right) + \left((-1)^n - \frac{2\alpha}{\alpha+1}(-\alpha)^n g\right).
\end{aligned}$$

Therefore on $t \in [n\pi, (n+1)\pi]$, we have

$$x(t) = x_1(t) + x_2(t),$$

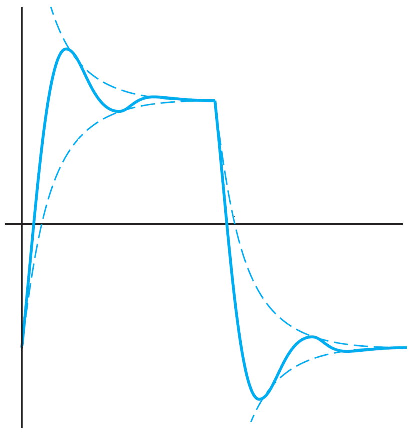
where

$$x_1(t) = \frac{\alpha-1}{\alpha+1}\left(\cos(4t)e^{-2t} + \frac{1}{2}\sin(4t)e^{-2t}\right),$$

is the damped **transient part**, and

$$x_2(t) = (-1)^n\left(1 - \frac{2\alpha}{\alpha+1}\left(\cos(4t)e^{-2(t-n\pi)} + \frac{1}{2}\sin(4t)e^{-2(t-n\pi)}\right)\right),$$

is the **steady-state part**. Note that $x(t)$ is 2π-periodic, the cycle $[0, 2\pi]$ is shown in the following figure.



■

Exercise 59. *Solve the initial value problem*

$$y'' + y = \sum_{n\geqslant 1}\delta(x - n\pi), \quad y(0) = 0, \quad y'(0) = 0.$$

Exercise 60. *In all of the examples in this section, justify the prediction of Final Value Theorem.*

6.8 The Method of Undetermined Coefficients

Consider equation (41). If $R(x)$ is of the special form in the left column of the following table, then it is guaranteed that there is a particular solution of (41) of the form in the corresponding right column of the same table.

$R(x)$	y
d-poly	x^m (d-poly)
e^{ax}	$Ax^m e^{ax}$
$\cos bx$ or $\sin bx$	$x^m(A\cos bx + B\sin bx)$
(d-poly)$e^{ax}\cos bx$ or (d-poly)$e^{ax}\sin bx$	$x^m e^{ax}$((d-poly)$\cos bx$+(another d-poly)$\sin bx$)

The table needs some explanation:

1. "d-poly" stands for "a polynomial of degree d".
2. Integer m in different rows of the second column, respectively, equals the order of $r = 0$, $r = a$, $r = ib$ or $r = a + bi$ as the root of the characteristic polynomial (42).
3. Clearly the last row contains the other rows as special case.

Example 43. Consider the equation

$$y''' - y'' - y' + y = (12x + 14)e^x.$$

Since $r = 1$ is a repeated root with multiplicity $m = 2$ of the characteristic polynomial

$$r^3 - r^2 - r + 1 = (r - 1)^2(r + 1),$$

the equation has a particular solution of the form

$$y = e^x x^2 (Ax + B) = e^x \left(Ax^3 + Bx^2\right),$$

where A and B are constants. Satisfying this indefinite form into the equation gives

$$\begin{aligned}(12x + 14)e^x &= y - y' - y'' + y''' \\ &= e^x \left(Ax^3 + Bx^2\right) \\ &- e^x \left(Ax^3 + (3A + B)x^2 + 2Bx\right) \\ &- e^x \left(Ax^3 + (6A + B)x^2 + (6A + 4B)x + 2B\right) \\ &+ e^x \left(Ax^3 + (9A + B)x^2 + (18A + 6B)x + 6A + 6B\right),\end{aligned}$$

hence after eliminating e^x, and equating similar powers of x, we have

$$\begin{cases} 12A = 12 \\ 6A + 4B = 14 \end{cases},$$

hence $A = 1$ and $B = 2$. Therefore $y = e^x\left(x^3 + 2x^2\right)$ is a particular solution. ■

Exercise 61. *Find a particular solution of each of the following equations: (a)* $y''' + 2y'' = x^4$. *(b)* $y'' + 2y' + y = e^{-x} + xe^x$.

6.9 Resonance

Recall Example 9, where we modeled RLC circuit and mass-spring-dashpot system with nonhomogeneous constant-coefficient second order ODE

$$y'' + ay' + by = R(t). \tag{47}$$

Electrical and mechanical engineers study this equation in detail. Here I just want to mention the important phenomenon of *resonance* occurring to this equation when the exciter is a simple sinusoidal, say

$$R(t) = A\cos(\omega t + \varphi) = B\cos(\omega t) + C\sin(\omega t),$$

where amplitude A, angular frequency ω, and initial phase φ are constants. Resonance happens when $i\omega$ is very close to a root of characteristic polynomial $r^2 + ar + b$, namely when

$$a \approx 0, \quad \omega \approx \sqrt{b}.$$

If this happens then by Section 6.8, the solution of (47) is of the form

$$y = \underbrace{C_1e^{r_1 t} + C_2e^{r_2 t}}_{y_1} + \underbrace{t\left(D\cos(\omega t) + E\sin(\omega t)\right)}_{y_2},$$

where

$$r_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}, \quad r_2 = \frac{-a - \sqrt{a^2 - 4b}}{2},$$

are both either negative real, or complex with negative real part. Thus after a while, y_1 fades, and $y \approx y_2$ (Exercise 56). However because of t multiplier, y_2 becomes unbounded at some times. This is called **resonance**, which must seriously be taken into account in the design of bridges or buildings, avoiding catastrophic failure of the structure.

More generally in equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = A\cos(\omega t + \varphi),$$

if $i\omega$ is a characteristic root, namely satisfies

$$(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \cdots + a_1(i\omega) + a_0 = 0,$$

then resonance happens.

6.10 Euler-Cauchy Equation

An equation of the form

$$\boxed{x^2\frac{d^2y}{dx^2} + ax\frac{dy}{dx} + by = R(x)},$$

where a and b are constants, and $g(x)$ is scalar-valued function of scalar x, is called **Cauchy-Euler** (or **equidimensional equation**). It reduces to a constant coefficient linear equation by changing independent variable x to $z = \log x$. Since

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = \frac{dy}{dz}\frac{1}{x},$$

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

the new equation is

$$\frac{d^2y}{dz^2} + (a-1)\frac{dy}{dz} + by = R\left(e^z\right).$$

This trick equally works for the higher-order equation

$$x^n\frac{d^ny}{dx^n} + a_{n-1}x\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_0y = g(x), \tag{48}$$

which we now explain. Introducing change of variable $z = \log x$, and differential operators

$$D_x := \frac{d}{dx}, \quad D_z := \frac{d}{dz},$$

exactly as before one could show that

$$D_x = \frac{1}{x}D_z, \quad D_x^2 = \frac{1}{x^2}D_z(D_z-1), \quad \ldots, \quad D_x^n = \frac{1}{x^n}\prod_{0\leqslant k\leqslant n-1}(D_z-k).$$

This way equation (48) reduces to constant coefficient linear equation

$$\left(a_0 + a_1D_z + a_2D_z(D_z-1) + \cdots + \prod_{0\leqslant k\leqslant n-1}(D_z-k)\right)y = R\left(e^z\right).$$

6.11 Some Qualitative Analysis

To be added. ???

7 Power Series Methods

The solutions of most differential equations appearing in practice can *not* be expressed by elementary functions and their integrals.[14] *Solving* such differential equations then means to decide on appropriate simple building block functions, and then represent the solution in terms of these blocks. This is analogous to building new words to make more advanced sentences. In this chapter we use power functions x^n, $n = 0, 1, 2, \ldots$, or their shifted versions

$$(x - x_0)^n, \quad x_0 \in \mathbb{R}, \quad n = 0, 1, 2, \ldots,$$

(by convention, we consider $(x - x_0)^0$ as constant function 1.) as our building blocks, and we try to represent functions by *infinite* series like

$$\boxed{\sum_{n \geqslant 0} a_n (x - x_0)^n}, \tag{49}$$

with scalar sequences (a_n) (these are called **power series**), or generalizations like

$$\boxed{|x - x_0|^r \sum_{n \geqslant 0} a_n (x - x_0)^n}, \tag{50}$$

with real numbers r (these are called **Frobenius series**).

Exercise 62. *Geometric series* $\frac{1}{1-x} = \sum_{n \geqslant 0} x^n$*, valid for* $|x| < 1$*, is maybe the most important power series. Justify it.*

Exercise 63. *Recall the construction of basic elementary functions namely polynomials,* $(1 + x)^\alpha$*,* e^x*,* $\log(1 + x)$*,* $\sin x$*,* $\cos x$*,* $\arctan x$ *by power series.*

A scalar-valued function $f(x)$ of a real variable which equals a power series like (49) on some interval $|x - x_0| < R$, $R > 0$, is called **(real) analytic at** x_0. A function defined on $|x - x_0| < R$ which agrees on $0 < |x - x_0| < R$ with some analytic function at x_0, is said to have an **analytic extension** to $|x - x_0| < R$; and with an abuse of language, these functions are considered to be analytic at x_0. A function is called **analytic on an open interval** if it is analytic at all points of that interval.

[14] Even conversely, elementary functions could be defined by differential equation. For example, $\sin x$ is the unique solution of the initial value problem

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

or the power function $y = (1 + x)^\alpha$ is the unique solution of the initial value problem

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

Example 44. Consider

$$f(x) := e^x = \sum_{n\geqslant 0} \frac{x^n}{n|}, \quad g(x) := \frac{\sin x}{x} = \frac{1 - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots}{x}, \quad h(x) := \sum_{n\geqslant 1} n^n x^n.$$

Function $f(x)$ is analytic at $x = 0$, because the series converges for each $x \in \mathbb{R}$. Function $g(x)$ is considered to be analytic at $x = 0$, because it equals $1 - \frac{x^2}{3!} + \frac{x^4}{5!} - + \cdots$ convergent for each $x \in \mathbb{R}$. Function $h(x)$ is not analytic at $x = 0$, because root test shows that it is convergent only at $x = 0$. ■

Remark 3 (Optional)**.** It is a non-trivial fact [19, Proposition 1.2.3] that a power series $\sum a_n(x - x_0)^n$ convergent on $|x - x_0| < R$, $R > 0$, is analytic on $|x - x_0| < R$. The converse is also true: an analytic function on $|x - x_0| < R$, $R > 0$, has a power series representation $\sum a_n(x - x_0)^n$ convergent on $|x - x_0| < R$. (reference ???) ■

Those properties of analytic functions that we need are gathered in the following propositions.

Proposition 9. *(a) Polynomials, e^x, $\sin x$, $\cos x$ are analytic on $\mathbb{R}$. Rational functions (namely quotient of polynomials) are analytic on $\mathbb{R}$ except those roots of the denumerator whose multiplicity is strictly bigger than their multiplicity as (potential) roots of numerator. $\log(1 + x)$ and $(1 + x)^\alpha$ ($\alpha \in \mathbb{R}$) are analytic on $|x| < 1$.*

(b) Let f and g be analytic at x_0. Then $f \pm g$, $f \times g$ are analytic at x_0. f/g is analytic at x_0 if $g(x_0) \neq 0$.

(c) If f is analytic at x_0, and g is analytic at $f(x_0)$, then $g \circ f$ is analytic at x_0.

Proposition 10. *Consider the power series $\sum_{n\geqslant 0} a_n(x - x_0)^n$, and let R, called **the radius of convergence**, be either of the following limits (in case they exist or equal ∞)*

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| \quad \text{or} \quad R = \lim_{n\to\infty} |a_n|^{-\frac{1}{n}}.$$

(a) If $0 < R < \infty$, then the power series converges for $|x - x_0| < R$, and diverges for $|x - x_0| > R$.
(b) If $R = 0$, then the power series diverges for all $x \in \mathbb{R}$ except $x = x_0$.[15]
(c) If $R = \infty$, then the power series converges for all $x \in \mathbb{R}$.

Proposition 11. *Let f be analytic at x_0 with representation*

$$f(x) = \sum_{n\geqslant 0} a_n(x - x_0)^n,$$

[15]Note that in this case the power series is useless, and does not define an analytic function at x_0.

valid on $|x - x_0| < R$*, for some positive* R*. Then:*

(a)

$$f'(x) = \sum_{n\geqslant 1} n a_n (x - x_0)^{n-1}, \qquad \int_{x_0}^{x} f(\xi)d\xi = \sum_{n\geqslant 0} \frac{a_n}{n+1}(x - x_0)^{n+1},$$

valid on the same interval.

(b) f *is smooth on the same interval, and* $a_n = \frac{f^{(n)}(x_0)}{n!}$ *for each* n*. These are called* ***Taylor coefficients*** *of* f *at* x_0*.*

(c) If $f(x) = \sum_{n\geqslant 0} b_n (x - x_0)^n$ *valid on the same interval, then* $a_n = b_n$ *for each* n*.*

Part (a) is [1, vol. I, p. 432]; other parts follow immediately.

Exercise 64. *There exists smooth non-analytic functions. Prove that the following function is smooth on* $\mathbb{R}$ *but not analytic at* $x = 0$*.*

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}.$$

(Hint. Prove that $f^{(n)}(0) = 0$ *for* $n = 0, 1, 2, \ldots$*.)*

In Sections 7.1 and 7.2, we concentrate on linear second-order differential equations of the form

$$y'' + P(x)y' + Q(x)y = 0, \tag{51}$$

where P and Q are real-valued functions of real variable x which ranges on some interval $I = (x_0 - R, x_0 + R)$ or its punctured version $I = (x_0 - R, x_0 + R) \setminus \{x_0\}$, for some real x_0 and positive real R. Some of the most important such equations with *non-elementary* solutions are the **Airy equation**

$$y'' + xy = 0, \tag{52}$$

the **Bessel equation**

$$x^2 y'' - xy' + \left(x^2 - p^2\right) y = 0, \tag{53}$$

where p is a real constant,[16] the **Legendre equation**

$$(1 - x^2)y'' - 2xy' + p(p+1) = 0, \tag{54}$$

where p is a real constant,[17], the **Hermite Equation**

$$y'' - 2xy' + 2py = 0, \tag{55}$$

where p is a real constant.[18] Bessel and Legendre functions appear in the study of Laplacian in cylindrical and spherical coordinates. Hermit functions appear in the study of quantum oscillator.

[16] When p is half of an odd integer, the Bessel Equation has one elementary solution.

[17] When p is an integer, the Legendre equation has one polynomial solution, and another non-elementary one.

[18] When p is a nonnegative integer, the Hermite Equation has one polynomial solution, and another non-elementary one.

7.1 Linear Second-Order Ordinary Differential Equation: Nonsingular Points

Now we study equation (51) around *good* point x_0, in the following sense. A point x_0 is called **nonsingular** (or **ordinary**) for equation (51), defined on $I = (x_0 - R, x_0 + R)$, if both P and Q are analytic at x_0 (or have an analytic extension to I), otherwise x_0 is a **singular** point.

Theorem 10. *Consider equation (51) on $I = (x_0 - R, x_0 + R)$, $R > 0$, and assume*

$$P(x) = \sum_{n \geqslant 0} p_n (x - x_0)^n, \quad Q(x) = \sum_{n \geqslant 0} q_n (x - x_0)^n,$$

valid on I. Then (51) has two independent solutions

$$y_1(x) = \sum_{n \geqslant 0} a_n (x - x_0)^n, \quad y_2(x) = \sum_{n \geqslant 0} b_n (x - x_0)^n,$$

valid and satisfying equation on I.

Easy proof could be found in [1, vol. II, p. 169], or [23, p. 208].

Remark 4 (Optional)**.** A general theme is that a system of ODEs or PDEs with analytic data has analytic solutions. Refer to [11, vol. I, 10.5.3] or [24, vol II, 6.4]. However Hans Lewy found a linear PDE with smooth data having no solution.

■

Example 45. Let us use this theorem to solve

$$y'' + y = 0, \tag{56}$$

around nonsingular point $x = 0$. Replacing y by $\sum_{n \geqslant 0} a_n x^n$, and y'' by

$$\frac{d}{dx}\frac{d}{dx} \sum_{n \geqslant 0} a_n x^n = \frac{d}{dx} \sum_{n \geqslant 1} n a_n x^{n-1} = \sum_{n \geqslant 2} n(n-1) a_n x^{n-2},$$

we have

$$\begin{aligned} 0 = \sum_{n \geqslant 2} n(n-1) a_n x^{n-2} + \sum_{n \geqslant 0} a_n x^n &= \sum_{n \geqslant 0} (n+2)(n+1) a_{n+2} x^n + \sum_{n \geqslant 0} a_n x^n \\ &= \sum_{n \geqslant 0} \left((n+2)(n+1) a_{n+2} + a_n \right) x^n, \end{aligned}$$

hence

$$(n+2)(n+1) a_{n+2} + a_n = 0,$$

for all $n \geqslant 0$. This recursively determines a_n, $n \geqslant 2$, in terms of a_0 and a_1, as follows:

$$a_2 = -\frac{1}{2}a_0, \quad a_3 = -\frac{1}{3!}a_1, \quad a_4 = -\frac{1}{4\times 3}a_2 = \frac{1}{4!}a_0, \quad a_5 = \frac{1}{5!}a_1, \quad \text{etc.}$$

Therefore, we have shown that

$$y = a_0\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \cdots\right) + a_1\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots\right),$$

satisfies (56). Notice that the powers series in parentheses are $\cos x$ and $\sin x$, respectively. ■

Exercise 65. *BR, page 116. deriving properties of sin and cos with y"+y=0. ???*

Exercise 66. *Determine a lower bound for the radius of convergence of series solutions of the differential equation*

$$(1 + x^2) + 2xy + e^x y = 0,$$

around the point $x = -\frac{1}{2}$.

Example 46. As another example for the application of Theorem 10, we solve

$$y'' + \frac{1}{2x+3}y' + e^x y = 0, \tag{57}$$

around nonsingular point $x = 1$. It is slightly easier to change independent variable x to $t = x - 1$. Then the equivalent problem is to solve

$$\frac{d^2y}{dt^2} + \frac{1}{2t+5}\frac{dy}{dt} + e^{t+1}y = 0,$$

around nonsingular point $t = 0$. Replacing y by $\sum_{n\geqslant 0} a_n t^n$, and using the following power series representations of coefficients

$$\frac{1}{2t+5} = \frac{1}{5}\frac{1}{1+\frac{2t}{5}} = \frac{1}{5}\sum_{n\geqslant 0}\alpha^n t^n, \quad \alpha = -\frac{2}{5}, \quad |t| < \frac{5}{2}$$

and

$$e^{t+1} = \sum_{n\geqslant 0}\frac{e}{n!}t^n, \quad t \in \mathbb{R}$$

we have

$$0 = \sum_{n\geqslant 2} n(n-1)a_n t^{n-2} + \left(\frac{1}{5}\sum_{n\geqslant 0}\alpha^n t^n\right)\left(\sum_{n\geqslant 1} n a_n t^{n-1}\right) + \left(\sum_{n\geqslant 0}\frac{e}{n!}t^n\right)\left(\sum_{n\geqslant 0} a_n t^n\right),$$

or equivalently

$$0 = 2a_2 + 6a_3 t + \cdots + \frac{1}{5}\left(1 + \alpha t + \alpha^2 t^2 + \cdots\right)\left(a_1 + 2a_2 t + \cdots\right) + \\ + e\left(1 + t + \frac{1}{2}t^2 + \cdots\right)\left(a_0 + a_1 t + a_2 t^2 + \cdots\right).$$

Comparing coefficients of similar t^n on both sides, we have

$$0 = 2a_2 + \frac{1}{5}a_1 + ea_0, \quad 0 = 6a_3 + \frac{1}{5}(2a_2 + \alpha a_1) + e(a_1 + a_0), \quad \text{etc.},$$

hence

$$a_2 = -\frac{e}{2}a_0 - \frac{1}{10}a_1, \quad a_3 = -\frac{1}{15}a_2 - \frac{\alpha + 5e}{30}a_1 - \frac{e}{6}a_0, \quad \text{etc.},$$

which recursively determines a_n, $n \geqslant 2$, in terms of a_0 and a_1. What Theorem 10 guarantees is that, then

$$y = \sum_{n \geqslant 0} a_n t^n = \sum_{n \geqslant 0} a_n (x - 1)^n,$$

converges at least for $|x - 1| < \frac{5}{2}$, and solves (46). ■

Remark 5. Theorem 10, gives a lower bound for R where the power series representation of y_1 and y_2 are valid on $|x - x_0| < R$, and satisfy equation (51). You could show that the Legendre equation (54), when p is nonnegative integer, has a polynomial solution (of degree p); although the other independent solution has radius of convergence 1. ■

Exercise 67. *Consider the differential equation*

$$x^2 y' - y + x = 0,$$

around nonsingular point $x = 0$. Prove that it has no analytic solution at $x = 0$. [Hint. Assume it has analytic solution $y = \sum_{n \geqslant 0} a_n x^n$, and show that $a_n = (n-1)!$ for $n \geqslant 1$. Continue.]

7.2 Linear Second-Order Ordinary Differential Equation: Regular Nonsingular Points

Now we study equation (51) around *bad but not that much bad* point x_0, in the following sense. A point x_0 is called **regular singular** for equation (51), defined on $I = (x_0 - R, x_0 + R) \setminus \{x_0\}$, if either P or Q or both has no analytic extension to $(x_0 - R, x_0 + R)$, but both $(x - x_0)P(x)$ and $(x - x_0)^2 Q(x)$ have analytic extension to $(x_0 - R, x_0 + R)$. Why we study such points? Firstly because many important points for practical equations are of this type, and secondly because we can completely analyze them by the following theorem.

Theorem 11 (Frobenius)**.** *Consider equation (51) on* $I = (x_0 - R, x_0 + R) \setminus \{x_0\}$, $R > 0$, *and assume*

$$(x - x_0)P(x) = \sum_{n \geqslant 0} p_n (x - x_0)^n, \quad (x - x_0)^2 Q(x) = \sum_{n \geqslant 0} q_n (x - x_0)^n,$$

valid on I.[19] *Let* r_1 *and* r_2 *be roots of the* ***indicial equation***

$$\boxed{r(r-1) + p_0 r + q_0 = 0}. \tag{58}$$

(a) If $r_1 - r_2$ *is not an integer, then (51) has two independent solutions*

$$\boxed{y_1(x) = |x - x_0|^{r_1} \sum_{n \geqslant 0} a_n (x - x_0)^n, \quad (a_0 = 1)}, \tag{59}$$

$$y_2(x) = |x - x_0|^{r_2} \sum_{n \geqslant 0} b_n (x - x_0)^n, \quad (b_0 = 1),$$

valid and satisfying equation for $0 < |x - x_0| < R$.

(b) If $r_1 - r_2$ *is a nonnegative integer, then (51) has two independent solutions*

$$y_1(x) = |x - x_0|^{r_1} \sum_{n \geqslant 0} a_n (x - x_0)^n, \quad (a_0 = 1),$$

$$\boxed{y_2(x) = |x - x_0|^{r_2} \sum_{n \geqslant 0} b_n (x - x_0)^n + C y_1(x) \log |x - x_0|, \quad (b_0 = 1)}, \tag{60}$$

valid and satisfying equation for $0 < |x - x_0| < R$. *Furthermore,* $C \neq 0$ *when* $r_1 = r_2$.[20]

Proof could be found in [8, p. 177] or [4, p. 282-5]. For the corresponding theorem for higher-order (instead of second-order) equations refer [10, section 4.8].

Remark 6 (Optional)**.** Most references, equivalently, instead of equation (51), state Frobenius theorem for equation

$$x^2 y'' + x p(x) y' + q(x) y = 0,$$

assuming that p and q be analytic at x_0. In our formulation we followed [4]. ■

[19] More strictly, we should have said that $(x - x_0)P(x)$ and $(x - x_0)^2 Q(x)$ have analytic extensions to $(x_0 - R, x_0 + R)$ given by power series on the right hand side.

[20] [8, p. 165] says that, when $r_1 \neq r_2$, the second independent solution y_2 could have the form

$$y_2(x) = |x - x_0|^{r_1 + 1} \sum_{n \geqslant 0} b_n (x - x_0)^n + y_1(x) \log |x - x_0|.$$

Remark 7. As a motivation for indicial equation (58), *power singularity* in (59), and *log singularity* in (60), let us study the simplest differential equation featuring regular singular points, namely Euler-Cauchy ODE

$$y'' + \frac{p}{x}y' + \frac{q}{x^2}y = 0, \tag{61}$$

where p and q are scalars. This was studied in Section 6.10, but here we follow an alternative approach. The power function $y = x^r$, $r \in \mathbb{C}$, satisfies (61) exactly when

$$0 = r(r-1)x^{r-2} + prx^{r-2} + qx^{r-2} = x^{r-2}\left(r(r-1) + pr + q\right),$$

which happens exactly when

$$r(r-1) + pr + q = 0.$$

Now let r_1 and r_2 be the roots of this quadratic equation. If $r_1 \neq r_2$, then $y = C_1x^{r_1} + C_2x^{r_2}$ is the general solution of (61) on $(0,\infty)$; otherwise if $r_1 = r_2 = \frac{1-p}{2}$, then

$$y = C_1x^{r_1} + C_2x^{r_1}\int x^{-2r_1}e^{-\int \frac{p}{x}dx}dx = C_1x^{r_1} + C_2x^{r_1}\int x^{-2r_1-p}dx = C_1x^{r_1} + C_2x^{r_1}\log(x).$$

■

Example 47. Let us try to solve Bessel equation (53) around $x = 0$. (This is the most important point in practical applications of the Bessel equation.) Rewriting our equation as

$$y'' + \frac{1}{x}y' + \left(1 - \frac{p^2}{x^2}\right)y = 0,$$

defined on $(-\infty,\infty)\setminus\{0\}$, the desired point $x = 0$ turns out to be regular singular, because both $\frac{1}{x}$ and $1 - \frac{p^2}{x^2}$ are non-analytic at $x = 0$, but after multiplication by x and x^2, both become analytic:

$$x\frac{1}{x} = 1, \quad x^2\left(1 - \frac{p^2}{x^2}\right) = -p^2 + x^2.$$

Indicial equation

$$r(r-1) + r - p^2 = 0,$$

has roots $r_1 = p$ and $r_2 = -p$. Without loss og generality we assume $p \geqslant 0$. By Theorem 11, our equation has a solution of the form

$$y = x^p\sum_{n\geqslant 0} a_nx^n = \sum_{n\geqslant 0} a_nx^{n+p}, \quad a_0 = 1,$$

valid and satisfying equation on $x > 0$. Plugging this into Bessel equation,

$$\begin{aligned}
0 &= x^2 \sum_{n\geqslant 0} (n+p)(n+p-1)a_n x^{n+p-2} + x \sum_{n\geqslant 0} (n+p)a_n x^{n+p-1} + (x^2-p^2) \sum_{n\geqslant 0} a_n x^{n+p} \\
&= x^p \left(\sum_{n\geqslant 0} (n+p)(n+p-1)a_n x^n + \sum_{n\geqslant 0} (n+p)a_n x^n + (x^2-p^2) \sum_{n\geqslant 0} a_n x^n \right) \\
&= x^p \left(\sum_{n\geqslant 0} n(n+2p)a_n x^n + \sum_{n\geqslant 0} a_n x^{n+2} \right) \\
&= x^p \left((1+2p)a_1 x + \sum_{n\geqslant 0} \left((n+2)(n+2+2p)a_{n+2} + a_n\right) x^{n+2} \right),
\end{aligned}$$

hence

$$a_1 = 0, \quad a_{n+2} = \frac{-a_n}{(n+2)(n+2+2p)},$$

for each $n \geqslant 0$. Therefore

$$y_1 = x^p \left(1 + \sum_{n\geqslant 1} \frac{(-1)^n}{n!(1+p)(2+p)\cdots(n+p)} \left(\frac{x}{2}\right)^{2n} \right).$$

In mathematical tradition $\frac{1}{2^p p!} y_1$ is denoted by $J_p(x)$, and is called the **Bessel function of first kind of order** p. Hence we have shown that

$$J_p(x) = \sum_{n\geqslant 0} \frac{(-1)^n}{n!(n+p)!} \left(\frac{x}{2}\right)^{2n+p},$$

is a solution of Bessel equation on $(0,\infty)$. Further use of Theorem 11, gives another independent solution. Interested reader can refer to [8, p. 168-178]. The end result is as follows. The general solution to Bessel equations on $(0,\infty)$ is

$$y = \begin{cases} C_1 J_p(x) + C_2 J_{-p}(x), & p \notin \mathbb{Z} \\ C_1 J_n(x) + C_2 K_n(x), & p = n \in \mathbb{Z}_{\geqslant 0} \end{cases},$$

where $K_n(x)$, the **Bessel function of second kind of order** n, is

$$K_n(x) = J_n(x)\log x - \frac{1}{2}\sum_{m=0}^{n-1} \frac{(n-m-1)!}{m!} \left(\frac{x}{2}\right)^{2m-n} - \frac{1}{2}\sum_{m\geqslant 0} \frac{(-1)^m (h_m + h_{m+n})}{m!(m+n)!} \left(\frac{x}{2}\right)^{2m+n},$$

where $h_0 = 0$ and $h_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m}$ for m positive integer. Notice Figure 6. On way to distinguish between Bessel functions of first and second kinds is their behavior at origin:

$$J_0(0+) = 1, \quad J_p(0+) = 0, \quad J_{-p}(0+) = \infty, \quad K_0(0+) = K_p(0+) = -\infty,$$

for $p > 0$. ■

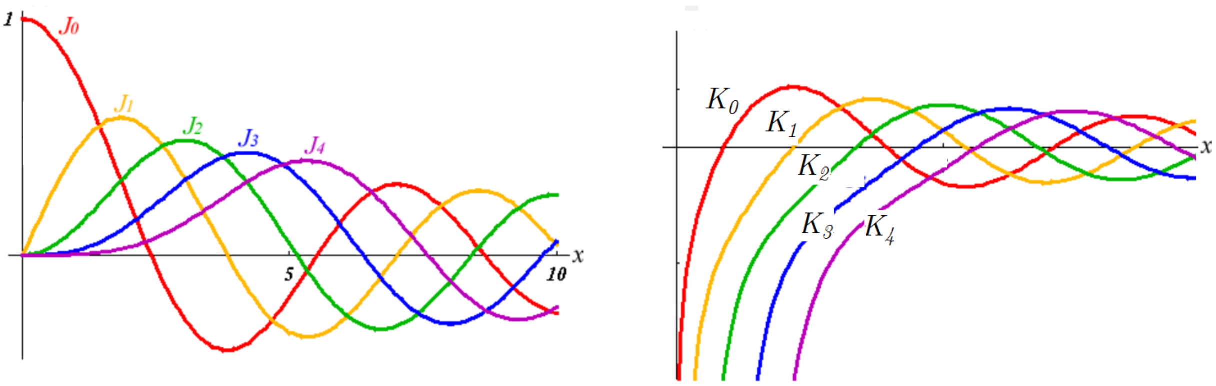


Figure 6: Bessel Functions.

7.3 First-Order Ordinary Differential Equations

We now want to solve first-order equations $y' = F(x, y)$ using power series method. First we need to make sense of F to be analytic. Let $F(x, y)$ be a real-valued function of real variables x and y, defined on some open rectangle around point $P_0 = (x_0, y_0)$. This function is **(real) analytic at** P_0 if it equals a double power series like

$$\sum_{m,n\geqslant 0} a_{mn}(x-x_0)^m(y-y_0)^n = a_{00} + a_{10}(x-x_0) + a_{01}(y-y_0) + a_{11}(x-x_0)(y-y_0) + \cdots,$$

on some (nonempty) open rectangle around P_0.

We have the following uniqueness result:

Proposition 12. *If two double power series*

$$\sum_{m,n\geqslant 0} a_{mn}(x-x_0)^m(y-y_0)^n \quad \textit{and} \quad \sum_{m,n\geqslant 0} b_{mn}(x-x_0)^m(y-y_0)^n,$$

are equal on some open rectangle around (x_0, y_0)*, then* $a_{mn} = b_{mn}$ *for each* m *and* n.

Theorem 12. *Let* $F(x, y)$ *be an analytic function at point* (x_0, y_0)*. Then the unique local solution* $y = y(x)$ *of the initial value problem*

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

is analytic at x_0.

Easy proof could be found in [4, p. 127-8].

Example 48. To solve equation $y' = y^2 + x$ around $(0,0)$, we plug $y = \sum_{n\geqslant 0} a_n x^n$ into it. Therefore

$$a_1 + 2a_2x + 3a_3x^2 + \cdots = x + \left(a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots\right)^2,$$

hence

$$a_1 = a_0^2, \quad 2a_2 = 1 + 2a_0a_1, \quad 3a_3 = 2a_0a_2 + a_1^2, \quad \text{etc.},$$

which gives all a_n, $n \geqslant 1$, in terms of a_0:

$$a_1 = a_0^2, \quad a_2 = \frac{1}{2} + a_0^3, \quad a_3 = \frac{1}{3}a_0 + a_0^4, \quad \text{etc..}$$

Therefore

$$y = a_0\left(1 + a_0^2x + \left(\frac{1}{2} + a_0^3\right)x^2 + \left(\frac{1}{3}a_0 + a_0^4\right)x^3 + \cdots\right),$$

solves our equation. ■

Remark 8 (Optional)**.** Equations $y' = y^2 \pm x$ or $y' = y^2 \pm x^2$ are closely related to Bessel equation! This is already reflected in computer-generated solution (12). Refer to [25, p. 126]. ■

8 Systems of First-Order Ordinary Differential Equations I: Basics and Linear Equations

As we have seen in Chapter 3, some dynamic phenomena are described by systems of ordinary differential equations. In this chapter and the next we study such equations. We confine ourself to first-order systems of the form

$$\boxed{\frac{dx_i}{dt} = F_i(t, x_1, \dots, x_n), \quad i = 1, \dots, n}, \tag{62}$$

where $F_i(t, x_1, \dots, x_n)$, $i = 1, \dots, n$, are real-valued functions of real variables $t, x_1, \dots, x_n$. These are among the most important objects studied in analysis and geometry. We usually think of t as time, and x_i, $i = 1, \dots, n$, are called **state variables**. The evolution of system (62) through time is called a **flow**.

Many higher-order systems appearing in practice reduce to first-order system (62) by a trick we explain with an example. Consider the system

$$\begin{cases} y'' = F(t, y, y', z, z', z'') \\ z''' = G(t, y, y', z, z', z'') \end{cases},$$

where F and G are real-valued functions of six real variables, and primes denote differentiation with respect to *time* variable t. Introducing new variables

$$x_1 := y, \quad x_2 := y', \quad x_3 := z, \quad x_4 := z', \quad x_5 := z'',$$

our second-order system reduces to the first-order one:

$$\begin{cases} x_1' = x_2 \\ x_2' = F(t, x_1, x_2, x_3, x_4, x_5) \\ x_3' = x_4 \\ x_4' = x_5, \\ x_5' = G(t, x_1, x_2, x_3, x_4, x_5) \end{cases}.$$

Exercise 68. *(a) Find a first-order system of the form (62) describing linear ODE (35).*

(b) Reduce the governing equations of N-body problem into a system of first-order equations. To save life denote the distance between i-th and j-th bodies

$$\left((x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2\right)^{\frac{1}{2}},$$

by r_{ij}.

(c) Can you find a first-order system of the form (62) describing system (67) below?

Besides from Section 8.1, in this chapter we concentrate on linear ODE systems:

$$\boxed{\frac{dx_i}{dt} = P_{i1}(t)x_1 + \cdots + P_{in}(t)x_n + R_i(t), \quad i = 1, \ldots, n}, \tag{63}$$

where P_{ij} and R_i are real-valued functions of t ranging on some open interval I. Some aspects of nonlinear equation are studied in the next chapter. If all $R_i \equiv 0$, then (63) is called **homogeneous**. The corresponding homogeneous equation to (63) is

$$\frac{dx_i}{dt} = P_{i1}(t)x_1 + \cdots + P_{in}(t)x_n, \quad i = 1, \ldots, n. \tag{64}$$

Life is saved denoting linear system (63) in matrix notation

$$\boxed{\frac{dx}{dt} = P(t)x + R(t)}, \tag{65}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad P(t) = \begin{bmatrix} P_{11}(t) & P_{12}(t) & \ldots & P_{1n}(t) \\ P_{21}(t) & P_{22}(t) & & P_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1}(t) & P_{n2}(t) & \ldots & P_{nn}(t) \end{bmatrix}, \quad R = \begin{bmatrix} R_1(t) \\ R_2(t) \\ \vdots \\ R_n(t) \end{bmatrix}.$$

Note that here x and R are vector-valued functions of t, and P a matrix-valued one. Corresponding homogeneous system is then

$$\frac{dx}{dt} = P(t)x. \tag{66}$$

Example 49. Let us solve the system

$$\begin{cases} x'' + x + y' - 2y = t \\ x'' - x + y' + y = 3 \end{cases}. \tag{67}$$

Representing the differential operator $\frac{d}{dt}$ by D, our system now looks like the *algebraic* one:

$$\begin{cases} (D^2+1)x + (D-2)y = t \\ (D^2-1)x + (D+1)y = 3 \end{cases},$$

and we could use Cramer's rule to deduce

$$\begin{vmatrix} D^2+1 & D-2 \\ D^2-1 & D+1 \end{vmatrix} x = \begin{vmatrix} t & D-2 \\ 3 & D+1 \end{vmatrix}, \quad \begin{vmatrix} D^2+1 & D-2 \\ D^2-1 & D+1 \end{vmatrix} y = \begin{vmatrix} D^2+1 & t \\ D^2-1 & 3 \end{vmatrix},$$

or equivalently

$$\left(3D^2+2D-1\right)x=t+7, \quad \left(3D^2+2D-1\right)y=t+3.$$

Applying methods of Chapter 6, we get

$$\begin{cases} x=C_1e^{\frac{t}{3}}+C_2e^{-t}-t-9 \\ y=C_3e^{\frac{t}{3}}+C_4e^{-t}-t-5 \end{cases}. \tag{68}$$

Are we done? So far we have shown that solutions of (67) are of the form (68), and it remains to answer whether for all choices of constants C_1, C_2, C_3, C_4, functions in (68) satisfy (67). Plugging (68) into (67), we have

$$\begin{cases} \left(\frac{10}{9}C_1-\frac{5}{3}C_3\right)e^{\frac{t}{3}}+(2C_2-3C_4)e^{-t}=0 \\ \left(\frac{-8}{9}C_1+\frac{4}{3}C_3\right)e^{\frac{t}{3}}=0 \end{cases}, \tag{69}$$

which gives

$$C_3=\frac{2}{3}C_1, \quad C_4=\frac{2}{3}C_2.$$

Therefore in matrix notations

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} C_1e^{\frac{t}{3}}+C_2e^{-t}-t-9 \\ \frac{2}{3}C_1e^{\frac{t}{3}}+\frac{2}{3}e^{-t}-t-5 \end{bmatrix}=C_1e^{\frac{1}{3}t}\begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}+C_2e^{-t}\begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}-\begin{bmatrix} t+9 \\ t+5 \end{bmatrix}, \tag{70}$$

is our general solution. ■

We wanted to follow a systematic method in previous example. One might be able to solve special systems with easier methods, as shown in the following exercise.

Exercise 69. *Solve system (67) in the following way. Subtract two equation to get*

$$y=\frac{2}{3}(2x-t+3). \quad (\dagger)$$

In first equation, replace y with right hand side of $(\dagger)$ to get a second-order ODE satisfied by x, which gives the general solution for x, with two constants. Then find y with $(\dagger)$.

8.1 The Existence and Uniqueness of Solutions

Here comes the most fundamental theorem in whole course, the generalization of Theorem 6.

Theorem 13 (Picard-Lindelöf-Peano). *Let $F_i(t, x_1, \ldots, x_n)$, $i = 1, \ldots, n$, be real-valued functions defined on an open box U containing the point $(a, b_1, \ldots, b_n)$.*

(a) If all F_i are continuous on U, then the initial value problem

$$\frac{dx_i}{dt} = F_i(t, x_1, \ldots, x_n), \quad x_i(a) = b_i, \quad i = 1, \ldots, n,$$

has a local solution around $t = a$, namely there are functions $x_i(t)$, $i = 1, \ldots, n$, on some open interval I containing a such that

$$\frac{dx_i}{dt} = F_i(t, x_1(t), \ldots, x_n(t)), \quad x_i(a) = b_i, \quad i = 1, \ldots, n, \tag{71}$$

for each $t \in I$.

(b) If all F_i and $\frac{\partial F_i}{\partial x_j}$ are continuous on U, then (71) has a unique local solution around $t = a$, namely any two n-tuple of solutions match on some interval around $t = a$.

(c) If all F_i are analytic then the system (62) has only analytic solutions.

Part (a) is proved in most general references for differential equations or even real analysis books, say [23, chapter 13], [1, vol. II, Theorem 7.19]. Part (b) is proved in [17, p. 8-11] or [4, chapter 6]. Part (c) is proved in [11, vol. I, 10.5.3].

Exercise 70. *How many constants does the general solution of the system*

$$\begin{cases} x^{(5)} + e^t x^{(3)} + x - y^{(4)} - y + z^{(5)} + 8z = t^3 \sin t \\ e^{-t} x^{(4)} + y^{(5)} - 7z = t + 5 \\ x^{(3)} y + \arctan\left(x y^{(3)}\right) - z^{(6)} = 3t^2 \end{cases},$$

contain?

8.2 Linear Systems

Theorems 8 and 9 find the following generalization in the context of systems of equations.

Theorem 14. *In (63), let $P_{ij}(t)$, $i, j = 1, \ldots, n$, and $R(t)$ be continuous scalar-valued functions of real variable t ranging on open interval I. For any $t_0 \in I$, and any scalars $A_1, \ldots, A_n$, there are unique scalar-valued functions $x_i(t)$, $i = 1, \ldots, n$, defined on whole $t \in I$ satisfying*

$$\frac{dx_i(t)}{dt} = P_{i1}(t)x_1(t) + \cdots + P_{in}(t)x_n(t) + R_i(t), \quad x_i(t_0) = A_i, \quad i = 1, \ldots, n,$$

for each $t \in I$.

This is proved in [23, chapter 13] or [4, chapter 6].

Theorem 15. *Let* x^i*,* $i = 1, \dots, n$*, be solutions of homogeneous system (64). (Note that each* x^i *is a vector-valued function of real variable* t *ranging on some open interval.)*

(a) If x^i*,* $i = 1, \dots, n$*, are* ***(linearly) independent***[21]*, then*

$$\boxed{x = C_1 x^1 + \cdots + C_n x^n},$$

is the general solution of (64). This general solution could be concisely written as

$$\boxed{x = \Psi C},$$

where $\Psi := \left[x^1, \dots, x^n\right]$ *is the* ***fundamental matrix*** *(for each* $i = 1, \dots, n$*, its* i*-th column is* x^i*), and* $C = [C_1, \dots, C_n]^t$ *is a vector of constants.*

(b) x^i*,* $i = 1, \dots, n$*, are independent if and only if their* ***Wronkian***

$$W(x^1, \dots, x^n) = \det\left[x^1, \dots, x^n\right],$$

is nowhere-zero, if and only if $W(x^1, \dots, x^n)$ *is nonzero at least in one point.*

(c) If x^i*,* $i = 1, \dots, n$*, are independent, and* u *is a particular solution of inhomogeneous system (63), then*

$$\boxed{x = C_1 x^1 + \cdots + C_n x^n + u},$$

is the general solution of (63).

This is proved with arguments similar Theorem 9, and reveals the structure of general solutions to linear systems (63).

8.3 The Method of Variation of Parameters

Following the idea of Section 6.5, one can find a particular solution of inhomogeneous system (65) having at hand the general solution $x = \Psi C$ of the corresponding homogeneous system (66). We try to find vector-valued function $u(t) = \left[u^1(t), \dots, u^n(t)\right]^t$ such that $x = \Psi u$ satisfies (65), namely

$$\Psi' u + \Psi u' = P \Psi u + R,$$

or equivalently $\Psi u' = R$, which has solution $u = \int \Psi^{-1} R$, hence

$$\boxed{x = \Psi \int \Psi^{-1} R dt},$$

satisfies (65).

[21] Namely there exists no constants C_i, $i = 1, \dots, n$, at least one nonzero, such that $\sum C_i x^i \equiv 0$.

8.4 Linear First-Order Constant-Coefficient Systems

In this section we solve constant-coefficient linear system of ODEs:

$$\frac{dx_i}{dt} = a_{i1}x_1 + \cdots + a_{in}x_n + g_i(t), \quad i = 1, \ldots, n, \tag{72}$$

where a_{ij}, $i, j = 1, \ldots, n$, are constant, g_i, $i = 1, \ldots, n$, are scalar-valued function of t. We save life denoting this in matrix notation

$$\boxed{\frac{dx}{dt} = Ax + g}, \tag{73}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, \quad g = \begin{bmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{bmatrix}.$$

We give different methods to solve this, and apply them all to the following simple example, so that you can compare methods.

$$\begin{cases} \frac{dx}{dt} = 5x + 8y + t \\ \frac{dy}{dt} = -2x - 3y + 1 \\ x(0) = 16, \quad y(0) = -10 \end{cases}. \tag{74}$$

8.4.1 Elimination

Computing $y = \frac{1}{8}(x' - 5x - t)$ from the first equation, and plugging into the second, we get:

$$x'' - 2x' + x = 3t + 9,$$

which is solved as:

$$x = C_1e^t + C_2te^t + 3t + 15,$$

using the method of undetermined coefficients. Then:

$$y = \frac{1}{8}(x' - 5x - t) = \frac{-4C_1 + C_2}{8}e^t - \frac{C_2}{2}te^t - 2t - 9.$$

Enforcing $x(0) = 16$ and $y(0) = -10$, we get $C_1 = 1$ and $C_2 = -4$. In matrix notation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 - 4t \\ -1 + 2t \end{bmatrix} e^t + \begin{bmatrix} 3t + 15 \\ -2t - 9 \end{bmatrix}.$$

8.4.2 Using Laplace Transform and Matrix Exponentials

Applying Laplace transform componentwisely to equation (73), we get

$$sX - x_0 = AX + G,$$

where $x_0 = x(0)$ is the initial value of x, and $X(s)$ and $G(s)$ are Laplace transforms of $x(t)$ and $g(t)$, respectively. Equivalently

$$(sI - A)X = x_0 + G,$$

which is solved as

$$X = (sI - A)^{-1}(x_0 + G),$$

hence

$$\boxed{x(t) = \mathcal{L}^{-1}\left((sI - A)^{-1}(x_0 + G)\right)}, \tag{75}$$

solves (73).

Let us apply this formula to system (74). We have

$$X = (sI - A)^{-1}(x_0 + G) = \begin{bmatrix} s-5 & -8 \\ 2 & s+3 \end{bmatrix}^{-1} \begin{bmatrix} 16 + \frac{1}{s^2} \\ -10 + \frac{1}{s} \end{bmatrix} = \frac{1}{(s-1)^2} \begin{bmatrix} s+3 & 8 \\ -2 & s-5 \end{bmatrix} \begin{bmatrix} \frac{16s^2+1}{s^2} \\ \frac{-10s+1}{s} \end{bmatrix} =$$
$$= \frac{1}{(s-1)^2 s^2} \begin{bmatrix} 16s^3 - 32s^2 + 9s + 3 \\ -10s^3 + 19s^2 - 5s - 2 \end{bmatrix} = \begin{bmatrix} \frac{15}{s} + \frac{3}{s^2} + \frac{1}{s-1} + \frac{-4}{(s-1)^2} \\ \frac{-9}{s} + \frac{-2}{s^2} + \frac{-1}{s-1} + \frac{2}{(s-1)^2} \end{bmatrix}.$$

Therefore

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 15 + 3t + e^t - 4te^t \\ -9 - 2t - e^t + 2te^t \end{bmatrix} = \begin{bmatrix} 1 - 4t \\ -1 + 2t \end{bmatrix} e^t + \begin{bmatrix} 3t + 15 \\ -2t - 9 \end{bmatrix}.$$

To find the time version of (75), recall matrix exponential function e^{At} of Section 1.5. Since

$$\frac{d}{dt}\left(e^{At}\right) = \frac{d}{dt}\left(I + At + A^2\frac{t^2}{2!} + A^3\frac{t^3}{3!} + \cdots\right) = A + A^2 t + A^3\frac{t^2}{2!} + \cdots =$$
$$= A\left(I + At + A^2\frac{t^2}{2!} + \cdots\right) = Ae^{At},$$

the unique solution of

$$\frac{dx}{dt} = Ax, \quad x(0) = x_0,$$

is $x = e^{At}x_0$. Comparing with (75) shows that

$$\boxed{e^{At} = \mathcal{L}^{-1}\left((sI - A)^{-1}\right)}.$$

Now we can derive the time version of (75):

$$\boxed{x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}g(\tau)d\tau}, \tag{76}$$

which exactly looks like scalar case (24)! This is maybe the most beautiful formula in this course.

Exercise 71. *In equation (73) let the characterstic polynomial of A be*

$$\det(tI - A) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0.$$

Prove that each entry of x_i of x satisfies the following differential equation

$$\left(\frac{d^n}{dt^n} + c_{n-1}\frac{d^{n-1}}{dt^{n-1}} + \cdots + c_1\frac{d}{dt} + c_0\right)x_i = 0.$$

8.4.3 Using Linear Algebra

In this section we give three alternative methods to solve (73). The first one is conditional, and only applicable when matrix A is **diagonalizable** in the sense that there exists an invertible matrix T with $T^{-1}AT$ diagonal. The second and third methods are are always applicable. These methods are based on corrsponding theorems of linear algebra which we now state.

Theorem 16. *An square matrix is diagonizable if and only it has a basis of eigenvectors. In more detail, if $T^{-1}AT = D$ is diagonal, then each k-th column of T is an eigenvector of A corresponding to eigenvalue D_{kk}; and conversely, if A has a basis on eigenvectors, then putting these eigenvectors into columns of the matrix T, and putting their corresponding eigenvalues into diagonal entries of the diagonal matrix D, gives $A = PDP^{-1}$.*

Theorem 17 (Schur)**.** *For every square matrix A there is an invertible matrix T such that $T^{-1}AT$ is upper triangular.*

Theorem 18 (Jordan)**.** *For every square matrix A there is an invertible matrix T such that $T^{-1}AT$ is block diagonal, where each block, called a **Jordan block**, has equal entries in main diagonal, and ones in the first diagonal above the main diagonal, and zeros elsewhere.*

Jordan decomposition is certainly more special that Schur decomposition. Algorithms for them is beyond this course.

Using Theorem 16. If A is diagonalizable as $A = TDT^{-1}$, then changing variable $y = T^{-1}x$, our new equation becomes

$$Ty' = TDy + g,$$

or equivalently

$$y' = Dy + h, \quad h = T^{-1}g,$$

which is a collection of n first-order constant coefficient equations. Here is an example. The matrix

$$\begin{bmatrix} 5 & 8 \\ -2 & -3 \end{bmatrix}$$

of system (74) is not diagonalizable (Exercise 73), so instead consider the following system:

$$\begin{cases} \frac{dx_1}{dt} = x_1 + 2x_2 + t \\ \frac{dy}{dt} = 2x_1 + x_2 + 1 \\ x_1(0) = -1, \quad x_2(0) = 2 \end{cases}. \tag{77}$$

The matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},$$

has eigenvalues -1 and 3 with correspondng eigenvectors $[1, -1]^t$ and $[1, 1]^t$, respectively, hence

$$A = TDT^{-1}, \quad T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}.$$

Change of variables

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y = T^{-1}x = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

our new system is

$$\begin{cases} y_1' = -y_1 + \frac{t-1}{2} \\ y_2' = 3y_2 + \frac{t+1}{2} \\ y_1(0) = -\frac{3}{2}, \quad y_2(0) = \frac{1}{2} \end{cases}.$$

Using Theorem 17. If A is upper triangulated by $A = TUT^{-1}$, then changing variable $y = T^{-1}x$, exactly as before, our new equation becomes

$$y' = Uy + h, \quad h = T^{-1}g,$$

which reversely look like

$$y_n' = U_{nn}y_n + h_n, \quad y_{n-1}' = U_{n-1,n-1}y_{n-1} + U_{n-1,n}y_n + h_{n-1}, \quad \dots,$$

which are recurvely solved as

$$y_n = \frac{1}{D - U_{nn}} h_n = e^{U_{nn}t} \frac{1}{D} e^{-U_{nn}t} h_n,$$

$$y_{n-1} = \frac{1}{D - u_{n-1,n-1}} \left(u_{n-1,n} y_n + h_{n-1} \right) = e^{u_{n-1,n-1} t} \frac{1}{D} e^{-u_{n-1,n-1} t} \left(u_{n-1,n} y_n + h_{n-1} \right),$$

etc.

Using Theorem 18. Proceed exactly as the previous upper triangular case.

Remark 9 (Optional)**.** Formula (76) shows that solving (73) is equivalent to computing matrix exponential. This is done analytically using Jordan decomposition in [9, Theorem 3.10]. ■

Exercise 72. *(a) Diagonalize the matrix in Exercise 14. (b) Which of the matrices in Example 3 are diagonalizable?*

Exercise 73. *(a) Show that for each scalar a, the matrix $\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}$ is not diagonalizable. (b) Show that the matrix $\begin{bmatrix} 5 & 8 \\ -2 & -3 \end{bmatrix}$ is not diagonalizable.*

8.4.4 An Explicit Fundamental System of Solutions for Equation $x' = Ax$

Putting Structure Theorem 15 and the method of variation of parameter (Section 8.3) together, to solve (73), it suffices to find n independent solutions of homogeneous equation

$$\frac{dx}{dt} = Ax. \tag{78}$$

But how does these independent solutions look like? Let us gather ideas by an example.

Example 50. Let us solve the system

$$\begin{cases} \frac{dx}{dt} = x + y \\ \frac{dy}{dt} = 4x - 2y \end{cases},$$

by elimination method. From the first equation we find $y = x' - x$, and plugging this into the second gives

$$x'' + x' - 6x = 0,$$

which has general solution

$$x = C_1 e^{-3t} + C_2 e^{2t},$$

therefore

$$y = x' - x = -3C_1 e^{-3t} + 2C_2 e^{2t} - C_1 e^{-3t} - C_2 e^{2t} = -C_1 e^{-3t} + C_2 e^{2t}.$$

Therefore in matrix notation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} C_1 e^{-3t} + C_2 e^{2t} \\ -4C_1 e^{-3t} + C_2 e^{2t} \end{bmatrix} = C_1 e^{-3t} \begin{bmatrix} 1 \\ -4 \end{bmatrix} + C_2 e^{2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{79}$$

is our general solution. ■

So we guess solutions of (78) might look like

$$x = e^{\lambda t} u,$$

where λ is a scalar, and u a constant *nonzero* vector. (We neglect $u = 0$, because we are looking for a fundamental system.) Let us check this guess. The function $x = e^{\lambda t} u$ satisfies $x' = Ax$ exactly when

$$\lambda e^{\lambda t} u = A e^{\lambda t} u,$$

or equivalently

$$Au = \lambda u,$$

namely exactly when λ is an eigenvalue of A with corresponding eigenvector u. To continue our plan of finding a fundamental system of solution, let A has eigenvalues λ_k, $k = 1, \dots, n$, counted with multiplicity, with corresponding eigenvectors u^k, $k = 1, \dots, n$. Then our analysis above shows that all $e^{\lambda_k t} u^k$, $k = 1, \dots, n$, are solutions. But are they independent? To answer this, we compute Wronskian:

$$W\left(e^{\lambda_1 t} u^1, \dots, e^{\lambda_n t} u^n\right) = e^{(\lambda_1 + \cdots + \lambda_n)t} \det\left[u^1, \dots, u^n\right].$$

Thus we win if we could find a basis of eigenvectors for A, and from Section 1.4.4, we know that this is possible exactly when A is diagonalizable.

For example, when all roots of the characteristic polynomial of A are simple, then by Proposition 2, corresponding vectors $u^1, \dots, u^n$ are a basis, hence

$$e^{\lambda_1 t} u^1, \quad \dots, \quad e^{\lambda_n t} u^n,$$

is a fundamental system of solutions of $x' = Ax$. The main problem is when the characteristic equation has repeated roots, and for some repeated eigenvalues we can not extract independent eigenvector to the number of their multiplicity. For example, let λ be an eigenvalue of multiplicity 2, which has only one independent eigenvector u. Let us think of multiplicity 2 case as limit degenerate case of two simple eigenvalues λ and $\lambda + h$ where h is very small. We could also assume that eigenvector corresponding to $\lambda + h$ is $u + v$ where v is also negligible. From our previous analysis we know that both $e^{\lambda t} u$ and

$$e^{(\lambda+h)t}(u + v) = e^{\lambda t} e^{ht}(u + v) \approx e^{\lambda t}(1 + ht)(u + v) \approx e^{\lambda t}(u + v + htu),$$

are solutions of $x' = Ax$, hence also their difference divided by scalar h, namely

$$e^{\lambda t}\left(\frac{v}{h} + tu\right),$$

which is of the form

$$e^{\lambda t}\left(w + tu\right).$$

For higher multiplicities, similar arguments gives solutions of the form

$$e^{\lambda t}\left(w_0 + tw_1 + \frac{t^2}{2!}w_2 + \cdots + \frac{t^m}{m!}w_m\right).$$

We have the following structure theorem, proved in [9, Theorem 3.7].

Theorem 19. *Let A be a square matrix of order n. Let λ_i, $i = 1, \ldots, k$, be all distinct eigenvalues of A, with corresponding multiplicities m_i, $i = 1, \ldots, k$. Note that $\sum m_i = n$. By a linear algebra theorem*[22]*, for each i, the* ***generalized eigenvalue space corresponding to*** *λ_i*

$$\mathcal{E}_i := \{v \in \mathbb{C}^n : (A - \lambda_i)^{m_i} v = 0\},$$

is of dimension m_i, so let u_{ij}, $j = 1, \ldots, m_i$, be a basis for it. Then the collection

$$\left\{e^{\lambda_i t} p_{ij} : i = 1, \ldots, k, \quad j = 1, \ldots, m_i\right\},$$

where

$$\boxed{p_{ij} = u_{ij} + t(A - \lambda_i)u_{ij} + \frac{t^2}{2!}(A - \lambda_i)^2 u_{ij} + \cdots + \frac{t^{m_i - 1}}{(m_i - 1)!}(A - \lambda_i)^{m_i - 1} u_{ij}},$$

is a fundamental system of solutions for (78).

Example 51. Let

$$A = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix}.$$

Then A has eigenvalue $\lambda = -2$ of multiplicity 3. Since $(A + 2)^3 = 0$, the generalized eigenvalue space is whole $\mathbb{C}^3$, with basis say

$$u_1 = [1, 0, 0]^t, \quad u_2 = [0, 1, 0]^t, \quad u_1 = [0, 0, 1]^t.$$

[22] [18, p. 237].

So a fundamental system for $x' = Ax$ consists of

$$e^{-2t}\left(u_1 + t(A+2)u_1 + \frac{t^2}{2}(A+2)^2u_1\right) = e^{-2t}\begin{bmatrix}1\\0\\0\end{bmatrix},$$

$$e^{-2t}\left(u_2 + t(A+2)u_2 + \frac{t^2}{2}(A+2)^2u_2\right) = e^{-2t}\begin{bmatrix}t\\1\\0\end{bmatrix},$$

$$e^{-2t}\left(u_3 + t(A+2)u_3 + \frac{t^2}{2}(A+2)^2u_3\right) = e^{-2t}\begin{bmatrix}\frac{t^2}{2}\\t\\1\end{bmatrix}.$$

∎

9 Systems of Ordinary Differential Equations II: Nonlinear Equations

We follow [23, chapter 11].

10 First Touches on Partial Differential Equations

In this chapter we discuss the first most important analytic/numerical/qualitative methods to analyze some special classes of PDEs. More specifically, we study heat, wave and Laplace equations which respectively describe diffusion, oscillation and steady-state phenomena in physics. The crucial issue of existence, uniqueness and regularity (namely smoothness) of solutions to PDEs is beyond the scope of this course, and the interested reader could refer to [6] or [24].

10.1 Linear First-Order Partial Differential Equations

To be added. ???

10.2 Heat Equation in Infinite Rod

Recall that the heat propagation in a long rod

$$\begin{cases} \text{PDE}: & u_t = u_{xx}, \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{IC}: & u(x,0) = f(x), \quad -\infty < x < \infty \end{cases},$$

was solved in Exercise 52 by convolution integral

$$u(x,t) = \int_{\mathbb{R}} G(x-\xi,t) f(\xi) d\xi, \quad G(x,t) = \frac{1}{\sqrt{4\pi}} t^{-\frac{1}{2}} e^{-\frac{x^2}{4t}}. \tag{80}$$

We now use this solution to solve the following problem

$$\begin{cases} \text{PDE}: & u_t = u_{xx} + h(x,t), \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{IC}: & u(x,0) = 0, \quad -\infty < x < \infty \end{cases}. \tag{81}$$

The idea is to represent $h(x,t)$, for each x, as linear combinations of unit impulses

$$h(x,t) = \int_0^\infty h(x,\tau)\delta(t-\tau) d\tau.$$

By linearity, if $u^\tau(x,t)$, for each $\tau \geqslant 0$, solves

$$\begin{cases} \text{PDE}: & u_t = u_{xx} + h(x,\tau)\delta(t-\tau), \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{IC}: & u(x,0) = 0, \quad -\infty < x < \infty \end{cases}. \tag{82}$$

then

$$u(x,t) = \int_0^\infty u^\tau(x,t) d\tau,$$

solves (81). For $0 \leqslant t < \tau$, $u^{\tau}(x,t)$ solves

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

hence $u^{\tau}(x,t) = 0$. Specially $u^{\tau}(x,\tau-) = 0$. Integrating (82) from $t = \tau-$ to $t = \tau+$, we get $u^{\tau}(x,\tau+) = h(x,\tau)$. Therefore for $t > \tau$, $u^{\tau}(x,t)$ solves

$$u_t = u_{xx}, \quad u(x,\tau+) = h(x,\tau),$$

hence, by Exercise 52, $u^{\tau}(x,t) = \int_{\mathbb{R}} G(x-\xi, t-\tau)h(\xi,\tau)d\xi$. The whole analysis shows that

$$u(x,t) = \int_0^{\infty} u^{\tau}(x,t)d\tau = \int_0^t \left(\int_{\mathbb{R}} G(x-\xi, t-\tau)h(\xi,\tau)d\xi \right) d\tau,$$

or equivalently

$$u(x,t) = \int_0^t \int_{\mathbb{R}} G(x-\xi, t-\tau)h(\xi,\tau)d\xi d\tau, \quad G(x,t) = \frac{1}{\sqrt{4\pi}} t^{-\frac{1}{2}} e^{-\frac{x^2}{4t}}, \tag{83}$$

solves (81). This is another instance of **Duhamel's Principle**.

Finally, adding (80) and (95), we have:

Theorem 20. *The solution of*

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx} + h(x,t), \quad 0 < t < \infty, \ -\infty < x < \infty \\ \text{IC}: & u(x,0) = f(x), \ -\infty < x < \infty \end{cases},$$

is given by

$$\boxed{u(x,t) = \int_{\mathbb{R}} G(x-\xi, t)f(\xi)d\xi + \int_0^t \int_{\mathbb{R}} G(x-\xi, t-\tau)h(\xi,\tau)d\xi d\tau},$$

where

$$\boxed{G(x,t) = \frac{1}{\sqrt{4\pi}} t^{-\frac{1}{2}} e^{-\frac{x^2}{4t}}},$$

is called the ***heat kernel of*** $\mathbb{R}$ *or the* ***fundamental solution of one-dimensional heat equation****.*

10.3 Heat Equation in Finite Rod

We now consider the heat propagation in a finite rod in the form of the problem posed at the end of Example 18:

$$\begin{cases} \text{PDE}: & u_t = u_{xx}, \quad 0 < t < \infty, \quad 0 < x < \pi \\ \text{BCs}: & u(0,t) = u(\pi,t) = 0, \quad 0 \leqslant t < \infty \\ \text{IC}: & u(x,0) = 1, \quad 0 < x < \pi \end{cases}. \tag{84}$$

10.3.1 Numerical Solution

First we give a numerical solution. Recall from calculus the following formulas for the first and second derivative of a scalar-valued function $f(z)$ of a real variable:

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}, \quad f''(a) = \lim_{h\to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

Therefore, to approximate $f'(a)$ and $f''(a)$, using the values of f at point a, its right neighbor $a+h$, and left neighbor $a-h$, with h small, we could use

$$f'(a) \approx \frac{f(a+h) - f(a)}{h}, \quad f''(a) \approx \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

We return back to our problem. Fixing a large natural number N, and a small time step δt, gives us grid points

$$(x_i, t_j), \quad i = 0, \ldots, N, \quad j = 0, 1, \ldots,$$

where

$$x_i = i\delta x, \quad \delta x = \frac{\pi}{N}, \quad t_j = j\delta t.$$

Setting

$$u_{i,j} := u(x_i, t_j),$$

our discretized heat equation will be

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\delta x^2},$$

which we write as

$$u_{i,j+1} = u_{i,j} + \alpha \left(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}\right), \qquad i = 1, \ldots, N-1, \quad j = 0, 1, \ldots, \tag{85}$$

with constant $\alpha = \frac{\delta t}{\delta x^2}$. From boundary and initial conditions, we know the values

$$u_{0,j}, \quad u_{N,j}, \quad u_{i,0}, \quad i = 0, \ldots, N, \quad j = 0, 1, \ldots. \tag{86}$$

It is clear one can use (85) and (86), to iteratively find

$$u_{1,1}, \quad u_{2,1}, \quad \ldots, \quad u_{N-1,1}, u_{1,2}, \quad u_{2,2}, \quad \ldots, \quad u_{N-1,2}, \quad \text{etc.}$$

10.3.2 Analytic Solution

We set out to give an anlytic solution for problem (84). This will take three steps.

Step I: Finding Modes. Observe that both PDE and BCs (but not IC) are *linear*, in the sense that if several functions satisfy them then any linear combination of them also satisfy both. The idea is then to find enough nonzero functions satisfying both PDE and BCs (theses are called **modes**), and then try to make a linear combination of modes satisfy IC. Let us see where this idea leads us.

We start looking for *separable* modes, namely nonzero functions of the form $u(x,t) = X(x)T(t)$ satisfying both PDE and BCs. This functions satisfying PDE means $XT' = X''T$, or equivalently

$$\frac{T'}{T} = \frac{X''}{X}.$$

Since the left hand side of this equation is a function of t, and the right hand side a function of x, so both equal a constant say λ. Since we want u satisfy BCs, we require X satisfy $X(0) = X(\pi) = 0$. Three cases appear.

Case $\lambda > 0$. Then $X = Ae^{\sqrt{\lambda}x} + Be^{\sqrt{-\lambda}x}$. Enforcing $X(0) = X(\pi) = 0$, as you can easily check, makes both A and B zero, hence no modes exists in this case.

Case $\lambda = 0$. Then $X = A + Bx$, and again no modes exists in this case.

Case $\lambda < 0$. Then

$$X = A\sin\left(\sqrt{\lambda}x\right) + BA\cos\left(\sqrt{\lambda}x\right).$$

This functions satisfies $X(0) = X(\pi) = 0$, for $B = 0$ and $\lambda = n$ integer.

Separable modes we found are

$$u_n(x,t) = e^{-n^2 t}\sin(nx), \quad n \in \mathbb{Z},$$

therefore based on what we have discussed in the first paragraph of this example, any linear combination

$$u(x,t) = \sum_{n\in\mathbb{Z}} C_n e^{-n^2 t}\sin(nx),$$

satisfies PDE and BCs, and we hope to find constants C_n such that it also satisfies IC. Here is a minor simplification. Since the right of our latter equation could be written as

$$(C_1 - C_{-1})e^{-t}\sin x + (C_2 - C_{-2})e^{-4t}\sin 2x + (C_3 - C_{-3})e^{-9t}\sin 3x + \cdots,$$

we have equivalent linear combination

$$u(x,t) = \sum_{n\geqslant 1} D_n e^{-n^2 t}\sin(nx), \tag{87}$$

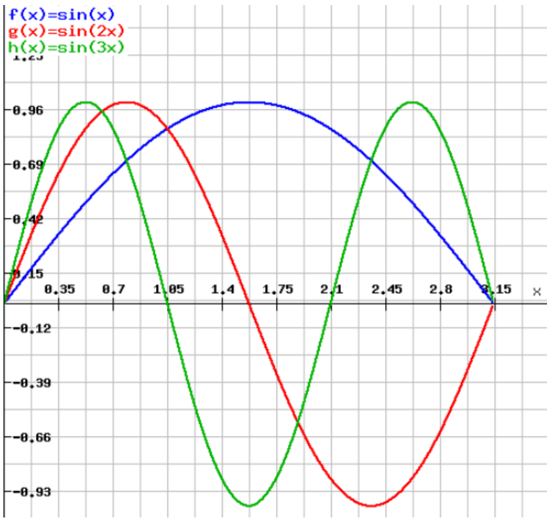
and it only remains to find constants D_n such that IC is fulfilled:

$$1 = \sum_{n \geqslant 1} D_n \sin(nx), \quad \text{for all } x \in (0,1). \tag{88}$$

Step II: Finding the Contributions of Different Modes Namely Coefficients D_n *in (88).* This seems strange at first, namely expressing constant function 1 as linear combination of *harmonics*

$$\sin nx, \quad n = 1, 2, \ldots,$$

shown in the following figure.



It seems like these harmonics are *unrelated* or *orthogonal* in the sense that the integral of their product on $[0, \pi]$ is zero:

$$\int_0^\pi \sin(nx)\sin(mx)\,dx = 0, \quad m \neq n.$$

This is confirmed by the following computation:

$$\int_0^\pi \sin(nx)\sin(mx)\,dx = \int_0^\pi (\cos((n-m)x) + \cos((n+m)x))\,dx =$$

$$= \left[\frac{\sin((n-m)x)}{n-m} + \frac{\sin((n+m)x)}{n+m}\right]_0^\pi = 0.$$

To use this observation, we multiply both sides of (88) by $\sin mx$, and integrating on $[0, \pi]$, hence to get

$$\int_0^\pi \sin(mx)\,dx = D_m \int_0^\pi \sin^2(mx)\,dx,$$

hence

$$D_m = \frac{\int_0^\pi \sin(mx)\,dx}{\int_0^\pi \sin^2(mx)\,dx} = \frac{\left[\frac{\cos(mx)}{m}\right]_0^\pi}{\int_0^\pi \frac{1-\cos(2mx)}{2}\,dx} = \frac{\begin{cases} \frac{2}{m}, & m \text{ odd} \\ 0, & m \text{ even} \end{cases}}{\left[\frac{x}{2} - \frac{\sin(2mx)}{4m}\right]_0^\pi} = \begin{cases} \frac{4}{m\pi}, & m \text{ odd} \\ 0, & m \text{ even} \end{cases}.$$

What we have done in this step is that, if (88) holds then D_n better be given by

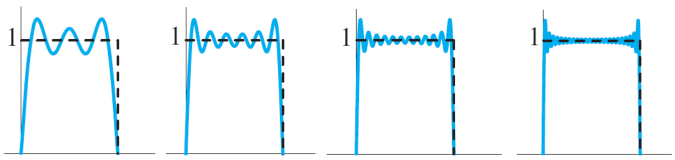
$$D_n = \begin{cases} \frac{4}{n\pi}, & n \text{ odd} \\ 0, & n \text{ even} \end{cases}.$$

We are left with the question of the validity of epresentation

$$1 = \frac{4}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots\right), \quad \text{for all } x \in (0, 1), \tag{89}$$

namely does the infinite series on the right hand side converges to constant function 1 on for each $0 < x < \pi$?

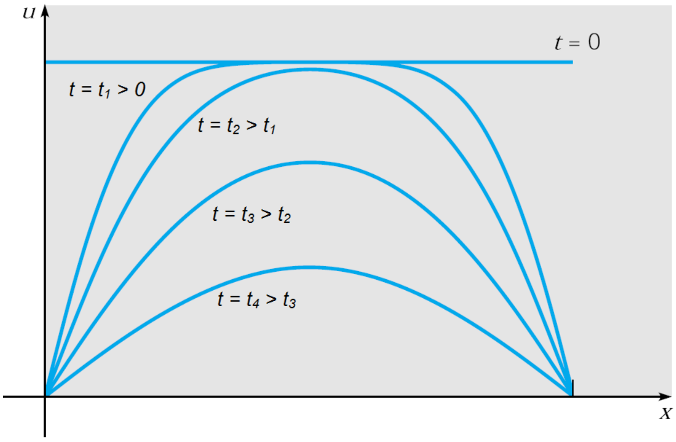
Step III: Validity of Representation (89). From the first time this question was asked, it took mathematics a long while to answer it as YES! This is a technical issue beyond this course, but discussed in optional Section 10.8. As mentioned there, as far as practical engineering applications are concerned, there is no problem in representations like (89). The following figure shows how good the partial sums of the series of the right hand side of (89) approximates constant function 1, where form left to right we have used partial sums with respectively 3, 6, 12 and 24 terms.



These steps overally show that

$$u(x, t) = \frac{4}{\pi}\left(e^{-t}\sin x + \frac{1}{3}e^{-9t}\sin 3x + \frac{1}{5}e^{-25t}\sin 5x + \cdots\right),$$

solves our problem. The following figure shows $u(x, t)$ for several different fixed times.



Remark 10 (Optional)**.** A very important feature of heat equation is that with whatever rough IC you start with, the solution becomes *immediately* smooth. This turn out to be very important in pure mathematics, and gives rise to *heat equation method* in geometric analysis. The same *smoothing* phenomenon completely fails for wave equation, where the singularities of initial conditions generally propagate through time. Existence and uniqueness results for heat and wave equations could be found in [6, chapter 10] and [24, section 6.1]. ■

Exercise 74 (An Analogue Integrator)**.** *Consider the following system describing temperature distribution in a rod of finite length* π *with lateral surfaces and two ends completely insulated.*

$$\begin{cases} \text{PDE}: & u_t = u_{xx}, \quad 0 < t < \infty, \quad 0 < x < \pi \\ \text{BCs}: & u_x(0,t) = u_x(\pi,t) = 0, \quad 0 \leqslant t < \infty\,, \\ \text{IC}: & u(x,0) = f(x), \quad 0 < x < \pi \end{cases}$$

where initial distribution $f(x)$ *is known. Intuition says that eternal distribution is* $u(x,\infty) = C$*, a constant function. (a) Justify the boundary conditions. (b) Prove that intuition is true. (c) Find* C*. [Hint: Modes are* $e^{-n^2t}\cos(nx)$*,* $n = 0,1,2,\ldots$*. Answer of (c) is* $\frac{1}{\pi}\int_0^{\pi} f(x)dx$*.]*

10.4 Fourier Series

Series appeared in (88) is an special case of **Fourier series**, which we now explain. Let $f(x)$ be scalar-valued function of real variable x, which is T**-periodic**, namely there is some positive constant T, called **period**, such that $f(x+T) = f(x)$ for all x. Set $\omega_0 = \frac{2\pi}{T}$. The idea of Fourier series is to represent $f(x)$ by a linear combination of the most fundamental T-periodic functions:

$$1, \quad \cos(n\omega_0 x)\,, \quad \sin(n\omega_0 x)\,, \quad n = 1,2,\ldots,$$

namely an infinite series representation

$$\boxed{f(x) = \frac{1}{2}a_0 + \sum_{n \geqslant 0} a_n \cos(n\omega_0 x) + b_n \sin(n\omega_0 x)}.$$

To find constant a_n and b_n we use orthogonality relations

$$\int_I \cos(m\omega_0 x)\sin(n\omega_0 x)dx = 0,$$

$$\int_I \cos(m\omega_0 x)\cos(n\omega_0 x)dx = \int_I \sin(m\omega_0 x)\sin(n\omega_0 x)dx = \begin{cases} \frac{T}{2} & m = n \\ 0 & m \neq n \end{cases},$$

where m and n are positive integers, and I is any interval of length T. Following exactly as before, we get

$$\boxed{a_n = \frac{2}{T}\int_I f(x)\cos(k\omega_0 x)dx, \quad n = 0, 1, 2, \dots},$$

$$\boxed{b_n = \frac{2}{T}\int_I f(x)\sin(k\omega_0 x)dx, \quad n = 1, 2, 3, \dots}.$$

Exercise 75. *Express sawtooth* 2*-periodic wave, given by*

$$\begin{cases} x+1, & -1 \leqslant x \leqslant 0 \\ -x+1, & 0 \leqslant x \leqslant 1 \end{cases},$$

in cycle $[-1, 1]$*, by its Fourier series.*

10.5 Wave Equation in Infinite String

A long string satisfies the wave equation together with initial conditions describing its initial position and velocity:

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx}, \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{ICs}: & u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad -\infty < x < \infty \end{cases}. \tag{90}$$

To solve this problem, we consider the change of independent variables (x, t) to (α, β) with $\alpha = x - t$ and $\beta = x + t$. Chain rule for partial differentiation gives

$$u_t = u_\alpha \alpha_t + u_\beta \beta_t = -u_\alpha + u_\beta,$$

$$u_{tt} = \frac{\partial}{\partial t}(-u_\alpha + u_\beta) = (-u_\alpha + u_\beta)_\alpha \alpha_t + (-u_\alpha + u_\beta)_\beta \beta_t = u_{\alpha\alpha} - 2u_{\alpha\beta} + u_{\beta\beta},$$

and similarly $u_{xx} = u_{\alpha\alpha} + 2u_{\alpha\beta} + u_{\beta\beta}$. Therefore $u_{tt} = u_{xx}$ becomes $u_{\alpha\beta} = 0$, which is easily solved as

$$u(x,t) = F(\alpha) + G(\beta) = F(x-t) + G(x+t), \tag{91}$$

where $F(z)$ and $G(z)$ are arbitrary smooth real-valued functions of real variable z. Imposing initial conditions, we get $F(x) + G(x) = f(x)$ and $-F'(x) + G'(x) = g(x)$. Integrating the second equation, and solving for F and G, we get

$$F(z) = \frac{1}{2}\left(f(z) - \int_0^z g(\zeta)d\zeta - C\right), \quad G(z) = \frac{1}{2}\left(f(z) + \int_0^z g(\zeta)d\zeta + C\right).$$

where C is a constant. Therefore

$$u(x,t) = \frac{1}{2}\left(f(x-t) - \int_0^{x-t} g(\zeta)d\zeta + C_1\right) + \frac{1}{2}\left(f(x+t) + \int_0^{x+t} g(\zeta)d\zeta - C_1\right),$$

or equivalently

$$\boxed{u(x,t) = \frac{1}{2}(f(x-t) + f(x+t)) + \frac{1}{2}\int_{x-t}^{x+t} g(\xi)d\xi}. \tag{92}$$

Note that the graph of $f(x-t)$ as a function of x for some fixed t is the the graph of $f(x)$ shifted t units to the right. Therefore one could think of $f(x-t)$ as a **traveling wave** to the right, namely positive x. Similarly, $f(x+t)$ represents a wave traveling to the left.

Next we use (92) to solve

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx} + h(x,t), \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{IC}: & u(x,0) = u_t(x,0) = 0, \quad -\infty < x < \infty \end{cases}. \tag{93}$$

The idea is to represent $h(x,t)$, for each x, as linear combinations of unit impulses

$$h(x,t) = \int_0^\infty h(x,\tau)\delta(t-\tau)d\tau.$$

By linearity, if $u^\tau(x,t)$, for each $\tau \geqslant 0$, solves

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx} + h(x,\tau)\delta(t-\tau), \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{IC}: & u(x,0) = u_t(x,0) = 0, \quad -\infty < x < \infty \end{cases}, \tag{94}$$

then

$$u(x,t) = \int_0^\infty u^\tau(x,t)d\tau,$$

solves (93). For $0 \leqslant t < \tau$, $u^\tau(x,t)$ solves

$$u_{tt} = u_{xx}, \quad u(x,0) = u_t(x,0) = 0,$$

hence $u^\tau(x,t) = 0$. Specially $u^\tau(x,\tau-) = \frac{\partial}{\partial t}u^\tau(x,\tau-) = 0$. Firstly, note that $u^\tau(x,t)$ does not jump at $t = \tau$, because otherwise $\delta''(t-\tau)$ appears in $\frac{\partial^2}{\partial t^2}u^\tau(x,t)$; however the PDE in (94) witnesses that this is not the case. Therefore $u^\tau(x,\tau+) = 0$. Secondly, integrating the PDE in (94) from $t = \tau-$ to $t = \tau+$, we get $\frac{\partial}{\partial t}u^\tau(x,\tau+) = h(x,\tau)$. Therefore for $t > \tau$, $u^\tau(x,t)$ solves

$$u_{tt} = u_{xx}, \quad u(x,\tau+) = 0, \quad \frac{\partial}{\partial t}u^\tau(x,\tau+) = h(x,\tau),$$

hence, by (92)

$$u^\tau(x,t) = \frac{1}{2}\int_{x-(t-\tau)}^{x+(t-\tau)} h(\xi,\tau)\,d\xi.$$

The whole analysis shows that

$$u(x,t) = \int_0^\infty u^\tau(x,t)\,d\tau = \int_0^t \left(\frac{1}{2}\int_{x-(t-\tau)}^{x+(t-\tau)} h(\xi,\tau)\,d\xi\right) d\tau,$$

or equivalently

$$u(x,t) = \frac{1}{2}\int_0^t \int_{x-(t-\tau)}^{x+(t-\tau)} h(\xi,\tau)\,d\xi\,d\tau, \tag{95}$$

solves (93). This is another instance of **Duhamel's Principle**.

Finally, adding (92) and (95), we have:

Theorem 21. *The solution of*

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx} + h(x,t), \quad 0 < t < \infty, \quad -\infty < x < \infty \\ \text{ICs}: & u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad -\infty < x < \infty \end{cases},$$

is given by

$$\boxed{u(x,t) = \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(\xi)\,d\xi + \frac{1}{2}\int_0^t \int_{x-(t-\tau)}^{x+(t-\tau)} h(\xi,\tau)\,d\xi\,d\tau}.$$

10.6 Wave Equation in Finite String

Let us now study the dynamics of finite length string with fixed ends via its mathematical model

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx}, \quad 0 < t < \infty, \quad 0 < x < L \\ \text{BCs}: & u(0,t) = u(L,t) = 0, \quad 0 \leqslant t < \infty \\ \text{ICs}: & u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad 0 < x < L \end{cases}. \tag{96}$$

Either using formula (91), or the method of separation of variables and Fourier series of Section 10.3.2, we get the following solution.

Theorem 22. *In problem (96), first make* f *and* g *odd on* $[-L, L]$*, and then make them* $2L$*-periodic, namely extend* f *and* g *to the functions* f^* *and* g^* *satisfying*

$$f^*(z) = -f^*(-z) = f^*(z+2L), \quad g^*(z) = -g^*(-z) = g^*(z+2L).$$

Then the solution to problem (96) is given by

$$u(x,t) = \frac{1}{2}\left(f^*(x-t) + f^*(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g^*(\xi)d\xi.$$

Example 52. Figure 7 draws the solution $u(x,t)$ to problem (96) at some selected times, for $f(z)$ a triangular pulse, and $g(z) \equiv 0$. Note that $f(z)$ is not smooth at $z = \frac{L}{2}$, and this singularity *propagates* through time, and the wave equation, in contrast with heat equation, has no smoothing effect on the initial data. ■

Exercise 76. *Consider the following problem which describes wave propagation in a semi-infinite string with one end fixed.*

$$\begin{cases} \text{PDE}: & u_{tt} = u_{xx}, \quad 0 < t < \infty, \quad 0 < x < \infty \\ \text{BCs}: & u(0,t) = f(t), \quad u(\infty,t) = 0, \quad 0 \leqslant t < \infty \\ \text{ICs}: & u(x,0) = u_t(x,0) = 0 \quad 0 < x < \infty \end{cases}. \tag{97}$$

(a) Solve this problem using formula (91). (b) Solve this problem applying Laplace transform with respect to t*. [Answer.* $u(x,t) = f(t-x)H(t-x)$*, where* $H(z)$ *is the unit step function.]*

10.7 Laplace Equation

Laplace equation, in dimension two, is

$$u_{xx} + u_{yy} = 0, \tag{98}$$

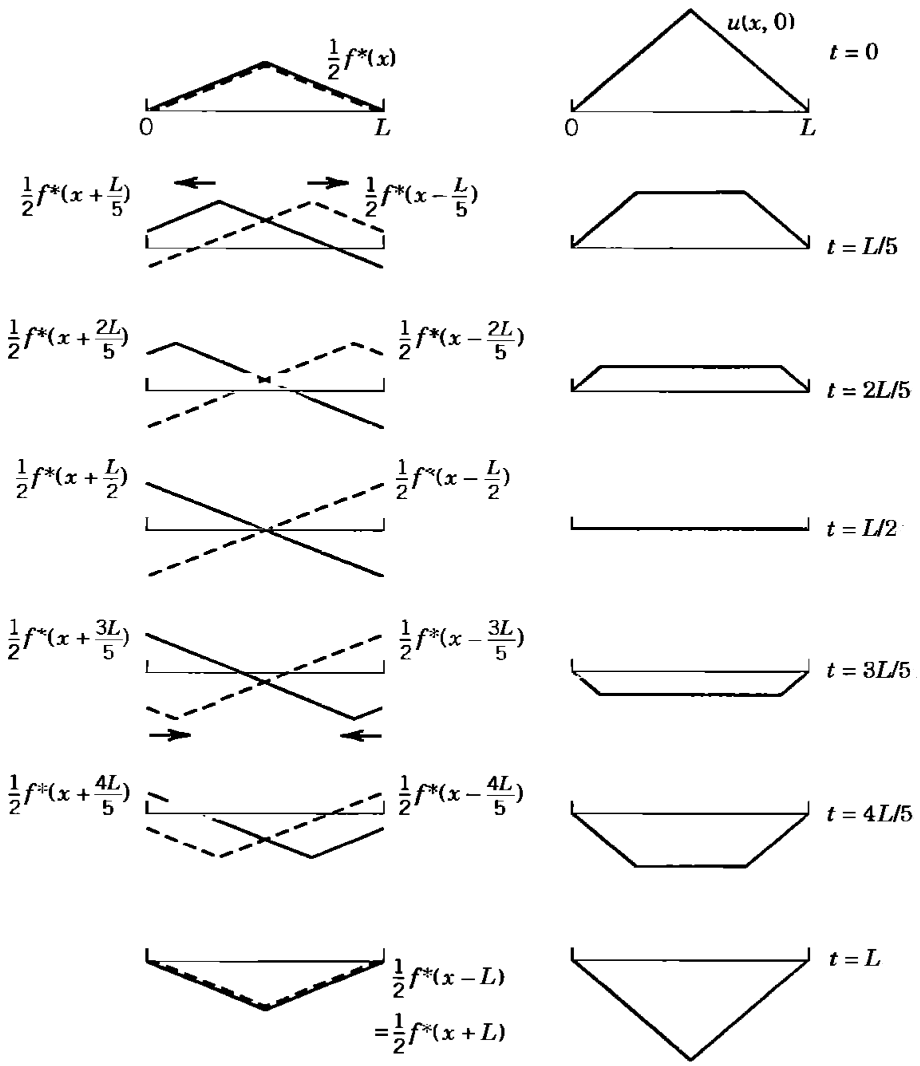


Figure 7: The solution to problem (96), for $f(z)$ a triangular pulse, and $g(z) \equiv 0$. Taken from [20, p. 546].

where u is a scalar-valued function of real variables x and y. Such u is called a **harmonic function**. Steady-state temperature distribution, electric potential in charge-free space, etc. are examples of harmonic functions.

There are some simple geometric intuitions behind (98), formulated in the form of *mean value theorem* and *maximum (or minimum) principle*, which we now discuss. If any of the second derivatives u_{xx} or u_{yy} has a positive sign, the other must be negative. Geometrically, this means if u has positive curvature in any coordinate direction, it must have negative curvature in the other direction. Therefore all the critical points of u (points where $u_x = u_y = 0$) must be saddle points, not maxima or minima. Therefore the extrema of a harmonic function must be on the boundary of the region.

Recall that the Dirichlet problem for upper-half plane

$$\begin{cases} \text{PDE}: & u_{xx} + u_{yy} = 0, \quad -\infty < x < \infty, \quad 0 < y < \infty \\ \text{BC}: & u(x,0) = f(x), \quad -\infty < x < \infty \end{cases}, \tag{99}$$

was solved in Exercise 53 by convolution integral

$$u(x,y) = \int_{-\infty}^{\infty} \frac{1}{\pi} \frac{y}{(x-\xi)^2 + y^2} f(\xi)\, d\xi. \tag{100}$$

This solution is not unique, since adding a constant times y or $3x^2y - y^3$ is agian a solution, however if f is bounded, then (100) is the unique bounded solution of (99). There is an elegent probabilistic interpretation of (100) which we now describe. ???

10.8 The Convergence of Fourier Series (Optional)

Let $f(x)$ be real-valued function of real variable x defined on some interval $I = [a,b]$. Set

$$T = b - a, \quad \omega_0 = \frac{2\pi}{T},$$

and

$$S_n(x) = \frac{1}{2}a_0 + \sum_{1 \leqslant k \leqslant n} a_k \cos(k\omega_0 x) + b_k \sin(k\omega_0 x), \quad n = 0, 1, 2, \ldots,$$

where

$$a_k = \frac{2}{T}\int_I f(x)\cos(k\omega_0 x)\,dx, \quad b_k = \frac{2}{T}\int_I f(x)\sin(k\omega_0 x)\,dx, \quad k = 0, 1, 2, \ldots,$$

and

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \cdots + S_n(x)}{n+1}, \quad n = 0, 1, 2, \ldots,$$

and

$$F(x) = \frac{f(x+) + f(x-)}{2}.$$

Theorem 23. *With notations as above, We have:*

(a) If f is piecewise continuously differentiable on I[23]*, or more generally of bounded variation on I*[24]*, then $S_n(x) \to F(x)$ as $n \to \infty$. (**Dirichlet-Jordan Theorem** [2, 11.12])*

*(b) If $\int_I |f|$ exists in Lebesgue sense, then $\sigma_n(x) \to F(x)$ for each $x \in I$ where $F(x)$ exists. (**Fejér Theorem** [2, 11.15])*

*(c) If $\int_I |f|^2$ exists in Lebesgue sense, then $S_n(x) \to f(x)$ for almost every $x \in I$. (**Carleson Theorem** [15, 3.6.14])*

*(d) If $\int_I |f|^2$ exists in Lebesgue sense, then $\int_I |f - S_n|^2 \to 0$ as $n \to \infty$. (**Riesz-Fischer Theorem** [14, 8.20], [2, 11.16])*

Note that for each famous summability method, namely, *pointwise, Cesaro, almost everywhere, and square mean (or L^2)*, we brought a convergence result in latter theorem. Easy part (d) suffices for all engineering applications, because physical systems respond to the energy of the signals, and from this perspective the function and its Fourier series are indistinguishable. Carleson Theorem is surely among the deepest result in analysis.

[23] Namely, f' is continuous on I, except for finitely many points $x_1, \ldots, x_N$, and at each x_j we have $f'_-(x_j) = \lim_{x \to x_j -} f'(x)$ and $f'_+(x_j) = \lim_{x \to x_j +} f'(x)$.

[24] Namely there exists finite real M such that $\sum_{1 \leqslant k \leqslant N} |f(x_k) - f(x_{k-1})| < M$ for each $a = x_0 < x_1 < \cdots < x_N = b$.

11 Sturm-Liouville Eigenvalue Problem

Recall that in Example 10.3.2, during the the process of solving one-dimensional heat equation by the method of separation of variables, the *eigenvalue problem*

$$\frac{d^2X}{dx^2} = \lambda X, \quad X(0) = X(\pi) = 0,$$

showed up. A common general framework for many such eigenvalue problems appearing in solving second-order linear PDEs is **Sturm-Liouville Boundary Value Problem**:

$$\boxed{\begin{cases} \text{DE}: & \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y + \lambda r(x)y = 0, \quad a < x < b \\ \text{BCs}: & \alpha_1 y(a) + \alpha_2 y'(a) = 0, \quad \beta_1 y(b) + \beta_2 y'(b) = 0 \end{cases}}, \tag{101}$$

where $p(x)$, $q(x)$ and $(x)r$ are real-valued functions of real variable x ranging on (a, b), α_1 and α_2 are real constants not both zero, and β_1 and β_2 are also real constants not both zero.

Theorem 24. *Consider Sturm-Liouville problem (101) with* ***regularity assumptions****:*

1. *a and b are finite;*
2. *p, p', q and r are continuous on $[a, b]$;*
3. *p and q are strictly positive on $[a, b]$.*

For any two functions f and g on $[a, b]$, define ***inner-product*** *and* ***norm*** *according to*

$$\boxed{\langle f, g\rangle := \int_a^b r(x)f(x)g(x)\,dx, \quad \|f\| := \sqrt{\langle f, f\rangle}},$$

in case the integrals exist. Then:

1. *Eigenvalues are real and countably infinite, say ordered as $\lambda_0 < \lambda_1 < \lambda_2 < \cdots$. Also $\lambda_n \to \infty$ as $n \to \infty$.*
2. *Corresponding to each eigenvalue λ_n is a unique (up to a nonzero multiplicative constant) eigenfunction y_n. Also y_n has exactly n zeros in (a, b).*
3. *Eigenvalues corresponding to different eigenvalue are* ***orthogonal*** *in the sense that $\langle y_m, y_n\rangle = 0$ whenever $m \neq n$.*

4. *Each function* f *on* $[a,b]$ *with* $\|f\| < \infty$*, has* ***eigenvalue expansion***

$$\boxed{f = \sum_{n=0}^{\infty} \frac{\langle f, y_n \rangle}{\langle y_n, y_n \rangle} y_n},$$

in the sense that

$$\lim_{N\to\infty} \left\| f - \sum_{n=0}^{N} \frac{\langle f, y_n \rangle}{\langle y_n, y_n \rangle} y_n \right\| = 0.$$

5. *If* f *is piecewise continuously differentiable on* $[a,b]$[25]*, or more generally of bounded variation on* I[26]*, then*

$$\sum_{n=0}^{\infty} \frac{\langle f, y_n \rangle}{\langle y_n, y_n \rangle} y_n(x) = \frac{f(x+) + f(x-)}{2},$$

for $a < x < b$.

For proof refer to [25, Theorem 5.11] or [4, chapters 10-11]. Last part is proved in [26, vol. I, Theorem 1.9].

Remark 11. In Sturm-Liouville Problem, if interval $[a,b]$ is unbounded, or if p or q vanish at endpoints a or b, then there might be a continuum of eigenvalues. The analysis of these problems, which very much appear in practice, are ways more harder than regular Sturm-Liouville Problem studied in Theorem 24. Refer to [27, Theorem 1012.1] or [26]. ■

Example 53. Let us study eigenvalue problem

$$\begin{cases} y'' + \lambda y = 0, & 0 < x < 1 \\ y'(0) = y(0), & y(1) = 0 \end{cases}.$$

By Theorem 24, we know that all eigenvalues of real, so we consider three cases.
Case I: $\lambda = \alpha^2$, $\alpha > 0$. Applying boundary conditions to the general solution

$$y = C_1 \cos \alpha x + C_2 \sin \alpha x,$$

gives

$$C_2 \alpha = C_1, \quad C_1 \cos \alpha + C_2 \sin \alpha = 0,$$

[25]Namely, f' is continuous on $[a,b]$, except for finitely many points $x_1, \ldots, x_N$, and at these points, left and right derivatives exist and equals corresponding limit of $f'(t)$.
[26]Namely there exists finite real M such that $\sum_{1 \leqslant k \leqslant N} |f(x_k) - f(x_{k-1})| < M$ for each $a = x_0 < x_1 < \cdots < x_N = b$.

which gives nonzero solution exactly when

$$\tan \alpha = -\alpha. \tag{102}$$

Intersecting the graph of functions $\tan \alpha$ and $-\alpha$ we get countably infinite roots

$$\alpha_0 \approx 2.0288 < \alpha_1 \approx 4.9132 < \alpha_2 \approx 7.9787 < \cdots,$$

tending to infinity.

Case II: $\lambda = 0$. *Case III:* $\lambda = -\alpha^2$, $\alpha > 0$. We leave it an exercise that in these two cases no eigenvalue exists.

Thus our problem has countably infinite eigenvalues Our whole
and corresponding eigenfunctions

$$\lambda_0 = \alpha_0^2 < \lambda_1 = \alpha_1^2 < \lambda_2 = \alpha_2^2 < \cdots,$$

and corresponding eigenfunctions

$$y_n = \alpha_n \cos \alpha_n x + \sin \alpha_n x, \quad n = 0, 1, 2, \ldots.$$

■

Example 54. Let us try to solve eigenvalue problem

$$\begin{cases} x^2 y'' - \lambda(xy' - y) = 0, & 1 < x < 2 \\ y(1) = 0, \quad y(2) - y'(2) = 0 \end{cases},$$

which is not of Sturm-Liouville type, so we better not expect assertions of Theorem 24 hold.

The equation is equidimesional studied in Section 6.10. The change of variables $z = \log x$ gives new constant coefficient equation

$$(D_z(D_z - 1) - \lambda D_z + \lambda)\, y = 0,$$

or equivalently

$$\left(D_z^2 - (\lambda + 1) D_z + \lambda\right) y = 0.$$

Characteristic roots are

$$r = \frac{\lambda + 1 \pm \sqrt{(\lambda+1)^2 - 4\lambda}}{2} = \frac{\lambda + 1 \pm \sqrt{(\lambda - 1)^2}}{2} = 1 \text{ or } \lambda,$$

so we split into two cases.

Case I: $\lambda = 1$. General solution

$$y = C_1 e^t + C_2 t e^t = C_1 x + C_2 x \log x,$$

should satisfy boundary conditions, hence

$$C_1 = 0, \quad 2C_1 + C_2 2 \log 2 = C_1 + C_2(\log 2 + 1),$$

which leads to $C_1 = C_2 = 0$. Therefore we have no eigenfunction in this case.

Case II: $\lambda \neq 1$. General solution

$$y = C_1 e^t + C_2 e^{\lambda t} = C_1 x + C_2 x^{\lambda},$$

should satisfy boundary conditions

$$C_1 + C_2 = 0, \quad C_1 + C_2 2^{\lambda - 1} (2 - \lambda) = 0,$$

which has nontrivial solution exactly when

$$2^{1-\lambda} = 2 - \lambda. \tag{103}$$

The *real* solutions of this equation are x-coordinates of intersection points of plane curves

$$y = 2^{x-1}, \quad y = \frac{1}{2 - x},$$

which are $(0, 2^{-1})$ and $(1, 1)$, hence we find roots $\lambda = 0$ and $\lambda = 1$. We can only accept $\lambda = 0$, since we are assuming $\lambda \neq 1$. The corresponding eigenfunction is

$$y = C_1 x + C_2 x^{\lambda} = C_1 x - C_1 = C_1(x - 1).$$

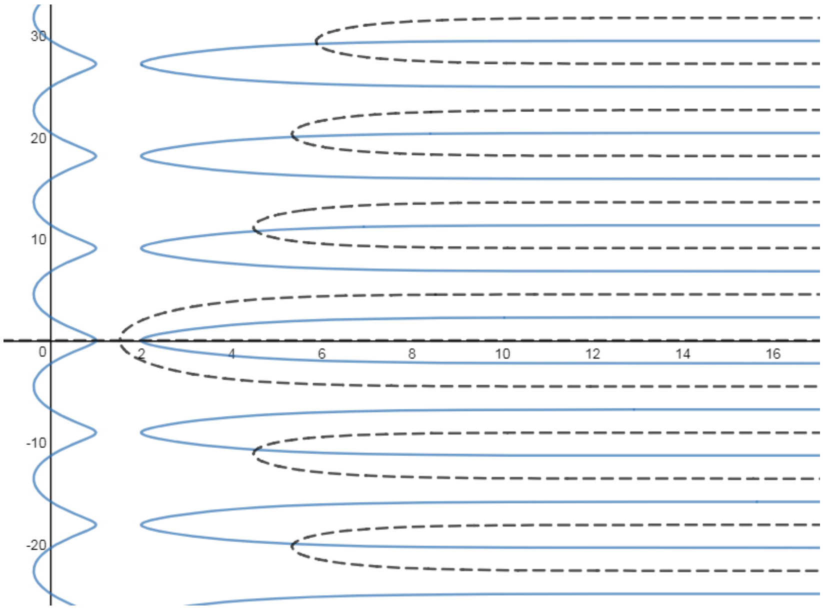
To find *complex* roots of (103), assuming $2 - \lambda = x + iy$, with real x and y, equation (103) becomes

$$x + iy = 2^{x-1+iy} = e^{(x-1+iy)\log 2} = 2^{x-1}(\cos(y \log 2) + i \sin(y \log 2)),$$

or equivalently

$$x = 2^{x-1} \cos(y \log 2), \quad y = 2^{x-1} \sin(y \log 2).$$

The following figure shows these two curves in xy-plane. It seems like we have infinitely many complex eigenvalues.



■

Exercise 77. *Consider eigenvalue problem*

$$\begin{cases} y'' + y' + \lambda(y' + y) = 0, & 0 < x < 1 \\ y'(0) = 0, \quad y(1) = 0 \end{cases}.$$

(a) Determine all real eigenvalues.

(b) (Optional) Are there any complex eigenvalues? Use a plotter say https://www.desmos.com/calculator.

[Answer. (a) Nothing. (b) Infinitely many.]

References

[1] Apostol, T. M., *Calculus*, Second Edition, Volumes I-II, John Wiley and Sons, 1967. 9, 45, 46, 50, 67, 88, 89, 100

[2] Apostol, T. M., *Mathematical Analysis*, Second Edition, Addison-Wesley, 1974. 124

[3] Aubin, T. , *A Course in Differential Geometry*, American Mathematical Society, 2000. 54

[4] Birkhoff, G. and G.-C. Rota, *Ordinary Differential Equations*, Fourth Edition, John Wiley and Sons, 1989. 45, 67, 92, 95, 100, 101, 126

[5] Boyce, W. E. and R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, Seventh Edition, John Wiley and Sons, 2001. 1

[6] Brezis, H., *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010. 111, 117

[7] Churchill, R. V., *Operational Calculus*, Third Edition, McGraw-Hill, 1972. 66

[8] Coddington, E. A., *An Introduction to Ordinary Differential Equations*, Dover Publications, 1961. 92, 94

[9] Coddington, E. A. and R. Carleson, *Linear Ordinary Differential Equations*, SIAM, 1997. 106, 108

[10] Coddington, E. A. and N. Levinson, *Theory of Ordinary Differential Equations*, Fourth Edition, McGraw-Hill, 1955. 92

[11] Dieudonne, J., *Treatise on Analysis*, Volumes I-XII, Academic Press, 1969-93. 89, 100

[12] Doetsch, G., *Introduction to the Theory and Application of the Laplace Tansform*, Springer, 1974. 66

[13] Edwards, C. H., D. E. Penney and D. T. Calvis, *Differential Equations and Boundary Value Problems: Computing and Modeling*, Fifth Edition, Pearson Education Inc., 2015. 1, 59, 79, 81

[14] Folland, G. B., *Real Analysis: Modern Techniques and Their Applications*, Second Edition, John Wiley and Sons, 1999. 124

[15] Grafakos, L., *Classical Fourier Analysis*, Second Edition, Springer, 2008. 124

[16] Hardy, G. H., *The Integration of Functions of a Single Real Variable*, Second Edition, Cambridge University Press, 1916. 9

[17] Hartman, P., *Ordinary Differential Equations*, Second Edition, SIAM, 2002. 45, 100

[18] Hoffman, K. and R. Kunze, *Linear Algebra*, Second edition, Prentice-Hall, 1971. 108

[19] Krantz, S. G. and H. R. Parks, *A Primer of Real Analytic Function*, Second Edition, Birkhäuser, 2002. 87

[20] Kreyszig, E., *Advanced Engineering Mathematics*, Ninth Edition, John Wiley and Sons, 2006. 122

[21] Oppenheim, A. V., A. S. Willksy and H. Nawab, *Signals and Systems*, Second Edition, Prentice Hall, 1997. 56, 58, 63

[22] Ramo, S., J. H. Whinnery and T. Van Duzer, *Fields and Waves in Communication Electronics*, Second Edition, John Wiley and Sons, 1984. 30

[23] Simmons, G. F., *Differential Equations with Applications and Historical Notes*, Second Edition, McGraw-Hill, 1991. 1, 25, 34, 45, 62, 67, 89, 100, 101, 110

[24] Taylor, M. E., *Partial Differential Equations*, Second Edition, Volumes I-III, Springer, 2011. 40, 54, 89, 111, 117

[25] Teschl, G., *Ordinary Differential Equations and Dynamical Systems*, American Mathematical Society, 2011. 96, 126

[26] Titchmarch, E. C., *Eigenfunction Expansions*, Second Edition, Volumes I-II, Oxford University Press, 1962. 126

[27] Zettl, A., *Sturm-Liouville Theory*, American Mathematical Society, 2000. 126