An introduction to ordinary differential equations
Lecture notes, Fall 2024
(Version: September 25, 2024)

Pu-Zhao Kow

Department of Mathematical Sciences, National Chengchi University
Email address: pzkow@g.nccu.edu.tw

Preface

This lecture note is prepared mainly based on [**BD22**, **HS99**] for the course *Differential Equations*, for *graduate levels*, during Fall 2024 (113-1). In next semester, we will more focus on the partial differential equations (PDE), and we will use the lecture note [**Kow24**]. There is another Differential Equations course for undergraduate levels. The lecture note may updated during the course.

Title. Differential Equations

Lectures. Thursday (13:10–16:00)

Language. Chinese and English. Materials will be prepared in English

Instructor. Pu-Zhao Kow (Email: `pzkow@g.nccu.edu.tw`)

Office hour. Thursday (16:10–17:00)

Completion. Homework Assignments 60% (must be prepared using LaTeX), Midterm presentation 20%, Final presentation 20%

We will focus in pointing the relation of ordinary differential equations (ODE) with other (mathematical) fields, especially the partial differential equations (PDE), rather than go through all boring details in class. One can choose to present the proof of some theorems in this lecture note which the proof is skipped, or other interesting topics (more credit will be earned), during the midterm or final presentations.

Contents

CHAPTER 1

Introduction

1.1. Some mathematical models

In order to motivate this course, we begin with some examples from [**BD22**, **Che16**]. The most simplest and important example which can be modeled by ordinary differential equations (ODE) is a relaxation process, i.e. the system starts from state and eventual reaches an equilibrium state.

EXAMPLE 1.1.1 (A falling object). We now consider an object with mass m falling from height y_0 at time $t = 0$. Let $v(t)$ be its velocity at time t. According to physics law, we know that the acceleration of the object at time t is the rate of change of velocity $v(t)$, that is,

$$a(t) = \frac{\mathrm{d}}{\mathrm{d}t}v(t) \equiv v'(t).$$

According to *Newton's second law*, the net force F exerted on the on the object is expressed by the equation

$$F(t) = ma(t) = mv'(t). \tag{1.1.1}$$

Next, we consider the forces that act on the object as it falls. The gravity exerts a force equal to the weight of the object given by mg, where g is the acceleration due to the gravity. The drag force due to air resistance has the magnitude $\gamma v(t)$, where γ is a constant called the drag coefficient. Therefore the net force is given by

$$F(t) = mg - \gamma v(t). \tag{1.1.2}$$

Combining (1.1.1) and (1.1.2), we reach the ODE

$$mv'(t) = mg - \gamma v(t) \quad \text{for } t > 0.$$

EXAMPLE 1.1.2 (Heating (or cooling) of an object). We now consider an object, with initial temperature T_0 at time $t = 0$, which is taken out of refrigerator to defrost. Let $T(t)$ be its temperature at time t. Suppose that the room temperature is given by K. The *Newton's law of cooling/heating* says that the rate change of temperature $T(t)$ is proportional to the difference between $T(t)$ and K, more precisely,

$$T'(t) = -\alpha(T(t) - K),$$

where $\alpha > 0$ is a conductivity coefficient.

EXAMPLE 1.1.3 (Population growth model). We first describe the population model proposed by Malthus (1766–1834). Let $y(t)$ be the population (in a "large" area) at time t. He built a model based on the following hypothesis:

$$y'(t) = \text{births} - \text{deaths} + \text{migration},$$

and he assume that the births and the deaths are proportion to the current population $y(t)$, that is,

$$\text{births} - \text{deaths} = ry(t),$$

where the constant $r \in \mathbb{R}$ is called the net growth rate. If there is no migration at all, then the model reads

$$y'(t) = ry(t),$$

which is called the *simple population growth model*. Suppose that the initial population is y_0 at time $t = 0$. In fact, the unique solution is

$$y(t) = y_0 e^{rt},$$

which is not make sense, since the environment limitation is not taken into account. With this consideration, we should expect that there is an environment carrying capacity K such that

$$y'(t) > 0 \text{ when } y(t) < K, \quad y'(t) < 0 \text{ when } y(t) > K$$

due to a competition of resource. Verhulst (1804–1849) proposed another model which take the limit of environment into consideration (i.e. in a "small" area):

$$y'(t) = ry\left(1 - \frac{y(t)}{K}\right),$$

which is called the *logistic population model*. Note the above simple population growth model formally corresponds to the case when $K = +\infty$. See also [BD22, Section 2.5] for further explainations.

1.2. Classification of ODE

We say that the system of equations of the form

$$\boldsymbol{F}(t, \boldsymbol{u}(t), \boldsymbol{u}''(t), \cdots, \boldsymbol{u}^{(m)}(t)) = 0,$$

or in equation form

$$\begin{cases} F_1(t, \boldsymbol{u}(t), \boldsymbol{u}''(t), \cdots, \boldsymbol{u}^{(m)}(t)) = 0, \\ \vdots \\ F_\ell(t, \boldsymbol{u}(t), \boldsymbol{u}''(t), \cdots, \boldsymbol{u}^{(m)}(t)) = 0, \end{cases}$$

the *ordinary differential equation* (*ODE*) of order m, where we write $\boldsymbol{u}^{(k)} := (u_1^{(k)}, \cdots, u_n^{(k)})$ the k^{th}-order derivative of $\boldsymbol{u} = (u_1, \cdots, u_n)$. Throughout this note, we use the bold font to emphasize the vector-valued functions.

Example 1.2.1. For example,

$$u''' + 2e^t u'' + uu' = t^4$$

is a third order ODE.

In many cases, we only consider the system of ODE of the form (with $n = \ell$)

$$\boldsymbol{u}^{(m)}(t) = \boldsymbol{f}(t, \boldsymbol{u}(t), \boldsymbol{u}''(t), \cdots, \boldsymbol{u}^{(m-1)}(t)).$$

Otherwise, for example, the equation

$$(u')^2 + tu' + 4u = 0$$

leads to two equations

$$u' = \frac{-t + \sqrt{t^2 - 16u}}{2}, \quad u' = \frac{-t - \sqrt{t^2 - 16u}}{2}.$$

An ODE is said to be *linear* if it takes the form

$$a_0(t)\boldsymbol{u}^{(n)} + a_1(t)\boldsymbol{u}^{(n-1)} + \cdots + a_n(t)\boldsymbol{u} = \boldsymbol{g}(t),$$

for some matrices $a_0, \cdots, a_n$, otherwise we say that the ODE is *nonlinear*. If $\boldsymbol{g}(t) \equiv \boldsymbol{0}$, then we say that the ODE is *homogeneous*, otherwise *inhomogeneous*.

CHAPTER 2

First order nonlinear ODE

This chapter deals with first order nonlinear ODE of the form

$$\boldsymbol{u}'(t) = \boldsymbol{f}(t, \boldsymbol{u}), \quad \boldsymbol{u}(t_0) = \boldsymbol{u}_0 \tag{2.0.1}$$

for some vector $\boldsymbol{u}_0 \in \mathbb{R}^n$. We again remind the readers that we use the notation

$$\boldsymbol{u}(t) = (u_1(t), \cdots, u_n(t)),$$
$$\boldsymbol{f}(t) = (f_1(t), \cdots, f_n(t)),$$

and we also use the notation

$$|\boldsymbol{u}(t)| = \max\{|u_1(t)|, \cdots, |u_n(t)|\}.$$

2.1. Well-posedness of ODE

If $\boldsymbol{f} \equiv \boldsymbol{0}$, then (2.0.1) reads

$$\boldsymbol{u}'(t) = \boldsymbol{0}, \quad \boldsymbol{u}(t_0) = \boldsymbol{u}_0,$$

and one easily sees that the constant function $\boldsymbol{u} = \boldsymbol{u}_0$ is a solution which is valid for all $t \in \mathbb{R}$. We first state the fundamental existence theorem when $\boldsymbol{f} \not\equiv \boldsymbol{0}$.

THEOREM 2.1.1 ([**HS99**, Theorem I-2-5]). *Let $a > 0$ and $b > 0$. If $\boldsymbol{f} = \boldsymbol{f}(t, \boldsymbol{y})$ is a (real-valued) continuous function on a closed cylinder*

$$\mathcal{R} = \{(t, \boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \le a, |\boldsymbol{y} - \boldsymbol{u}_0| \le b\}$$

such that

$$M := \max_{(t,\boldsymbol{y}) \in \mathcal{R}} |\boldsymbol{f}(t, \boldsymbol{y})| > 0,$$

then there exists a function $\boldsymbol{u} \in (C^1((t_0 - \alpha, t_0 + \alpha)))^n$ with $\alpha = \min\{a, \frac{b}{M}\}$ satisfying (2.0.1) in $(t_0 - \alpha, t_0 + \alpha)$.

However, the uniqueness does not hold true in general without further assumption on $\boldsymbol{f}$. We demonstrate this in the following few examples.

EXAMPLE 2.1.2. We define the function

$$u(t) := \begin{cases} 0 & , t \le 3, \\ \left(\frac{2}{5}(t^2 - 9)\right)^{5/4} & , t > 3. \end{cases}$$

By using left and right limits, it is not difficult to check that $u \in C(\mathbb{R})$. By using elementary calculus, one computes that

$$u'(t) = \left(\frac{2}{5}\right)^{1/4} t(t^2 - 9)^{1/4} \quad \text{for } t > 3,$$
$$u'(t) = 0 \quad \text{for } t < 3.$$

Since

$$\lim_{h\to 0_+}\frac{u(3+h)-u(3)}{h}=\lim_{h\to 0_+}\frac{1}{h}\left(\frac{2}{5}((3+h)^2-9)\right)^{5/4}$$
$$=\lim_{h\to 0_+}\frac{1}{h}\left(\frac{2}{5}h(h+6)\right)^{5/4}=\lim_{h\to 0_+}h^{1/5}\left(\frac{2}{5}(h+6)\right)^{5/4}=0$$

and

$$\lim_{h\to 0_-}\frac{u(3+h)-u(3)}{h}=\lim_{h\to 0_-}\frac{0}{h}=0,$$

then

$$u'(3):=\lim_{h\to 0}\frac{u(3+h)-u(3)}{h}=0.$$

Now we also see that

$$\lim_{t\to 3_+}u'(t)=\lim_{t\to 3_+}\left(\frac{2}{5}\right)^{1/4}t(t^2-9)^{1/4}=0,$$
$$\lim_{t\to 3_-}u'(t)=\lim_{t\to 3_+}0=0,$$

which concludes that $u'\in C(\mathbb{R})$, and thus $u\in C^1(\mathbb{R})$. One can easily check that

$$\begin{cases}u'(t)=f(t,u(t)) \text{ for all } t\in\mathbb{R}, & u(t_0)=0\\ \text{with } f(t,y)=ty^{1/5} \text{ and } t_0=3.\end{cases}\tag{2.1.1}$$

Note that f is continuous in $\mathbb{R}\times\mathbb{R}$, and hence the assumptions in Theorem 2.1.1 satisfy. Since $u\equiv 0$ is also another solution of (2.1.1), one sees that the solution of initial value problem (2.1.1) is *not unique*.

EXERCISE 2.1.3 ([**HS99**, Example III-1-1]). Verify that the initial-value problem

$$u'(t)=(u(t))^{1/3} \text{ for all } t\in\mathbb{R}, \quad u(t_0)=0$$

has at least two nontivial $C^1(\mathbb{R})$-solutions:

$$u(t)=\begin{cases}0 & ,t\le t_0,\\ \left(\frac{2}{3}(t-t_0)\right)^{3/2} & ,t>t_0,\end{cases}$$

and

$$u(t)=\begin{cases}0 & ,t\le t_0,\\ -\left(\frac{2}{3}(t-t_0)\right)^{3/2} & ,t>t_0.\end{cases}$$

EXERCISE 2.1.4 ([**HS99**, Example III-1-3]). Verify that the initial-value problem

$$u'(t)=\sqrt{|u(t)|} \text{ for all } t\in\mathbb{R}, \quad u(t_0)=0$$

has at least one nontivial $C^1(\mathbb{R})$-solution:

$$u(t)=\begin{cases}-\frac{1}{4}(t-t_0)^2 & ,t\le t_0,\\ \frac{1}{4}(t-t_0)^2 & ,t>t_0.\end{cases}$$

We now state a sufficient condition to guarantee also the uniqueness of the solution.

THEOREM 2.1.5 (Fundamental theorem of ODE [**HS99**, Theorem I-1-4]). *Suppose that all assumptions in Theorem 2.1.1 hold. If we additionally assume that*

$$|\boldsymbol{f}(t,\boldsymbol{y}_1)-\boldsymbol{f}(t,\boldsymbol{y}_2)|\le L|\boldsymbol{y}_1-\boldsymbol{y}_2| \tag{2.1.2}$$

whenever $(t,\boldsymbol{y}_1)$ and $(t,\boldsymbol{y}_2)$ are in $\mathcal{R}$, then the solution described in Theorem 2.1.1 is the unique $(C^1((t_0-\alpha,t_0+\alpha)))^n$ solution.

REMARK 2.1.6. See also Theorem 2.1.10 below.

EXERCISE 2.1.7. Verify that the ODEs in Example 2.1.2, Exercise 2.1.3 and Exercise 2.1.4 do not satisfy the Lipschitz condition (2.1.2).

EXERCISE 2.1.8. Under the assumptions of Theorem 2.1.5, show that the initial value problem (2.0.1) is equivalent to the integral equation

$$\boldsymbol{u}(t)=\boldsymbol{u}_0+\int_{t_0}^{t}\boldsymbol{f}(s,\boldsymbol{u}(s))\,\mathrm{d}s$$

for all $t\in(t_0-\alpha,t_0+\alpha)$.

By using Exercise 2.1.8, under the assumptions of Theorem 2.1.5, if $\boldsymbol{u}^1\in C^1((t_0-\alpha,t_0+\alpha))$ and $\boldsymbol{u}^2\in C^1((t_0-\alpha,t_0+\alpha))$ are the unique solution of (2.0.1) corresponding to initial data $\boldsymbol{u}_0^1$ and $\boldsymbol{u}_0^2$ respectively, then one sees that

$$\boldsymbol{u}^1(t)-\boldsymbol{u}^2(t)=\boldsymbol{u}_0^1-\boldsymbol{u}_0^2+\int_{t_0}^{t}\left(\boldsymbol{f}(s,\boldsymbol{u}^1(s))-\boldsymbol{f}(s,\boldsymbol{u}^2(s))\right)\,\mathrm{d}s.$$

By using the Lipschitz condition (2.1.2), one sees that

$$\begin{aligned}|\boldsymbol{u}^1(t)-\boldsymbol{u}^2(t)|&\le|\boldsymbol{u}_0^1-\boldsymbol{u}_0^2|+\int_{t_0}^{t}\left|\boldsymbol{f}(s,\boldsymbol{u}^1(s))-\boldsymbol{f}(s,\boldsymbol{u}^2(s))\right|\,\mathrm{d}s\\&\le|\boldsymbol{u}_0^1-\boldsymbol{u}_0^2|+L\int_{t_0}^{t}|\boldsymbol{u}^1(s)-\boldsymbol{u}^2(s)|\,\mathrm{d}s\quad\text{for all } t\in(t_0-\alpha,t_0+\alpha).\end{aligned} \tag{2.1.3}$$

If $t\le t_0$, then one immediately reach

$$|\boldsymbol{u}^1(t)-\boldsymbol{u}^2(t)|\le|\boldsymbol{u}_0^1-\boldsymbol{u}_0^2|. \tag{2.1.4}$$

If $t\ge t_0$, we will use the following useful lemma.

LEMMA 2.1.9 (Gronwall [**HS99**, Lemma I-1-5]). *If $g\in C([t_0,t_1])$ satisfies the inequality*

$$0\le g(t)\le K+L\int_{t_0}^{t}g(s)\,\mathrm{d}s\quad\textit{for all } t\in[t_0,t_1], \tag{2.1.5}$$

then

$$0\le g(t)\le Ke^{L(t-t_0)}\quad\textit{for all } t\in[t_0,t_1].$$

PROOF. Set $v(t):=\int_{t_0}^{t}g(s)\,\mathrm{d}s$, from (2.1.5) we have

$$\frac{\mathrm{d}v}{\mathrm{d}t}\le K+Lv(t)\text{ for all }t\in(t_0,t_1),\quad v(t_0)=0.$$

One sees that (this technique is called the method of integrating factors)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-L(t-t_0)}v(t)\right) &= e^{-L(t-t_0)}\frac{\mathrm{d}v}{\mathrm{d}t} - Le^{-L(t-t_0)}v(t)\\ &= e^{-L(t-t_0)}\left(\frac{\mathrm{d}v}{\mathrm{d}t} - Lv(t)\right) \leq Ke^{-L(t-t_0)} \quad \text{for all } t \in (t_0, t_1).\end{aligned}$$

Now the fundamental theorem of calculus implies (i.e. integrating the above inequality from t_0 to $\tau \in (t_0, t_1)$)

$$e^{-L(\tau-t_0)}v(\tau) \leq K\int_{t_0}^{\tau} e^{-L(t-t_0)}\,\mathrm{d}t = \frac{K}{L}(1 - e^{-L(\tau-t_0)}) \quad \text{for all } \tau \in (t_0, t_1),$$

which implies

$$\int_{t_0}^{t} g(s)\,\mathrm{d}s = v(t) \leq \frac{K}{L}(e^{L(t-t_0)} - 1).$$

Finally, plugging this into (2.1.5) to see that

$$g(t) \leq K + L\int_{t_0}^{t} g(s)\,\mathrm{d}s \leq K + K(e^{L(t-t_0)} - 1) = Ke^{L(t-t_0)}$$

for all $t \in (t_0, t_1)$, and our result follows from the continuity of g. □

In view of (2.1.4) and (2.1.3), we now choose any $t_0 < t_1 < t_0+\alpha$ and $g(t) = |\boldsymbol{u}^1(t)-\boldsymbol{u}^2(t)|$ as well as $K = |\boldsymbol{u}_0^1 - \boldsymbol{u}_0^2|$ in Lemma 2.1.9 to see that:

THEOREM 2.1.10 (Dependence on data). *If all assumptions in Theorem 2.1.5 hold, then the stability estimate*

$$|\boldsymbol{u}^1(t) - \boldsymbol{u}^2(t)| \leq |\boldsymbol{u}_0^1 - \boldsymbol{u}_0^2|e^{\alpha L} \quad \textit{for all } t \in (t_0 - \alpha, t_0 + \alpha)$$

hold, where $\boldsymbol{u}^1 \in C^1((t_0 - \alpha, t_0 + \alpha))$ *and* $\boldsymbol{u}^2 \in C^1((t_0 - \alpha, t_0 + \alpha))$ *are the unique solution of* (2.0.1) *corresponding to initial data* $\boldsymbol{u}_0^1$ *and* $\boldsymbol{u}_0^2$ *respectively.*

REMARK 2.1.11. We refer to [**HS99**, Chapter II] for further generalizations of Theorem 2.1.10

EXERCISE 2.1.12. Show that the initial value problem

$$\begin{aligned}u'(t) &= \frac{1}{(3-(t-1)^2)(9-(u-5)^2)} \text{ for all } t \in (1-\sqrt{2}, 1+\sqrt{2}),\\ u(1) &= 5,\end{aligned}$$

has a unique $C^1((1-\sqrt{2}, 1+\sqrt{2}))$-solution.

EXERCISE 2.1.13. Show that the initial value problem

$$\begin{aligned}u'(t) &= \frac{1}{(1+(t-4)^2)(5+(u-3)^2)} \text{ for all } t \in \mathbb{R},\\ u(4) &= 3,\end{aligned}$$

has a unique $C^1(\mathbb{R})$-solution.

2.2. Some techniques for solving the equation

We now consider a single equation of ODE:

$$u'(t) = f(t, u(t)), \quad u(t_0) = u_0. \tag{2.2.1}$$

Unfortunately, there is no universally applicable method for solving solution(s) $u \in C^1$ for the equation (2.2.1) . We now exhibit some methods which can help to solve some certain class of ODE.

DEFINITION 2.2.1. The ODE (2.2.1) is said to be *separable* if it can be expressed in the form of

$$M(t) + N(u(t))u'(t) = 0, \tag{2.2.2}$$

for some continuous functions M and N, or sometimes we abuse the notation by writing $M(t)\,\mathrm{d}t + N(u)\,\mathrm{d}u = 0$.

For $u \in C^1$, we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{u_0}^{u(t)} N(z)\,\mathrm{d}z\right) = N(u(t))u'(t),$$

then one can rewrite (2.2.2) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{u_0}^{u(t)} N(z)\,\mathrm{d}z\right) = -M(t).$$

Integrate both sides with respect to the variable t from t_0 to τ, we see that

$$\int_{u_0}^{u(\tau)} N(z)\,\mathrm{d}z = \int_{t_0}^{\tau} \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{u_0}^{u(t)} N(z)\,\mathrm{d}z\right)\,\mathrm{d}t = -\int_{t_0}^{\tau} M(t)\,\mathrm{d}t, \tag{2.2.3}$$

which solves the ODE implicitly.

EXAMPLE 2.2.2. We now want to find the general solution of

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{t^2}{1-u^2}. \tag{2.2.4}$$

We see that, if we impose the initial condition $u(t_0) = u_0 \neq 1$, the conditions in Theorem 2.1.5 hold, therefore there exists a unique solution $u \in C^1$ near t_0. Note that $u_0 \neq 1$ and the continuity of u impliy that $u(t) \neq 1$ for all t near t_0, therefore (2.2.4) make sense. Now we use chain rule to see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(u(t) - \frac{1}{3}(u(t))^3\right) = (1-u^2)\frac{\mathrm{d}u}{\mathrm{d}t} = t^2,$$

and thus

$$u(\tau) - \frac{1}{3}(u(\tau))^3 - \left(u_0 - \frac{1}{3}u_0^3\right) = \int_{t_0}^{\tau} \frac{\mathrm{d}}{\mathrm{d}t}\left(u(t) - \frac{1}{3}(u(t))^3\right)\,\mathrm{d}t = \int_{t_0}^{\tau} t^2\,\mathrm{d}t = \frac{1}{3}\tau^3 - \frac{1}{3}t_0^3,$$

which gives

$$-t^3 + 3u(t) - (u(t))^3 = -t_0^3 + 3u_0 - u_0^3 \quad \text{for all } t \text{ near } t_0,$$

or in the form of depressed cubic form

$$(u(t))^3 - 3u(t) + (t^3 - t_0^3 + 3u_0 - u_0^3) = 0,$$

and thus u can be expressed by using Cardano's formula[1] case by case.

EXERCISE 2.2.3. Do the same thing for the ODE

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{3t^2 + 4t + 2}{2(u-1)}.$$

EXERCISE 2.2.4. Do the same thing for the ODE

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{4t - t^3}{4 + u^3}.$$

We now consider the following ODE:

$$M(t, u(t)) + N(t, u(t))u'(t) = 0, \quad u(t_0) = u_0. \tag{2.2.5}$$

Note that (2.2.1) and (2.2.2) are both special case of (2.2.5). We now want to solve (2.2.5) under some sufficient conditions.

Assume that $M = M(t, y)$, $N = N(t, y)$, $\partial_y M$ and $\partial_t N$ are continuous in an open rectangle $(t_1, t_2) \times (y_1, y_2)$ and $u \in C^1((t_1, t_2))$ satisfies $y_1 < u(t) < y_2$ for all $t \in (y_1, y_2)$. For each $\psi \in C^1((t_1, t_2) \times (y_1, y_2))$, by using chain rule one sees that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t, u(t))\right) = \partial_t \psi(t, u(t)) + \partial_y \psi(t, u(t))u'(t). \tag{2.2.6}$$

Comparing this equality with the ODE (2.2.5), it is natural to find ψ such that $\partial_y \psi = N$ in $(t_1, t_2) \times (y_1, y_2)$, which can be achieved by choosing

$$\psi(t, y) := \int_{u_0}^{y} N(t, z)\,\mathrm{d}z \quad \text{for all } (t, y) \in (t_1, t_2) \times (y_1, y_2).$$

We need the following lemma for further computations (one way to prove this is to utilize the Lebesgue dominated convergence theorem):

LEMMA 2.2.5 ([**Str08**, Theorem 1 in Appendix A.3]). *Suppose that $f(t, y)$ and $\partial_t f(t, y)$ are continuous in the closed rectangle $[s_1, s_2] \times [z_1, z_2]$, then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{z_1}^{z_2} f(t, y)\,\mathrm{d}y\right) = \int_{z_1}^{z_2} \partial_t f(t, y)\,\mathrm{d}y \quad \textit{for all } t \in [s_1, s_2].$$

Now the above lemma guarantees that

$$\partial_t \psi(t, y) = \int_{u_0}^{y} \partial_t N(t, z)\,\mathrm{d}z \quad \text{for all } (t, y) \in (t_1, t_2) \times (y_1, y_2).$$

Now if

$$\partial_t N = \partial_y M \quad \text{in } (t_1, t_2) \times (y_1, y_2), \tag{2.2.7}$$

then we reach

$$\partial_t \psi(t, y) = \int_{u_0}^{y} \partial_z M(t, z)\,\mathrm{d}z = M(t, y) - M(t, u_0) \quad \text{for all } (t, y) \in (t_1, t_2) \times (y_1, y_2).$$

Now from (2.2.6), and consequently by (2.2.5), we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t, u(t))\right) = M(t, u(t)) - M(t, u_0) + N(t, u(t))u'(t) = -M(t, u_0),$$

[1] https://en.wikipedia.org/wiki/Cubic_equation

thus

$$\psi(\tau, u(\tau)) - \psi(t_0, u_0) = \int_{t_0}^{\tau} \frac{\mathrm{d}}{\mathrm{d}t} \left(\psi(t, u(t))\right) \, \mathrm{d}t = -\int_{t_0}^{\tau} M(t, u_0) \, \mathrm{d}t \tag{2.2.8}$$

which solves u implicitly. In view of the above ideas, it is natural to introduce the following definition.

DEFINITION 2.2.6. The ODE (2.2.5) is said to be exact if (2.2.7) holds.

REMARK 2.2.7. If the ODE is separable (in the sense of Definition 2.2.1), then it also exact. In this case, (2.2.8) reduces to (2.2.3).

EXAMPLE 2.2.8. We now want to solve the ODE

$$(u(t)\cos t + 2te^{u(t)}) + (\sin t + t^2 e^{u(t)} - 1)u'(t) = 0$$

with suitable initial condition $u(t_0) = u_0$. In view of the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\psi(t, u(t))\right) = \partial_t \psi(t, u(t)) + \partial_y \psi(t, u(t)) u'(t), \tag{2.2.9}$$

it is natural to choose

$$\psi(t, y) = \int_{u_0}^{y} (\sin t + t^2 e^z - 1) \, \mathrm{d}z = (y \sin t + t^2 e^y - y) - (u_0 \sin t + t^2 e^{u_0} - u_0)$$

so that $\partial_y \psi = \sin t + t^2 e^y - 1$. Now we compute

$$\partial_t \psi(t, y) = y \cos t + 2te^y - 2te^{u_0},$$

and from (2.2.9), and consequently the ODE, we see that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\psi(t, u(t))\right) = (u(t)\cos t + 2te^{u(t)} - 2te^{u_0}) + (\sin t + t^2 e^{u(t)} - 1)u'(t) = -2te^{u_0},$$

thus

$$\begin{aligned} t_0^2 e^{u_0} - \cancel{\tau^2 e^{u_0}} &= -\int_{t_0}^{\tau} 2te^{u_0} \, \mathrm{d}t = \int_{t_0}^{\tau} \frac{\mathrm{d}}{\mathrm{d}t} \left(\psi(t, u(t))\right) \, \mathrm{d}t = \psi(\tau, u(\tau)) - \psi(t_0, u_0) \\ &= (u(\tau)\sin\tau + \tau^2 e^{u(\tau)} - u(\tau)) - (u_0 \sin t_0 + \cancel{\tau^2 e^{u_0}} - u_0), \end{aligned}$$

that is,

$$u(t)\sin t + t^2 e^{u(t)} - u(t) = C := t_0^2 e^{u_0} + u_0 \sin t_0 - u_0.$$

We now want to deal with the ODE (2.2.5) which is not necessarily to be exact in the sense of Definition 2.2.6. The idea is quite simple: we want to multiply an integrating factor $\mu(t, y)$ so that

$$\tilde{M}(t, u(t)) + \tilde{N}(t, u(t))u'(t) = 0, \quad u(t_0) = u_0$$

is exact, where $\tilde{M}(t, y) = \mu(t, y)M(t, y)$ and $\tilde{N}(t, y) = \mu(t, y)N(t, y)$. Now (2.2.7) reads

$$\partial_t \mu N + \mu \partial_t N = \partial_t \tilde{N} = \partial_y \tilde{M} = \partial_y \mu M + \mu \partial_y M,$$

that is,

$$M\partial_y \mu - N\partial_t \mu + (\partial_y M - \partial_t N)\mu = 0. \tag{2.2.10}$$

This is nothing but just a transport equation, which will be discussed in Section 2.3 below. It is remarkable to mention that, one can directly check that if

$$k := \frac{\partial_y M - \partial_t N}{N} \text{ is a function of } t \text{ only},$$

the integrating factor μ is also a function of t only (which is independent of y), and it satisfies the linear ODE

$$\mu'(t) = k(t)\mu(t),$$

so that for each K with $K' = k$ one has

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{-K(t)}\mu(t)) = -k(t)e^{-K(t)}\mu(t) + e^{-K(t)}\mu'(t) = 0.$$

This can be achieved by choosing

$$\mu(t) := e^{K(t)}.$$

EXAMPLE 2.2.9. We now want to solve the ODE $(3tu+u^2)+(t^2+tu)u' = 0$ with suitable initial condition $u(t_0) = u_0$. We want to multiply an integrating factor $\mu = \mu(t,y)$ so that

$$\begin{aligned}
&\partial_t\left(\mu(t,y)(t^2+ty)\right) = \partial_y\left(\mu(t,y)(3ty+y^2)\right)\\
\iff &\partial_t\mu(t,y)(t^2+ty) + \mu(t,y)(2t+y)\\
&\qquad = \partial_y\mu(t,y)(3ty+y^2) + \mu(t,y)(3t+2y)\\
\iff &(t\partial_t\mu(t,y) - \mu(t,y))(t+y) = y\partial_y\mu(t,y)(3t+y).
\end{aligned}$$

Note that the choice $\mu(t,y) = t$ fulfills the above requirement. We only need to find a μ, no need to find its general solution. We now write the ODE as

$$(3t^2u + tu^2) + (t^3 + t^2u)u' = 0, \quad u(t_0) = u_0. \tag{2.2.11}$$

In view of the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t,u(t))\right) = \partial_t\psi(t,u(t)) + \partial_y\psi(t,u(t))u'(t), \tag{2.2.12}$$

it is natural to choose ψ such that $\partial_y(t,y) = t^3 + t^2y$. One way to achieve this is take

$$\psi(t,y) = \int_{u_0}^{y}(t^3+t^2z)\,\mathrm{d}z = \left(t^3y + \frac{1}{2}t^2y^2\right) - \left(t^3u_0 + \frac{1}{2}t^2u_0^2\right).$$

We now compute that

$$\partial_t\psi(t,y) = \left(3t^2y + ty^2\right) - \left(3t^2u_0 + tu_0^2\right).$$

Now from (2.2.12), and consequently from (2.2.11), we reach

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t,u(t))\right) = \left(3t^2u + tu^2\right) - \left(3t^2u_0 + tu_0^2\right) + (t^3+t^2u)u' = 3t^2u_0 - tu_0^2,$$

thus

$$\begin{aligned}
&\left(\tau^3u(\tau) + \frac{1}{2}\tau^2(u(\tau))^2\right) - \cancel{\left(\tau^3u_0 + \frac{1}{2}\tau^2u_0^2\right)}\\
&\quad = \psi(\tau,u(\tau)) - \psi(t_0,u_0) = \int_{t_0}^{\tau}\frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t,u(t))\right)\,\mathrm{d}t\\
&\quad = \int_{t_0}^{\tau}(3t^2u_0 + tu_0^2)\,\mathrm{d}t = \cancel{\left(\tau^3u_0 + \frac{1}{2}\tau^2u_0^2\right)} - \left(t_0^3u_0 + \frac{1}{2}t_0^2u_0^2\right),
\end{aligned}$$

which concludes that

$$\tau^2(u(\tau))^2 + 2\tau^3u(\tau) = -2t_0^3u_0 - t_0^2u_0^2.$$

2.3. From ODE to PDE

2.3.1. Linear equations. We now give an application of ODE in the theory of partial differential equations (PDE). We begin our discussions from a simple model. Given a horizontal pipe of fixed cross section in the (positive) x-direction. Suppose that there is a fluid flowing at a constant rate c ($c = 0$ means the fluid is stationary; $c > 0$ means flowing toward right, otherwise towards left). We now assume that there is a substance is suspended in the water. Fix a point at the pipe, and we set the point as the origin 0, and let $u(t, x)$ be the concentration of such substance. The amount of pollutant in the interval $[0, y]$ at time t is given by

$$\int_0^y u(t, x)\,\mathrm{d}x.$$

At the later time $t + \tau$, the same molecules of pollutant moved by the displacement $c\tau$, and this means

$$\int_0^y u(t, x)\,\mathrm{d}x = \int_{c\tau}^{y+c\tau} u(t + \tau, x)\,\mathrm{d}x.$$

If u is continuous, by using the fundamental theorem of calculus, by differentiating the above equation with respect to y, one sees that

$$u(t, y) = u(t + \tau, y + c\tau) \quad \text{for all } y \in \mathbb{R}. \tag{2.3.1}$$

If we further assume $u \in C^1$, then differentiating (2.3.1) with respect to τ, we reach the following *transport equation*:

$$0 = u(t + \tau, y + c\tau)|_{\tau=0} = \partial_t u(t, x) + c\partial_x u(t, x) \quad \text{for all } (t, x) \in \mathbb{R} \times \mathbb{R}.$$

We now consider the transport equation with variable coefficient equation of the form

$$\partial_t u + c(t, x)\partial_x u = 0, \quad u(0, x) = f(x), \tag{2.3.2}$$

where $f \in C^1(\mathbb{R})$ and $c = c(t, x)$ satisfies all assumption in Theorem 2.1.5. Given any $s \in \mathbb{R}$ and we consider a curve $x = \gamma_s(t)$, where γ solves the ODE

$$\gamma_s'(t) = c(t, \gamma_s(t)), \quad \gamma_s(0) = s. \tag{2.3.3}$$

We now restrict u on a curve $x = \gamma_s(t)$, and one sees that

$$\begin{aligned}\partial_t \left(u|_{\gamma_s}(t)\right) = \partial_t \left(u(t, \gamma_s(t))\right) &= \left(\partial_t u + \gamma'(t)\partial_x u\right)|_{x=\gamma_s(t)} \\ &= \left(\partial_t u + c(t, x)\partial_x u\right)|_{x=\gamma_s(t)} = 0.\end{aligned}$$

This means that u is constant along the characteristic curve γ_s. Hence

$$u(t, \gamma_s(t)) = u(0, \gamma_s(0)) = f(\gamma_s(0)) = f(s). \tag{2.3.4}$$

For later convenience, we write $\gamma(t, s) = \gamma_s(t)$. Fix $x \in \mathbb{R}$ and we now want to solve the equation $x = \gamma(t, s)$. From $\gamma(0, x) = x$, and since $\partial_s\gamma(0, x) = (\partial_s\gamma_s(0))|_{s=x} = 1 \neq 0$, then we can apply the implicit function theorem [**Apo74**, Theorem 13.7] to guarantee that there exist an open neighborhood $U_x \subset \mathbb{R}$ of 0 and $g_x \in C^1(U_x)$ such that $g_x(0) = x$ and $x = \gamma(t, s)|_{s=g_x(t)}$ for all $t \in U_x$. In other words, we found a solution $s = g_x(t) \equiv g(x, t)$ of the equation $x = \gamma(t, s)$ in U_x. Plugging this solution into (2.3.4), we conclude

$$u(t, x) = f(g(x, t)) \quad \text{for all } x \text{ with } \partial_s\gamma(0, x) \neq 0 \text{ and } t \in U_x. \tag{2.3.5}$$

This completes the local existence proof. Uniqueness follows from the fact that u is constant along the characteristic curve γ.

EXAMPLE 2.3.1. Given any $f \in C^1(\mathbb{R})$, let us now consider (2.3.2) with $c =$ constant. In this case, (2.3.3) reads $\gamma'(t) = c$. For each $s \in \mathbb{R}$, it is easy to see that the solution of $\gamma_s'(t) = c$ with $\gamma_s(0) = s$ is

$$\gamma(t, s) \equiv \gamma_s(t) = ct + s.$$

For each $x \in \mathbb{R}$, the solution of $x = \gamma(t, s)$ is clearly given by $s = g(x, t) \equiv x - ct$, and thus from (2.3.5) we conclude that

$$u(t, x) = f(x - ct).$$

EXAMPLE 2.3.2. Given any $f \in C^1(\mathbb{R})$, we now want to solve $\partial_t u + x\partial_x u = 0$ with $u(0, x) = f(x)$ for all $x \in \mathbb{R}$. Write $c(t, x) = x$, and for each $s \in \mathbb{R}$ we consider the ODE

$$\gamma_s'(t) = c(t, \gamma_s(t)) \equiv \gamma_s(t), \quad \gamma_s(0) = s.$$

By using the integrating factor, one can easily see that the solution of the ODE is

$$\gamma_s(t) = e^t s.$$

For each $x \in \mathbb{R}$, the solution of $x = \gamma_s(t)$ is given by $s = g(x, t) \equiv e^{-t}x$, and thus from (2.3.5) we conclude that

$$u(t, x) = f(g(x, t)) = f(e^{-t}x).$$

EXAMPLE 2.3.3. Given any $f \in C^1(\mathbb{R})$, we now want to solve $\partial_t u + 2tx^2\partial_x u = 0$ with $u(0, x) = f(x)$ for all $x \in \mathbb{R}$. Write $c(t, x) = 2tx^2$, and for each $s \neq 0$ we consider the ODE

$$\gamma_s'(t) = 2t(\gamma_s(t))^2, \quad \gamma_s(0) = s^{-1}.$$

By using the method of separation of variables, one can easily see that the solution of the ODE is

$$\gamma_s(t) = (s - t^2)^{-1},$$

which is valid

$$\begin{cases} \text{for all } t \in \mathbb{R} & \text{when } s < 0, \\ \text{for all } t^2 < s & \text{when } s > 0, \end{cases} \tag{2.3.6}$$

but *the ODE is not solvable* when $s = 0$. When $s \neq 0$, the solution of $x = \gamma_s(t)$ is given by $s = t^2 + \frac{1}{x}$, and thus from (2.3.5) we conclude that

$$u(t, x) = f(s^{-1}) = f\left(\frac{x}{1 + t^2 x}\right) \quad \text{for all } x \neq -t^{-2}.$$

We now summarize the above ideas in the following algorithm:

Algorithm 1 Solving $\partial_t u + c(t, x)\partial_x u + d(t, x)u = F(t, x)$ with $u(0, x) = f(x)$

1: Solve the ODE $\gamma_s'(t) = c(t, \gamma_s(t))$ with given $\gamma_s(0)$ for any suitable parameter s.
2: Compute $\partial_t(u(t, \gamma_s(t)))$.
3: Rewrite the identity $x = \gamma_s(t)$ in the form of $s = g(x, t)$.
4: Identify the domain for which $u(t, x) = f(g(x, t))$ solves $\partial_t u + c(t, x)\partial_x u = 0$.

EXERCISE 2.3.4. Given any $f \in C^1(\mathbb{R})$, solve the equation $(1 + t^2)\partial_t u + \partial_x u = 0$ with $u(0, x) = f(x)$ and identify the range of x.

EXERCISE 2.3.5. Given any $f \in C^1(\mathbb{R})$, solve the equation $t\partial_t u + x\partial_x u = 0$ with $u(0, x) = f(x)$ and identify the range of x.

EXERCISE 2.3.6. Solve the equation $x\partial_t u + t\partial_x u = 0$ with $u(0, x) = e^{-x^2}$.

2.3.2. Quasilinear equations. The ideas in previous subsection can be extend for quasilinear equation of the form

$$a(x,y,u)\partial_x u + b(x,y,u)\partial_y u = c(x,y,u). \tag{2.3.7}$$

Here we follow the approach in [**Joh78**, Sections 1.4–1.6]. We write (2.3.7) as

$$(a,b,c)\cdot(\partial_x u, \partial_y u, -1) = 0.$$

We represent the function u by a surface $z = u(x,y)$ in $\mathbb{R}^3$, and we write

$$(a,b,c)\cdot\left(\frac{\mathrm{d}z}{\mathrm{d}x}, \frac{\mathrm{d}z}{\mathrm{d}y}, -1\right) = 0.$$

Note that $\left(\frac{\mathrm{d}z}{\mathrm{d}x}, \frac{\mathrm{d}z}{\mathrm{d}y}, -1\right)$ is the normal vector of the surface, thus (a,b,c) is a tangent vector. We now consider a "regular" curve $(x(t), y(t), z(t))$ in that surface, and now we see that $(x'(t), y'(t), z'(t))$ is a tangent vector at the point $(x(t), y(t), z(t))$. This suggests us to consider the characteristic ODE:

$$\begin{cases} x'(t) = a(x(t), y(t), z(t)), \\ y'(t) = b(x(t), y(t), z(t)), \\ z'(t) = c(x(t), y(t), z(t)), \end{cases} \tag{2.3.8}$$

which is a special case of the ODE (2.0.1). Here, the system is even autonomous, i.e. the coefficients are independent of variable t does not appear explicitly. If we assume that $a, b, c \in C^1$, then one can apply Theorem 2.1.1 to ensure the existence of characteristic curve $(x(t), y(t), z(t))$ which is C^1. We now prove that the above choice of the characteristic ODE really describes the surface $z = u(x,y)$.

LEMMA 2.3.7. *Assume that $a, b, c \in C^1$ near $(x_0, y_0, z_0) \in S$, where S is the surface described by $z = u(x,y)$. If γ is a C^1 curve described by $(x(t), y(t), z(t))$ with $(x(t_0), y(t_0), z(t_0)) = (x_0, y_0, z_0)$, then γ lies completely on S.*

PROOF. For convenience, we write $U(t) := z(t) - u(x(t), y(t))$ so that $U(t_0) = 0$ since $(x_0, y_0, z_0) \in S$. Using chain rule and from (2.3.8) one sees that

$$\begin{aligned} U'(t) &= z'(t) - (\partial_x u)x'(t) - (\partial_y u)y'(t) \\ &= c(x,y,z) - \partial_x u(x,y)a(x,y,z) - \partial_y u(x,y)b(x,y,z) \\ &= c(x,y,U - u(x,y)) - \partial_x u(x,y)a(x,y,U-u(x,y)) \\ &\quad - \partial_y u(x,y)b(x,y,U-u(x,y)). \end{aligned} \tag{2.3.9}$$

From (2.3.7), we see that $U \equiv 0$ is a solution of the ODE (2.3.9). By using the fundamental theorem of ODE (Theorem 2.1.5), we see that $U \equiv 0$ is the unique solution of the ODE (2.3.9), which concludes our lemma. □

We now want to solve the Cauchy problem for (2.3.7) with the Cauchy data

$$h(s) = u(f(s), g(s)) \quad \text{for some } f, g, h \in C^1 \text{ near } s_0. \tag{2.3.10}$$

Note that the initial value problem we previous considered is simply the special case when $f(s) \equiv x_0$ and $g(s) = s$. Now the characteristic ODE (2.3.8) (with suitable parameterization)

reads

$$\begin{cases} \partial_t X(s,t) = a(X(s,t),Y(s,t),Z(s,t)), \\ \partial_t Y(s,t) = b(X(s,t),Y(s,t),Z(s,t)), \\ \partial_t Z(s,t) = c(X(s,t),Y(s,t),Z(s,t)), \\ \quad \text{with initial conditions} \\ X(s,0) = f(s), \quad Y(s,0) = g(s), \quad Z(s,0) = h(s). \end{cases} \tag{2.3.11}$$

If $a, b, c \in C^1$ near $(f(s_0), g(s_0), h(s_0))$, thus the fundamental theorem of ODE (Theorem 2.1.5) guarantees that there exists a unique solution of (2.3.11):

$$(X(s,t), Y(s,t), Z(s,t))$$

which is C^1 for (s,t) near $(s_0, 0)$. If

$$\det \begin{pmatrix} f'(s_0) & g'(s_0) \\ a(x_0,y_0,z_0) & b(x_0,y_0,z_0) \end{pmatrix} = \det \begin{pmatrix} \partial_s X(s_0,0) & \partial_s Y(s_0,0) \\ \partial_t X(s_0,0) & \partial_t Y(s_0,0) \end{pmatrix} \neq 0,$$

then we can use implicit function theorem [**Apo74**, Theorem 13.7] to guarantee that there exists a unique solution $(s,t) = (S(x,y), T(x,y))$ of

$$x = X(S(x,y), T(x,y)), \quad y = Y(S(x,y), T(x,y))$$

of class C^1 is a neighborhood of (x_0, y_0) and satisfying

$$S(x_0, y_0) = s_0, \quad T(x_0, y_0) = 0,$$

so that we finally we conclude that the local solution of the Cauchy problem for (2.3.7) with the Cauchy data (2.3.10) is given by

$$u(x,y) = Z(S(x,y), T(x,y)).$$

The above arguments can be readily extend for higher dimensional case:

THEOREM 2.3.8. *We now consider the Cauchy problem*

$$\sum_{i=1}^{n} a_i(x_1, \cdots, x_n, u) u_{x_i} = c(x_1, \cdots, x_n, u)$$

with Cauchy data

$$h(s_1, \cdots, s_{n-1}) = u(f_1(s_1, \cdots, s_{n-1}), \cdots, f_n(s_1, \cdots, s_{n-1}))$$

for some $f, \cdots, f_n, h \in C^1$ *near* $(s_1^0, \cdots, s_{n-1}^0)$. *If* $a_1, \cdots, a_n, c \in C^1$ *near* $(f_1(s_0), \cdots, f_n(s_0), h(s_0))$ *such that*

$$\det \begin{pmatrix} \partial_{s_1} f_1(s_1^0, \cdots, s_{n-1}^0) & \cdots & \partial_{s_1} f_n(s_1^0, \cdots, s_{n-1}^0) \\ \vdots & & \vdots \\ \partial_{s_n} f_1(s_1^0, \cdots, s_{n-1}^0) & \cdots & \partial_{s_n} f_n(s_1^0, \cdots, s_{n-1}^0) \\ a_1(x_1^0, \cdots, x_n^0, z^0) & \cdots & a_n(x_1^0, \cdots, x_n^0, z^0) \end{pmatrix} \neq 0,$$

where $x_i^0 = f_i(s_1^0, \cdots, s_{n-1}^0)$ *for all* $i = 1, \cdots, n$ *and* $z^0 = h(s_1^0, \cdots, s_{n-1}^0)$, *then there exists a unique* C^1 *solution* $u = u(x_1, \cdots, x_n)$ *near* $(x_1^0, \cdots, x_n^0, z^0)$.

REMARK 2.3.9. The corresponding characteristic ODE is

$$\begin{aligned}\partial_t x_i(s_1,\cdots,s_{n-1},t) &= a_i(x_1,\cdots,x_n,z) \quad \text{for } i=1,\cdots,n,\\ \partial_t z(s_1,\cdots,s_{n-1},t) &= c(x_1,\cdots,x_n,z),\end{aligned}$$

with initial condition

$$\begin{aligned}x_i(s_1,\cdots,s_{n-1},0) &= f_i(s_1,\cdots,s_{n-1}) \quad \text{for } i=1,\cdots,n,\\ z(s_1,\cdots,s_{n-1},0) &= h(x_1,\cdots,x_n,z).\end{aligned}$$

EXAMPLE 2.3.10. We now want to solve the initial value problem

$$u\partial_x u + \partial_y u = 0, \quad u(x,0) = h(x).$$

The characteristic ODE is

$$\partial_t x(s,t) = z, \quad \partial_t y(s,t) = 1, \quad \partial_t z(s,t) = 0$$

with initial condition

$$x(s,0) = s, \quad y(s,0) = 0, \quad z(s,0) = h(s).$$

Solving the ODE yields

$$x = s + zt, \quad y = t, \quad z = h(s).$$

Eliminating s,t yields the implicit equation

$$u(x,y) = h(x - u(x,y)y).$$

It is interesting to see that

$$\partial_x u = h'(x - u(x,y)y)(1 - y\partial_x u),$$

then

$$\partial_x u + yh'(x - u(x,y)y)\partial_x u = h'(x - u(x,y)y),$$

and this implies

$$\partial_x u(x,y) = \frac{h'(x - u(x,y)y)}{1 + yh'(x - u(x,y)y)}.$$

We see, for example when $h(z) = -z$, that

$$\partial_x u(x,y) = \frac{-1}{1-y}.$$

and this quantity will blow up at $y = 1$, which means that there cannot exist a strict solution u of class C^1 beyond $y = 1$. This type of behavior is typical for a nonlinear partial differential equation. In general, we need to consider "weak" solution to study the PDE, but we will not going to go too far beyond this point.

EXERCISE 2.3.11. Solve the initial value problem

$$xu\partial_x u - \partial_y u = 0, \quad u(x,0) = x.$$

REMARK 2.3.12. For the general first order equation, we refer [**Joh78**, Sections 1.7] for details. Here we will not going to discuss here.

CHAPTER 3

Linear ODE

We now study the structure of solutions of a linear system

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y} + \boldsymbol{b}(t),$$

where the entries of the $n \times n$ matrix $A(t)$ are complex-valued continuous functions of a real independent variable t, and $\boldsymbol{b}(t)$ is a complvex-valued continuous function. Well-posedness of the ODE can be guaranteed by the fundamental theorem of ODE (Theorem 2.1.5). We will follow the approach in some parts of [**HS99**, Chapter IV].

3.1. Homogeneous ODE with constant coefficients

The main theme of this section is to show that the unique (guaranteed by the fundamental theorem of ODE in Theorem 2.1.5) matrix-valued solution $Y = Y(t)$ of

$$Y'(t) = AY(t) \text{ for all } t \in \mathbb{R}, \quad Y(0) = I, \tag{3.1.1}$$

which is called the *fundamental matrix solution*, where I is the $n \times n$ identity matrix, takes the form

$$Y(t) = \exp(tA) \quad \text{for all } t \in \mathbb{R}. \tag{3.1.2}$$

If this is the case, then the unique solution of

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{p}$$

is exactly

$$\boldsymbol{y}(t) = \exp((t - t_0)A)\boldsymbol{p} \quad \text{for all } t \in R.$$

If $n = 1$, the above discussions are trivial, we are interested in the case when $n \geq 2$.

Let $\mathbb{C}^{n\times n}$ denote the set of all $n \times n$ matrices whose entries are complex numbers. The set of all inverible matix with entries in $\mathbb{C}$ is denoted by $\mathrm{GL}(n, \mathbb{C})$, which also can be characterized by

$$\mathrm{GL}(n, \mathbb{C}) = \{A \in \mathbb{C}^{n\times n} : \det A \neq 0\}.$$

The collection $\mathrm{GL}(n, \mathbb{C})$ is known as the *general linear group* of order n. For any $A \in \mathbb{C}^{n\times n}$, we define the *Hilbert-Schmidt* norm

$$\|A\| := \left(\sum_{j,k=1}^{n} |A_{jk}|^2 \right)^{1/2}, \tag{3.1.3}$$

where A_{jk} is the of A on the j^{th} row and the k^{th} column. The trace of $A \in \mathbb{C}^{n\times n}$ is defined by

$$\mathrm{tr}\,(A) := \sum_{j=1}^{n} A_{jj}.$$

Exercise 3.1.1. Let $A, B \in \mathbb{C}^{n\times n}$, show that

(a) $\|X\| = (\operatorname{tr}(A^*A))^{1/2}$, where A^* is the *conjugate transpose* (or *adjoint*) of $A \in \mathbb{C}^{n\times n}$.
(b) $\|X+Y\| \le \|X\| + \|Y\|$,
(c) $\|XY\| \le \|X\|\|Y\|$.

DEFINITION 3.1.2. Let $\{A_m\}$ be a sequence of complex matrices in $\mathbb{C}^{n\times n}$. We say that A_m converges to matrix A if

$$\lim_{m\to\infty} (A_m)_{jk} = A_{jk} \quad \text{for all } 1 \le j,k \le n.$$

EXERCISE 3.1.3. Show that A_m converges to Aif and only if $\lim_{m\to\infty} \|A_m - A\| = 0$.

We first need to prove the following lemma.

LEMMA 3.1.4 ([**Hal15**, Proposition 2.1]). *For each $A \in \mathbb{C}^{n\times n}$, we define A^m be the repeated matrix product of X with itself and $A^0 = I$. Then the series*

$$\exp(A) := \sum_{m=0}^{\infty} \frac{A^m}{m!} \tag{3.1.4}$$

converges absolutely (in the sense of Exercise 3.1.3). In addition, the function

$$A \in \mathbb{C}^{n\times n} \mapsto e^A \in \mathbb{C}^{n\times n}$$

is a continuous function (with respect to the Hilbert-Schmidt norm (3.1.3)).

REMARK 3.1.5. Lemma 3.1.4 can be rephrase as: the radius of converge of the power series (3.1.4) is $+\infty$.

PROOF OF LEMMA 3.1.4. By using Exercise 3.1.1(c), we see that

$$\|X^m\| \le \|X\|^m \quad \text{for all } m \in \mathbb{N},$$

hence

$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \le \|I\| + \sum_{m=1}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} < +\infty,$$

which shows that (3.1.4) converges absolutely. On the other hand, we see that

$$\sup_{\|A\|\le R} \left\| \exp(A) - \sum_{m=0}^{N} \frac{A^m}{m!} \right\| = \sup_{\|A\|\le R} \left\| \sum_{m=N+1}^{\infty} \frac{A^m}{m!} \right\| \le \sum_{m=N+1}^{\infty} \frac{R^m}{m!},$$

hence

$$\sup_{\|A\|\le R} \left\| \exp(A) - \sum_{m=0}^{N} \frac{A^m}{m!} \right\| \to 0 \quad \text{as } N \to +\infty,$$

thus (3.1.4) converges uniformly on the closed ball $\{A \in \mathbb{C}^{n\times n} : \|A\| \le R\}$, and thus $A \mapsto e^A$ is continuous on the open ball $\{A \in \mathbb{C}^{n\times n} : \|A\| < R\}$. Since this holds true for all $R > 0$, hence we conclude our result. □

It is easy to see that

$$\begin{cases} \exp(0) = I, \quad (\exp(A))^* = \exp(A^*), \\ \exp(BAB^{-1}) = B\exp(A)B^{-1} \text{ for all } B \in \mathrm{GL}(n,\mathbb{C}). \end{cases} \tag{3.1.5}$$

The following lemma is crucial (see also [**Hal15**, Theorem 5.1] for a generalization).

LEMMA 3.1.6 ([Hal15, Proposition 2.3]). *If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ are commute (i.e. $AB = BA$), then*

$$\exp(A+B) = \exp(A)\exp(B) = \exp(B)\exp(A). \tag{3.1.6}$$

PROOF. We simply multiply the two power series $\exp(A)$ and $\exp(B)$ term by term, which is permitted because both series converge absolutely (by Lemma 3.1.4). We also able to rearrange the terms (since A and B are commute), so we can collect terms where the power of A plus the power of B equals to m:

$$\begin{aligned}\exp(A)\exp(B) &= \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{A^k}{k!}\frac{B^{m-k}}{(m-k)!}\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}A^kB^{m-k}\\ &= \sum_{m=0}^{\infty}\frac{(A+B)^m}{m!} = \exp(A+B),\end{aligned}$$

which conclude our lemma. □

EXAMPLE 3.1.7. In general, the identity (3.1.6) does not hold true for those $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ which are not commute. For example, we choose

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

One sees that

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = BA.$$

Since $A^2 = A$ and $B^2 = B$, then we see that

$$\begin{aligned}\exp(A) &= I + A\sum_{m=1}^{\infty}\frac{1}{m!} = I + (e-1)A = \begin{pmatrix} 1 & e-1 \\ 0 & 1 \end{pmatrix},\\ \exp(B) &= I + B\sum_{m=1}^{\infty}\frac{1}{m!} = I + (e-1)B = \begin{pmatrix} 1 & 0 \\ e-1 & 1 \end{pmatrix},\end{aligned}$$

thus

$$\exp(A)\exp(B) = \begin{pmatrix} 1+(e-1)^2 & e-1 \\ e-1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & e-1 \\ e-1 & 1+(e-1)^2 \end{pmatrix} = \exp(B)\exp(A).$$

Since $(A+B)^2 = I$, hence we see that

$$\begin{aligned}\exp(A+B) &= \sum_{m:\text{even}}\frac{1}{m!}I + \sum_{m:\text{odd}}\frac{1}{m!}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\\ &= \sum_{m=0}^{\infty}\frac{1}{(2m)!}I + \sum_{m=0}^{\infty}\frac{1}{(2m+1)!}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\\ &= \cosh(1)I + \sinh(1)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{pmatrix}\end{aligned}$$

where

$$\cosh x = \frac{e^{-x} + e^{x}}{2}, \quad \sinh x = \frac{e^{x} - e^{-x}}{2}.$$

In fact, $\cosh(1) \approx 1.5431$ and $\sinh(1) \approx 1.1752$, and we see that all

$$\exp(A+B), \quad \exp(A)\exp(B), \quad \exp(B)\exp(A)$$

are not equal. It is also interesting to compare Lemma 3.1.6 with the Lie product formula (Theorem 3.1.51) below.

The following are immediate consequences of Lemma 3.1.6.

COROLLARY 3.1.8. *Given any $A \in \mathbb{C}^{n\times n}$, one has*

$$\exp(\alpha A)\exp(\beta A) = \exp((\alpha+\beta)A) \quad \textit{for all } \alpha, \beta \in \mathbb{C}.$$

COROLLARY 3.1.9. *Given any $A \in \mathbb{C}^{n\times n}$, one has $\exp(A) \in \mathrm{GL}(n,\mathbb{C})$ with*

$$(\exp(A))^{-1} = \exp(-A).$$

It is worth to mention the following theorem, despite we do not use it in this course.

THEOREM 3.1.10 ([**Hal15**, Theorem 2.10]). *Each $A \in \mathrm{GL}(n,\mathbb{C})$ can be expressed as $\exp(B)$ for some $B \in \mathbb{C}^{n\times n}$. In other words, the mapping $B \in \mathbb{C}^{n\times n} \to \exp(B) \in \mathrm{GL}(n,\mathbb{C})$ is surjective.*

We now verify (3.1.2) solves (3.1.1):

THEOREM 3.1.11. *Let $A \in \mathbb{C}^{n\times n}$. Then $t \mapsto \exp(tA)$ is a smooth curve in $\mathbb{C}^{n\times n}$ and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA) = A\exp(tA) = \exp(tA)A.$$

PROOF. It is well-known that one can differentiate a power series term by term within its radius of convergence (see e.g. [**Pug15**, Theorem 12 in Chapter 4]). In view of Remark 3.1.5, our theorem immediate follows by differentiating the power series $\exp(tA)$. □

If we take A with $\|A\| = 1$, we see that Theorem 3.1.11 is nothing by just a directional derivative. For each fixed $1 \le j_0, k_0 \le n$, if we choose

$$A = \begin{cases} 1 & , j = j_0 \text{ and } k = k_0, \\ 0 & \text{otherwise}, \end{cases}$$

then the derivative in Theorem 3.1.11 is simply a partial derivative. In fact, the matrix exponential map is (total) differentiable:

THEOREM 3.1.12 ([**Hal15**, Theorem 2.16]). *The matrix exponential map $\exp : \mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2} \to \mathrm{GL}(n,\mathbb{C})$ is an infinitely differentiable map.*

PROOF. Fix any $A \in \mathbb{C}^{n\times n}$. Note that for each j and k, the quantity $(A^m)_{jk}$ is a homogeneous polynomial of degree m in the entries of A. Thus, the series for the function $(A^m)_{jk}$ has the form of a multivariable power series on $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$. Since the series converges on all $\mathbb{R}^{2n^2}$ (more precisely, the radius of convergence $= \infty$), it is permissible to differentiate the series term by term as many times as we wish (see e.g. [**Pug15**, Theorem 12 in Chapter 4]). □

3.1.1. Computations of the exponential. For $A \in \mathbb{C}^{n\times n}$, we denote $A^{(jk)} \in \mathbb{C}^{(n-1)\times(n-1)}$ be the matrix obtained from A by crossing out j^{th} row and k^{th} column. We now define the determinant by induction.

DEFINITION 3.1.13. For $A \in \mathbb{C}^{1\times 1} \cong \mathbb{C}$, we simply define $\det(A) := A$. For each $A \in \mathbb{C}^{n\times n}$, we define

$$\det(A) := \sum_{k=1}^{n} (-1)^{1+k} A_{1k} \det(A^{(1k)}).$$

THEOREM 3.1.14 (Cofactor expansion [Tre17, Theorem 5.1]). *For each $A \in \mathbb{C}^{n\times n}$, for each fixed $j = 1, \cdots, n$, we have*

$$\det(A) = \sum_{k=1}^{n} (-1)^{j+k} A_{jk} \det(A^{(jk)}),$$

that is, the determinant is independent of j. In addition, we have

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+k} A_{jk} \det(A^{(jk)}),$$

that is, the determinant is independent of k.

Let S_n is the set of permutation on the indices $\{1, \cdots, n\}$, that is,

$$S_n = \{\text{bijective function } \sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}\}.$$

THEOREM 3.1.15 ([Tre17, (4.2)]). *The* determinant *of $A \in \mathbb{C}^{n\times n}$ can be computed by*

$$\det(A) = \sum_{\sigma \in S_n} c_\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n},$$

for some constant $c_\sigma \in \{-1, 1\}$.

REMARK 3.1.16. For those who familiar with abstract algebra, here we also remark that c_σ is exact the sign of the permutation $\sigma \in S_n$, denoted by $\operatorname{sign}(\sigma)$. In addition, it is also related to determinant via the formula

$$\operatorname{sign}(\sigma) = \det \begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \cdots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix},$$

where $\boldsymbol{e}_j$ is the j^{th} column of I.

It is important to mention the following properties:

THEOREM 3.1.17 ([Tre17, Theorem 3.4 and Theorem 3.5]). *Given any $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$, one has*

$$\det(A) = \det(A^\intercal), \quad \det(AB) = \det(A)\det(B).$$

A matrix $A \in \mathbb{C}^{n\times n}$ is said to be *diagonal* if $a_{jk} = 0$ for all $j \neq k$. We denote $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ the diagonal matrix with entries $\lambda_1, \cdots, \lambda_n$ on the main diagonal. A matrix $A \in \mathbb{C}^{n\times n}$ is said to be diagonalizable if there exists a matrix $P \in \mathrm{GL}(n, \mathbb{C})$ such that $P^{-1}AP$ is diagonal. If we write

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \cdots, \lambda_n),$$

then by using (3.1.5) we can easily compute its exponential of a diagonalizable matrix A as

$$\begin{aligned}\exp(A) &= \exp(P\mathrm{diag}\,(\lambda_1,\cdots,\lambda_n)P^{-1})\\ &= P\exp(\mathrm{diag}\,(\lambda_1,\cdots,\lambda_n))P^{-1} = P\mathrm{diag}\,(e^{\lambda_1},\cdots,e^{\lambda_n}))P^{-1}.\end{aligned} \tag{3.1.7}$$

A nontrivial vector $\boldsymbol{p}$ is said to be an eigenvector of A with eigenvalue λ if $A\boldsymbol{p} = \lambda\boldsymbol{p}$. Since

$$\lambda \text{ is an eigenvalue of } A \iff \det(A-\lambda I) = 0,$$

this suggests us to consider the characteristic polynomial

$$p(z) = \det(zI - A) = z^n + \sum_{j=0}^{n-1} c_j z^j$$

for some $c_0,\cdots,c_j \in \mathbb{C}$. By using the fundamemtal theorem of algebra (see e.g. [**Kow23**]), there has exactly n complex roots. We also define

$$p(B) := B^n + \sum_{j=0}^{n-1} c_j B^j \quad \text{for all } B \in \mathbb{C}^{n\times n}. \tag{3.1.8}$$

It is important to mention the following theorem regarding the characteristic polynomial (3.1.8):

THEOREM 3.1.18 (Cayley-Hamilton). *If $A \in \mathbb{C}^{n\times n}$, then $p(A) = 0$.*

A complex number λ is called a root of p if $p(\lambda) = 0$. The multiplicity of this root is called the *algebraic multiplicity* of the eigenvalue λ. There is another notion of multiplicity of an eigenvalue: the dimension of the *eigenspace*

$$\ker(\lambda I - A) := \{\boldsymbol{p} \in \mathbb{R}^n : (\lambda I - A)\boldsymbol{p} = \boldsymbol{0}\}$$

is called the *geometric multiplicity* of the eigenvalue λ.

THEOREM 3.1.19 ([**Tre17**, Theorem 2.8]). *Let $A \in \mathbb{C}^{n\times n}$. Then $A \in \mathrm{GL}(n,\mathbb{C})$ if and only if for each eigenvalue λ the dimension of the eigenspace $\ker(A-\lambda I)$ coincides with its algebraic multiplicity.*

EXAMPLE 3.1.20 (A nondiagonalizable matrix). We consider the matrix

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

Its characteristic polynomial is

$$p(z) = \det(zI - A) = \det\begin{pmatrix} z-1 & -1 \\ 0 & z-z \end{pmatrix} = (z-1)^2,$$

so A has an eigenvalue 1 of algebraic multiplicity 2. However, one sees that

$$\dim\ker(I-A) = \dim\ker\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = 1,$$

which shows that A is not diagonalizable.

A set of vectors $\{\boldsymbol{p}_1,\cdots,\boldsymbol{p}_k\}$ is said to be linear independent if

$$\sum_{i=1}^{k} c_i\boldsymbol{p}_i = 0 \implies c_i = 0 \text{ for all } i = 1,\cdots,k.$$

LEMMA 3.1.21 ([**HS99**, Lemma IV-1-2]). *A matrix $A \in \mathbb{C}^{n\times n}$ is diagonalizable if and only if A has n linearly independent eigenvectors $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$.*

From Lemma 3.1.21, we immediately obtain the following corollary.

COROLLARY 3.1.22. *If $A \in \mathbb{C}^{n\times n}$ has n distinct eigenvalues, then $A \in \mathrm{GL}(n, \mathbb{C})$.*

EXERCISE 3.1.23. Show that the set of diagonalizable $n \times n$ matrix is a proper subset (i.e. a subset which is not equal) of $\mathbb{C}^{n\times n}$ for all $n \geq 2$. (**Hint.** modify the ideas in Example 3.1.20).

LEMMA 3.1.24. *The set of diagonalizable $n \times n$ matrix is dense in $\mathbb{C}^{n\times n}$ (in the sense of Exercise 3.1.3). In other words, given any $A \in \mathbb{C}^{n\times n}$, there exists a sequence of diagonalizable matrix B_k which is converges to A.*

PROOF. By using Lemma 3.1.30, there exists $P \in \mathrm{GL}(n, \mathbb{C})$ such that $B = P^{-1}AP$ is upper triangular. If we can show that there exists a sequence of diagonalizable matrix $\tilde{B}_k$ which is converges to B, then from Exercise 3.1.1 and Exercise 3.1.3 we have

$$\begin{aligned}\limsup_{k\to+\infty} \|P\tilde{B}_k P^{-1} - A\| &= \limsup_{k\to+\infty} \|P(\tilde{B}_k - B)P^{-1}\| \\ &\leq \|P\|\|P^{-1}\| \lim_{k\to+\infty} \|\tilde{B}_k - B\| = 0,\end{aligned}$$

which concludes the lemma with $\tilde{B}_k = P\tilde{B}_k P^{-1}$. This can be done by setting

$$\tilde{B}_k := B + \mathrm{diag}(\epsilon_{k,1}, \cdots, \epsilon_{k,n})$$

where the quantities $\epsilon_{k,j}$ are chosen in such a way that n numbers $b_{11} + \epsilon_{k,1}, \cdots, b_{nn} + \epsilon_{k,n}$ are distinct and $\epsilon_{k,j} \to 0$ for all $j = 1, \cdots, n$, because by Corollary 3.1.22 we see that $\tilde{B}_k$ are diagonalizable matrices. □

EXERCISE 3.1.25. For each $A \in \mathbb{C}^{n\times n}$, show that $\det(\exp(A)) = e^{\mathrm{tr}\,(A)}$. Use to show that $\mathrm{tr}\,(A) = \lambda_1 + \cdots + \lambda_n$, where $\lambda_j \in \mathbb{C}$ are eigenvalues (may identical) of A.

However, in practical, it is not easy to check whether a matrix A is diagonalizable or not. There are some sufficient conditions which are relatively easy to check.

DEFINITION 3.1.26 ([**Tre17**, Section 6]). A matrix $U \in \mathbb{C}^{n\times n}$ is called *unitary* if $U^*U = I$. The set of unitary matrix is defined by $\mathrm{U}(n, \mathbb{C})$, which is also called the *unitary group*.

LEMMA 3.1.27. *If $U \in \mathbb{C}^{n\times n}$ is unitary, then $U \in \mathrm{GL}(n, \mathbb{C})$ with $U^{-1} = U^*$, which is also unitary. In other words, $\mathrm{U}(n, \mathbb{C}) \subset \mathrm{GL}(n, \mathbb{C})$.*

A matrix $A \in \mathbb{C}^{n\times n}$ is called *normal* if $A^*A = AA^*$. A matrix $A \in \mathbb{C}^{n\times n}$ is called *Hermitian* (or *self-adjoint*) if $A = A^*$. We write

$$(\boldsymbol{u}, \boldsymbol{v})_{\mathbb{C}^{n\times n}} := \boldsymbol{v}^*\boldsymbol{u} \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n.$$

It is important to notice that

$$\overline{(\boldsymbol{v}, \boldsymbol{u})_{\mathbb{C}^{n\times n}}} = ((\boldsymbol{v}, \boldsymbol{u})_{\mathbb{C}^{n\times n}})^* = (\boldsymbol{u}^*\boldsymbol{v})^* = \boldsymbol{v}^*\boldsymbol{u} = (\boldsymbol{u}, \boldsymbol{v})_{\mathbb{C}^{n\times n}} \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n. \tag{3.1.9}$$

LEMMA 3.1.28. *A is Hermitian if and only if $(A\boldsymbol{u}, \boldsymbol{v})_{\mathbb{C}^{n\times n}} = (\boldsymbol{u}, A\boldsymbol{v})_{\mathbb{C}^{n\times n}}$ for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$.*

PROOF. The lemma easily hold from the identities

$$(\boldsymbol{u}, A\boldsymbol{v})_{\mathbb{C}^{n\times n}} = (A\boldsymbol{v})^*\boldsymbol{u} = \boldsymbol{v}^*A^*\boldsymbol{u}$$

and

$$(A\boldsymbol{u}, \boldsymbol{v})_{\mathbb{C}^{n\times n}} = \boldsymbol{v}^*A\boldsymbol{u}$$

for all $A \in \mathbb{C}^{n\times n}$ as well as for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$. □

In fact, all normal matrices are unitary diagonalizable:

THEOREM 3.1.29 ([**Tre17**, Theorem 2.4]). *If $A \in \mathbb{C}^{n\times n}$ is normal, then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $D := U^*AU$ is diagonal. If $A \in \mathbb{C}^{n\times n}$ is Hermitian, then D is real-valued.*

We still can simplify arbitrary matrix by using unitary matrices. A matrix $A \in \mathbb{C}^{n\times n}$ is said to be upper triangular if $a_{jk} = 0$ for $j > k$.

LEMMA 3.1.30 (Schur representation [**Tre17**, Theorem 1.1 and Theorem 1.2]). *For each $A \in \mathbb{C}^{n\times n}$, there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular. If $A \in \mathbb{R}^{n\times n}$ and all its eigenvalues are real, then we can choose $U \in \mathbb{R}^{n\times n}$.*

In view of the power series of $\exp(A)$, it is also natural to study the following class of matrix.

DEFINITION 3.1.31. A matrix $N \in \mathbb{C}^{n\times n}$ is said to be *nilpotent* if $N^n = 0$.

REMARK 3.1.32. If N is nilpotent, then $\exp(N)$ is simply a finite sum. One can directly verify that $\exp(N)$ is unipotent (i.e. $\exp(N) - I$ is nilpotent)

LEMMA 3.1.33. *A matrix $N \in \mathbb{C}^{n\times n}$ is nilpotent if and only if all eigenvalues of N are zero.*

PROOF. Let λ be an eigenvalue of a nilpotent matrix N with eigenfunction $\boldsymbol{p} \not\equiv 0$, i.e.

$$N\boldsymbol{p} = \lambda\boldsymbol{p}.$$

Then $0 = N^n\boldsymbol{p} = \lambda^n\boldsymbol{p}$, which implies $\lambda^n = 0$. Since $|\lambda|^n = |\lambda^n| = 0$, then we conclude that $\lambda = 0$. The converse can be easily verified as well. □

By using Schur's representation (Lemma 3.1.30) and Lemma 3.1.33, we reach the following lemma.

LEMMA 3.1.34. *A matrix $N \in \mathbb{C}^{n\times n}$ (resp. $N \in \mathbb{R}^{n\times n}$) is nilpotent if and only if there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ (resp. $U \in \mathbb{R}^{n\times n}$) such that $T = U^*NU$ is upper triangle with $T_{jj} = 0$ for all $j = 1, \cdots, n$.*

We already discuss a relation between the set of diagonalizable matrix and $\mathbb{C}^{n\times n}$ in Lemma 3.1.24. The following theorem gives another relation between them.

THEOREM 3.1.35 (Jordan-Chevalley decomposition [**HS99**, Theorem IV-1-11]). *Let $A \in \mathbb{C}^{n\times n}$. Then there exist a diagonalizable matrix $D \in \mathbb{C}^{n\times n}$ and a nilpotent matrix $N \in \mathbb{C}^{n\times n}$ such that*

$$A = D + N \quad \text{and} \quad DN = ND, \tag{3.1.10}$$

and the decomposition (3.1.10) is unique. If $A \in \mathbb{R}^{n\times n}$, then $D \in \mathbb{R}^{n\times n}$ and $N \in \mathbb{C}^{n\times n}$.

Since the trace operator is linear and $\operatorname{tr}(N) = 0$ (see Lemma 3.1.33), we immediately reach the following corollary.

COROLLARY 3.1.36. *If $A \in \mathbb{C}^{n\times n}$ satisfies $\operatorname{tr}(A) = 0$, then the diagonalizable matrix D in (3.1.10) satisfies $\operatorname{tr}(D) = 0$.*

EXERCISE 3.1.37. Let

$$A_1 = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

Compute $\exp(A_1)$, $\exp(A_2)$ and $\exp(A_3)$.

EXERCISE 3.1.38. Show that for any $a, b, d \in \mathbb{C}$ that

$$\exp\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} e^a & b\frac{e^a - e^d}{a-d} \\ 0 & e^d \end{pmatrix}.$$

Since

$$\lim_{a\to d} \frac{e^a - e^d}{a-d} = e^a,$$

we simply interpret $\frac{e^a-e^d}{a-d}$ as e^a when $d = a$. (**Hint.** Show that

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}^m = \begin{pmatrix} a^m & b\frac{a^m - d^m}{a-d} \\ 0 & b^m \end{pmatrix}$$

for all $m \in \mathbb{N}$ and $a \neq d$.)

We now exhibit the general algorithm to compute the decomposition in Theorem 3.1.35.

Algorithm 2 Computation of D and N in Theorem 3.1.35

1: Input a matrix $A \in \mathbb{C}^{n\times n}$.
2: Compute the characteristic polynomial $p(z) = (zI - A)$.
3: Decompose $1/p(z)$ into partial fractions

$$\frac{1}{p(z)} = \sum_{j=1}^{k} \frac{Q_j(z)}{(z-\lambda_j)^{m_j}},$$

where for each j the quantity $Q_j(z)$ is a nonzero polynomial with $\deg(Q_j) \leq m_j - 1$ and $\lambda_1, \cdots, \lambda_k$ are distinct zeros of p (i.e. eigenvalues of A).
4: For each $j = 1, \cdots, n$, we define

$$P_j(A) := Q_j(A) \prod_{\ell \neq j} (A - \lambda_j I)^{m_\ell}.$$

5: Output $D = \lambda_1 P_1(A) + \cdots + \lambda_k P_k(A)$ and $N = A - D$.

REMARK 3.1.39. There always exists a unique decomposition in Step 3 of Algorithm 2, see e.g. [**Kow23**] for a proof. This algorithm highlighted the proof of Theorem 3.1.35. Here we also highlight that $P_j(A)P_k(A) = 0$ for all $j \neq k$ and $P_j(A)^2 = P_j(A)$ for all $j = 1, \cdots, n$, see [**HS99**, Lemma IV-1-9].

Combining Lemma 3.1.6 and Theorem 3.1.35, we reach

$$\begin{aligned}\exp(A) &= \exp(D)\exp(N) = \exp(\lambda_1 P_1(A))\cdots\exp(\lambda_k P_k(A))\exp(N)\\ &= \exp(N)\exp(D) = \exp(N)\exp(\lambda_1 P_1(A))\cdots\exp(\lambda_k P_k(A)).\end{aligned}$$

EXAMPLE 3.1.40 ([**HS99**, Example IV-1-18]). The characteristic polynomial of the matrix

$$A = \begin{pmatrix} 252 & 498 & 4134 & 698 \\ -234 & -465 & -3885 & -656 \\ 15 & 30 & 252 & 42 \\ -10 & -20 & -166 & -25 \end{pmatrix}$$

is $p(z) = (z-4)^2(z-3)^2$. One can compute

$$\begin{aligned}\frac{1}{p(z)} &= \frac{1}{(z-4)^2} - \frac{2}{z-4} + \frac{1}{(z-3)^2} + \frac{2}{z-3}\\ &= \frac{1-2(z-4)}{(z-4)^2} + \frac{1+2(z-3)}{(z-3)^2}\end{aligned}$$

Accordingly, we set

$$\begin{aligned}P_1(z) &:= (1-2(z-4))(z-3)^2, \quad \lambda_1 = 4,\\ P_2(z) &:= (1+2(z-3))(z-4)^2, \quad \lambda_2 = 3,\end{aligned}$$

and we compute

$$P_1(A) = \begin{pmatrix} -1 & -2 & 134 & 198 \\ 1 & 2 & -125 & -186 \\ 0 & 0 & 9 & 12 \\ 0 & 0 & -6 & -8 \end{pmatrix}, \quad P_2(A) = \begin{pmatrix} 2 & 2 & -134 & -198 \\ -1 & -1 & 125 & 186 \\ 0 & 0 & -8 & -12 \\ 0 & 0 & 6 & 9 \end{pmatrix}.$$

Therefore

$$S = \lambda_1 P_1(A) + \lambda_2 P_2(A) = \begin{pmatrix} 2 & -2 & 134 & 198 \\ 1 & 5 & -125 & -186 \\ 0 & 0 & 12 & 12 \\ 0 & 0 & -6 & -5 \end{pmatrix}$$

and

$$N = A - S = \begin{pmatrix} 250 & 500 & 4000 & 500 \\ -235 & -470 & -3760 & = 470 \\ 15 & 30 & 240 & 30 \\ -10 & -20 & -160 & -20 \end{pmatrix}.$$

EXERCISE 3.1.41. Decompose $A = \begin{pmatrix} 3 & 4 & 3 \\ 2 & 7 & 4 \\ -4 & 8 & 3 \end{pmatrix}$ by using Algorithm 2.

EXERCISE 3.1.42 (Bochner's subordination). Let $0 < s < 1$, by using the integration by parts on $\Gamma(1-s)$, where Γ is the gamma function, show that

$$\lambda^s = \frac{1}{|\Gamma(-s)|}\int_0^\infty (1-e^{t\lambda})t^{-1-s}\,\mathrm{d}t \quad \text{for all } \lambda > 0.$$

DEFINITION 3.1.43. A Hermitian matrix $A \in \mathbb{C}^{n\times n}$ is said to *positive definite*, denoted by $A \succ 0$, if

$$\boldsymbol{p}^* A \boldsymbol{p} > 0 \quad \text{for all } \boldsymbol{p} \in \mathbb{C}^n \setminus \{\boldsymbol{0}\}.$$

By using Theorem 3.1.29, we see that $A \succ 0$ if and only if

$$A = U \operatorname{diag}(\lambda_1, \cdots, \lambda_n) U^*$$

for some $\lambda_1 > 0, \cdots, \lambda_n > 0$ and unitary $U \in \mathbb{C}^{n\times n}$, and accordingly we define

$$A^s := U \operatorname{diag}(\lambda_1^s, \cdots, \lambda_n^s) U^*.$$

Now using the Bochner's subordination (Exercise 3.1.42) and (3.1.7), we can compute A^s via the formula

$$A^s = \frac{1}{|\Gamma(-s)|} \int_0^\infty (1 - \exp(tA)) t^{-1-s} \, \mathrm{d}t \quad \text{for all } A \succ 0, \tag{3.1.11}$$

which gives an application of the matrix fundamental solution (3.1.2).

REMARK 3.1.44. The Fourier transform suggests us to (formally) replace A by $-\Delta$ in (3.1.11), and we reach the fractional Laplacian

$$(-\Delta)^s := \frac{1}{|\Gamma(-s)|} \int_0^\infty (1 - e^{-t\Delta}) t^{-1-s} \, \mathrm{d}t.$$

One can see e.g. [**Kwa17**] for an introduction.

3.1.2. The matrix logarithm. We now wish to define a matrix logarithm, which should be an inverse function to the matrix exponential. One simplest way to define the matrix logarithm is by a power series. We recall the following fact concerning the principal branch of complex logarithm (see e.g. [**Kow23**] for more details):

LEMMA 3.1.45. *The radius of convergence of the complex power series*

$$\log z := \sum_{m=1}^\infty (-1)^{m+1} \frac{(z-1)^m}{m} \tag{3.1.12}$$

is 1, in other words, the series (3.1.12) is defined and holomorphic in a circle of radius 1 about $z = 1$. In addition, we have

$$e^{\log z} = z \quad \textit{for all } z \textit{ with } |z - 1| < 1.$$

Moreover, we have

$$|e^u - 1| < 1 \quad \textit{and} \quad \log e^u = u \quad \textit{for all } u \textit{ with } |u| < \log 2.$$

Based on the above lemma, we now can define the matrix logarithm by the following theorem.

THEOREM 3.1.46. [**Hal15**, Theorem 2.8] *The matrix logarithm*

$$\log(A) := \sum_{m=1}^\infty (-1)^{m+1} \frac{(A - I)^m}{m}$$

is defined and continuous on the set of all matrices $A \in \mathbb{C}^{n\times n}$ with $\|A - I\| < 1$. In addition, we have

$$\exp(\log(A)) = A$$

for all matrices $A \in \mathbb{C}^{n\times n}$ *with* $\|A - I\| < 1$. *Moreover, we have* $\|\exp(B) - 1\| < 1$ *and*

$$\log(\exp(B)) = B$$

for all matrices $B \in \mathbb{C}^{n\times n}$ *with* $\|B\| < \log 2$.

PROOF. By using Exercise 3.1.1, we have $\|(A-I)^m\| \leq \|A-I\|^m$, by using the similar arguments in Lemma 3.1.4, one can show that $\log(A)$ is defined and continuous on the set of all matrices $A \in \mathbb{C}^{n\times n}$ with $\|A - I\| < 1$. We left the details for readers as an exercise.

Let $A \in \mathbb{C}^{n\times n}$ with $\|A - I\| < 1$. By using Lemma 3.1.24, one can find a sequence of diagonalizable matrix $A_k \in \mathbb{C}^{n\times n}$ such that $\|A_k - A\| \to 0$ as $k \to \infty$. Since $\|A_k - I\| < 1$ for all sufficiently large k, then we know that

$$\log(A_k) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A_k - I)^m}{m}$$

is defined and continuous for all sufficiently large k. We write

$$A_k = Q_k \mathrm{diag}(\lambda_{k,1}, \cdots, \lambda_{k,n}) Q_k^{-1},$$

and we see that

$$(A_k - I)^m = Q_k \mathrm{diag}\left((\lambda_{k,1} - 1)^m, \cdots, (\lambda_{k,n} - 1)^m\right) Q_k^{-1},$$

thus

$$\begin{aligned}\log(A_k) &= Q_k \left(\sum_{m=1}^{\infty} (-1)^{m+1} \frac{\mathrm{diag}\left((\lambda_{k,1} - 1)^m, \cdots, (\lambda_{k,n} - 1)^m\right)}{m} \right) Q_k^{-1} \\ &= Q_k \mathrm{diag}\left(\log(\lambda_{k,1}), \cdots, \log(\lambda_{k,n})\right) Q_k^{-1}.\end{aligned}$$

Now by using (3.1.5) and Lemma 3.1.45 we see that

$$\begin{aligned}\exp(\log(A_k)) &= Q_k \exp \mathrm{diag}\left(\log(\lambda_{k,1}), \cdots, \log(\lambda_{k,n})\right) Q_k^{-1} \\ &= Q_k \mathrm{diag}\left(\exp\log(\lambda_{k,1}), \cdots, \exp\log(\lambda_{k,n})\right) Q_k^{-1} \\ &= Q_k \mathrm{diag}(\lambda_{k,1}, \cdots, \lambda_{k,n}) Q_k^{-1} = A_k.\end{aligned}$$

Finally, by continuity of the mapping $A \mapsto \log(A)$ and $B \mapsto \exp(B)$, we conclude

$$\exp(\log(A)) = A$$

by taking $k \to +\infty$.

Now, if $\|B\| < \log 2$, then using Exercise 3.1.1 we see that

$$\|\exp(B) - I\| = \left\| \sum_{m=1}^{\infty} \frac{B^m}{m!} \right\| \leq \sum_{m=1}^{\infty} \frac{\|B\|^m}{m!} = e^{\|B\|} - 1 < 1,$$

thus $\log(\exp(B))$ is well-defined. The proof of $\log(\exp(B)) = B$ is very similar to the proof of $\exp(\log(A)) = A$, therefore we left the details for readers as an exercise. □

REMARK 3.1.47. If A is unipotent (i.e. $A - I$ is nilpotent), then $\log(A)$ is simply a finite sum, which can be defined without the assumption $\|A - I\| < 1$. In this case, one can easily verify that $\log(A)$ is nilpotent. See also Remark 3.1.32.

EXERCISE 3.1.48. Show that:

(a) If A is unipotent, then $\exp(\log(A)) = A$.
(b) If B is nilpotent, then $\log(\exp(B)) = B$.

(**Hint.** Let $A(t) := I + t(A - I)$ and show that $\exp(\log(A(t)))$ depends polynomially on t and that $\exp(\log(A(t))) = A(t)$ for all sufficiently small t)

EXERCISE 3.1.49. Show that there exists a constant $c > 0$ such that

$$\| \log(I + X) - X\| \leq c\|X\|^2 \tag{3.1.13}$$

holds true for all $X \in \mathbb{C}^{n\times n}$ with $\|X\| \leq 1/2$.

REMARK 3.1.50. We may restate (3.1.13) by saying that

$$\log(I + X) = X + O(\|X\|^2),$$

where $O(\|X\|^2)$ denotes a quantity of order $\|X\|^2$, i.e. a quantity that is bounded by a constant times $\|X\|^2$ for all sufficiently small values of $\|X\|$.

3.1.3. One parameter subgroup, Lie group and Lie algebra. We now exhibit a result involving the exponential of a matrix that will be important in the study of Lie algbras.

THEOREM 3.1.51 (Lie product formula [**Hal15**, Theorem 2.11]). *For each $A, B \in \mathbb{C}^{n\times n}$, we have*

$$\exp(A + B) = \lim_{m\to\infty} (\exp(A/m) \exp(B/m))^m.$$

PROOF. By multiplying the power series for $\exp(A/m)$ and $\exp(B/m)$, one sees that

$$\exp(A/m) \exp(B/m) = I + \frac{A}{m} + \frac{B}{m} + O(m^{-2}).$$

Now since $\exp(A/m) \exp(B/m) \to I$ as $m \to \infty$, then $\log(\exp(A/m) \exp(B/m))$ is well-defined for all sufficiently large m. By using Exercise 3.1.49 (with $X = \frac{A}{m} + \frac{B}{m} + O(m^{-2})$), we see that

$$\begin{aligned}\log\left(\exp(A/m) \exp(B/m)\right) &= \log\left(I + \frac{A}{m} + \frac{B}{m} + O(m^{-2})\right)\\ &= \frac{A}{m} + \frac{B}{m} + O(m^{-2}) + O\left(\left\|\frac{A}{m} + \frac{B}{m} + O(m^{-2})\right\|^2\right)\\ &= \frac{A}{m} + \frac{B}{m} + O(m^{-2}).\end{aligned}$$

Now Theorem 3.1.46 guarantees that

$$\exp(A/m) \exp(B/m) = \exp\left(\frac{A}{m} + \frac{B}{m} + O(m^{-2})\right),$$

therefore,

$$(\exp(A/m) \exp(B/m))^m = \exp\left(A + B + O(m^{-1})\right).$$

By the continuity of the exponential, we conclude our theorem by taking $m \to +\infty$. □

REMARK 3.1.52. There is a version of this result, known as the Trotter product formula, which holds for suitable unbounded operators on an infinite-dimensional Hilbert space, see e.g. [**Hal13**, Theorem 20.1].

DEFINITION 3.1.53. We call $\{Y(t)\}_{t\in\mathbb{R}}$ a *one parameter subgroup of* $\mathrm{GL}(n, \mathbb{C})$ if

(a) $Y : \mathbb{R} \to \mathrm{GL}(n, \mathbb{C})$ is continuous;
(b) $Y(0) = I$; and

(c) $Y(t+s) = Y(t)A(s)$ for all $t, s \in \mathbb{R}$.

EXAMPLE 3.1.54. The fundamental matrix solution $\{\exp(tA)\}_{t\in\mathbb{R}}$ given in (3.1.2) forms a one parameter subgroup, where (a) is verified by Theorem 3.1.11, (b) can be found in the basic properties (3.1.5), and (c) is a special case of Corollary 3.1.8.

We now proof the following lemma.

LEMMA 3.1.55. *Let* $0 < \epsilon < \log 2$, *let* $B_{\epsilon/2} := \{A \in \mathbb{C}^{n\times n} : \|A\| < \epsilon/2\}$ *and let*

$$\exp(B_{\epsilon/2}) := \{\exp(A) \in \mathbb{C}^{n\times n} : \|A\| < \epsilon/2\}.$$

Given any $A \in \exp(B_{\epsilon/2})$, *there exists a unique* $\sqrt{A} \in \exp(B_{\epsilon/2})$ *such that* $\sqrt{A}^2 = A$, *which is given by*

$$\sqrt{A} = \exp\left(\frac{1}{2}\log(A)\right).$$

PROOF. By using Lemma 3.1.6, one sees that

$$\sqrt{A}^2 = \exp\left(\frac{1}{2}\log(A) + \frac{1}{2}\log(A)\right) = \exp(\log(A)) = A.$$

To establish uniqueness, suppose $B \in \exp(B_{\epsilon/2})$ satisfies $B^2 = A$. Now using Lemma 3.1.6 we see that

$$\begin{aligned}\exp(2\log(B)) &= \exp(\log(B) + \log(B)) \\ &= \exp(\log B)\exp(\log B) = B^2 = A\end{aligned}$$

Since $B \in \exp(B_{\epsilon/2})$, then we can check that $\|2\log(B)\| < \epsilon < \log 2$. Now we can use Theorem 3.1.46 to see that

$$\log(A) = \log(\exp(2\log(B))) = 2\log(B),$$

thus $\log(B) = \frac{1}{2}\log(A)$. Finally, again using Theorem 3.1.46 we conclude that

$$B = \exp(\log(B)) = \exp\left(\frac{1}{2}\log(A)\right) = \sqrt{A},$$

which conclude the uniqueness. □

We now show that Example 3.1.54 already exhibit all one parameter subgroup.

THEOREM 3.1.56 ([**Hal15**, Theorem 2.14]). *If* $\{Y(t)\}_{t\in\mathbb{R}}$ *is a one parameter subgroup of* $\mathrm{GL}(n,\mathbb{C})$, *then there exists a unique* $A \in \mathbb{C}^{n\times n}$ *such that*

$$Y(t) = e^{tA} \quad \textit{for all } t \in \mathbb{R}.$$

PROOF. We first prove the uniqueness of such A. Suppose that $e^{tA} = e^{tB}$ for all $t \in \mathbb{R}$, by using Theorem 3.1.11, we can differentiate the power series of matrix exponential term-by-term and see that

$$A = \frac{\mathrm{d}}{\mathrm{d}t}e^{tA}\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}e^{tB}\bigg|_{t=0} = B.$$

Since the function $\log : \exp(B_{\epsilon/2}) \to B_{\epsilon/2}$ is bijective and continuous, then we see that $\exp(B_{\epsilon/2})$ is an open set in $\mathrm{GL}(n,\mathbb{C})$. Since $Y(0) = I \in \exp(B_{\epsilon/2})$ and $t \mapsto Y(t)$ is continuous, then there exists $t_0 > 0$ such that

$$Y(t) \in \exp(B_{\epsilon/2}) \quad \text{for all } t \in [-t_0, t_0].$$

Now we define

$$A := \frac{1}{t_0}\log(Y(t_0)),$$

so that $t_0 A = \log(Y(t_0))$. Now we see that $t_0 A \in B_{\epsilon/2}$ and thus Theorem 3.1.46 allows us to apply matrix exponential on the identity $t_0 A = \log(Y(t_0))$ to see that

$$\exp(t_0 A) = \exp(\log(Y(t_0))) = Y(t_0).$$

Now $Y(t_0/2)$ is again in $\exp(B_{\epsilon/2})$ and by Definition 3.1.53(c) we have $Y(t_0/2)^2 = Y(t_0)$. By using Lemma 3.1.55 we see that

$$Y(t_0/2) = \sqrt{Y(t_0)} = \exp\left(\frac{1}{2}\log(Y(t_0))\right) = \exp(t_0 A/2)$$

Applying this argument repeatedly, we conclude that

$$Y(t_0/2^k) = \exp(t_0 A/2^k) \quad \text{for all } k \in \mathbb{N}.$$

Now by Definition 3.1.53(c) and Lemma 3.1.6 we see that

$$Y(mt_0/2^k) = Y(t_0/2^k)^m = \exp(mt_0 A/2^k) \quad \text{for all } k \in \mathbb{N} \text{ and } m \in \mathbb{Z},$$

in other words,

$$Y(t) = \exp(tA) \quad \text{for all } t \in \mathbb{R} \text{ of the form } t = \frac{m}{2^k}t_0.$$

Since the set $\{t \in \mathbb{R} : t = \frac{m}{2^k}t_0 \text{ for some } k \in \mathbb{N} \text{ and } m \in \mathbb{Z}\}$ is dense in $\mathbb{R}$, by continuity of $t \mapsto Y(t)$ and $t \mapsto \exp(tA)$, it follows that $Y(t) = \exp(tA)$ for all $t \in \mathbb{R}$, which conclude our theorem. □

Theorem 3.1.56 says that there is a one-to-one correspondence between $\mathbb{C}^{n\times n}$ and the collection of one parameter subgroups of $\mathrm{GL}(n,\mathbb{C})$. This also suggests us to extend Definition 3.1.53 as follows:

DEFINITION 3.1.57. Let $\mathcal{G}$ be a *matrix Lie group*, i.e. a subgroup of $\mathrm{GL}(n,\mathbb{C})$ with respect to matrix multiplication. In other words, $\mathcal{G}$ is a subset of $\mathrm{GL}(n,\mathbb{C})$ satisfying

- $AB \in \mathcal{G}$ for all $A \in \mathcal{G}$ and $B \in \mathcal{G}$;
- $I \in \mathcal{G}$; and
- $A^{-1} \in \mathcal{G}$ for all $A \in \mathcal{G}$.

We call $\{Y(t)\}_{t\in\mathbb{R}}$ a *one parameter subgroup of* $\mathcal{G}$ if

(a) $Y : \mathbb{R} \to G$ is continuous;
(b) $Y(0) = I$; and
(c) $Y(t+s) = Y(t)A(s)$ for all $t, s \in \mathbb{R}$.

EXAMPLE 3.1.58. The trivial subgroup $\{Y(t) = I\}_{t\in\mathbb{R}}$ is a one parameter subgroup of $\mathcal{G}$, which means the existence of one parameter subgroup holds.

We now able to give some examples. By using Lemma 3.1.27, it is easy to check that the unitary group $\mathrm{U}(n,\mathbb{C})$ is a matrix Lie group.

THEOREM 3.1.59. *If $\{Y(t)\}_{t\in\mathbb{R}}$ is a one parameter subgroup of $\mathrm{U}(n,\mathbb{C})$, then there exists a unique $A \in \mathbb{C}^{n\times n}$ which is skew-Hermitian (i.e. $A^* = -A$) such that*

$$Y(t) = \exp(tA) \quad \textit{for all } t \in \mathbb{R}.$$

PROOF. It is easy to check that

$$\begin{aligned}\exp(tA^*) = \exp((tA)^*) &= I + (tA)^* + \frac{((tA)^*)^2}{2!} + \cdots \\ &= \left(I + tA + \frac{(tA)^2}{2!} + \cdots\right)^* = (\exp(tA))^* \quad \text{for all } t \in \mathbb{R}.\end{aligned}$$

If $A \in \mathbb{C}^{n\times n}$ is skew-Hermition, then by Corollary 3.1.9 we see that

$$(\exp(tA))^* = \exp(tA^*) = \exp(-tA) = (\exp(tA))^{-1} \quad \text{for all } t \in \mathbb{R},$$

which shows that $\exp(tA) \in \mathrm{U}(n, \mathbb{C})$ for all $t \in \mathbb{R}$, that is, $\{\exp(tA)\}_{t\in\mathbb{R}}$ forms a one parameter subgroup of $\mathrm{U}(n, \mathbb{C})$. Now let $\{Y(t)\}_{t\in\mathbb{R}}$ be a one parameter subgroup of $\mathrm{U}(n, \mathbb{C})$. Since $\{Y(t)\}_{t\in\mathbb{R}}$ is also a one parameter subgroup of $\mathrm{GL}(n, \mathbb{C})$, then by Theorem 3.1.56 there exists a matrix $B \in \mathbb{C}^{n\times n}$ such that

$$Y(t) = \exp(tB) \quad \text{for all } t \in \mathbb{R}.$$

Since $\exp(tB) = Y(t) \in \mathrm{U}(n, \mathbb{C})$ for all $t \in \mathbb{R}$, then by Corollary 3.1.9 we see that

$$\exp(tB^*) = (\exp(tB))^* = (\exp(tB))^{-1} = \exp(-tB).$$

Now by Theorem 3.1.46 we see that

$$tB^* = \log(\exp(tB^*)) = \log(\exp(-tB)) = -tB \quad \text{for all } t \text{ with small } |t|,$$

which shows that B is skew-Hermitian. □

It is not difficult to see that the *special linear group* $\mathrm{SL}(n, \mathbb{C}) := \{A \in \mathbb{C}^{n\times n} : \det(A) = 1\}$ is a matrix Lie group.

THEOREM 3.1.60. *If $\{Y(t)\}_{t\in\mathbb{R}}$ is a one parameter subgroup of $\mathrm{SL}(n, \mathbb{C})$, then there exists a unique $A \in \mathbb{C}^{n\times n}$ with $\mathrm{tr}\,(A) = 0$ such that*

$$Y(t) = \exp(tA) \quad \textit{for all } t \in \mathbb{R}.$$

REMARK 3.1.61. See also Lemma 3.1.36. It is interesting to mention that, if $A \in \mathbb{C}^{n\times n}$ satisfies $\mathrm{tr}\,(A) = 0$, then there exist matrices $X \in \mathbb{C}^{n\times n}$ and $Y \in \mathbb{C}^{n\times n}$ so that $A = XY - YX$, where X is Hermitian and $\mathrm{tr}\,(Y) = 0$.

PROOF. Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{tr}\,(A) = 0$. By using Exercise 3.1.25, one can check that

$$\det(\exp(tA)) = e^{\mathrm{tr}\,(tA)} = e^0 = 1 \quad \text{for all } t \in \mathbb{R},$$

that is, $\{\exp(tA)\}_{t\in\mathbb{R}}$ forms a one parameter subgroup of $\mathrm{SL}(n, \mathbb{C})$. Now let $\{Y(t)\}_{t\in\mathbb{R}}$ be a one parameter subgroup of $\mathrm{SL}(n, \mathbb{C})$. By Theorem 3.1.56 there exists a matrix $A \in \mathbb{C}^{n\times n}$ such that

$$Y(t) = \exp(tA) \quad \text{for all } t \in \mathbb{R}.$$

Again by Exercise 3.1.25 we see that

$$1 = \det(Y(1)) = e^{\mathrm{tr}\,(A)},$$

thus $\mathrm{tr}\,(A) = \log(e^{\mathrm{tr}\,(A)}) = 0$ which conclude our theorem. □

One also can check that the special unitary group $\mathrm{SU}(n, \mathbb{C}) := \mathrm{U}(n, \mathbb{C}) \cap \mathrm{SL}(n, \mathbb{C})$ is also a matrix Lie group. Imitate the proof of Theorem 3.1.59 and Theorem 3.1.60, one can easily check the following corollary.

COROLLARY 3.1.62. *If $\{Y(t)\}_{t\in\mathbb{R}}$ is a one parameter subgroup of $\mathrm{SU}(n,\mathbb{C})$, then there exists a unique skew-Hermitian matrix $A\in\mathbb{C}^{n\times n}$ with $\mathrm{tr}\,(A)=0$ such that*

$$Y(t)=\exp(tA)\quad \textit{for all } t\in\mathbb{R}.$$

We now define the following sets:

$$\begin{aligned}\mathfrak{gl}(n,\mathbb{C})&:=\mathbb{C}^{n\times n},\\ \mathfrak{u}(n,\mathbb{C})&:=\{A\in\mathbb{C}^{n\times n}:A^*=-A\},\\ \mathfrak{sl}(n,\mathbb{C})&:=\{A\in\mathbb{C}^{n\times n}:\mathrm{tr}\,(A)=0\},\\ \mathfrak{su}(n,\mathbb{C})&:=\mathfrak{u}(n,\mathbb{C})\cap\mathfrak{sl}(n,\mathbb{C}).\end{aligned}$$

We point out that Theorem 3.1.56, Theorem 3.1.59, Theorem 3.1.60 and Corollary 3.1.62 say that one has the following one-to-one correspondence:

$$\begin{aligned}&\mathrm{GL}(n,\mathbb{C})\leftrightarrow\mathfrak{gl}(n,\mathbb{C}),\\ &\mathrm{U}(n,\mathbb{C})\leftrightarrow\mathfrak{u}(n,\mathbb{C}),\\ &\mathrm{SL}(n,\mathbb{C})\leftrightarrow\mathfrak{sl}(n,\mathbb{C}),\\ &\mathrm{SU}(n,\mathbb{C})\leftrightarrow\mathfrak{su}(n,\mathbb{C}).\end{aligned}$$

Now it is also natural to introduce the following terminologies:

DEFINITION 3.1.63. Let G be a matrix Lie group (see Definition 3.1.57). The *Lie algebra* of G, denoted $\mathfrak{g}$, is the set of all matrices X such that $e^{tX}\in G$ for all $t\in\mathbb{R}$.

We now summarize the above examples in the following table, see [**Hal15**] for more examples.

matrix Lie group G	Lie algebra $\mathfrak{g}$
$\mathrm{GL}(n,\mathbb{C})$	$\mathfrak{gl}(n,\mathbb{C})$
$\mathrm{U}(n,\mathbb{C})$	$\mathfrak{u}(n,\mathbb{C})$
$\mathrm{SL}(n,\mathbb{C})$	$\mathfrak{sl}(n,\mathbb{C})$
$\mathrm{SU}(n,\mathbb{C})$	$\mathfrak{su}(n,\mathbb{C})$

TABLE 1. Some examples of Lie algebra $\mathfrak{g}$ of matrix Lie group G

We define the *commutator* by

$$[A,B]:=AB-BA\quad \text{for all } A,B\in\mathbb{C}^{n\times n}.$$

Note that A and B are commute if and only if $[A,B]=0$. The following theorem exhibit some basic properties of Lie algebra.

THEOREM 3.1.64 ([**Hal15**, Theorem 3.20]). *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. The following holds:*

(a) *$A\mathfrak{g}A^{-1}\subset\mathfrak{g}$ for all $A\in G$.* [1]
(b) *$\mathfrak{g}$ is $\mathbb{R}$-linear, that is, $aA+bB\in\mathfrak{g}$ for all $A,B\in\mathfrak{g}$ and $a,b\in\mathbb{R}$.*
(c) *$[A,B]\in\mathfrak{g}$ for all $A,B\in\mathfrak{g}$.*

The proof of Theorem 3.1.64(a) is easy, which we left the details for readers as an exercise.

[1] In other words, for each $B\in\mathfrak{g}$, one has $ABA^{-1}\in\mathfrak{g}$ for all $A\in G$.

PROOF OF THEOREM 3.1.64(B). By definition, for each $A \in \mathfrak{g}$, it is easy to check that $tA \in \mathfrak{g}$ for all $t \in \mathbb{R}$. For each $A, B \in G$ and for each $m \in \mathbb{N}$, we see that

$$(\exp(tA/m)\exp(tB/m))^m \in G.$$

Now using the Lie product formula (Theorem 3.1.51), we conclude that

$$\exp(t(A+B)) = \lim_{m\to\infty}(\exp(tA/m)\exp(tB/m))^m \in G,$$

which conclude (b). □

PROOF OF THEOREM 3.1.64(C). By using (a), one sees that $\exp(tA)B\exp(-tA) \in \mathfrak{g}$ for all $t \in \mathbb{R}$, and by (b) we see that

$$\frac{\exp(tA)B\exp(-tA) - B}{t} \in \mathfrak{g} \quad \text{for all } t \in \mathbb{R}\setminus\{0\},$$

thus

$$\frac{\mathrm{d}}{\mathrm{d}t}(\exp(tA)B\exp(-tA))\Big|_{t=0} = \lim_{t\to 0}\frac{\exp(tA)B\exp(-tA) - B}{t} \in \mathfrak{g}.$$

Now using product rule we see that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\exp(tA)B\exp(-tA))\Big|_{t=0} \\
&\quad = A\exp(tA)B\exp(-tA) + \exp(tA)B(-A\exp(-tA))\Big|_{t=0} \\
&\quad = AB - BA = [A,B]
\end{aligned}$$

which complete the proof of (c). □

DEFINITION 3.1.65. Let G and H be matrix Lie groups. A map $\Phi : G \to H$ is called a *Lie group homomorphism* if:

(a) $\Phi : G \to H$ is a group homomorphism[2];
(b) $\Phi : G \to H$ is continuous.

In addition, if $\Phi : G \to H$ is bijective and its inverse $\Phi^{-1} : H \to G$ is continuous, then Φ is called a *Lie group isomorphism*.

The following theorem tells us that a Lie group homomorphism between two Lie groups gives rise in a natural way to a map between the corresponding Lie algebras.

THEOREM 3.1.66 ([**Hal15**, Theorem 3.28]). *Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then there exists a unique $\mathbb{R}$-linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\Phi(\exp(A)) = \exp(\phi(A)) \quad \textit{for all } A \in \mathfrak{g},$$

which satisfies

(a) $\phi(BAB^{-1}) = \Phi(B)\phi(A)\phi(B)^{-1}$ *for all* $A \in \mathfrak{g}$ *and* $B \in G$;
(b) $\phi([A,B]) = [\phi(A), \phi(B)]$ *for all* $A, B \in \mathfrak{g}$.
(c) $\phi(A) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi(\exp(tA))\big|_{t=0}$ *for all* $A \in \mathfrak{g}$.

[2]This means that $\Phi(A)\Phi(B) = \Phi(AB)$ for all $A, B \in G$. Here $\Phi(A)\Phi(B)$ is matrix multiplication in H, while AB is matrix multiplication in G.

If, in addition, $\Phi : G \to H$ is a Lie group isomorphism, then such $\mathbb{R}$-linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is bijective.

DEFINITION 3.1.67. We call such mapping $\phi : \mathfrak{g} \to \mathfrak{h}$ the *associated Lie algebra homomorphism* of the Lie group homomorphism $\Phi : G \to H$.

The proof of Theorem 3.1.66 is similar to Theorem 3.1.64, here we left the details for readers as an exercise.

DEFINITION 3.1.68. Let G be a matrix Lie group. A *representation* of a matrix Lie group G is a Lie group homomorphism

$$\Pi : G \to \mathrm{GL}(n, \mathbb{C}).$$

A representation of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism

$$\pi : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{C}).$$

In this note, we only deal with $\mathbb{C}^{n\times n}$. The notations in this section can be extend to abstract vector spaces, and this is related to the *group representation theory*, see e.g. the monograph [**Hal15**] for further details.

3.2. Homogeneous ODE with variable coefficients

In this section, we explain the basic results concerning the structure of solutions of a homogeneous system of ODE given by

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{p} = (p_1, \cdots, p_n). \tag{3.2.1}$$

We assume that A is continuous near t_0. By using the fundamental theorem of ODE (Theorem 2.1.5), there exist a unique C^1-solution Y such that

$$Y'(t) = A(t)Y(t) \text{ near } t_0, \quad Y(t_0) = I. \tag{3.2.2}$$

By using $\det(Y(t_0)) = 1$ and the continuity of $\det : \mathbb{C}^{n\times n} \to \mathbb{C}$, one sees that $\det(Y(t)) \neq 0$ for all t near t_0, i.e.

$$Y(t) \in \mathrm{GL}(n, \mathbb{C}) \quad \text{for all } t \text{ near } t_0.$$

We call such $Y(t)$ a *fundamental matrix solution* near t_0. Furthermore, the columns $\boldsymbol{y}_1(t), \cdots, \boldsymbol{y}_n(t)$ of $Y(t)$, which forms a linearly independent set in $\mathbb{C}^n$ near t_0, are called the *fundamental set of n linearly independent solution* near t_0. We see that

$$Y(t)\boldsymbol{p} = \sum_{j=1}^{n} p_j \boldsymbol{y}_j$$

is the unique solution of (3.2.1) near t_0. By arbitrariness of $\boldsymbol{p}$, we can rephrase the above in the following theorem.

THEOREM 3.2.1. *If A is continuous near t_0, then the set of all solutions of the ODE $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ near t_0 forms an n-dimensional vector space over $\mathbb{C}$.*

Similarly, by using the fundamental theorem of ODE (Theorem 2.1.5), there exist a unique C^1-solution Y such that

$$Z'(t) = -Z(t)A(t) \text{ near } t_0, \quad Z(t_0) = I, \tag{3.2.3}$$

which satisfies

$$Z(t) \in \mathrm{GL}(n, \mathbb{C}) \quad \text{for all } t \text{ near } t_0.$$

By using the chain rule, one sees that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(Z(t)Y(t)) &= Z'(t)Y(t) + Z(t)Y'(t)\\ &= -Z(t)A(t)Y(t) + Z(t)A(t)Y(t) = 0\end{aligned}$$

and $Z(t_0)Y(t_0) = I$. Hence we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}(Z(t)Y(t) - I) = 0, \quad (Z(t)Y(t) - I)\Big|_{t=t_0} = 0.$$

By using (uniqueness part in) the fundamental theorem of ODE (Theorem 2.1.5), we now see that

$$Z(t)Y(t) - I = 0 \text{ for all } t \text{ near } t_0,$$

and hence

$$Y(t) = Z(t)^{-1} \in \mathrm{GL}(n, \mathbb{C}) \text{ for all } t \text{ near } t_0.$$

We can refer (3.2.3) be the *adjoint problem* of (3.2.2).

We now want to compute $Y(t)$ in terms of $A(t)$. In this case when $A(t) \in \mathbb{C}^{1\times 1} \cong \mathbb{C}$, i.e. the ODE (3.2.2) is scalar ($n = 1$), then by using the fundamental theorem of calculus, we can easily obtain

$$Y(t) = \exp\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right) \quad \text{for all } t \text{ near } t_0. \tag{3.2.4}$$

When $n \geq 2$, the situation become tricky: By using the product rule, we see that

$$\begin{aligned}Y'(t) &= \left(\exp\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right)\right)'\\ &= \left(I + \left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right) + \frac{1}{2!}\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right)^2 + \cdots\right)'\\ &= A(t) + \frac{1}{2!}\left(\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right)^2\right)' + \cdots\\ &= A(t) + \frac{1}{2!}\left(A(t)\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right) + \left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right)A(t)\right) + \cdots.\end{aligned}$$

If

$$A(t) \text{ and } \int_{t_0}^{t} A(s)\,\mathrm{d}s \text{ are commute}, \tag{3.2.5}$$

then we reach

$$Y'(t) = A(t) + A(t)\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right) + \frac{1}{2!}A(t)\left(\int_{t_0}^{t} A(s)\,\mathrm{d}s\right)^2 + \cdots = A(t)Y(t).$$

In addition, by using Exercise 3.1.25, from (3.2.4) we see that

$$\det(Y(t)) = \exp\left(\int_{t_0}^{t} \mathrm{tr}\,(A(s))\,\mathrm{d}s\right) \quad \text{for all } t \text{ near } t_0. \tag{3.2.6}$$

EXAMPLE 3.2.2. $A(t) = \mathrm{diag}\,(\lambda_1(t), \cdots, \lambda_n(t))$ for some scalar functions $\lambda_1, \cdots, \lambda_n$ which are continuous near t_0.

We remind the readers that the existence of unique fundamental matrix solution $Y(t)$ does not require the additional assumption (3.2.5). The requirement (3.2.5) is quite restrictive: In general we do not know the explicit formula of the unique fundamental matrix solution $Y(t)$. Due to this reason, it is worth to mention the following theorem.

THEOREM 3.2.3 (Abel's formula [**HS99**, Remark IV-2-7(4)]). *If A is continuous near t_0, then the fundamental matrix solution $Y(t)$ satisfying (3.2.2) satisfies (3.2.6).*

DEFINITION 3.2.4. The quantity $W(t) := \det(Y(t))$ is called the *Wronskian* and the Abel formula (3.2.6) reads

$$W(t) = \exp\left(\int_{t_0}^{t} \operatorname{tr}\left(A(s)\right) \mathrm{d}s\right) \quad \text{for all } t \text{ near } t_0. \tag{3.2.7}$$

PROOF OF THEOREM 3.2.3. Using the product rule and Theorem 3.1.15, we see that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\det(Y(t))) \\
&\quad = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \left(\frac{\mathrm{d}}{\mathrm{d}t} Y_{\sigma(1),1}(t)\right) Y_{\sigma(2),2}(t)\cdots Y_{\sigma(n),n}(t) \\
&\qquad \vdots \\
&\qquad + \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) Y_{\sigma(1),1}(t) Y_{\sigma(2),2}(t)\cdots \left(\frac{\mathrm{d}}{\mathrm{d}t} Y_{\sigma(n),n}(t)\right) \\
&\quad = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \left(\frac{\mathrm{d}}{\mathrm{d}t} Y(t)\right)_{\sigma(1),1} Y_{\sigma(2),2}(t)\cdots Y_{\sigma(n),n}(t) \\
&\qquad \vdots \\
&\qquad + \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) Y_{\sigma(1),1}(t) Y_{\sigma(2),2}(t)\cdots \left(\frac{\mathrm{d}}{\mathrm{d}t} Y(t)\right)_{\sigma(n),n}
\end{aligned}$$

Since $Y'(t) = A(t)Y(t)$, then we reach

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\det(Y(t))) \\
&\quad = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\,(A(t)Y(t))_{\sigma(1),1}\, Y_{\sigma(2),2}(t)\cdots Y_{\sigma(n),n}(t) \\
&\quad\quad \vdots \\
&\quad\quad + \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) Y_{\sigma(1),1}(t) A_{\sigma(2),2}(t)\cdots (A(t)Y(t))_{\sigma(n),n} \\
&\quad = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(\boldsymbol{a}_{\sigma(1)}(t)\boldsymbol{y}_1(t))(\boldsymbol{e}_{\sigma(2)}^{\intercal}(t)\boldsymbol{y}_2(t))\cdots(\boldsymbol{e}_{\sigma(n)}^{\intercal}(t)\boldsymbol{y}_n(t)) \\
&\quad\quad \vdots \\
&\quad\quad + \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(\boldsymbol{e}_{\sigma(1)}^{\intercal}(t)\boldsymbol{y}_1(t))\cdots(\boldsymbol{e}_{\sigma(n-1)}^{\intercal}(t)\boldsymbol{y}_{n-1}(t))(\boldsymbol{a}_{\sigma(n)}(t)\boldsymbol{y}_n(t)) \\
&\quad = \det\left(\begin{pmatrix} \boldsymbol{a}_1(t) \\ \boldsymbol{e}_2^{\intercal}(t) \\ \vdots \\ \boldsymbol{e}_n^{\intercal}(t) \end{pmatrix} Y(t)\right) + \cdots + \det\left(\begin{pmatrix} \boldsymbol{e}_1^{\intercal}(t) \\ \vdots \\ \boldsymbol{e}_{n-1}^{\intercal}(t) \\ \boldsymbol{a}_n(t) \end{pmatrix} Y(t)\right)
\end{aligned}$$

where $\boldsymbol{a}_j$ be the j^{th} row of A and $\boldsymbol{y}_j$ is the j^{th} column of Y. Now we use Theorem 3.1.17 to see that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\det(Y(t))) &= \det\begin{pmatrix} \boldsymbol{a}_1(t) \\ \boldsymbol{e}_2^{\intercal}(t) \\ \vdots \\ \boldsymbol{e}_n^{\intercal}(t) \end{pmatrix} \det(Y(t)) + \cdots \begin{pmatrix} \boldsymbol{e}_1^{\intercal}(t) \\ \vdots \\ \boldsymbol{e}_{n-1}^{\intercal}(t) \\ \boldsymbol{a}_n(t) \end{pmatrix} \det(Y(t)) \\
&= a_{11}(t)\det(Y(t)) + \cdots + a_{nn}(t)\det(Y(t)) \\
&= \operatorname{tr}(A(t))\det(Y(t)),
\end{aligned}$$

which shows that $\det(Y(t))$ satisfies the *scalar* ODE of the form (3.2.2). Since $\det(Y(t_0)) = 1$, by using the fundamental theorem of ODE (Theorem 2.1.5), we conclude that the unique solution of the scalar ODE is given by (3.2.6) and we complete the proof of the theorem. □

3.3. Nonhomogeneous equations

In this section, we explain how to solve an initial-value problem

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{b}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{p}. \tag{3.3.1}$$

Here we assume that both A and $\boldsymbol{b}$ are continuous for all t near t_0. Let $Y(t) = Y(t;t_0) \in \mathrm{GL}(n,\mathbb{C})$ be the fundamental matrix solution satisfying

$$Y'(t) = A(t)Y(t), \quad Y(t_0) = I,$$

which was mentioned in previous sections (Section 3.1 and Section 3.2).

By plugging the anzats

$$\boldsymbol{y}(t) = Y(t)\boldsymbol{z}(t) \quad \text{for all } t \text{ near } t_0,$$

into (3.3.1) and by using the product rule, we see that

$$\begin{aligned} &\cancel{A(t)Y(t)\boldsymbol{z}(t)} + Y(t)\boldsymbol{z}'(t) = Y'(t)\boldsymbol{z}(t) + Y(t)\boldsymbol{z}'(t) \\ &\quad = \boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{b}(t) = \cancel{A(t)Y(t)\boldsymbol{z}(t)} + \boldsymbol{b}(t). \end{aligned}$$

Mutiply the above equation by $(Y(t))^{-1} \in \mathrm{GL}(n, \mathbb{C})$, we now see that

$$\boldsymbol{z}'(t) = (Y(t))^{-1}\boldsymbol{b}(t), \quad \boldsymbol{z}(t_0) = \boldsymbol{p}, \tag{3.3.2}$$

and its unique solution (guaranteed by the fundamental theorem of ODE (Theorem 2.1.5)) is given by

$$\boldsymbol{z}(t) = \boldsymbol{p} + \int_{t_0}^{t} (Y(s))^{-1}\boldsymbol{b}(s)\,\mathrm{d}s \quad \text{for all } t \text{ near } t_0.$$

Now we see that the unique solution $Y(t)$ of (3.3.1) is given by

$$\boldsymbol{y}(t) = Y(t; t_0)\left(\boldsymbol{p} + \int_{t_0}^{t} (Y(s; t_0))^{-1}\boldsymbol{b}(s)\,\mathrm{d}s\right) \quad \text{for all } t \text{ near } t_0.$$

Remark 3.3.1. The numerical computation of $(Y(t))^{-1}\boldsymbol{b}(t)$ is quite fundamental, but keep in mind that one never compute the inverse of $(Y(t))^{-1}$ directly (the computation is both inaccuate and slow). In practical, this is computed via iterative algorithm such as GMRES, conjugate gradient, etc. One can refer to the monograph [**TB22**] for a nice introduction of numerical linear algebra. Here we recall that in fact $(Y(t))^{-1}$ is the fundamental matrix solution of the adjoint problem (3.2.3), which allows us to compute $(Y(t))^{-1}$ by solving an ODE, which is much better than compute $(Y(t))^{-1}\boldsymbol{b}(t)$ via numerical method.

In fact, the above formula can be further simplified and we now label the subscript for clarification. For each fixed s near t_0, we now see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(Y(t; t_0)(Y(s; t_0))^{-1}\right) = A(t)\left(Y(t; t_0)(Y(s; t_0))^{-1}\right), \quad \left(Y(t; t_0)(Y(s; t_0))^{-1}\right)\Big|_{t=s} = I,$$

and t the fundamental theorem of ODE (Theorem 2.1.5) says that

$$Y(t; s) = Y(t; t_0)(Y(s; t_0))^{-1},$$

and we now conclude that the unique solution of (3.3.1) is given by

$$\boldsymbol{y}(t) = Y(t; t_0)\boldsymbol{p} + \int_{t_0}^{t} Y(t; s)\boldsymbol{b}(s)\,\mathrm{d}s \quad \text{for all } t \text{ near } t_0. \tag{3.3.3}$$

It is worth to mention that the solution formula (3.3.3) do not involve $Y(t)^{-1}$, as well as the adjoint problem (3.2.3), at all.

Remark 3.3.2. When $A(t) \equiv A$ is a constant matrix, we see that

$$Y(t; s) = \exp((t - s)A) \quad \text{for all } t, s \in \mathbb{R},$$

and now (3.3.3) reads

$$\boldsymbol{y}(t) = \exp((t - t_0)A)\boldsymbol{p} + \int_{t_0}^{t} \exp((t - s)A)\boldsymbol{b}(s)\,\mathrm{d}s.$$

Note that the term $\int_{t_0}^{t} \exp((t - s)A)\boldsymbol{b}(s)\,\mathrm{d}s$ is the convolution.

EXAMPLE 3.3.3. Let us solve the initial value problem

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{b}(t), \quad \boldsymbol{y}(0) = \boldsymbol{p} \tag{3.3.4}$$

with

$$A = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 3 & 2 & 1 \end{pmatrix}, \quad \boldsymbol{b}(t) = \begin{pmatrix} 2 \\ 0 \\ t \end{pmatrix}, \quad \boldsymbol{p} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

The fundamental matrix solution is given by

$$Y(t) = \exp(tA) = \begin{pmatrix} e^{-2t} & te^{-2t} & 0 \\ 0 & e^{-2t} & 0 \\ e^t - e^{-2t} & e^t - (1+t)e^{-2t} & e^t \end{pmatrix}. \tag{3.3.5}$$

Therefore, the solution of (3.3.4) is given by

$$\boldsymbol{y}(t) = \begin{pmatrix} 1 + te^{-2t} \\ e^{-2t} \\ -4 - t + 5e^t - (1+t)e^{-2t} \end{pmatrix}.$$

EXERCISE 3.3.4. Prove (3.3.5).

3.4. Higher order linear ODE

In this section, we explain how to solve the initial value problem of an n^{th} order linear ODE

$$\begin{cases} u^{(n)} + a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u = b(t) & \text{for all } t \text{ near } t_0, \\ u(t_0) = p_1, u'(t_0) = p_2, \cdots, u^{(n-1)}(t_0) = p_n. \end{cases} \tag{3.4.1}$$

We first prove the following fundamental result.

THEOREM 3.4.1. *If the coefficients $a_1, \cdots, a_n, b$ are continuous near t_0, then there exists a unique C^n-solution u of (3.4.1) near t_0.*

PROOF. **Existence of solution.** We now define

$$\begin{aligned} A(t) &= \begin{pmatrix} \mathbf{0}_{n-1} & I_{n-1} \\ -a_n(t) & \big(-a_{n-1}(t), \cdots, -a_1(t)\big) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n(t) & -a_{n-1}(t) & -a_{n-2}(t) & \cdots & -a_1(t) \end{pmatrix} \end{aligned} \tag{3.4.2}$$

where $\mathbf{0}_{n-1} \in \mathbb{R}^{n-1}$ is the zero vector, $I_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the identity matrix, and

$$\boldsymbol{b}(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}, \quad \boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}. \tag{3.4.3}$$

In previous section (Section 3.3), we have showed that there exists a unique C^1 solution $\boldsymbol{y}(t)$ of

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{b}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{p}. \tag{3.4.4}$$

For each t and s, which are close to t_0, let

$$Y(t;s) = \begin{pmatrix} \boldsymbol{y}_1(t;s) \\ \vdots \\ \boldsymbol{y}_{n-1}(t;s) \\ \boldsymbol{y}_n(t;s) \end{pmatrix}$$

be the fundamental matrix solution satisfying $Y(s) = I$ and

$$\begin{aligned} \begin{pmatrix} \boldsymbol{y}_1'(t;s) \\ \vdots \\ \boldsymbol{y}_{n-1}'(t;s) \\ \boldsymbol{y}_n'(t;s) \end{pmatrix} &= Y'(t;s) = A(t)Y(t;s) \\ &= \begin{pmatrix} \mathbf{0}_{n-1} & I_{n-1} \\ -a_n(t) & \big(-a_{n-1}(t), \cdots, -a_1(t)\big) \end{pmatrix} Y(t;s) \\ &= \begin{pmatrix} \boldsymbol{y}_2(t;s) \\ \vdots \\ \boldsymbol{y}_n(t;s) \\ -a_n(t)\boldsymbol{y}_1(t;s) - \cdots - a_1\boldsymbol{y}_n(t;s) \end{pmatrix}, \end{aligned}$$

and we see that

$$\boldsymbol{y}_j'(t;s) = \boldsymbol{y}_{j+1}(t;s) \quad \text{for all } j = 1, \cdots, n-1. \tag{3.4.5}$$

From (3.3.3), we now have

$$\boldsymbol{y}(t) = Y(t;t_0)\boldsymbol{p} + \int_{t_0}^{t} Y(t;s)\boldsymbol{b}(s)\,\mathrm{d}s. \tag{3.4.6}$$

We now define

$$u(t) := y_1(t) = \boldsymbol{y}_1(t;t_0)\boldsymbol{p} + \int_{t_0}^{t} \boldsymbol{y}_1(t;s)\boldsymbol{b}(s)\,\mathrm{d}s \in C^1 \text{ near } t_0 \tag{3.4.7}$$

Now from (3.4.5) we see that $u \in C^n$ near t_0 and

$$\boldsymbol{y}(t) = \begin{pmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{pmatrix}, \tag{3.4.8}$$

then we see that

$$\begin{aligned} &A(t)\boldsymbol{y}(t)+\boldsymbol{b}(t)\\ &=\begin{pmatrix} \mathbf{0}_{n-1} & I_{n-1}\\ -a_n(t) & \big(-a_{n-1}(t),\cdots,-a_1(t)\big)\end{pmatrix}\begin{pmatrix} u(t)\\ \vdots\\ u^{(n-2)}(t)\\ u^{(n-1)}(t)\end{pmatrix}+\begin{pmatrix} 0\\ \vdots\\ 0\\ b(t)\end{pmatrix}\\ &=\begin{pmatrix} u'(t)\\ \vdots\\ u^{(n-2)}(t)\\ -a_1(t)u^{(n-1)}(t)-a_2(t)u^{(n-1)}(t)-\cdots-a_n(t)u(t)+b(t)\end{pmatrix}, \end{aligned}$$

and hence (3.4.4) gives (3.4.1).

Uniqueness. if $u\in C^n$ is a solution of (3.4.1) near t_0, then the C^1-function $\boldsymbol{y}$ given in (3.4.8) satisfies (3.4.4). □

REMARK 3.4.2. We recall (Definition 3.2.4) the Wronskian is defined by $W(t):=\det(Y(t))$, and using the Abel's formula (Theorem 3.2.3), we see that

$$W(t)=\exp\left(\int_{t_0}^t \operatorname{tr}\left(A(s)\right)\mathrm{d}s\right)=\exp\left(-\int_{t_0}^t a_1(s)\,\mathrm{d}s\right).$$

REMARK 3.4.3 (2-dimensional case [**HS99**, Remark IV-7-2 and Remark IV-7-3]). In practical computation, the most difficult part is to compute $\boldsymbol{y}_1(t;s)=\boldsymbol{y}_1(t)(\boldsymbol{y}_1(s))^{-1}$. When $n=2$, the computation of $(\boldsymbol{y}_1(s))^{-1}$ can be further simplified. As mentioned in the proof, one has

$$\boldsymbol{y}_1(t)=\begin{pmatrix}\phi_1(t) & \phi_2(t)\\ \phi_1'(t) & \phi_2'(t)\end{pmatrix}$$

where $\phi_1(t)$ and $\phi_2(t)$ are two linearly independent solutions of the associated homogeneous equation

$$v''+a_1(t)v'+a_2(t)v=0. \tag{3.4.9}$$

We see that

$$(\boldsymbol{y}_1(s;t_0))^{-1}=\frac{1}{W(t)}\begin{pmatrix}\phi_2'(t) & -\phi_2(t)\\ -\phi_1'(t) & \phi_1(t)\end{pmatrix}.$$

In this case, the Wronskian reads

$$W(t)=\phi_1(t)\phi_2'(t)-\phi_1'(t)\phi_2(t) \tag{3.4.10}$$

Therefore from (3.4.7) we know that the unique solution u of

$$\begin{cases} u''+a_1(t)u'+a_2(t)u=b(t) & \text{for all } t \text{ near } t_0,\\ u(t_0)=p_1, u'(t_0)=p_2,\end{cases}$$

is given by

$$u(t):=p_1\phi_1(t)+p_2\phi_2(t)-\phi_1(t)\int_{t_0}^t\frac{\phi_2(s)}{W(s)}b(s)\,\mathrm{d}s+\phi_2(t)\int_{t_0}^t\frac{\phi_1(s)}{W(s)}b(s)\,\mathrm{d}s.$$

We now write (3.4.10) as

$$\left(\frac{\phi_2(t)}{\phi_1(t)}\right)' = \frac{\phi_2'(t)}{\phi_1(t)} - \frac{\phi_1'(t)}{(\phi_1(t))^2}\phi_2(t) = \frac{W(t)}{(\phi_1(t))^2}.$$

Let $\tilde{W}$ be any function such that $\tilde{W}'(t) = \frac{W(t)}{(\phi_1(t))^2}$, and one sees that $\phi_2(t) = \phi_1(t)\tilde{W}(t)$ satisfies the above ODE. Now using Corollary 3.4.5. we conclude that the solution set of (3.4.9) is a 2-dimensional vector space with basis

$$\{\phi_1(t), \phi_1(t)\tilde{W}(t)\}.$$

REMARK 3.4.4 (3-dimensional case [**HS99**, Remark IV-7-4]). We write

$$Y(t) = \begin{pmatrix} \phi(t) & \phi_1(t) & \phi_2(t) \\ \phi'(t) & \phi_1'(t) & \phi_2'(t) \\ \phi''(t) & \phi_1''(t) & \phi_2''(t) \end{pmatrix},$$

and we see that $\phi(t)$, $\phi_1(t)$ and $\phi_2(t)$ are lienarly independent solutions of the homogeneous equation

$$u''' + a_1(t)u'' + a_2(t)u' + a_3(t)u = 0.$$

By using the cofactor expansion of determinant (Theorem 3.1.14), the Wronskian reads

$$W(t) = \phi''(t)\det\begin{pmatrix} \phi_1(t) & \phi_2(t) \\ \phi_1'(t) & \phi_2'(t) \end{pmatrix} - \phi'(t)\det\begin{pmatrix} \phi_1(t) & \phi_2(t) \\ \phi_1''(t) & \phi_2''(t) \end{pmatrix} + \phi(t)\det\begin{pmatrix} \phi_1'(t) & \phi_2'(t) \\ \phi_1''(t) & \phi_2''(t) \end{pmatrix},$$

and we write

$$\phi''(t) + \frac{A_1(t)}{A_0(t)}\phi'(t) + \frac{A_2(t)}{A_0(t)}\phi(t) = \frac{W(t)}{A_0(t)}$$

where

$$A_0(t) = \det\begin{pmatrix} \phi_1(t) & \phi_2(t) \\ \phi_1'(t) & \phi_2'(t) \end{pmatrix},$$
$$A_1(t) = \det\begin{pmatrix} \phi_1(t) & \phi_2(t) \\ \phi_1''(t) & \phi_2''(t) \end{pmatrix},$$
$$A_2(t) = \det\begin{pmatrix} \phi_1'(t) & \phi_2'(t) \\ \phi_1''(t) & \phi_2''(t) \end{pmatrix}.$$

This means that the general solution of ϕ can be expressed in terms of ϕ_1, ϕ_2 and W.

The proof of Theorem 3.4.1 itself gives an algorithm to compute the unique solution. Since $Y(t;s) \in \mathrm{GL}(n,\mathbb{C})$, then we also have the following corollary.

COROLLARY 3.4.5. *If the coefficients $a_1, \cdots, a_n$ are continuous near t_0, then the solution set*

$$\{u \in C^n \text{ near } t_0 : u^{(n)} + a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u = 0 \text{ near } t_0\}$$

forms an n-dimensional vector space.

EXAMPLE 3.4.6. Let us solve the initial value problem

$$u''' - 2u'' - 5u + 6u = 3t, \quad u(0) = 1, u'(0) = 2, u''(0) = 0. \tag{3.4.11}$$

In this case, the matrix A given in (3.4.2) and the vectors $\boldsymbol{b}$ as well as $\boldsymbol{p}$ given in (3.4.3) read

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{pmatrix}, \quad \boldsymbol{b}(t) = \begin{pmatrix} 0 \\ 0 \\ 3t \end{pmatrix}, \quad \boldsymbol{p} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

We can compute

$$\begin{aligned} Y(t) &= \exp(tA) \\ &= -\frac{e^t}{6}\begin{pmatrix} -6 & -1 & 1 \\ -6 & -1 & 1 \\ -6 & -1 & 1 \end{pmatrix} + \frac{e^{-2t}}{15}\begin{pmatrix} 3 & -4 & 1 \\ -6 & 8 & -2 \\ 12 & -16 & 4 \end{pmatrix} + \frac{e^{3t}}{10}\begin{pmatrix} -2 & 1 & 1 \\ -6 & 3 & 3 \\ -18 & 9 & 9 \end{pmatrix}. \end{aligned} \tag{3.4.12}$$

By a direct but long (and boring) computations we reach

$$\begin{aligned} &\int_0^t Y(s)^{-1}\boldsymbol{b}(s)\,\mathrm{d}s \\ &= -\frac{1}{2}\begin{pmatrix} 1-(t+1)e^{-t} \\ 1-(t+1)e^{-t} \\ 1-(t+1)e^{-t} \end{pmatrix} + \frac{1}{20}\begin{pmatrix} 1+(2t-1)e^{2t} \\ -2(1+(2t-1)e^{2t}) \\ -4(1+(2t-1)e^{2t}) \end{pmatrix} + \frac{1}{30}\begin{pmatrix} 1+(3t+1)e^{-3t} \\ 3(1+(3t+1)e^{-3t}) \\ 9(1+(3t+1)e^{-3t}) \end{pmatrix}, \end{aligned}$$

and hence

$$\boldsymbol{y}(t) = \frac{1}{60}\begin{pmatrix} 30t+25-17e^{-2t}+50e^t+2e^{3t} \\ 2(15+17e^{-2t}+25e^t+3e^{3t}) \\ 2(-34e^{-2t}+25e^t+9e^{3t}) \end{pmatrix}.$$

We finally conclude from (3.4.7) that

$$u(t) = \frac{1}{60}(30t+25-17e^{-2t}+50e^t+2e^{3t})$$

is the unique solution of (3.4.11).

EXERCISE 3.4.7. Proof (3.4.12) using Algorithm 2. (**Hint.** See Exercise 3.1.40)

Let $u \in C^n$ be the unique solution of (3.4.1). Let $v \in C^n$ be any function satisfying

$$v^{(n)} + a_1(t)v^{(n-1)} + \cdots + a_{n-1}(t)v' + a_n(t)v = b(t) \text{for all } t \text{ near } t_0,$$

then one sees that the solution $w = u - v$ is the unique solution to

$$\begin{cases} w^{(n)} + a_1(t)w^{(n-1)} + \cdots + a_{n-1}(t)w' + a_n(t)w = 0 & \text{for all } t \text{ near } t_0, \\ w(t_0) = p_1 - v(t_0), u'(t_0) = p_2 - v'(t_0), \cdots, u^{(n-1)}(t_0) = p_n - v^{(n-1)}(t_0). \end{cases} \tag{3.4.13}$$

For the case when $a_1, \cdots, a_n$ are constants, there is an efficient way (based on Corollary 3.4.5) to compute the solution w of (3.4.13). By plugging the anzats $w(t) = e^{\lambda t}$ into the equation

$$w^{(n)} + a_1 w^{(n-1)} + \cdots + a_{n-1}w' + a_n w = 0, \tag{3.4.14}$$

we reach the equation

$$P(\lambda) := \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0.$$

It is important to obseve that (3.4.14) can be rewritten as

$$P\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) w = 0.$$

By using the fundamental theorem of algebra (see e.g. [**Kow23**] for a proof), there exists distinct $\lambda_1, \cdots, \lambda_k \in \mathbb{C}$ such that

$$P(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k} \quad \text{with } m_1 + \cdots + m_k = n.$$

For any differentiable function g, we see that

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} - \lambda_j\right)(e^{\lambda_j t} g(t)) = e^{\lambda_j t} g'(t),$$

and inductively one can show

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} - \lambda_j\right)^{m_j} (e^{\lambda_j t} g(t)) = e^{\lambda_j t} g^{(m_j)}(t).$$

For each $c_1, \cdots, c_{m_j} \in \mathbb{C}$, we now choose $g(t) = c_1 + c_2 t + \cdots + c_{m_j} t^{m_j}$, we see that $g^{(m_j)}(t) \equiv 0$, and hence

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} - \lambda_j\right)^{m_j} (c_1 e^{t\lambda_j} + c_2 t e^{t\lambda_j} + \cdots + c_{m_j} t^{m_j - 1} e^{t\lambda_j}) = 0,$$

thus

$$P\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)(c_1 e^{t\lambda_j} + c_2 t e^{t\lambda_j} + \cdots + c_{m_j} t^{m_j - 1} e^{t\lambda_j}) = 0.$$

Since the above argument works for all $j = 1, \cdots, k$, combining with Corollary 3.4.5, we reach the following theorem.

THEOREM 3.4.8. *The solution set of* (3.4.14) *is a* $\mathbb{C}$*-vector space with basis*

$$\bigcup_{j=1}^{k} \{e^{t\lambda_j}, t e^{t\lambda_j}, \cdots, t^{m_j - 1} e^{t\lambda_j}\}.$$

EXAMPLE 3.4.9. We now revisit Example 3.4.6. We first note that

$$v(t) := \frac{1}{60}(30t + 25) = \frac{t}{2} + \frac{5}{12}$$

is a particular solution to $v''' - 2v'' - 5v + 6v = 3t$. We now see that $w = u - v$ solves

$$w''' - 2w'' - 5w + 6w = 0, \quad w(0) = \frac{7}{12}, u'(0) = \frac{3}{2}, w''(0) = 0.$$

We now consider

$$P(\lambda) := \lambda^3 - 2\lambda^2 - 5\lambda + 6 = 0$$

One can check that the roots of P are $-2, 1, 3$. Hence the general solution of w is

$$w(t) = c_1 e^{-2t} + c_2 e^t + c_3 e^{3t}.$$

Now we see that

$$\begin{aligned} \frac{7}{12} &= w(0) = c_1 + c_2 + c_3, \\ \frac{3}{2} &= w'(0) = -2c_1 + c_2 + 3c_3, \\ 0 &= w''(0) = 4c_1 + c_2 + 9c_3. \end{aligned}$$

Solving the above system gives $c_1 = -\frac{17}{60}$, $c_2 = \frac{5}{6}$ and $c_3 = \frac{1}{30}$, and we reach

$$w(t) = \frac{1}{60}(-17e^{-2t} + 50e^t + 2e^{3t}).$$

Finally, we conclude that

$$u(t) = v(t) + w(t) = \frac{1}{60}(30t + 25 - 17e^{-2t} + 50e^{t} + 2e^{3t})$$

is the unique solution of (3.4.11).

3.5. Strum-Liouville eigenvalue problem

Let $a, b, \theta_1, \theta_2 \in \mathbb{R}$ with $a < b$, let $q \in C([a,b])$ is real-valued and let $p \in C^1((a,b)) \cap C([a,b])$ is real-valued such that

$$p(t) > 0 \quad \text{for all } t \in [a,b].$$

We define the linear operator $\mathcal{L} : C^2((a,b)) \cap C^1([a,b]) \to C((a,b))$ by

$$(\mathcal{L}[u])(t) := \frac{\mathrm{d}}{\mathrm{d}t}\left(p(t)\frac{\mathrm{d}u}{\mathrm{d}t}\right) + q(t)u$$

In this section, we consider the eigenvalue problem

$$\mathcal{L}[u] = -\lambda u \quad \text{for all } t \in (a,b) \tag{3.5.1a}$$

subject to the boundary conditions

$$\begin{cases} u(a)\cos\theta_1 - p(a)u'(a)\sin\theta_1 = 0, \\ u(b)\cos\theta_2 - p(b)u'(b)\sin\theta_2 = 0. \end{cases} \tag{3.5.1b}$$

It is easy to see that (3.5.1a)–(3.5.1b) has a solution $u \equiv 0$, which is called the *trivial solution*. We see that the boundary value problem (3.5.1a)–(3.5.1b) is over-determined, and we expect that in general the nonexistence of nontrivial solution (i.e. $u \not\equiv 0$) without any further assumptions. We are interested in the following object:

Definition 3.5.1. If there exists $\lambda \in \mathbb{C}$ and a nontrivial solution $u \in C^2((a,b)) \cap C^1([a,b])$ of (3.5.1a)–(3.5.1b), then we say that such λ is a *Strum-Liouville eigenvalue* and such nontrivial solution u is called the corresponding *Strum-Liouville eigenfunction*. The boundary value problem (3.5.1a)–(3.5.1b) is called the *Strum-Liouville eigenvalue problem*.

We define

$$(u,v)_{L^2(a,b)} := \int_a^b u(t)\overline{v(t)}\,\mathrm{d}t,$$

and

$$\|u\|_{L^2(a,b)} := (u,u)_{L^2(a,b)}^{1/2} = \left(\int_a^b |u(t)|^2\,\mathrm{d}t\right)^{1/2}.$$

We remind the inner product is skew-Hermitian:

$$(v,u)_{L^2(a,b)} = \int_a^b v(t)\overline{u(t)}\,\mathrm{d}t = \overline{(u,v)_{L^2(a,b)}}.$$

one can compare this with the the inner product $(\cdot,\cdot)_{\mathbb{C}^{n\times n}}$ given in (3.1.9). We first observe the following lemma.

LEMMA 3.5.2. *The operator* $\mathcal{L} : C^2((a,b)) \to C((a,b))$ *is* Hermitian *or* self-adjoint *in the sense of*

$$(\mathcal{L}[u], v)_{L^2(a,b)} = (u, \mathcal{L}[v])_{L^2(a,b)} \tag{3.5.2}$$

for all $u, v \in C^2((a,b)) \cap C^1([a,b])$ *both satisfy the boundary condition* (3.5.1b). *In addition, if* λ *is a Strum-Liouville eigenvalue, then* $\lambda \in \mathbb{R}$.

REMARK 3.5.3. It is interesting to compare (3.5.2) with a characterization of Hermitian matrix in Lemma 3.1.28, and compare the second statement with the result for Hermitian matrix in Theorem 3.1.29.

PROOF OF LEMMA 3.5.2. By using the integration by parts, we see that

$$\begin{aligned}&(\mathcal{L}[u], v)_{L^2(a,b)} \\ &\quad = p(t)u'(t)\overline{v(t)}\Big|_{t=a}^{t=b} - \int_a^b p(t)u'(t)\overline{v'(t)}\,\mathrm{d}t + \int_a^b q(t)u(t)\overline{v(t)}\,\mathrm{d}t.\end{aligned}$$

We now treat the term $p(b)u'(b)\overline{v(b)}$ into two cases:

Case 1. If $\theta_2 \notin \pi\mathbb{Z}$, then $\sin\theta_2 \neq 0$, and we now can write

$$p(b)u'(b)\overline{v(b)} = \overbrace{(p(b)u'(b)\sin\theta_2)}^{=0}\frac{\overline{v(b)}}{\sin\theta_2} = 0.$$

Case 2. Otherwise, if $\theta_2 \in \pi\mathbb{Z}$, then $\sin\theta_2 = 0$ and $\cos\theta_2 \in \{-1, 1\}$. Now from (3.5.1b) we see that

$$u(b) = v(b) = 0,$$

and thus $p(b)u'(b)\overline{v(b)} = 0$.

Similar arguments show that $p(a)u'(a)\overline{v(a)} = 0$. We now see that

$$\begin{aligned}&\int_a^b \mathcal{L}[u](t)\overline{v(t)}\,\mathrm{d}t \\ &\quad = -\int_a^b p(t)u'(t)\overline{v'(t)}\,\mathrm{d}t + \int_a^b q(t)u(t)\overline{v(t)}\,\mathrm{d}t.\end{aligned}$$

Since both p and q are real-valued, interchanging the role of u and $\overline{v}$ we see that

$$\begin{aligned}&\int_a^b u(t)\overline{\mathcal{L}[v](t)}\,\mathrm{d}t \\ &\quad = -\int_a^b p(t)u'(t)\overline{v'(t)}\,\mathrm{d}t + \int_a^b q(t)u(t)\overline{v(t)}\,\mathrm{d}t,\end{aligned}$$

and combining the above two equations, we conclude that $\mathcal{L}$ is Hermitian.

Now let $\lambda \in \mathbb{C}$ be an eigenvalue with eigenfunction u. We see that

$$\lambda\|u\|_{L^2(a,b)}^2 = (\mathcal{L}[u], v)_{L^2(a,b)} = (u, \mathcal{L}[v])_{L^2(a,b)} = \overline{\lambda}\|u\|_{L^2(a,b)}^2.$$

Since $\|u\|_{L^2(a,b)}^2 \neq 0$, we conclude that $\lambda = \overline{\lambda}$. □

In fact, we have the following theorems.

THEOREM 3.5.4 ([**HS99**, Theorem VI-3-11, Theorem VI-4-1 and Theorem VI-4-4]). *Let $a, b, \theta_1, \theta_2 \in \mathbb{R}$ with $a < b$, let $q \in C([a,b])$ is real-valued and let $p \in C^1((a,b)) \cap C([a,b])$ is real-valued such that*

$$p(t) > 0 \quad \text{for all } t \in [a,b].$$

Then:

(a) *there exists a countable sequence of eigenvalues $\lambda_1 < \lambda_2 < \lambda_3 < \cdots \to +\infty$ of the Strum-Liouville eigenvalue problem (3.5.1a)–(3.5.1b).*

(b) *Let λ_i and λ_j be two distinct eigenvalues, then the corresponding eigenfunctions u_i and u_j are orthogonal in $L^2(a,b)$, that is,*

$$(u_i, u_j)_{L^2(a,b)} = 0.$$

(c) *For every $u \in C^2((a,b)) \cap C^1([a,b])$ satisfying the boundary condition (3.5.1b), the series*

$$\sum_{j=1}^{+\infty} (f, u_j)_{L^2(a,b)} u_j$$

converges to u in $L^\infty(a,b)$, provided the eigenfunctions are normalized to $\|u_j\|_{L^2(a,b)} = 1$.

EXAMPLE 3.5.5. For simplicity, we put $a = 0$ and $b = \pi$. We now choose $p(t) \equiv 1$ and $q(t) \equiv 0$, and now the Strum-Liouville problem (3.5.1a)–(3.5.1b) reads

$$u''(t) = -\lambda u(t) \quad \text{for all } t \in (0, \pi) \tag{3.5.3a}$$

subject to the boundary conditions

$$\begin{cases} u(a)\cos\theta_1 - u'(a)\sin\theta_1 = 0, \\ u(b)\cos\theta_2 - u'(b)\sin\theta_2 = 0. \end{cases} \tag{3.5.3b}$$

By choosing $\theta_1 = \theta_2 = 0$, we reach an orthogonal sequence $\{\sin(nt)\}_{n\in\mathbb{N}}$ of eigenfunctions. By choosing $\theta_1 = \theta_2 = \frac{\pi}{2}$, we reach an orthogonal sequence $\{\cos(nt)\}_{n=0}^{\infty}$ of eigenfunctions. This induces Fourier series [**Kow22**]. If we still have time, we can continue the course by using the material [**Kow22**].

Bibliography

[Apo74] T. M. Apostol. *Mathematical analysis*. Addison-Wesley Publishing Co., second edition, 1974. MR0344384, Zbl:0309.26002.

[BD22] W. E. Boyce and R. C. DiPrima. *Elementary differential equations and boundary value problems*. John Wiley & Sons, Inc., Hoboken, NJ, 12th edition, 2022. MR0179403, Zbl:1492.34001.

[Che16] I-L. Chern. *Mathematical modeling and ordinary differential equations*. Lecture notes. National Taiwan University, 2016. https://www.math.ntu.edu.tw/~chern/notes/ode2015.pdf.

[Hal15] B. Hall. *Lie groups, Lie algebras, and representations. An elementary introduction*, volume 222 of *Grad. Texts in Math.* Springer, Cham, second edition, 2015. MR3331229, Zbl:1316.22001, doi:10.1007/978-3-319-13467-3.

[Hal13] B. C. Hall. *Quantum theory for mathematicians*, volume 267 of *Grad. Texts in Math.* Springer, New York, NY, 2013. Zbl:1273.81001, doi:10.1007/978-1-4614-7116-5.

[HS99] P.-F. Hsieh and Y. Sibuya. *Basic theory of ordinary differential equations*. Universitext. Springer-Verlag, New York, 1999. MR1697415, doi:10.1007/978-1-4612-1506-6.

[Joh78] F. John. *Partial differential equations*, volume 1 of *Appl. Math. Sci.* Springer-Verlag, New York-Berlin, third edition, 1978. MR0514404, Zbl:0426.35002.

[Kow22] P.-Z. Kow. *Fourier analysis and distribution theory*. University of Jyväskylä, 2022. https://puzhaokow1993.github.io/homepage.

[Kow23] P.-Z. Kow. *Complex Analysis*. National Chengchi University, Taipei, 2023. https://puzhaokow1993.github.io/homepage.

[Kow24] P.-Z. Kow. *An introduction to partial differential equations and functional analysis*. National Chengchi University, Taipei, 2024. https://puzhaokow1993.github.io/homepage.

[Kwa17] M. Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fractional Calculus and Applied Analysis*, 20(1):7–51, 2017. MR3613319, Zbl:1375.47038, doi:10.1515/fca-2017-0002, arXiv:1507.07356.

[Pug15] C. C. Pugh. *Real mathematical analysis*. Undergrad. Texts Math. Springer, Cham, second edition, 2015. MR3380933, Zbl:1329.26003, doi:10.1007/978-3-319-17771-7.

[Str08] W. A. Strauss. *Partial differential equations: An introduction*. John Wiley & Sons, Ltd., Chichester, second edition, 2008. MR2398759, Zbl:1160.35002.

[TB22] L. N. Trefethen and D. III Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 25th anniversary edition, 2022. MR4713493, Zbl:1510.65092.

[Tre17] S. Treil. *Linear algebra done wrong*. Brown University, 2017. https://www.math.brown.edu.