

# Notes: MATH273 (F25) Basic Theory of Ordinary Differential Equations

Lecturer: Jiajie Chen

Notes by: Aden Chen

Tuesday 11<sup>th</sup> November, 2025

# Contents

<b>1</b>	<b>Motivation, Preview of Application</b>	<b>3</b>
<b>2</b>	<b>Basic Definitions and Examples</b>	<b>4</b>
2.1	First Order Linear ODEs . . . . .	5
<b>3</b>	<b>Separation of Variables</b>	<b>8</b>
3.1	Generalization . . . . .	10
3.2	. . . . .	12
<b>4</b>	<b>Second Order Linear ODEs</b>	<b>13</b>
4.1	Homogeneous Second Order Linear ODEs with Constant Coefficients	13
4.2	Series Solution to Homogeneous Second Order Linear ODEs . . . . .	16
4.3	Non-Homogeneous Second Order Linear ODEs . . . . .	19
<b>5</b>	<b>First order ODE system</b>	<b>21</b>
5.1	Motivation and Setup . . . . .	21
5.2	Solving the Homogeneous First Order Linear ODE System . . . . .	22
5.3	Finding Solutions Explicitly . . . . .	25
5.4	Matrix Exponential . . . . .	29
5.5	Nonhomogeneous ODE Systems . . . . .	32
<b>6</b>	<b>The Theory of Existence and Uniqueness</b>	<b>35</b>
<b>7</b>	<b>Quantitative Estimates</b>	<b>37</b>
7.1	Extension . . . . .	42
7.2	Finite Time Blowup . . . . .	44
<b>8</b>	<b>Autonomous ODE, Stability, and Phase Portrait/Plane</b>	<b>45</b>
8.1	Phase Portrait of 2 x 2 Systems . . . . .	45
8.2	Stability . . . . .	47


# 1 Motivation, Preview of Application

*Example 1.1* ((Stochastic) gradient descent). We are interested in solving

$$\min_{x \in D} g(x)$$

where  $g$  represents a cost, and  $D \subset \mathbb{R}^n$ . The FOC is  $\nabla g(x) = 0$ . If  $g$  is nonlinear and  $n \gg 1$ , this is a very hard problem. We can however always consider the ODE

$$\frac{d}{dt}x(t) = -\nabla g(x(t)),$$

where  $g$  is given and  $x(t)$  is unknown. If  $\mathbb{R}^{n \times n} \ni \nabla^2 g > 0$  (is positive definite) and  $x(t_0) = x_0$ , then  $x(t) \rightarrow x_*$ , where  $x_* := \arg \min_{x \in D} g(x)$ . 

## 2 Basic Definitions and Examples

**Definition 2.1** (Differential Equation). A **differential equation** is an equation that relates a function  $y$  and its derivatives. A general representation is

$$F \left[ x, y, \partial_i y, \partial_i^2 y, \dots, \partial_i^{(n)} y \right] = 0.$$

*Problem 2.2* (Heat Equation). Consider  $u(t, x)$ , where  $t \in \mathbb{R}$  and  $x \in \mathbb{R}^d$ .

$$\partial_t u(t, x) = \Delta u(t, x) = \sum_{i=1}^n \partial_{x_i}^2 u(t, x).$$

This is a second order differential equation.

*Problem 2.3.*

$$\frac{d^2}{dt^2} u + \frac{d}{dt} u = u.$$

This is a second order ODE.

- **ODEs** contain only derivatives on one variable.
- **PDEs** can contain multiple partial derivatives.

**Definition 2.4** (Order of a Differential Equation). The **order** of a differential equation is the order of the highest order derivative that appears in the equation.

**Definition 2.5** (Linear and Nonlinear ODEs). We say an ODE is **linear** if

$$F \left[ x, y(x), y'(x), \dots, y^{(n)}(x) \right]$$

depends on  $y, y', \dots, y^{(n)}$  linearly. Note that we allow  $F$  to depend on  $x$  nonlinearly.

An **nonlinear** ODE is one that is not linear.

Note that we may always represent a linear ODE as

$$a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \dots + a_1(x)y'(x) + a_0(x)y(x) = b(x),$$

where each  $a_i(x)$  can be nonlinear in  $x$ .

In general, linear ODEs are fully solvable by hand. For nonlinear ODEs, we may only be able to solve some special cases.

*Problem 2.6.*

$$\frac{d}{dt} y = c,$$

where  $c$  is a constant. Integrating, we get  $y = ct + b$ , where  $b$  is an arbitrary constant.

$$\frac{d^2}{dt^2} y = 0.$$

Integrating twice, we get  $y = at + b$ , where  $a, b$  are arbitrary constants.

These (non unique) are called **general solutions**. We thus sometimes prescribe also an initial condition (IC) to further determine the solution. An example of an IC for the second order ODE above is

$$y(0) = y_0, \quad y'(0) = v_0,$$

In these examples, note in particular that we have uniqueness results given the ICs.

## 2.1 First Order Linear ODEs

We may represent a first order linear ODE as

$$a_1(x)y'(x) + a_0(x)y(x) = b(x).$$

The general method is to rewrite the ODE as

$$\frac{d}{dt}[y(t)] = f(t)$$

and integrate both sides.

When  $a_1 \neq 0$ , we can rewrite the ODE as

$$y'(x) + p(x)y(x) = g(x),$$

with  $p, g$  given. The particular case  $b = 0$  ( $g = 0$ ) is considered first:

### 2.1.1 The case $b = 0$

*Problem 2.7.* Consider

$$\frac{d}{dt}y(t) = a(t)y(t). \tag{1}$$

Assuming  $y(t) \neq 0$ , we may rewrite this as

$$a(t) = \frac{1}{y(t)} \frac{d}{dt}y(t) = \frac{d}{dt}[\log |y(t)|].$$

Integrating, we get

$$\log |y(t)| = \int a(t) dt + C,$$

and so

$$y(t) = \pm e^C \exp\left(\int a(t) dt\right) = C' \exp\left(\int a(t) dt\right),$$

where  $C'$  is an arbitrary constant (the case  $C' = 0$  is attained when  $y = 0$ ).

### 2.1.2 The Integrating Factor

*Problem 2.8.* Consider

$$y'(x) + p(x)y(x) = g(x). \quad (2)$$

Observe that for each  $\mu(t) \neq 0$ , the ODE is equivalent to

$$\mu y' + \mu p y = \mu g.$$

Let's guess that the left hand side can be written as  $\frac{d}{dt}[a(t)y(t)]$  for some  $a$ . It follows that

$$\frac{d}{dt}[a(t)y(t)] = a'(t)y(t) + a(t)y'(t) = \mu y' + \mu p y \implies \begin{cases} a = \mu, \\ \mu' = \mu p. \end{cases}$$

The function  $\mu$  is called the **integrating factor**. It suffices to find one  $\mu$  such that  $\mu' = \mu p$ . A  $\mu$  is given by the previous case:

$$\mu(t) = \exp\left(\int_{t_0}^t p(s) \, ds\right).$$

We now reduced the ODE to

$$\frac{d}{dt}[\mu(t)y(t)] = \mu(t)g(t),$$

which can be solved by integrating and dividing by  $\mu$ :

$$y(t) = \frac{1}{\mu(t)} \left( \int_{t_0}^t \mu(s)g(s) \, ds + C \right).$$

*Example 2.9.*

$$y' + y = e^t.$$

We seek a  $\mu$  such that

$$\frac{d}{dt}[\mu y] = \mu y' + \mu y = \mu e^t.$$

This gives

$$\begin{cases} \mu' = \mu, \\ \mu = e^t. \end{cases}$$

Using this choice of  $\mu$  we rewrite the ODE as

$$\frac{d}{dt}[e^{ty}] = e^t e^t = e^{2t}$$

$$e^{ty} = \frac{1}{2}e^{2t} + C,$$

from which  $y = \frac{1}{2}e^t + ce^{-t}$ .



Determining C: Suppose we are given

$$y(t_0) = y_0.$$

Then

$$\mu y(t) = \int_{t_0}^t \mu g(s) \, ds + c.$$

Taking  $t = t_0$ , we get

$$\mu y(t_0) = c$$

Since (given our choice of  $\mu$ )

$$\mu(t_0) = 1,$$

we have  $C = y(t_0) = y_0$ , which we can plug back in the general solutions obtained above.

### 3 Separation of Variables

Recall that first order ODEs can be represented as

$$y'(t) + p(t)y(t) = g(t).$$

Using the implicit function theorem, we can in principle rewrite this as

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

and then

$$M(x, y) + N(x, y)y' = 0.$$

Question: for which  $M, N$  can we solve this ODE?

Recall that last lecture we solved

$$\frac{d}{dt}[y(t)] = g(t)$$

with  $y$  unknown.

A first special case (separation of variables) is when  $M = M(x)$  and  $N = N(y)$ :

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

**Proof (Formal Derivation).** If we formally treat  $dx$  and  $dy$  as differentials, we can rewrite the above as

$$N(y)dy = -M(x)dx.$$

In this view the variables  $x$  and  $y$  are separated. Integrating both sides, we obtain

$$\int N(y)dy = - \int M(x)dx + C.$$

If we can find  $n$  and  $m$  such that  $n' = N$  and  $m' = M$ , then we have

$$n(y) = -m(x) + C,$$

from which we can solve for  $y$ . □

**Proof (Rigorous Derivation).** We integrate over  $x$  to get

$$\int M(x) dx + \int N(y) \frac{dy}{dx} dx = 0.$$

With a change of variables we have

$$\int M(x) dx + \int N(y) dy = C.$$

□

### 3.0.1 Examples

*Example 3.1.*


$$x + y \frac{dy}{dx} = 0.$$

We have  $M = 1$  and  $N = y$ . Integrating, we get

$$\frac{x^2}{2} + C + \int y \frac{dy}{dx} dx = 0$$

and so

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

With additional initial conditions we can determine  $y(x)$ . 

*Example 3.2.*

$$y + e^x \frac{dy}{dx} = 0.$$

Dividing by  $ye^x$  (assuming  $y \neq 0$ ), we get

$$\frac{1}{y} \frac{dy}{dx} = -e^{-x}.$$

and then

$$-e^{-x} + C + \log|y| = 0.$$



More generally, suppose the dependence of  $M$  and  $N$  on  $(x, y)$  can be separated in the following sense:


$$M_1(x)M_2(y) + N_1(x)N_2(y) \frac{dy}{dx} = 0.$$

Again dividing both sides, we get

$$\frac{M_1}{N_1} + \frac{N_2}{M_2} \frac{dy}{dx} = 0.$$

*Example 3.3.*

$$e^{x+y} + xy \frac{dy}{dx} = 0.$$

Use above. 

Warning: this method does not work for the following ODE:

$$M_1(x)M_2(y) + Z_1(x)Z_2(y) + N_1(x)N_2(y) \frac{dy}{dx} = 0$$

### 3.1 Generalization

The method of integrating both sides cannot be pushed much further beyond the following case:

$$\frac{d}{dx} [\varphi(x, y(x))] = g(x).$$

Integration gives

$$\varphi(x, y(x)) = \int g(x) dx + c.$$

In principle we can solve for  $y$  by the implicit function theorem.

The ODE above can be written equivalently as

$$\frac{d}{dx} [\tilde{\varphi}(x, y(x))] := \frac{d}{dx} \left[ \varphi(x, y(x)) - \int g(x) dx \right] = 0.$$

Thus we can without loss of generality suppose  $g = 0$ . We turn next thus to the ODE

$$\frac{d}{dx} \varphi(x, y(x)) = 0.$$

For which  $M, N$  can we convert  $M(x, y) + N(x, y) \frac{dy}{dx} = 0$  to the above form?

$$\frac{d}{dx} \varphi(x, y(x)) = \partial_1 \varphi + \partial_2 \varphi \frac{dy}{dx}.$$

This implies

$$\begin{cases} M(x, y) = \partial_1 \varphi(x, y), \\ N(x, y) = \partial_2 \varphi(x, y). \end{cases}$$

**Definition 3.4.** We say  $M + Ny' = 0$  is an **exact equation** if there exists  $\varphi$  such that

$$M = \partial_1 \varphi, \quad N = \partial_2 \varphi.$$

What is the minimum requirement for  $M, N$  to be an exact equation? If  $\varphi$  is sufficiently smooth ( $C^2$ ), then we would expect

$$\partial_2 M = \partial_2 \partial_1 \varphi = \partial_1 \partial_2 \varphi = \partial_1 N.$$

This in fact is also sufficient:

**Theorem 3.5.** Suppose  $M, N, \partial_2 M, \partial_1 N$  are continuous in the box  $B = [a, b] \times [c, d]$  and  $(x, y) \in B$ . Then the equation  $M(x, y) + N(x, y)y' = 0$  is exact if and only if

$$\partial_2 M(x, y) = \partial_1 N(x, y).$$

That is, there exists  $\varphi$  such that


$$M = \partial_1 \varphi, \quad N = \partial_2 \varphi.$$

*Example 3.6.* The equation  $M(x) dx + N(y) dy = 0$  is exact, with  $\partial_2 M = 0 = \partial_1 N$ . But observe also that

$$\partial_2 [M(x) + y] = 1 = \partial_1 [N(y) + x].$$

So we can solve the ODE

$$(M(x) + y) + (N(y) + x) y' = 0,$$

which is not separable. 

**Proof.** That exactness implies  $\partial_2 M = \partial_1 N$  is easy and shown above.

It remains to prove that  $\partial_2 M = \partial_1 N$  implies exactness. To that end we construct  $\varphi$  as follows:

Step 1: construct  $\varphi$  so that  $\partial_1 \varphi = M(x, y)$ . We set

$$\varphi(x, y) = \int_{x_0}^x M(s, y) ds + h(y),$$

where  $h$  is to be determined.

Step 2: determine  $h$  so that  $\partial_2 \varphi = N(x, y)$ . Note that

$$\begin{aligned} \partial_2 \varphi(x, y) &= \int_{x_0}^x \partial_2 M(s, y) ds + h'(y) \\ &= \int_{x_0}^x \partial_1 N(s, y) ds + h'(y) \\ &= N(x, y) - N(x_0, y) + h'(y), \end{aligned}$$

from which we can specify  $h$  by

$$h(y) = \int_{y_0}^y N(x_0, s) ds + C.$$

In sum,  $\varphi$  is given by

$$\varphi(x, y) = \int_{x_0}^x M(s, y) ds + \int_{y_0}^y N(x_0, s) ds + C.$$

This completes the proof. □

*Example 3.7.*

$$y^2 x + (1 + x^2 y) y' = 0.$$

We have

$$\partial_2 M = 2xy = \partial_1 N,$$

and may thus set

$$\varphi(x, y) = \int y^2 x dx = \frac{x^2 y^2}{2} + h(y).$$

$$\partial_2 \varphi = x^2 y + h'(y) = 1 + x^2 y \implies h(y) = y + C.$$

Finally, we can rewrite the original ODE as

$$\frac{d}{dx} \left[ \frac{x^2 y^2}{2} + y \right] = 0 \implies \frac{x^2 y^2}{2} + y = C,$$

with

$$\varphi(x, y) = \frac{x^2 y^2}{2} + y + C'.$$

Suppose we have the IC  $y(0) = 1$ . Then

$$C = \frac{0^2 1^2}{2} + 1 = 1.$$



### 3.2

We consider further ODEs of the form

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

We can equivalently write this as

$$\mu M + \mu N \frac{dy}{dx} = 0$$

for some  $\mu \neq 0$ . This is exact when

$$\partial_2(\mu M) = \partial_1(\mu N).$$

The goal, thus, is to find  $\mu$  such that the above is true, when the original ODE might not be exact. If  $\mu(x, y) = \mu(x)$  or  $\mu(x, y) = \mu(y)$ , then we need not deal with mixed partials.

Let's begin with  $\mu = \mu(x)$ : We would like to solve

$$\partial_2(\mu M) = \mu \partial_2 M = \mu' N + \mu \partial_1 N = \partial_1(\mu N),$$

or equivalently

$$\frac{\mu'}{\mu} = \frac{\partial_2 M - \partial_1 N}{N}.$$

This approach works when the right hand side is a function of  $x$  only.

A similar condition can be derived for  $\mu = \mu(y)$ .

## 4 Second Order Linear ODEs

A second order linear ODE can be written as

$$F(t, y, y', y'') = 0,$$

and by the inverse function theorem, as

$$y'' = f(t, y, y').$$

**Definition 4.1.** We say this ODE is **linear** if  $F$  depends on  $y, y', y''$  linearly (note again that we do not require linearity in  $t$ ). Thus a second order linear ODE can be written as


$$P(t)y'' + Q(t)y' + R(t)y = G(t).$$

In the case that  $P(t) \neq 0$ , we can rewrite this as


$$y'' + p(t)y' + q(t)y = g(t),$$

*Example 4.2.*  $y'' = 0$ . The general solution is  $y(t) = c_1t + c_2$ . We need 2 ICs to determine  $c_1, c_2$ , for example

$$\begin{cases} y(t_0) = y_0 \\ y'(t_0) = v_0 \end{cases}.$$

In general, for an  $n^{\text{th}}$  order ODE we need  $n$  ICs to determine a unique solution. 

**Definition 4.3.** We say the ODE is **homogeneous** if  $G = 0$  and **nonhomogeneous** otherwise.

*Remark 4.4* (Property of homogeneous ODEs). If  $y$  solves  $y'' + p(t)y' + q(t)y = 0$ , then  $ay$  solves the same ODE for any  $a \in \mathbb{Z}$ . 

We start with the homogeneous case.

### 4.1 Homogeneous Second Order Linear ODEs with Constant Coefficients

$$ay'' + by' + cy = 0, \quad a, b, c \in \mathbb{R}.$$

#### 4.1.1 The Ansatz of Polynomials

We assume first that  $y(t) = \sum_{j=0}^n a_j t^j$ . Plugging this into the ODE, we get terms involving  $t^n$  which cannot be canceled.

#### 4.1.2 Recall

If  $a \equiv 0$ , then this reduces to  $by' + cy = 0$ . This can be written as one of the following:

$$y' + \frac{c}{b}y = 0, \quad b \frac{y'}{y} + c = 0.$$

And in either case we will get  $y(t) = e^{-\frac{c}{b}t} \cdot c_0$ .

### 4.1.3 The Ansatz of Exponentials

Inspired by the first order case, we now try the ansatz  $y(t) = c_0 e^{\lambda t}$ . Plugging into the ODE, we get

$$c_0 [a\lambda^2 e^{\lambda t} + b\lambda e^{\lambda t} + c e^{\lambda t}] = 0,$$

which reduces the original ODE to the following:

$$a\lambda^2 + b\lambda + c = 0.$$

### 4.1.4 The Operator $L$

Define the operator  $L$  as

$$(Ly)(t) = P(t)y'' + Q(t)y' + R(t)y.$$

*Example 4.5.* For any constants  $c_1, c_2$  and functions  $y_1, y_2$ . Note that

$$\begin{aligned} L[c_1 y_1 + c_2 y_2] &= P(t)[c_1 y_1 + c_2 y_2]'' + Q(t)[c_1 y_1 + c_2 y_2]' + R(t)[c_1 y_1 + c_2 y_2] \\ &= c_1 L[y_1] + c_2 L[y_2], \end{aligned}$$

and so the operator  $L$  is linear. 

A solution  $y$  to the ODE  $P(t)y'' + Q(t)y' + R(t)y = 0$  then can equivalently be written as  $Ly = 0$ . Now note that by linearity, we have if  $Ly_1 = Ly_2 = 0$  for two “different solutions”  $y_1$  and  $y_2$ , then since

$$L[c_1 y_1 + c_2 y_2] = c_1 Ly_1 + c_2 Ly_2 = 0,$$

the general solution can be obtained as

$$y = c_1 y_1 + c_2 y_2.$$


This technique of obtaining the general solution is called **linear superposition**. It turns out that the correct notion of solutions being “different” is linear independence.

*Example 4.6.*

$$y'' - 5y' + 6y = 0.$$

We can solve  $\lambda^2 - 5\lambda + 6 = 0$  to get

$$\lambda_1 = 2, \quad \lambda_2 = 3.$$

Thus the first solution is  $y_1 = e^{2t}$  and the second solution is  $y_2 = e^{3t}$ . 

#### 4.1.5 Three Cases of Obtaining the General Solution

The solution of the characteristic polynomial can be classified into three cases:

- (i) Two real roots. In this case the general solution is

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

- (ii) Two different complex roots that are complex conjugates  $\lambda \pm i\mu$  (since all coefficients are real). Recall that

$$e^z := \sum_{k \geq 0} \frac{z^k}{k!}, \quad z \in \mathbb{C}, \quad e^{i\theta} = \cos \theta + i \sin \theta, \quad \theta \in \mathbb{R}.$$

Thus noting that

$$\begin{aligned} e^{(\lambda+i\mu)t} &= e^{\lambda t} [\cos \mu t + i \sin \mu t] \\ e^{(\lambda-i\mu)t} &= e^{\lambda t} [\cos \mu t - i \sin \mu t], \end{aligned}$$

we see two ways to obtain a real solution:


- Choose  $c_1 = c_2 \in \mathbb{R}$  to get a multiple of  $y(t) = e^{\lambda t} \cos \mu t$ .
- Choose  $c_1 = -c_2 \in i\mathbb{R}$  to get a multiple of  $y(t) = e^{\lambda t} \sin \mu t$ .

The general solution is then a linear combination of the above two:

$$y(t) = e^{\lambda t} [c_1 \cos \mu t + c_2 \sin \mu t], \quad c_1, c_2 \in \mathbb{R}.$$

*Example 4.7.*

$$y'' + y = 0.$$

Solving the characteristic polynomial gives  $\lambda_1 = i$  and  $\lambda_2 = -i$ . Two real solutions are  $y_1(t) = \cos t$  and  $y_2(t) = \sin t$ . 

- (iii) One real root with a multiplicity two. For the characteristic polynomial  $a\lambda^2 + b\lambda + c = 0$ , we have  $\lambda = \lambda_1 = \lambda_2 = -b/2a$  and  $4ac = b^2$ . This gives a solution  $y(t) = e^{\lambda t}$ . We seek another solution  $y_2$  using the so called **reduction of order** method. We try the ansatz  $y_2(t) = \mu(t)y_1(t)$ .

*Claim 4.8.*  $\mu$  solves a first order ODE.

**Proof.** Suppose  $y_1$  solves  $P y_1'' + Q y_1' + R y_1 = 0$ . If  $y_2 = \mu y_1$  satisfies the same ODE, then

$$P(\mu y_1)'' + Q(\mu y_1)' + R(\mu y_1) = 0.$$

In general after expanding the left hand side, we get  $\sum_{i=0}^2 a_i(t) \mu^{(i)}(t) = 0$ . We will show  $a_0 = 0$ . Expanding, we get

$$P [\mu'' y_1 + 2\mu' y_1' + \mu y_1''] + Q [\mu' y_1 + \mu y_1'] + R \mu y_1 = 0.$$

Note that the  $\mu$ -terms sum to  $\mu [Py_1'' + Qy_1' + Ry_1] = 0$ . Thus  $\mu$  solves the ODE involving  $\mu'$  and  $\mu''$

$$\mu'' Py_1 + \mu' [2Py_1' + Qy_1] = 0.$$

This can be solved by separation of variables:

$$\frac{\mu''}{\mu'} = -\frac{2Py_1' + Qy_1}{Py_1}.$$

□

*Example 4.9.* Suppose  $P \equiv a$ ,  $Q \equiv b$ , and  $R \equiv c$ . We have  $y_1 = e^{\lambda t}$  where  $\lambda := -b/2a$ . We have that  $\mu$  defined above solved

$$\mu'' ae^{\lambda t} + \mu' [2a(e^{\lambda t})' + be^{\lambda t}] = 0.$$

The term in the bracket evaluates to 0 by  $2a\lambda + b = 0$ . Thus  $\mu'' = 0$  and so  $\mu(t) = t + C$ . Thus the general solution is

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



## 4.2 Series Solution to Homogeneous Second Order Linear ODEs

Consider the ODE

$$Py'' + Qy' + Ry = 0.$$

We will use the ansatz  $y(x) = \sum_{n \geq 0} a_n (x - x_0)^n$ .

*Remark 4.10.* Recall the following facts about power series  $\sum_{n \geq 0} a_n (x - x_0)^n$ :

- The root test for convergence: Let  $\mu := \limsup_{n \rightarrow \infty} |a_n|^{1/n}$ . If  $|x - x_0| < \mu^{-1}$ , then the power series converges absolutely.
- Within the radius of convergence, we can differentiate and integrate the power series term by term.

$$y'(x) = \sum_{n \geq 1} n a_n (x - x_0)^{n-1}.$$

Similarly we can compute the  $k^{\text{th}}$  derivative:

$$y^{(k)}(x) = \sum_{n \geq k} n(n-1) \dots (n-k+1) a_n (x - x_0)^{n-k}.$$

The resulting power series have the same radius of convergence.

- We say  $y$  is analytic in  $B(x_0, \mu^{-1})$ .



*Example 4.11* (Airy's Equation). We seek the solution to the ODE  $y'' - xy = 0$  near  $x_0 = 0$ .

$$\begin{aligned} y &= \sum_{n \geq 0} a_n x^n, \\ y'' &= \sum_{n \geq 2} a_n n(n-1) x^{n-2} = \sum_{n \geq 0} a_{n+2} (n+2)(n+1) x^n, \\ xy &= \sum_{n \geq 0} a_n x^{n+1} = \sum_{n \geq 1} a_{n-1} x^n. \end{aligned}$$

We collect terms to get

$$\begin{aligned} x^0 : & \quad a_2 \cdot 2 \cdot 1 = 0 \\ x_n, n \geq 1 : & \quad a_{n+2} (n+2)(n+1) - a_{n-1} = 0. \end{aligned}$$

This gives  $a_2 = 0$  and

$$a_{m+3} = \frac{a_m}{(m+3)(m+2)}, \quad m := n-1 \geq 0$$

Equivalently, for  $i = 0, 1, 2$ , with  $m+3 = 3k+i$  we have

$$a_{3k+i} = \frac{a_{3k+i-3}}{(3k+i)(3k+i-1)}.$$

This can be solved using iterative substitution (See remark below). With

$$b_k = a_{3k+i}, \quad c_k = \frac{1}{(3k+i)(3k+i-1)},$$

we get


$$a_{3k+i} = a_i \prod_{j=1}^k \frac{1}{(3j+i)(3j+i-1)}$$

with  $a_2 = 0$  and  $a_0, a_1$  free. Thus the general solution is

$$y(x) = a_0 \sum_{i \geq 0} A_i x^{3i} + a_1 \sum_{k \geq 0} B_k x^{3k+1},$$

where

$$A_k = \prod_{j=1}^k \frac{1}{(3j)(3j-1)}, \quad B_k = \prod_{j=1}^k \frac{1}{(3j+1)(3j)}.$$

Note in particular that  $|A_k|, |B_k| \leq 1$  for each  $k$ , and thus  $\mu := \limsup[\cdot]^{1/n} \leq 1$ . Thus the solution is analytic on a neighborhood of 0 with radius of convergence at least 1. 

**Remark 4.12.** Suppose  $b_k = c_k b_{k-1}$  for  $k \geq 1$  and  $c_k$  and  $b_0$  are given. Then we have the following solution by iterative substitution:

$$\begin{aligned} b_k &= c_k [c_{k-1} b_{k-2}] = c_k c_{k-1} [c_{k-2} b_{k-3}] = \cdots \\ &= c_k c_{k-1} \cdots c_1 b_0 = b_0 \prod_{i=1}^k c_i. \end{aligned}$$



**Example 4.13** (Airy's Equation, around  $x_0 = 1$ ). We seek the solution to the ODE  $y'' - xy = 0$  near  $x_0 = 1$ .


$$\begin{aligned} y &= \sum_{n \geq 0} a_n (x-1)^n, \\ y'' &= \sum_{n \geq 2} a_n n(n-1) (x-1)^{n-2} = \sum_{n \geq 0} a_{n+2} (n+2)(n+1) (x-1)^n, \\ xy &= (x-1+1)y \\ &= \sum_{n \geq 0} a_n (x-1)^{n+1} + \sum_{n \geq 0} a_n (x-1)^n = \sum_{n \geq 1} a_{n-1} (x-1)^n + \sum_{n \geq 0} a_n (x-1)^n. \end{aligned}$$


We collect terms:

$$\begin{aligned} (x-1)^0 : & \quad a_2 \cdot 2 \cdot 1 - a_0 = 0 \\ (x-1)^n, n \geq 1 : & \quad a_{n+2} (n+2)(n+1) - a_{n-1} - a_n = 0. \end{aligned}$$

This gives  $a_2 = a_0/2$  and

$$a_3 = \frac{a_0 + a_2}{3 \cdot 2}, \quad a_4 = \frac{a_2 + a_1}{4 \cdot 3} = \frac{a_0}{2 \cdot 3 \cdot 4} + \frac{a_1}{3 \cdot 4}, \quad \dots$$

In this case it is hard to obtain a closed form for  $a_n$ . We note that the general solution is determined by  $a_0, a_1$ . 

**Remark 4.14.** This power series is useful for numerical approximation of the solution. We can truncate the series at some  $N$  and use the first  $N$  terms to approximate the solution. 

**Theorem 4.15** (5.3.1, BDM 9<sup>th</sup> edition). Consider  $P(x)y'' + Q(x)y' + R(x)y = 0$ , where we assume  $P, Q, R$  are analytic near  $x_0$  with convergence radius  $R$  and write

$$P(x) = \sum P_j (x-x_0)^j, \forall |x-x_0| < R$$

and similarly for  $Q$  and  $R$ . If  $P(x_0) \neq 0$  in the ball  $B(x_0, R)$ , we can consider

$$y'' + \frac{Q}{P}y' + \frac{R}{P}y = 0.$$

Then there exists a power series solution  $y$  to the ODE of the form

$$y(x) = \sum_{n \geq 0} a_n (x-x_0)^n = a_0 y_1 + a_1 y_2$$

with convergence radius at least  $R$ .

**Proof.** Omitted. □

*Example 4.16.* Examples of analytic functions:

- Polynomials.  $R = +\infty$ .
- $e^x$ .  $R = +\infty$ .
- $\log(1+x)$ .
- $\sin x, \cos x$ .

Examples of non-analytic functions:

- Any non-smooth function. E.g.,  $|x|^{1/2}$ .



### 4.3 Non-Homogeneous Second Order Linear ODEs

$$Py'' + Qy' + Ry = G.$$

Observation: if  $\varphi$  and  $\psi$  are two solutions to the ODE above, then  $\varphi - \psi$  solves the corresponding homogeneous ODE  $Py'' + Qy' + Ry = 0$ .

**Proposition 4.17.** *Suppose that  $y_0$  solves  $Py'' + Qy' + Ry = G$  and  $y_1, y_2$  are two different solutions to  $Py'' + Qy' + Ry = 0$ . Then the general solution to the non-homogeneous ODE is*

$$y = y_0 + c_1y_1 + c_2y_2, \quad \forall c_1, c_2 \in \mathbb{R}.$$

*That is, the general solution is a particular solution plus the general solution to the corresponding homogeneous ODE.*

With this in mind, we see that we need only find one solution  $y_0$  to the non-homogeneous ODE.

#### 4.3.1 Variation of Parameters / Constants

Assume that  $y_1, y_2$  are two different solutions to the corresponding homogeneous ODE  $y'' + py' + qy = 0$ . Recall that the goal is a particular solution  $y_0$  to the non-homogeneous ODE  $y'' + py' + qy = g$ . We try the ansatz

$$y_0(t) = \mu_1(t)y_1(t) + \mu_2(t)y_2(t),$$

where functions  $\mu_1, \mu_2$  are to be determined. Plugging into the  $y'' + py' + qy$ , we will get in general terms involving  $\mu_i, \mu'_i, \mu''_i$ . We will select  $\mu_i$  in a way that the  $\mu_i$  and  $\mu''_i$  terms vanish.

Note that

$$y' = \mu'_1y_1 + \mu_1y'_1 + \mu'_2y_2 + \mu_2y'_2 = \mu_1y'_1 + \mu_2y'_2,$$

where the last equality results after we *impose the restriction*  $\mu'_1 y_1 + \mu'_2 y_2 = 0$ . Now,

$$y'' = \mu'_1 y'_1 + \mu'_2 y'_2 + \mu_1 y''_1 + \mu_2 y''_2$$

and so

$$y'' + py' + qy = \mu'_1 y'_1 + \mu'_2 y'_2 + \mu_1 (y''_1 + py'_1 + qy_1) + \mu_2 (y''_2 + py'_2 + qy_2),$$

where  $y_1$  and  $y_2$  solve the homogeneous ODE, and so the last two terms vanish. If we set

$$\begin{cases} \mu'_1 y_1 + \mu'_2 y_2 = 0, \\ \mu'_1 y'_1 + \mu'_2 y'_2 = g, \end{cases} \iff \begin{pmatrix} y_1 & y_2 \\ y'_1 & y'_2 \end{pmatrix} \begin{pmatrix} \mu'_1 \\ \mu'_2 \end{pmatrix} = \begin{pmatrix} 0 \\ g \end{pmatrix}$$

we get a  $2 \times 2$  linear system for  $\mu'_1, \mu'_2$ . This is solvable if and only the matrix is invertible:

$$\begin{pmatrix} \mu_2 \\ \mu_2 \end{pmatrix} (x) = \int_{x_0}^x A^{-1}(t) \begin{pmatrix} 0 \\ g(t) \end{pmatrix} dt.$$

## 5 First order ODE system

### 5.1 Motivation and Setup

The ODE

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

can be rewritten as a first order system of ODEs. Write for each  $j$ ,  $x_j(t) := y^{(j-1)}(t)$  so that we have

$$y^{(j)} = \frac{d}{dt} x_j.$$

Now the original ODE can be rewritten as

$$a_n x'_n(t) + a_{n-1} x_n(t) + \cdots + a_1 x_2(t) + a_0 x_1(t) = g(t).$$

This gives the system

$$\begin{cases} x'_1 = x_2 \\ x'_2 = x_3 \\ \vdots \\ x'_{n-1} = x_n \\ x'_n = \frac{1}{a_n} [g(t) - a_{n-1}x_n - \cdots - a_1x_2 - a_0x_1] \end{cases}.$$

This can be summarized as

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \cdots & -\frac{a_{n-1}}{a_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{g(t)}{a_n} \end{pmatrix}.$$

Let's for now abstract away from the above construction and consider the general first order system. With  $F_1, \dots, F_n$  given and  $x_1(t), \dots, x_n(t)$  unknown, we consider the system

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} F_1(t, x_1, \dots, x_n) \\ F_2(t, x_1, \dots, x_n) \\ \vdots \\ F_n(t, x_1, \dots, x_n) \end{pmatrix},$$


which can be summarized as

$$\frac{d}{dt} \mathbf{x} = \mathbf{F}(t, \mathbf{x}).$$

This is a linear system if we can write

$$\mathbf{F}(t, \mathbf{x}) = A(t)\mathbf{x} + \mathbf{b}(t)$$

for  $A(t) \in \mathbb{R}^{n \times n}$  and  $\mathbf{b}(t) \in \mathbb{R}^{n \times 1}$ .

*Example 5.1.* The first order ODE system that arises from the  $n^{\text{th}}$  order ODE is linear. 

**Definition 5.2.** We say the system  $\mathbf{F}(t, \mathbf{x}) = A(t)\mathbf{x} + \mathbf{b}(t)$  is **homogeneous** if  $\mathbf{b} \equiv 0$  and **non-homogeneous** otherwise.

*Example 5.3.*

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

Let  $x_1 := y$  and  $x_2 := y'$ . Then we have

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ g \end{pmatrix}.$$



Initial conditions for the first order system can be written as


$$\mathbf{x}(t_0) = \mathbf{x}_0 \in \mathbb{R}^{n \times 1}.$$

**Theorem 5.4** (Existence and Uniqueness). *Suppose  $F_j$  is continuous and  $|\partial_{x_i} F_j| \leq M$  (Lipschitz) for  $t \in (a_0, b_0) =: I_0$  and  $x_i \in (a_i, b_i) =: I_i$ . Then for each  $t_0 \in I_0$ ,  $x_{0,i} \in I_i$  there exists  $\delta > 0$  and a unique solution  $\mathbf{x}(t)$  to the ODE system on  $(t_0 - \delta, t_0 + \delta)$  with  $\mathbf{x}(t_0) = \mathbf{x}_0$  such that  $|\mathbf{x}(t) - \mathbf{x}_0| < \delta$ .*

Note that for the special case

$$\frac{d}{dt} \mathbf{x}(t) = A(t)\mathbf{x} + \mathbf{b}, \quad \mathbf{x}, \mathbf{b} \in \mathbb{R}^{n \times 1}, \quad A \in \mathbb{R}^{n \times n},$$

it is important to restrict  $A$  and  $\mathbf{b}$  to be real, since complex numbers can roughly be identified as two real numbers, and in those cases solutions may not be unique.

*Example 5.5.* Consider  $F = A(\mathbf{x})\mathbf{x} + \mathbf{b}$ . If  $|A| \leq C$ , then there exists unique solution in a small neighborhood. 

## 5.2 Solving the Homogeneous First Order Linear ODE System

Consider

$$\frac{d}{dt} \mathbf{x}(t) = A(t)\mathbf{x}(t)$$

Suppose  $\mathbf{x}^{(1)} = (x_1^{(1)}, \dots, x_n^{(1)})'$ ,  $\dots$ ,  $\mathbf{x}^{(j)} = (x_1^{(j)}, \dots, x_n^{(j)})'$  are  $j$  solutions to the ODE. Then by linearity, so are  $\sum_j c_j \mathbf{x}^{(j)}$ .

How do we differentiate different solutions?

**Definition 5.6.** We say  $\mathbf{x}^1, \dots, \mathbf{x}^n$  are linear independent if the **Wronski matrix** is not singular, i.e., if the **Wronskian**

$$W(t) := \det \begin{pmatrix} x^1 & \dots & x^n \end{pmatrix}$$

is not identically zero.

Note that if  $W(t_0) = 0$ , then  $x^{(j_0)}(t_0) = \sum_{j \neq j_0} c_j x^j(t_0)$ .

**Theorem 5.7.**

$$\frac{d}{dt} W(t) = \text{tr}(A(t)) W(t).$$

**Proof.** Write

$$M(t) = \begin{pmatrix} x^1(t) & x^2(t) & \dots & x^n(t) \end{pmatrix} = \begin{pmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \dots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \dots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \dots & x_n^{(n)}(t) \end{pmatrix} =: \begin{pmatrix} \mathbf{b}_1(t) \\ \mathbf{b}_2(t) \\ \vdots \\ \mathbf{b}_n(t) \end{pmatrix}.$$

We claim that

$$\frac{d}{dt} \det \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix} = \det \begin{pmatrix} b'_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} + \det \begin{pmatrix} b_1 \\ b'_2 \\ \dots \\ b_n \end{pmatrix} + \dots + \det \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b'_n \end{pmatrix}.$$

To see this we recall

$$\frac{d}{dt} (C_1 \dots C_n) = \sum C_1 \dots C_{k-1} C'_k C_{k+1} \dots C_n$$

and

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum a_{1i_1} a_{2i_2} \dots a_{ni_n} \text{sgn}(\sigma).$$

Now,

$$\frac{d}{dt} W(t) = \frac{d}{dt} \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix} = \sum_{k=1}^n \det \begin{pmatrix} b_1 \\ \dots \\ b'_k \\ \dots \\ b_n \end{pmatrix}.$$

Recalling  $\frac{d}{dt} x^i = Ax^i$  for each  $i$  and fixing  $k$ , we have  $\frac{d}{dt} x_k^i = [Ax^i]_k = \sum_j A_{kj} x_j^i$  for each  $i$ . Thus by stacking we have

$$\frac{d}{dt} b_k = \frac{d}{dt} \begin{pmatrix} X_k^{(1)} & \dots & X_k^{(n)} \end{pmatrix} = \sum_j A_{kj} \begin{pmatrix} X_j^{(1)} & \dots & X_j^{(n)} \end{pmatrix} = \sum_j A_{kj} b_j.$$

Now,

$$\begin{aligned} \frac{d}{dt} W(t) &= \sum_{k=1}^n \det \begin{pmatrix} b_1 \\ \dots \\ b'_k \\ \dots \\ b_n \end{pmatrix} = \sum_{k=1}^n \det \begin{pmatrix} b_1 \\ \dots \\ \sum_j A_{kj} b_j \\ \dots \\ b_n \end{pmatrix} \\ &= \sum_{k=1}^n \det \begin{pmatrix} b_1 \\ \dots \\ A_{kk} b_k \\ \dots \\ b_n \end{pmatrix} = \sum_{k=1}^n A_{kk} \det \begin{pmatrix} b_1 \\ \dots \\ b_k \\ \dots \\ b_n \end{pmatrix} = \sum_k A_{kk} W(t). \end{aligned}$$

□

**Corollary 5.8.**

- $W(t_0) = 0$  for some  $t$  if and only if  $W(t) = 0$  for each  $t \in I$ .
- $W(t_0) \neq 0$  for some  $t_0 \in I$  if and only if  $W(t) \neq 0$  for each  $t \in I$ .

In particular, it suffices to check  $W(t_0)$  at any  $t_0$ .

**Proof.** Note that

$$W(t) = W(t_0) \exp \left( \int_{t_0}^t \text{tr}(A(s)) \, ds \right),$$

where the last term is never zero. Write  $\mu := \text{tr} \circ A$ . We have if  $\mu(t) \neq 0$  for each  $t$ , then  $W(t_0) = 0$  if and only if  $W(t) = 0$  for each  $t \in I$ .

$W(t_0) \neq 0$  if and only if  $W(t) \neq 0$ . □

**Theorem 5.9.** Suppose that  $x^1, \dots, x^n$  are  $n$  linearly independent real solutions to the homogeneous ODE system  $\frac{d}{dt}x = A(t)x$ . Then any solution  $x$  to the ODE can be written uniquely as

$$x(t) = \sum_{j=1}^n c_j x^{(j)}(t), \quad c_j \in \mathbb{R}.$$

In particular, this tells us that the space of solutions to the homogeneous ODE system (or any  $n^{\text{th}}$  order ODE) is an  $n$ -dimensional vector space.

**Proof.** For each  $c$  write

$$y_c(t) = x(t) - \sum_i c_i x^i(t).$$

By previous results,  $y_c$  solves the homogeneous ODE.

Now fix  $t_0$ , we seek  $c_i$  such that

$$(x^1(t_0) \quad x^2(t_0) \quad \dots \quad x^n(t_0)) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = x(t_0).$$

Since  $x^1, \dots, x^n$  are linearly independent, the Wronski matrix

$$M(t_0) := \begin{pmatrix} x^1(t_0) & x^2(t_0) & \dots & x^n(t_0) \end{pmatrix}$$

is invertible, and we can obtain  $\mathbf{c}$  as  $M^{-1}(t_0)x(t_0)$ .

In particular we have  $y_c(t_0) = 0$  and  $y_c$  solves the homogeneous ODE. Since 0 is also a solution to the homogeneous ODE, by uniqueness we have  $y_c(t) \equiv 0$ , which gives  $x(t) \equiv \sum c_i x^i$ .  $\square$

**Theorem 5.10.** Suppose that  $X^{(i)}$  solves  $\frac{d}{dt}X^{(i)} = A(t)X^{(i)}$  with the IC  $X^{(i)}(t_0) = \mathbf{e}_i$  for  $1 \leq i \leq n$ . Then  $X^{(i)}$  are linearly independent solutions.

**Proof.** The existence and uniqueness theorem guarantees that  $X^{(i)}$  exist and are unique. To check that they are independent, we need to check that


$$W(t) = \det \begin{pmatrix} X^{(1)}(t) & X^{(2)}(t) & \dots & X^{(n)}(t) \end{pmatrix}$$

is not identically zero. Recall that to do this it suffices to check  $W(t_0) \neq 0$ :

$$W(t_0) = \det I = 1 \neq 0.$$


But note of course that we can pick any  $n$  linearly independent initial conditions in  $\mathbb{R}^n$ .  $\square$

*Remark 5.11.* The preceding two theorems together implies that there are exactly  $n$  linearly independent solutions to the homogeneous ODE system.

Note that the theorem above also gives a method to find  $n$  linearly independent solutions. 

*Example 5.12.*

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

This is a second order ODE which can be rewritten as a  $2 \times 2$  first order ODE system. The results above imply that there are exactly two linearly independent solutions to the ODE. The linear independence turns out to be equivalent to  $y_1 \neq cy_2$  in this case. 

### 5.3 Finding Solutions Explicitly

Recall that for  $n = 1$ , we have that the ODE

$$\frac{d}{dt}x(t) = a(t)x(t), \quad \in \mathbb{R}$$

has solution

$$x(t) = x(t_0) \exp \left( \int_{t_0}^t a(s) ds \right).$$

In the special case that  $A(t)$  is diagonal with  $n \geq 2$ , we may easily generalize the result above:  $\frac{d}{dt}\mathbf{x} = A\mathbf{x}$  has solution

$$\frac{d}{dt}X_i = A_{ii}X_i, \quad 1 \leq i \leq n.$$

We say that the solution is **decoupled**.

For more general  $A(t)$ , we can only restrict to the special case  $A(t) \equiv A$  is constant. To see why further generalization is hard, consider the case  $n = 2$  and

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

Recall that we can only solve this explicitly if  $p, q$  are constant.

### 5.3.1 The case $A$ is constant

Recall that for  $n = 1$ , we know that the ODE

$$\frac{d}{dt}x(t) = ax(t), \quad a \in \mathbb{R}$$

has solution  $x(t) = ce^t$ .

We seek to generalize this to the case

$$\frac{d}{dt}\mathbf{x}(t) = A\mathbf{x}(t), \quad \mathbf{a} \in \mathbb{R}^{n \times n}.$$

The examples above for  $n = 1$  motivates the ansatz

$$\mathbf{x}(t) = e^{\lambda t} \mathbf{v}, \quad \mathbf{v} \in \mathbb{R}^{n \times 1}, \lambda \in \mathbb{R}.$$

Note that

$$\begin{aligned} \frac{d}{dt} \left( e^{\lambda t} \mathbf{v} \right) &= \lambda e^{\lambda t} \mathbf{v} \\ A(e^{\lambda t} \mathbf{v}) &= e^{\lambda t} A\mathbf{v}. \end{aligned}$$

Thus  $\mathbf{x}(t)$  solves the ODE system if and only if  $\lambda \mathbf{v} = A\mathbf{v}$ , or if and only if  $(\lambda, \mathbf{v})$  is an eigenvalue-eigenvector pair of  $A$ .

*Remark 5.13.*

- Special linear combinations of  $x_i(t)$  are solutions to a corresponding  $1 \times 1$  ODE.
- After projecting in the direction of the eigenvectors, the ODE system decouples.



We have the following cases:

- $A$  has  $n$  distinct real eigenvalues.
- $A$  has complex eigenvectors, but all eigenvalues are distinct.
- Repeated eigenvalues.
- $n$  linearly independent real eigenvectors.

The first three cases are disjoint, while the last case can overlap with the first three.

### 5.3.2 Case (i): A has n distinct real eigenvalues

Recall that when eigenvalues are real, so are the eigenvectors. Then we have the following  $n$  solutions:

$$\mathbf{x}^{(i)} = e^{\lambda_i t} \mathbf{v}_i, \quad 1 \leq i \leq n.$$

It can be shown that if  $\lambda_1, \dots, \lambda_n$  are distinct, then  $\mathbf{v}_1, \dots, \mathbf{v}_n$  are linearly independent. This gives  $W(0) = \det(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0$  and so the solutions above are linearly independent.

### 5.3.3 Case (iv): A has n linearly independent real eigenvectors

Note that case (iv) includes case (i). This is a strict subset:

*Example 5.14.*

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



Since  $\mathbf{v}_i$  are real, so are  $\lambda_i$ . We have

$$A (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n) = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \iff AP = P\Lambda.$$

We have  $\mathbf{v}_i$  are linearly independent if and only if  $P$  is invertible. In such case we have  $A = P\Lambda P^{-1}$ , or  $A$  is diagonalizable.

Recall the following:

**Theorem 5.15** (Symmetric Matrix). *If  $A = A^T \in \mathbb{R}^{n \times n}$ , then  $A = Q\Lambda Q^T$ , where  $Q \in \mathbb{R}^{n \times n}$ ,  $QQ^T = Q^T Q = I_n$ ,  $Q^{-1} = Q^T$ , and  $\Lambda$  is diagonal with real eigenvalues.*

Thus if  $A$  is symmetric, we can find  $n$  linearly independent solutions to the ODE system.

### 5.3.4 Case (ii): A has complex eigenvalues, but all eigenvalues are distinct

Note that if  $(\lambda, \mathbf{v})$  is an eigenvalue-eigenvector pair, then so is  $(\bar{\lambda}, \bar{\mathbf{v}})$ , since  $A\mathbf{v} = \lambda\mathbf{v} \iff A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$  (Note that this relies on  $A$  being real). We seek two real solutions from  $e^{\lambda t} \mathbf{v}$  and  $e^{\bar{\lambda} t} \bar{\mathbf{v}}$ . Suppose then that  $\lambda = \alpha + i\beta$  and  $\mathbf{v} = a + i\mathbf{b}$  for  $\alpha, \beta \in \mathbb{R}$ ,  $a, \mathbf{b} \in \mathbb{R}^{n \times 1}$ . Then

$$\begin{aligned} e^{\lambda t} \mathbf{v} &= e^{(\alpha+i\beta)t} (a + i\mathbf{b}) \\ &= e^{\alpha t} (\cos(\beta t) + i \sin(\beta t)) (a + i\mathbf{b}) \\ &= e^{\alpha t} [(\cos(\beta t)a - \sin(\beta t)\mathbf{b}) + i(\sin(\beta t)a + \cos(\beta t)\mathbf{b})] \end{aligned}$$

and similarly,

$$e^{\bar{\lambda} t} \bar{\mathbf{v}} = e^{\alpha t} [(\cos(\beta t)a - \sin(\beta t)\mathbf{b}) - i(\sin(\beta t)a + \cos(\beta t)\mathbf{b})].$$

From this we see that

$$\begin{aligned}\operatorname{Re} e^{\lambda t} v &= \frac{e^{\lambda t} v + e^{\bar{\lambda} t} \bar{v}}{2} = e^{\alpha t} (\cos(\beta t) a - \sin(\beta t) b) \\ \operatorname{Im} e^{\lambda t} v &= \frac{e^{\lambda t} v - e^{\bar{\lambda} t} \bar{v}}{2i} = e^{\alpha t} (\sin(\beta t) a + \cos(\beta t) b)\end{aligned}$$

are two real solutions to the ODE system.

Thus each complex eigenvalue-eigenvector pair gives two real solutions to the ODE system.

### 5.3.5 Case (iii): A has Repeated eigenvalues

*Example 5.16.*

$$\frac{d}{dt} \mathbf{x} = \begin{pmatrix} \lambda & 1 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix} \mathbf{x}.$$

The eigenvalue is  $\lambda$  with multiplicity 3. This matrix is not diagonalizable, since

$$(A - \lambda I) \mathbf{v} = \begin{pmatrix} & 1 \\ & & 1 \\ & & & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = 0 \iff v_2 = v_3 = 0,$$

and so each eigenvector is of the form  $\mathbf{v} = (v_1 \ 0 \ 0)^\top$ .

The methods discussed above thus does not work. Note, however, that  $x_3$  can be solved easily:  $x_3(t) = c_3 e^{\lambda t}$ . Then the restriction on  $x_2$  becomes  $x_2' = \lambda x_2 + c_3 e^{\lambda t}$ . Using the integrating factor  $e^{-\lambda t}$ , we have

$$\frac{d}{dt} (e^{-\lambda t} x_2) = c_3,$$

which gives  $x_2(t) = e^{\lambda t} (c_2 + c_3 t)$ . Finally, the restriction on  $x_1$  becomes  $x_1' = \lambda x_1 + e^{\lambda t} (c_2 + c_3 t)$ . Using the integrating factor  $e^{-\lambda t}$  again, we have

$$\frac{d}{dt} (e^{-\lambda t} x_1) = c_2 + c_3 t,$$

which gives  $x_1(t) = e^{\lambda t} \left( c_1 + c_2 t + \frac{c_3 t^2}{2} \right)$ . Thus the general solution to the ODE system is

$$\mathbf{x}(t) = e^{\lambda t} \begin{pmatrix} c_1 + c_2 t + \frac{c_3 t^2}{2} \\ c_2 + c_3 t \\ c_3 \end{pmatrix} = c_1 e^{\lambda t} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 e^{\lambda t} \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix} + c_3 e^{\lambda t} \begin{pmatrix} \frac{t^2}{2} \\ t \\ 1 \end{pmatrix}.$$



## 5.4 Matrix Exponential

Consider the  $n = 1$  case  $x' = ax$ ,  $a \in \mathbb{R}$ . We have solution  $x(t) = e^{at}x_0$ . We seek to generalize this to the case  $n \geq 2$  where

$$\mathbf{x}(t) = e^{At} \mathbf{x}_0.$$

But of course we need first to make sense of  $e^{At}$  for  $A \in \mathbb{R}^{n \times n}$ . The hope is that a definition will be consistent with  $(e^{At})' = Ae^{At}$ .

**Definition 5.17.**

$$\exp[A] := \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathbb{R}^{n \times n},$$

if the series converges.

*Example 5.18.*

- If  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ , we have

$$\exp[A] = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$


- $A = \begin{pmatrix} & \beta \\ -\beta & \end{pmatrix} = \beta J$ , where  $J = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ . We have  $A^k = \beta^k J^k$ . Note that  $J^2 = -I$ ,  $J^3 = -J$ ,  $J^4 = I$ . Thus we have

$$A^{4k+i} = \beta^{4k+i} J^i, \quad i = 0, 1, 2, 3,$$

and then

$$\exp[A] = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix}.$$



*Remark 5.19.* In general  $AB \neq BA$ . If however  $A = P^a$  and  $B = P^b$ , then  $AB = BA$ . 

For  $\exp[A]$  to be well-defined, we require  $\sum_{k \geq 0} (A^k)_{ij}/k!$  to converge for each  $1 \leq i, j \leq n$ . Let's suppose  $\max_{i,j} |A_{ij}| \leq a$ .

**Proposition 5.20.**

$$|(A^k)_{ij}| \leq (na)^{k-1} a.$$

**Proof.** We use induction. The case  $k = 1$  is clear. Now suppose the condition holds for  $k \geq 1$ . From  $(A^{k+1})_{ij} = (A^k A)_{ij} = \sum_{l=1}^n (A^k)_{il} A_{lj}$  we have

$$|(A^{k+1})_{ij}| \leq \sum_{l=1}^n |(A^k)_{il}| |A_{lj}| \leq \sum_{l=1}^n (na)^{k-1} a^2 = (na)^k a.$$

□

In light of the proposition above, we have for each  $N$ .

$$\sum_{k=0}^N \frac{|(A^k)_{ij}|}{k!} \leq \sum_{k=0}^{\infty} \frac{(na)^{k-1}a}{k!} = \sum_{k=0}^{\infty} \frac{1}{n} \frac{(na)^k}{k!} = \frac{e^{na}}{n} < \infty$$

Thus  $\exp[A]$  is well-defined for all  $A \in \mathbb{R}^{n \times n}$ .

**Lemma 5.21.**

- (i) If  $B = T^{-1}AT$ , then  $\exp[B] = T^{-1} \exp[A]T$ .
- (ii) If  $AB = BA$ , then  $\exp[A + B] = \exp[A] \cdot \exp[B]$ .
- (iii)  $(\exp[A])^{-1} = \exp[-A]$ .

**Proof.**

- (i) Note only that  $B^k = T^{-1}A^kT$ .
- (ii) Note that

$$\begin{aligned} \exp[A + B] &= \sum_{k=0}^{\infty} \frac{(A + B)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \left( \sum_j \binom{k}{j} A^j B^{k-j} \right) \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{A^j B^{k-j}}{j!(k-j)!} = \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} \frac{A^j}{j!} \frac{B^{k-j}}{(k-j)!} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^j}{j!} \frac{B^k}{k!} \\ &= \exp[A] \exp[B]. \end{aligned}$$

- (iii) Applying property (ii) with  $B = -A$  gives  $\exp[A] \exp[-A] = \exp[0] = I$ .  
Similarly,  $\exp[-A] \exp[A] = I$ .

□

**Proposition 5.22.** Suppose  $Av = \lambda v$ . Then  $\exp[A]v = e^{\lambda}v$ .

**Proof.** Use the fact that  $A^k v = \lambda^k v$ .

□

**Proposition 5.23.**

$$\frac{d}{dt} \exp[tA] = A \exp[tA] = \exp[tA]A.$$

**Proof.** A previous calculation shows that if  $\max |(A)_{ij}| \leq a$ , then  $\sum(\dots) \leq e^{nat}/n < \infty$ . So  $t \mapsto \exp[tA]$  behaves like a power series with radius of convergence  $\infty$ .

$$\begin{aligned} \frac{d}{dt} \left( \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} \right) &= \sum_{k=0}^{\infty} \frac{d}{dt} \frac{(tA)^k}{k!} = \sum_{k=1}^{\infty} \frac{A^k}{k!} k t^{k-1} = \sum_{k=1}^{\infty} \frac{A^k}{(k-1)!} t^{k-1} \\ &= A \sum_{k=1}^{\infty} \frac{(tA)^{k-1}}{(k-1)!} = A \exp[tA]. \end{aligned}$$

But in the last line we may as well place  $A$  at the end to obtain  $\frac{d}{dt} = \exp[tA]A$ .

□

**Theorem 5.24.** Suppose that  $A \in \mathbb{R}^{n \times n}$ . Then the solution to

$$\frac{d}{dt}x = Ax, \quad x(t_0) = x_0 \in \mathbb{R}^{n \times 1}$$

is given by

$$x(t) = e^{A(t-t_0)}x_0.$$

**Proof.** Note that

$$\frac{d}{dt}x(t) = \frac{d}{dt} \left( e^{A(t-t_0)}x_0 \right) = A e^{A(t-t_0)}x_0 = Ax(t).$$

At  $t = t_0$  we have

$$x(t_0) = e^{A \cdot 0}x_0 = Ix_0 = x_0.$$

By the existence and uniqueness theorem, this is the unique solution.  $\square$

**Remark 5.25.**

- (i)  $e^{At}$  is called the **Fundamental matrix** of the ODE system.
- (ii) If  $x_0 = v$ ,  $Av = \lambda v$ , then  $x(t) = e^{At}v = e^{\lambda t}v$ . In this connection we see that the eigenvalue-eigenvector method is a special case of the matrix exponential method. More generally, see next point:
- (iii) Suppose now  $A$  has  $n$  linearly independent eigenvectors  $v_1, \dots, v_n \in \mathbb{C}^n$  with corresponding eigenvalues  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  (where note that we allow  $\mathbb{C}$ ). Then by writing  $P = (v_1, \dots, v_n)$ , we have  $AP = P\Lambda$ , and so  $A$  is diagonalizable with  $A = P\Lambda P^{-1}$ . In particular,

$$e^{At} = P e^{\Lambda t} P^{-1} = P \text{diag} \left( e^{\lambda_1 t}, \dots, e^{\lambda_n t} \right) P^{-1}.$$

Thus

$$e^{At} P e_i = e^{\lambda_i t} v_i.$$



**Theorem 5.26 (Jordan Normal Form).** Suppose that  $A \in \mathbb{C}^{n \times n}$ . Then there exists  $U, J \in \mathbb{C}^{n \times n}$  such that

$$A = UJU^{-1},$$

where  $J$  is given by

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}, \quad J_j = \begin{pmatrix} \lambda_j & 1 & & \\ & \lambda_j & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{pmatrix} \in \mathbb{C}^{m_j \times m_j}, \quad \sum m_j = n.$$

The columns of  $U$  are called generalized eigenvectors and satisfy

$$(A - \lambda I)^k u_i = 0, \quad k \leq n.$$

Given a Jordan normal form decomposition  $A = UJU^{-1}$ , we have

$$e^{At} = U \exp[Jt] U^{-1},$$

Here,

$$J^k = \begin{pmatrix} J_1^k & & \\ & \ddots & \\ & & J_m^k \end{pmatrix}, \quad J_j^k = \begin{pmatrix} \lambda_j & 1 & & \\ & \lambda_j & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{pmatrix}^k.$$

If  $J_i = \alpha_i I + N$ , where  $N \in \mathbb{R}^{l \times l}$  is the nilpotent matrix with 1 on the superdiagonal and 0 elsewhere, then

$$e^{J_i t} = e^{\alpha_i t} e^{Nt} = \exp[\alpha_i t I] \exp(Nt),$$

since  $IN = NI$ . We have  $\exp(\alpha_i t I) = e^{\alpha_i t} I$  and  $N^k$  is the matrix with 1 on the  $k^{\text{th}}$  superdiagonal and 0 elsewhere (Ex.). Thus

$$\exp[Nt] = I + Nt + \frac{(Nt)^2}{2!} + \cdots + \frac{(Nt)^{l-1}}{(l-1)!} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{l-1}}{(l-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{l-2}}{(l-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

*Example 5.27.*

$$\frac{d}{dt}x = \begin{pmatrix} \lambda & 1 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}x$$

has solution

$$e^{At} = e^{\lambda t} e^{Nt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$



## 5.5 Nonhomogeneous ODE Systems

Consider the nonhomogeneous ODE system

$$\frac{d}{dt}x = A(t)x + G(t), \quad x(t_0) = x_0.$$

Recall that in the  $n = 1$  case ( $y'' + py' + qy = g(t)$ ), we first find solutions  $y_1$  and  $y_2$  to the corresponding homogeneous ODE, and then use the ansatz  $y = \mu_1 y_1 + \mu_2 y_2$  with some clever restrictions on  $\mu_1, \mu_2$  (variation of parameters/constants).

To generalize to nonhomogeneous ODE systems, we first recall the case  $n = 1$

$$\frac{d}{dt}x = ax + g(t)$$

with  $a$  constant. Using the integrating factor  $e^{-at}$ , we have

$$\frac{d}{dt} (e^{-at} x) = e^{-at} g(t)$$

and so

$$e^{-at} x(t) = e^{-at_0} x(t_0) + \int_{t_0}^t e^{-as} g(s) ds,$$

giving

$$x(t) = e^{a(t-t_0)} x_0 + \int_{t_0}^t e^{a(t-s)} g(s) ds$$

or equivalently

$$x(t) = e^{a(t-t_0)} \left[ x_0 + \int_{t_0}^t e^{-a(s-t_0)} g(s) ds \right].$$

Note that the first term  $e^{a(t-t_0)} x_0$  solves the homogeneous ODE with the given IC, and for fixed  $s$ , the term  $e^{a(t-s)} g(s)$  solves the homogeneous ODE with IC  $g(s)$  at  $t = s$ . We may think of them as solutions to the following two ODEs:

$$\begin{cases} \frac{d}{dt} x_1 = ax_1, & x_1(t_0) = x_0 \\ \frac{d}{dt} x_2 = g(t), & x_2(t_0) = 0 \end{cases}$$

**Theorem 5.28.** For  $n \geq 2$ , the solution to  $\frac{d}{dt} x = Ax + G(t)$ ,  $x(t_0) = x_0$  is given by

$$\begin{aligned} x(t) &= e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-s)} G(s) ds \\ &= e^{A(t-t_0)} \left[ x_0 + \int_{t_0}^t e^{-A(s-t_0)} G(s) ds \right]. \end{aligned}$$

**Proof.** Define

$$c(t) := x_0 + \int_{t_0}^t e^{-A(s-t_0)} G(s) ds.$$

Then the proposed solution is  $x(t) = e^{A(t-t_0)} c(t)$ . In like of existence and uniqueness, we just need to verify that this solves the ODE with the IC. We have

$$\frac{d}{dt} [e^{A(t-t_0)} c(t)] = A e^{A(t-t_0)} c(t) + e^{A(t-t_0)} \dot{c}(t).$$

Now, using the fact that  $\frac{d}{dt} \left( \int_{t_0}^t b(s) ds \right) = b(t)$ , we have

$$\dot{c}(t) = e^{-A(s-t_0)} G(t) \Big|_{s=t} = e^{-A(t-t_0)} G(t),$$

Thus

$$\begin{aligned} \frac{d}{dt} [e^{A(t-t_0)} c(t)] &= A e^{A(t-t_0)} c(t) + e^{A(t-t_0)} \dot{c}(t) \\ &= Ax + G(t). \end{aligned}$$

□

*Remark 5.29.* From this result we have **Duhamel's formula**: The differential equation

$$\partial_t f = Lf + g(t, x), \quad f|_{t=0} = f_0$$

where  $L$  is a  $t$ -independent linear operator (e.g.,  $\partial_x$ ), has solution

$$f = e^{Lt} f_0 + \int_0^t e^{L(t-s)} g(s, x) \, ds.$$



## 6 The Theory of Existence and Uniqueness

We focus on the case  $n = 1$ , but the proof generalizes to  $n \geq 2$  easily.

**Theorem 6.1** (Existence and Uniqueness). *Consider the differential equation*

$$\frac{d}{dt}y = f(t, y(t)), \quad y(t_0) = y_0$$

*in the region  $R : |t - t_0| \leq a, |y - y_0| \leq b$ .*

*We assume that  $f$  is continuous in  $R$  and Lipschitz in  $y$  with Lipschitz constant  $L$ . Then, there exists  $h > 0$  such that the ODE admits a unique  $C^1$  solution for  $|t - t_0| \leq h$ .*

**Proof.** The idea is to construct a sequence  $\varphi_0, \varphi_1, \dots$  so that  $\varphi_n \rightarrow \varphi$ . One idea is to define  $\varphi_0 = y_0$ ,  $\varphi_1(t) = y_0 + f(t_0, y_0)(t - t_0)$ , and so on. But this requires  $f$  to be differentiable.

Alternatively, we may integrate:

$$\varphi(t) := t_0 + \int_{t_0}^t f(s, \varphi(s)) \, ds.$$

Note that differencing both sides gives  $\frac{d}{dt}\varphi(t) = f(t, \varphi(t))$ . Again we set  $\varphi_0 = y_0$ . For  $n \geq 0$ , we define

$$\varphi_{n+1}(t) := y_0 + \int_{t_0}^t f(s, \varphi_n(s)) \, ds$$

and show that  $\varphi_n$  converges. This method is called **Picaro iteration**.<sup>1</sup>

We will show first that  $|\varphi_n(t) - y_0| \leq b$  in a small neighborhood of  $t_0$ . From  $f$  being continuous, we know that  $|f| \leq M$  in  $R$ . Thus

$$|\varphi_{n+1}(t) - y_0| \leq \int_{t_0}^t |f(s, \varphi_n(s))| \, ds \leq M|t - t_0|.$$

And so by setting  $h \leq b/M$ , we have  $|\varphi_n(t) - y_0| \leq b$  for all  $n$  and  $|t - t_0| \leq h$ .

Next, we show that  $|\varphi_{n+1} - \varphi_n| \rightarrow 0$  as  $n \rightarrow \infty$ . Note that

$$\begin{aligned} |\varphi_{n+1}(t) - \varphi_n(t)| &= \left| \int_{t_0}^t f(s, \varphi_n(s)) - f(s, \varphi_{n-1}(s)) \, ds \right| \\ &\leq L \int_{t_0}^t |\varphi_n(s) - \varphi_{n-1}(s)| \, ds. \end{aligned}$$

Note recall that we have the uniform in time bound

$$|\varphi_1 - \varphi_0| = |\varphi_1 - y_0| \leq Mh.$$

Iterating the inequality above gives

$$|\varphi_{n+1}(t) - \varphi_n(t)| \leq (Lh)^n Mh$$

---

<sup>1</sup>Another way is to use a fixed point theorem on a suitable function space.

which converges if we choose  $h$  such that  $Lh < 1$ . We define thus

$$\varphi := \lim_{n \rightarrow \infty} \text{ved} \varphi_n = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} (\varphi_{i+1} - \varphi_i) + \varphi_0.$$

Note that the series on the right converges uniformly in light of the bound above. In particular,

$$|\varphi - \varphi_n| \leq \sum_{i \geq n} |\varphi_{i+1} - \varphi_i| \leq Mh \frac{(Lh)^n}{1 - Lh}.$$

To see  $\varphi$  solves the ODE, note that from (exercise)

$$\varphi(t) = y_0 + \int_{t_0}^t f(s, \varphi(s)) \, ds,$$

we have

$$\frac{d}{dt} \varphi(t) = f(t, \varphi(t)) \implies \varphi \in C^1.$$

It remains to show uniqueness. Suppose that  $\psi$  is another solution to the ODE with the same IC. Note that

$$\varphi(t) = t_0 + \int_{t_0}^t f(s, \varphi(s)) \, ds, \quad \psi(t) = t_0 + \int_{t_0}^t f(s, \psi(s)) \, ds.$$

Thus

$$d(t) := \varphi(t) - \psi(t) = \int_{t_0}^t [f(s, \varphi(s)) - f(s, \psi(s))] \, ds.$$

The Lipschitz condition gives

$$|d(t)| \leq L \int_{t_0}^t |d(s)| \, ds \implies |d'(t)| \leq LD'(t), \quad \text{where } D(t) := \int_{t_0}^t |d(s)| \, ds.$$

Using the integrating factor  $e^{-Lt}$  we get

$$\frac{d}{dt} \left( e^{-Lt} D(t) \right) \leq 0, \quad D(t) \geq 0$$

which gives

$$e^{-Lt} D(t) \leq e^{-Lt_0} D(t_0) = 0$$

and in turn

$$D(t) = 0, d = 0, \varphi = \psi.$$

□

*Remark 6.2.* The assumption  $f \in C^0$  already gives existence (Lipschitz is not required). The Lipschitz assumption gives uniqueness (no continuity required). ☕

## 7 Quantitative Estimates

$$\frac{d}{dt}x = f(t, x), \quad x(t_0) = x_0.$$

Recall that

- $f \in C^0 \implies$  Existence
- $f$  is Lipschitz in  $x \implies$  Uniqueness

Recall that with

$$|\varphi(t) - \psi(t)| \leq L \int_{t_0}^t |\varphi(s) - \psi(s)| \, ds,$$

we have

$$d(t) := |\varphi(t) - \psi(t)| \leq L \int_{t_0}^t d(s) \, ds.$$

We generalize this result:

**Lemma 7.1** (Gronwall's Inequality). *For given  $\alpha, \beta$ , assume  $\beta(t) \geq 0$ ,  $\psi(t) \leq \alpha(t) + \int_0^t \beta(s)\psi(s) \, ds$  and define*

$$B(s) := \exp\left(\int_0^s \beta(t) \, dt\right).$$

Then,

$$\psi(t) \leq \alpha(t) + \int_0^t \alpha(s)\beta(s) \frac{B(t)}{B(s)} \, ds.$$

If  $\alpha \in C^1$ , the right hand side may be written by integration by parts as

$$\alpha(0)B(t) + \int_0^t \alpha'(s) \frac{B(t)}{B(s)} \, ds.$$

**Proof.** We write

$$A(t) := \int_0^t \beta(s)\psi(s) \, ds$$

and seek an ODE inequality for  $A$ . The idea is to solve an ODE inequality for  $A$  of the form  $\frac{d}{dt}[\text{unknown}] \leq \text{known}$  (note that this works only if the RHS is known). Note that we have  $A'(t) = \beta(t)\psi(t)$ . Using the assumptions that  $\psi(t) \leq \alpha(t) + A(t)$  and  $\beta \geq 0$ , we have

$$\frac{d}{dt}A(t) = \psi\beta \leq \alpha\beta + A\beta.$$

This is a first order linear “ODE inequality” in  $A$ . Using the integrating factor  $1/B(t)$ , we have

$$\frac{d}{dt} \left( \frac{A(t)}{B(t)} \right) \leq \alpha(t)\beta(t) \frac{1}{B(t)},$$

where we note that

$$B^{-1} = \exp\left(-\int_0^t \beta \, dt\right), \quad \dot{B}^{-1} = -\beta B^{-1}.$$

Integrating both sides from 0 to  $t$  gives

$$A(t)B^{-1}(t) \leq \int_0^t \alpha(s)\beta(s)B^{-1}(s) \, ds + A(0)B^{-1}(0),$$

where the last term is zero since  $A(0) = 0$  by definition. Now we have

$$\psi(t) \leq \alpha(t) + A(t) \leq \alpha(t) + \int_0^t \alpha(s)\beta(s) \frac{B(t)}{B(s)} \, ds.$$

□

*Remark 7.2.* Applications:

- $\alpha = 0, \beta = L, \psi = d$  and we recover the motivating example.
- $\alpha = 0$  implies  $\psi \leq 0$ .
- $\alpha(t) = \alpha_0 + \alpha_1 t, \beta = L$ . Then

$$B(s) = \exp\left(\int_0^s L \, ds\right) = e^{Ls}$$

and so

$$\psi(t) \leq \alpha_0 e^{Lt} + \alpha_1 \int_0^t e^{L(t-s)} \, ds.$$

In some sense this is similar to the Duhamel's formula for the ODE system

$$\begin{cases} \frac{d}{dt}x = Lx + \alpha_1 \\ x(0) = \alpha_0 \end{cases} \implies x(t) = e^{Lt}\alpha_0 + \alpha_1 \int_0^t e^{L(t-s)} \, ds.$$

☕

Consider now a small perturbation in  $x_0$  or  $f$  in the ODE

$$\frac{d}{dt}x = f(t, x), \quad x(t_0) = x_0$$

so that it becomes

$$\frac{d}{dt}y = g(t, y), \quad y(t_0) = y_0.$$

How do these errors affect  $x$ ? The hope is that  $x$  does not change too much. Otherwise the model is not robust.

**Theorem 7.3** (Continuous Dependence in IC, parameters, etc.). *Suppose  $f$  is Lipschitz in the domain  $D$  of interest with Lipschitz constant  $L$  and denote*

$$M := \max_{(t,x) \in D} |f(t,x) - g(t,x)|.$$

*Then,*

$$|x(t) - y(t)| \leq e^{L|t-t_0|} |x_0 - y_0| + \frac{M}{L} \left( e^{L|t-t_0|} - 1 \right),$$

*where the first part is due to the IC error and the second part is due to the ODE error.*

**Proof.** Note that

$$\frac{d}{dt} [x(t) - y(t)] = f(t, x(t)) - g(t, y(t)).$$

Integrating,

$$|x(t) - y(t)| \leq |x_0 - y_0| + \int_{t_0}^t |f(s, x(s)) - g(s, y(s))| \, ds.$$

The integrand may be bounded as

$$\begin{aligned} |f(s, x(s)) - g(s, y(s))| &\leq |f(s, x(s)) - f(s, y(s))| + |f(s, y(s)) - g(s, y(s))| \\ &\leq L|x(s) - y(s)| + M. \end{aligned}$$

Observe that

$$|x(t) - y(t)| \leq |x_0 - y_0| + \int_{t_0}^t [L|x(s) - y(s)| + M] \, ds.$$

Setting  $\alpha := |x_0 - y_0| + M(t - t_0)$  and  $\psi(t) := |x(t) - y(t)|$ , we may apply Gronwall's inequality with  $\beta = L$  to get

$$\begin{aligned} |x(t) - y(t)| &\leq e^{L(t-t_0)} |x_0 - y_0| + \int_{t_0}^t M L e^{L(t-s)} \, ds \\ &\leq e^{L|t-t_0|} |x_0 - y_0| + \frac{M}{L} \left( e^{L|t-t_0|} - 1 \right). \end{aligned}$$

□

**Corollary 7.4.**

- If  $f = g$ ,  $x_0 = y_0$ , then  $x = y$ .
- If  $f = g$  then

$$|x(t) - y(t)| \leq e^{L|t-t_0|} |x_0 - y_0|.$$

*That is, we have continuous (and in fact Lipschitz) dependence on the IC. In other words, if we define  $\Phi(t, x_0)$  to be the solution of*

$$\frac{d}{dt} \Phi = f(t, \Phi), \quad \Phi(t_0, x_0) = x_0,$$

*then  $\Phi$  is Lipschitz in  $x_0$ .*

- If  $x_0 = y_0$  and say  $|f - g| \leq \varepsilon$ , then

$$|x(t) - y(t)| \leq \frac{\varepsilon}{L} \left( e^{L|t-t_0|} - 1 \right).$$

In particular, if  $f_\alpha(t, x) := \alpha f_0 + (1 - \alpha)f_1$ , then

$$|f_\alpha - f_\beta| \leq |\alpha - \beta| (|f_1| + |f_0|).$$

*Example 7.5.* We perturb the ODE

$$\frac{d}{dt}x = \frac{x^2}{1 + t^2 + x^2}, \quad x(0) = 0$$

to get

$$\frac{d}{dt}y = \frac{y^2 + \varepsilon}{1 + t^2 + y^2} + \varepsilon, \quad y(0) = \varepsilon.$$

We have  $|f - g| \leq \varepsilon$  and  $|x_0 - y_0| \leq \varepsilon$ , and we may check that  $|\partial_x f(t, x)| \leq L$  for any  $t, x$ . Thus

$$|x(t) - y(t)| \leq e^{Lt} \varepsilon + \frac{\varepsilon}{L} \left( e^{Lt} - 1 \right) = \varepsilon \left( e^{Lt} + \frac{e^{Lt} - 1}{L} \right).$$



Now consider

$$y' \leq F(t, y(t)), \quad t \in [a, b].$$

**Theorem 7.6** (Comparison). Suppose  $F$  is Lipschitz. Let  $f, g \in C^1$  are such that

$$f' \leq F(t, f(t)), \quad g' = F(t, g(t)).$$

If  $f(a) \leq g(a)$ , then

$$f(t) \leq g(t), \quad t \in [a, b].$$

**Proof.** Suppose for contradiction that  $f(t_1) > g(t_1)$ . Let  $\Omega := \{t : f(t) \leq g(t)\}$ . We have  $a \in \Omega$ . Let  $t_0 := \max\{t \in [a, t_0] \cap \Omega\}$ . We know  $f(t_0) = g(t_0)$  and  $f(t) \leq g(t)$  for each  $t \in \Omega$ . Moreover,  $f(t) > g(t)$  for each  $t \in (t_0, t_1]$ . But this and continuity implies that  $f(t_0) \geq g(t_0)$ . Thus  $f(t_0) = g(t_0)$ .

Now, from assumption we have

$$f'(t) - g'(t) \leq F(t, f(t)) - F(t, g(t))$$

By integrating both sides,

$$\begin{aligned} f(t) - g(t) &\leq f(t_0) - g(t_0) + \int_{t_0}^t F(s, f(s)) - F(s, g(s)) \, ds \\ &\leq \int_{t_0}^t L(f(s) - g(s)) \, ds, \end{aligned}$$

where the second inequality is justified from  $f(t) \geq g(t)$  on  $(t_0, t_1]$ . By Gronwall's inequality, we have

$$f(t) - g(t) \leq 0, \quad t \in [t_0, t_1]$$

a contradiction. □

Now, if  $F$  is not Lipschitz, this no longer holds true:

*Example 7.7.*

$$\frac{d}{dt}y = y^{\frac{1}{3}}, \quad y(0) = 0$$

has solution  $g(t) \equiv 0$ . Now try the ansatz  $f(t) = ct^\alpha$ . We have


$$f'(t) = c\alpha t^{\alpha-1}, \quad f^{\frac{1}{3}}(t) = c^{\frac{1}{3}}t^{\frac{\alpha}{3}},$$


giving

$$\alpha = \frac{3}{2}, \quad c = \pm \left(\frac{3}{2}\right)^{\frac{3}{2}}.$$

In particular, a solution is

$$f(t) = \begin{cases} ct^{\frac{3}{2}}, & t \geq 0 \\ -c(-t)^{\frac{3}{2}}, & t < 0 \end{cases},$$

which crosses  $g$  at  $t = 0$  and is  $C^1$  (since in particular both parts of  $f$  is  $C^1$  at 0). 

*Remark 7.8.* Another version of the comparison theorem is as follows:  $f' < F(t, f)$  with strict inequality and  $f(a) < g(a)$ ,  $f$  is continuous (Lipschitz not required??). 

**Theorem 7.9.**

$$\frac{d}{dt}x = f(t, x), \quad \frac{d}{dt}y = g(t, y).$$

If  $f(t, z) \leq g(t, z)$  for any  $t, z$  in the domain of interest  $D$ , and if  $f$  or  $g$  is Lipschitz in  $D$ , and that  $x(a) \leq y(a)$ , then  $x(t) \leq y(t)$  for each  $t \in [a, b]$ .

**Proof.** Suppose  $g$  is Lipschitz (the other case is similar). We have

$$\frac{d}{dt}x(t) = f(t, x) \leq g(t, x(t)), \quad \frac{d}{dt}y(t) = g(t, y).$$

By the comparison theorem, we have  $x(t) \leq y(t)$  for each  $t \in [a, b]$ . □

*Example 7.10.* Consider

$$\frac{d}{dt}x_\alpha(t) = f(t, x_\alpha(t)) + \alpha, \quad x_\alpha(t_0) = x_0.$$

Then the solution satisfies

$$x_\alpha(t) \leq x_\beta(t), \quad t \in [a, b], \quad \text{if } \alpha \leq \beta.$$




*Example 7.11.*

$$\frac{d}{dt}x(t) = P_n(x, t), \quad t \in [0, 1], \quad x(0) = x_0,$$

where  $P_n$  is a polynomial such that  $P_n(x, t) \leq Ce^x$  for some constant  $C > 0$ . We know  $x$  is bounded above by the solution to

$$\frac{d}{dt}y = Ce^y, \quad y(0) = x_0.$$

Of course, this works for any upper bounding function. 


## 7.1 Extension

In this section we will always assume that  $f \in C(\mathbb{R}, \mathbb{R})$  is Lipschitz in the second argument in any bounded region.

*Example 7.12 (of Blowup).*

$$\frac{d}{dt}y = y^2, \quad y(0) = 1.$$

$$\frac{d}{dt}y^{-1} = -1 \implies y^{-1}(t) = 1 - t \implies y = \frac{1}{1-t}.$$

But observe that the solution exists up to  $1^-$ :  $\lim_{t \rightarrow 1^-} y(t) = +\infty$ . 

How can we detect/characterize blowup?

**Lemma 7.13 (Gluing).** *Suppose  $f \in C(\mathbb{R}, \mathbb{R})$  and  $f(t, x)$  is Lipschitz in any bounded region:*

$$|f(t, x_1) - f(t, x_2)| \leq L_D |x_1 - x_2|, \quad (t, x_1), (t, x_2) \in D \subset \mathbb{R}^2.$$

Suppose  $\varphi_i(t)$  solves

$$\frac{d}{dt}x = f(t, x), \quad x(t_0) = x_0$$

on  $t \in I_i := (a_i, b_i)$ . If  $t_0 \in I_1 \cap I_2 = (a, b)$  and  $\varphi_1(t_0) = \varphi_2(t_0)$ . Then  $\varphi_1(t) = \varphi_2(t)$ ,  $t \in I_1 \cap I_2$  and

$$\varphi(t) := \begin{cases} \varphi_1(t), & t \in I_1 \\ \varphi_2(t), & t \in I_2 \end{cases}$$

is a  $C^1$  solution to the ODE on  $I_1 \cup I_2$ .

**Proof.** The hard part is to show  $\varphi_1 = \varphi_2$  on  $I_1 \cap I_2$ .

Apply existence and uniqueness on to the ODE with IC at  $t_0$  to get  $\varphi_1 = \varphi_2$  on  $t_0 \pm h$ . Write

$$J_- := \{p : \varphi_1(t) = \varphi_2(t), t \in [p, t_0 + h] \cap (a, b)\}.$$

Then we can prove that  $a = \inf J_- =: A$ . Suppose not, then  $A > a$ . There exists a sequence  $P_n \rightarrow A$ , each  $P_n \in J$ , so that  $\varphi_1 = \varphi_2$  on each  $[P_n, t_0 + h]$ . By continuity,  $\varphi_1(P_n) = \varphi_2(P_n)$ . Again by continuity,  $\varphi_1(A) = \varphi_2(A)$ . But by existence and uniqueness,  $\varphi_1 = \varphi_2$  on  $[A - h', t_0 + h]$  for some  $h' > 0$ , contradicting the definition of  $A$ .

Alternatively:

Set  $J := \{t \in I_1 \cap I_2 : \varphi_1(t) = \varphi_2(t)\}$ . The goal is to show  $J = I_1 \cap I_2$ ,  $J \neq \emptyset$ . By continuity,  $J$  is relatively closed. But since it is also relatively open (by local existence and uniqueness), we have  $J = I_1 \cap I_2$ .

Yet another alternative: Using Gronwall's inequality, from

$$|\varphi_1(t) - \varphi_2(t)| \leq \int_{t_0}^t L |\varphi_1(s) - \varphi_2(s)| \, ds,$$

we have  $\varphi_1 = \varphi_2$  on  $I_1 \cap I_2$ .

We now know that  $\varphi$  is well-defined. It is  $C^1$  since each  $\varphi_i$  is  $C^1$  on  $I_i$ , and it solves the ODE since each  $\varphi_i$  does so on  $I_i$ .  $\square$

What happens if  $f$  is not Lipschitz? Well then  $\varphi_1$  may not be identically equal to  $\varphi_2$  on  $I_1 \cap I_2$ . Then there are multiple ways to choose  $\varphi$ , and the second statement is ambiguous.

**Theorem 7.14** (Blowup Criterion). *Suppose  $f \in C(\mathbb{R}, \mathbb{R})$  is Lipschitz in  $x$  in any bounded region. Suppose that  $\varphi$  is a  $C^1$  solution to  $\frac{d}{dt}x = f(t, x)$  on  $(t_-, t_+)$ . The solution can be extended to  $(t_-, t_+ + \varepsilon)$  for some  $\varepsilon > 0$  if and only if one of the following equivalent conditions hold:*

(i) *There exists a sequence  $\{t_n\}$  such that  $t_n \rightarrow (t_+)^-$  and  $\varphi(t_n) \rightarrow \varphi_0 \neq \pm\infty$ . Note that this is true for oscillating behavior but not monotonic blowup.*

(ii)  $\limsup_{t \rightarrow (t_+)^-} |\varphi(t)| < +\infty$ .

*In particular, the solution cannot be extended beyond  $t_+$  if  $\lim_{t \rightarrow (t_+)^-} |\varphi(t)| = +\infty$ .*

**Proof.** Note that the forward direction is implied by the gluing lemma. For the converse, suppose  $|\varphi(t_n) - \varphi_0| \leq 1$  for each  $n \geq N$ . The idea is to apply existence and uniqueness at  $(t_n, \varphi(t_n))$  and hope that we get a solution on  $(t_n - h, t_n + h)$  with  $t_n + h > t_+$ .

Write

$$\begin{aligned} S &:= [t_+ - 1, t_+ + 1] \times [\varphi_0 - 1, \varphi_0 + 1] \\ D &:= [t_+ - 2, t_+ + 2] \times [\varphi_0 - 2, \varphi_0 + 2] \end{aligned}$$

We will control  $(t_n, \varphi(t_n))$  to be in  $S$ , and  $\varphi$  to be in  $D$ . By  $f$  being continuous and Lipschitz in  $x$  in  $D$ , we have

$$\begin{cases} |f| \leq M \\ |f(t, x) - f(t, y)| \leq L|x - y| \end{cases}$$

in  $D$  for some  $L, M > 0$ . In particular,  $|\varphi'| \leq M$  in  $D$ . Applying existence and uniqueness at  $(t_n, \varphi(t_n))$  gives a unique solution  $\psi$  such that  $\psi(t_n) = \varphi(t_n)$  and with domain  $t_n - h, t_n + h$ , where  $h$  can be chosen as

$$0 < h < \min \left\{ 1, \frac{1}{M}, \frac{1}{L} \right\}.$$

Since  $h$  is independent of  $n$ , we may choose  $n$  large enough so that  $t_n + h > t_+$ .  $\square$


Q: which ODEs admit global solutions?

**Proposition 7.15.** *Consider  $x' = f(t, x)$  with  $|f| \leq M$  for each  $(t, x)$ ,  $f \in C(\mathbb{R}, \mathbb{R})$  and Lipschitz in  $x$  in any bounded region. Then the solution exists globally and is unique.*

**Proof.** Assume the solution exists on  $(t_-, t_+)$ . Note that

$$|x(t) - x(t_0)| = \left| \int_{t_0}^t f(s, x(s)) \, ds \right| \leq M|t - t_0| < \infty.$$

By the blowup criterion, we know that  $x$  can be extended beyond  $t_+$  and  $t_-$ .  $\square$

*Remark 7.16.* The idea of existence-uniqueness / blowup criterion applies also to some time evolution PDEs e.g.,  $\partial_t u = N(u, \partial u, \dots)$ . We similarly have local existence-uniqueness, and we cannot extend beyond  $t_+$  if and only if  $\lim_{t \rightarrow (t_+)^-} \|u(t, x)\|_Y = \infty$ . 

## 7.2 Finite Time Blowup

Consider the Riccati ODE  $y' = y^p$ ,  $y(0) = 1$ .


- $p = 1$ . The unique solution is  $y(t) = e^t$ .
- $p > 1$ . By the comparison principle we have  $y \geq e^t$  since  $y(0) \geq 1$ . By separation of variables, we have

$$\frac{y'}{y^p} = 1 \implies y^{1-p}(t) = 1 + (1-p)t.$$

Note that


$$y^{p-1}(t) = \frac{1}{1 - (p-1)t} \rightarrow +\infty \text{ as } t \rightarrow \left(\frac{1}{p-1}\right)^-.$$


So we have blow up at  $t_+ = 1/(p-1) < \infty$ .


*Example 7.17.* Consider  $y' = e^y \geq y^2$ . By comparison with  $y' = y^2$ , we have finite time blowup. Consider  $y' = y^n + \dots + 1 + e^t$ ,  $y(0) = 1$ . By comparison with  $y' = y^2$ , we have finite time blowup. 

*Example 7.18.*  $y' = y^2$ ,  $y(0) = y_0$ .

- If  $y_0 > 0$ , we have finite time blowup.
- If  $y_0 = 0$ , we have  $y \equiv 0$ .
- If  $y_0 < 0$ , we have  $y = y_0/(1 - ty_0)$ . Note that  $1 - ty_0 > 0$  for  $t > 0$ . Thus  $y_0 \sim -1/t$  as  $t \rightarrow +\infty$ .

From this we see nonlinear stabilization: negative IC gives global solution decaying to 0, while positive IC gives finite time blowup. 

*Example 7.19.* Consider  $y' = y^2 + \varepsilon$  for  $\varepsilon > 0$  and  $y(0) = y_0$ . Note that  $y(t) \geq y_0 + \varepsilon t$  and so  $y(t_0) > 0$  for  $t_0$  large. Then, we have  $\dot{y} > y^2$  and thus blowup. 


*Example 7.20.* Consider  $y' = y^2 - \varepsilon^2$ . We have  $y_1 \equiv \varepsilon$  and  $y_2 \equiv -\varepsilon$ . For  $y_0 > \varepsilon$  we again have  $y' \geq y^2$  and so blowup. 

## 8 Autonomous ODE, Stability, and Phase Portrait/Plane

Recall that a general first order ODE is  $x' = F(t, x)$ , where  $x, F \in \mathbb{R}^{n \times 1}$ . An **autonomous** ODE is one where  $F$  does not depend on  $t$ :

$$x' = F(x).$$

A special solution is the constant solution  $x(t) \equiv x_*$ . We say  $x_*$  is in **equilibrium** or a **critical point** if  $F(x_*) = 0$ .

*Example 8.1.*  $x' = x^2 - x + t$  has no constant solution. 


We call  $(t, x(t))$  the **trajectory** of the solution  $x(t)$ . For  $n = 2$ ,

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix},$$

we call the  $x_1$ - $x_2$  plane the **phase plane**. We may think of  $x_2$  as a function of  $x_1$ ,  $x_2 = x_2(x_1)$ , to remove the dependency on  $t$ : By the chain rule,


$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{f_1}{f_2}(x_1, x_2).$$

The representative set of trajectories is called the **phase portrait**.

*Example 8.2.* The ODE  $x' = x^2 - x =: f$  has critical points at  $x_* = 0, 1$ . 

*Example 8.3.* The system

$$\begin{cases} x' = x^2 - 2y \\ y' = x + y \end{cases}.$$

Note that  $x^2 = 2y = -2x$  gives  $x = 0$  or  $x = -2$ . This gives the critical points  $(0, 0)$  and  $(-2, 2)$ . 

### 8.1 Phase Portrait of 2 x 2 Systems

Consider the linear  $2 \times 2$  system:

$$x' = Ax, \quad A \in \mathbb{R}^{2 \times 2}.$$

The equilibrium is at  $x_* = 0$  and the real eigen-vectors associated with the eigenvalue 0. But what happens as  $t \rightarrow \infty$  for a general solution  $x$ ?

*Remark 8.4.* Recall first that for

$$z(t) := e^{\alpha t} (\cos \beta t + i \sin \beta t),$$

we have

- (i) If  $\alpha > 0$ , then  $|z(t)| \rightarrow \infty$  as  $t \rightarrow \infty$ .
- (ii) If  $\alpha = 0$ , then  $z(t)$  is bounded, periodic, and in particular does not converge.
- (iii) If  $\alpha < 0$ , then  $z(t) \rightarrow 0$  exponentially fast as  $t \rightarrow \infty$ .



We consider the following three cases:

**Case 1:**

$\lambda_1 < \lambda_2$  real. The general solution is

$$x(t) = \underbrace{e^{\lambda_1 t}}_{\text{amplitude}} \underbrace{\left( e^{(\lambda_1 - \lambda_2)t} c_1 v_1 + c_2 v_2 \right)}_{\text{direction}}.$$

(i) If  $\lambda_1 < 0$ , then any solution of the form

$$x(t) = c_1 e^{\lambda_1 t} v_1$$

will converge to 0 as  $t \rightarrow \infty$ , remaining on the line spanned by  $v_1$  on the way.

(ii) **(Sink)**  $\lambda_1 < \lambda_2 < 0$ , then  $x(t) \approx e^{\lambda_2 t} c_2 v_2$  for large  $t$ . Thus  $x(t) \rightarrow 0$  as  $t \rightarrow \infty$  (converging in the  $v_1$  direction faster).

(iii) **(Saddle)**  $\lambda_1 < 0 < \lambda_2$ . We have  $x \rightarrow 0$  only if  $x \equiv 0$ .

(iv) **(Source)**  $0 < \lambda_1 < \lambda_2$ , then  $x(t) \rightarrow \infty$  (diverging in the  $v_2$  direction faster) as  $t \rightarrow \infty$  unless  $x \equiv 0$ .

These can be visualized in a phase portrait.

**Case 2:**

$\lambda = \alpha \pm i\beta$ ,  $v = a \pm ib$ ,  $b \neq 0$ . The general solution is

$$\begin{aligned} x(t) &= c_1 \operatorname{Re}(e^{\lambda t} v) + c_2 \operatorname{Im}(e^{\lambda t} v) \\ &= e^{\alpha t} [c_1 (\cos \beta t a - \sin \beta t b) + c_2 (\sin \beta t a + \cos \beta t b)] \\ &= e^{\alpha t} V(t), \end{aligned}$$

where

$$V(t) := c_1 (\cos \beta t a - \sin \beta t b) + c_2 (\sin \beta t a + \cos \beta t b)$$

is periodic in  $t$ .

(i) If  $\alpha < 0$ ,  $|x| \rightarrow 0$ .

(ii) If  $\alpha = 0$ , then  $x(t) = V(t)$  is bounded and periodic (but does not converge).

(iii) If  $\alpha > 0$ , then  $|x| \rightarrow \infty$ .

**Case 3:**

One real eigen-value with two linearly independent eigen-vectors. General solution

$$x(t) = e^{\lambda t} \underbrace{(c_1 v_1 + c_2 (t v_1 + v_2))}_{\text{constant}}.$$

The trajectory is along a straight line.

#### Case 4:

One eigenvector  $v_1$  and one generalized eigen-vector  $v_2$  such that  $(A - \lambda)^2 v_2 = 0$ .  
General solution:

$$\begin{aligned} x(t) &= e^{\lambda t} (c_1 v_1 + c_2 t v_2) \\ &= e^{\lambda t} t \left( \frac{1}{t} c_1 v_1 + c_2 v_2 \right). \end{aligned}$$

For large  $t$ , the trajectory is approximately along the line spanned by  $v_2$ .


- (i) If  $\lambda > 0$ ,  $|x| \rightarrow \infty$  exponentially fast.
- (ii) If  $\lambda = 0$ ,  $|x| \sim t$ .
- (iii)  $\lambda < 0$ ,  $|x| \rightarrow 0$  exponentially fast.

Q: How smooth is  $x_2$  as a function of  $x_1$ ? Obviously Case 3 is  $C^\infty$ .

Nonlinear phase plane/portrait is much harder:

Example 8.5.

$$\begin{cases} x'_1 = f_1(x_1, x_2) \\ x'_2 = f_2(x_1, x_2) \end{cases}.$$

Suppose for example that  $f_i$  are polynomials. We may have a few critical points, say  $(0, 0)$  and  $x_c$ . A hard problem: find a  $C^\infty$  trajectory starting from 0 and passing  $x_c$ . 

Note that by existence and uniqueness, trajectories cannot cross (except at critical points).

## 8.2 Stability

We say the critical point  $x_*$  is

- **stable** if for any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that if  $|x_0 - x_*| < \delta$ ,  $|x(t) - x_*| < \varepsilon$  for each  $t \geq t_0$ .
- **asymptotically stable** if there exists  $\delta > 0$  such that if  $|x_0 - x_*| < \delta$ , then  $\lim_{t \rightarrow \infty} x(t) = x_*$ .
- **unstable** if it is not stable.

We summarize the stability of the critical point  $(0, 0)$  for the  $2 \times 2$  linear systems:

- (i) If  $\operatorname{Re} \lambda_i < 0$  for  $i = 1, 2$ , then both stable and asymptotically stable (sink).
  - (ii) If  $\operatorname{Re} \lambda_i > 0$  for some  $i$ , then unstable.
  - (iii) If  $\operatorname{Re} \lambda_i \leq 0$  for all  $i$  and  $\operatorname{Re} \lambda_i = 0$  for some  $i$ , then not asymptotically stable.
- Moreover,

- (i) if  $\lambda_1 \neq \lambda_2$ , then stable ( $|e^{\lambda_1 t}| = |e^{\lambda_2 t}| \leq 1$ ).
- (ii) if  $0 = \lambda_1 = \lambda_2$ , 2 eigenvectors, stable.
- (iii)  $0 = \lambda_1 = \lambda_2$ , where  $\lambda_2$  is the generalized eigen-vector, then unstable (since  $|x(t)| \sim t$ ).