MA3220 Ordinary Differential Equations

Wong Yan Loi

November 20, 2007

Contents

1 First Order Differential Equations **5**
1.1 Introduction 5
1.2 Exact Equations, Integrating Factors 8
1.3 First Order Linear Equations 10
1.4 First Order Implicit Equations 13

2 Higher Order Linear Equations **17**
2.1 General Theory 17
2.2 Linear Equations with Constant Coefficients 22
2.3 Operator methods 29
2.4 Exact 2nd order Equations 32
2.5 The adjoint differential equation and integrating factor 32

3 Linear Differential Systems **35**
3.1 Linear Systems 35
3.2 Homogeneous Linear Systems 36
3.3 Non-Homogeneous Linear Systems 39
3.4 Homogeneous Linear Systems with Constant Coefficients 40
3.5 Higher Order Linear Equations 50
3.6 Appendix 1: Proof of Lemma 3.12 54

4 Power Series Solutions **57**
4.1 Power Series 57
4.2 Series Solutions of First Order Equations 60
4.3 Second Order Linear Equations and Ordinary Points 61
4.4 Regular singular points and the method of Frobenius 65
4.9 Bessel's equation 68

5 Fundamental Theory of ODEs **77**
5.1 Existence-Uniqueness Theorem 77
5.2 The method of successive approximations 78
5.3 Convergence of the successive approximations 81
5.4 Non-local Existence of Solutions 84

5.5 Gronwall's Inequality and Uniqueness of Solution 87
5.6 Existence and Uniqueness of Solutions to Systems 90

Chapter 1

First Order Differential Equations

1.1 Introduction

1. Ordinary differential equations.

An *ordinary differential equation* (ODE for short) is a relation containing one real variable x, the real dependent variable y, and some of its derivatives y', y'', $\cdots$, $y^{(n)}$, $\cdots$, with respect to x.

The *order* of an ODE is defined to be the order of the highest derivative that occurs in the equation. Thus, an n-th order ODE has the general form

$$F(x, y, y', \cdots, y^{(n)}) = 0. \tag{1.1.1}$$

We shall always assume that (1.1.1) can be solved explicitly for $y^{(n)}$ in terms of the remaining $n+1$ quantities as

$$y^{(n)} = f(x, y, y', \cdots, y^{(n-1)}), \tag{1.1.2}$$

where f is a known function of x, y, y', $\cdots$, $y^{(n-1)}$.

An n-th order ODE is *linear* if it can be written in the form

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_n(x)y = r(x). \tag{1.1.3}$$

The functions $a_j(x)$, $0 \le j \le n$ are called *coefficients* of the equation. We shall always assume that $a_0(x) \not\equiv 0$ in any interval in which the equation is defined. If $r(x) \equiv 0$, (1.1.3) is called a *homogeneous equation*. If $r(x) \not\equiv 0$, (1.1.3) is said to be a *non-homogeneous* equation, and $r(x)$ is called the *non-homogeneous term*.

2. Solutions.

A functional relation between the dependent variable y and the independent variable x that satisfies the given ODE in some interval J is called a *solution* of the given ODE on J.

A *general solution* of an n-th order ODE depends on n arbitrary constants, i.e. the solution y depends on x and n real constants c_1, $\cdots$, c_n.

A first order ODE may be written as

$$F(x, y, y') = 0. \tag{1.1.4}$$

In this chapter we consider only first order ODE. The function $y = \phi(x)$ is called an *explicit solution* of (1.1.4) in the interval J provided

$$F(x, \phi(x), \phi'(x)) = 0 \qquad \text{for all } x \text{ in } J. \tag{1.1.5}$$

A relation of the form $\psi(x, y) = 0$ is said to be an *implicit solution* of (1.1.4) provided it determines one or more functions $y = \phi(x)$ which satisfy (1.1.5). The pair of equations

$$x = x(t), \quad y = y(t) \tag{1.1.6}$$

is said to be a *parametric solution* of (1.1.4) if

$$F(x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)}) = 0.$$

Example. Consider the ODE $x + yy' = 0$ for $x \in (-1, 1)$.
$x^2 + y^2 = 1$ is an implicit solution while $x = \cos t, y = \sin t$, $t \in (0, \pi)$ is a parametric solution.

3. Integral curves.
The solutions of an ODE

$$y' = f(x, y) \tag{1.1.7}$$

represent a one-parameter family of curves in the xy-plane. These are called *integral curves*.

In other words, if $y = y(x)$ is a solution to (1.1.7), then vector field $\mathbf{F}(x, y) = \langle 1, f(x, y)\rangle$ is tangent to the curve $\mathbf{r}(x) = \langle x, y(x)\rangle$ at every point (x, y) since $\mathbf{r}'(x) = \mathbf{F}(x, y)$.

4. Elimination of constants: formation of ODE.

Example. The family of functions $y = Ae^x + B\sin x$ satisfies the ODE: $y'''' - y = 0$ when the constants A and B are eliminated using the derivatives.

5. Separable equations.
Typical separable equation can be written as

$$y' = \frac{f(x)}{g(y)}, \qquad \text{or} \qquad g(y)dy = f(x)dx. \tag{1.1.8}$$

The solution is given by

$$\int g(y)dy = \int f(x)dx + c.$$

Example. Solve $y' = -2xy$, $y(0) = 1$.
Ans: $y = e^{-x^2}$.

The equation $y' = f(\frac{y}{x})$ can be reduced to a separable equation by letting $u = \frac{y}{x}$, i.e. $y = xu$. So $f(u) = y' = u + xu'$,

$$\int \frac{du}{f(u) - u} = \int \frac{dx}{x} + c.$$

Example. Solve $2xyy' + x^2 - y^2 = 0$.

Ans: $x^2 + y^2 = cx$.

6. Homogeneous equations.

A function is called *homogeneous of degree* n if $f(tx, ty) = t^n f(x, y)$ for all x, y, t.
For example $\sqrt{x^2 + y^2}$ and $x + y$ are homogeneous of degree 1, $x^2 + y^2$ is homogeneous of degree 2 and $\sin(x/y)$ is homogeneous of degree 0.
The ODE $M(x, y) + N(x, y)y' = 0$ is said to be *homogeneous of degree* n if both $M(x, y)$ and $N(x, y)$ are homogeneous of degree n.
If we write the above DE as $y' = f(x, y)$, where $f(x, y) = -M(x, y)/N(x, y)$. Then $f(x, y)$ is homogeneous of degree 0. To solve the DE

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

where f is homogeneous of degree 0, we use the substitution $y = zx$. Then

$$\frac{dy}{dx} = z + x\frac{dz}{dx}.$$

Thus the DE becomes

$$z + x\frac{dz}{dx} = f(x, zx) = x^0 f(1, z) = f(1, z).$$

Consequently, the variables can be separated to yield

$$\frac{dz}{f(1, z) - z} = \frac{dx}{x},$$

and integrating both sides will give the solution.

Example. Solve $y' = \frac{x+y}{x-y}$.
Ans: $\tan^{-1}(y/x) = \ln\sqrt{x^2 + y^2} + c$.

Example. An equation in the form

$$y' = \frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}.$$

can be reduced to a homogeneous equation by a suitable substitution $x = z + h, y = w + k$ when $a_1b_2 \neq a_2b_1$, where h and k are solutions of the system of linear equations $a_1h + b_1k + c_1 = 0, a_2h + b_2k + c_2 = 0$.

Example. Solve $y' = \frac{x+y-2}{x-y}$.
Ans: $\tan^{-1}\left(\frac{y-1}{x-1}\right) = \ln\sqrt{(x-1)^2 + (y-1)^2} + c$.

Example. Solve $(x + y + 1) + (2x + 2y + 1)y' = 0$.
Ans: $x + 2y + \ln|x + y| = c, x + y = 0$.

1.2 Exact Equations, Integrating Factors

1. Exact equations.

We can write a first order ODE in the following form

$$M(x,y)dx + N(x,y)dy = 0. \tag{1.2.1}$$

(1.2.1) is called *exact* if there exists a function $u(x,y)$ such that

$$M(x,y)dx + N(x,y)dy = du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy.$$

Once (1.2.1) is exact, the general solution is given by

$$u(x,y) = c.$$

Theorem 1.1 *1.1 Assume M and N together with their first partial derivatives are continuous in the rectangle S: $|x-x_0| < a$, $|y-y_0| < b$. A necessary and sufficient condition for (1.2.1) to be exact is*

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \qquad \text{for all } (x,y) \text{ in } S. \tag{1.2.2}$$

When (1.2.2) is satisfied, a general solution of (1.2.1) is given by $u(x,y) = c$, where

$$u(x,y) = \int_{x_0}^{x} M(s,y)ds + \int_{y_0}^{y} N(x_0,t)dt \tag{1.2.3}$$

and c is an arbitrary constant.

Proof. The first part is by Green's Theorem. For the second part, we have $u(x,y) = \int_{x_0}^{x} M(s,y)ds + \phi(y)$. Substituting $x = x_0$ we have $u(x_0,y) = \phi(y)$. Therefore, $N(x_0,y) = \frac{du(x_0,y)}{dy} = \phi'(y)$ and $\phi(y) = \int_{y_0}^{y} N(x_0,t)dt$.

Remark. In Theorem 1.1, the rectangle S can be replaced by any region which does not include any "hole".

Example. Solve $(x^3 + 3xy^2)dx + (3x^2y + y^3)dy = 0$.

Ans: $x^4 + 6x^2y^2 + y^4 = c$.

2. Integrating factors.

A non-zero function $\mu(x,y)$ is an *integrating factor* of (1.2.1) if the equivalent differential equation

$$\mu(x,y)M(x,y)dx + \mu(x,y)N(x,y)dy = 0 \tag{1.2.4}$$

is exact.

If μ is an integrating factor of (1.2.1) then $(\mu M)_y = (\mu N)_x$, i.e.

$$N\mu_x - M\mu_y = \mu(M_y - N_x). \tag{1.2.5}$$

One may look for an integrating factor of the form $\mu = \mu(v)$, where v is a known function of x and y. Plugging into (1.2.5) we find

$$\frac{1}{\mu}\frac{d\mu}{dv} = \frac{M_y - N_x}{Nv_x - Mv_y}. \tag{1.2.6}$$

If $\frac{M_y - N_x}{Nv_x - Mv_y}$ is a function of v alone, say, $\phi(v)$, then

$$\mu = e^{\int^v \phi(v)dv}$$

is an integrating factor of (1.2.1).

Let $v = x$. If $\frac{M_y - N_x}{N}$ is a function of x alone, say, $\phi_1(x)$, then $e^{\int^x \phi_1(x)dx}$ is an integrating factor of (1.2.1).

Let $v = y$. If $-\frac{M_y - N_x}{M}$ is a function of y alone, say, $\phi_2(y)$, then $e^{\int^y \phi_2(y)dy}$ is an integrating factor of (1.2.1).

Let $v = xy$. If $\frac{M_y - N_x}{yN - xM}$ is a function of $v = xy$ alone, say $\phi_3(xy)$, then $e^{\int^{xy} \phi_3(v)dv}$ is an integrating factor of (1.2.1).

Example Solve $(x^2y + y + 1) + x(1 + x^2)y' = 0$.
Ans: $xy + \tan^{-1} x = c$.

Example Solve $(y - y^2) + xy' = 0$
Ans: $y = (1 - cx)^{-1}$.

Example Solve $(xy^3 + 2x^2y^2 - y^2) + (x^2y^2 + 2x^3y - 2x^2)y' = 0$
Ans: $e^{xy}(1/x + 2/y) = c$.

3. Find integrating factors by inspection.
The following are some differential formulas that are often useful.

$$d(\frac{x}{y}) = \frac{ydx - xdy}{y^2}$$

$$d(xy) = xdy + ydx$$

$$d(x^2 + y^2) = 2xdx + 2ydy$$

$$d(\tan^{-1} \frac{x}{y}) = \frac{ydx - xdy}{x^2 + y^2}$$

$$d(\log \frac{x}{y}) = \frac{ydx - xdy}{xy}$$

We see that the very simple ODE $ydx - xdy = 0$ has $1/x^2, 1/y^2, 1/(x^2+y^2)$ and $1/xy$ as integrating factors.

1.3 First Order Linear Equations

1. Homogeneous equations.
A first order homogeneous linear equation is of the form

$$y' + p(x)y = 0, \tag{1.3.1}$$

where $p(x)$ is a continuous function on an interval J. Let $P(x) = \int_a^x p(s)ds$. Multiplying (1.3.1) by $e^{P(x)}$, we get

$$\frac{d}{dx}[e^{P(x)}y] = 0,$$

so $e^{P(x)}y = c$. The general solution of (1.3.1) is given by

$$\begin{aligned} y(x) &= ce^{-P(x)}, \qquad \text{where} \\ P(x) &= \int_a^x p(s)ds. \end{aligned} \tag{1.3.2}$$

2. Non-homogeneous equations.
Now consider a first order non-homogeneous linear equation

$$y' + p(x)y = q(x), \tag{1.3.3}$$

where $p(x)$ and $q(x)$ are continuous functions on an interval J. Let $P(x) = \int_a^x p(s)ds$. Multiplying (1.3.3) by $e^{P(x)}$ we get

$$\frac{d}{dx}[e^{P(x)}y] = e^{P(x)}q(x).$$

Thus

$$e^{P(x)}y(x) = \int_a^x e^{P(t)}q(t)dt + c.$$

The general solution is given by

$$\begin{aligned} y(x) &= e^{-P(x)}[\int_a^x e^{P(t)}q(t)dt + c], \qquad \text{where} \\ P(x) &= \int_a^x p(s)ds. \end{aligned} \tag{1.3.4}$$

Example. Solve $y' - y = e^{2x}$.
Ans: $y = ce^x + e^{2x}$.

3. The Bernoulli equation.
An ODE in the form

$$y' + p(x)y = q(x)y^n, \tag{1.3.5}$$

where $n \neq 0, 1$, is called the *Bernoulli equation*. The functions $p(x)$ and $q(x)$ are continuous functions on an interval J.
Let $u = y^{1-n}$. Substituting into (1.3.5) we get

$$u' + (1-n)p(x)u = (1-n)q(x). \tag{1.3.6}$$

This is a first order linear ODE.

Example. Solve $xy' + y = x^4y^3$.
Ans: $\frac{1}{y^2} = -x^4 + cx^2$.

4. The Riccati equation.
An ODE of the form

$$y' = P(x) + Q(x)y + R(x)y^2 \tag{1.3.7}$$

is called the *Riccati equation*. The functions $P(x)$, $Q(x)$, $R(x)$ are continuous on an interval J. In general, the Riccati equation cannot be solved by a sequence of integrations. However, if a particular solution is known, then (1.3.7) can be reduced to a linear equation, and thus is solvable.

Theorem 1.2 *Let $y = y_0(x)$ be a particular solution of the Riccati equation (1.3.7). Set*

$$\begin{aligned} H(x) &= \int_{x_0}^{x} [Q(t) + 2R(t)y_0(t)]dt, \\ Z(x) &= e^{-H(x)}\Big[c - \int_{x_0}^{x} e^{H(t)}R(t)dt\Big], \end{aligned} \tag{1.3.8}$$

where c is an arbitrary constant. Then the general solution is given by

$$y = y_0(x) + \frac{1}{Z(x)}. \tag{1.3.9}$$

Proof. In (1.1.7) we let $y = y_0(x) + u(x)$ to get

$$y_0' + u' = P + Q(y_0 + u) + R(y_0 + u)^2.$$

Since y_0 satisfies (1.3.7), we have

$$y_0' = P + Qy_0 + Ry_0^2.$$

From these two equalities we get

$$u' = (Q + 2Ry_0)u + Ru^2. \tag{1.3.10}$$

This is a Bernoulli equation with $n = 2$. Set $Z = u^{-1}$ and reduce (1.3.10) to

$$Z' + (Q + 2Ry_0)Z = -R. \tag{1.3.11}$$

(1.3.11) is a linear equation and the solution is given by (1.3.8). □

Example. Solve $y' = y/x + x^3y^2 - x^5$. Note $y_1 = x$ is a solution.
Ans: $ce^{2x^5/5} = \frac{y-x}{y+x}$.

From (1.3.8), (1.3.9), the general solution y of the Riccati equation (1.3.7) can be written as

$$y = \frac{cF(x) + G(x)}{cf(x) + g(x)}, \tag{1.3.12}$$

where

$$f(x) = e^{-H(x)},$$
$$g(x) = -e^{-H(x)} \int_{x_0}^{x} e^{H(t)} R(t)dt,$$
$$F(x) = y_0(x)f(x), \quad G(x) = y_0 g(x) + 1.$$

Given four distinct functions $p(x)$, $q(x)$, $r(x)$, $s(x)$, we define the cross ratio by

$$\frac{(p-q)(r-s)}{(p-s)(r-q)}.$$

Property 1. The cross ratio of four distinct particular solutions of a Riccati equation is independent of x.

Proof. From (1.3.12), the four solutions can be written as

$$y_j(x) = \frac{c_j F(x) + G(x)}{c_j f(x) + g(x)}.$$

Computations show that

$$\frac{(y_1 - y_2)(y_3 - y_4)}{(y_1 - y_4)(y_3 - y_2)} = \frac{(c_1 - c_2)(c_3 - c_4)}{(c_1 - c_4)(c_3 - c_2)}.$$

The right hand is independent of x. □

As a consequence we get

Property 2. Suppose y_1, y_2, y_3 are three distinct particular solutions of a Riccati equation (1.3.7). Then the general solution is given by

$$\frac{(y_1 - y_2)(y_3 - y)}{(y_1 - y)(y_3 - y_2)} = c, \tag{1.3.13}$$

where c is an arbitrary constant.

Property 3. Suppose that y_1 and y_2 are two distinct particular solutions of a Riccati equation (1.3.7), then its general solution is given by

$$\ln\left|\frac{y - y_1}{y - y_2}\right| = \int [y_1(x) - y_2(x)]R(x)dx + c, \tag{1.3.14}$$

where c is an arbitrary constant.

Proof. y and y_j satisfy (1.3.7). So

$$y' - y_j' = (y - y_j)[Q + R(y + y_j)],$$
$$\frac{y' - y_j'}{y - y_j} = Q + R(y + y_j).$$

Thus

$$\frac{y' - y_1'}{y - y_1} - \frac{y' - y_2'}{y - y_2} = R(y_1 - y_2).$$

Integrating yields (1.3.14). □

Example. Solve $y' = e^{-x}y^2$. Note $y_1 = e^x$ and $y_2 = 0$ are 2 solutions.
Ans: $y = ce^x/(c + e^x)$.

1.4 First Order Implicit Equations

In the above we discussed first order explicit equations, i.e. equations in the form $y' = f(x, y)$. In this section we discuss solution of some first order explicit equations

$$F(x, y, y') = 0 \tag{1.4.1}$$

which are not solvable in y'.

1. Method of differentiation.
Consider an equations solvable in y:

$$y = f(x, y'). \tag{1.4.2}$$

Let $p = y'$. Differentiating $y = f(x, p)$ we get

$$[f_x(x, p) - p]dx + f_p(x, p)dp = 0. \tag{1.4.3}$$

This is a first order explicit equation in x and p. If $p = \phi(x)$ is a solution of (1.4.3), then

$$y = f(x, \phi(x))$$

is a solution of (1.4.2).

Example. Clairaut's equation

$$y = xy' + f(y'), \tag{1.4.4}$$

where f has continuous second order derivative and $f''(p) \neq 0$.
Let $p = y'$. We have $y = xp + f(p)$. Differentiating we get

$$[x + f'(p)]p' = 0.$$

When $p' = 0$ we have $p = c$ and (1.4.4) has a general solution

$$y = cx + f(c).$$

When $x + f'(p) = 0$ we get a solution of (1.4.4) given by parameterized equations

$$x = -f'(p), \quad y = px + f(p).$$

2. Method of parameterization.
This method can be used to solve equations where either x or y is missing. Consider

$$F(y, y') = 0, \tag{1.4.5}$$

where x is missing. Let $p = y'$ and write (1.4.5) as

$$F(y, p) = 0.$$

It determines a family of curves in yp plane. Let $y = g(t)$, $p = h(t)$ be one of the curves, i.e. $F(g(t), h(t)) = 0$. Since

$$dx = \frac{dy}{y'} = \frac{dy}{p} = \frac{g'(t)dt}{h(t)},$$

we have $x = \int_{t_0}^{t} \frac{g'(t)}{h(t)} dt + c$. The solutions of (1.4.5.) are given by

$$x = \int_{t_0}^{t} \frac{g'(t)}{h(t)} dt + c, \quad y = g(t).$$

This method can also be applied to the equations

$$F(x, y') = 0,$$

where y is missing.

Example. Solve $y^2 + y'^2 - 1 = 0$.
Ans: $y = \cos(c - x)$.

3. Reduction of order.
Consider the equation

$$F(x, y', y'') = 0, \tag{1.4.6}$$

where y is missing. Let $p = y'$. Then $y'' = p'$. Write (1.4.6) as

$$F(x, p, p') = 0. \tag{1.4.7}$$

It is a first order equation in x and p. If $p = \phi(x, c_1)$ is a general solution of (1.4.7), then the general solution of (1.4.6) is

$$y = \int_{x_0}^{x} \phi(t, c_1) dt + c_2.$$

Example. Solve $xy'' - y' = 3x^2$.
Ans: $y = x^3 + c_1 x^2 + c_2$.

Consider the equation

$$F(y, y', y'') = 0, \tag{1.4.8}$$

where x is missing. Let $p = y'$. Then $y'' = \frac{dp}{dx} = \frac{dp}{dy}\frac{dy}{dx} = \frac{dp}{dy}p$. Write (1.4.8) as

$$F(y, p, p\frac{dp}{dy}) = 0. \tag{1.4.9}$$

It is a first order equation in y and p. If $p = \psi(y, c_1)$ is a general solution of (1.4.9), then we solve the equation

$$y' = \psi(y, c_1)$$

to get a general solution of (1.4.8).

Example. Solve $y'' + k^2 y = 0$, where k is a positive constant.
Ans: $y = c_1 \sin(kx) + c_2 \cos(kx)$.

Chapter 2

Higher Order Linear Equations

2.1 General Theory

Consider n-th order linear equation

$$y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_{n-1}(x)y' + a_n(x)y = f(x), \tag{2.1.1}$$

where $y^{(k)} = \frac{d^k y}{dx^k}$. Throughout this section we assume that $a_j(x)$'s and $f(x)$ are continuous functions defined on the interval (a,b). When $f(x) \not\equiv 0$, (2.1.1) is called a non-homogeneous equation. The associated homogeneous equation is

$$y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_{n-1}(x)y' + a_n(x)y = 0. \tag{2.1.2}$$

Let us begin with the initial value problem:

$$\begin{cases} y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_n(x)y = f(x), \\ y(x_0) = y_0, \\ y'(x_0) = y_1, \\ \cdots\cdots\cdots \\ y^{(n-1)}(x_0) = y_{n-1}. \end{cases} \tag{2.1.3}$$

Theorem 2.1 *Assume that $a_1(x)$, $\cdots$, $a_n(x)$ and $f(x)$ are continuous functions defined on the interval (a,b). Then for any $x_0 \in (a,b)$ and for any numbers y_0, $\cdots$, y_{n-1}, the initial value problem (2.1.3) has a unique solution defined on (a,b).*
Especially if $a_j(x)$'s and $f(x)$ are continuous on $\mathbb{R}$ then for any x_0 and y_0, $\cdots$, y_{n-1}, the initial value problem (2.1.3) has a unique solution defined on $\mathbb{R}$.

Proof of this theorem will be given in later chapter.

Corollary 2.2 *Let $y = y(x)$ be a solution of the homogeneous equation (2.1.2) in an interval (a,b). Assume that there exists $x_0 \in (a,b)$ such that*

$$y(x_0) = 0,\ y'(x_0) = 0,\ \cdots,\ y^{(n-1)}(x_0) = 0. \tag{2.1.4}$$

Then $y(x) \equiv 0$ on (a,b).

Proof. y is a solution of the initial value problem (2.1.2), (2.1.4). From Theorem 2.1, this problem has a unique solution. Since $\phi(x) \equiv 0$ is also a solution of the problem, we have $y(x) \equiv 0$ on (a, b).
□

In the following we consider the general solutions of (2.1.1) and (2.1.2).
Given continuous functions $a_j(x)$, $j = 0, 1, \cdots, n$ and $f(x)$, define an operator L by

$$L[y] = a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_n(x)y. \tag{2.1.5}$$

Property 1. $L[cy] = cL[y]$ for any constant c.
Property 2. $L[u + v] = L[u] + L[v]$.

Proof. Let's verify Property 2.

$$\begin{aligned} & L[u+v] \\ = \ & a_0(x)(u+v)^{(n)} + a_1(x)(u+v)^{(n-1)} + \cdots + a_n(x)(u+v) \\ = \ & \big(a_0(x)u^{(n)} + a_1(x)u^{(n-1)} + \cdots + a_n(x)u\big) + \big(a_0(x)v^{(n)} + a_1(x)v^{(n-1)} + \cdots + a_n(x)v\big) \\ = \ & L[u] + L[v]. \end{aligned}$$

Definition. An operator satisfying Properties 1 and 2 is called a *linear operator*.

The differential operator L defined in (2.1.4) is a linear operator.
Note that (2.1.1) and (2.1.2) can be written as

$$L[y] = f(x), \tag{2.1.1'}$$

and

$$L[y] = 0. \tag{2.1.2'}$$

From Properties 1 and 2 we get the following conclusion.

Theorem 2.3 *(1) If y_1 and y_2 are solutions of the homogeneous equation (2.1.2) in an interval (a, b), then for any constants c_1 and c_2,*

$$y = c_1y_1 + c_2y_2$$

is also a solution of (2.1.2) in the interval (a, b).
(2) If y_p is a solution of (2.1.1) and y_h is a solution of (2.1.2) on an interval (a, b), then

$$y = y_h + y_p$$

is also a solution of (2.1.1) in the interval (a, b).

Proof. (1) As $L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2] = 0$, we see that $c_1y_1 + c_2y_2$ is also a solution of (2.1.2).
(2) $L[y_h + y_p] = L[y_h] + L[y_p] = 0 + f(x)$, we see that $y = y_h + y_p$ is a solution of (2.1.1).

In order to discuss structures of solutions, we need the following definition.

Definition. Functions $\phi_1(x)$, $\cdots$, $\phi_k(x)$ are *linearly dependent* on (a, b) if there exists constants c_1, $\cdots$, c_k, not all zero, such that

$$c_1\phi_1(x) + \cdots + c_k\phi_k(x) = 0$$

for all $x \in (a, b)$. A set of functions are *linearly independent* on (a, b) if they are not linearly dependent on (a, b).

Lemma 2.4 *Functions $\phi_1(x)$, $\cdots$, $\phi_k(x)$ are linearly dependent on (a, b) if and only if the following vector-valued functions*

$$\begin{pmatrix} \phi_1(x) \\ \phi_1'(x) \\ \cdots \\ \phi_1^{(n-1)}(x) \end{pmatrix}, \cdots, \begin{pmatrix} \phi_k(x) \\ \phi_k'(x) \\ \cdots \\ \phi_k^{(n-1)}(x) \end{pmatrix} \tag{2.1.6}$$

are linearly dependent on (a, b).

Proof. "$\Longleftarrow$" is obvious. To show "$\Longrightarrow$", assume that ϕ_1, $\cdots$, ϕ_k are linearly dependent on (a, b). There exists constants c_1, $\cdots$, c_k, not all zero, such that, for all $x \in (a, b)$,

$$c_1\phi_1(x) + \cdots + c_k\phi_k(x) = 0.$$

Differentiating this equality successively we find that

$$\begin{aligned} &c_1\phi_1'(x) + \cdots + c_k\phi_k'(x) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &c_1\phi_1^{(n-1)}(x) + \cdots + c_k\phi_k^{(n-1)}(x) = 0. \end{aligned}$$

Thus

$$c_1 \begin{pmatrix} \phi_1(x) \\ \phi_1'(x) \\ \cdots \\ \phi_1^{(n-1)}(x) \end{pmatrix} + \cdots + c_k \begin{pmatrix} \phi_k(x) \\ \phi_k'(x) \\ \cdots \\ \phi_k^{(n-1)}(x) \end{pmatrix} = \mathbf{0}$$

for all $x \in (a, b)$. Hence the k vector-valued functions are linearly dependent on (a, b). □

Recall that, n vectors in $\mathbb{R}^n$ are linearly dependent if and only if the determinant of matrix formed by these vectors is zero.

Definition. The Wronskian of n functions $\phi_1(x)$, $\cdots$, $\phi_n(x)$ is defined by

$$W(\phi_1, \cdots, \phi_n)(x) = \begin{vmatrix} \phi_1(x) & \cdots & \phi_n(x) \\ \cdots & \cdots & \cdots \\ \phi_1^{(n-1)}(x) & \cdots & \phi_n^{(n-1)}(x) \end{vmatrix}. \tag{2.1.7}$$

Note that Wronskian of $\phi_1, \cdots, \phi_n$ is the determinant of the matrix formed by the vector-valued functions given in (2.1.6).

Theorem 2.5 *Let $y_1(x), \cdots, y_n(x)$ be n solutions of (2.1.2) on (a,b) and let $W(x)$ be their Wronskian.*
(1) $y_1(x), \cdots, y_n(x)$ are linearly dependent on (a,b) if and only if $W(x) \equiv 0$ on (a,b).
(2) $y_1(x), \cdots, y_n(x)$ are linearly independent on (a,b) if and only if $W(x)$ does not vanish on (a,b).

Corollary 2.6 *(1) The Wronskian of n solutions of (2.1.2) is either identically zero, or nowhere zero.*
(2) n solutions $y_1, \cdots, y_n$ of (2.1.2) are linearly independent on (a,b) if and only if the set of vectors

$$\begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix}, \cdots, \begin{pmatrix} y_n(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix}$$

are linearly independent for some $x_0 \in (a,b)$.

Proof of Theorem 2.5. Let $y_1, \cdots, y_n$ be solutions of (2.1.2) on (a,b), and let $W(x)$ be their Wronskian.

Step 1. We first show that, if $y_1, \cdots, y_n$ are linearly dependent on (a,b), then $W(x) \equiv 0$.
Since these solutions are linearly dependent, from Lemma 2.3, n vector-valued functions

$$\begin{pmatrix} y_1(x) \\ y_1'(x) \\ \cdots \\ y_1^{(n-1)}(x) \end{pmatrix}, \cdots, \begin{pmatrix} y_n(x) \\ y_n'(x) \\ \cdots \\ y_n^{(n-1)}(x) \end{pmatrix}$$

are linearly dependent on (a,b). Thus for all $x \in (a,b)$, the determinant of the matrix formed by these vectors, namely, the Wronskian of $y_1, \cdots, y_n$, is zero.

Step 2. Now, assume that the Wronskian $W(x)$ of n solutions $y_1, \cdots, y_n$ vanishes at $x_0 \in (a,b)$. We shall show that $y_1, \cdots, y_n$ are linearly dependent on (a,b).
Since $W(x_0) = 0$, the n vectors

$$\begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix}, \cdots, \begin{pmatrix} y_n(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix}$$

are linearly dependent. Thus there exist n constants $c_1, \cdots, c_n$, not all zero, such that

$$c_1 \begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix} + \cdots + c_n \begin{pmatrix} y_1(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix} = \mathbf{0} \tag{2.1.8}$$

Define

$$y_0(x) = c_1 y_1(x) + \cdots + c_n y_n(x).$$

From Theorem 2.3, y_0 is a solution of (2.1.2). From (2.1.8), y_0 satisfies the initial conditions

$$y(x_0) = 0,\ y'(x_0) = 0, \cdots, y^{(n-1)}(x_0) = 0. \tag{2.1.9}$$

From Corollary 2.2, we have $y_0 \equiv 0$, namely,

$$c_1 y_1(x) + \cdots + c_n y_n(x) = 0$$

for all $x \in (a, b)$. Thus $y_1, \cdots, y_n$ are linearly dependent on (a, b). □

Example. Consider the differential equation $y'' - \frac{1}{x} y' = 0$ on the interval $(0, \infty)$. Both $\phi_1(x) = 1$ and $\phi_2(x) = x^2$ are solutions of the differential equation. $W(\phi_1, \phi_2)(x) = \begin{vmatrix} 1 & x^2 \\ 0 & 2x \end{vmatrix} = 2x \neq 0$ for $x > 0$. Thus ϕ_1 and ϕ_2 are linearly independent solutions.

Theorem 2.7 *Let $a_1(x)$, $\cdots$, $a_n(x)$ and $f(x)$ be continuous on the interval (a, b). The homogeneous equation (2.1.2) has n linearly independent solutions on (a, b).*
Let $y_1, \cdots, y_n$ be n linearly independent solutions of (2.1.2) defined on (a, b). The general solution of (2.1.2) is given by

$$y(x) = c_1 y_1(x) + \cdots + c_n y_n(x), \tag{2.1.10}$$

where $c_1, \cdots, c_n$ are arbitrary constants.

Proof. (1) Fix $x_0 \in (a, b)$. For $k = 1, 2, \cdots, n$, let y_k be the solution of (2.1.2) satisfying the initial conditions

$$y_k^{(j)}(x_0) = \begin{cases} 0 & \text{if} \quad j \neq k - 1, \\ 1 & \text{if} \quad j = k - 1. \end{cases}$$

The n vectors

$$\begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix}, \cdots, \begin{pmatrix} y_n(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix}$$

are lineally independent since they form the identity matrix. From Corollary 2.6, $y_1, \cdots, y_n$ are linearly independent on (a, b). From Theorem 2.3, for any constants $c_1, \cdots, c_n$, $y = c_1 y_1 + \cdots + c_n y_n$ is a solution of (2.1.2).

(2) Now let $y_1, \cdots, y_n$ be n linearly independent solutions of (2.1.2) on (a, b). We shall show that the general solution of (2.1.2) is given by

$$y = c_1 y_1 + \cdots + c_n y_n. \tag{2.1.11}$$

Given a solution $\tilde{y}$ of (2.1.2), and fix $x_0 \in (a, b)$. Since $y_1, \cdots, y_n$ are linearly independent on (a, b), the vectors

$$\begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix}, \cdots, \begin{pmatrix} y_n(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix}$$

are linearly independent vectors. They form a basis for $\mathbb{R}^n$. Thus the vector

$$\begin{pmatrix} \tilde{y}(x_0) \\ \tilde{y}'(x_0) \\ \cdots \\ \tilde{y}^{(n-1)}(x_0) \end{pmatrix}$$

can be represented as a linear combination of the n vectors, namely, there exist n constants $\tilde{c}_1, \cdots, \tilde{c}_n$ such that

$$\begin{pmatrix} \tilde{y}(x_0) \\ \tilde{y}'(x_0) \\ \cdots \\ \tilde{y}^{(n-1)}(x_0) \end{pmatrix} = \tilde{c}_1 \begin{pmatrix} y_1(x_0) \\ y_1'(x_0) \\ \cdots \\ y_1^{(n-1)}(x_0) \end{pmatrix} + \cdots + \tilde{c}_n \begin{pmatrix} y_n(x_0) \\ y_n'(x_0) \\ \cdots \\ y_n^{(n-1)}(x_0) \end{pmatrix}.$$

Let

$$\phi(x) = \tilde{y}(x) - [\tilde{c}_1 y_1(x) + \cdots + \tilde{c}_n y_n(x)].$$

$\phi(x)$ is a solution of (2.1.2) and satisfies the initial conditions (2.1.4) at $x = x_0$. By Corollary 2.2, $\phi(x) \equiv 0$ on (a, b). Thus

$$\tilde{y}(x) = \tilde{c}_1 y_1(x) + \cdots + \tilde{c}_n y_n(x).$$

So (2.1.11) gives a general solution of (2.1.2). □

Any set of n linearly independent solutions is called a *fundamental set of solutions*.

Now we consider the non-homogeneous equation (2.1.1). We have

Theorem 2.8 *Let y_p be a particular solution of (2.1.1), and $y_1, \cdots, y_n$ be a fundamental set of solutions for the associated homogeneous equation (2.1.2). The general solution of (2.1.1) is given by*

$$y(x) = c_1 y_1(x) + \cdots + c_n y_n(x) + y_p(x). \tag{2.1.12}$$

Proof. Let y be a solution of the non-homogeneous equation. Then $y - y_p$ is a solution of the homogeneous equation. Thus $y(x) - y_p(x) = c_1 y_1(x) + \cdots + c_n y_n(x)$. □

2.2 Linear Equations with Constant Coefficients

Let us begin with second order linear equation with constant coefficients

$$y'' + ay' + by = 0, \tag{2.2.1}$$

where a and b are constants. We look for a solution of the form $y = e^{\lambda x}$. Plugging into (2.2.1) we find that, $e^{\lambda x}$ is a solution of (2.2.1) if and only if

$$\lambda^2 + a\lambda + b = 0. \tag{2.2.2}$$

(2.2.2) is called the *auxiliary equation* or *characteristic equation* of (2.2.1). The roots of (2.2.2) are called *characteristic values* (or eigenvalues):

$$\lambda_1 = \frac{1}{2}(-a + \sqrt{a^2 - 4b}),$$
$$\lambda_2 = \frac{1}{2}(-a - \sqrt{a^2 - 4b}).$$

1. If $a^2 - 4b > 0$, (2.2.2) has two distinct real roots λ_1, λ_2, and the general solutions of (2.2.1) is

$$y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$

2. If $a^2 - 4b = 0$, (2.2.2) has one real roots λ (we may say that (2.2.2) has two equal roots $\lambda_1 = \lambda_2$). The general solution of (2.2.2) is

$$y = c_1 e^{\lambda x} + c_2 x e^{\lambda x}.$$

3. If $a^2 - 4b < 0$, (2.2.2) has a pair of complex conjugate roots

$$\lambda_1 = \alpha + i\beta, \quad \lambda_2 = \alpha - i\beta.$$

The general solution of (2.2.2) is

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

Example. Solve $y'' + y' - 2y = 0$, $y(0) = 4$, $y'(0) = -5$.
Ans: $\lambda_1 = 1$, $\lambda_2 = -2$, $y = e^x + 3e^{-2x}$.

Example. Solve $y'' - 4y' + 4y = 0$, $y(0) = 3$, $y'(0) = 1$.
Ans: $\lambda_1 = \lambda_2 = 2$, $y = (3 - 5x)e^{2x}$.

Example. Solve $y'' - 2y' + 10y = 0$.
Ans: $\lambda_1 = 1 + 3i$, $\lambda_2 = 1 - 3i$,
$y = e^x(c_1 \cos 3x + c_2 \sin 3x)$.

Now we consider n-th order homogeneous linear equations with constant coefficients

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0, \tag{2.2.3}$$

where $a_1, \cdots, a_n$ are real constants.
$y = e^{\lambda x}$ is a solution of (2.2.3) if and only if λ satisfies

$$\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0. \tag{2.2.4}$$

The solutions of (2.2.4) are called *characteristic values* or eigenvalues for the equation (2.2.3).

Let $\lambda_1, \cdots, \lambda_s$ be the distinct eigenvalues for (2.2.3). Then we can write

$$\begin{aligned}&\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \\ &= (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s},\end{aligned} \tag{2.2.5}$$

where $m_1, \cdots, m_s$ are positive integers and

$$m_1 + \cdots + m_s = n.$$

We call them the *multiplicity* of the eigenvalues $\lambda_1, \cdots, \lambda_s$ respectively.

Lemma 2.9 *Assume λ is an eigenvalue of (2.2.3) of multiplicity m.*
(i) $e^{\lambda x}$ is a solution of (2.2.3).
(ii) If $m > 1$, then for any positive integer $1 \le k \le m-1$, $x^k e^{\lambda x}$ is a solution of (2.2.3).
(iii) If $\lambda = \alpha + i\beta$, then

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

are solutions of (2.2.3), where $0 \le k \le m-1$.

Theorem 2.10 *Let $\lambda_1, \cdots, \lambda_s$ be the distinct eigenvalues for (2.2.3), with multiplicity $m_1, \cdots, m_s$ respectively. Then (2.2.3) has a fundamental set of solutions*

$$\begin{aligned}&e^{\lambda_1 x},\ xe^{\lambda_1 x},\ \cdots,\ x^{m_1-1}e^{\lambda_1 x}; \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots; \\ &e^{\lambda_s x},\ xe^{\lambda_s x},\ \cdots,\ x^{m_s-1}e^{\lambda_s x}.\end{aligned} \tag{2.2.6}$$

Proof of Lemma 2.9 and 2.10

Consider the n-th order linear equation with constant coefficients

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0, \tag{A1}$$

where $y^{(k)} = \frac{d^k y}{dx^k}$. Let $L(y) = y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y$ and $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$. Note that p is a polynomial in z of degree n. Then we have

$$L(e^{zx}) = p(z)e^{zx} \tag{A2}$$

Before we begin the proof, let's observe that $\frac{\partial^2}{\partial z \partial x}e^{zx} = \frac{\partial^2}{\partial x \partial z}e^{zx}$ by Clairaut's theorem because e^{zx} is differentiable in (x, z) as a function of two variable and all the higher order partial derivatives exist and continuous. That means we can interchange the order of differentiation with respect to x and z as we wish. Therefore $\frac{d}{dz}L(e^{zx}) = L(\frac{d}{dz}e^{zx})$. For instance, one may verify directly that

$$\frac{d}{dz}\frac{d^k}{dx^k}(e^{zx}) = xz^k e^{zx} + kz^{k-1}e^{zx} = \frac{d^k}{dx^k}(\frac{d}{dz}e^{zx}).$$

Here one may need to use Leibniz's rule of taking the k-th derivative of a product of two functions:

$$(u \cdot v)^{(k)} = \sum_{i=0}^{k}\binom{k}{i}u^{(i)}v^{(k-i)}. \tag{A3}$$

More generally, $\frac{d^k}{dz^k}L(e^{zx}) = L(\frac{d^k}{dz^k}e^{zx})$. (Strictly speaking, partial derivative notations should be used.) Now let's prove our results.

(1) If λ is a root of p, then $L(e^{\lambda x}) = 0$ by (A2) so that $e^{\lambda x}$ is a solution of (A1).
(2) If λ is a root of p of multiplicity m, then $p(\lambda) = 0, p'(\lambda) = 0, p''(\lambda) = 0, \ldots, p^{(m-1)}(\lambda) = 0$. Now for $k = 1, \ldots, m-1$, differentiating (A2) k times with respect to z, we have

$$L(x^k e^{zx}) = L(\frac{d^k}{dz^k}e^{zx}) = \frac{d^k}{dz^k}L(e^{zx}) = \frac{d^k}{dz^k}(p(z)e^{zx}) = \sum_{i=0}^{k}\binom{k}{i}p^{(i)}(z)x^{k-i}e^{zx}.$$

Thus $L(x^k e^{\lambda x}) = 0$ and $x^k e^{\lambda x}$ is solution of (A1).
(3) Let $\lambda_1, \cdots, \lambda_s$ be the distinct roots of p, with multiplicity $m_1, \cdots, m_s$ respectively. Then we wish to prove that

$$\begin{array}{l} e^{\lambda_1 x},\ xe^{\lambda_1 x},\ \cdots,\ x^{m_1-1}e^{\lambda_1 x}; \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots ; \\ e^{\lambda_s x},\ xe^{\lambda_s x},\ \cdots,\ x^{m_s-1}e^{\lambda_s x}. \end{array} \tag{A4}$$

are linearly independent over $\mathbb{R}$. To prove this, suppose for all x in $\mathbb{R}$

$$\begin{array}{l} c_{11}e^{\lambda_1 x} + \ c_{12}xe^{\lambda_1 x} \ + \cdots + \ c_{1m_1}x^{m_1-1}e^{\lambda_1 x} \\ + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots + \\ c_{s1}e^{\lambda_s x} + \ c_{s2}xe^{\lambda_s x} \ + \cdots + \ c_{sm_s}x^{m_s-1}e^{\lambda_s x} = 0. \end{array}$$

Let's write this as

$$P_1(x)e^{\lambda_1 x} + P_2(x)e^{\lambda_2 x} + \cdots + P_s(x)e^{\lambda_s x} = 0,$$

for all x in $\mathbb{R}$, where $P_i(x) = c_{i1} + c_{i2}x + \cdots + c_{im_i}x^{m_i-1}$. We need to prove $P_i(x) \equiv 0$ for all i. Assume that one of the P_i's is not identically zero. By re-labelling the P_i's, we may assume that $P_s(x) \not\equiv 0$. Dividing the above equation by $e^{\lambda_1 x}$, we have

$$P_1(x) + P_2(x)e^{(\lambda_2-\lambda_1)x} + \cdots + P_s(x)e^{(\lambda_s-\lambda_1)x} = 0,$$

for all x in $\mathbb{R}$. Upon differentiating this equation sufficiently many times (at most m_1 times since $P_1(x)$ is a polynomial of degree $m_1 - 1$), we can reduce $P_1(x)$ to 0. Note that in this process, the degree of the resulting polynomial multiplied by $e^{(\lambda_i-\lambda_1)x}$ remains unchanged. Therefore, we get

$$Q_2(x)e^{(\lambda_2-\lambda_1)x} + \cdots + Q_s(x)e^{(\lambda_s-\lambda_1)x} = 0,$$

where deg $Q_i =$ deg P_i. Cancelling the term $e^{\lambda_1}x$ we have

$$Q_2(x)e^{\lambda_2 x} + \cdots + Q_s(x)e^{\lambda_s x} = 0.$$

Repeating this procedure, we arrive at

$$R_s(x)e^{\lambda_s x} = 0,$$

where deg $R_s =$ deg P_s. Hence $R_s(x) \equiv 0$ on $\mathbb{R}$, contradicting the fact that deg $R_s =$ deg P_s and P_s is not identically zero. Thus all the $P_i(x)$ are identically zero. That means all c_{ij}'s are zero and the functions in (A4) are linearly independent.

Remark. If (2.2.3) has a complex eigenvalue $\lambda = \alpha + i\beta$, then $\bar{\lambda} = \alpha - i\beta$ is also an eigenvalue. Thus both $x^k e^{(\alpha+i\beta)x}$ and $x^k e^{(\alpha-i\beta)x}$ appear in (2.2.6), where $0 \le k \le m-1$. In order to obtain a fundamental set of real solutions, the pair of solutions $x^k e^{(\alpha+i\beta)x}$ and $x^k e^{(\alpha-i\beta)x}$ in (2.2.6) should be replaced by $x^k e^{\alpha x}\cos(\beta x)$ and $x^k e^{\alpha x}\sin(\beta x)$.

In the following we discuss solution of non-homogeneous equations. For simplicity we consider

$$y'' + by' + cy = f(x), \tag{2.2.7}$$

where b and c are real constants. The associated homogeneous equation is (2.2.1), and the characteristic equation of (2.2.1) is (2.2.2). We shall look for a particular solution of (2.2.7). This method works in general even if a and b are functions of x. Also the method applies to higher order equations.

1. Methods of variation of parameters.

Let y_1 and y_2 be two linearly independent solutions of the associated homogeneous equation (2.2.1) and $W(x)$ be their Wronskian. Look for a particular solution of (2.2.7) in the form

$$y_p = u_1 y_1 + u_2 y_2,$$

where u_1 and u_2 are functions to be determined. Suppose

$$u_1' y_1 + u_2' y_2 = 0.$$

Plugging y_p into (2.2.7) we get

$$u_1' y_1' + u_2' y_2' = f.$$

Hence u_1' and u_2' satisfy

$$\begin{cases} u_1' y_1 + u_2' y_2 = 0, \\ u_1' y_1' + u_2' y_2' = f. \end{cases} \tag{2.2.8}$$

Solving it, we find that

$$u_1' = -\frac{y_2}{W} f, \qquad u_2' = \frac{y_1}{W} f.$$

Integrating yields

$$\begin{aligned} u_1(x) &= -\int_{x_0}^{x} \frac{y_2(t)}{W(t)} f(t)dt, \\ u_2(x) &= \int_{x_0}^{x} \frac{y_1(t)}{W(t)} f(t)dt. \end{aligned} \tag{2.2.9}$$

Example. Solve the differential equation. $y'' + y = \sec x$.

Solution. A basis for the solutions of the homogeneous equation consists of $y_1 = \cos x$ and $y_2 = \sin x$. Now $W(y_1, y_2) = \cos x \cos x - (-\sin x)\sin x = 1$. Thus $u_1 = -\int \sin x \sec x\, dx = \ln|\cos x| + c_1$ and $u_2 = \int \cos x \sec x\, dx = x + c_2$. From this, a particular solution is given by $y_p = \cos x \ln|\cos x| + x \sin x$. Therefore, the general solution is $y = c_1 \cos x + c_2 \sin x + \cos x \ln|\cos x| + x \sin x$.

The method of variation of parameters can also be used to find another solution of a second order homogeneous linear differential equation when one solution is given. Suppose z is a known solution of the equation

$$y'' + P(x)y' + Q(x)y = 0.$$

We assume $y = vz$ is a solution so that

$$\begin{aligned} 0 &= (vz)'' + P(vz)' + Q(vz) = (v''z + 2v'z' + vz'') + P(v'z + vz') + Qvz \\ &= (v''z + 2v'z' + Pv'z) + v(z'' + Pz' + Qz) = v''z + v'(2z' + Pz). \end{aligned}$$

That is

$$\frac{v''}{v'} = -2\frac{z'}{z} - P.$$

An integration gives $v' = z^{-2}e^{-\int P\,dx}$ and $v = \int z^{-2}e^{-\int P\,dx}\,dx$. We leave it as an exercise to show that z and vz are linearly independent solutions by computing their Wronskian.

Example. Given $y_1 = x$ is a solution of $x^2y'' + xy' - y = 0$, find another solution.
Solution. Let's write the DE in the form $y'' + \frac{1}{x}y' - \frac{1}{x^2}y = 0$. Then $P(x) = 1/x$. Thus a second linearly independent solution is given $y = vx$, where

$$v = \int x^{-2}e^{-\int 1/x\,dx}\,dx = \int x^{-2}x^{-1}\,dx = -\frac{1}{x^2}.$$

Therefore the second solution is $y = -x^{-1}$ and the general solution is $y = c_1x + c_2x^{-1}$.

2. Method of undetermined coefficients.
Case 1. $f(x) = P_n(x)e^{\alpha x}$, where $P_n(x)$ is a polynomial of degree $n \geq 0$.
We look for a particular solution in the form

$$y = Q(x)e^{\alpha x},$$

where $Q(x)$ is a polynomial. Plugging it into (2.2.7) we find

$$Q'' + (2\alpha + b)Q' + (\alpha^2 + b\alpha + c)Q = P_n(x). \tag{2.2.10}$$

Subcase 1.1. If $\alpha^2 + b\alpha + c \neq 0$, namely, α is not a root of (2.2.2), we choose $Q = R_n$, a

polynomial of degree n, and

$$y = R_n(x)e^{\alpha x}.$$

The coefficients of R_n can be determined by comparing the terms of same power in the two sides of (2.2.10). Note that in this case both sides of (2.2.10) are polynomials of degree n.

Subcase 1.2. If $\alpha^2 + b\alpha + c = 0$ but $2\alpha + b \neq 0$, namely, α is a simple root of (2.2.2), then (2.2.10) is reduced to

$$Q'' + (2\alpha + b)Q' = P_n. \tag{2.2.11}$$

We choose Q to be a polynomial of degree $n+1$. Since the constant term of Q does not appear in (2.2.11), we may choose $Q(x) = xR_n(x)$, where $R_n(x)$ is a polynomial of degree n.

$$y = xR_n(x)e^{\alpha x}.$$

Subcase 1.3 If $\alpha^2 + b\alpha + c = 0$ and $2\alpha + b = 0$, namely, α is a root of (2.2.2) with multiplicity 2, then (2.2.10) is reduced to

$$Q'' = P_n. \tag{2.2.12}$$

We choose $Q(x) = x^2 R_n(x)$, where $R_n(x)$ is a polynomial of degree n.

$$y = x^2 R_n(x)e^{\alpha x}.$$

Example. Find the general solution of $y'' - y' - 2y = 4x^2$.
Ans: $y = c_1e^{2x} + c_2e^{-x} - 3 + 2x - 2x^2$.

Example. Find a particular solution of $y''' + 2y'' - y' = 3x^2 - 2x + 1$.
Ans: $y = -27x - 5x^2 - x^3$.

Example. Solve $y'' - 2y' + y = xe^x$.
Ans: $y = c_1e^x + c_2xe^x + \frac{1}{6}x^3e^x$.

Case 2. $f(x) = P_n(x)e^{\alpha x}\cos(\beta x)$ or $f(x) = P_n(x)e^{\alpha x}\sin(\beta x)$, where $P_n(x)$ is a polynomial of degree $n \geq 0$.
We first look for a solution of

$$y'' + by' + cy = P_n(x)e^{(\alpha + i\beta)x}. \tag{2.2.13}$$

Using the method in Case 1 we obtain a complex-valued solution

$$z(x) = u(x) + iv(x),$$

where $u(x) = \Re(z(x))$, $v(x) = \Im(z(x))$. Substituting $z(x) = u(x) + iv(x)$ into (2.2.13) and taking the real and imaginary parts, we can show that $u(x) = \Re(z(x))$ is a solution of

$$y'' + by' + cy = P_n(x)e^{\alpha x}\cos(\beta x), \tag{2.2.14}$$

and $v(x) = \Im(z(x))$ is a solution of

$$y'' + by' + cy = P_n(x)e^{\alpha x}\sin(\beta x). \tag{2.2.15}$$

Example. Solve $y'' - 2y' + 2y = e^x\cos x$.
Ans: $y = c_1e^x\cos x + c_2e^x\sin x + \frac{1}{2}xe^x\sin x$.

The following conclusions will be useful.

Theorem 2.11 *Let y_1 and y_2 be particular solutions of the equations*

$$y'' + ay' + by = f_1(x)$$

and

$$y'' + ay' + by = f_2(x)$$

respectively, then $y_p = y_1 + y_2$ is a particular solution of

$$y'' + ay' + by = f_1(x) + f_2(x).$$

Proof. Exercise.

Example. Solve $y'' - y = e^x + \sin x$.
Solution. A particular solution for $y'' - y = e^x$ is given by $y_1 = \frac{1}{2}xe^x$. Also a particular solution for $y'' - y = \sin x$ is given by $y_2 = -\frac{1}{2}x \sin x$. Thus $\frac{1}{2}(xe^x - \sin x)$ is a particular solution of the given differential equation. The general solution of the corresponding homogeneous differential equation is given by $c_1e^{-x} + c_2e^x$. Hence the general solution of the given differential equation is $c_1e^{-x} + c_2e^x + \frac{1}{2}(xe^x - \sin x)$.

2.3 Operator methods

Let x denote independent variable, and y dependent variable. Introduce

$$Dy = \frac{d}{dx}y, \quad D^n y = \frac{d^n}{dx^n}y = y^{(n)}.$$

We define $D^0 y = y$. Given a polynomial $L(x) = \sum_{j=0}^{n} a_j x^j$, where a_j's are constants, we define a differential operator $L(D)$ by

$$L(D)y = \sum_{j=0}^{n} a_j D^j y.$$

Then the equation

$$\sum_{j=0}^{n} a_j y^{(j)} = f(x) \tag{2.3.1}$$

can be written as

$$L(D)y = f(x). \tag{2.3.2}$$

Let $L(D)^{-1}f$ denote any solution of (2.3.2). We have

$$D^{-1}D = DD^{-1} = D^0,$$
$$L(D)^{-1}L(D) = L(D)L(D)^{-1} = D^0.$$

However, $L(D)^{-1}f$ is not unique.

To see the above properties, first recall that $D^{-1}f$ means a solution of $y' = f$. Thus $D^{-1}f = \int f$. Hence it follows that $D^{-1}D = DD^{-1}$= identity operator D^0.

For the second equality, note that a solution of $L(D)y = L(D)f$ is simply f. Thus by definition of $L(D)^{-1}$, we have $L(D)^{-1}(L(D)f) = f$. This means $L(D)^{-1}L(D) = D^0$. Lastly, since $L(D)^{-1}f$ is a solution of $L(D)y = f(x)$, it is clear that $L(D)(L(D)^{-1}f) = f$. In other words, $L(D)L(D)^{-1} = D^0$.

More generally, we have:

$$\begin{aligned}
&1. \quad D^{-1}f(x) = \int f(x)dx + C,\\
&2. \quad (D-a)^{-1}f(x) = Ce^{ax} + e^{ax}\int e^{-ax}f(x)dx,\\
&3. \quad L(D)(e^{ax}f(x)) = e^{ax}L(D+a)f(x),\\
&4. \quad L(D)^{-1}(e^{ax}f(x)) = e^{ax}L(D+a)^{-1}f(x).
\end{aligned} \tag{2.3.3}$$

Proof. Property 2 is just the solution of the first order linear ODE. To prove Property 3, first observe that $(D-r)(e^{ax}f(x)) = e^{ax}D(f(x)) + ae^{ax}f(x) - re^{ax}f(x) = e^{ax}(D+a-r)(f(x))$. Thus $(D-s)(D-r)(e^{ax}f(x)) = (D-s)[e^{ax}(D+a-r)(f(x))] = e^{ax}(D+a-s)(D+a-r)(f(x))$. Now we may write $L(D) = (D-r_1)\cdots(D-r_n)$. Then $L(D)(e^{ax}f(x)) = e^{ax}L(D+a)f(x)$. This says that we can move the factor e^{ax} to the left of the operator $L(D)$ if we replace $L(D)$ by $L(D+a)$.

To prove Property 4, apply $L(D)$ to the right hand side. We have

$$L(D)[e^{ax}L(D+a)^{-1}f(x)] = e^{ax}[L(D+a)(L(D+a)^{-1}f(x))] = e^{ax}f(x).$$

Thus $L(D)^{-1}(e^{ax}f(x)) = e^{ax}L(D+a)^{-1}f(x)$. □

Let $L(x) = (x-r_1)\cdots(x-r_n)$. The solution of (2.3.2) is given by

$$y = L(D)^{-1}f(x) = (D-r_1)^{-1}\cdots(D-r_n)^{-1}f(x). \tag{2.3.4}$$

Then we obtain the solution by successive integration. Moreover, if $r_j's$ are distinct, we can write

$$\frac{1}{L(x)} = \frac{A_1}{x-r_1} + \cdots + \frac{A_n}{x-r_n},$$

where $A_j's$ can be found by the method of partial fractions. Then the solution is given by

$$y = [A_1(D-r_1)^{-1} + \cdots + A_n(D-r_n)^{-1}]f(x). \tag{2.3.5}$$

Next consider the case of repeated roots. Let the multiple root be equal to m and the equation to be solved is

$$(D-m)^n y = f(x) \tag{2.3.6}$$

To solve this equation, let us assume a solution of the form $y = e^{mx}v(x)$, where $v(x)$ is a function of x to be determined. One can easily verify that $(D-m)^n e^{mx}v = e^{mx}D^n v$. Thus equation (2.3.6) reduces to

$$D^n v = e^{-mx}f(x) \tag{2.3.7}$$

If we integrate (2.3.7) n times, we obtain

$$v = \int \int \cdots \int \int e^{-mx} f(x)\, dx \cdots dx + c_0 + c_1 x + \cdots + c_{n-1} x^{n-1} \tag{2.3.8}$$

Thus we see that

$$(D-m)^{-n} f(x) = e^{mx} \left[\iint \cdots \iint e^{-mx} f(x) dx \cdots\, dx + c_0 + c_1 x + \cdots + c_{n-1} x^{n-1} \right] \tag{2.3.9}$$

Example Solve $(D^2 - 3D + 2)y = xe^x$.
Solution First $\frac{1}{D^2-3D+2} = \frac{1}{D-2} - \frac{1}{D-1}$. Therefore

$$\begin{aligned} y &= (D^2 - 3D + 2)^{-1}(xe^x) \\ &= (D-2)^{-1}(xe^x) - (D-1)^{-1}(xe^x) \\ &= e^{2x} D^{-1}(e^{-2x} xe^x) - e^x D^{-1}(e^{-x} xe^x) \\ &= e^{2x} D^{-1}(e^{-x} x) - e^x D^{-1}(x) \\ &= e^{2x}(-xe^{-x} - e^{-x} + c_1) - e^x(\tfrac{1}{2}x^2 + c_2) \\ &= -e^x(\tfrac{1}{2}x^2 + x + 1) + c_1 e^{2x} + c_2 e^x. \end{aligned}$$

Example. Solve $(D^3 - 3D^2 + 3D - 1)y = e^x$.
Solution. The DE is equivalent to $(D-1)^3 y = e^x$. Therefore,

$$y = (D-1)^{-3} e^x = e^x \left[\iiint e^{-x} e^x\, dx + c_0 + c_1 x + c_2 x^2 \right] = e^x \left[\frac{1}{6} x^3 + c_0 + c_1 x + c_2 x^2 \right].$$

If $f(x)$ is a polynomial in x, then $(1-D)(1 + D + D^2 + D^3 + \cdots)f = f$. Thus $(1-D)^{-1}(f) = (1 + D + D^2 + D^3 + \cdots)f$. Therefore, if f is a polynomial, we may formally expand $(D-r)^{-1}$ into power series in D and applying it to f. If the degree of f is n, then it is only necessary to expand $(D-r)^{-1}$ up to D^n.

Example. Solve $(D^4 - 2D^3 + D^2)y = x^3$.
Solution. We have

$$\begin{aligned} y &= (D^4 - 2D^3 + D^2)^{-1} f = \tfrac{1}{D^2(1-D)^2} x^3 \\ &= D^{-2}(1 + 2D + 3D^2 + 4D^3 + 5D^4 + 6D^5)x^3 \\ &= D^{-2}(x^3 + 6x^2 + 18x + 24) \\ &= D^{-1}(\tfrac{x^4}{4} + 2x^3 + 9x^2 + 24x) \\ &= \tfrac{x^5}{20} + \tfrac{x^4}{2} + 3x^3 + 12x^2. \end{aligned}$$

Therefore, the general solution is $y = (A + Bx)e^x + (C + Dx) + \frac{x^5}{20} + \frac{x^4}{2} + 3x^3 + 12x^2$.

2.4 Exact 2nd order Equations

The general 2nd order linear differential equation is of the form

$$p_0(x)y'' + p_1(x)y' + p_2(x)y = f(x) \tag{A1}$$

The equation can be written as

$$(p_0y' - p_0'y)' + (p_1y)' + (p_0'' - p_1' + p_2)y = f(x) \tag{A2}$$

It is said to be exact if

$$p_0'' - p_1' + p_2 \equiv 0. \tag{A3}$$

In the event that the equation is exact, a first integral to (A1) is

$$p_0(x)y' - p_0'(x)y + p_1(x)y = \int f(x)\,dx + C_1.$$

Example. Find the general solution of the DE

$$\frac{1}{x}y'' + (\frac{1}{x} - \frac{2}{x^2})y' - (\frac{1}{x^2} - \frac{2}{x^3})y = e^x.$$

Solution. Condition (A3) is fulfilled. The first integral is

$$\frac{1}{x}y' + \frac{1}{x^2}y + (\frac{1}{x} - \frac{2}{x^2})y = e^x + C_1.$$

That is

$$y' + (1 - \frac{1}{x})y = xe^x + C_1x.$$

From the last equation, the general solution is found to be

$$y = \frac{1}{2}xe^x + C_1x + C_2xe^{-x}.$$

2.5 The adjoint differential equation and integrating factor

If (A1) is multiplied by a function $v(x)$ so that the resulting equation is exact, then $v(x)$ is called an integrating factor of (A1). That is

$$(p_0v)'' - (p_1v)' + p_2v = 0. \tag{A4}$$

This is a differential equation for v, which is, more explicitly,

$$p_0(x)v'' + (2p_0'(x) - p_1(x))v' + (p_0''(x) - p_1'(x) + p_2(x))v = 0. \tag{A5}$$

Equation (A5) is called the *adjoint* of the given differential equation (A1). A function $v(x)$ is thus an integrating factor for a given differential equation, if and only if it is a solution of the adjoint

equation. Note that the adjoint of (A5) is in turn found to be the associated homogeneous equation of (A1), thus each is the adjoint of the other.
In this case, a first integral to (A1) is

$$v(x)p_0(x)y' - (v(x)p_0(x))'y + v(x)p_1(x)y = \int v(x)f(x)\,dx + C_1.$$

Example. Find the general solution of the DE

$$(x^2 - x)y'' + (2x^2 + 4x - 3)y' + 8xy = 1.$$

Solution. The adjoint of this equation is

$$(x^2 - x)v'' - (2x^2 - 1)v + (4x - 2)v = 0.$$

By the trial of x^m, this equation is found to have x^2 as a solution. Thus x^2 is an integrating factor of the given differential equation. Multiplying the original equation by x^2, we obtain

$$(x^4 - x^3)y'' + (2x^4 + 4x^3 - 3x^2)y' + 8x^3y = x^2.$$

Thus a first integral to it is

$$(x^4 - x^3)y' - (4x^3 - 3x^2)y + (2x^4 + 4x^3 - 3x^2)y = \int x^2\,dx + C.$$

After simplification, we have

$$y' + \frac{2x}{x-1}y = \frac{1}{3(x-1)} + \frac{C}{x^3(x-1)}.$$

An integrating factor for this first order linear equation is $e^{2x}(x-1)^2$. Thus the above equation becomes

$$e^{2x}(x-1)^2y = \frac{1}{3}\int (x-1)e^{2x}\,dx + C\int \frac{e^{2x}(x-1)}{x^3}\,dx + C_2.$$

That is

$$e^{2x}(x-1)^2y = \frac{1}{3}\left[\frac{x}{2} - \frac{3}{4}\right]e^{2x} + C\frac{e^{2x}}{2x^2} + C_2.$$

Thus the general solution is

$$y = \frac{1}{(x-1)^2}\left(\frac{x}{6} - \frac{1}{4} + \frac{C_1}{x^2} + C_2e^{-2x}\right).$$

Exercise. Solve the following differential equation by finding an integrating factor of it.

$$y'' + \frac{4x}{2x-1}y' + \frac{8x-8}{(2x-1)^2}y = 0.$$

[**Answer:** $y = \frac{C_1}{2x-1} + \frac{C_2x}{2x-1}e^{-2x}$.]

Solution. The adjoint equation is

$$v'' - \frac{4x}{2x-1}v' + \frac{4}{2x-1}v = 0,$$

or equivalently

$$(2x-1)v'' - 4xv' + 4v = 0.$$

An obvious solution is $v = x$. Therefore $v = x$ is an integrating factor of the original differential equation. Thus

$$xy'' + \frac{4x^2}{2x-1}y' + \frac{(8x-8)x}{(2x-1)^2}y = 0$$

is exact. The first integral is

$$xy' - y + \frac{4x^2}{2x-1}y = C_1,$$

or equivalently,

$$y' + \frac{4x^2-2x+1}{x(2x-1)}y = \frac{C_1}{x}.$$

That is

$$y' + \left(2 - \frac{1}{x} + \frac{2}{2x-1}\right)y = \frac{C_1}{x}.$$

Thus $e^{\int 2-\frac{1}{x}+\frac{2}{2x-1}\,dx} = e^{2x}\left(\frac{2x-1}{x}\right)$ is an integrating factor of this first order equation. Multiplying by this factor, we have

$$e^{2x}\left(\frac{2x-1}{x}\right)y = C_1\int \frac{e^{2x}(2x-1)}{x^2}\,dx + C_2.$$

That is

$$e^{2x}\left(\frac{2x-1}{x}\right)y = C_1\frac{e^{2x}}{x} + C_2,$$

or equivalently

$$y = \frac{C_1}{2x-1} + \frac{C_2 x}{2x-1}e^{-2x}.$$

Chapter 3

Linear Differential Systems

3.1 Linear Systems

The following system is called a linear differential system of first order in *normal form*:

$$\begin{cases} x_1' = a_{11}(t)x_1 + \cdots + a_{1n}(t)x_n + g_1(t), \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n' = a_{n1}(t)x_1 + \cdots + a_{nn}(t)x_n + g_n(t), \end{cases}$$

where $a_{ij}(t)$ and $g_j(t)$ are continuous functions of t and $'$ denotes differentiation with respect to t. Denote

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \cdots \\ x_n(t) \end{pmatrix}, \quad \mathbf{g}(t) = \begin{pmatrix} g_1(t) \\ \cdots \\ g_n(t) \end{pmatrix}, \quad \mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \cdots & \cdots & \cdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

We call $\mathbf{g}(t)$ a continuous vector field, or a continuous vector-valued function, if all its components are continuous functions. We call $\mathbf{A}(t)$ a continuous matrix, or a continuous matrix-valued function, if all its entries are continuous functions. Define

$$\mathbf{x}' = \begin{pmatrix} x_1'(t) \\ \cdots \\ x_n'(t) \end{pmatrix}, \quad \int \mathbf{x}(t)dt = \begin{pmatrix} \int x_1(t)dt \\ \cdots \\ \int x_n(t)dt \end{pmatrix}.$$

Then the linear system can be written in the *matrix form*:

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{g}(t). \tag{3.1.1}$$

When $\mathbf{g} \equiv 0$, (3.1.1) is reduced to

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x}. \tag{3.1.2}$$

(3.1.2) is called a homogeneous differential system, and (3.1.1) is called a non-homogeneous system if $\mathbf{g}(t) \not\equiv \mathbf{0}$. We shall also call (3.1.2) the *homogeneous system associated with* (3.1.1), or the *associated homogeneous system*.

Example.

$$\begin{array}{rcl} x_1' &=& 2x_1 + 3x_2 + 3t, \\ x_2' &=& -x_1 + x_2 - 7\sin t \end{array}$$

is equivalent to

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 3t \\ -7\sin t \end{pmatrix}.$$

Example. Given a second order system

$$\begin{cases} \dfrac{d^2x}{dt^2} = x + 2y + 3t, \\ \dfrac{d^2y}{dt^2} = 4x + 5y + 6t, \end{cases}$$

it can be expressed into an equivalent first order differential system by introducing more variables. For this example, let $u = x'$ and $v = y'$. Then we have

$$\begin{array}{rcl} x' &=& u \\ u' &=& x + 2y + 3t \\ y' &=& v \\ v' &=& 4x + 5y + 6t \end{array}.$$

Next, we begin with the initial value problem

$$\begin{cases} \mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{g}(t), \\ \mathbf{x}(t_0) = \mathbf{x}_0, \end{cases} \tag{3.1.3}$$

where $\mathbf{x}_0$ is a constant vector. Similar to Theorem 2.1 we can show the following theorem.

Theorem 3.1 *Assume that $\mathbf{A}(t)$ and $\mathbf{g}(t)$ are continuous on an open interval $a < t < b$ containing t_0. Then, for any constant vector $\mathbf{x}_0$, (3.1.3) has a solution $\mathbf{x}(t)$ defined on this interval. This solution is unique.*
Especially, if $\mathbf{A}(t)$ and $\mathbf{g}(t)$ are continuous on $\mathbb{R}$, then for any $t_0 \in \mathbb{R}$ and $\mathbf{x}_0 \in \mathbb{R}^n$, (3.1.3) has a unique solution $\mathbf{x}(t)$ defined on $\mathbb{R}$.

3.2 Homogeneous Linear Systems

In this section we assume $\mathbf{A} = \Big(a_{ij}(t)\Big)$ is a continuous n by n matrix-valued function defined on the interval (a, b). We shall discuss the structure of the set of all solutions of (3.1.2).

Lemma 3.2 *Let $\mathbf{x}(t)$ and $\mathbf{y}(t)$ be two solutions of (3.1.2) on (a, b). Then for any numbers c_1, c_2, $\mathbf{z}(t) = c_1\mathbf{x}(t) + c_2\mathbf{y}(t)$ is also a solution of (3.1.2) on (a, b).*

Definition $\mathbf{x}_1(t), \cdots, \mathbf{x}_r(t)$ are *linearly dependent* in (a, b), if there exists numbers $c_1, \cdots, c_r$, not all zero, such that

$$c_1\mathbf{x}_1(t) + \cdots + c_r\mathbf{x}_r(t) = \mathbf{0} \qquad \text{for all} \quad t \in (a, b).$$

$\mathbf{x}_1(t), \cdots, \mathbf{x}_r(t)$ are *linearly independent* on (a, b) if they are not linearly dependent.

Lemma 3.3 *A set of solutions* $\mathbf{x}_1(t), \ \cdots, \ \mathbf{x}_r(t)$ *of (3.1.2) are linearly dependent on* (a, b) *if and only if* $\mathbf{x}_1(t_0), \ \cdots, \ \mathbf{x}_r(t_0)$ *are linearly dependent vectors for any fixed* $t_0 \in (a, b)$.

Proof. Obviously "$\Longrightarrow$" is true. We show "$\Longleftarrow$". Suppose that, for some $t_0 \in (a, b)$, $\mathbf{x}_1(t_0), \cdots$, $\mathbf{x}_r(t_0)$ are linearly dependent. Then there exists constants $c_1, \ \cdots, \ c_r$, not all zero, such that

$$c_1\mathbf{x}_1(t_0) + \cdots + c_r\mathbf{x}_r(t_0) = \mathbf{0}.$$

Let $\mathbf{y}(t) = c_1\mathbf{x}_1(t) + \cdots + c_r\mathbf{x}_r(t)$. Then $\mathbf{y}(t)$ is the solution of the initial value problem

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x}, \quad \mathbf{x}(t_0) = \mathbf{0}.$$

Since $\mathbf{x}(t) = \mathbf{0}$ is also a solution of the initial value problem, by the uniqueness we have $\mathbf{y}(t) \equiv \mathbf{0}$ on (a, b), i.e.

$$c_1\mathbf{x}_1(t) + \cdots + c_r\mathbf{x}_r(t) \equiv \mathbf{0}$$

on (a, b). Since c_j's are not all zero, $\mathbf{x}_1(t), \cdots, \mathbf{x}_r(t)$ are linearly dependent on (a, b). $\square$

Theorem 3.4 *(i) (3.1.2) has* n *linearly independent solutions.*
(ii) Let $\mathbf{x}_1, \ \cdots, \ \mathbf{x}_n$ *be any set of* n *linearly independent solutions of (3.1.2) on* (a, b). *Then the general solution of (3.1.2) is given by*

$$\mathbf{x}(t) = c_1\mathbf{x}_1(t) + \cdots + c_n\mathbf{x}_n(t), \tag{3.2.1}$$

where c_j*'s are arbitrary constants.*

Proof. (i) Let $\mathbf{e}_1, \ \cdots, \ \mathbf{e}_n$ be a set of linearly independent vectors in $\mathbb{R}^n$. Fix $t_0 \in (a, b)$. For each j from 1 to n, consider the initial value problem

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x}, \quad \mathbf{x}(t_0) = \mathbf{e}_j.$$

From Theorem 3.1, there exists a unique solution $\mathbf{x}_j(t)$ defined on (a, b). From Lemma 3.3, $\mathbf{x}_1(t), \ \cdots, \ \mathbf{x}_n(t)$ are linearly independent on (a, b).

(ii) Now let $\mathbf{x}_1(t), \ \cdots, \ \mathbf{x}_n(t)$ be any set of n linearly independent solutions of (3.1.2) on (a, b). Fix $t_0 \in (a, b)$. From Lemma 3.3, $\mathbf{x}_1(t_0), \cdots, \mathbf{x}_n(t_0)$ are linearly independent vectors. Let $\tilde{\mathbf{x}}(t)$ be any solution of (3.2.1). Then $\tilde{\mathbf{x}}(t_0)$ can be represented by a linear combination of $\mathbf{x}_1(t_0), \cdots$, $\mathbf{x}_n(t_0)$, namely, there exists n constants $\tilde{c}_1, \cdots, \tilde{c}_n$ such that

$$\tilde{\mathbf{x}}(t_0) = \tilde{c}_1\mathbf{x}_1(t_0) + \cdots + \tilde{c}_n\mathbf{x}_n(t_0).$$

As in the proof of Lemma 3.3, we can show that

$$\tilde{\mathbf{x}}(t) = \tilde{c}_1\mathbf{x}_1(t) + \cdots + \tilde{c}_n\mathbf{x}_n(t).$$

Thus $c_1\mathbf{x}_1(t) + \cdots + c_n\mathbf{x}_n(t)$ is the general solution of (3.1.2). $\square$

Recall that, n vectors

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ \cdots \\ a_{n1} \end{pmatrix}, \quad \cdots\cdots, \quad \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ \cdots \\ a_{nn} \end{pmatrix}$$

are linearly dependent if and only if the determinant

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = 0.$$

In order to check whether n solutions are linearly independent, we need the following notation.

Definition. The Wronskian of n vector-valued functions

$$\mathbf{x}_1(t) = \begin{pmatrix} x_{11}(t) \\ \cdots \\ x_{n1}(t) \end{pmatrix}, \quad \cdots, \quad \mathbf{x}_n(t) = \begin{pmatrix} x_{1n}(t) \\ \cdots \\ x_{nn}(t) \end{pmatrix}$$

is the determinant

$$W(t) \equiv W(\mathbf{x}_1, \cdots, \mathbf{x}_n)(t) = \begin{vmatrix} x_{11}(t) & x_{12}(t) & \cdots & x_{1n}(t) \\ x_{12}(t) & x_{22}(t) & \cdots & x_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}(t) & x_{n2}(t) & \cdots & x_{nn}(t) \end{vmatrix}.$$

Using Lemma 3.3 we can show that

Theorem 3.5 *(i) The Wronskian of n solutions of (3.1.2) is either identically zero or nowhere zero in (a, b).*
(ii) n solutions of (3.1.2) are linearly dependent in (a, b) if and only if their Wronskian is identically zero in (a, b).

Definition. A set of n linearly independent solutions of (3.1.2) is called a *fundamental set of solutions*, or a basis of solutions. Let

$$\mathbf{x}_1(t) = \begin{pmatrix} x_{11}(t) \\ \cdots \\ x_{n1}(t) \end{pmatrix}, \quad \cdots, \quad \mathbf{x}_n(t) = \begin{pmatrix} x_{1n}(t) \\ \cdots \\ x_{nn}(t) \end{pmatrix}$$

be a fundamental set of solutions of (3.1.2) on (a, b). The matrix-valued function

$$\Phi(t) = \begin{pmatrix} x_{11}(t) & x_{12}(t) & \cdots & x_{1n}(t) \\ x_{12}(t) & x_{22}(t) & \cdots & x_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}(t) & x_{n2}(t) & \cdots & x_{nn}(t) \end{pmatrix}$$

is called a *fundamental matrix* of (3.1.2) on (a, b).

Remark. (i) From Theorem 3.5, a fundamental matrix is non-singular for all $t \in (a, b)$.
(ii) A fundamental matrix $\Phi(t)$ satisfies the following *matrix equation*:

$$\Phi' = \mathbf{A}(t)\Phi. \tag{3.2.2}$$

(iii) Let $\Phi(t)$ and $\Psi(t)$ are two fundamental matrices defined on (a, b). Then there exists a constant, non-singular matrix $\mathbf{C}$ such that

$$\Psi(t) = \Phi(t)\mathbf{C}.$$

Theorem 3.6 *Let $\Phi(t)$ be a fundamental matrix of (3.1.2) on (a, b). Then the general solution of (3.1.2) is given by*

$$\mathbf{x}(t) = \Phi(t)\mathbf{c}, \tag{3.2.3}$$

where $\mathbf{c} = \begin{pmatrix} c_1 \\ \cdots \\ c_n \end{pmatrix}$ *is an arbitrary constant vector.*

3.3 Non-Homogeneous Linear Systems

In this section we consider the solutions of non-homogeneous system (3.1.1), where $\mathbf{A} = \Big(a_{ij}(t)\Big)$ is a continuous n by n matrix-valued function and $\mathbf{g}(t)$ is a continuous vector-valued function, both defined on the interval (a, b).

Theorem 3.7 *Let $\mathbf{x}_p(t)$ be a particular solution of (3.1.1), and $\Phi(t)$ be a fundamental matrix of the associated homogeneous system (3.1.2). Then the general solution of (3.1.1) is given by*

$$\mathbf{x}(t) = \Phi(t)\mathbf{c} + \mathbf{x}_p(t), \tag{3.3.1}$$

where $\mathbf{c}$ is an arbitrary constant vector.

Proof. For any constant vector $\mathbf{c}$, $\mathbf{x}(t) = \Phi(t)\mathbf{c} + \mathbf{x}_p(t)$ is a solution of (3.1.1). On the other hand, let $\mathbf{x}(t)$ be a solution of (3.1.1) and set $\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}_p(t)$. Then $\mathbf{y}' = \mathbf{A}(t)\mathbf{y}$. From (3.2.3), there exists a constant vector $\tilde{\mathbf{c}}$ such that $\mathbf{y}(t) = \Phi(t)\tilde{\mathbf{c}}$. So $\mathbf{x}(t) = \Phi(t)\tilde{\mathbf{c}} + \mathbf{x}_p(t)$. Thus (3.3.1) gives a general solution of (3.1.1). □

Method of variation of parameters.
Let $\mathbf{\Phi}$ be a fundamental matrix of (3.1.2). We look for a particular solution of (3.1.1) in the form

$$\mathbf{x}(t) = \Phi(t)\mathbf{u}(t), \qquad \mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ \cdots \\ u_n(t) \end{pmatrix}.$$

Plugging into (3.1.1) we get

$$\mathbf{\Phi}'\mathbf{u} + \mathbf{\Phi}\mathbf{u}' = \mathbf{A}\mathbf{\Phi}\mathbf{u} + \mathbf{g}.$$

From (3.2.2), $\mathbf{\Phi}' = \mathbf{A\Phi}$. So $\mathbf{\Phi u}' = \mathbf{g}$, and thus $\mathbf{u}' = \mathbf{\Phi}^{-1}\mathbf{g}$.

$$\mathbf{u}(t) = \int_{t_0}^{t} \mathbf{\Phi}^{-1}(s)\mathbf{g}(s)ds + \mathbf{c}. \tag{3.3.2}$$

Choosing $\mathbf{c} = \mathbf{0}$, we get a particular solution:

$$\mathbf{x}_p(t) = \mathbf{\Phi}(t) \int_{t_0}^{t} \mathbf{\Phi}^{-1}(s)\mathbf{g}(s)ds.$$

So we obtain the following:

Theorem 3.8 *The general solution of (3.1.1) is given by*

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{c} + \mathbf{\Phi}(t) \int_{t_0}^{t} \mathbf{\Phi}^{-1}(s)\mathbf{g}(s)ds, \tag{3.3.3}$$

where $\Phi(t)$ is a fundamental matrix of the associated homogeneous system (3.1.2).

Example. Solve $\begin{cases} x_1' &= 3x_1 - x_2 + t \\ x_2' &= 2x_1 + t \end{cases}$.

3.4 Homogeneous Linear Systems with Constant Coefficients

Consider a homogeneous linear system

$$\mathbf{x}' = \mathbf{A}\mathbf{x}, \tag{3.4.1}$$

where $\mathbf{A} = (a_{ij})$ is a constant n by n matrix.

Let us try to find a solution of (3.4.1) in the form $\mathbf{x}(t) = e^{\lambda t}\mathbf{k}$, where $\mathbf{k}$ is a constant vector, $\mathbf{k} \neq \mathbf{0}$. Plugging it into (3.4.1) we find

$$\mathbf{A}\mathbf{k} = \lambda\mathbf{k}. \tag{3.4.2}$$

Definition. Assume that a number λ and a vector $\mathbf{k} \neq \mathbf{0}$ satisfy (3.4.2), then we call λ an eigenvalue of $\mathbf{A}$, and $\mathbf{k}$ an eigenvector associated with λ.

Lemma 3.9 *λ is an eigenvalue of $\mathbf{A}$ if and only if*

$$\det\,(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{3.4.3}$$

(where $\mathbf{I}$ is the $n \times n$ unit matrix), namely,

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Remark. Let $\mathbf{A}$ be an n by n matrix and λ_1, λ_2, $\cdots$, λ_k be the distinct roots of (3.4.3). Then there exist positive integers m_1, m_2, $\cdots$, m_k, such that

$$\det(\mathbf{A}-\lambda\mathbf{I}) = (-1)^n(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k},$$

and

$$m_1 + m_2 + \cdots + m_k = n.$$

m_j is called the *algebraic multiplicity* (or simply *multiplicity*) of the eigenvalue λ_j. The number of linearly independent eigenvectors of $\mathbf{A}$ associated with λ_j is called the *geometric multiplicity* of the eigenvalue λ_j and is denoted by $\mu(\lambda_j)$. We always have

$$\mu(\lambda_j) \le m_j.$$

If $\mu(\lambda_j) = m_j$ then we say that the eigenvalue λ_j is *quasi-simple.* Especially if $m_j = 1$ we say that λ_j is a *simple* eigenvalue. Note that in this case λ_j is a simple root of (3.4.3).

Theorem 3.10 *If λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{k}$ is an associated eigenvector, then*

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{k}$$

is a solution of (3.4.1).
Let $\mathbf{A}$ be a real matrix. If λ is a complex eigenvalue of $\mathbf{A}$, and $\mathbf{k}$ is an eigenvector associated with λ, then

$$\mathbf{x}_1 = \Re(e^{\lambda t}\mathbf{k}), \quad \mathbf{x}_2 = \Im(e^{\lambda t}\mathbf{k})$$

are two linearly independent real solutions of (3.4.1).

In the following we always assume that $\mathbf{A}$ is a real matrix.

Theorem 3.11 *If $\mathbf{A}$ has n linearly independent eigenvectors $\mathbf{k}_1$, $\cdots$, $\mathbf{k}_n$ associated with eigenvalues λ_1, $\cdots$, λ_n respectively, then*

$$\Phi(t) = (e^{\lambda_1 t}\mathbf{k}_1, \cdots, e^{\lambda_n t}\mathbf{k}_n)$$

is a fundamental matrix of (3.4.1), and the general solution is given by

$$\mathbf{x}(t) = \Phi(t)\mathbf{c} = c_1 e^{\lambda_1 t}\mathbf{k}_1 + \cdots + c_n e^{\lambda_n t}\mathbf{k}_n, \tag{3.4.4}$$

where $\mathbf{c} = \begin{pmatrix} c_1 \\ \cdots \\ c_n \end{pmatrix}$ is an arbitrary constant vector.

Proof. We only need to show $\det\Phi(t) \neq 0$. Since $\mathbf{k}_1$, $\cdots$, $\mathbf{k}_n$ are linearly independent, so $\det\Phi(0) \neq 0$. From Theorem 3.5 we see that $\det\Phi(t) \neq 0$ for any t. Hence $\Phi(t)$ is a fundamental matrix. □

Remark. Under the conditions of Theorem 3.11, the eigenvalues $\lambda_1, \cdots, \lambda_n$ of $\mathbf{A}$ need not to be distinct. In fact we only assume that all the eigenvalues of $\mathbf{A}$ are quasi-simple.
If $\mathbf{A}$ has n distinct eigenvalues $\lambda_1, \cdots, \lambda_n$, and let $\mathbf{k}_1, \cdots, \mathbf{k}_n$ be the associated eigenvectors, then they are linearly independent. Hence the general solution is given by (3.4.4).

Example. $\mathbf{x}' = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix} \mathbf{x}$.
$\mathbf{A} = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix}$ has eigenvalues $\lambda_1 = -2$, and $\lambda_2 = -4$.
For $\lambda_1 = -2$ we find an eigenvector $\mathbf{k}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
For $\lambda_2 = -4$ we find an eigenvector $\mathbf{k}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
The general solution is given by

$$\mathbf{x}(t) = c_1 e^{-2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-4t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Example. Solve the system

$$\mathbf{x}' = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix} \mathbf{x} + e^{-2t} \begin{pmatrix} -6 \\ 2 \end{pmatrix}.$$

We first solve the associated homogeneous system

$$\mathbf{x}' = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix} \mathbf{x}$$

and find two linearly independent solutions $\mathbf{x}_1(t) = \begin{pmatrix} e^{-2t} \\ e^{-2t} \end{pmatrix}$, $\mathbf{x}_2(t) = \begin{pmatrix} e^{-4t} \\ -e^{-4t} \end{pmatrix}$. The fundamental matrix is

$$\mathbf{\Phi} = (\mathbf{x}_1(t), \mathbf{x}_2(t)) = \begin{pmatrix} e^{-2t} & e^{-4t} \\ e^{-2t} & -e^{-4t} \end{pmatrix}.$$

$$\mathbf{\Phi}^{-1} = \frac{1}{-2e^{-6t}} \begin{pmatrix} -e^{-4t} & -e^{-4t} \\ -e^{-2t} & e^{-2t} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{2t} & e^{2t} \\ e^{4t} & -e^{4t} \end{pmatrix},$$

Let $\mathbf{g}(t) = e^{-2t} \begin{pmatrix} -6 \\ 2 \end{pmatrix}$. Then

$$\mathbf{\Phi}^{-1}(t)\mathbf{g}(t) = \begin{pmatrix} -2 \\ -4e^{2t} \end{pmatrix},$$

$$\mathbf{u}(t) = \int_0^t \Phi^{-1}(s)\mathbf{g}(s)ds = \begin{pmatrix} -2t \\ -2e^{2t} + 2 \end{pmatrix},$$

$$\Phi(t)\mathbf{u}(t) = 2e^{-2t}\begin{pmatrix} -t-1 \\ -t+1 \end{pmatrix} + 2e^{-4t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The general solution is

$$\mathbf{x} = c_1 e^{-2t}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-4t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} - 2te^{-2t}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2e^{-2t}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Example. Solve $\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}\mathbf{x}$.

$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}$ has eigenvalues $\pm 2i$.

For $\lambda = 2i$ we find an eigenvector $\mathbf{k} = \begin{pmatrix} 1 \\ 2i \end{pmatrix}$.

$$e^{2it}\begin{pmatrix} 1 \\ 2i \end{pmatrix} = (\cos 2t + i\sin 2t)\begin{pmatrix} 1 \\ 2i \end{pmatrix} = \begin{pmatrix} \cos 2t \\ -2\sin 2t \end{pmatrix} + i\begin{pmatrix} \sin 2t \\ 2\cos 2t \end{pmatrix}.$$

The general solution is given by

$$\mathbf{x}(t) = c_1\begin{pmatrix} \cos 2t \\ -2\sin 2t \end{pmatrix} + c_2\begin{pmatrix} \sin 2t \\ 2\cos 2t \end{pmatrix}.$$

Example. Solve

$$\begin{cases} x' = -3x + 4y - 2z, \\ y' = x + z, \\ z' = 6x - 6y + 5z. \end{cases}$$

$\mathbf{A} = \begin{pmatrix} -3 & 4 & -2 \\ 1 & 0 & 1 \\ 6 & -6 & 5 \end{pmatrix}$ has eigenvalues $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = -1$.

For $\lambda_1 = 2$ we find an eigenvector $\mathbf{k}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

For $\lambda_2 = 1$ we find an eigenvector $\mathbf{k}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

For $\lambda_3 = -1$ we find an eigenvector $\mathbf{k}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

The general solution is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 e^{2t}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + c_2 e^{t}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_3 e^{-t}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

namely

$$\begin{aligned} x(t) &= c_2 e^t + c_3 e^{-t}, \\ y(t) &= c_1 e^{2t} + c_2 e^t, \\ z(t) &= 2c_1 e^{2t} - c_3 e^{-t}. \end{aligned}$$

Example. Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ -1 & 2 & 3 \end{pmatrix}.$$

$\mathbf{A}$ has eigenvalues $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = 3 \pm i$.

For $\lambda_1 = 2$ we find an eigenvector $\mathbf{k}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

For $\lambda_2 = 3 + i$ we find an eigenvector $\mathbf{k}_2 = \begin{pmatrix} 1 \\ 1+i \\ 2-i \end{pmatrix}$.

We have

$$\begin{aligned} e^{(3+i)t}\mathbf{k}_2 &= e^{3t} \begin{pmatrix} \cos t + i \sin t \\ \cos t - \sin t + i(\cos t + \sin t) \\ 2\cos t + \sin t + i(2\sin t - \cos t) \end{pmatrix}, \\ \Re(e^{(3+i)t}\mathbf{k}_2) &= e^{3t} \begin{pmatrix} \cos t \\ \cos t - \sin t \\ 2\cos t + \sin t \end{pmatrix} \\ \Im(e^{(3+i)t}\mathbf{k}_2) &= e^{3t} \begin{pmatrix} \sin t \\ \cos t + \sin t \\ 2\sin t - \cos t \end{pmatrix}. \end{aligned}$$

The general solution is

$$\mathbf{x}(t) = c_1 e^{2t} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 e^{3t} \begin{pmatrix} \cos t \\ \cos t - \sin t \\ 2\cos t + \sin t \end{pmatrix} + c_3 e^{3t} \begin{pmatrix} \sin t \\ \cos t + \sin t \\ 2\sin t - \cos t \end{pmatrix}.$$

Example. Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where

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

We have

$$\det(\mathbf{A} - \lambda\mathbf{I}) = -(\lambda - 3)^2(\lambda + 3).$$

$\mathbf{A}$ has eigenvalues $\lambda_1 = \lambda_2 = 3$, $\lambda_3 = -3$ (We may say that, $\lambda = 3$ is an eigenvalue of algebraic multiplicity 2, and $\lambda = -3$ is a simple eigenvalue).
For $\lambda = 3$ we solve the equation $\mathbf{Ak} = 3\mathbf{k}$, namely

$$\begin{pmatrix} -2 & -2 & 2 \\ -2 & -2 & 2 \\ 2 & 2 & -2 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \mathbf{0}.$$

The solution is $\mathbf{k} = \begin{pmatrix} v-u \\ u \\ v \end{pmatrix}$. So we find two eigenvectors $\mathbf{k}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ and $\mathbf{k}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

For $\lambda_3 = -3$ we find an eigenvector $\mathbf{k}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

The general solution is given by

$$\mathbf{x}(t) = c_1 e^{3t} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 e^{3t} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c_3 e^{-3t} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Now we consider the solutions of (3.4.1) associated with a multiple eigenvalue λ, with geometric multiplicity $\mu(\lambda)$ less than the algebraic multiplicity.

Lemma 3.12 *Assume λ is an eigenvalue of $\mathbf{A}$ with algebraic multiplicity $m > 1$. Then the following system*

$$(\mathbf{A} - \lambda\mathbf{I})^m \mathbf{v} = \mathbf{0} \tag{3.4.5}$$

has exactly m linearly independent solutions.

By direct computations we can prove the following theorem.

Theorem 3.13 *Assume that λ is an eigenvalue of $\mathbf{A}$ with algebraic multiplicity $m > 1$. Let $\mathbf{v}_0 \neq \mathbf{0}$ be a solutions of (3.4.5). Define*

$$\mathbf{v}_l = (\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_{l-1}, \quad l = 1, 2, \cdots, m-1, \tag{3.4.6}$$

and let

$$\mathbf{x}(t) = e^{\lambda t} \left[\mathbf{v}_0 + t\mathbf{v}_1 + \frac{t^2}{2}\mathbf{v}_2 + \cdots + \frac{t^{m-1}}{(m-1)!}\mathbf{v}_{m-1} \right]. \tag{3.4.7}$$

Then $\mathbf{x}(t)$ is a solution of (3.4.1).
Let $\mathbf{v}_0^{(1)}, \cdots, \mathbf{v}_0^{(m)}$ be m linearly independent solutions of (3.4.5). They generate m linearly independent solutions of (3.4.1) via (3.4.6) and (3.4.7).

Remark. In (3.4.6), we always have

$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{v}_{m-1}=\mathbf{0}.$$

If $\mathbf{v}_{m-1}\neq\mathbf{0}$ then $\mathbf{v}_{m-1}$ is an eigenvector of $\mathbf{A}$ associated with the eigenvalue λ.

In practice, to find the solutions of (3.4.1) associated with an eigenvalue λ of multiplicity m, we first solve (3.4.5) and find m linearly independent solutions

$$\mathbf{v}_0^{(1)},\quad \mathbf{v}_0^{(2)},\quad \cdots,\quad \mathbf{v}_0^{(m)}.$$

For each of these vectors, say $\mathbf{v}_0^{(k)}$, we compute the iteration sequence

$$\mathbf{v}_l^{(k)}=(\mathbf{A}-\lambda\mathbf{I})\mathbf{v}_{l-1}^{(k)},\quad l=1,2,\cdots$$

There is an integer $0\le j\le m-1$ (j depends on the choice of $\mathbf{v}_0^{(k)}$) such that

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

Thus $\mathbf{v}_j$ is an eigenvector of $\mathbf{A}$ associated with the eigenvalue λ. Then the iteration stops and yields a solution

$$\mathbf{x}^{(k)}(t)=e^{\lambda t}\left[\mathbf{v}_0^{(k)}+t\mathbf{v}_1^{(k)}+\frac{t^2}{2}\mathbf{v}_2^{(k)}+\cdots+\frac{t^j}{j!}\mathbf{v}_j^{(k)}\right]. \tag{3.4.8}$$

Example. Solve $\mathbf{x}'=\mathbf{A}\mathbf{x}$, where

$$\mathbf{A}=\begin{pmatrix}-1&1&0\\0&-1&4\\1&0&-4\end{pmatrix}.$$

From $\det(\mathbf{A}-\lambda\mathbf{I})=-\lambda(\lambda+3)^2=0$ we find eigenvalues $\lambda_1=-3$ with multiplicity 2, and $\lambda_2=0$ simple.
For the double eigenvalue $\lambda_1=-3$ we solve

$$(\mathbf{A}+3\mathbf{I})^2\mathbf{v}=\begin{pmatrix}4&4&4\\4&4&4\\1&1&1\end{pmatrix}\mathbf{v}=\mathbf{0},$$

and find two linearly independent solutions $\mathbf{v}_0^{(1)}=\begin{pmatrix}1\\0\\-1\end{pmatrix}$, $\mathbf{v}_0^{(2)}=\begin{pmatrix}0\\1\\-1\end{pmatrix}$. Plugging $\mathbf{v}_0^{(1)}$, $\mathbf{v}_0^{(2)}$ into (3.4.6), (3.4.7) we get

$$\mathbf{v}_1^{(1)}=(\mathbf{A}+3\mathbf{I})\mathbf{v}_0^{(1)}=\begin{pmatrix}2\\-4\\2\end{pmatrix},$$

$$\mathbf{x}^{(1)}=e^{-3t}(\mathbf{v}_0^{(1)}+t\mathbf{v}_1^{(1)})=e^{-3t}\left[\begin{pmatrix}1\\0\\-1\end{pmatrix}+t\begin{pmatrix}2\\-4\\2\end{pmatrix}\right],$$

$$\mathbf{v}_1^{(2)} = (\mathbf{A} + 3\mathbf{I})\mathbf{v}_0^{(2)} = \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix},$$

$$\mathbf{x}^{(2)} = e^{-3t}(\mathbf{v}_0^{(2)} + t\mathbf{v}_1^{(2)}) = e^{-3t}\left[\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + t\begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}\right].$$

For the simple eigenvalue $\lambda_2 = 0$ we find an eigenvector $\mathbf{k}_3 = \begin{pmatrix} 4 \\ 4 \\ 1 \end{pmatrix}$.

So the general solution is

$$\begin{aligned}\mathbf{x}(t) &= c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + c_3\mathbf{k}_3 \\ &= c_1e^{-3t}\left[\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t\begin{pmatrix} 2 \\ -4 \\ 2 \end{pmatrix}\right] + c_2e^{-3t}\left[\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + t\begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}\right] + c_3\begin{pmatrix} 4 \\ 4 \\ 1 \end{pmatrix}.\end{aligned}$$

Example. Solve the system

$$\begin{cases} x' = 2x + y + 2z, \\ y' = -x + 4y + 2z, \\ z' = 3z. \end{cases}$$

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ -1 & 4 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

The eigenvalue is $\lambda = 3$ with multiplicity 3. Solving

$$(\mathbf{A} - 3\mathbf{I})^3\mathbf{v} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{v} = \mathbf{0},$$

we obtain 3 obvious linearly independent solutions

$$\mathbf{v}_0^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{v}_0^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_0^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Plugging $\mathbf{v}_0^{(j)}$ into (3.4.6), (3.4.7) we get

$$\mathbf{v}_1^{(1)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_0^{(1)} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix},$$

$$\mathbf{v}_2^{(1)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_1^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\mathbf{x}^{(1)} = e^{3t}(\mathbf{v}_0^{(1)} + t\mathbf{v}_1^{(1)}) = e^{3t}\left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}\right];$$

$$\mathbf{v}_1^{(2)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_0^{(2)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$\mathbf{v}_2^{(2)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_1^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\mathbf{x}^{(2)} = e^{3t}(\mathbf{v}_0^{(2)} + t\mathbf{v}_1^{(2)}) = e^{3t}\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right];$$

$$\mathbf{v}_1^{(3)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_0^{(3)} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix},$$

$$\mathbf{v}_2^{(3)} = (\mathbf{A} - 3\mathbf{I})\mathbf{v}_1^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\mathbf{x}^{(3)} = e^{3t}(\mathbf{v}_0^{(3)} + t\mathbf{v}_1^{(3)}) = e^{3t}\left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}\right].$$

The general solution is

$$\begin{aligned}\mathbf{x}(t) &= c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + c_3\mathbf{x}^{(3)} \\ &= c_1 e^{3t}\left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}\right] + c_2 e^{3t}\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right] \\ &+ c_3 e^{3t}\left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}\right].\end{aligned}$$

Remark. It is possible to reduce the number of constant vectors in the general solution of $\mathbf{x}' = \mathbf{A}\mathbf{x}$ by using a basis for the Jordan canonical form of $\mathbf{A}$. We will not go into the details of the Jordan canonical form. However the following algorithm usually works well if the size of $\mathbf{A}$ is small.

Consider an eigenvalue λ of $\mathbf{A}$ with algebraic multiplicity m.

Start with $r = m$. Let $\mathbf{v}$ be a vector such that $(\mathbf{A}-\lambda\mathbf{I})^r\mathbf{v} = \mathbf{0}$ but $(\mathbf{A}-\lambda\mathbf{I})^{r-1}\mathbf{v} \neq \mathbf{0}$. [$\mathbf{v}$ is called a generalized eigenvector of rank r associated with the eigenvalue λ. If no such $\mathbf{v}$ exists, reduce r by 1.] Then

$$\mathbf{u}_r = \mathbf{v}, \mathbf{u}_{r-1} = (\mathbf{A}-\lambda\mathbf{I})\mathbf{v}, \mathbf{u}_{r-2} = (\mathbf{A}-\lambda\mathbf{I})^2\mathbf{v}, \ldots, \mathbf{u}_1 = (\mathbf{A}-\lambda\mathbf{I})^{r-1}\mathbf{v},$$

form a chain of linearly independent solutions of (3.4.5) with $\mathbf{u}_1$ being the base eigenvector corresponding to the eigenvalue λ. This gives r independent solutions of $\mathbf{x}' = \mathbf{A}\mathbf{x}$:

$$\begin{aligned}
\mathbf{x}_1(t) &= \mathbf{u}_1 e^{\lambda t}, \\
\mathbf{x}_2(t) &= (\mathbf{u}_1 t + \mathbf{u}_2)e^{\lambda t}, \\
\mathbf{x}_3(t) &= (\tfrac{1}{2}\mathbf{u}_1 t^2 + \mathbf{u}_2 t + \mathbf{u}_3)e^{\lambda t}, \\
&\;\,\cdot \\
&\;\,\cdot \\
&\;\,\cdot \\
\mathbf{x}_r(t) &= (\tfrac{1}{(r-1)!}\mathbf{u}_1 t^{r-1} + \cdots + \tfrac{1}{2!}\mathbf{u}_{r-2}t^2 + \mathbf{u}_{r-1}t + \mathbf{u}_r)e^{\lambda t}.
\end{aligned}$$

Repeat this procedure by finding another $\mathbf{v}$ which is not in the span of the previous chains of vectors. Also do this for each eigenvalue of $\mathbf{A}$. Results of linear algebra shows that
(1) Any chain of generalized eigenvectors constitutes a linearly independent set of vectors.
(2) If two chains of generalized eigenvectors are based on linearly independent eigenvectors, then the union of these vectors is a linearly independent set of vectors (whether the two base eigenvectors are associated with different eigenvalues or with the same eigenvalue).

Example. Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where $\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$.

$\mathbf{A}$ has an eigenvalue $\lambda = 3$ of multiplicity 4. Direct calculation gives $(\mathbf{A}-3\mathbf{I}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$,

$(\mathbf{A}-3\mathbf{I})^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$, $(\mathbf{A}-3\mathbf{I})^3 = \mathbf{0}$, and $(\mathbf{A}-3\mathbf{I})^4 = \mathbf{0}$.

It can be seen that $\mathbf{v_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{v_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ are two linearly independent eigenvectors of $\mathbf{A}$.

Together with $\mathbf{v_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{v_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, they form a basis of $\{\mathbf{v} \mid (\mathbf{A}-3\mathbf{I})^4\mathbf{v} = \mathbf{0}\} = \mathbb{R}^4$.

Note that $(\mathbf{A}-3\mathbf{I})\mathbf{v_2} = \mathbf{v_3}$, and $(\mathbf{A}-3\mathbf{I})\mathbf{v_3} = \mathbf{v_4}$. Hence $\{\mathbf{v_4}, \mathbf{v_3}, \mathbf{v_2}\}$ forms a chain of generalized eigenvectors associated with the eigenvalue 3. $\{\mathbf{v_1}\}$ alone is another chain. Therefore the general solution is

$$\mathbf{x}(t) = e^{3t}\left(c_1\mathbf{v_1} + c_2(\mathbf{v_2} + t\mathbf{v_3} + \frac{t^2}{2}\mathbf{v_4}) + c_3(\mathbf{v_3} + t\mathbf{v_4}) + c_4\mathbf{v_4}\right).$$

That is

$$\mathbf{x}(t) = \begin{pmatrix} c_1e^{3t} \\ c_2e^{3t} \\ (c_2t + c_3)e^{3t} \\ (\frac{c_2t^2}{2} + c_3t + c_4)e^{3t} \end{pmatrix}.$$

Exercise. Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{pmatrix} 7 & 5 & -3 & 2 \\ 0 & 1 & 0 & 0 \\ 12 & 10 & -5 & 4 \\ -4 & -4 & 2 & -1 \end{pmatrix}.$$

Ans: $\mathbf{x}(t) = c_1e^{-t}\begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} + c_2e^{t}\begin{pmatrix} 1 \\ -2 \\ 0 \\ 2 \end{pmatrix} + c_3e^{t}\begin{pmatrix} -1 \\ 0 \\ -2 \\ 0 \end{pmatrix} + c_4e^{t}\begin{pmatrix} -t \\ 0 \\ 1-2t \\ 1 \end{pmatrix}.$

3.5 Higher Order Linear Equations

Consider n-th order linear equation

$$y^{(n)} + a_1(t)y^{(n-1)} + \cdots + a_{n-1}(t)y' + a_n(t)y = f(t), \tag{3.5.1}$$

where $y^{(k)} = \frac{d^ky}{dt^k}$. Throughout this section we assume that $a_j(t)$'s and $f(t)$ are continuous functions defined on the interval (a, b). When $f(t) \not\equiv 0$, (3.5.1) is called a non-homogeneous equation. The associated homogeneous equation is

$$y^{(n)} + a_1(t)y^{(n-1)} + \cdots + a_{n-1}(t)y' + a_n(t)y = 0. \tag{3.5.2}$$

The general theory of solutions of (3.5.1) and (3.5.2) can be established by applying the results in the previous sections to the equivalent systems.
We begin with the initial value problem

$$\begin{cases} y^{(n)} + a_1(t)y^{(n-1)} + \cdots + a_n(t)y = f(t), \\ y(t_0) = y_0, \\ y'(t_0) = y_1, \\ \cdots\cdots\cdots \\ y^{(n-1)}(t_0) = y_{n-1}. \end{cases} \tag{3.5.3}$$

Theorem 3.14 *Assume that $a_1(t)$, $\cdots$, $a_n(t)$ and $f(t)$ are continuous functions defined on the interval (a,b). Then for any $t_0 \in (a,b)$ and for any numbers y_0, $\cdots$, y_{n-1}, the initial value problem (3.5.3) has a unique solution defined on (a,b).*
Especially if $a_j(t)$'s and $f(t)$ are continuous on $\mathbb{R}$, then for any t_0 and y_0, $\cdots$, y_{n-1}, the initial value problem (3.5.3) has a unique solution defined on $\mathbb{R}$.

Next we consider the structure of solutions of the homogeneous equation (3.5.2).

Definition. Functions $\phi_1(t)$, $\cdots$, $\phi_r(t)$ are *linearly dependent* on (a,b) if there exists constants c_1, $\cdots$, c_r, not all zero, such that

$$c_1\phi_1(t) + \cdots + c_r\phi_r(t) = 0$$

for all $t \in (a,b)$. A set of functions are *linearly independent* on (a,b) if they are not linearly dependent on (a,b).

Lemma 3.15 *Functions $\phi_1(t), \cdots, \phi_r(t)$ are linearly dependent on (a,b) if and only if the following vector-valued functions*

$$\begin{pmatrix} \phi_1(t) \\ \phi_1'(t) \\ \cdots \\ \phi_1^{(n-1)}(t) \end{pmatrix}, \cdots, \begin{pmatrix} \phi_r(t) \\ \phi_r'(t) \\ \cdots \\ \phi_r^{(n-1)}(t) \end{pmatrix}$$

are linearly dependent on (a,b).

Proof. "$\Longleftarrow$" is obvious. To show "$\Longrightarrow$", assume that ϕ_1, $\cdots$, ϕ_r are linearly dependent on (a,b). There exists constants c_1, $\cdots$, c_r, not all zero, such that

$$c_1\phi_1(t) + \cdots + c_r\phi_r(t) = 0$$

for all $t \in (a,b)$. Differentiate this equality successively we find

$$\begin{aligned} &c_1\phi_1'(t) + \cdots + c_r\phi_r'(t) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &c_1\phi_1^{(n-1)}(t) + \cdots + c_r\phi_r^{(n-1)}(t) = 0. \end{aligned}$$

Thus

$$c_1 \begin{pmatrix} \phi_1(t) \\ \phi_1'(t) \\ \cdots \\ \phi_1^{(n-1)}(t) \end{pmatrix} + \cdots + c_r \begin{pmatrix} \phi_r(t) \\ \phi_r'(t) \\ \cdots \\ \phi_r^{(n-1)}(t) \end{pmatrix} = \mathbf{0}$$

for all $t \in (a,b)$. Hence the r vector-valued functions are linearly dependent on (a,b). $\square$

The Wronskian of n functions $\phi_1(t)$, $\cdots$, ϕ_n is defined by

$$W(\phi_1, \cdots, \phi_n)(t) = \begin{vmatrix} \phi_1(t) & \cdots & \phi_n(t) \\ \cdots & \cdots & \cdots \\ \phi_1^{(n-1)}(t) & \cdots & \phi_n^{(n-1)}(t) \end{vmatrix}. \qquad (3.5.4)$$

From Lemma 3.15 and Theorem 3.5 we get

Proposition 3.5.1 *Let $y_1(t)$, $\cdots$, $y_n(t)$ be n solutions of (3.5.2) on (a, b). They are linearly independent on (a, b) if and only if their Wronskian $W(t) \equiv W(y_1, \cdots, y_n)(t)$ does not vanish on (a, b).*

Theorem 3.16 *Let $a_1(t)$, $\cdots$, $a_n(t)$ be continuous on the interval (a, b). The homogeneous equation (3.5.2) has n linearly independent solutions on (a, b).*
Let $y_1, \cdots, y_n$ be n linearly independent solutions of (3.5.2) defined on (a, b). The general solution of (3.5.2) is given by

$$y(t) = c_1 y_1(t) + \cdots + c_n y_n(t), \tag{3.5.5}$$

where $c_1, \cdots, c_n$ are arbitrary constants.

Any set of n linearly independent solutions is called a *fundamental set of solutions*.
Now we consider the non-homogeneous equation (3.5.1). We have

Theorem 3.17 *Let y_p be a particular solution of (3.5.1), and $y_1, \cdots, y_n$ be a fundamental set of solutions for the associated homogeneous equation (3.5.2). The general solution of (3.5.1) is given by*

$$y(t) = c_1 y_1(t) + \cdots + c_n y_n(t) + y_p(t). \tag{3.5.6}$$

From (3.3.3) we can derive the variation of parameter formula for higher order equations. Consider a second order equation

$$y'' + p(t)y' + q(t)y = f(t). \tag{3.5.7}$$

Let $x_1 = y$, $x_2 = y'$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then (3.5.7) is written as

$$\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ f \end{pmatrix} \tag{3.5.8}$$

Assume $y_1(t)$ and $y_2(t)$ are two linearly independent solutions of the associated homogeneous equation

$$y'' + py' + qy = 0.$$

We look for a solution of (3.5.7) in the form

$$y = u_1 y_1 + u_2 y_2.$$

Choose a fundamental matrix $\Phi(t) = \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$. The corresponding solution of (3.5.8) is in the form

$$\mathbf{x} = \begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} y_1 u_1 + y_2 u_2 \\ y_1' u_1 + y_2' u_2 \end{pmatrix} \tag{3.5.9}$$

Recall that, if $ad - bc \neq 0$, then

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

Thus

$$\Phi(t)^{-1} = \frac{1}{W(t)}\begin{pmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{pmatrix},$$

where $W(t)$ is the Wronskian of $y_1(t)$, $y_2(t)$. Using (3.3.3) we can derive

$$u_1(t) = -\int \frac{y_2(t)}{W(t)} f dt, \qquad u_2(t) = \int \frac{y_1(t)}{W(t)} f(t) dt. \tag{3.5.10}$$

Note that, (3.5.9) implies

$$y_1'u_1 + y_2'u_2 = y' = y_1'u_1 + y_2'u_2 + y_1u_1' + y_2u_2'.$$

Hence

$$y_1u_1' + y_2u_2' = 0. \tag{i}$$

Plugging $y' = y_1'u_1 + y_2'u_2$ into (3.5.7) we find

$$y_1'u_1' + y_2'u_2' = f. \tag{ii}$$

Solving (i) (ii) we find

$$u_1' = -\frac{y_2}{W} f, \quad u_2' = \frac{y_1}{W} f.$$

Again we get (3.5.10).

Now we consider linear equations with constant coefficients

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = f(t), \tag{3.5.11}$$

and the associated homogeneous equation

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0, \tag{3.5.12}$$

where $a_1, \cdots, a_n$ are real constants. Recall that (3.5.12) is equivalent to a system

$$\mathbf{x}' = \mathbf{A}\mathbf{x},$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

The equation for the eigenvalues of $\mathbf{A}$ is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0. \tag{3.5.13}$$

The solutions of (3.5.13) are called *characteristic values* or eigenvalues for the equation (3.5.12).

Let $\lambda_1, \cdots, \lambda_s$ be the distinct eigenvalues for (3.5.12). Then we can write

$$\begin{aligned}&\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \\ &= (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s},\end{aligned}$$

where $m_1, \cdots, m_s$ are positive integers and

$$m_1 + \cdots + m_s = n.$$

We call them the multiplicity of the eigenvalues $\lambda_1, \cdots, \lambda_s$ respectively.

Lemma 3.18 *Assume λ is an eigenvalue of (3.5.12) of multiplicity m.*
(i) $e^{\lambda t}$ is a solution of (3.5.12).
(ii) If $m > 1$, then for any positive integer $1 \le k \le m-1$, $t^k e^{\lambda t}$ is a solution of (3.5.12).
(iii) If $\lambda = \alpha + i\beta$, then $t^k e^{\alpha t}\cos(\beta t),\ t^k e^{\alpha t}\sin(\beta t)$ are solutions of (3.5.12), where $0 \le k \le m-1$.

Theorem 3.19 *Let $\lambda_1, \cdots, \lambda_s$ be the distinct eigenvalues for (3.5.12), with multiplicity $m_1, \cdots, m_s$ respectively. Then (3.5.12) has a fundamental set of solutions*

$$\begin{aligned}&e^{\lambda_1 t},\ te^{\lambda_1 t},\ \cdots,\ t^{m_1-1}e^{\lambda_1 t};\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots;\\ &e^{\lambda_s t},\ te^{\lambda_s t},\ \cdots,\ t^{m_s-1}e^{\lambda_s t}.\end{aligned} \tag{3.5.14}$$

Proof. One way to prove Theorem 3.19 is to find a fundamental matrix of the equivalent system, such that the first row is given by the functions in (3.5.14).
Another way to prove Theorem 3.19 is to show that each function in (3.5.14) is a solution of (3.5.12), and they are linearly independent. □

Remark. If (3.5.12) has a complex eigenvalue $\lambda = \alpha + i\beta$, then $\bar{\lambda} = \alpha - i\beta$ is also an eigenvalue. Thus both $t^k e^{(\alpha+i\beta)t}$ and $t^k e^{(\alpha-i\beta)t}$ appear in (3.5.14), where $0 \le k \le m-1$. In order to obtain a fundamental set of real solutions, the pair of solutions $t^k e^{(\alpha+i\beta)t}$ and $t^k e^{(\alpha-i\beta)t}$ in (3.5.14) should be replaced by $t^k e^{\alpha t}\cos(\beta t)$ and $t^k e^{\alpha t}\sin(\beta t)$ respectively.

3.6 Appendix 1: Proof of Lemma 3.12

Lemma 3.12 Let A be an $n \times n$ complex matrix and λ an eigenvalue of A with algebraic multiplicity m. Then

$$\dim\{\mathbf{x} \in \mathbb{C}^n \mid (\lambda I - A)^m \mathbf{x} = \mathbf{0}\} = m.$$

Proof. The proof consists of several steps. Let $T = \{\mathbf{x} \in \mathbb{C}^n \mid (\lambda I - A)^m \mathbf{x} = \mathbf{0}\}$. The space T is called the generalized eigenspace corresponding to the eigenvalue λ.

Step 1. T is a subspace of $\mathbb{C}^n$. This is just direct verification.

Step 2. T is invariant under A meaning $A[T] \subseteq T$. This is because if we take a vector $\mathbf{x}$ in T, then $(\lambda I - A)^m\mathbf{x} = \mathbf{0}$ so that $A(\lambda I - A)^m\mathbf{x} = \mathbf{0}$, which is the same as $(\lambda I - A)^m(A\mathbf{x}) = \mathbf{0}$. Therefore, $A\mathbf{x}$ is also in T.

Step 3. By Step 2 which says that $A[T] \subseteq T$, we may consider A as a linear transformation on the subspace T. In other words, $A : T \longrightarrow T$. Let μ be an eigenvalue of $A : T \longrightarrow T$. That is $A\mathbf{v} = \mu\mathbf{v}$. Then $\mathbf{v} \in T$ implies that $(\lambda I - A)^m\mathbf{v} = \mathbf{0}$. Since $A\mathbf{v} = \mu\mathbf{v}$, this simplifies to $(\lambda - \mu)^m\mathbf{v} = \mathbf{0}$. Being an eigenvector, $\mathbf{v} \neq \mathbf{0}$ so that $\mu = \lambda$. Therefore, all eigenvalues of $A : T \longrightarrow T$ are equal to λ.

Step 4. Let $\dim T = r$. Certainly $r \leq n$. Then by Step 3, the characteristic polynomial of $A : T \longrightarrow T$ is $(\lambda - z)^r$. Since T is an invariant subspace of $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, one can choose a basis of T and then extend it to a basis of $\mathbb{C}^n$ so that with respect to this new basis of $\mathbf{C}^n$, the matrix A is similar to a matrix where the upper left hand $r \times r$ submatrix represents A on T and the lower left hand $(n - r) \times r$ submatrix is the zero matrix. From this, we see that $(\lambda - z)^r$ is a factor of the characteristic polynomial of $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$. Hence $r \leq m$.

We also need the Cayley-Hamilton Theorem in the last step of the proof.

Cayley-Hamilton Theorem Let $p(z)$ be the characteristic polynomial of an $n \times n$ matrix A. Then $p(A) = \mathbf{0}$.

Step 5. Let $p(z) = (\lambda_1 - z)^{m_1} \cdots (\lambda_k - z)^{m_k}$ be the characteristic polynomial of A, where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$. For each i from 1 to k, let

$$T_i = \{\mathbf{x} \in \mathbb{C}^n \mid (\lambda_i I - A)^{m_i}\mathbf{x} = \mathbf{0}\}.$$

By Step 4, we have $\dim T_i \leq m_i$.

Let $f_i(z) = (\lambda_i - z)^{m_i}$ and $g_i(z) = f_1(z) \cdots \hat{f}_i(z) \cdots f_k(z)$, where $\hat{f}_i(z)$ means the polynomial $f_i(z)$ is omitted. Note that $f_i(z)g_i(z) = p(z)$ for all i.

Resolving $\frac{1}{(\lambda_1 - z)^{m_1} \cdots (\lambda_k - z)^{m_k}}$ into partial fractions, we have the identity

$$\frac{1}{(\lambda_1 - z)^{m_1} \cdots (\lambda_k - z)^{m_k}} \equiv \frac{h_1(z)}{(\lambda_1 - z)^{m_1}} + \frac{h_2(z)}{(\lambda_2 - z)^{m_2}} + \cdots + \frac{h_k(z)}{(\lambda_k - z)^{m_k}},$$

where $h_1(z), \ldots, h_k(z)$ are polynomials in z. Finding the common denominator of the right hand side and equate the numerators on both sides, we have $1 \equiv g_1(z)h_1(z) + \cdots + g_k(z)h_k(z)$. Substituting the matrix A into this polynomial identity, we have

$$g_1(A)h_1(A) + \cdots + g_k(A)h_k(A) = I,$$

where I is the identity $n \times n$ matrix.

Now for any $\mathbf{x} \in \mathbb{C}^n$, we have

$$g_1(A)h_1(A)\mathbf{x} + \cdots + g_k(A)h_k(A)\mathbf{x} = \mathbf{x}.$$

Note that each $g_i(A)h_i(A)\mathbf{x}$ is in T_i because $f_i(A)[g_i(A)h_i(A)\mathbf{x}] = p(A)h_i(A)\mathbf{x} = \mathbf{0}$ by the Cayley-Hamilton Theorem. This shows that any vector in $\mathbb{C}^n$ can be expressed as a sum of vectors where the i-summand is in T_i. In other words,

$$\mathbb{C}^n = T_1 + T_2 + \cdots + T_k.$$

Consequently, $m_1 + \cdots m_k = n \le \text{dim}T_1 + \cdots + \text{dim}T_k \le m_1 + \cdots + m_k$ so that $\text{dim}T_i = m_i$.

Remarks

1. In fact

$$\mathbb{C}^n = T_1 \oplus \cdots \oplus T_k.$$

2. If A is a real matrix and λ is a real eigenvalue of A of algebraic multiplicity m, then T is a real vector space and the real dimension of T is m. This is because for any set of real vectors in $\mathbb{R}^n$, it is linearly independent over $\mathbb{R}$ if and only if it is linearly independent over $\mathbb{C}$.

3. If A is a real matrix and λ is a complex eigenvalue of A of algebraic multiplicity m, then $\overline{\lambda}$ is also an eigenvalue of A of algebraic multiplicity m. In this case, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a basis over $\mathbb{C}$ of T_λ, where T_λ is the generalized eigenspace corresponding to λ, then $\{\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_m\}$ is a basis over $\mathbb{C}$ of $T_{\overline{\lambda}}$. It can be shown that the $2m$ real vectors $\Re\mathbf{v}_1, \ldots, \Re\mathbf{v}_m, \Im\mathbf{v}_1, \ldots, \Im\mathbf{v}_m$ are linearly independent over $\mathbb{R}$ and form a basis of $(T_\lambda \oplus T_{\overline{\lambda}}) \cap \mathbb{R}^n$.

Chapter 4

Power Series Solutions

4.1 Power Series

An infinite series of the form

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots \tag{4.1.1}$$

is a power series in $x - x_0$. In what follows, we will be focusing mostly at the point $x_0 = 0$. That is

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots \tag{4.1.2}$$

(4.1.2) is said to *converge* at a point x if the limit $\lim_{m\to\infty} \sum_{n=0}^{m} a_n x^n$ exists, and in this case the *sum* of the series is the value of this limit. It is obvious that (4.1.2) always converges at $x = 0$. It can be showed that each power series like (4.1.2) corresponds to a positive real number R, called the *radius of convergence*, with the property that the series converges if $|x| < R$ and diverges if $|x| > R$. It is customary to put R equal to 0 when the series converges only at $x = 0$, and equal to ∞ when it converges for all x. In many important cases, R can be found by the ratio test as follow.
If each $a_n \neq 0$ in (4.1.2), and if for a fixed point $x \neq 0$ we have

$$\lim_{n\to\infty} \left| \frac{a_{n+1}x^{n+1}}{a_n x^n} \right| = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| |x| = L,$$

then (4.1.2) converges for $L < 1$ and diverges if $L > 1$. It follows from this that

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|$$

if this limit exists (we put $R = \infty$, if $|a_n/a_{n+1}| \longrightarrow \infty$)
The interval $(-R, R)$ is called the *interval of convergence* in the sense that inside the interval the series converges and outside the interval the series diverges.

Consider the following power series

$$\sum_{n=0}^{\infty} n! x^n = 1 + x + 2!x^2 + 3!x^3 + \cdots \tag{4.1.3}$$

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{4.1.4}$$

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots \tag{4.1.5}$$

It is easy to verify that (4.1.3) converges only at $x = 0$. Thus $R = 0$. For (4.1.4), it converges for all x so that $R = \infty$. For (4.1.5), the power series converges for $|x| < 1$ and $R = 1$.

Suppose that (4.1.2) converges for $|x| < R$ with $R > 0$, and denote its sum by $f(x)$. That is

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots \tag{4.1.6}$$

Then one can prove that f is continuous and has derivatives of all orders for $|x| < R$. Also the series can be differentiated termwise in the sense that

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + \cdots ,$$

$$f''(x) = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} = 2a_2 + 3 \cdot 2a_3 x + \cdots ,$$

and so on. Furthermore, the resulting series are still convergent for $|x| < R$. These successive differentiated series yield the following basic formula relating a_n to with $f(x)$ and its derivatives.

$$a_n = \frac{f^n(0)}{n!} \tag{4.1.7}$$

Moreover, (4.1.6) can be integrated termwise provided the limits of integration lie inside the interval of convergence.

If

$$g(x) = \sum_{n=0}^{\infty} b_n x^n = b_0 + b_1 x + b_2 x^2 + \cdots \tag{4.1.8}$$

is another power series with interval of convergence $|x| < R$, then (4.1.6) and (4.1.8) can be added or subtracted termwise:

$$f(x) \pm g(x) = \sum_{n=0}^{\infty} (a_n \pm b_n) x^n = (a_0 \pm b_0) + (a_1 \pm b_1)x + (a_2 \pm b_2)x^2 + \cdots \tag{4.1.9}$$

They can also be multiplied like polynomials, in the sense that

$$f(x)g(x) = \sum_{n=0}^{\infty} c_n x^n,$$

where $c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$.

Suppose two power series (4.1.6) and (4.1.8) converge to the same function so that $f(x) = g(x)$ for $|x| < R$, then (4.1.7) implies that they have the same coefficients, $a_n = b_n$ for all n. In particular, if $f(x) = 0$ for all $|x| < R$, then $a_n = 0$, for all n.

Let $f(x)$ be a continuous function that has derivatives of all orders for $|x| < R$. Can it be represented by a power series? If we use (4.1.7), it is natural to expect

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = f(0) + f'(0)x + \frac{f''(0)}{2!} x^2 + \cdots \tag{4.1.10}$$

to hold for all $|x| < R$. Unfortunately, this is not always true. Instead, one can use Taylor's expansion for $f(x)$:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k \; + \; R_n(x),$$

where the remainder $R_n(x)$ is given by

$$R_n(x) = \frac{f^{(n+1)}(\overline{x})}{(n+1)!} x^{n+1}$$

for some point $\overline{x}$ between 0 and x. To verify (4.1.6), it suffices to show that $R_n(x) \longrightarrow 0$ as $n \longrightarrow \infty$.

Example The following familiar expansions are valid for all x.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

A function $f(x)$ with the property that a power series expansion of the form

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \tag{4.1.11}$$

is valid in some interval containing the point x_0 is said to be *analytic* at x_0. In this case, a_n is necessarily given by

$$a_n = \frac{f^{(n)}(x_0)}{n!},$$

and (4.1.11) is called the Taylor series of $f(x)$ at x_0.
Thus $e^x, \sin x, \cos x$ are analytic at all points. Concerning analytic functions, we have the following basic results.

1. Polynomials, $e^x, \sin x, \cos x$ are analytic at all points.
2. If $f(x)$ and $g(x)$ are analytic at x_0, then $f(x) \pm g(x)$, $f(x)g(x)$, and $f(x)/g(x)$ [provided $g(x_0) \neq 0$] are also analytic at x_0.
3. If $f(x)$ is analytic at x_0, and $f^{-1}(x)$ is a continuous inverse, then $f^{-1}(x)$ is analytic at $f(x_0)$ if $f'(x_0) \neq 0$.
4. If $g(x)$ is analytic at x_0 and $f(x)$ is analytic at $g(x_0)$, then $f(g(x))$ is analytic at x_0.
5. The sum of a power series is analytic at all points inside the interval of convergence.

4.2 Series Solutions of First Order Equations

A first order differential equation $y' = f(x, y)$ can be solved by assuming that it has a power series solution. Let's illustrate this with two familiar examples.

Example. Consider the differential equation $y' = y$. We assume it has a power series solution of the form

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots \tag{4.2.1}$$

that converges for $|x| < R$. That is the equation $y' = y$ has a solution which is analytic at the origin. Then

$$y' = a_1 + 2a_2 x + \cdots + n a_n x^{n-1} + \cdots \tag{4.2.2}$$

has the same interval of convergence. Since $y' = y$, the series (4.2.1) and (4.2.2) have the same coefficients. That is

$$(n+1)a_{n+1} = a_n, \qquad \text{all for } n = 0, 1, 2, \ldots$$

Thus $a_n = \frac{1}{n} a_{n-1} = \frac{1}{n(n-1)} a_{n-2} = \cdots = \frac{1}{n!} a_0$. Therefore

$$y = a_0 \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots \right),$$

where a_0 is an arbitrary constant. In this case, we recognize this as the power series of e^x. Thus the general solution is $y = a_0 e^x$.

Example. The function $y = (1 + x)^p$, where p is a real constant satisfies the differential equation

$$(1 + x)y' = py, \quad y(0) = 1. \tag{4.2.3}$$

As before, we assume it has a power series solution of the form

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

with positive radius of convergence. Then

$$\begin{array}{rcll} y' & = & a_1 + & 2a_2 x + 3a_3 x^2 + \cdots + \quad (n+1)a_{n+1} x^n + \cdots, \\ xy' & = & & a_1 x + 2a_2 x^2 + \cdots + \quad n a_n x^n + \cdots, \\ py & = & pa_0 + & pa_1 x + pa_2 x^2 + \cdots + \quad pa_n x^n + \cdots, \end{array}$$

Using (4.2.3) and equating coefficients, we have

$$(n+1)a_{n+1} + n a_n = p a_n, \qquad \text{for all } n = 0, 1, 2, \ldots$$

That is

$$a_{n+1} = \frac{p-n}{n+1} a_n,$$

so that

$$a_1 = p, a_2 = \frac{p(p-1)}{2}, a_3 = \frac{p(p-1)(p-2)}{2 \cdot 3}, ..., a_n = \frac{p(p-1)\cdots(p-n+1)}{n!}.$$

In other words,

$$y = 1 + px + \frac{p(p-1)}{2}x^2 + \frac{p(p-1)(p-2)}{2\cdot 3}x^3 + \cdots + \frac{p(p-1)\cdots(p-n+1)}{n!}x^n + \cdots .$$

By ratio test, this series converges for $|x| < 1$. Since (4.2.3) has a unique solution, we conclude that

$$(1+x)^p = 1 + px + \frac{p(p-1)}{2}x^2 + \frac{p(p-1)(p-2)}{2\cdot 3}x^3 + \cdots + \frac{p(p-1)\cdots(p-n+1)}{n!}x^n + \cdots ,$$

for $|x| < 1$, This is just the binomial series of $(1+x)^p$.

4.3 Second Order Linear Equations and Ordinary Points

Consider the homogeneous second order linear differential equation

$$y'' + P(x)y' + Q(x)y = 0 \tag{4.3.1}$$

Definition The point x_0 is said to be an *ordinary point* of (4.3.1) if $P(x)$ and $Q(x)$ are analytic at x_0. If at $x = x_0$, $P(x)$ and/or $Q(x)$ are not analytic, then x_0 is said to be a *singular point* of (4.3.1). A singular point x_0 at which the functions $(x-x_0)P(x)$ and $(x-x_0)^2Q(x)$ are analytic is called a *regular singular point* of (4.3.1). If a singular point x_0 is not a regular singular point, then it is called an *irregular singular point*.

Example. If $P(x)$ and $Q(x)$ are constant, then every point is an ordinary point of (4.3.1).

Example. Consider the equation $y'' + xy = 0$. Since the function $Q(x) = x$ is analytic at every point, every point is an ordinary point.

Example. In the Cauchy-Euler equation $y'' + \frac{a_1}{x}y' + \frac{a_2}{x^2}y = 0$, where a_1 and a_2 are constants, the point $x = 0$ is a singular point, but every other point is an ordinary point.

Example. Consider the differential equation

$$y'' + \frac{1}{(x-1)^2}y' + \frac{8}{x(x-1)}y = 0.$$

The singular points are 0 and 1. At the point 0, $xP(x) = x(1-x)^{-2}$ and $x^2Q(x) = -8x(1-x)^{-1}$, which are analytic at $x = 0$, and hence the point 0 is a regular singular point. At the point 1, we have $(x-1)P(x) = 1/(x-1)$ which is not analytic at $x = 1$, and hence the point 1 is an irregular singular point.

To discuss the behavior of the singularities at infinity, we use the transformation $x = 1/t$, which converts the problem to the behavior of the transformed equation near the origin. Using the substitution $x = 1/t$, (4.3.1) becomes

$$\frac{d^2y}{dt^2} + \left(\frac{2}{t} - \frac{1}{t^2}P(\frac{1}{t})\right)\frac{dy}{dt} + \frac{1}{t^4}Q(\frac{1}{t})y = 0 \tag{4.3.2}$$

We define the point at infinity to be an ordinary point, a regular singular point, or an irregular singular point of (4.3.1) according as the origin of (4.3.2) is an ordinary point, a regular singular point, or an irregular singular point.

Example. Consider the differential equation

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

The substitution $x = 1/t$ transforms the equation into

$$\frac{d^2y}{dt^2} + \left(\frac{3-t}{2t}\right)\frac{dy}{dt} + \frac{1}{2t}y = 0.$$

Hence the point at infinity is a regular singular point of the original differential equation.

Theorem 4.1 *Let x_0 be an ordinary point of the differential equation*

$$y'' + P(x)y' + Q(x)y = 0,$$

and let a_0 and a_1 be arbitrary constants. Then there exists a unique function $y(x)$ that is analytic at x_0, is a solution of the differential equation in an interval containing x_0, and satisfies the initial conditions $y(x_0) = a_0, y'(x_0) = a_1$. Furthermore, if the power series expansions of $P(x)$ and $Q(x)$ are valid on an interval $|x - x_0| < R$, $R > 0$, then the power series expansion of this solution is also valid on the same interval.

Example. Find two linearly independent solutions of $y'' - xy' - x^2y = 0$
Ans. $y_1(x) = 1 + \frac{1}{12}x^4 + \frac{1}{90}x^6 + \frac{3}{1120}x^8 + \cdots$ and $y_2(x) = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \frac{13}{1008}x^7 + \cdots$.

Example. Using power series method, solve the initial value problem $(1 + x^2)y'' + 2xy' - 2y = 0$, $y(0) = 0, y'(0) = 1$.
Ans. $y = x$.

Legendre's equation

$$(1 - x^2)y'' - 2xy' + p(p + 1)y = 0,$$

where p is a constant called the order of Legendre's equation.
That is $P(x) = -\frac{2x}{1-x^2}$ and $Q(x) = \frac{p(p+1)}{1-x^2}$. The origin is an ordinary point, and we expect a solution of the form $y = \sum a_n x^n$. Thus the left hand side of the equation becomes

$$(1 - x^2)\sum_{n=0}^{\infty}(n + 1)(n + 2)a_{n+2}x^n - 2x\sum_{n=0}^{\infty}(n + 1)a_{n+1}x^n + p(p + 1)\sum_{n=0}^{\infty}a_n x^n,$$

or

$$\sum_{n=0}^{\infty}(n + 1)(n + 2)a_{n+2}x^n - \sum_{n=2}^{\infty}(n - 1)na_n x^n - \sum_{n=1}^{\infty}2na_n x^n + \sum_{n=0}^{\infty}p(p + 1)a_n x^n.$$

The sum of these series is required to be zero, so the coefficient of x^n must be zero for every n. This gives

$$(n+1)(n+2)a_{n+2} - (n-1)na_n - 2na_n + p(p+1)a_n = 0,$$

for $n = 2, 3, \ldots$ In other words,

$$a_{n+2} = -\frac{(p-n)(p+n+1)}{(n+1)(n+2)}a_n.$$

This recursion formula enables us to express a_n in terms of a_0 or a_1 according as n is even or odd. In fact, for $m > 0$, we have

$$a_{2m} = (-1)^m \frac{p(p-2)(p-4)\cdots(p-2m+2)(p+1)(p+3)\cdots(p+2m-1)}{(2m)!}a_0,$$

$$a_{2m+1} = (-1)^m \frac{(p-1)(p-3)\cdots(p-2m+1)(p+2)(p+4)\cdots(p+2m)}{(2m+1)!}a_1.$$

With that, we get two linearly independent solutions

$$y_1(x) = \sum_{m=0}^{\infty} a_{2m}x^{2m} \quad \text{and} \quad y_2(x) = \sum_{m=0}^{\infty} a_{2m+1}x^{2m+1},$$

and the general solution is given by

$$\begin{aligned} y = a_0 &\left[1 - \frac{p(p+1)}{2!}x^2 + \frac{p(p-2)(p+1)(p+3)}{4!}x^4 \right. \\ &\left. - \frac{p(p-2)(p-4)(p+1)(p+3)(p+5)}{6!}x^6 + \cdots \right] \\ &+ a_1 \left[x - \frac{(p-1)(p+2)}{3!}x^3 + \frac{(p-1)(p-3)(p+2)(p+4)}{5!}x^5 \right. \\ &\left. - \frac{(p-1)(p-3)(p-5)(p+2)(p+4)(p+6)}{7!}x^7 + \cdots \right]. \end{aligned}$$

When p is not an integer, the series representing y_1 and y_2 have radius of convergence $R = 1$. For example,

$$\left|\frac{a_{2n+2}x^{2n+2}}{a_{2n}x^{2n}}\right| = \left|-\frac{(p-2n)(p+2n+1)}{(2n+1)(2n+2)}\right| |x^2| \longrightarrow |x|^2$$

as $n \longrightarrow \infty$, and similarly for the second series. In fact, by Theorem 4.1, and the familiar expansion

$$\frac{1}{1-x^2} = 1 + x^2 + x^4 + \cdots, \quad R = 1,$$

that $R = 1$ for both $P(x)$ and $Q(x)$. Thus, we know any solution of the form $y = \sum a_n x^n$ must be valid at least for $|x| < 1$.

The functions defined in the series solution of Legendre's equation are called Legendre functions. When p is a nonnegative integer, one of these series terminates and becomes a polynomial in x.

For instance, if $p = n$ is an even positive integer, the series representing y_1 terminates and y_1 is a polynomial of degree n. If $p = n$ is odd, y_2 again is a polynomial of degree n. These are called Legendre polynomials $P_n(x)$ and they give particular solutions to Legendre's equation

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0,$$

where n is a nonnegative integer. It is customary to choose the arbitrary constants a_0 or a_1 so that the coefficient of x^n in $P_n(x)$ is $(2n)!/[2^n(n!)^2]$ so that $P_n(1) = 1$. Then

$$P_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}.$$

The six Legendre polynomials are

$$\begin{aligned}
&P_0 = 1, && P_1(x) = x\\
&P_2(x) = \tfrac{1}{2}(3x^2-1), && P_3(x) = \tfrac{1}{2}(5x^3-3x)\\
&P_4(x) = \tfrac{1}{8}(35x^4-30x^2+3), && P_5(x) = \tfrac{1}{8}(63x^5-70x^3+15x)
\end{aligned}$$

There is also a Rodrigues' formula for the Legendre polynomial given by

$$P_n(x) = \frac{1}{n!2^n}\frac{d^n}{dx^n}(x^2-1)^n.$$

Hermite's equation $y'' - 2xy' + 2py = 0$, where p is a constant. The general solution of Hermite's equation is $y(x) = a_0y_1(x) + a_1y_2(x)$, where

$$y_1(x) = 1 - \frac{2p}{2!}x^2 + \frac{2^2p(p-2)}{4!}x^4 - \frac{2^3p(p-2)(p-4)}{6!}x^6 + \cdots,$$

$$y_2(x) = x - \frac{2(p-1)}{3!}x^3 + \frac{2^2(p-1)(p-3)}{5!}x^5 - \frac{2^3(p-1)(p-3)(p-5)}{7!}x^7 + \cdots.$$

By Theorem 4.1, both series for y_1 and y_2 converge for all x. Note that y_1 is a polynomial if p is an even integer, whereas y_2 is a polynomial if p is an odd integer.

The Hermite polynomial of degree n denoted by $H_n(x)$ is the nth-degree polynomial solution of Hermite's equation, multiplied by a suitable constant so that the coefficient of x^n is 2^n. The first six Hermite's polynomials are

$$\begin{aligned}
&H_0(x) = 1, && H_1(x) = 2x,\\
&H_2(x) = 4x^2-2, && H_3(x) = 8x^3-12x,\\
&H_4(x) = 16x^4-48x^2+12, && H_5(x) = 32x^5-160x^3+120x
\end{aligned}$$

A general formula for the Hermite polynomials is

$$H_n = (-1)^n e^{x^2}\frac{d^n}{dx^n}\left(e^{-x^2}\right).$$

4.4 Regular singular points and the method of Frobenius

Consider the second order linear homogeneous differential equation

$$x^2y'' + xp(x)y' + q(x)y = 0, \qquad (4.4.1)$$

where $p(x)$ and $q(x)$ are analytic at $x = 0$. In other words, 0 is a regular singular point of (4.4.1). Let $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3 + \cdots$, and $q(x) = q_0 + q_1x + q_2x^2 + q_3x^3 + \cdots$. Suppose (4.4.1) has a series solution of the form

$$y = x^r\sum_{n=0}^{\infty} a_nx^n = \sum_{n=0}^{\infty} a_nx^{n+r} \qquad (4.4.2)$$

An infinite series of the form (4.4.2) is called a Frobenius series, and the method that we are going to describe is called the method of Frobenius. We may assume $a_0 \neq 0$ because the series must have a first nonzero term. Termwise differentiation gives

$$y' = \sum_{n=0}^{\infty} a_n(n + r)x^{n+r-1}, \qquad (4.4.3)$$

and

$$y'' = \sum_{n=0}^{\infty} a_n(n + r)(n + r - 1)x^{n+r-2}. \qquad (4.4.4)$$

Substituting the series of y, y' and y'' into (4.4.1) yields

$$[r(r-1)a_0x^r + (r+1)ra_1x^{r+1} + \cdots] + [p_0x + p_1x^2 + \cdots]\cdot[ra_0x^{r-1} + (r+1)a_1x^r + \cdots]$$
$$+[q_0 + q_1x + \cdots]\cdot[a_0x^r + a_1x^{r+1} + \cdots] = 0. \qquad (4.4.5)$$

The lowest power of x in (4.4.5) is x^r. If (4.4.5) is to be satisfied identically, the coefficient $r(r-1)a_0 + p_0ra_0 + q_0a_0$ of x^r must vanish. As $a_0 \neq 0$, it follows that r satisfies the quadratic equation

$$r(r-1) + p_0r + q_0 = 0. \qquad (4.4.6)$$

This is the same equation obtained with the Cauchy-Euler equation. Equation (4.4.6) is called the **indicial equation** of (4.4.1) and its two roots (possibly equal) are the **exponents** of the differential equation at the regular singular point $x = 0$.

Let r_1 and r_2 be the roots of the indicial equation. If $r_1 \neq r_2$, then there are two possible Frobenius solutions and they are linearly independent. Whereas $r_1 = r_2$, there is only one possible Frobenius series solution. The second one cannot be a Frobenius series and can only be found by other means.

Example. Find the exponents in the possible Frobenius series solutions of the equation

$$2x^2(1 + x)y'' + 3x(1 + x)^3y' - (1 - x^2)y = 0.$$

Solution. Clearly $x = 0$ is a regular singular point since $p(x) = \frac{3}{2}(1 + x)^2$ and $q(x) = -\frac{1}{2}(1 - x)$ are polynomials. Rewrite the equation in the standard form:

$$y'' + \frac{\frac{3}{2}(1 + 2x + x^2)}{x}y' + \frac{-\frac{1}{2}(1 - x)}{x^2}y = 0.$$

We see that $p_0 = \frac{3}{2}$ and $q_0 = -\frac{1}{2}$. Hence the indicial equation is

$$r(r-1) + \frac{3}{2}r - \frac{1}{2} = r^2 + \frac{1}{2}r - \frac{1}{2} = (r+1)(r-\frac{1}{2}) = 0,$$

with roots $r_1 = \frac{1}{2}$ and $r_2 = -1$. The two possible Frobenius series solutions are of the forms

$$y_1(x) = x^{\frac{1}{2}} \sum_{n=0}^{\infty} a_n x^n \quad \text{and} \quad y_2(x) = x^{-1} \sum_{n=0}^{\infty} a_n x^n.$$

Once the exponents r_1 and r_2 are known, the coefficients in a Frobenius series solution can be found by substitution of the series (4.4.2),(4.4.3) and (4.4.4) into the differential equation (4.4.1). If r_1 and r_2 are complex conjugates, we always get two linearly independent solutions. We shall restrict our attention for real solutions of the indicial equation and seek solutions only for $x > 0$. The solutions on the interval $x < 0$ can be studied by changing the variable to $t = -x$ and solving the resulting equation for $t > 0$.

Let's work out the recursion relations for the coefficients. By (4.4.3), we have

$$\begin{aligned}
\tfrac{1}{x}p(x)y' &= \frac{1}{x}\left(\sum_{n=0}^{\infty} p_n x^n\right)\left[\sum_{n=0}^{\infty} a_n(n+r)x^{n+r-1}\right] \\
&= x^{r-2}\left(\sum_{n=0}^{\infty} p_n x^n\right)\left[\sum_{n=0}^{\infty} a_n(n+r)x^{n}\right] \\
&= x^{r-2}\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n} p_{n-k}a_k(r+k)\right]x^n \\
&= x^{r-2}\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n-1} p_{n-k}a_k(r+k) + p_0 a_n(r+n)\right]x^n.
\end{aligned}$$

Also we have

$$\begin{aligned}
\tfrac{1}{x^2}q(x)y &= \frac{1}{x^2}\left(\sum_{n=0}^{\infty} q_n x^n\right)\left(\sum_{n=0}^{\infty} a_n x^{r+n}\right) \\
&= \frac{1}{x^{r-2}}\left(\sum_{n=0}^{\infty} q_n x^n\right)\left(\sum_{n=0}^{\infty} a_n x^{n}\right) \\
&= x^{r-2}\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} q_{n-k}a_k\right)x^n \\
&= x^{r-2}\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n-1} q_{n-k}a_k + q_0 a_n\right)x^n.
\end{aligned}$$

Substituting these into the differential equation (4.4.1) and cancelling the term x^{r-2}, we have

$$\sum_{n=0}^{\infty}\left\{a_n[(r+n)(r+n-1)+(r+n)p_0+q_0] + \sum_{k=0}^{n-1} a_k[(r+k)p_{n-k}+q_{n-k}]\right\}x^n = 0.$$

Thus, equating the coefficients to zero, we have for $n \geq 0$,

$$a_n[(r+n)(r+n-1)+(r+n)p_0+q_0]+\sum_{k=0}^{n-1} a_k[(r+k)p_{n-k}+q_{n-k}]=0. \tag{4.4.7}$$

When $n = 0$, we get $r(r-1)+rp_0+q_0 = 0$, which is true because r is a root of the indicial equation. Then a_n can be determined by (4.4.7) recursively provided

$$(r+n)(r+n-1)+(r+n)p_0+q_0 \neq 0.$$

This would be the case if the two roots of the indicial equation do not differ by an integer. Suppose $r_1 > r_2$ are the two roots of the indicial equation with $r_1 = r_2 + N$ for some positive integer N. If we start with the Frobenius series with the smaller exponent r_2, then at the N-th step the process breaks off because the coefficient a_N in (4.4.7) is zero. In this case, only the Frobenius series solution with the larger exponent exists. The other solution cannot be a Frobenius series.

Theorem 4.2 *Assume that $x = 0$ is a regular singular point of the differential equation (4.4.1) and that the power series expansions of $p(x)$ and $q(x)$ are valid on an interval $|x| < R$ with $R > 0$. Let the indicial equation (4.4.6) have real roots r_1 and r_2 with $r_1 \geq r_2$. Then (4.4.1) has at least one solution*

$$y_1 = x^{r_1}\sum_{n=0}^{\infty} a_n x^n, \quad (a_0 \neq 0) \tag{4.4.8}$$

on the interval $0 < x < R$, where a_n are determined in terms of a_0 by the recursion formula (4.4.7) with r replaced by r_1, and the series $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < R$. Furthermore, if $r_1 - r_2$ is not zero or a positive integer, then equation (4.4.1) has a second independent solution

$$y_1 = x^{r_2}\sum_{n=0}^{\infty} a_n x^n, \quad (a_0 \neq 0) \tag{4.4.9}$$

on the same interval, where a_n are determined in terms of a_0 by the recursion formula (4.4.7) with r replaced by r_2, and again the series $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < R$.

Remark. (1) If $r_1 = r_2$, then there cannot be a second Frobenius series solution. (2) If $r_1 - r_2 = n$ is a positive integer and the summation of (4.4.7) is nonzero, then there cannot be a second Frobenius series solution. (3) If $r_1 - r_2 = n$ is a positive integer and the summation of (4.4.7) is zero, then a_n is unrestricted and can be assigned any value whatever. In particular, we can put $a_n = 0$ and continue to compute the coefficients without difficulties. Hence, in this case, there does exist a second Frobenius series solution. In many cases of (1) and (2), it is possible to determine a second solution by the method of variation of parameters. For instance a second solution for the Cauchy-Euler equation for the case where its indicial equation has equal roots is given by $x^r \ln x$.

Example. Find two linearly independent Frobenius series solutions of the differential equation $2x^2y'' + x(2x+1)y' - y = 0$.
Ans. $y_1 = x(1 - \frac{2}{5}x + \frac{4}{35}x^2 + \cdots),\ y_2 = x^{-\frac{1}{2}}(1 - x + \frac{1}{2}x^2 + \cdots)$.

Example. Find the Frobenius series solutions of $xy'' + 2y' + xy = 0$.

Solution. Rewrite the equation in the standard form $x^2y'' + 2xy' + x^2y = 0$. We see that $p(x) = 2$ and $q(x) = x^2$. Thus $p_0 = 2$ and $q_0 = 0$ and the indicial equation is $r(r-1) + 2r = r(r+1) = 0$ so that the exponents of the equation are $r_1 = 0$ and $r_2 = -1$. In this case, $r_1 - r_2$ is an integer and we may not have two Frobenius series solutions. We know there is a Frobenius series solution corresponding to $r_1 = 0$. Let's consider the possibility of the solution corresponding to the smaller exponent $r_2 = -1$. Let's begin with $y = x^{-1}\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n x^{n-1}$. Substituting this into the given equation, we obtain

$$\sum_{n=0}^{\infty}(n-1)(n-2)c_n x^{n-2} + 2\sum_{n=0}^{\infty}(n-1)c_n x^{n-2} + \sum_{n=0}^{\infty} c_n x^n = 0,$$

or equivalently

$$\sum_{n=0}^{\infty} n(n-1)c_n x^{n-2} + \sum_{n=0}^{\infty} c_n x^n = 0,$$

or

$$\sum_{n=0}^{\infty} n(n-1)c_n x^{n-2} + \sum_{n=2}^{\infty} c_{n-2} x^{n-2} = 0.$$

The cases $n = 0$ and $n = 1$ reduce to $0 \cdot c_0 = 0$ and $0 \cdot c_1 = 0$. Thus c_0 and c_1 are arbitrary and we can expect to get two linearly independent Frobenius series solutions. Equating coefficients, we obtain the recurrence relation

$$c_n = -\frac{c_{n-2}}{n(n-1)}, \quad \text{for } n \geq 2.$$

It follows from this that for $n \geq 1$.

$$c_{2n} = \frac{(-1)^n c_0}{(2n)!} \text{ and } c_{2n+1} = \frac{(-1)^n c_1}{(2n+1)!}.$$

Therefore, we have

$$y = x^{-1}\sum_{n=0}^{\infty} c_n x^n = \frac{c_0}{x}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}x^{2n} + \frac{c_1}{x}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n+1}.$$

We recognize this general solution as

$$y = \frac{1}{x}(c_0 \cos x + c_1 \sin x).$$

If we begin with the larger exponent, we will get the solution $(\sin x)/x$.

4.9 Bessel's equation

The second order linear homogeneous differential equation

$$x^2y'' + xy' + (x^2 - p^2)y = 0, \tag{4.5.1}$$

where p is a constant is called Bessel's equation. Its general solution is of the form

$$y = c_1 J_p(x) + c_2 Y_p(x). \tag{4.5.2}$$

The function $J_p(x)$ is called the Bessel function of order p of the first kind and the $Y_p(x)$ is the Bessel function of order p of the second kind. These functions have been tabulated and behave somewhat like the trigonometric functions of damped amplitude. If we let $y = u/\sqrt{x}$, we obtain

$$\frac{d^2u}{dx^2} + \left(1 - \frac{p^2 - \frac{1}{4}}{x^2}\right) u = 0. \tag{4.5.3}$$

In the special case in which $p = \pm\frac{1}{2}$, this equation becomes

$$\frac{d^2u}{dx^2} + u = 0.$$

Hence $u = c_1 \sin x + c_2 \cos x$ and

$$y = c_1 \frac{\sin x}{\sqrt{x}} + c_2 \frac{\cos x}{\sqrt{x}}. \tag{4.5.4}$$

Also we see that as $x \longrightarrow \infty$ in (4.5.3), and p is finite, we would expect the solution of (4.5.1) to behave as (4.5.4).

It is easy to see that $x = 0$ is a regular singular point of Bessel's equation. Here $p(x) = 1$ and $q(x) = -p^2 + x^2$. Thus the indicial equation is $r(r-1) + r - p^2 = r^2 - p^2 = 0$. Therefore, the exponents are $\pm p$. Let r be either $-p$ or p. If we substitute $y = \sum_{m=0}^{\infty} c_m x^{m+r}$ into Bessel's equation, we find in the usual manner that $c_1 = 0$ and that for $m \geq 2$,

$$[(m+r)^2 - p^2]c_m + c_{m-2} = 0 \tag{4.5.5}$$

The case $r = p \geq 0$**.** If we use $r = p$ and write a_m in place of c_m, then (4.5.5) yields the recursion formula

$$a_m = -\frac{a_{m-2}}{m(2p+m)} \tag{4.5.6}$$

As $a_1 = 0$, it follows that $a_m = 0$ for all odd values of m. The first few even coefficients are

$$a_2 = -\frac{a_0}{2(2p+2)} = -\frac{a_0}{2^2(p+1)},$$

$$a_4 = -\frac{a_2}{4(2p+4)} = \frac{a_0}{2^4 \cdot 2(p+1)(p+2)},$$

$$a_6 = -\frac{a_4}{6(2p+6)} = -\frac{a_0}{2^6 \cdot 2 \cdot 3(p+1)(p+2)(p+3)}.$$

In general, one can show that

$$a_{2m} = \frac{(-1)^m a_0}{2^{2m} m!(p+1)(p+2)\cdots(p+m)}.$$

Thus we have a solution associated with the larger exponent p

$$y_1 = a_0 \sum_{m=0}^{\infty} \frac{(-1)^m}{2^{2m} m!(p+1)(p+2)\cdots(p+m)} x^{2m+p}.$$

If $p = 0$, this is the only Frobenius series solution. In this case, if we choose $a_0 = 1$, we get a solution of Bessel's equation of order 0 given by

$$J_0(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m}(m!)^2} = 1 - \frac{x^2}{4} + \frac{x^4}{64} - \frac{x^6}{2304} + \cdots .$$

This special function $J_0(x)$ is called the Bessel function of order zero of the first kind. A second linearly independent solution can be obtained by other means, but it is not a Frobenius series.

The case $r = -p < 0$**.** Our theorem does not guarantee the existence of a Frobenius solution associated with the smaller exponent. However, as we shall see, it does have a second Frobenius series solution so long as p is not an integer. Let's write b_m in place of c_m in (4.5.5). Thus we have $b_1 = 0$ and for $m \geq 2$,

$$m(m-2p)b_m + b_{m-2} = 0 \tag{4.5.7}$$

Note that there is a potential problem if it happens that $2p$ is a positive integer, or equivalently if p is a positive integer or an odd integral multiple of $\frac{1}{2}$. Suppose $p = k/2$ where k is an odd positive integer. Then for $m \geq 2$, (4.5.7) becomes

$$m(m-k)b_m = -b_{m-2} \tag{4.5.8}$$

Recall $b_1 = 0$ so that $b_3 = 0$, $b_5 = 0$, $\cdots$, $b_{k-2} = 0$ by (4.5.8). Now in order to satisfy (4.5.8) for $m = k$, we can simply choose $b_k = 0$. Subsequently all $b_m = 0$ for all odd values of m. [If we let b_k to be arbitrary and non-zero, the subsequent solution so obtained is just $b_k y_1(x)$. Thus no new solution arises in this situation.]

So we only have to work out b_m in terms of b_0 for even values of m. In view of (4.5.8), it is possible to solve b_m in terms of b_{m-2} since $m(m-k) \neq 0$ as m is always even while k is odd. The result is the same as before except we should replace p by $-p$. Thus in this case, we have a second solution

$$y_2 = a_0 \sum_{m=0}^{\infty} \frac{(-1)^m}{2^{2m} m!(-p+1)(-p+2)\cdots(-p+m)} x^{2m-p}.$$

Since $p(x) = 1$ and $q(x) = x^2 - p^2$ are just polynomials. The series representing y_1 and y_2 converge for all $x > 0$. If $p > 0$, then the first term in y_1 is $a_0 x^p$, whereas the first term in y_2 is $b_0 x^{-p}$. Hence $y_1(0) = 0$, but $y_1(0) \longrightarrow \pm\infty$ as $x \longrightarrow 0$, so that y_1 and y_2 are linearly independent. So we have two linearly independent solutions as long as p is not an integer.

If $p = n$ is an nonnegative integer and we take $a_0 = \frac{1}{2^n n!}$, the solution y_1 becomes

$$J_n = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+n)!} \left(\frac{x}{2}\right)^{2m+n}.$$

J_n is called the Bessel function of the first kind of integral order n.

Remarks.

1. If p is not an integer, the factorials in J_p can be replaced by the so called Gamma functions and the general solution is $Y = c_1 J_p + c_2 J_{-p}$. If p is an integer, (4.5.7) can still be used to get a solution

J_{-p}, but it turns out it is just $(-1)^p J_p$, so there is only one Frobenius series solution. A second solution can be obtained by considering the function

$$Y_p(x) = \frac{J_p(x)\cos p\pi - J_{-p}(x)}{\sin p\pi}.$$

If p is not an integer, Y_p is a solution of Bessel's equation of order p as it is a linear combination of J_p and J_{-p}. If p approaches to an integer, the expression of Y_p gives an indeterminate form as both the numerator and denominator approach zero. To get a second solution when $p = n$ is an integer, we take limit as p tends to n to get a solution Y_n.

$$Y_n(x) = \lim_{p\to n} \frac{J_p(x)\cos p\pi - J_{-p}(x)}{\sin p\pi}.$$

Y_n is called a Bessel function of the second kind, and it follows that $y = c_1 J_p + c_2 Y_p$ is the general solution of Bessel's equation in all cases, whether p is an integer or not.

2. The case $r_1 = r_2$. Let $L(y) = x^2y'' + xp(x)y' + q(x)y$. We are solving $L(y) = 0$ by taking a series solution of the form $y(x) = x^r \sum_{n=0}^{\infty} a_n x^n$. If we treat r as a variable, then a_n's are functions of r. That is $y(x, r) = x^r \sum_{n=0}^{\infty} a_n(r)x^n$. Substituting this into $L(y)$ and requires it to be a solution, we get (4.4.7), which can be used to determine $a_n(r)$ recursively provided

$$(r + n)(r + n - 1) + (r + n)p_0 + q_0 \neq 0.$$

When r is near the double root $r_1 = r_2$, this expression is nonzero so that all a_n can be determined from (4.4.7). This means

$$L(y(x, r)) = a_0(r - r_1)^2 x^r.$$

So if $a_0 \neq 0$, we take $r = r_1$, we get one Frobenius series solution $y_1(x)$. Now let's differentiate the above equation with respect to r. We get

$$L(\frac{\partial y}{\partial r}) = \frac{\partial}{\partial r}L(y) = a_0[(r - r_1)^2 x^r \ln x + 2(r - r_1)x^r].$$

Evaluating at $r = r_1$, we obtain

$$L(\left.\frac{\partial y}{\partial r}\right|_{r=r_1}) = \left.\frac{\partial}{\partial r}L(y)\right|_{r=r_1} = 0.$$

Consequently, we have the second solution

$$y_2(x) = \frac{\partial y}{\partial r}(x, r_1) = x^{r_1}\ln x \sum_{n=0}^{\infty} a_n(r_1)x^n + x^{r_1}\sum_{n=0}^{\infty} a'_n(r_1)x^n = y_1(x)\ln x + x^{r_1}\sum_{n=1}^{\infty} a'_n(r_1)x^n.$$

Note that the sum in the last expression starts at $n = 1$ because a_0 is a constant and $a'_0 = 0$.
If we apply this method to Bessel's equation of order $p = 0$, we get by choosing $a_0 = 1$ the solutions

$$y_1(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}, \quad \text{and}$$

$$y_2(x) = y_1(x)\ln x - \sum_{n=1}^{\infty} \frac{(-1)^n H(n)}{(n!)^2}\left(\frac{x}{2}\right)^{2n},$$

where $H(n) = \sum_{k=1}^{n} \frac{1}{k}$.

3. The case $r_1 - r_2$ is a nonnegative integer
Consider

$$x^2 y'' + xp(x)y' + q(x)y = 0, \quad x > 0,$$

where $p(x) = \sum_{n=0}^{\infty} p_n x^n$ and $q(x) = \sum_{n=0}^{\infty} q_n x^n$. Let r_1 and r_2 be the roots (exponents) with $r_1 \geq r_2$ of the indicial equation $r(r-1) + p_0 r + q_0 = 0$.
Write $r_1 = r_2 + m$, where m is a nonnegative integer. Let $y_1 = x^{r_1}\sum_{n=0}^{\infty} a_n x^n$ be a Frobenius series solution corresponding to the larger exponent r_1. For simplicity, we take $a_0 = 1$.
Let $u = x^{r_2}\sum_{n=0}^{\infty} b_n x^n$ and make a change of variable:

$$y(x) = u(x) - b_m y_1(x)\ln x.$$

We get

$$x^2 u'' + xp(x)u' + q(x)u = b_m[2xy_1' + (p(x)-1)y_1].$$

Now let's substitute $u = x^{r_2}\sum_{n=0}^{\infty} b_n x^n$ to see if we can determine the b_n's. Note that the first term in the power series expansion of $b_m[2xy_1' + (p(x)-1)y_1]$ is mb_m, with $m \geq 0$.
Hence after substituting the power series of u into the above equation, we have

$$(r_2(r_2-1) + p_0 r_2 + q_0)b_0 x^{r_2} + A_1 x^{r_2+1} + \cdots + A_m x^{r_2+m} + \cdots = mb_m x^{r_1} + \cdots. \quad (4.5.9)$$

The first term on the left hand side is 0 as r_2 is a root of the indicial equation. This means b_0 can be arbitrary. The coefficients $A_1, A_2, \ldots$ are given by the main recurrence relation (4.4.7). Thus by equating $A_1, \ldots, A_{m-1}$ to 0, one can determine $b_1, \ldots b_{m-1}$. The next term on the left hand side of (4.5.9) is the coefficient A_m of x^{r_1}. In the expression of A_m given by (4.4.7), the coefficient of b_m is 0. Previously, this forbids the determination of b_m and possibly runs into a contradiction. Now on the right hand side of (4.5.9), if $m > 0$, then one can determine b_m by equating the coefficients of x^{r_1} on both sides. From then on, all the subsequent b_n's can be determined and we get a solution of the form $y(x) = u(x) - b_m y_1(x)\ln x$. Note that if $b_m = 0$ in this determination, then a second Frobenius series solution in fact can be obtained with the smaller exponent r_2.

Example. Consider $x^2 y'' + xy = 0$. Here $p(x) = 0, q(x) = x$. The exponents are 0 and 1. Hence $m = 1$. Corresponding to the exponent 1, the recurrence relation is $n(n+1)a_n + a_{n-1} = 0$ for $n \geq 0$.
We have the solution

$$y_1 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} n}{(n!)^2} x^n = x - \frac{1}{2}x^2 + \frac{1}{12}x^3 - \cdots.$$

Now $b_1[2xy_1' + (p(x)-1)y_1] = b_1(2x(1 - x + \frac{1}{4}x^2 - \cdots) - (x - \frac{1}{2}x^2 + \frac{1}{12}x^3 - \cdots)] = b_1[x - \frac{3}{2}x^2 + \frac{5}{12}x^3 - \cdots]$.

Substituting $u = x^0 \sum_{n=0}^{\infty} b_n x^n$ into $x^2u'' + xu = b_1[2xy_1' + (p(x) - 1)y_1]$, we get

$$0\cdot(0-1)b_0+[(1)(0)b_1+b_0]x+[(2)(1)b_2+b_1]x^2+[(3)(2)b_3+b_2]x^3+\cdots = b_1[x-\frac{3}{2}x^2+\frac{5}{12}x^3-\cdots].$$

Comparing coefficients, we have $b_0 = b_1$, $2b_2+b_1 = -\frac{3}{2}b_1$ and $6b_3+b_2 = \frac{5}{12}b_1, \cdots$. Thus $b_1 = b_0$, $b_2 = -\frac{5}{4}b_0$, $b_3 = \frac{5}{18}b_0, \ldots$. Therefore $u = b_0(1 + x - \frac{5}{4}x^2 + \frac{5}{18}x^3 - \cdots)$. By taking $b_0 = 1$, we get the solution $y = (1 + x - \frac{5}{4}x^2 + \frac{5}{18}x^3 - \cdots) - y_1(x)\ln x$.

If $m = 0$, then $r_1 = r_2$ and the first terms on both sides of (4.5.9) are 0. Thus we can continue to determine the rest of b_n's. In this case, the $\ln$ term is definitely present.

Exercise. Find the general solution of the differential equation

$$x^2(1 + x^2)y'' - x(1 + 2x + 3x^2)y' + (x + 5x^2)y = 0.$$

[**Answer.** $y_1 = x^2(1 + x + \frac{1}{2}x^2 + \cdots),\ y_2 = (1 + x + 2x^2 + \frac{8}{3}x^3 + \cdots) - 2y_1 \ln x$.]

Appendix 2 Some Properties of the Legendre Polynomials

The Legendre polynomial $P_n(x)$ is a polynomial of degree n satisfying Legendre's equation

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0,$$

where n is a nonnegative integer. It is normalized so that the coefficient of x^n is $(2n)!/[2^n(n!)^2]$. Explicitly it is given by

$$P_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}.$$

There is also a Rodrigues' formula for the Legendre polynomial given by

$$P_n(x) = \frac{1}{n!2^n}\frac{d^n}{dx^n}(x^2-1)^n.$$

Note that in Rodrigues' formula, the coefficient of x^n is $(2n)!/[2^n(n!)^2]$. We can use Rodrigues' formula to show that $P_n(1) = 1$. By this formula, we have $2^nP_n(1)$ is the coefficient of $(x-1)^n$ in the Taylor polynomial expansion of $(x^2-1)^n$ at $x = 1$. As $(x^2-1)^n = (x-1)^n(x-1+2)^n = (x-1)^n[(x-1)^n + n(x-1)^{n-1}2 + \cdots + 2^n]$, it is clear that the coefficient of $(x-1)^n$ is 2^n. Thus $P_n(1) = 1$.

The Legendre polynomial $P_n(x)$ has the generating function $\phi(Z) = (1-2xZ+Z^2)^{-\frac{1}{2}} = (1 + Z^2 - 2xZ)^{-\frac{1}{2}}$. That is $P_n(x)$ is the coefficient of Z^n in the expansion of ϕ. To see this, let's write

$$\phi(Z) = \sum_{n=0}^{\infty} A_n Z^n, \quad -1 \le x \le 1 \ \text{ and } \ |Z| < 1. \tag{A.1}$$

Using Binomial expansion,

$$(1+Z^2-2xZ)^{-\frac{1}{2}} = 1 - \frac{1}{2}(Z^2-2xZ) + \frac{(-\frac{1}{2})(-\frac{1}{2}-1)}{2!}(Z^2-2xZ)^2 + \cdots,$$

it is clear that A_n is a polynomial of degree n. If we let $x = 1$, we obtain

$$\phi(1) = (1-2Z+Z^2)^{-\frac{1}{2}} = (1-Z)^{-1} = 1 + Z + Z^2 + Z^3 + \cdots, \quad |Z| < 1.$$

Hence $A_n(1) = 1$ for all n. Now, if we can show that A_n satisfies Legendre's equation, it will be identical with $P_n(x)$ as the A_n's are the only polynomials of degree n that satisfy the equation and have the value 1 when $x = 1$. Differentiating ϕ with respect to Z and x, we obtain

$$(1-2Zx+Z^2)\frac{\partial\phi}{\partial Z} = (x-Z)\phi, \tag{A.2}$$

$$Z\frac{\partial\phi}{\partial Z} = (x-Z)\frac{\partial\phi}{\partial x}. \tag{A.3}$$

Substituting (A.1) into (A.2) and equating the coefficients of Z^{n-1}, we obtain

$$nA_n - (2n-1)xA_{n-1} + (n-1)A_{n-2} = 0 \tag{A.4}$$

Also substituting (A.1) into (A.3) and equating the coefficients of Z^{n-1}, we obtain

$$x\frac{dA_{n-1}}{dx} - \frac{dA_{n-2}}{dx} = (n-1)A_{n-1} \tag{A.5}$$

In (A.5), replace n by $n+1$ to get

$$x\frac{dA_n}{dx} - \frac{dA_{n-1}}{dx} = nA_n \tag{A.6}$$

Now differentiate (A.4) with respect to x and eliminate dA_{n-2}/dx by (A.5), we have

$$\frac{dA_n}{dx} - x\frac{dA_{n-1}}{dx} = nA_{n-1} \tag{A.7}$$

We now multiply (A.6) by $-x$ and add it to (A.7) and obtain

$$(1-x^2)\frac{dA_n}{dx} = n(A_{n-1} - xA_n) \tag{A.8}$$

Differentiating (A.8) with respect to x and simplifying the result by (A.6), we finally obtain

$$(1-x^2)\frac{d^2A_n}{dx^2} - 2x\frac{dA_n}{dx} + n(n+1)A_n = 0 \tag{A.9}$$

This shows that A_n is a solution of Legendre's equation. Using this generating function and Legendre's equation, it can be shown that $P_n(x)$ satisfy the following orthogonal relations.

$$\int_{-1}^{1} P_m(x)P_n(x)\,dx = \begin{cases} 0 & \text{if } m \neq n \\ \frac{2}{2n+1} & \text{if } m = n \end{cases} . \tag{A.10}$$

Chapter 5

Fundamental Theory of ODEs

5.1 Existence-Uniqueness Theorem

We consider the initial value problem

$$\frac{dx}{dt} = f(t,x), \quad x(t_0) = x_0, \tag{5.1.1}$$

Definition. Let G be a subset in $\mathbb{R}^2$. $f(t,x) : G \to \mathbb{R}$ is said to satisfy the Lipschitz condition with respect to x in G if there exists a constant $L > 0$ such that, for any $(t,x_1),(t,x_2) \in G$,

$$|f(t,x_1) - f(t,x_2)| \le L|x_1 - x_2|.$$

L is called a Lipschitz constant.

Theorem 5.1 (Picard) *Let $f(t,x)$ be continuous on the rectangle*

$$R : |t-t_0| \le a,\ |x-x_0| \le b \quad (a,b>0),$$

and let

$$|f(t,x)| \le M$$

for all $(t,x) \in R$. Furthermore, assume f satisfies a Lipschitz condition with constant L in R. Then there is a unique solution to the initial value problem

$$\frac{dx}{dt} = f(t,x), \ \ x(t_0) = x_0$$

on the interval $I = [t_0 - \alpha, t_0 + \alpha]$, where $\alpha = \min\{a, b/M\}$.

Proof of the existence of solution will be given in section 5.2 and 5.3. The uniqueness of solution will be proved in section 5.5.

Example 1. Let $f(t,x) = x^2 e^{-t^2} \sin t$ be defined on

$$G = \{(t,x) \in \mathbb{R}^2 : 0 \le x \le 2\}.$$

Let $(t,x_1),(t,x_2) \in G$.

$$
\begin{aligned}
& |f(t,x_1) - f(t,x_2)| \\
= \; & |x_1^2 e^{-t^2} \sin t - x_2^2 e^{-t^2} \sin t| \\
= \; & |e^{-t^2} \sin t||x_1 + x_2||x_1 - x_2| \\
\le \; & (1)(4)|x_1 - x_2|
\end{aligned}
$$

Thus we may take $L = 4$ and f satisfies a Lipschitz condition in G with Lipschitz constant 4.
Example 2. Let $f(t,x) = t\sqrt{x}$ be defined on

$$G = \{(t,x) \in \mathbb{R}^2 : 0 \le t \le 1, 0 \le x \le 1\}.$$

Consider the two points $(1,x), (1,0) \in G$. We have $|f(1,x) - f(1,0)| = \sqrt{x} = \frac{1}{\sqrt{x}}|x - 0|$. However, as $x \to 0^+$, $\frac{1}{\sqrt{x}} \to +\infty$, so that f cannot satisfy the Lipschitz condition with any finite constant $L > 0$ on G.

Proposition 5.1.1 *Suppose $f(t,x)$ has a continuous partial derivative $f_x(t,x)$ on a rectangle $R = \{(t,x) \in \mathbb{R}^2 : a_1 \le t \le a_2, b_1 \le x \le b_2\}$ in the tx-plane. Then f satisfies a Lipschitz condition on R.*

Proof. Since $f_x(t,x)$ is continuous on R, it attains its maximum value in R by the extreme value theorem. Let K be the maximum value of $|f_x(t,x)|$ on R. By Mean Value Theorem, we have

$$|f(t,x_1) - f(t,x_2)| = |f_x(t,c)||x_1 - x_2|,$$

for some c between x_1 and x_2.
Therefore,

$$|f(t,x_1) - f(t,x_2)| \le K|x_1 - x_2|$$

for all $(t,x_1), (t,x_2) \in R$. Thus, f satisfies a Lipschitz condition in G with Lipschitz constant K.
Example 3. Let $f(t,x) = x^2$ be defined on

$$G = \{(t,x) \in \mathbb{R}^2 : 0 \le t \le 1\}.$$

First

$$|f(t,x_1) - f(t,x_2)| = |x_1^2 - x_2^2| = |x_1 + x_2||x_1 - x_2|.$$

Since x_1 and x_2 can be arbitrarily large, f cannot satisfy the Lipschitz condition on G. If we replace G by any closed and bounded region, then f will satisfy the Lipschitz condition.

5.2 The method of successive approximations

We will give the proof of Theorem 5.1 in several steps. Let's fix $f(t,x)$ to be a continuous function defined on the rectangle

$$R : |t - t_0| \le a, \; |x - x_0| \le b \; (a, b > 0).$$

The objective is to show that on some interval I containing t_0, there is a solution ϕ to (5.1.1). The first step will be to show that the initial value problem (5.1.1) is equivalent to an integral equation, namely

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds. \tag{5.2.1}$$

By a solution of this equation on I is meant a continuous function ϕ on I such that $(t, \phi(t))$ is in R for all $t \in I$, and

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds.$$

Theorem 5.2 *A function ϕ is a solution of the initial value problem (5.1.1) on an interval I if and only if it is a solution of the integral equation (5.1.2) on I.*

Proof. Suppose ϕ is a solution of the initial value problem on I. Then

$$\phi'(t) = f(t, \phi(t)) \tag{5.2.2}$$

on I. Since ϕ is continuous on I, and f is continuous on R, the function $f(t, \phi(t))$ is continuous on I. Integrating (5.2.2) from t_0 to t we obtain

$$\phi(t) - \phi(t_0) = \int_{t_0}^{t} f(s, \phi(s))\, ds.$$

Since $\phi(t_0) = x_0$, we see that ϕ is a solution of (5.2.1).
Conversely, suppose ϕ satisfies (5.2.1). Differentiating we find, using the fundamental theorem of Calculus, that $\phi'(t) = f(t, \phi(t))$ for all $t \in I$. Moreover, from (5.2.1), it is clear that $\phi(t_0) = x_0$ and thus ϕ is a solution of (5.1.1).

As a first approximation to the solution of (5.2.1), we consider ϕ_0 defined by $\phi_0(t) = x_0$. This function satisfies the initial condition $\phi_0(t_0) = x_0$, but does not in general satisfy (5.2.1). However, if we compute

$$\phi_1(t) = x_0 + \int_{t_0}^{t} f(s, \phi_0(s))\, ds = x_0 + \int_{t_0}^{t} f(s, x_0)\, ds,$$

we might expect ϕ_1 is a closer approximation to a solution than ϕ_0. In fact, if we continue the process and define successively

$$\phi_0(t) = x_0, \qquad \phi_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s, \phi_k(s))\, ds, \quad k = 0, 1, 2, \ldots \tag{5.2.3}$$

we might expect, on taking the limit as $k \to \infty$, that we would obtain $\phi_k(t) \to \phi(t)$, where ϕ would satisfy

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds.$$

Thus ϕ would be our desired solution.

We call the functions $\phi_0, \phi_1, \phi_2 \cdots$ defined by (5.2.3) *successive approximations* to a solution of the integral equation (5.2.1), or the initial value problem (5.1.1).

Example. Consider the initial value problem $x' = tx,\ \ x(0) = 1$.
The integral equation corresponding to this problem is

$$x(t) = 1 + \int_{0}^{t} s \cdot x(s)\, ds,$$

and the successive approximations are given by

$$\phi_0(t) = 1, \qquad \phi_{k+1}(t) = 1 + \int_0^t s\phi_k(s)\,ds, \quad k = 0, 1, 2, \ldots .$$

Thus $\phi_1(t) = 1 + \int_0^t s\,ds = 1 + \frac{t^2}{2}$, $\phi_2(t) = 1 + \int_0^t s(1 + \frac{s^2}{2})\,ds = 1 + \frac{t^2}{2} + \frac{t^4}{2\cdot 4}$, and it may be established by induction that

$$\phi_k(t) = 1 + \left(\frac{t^2}{2}\right) + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \cdots + \frac{1}{k!}\left(\frac{t^2}{2}\right)^k .$$

We recognize $\phi_k(x)$ as a partial sum for the series expansion of the function $\phi(t) = e^{t^2/2}$. We know that this series converges for all t and this means that $\phi_k(t) \to \phi(t)$ as $k \to \infty$, for all $x \in \mathbb{R}$. Indeed ϕ is a solution of this initial value problem.

Theorem 5.3 *Suppose $|f(t,x)| \le M$ for all $(t,x) \in R$. Then the successive approximations ϕ_k, defined by (5.2.3), exist as continuous functions on*

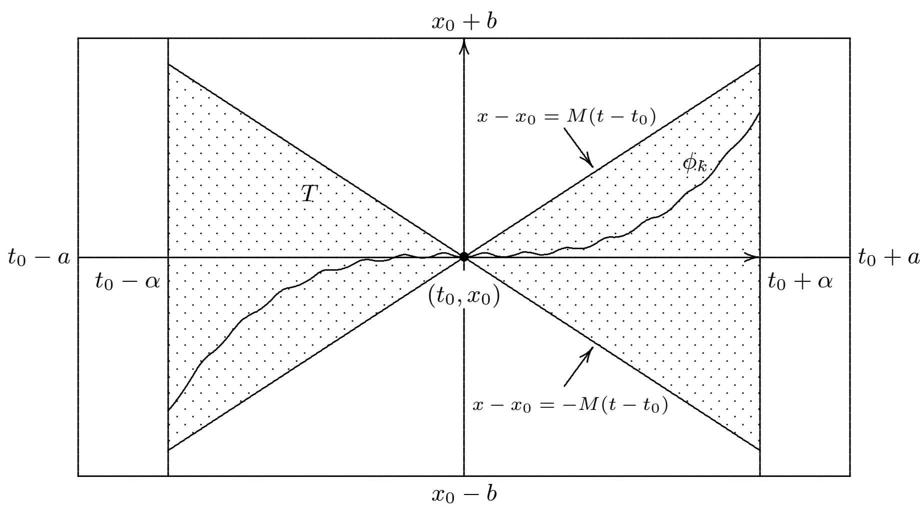
$$I: \;\; |t - t_0| \le \alpha = \min\{a, b/M\},$$

and $(t, \phi_k(t))$ is in R for $t \in I$. Indeed, the ϕ_k's satisfy

$$|\phi_k(t) - x_0| \le M|t - t_0| \tag{5.2.4}$$

for all $t \in I$.

Note: Since for $t \in I$, $|t - t_0| \le b/M$, the inequality (5.2.4) implies that $|\phi_k(t) - x_0| \le b$ for all $t \in I$, which shows that the points $(t, \phi_k(t))$ are in R for $t \in I$.
The geometric interpretation of the inequality (5.2.4) is that the graph of each ϕ_k lies in the region T in R bounded by the two lines $x - x_0 = M(t - t_0),\; x - x_0 = -M(t - t_0)$, and the lines $t - t_0 = \alpha,\; t - t_0 = -\alpha$.



Proof. We prove it by induction on k. Clearly ϕ_0 exists on I as a continuous function, and satisfies (5.2.4) with $k = 0$. Now suppose the theorem has been proved for the functions $\phi_0, \phi_1, \ldots, \phi_k$, with $k \geq 0$. We shall prove that it is valid for ϕ_{k+1}. By induction hypothesis, the point $(t, \phi_k(t))$ is in R for $t \in I$. Thus the function $f(t, \phi_k(t))$ exists for $t \in I$ and is continuous on I. Therefore, ϕ_{k+1}, which is given by

$$\phi_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s, \phi_k(s))\, ds,$$

exists as a continuous function on I. Moreover,

$$|\phi_{k+1}(t) - x_0| \leq \left| \int_{t_0}^{t} |f(s, \phi_k(s))|\, ds \right| \leq M|t - t_0|,$$

which shows that ϕ_{k+1} satisfies (5.2.4).

5.3 Convergence of the successive approximations

We now prove the main existence theorem

Theorem 5.4 *Let $f(t, x)$ be continuous on the rectangle*

$$R : |t - t_0| \leq a,\ |x - x_0| \leq b \quad (a, b > 0),$$

and let

$$|f(t, x)| \leq M$$

for all $(t, x) \in R$. Furthermore, assume f satisfies a Lipschitz condition with constant L in R. Then the successive approximations

$$\phi_0(t) = x_0, \qquad \phi_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s, \phi_k(s))\, ds, \quad k = 0, 1, 2, \ldots$$

converges uniformly on the interval $I = [t_0 - \alpha, t_0 + \alpha]$ with $\alpha = \min\{a, b/M\}$, to a solution of the initial value problem $\dfrac{dx}{dt} = f(t, x),\ \ x(t_0) = x_0$ on I.

Proof. (a) *Convergence of* $\{\phi_k(t)\}$. The key to the proof is the observation that ϕ_k may be written as

$$\phi_k = \phi_0 + (\phi_1 - \phi_0) + (\phi_2 - \phi_1) + \cdots + (\phi_k - \phi_{k-1}),$$

and hence $\phi_k(t)$ is a partial sum for the series

$$\phi_0(t) + \sum_{p=1}^{\infty} [\phi_p(t) - \phi_{p-1}(t)]. \tag{5.3.1}$$

Therefore to show that the sequence $\{\phi_k(t)\}$ converges uniformly is equivalent to show that the series (5.3.1) converges uniformly.

By Theorem 5.3, the functions ϕ_k all exist as continuous functions on I, and $(t, \phi_p(t))$ is in R for $t \in I$. Moreover,

$$|\phi_1(t) - \phi_0(t)| \leq M|t - t_0|, \tag{5.3.2}$$

for $t \in I$. Next consider the difference of ϕ_2 and ϕ_1. We have

$$\phi_2(t) - \phi_1(t) = \int_{t_0}^{t} [f(s, \phi_1(s)) - f(s, \phi_0(s))]\, ds.$$

Therefore

$$|\phi_2(t) - \phi_1(t)| \le \left| \int_{t_0}^{t} |f(s, \phi_1(s)) - f(s, \phi_0(s))|\, ds \right|,$$

and since f satisfies the Lipschitz condition

$$|f(t, x_1) - f(t, x_2)| \le L|x_1 - x_2|,$$

we have

$$|\phi_2(t) - \phi_1(t)| \le L \left| \int_{t_0}^{t} |\phi_1(s) - \phi_0(s)|\, ds \right|.$$

Using (5.3.2), we obtain

$$|\phi_2(t) - \phi_1(t)| \le ML \left| \int_{t_0}^{t} |s - t_0|\, ds \right|.$$

Thus if $t \ge t_0$,

$$|\phi_2(t) - \phi_1(t)| \le ML \int_{t_0}^{t} (s - t_0)\, ds = \frac{ML(t - t_0)^2}{2}.$$

The same result is valid in case $t \le t_0$.

We shall prove by induction that

$$|\phi_p(t) - \phi_{p-1}(t)| \le \frac{ML^{p-1}|t - t_0|^p}{p!} \qquad (5.3.3)$$

for all $t \in I$.

We have proved this for $p = 1$ and $p = 2$. Let's assume $t \ge t_0$. The proof is similar for $t \le t_0$. Assume (5.3.3) is true for $p = m$. Using the definition of ϕ_{m+1} and ϕ_m, we have

$$\phi_{m+1}(t) - \phi_m(t) = \int_{t_0}^{t} [f(s, \phi_m(s)) - f(s, \phi_{m-1}(s))]\, ds,$$

and thus

$$|\phi_{m+1}(t) - \phi_m(t)| \le \left| \int_{t_0}^{t} |f(s, \phi_m(s)) - f(s, \phi_{m-1}(s))|\, ds \right|.$$

Using the Lipschitz condition, we get

$$|\phi_{m+1}(t) - \phi_m(t)| \le L \left| \int_{t_0}^{t} |\phi_m(s) - \phi_{m-1}(s)|\, ds \right|.$$

By induction hypothesis, we obtain

$$|\phi_{m+1}(t) - \phi_m(t)| \le \frac{ML^m}{m!} \int_{t_0}^{t} |s - t_0|^m\, ds = \frac{ML^m|t - t_0|^{m+1}}{(m+1)!}.$$

Thus, (5.3.3) is true for all positive integer p.

Since $|t - t_0| \le \alpha$ for all $t \in I$, we can further deduce from (5.3.3) that

$$|\phi_p(t) - \phi_{p-1}(t)| \leq \frac{ML^{p-1}\alpha^p}{p!} = \frac{M}{L}\frac{(L\alpha)^p}{p!}. \tag{5.3.4}$$

Since the series $\sum_{p=1}^{\infty} \frac{M}{L}\frac{(L\alpha)^p}{p!}$ converges to $\frac{M}{L}(e^{L\alpha} - 1)$, we have by Weierstrass M-test that the series

$$\phi_0(t) + \sum_{p=1}^{\infty}[\phi_p(t) - \phi_{p-1}(t)]$$

converges absolutely and uniformly on I. Thus the sequence of partial sum which is $\phi_k(t)$ converges uniformly on I to a limit $\phi(t)$. Next we shall show that this limit ϕ is a solution of the integral equation (5.2.1).

(b) *Properties of the limit* ϕ. Since each ϕ_k is continuous on I and the sequence converges uniformly to ϕ, the function ϕ is also continuous on I. Now if t_1 and t_2 are in I, we have

$$|\phi_{k+1}(t_1) - \phi_{k+1}(t_2)| = \left| \int_{t_2}^{t_1} f(s, \phi_k(s))\, ds \right| \leq M|t_1 - t_2|,$$

which implies, by letting $k \to \infty$,

$$|\phi(t_1) - \phi(t_2)| \leq M|t_1 - t_2|. \tag{5.3.5}$$

It also follows from (5.3.5) that the function ϕ is continuous on I. In fact ϕ is uniformly continuous on I. Letting $t_1 = t, t_2 = t_0$ in (5.3.5), we see that

$$|\phi(t) - \phi(t_0)| \leq M|t - t_0|$$

which implies that the points $(t, \phi(t))$ are in R for all $t \in I$.

(c) *Estimate for* $|\phi(t) - \phi_k(t)|$. We have

$$\phi(t) = \phi_0(t) + \sum_{p=1}^{\infty}[\phi_p(t) - \phi_{p-1}(t)],$$

and

$$\phi_k(t) = \phi_0(t) + \sum_{p=1}^{k}[\phi_p(t) - \phi_{p-1}(t)].$$

Using (5.3.4), we have

$$\begin{aligned}
|\phi(t) - \phi_k(t)| &= \left| \sum_{p=k+1}^{\infty} [\phi_p(t) - \phi_{p-1}(t)] \right| \\
&\leq \sum_{p=k+1}^{\infty} |\phi_p(t) - \phi_{p-1}(t)| \\
&\leq \sum_{p=k+1}^{\infty} \frac{M}{L}\frac{(L\alpha)^p}{p!} \\
&\leq \frac{M}{L}\frac{(L\alpha)^{k+1}}{(k+1)!}\sum_{p=0}^{\infty}\frac{(L\alpha)^p}{p!} \\
&\leq \frac{M}{L}\frac{(L\alpha)^{k+1}}{(k+1)!}e^{L\alpha}.
\end{aligned}$$

Letting $\epsilon_k = \frac{(L\alpha)^{k+1}}{(k+1)!}$, we see that $\epsilon_k \to 0$ as $k \to \infty$ as ϵ_k is a general term for the series $e^{L\alpha}$. In terms of ϵ_k, we may rewrite the above inequality as

$$|\phi(t) - \phi_k(t)| \leq \frac{M}{L} e^{L\alpha} \epsilon_k, \quad \text{and } \ \epsilon_k \to 0 \text{ as } k \to \infty \tag{5.3.6}$$

(d) *The limit ϕ is a solution*. To complete the proof we must show that

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds,$$

for all $t \in I$. Note that since ϕ is continuous, the integrand $f(s, \phi(s))$ of the right hand side is continuous on I. Since

$$\phi_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s, \phi_k(s))\, ds,$$

we get the result by taking limit on both sides as $k \to \infty$ provided we can show

$$\int_{t_0}^{t} f(s, \phi_k(s))\, ds \to \int_{t_0}^{t} f(s, \phi(s))\, ds, \ \text{ as } k \to \infty.$$

Now
$$\begin{aligned}
\left| \int_{t_0}^{t} f(s, \phi(s))\, ds - \int_{t_0}^{t} f(s, \phi_k(s))\, ds \right| &\leq \left| \int_{t_0}^{t} |f(s, \phi(s)) - f(s, \phi_k(s))|\, ds \right| \\
&\leq L \left| \int_{t_0}^{t} |\phi(s) - \phi_k(s)|\, ds \right| \\
&\leq M e^{L\alpha} \epsilon_k |t - t_0| \qquad \text{by (5.3.6)} \\
&\leq M\alpha e^{L\alpha} \epsilon_k \ \to \ 0 \text{ as } k \to \infty.
\end{aligned}$$

This completes the proof of the Theorem 5.4.

Example. Consider the initial value problem $x' = (\sin t)x^2$, $x(0) = \frac{1}{2}$.
Let $f(t, x) = (\sin t)x^2$ be defined on

$$R = \{(t, x) : |t| \leq 1, |x - \frac{1}{2}| \leq \frac{1}{2}\}.$$

$|f(t, x)| = |(\sin t)x^2| \leq 1$. Thus we may take $M = 1$. Therefore by Theorem 5.4 a solution exists on $[-\alpha, \alpha]$ where $\alpha = \min\{1, \frac{1}{2}\} = \frac{1}{2}$. In fact $x(t) = (1 + \cos t)^{-1}$ is a solution defined on the maximal domain $(-\pi, \pi)$.

Exercise. Consider the initial value problem $x' = tx + x^{10}$, $x(0) = \frac{1}{10}$. Show that a solution of this problem exists for $|t| \leq \frac{1}{2}$.

5.4 Non-local Existence of Solutions

Theorem 5.5 *Let $f(t, x)$ be a continuous function on the strip $S = \{(t, x) \in \mathbb{R}^2 : |t - t_0| \leq a\}$, where a is a given positive number, and f satisfies the Lipschitz condition with respect to S. Then the initial value problem*

$$x'(t) = f(t, x), \ \ x(t_0) = x_0,$$

where $(t_0, x_0) \in S$ has a unique solution on the entire interval $[-a + t_0, a + t_0]$.

Remark. If f is bounded on S, the result can be deduced from Picard's Theorem. If f is not necessarily bounded, the proof is slightly different.

Proof of Theorem 5.5. First note that the given region S is not bounded above or below. Hence $f(t,x)$ needs not be bounded in S. However, as in Theorem 5.4, we shall consider the series

$$\phi_0(t) + \sum_{p=1}^{\infty}(\phi_p(t) - \phi_{p-1}(t))$$

whose n-th partial sum is $\phi_n(t)$ and $\phi_n(t) \to \phi(t)$ giving the solution of the initial value problem. Since $f(t,x)$ is not bounded in S, we adopt a different method of estimating different terms of the series. Let $M_0 = |x_0|$ and $M_1 = \max|\phi_1(t)|$. The fact that M_1 exists can be seen as follows. Since $f(t,x)$ is continuous in S, for a fixed x_0, $f(t,x_0)$ is a continuous function on $|t-t_0| \le a$.
Thus $\phi_1(t) = x_0 + \int_{t_0}^{t} f(s,x_0)\,ds$ is a continuous function in this interval so that $|\phi_1(t)|$ attains its maximum in this interval. We take it to be M_1 and let $M = M_0 + M_1$.

Thus, $|\phi_0(t)| = |x_0| \le M$ and $|\phi_1(t) - \phi_0(t)| \le M$. If $t_0 \le t \le t_0 + a$, then we have

$$\begin{aligned}|\phi_2(t) - \phi_1(t)| &= \left|\int_{t_0}^{t}[f(s,\phi_1(s)) - f(s,\phi_0(s))]\,ds\right| \le \int_{t_0}^{t}|f(s,\phi_1(s)) - f(s,\phi_0(s))|\,ds \\ &\le L\int_{t_0}^{t}|\phi_1(s) - \phi_0(s)|\,ds \le LM(t-t_0), \text{ where } L \text{ is the Lipschitz constant.}\end{aligned}$$

Now

$$\begin{aligned}|\phi_3(t) - \phi_2(t)| &= \left|\int_{t_0}^{t}[f(s,\phi_2(s)) - f(s,\phi_1(s))]\,ds\right| \le \int_{t_0}^{t}|f(s,\phi_2(s)) - f(s,\phi_1(s))|\,ds \\ &\le L\int_{t_0}^{t}|\phi_2(s) - \phi_1(s)|\,ds \le L^2M\int_{t_0}^{t}|(s-t_0)|\,ds = \frac{L^2M}{2}(t-t_0)^2.\end{aligned}$$

Hence, in general, we can prove by induction that

$$|\phi_n(t) - \phi_{n-1}(t)| \le \frac{L^{n-1}M(t-t_0)^{n-1}}{(n-1)!}.$$

Similar argument is true for the interval $t_0 - a \le t \le t_0$. Hence for every t with $|t-t_0| \le a$,

$$|\phi_n(t) - \phi_{n-1}(t)| \le \frac{L^{n-1}M(t-t_0)^{n-1}}{(n-1)!} \le \frac{L^{n-1}M}{(n-1)!}a^{n-1}.$$

Thus

$$|\phi_0(t)| + \sum_{n=1}^{\infty}|\phi_n(t) - \phi_{n-1}(t)| \le M\sum_{n=1}^{\infty}\frac{(La)^{n-1}}{(n-1)!}.$$

Hence each term on the left hand side of the above equation is less than the corresponding term of the convergent series of positive constants. Hence, by Weierstrass M-test, the series on the left converges uniformly on the whole interval $|t-t_0| \le a$ and let's denote its limit by $\phi(t)$.

Next we show that $\phi(t)$ is a solution of the initial value problem. We need to show that $\phi(t)$ satisfies the integral equation

$$\phi(t) - x_0 - \int_{t_0}^{t} f(s,\phi(s))\,ds = 0. \tag{5.4.1}$$

We know that

$$\phi_n(t) - x_0 - \int_{t_0}^{t} f(s, \phi_{n-1}(s))\, ds = 0. \tag{5.4.2}$$

Substituting the value of x_0 in (5.4.2) into the left hand side of (5.4.1), we get

$$\phi(t) - x_0 - \int_{t_0}^{t} f(s, \phi(s))\, ds = \phi(t) - \phi_n(t) - \int_{t_0}^{t} f(s, \phi(s)) - f(s, \phi_{n-1}(s))\, ds.$$

Thus we obtain

$$\begin{aligned}\left|\phi(t) - x_0 - \textstyle\int_{t_0}^{t} f(s, \phi(s))\, ds\right| &\le |\phi(t) - \phi_n(t)| + \textstyle\int_{t_0}^{t} |f(s, \phi(s)) - f(s, \phi_{n-1}(s))|\, ds \\ &\le |\phi(t) - \phi_n(t)| + L \textstyle\int_{t_0}^{t} |\phi(s) - \phi_{n-1}(s)|\, ds \end{aligned} \tag{5.4.3}$$

Since $\phi_n(t) \to \phi(t)$ uniformly for $t \in [t_0 - a, t_0 + a]$, the right hand side of (5.4.3) tends to zero as $n \to \infty$. Hence

$$\phi(t) - x_0 - \int_{t_0}^{t} f(s, \phi(s))\, ds = 0. \tag{5.2.4}$$

The uniqueness of solution will be proved in section 5.5.

Corollary 5.6 *Let $f(t, x)$ be a continuous function defined on $\mathbb{R}^2$. Suppose that for any $a > 0$, f satisfies the Lipschitz condition with respect to $S = \{(t, x) \in \mathbb{R}^2 : |t| \le a\}$ with $(t_0, x_0) \in S$. Then the initial value problem*

$$x'(t) = f(t, x), \;\; x(t_0) = x_0$$

has a unique solution on the entire $\mathbb{R}$.

Proof. If t is any real number, there is an $a > 0$ such that t is contained in $[t_0 - a, t_0 + a]$. For this a, the function f satisfies the condition of Theorem 5.5 on the strip

$$\{(t, x) \in \mathbb{R}^2 : |t - t_0| \le a\}.$$

Since this strip is contained in the strip

$$\{(t, x) \in \mathbb{R}^2 : |t| \le a + |t_0|\}.$$

Thus there is a unique solution $\phi(t)$ to the initial value problem for all $t \in \mathbb{R}$.

Example. Consider the initial value problem $x' = \sin(tx), \; x(0) = 1$.

Let $f(t, x) = \sin(tx)$. Let $a > 0$. Using the mean value theorem, we have for any $t \in [-a, a]$, $|f(t, x_1) - f(t, x_2)| = |\sin(tx_1) - \sin(tx_2)| = |t \cos(t\zeta)(x_1 - x_2)| \le |t||x_1 - x_2| \le a|x_1 - x_2|$. Thus f satisfies a Lipschitz condition on the strip $S = \{t \in \mathbb{R} \mid |t| \le a\}$, and there exists a solution on the entire $\mathbb{R}$.

Exercise. Show that the initial value problem $x' = \dfrac{x^3 e^t}{1 + x^2} + t^2 \cos x, \; x(0) = 1$ has a solution on $\mathbb{R}$.

5.5 Gronwall's Inequality and Uniqueness of Solution

Theorem 5.7 *Let f, g, and h be continuous nonnegative functions defined for $t \geq t_0$. If*

$$f(t) \leq h(t) + \int_{t_0}^{t} g(s)f(s)\,ds, \ \ t \geq t_0,$$

then

$$f(t) \leq h(t) + \int_{t_0}^{t} g(s)h(s)e^{\int_s^t g(u)\,du}\,ds, \ \ t \geq t_0.$$

Proof. First we are given

$$f(t) \leq h(t) + \int_{t_0}^{t} g(s)f(s)\,ds \tag{5.5.1}$$

Let $z(t) = \int_{t_0}^{t} g(s)f(s)\,ds$. Then for $t \geq t_0$,

$$z'(t) = g(t)f(t) \tag{5.5.2}$$

Since $g(t) \geq 0$, multiplying both sides of (5.5.1) by $g(t)$ and using (5.5.2), we get

$$z'(t) \leq g(t)[h(t) + z(t)]$$

which gives

$$z'(t) - g(t)z(t) \leq g(t)h(t).$$

This is a first order differential inequality which can be solved by finding an integrating factor $e^{-\int_{t_0}^{t} g(u)\,du}$. Hence the solution is

$$z(t)e^{-\int_{t_0}^{t} g(u)\,du} \leq \int_{t_0}^{t} g(s)h(s)e^{-\int_{t_0}^{s} g(u)\,du}\,ds$$

Or equivalently,

$$z(t) \leq \int_{t_0}^{t} g(s)h(s)e^{-\int_{t_0}^{s} g(u)\,du}e^{\int_{t_0}^{t} g(u)\,du}\,ds = \int_{t_0}^{t} g(s)h(s)e^{\int_s^t g(u)\,du}\,ds \tag{5.5.3}$$

Substituting for $z(t)$ in (5.5.3), we get

$$\int_{t_0}^{t} g(s)f(s)\,ds \leq \int_{t_0}^{t} g(s)h(s)e^{\int_s^t g(u)\,du}\,ds \tag{5.5.4}$$

From (5.5.1), we can replace the left side of (5.5.4) by the lesser inequality to obtain

$$f(t) - h(t) \leq \int_{t_0}^{t} g(s)h(s)e^{\int_s^t g(u)\,du}\,ds.$$

Theorem 5.8 *(Gronwall's Inequality) Let f and g be continuous nonnegative functions for $t \geq t_0$. Let k be any nonnegative constant. If*

$$f(t) \leq k + \int_{t_0}^{t} g(s)f(s)\,ds, \quad \text{for } \ t \geq t_0,$$

then

$$f(t) \leq ke^{\int_{t_0}^{t} g(s)\,ds}, \quad \text{for } \ t \geq t_0.$$

Corollary 5.9 *Let f be a continuous nonnegative function for $t \geq t_0$ and k a nonnegative constant. If*

$$f(t) \leq k \int_{t_0}^{t} f(s)\, ds$$

for all $t \geq t_0$, then $f(t) \equiv 0$ for all $t \geq t_0$.

Proof. For any $\epsilon > 0$, we can rewrite the given hypothesis as

$$f(t) \leq \epsilon + k \int_{t_0}^{t} f(s)\, ds,$$

for all $t \geq t_0$. Hence applying Gronwall's inequality, we have

$$f(t) \leq \epsilon e^{\int_{t_0}^{t} k\, ds},$$

for all $t \geq t_0$, which gives $f(t) \leq \epsilon e^{k(t-t_0)}$, for all $t \geq t_0$. Since ϵ is arbitrary, we get $f(t) \equiv 0$ by taking limit as $\epsilon \to 0^+$.

Remark. Similar results hold for $t \leq t_0$ when we all the integrals are integrated from t to t_0. For example, in Corollary 5.9, if

$$f(t) \leq k \int_{t}^{t_0} f(s)\, ds$$

for all $t \leq t_0$, then $f(t) \equiv 0$ for all $t \leq t_0$.

Corollary 5.10 *Let $f(t,x)$ be a continuous function which satisfies a Lipschitz condition on R with a Lipschitz constant L, where R is either a rectangle or a strip. If ϕ and φ are two solutions of*

$$x' = f(t,x), \quad x(t_0) = x_0,$$

on an interval I containing t_0, then $\phi(t) = \varphi(t)$ for all $t \in I$.

Proof. Let $I = [t_0 - \alpha, t_0 + \alpha]$. For $t \in [t_0, t_0 + \alpha]$, we have

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds,$$

and

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds.$$

Thus

$$|\phi(t) - \varphi(t)| \leq \int_{t_0}^{t} |f(s,\phi(s)) - f(s,\varphi(s))|\, ds \leq L \int_{t_0}^{t} |\phi(s) - \varphi(s)|\, ds.$$

By Corollary 5.9, $|\phi(t) - \varphi(t)| \equiv 0$ for $t \in [t_0, t_0 + \alpha]$. Thus $\phi(t) = \varphi(t)$ for $t \in [t_0, t_0 + \alpha]$. Similarly, $\phi(t) = \varphi(t)$ for $t \in [t_0 - \alpha, t_0]$.

Remark. If we only assume that $f(t,x)$ is a continuous function, we can still show that (5.1.1) has at least one solution, but the solution may not be unique.

Theorem 5.11 *(Peano) Assume G is an open subset of $\mathbb{R}^2$ containing (t_0, x_0) and $f(t, x)$ is continuous in G. Then there exists $a > 0$ such that (5.1.1) has at least one solution on the interval $[t_0 - a, t_0 + a]$.*

Example. Consider the initial value problem $x' = x^{2/3}$, $x(0) = 0$.
We find that $x(t) = 0$ and $x(t) = \frac{1}{27}t^3$ are both solutions.

Example. Suppose $\phi(t)$ is a solution to the initial value problem

$$x' = \frac{x^3 - x}{1 + t^2x^2}, \qquad \phi(0) = \frac{1}{2}.$$

Show that $0 < \phi(t) < 1$ for all t for which $\phi(t)$ is defined.

Solution. Let $\phi(t)$ be a solution defined on a domain J to the initial value problem. Suppose there exists $s \in J$ such that $\phi(s) \geq 1$. Since $\phi(t)$ is continuous and $\phi(0) = 1/2$, we have by Intermediate value theorem, that $\phi(s_0) = 1$ for some $s_0 \in (0, s)$. We may take s_0 to be the least value in $(0, s)$ such that $\phi(s_0) = 1$. In other words, $\phi(t) < 1$ for all $t \in (0, s_0)$ and $\phi(s_0) = 1$.

Now consider the initial value problem

$$x' = \frac{x^3 - x}{1 + t^2x^2}, \qquad x(s_0) = 1.$$

The function $f(t, x) = \frac{x^3 - x}{1+t^2x^2}$ satisfies the conditions of the Existence and Uniqueness Theorem. Thus there is a unique solution defined on an interval $I = [s_0 - \alpha, s_0 + \alpha]$ for some $\alpha > 0$. The above function $\phi(t)$ defined on J is a solution to this initial value problem, and it has the property that $\phi(t) < 1$ for all $t < s_0$. However, $\varphi(t) \equiv 1$ is clearly a solution to this initial value problem on I. But φ and ϕ are different solutions to the initial value problem contradicting the uniqueness of the solution. Consequently, $\phi(t) < 1$ for all $t \in J$. Similarly, $\phi(t) > 0$ for all $t \in J$.

Corollary 5.12 *Let $f(t, x)$ be a continuous function which is defined either on a strip*

$$R = \{(t, x) \in \mathbb{R}^2 \mid |t - t_0| \leq a\},$$

or a rectangle

$$R = \{(t, x) \in \mathbb{R}^2 \mid |t - t_0| \leq a, |x - x_0| \leq b\}.$$

Assume f satisfies a Lipschitz condition on R with a Lipschitz constant L. Let ϕ and φ be solutions defined on $I = [-a + t_0, t_0 + a]$ of $x' = f(t, x)$ satisfying the initial condition $x(t_0) = x_0$ and $x(t_0) = x_1$ respectively on I, then

$$|\phi(t) - \varphi(t)| \leq |x_0 - x_1|e^{L|t-t_0|}$$

for all $t \in I$.

Remark. In particular

$$|\phi(t) - \varphi(t)| \leq |x_0 - x_1|e^{La},$$

for all $t \in I$. Thus if the initial values x_0 and x_1 are close, the resulting solutions ϕ and φ are also close.

The proof of this corollary is by Gronwall's lemma and is left as an exercise.

5.6 Existence and Uniqueness of Solutions to Systems

Consider a system of differential equations

$$\begin{cases} x_1' = f_1(t, x_1, \cdots, x_n), \\ x_2' = f_2(t, x_1, \cdots, x_n), \\ \qquad \cdots\cdots\cdots\cdots\cdots \\ x_n' = f_n(t, x_1, \cdots, x_n), \end{cases}$$

where $x_j' = \frac{dx_j}{dt}$. Let us introduce notations

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}, \ \mathbf{x}' = \begin{pmatrix} x_1' \\ \cdots \\ x_n' \end{pmatrix}, \ \mathbf{f}(t, \mathbf{x}) = \begin{pmatrix} f_1(t, \mathbf{x}) \\ \cdots \\ f_n(t, \mathbf{x}) \end{pmatrix}.$$

Then the system can be written in a vector form:

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}). \tag{5.6.1}$$

Differential equations of higher order can be reduced to equivalent systems. Let us consider

$$\frac{d^n y}{dt^n} + F(t, y, y', \cdots, \frac{d^{n-1} y}{dt^{n-1}}) = 0. \tag{5.6.2}$$

Let

$$x_1 = y, \quad x_2 = \frac{dy}{dt}, \quad \cdots, \quad x_n = \frac{d^{n-1} y}{dt^{n-1}}.$$

Then (5.6.2) is equivalent to the following system

$$\begin{cases} x_1' = x_2, \\ x_2' = x_3, \\ \cdots\cdots\cdots \\ x_n' = -F(t, x_1, x_2, \cdots, x_n). \end{cases}$$

It can be written in the form of (5.6.1) if we let

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}, \quad \mathbf{f}(t, \mathbf{x}) = \begin{pmatrix} x_2, \\ x_3, \\ \cdots \\ -F(t, x_1, \cdots, x_n) \end{pmatrix}.$$

Recall that for a vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, its magnitude $|\mathbf{x}|$ is defined to be

$$|\mathbf{x}| = |x_1| + |x_2| + \cdots + |x_n|.$$

The following 2 properties of the magnitude can be proved easily.
1. Triangle Inequality $|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$.
2. If $A = (a_{ij})$ is an $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$, then $|A\mathbf{x}| \leq |A||\mathbf{x}|$ where $|A| = \sum_{i,j}^n |a_{ij}|$.

Definition. Let G be a subset in $\mathbb{R}^{1+n}$. $\mathbf{f}(t,\mathbf{x}) : G \to \mathbb{R}^n$ is said to satisfy the Lipschitz condition with respect to $\mathbf{x}$ in G if there exists a constant $L > 0$ such that, for all $(t,\mathbf{x}), (t,\mathbf{y}) \in G$,

$$|\mathbf{f}(t,\mathbf{x}) - \mathbf{f}(t,\mathbf{y})| \leq L|\mathbf{x}-\mathbf{y}|.$$

Example. Let $\mathbf{f} : \mathbb{R}^{1+2} \longrightarrow \mathbb{R}^2$ be given by

$$\mathbf{f}(t,\mathbf{x}) = \begin{pmatrix} 2x_2\cos t \\ x_1 \sin t \end{pmatrix}, \qquad \text{where } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then $|\mathbf{f}(t,\mathbf{x}) - \mathbf{f}(t,\mathbf{y})| = \left| \begin{pmatrix} 2x_2\cos t \\ x_1 \sin t \end{pmatrix} - \begin{pmatrix} 2y_2\cos t \\ y_1 \sin t \end{pmatrix} \right|$
$= |2\cos t(x_2 - y_2)| + |\sin t(x_1 - y_1)|$
$\leq 2|x_2 - y_2| + |x_1 - y_1|$
$\leq 2(|x_2 - y_2| + |x_1 - y_1|)$
$= 2|\mathbf{x} - \mathbf{y}|$.

Thus $\mathbf{f}$ satisfies the Lipschitz condition with respect to $\mathbf{x}$ in $\mathbb{R}^3$ with Lipschitz constant 2.

Theorem 5.13 *Suppose $\mathbf{f}$ is defined on a set $G \subset \mathbb{R}^{1+n}$ of the form*

$$|t - t_0| \leq a, \ |\mathbf{x} - \mathbf{x}_0| \leq b, \quad (a, b > 0)$$

or of the form

$$|t - t_0| \leq a, \ |\mathbf{x}| < \infty, \quad (a > 0).$$

If $\partial\mathbf{f}/\partial x_k$ $(k = 1, \ldots, n)$ exists, is continuous on G, and there is a constant $L > 0$ such that

$$\left| \frac{\partial \mathbf{f}}{\partial x_k} \right| \leq L, \quad (k = 1, \ldots, n),$$

for all $(t,\mathbf{x}) \in G$, then $\mathbf{f}$ satisfies a Lipschitz condition on G with Lipschitz constant L.

Proof. Let $\mathbf{f}(t,\mathbf{x}) = \begin{pmatrix} f_1(t,\mathbf{x}) \\ f_2(t,\mathbf{x}) \\ \vdots \\ f_n(t,\mathbf{x}) \end{pmatrix}$, where each $f_i(t,\mathbf{x}) : \mathbb{R}^{1+n} \longrightarrow \mathbb{R}$.

Thus

$$\frac{\partial \mathbf{f}}{\partial x_k} = \begin{pmatrix} \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_k} \\ \vdots \\ \frac{\partial f_n}{\partial x_k} \end{pmatrix}.$$

Let $(t,\mathbf{x}),(t,\mathbf{y}) \in G \subseteq \mathbb{R}^{1+n}$. Define $\mathbf{F} : [0,1] \longrightarrow \mathbb{R}^n$ by

$$\mathbf{F}(s) = \mathbf{f}(t, s\mathbf{x} + (1-s)\mathbf{y}) = \mathbf{f}(t, \mathbf{y} + s(\mathbf{x} - \mathbf{y})).$$

The point $s\mathbf{x} + (1-s)\mathbf{y}$ lies on the segment joining $\mathbf{x}$ and $\mathbf{y}$, hence the point $(t, s\mathbf{x} + (1-s)\mathbf{y})$ is in G.

Now $\mathbf{F}'(s) = \sum_{k=1}^{n} \frac{\partial \mathbf{f}}{\partial x_k}\frac{dx_k}{ds} = \sum_{k=1}^{n} \begin{pmatrix} \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_k} \\ \vdots \\ \frac{\partial f_n}{\partial x_k} \end{pmatrix} (x_k - y_k).$

Therefore,

$$|\mathbf{F}'(s)| \leq \sum_{k=1}^{n} \left|\frac{\partial \mathbf{f}}{\partial x_k}\right| |x_k - y_k| \leq L \sum_{k=1}^{n} |x_k - y_k| = L|\mathbf{x} - \mathbf{y}|,$$

for $s \in [0,1]$.
Since

$$\mathbf{f}(t,x) - \mathbf{f}(t,y) = \mathbf{F}(1) - \mathbf{F}(0) = \int_0^1 \mathbf{F}'(s)\, ds$$

we have $|\mathbf{f}(t,x) - \mathbf{f}(t,y)| \leq L|\mathbf{x} - \mathbf{y}|$.

Theorem 5.14 (Picard) *Let $\mathbf{f}(t,x)$ be continuous on the set*

$$R : |t - t_0| \leq a,\ |\mathbf{x} - \mathbf{x}_0| \leq b \quad (a, b > 0),$$

and let

$$|\mathbf{f}(t,x)| \leq M$$

for all $(t,\mathbf{x}) \in R$. Furthermore, assume $\mathbf{f}$ satisfies a Lipschitz condition with constant L in R. Then there is a unique solution to the initial value problem

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t,\mathbf{x}),\ \ \mathbf{x}(t_0) = \mathbf{x}_0$$

on the interval $I = [t_0 - \alpha, t_0 + \alpha]$, where $\alpha = \min\{a, b/M\}$.

Theorem 5.15 *Let $\mathbf{f}(t,x)$ be a continuous function on the strip $S = \{(t,\mathbf{x}) \in \mathbb{R}^{n+1} : |t - t_0| \leq a\}$, where a is a given positive number, and $\mathbf{f}$ satisfies the Lipschitz condition with respect to S. Then the initial value problem*

$$\mathbf{x}'(t) = \mathbf{f}(t,\mathbf{x}),\ \ \mathbf{x}(t_0) = \mathbf{x}_0,$$

where $(t_0, \mathbf{x}_0) \in S$ has a unique solution on the entire interval $[-a + t_0, a + t_0]$.

Corollary 5.16 *Let $\mathbf{f}(t,\mathbf{x})$ be a continuous function defined on $\mathbb{R}^{n+1}$. Suppose that for any $a > 0$, $\mathbf{f}$ satisfies the Lipschitz condition with respect to $S = \{(t,\mathbf{x}) \in \mathbb{R}^{n+1} : |t| \leq a\}$ with $(t_0, \mathbf{x}_0) \in S$. Then the initial value problem*

$$\mathbf{x}'(t) = \mathbf{f}(t,\mathbf{x}),\ \ \mathbf{x}(t_0) = \mathbf{x}_0$$

has a unique solution on the entire $\mathbb{R}$.

The proofs carry over directly from those for Theorem 5.1 and 5.5 and Corollary 5.6 using the method of successive approximations. That is the successive approximations

$$\phi_0(t) = \mathbf{x}_0, \qquad \phi_{k+1}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \phi_k(s))\, ds, \quad k = 0, 1, 2, \ldots$$

converge uniformly on the interval $I = [t_0 - \alpha, t_0 + \alpha]$ with $\alpha = \min\{a, b/M\}$, to a solution of the initial value problem $\dfrac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}),\ \ \mathbf{x}(t_0) = \mathbf{x}_0$ on I.

Example. Find the first 5 successive approximations to the initial value problem

$$x'' = -e^t x,\ x(0) = 1, x'(0) = 0.$$

The initial value problem is equivalent to the following initial value problem of differential system.

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}' = \begin{pmatrix} y(t) \\ -e^t x(t) \end{pmatrix}, \qquad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We start with

$$\begin{pmatrix} x_0(t) \\ y_0(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{for all} \quad t \in \mathbb{R}.$$

Then

$$\begin{pmatrix} x_1(t) \\ y_1(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} 0 \\ -e^s \times 1 \end{pmatrix} ds = \begin{pmatrix} 1 \\ 1 - e^t \end{pmatrix}.$$

$$\begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} 1 - e^s \\ -e^s \end{pmatrix} ds = \begin{pmatrix} 2 + t - e^{-t} \\ 1 - e^t \end{pmatrix}.$$

$$\begin{pmatrix} x_3(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} 1 - e^s \\ -e^s(2 + s - e^s) \end{pmatrix} ds = \begin{pmatrix} 2 + t - e^{-t} \\ \frac{1}{2} - e^t - te^t + \frac{1}{2}e^{2t} \end{pmatrix}.$$

$$\begin{pmatrix} x_4(t) \\ y_4(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} \frac{1}{2} - e^s - se^s + \frac{1}{2}e^{2s} \\ -e^s(2 + s - e^s) \end{pmatrix} ds = \begin{pmatrix} \frac{3}{4} + \frac{t}{2} - te^t + \frac{1}{4}e^{2t} \\ \frac{1}{2} - e^t - te^t + \frac{1}{2}e^{2t} \end{pmatrix}.$$

Example. Consider the linear differential system $\mathbf{x}' = A\mathbf{x}$, where $A = (a_{ij})$ is an $n \times n$ constant matrix. Let $\mathbf{f}(t, \mathbf{x}) = A\mathbf{x}$. For any $a > 0$ and for all $|t| < a$, we have $|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| = |A(\mathbf{x}_1 - \mathbf{x}_2)| \le |A||\mathbf{x}_1 - \mathbf{x}_2|$, where $|A| = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|$, so that $\mathbf{f}$ satisfies the Lipschitz condition on the strip $S = \{(t, \mathbf{x}) \in \mathbb{R}^{n+1} : |t| \le a\}$. Therefore the system has a unique solution for any initial value and is defined on the entire $\mathbb{R}$.

Example. Let $\mathbf{x}' = A(t)\mathbf{x}$, where $A(t) = (a_{ij}(t))$ is an $n \times n$ matrix of continuous functions defined on a closed interval I. Let $|a_{ij}(t)| \le K$ for all $t \in I$ and all $i, j = 1, \ldots n$.
Thus if $\mathbf{f}(t, \mathbf{x}) = A(t)\mathbf{x}$, then

$$\frac{\partial \mathbf{f}}{\partial x_k} = \begin{pmatrix} a_{1k}(t) \\ a_{2k}(t) \\ \vdots \\ a_{nk}(t) \end{pmatrix},$$

which is independent of $\mathbf{x}$.
Therefore,

$$\left|\frac{\partial \mathbf{f}}{\partial x_k}\right| = \sum_{i=1}^{n} |a_{ik}(t)| \leq nK \equiv L, \qquad \text{for all } t \in I \text{ and } k = 1, \ldots, n.$$

By Theorem 5.13 the function $\mathbf{f}$ satisfies a Lipschitz condition on the strip

$$S = \{(t, \mathbf{x}) \in \mathbb{R}^{1+n} \mid t \in I\}$$

with Lipschitz constant L. Thus by Theorem 5.15, the system $\mathbf{x}' = A(t)\mathbf{x}$ has a unique solution for any initial value in S and is defined on all of I.

Bibliography

[1] Ravi P. Agarwal and Ramesh C. Gupta, *Essentials of ordianry differential equations*, McGraw-Hill (1991)

[2] Earl A. Coddington, *An introduction to ordinary differential equations*, Dover (1961)

[3] George F. Simmons, *Differential equations with applications and historical notes*, 2nd edition, McGraw-Hill (1991)