Lecture Note on ODEs and Dynamics: Real and Complex

Anthony Hong*

March 24, 2024

*Thanks to Professor McCarthy for his inspirational teaching in course Math 350 Dynamical Systems and Chaos at WUSTL

Figure 1: AI's understanding of vector fields on sphere.

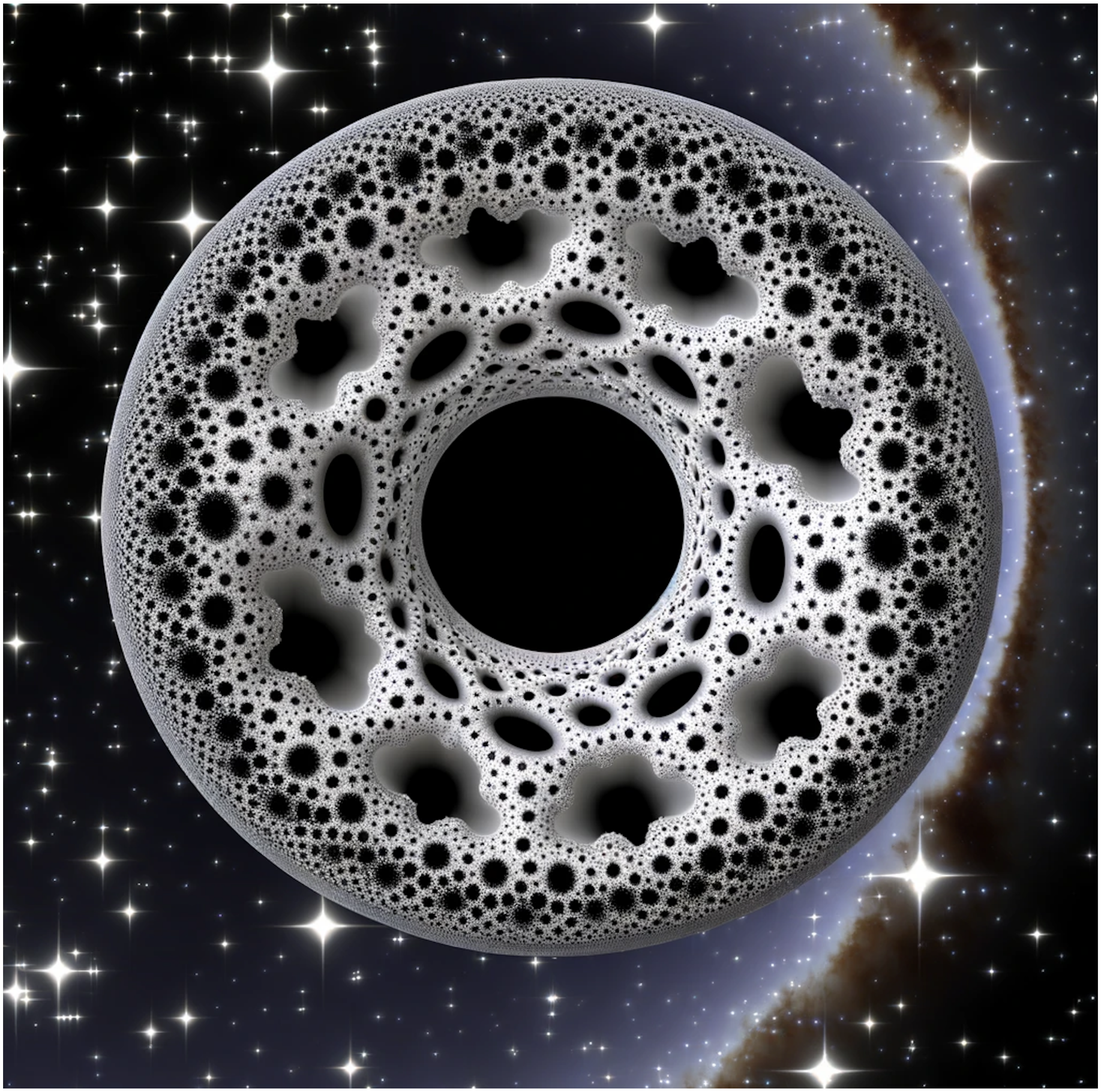

Figure 2: AI's understanding of a torus.

Figure 3: AI's understanding of fractals.

Contents

Introduction: ODEs and Dynamics

"Science is a differential equation. Religion is a boundary condition." (Alan Turing, 1954)

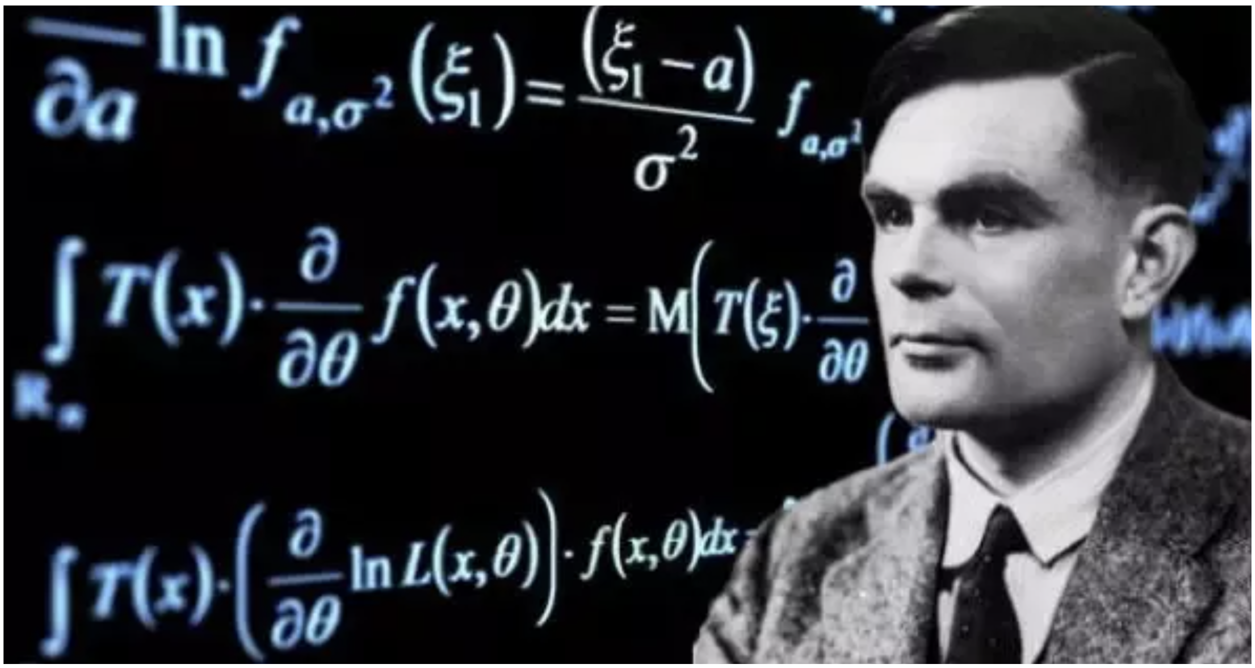

Figure 4: Alan Turing

In "The Scientific Outlook" (1931), Bertrand Russell articulates a compelling argument: "Ordinary language is totally unsuited for expressing what physics really asserts, since the words of everyday life are not sufficiently abstract. Only mathematics and mathematical logic can say as little as the physicist means to say." This perspective underscores the intrinsic limitations of everyday language in conveying the nuanced and abstract principles of physics, highlighting the indispensable role of mathematics and mathematical logic in this realm. Galileo Galilei, too, recognized the mathematical underpinnings of the universe, expressing awe at its beautifully elegant equations and asserting that the universe's laws are delineated in the language of mathematics. This realization gains further significance in the study of dynamic natural phenomena, which extend beyond the reach of algebra, designed primarily for static situations. The exploration of these phenomena necessitates the use of differential equations, which relate unknown functions to their derivatives, offering insights into the complex interplay of changing quantities in the universe.

In fact, differential equations are just one, nonetheless important, aspect of **dynamical system**. A **dynamical system** is a semigroup G acting on a space M. That is, there is a map

$$\begin{array}{rcl} \Phi : U \subseteq G \times M & \rightarrow & M \\ (t, x) & \mapsto & \Phi_t(x) \end{array}$$

with $\mathrm{proj}_2(U) = X$ (where proj_2 is the 2nd projection map) and for any x in X:

$$\begin{aligned} &\Phi(0, x) = x \\ &\Phi\left(t_2, \Phi\left(t_1, x\right)\right) = \Phi\left(t_2 + t_1, x\right), \end{aligned}$$

for $t_1, t_2 + t_1 \in I(x)$ and $t_2 \in I(\Phi(t_1, x))$, where we have defined the set $I(x) := \{t \in T : (t, x) \in U\}$ for any x in X. If G is a group, we will speak of an invertible dynamical system. We are mainly interested in **discrete dynamical systems** where

$$G = \mathbb{N}_0 \quad \text{or} \quad G = \mathbb{Z}$$

and in continuous dynamical systems where

$$G = \mathbb{R}^+ \quad \text{or} \quad G = \mathbb{R}.$$

Example 0.0.1: The prototypical example of a discrete dynamical system is an iterated map. Let f map an interval I into itself and consider

$$\Phi_n = f^n = f \circ f^{n-1} = \underbrace{f \circ \cdots \circ f}_{n \text{ times}}, \quad G = \mathbb{N}_0.$$

Clearly, if f is invertible, so is the dynamical system if we extend this definition for $n \in \mathbb{Z}$ in the usual way.

Example 0.0.2: The prototypical example of a continuous dynamical system is the flow of an autonomous differential equation

$$\Phi_t, \quad G = \mathbb{R},$$

We will return to this point in chapter 5.

Part I

Real Systems

Chapter 1

Initial Value Problems

Let's see some examples of differential equations.

$$\frac{\mathrm{d}y}{\mathrm{d}x}+2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\sin x,\quad \frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}+\frac{g}{l}\sin\varphi=0,\quad \frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}$$

The first one is a second order linear constant-coefficient ODE, the second is a nonlinear ODE, as φ is the dependent variable with respect to the independent one t, and the third is a partial differential equation (PDE). We formalize the definition.

Let $U\subseteq\mathbb{R}^m, V\subseteq\mathbb{R}^n$ and $k\in\mathbb{N}_0$. Then $C^k(U,V)$ denotes the set of functions $U\to V$ having continuous derivatives up to order k. In addition, we will abbreviate $C(U,V)=C^0(U,V)$ and $C^k(U)=C^k(U,\mathbb{R})$.

A classical ordinary differential equation (ODE) is a relation of the form

$$F\left(t,x,x^{(1)},\ldots,x^{(k)}\right)=0 \tag{1.1}$$

for the unknown function $x\in C^k(J), J\subseteq\mathbb{R}$. Here $F\in C(U)$ with U an open subset of $\mathbb{R}^{k+2}$ and

$$x^{(k)}(t)=\frac{\mathrm{d}^k x(t)}{\mathrm{d}t^k},\quad k\in\mathbb{N}_0, \tag{1.2}$$

are the ordinary derivatives of x. One frequently calls t the independent and x the dependent variable. The highest derivative appearing in F is called the order of the differential equation. A solution of the ODE (1.1) is a function $\phi\in C^k(I)$, where $I\subseteq J$ is an interval, such that

$$F\left(t,\phi(t),\phi^{(1)}(t),\ldots,\phi^{(k)}(t)\right)=0,\quad \text{for all } t\in I. \tag{1.3}$$

This implicitly implies $\left(t,\phi(t),\phi^{(1)}(t),\ldots,\phi^{(k)}(t)\right)\in U$ for all $t\in I$. Unfortunately there is not too much one can say about general differential equations in the above form (1.1). Hence we will assume that one can solve F for the highest derivative, resulting in a differential equation of the form

$$x^{(k)}=f\left(t,x,x^{(1)},\ldots,x^{(k-1)}\right).$$

By the implicit function theorem this can be done at least locally near some point $(t,y)\in U$ if the partial derivative with respect to the highest derivative does not vanish at that point, $\frac{\partial F}{\partial y_k}(t,y)\neq 0$. This is the type of differential equations we will consider from now on.

We have seen in the previous section that the case of real-valued functions is not enough and we should admit the case $x:\mathbb{R}\to\mathbb{R}^n$. This leads us to **systems of ordinary differential equations**

$$\begin{aligned} x_1^{(k)}&=f_1\left(t,x,x^{(1)},\ldots,x^{(k-1)}\right),\\ &\vdots\\ x_n^{(k)}&=f_n\left(t,x,x^{(1)},\ldots,x^{(k-1)}\right).\end{aligned}$$

Such a system is said to be **linear**, if it is of the form

$$x_i^{(k)} = g_i(t) + \sum_{l=1}^{n} \sum_{j=0}^{k-1} f_{i,j,l}(t) x_l^{(j)}.$$

It is called **homogeneous**, if $g_i(t) = 0$.

Remark 1.0.1: Any system can always be reduced to a first-order system by changing to the new set of dependent variables $y = \left(x, x^{(1)}, \ldots, x^{(k-1)}\right)$. This yields the new first-order system

$$\begin{aligned} \dot{y}_1 &= y_2, \\ &\vdots \\ \dot{y}_{k-1} &= y_k, \\ \dot{y}_k &= f(t, y). \end{aligned}$$

We can even add t to the dependent variables $z = (t, y)$, making the right hand side independent of t

$$\begin{aligned} \dot{z}_1 &= 1, \\ \dot{z}_2 &= z_3, \\ &\vdots \\ \dot{z}_k &= z_{k+1}, \\ \dot{z}_{k+1} &= f(z). \end{aligned}$$

Such a system, where f does not depend on t, is called **autonomous**. In particular, it suffices to consider the case of autonomous first-order systems which we will frequently do.

If plugging in $x = \varphi(t)$ solves (1.1), we then call $x = \varphi(t)$ the **(explicit) solution** of the ODE. Likewise, the **implicit solution** of the form $\Phi(t, x) = 0$ also solves the ODE in a sense that it determines a solution $x = \varphi(t)$. It is usually a loss of information or difficulty to write explicitly that requires one to write the implicit form instead.

Often when we integrate a derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ we find that there is a constant C along with it (an immediate consequence of the Fundamental Theorem of Calculus):

$$\int \frac{\mathrm{d}y}{\mathrm{d}x} \mathrm{d}x = y(x) + C, C \in \mathbb{R}.$$

As C varies, we can get multiple solutions, sometimes in a geometric sense when graphed on plane, called integral curves or solution curves. The solution of the following form

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

determines a class of solution due to the varying constants and is therefore called the **general solution**.

Correspondingly, we have **particular solution** when we actually settle down the constants in a general solution, according to the **initial value condition (IC)** or **boundary value condition (BC)** given. Differential equation along with initial data or boundary data is as a whole called **initial value problem (Cauchy problem)** or **boundary value problem (Dirichlet problem)**. For example, let's say we have the following initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y, y \geqslant 0, y(0) = 1$$

Then one can easily obtain a general solution

$$y(x, C) = Ce^x.$$

Plugging in initial data, we determine the constant and particular solution then

$$Ce^0 = C = 1 \Rightarrow y(x) = e^x.$$

We will present fundamental theorems of first-order ODEs in sketch. Consider the following initial value problem (IVP)

$$\dot{x} = f(t, x), \quad x(t_0) = x_0. \tag{1.4}$$

We suppose $f \in C(U, \mathbb{R}^n)$, where U is an open subset of $\mathbb{R}^{n+1}$ and $(t_0, x_0) \in U$

Suppose f is locally **Lipschitz continuous** in the second argument, uniformly with respect to the first argument, that is, for every compact set $V \subset U$ the following number

$$L = \sup_{(t,x) \neq (t,y) \in V} \frac{|f(t,x) - f(t,y)|}{|x - y|}$$

(which depends on V) is finite.

Theorem 1.0.1 (Picard-Lindelöf Theorem)**.** *Suppose $f \in C(U, \mathbb{R}^n)$, where U is an open subset of $\mathbb{R}^{n+1}$, and $(t_0, x_0) \in U$. If f is locally Lipschitz continuous in the second argument, uniformly with respect to the first, then there exists a unique local solution $\bar{x}(t)$ of the IVP (1.4).*

Proof. The proof is by using **Picard iteration**. See [11] Chapter 2 for more details. □

We also have

Theorem 1.0.2 (Peano Theorem)**.** *Suppose f is continuous on $V = [t_0, t_0 + T] \times B_\delta(x_0)$ and denote its maximum by M. Then there exists at least one solution of the initial value problem (2.11) for $t \in [t_0, t_0 + T_0]$, where $T_0 = \min\left\{T, \frac{\delta}{M}\right\}$. The analogous result holds for the interval $[t_0 - T, t_0]$.*

In many cases, f will be even differentiable. In particular, recall that $f \in C^1(U, \mathbb{R}^n)$ implies that f is locally Lipschitz continuous.

***Exercise* 1.0.1:** *Show that $f \in C^1(\mathbb{R})$ is locally Lipschitz continuous. In fact, show that*

$$|f(y) - f(x)| \leq \sup_{\varepsilon \in [0,1]} |f'(x + \varepsilon(y - x))| \, |x - y|.$$

Generalize this result to $f \in C^1(\mathbb{R}^m, \mathbb{R}^n)$.

Solution. Lemma 9.1.4 and corollary 9.1.1 conclude the above statement without proof. For a proof, apply [8] Theorem 9.19 (where one uses continuity of f' to get boundedness of the norm of f') to balls centered at each point, which are convex.

Lemma 1.0.1. Suppose $f \in C^k(U, \mathbb{R}^n), k \geq 1$, where U is an open subset of $\mathbb{R}^{n+1}$, and $(t_0, x_0) \in U$. Then the local solution $\bar{x}$ of the IVP (2.11) is C^{k+1}.

Proof. Let $k = 1$. Then $\bar{x}(t) \in C^1$ by Picard Lindelöf Theorem 1.0.1. Moreover, using $\dot{\bar{x}}(t) = f(t, \bar{x}(t)) \in C^1$ we infer $\bar{x}(t) \in C^2$. The rest follows from induction. □

Therefore, if f is smooth, i.e., infinitly differentiable, we have

Theorem 1.0.3 (Fundamental Theorem for ODEs)**.** *Let $J \subseteq \mathbb{R}$ be an open interval and $U \subseteq \mathbb{R}^n$ be an open subset, and let $f : J \times U \to \mathbb{R}^n$ be a smooth vector-valued function.*

$$\dot{x}^i(t) = f^i(t, x(t)), \qquad i = 1, \cdots, n \tag{1}$$

$$x^i(t_0) = x_0^i, \qquad i = 1, \cdots, n \tag{2}$$

Then,

(a) Existence: For any $s_0 \in J$ and $x_0 \in U$, there exist an open interval $J_0 \subseteq J$ containing s_0 and an open subset $U_0 \subseteq U$ containing x_0, such that for each $t_0 \in J_0$ and $x_0 = \left(x_0^1, \ldots, x_0^n\right) \in U_0$, there is a C^1 map $x : J_0 \to U$ that solves (1)-(2).

(b) *Uniquenes: Any two differentiable solutions to (1)-(2) agree on their common domain.*

(c) *Smoothness: Let J_0 and U_0 be as in (a), and define a map $\theta : J_0 \times J_0 \times U_0 \to U$ by letting $\theta(t, t_0, x_0) = x(t)$, where $x : J_0 \to U$ is the unique solution to (1)-(2). Then θ is smooth.*

Remark 1.0.2: Note that (1.4) is the same as (1) and (2) in Theorem 1.0.3, which is a first-order system of ODE (need not to be linear).

We have a counterpart for implicitly defined first-order ODE, which is a consequence of the above Picard-Lindelöf 1.0.1 and the implicit function theorem 9.1.2 that locally converts the implicit form to an explicit form.

***Exercise* 1.0.2:** *Write the Existence and Uniqueness theorem for implicitly defined first-order ODE.*

Extensibility of Solutions

See [11] for proofs of the following claims.

Suppose that solutions of the IVP (1.4) exist locally and are unique (e.g., f is Lipschitz). Let ϕ_1, ϕ_2 be two solutions of the IVP (1.4) defined on the open intervals I_1, I_2, respectively. Let $I = I_1 \cap I_2 = (T_-, T_+)$and let (t_-, t_+)be the maximal open interval on which both solutions coincide. I claim that $(t_-, t_+) = (T_-, T_+)$. In fact, if $t_+ < T_+$, both solutions would also coincide at $t_+$by continuity. Next, considering the IVP with initial condition $x(t_+) = \phi_1(t_+) = \phi_2(t_+)$shows that both solutions coincide in a neighborhood of $t_+$by Theorem 1.0.1. This contradicts maximality of $t_+$and hence $t_+ = T_+$. Similarly, $t_- = T_-$. Moreover, we get a solution

$$\phi(t) = \begin{cases} \phi_1(t), & t \in I_1 \\ \phi_2(t), & t \in I_2 \end{cases}$$

defined on $I_1 \cup I_2$. In fact, this even extends to an arbitrary number of solutions and in this way we get a (unique) solution defined on some maximal interval.

Theorem 1.0.4. *Suppose the IVP (1.4) has a unique local solution (e.g. the conditions of Theorem 1.0.1 are satisfied). Then there exists a unique maximal solution defined on some maximal interval $I_{(t_0,x_0)} = (T_-(t_0, x_0), T_+(t_0, x_0))$.*

Remark 1.0.3: If we drop the requirement that f is Lipschitz, we still have existence of solutions (see Theorem 1.0.2), but we already know that we loose uniqueness. Even without uniqueness, two given solutions of the IVP (1.4) can still be glued together at t_0 (if necessary) to obtain a solution defined on $I_1 \cup I_2$. Furthermore, Zorn's lemma ensures existence of maximal solutions in this case.

Now let us look at how we can tell from a given solution whether an extension exists or not.

Lemma 1.0.2. Let $\phi(t)$ be a solution of (1.4) defined on the interval (t_-, t_+). Then there exists an extension to the interval $(t_-, t_+ + \varepsilon)$ for some $\varepsilon > 0$ if and only if there exists a sequence $t_n \in (t_-, t_+)$such that

$$\lim_{n \to \infty} (t_n, \phi(t_n)) = (t_+, y) \in U.$$

Similarly for t_-.

Our final goal is to show that solutions exist for all $t \in \mathbb{R}$ if $f(t, x)$ grows at most linearly with respect to x. But first we need a better criterion which does not require a complete knowledge of the solution.

Theorem 1.0.1. *Let $\phi(t)$ be a solution of (1.4) defined on the interval (t_-, t_+). Suppose there is a compact set $[t_0, t_+] \times C \subset U$ such that $\phi(t_n) \in C$ for some sequence $t_n \in [t_0, t_+)$converging to t_+. Then there exists an extension to the interval $(t_-, t_+ + \varepsilon)$ for some $\varepsilon > 0$.*

In particular, if there is such a compact set C for every $t_+ > t_0$(C might depend on t_+), then the solution exists for all $t > t_0$. Similarly for t_-.

Proof. Let $t_n \to t_+$. By compactness $\phi(t_n)$ has a convergent subsequence and the claim follows from the previous lemma. □

The logical negation of this result is also of interest.

Theorem 1.0.2. *Let $I_{(t_0,x_0)} = (T_-(t_0,x_0), T_+(t_0,x_0))$ be the maximal interval of existence of a solution starting at $x(t_0) = x_0$. If $T_+ = T_+(t_0,x_0) < \infty$, then the solution must eventually leave every compact set C with $[t_0, T_+] \times C \subset U$ as t approaches T_+. In particular, if $U = \mathbb{R} \times \mathbb{R}^n$, the solution must tend to infinity as t approaches T_+.*

Now we come to the proof of our anticipated result.

Theorem 1.0.5. *Suppose $U = \mathbb{R} \times \mathbb{R}^n$ and for every $T > 0$ there are constants $M(T), L(T)$ such that*

$$|f(t,x)| \leq M(T) + L(T)|x|, \quad (t,x) \in [-T,T] \times \mathbb{R}^n.$$

Then all solutions of the IVP (1.4) are defined for all $t \in \mathbb{R}$.

Proof. Using the above estimate for f we have ($t_0 = 0$ without loss of generality)

$$|\phi(t)| \leq |x_0| + \int_0^t (M + L|\phi(s)|)ds, \quad t \in [0,T] \cap I.$$

Setting $\psi(t) = \frac{M}{L} + |\phi(t)|$ and applying Gronwall's inequality (see [11] Lemma 2.7) shows

$$|\phi(t)| \leq |x_0| \mathrm{e}^{LT} + \frac{M}{L}\left(\mathrm{e}^{LT} - 1\right).$$

Thus ϕ lies in a compact ball and the result follows by the previous lemma. Again, let me remark that it suffices to assume

$$|f(t,x)| \leq M(t) + L(t)|x|, \quad x \in \mathbb{R}^n,$$

where $M(t), L(t)$ are locally integrable.

***Exercise* 1.0.3:** *Show that above theorem is false (in general) if the estimate is replaced by*

$$|f(t,x)| \leq M(T) + L(T)|x|^\alpha$$

with $\alpha > 1$.

Chapter 2

General ODEs

We will talk about methods for general ODEs in this chapter, first-order ODEs, order reduction and power series method for ODE of higher orders. In the next chapter, we will present general theory of linear ODEs.

2.1 First-Order ODEs

The general form of a first-order differential equation is

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(x,y) \tag{2.1}$$

where x and y are real variables, and F is a real-valued function of x and y. We will examine several basics methods to solve them.

2.1.1 Method of Separation of Variables

Equations of the following form are called **separable**

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)\varphi(y),$$

where $f(x)$ and $\varphi(y)$ are continuous functions of x and y respectively. Separation allows us to integrate the functions separately, returning the functions to their antiderivative forms. If $\varphi(y) \neq 0$, then

$$\frac{\mathrm{d}y}{\varphi(y)} = f(x)\mathrm{d}x \Rightarrow \int \frac{\mathrm{d}y}{\varphi(y)} = \int f(x)\mathrm{d}x$$

Example 2.1.1: Solve the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y(-a+bx)}{x(d-cy)} \quad (x \geqslant 0, y \geqslant 0, a, b, c, d \in \mathbb{R})$$

solution:

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{y}{d-cy}\frac{-a+bx}{x} = \varphi(y)f(x)\\ &\Rightarrow \int \frac{\mathrm{d}y}{\frac{y}{d-cy}} = \int \frac{-a+bx}{x}\mathrm{d}x\\ &\Rightarrow d\ln|y| - cy + C_1 = bx - a\ln|x| + C_2\end{aligned}$$

where C_1 and C_2 are two constants. As $x, y \geqslant 0$, we have $|x| = x, |y| = y$, and

$$\widetilde{C}e^{d\ln y + a\ln x - cy - bx} = e^0 = 1$$

or

$$\left(e^{\ln y}\right)^d \cdot \left(e^{\ln x}\right)^a \cdot e^{-cy-bx} = y^d \cdot x^a \cdot e^{-cy-bx} = \widetilde{C},$$

where we use $\widetilde{C}$ to denote the constant however many times for which it is incorporated with other constants.

Example 2.1.2: The link gives a nice review about the derivation of the **logistic differential equation**.

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - \frac{N}{N_m}\right), N(t_0) = N_0, N(t) \geqslant 0$$

solution: We use method of separation of variables.

$$\begin{aligned} \frac{\mathrm{d}N}{\mathrm{d}t} &= rN\left(\frac{N_m - N}{N_m}\right) \\ \frac{\mathrm{d}N}{N(N_m - N)} &= \frac{r\mathrm{d}t}{N_m} \\ \int \frac{\mathrm{d}N}{N(N_m - N)} &= \int \frac{r\mathrm{d}t}{N_m} \end{aligned}$$

Let

$$\int \frac{\mathrm{d}N}{N(N_m - N)} = \int \left(\frac{A}{N_m - N} + \frac{B}{N}\right)\mathrm{d}N$$

we have

$$\begin{cases} A - B = 0 \\ BN_m = 1 \end{cases}, \text{ or } \begin{cases} A = B \\ B = \frac{1}{N_m} \end{cases}$$

Thus,

$$\begin{aligned} &\int \frac{r\mathrm{d}t}{N_m} = \frac{r}{N_m}t + C_2 \\ =&\int \frac{\mathrm{d}N}{N(N_m - N)} = \int \frac{1}{N_m\left(\frac{1}{N_m - N} + \frac{1}{N}\right)}\mathrm{d}N \\ =&\frac{1}{N_m}\left(\ln|N| - \ln|N_m - N|\right) + C_1 \\ \text{or } &|N/(N_m - N)| = \widetilde{C}e^{rt} \end{aligned}$$

Notice that the population $N(t) \geqslant 0$ with environmental capacity $N_m \geqslant N(t)$, we have $|N/(N_m - N)| = N/(N_m - N)$. Hence,

$$N(t) = \frac{N_m}{1 + \widetilde{C}e^{-rt}}$$

Pugging in the IVC $N(t_0) = N_0$, we have

$$N(t) = \frac{N_m}{1 + \left(\frac{N_m}{N_0} - 1\right)e^{r(t_0 - t)}}$$

Families of Separable First Order ODE

We introduce two families of ODE that can be converted into a separable type.

The first family is when RHS is **homogeneous function**, i.e., $F(ax, ay) = F(x, y)$. For example, $F(x, y) = \frac{6y}{x}$ is a homogeneous function. To solve

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(x, y),$$

we apply a transformation $u = \frac{y}{x}$, or $y = ux$. Then

$$\frac{\mathrm{d}ux}{\mathrm{d}x} = u + x\frac{\mathrm{d}u}{\mathrm{d}x} = F(x, ux) = F(1, u)$$
$$\frac{\mathrm{d}u}{F(1,u) - u} = \frac{\mathrm{d}x}{x}$$

which becomes a separable type.

Example 2.1.3: Solve the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{x} + \tan\frac{y}{x}$$

solution:

$$\int \frac{\mathrm{d}u}{F(1,u) - u} = \int \frac{\mathrm{d}x}{x}$$
$$\int \frac{\mathrm{d}u}{u + \tan u - u} = \int \frac{\mathrm{d}x}{x}$$
$$\ln|\sin u| = \ln|x| + \widetilde{C}$$
$$\sin u = \sin\frac{y}{x} = \pm e^{\widetilde{C}}x$$

The equation also has $\tan u = 0$, or $\sin u = 0$, as a solution. Notice that $c = \pm e^{\widetilde{C}}$ has its range $(0, +\infty) \cup (-\infty, 0) = \mathbb{R} - \{0\}$, so the general solution is then

$$\sin\frac{y}{x} = cx, c \in \mathbb{R}$$

We will use the following simple algebraic identities for the second family.

Lemma 2.1.1.

L1. if $\frac{a}{b} = \frac{c}{d}$, then $\frac{a+b}{b} = \frac{c+d}{d}$. Trivially true.

L2. if $\frac{a}{b} = \frac{c}{d}$, then $\frac{a-b}{b} = \frac{c-d}{d}$. Trivially true.

L3. if $\frac{a}{b} = \frac{c}{d}$, then $\frac{a+b}{a-b} = \frac{c+d}{c-d}$. Divide the first two.

L4. if $\frac{a}{b} = \frac{c}{d}$, then $\frac{a}{b} = \frac{c}{d} = \frac{a \pm c}{b \pm d}$. Let $\frac{a}{b} = \frac{c}{d} = k$ and directly verify it.

The second family is of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}$$

There are three situations:

1. $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = k$. By L4, we immediately have $\frac{\mathrm{d}y}{\mathrm{d}x} = k$ and $y = kx + C$
2. $\frac{a_1}{a_2} = \frac{b_1}{b_2} = k \neq \frac{c_1}{c_2}$.
   Let $u = a_2x + b_2y$ then $\frac{a_1x+b_1y}{a_2x+b_2y} = k$ gives that $a_1x + b_1y = ku$. Then

   $$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2} = \frac{ku + c_1}{u + c_2}$$

   and thus

   $$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{d(a_2x + b_2y)}{\mathrm{d}x} = a_2 + b_2\frac{\mathrm{d}y}{\mathrm{d}x} = a_2 + b_2\left(\frac{ku + c_1}{u + c_2}\right) = a_2 + b_2k + b_2\frac{c_1 - c_2k}{u + c_2}$$

   which is a separable form.

3. $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}, \exists c_i \neq 0$ then

$$\begin{cases} a_1x + b_1y + c_1 = 0 \\ a_2x + b_2y + c_2 = 0 \end{cases}$$

represents two instersecting lines on the Oxy plane. Let the intersection point be (α, β) and set

$$\begin{cases} X = x - \alpha \\ Y = y - \beta \end{cases}$$

which yields

$$\begin{cases} a_1X + b_1Y = 0 \\ a_2X + b_2Y = 0 \end{cases}$$

Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a_1X + b_1Y}{a_2X + b_2Y} = g\left(\frac{Y}{X}\right)$$

which becomes a homogeneous equation we've discussed. Notice that if $c_1 = c_2 = 0$, we may substitute $u = y/x$ directly. Besides, equations of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\left(\frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}\right)$$

can be solved by the above method too.

Other families include

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(ax + by + c)$$

where we may use substitution $u = ax + by + c$, and

$$\frac{y}{x}\frac{\mathrm{d}y}{\mathrm{d}x} = f(xy)$$

where we may use substitution $u = xy$.

Example 2.1.4: solve

$$(x + y)\mathrm{d}y + (x - y)\mathrm{d}x = 0$$

solution:

$$\begin{aligned} (x + y)\mathrm{d}y &= (y - x)\mathrm{d}x \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{-1 + y/x}{1 + y/x} = g\left(\frac{y}{x}\right) \\ u = \frac{y}{x} &\Rightarrow \int \frac{\mathrm{d}u}{\frac{-u^2-1}{u+1}} = \int \frac{\mathrm{d}x}{x} \end{aligned}$$

Since

$$-\int \frac{u+1}{u^2+1}\mathrm{d}u = -\int \frac{u}{u^2+1}\mathrm{d}u - \int \frac{\mathrm{d}u}{u^2+1} = -\frac{1}{2}\ln|u^2+1| - \arctan u + C_1$$

and

$$\int \frac{\mathrm{d}x}{x} = \ln|x| + C_2$$

we have the implicit general solution

$$\ln\sqrt{x^2 + y^2} + \arctan\frac{y}{x} = \widetilde{C}$$

***Exercise* 2.1.1:** *Solve the following ODEs*

1. $x\frac{dy}{dx} - y + \sqrt{x^2 - y^2}$. *Hint: divide by* $\sqrt{x^2} = \mathrm{sgn}(x)x$.

2. $\frac{dy}{dx} = (x+y)^2$. *Hint: let* $u = x + y$.
3. $\frac{dy}{dx} = \frac{1}{(x+y)^2}$. *Hint: let* $u = x + y$.

***Exercise* 2.1.2:** *Solve the following ODEs*

1. $\frac{dy}{dx} = (x+1)^2 + (4y+1)^2 + 8xy + 1$.
2. $y(1 + x^2y^2)dx = xdy$.
3. $\frac{y}{x}\frac{dy}{dx} = \frac{2+x^2y^2}{x-x^2y^2}$. *Hint: the second and the third are of the family* $\frac{x}{y}\frac{dy}{dx} = f(xy)$.

2.1.2 Linear ODE and Method of Variation of Parameters

Equations of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = P(x)y + Q(x), \tag{2.2}$$

where $P(x)$ and $Q(x)$ are continuous functions over x, is called **first order linear ODE**. When $Q(x) \neq 0$, we call it **nonhomogeneous**, otherwise **homogeneous**. Notice that the previously defined homogeneity refers to the LHS being a homogeneous function, while the homogeneity there refers to the fact that the equation is written as a function of derivatives of the dependent variable y (including 0-order derivative), altogether equaling 0 rather than some function of x. Namely, the most general form

$$F\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^n y}{\mathrm{d}x^n}\right) = 0$$

becomes

$$F\left(y, \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^n y}{\mathrm{d}x^n}\right) = 0$$

rather than

$$F\left(y, \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^n y}{\mathrm{d}x^n}\right) = f(x)$$

We observe that the homogeneous first order linear ODE is of a separable type. For

$$\frac{\mathrm{d}y}{\mathrm{d}x} = P(x)y$$

We have

$$\int \frac{\mathrm{d}y}{y} = \int P(x)\mathrm{d}x$$
$$y(x) = C \cdot \exp\left(\int P(x)\mathrm{d}x\right)$$

We now conjecture that the nonhomogeneous one has a solution like

$$y(x) = C(x) \cdot \exp\left(\int P(x)\mathrm{d}x\right)$$

Then

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{C(x) \cdot \exp\left(\int P(x)\mathrm{d}x\right)}{\mathrm{d}x} \\ &= \frac{\mathrm{d}C(x)}{\mathrm{d}x}\exp\left(\int P(x)\mathrm{d}x\right) + \exp\left(\int P(x)\mathrm{d}x\right)C(x)P(x) \\ &= \frac{\mathrm{d}C(x)}{\mathrm{d}x}\exp\left(\int P(x)\mathrm{d}x\right) + P(x)y\end{aligned}$$

Since

$$\frac{\mathrm{d}y}{\mathrm{d}x} = P(x)y + Q(x)$$

we have

$$P(x)y + Q(x) = \frac{\mathrm{d}C(x)}{\mathrm{d}x}\exp\left(\int P(x)\mathrm{d}x\right) + P(x)y$$

and

$$\begin{aligned} Q(x) &= \frac{\mathrm{d}C(x)}{\mathrm{d}x}\exp\left(\int P(x)\mathrm{d}x\right) \\ \Rightarrow \frac{\mathrm{d}C(x)}{\mathrm{d}x} &= Q(x)\exp\left(-\int P(x)\mathrm{d}x\right) \\ \Rightarrow C(x) &= \int\left[Q(x)\exp\left(-\int P(x)\mathrm{d}x\right)\right]\mathrm{d}x + \widetilde{C} \end{aligned}$$

The solution to the equation (2.2) is then

$$y(x) = \left\{\int\left[Q(x)\exp\left(-\int P(x)\mathrm{d}x\right)\right]\mathrm{d}x + \widetilde{C}\right\}\exp\left(\int P(x)\mathrm{d}x\right) \tag{2.3}$$

Since we change the constant C by $C(x)$, we call the above method **method of variation of parameters**. Later on, we will see this method be applied to higher order ODE. Note that when calculating the integrals, we let all constants to be 0 and add $\widetilde{C}$ instead.

Example 2.1.5: Solve ODE

$$(x+1)\frac{\mathrm{d}y}{\mathrm{d}x} - ny = e^x(x+1)^n,$$

where n is a constant.
solution: the ODE is a first order linear ODE, as it can be written as

$$\frac{\mathrm{d}y}{\mathrm{d}x} = (\frac{n}{n+1})y + e^x(x+1)^n$$

Let $P(x) = \frac{n}{x+1}$ and $Q(x) = e^x(x+1)^n$. Since

$$\begin{aligned} \exp\left(\int P(x)\mathrm{d}x\right) &= e^{n\ln|x+1|} = |x+1|^n = \begin{cases} \mathrm{sgn}(x+1)(x+1)^n, n = 2k+1 \\ (x+1)^n, n = 2k \end{cases} \\ &= (-1)^{n-1}\mathrm{sgn}(x+1)(x+1)^n \end{aligned}$$

we have

$$\begin{aligned} y(x) &= \left\{\int\left[Q(x)\exp\left(-\int P(x)\mathrm{d}x\right)\right]\mathrm{d}x\widetilde{C}\right\}\exp\left(\int P(x)\mathrm{d}x\right) \\ &= \left\{\int\left[(-1)^{n-1}\mathrm{sgn}(x+1)e^x(x+1)^n\frac{1}{(x+1)^n}\right]\mathrm{d}x\widetilde{C}\right\}(-1)^{n-1}\mathrm{sgn}(x+1)(x+1)^n \\ &= [(-1)^{n-1}\mathrm{sgn}(x+1)]^2(x+1)^n\left(\int e^x\mathrm{d}x + \widetilde{C}\right) \\ &= (x+1)^n(e^x + \widetilde{C}) \end{aligned}$$

Example 2.1.6: This one is a bit different. One can regard x as a function of y to solve it.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{2x - y^2}$$

Let

$$\frac{\mathrm{d}x(y)}{\mathrm{d}y} = \left(\frac{2}{y}\right)x - y, P(y) = \frac{2}{y}, Q(y) = -y$$

then

$$\begin{aligned}
x &= \left[\int Q(y)\exp\left(-\int P(y)\mathrm{d}y\right)\mathrm{d}y + \widetilde{C}\right]\exp\left(-\int P(y)\mathrm{d}y\right) \\
&= \left[\int -ye^{-2\ln|y|}\mathrm{d}y + \widetilde{C}\right]e^{2\ln|y|} \\
&= (-\ln|y| + \widetilde{C})y^2
\end{aligned}$$

Example 2.1.7: Solve the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 6\frac{y}{x} - xy^2 \tag{2.4}$$

We often call the ODE of the following form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = P(x)y + Q(x)y^n$$

the n**-th order Bernoulli equation**, where $n \neq 0, 1$ and $P(x), Q(x)$ are continuous functions over x. If $y \neq 0$, we multiply the two sides by y^{-n}:

$$y^{-n}\frac{\mathrm{d}y}{\mathrm{d}x} = y^{1-n}P(x) + Q(x)$$

Let $z = y^{1-n}$ be the substitution, then

$$\frac{\mathrm{d}z}{\mathrm{d}x} = (1-n)y^{-n}\frac{\mathrm{d}y}{\mathrm{d}x} = (1-n)[y^{1-n}P(x) + Q(x)] = (1-n)zP(x) + (1-n)Q(x)$$

which is a nonhomogeneous first order linear ODE, solvable by method of variation of parameter. We now solve the equation (2.4). First observe that $y = 0$ is a solution, and then

$$\begin{aligned}
y^{-2}\frac{\mathrm{d}y}{\mathrm{d}x} &= 6\frac{y^{-1}}{x} - x \quad \text{multiply two sides by } y^{-2} \\
\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}(y^{-1})}{\mathrm{d}x} &= \frac{-6}{x}z + x \quad \text{letting } z = y^{-1}
\end{aligned}$$

Let $P(x) = -6/x$ and $Q(x) = x$. For the above first order linear ODE,

$$\frac{1}{y} = z = \left[\int Q(x)\exp\left(-\int P(x)\mathrm{d}x\right)\mathrm{d}x + \widetilde{C}\right]\exp\left(-\int P(x)\mathrm{d}x\right) = \frac{x^2}{8} + \frac{\widetilde{C}}{x^6}, \text{or } y = \frac{8x^6}{x^8 + \widetilde{C}}$$

***Exercise* 2.1.3:** *Solve the following ODEs:*

1. $\frac{dy}{dx} = \frac{y}{x+y^3}$.
2. $\frac{ds}{dt} = -s\cos t + \frac{1}{2}\sin 2t$.
3. $\frac{dy}{dx} = \frac{x^4+y^3}{xy^2}$.
4. $\frac{dy}{dx} = \frac{ay}{x} + \frac{x+1}{x}$. *Hint: The second to last is a Bernoulli eq. For the last one, discuss cases $a = 0$, $a = 1$, and $a \neq 0, 1$.*

2.1.3 Exact Equations and Integrating factor

Exact Equations

Consider a general first order ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y)$$

We may write it as $dy - f(x,y)dx = 0$ and compare it with $M(x,y)dx + N(x,y)dy = 0$. **Exact equation** is an ODE where the LHS of the above equation is exactly the total derivative of a bivariate function $u(x,y)$. Namely,

$$\mathrm{d}u = \frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y = M(x,y)\mathrm{d}x + N(x,y)\mathrm{d}y = 0 \tag{2.5}$$

So the general solution of the ODE is $u(x,y) = C$. In particular, we need to find a function $u(x,y)$ such that, by (2.5),

$$\frac{\partial u}{\partial x} = M(x,y), \frac{\partial u}{\partial y} = N(x,y)$$
$$\Rightarrow \frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right) \xlongequal{\text{clairaut's thm}} \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) = \frac{\partial N}{\partial x}$$

Thus $\partial M/\partial y = \partial N/\partial x$ is the necessary condition for the equation $M(x,y)dx + N(x,y)dy$ to be an exact equation. We now prove that it is also the sufficient condition for the equation to be exact. Consider

$$u(x,y) = \int M(x,y)\mathrm{d}x + \int \left[N - \frac{\partial}{\partial y}\int M(x,y)\mathrm{d}x\right]\mathrm{d}y$$

and verify that $u_x = M, u_y = N$ using $\partial M/\partial y = \partial N/\partial x$. Therefore, the general solution of the exact equation is

$$\int M(x,y)\mathrm{d}x + \int \left[N - \frac{\partial}{\partial y}\int M(x,y)\mathrm{d}x\right]\mathrm{d}y = C \tag{2.6}$$

Example 2.1.8:

$$2(3xy^2 + 2x^3)\mathrm{d}x + 3(2x^2y + y^2)\mathrm{d}y = 0$$

solution:

$$M(x,y) = 6xy^2 + 4x^3, N(x,y) = 6x^2y + 3y^2$$

Thus, the necessary and sufficient condition for it being exact $\partial M/\partial y = \partial N/\partial x (= 12xy)$ is thus satisfied. Then

$$u(x,y) = \int 6xy^2 + 4x^3\mathrm{d}x + \varphi(y) = 3x^2y^2 + x^4 + \varphi(y)$$

and

$$\frac{\partial u}{\partial y} = 6x^2y + \frac{\mathrm{d}\varphi(y)}{\mathrm{d}y} = 3(2x^2y + y^2) \Rightarrow \varphi'(y) = 3y^2, \varphi(y) = y^3 + C$$

Thus, the general solution is

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

Example 2.1.9:

$$\left[\frac{y^2}{(x-y)^2} - \frac{1}{x}\right]\mathrm{d}x + \left[\frac{1}{y} - \frac{x^2}{(x-y)^2}\right]\mathrm{d}y = 0$$

Notice that the equation can be written as two parts:

$$\left[\frac{y^2}{(x-y)^2} - \frac{1}{x}\right]\mathrm{d}x + \left[\frac{1}{y} - \frac{x^2}{(x-y)^2}\right]\mathrm{d}y = \frac{y^2\mathrm{d}x - x^2\mathrm{d}y}{(x-y)^2} + \frac{y\mathrm{d}x - x\mathrm{d}y}{xy}$$

By the procedures to find u, it is easy to see that

$$u_1 = -\frac{xy}{x-y} \text{ for } \mathrm{d}u_1 = \frac{y^2\mathrm{d}x - x^2\mathrm{d}y}{(x-y)^2}$$
$$u_2 = \ln\left|\frac{y}{x}\right| \text{ for } \mathrm{d}u_2 = \frac{y\mathrm{d}x - x\mathrm{d}y}{xy}$$

By linearity of d, we see that $d(u_1 + u_2) = 0$ and the general solution of the original ODE is therefore

$$\ln\left|\frac{y}{x}\right| - \frac{xy}{x-y} = C$$

In fact, the linearity of the differentiation as an operator implies the superposition of differential equation.

some common total derivatives:

- $$x\text{d}y + y\text{d}x = d(xy)$$
- $$\frac{y\text{d}x - x\text{d}y}{y^2} = \text{d}\left(\frac{x}{y}\right)$$
- $$\frac{-y\text{d}x + x\text{d}y}{x^2} = \text{d}\left(\frac{y}{x}\right)$$
- $$\frac{y\text{d}x - x\text{d}y}{xy} = \text{d}\left(\ln\left|\frac{x}{y}\right|\right)$$
- $$\frac{y\text{d}x - x\text{d}y}{x^2+y^2} = \text{d}\left(\arctan\frac{x}{y}\right)$$
- $$\frac{y\text{d}x - x\text{d}y}{x^2-y^2} = \frac{1}{2}\text{d}\left(\ln\left|\frac{x-y}{x+y}\right|\right)$$

Exercise* 2.1.4*: *Solve the following ODEs:*

1. $(x^2+y)dx + (x-2y)dy = 0$.
2. $\left(\cos x + \frac{1}{y}\right)dx + \left(\frac{1}{y} - \frac{x}{y^2}\right)dy = 0$. *Hint: for the the second one, find out a common total derivative from the above list.*

Integrating Factor

The exact equations can be solved by direct integration, so transforming a non-exact equation to an exact one becomes crucial for solving certain types of equations.

If there exists a function $\mu = \mu(x,y) \neq 0$ such that

$$\mu M(x,y)\text{d}x + \mu N(x,y)\text{d}y = 0 \tag{2.7}$$

becomes an exact equation, namely theres is a function v such that $\mu M\text{d}x + \mu N\text{d}y = \text{d}v$, we call the function μ the **Integrating factor**. The general solution of the equation is then $v(x,y) = C$.

There may exist multiple integrating factors for a single equation. For example,

$$ydx - xdy = 0$$

has factors $1/x^2$, $1/y^2$, $1/xy$, and $1/(x^2 \pm y^2)$. I can be proved that if there is a solution for the equation, then there exists an integrating factor, and such is not unique.

Since the necessary and sufficient condition of exactness of an equation $M\text{d}x + N\text{d}y = 0$ is $\partial M/\partial y = \partial N/\partial x$, we have

$$\frac{\partial(\mu M)}{\partial y} = \frac{\partial(\mu N)}{\partial x}$$

for (2.7) to be exact. Namely,

$$N\frac{\partial\mu}{\partial x} - M\frac{\partial\mu}{\partial y} = \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right)\mu$$

which is a first order partial differential equation that can be more difficult to solve. However, sometimes finding a particular solution is much easier under some special conditions. For example, the equation

$$M(x,y)\text{d}x + N(x,y)\text{d}y = 0 \tag{2.8}$$

has an integrating factor

$$\exp\left(\int \psi(x)\mathrm{d}x\right)$$

when

$$\psi(x) = \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right)/N$$

is a only a function of x.

Likewise, (2.8) has an integrating factor

$$\exp\left(\int \varphi(y)\mathrm{d}y\right)$$

when

$$\varphi(y) = \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right)/-M$$

is only a function of y.

Example 2.1.10: solve the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y} + \sqrt{1+\left(\frac{x}{y}\right)^2}, (y>0)$$

solution:
The equation can be written as, by multiplying y at both sides,

$$x\mathrm{d}x + y\mathrm{d}y = \sqrt{x^2+y^2}\mathrm{d}x$$

or

$$\frac{1}{2}\mathrm{d}(x^2+y^2) = \sqrt{x^2+y^2}\mathrm{d}x$$

The equation therefore has the integrating factor

$$\mu = \frac{1}{\sqrt{x^2+y^2}}$$

Then

$$\begin{aligned}
\frac{\mathrm{d}(x^2+y^2)}{2\sqrt{x^2+y^2}} &= \mathrm{d}x \\
\int \frac{\mathrm{d}(x^2+y^2)}{2\sqrt{x^2+y^2}} &= \int \mathrm{d}x \\
\sqrt{x^2+y^2} &= x + C \Rightarrow y^2 = C(C+2x)
\end{aligned}$$

Exercise 2.1.5: *Solves the following ODEs:*

1. $(e^x + 3y^2)dx + 2xydy = 0$.
2. $ydx - xdy = (x^2+y^2)dx$.
3. $[x\cos(x+y) + \sin(x+y)]dx + x\cos(x+y)dy = 0$.
4. $x(4ydx+2xdy) + y^3(3ydx+5xdy) = 0$. *Hint: for the last one, use the method of undetermined coefficients for the integrating factor $\mu = x^m y^n$. Plugging it into the necessary and sufficient condition of exact equation.*

2.1.4 First Order Implicit ODE and Its Parametrization

The common form of a first order implicit ODE is often written as

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

If we can solve the derivative $y' = f(x, y)$ from the above equation, we can use the methods in previous sections (separation, exact euqatio, ...).

If not, we may parametrize the equation in the following four types:

1. $$y = f\left(x, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

   introduce the parameter $p = dy/dx$, then

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

   differentiate both sides:

   $$p = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial f(x,p)}{\partial x} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial f}{\partial p}\frac{\partial p}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}x}$$

   Since $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial p}$ are functions about x and p, the above equation is a first order ODE about x and p.

   Example 2.1.11: solve

   $$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^3 + 2x\frac{\mathrm{d}y}{\mathrm{d}x} - y = 0$$

   solution: First notice that $y = 0$ is a solution. Then observe that the equation is of the form $y = f(x, dy/dx)$ where we can introduce $p = y'$

   $$y = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^3 + 2x\frac{\mathrm{d}y}{\mathrm{d}x} = p^3 + 2xp$$

   and differentiate it:

   $$p = 3p^2\frac{\mathrm{d}p}{\mathrm{d}x} + 2x\frac{\mathrm{d}p}{\mathrm{d}x} + 2p$$

   or

   $$3p^2\mathrm{d}p + 2x\mathrm{d}p + p\mathrm{d}x = 0$$

   When $p \neq 0$, we notice that p can act as an integrating factor. Multiply it,

   $$(3p^3 + 2xp)\mathrm{d}p + p^2\mathrm{d}x = 0$$

   Integrating it, we have

   $$\frac{3p^4}{4} + xp^2 = c \Rightarrow x = \frac{c - \frac{3}{4}p^4}{p^2}, y = p^3 + \frac{2\left(c - \frac{3}{4}p^4\right)}{p}$$

2. $$x = f\left(y, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

   The same as the first type. For $p = dy/dx$

   $$\frac{1}{p} = \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}y} \tag{2.9}$$

   We reexamine example 2.1.11:

Example 2.1.12: Solve the equation by the above method

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^3 + 2x\frac{\mathrm{d}y}{\mathrm{d}x} - y = 0$$

solution:
We need to express x in terms of y and dy/dx. It is easy to see

$$x = \frac{y - p^3}{2p}$$

for $p = y' \neq 0$. Then by formula (2.9)

$$\frac{1}{p} = \frac{1}{2p} + \left(\frac{-y}{2p^2} - p\right)\frac{\mathrm{d}p}{\mathrm{d}y}$$

and

$$\begin{aligned}
&p\mathrm{d}y + (2p^3 + y)\mathrm{d}p = 0 \\
&\mathrm{d}(yp) + 2p^3\mathrm{d}p = 0 \\
&\int \mathrm{d}(yp) + 2p^3\mathrm{d}p = yp + \frac{p^4}{2} = C \\
&y = \frac{c - p^4}{2p}, x = \frac{y - p^3}{2p} = \frac{c - 3p^4}{4p^2}
\end{aligned}$$

which is the parametrized general solution. Of course, $y = 0$ is still a solution of the equation.

3.

$$F\left(x, \frac{\mathrm{d}y}{\mathrm{d}x}\right) = 0 \tag{2.10}$$

Let

$$p = \frac{\mathrm{d}y}{\mathrm{d}x} = y'$$

From a geometric viewpoint, $F(x, p) = 0$ represents a curve on the Oxp plane. Parametrize it properly,

$$\begin{cases} x = \varphi(t) \\ p = \psi(t) \end{cases}$$

where t is the parameter. Notice that on any integral curve solved from the equation (2.10), the relation $\mathrm{d}y = p\mathrm{d}x$ is always true. Plugging the parametrization into the equation (2.10), we have

$$\mathrm{d}y = \psi(t)\varphi'(t)\mathrm{d}t$$

Integrating both sides, we have

$$y = \int \psi(t)\varphi'(t)\mathrm{d}t + c$$

so the general solution of the equation is

$$\begin{cases} x = \varphi(t) \\ y = \int \psi(t)\varphi'(t)\mathrm{d}t + c \end{cases}$$

Example 2.1.13: solve the ODE

$$x^3 + y'^3 - 3xy' = 0,$$

where $y' = dy/dx$. Let $y' = p = tx$, then

$$\begin{cases} x = \frac{3t}{1+t^3} \\ p = \frac{3t^2}{1+t^3} \end{cases}$$

Thus,

$$\mathrm{d}y = p\mathrm{d}x = \frac{3t^2}{1+t^3}\left(\frac{3t}{1+t^3}\right)' dt = \frac{9(1-2t^3)t^2}{(1+t^3)^3}\mathrm{d}t$$

Integrating it, we have

$$y = \int \frac{9(1-2t^3)t^2}{(1+t^3)^3}\mathrm{d}t = \frac{3}{2}\frac{1+4t^3}{(1+t^3)^2} + c$$

The general solution in its parametrized form is thus

$$\begin{cases} x = \frac{3t}{1+t^3} \\ y = \frac{3}{2}\frac{1+4t^3}{(1+t^3)^2} + c \end{cases}$$

4.

$$F\left(y, \frac{\mathrm{d}y}{\mathrm{d}x}\right) = 0$$

Let $p = y'$, and then

$$\begin{cases} y = \varphi(t) \\ p = \psi(t) \end{cases}$$

As $\mathrm{d}y = p\mathrm{d}x \Rightarrow \varphi'(t)\mathrm{d}t = \psi(t)\mathrm{d}x$, we have

$$\mathrm{d}x = \frac{\varphi'(t)}{\psi(t)}\mathrm{d}t$$

and

$$x = \int \frac{\varphi'(t)}{\psi(t)}\mathrm{d}t + c$$

Similarly, the general solution of this equation is

$$\begin{cases} x = \int \frac{\varphi'(t)}{\psi(t)}\mathrm{d}t + c \\ y = \varphi(t) \end{cases}$$

Besides, it is easy to see that $y = k$ is also a solution to $F(y, 0) = 0$ and is thus a solution to the equation.

Example 2.1.14: solve the ODE

$$y^2(1-y') = (2-y')^2$$

solution:
let $2 - y' = yt$, or $y' = 2 - yt$, thus

$$y^2(1-y') = y^2(yt-1) = (yt)^2$$

and

$$y = \frac{1}{t} + t, y' = 1 - t^2$$

Thus,

$$\mathrm{d}x = \frac{\mathrm{d}y}{y'} = -\frac{1}{t^2}\mathrm{d}t$$

Integrating it we have $x = 1/t + c$ and

$$y = x + \frac{1}{x-c} - c$$

. Also notice that $F(y, y' = 0) = 0$ also has solution $y = \pm 2$.

Exercise 2.1.6: *solve the following ODEs:*

1. $xy'^3 = 1 + y'$.
2. $y = y'^2 e^{y'}$.
3. *Let a be a constant.*

$$y(1+y'^2) = 2a$$

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

2.2 Reduction of Order

We now consider ODEs of higher orders (2.11) and the method of **reduction of order**.

$$F(t,x,x',\cdots,x^{(n)})=0. \tag{2.11}$$

There are three types of such equation that are reducible.

(1) When the equation (2.11) does not contain $x,x',\cdots,x^{(k-1)}$ for some $1\leq k\leq n$. That is, the equation is of the form

$$F(t,x^{(k)},x^{(k+1)},\cdots,x^{(n)})=0. \tag{2.12}$$

In this case, we let $y=x^{(k)}$.

Example 2.2.1: Consider the ODE

$$x^{(5)}-\frac{1}{t}x^{(4)}=0.$$

Solution. Let $y=x^{(4)}$. Then

$$y'-\frac{1}{t}y=0.$$

The solution is $y=ct$. Thus, the solution of the original equation is

$$x=\sum_{j=0}^{5}c_jt^j$$

with $c_4=0$.

(2) When the equation (2.11) does not contain independent variable t. That is, the equation is of the form

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

This is an autonomous euqation. Letting $y=x'$ can reduce the order to $n-1$: notice that

$$\begin{aligned}
&\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}=\frac{\mathrm{d}y}{\mathrm{d}x}y\\
&\frac{\mathrm{d}^3x}{\mathrm{d}t^3}=y\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+y^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\\
&\cdots\cdots\\
&x^{(n)}=G\left(y,\frac{\mathrm{d}y}{\mathrm{d}x},\cdots,\frac{\mathrm{d}^{(n-1)}y}{\mathrm{d}x^{n-1}}\right).
\end{aligned}$$

Plug $x',x'',\cdots,x^{(k)}$ into F to get

$$K\left(x,y,\frac{\mathrm{d}y}{\mathrm{d}x},\cdots,\frac{\mathrm{d}^{(n-1)}y}{\mathrm{d}x^{n-1}}\right)=0.$$

Example 2.2.2: Consider the ODE

$$xx''+(x')^2=0.$$

Solution. Plug $x'=y,x''=\frac{\mathrm{d}y}{\mathrm{d}x}y$ into the equation to get

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x}xy+y^2&=0\\
\frac{\mathrm{d}y}{\mathrm{d}x}&=-\frac{y}{x}\quad\text{and } y=0 \text{ is also a solution.}\\
y&=\frac{c}{x}\text{ or } y=0
\end{aligned}$$

Therefore, the general solution of original equation is

$$x^2=c_1t+c_2.$$

(3) Liouvelle's method for homogeneous linear ODE

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \cdots + a_1(t)x = 0$$

We shall only present the two dimensional case:

$$y'' + P(x)y' + Q(x)y = 0 \tag{2.13}$$

If one knows a nowhere-zero solution y_1 of (2.13), then one could find another solution, independent of y_1. We try to find function $u(x)$ such that $y = y_1 u$ also satisfies equation (2.13), namely

$$\begin{aligned} 0 = Q\left(y_1 u\right) + P\left(y_1' u + y_1 u'\right) + \left(y_1'' u + 2y_1' u' + y_1 u''\right) \\ = \left(Qy_1 + Py_1' + y_1''\right) u + Py_1 u' + 2y_1' u' + y_1 u'' = \left(Py_1 + 2y_1'\right) u' + y_1 u''. \end{aligned}$$

Therefore we should have

$$\frac{du'}{u'} = -\frac{Py_1 + 2y_1'}{y_1} dx$$

which after integration becomes

$$\log u' = -\int P dx - 2\log y_1$$

hence

$$u = \int y_1^{-2} e^{-\int P dx}.$$

Retracing back, our analysis shows that

$$y_2 = y_1 \int y_1^{-2} e^{-\int P dx} dx,$$

is another solution to equation (2.13). To check independency of y_1 and y_2, we compute their Wronskian

$$W\left(y_1, y_2\right) = y_1\left(y_1' u + y_1 u'\right) - y_1' y_1 u = y_1^2 u' = e^{-\int P dx},$$

which never vanishes. We gather what we have proved in the following proposition.

Proposition 2.2.1. If y_1 is a nowhere-zero solution of the homogeneous equation (2.13), then another independent solution is given by

$$y_2 = y_1 \int y_1^{-2} e^{-\int P dx} dx.$$

***Exercise* 2.2.1:** *(a) Find the general solution of $xy'' - (x+2)y' + 2y = 0$ knowing that $y = e^x$ is one particular solution.*

(b) Find the general solution of $\left(1 - x^2\right) y'' - 2xy' + 2y = 0$ knowing that $y = x$ is one particular solution.

2.3 Power Series Method

Thanks to Jiří Lebl's latex note

Note: 1 or 1.5 lecture, §8.1 in [EP],§5.1 in [BD]

Many functions can be written in terms of a power series

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k.$$

If we assume that a solution of a differential equation is written as a power series, then perhaps we can use a method reminiscent of undetermined coefficients. That is, we will try to solve for the numbers a_k. Before we carry out this process, we review some results and concepts about power series.

As we said, a **power series** is an expression such as

$$\sum_{k=0}^{\infty} a_k (x-x_0)^k = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + a_3(x-x_0)^3 + \cdots, \tag{2.14}$$

where $a_0, a_1, a_2, \ldots, a_k, \ldots$ and x_0 are constants. Let

$$S_n(x) = \sum_{k=0}^{n} a_k (x-x_0)^k = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + a_3(x-x_0)^3 + \cdots + a_n(x-x_0)^n,$$

denote the so-called **partial sum**. If for some x, the limit

$$\lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} a_k (x-x_0)^k$$

exists, we say the series 2.14 **converges** at x. At $x = x_0$, the series always converges to a_0. When 2.14 converges at any other $x \neq x_0$, we say 2.14 is a **convergent power series**, and we write

$$\sum_{k=0}^{\infty} a_k (x-x_0)^k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k (x-x_0)^k.$$

If the series does not converge for any point $x \neq x_0$, we say that the series is **divergent**.

Example 2.3.1: The series

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

is convergent for any x. Recall that $k! = 1 \cdot 2 \cdot 3 \cdots k$ is the factorial. By convention we define $0! = 1$. You may recall that this series converges to e^x.

We say that 2.14 **converges absolutely** at x whenever the limit

$$\lim_{n\to\infty} \sum_{k=0}^{n} |a_k|\, |x-x_0|^k$$

exists. That is, the series $\sum_{k=0}^{\infty} |a_k|\, |x-x_0|^k$ is convergent. If 2.14 converges absolutely at x, then it converges at x. However, the opposite implication is not true.

Example 2.3.2: The series

$$\sum_{k=1}^{\infty} \frac{1}{k} x^k$$

converges absolutely for all x in the interval $(-1, 1)$. It converges at $x = -1$, as $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ converges (conditionally) by the alternating series test. The power series does not converge absolutely at $x = -1$, because $\sum_{k=1}^{\infty} \frac{1}{k}$ does not converge. The series diverges at $x = 1$.

If a power series converges absolutely at some x_1, then for all x such that $|x - x_0| \leq |x_1 - x_0|$ (that is, x is closer than x_1 to x_0) we have $|a_k(x-x_0)^k| \leq |a_k(x_1-x_0)^k|$ for all k. As the numbers $|a_k(x_1-x_0)^k|$ sum to some finite limit, summing smaller positive numbers $|a_k(x-x_0)^k|$ must also have a finite limit. Hence, the series must converge absolutely at x.

Theorem 2.3.1. *For a power series 2.14, there exists a number ρ (we allow $\rho = \infty$) called the* ***radius of convergence*** *such that the series converges absolutely on the interval $(x_0 - \rho, x_0 + \rho)$ and diverges for $x < x_0 - \rho$ and $x > x_0 + \rho$. We write $\rho = \infty$ if the series converges for all x.*

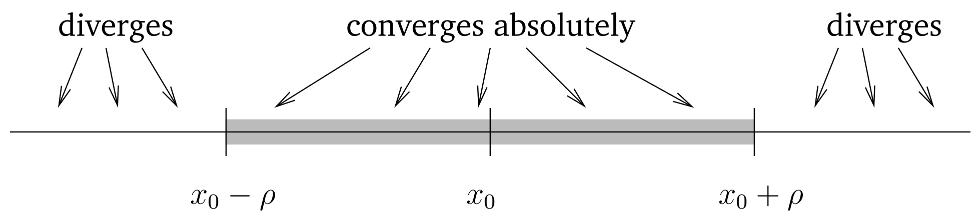

Figure 2.1: Convergence of a power series.

See 2.1. In 2.3.1 the radius of convergence is $\rho = \infty$ as the series converges everywhere. In 2.3.2 the radius of convergence is $\rho = 1$. We note that $\rho = 0$ is another way of saying that the series is divergent.

A useful test for convergence of a series is the **ratio test**. Suppose that

$$\sum_{k=0}^{\infty} c_k$$

is a series and the limit

$$L = \lim_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right|$$

exists. Then the series converges absolutely if $L < 1$ and diverges if $L > 1$.

We apply this test to the series 2.14. Let $c_k = a_k (x - x_0)^k$ in the test. Compute

$$L = \lim_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right| = \lim_{k\to\infty} \left| \frac{a_{k+1}(x-x_0)^{k+1}}{a_k (x-x_0)^k} \right| = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| |x - x_0|.$$

Define A by

$$A = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

Then the series 2.14 converges absolutely if $1 > L = A|x - x_0|$. If $A > 0$, then the series converges absolutely if $|x - x_0| < {}^1/_A$, and diverges if $|x - x_0| > {}^1/_A$. That is, the radius of convergence is ${}^1/_A$. If $A = 0$, then the series always converges.

A similar test is the **root test**. Suppose

$$L = \lim_{k\to\infty} \sqrt[k]{|c_k|}$$

exists. Then $\sum_{k=0}^{\infty} c_k$ converges absolutely if $L < 1$ and diverges if $L > 1$. We can use the same calculation as above to find A. Let us summarize.

Theorem 2.3.2 (Ratio and root tests for power series)**.** *Consider a power series*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

such that

$$A = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| \qquad \text{or} \qquad A = \lim_{k\to\infty} \sqrt[k]{|a_k|}$$

exists. If $A = 0$, then the radius of convergence of the series is ∞. Otherwise, the radius of convergence is ${}^1/_A$.

Example 2.3.3: Suppose we have the series

$$\sum_{k=0}^{\infty} 2^{-k} (x - 1)^k.$$

First we compute the limit in the ratio test,

$$A = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k\to\infty} \left| \frac{2^{-k-1}}{2^{-k}} \right| = \lim_{k\to\infty} 2^{-1} = 1/2.$$

Therefore the radius of convergence is 2, and the series converges absolutely on the interval $(-1, 3)$. And we could just as well have used the root test:

$$A = \lim_{k\to\infty} \lim_{k\to\infty} \sqrt[k]{|a_k|} = \lim_{k\to\infty} \sqrt[k]{|2^{-k}|} = \lim_{k\to\infty} 2^{-1} = 1/2.$$

Example 2.3.4: Consider

$$\sum_{k=0}^{\infty} \frac{1}{k^k} x^k.$$

Compute the limit for the root test,

$$A = \lim_{k\to\infty} \sqrt[k]{|a_k|} = \lim_{k\to\infty} \sqrt[k]{\left| \frac{1}{k^k} \right|} = \lim_{k\to\infty} \sqrt[k]{\left| \frac{1}{k} \right|^k} = \lim_{k\to\infty} \frac{1}{k} = 0.$$

So the radius of convergence is ∞: the series converges everywhere. The ratio test would also work here.

The root or the ratio test does not always apply. That is the limit of $\left|\frac{a_{k+1}}{a_k}\right|$ or $\sqrt[k]{|a_k|}$ might not exist. There exist more sophisticated ways of finding the radius of convergence, but those would be beyond the scope of this chapter. The two methods above cover many of the series that arise in practice. Often if the root test applies, so does the ratio test, and vice versa, though the limit might be easier to compute in one way than the other.

Functions represented by power series are called **analytic functions**. Not every function is analytic, although the majority of the functions you have seen in calculus are.

An analytic function $f(x)$ is equal to its **Taylor series*** near a point x_0. That is, for x near x_0 we have

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \tag{2.15}$$

where $f^{(k)}(x_0)$ denotes the k^{th} derivative of $f(x)$ at the point x_0.

For example, sine is an analytic function and its Taylor series around $x_0 = 0$ is given by

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

In 2.2 we plot $\sin(x)$ and the truncations of the series up to degree 5 and 9. You can see that the approximation is very good for x near 0, but gets worse for x further away from 0. This is what happens in general. To get a good approximation far away from x_0 you need to take more and more terms of the Taylor series.

One of the main properties of power series that we will use is that we can differentiate them term by term. That is, suppose that $\sum a_k (x - x_0)^k$ is a convergent power series. Then for x in the radius of convergence we have

$$\frac{d}{dx} \left[\sum_{k=0}^{\infty} a_k (x - x_0)^k \right] = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}.$$

Notice that the term corresponding to $k = 0$ disappeared as it was constant. The radius of convergence of the differentiated series is the same as that of the original.

Example 2.3.5: Let us show that the exponential $y = e^x$ solves $y' = y$. First write

$$y = e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

*Named after the English mathematician Sir Brook Taylor (1685–1731).

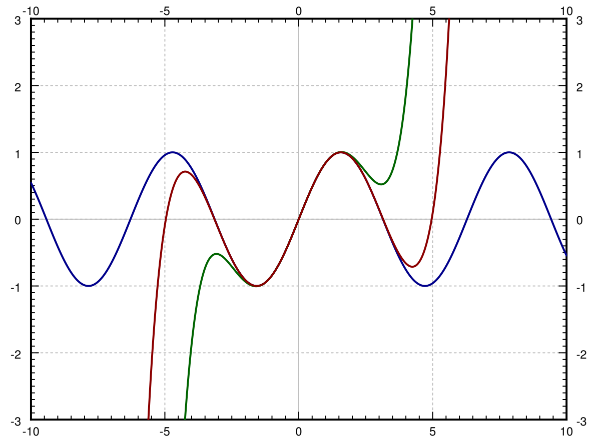

Figure 2.2: The sine function and its Taylor approximations around $x_0 = 0$ of 5^{th} and 9^{th} degree.

Now differentiate

$$y' = \sum_{k=1}^{\infty} k \frac{1}{k!} x^{k-1} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1}.$$

We **reindex** the series by simply replacing k with $k+1$. The series does not change, what changes is simply how we write it. After reindexing the series starts at $k = 0$ again.

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} = \sum_{k+1=1}^{\infty} \frac{1}{((k+1)-1)!} x^{(k+1)-1} = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

That was precisely the power series for e^x we started with, so we showed that $\frac{d}{dx}[e^x] = e^x$.

Convergent power series can be added and multiplied together, and multiplied by constants using the following rules. First, we can add series by adding term by term,

$$\left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right) + \left(\sum_{k=0}^{\infty} b_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} (a_k + b_k)(x-x_0)^k.$$

We can multiply by constants,

$$\alpha \left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} \alpha a_k (x-x_0)^k.$$

We can also multiply series together,

$$\left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right) \left(\sum_{k=0}^{\infty} b_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} c_k (x-x_0)^k,$$

where $c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$. The radius of convergence of the sum or the product is at least the minimum of the radii of convergence of the two series involved.

2.3.1 Power series for rational functions

Polynomials are simply finite power series. That is, a polynomial is a power series where the a_k are zero for all k large enough. We can always expand a polynomial as a power series about any point x_0 by writing the polynomial as a polynomial in $(x - x_0)$. For example, let us write $2x^2 - 3x + 4$ as a power series around $x_0 = 1$:

$$2x^2 - 3x + 4 = 3 + (x-1) + 2(x-1)^2.$$

In other words $a_0 = 3$, $a_1 = 1$, $a_2 = 2$, and all other $a_k = 0$. To do this, we know that $a_k = 0$ for all $k \geq 3$ as the polynomial is of degree 2. We write $a_0 + a_1(x-1) + a_2(x-1)^2$, we expand, and we solve for a_0, a_1, and a_2. We could have also differentiated at $x = 1$ and used the Taylor series formula 2.15.

Let us look at rational functions, that is, ratios of polynomials. An important fact is that a series for a function only defines the function on an interval even if the function is defined elsewhere. For example, for $-1 < x < 1$,

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots$$

This series is called the **geometric series**. The ratio test tells us that the radius of convergence is 1. The series diverges for $x \leq -1$ and $x \geq 1$, even though $\frac{1}{1-x}$ is defined for all $x \neq 1$.

We can use the geometric series together with rules for addition and multiplication of power series to expand rational functions around a point, as long as the denominator is not zero at x_0. Note that as for polynomials, we could equivalently use the Taylor series expansion 2.15.

Example 2.3.6: Expand $\frac{x}{1+2x+x^2}$ as a power series around the origin ($x_0 = 0$) and find the radius of convergence.

First, write $1 + 2x + x^2 = (1+x)^2 = \big(1-(-x)\big)^2$. Compute

$$\begin{aligned}
\frac{x}{1+2x+x^2} &= x\left(\frac{1}{1-(-x)}\right)^2 \\
&= x\left(\sum_{k=0}^{\infty} (-1)^k x^k\right)^2 \\
&= x\left(\sum_{k=0}^{\infty} c_k x^k\right) \\
&= \sum_{k=0}^{\infty} c_k x^{k+1},
\end{aligned}$$

where to get c_k, we use the formula for the product of series. We obtain, $c_0 = 1$, $c_1 = -1-1 = -2$, $c_2 = 1+1+1 = 3$, etc. Therefore

$$\frac{x}{1+2x+x^2} = \sum_{k=1}^{\infty} (-1)^{k+1} k x^k = x - 2x^2 + 3x^3 - 4x^4 + \cdots$$

The radius of convergence is at least 1. We use the ratio test

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty} \left|\frac{(-1)^{k+2}(k+1)}{(-1)^{k+1}k}\right| = \lim_{k\to\infty} \frac{k+1}{k} = 1.$$

So the radius of convergence is actually equal to 1.

When the rational function is more complicated, it is also possible to use method of partial fractions. For example, to find the Taylor series for $\frac{x^3+x}{x^2-1}$, we write

$$\frac{x^3+x}{x^2-1} = x + \frac{1}{1+x} - \frac{1}{1-x} = x + \sum_{k=0}^{\infty} (-1)^k x^k - \sum_{k=0}^{\infty} x^k = -x + \sum_{\substack{k=3 \\ k \text{ odd}}}^{\infty} (-2) x^k.$$

2.3.2 Series solutions of linear second order ODEs

Note: 1 or 1.5 lecture, §8.2 in [EP], §5.2 and §5.3 in [BD]

Suppose we have a linear second order homogeneous ODE of the form

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

Suppose that $p(x)$, $q(x)$, and $r(x)$ are polynomials. We will try a solution of the form

$$y = \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

and solve for the a_k to try to obtain a solution defined in some interval around x_0.

The point x_0 is called an **ordinary point** if $p(x_0) \neq 0$. That is, the functions

$$\frac{q(x)}{p(x)} \quad \text{and} \quad \frac{r(x)}{p(x)}$$

are defined for x near x_0. If $p(x_0) = 0$, then we say x_0 is a **singular point**. Handling singular points is harder than ordinary points and so we now focus only on ordinary points.

Example 2.3.7: Let us start with a very simple example

$$y'' - y = 0.$$

Let us try a power series solution near $x_0 = 0$, which is an ordinary point. Every point is an ordinary point in fact, as the equation is constant coefficient. We already know we should obtain exponentials or the hyperbolic sine and cosine, but let us pretend we do not know this.

We try

$$y = \sum_{k=0}^{\infty} a_k x^k.$$

If we differentiate, the $k = 0$ term is a constant and hence disappears. We therefore get

$$y' = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

We differentiate yet again to obtain (now the $k = 1$ term disappears)

$$y'' = \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}.$$

We reindex the series (replace k with $k + 2$) to obtain

$$y'' = \sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k.$$

Now we plug y and y'' into the differential equation

$$\begin{aligned} 0 = y'' - y &= \left(\sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k \right) - \left(\sum_{k=0}^{\infty} a_k x^k \right) \\ &= \sum_{k=0}^{\infty} \Big((k+2)(k+1) a_{k+2} x^k - a_k x^k \Big) \\ &= \sum_{k=0}^{\infty} \Big((k+2)(k+1) a_{k+2} - a_k \Big) x^k. \end{aligned}$$

As $y'' - y$ is supposed to be equal to 0, we know that the coefficients of the resulting series must be equal to 0. Therefore,

$$(k+2)(k+1) a_{k+2} - a_k = 0, \qquad \text{or} \qquad a_{k+2} = \frac{a_k}{(k+2)(k+1)}.$$

The equation above is called a **recurrence relation** for the coefficients of the power series. It did not matter what a_0 or a_1 was. They can be arbitrary. But once we pick a_0 and a_1, then all other coefficients are determined by the recurrence relation.

Let us see what the coefficients must be. First, a_0 and a_1 are arbitrary. Then,

$$a_2 = \frac{a_0}{2}, \quad a_3 = \frac{a_1}{(3)(2)}, \quad a_4 = \frac{a_2}{(4)(3)} = \frac{a_0}{(4)(3)(2)}, \quad a_5 = \frac{a_3}{(5)(4)} = \frac{a_1}{(5)(4)(3)(2)}, \quad \ldots$$

So for even k, that is $k = 2n$, we have

$$a_k = a_{2n} = \frac{a_0}{(2n)!},$$

and for odd k, that is $k = 2n + 1$, we have

$$a_k = a_{2n+1} = \frac{a_1}{(2n+1)!}.$$

Let us write down the series

$$y = \sum_{k=0}^{\infty} a_k x^k = \sum_{n=0}^{\infty} \left(\frac{a_0}{(2n)!} x^{2n} + \frac{a_1}{(2n+1)!} x^{2n+1} \right) = a_0 \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n} + a_1 \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1}.$$

We recognize the two series as the hyperbolic sine and cosine. Therefore,

$$y = a_0 \cosh x + a_1 \sinh x.$$

Of course, in general we will not be able to recognize the series that appears, since usually there will not be any elementary function that matches it. In that case we will be content with the series.

Example 2.3.8: Let us do a more complex example. Consider **Airy's equation***:

$$y'' - xy = 0,$$

near the point $x_0 = 0$. Note that $x_0 = 0$ is an ordinary point.

We try

$$y = \sum_{k=0}^{\infty} a_k x^k.$$

We differentiate twice (as above) to obtain

$$y'' = \sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2}.$$

We plug y into the equation

$$\begin{aligned} 0 = y'' - xy &= \left(\sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2} \right) - x \left(\sum_{k=0}^{\infty} a_k x^k \right) \\ &= \left(\sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2} \right) - \left(\sum_{k=0}^{\infty} a_k x^{k+1} \right). \end{aligned}$$

We reindex to make things easier to sum

$$\begin{aligned} 0 = y'' - xy &= \left(2a_2 + \sum_{k=1}^{\infty} (k+2)\,(k+1)\,a_{k+2} x^k \right) - \left(\sum_{k=1}^{\infty} a_{k-1} x^k \right) \\ &= 2a_2 + \sum_{k=1}^{\infty} \Big((k+2)\,(k+1)\,a_{k+2} - a_{k-1} \Big)\, x^k. \end{aligned}$$

*Named after the English mathematician Sir George Biddell Airy (1801–1892).

Again $y'' - xy$ is supposed to be 0, so $a_2 = 0$, and

$$(k+2)\,(k+1)\,a_{k+2} - a_{k-1} = 0, \qquad \text{or} \qquad a_{k+2} = \frac{a_{k-1}}{(k+2)(k+1)}.$$

We jump in steps of three. First, since $a_2 = 0$ we must have , $a_5 = 0$, $a_8 = 0$, $a_{11} = 0$, etc. In general, $a_{3n+2} = 0$.

The constants a_0 and a_1 are arbitrary and we obtain

$$a_3 = \frac{a_0}{(3)(2)}, \quad a_4 = \frac{a_1}{(4)(3)}, \quad a_6 = \frac{a_3}{(6)(5)} = \frac{a_0}{(6)(5)(3)(2)}, \quad a_7 = \frac{a_4}{(7)(6)} = \frac{a_1}{(7)(6)(4)(3)}, \quad \ldots$$

For a_k where k is a multiple of 3, that is $k = 3n$ we notice that

$$a_{3n} = \frac{a_0}{(2)(3)(5)(6)\cdots(3n-1)(3n)}.$$

For a_k where $k = 3n+1$, we notice

$$a_{3n+1} = \frac{a_1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}.$$

In other words, if we write down the series for y, it has two parts

$$\begin{aligned} y &= \left(a_0 + \frac{a_0}{6}x^3 + \frac{a_0}{180}x^6 + \cdots + \frac{a_0}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots\right) \\ &\quad + \left(a_1 x + \frac{a_1}{12}x^4 + \frac{a_1}{504}x^7 + \cdots + \frac{a_1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots\right) \\ &= a_0\left(1 + \frac{1}{6}x^3 + \frac{1}{180}x^6 + \cdots + \frac{1}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots\right) \\ &\quad + a_1\left(x + \frac{1}{12}x^4 + \frac{1}{504}x^7 + \cdots + \frac{1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots\right). \end{aligned}$$

We define

$$\begin{aligned} y_1(x) &= 1 + \frac{1}{6}x^3 + \frac{1}{180}x^6 + \cdots + \frac{1}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots, \\ y_2(x) &= x + \frac{1}{12}x^4 + \frac{1}{504}x^7 + \cdots + \frac{1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots, \end{aligned}$$

and write the general solution to the equation as $y(x) = a_0 y_1(x) + a_1 y_2(x)$. If we plug in $x = 0$ into the power series for y_1 and y_2, we find $y_1(0) = 1$ and $y_2(0) = 0$. Similarly, $y_1'(0) = 0$ and $y_2'(0) = 1$. Therefore $y = a_0 y_1 + a_1 y_2$ is a solution that satisfies the initial conditions $y(0) = a_0$ and $y'(0) = a_1$.

The functions y_1 and y_2 cannot be written in terms of the elementary functions that you know. See 2.3 for the plot of the solutions y_1 and y_2. These functions have many interesting properties. For example, they are oscillatory for negative x (like solutions to $y'' + y = 0$) and for positive x they grow without bound (like solutions to $y'' - y = 0$).

Sometimes a solution may turn out to be a polynomial.

Example 2.3.9: Let us find a solution to the so-called **Hermite's equation of order** n*:

$$y'' - 2xy' + 2ny = 0.$$

Let us find a solution around the point $x_0 = 0$. We try

$$y = \sum_{k=0}^{\infty} a_k x^k.$$

*Named after the French mathematician Charles Hermite (1822–1901).

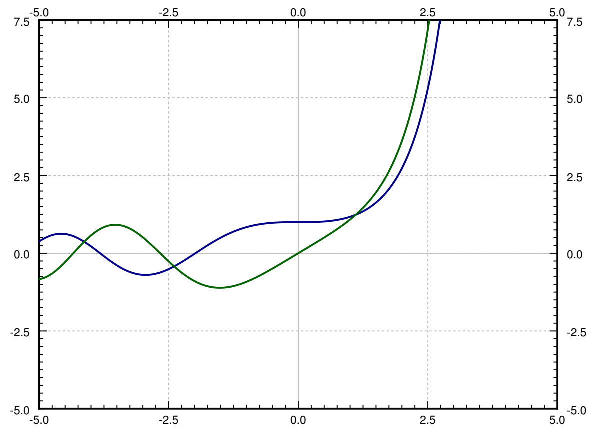

Figure 2.3: The two solutions y_1 and y_2 to Airy's equation.

We differentiate (as above) to obtain

$$y' = \sum_{k=1}^{\infty} k a_k x^{k-1},$$
$$y'' = \sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2}.$$

Now we plug into the equation

$$\begin{aligned}
0 = y'' &- 2xy' + 2ny \\
&= \left(\sum_{k=2}^{\infty} k(k-1)a_k x^{k-2}\right) - 2x\left(\sum_{k=1}^{\infty} k a_k x^{k-1}\right) + 2n\left(\sum_{k=0}^{\infty} a_k x^k\right) \\
&= \left(\sum_{k=2}^{\infty} k(k-1)a_k x^{k-2}\right) - \left(\sum_{k=1}^{\infty} 2k a_k x^k\right) + \left(\sum_{k=0}^{\infty} 2n a_k x^k\right) \\
&= \left(2a_2 + \sum_{k=1}^{\infty}(k+2)(k+1)a_{k+2}x^k\right) - \left(\sum_{k=1}^{\infty} 2k a_k x^k\right) + \left(2na_0 + \sum_{k=1}^{\infty} 2n a_k x^k\right) \\
&= 2a_2 + 2na_0 + \sum_{k=1}^{\infty}\big((k+2)(k+1)a_{k+2} - 2ka_k + 2na_k\big)x^k.
\end{aligned}$$

As $y'' - 2xy' + 2ny = 0$ we have

$$(k+2)(k+1)a_{k+2} + (-2k+2n)a_k = 0, \qquad \text{or} \qquad a_{k+2} = \frac{(2k-2n)}{(k+2)(k+1)}a_k.$$

This recurrence relation actually includes $a_2 = -na_0$ (which comes about from $2a_2 + 2na_0 = 0$). Again a_0 and a_1 are arbitrary.

$$a_2 = \frac{-2n}{(2)(1)}a_0, \qquad a_3 = \frac{2(1-n)}{(3)(2)}a_1,$$
$$a_4 = \frac{2(2-n)}{(4)(3)}a_2 = \frac{2^2(2-n)(-n)}{(4)(3)(2)(1)}a_0,$$
$$a_5 = \frac{2(3-n)}{(5)(4)}a_3 = \frac{2^2(3-n)(1-n)}{(5)(4)(3)(2)}a_1, \quad \ldots$$

Let us separate the even and odd coefficients. We find that

$$a_{2m} = \frac{2^m(-n)(2-n)\cdots(2m-2-n)}{(2m)!},$$
$$a_{2m+1} = \frac{2^m(1-n)(3-n)\cdots(2m-1-n)}{(2m+1)!}.$$

Let us write down the two series, one with the even powers and one with the odd.

$$y_1(x) = 1 + \frac{2(-n)}{2!}x^2 + \frac{2^2(-n)(2-n)}{4!}x^4 + \frac{2^3(-n)(2-n)(4-n)}{6!}x^6 + \cdots,$$
$$y_2(x) = x + \frac{2(1-n)}{3!}x^3 + \frac{2^2(1-n)(3-n)}{5!}x^5 + \frac{2^3(1-n)(3-n)(5-n)}{7!}x^7 + \cdots.$$

We then write

$$y(x) = a_0 y_1(x) + a_1 y_2(x).$$

We remark that if n is a positive even integer, then $y_1(x)$ is a polynomial as all the coefficients in the series beyond a certain degree are zero. If n is a positive odd integer, then $y_2(x)$ is a polynomial. For example, if $n = 4$, then

$$y_1(x) = 1 + \frac{2(-4)}{2!}x^2 + \frac{2^2(-4)(2-4)}{4!}x^4 = 1 - 4x^2 + \frac{4}{3}x^4.$$

2.3.3 Singular points and the method of Frobenius

Note: 1 or 1.5 lectures, §8.4 and §8.5 in [EP], §5.4–§5.7 in [BD]

While behavior of ODEs at singular points is more complicated, certain singular points are not especially difficult to solve. Let us look at some examples before giving a general method. We may be lucky and obtain a power series solution using the method of the previous section, but in general we may have to try other things.

Examples

Example 2.3.10: Let us first look at a simple first order equation

$$2xy' - y = 0.$$

Note that $x = 0$ is a singular point. If we try to plug in

$$y = \sum_{k=0}^{\infty} a_k x^k,$$

we obtain

$$0 = 2xy' - y = 2x\left(\sum_{k=1}^{\infty} k a_k x^{k-1}\right) - \left(\sum_{k=0}^{\infty} a_k x^k\right)$$
$$= a_0 + \sum_{k=1}^{\infty} (2k a_k - a_k)\, x^k.$$

First, $a_0 = 0$. Next, the only way to solve $0 = 2ka_k - a_k = (2k-1)\,a_k$ for $k = 1, 2, 3, \ldots$ is for $a_k = 0$ for all k. Therefore, in this manner we only get the trivial solution $y = 0$. We need a nonzero solution to get the general solution to the equation.

Let us try $y = x^r$ for some real number r. Consequently our solution—if we can find one—may only make sense for positive x. Then $y' = rx^{r-1}$. So

$$0 = 2xy' - y = 2xrx^{r-1} - x^r = (2r-1)x^r.$$

Therefore $r = 1/2$, or in other words $y = x^{1/2}$. Multiplying by a constant, the general solution for positive x is

$$y = Cx^{1/2}.$$

If $C \neq 0$, then the derivative of the solution "blows up" at $x = 0$ (the singular point). There is only one solution that is differentiable at $x = 0$ and that's the trivial solution $y = 0$.

Not every problem with a singular point has a solution of the form $y = x^r$, of course. But perhaps we can combine the methods. What we will do is to try a solution of the form

$$y = x^r f(x),$$

where $f(x)$ is an analytic function.

Example 2.3.11: Consider the equation

$$4x^2y'' - 4x^2y' + (1-2x)y = 0,$$

and again note that $x = 0$ is a singular point.

Let us try

$$y = x^r \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} a_k x^{k+r},$$

where r is a real number, not necessarily an integer. Again if such a solution exists, it may only exist for positive x. First let us find the derivatives

$$\begin{aligned} y' &= \sum_{k=0}^{\infty} (k+r)\, a_k x^{k+r-1}, \\ y'' &= \sum_{k=0}^{\infty} (k+r)\,(k+r-1)\, a_k x^{k+r-2}. \end{aligned}$$

Plugging into our equation we obtain

$$\begin{aligned} 0 &= 4x^2y'' - 4x^2y' + (1-2x)y \\ &= 4x^2 \left(\sum_{k=0}^{\infty} (k+r)\,(k+r-1)\, a_k x^{k+r-2} \right) - 4x^2 \left(\sum_{k=0}^{\infty} (k+r)\, a_k x^{k+r-1} \right) + (1-2x) \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) \\ &= \left(\sum_{k=0}^{\infty} 4(k+r)\,(k+r-1)\, a_k x^{k+r} \right) \\ &\qquad - \left(\sum_{k=0}^{\infty} 4(k+r)\, a_k x^{k+r+1} \right) + \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) - \left(\sum_{k=0}^{\infty} 2 a_k x^{k+r+1} \right) \\ &= \left(\sum_{k=0}^{\infty} 4(k+r)\,(k+r-1)\, a_k x^{k+r} \right) \\ &\qquad - \left(\sum_{k=1}^{\infty} 4(k+r-1)\, a_{k-1} x^{k+r} \right) + \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) - \left(\sum_{k=1}^{\infty} 2 a_{k-1} x^{k+r} \right) \\ &= 4r(r-1)\, a_0 x^r + a_0 x^r + \sum_{k=1}^{\infty} \Big(4(k+r)\,(k+r-1)\, a_k - 4(k+r-1)\, a_{k-1} + a_k - 2a_{k-1} \Big)\, x^{k+r} \\ &= \big(4r(r-1)+1\big)\, a_0 x^r + \sum_{k=1}^{\infty} \Big(\big(4(k+r)\,(k+r-1)+1\big)\, a_k - \big(4(k+r-1)+2\big)\, a_{k-1} \Big)\, x^{k+r}. \end{aligned}$$

To have a solution we must first have $\big(4r(r-1)+1\big)\, a_0 = 0$. Supposing that $a_0 \neq 0$ we obtain

$$4r(r-1)+1 = 0.$$

This equation is called the **indicial equation**. This particular indicial equation has a double root at $r = 1/2$.

OK, so we know what r has to be. That knowledge we obtained simply by looking at the coefficient of x^r. All other coefficients of x^{k+r} also have to be zero so

$$\big(4(k+r)\,(k+r-1)+1\big)\,a_k - \big(4(k+r-1)+2\big)\,a_{k-1} = 0.$$

If we plug in $r = 1/2$ and solve for a_k, we get

$$a_k = \frac{4(k+1/2-1)+2}{4(k+1/2)\,(k+1/2-1)+1}\,a_{k-1} = \frac{1}{k}\,a_{k-1}.$$

Let us set $a_0 = 1$. Then

$$a_1 = \frac{1}{1}a_0 = 1, \qquad a_2 = \frac{1}{2}a_1 = \frac{1}{2}, \qquad a_3 = \frac{1}{3}a_2 = \frac{1}{3\cdot 2}, \qquad a_4 = \frac{1}{4}a_3 = \frac{1}{4\cdot 3\cdot 2}, \qquad \cdots$$

Extrapolating, we notice that

$$a_k = \frac{1}{k(k-1)(k-2)\cdots 3\cdot 2} = \frac{1}{k!}.$$

In other words,

$$y = \sum_{k=0}^{\infty} a_k x^{k+r} = \sum_{k=0}^{\infty} \frac{1}{k!} x^{k+1/2} = x^{1/2} \sum_{k=0}^{\infty} \frac{1}{k!} x^k = x^{1/2} e^x.$$

That was lucky! In general, we will not be able to write the series in terms of elementary functions.

We have one solution, let us call it $y_1 = x^{1/2}e^x$. But what about a second solution? If we want a general solution, we need two linearly independent solutions. Picking a_0 to be a different constant only gets us a constant multiple of y_1, and we do not have any other r to try; we only have one solution to the indicial equation. Well, there are powers of x floating around and we are taking derivatives, perhaps the logarithm (the antiderivative of x^{-1}) is around as well. It turns out we want to try for another solution of the form

$$y_2 = \sum_{k=0}^{\infty} b_k x^{k+r} + (\ln x) y_1,$$

which in our case is

$$y_2 = \sum_{k=0}^{\infty} b_k x^{k+1/2} + (\ln x) x^{1/2} e^x.$$

We now differentiate this equation, substitute into the differential equation and solve for b_k. A long computation ensues and we obtain some recursion relation for b_k. The reader can (and should) try this to obtain for example the first three terms

$$b_1 = b_0 - 1, \qquad b_2 = \frac{2b_1 - 1}{4}, \qquad b_3 = \frac{6b_2 - 1}{18}, \qquad \ldots$$

We then fix b_0 and obtain a solution y_2. Then we write the general solution as $y = Ay_1 + By_2$.

2.3.4 The method of Frobenius

Before giving the general method, let us clarify when the method applies. Let

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

be an ODE. As before, if $p(x_0) = 0$, then x_0 is a singular point. If, furthermore, the limits

$$\lim_{x\to x_0} (x - x_0)\frac{q(x)}{p(x)} \qquad \text{and} \qquad \lim_{x\to x_0} (x - x_0)^2\frac{r(x)}{p(x)}$$

both exist and are finite, then we say that x_0 is a **regular singular point**.

Example 2.3.12: Often, and for the rest of this section, $x_0 = 0$. Consider

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

Write

$$\lim_{x\to 0} x\frac{q(x)}{p(x)} = \lim_{x\to 0} x\frac{x(1+x)}{x^2} = \lim_{x\to 0}(1+x) = 1,$$
$$\lim_{x\to 0} x^2\frac{r(x)}{p(x)} = \lim_{x\to 0} x^2\frac{(\pi+x^2)}{x^2} = \lim_{x\to 0}(\pi+x^2) = \pi.$$

So $x = 0$ is a regular singular point.

On the other hand if we make the slight change

$$x^2 y'' + (1+x)y' + (\pi + x^2)y = 0,$$

then

$$\lim_{x\to 0} x\frac{q(x)}{p(x)} = \lim_{x\to 0} x\frac{(1+x)}{x^2} = \lim_{x\to 0}\frac{1+x}{x} = \text{DNE}.$$

Here DNE stands for **does not exist**. The point 0 is a singular point, but not a regular singular point.

Let us now discuss the general **Method of Frobenius***. We only consider the method at the point $x = 0$ for simplicity. The main idea is the following theorem.

Theorem 2.3.3 (Method of Frobenius)**.** *Suppose that*

$$p(x)y'' + q(x)y' + r(x)y = 0 \tag{2.16}$$

has a regular singular point at $x = 0$, then there exists at least one solution of the form

$$y = x^r \sum_{k=0}^{\infty} a_k x^k.$$

*A solution of this form is called a **Frobenius-type solution**.*

The method usually breaks down like this:

(i) We seek a Frobenius-type solution of the form

$$y = \sum_{k=0}^{\infty} a_k x^{k+r}.$$

We plug this y into equation 2.16. We collect terms and write everything as a single series.

(ii) The obtained series must be zero. Setting the first coefficient (usually the coefficient of x^r) in the series to zero we obtain the **indicial equation**, which is a quadratic polynomial in r.

(iii) If the indicial equation has two real roots r_1 and r_2 such that $r_1 - r_2$ is not an integer, then we have two linearly independent Frobenius-type solutions. Using the first root, we plug in

$$y_1 = x^{r_1} \sum_{k=0}^{\infty} a_k x^k,$$

and we solve for all a_k to obtain the first solution. Then using the second root, we plug in

$$y_2 = x^{r_2} \sum_{k=0}^{\infty} b_k x^k,$$

and solve for all b_k to obtain the second solution.

*Named after the German mathematician Ferdinand Georg Frobenius (1849–1917).

(iv) If the indicial equation has a doubled root r, then there we find one solution

$$y_1 = x^r \sum_{k=0}^{\infty} a_k x^k,$$

and then we obtain a new solution by plugging

$$y_2 = x^r \sum_{k=0}^{\infty} b_k x^k + (\ln x) y_1,$$

into equation 2.16 and solving for the constants b_k.

(v) If the indicial equation has two real roots such that $r_1 - r_2$ is an integer, then one solution is

$$y_1 = x^{r_1} \sum_{k=0}^{\infty} a_k x^k,$$

and the second linearly independent solution is of the form

$$y_2 = x^{r_2} \sum_{k=0}^{\infty} b_k x^k + C(\ln x) y_1,$$

where we plug y_2 into 2.16 and solve for the constants b_k and C.

(vi) Finally, if the indicial equation has complex roots, then solving for a_k in the solution

$$y = x^{r_1} \sum_{k=0}^{\infty} a_k x^k$$

results in a complex-valued function—all the a_k are complex numbers. We obtain our two linearly independent solutions* by taking the real and imaginary parts of y.

The main idea is to find at least one Frobenius-type solution. If we are lucky and find two, we are done. If we only get one, we either use the ideas above or even a different method such as reduction of order to obtain a second solution.

2.3.5 Bessel functions

An important class of functions that arises commonly in physics are the **Bessel functions**†. For example, these functions appear when solving the wave equation in two and three dimensions. First consider **Bessel's equation** of order p:

$$x^2 y'' + x y' + \left(x^2 - p^2\right) y = 0.$$

We allow p to be any number, not just an integer, although integers and multiples of $1/2$ are most important in applications.

When we plug

$$y = \sum_{k=0}^{\infty} a_k x^{k+r}$$

into Bessel's equation of order p, we obtain the indicial equation

$$r(r-1) + r - p^2 = (r-p)(r+p) = 0.$$

We obtain two roots, $r_1 = p$ and $r_2 = -p$. If p is not an integer, then following the method of Frobenius and setting $a_0 = 1$, we find linearly independent solutions of the form

$$y_1 = x^p \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k} k! (k+p)(k-1+p) \cdots (2+p)(1+p)},$$

$$y_2 = x^{-p} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k} k! (k-p)(k-1-p) \cdots (2-p)(1-p)}.$$

*See Joseph L. Neuringera, **The Frobenius method for complex roots of the indicial equation**, International Journal of Mathematical Education in Science and Technology, Volume 9, Issue 1, 1978, 71–77.

†Named after the German astronomer and mathematician Friedrich Wilhelm Bessel (1784–1846).

Exercise* 2.3.1*:

a) Verify that the indicial equation of Bessel's equation of order p is $(r-p)(r+p)=0$.
b) Suppose p is not an integer. Carry out the computation to obtain the solutions y_1 and y_2 above.

Bessel functions are convenient constant multiples of y_1 and y_2. First we must define the **gamma function**

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

Notice that $\Gamma(1)=1$. The gamma function also has a wonderful property

$$\Gamma(x+1)=x\Gamma(x).$$

From this property, it follows that $\Gamma(n)=(n-1)!$ when n is an integer. So the gamma function is a continuous version of the factorial. We compute:

$$\begin{aligned}\Gamma(k+p+1)&=(k+p)(k-1+p)\cdots(2+p)(1+p)\Gamma(1+p),\\ \Gamma(k-p+1)&=(k-p)(k-1-p)\cdots(2-p)(1-p)\Gamma(1-p).\end{aligned}$$

Exercise* 2.3.2*: *Verify the identities above using $\Gamma(x+1)=x\Gamma(x)$.*

We define the **Bessel functions of the first kind** of order p and $-p$ as

$$\begin{aligned}J_p(x)&=\frac{1}{2^p\Gamma(1+p)}y_1=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,\Gamma(k+p+1)}\left(\frac{x}{2}\right)^{2k+p},\\ J_{-p}(x)&=\frac{1}{2^{-p}\Gamma(1-p)}y_2=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,\Gamma(k-p+1)}\left(\frac{x}{2}\right)^{2k-p}.\end{aligned}$$

As these are constant multiples of the solutions we found above, these are both solutions to Bessel's equation of order p. The constants are picked for convenience.

When p is not an integer, J_p and J_{-p} are linearly independent. When n is an integer we obtain

$$J_n(x)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,(k+n)!}\left(\frac{x}{2}\right)^{2k+n}.$$

In this case

$$J_n(x)=(-1)^nJ_{-n}(x),$$

and so J_{-n} is not a second linearly independent solution. The other solution is the so-called **Bessel function of second kind**. These make sense only for integer orders n and are defined as limits of linear combinations of $J_p(x)$ and $J_{-p}(x)$, as p approaches n in the following way:

$$Y_n(x)=\lim_{p\to n}\frac{\cos(p\pi)J_p(x)-J_{-p}(x)}{\sin(p\pi)}.$$

Each linear combination of $J_p(x)$ and $J_{-p}(x)$ is a solution to Bessel's equation of order p. Then as we take the limit as p goes to n, we see that $Y_n(x)$ is a solution to Bessel's equation of order n. It also turns out that $Y_n(x)$ and $J_n(x)$ are linearly independent. Therefore when n is an integer, we have the general solution to Bessel's equation of order n:

$$y=AJ_n(x)+BY_n(x),$$

for arbitrary constants A and B. Note that $Y_n(x)$ goes to negative infinity at $x=0$. Many mathematical software packages have these functions $J_n(x)$ and $Y_n(x)$ defined, so they can be used just like say $\sin(x)$ and $\cos(x)$. In fact, Bessel functions have some similar properties. For example, $-J_1(x)$ is a derivative of $J_0(x)$, and in general the derivative of $J_n(x)$ can be written as a linear combination of $J_{n-1}(x)$ and $J_{n+1}(x)$. Furthermore, these functions oscillate, although they are not periodic. See 2.4 for graphs of Bessel functions.

Example 2.3.13: Other equations can sometimes be solved in terms of the Bessel functions. For example, given a positive constant λ,

$$xy''+y'+\lambda^2xy=0,$$

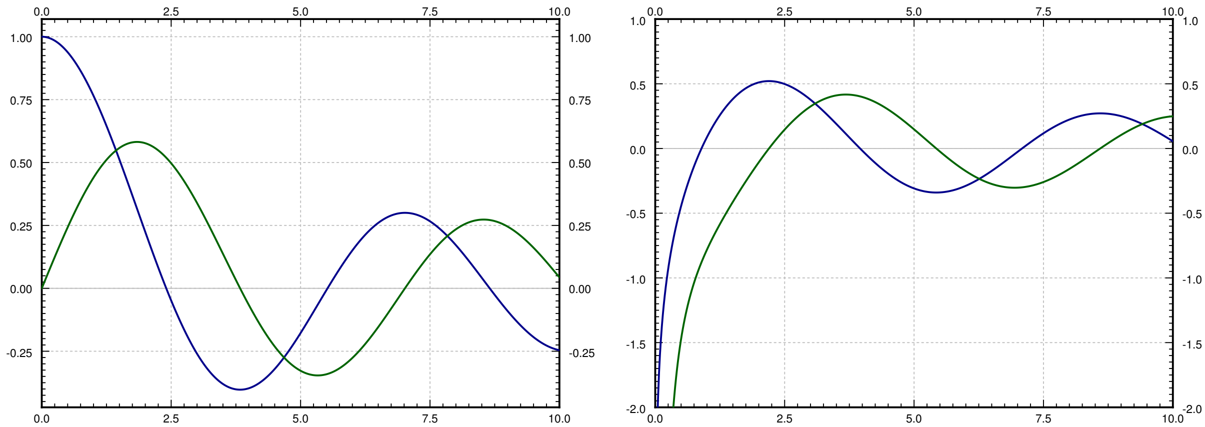

Figure 2.4: Plot of the $J_0(x)$ and $J_1(x)$ in the first graph and $Y_0(x)$ and $Y_1(x)$ in the second graph.

can be changed to $x^2y'' + xy' + \lambda^2x^2y = 0$. Then changing variables $t = \lambda x$, we obtain via chain rule the equation in y and t:

$$t^2y'' + ty' + t^2y = 0,$$

which we recognize as Bessel's equation of order 0. Therefore the general solution is $y(t) = AJ_0(t) + BY_0(t)$, or in terms of x:

$$y = AJ_0(\lambda x) + BY_0(\lambda x).$$

This equation comes up, for example, when finding the fundamental modes of vibration of a circular drum, but we digress.

Chapter 3

Linear ODEs

3.1 Linear ODEs and Linear Systems of ODEs

A linear ODE of n-th order is of the form

$$\sum_{i=0}^{n} a_i(t)x^{(i)} = a_n(t)x^{(n)} + a_{n-1}(t)x^{(n-1)} + \cdots + a_0(t)x = f(t). \tag{3.1}$$

where $a_i(t)$ and $f(t)$ are continuous functions over the interval $[a, b]$. The corresponding homogeneous linear ODE of n-th order is

$$\sum_{i=0}^{n} a_i(t)x^{(i)} = a_n(t)x^{(n)} + a_{n-1}(t)x^{(n-1)} + \cdots + a_0(t)x = 0. \tag{3.2}$$

and we hence say the the equation (3.1) is nonhomogeneous. A linear system of ODEs is of the following form

$$\begin{cases} \dot{x}_1 = a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + f_1(t), \\ \dot{x}_2 = a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + f_2(t), \\ \cdots\cdots\cdots\cdots \\ \dot{x}_n = a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t) \end{cases} \tag{3.3}$$

where $a_{ij}(t)$ and $f_i(t)$ are continuous functions over $[a, b]$. We can also write it as

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{f}(t),$$

with obvious meaning of each bold forms.

To use apply Theorem of existence and uniqueness 1.0.1 or 1.0.3 (both of them dealing with first-order system of ODE, see Remark 1.0.2) to (3.1), we need to convert it into a first-order system of ODE (which turns out to be linear too).

Proposition 3.1.1.

(1) n-th order linear ODE (3.1) can be converted into first-order linear system of ODE (3.3) via change of variables.

(2) Given IVP of n-th order linear ODE

$$\begin{cases} x^{(n)} + a_1(t)x^{(n-1)} + \cdots + a_{n-1}(t)x' + a_n(t)x = f(t), \\ x(t_0) = \eta_1, x'(t_0) = \eta_2, \cdots, x^{(n-1)}(t_0) = \eta_n, \end{cases} \tag{1}$$

where $a_i(t), f(t)$ are continuous over $[a, b]$, $t_0 \in [a, b]$, and η_i are given constants. Then the solution of it can be used to construct the solution of the IVP of the following first-order linear system of ODEs

$$\begin{cases} \dot{\boldsymbol{x}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n(t) & -a_{n-1}(t) & -a_{n-2}(t) & \cdots & -a_1(t) \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{bmatrix} \\ \boldsymbol{x}(t_0) = \boldsymbol{\eta} \end{cases}. \tag{2}$$

Conversely, given solution of (2), we can construct solution of (1).

Scholium. We emphasize that (1) is not just one aspect of (2). There exists first-order linear system of n ODE that cannot be converted into an n-th order linear ODE. For example,

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

proof of the proposition. (1) This is done just by the same process in Remark 1.0.1. For example, we use $y = (x, x', x'')$ to convert the equation

$$x''' + 4x'' - x' - 2x = 0$$

to be the system

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3, \\ y_3' = -4y_3 + y_2 + 2y_1 \end{cases}$$

(2) Let

$$x_1 = x, x_2 = x', x_3 = x'', \cdots, x_n = x^{(n-1)},$$

so

$$x_1' = x' = x_2, x_2' = x'' = x_3, \cdots, x_{n-1}' = x^{(n-1)} = x_n,$$
$$x_n' = x^{(n)} = -a_n(t)x_1 - a_{n-1}(t)x_2 - \cdots - a_1(t)x_n + f(t)$$

and

$$x_1(t_0) = x(t_0) = \eta_1, \quad x_2(t_0) = x'(t_0) = \eta_2, \cdots,$$
$$x_n(t_0) = x^{(n-1)}(t_0) = \eta_n.$$

Suppose $\psi(t)$ is a solution of (1) over $[a, b] \ni t_0$. Then $\psi(t), \psi'(t), \cdots, \psi^{(n)}(t)$ are continuous over $[a, b]$ and $\psi(t_0) = \eta_1, \psi'(t_0) = \eta_2, \cdots, \psi^{(n-1)}(t_0) = \eta_n$. Define over $[a, b]$,

$$\boldsymbol{\varphi}(t) = \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{bmatrix} = \begin{bmatrix} \psi(t) \\ \psi'(t) \\ \vdots \\ \psi^{(n-1)}(t) \end{bmatrix}.$$

Then $\boldsymbol{\varphi}(t_0)=\boldsymbol{\eta}$ and

$$\boldsymbol{\varphi}'(t)=\begin{bmatrix}\varphi_1'(t)\\ \varphi_2'(t)\\ \vdots\\ \varphi_{n-1}'(t)\\ \varphi_n'(t)\end{bmatrix}=\begin{bmatrix}\psi'(t)\\ \psi''(t)\\ \vdots\\ \psi^{(n-1)}(t)\\ \psi^{(n)}(t)\end{bmatrix}=\begin{bmatrix}\varphi_2(t)\\ \varphi_3(t)\\ \vdots\\ \varphi_n(t)\\ -a_1(t)\psi^{(n-1)}(t)-\cdots-a_n(t)\psi(t)+f(t)\end{bmatrix}$$

$$=\begin{bmatrix}\varphi_2(t)\\ \varphi_3(t)\\ \vdots\\ \varphi_n(t)\\ -a_n(t)\varphi_1(t)-\cdots-a_1(t)\varphi_n(t)+f(t)\end{bmatrix}$$

$$=\begin{bmatrix}0&1&0&\cdots&0\\ 0&0&1&\cdots&0\\ \vdots&\vdots&\vdots&&\vdots\\ 0&0&0&\cdots&1\\ -a_n(t)&-a_{n-1}(t)&-a_{n-2}(t)&\cdots&-a_1(t)\end{bmatrix}\begin{bmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \vdots\\ \varphi_{n-1}(t)\\ \varphi_n(t)\end{bmatrix}+\begin{bmatrix}0\\ 0\\ \vdots\\ 0\\ f(t)\end{bmatrix}$$

Therefore, $\boldsymbol{\varphi}(t)$ is a solution of (2).

Given solution $\boldsymbol{u}(t)$ of (2) on $[a,b]\ni t_0$, we let

$$\boldsymbol{u}(t)=\begin{bmatrix}u_1(t)\\ u_2(t)\\ \vdots\\ u_n(t)\end{bmatrix}$$

and define function $w(t)=u_1(t)$. By first equation of (2), we get

$$w'(t)=u_1'(t)=u_2(t)$$

The second equation of (2) reads

$$w''(t)=u_2'(t)=u_3(t)$$

and so on, until we get

$$w^{(n-1)}(t)=u_{n-1}'(t)=u_n(t)$$

from $(n-1)$-th equation. From the last equation, combined with the above observations, we get

$$\begin{aligned}w^{(n)}(t)&=u_n'(t)\\ &=-a_n(t)u_1(t)-a_{n-1}(t)u_2(t)-\cdots-a_2(t)u_{n-1}(t)-a_1(t)u_n(t)+f(t)\\ &=-a_1(t)w^{(n-1)}(t)-a_2(t)w^{(n-2)}(t)-\cdots-a_n(t)w(t)+f(t).\end{aligned}$$

Thus,

$$w^{(n)}(t)+a_1(t)w^{(n-1)}(t)+a_2(t)w^{(n-2)}(t)+\cdots+a_n(t)w(t)=f(t).$$

Besides,

$$w\left(t_0\right)=u_1\left(t_0\right)=\eta_1,\cdots,w^{(n-1)}\left(t_0\right)=u_n\left(t_0\right)=\eta_n,$$

so $w(t)$ is a solution of (1). □

Therefore, we have the theorem of existence and uniquess for equation (3.1):

Theorem 3.1.1 (The Existence and Uniquenes of Linear ODE). *If $a_i(t)$ and $f(t)$ are continuous functions over $[a,b]$, then for any $t_0\in[a,b]$ and initial conditions $x_0,x_0^{(1)},\cdots,x_0^{(n-1)}$, the equation (3.1) exists the only solution $x=\varphi(t)$ defined on the interval $[a,b]$ that satisfies the initial conditions:*

$$\varphi(t_0)=x_0,\frac{d\varphi(t_0)}{dt}=x_0^{(1)},\cdots,\frac{d^{n-1}\varphi(t_0)}{dt^{n-1}}=x_0^{(n-1)}$$

3.2 General Theory of Linear ODEs

3.2.1 The Structure of Solution

We first consider the homogeneous equation (3.2) and observe the following principle:

Theorem 3.2.1 (Superposition Principle). *If $x_1(t), \cdots, x_k(t)$ are k solutions of the equation (3.2), then the linear combination $\sum_{i=1}^{k} c_i x_i(t)$ is still a solution of (3.2).*

Proof. The theorem is immediate from the fact that $(cu)' = cu'$ and $(\sum u_i)' = \sum u'$. □

Before discussing the general solution of the euqation (3.2), some notions should be introduced.

Definition 3.2.1 (Linear (in)dependence). For functions $x_1(t), \ldots, x_k(t)$ defined on $[a, b]$, if there exists constants $c_1, \ldots, c_k$ not all zero such that

$$\forall t \in [a, b], \quad \sum_{i=1}^{k} c_i x_i(t) \equiv 0$$

we call these functions **linearly dependent**; otherwise **linearly independent**.

For example, functions $\cos t$ and $\sin t$ are linearly independent for any intervals, while functions $\cos^2 t$ and $\sin^2 t - 1$ are linearly dependent over any intervals. Functions $1, t, t^2, \ldots, t^n$ are linearly independent over any intervals, because it is not true that

$$c_0 + c_1 t + c_2 t^2 + \ldots + c_n t^n \equiv 0$$

holds for every t in a nondiscrete interval since Fundamental Theorem of Algebra implies that the equation has at most n distinct real roots.

Definition 3.2.2 (Wronskian Determinant). Consider functions $x_1(t), \ldots, x_k(t)$ that are $k-1$ times differentiable defined on $[a, b]$. Define their **Wronskian determinant** to be

$$W(t) \equiv W[x_1(t), \ldots, x_k(t)] \equiv \begin{vmatrix} x_1(t) & x_2(t) & \cdots & x_k(t) \\ x_1'(t) & x_2'(t) & \cdots & x_k'(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(k-1)}(t) & x_2^{(k-1)}(t) & \cdots & x_k^{(k-1)}(t) \end{vmatrix}$$

Theorem 3.2.2. *If functions $x_1(t), x_2(t), \cdots, x_n(t)$ are linearly dependent on the interval $[a, b]$, then the Wronskian determinant on the interval is 0.*

Proof. Since the functions $x_1(t), x_2(t), \cdots, x_n(t)$ are linearly dependent on the interval $[a, b]$, by definition, there exists not-all-zero coefficients c_i such that

$$\sum_{i=1}^{n} c_i x_1(t) = 0, \forall t \in [a, b]$$

We differentiate it $n-1$ times:

$$\begin{cases} c_1 x_1'(t) + c_2 x_2'(t) + \cdots + c_n x_n'(t) = 0, \\ c_1 x_1''(t) + c_2 x_2''(t) + \cdots + c_n x_n''(t) = 0, \\ \cdots\cdots \\ c_1 x_1^{(n-1)}(t) + c_2 x_2^{(n-1)}(t) + \cdots + c_n x_n^{(n-1)}(t) = 0. \end{cases} \quad t \in [a, b]$$

Notice that we shall regard the above system of linear equations as equations of unknowns "$c_1, c_2, \cdots, c_n$" with coefficient matrix

$$\boldsymbol{A}(t) = \begin{pmatrix} x_1(t) & x_2(t) & \cdots & x_n(t) \\ x_1'(t) & x_2'(t) & \cdots & x_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n-1)}(t) & x_2^{(n-1)}(t) & \cdots & x_n^{(n-1)}(t) \end{pmatrix}, t \in [a, b]$$

Namely,

$$\boldsymbol{A}(t)\boldsymbol{c} = \boldsymbol{0} \tag{3.4}$$

By theory in Linear Algebra, for every fixed $t_0 \in [a, b]$, we have that nonzero solution $\boldsymbol{c}$ of (3.4). Thus the determinant of coefficient matrix, which is $W(t)$, is zero over $[a, b]$. ☐

Remark 3.2.1: However, unlike linear algebra, the converse of this theorem is not true. See the following example.

Example 3.2.1: Define two functions

$$x_1(t) = \begin{cases} t^2, -1 \leqslant t < 0 \\ 0, 0 \leqslant t \leqslant 1 \end{cases}$$

and

$$x_2(t) = \begin{cases} 0, -1 \leqslant t < 0 \\ t^2, 0 \leqslant t \leqslant 1 \end{cases}$$

over the interval $[a, b]$. Their Wronskian determinant is

$$\begin{vmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{vmatrix} = \begin{cases} \begin{vmatrix} t^2 & 0 \\ 2t & 0 \end{vmatrix} = 0, -1 \leqslant t < 0 \\ \\ \begin{vmatrix} 0 & t^2 \\ 0 & 2t \end{vmatrix} = 0, 0 \leqslant t \leqslant 1 \end{cases}$$

However, the two functions are linearly independent, verified by solving $c_1x_1(t) + c_2x_2(t) = 0$ separately for two intervals.

However, if we add an additional condition to the above Theorem 3.2.2 that $x_1(t), x_2(t), \cdots, x_n(t)$ are solutions of homogeneous linear ODE (3.2), we have the following result (its contrapostive is the converse of the Theorem 3.2.2.)

Theorem 3.2.3. *If the homogeneous ODE (3.2) has solutions $x_1(t), x_2(t), \cdots, x_n(t)$ linearly dependent on the interval $[a, b]$, then $W(t) \neq 0\ \forall t \in [a, b]$.*

Proof. We prove by contradiction. Suppose there is a $t_0 \in [a, b]$ such that $W(t_0) = 0$. Consider the system of homogeneous linear equations with respect to $c_1, c_2, \cdots, c_n$:

$$\begin{cases} c_1x_1'(t_0) + c_2x_2'(t_0) + \cdots + c_nx_n'(t_0) = 0, \\ c_1x_1''(t_0) + c_2x_2''(t_0) + \cdots + c_nx_n''(t_0) = 0, \\ \cdots\cdots \\ c_1x_1^{(n-1)}(t_0) + c_2x_2^{(n-1)}(t_0) + \cdots + c_nx_n^{(n-1)}(t_0) = 0. \end{cases}$$

where its determinant of coefficient matrix $W(t_0) = 0$. Thus, the system has nonzero solution $\boldsymbol{c} = (c_1, \cdots, c_n)^T$. Now we use these constants to construct a function

$$x(t) \equiv \sum_{i=1}^{n} c_i x_i(t), t \in [a, b]$$

By the principle of superposition, the function $x(t)$ is a solution of the ODE (3.2). Notice that the system above, line by line, implies that

$$x(t_0) = 0, x'(t_0) = 0, \cdots, x^{(n-1)}(t_0) = 0 \tag{3.5}$$

In other words, we can say that $x(t)$ is the solution of the ODE (3.2) that satisfies the initial condition (3.5). However, $x = 0$ is a trivial solution of ODE (3.2) satisfying the initial condition (3.5) as well. By theorem of existence and uniqueness 3.1.1, the solution of ODE (3.2) is nonetheless unique. Namely,

$$x(t) = \sum_{i=1}^{n} c_i x_i(t) \equiv 0, t \in [a, b]$$

Since c_i's are not all zero, this contradicts to the linear independence of the solutions $x_i(t)$. □

According to theorem 3.2.2 and theorem 3.2.3, we see that the Wronskian determinant of the n solutions of the homogeneous linear ODE (3.2) is zero if solutions are linearly dependent or nonzero for over any nondiscrete interval when the coefficients are continuous if solutions are linear independent.

Now, consider n sets of initial conditions

$$\begin{cases} x_1(t_0) = 1, x_1'(t_0) = 0, & \cdots, \quad x_1^{(n-1)}(t_0) = 0; \\ x_2(t_0) = 0, x_2'(t_0) = 1, & \cdots, \quad x_2^{(n-1)}(t_0) = 0; \\ \cdots\cdots & \\ x_n(t_0) = 0, x_n'(t_0) = 0, & \cdots, \quad x_n^{(n-1)}(t_0) = 1. \end{cases}$$

each of them determining a solution of (3.2) by theorem of eixstence and uniqueness:

$$x_1(t), x_2(t), \cdots, x_n(t).$$

Since Wronskian determinant $W[x_1(t), x_2(t), \cdots, x_n(t)]$ at t_0 is $\det(I) = 1 \neq 0$, Theorem 3.2.2 implies that the n solutions $x_1(t), x_2(t), \cdots, x_n(t)$ are linearly independent. Therefore, we have the following theorems:

Theorem 3.2.4. *Homogeneous linear ODE of order n (3.2) has n linearly independent solutions.*

Theorem 3.2.5 (Strucutre of General Solution). *If $x_1(t), x_2(t), \cdots, x_n(t)$ are n linearly independent solutions of the equation (3.2), then the general solution of equation (3.2) can be expressed as*

$$x = \sum_{i=1}^{n} c_i x_i(t), \tag{3.6}$$

where c_i are any constants. Besides, the general solution includes all the solutions of the equation.

Proof. First, by the principle of superposition, we know (3.6) is the solution of equation (3.2), consisting of n abitrary constants. We point out that these constants are independent of each other. In fact, by regarding c_i as variables and $x_i(t)$ as constants, we have

$$\begin{aligned} \frac{\partial x}{\partial c_i} =& \frac{\partial x(c_1 x_1(t), \cdots, c_n x_n(t))}{\partial c_i} = \frac{\partial x(\xi_1, \cdots, \xi_n)}{\partial c_i} \\ =& \sum_{j=1}^{n} \frac{\partial x}{\partial \xi_j} \frac{\partial \xi_j}{\partial c_i} = \sum_{j=1}^{n} \frac{\partial(\xi_1 + \cdots + \xi_n)}{\partial \xi_j} \frac{\partial(c_j x_j(t))}{\partial c_i} \\ =& 1 \cdot 0 + \cdots + 1 \cdot x_i(t) + \cdots + 1 \cdot 0 = x_i(t), \end{aligned}$$

we have

$$\begin{vmatrix} \frac{\partial x}{\partial c_1} & \frac{\partial x}{\partial c_2} & \cdots & \frac{\partial x}{\partial x_n} \\ \frac{\partial x'}{\partial c_1} & \frac{\partial x'}{\partial c_2} & \cdots & \frac{\partial x'}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^{(n-1)}}{\partial c_1} & \frac{\partial x^{(n-1)}}{\partial c_2} & \cdots & \frac{\partial x^{(n-1)}}{\partial x_n} \end{vmatrix} = W[x_1(t), x_2(t), \cdots, x_n(t)] \neq 0, (t \in [a, b])$$

Therefore, 3.6 is the general solution of the equation (3.2). Now, we have proved that 3.6 constains all the solutions. By the solution is uniquely determined by the initial condition, which means that we only need to show: for any given initial condition

$$x(t_0) = x_0, x'(t_0) = x_0', \cdots, x^{(n-1)}(t_0) = x_0^{(n-1)} \tag{3.7}$$

we can determine the value of constants c_i to let 3.6 satisfies 3.7.

We let 3.6 satisfies the condition 3.7 to get the following system of linear equation with respect to $c_1, \cdots, c_n$:

$$\begin{cases} c_1x_1(t_0) + c_2x_2(t_0) + \cdots + c_nx_n(t_0) = x_0, \\ c_1x_1'(t_0) + c_2x_2'(t_0) + \cdots + c_nx_n'(t_0) = x_0', \\ \cdots\cdots \\ c_1x_1^{(n-1)}(t_0) + c_2x_2^{(n-1)}(t_0) + \cdots + c_nx_n^{(n-1)}(t_0) = x_0^{(n-1)}. \end{cases}$$

whose coefficient determinant is $W(t_0) \neq 0$ by Theorem 3.2.3. By the theory of Linear Algebra, the above system of linear equation has the only solution $\widetilde{c_1}, \cdots, \widetilde{c_n}$. Thus, we let the expression 3.6 take $\widetilde{c_1}, \cdots, \widetilde{c_n}$ as constants, and it then satisfies condition 3.7. □

Theorem 3.2.1. *All the solutions of the homogeneous linear ODE of order n form an n-dimensional vector space.*

3.2.2 Method of Variation of Parameters

We first point out two obvious relationships between equation (3.1) and equation (3.2):

1. if $x_1(t)$ is a solution of equation (3.1), and $x_2(t)$ is a solution of euqation (3.2), then $x_1(t) + x_2(t)$ is still a solution of equation (3.1).
2. the difference between any solution of equation (3.1) is a solution of equation (3.2).

Theorem 3.2.6. *Let $x_1(t), \cdots, x_n(t)$ be a basis of solution set of equation (3.2) and let $x_p(t)$ be a solution of equation (3.1). Then the general solution of equation (3.1) can be written as*

$$x(t) = x_p(t) + \sum_{i=1}^{n} c_ix_i(t) = x_p(t) + x_c(t)$$

*where $x_c(t)$ is called **complementary solution** of nonhomogeneous ODE (3.1) and $y_p(t)$ is called a **particular solution** of (3.1).*

Method of variation of parameters requires one to first know the basis (a set of linearly independent solutions) of equation (3.2) and then helps one to get a solution (and thus the general solution) of equation (3.4). It is not a particularly efficient method, and we shall only present the result in order 2. Consider

$$ay'' + by' + cy = f$$

first find a fundamental pair $\{y1, y2\}$ of solutions to the corresponding homogeneous equation

$$ay'' + by' + cy = 0$$

Then set

$$y = y_1c_1 + y_2c_2 \tag{3.8}$$

assuming that $c_1 = c_1(x)$ and $c_2 = c_2(x)$ are unknown functions whose derivatives satisfy the system

$$\begin{cases} y_1c_1' + y_2c_2' = 0 \\ y_1'c_1' + y_2'c_2' = f/a \end{cases}$$

Solve the system for c_1' and c_2'; integrate to get the formulas for u and v , and plug the results back into (3.8). That formula for y is your solution.

Example 3.2.2: Solve the equation

$$x'' + x = \frac{1}{\cos t}$$

where the fundamental pair $\{\sin t, \cos t\}$ is given. We follow the above recipe:
let

$$x(t) = c_1(t)\cos t + c_1(t)\sin t$$

and solve

$$\begin{cases}\cos t c_1' + \sin t c_2' = 0 \\ (\cos t)' c_1' + (\sin t)' c_2' = \frac{1}{\cos t}\end{cases}$$

or

$$\begin{cases}\cos t c_1' + \sin t c_2' = 0 \\ -\sin t c_1' + \cos t c_2' = \frac{1}{\cos t}\end{cases}$$

to get

$$\begin{cases}c_1'(t) = -\frac{\sin t}{\cos t} \\ c_2'(t) = 1\end{cases} \Rightarrow \begin{cases}c_1(t) = \ln|\cos t| + \gamma_1 \\ c_2'(t) = t + \gamma_2\end{cases}$$

Therefore, the general solution of the ODE is

$$x = \gamma_1 \cos t + \gamma_2 \sin t + \cos t \ln|\cos t| + t \sin t$$

where γ_i are any constants.

For more detials and examples, one may visit the link.

Exercise* 3.2.1*: *1. Given fundamental pairs $\{x_1, x_2\}$, solve the following ODEs:*

(1)

$$x'' - x = \cos t, x_1 = e^t, x_2 = e^{-t}$$

(2)

$$x'' + \frac{t}{1-t}x' - \frac{1}{1-t}x = t - 1, x_1 = t, x_2 = e^t$$

(3)

$$t^2 x'' - 4tx' + 6x = 36\frac{\ln t}{t}, x_1 = t^2, x_2 = t^3$$

(4)

$$t^2 x'' - 3tx' = 8x = 18t^2 \sin(\ln t), x_1 = t^2 \cos(2\ln t), x_2 = t^2 \sin 2\ln t$$

2. let $x_i(t)(i = 1, 2, \cdots, n)$ be any n solutions of the the homogeneous linear ODE (3.2), and let $W(t)$ be their Wronskian determinant. Prove that $W(t)$ satisfies the following first order linear ODE:

$$W' + a_1(t)W = 0$$

and thus

$$W(t) = W(t_0)\exp\left(-\int_{t_0}^{t} a_1(s)ds\right) t_0, t \in (a, b)$$

3.3 Constant Coefficient Linear ODEs

In this section, we focus on solving linear ODE with constant coefficients:

$$\sum_{i=0}^{n} a_i \frac{\mathrm{d}^i x}{\mathrm{d}t^i} = a_n \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x = f(t) \tag{3.9}$$

with its corresponding homogeneous equation

$$\sum_{i=0}^{n} a_i \frac{\mathrm{d}^i x}{\mathrm{d}t^i} = a_n \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x = 0 \tag{3.10}$$

where the coefficients $a_0, a_1, \cdots, a_n$ are real constants with $a_n \neq 0$. Besides, we introduce the **differential operator**

$$L := \sum_{i=0}^{n} a_i \frac{\mathrm{d}^i}{\mathrm{d}t^i} = a_n \frac{\mathrm{d}^n}{\mathrm{d}t^n} + a_{n-1} \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}}{\mathrm{d}t} + a_0$$

as a handful notation.

We also denote $D = \frac{\mathrm{d}}{\mathrm{d}t}$ and thus $D^n = \frac{\mathrm{d}^n}{\mathrm{d}t^n}$. Accordingly, we have

$$L \equiv \sum_{i=0}^{n} a_i D^i = a_n D^n + \cdots + a_1 D + a_0$$

It turns out that it is useful to think of the right-hand side of above euqation as a (formal) n-th degree polynomial in the "variable" D; it is a polynomial differential operator. For example, A first-degree polynomial operator with leading coefficient 1 has the form $D - a$, where a is a real number. It operates on a function $x = x(t)$ to produce

$$(D - a)x = Dx - ax = x' - ax$$

The important fact about such operators is that any two of them commute:

$$(D - a)(D - b)x = (D - b)(D - a)x$$

for any twice differentiable function $x = x(t)$. The proof of the above formula is the following computation:

$$\begin{aligned}
(D - a)(D - b)x &= (D - a)(x' - bx) \\
&= Dx' - ax' - bDx + abx \\
&= x'' - ax' - bx' + abx \\
&= D(x' - ax) - b(x' - ax) \\
&= (D - b)(D - a)x
\end{aligned}$$

We see from the proof that

$$(D - a)(D - b) = D^2 - (a + b)D + ab$$

Similarly, it can be shown by induction on the number of factors that an operator product of the form

$$(D - a_1)(D - a_2) \cdots (D - a_n)$$

expands–by multiplying out and collecting coefficients–in the same way as does an ordinary product of algebraic polynomials, regarding D as variables in polynomials. Consequently, the algebra of polynomial differential operators closely resembles the algebra of ordinary real polynomials.

By the theorem on structure of general solution of linear ODE, we see that we need to find n linearly independent solutions of equation (3.10) first. The tool we will look at for this constant-coefficient homogeneous linear ODE is called the **Euler Characteristic equation**.

3.3.1 The Characteristic Equation

Review the first order homogeneous linear ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} + ax = 0$$

we know that it has solution of the form $x = e^{-at}$, and the general solution of it is just $x = ce^{-at}$ with constant c. This inspires us to find exponential solutions

$$x = e^{\lambda t} \tag{3.11}$$

for equation (3.10), where λ is an undetermined coefficient.

Notice that

$$L[e^{\lambda t}] \equiv \sum_{i=0}^{n} a_i \frac{\mathrm{d}^i e^{\lambda t}}{\mathrm{d}t^i} = \sum_{i=0}^{n} a_i \lambda^i e^{\lambda t} \equiv F(\lambda) e^{\lambda t}$$

where $F(\lambda) := \sum_{i=0}^{n} a_i \lambda^i$ is an n-th order polynomial of λ. Since $e^{\lambda t} > 0$, we see that (3.11) solves (3.10) $\Leftrightarrow F(\lambda) = 0 \Leftrightarrow \lambda$ is the root of the algebraic equation

$$F(\lambda) \equiv \sum_{i=0}^{n} a_i \lambda^i = a_n \lambda^n + \cdots + a_1 \lambda + a_0 = 0$$

which is called the **Characteristic equation** or **Auxiliary equation**, whose roots are called **Characteristic roots**.

According to the fundamental theorem of algebra, every polynomial of order n has n zeros, though not necessarily distinct and not necessarily real. We first check the easiest situation:

Distinct Roots (Real and Complex)

Suppose that we have solved the characteristic equation with n distinct real roots

$$\lambda_1, \lambda_2, \cdots, \lambda_n$$

Then the functions

$$e^{\lambda_1 t}, e^{\lambda_2 t}, \cdots, e^{\lambda_n t}$$

are solutions of equation (3.10) and are linearly independent over $t \in [a, b]$ because

$$\begin{aligned} W(t) = &\begin{vmatrix} e^{\lambda_1 t} & e^{\lambda_2 t} & \cdots & e^{\lambda_n t} \\ \lambda_1 e^{\lambda_1 t} & \lambda_2 e^{\lambda_2 t} & \cdots & \lambda_n e^{\lambda_n t} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} e^{\lambda_1 t} & \lambda_2^{n-1} e^{\lambda_2 t} & \cdots & \lambda_n^{n-1} e^{\lambda_n t} \end{vmatrix} = e^{(\lambda_1 + \cdots + \lambda_n)t} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{vmatrix} \\ =& e^{(\lambda_1 + \cdots + \lambda_n)t} \det \text{Vandermonde}_{n \times n}(\lambda) (\text{where } V_{i,j} = \lambda_i^{j-1}) \\ =& e^{(\lambda_1 + \cdots + \lambda_n)t} \prod_{1 \leqslant j < i \leqslant n} (\lambda_i - \lambda_j) \\ \neq& 0 \text{ due to distinctness of roots and exp's positivity} \end{aligned}$$

We further divide the situation into two subcases:

1. All roots are real, then

$$e^{\lambda_1 t}, \cdots, e^{\lambda_n t}$$

are n linearly independent real solutions of equation (3.10) $L[x] = 0$, and the general solution is

$$x = c_1 e^{\lambda_1 t} + \cdots + c_n e^{\lambda_n t}$$

2. There are complex roots among λ_i, then the complex roots appear in pairs:
if $\lambda_1 = \alpha + i\beta$ is a characteristic root, then its conjugate $\lambda_2 = \alpha - i\beta$ is also a characteristic root. the equation (3.10) $L[x] = 0$ hence has two complex solutions:

$$\begin{aligned} e^{\lambda_1 t} =& e^{\alpha + i\beta} = e^{\alpha t}(\cos \beta t + i \sin \beta t) \\ e^{\lambda_2 t} =& e^{\alpha - i\beta} = e^{\alpha t}(\cos \beta t - i \sin \beta t) \end{aligned}$$

According to the following lemma, we see that their real and imaginary parts are also solutions of $L[x] = 0$, and we then also get two real solutions

$$e^{\alpha t} \cos \beta t, \ e^{\alpha t} \sin \beta t$$

Lemma 3.3.1. If all the coefficients $a_i(t)$ in the equation (3.2) are real-valued functions and $z(t) = \varphi(t) + i\psi(t)$ is a complex solution of the equation, then the real part and imaginary part of $z(t)$ and the conjugate $\bar{z}(t)$ are all solutions of the equation.
Furthermore, if the equation

$$a_n(t) \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}(t) \frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1(t) \frac{\mathrm{d}x}{\mathrm{d}t} + a_0(t) x = f(t) = u(t) + iv(t)$$

has complex solution $x = R(t) + iI(t)$, then $R(t)$ and $I(t)$ solves the following equations respectively:

$$a_n(t)\frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}(t)\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1(t)\frac{\mathrm{d}x}{\mathrm{d}t} + a_0(t)x = u(t)$$

and

$$a_n(t)\frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}(t)\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1(t)\frac{\mathrm{d}x}{\mathrm{d}t} + a_0(t)x = v(t)$$

Repeated Roots (Real and Complex)

Let us now consider the possibility that the characteristic equation

$$F(\lambda) \equiv \sum_{i=0}^{n} a_i\lambda^i = a_n\lambda^n + \cdots + a_1\lambda + a_0 = 0 \tag{3.12}$$

has repeated roots. For example, suppose that (3.12) has only two distinct roots, λ_0 of multiplicity 1 and λ_1 of multiplicity $k = n - 1 > 1$. Then (after dividing by a_n) (3.12) can be rewritten in the form

$$(\lambda - \lambda_1)^k(\lambda - \lambda_0) = 0$$

Similarly, the corresponding operator L can be written as

$$L = (D - \lambda_1)^k(D - \lambda_0)$$

because we notice that

$$F(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$$

is a the polynomial sharing the same coefficients with

$$L = a_nD^n + \cdots + a_1D + a_0$$

The two polynomials are formally the same.
Two solutions of the differential equation $L[x] = 0$ are certainly $x_0 = e^{\lambda_0 t}$ and $x_1 = e^{\lambda_1 t}$. This is, however, not sufficient; we need $k+1$ linearly independent solutions in order to construct a general solution, because the equation is of order $k+1$. To find the missing $k-1$ solutions, we note that

$$L[x] = (D - \lambda_0)[(D - \lambda_1)^k x] = 0$$

Consequently, every solution of the k-th order equation

$$(D - \lambda_1)^k x = 0 \tag{3.13}$$

will also be a solution of the original equation $L[x] = 0$. Hence our problem is reduced to that of finding a general solution of the differential equation in (3.13). The fact that $x_1 = e^{\lambda_1 t}$ is one solution of (3.13) suggests that we try the substitution

$$x(t) = u(t)x_1(t) = u(t)e^{\lambda_1 t}$$

where $u(t)$ is a function yet to be determined. Observe that

$$\begin{aligned}(D - \lambda_1)[ue^{\lambda_1 t}] =& D(ue^{\lambda_1 t}) - \lambda_1 ue^{\lambda_1 t}\\ =&(Du)e^{\lambda_1 t} + uD(e^{\lambda_1 t}) - \lambda_1 ue^{\lambda_1 t}\\ =&(Du)e^{\lambda_1 t}\end{aligned}$$

Upon k applications of this fact, it follows that

$$(D - \lambda_1)^k[ue^{\lambda_1 t}] = (D^k u)e^{\lambda_1 t}$$

for any sufficiently differentiable function $u(t)$. Hence $x = ue^{\lambda_1 t}$ will be a solution of equation (3.13) if and only if $D^k u = u^{(k)} = 0$. But this is so if and only if

$$u(t) = c_1 + c_2 t + c_3 t^2 + \cdots + c_k t^{k-1}$$

a polynomial of degree at most $k-1$. Hence our desired solution of equation (3.13) is

$$x(t) = ue^{\lambda_1 t} = (c_1 + c_2 t + c_3 t^2 + \cdots + c_k t^{k-1})e^{\lambda_1 t}$$

In particular, we see here the additional solutions $te^{\lambda_1 t}$, $t^2 e^{\lambda_1 t}$, $\cdots$, $t^{k-1}e^{\lambda_1 t}$ of the original ODE $L[x] = 0$.

The preceding analysis can be carried out with the operator $D - \lambda_1$ replaced with an arbitrary polynomial operator. When this is done, the result is a proof of the following theorem.

Theorem 3.3.1. *If the characteristic equation in (3.12) has a repeated root λ of multiplicity k, then the part of a general soution of the differential equation in (3.10) corresponding to λ is of the form*

$$(c_1 + c_2 t + c_3 t^2 + \cdots + c_k t^{k-1})e^{\lambda t}$$

We may notice that the k functions

$$e^{\lambda t}, te^{\lambda t}, \cdots, t^{k-1}e^{\lambda t}$$

are linearly independent on $\mathbb{R}$. Thus a root of multiplicity k corresponds to k linearly independent solutions of the differential equation. And the end of this section we will prove that k_i solutions contributed by i roots each of multiplicity k_i are linearly independent (Lemma 3.2.4).

Example 3.3.1: Find a general solution of the fifth-order differential equation regarding the function $y = y(x)$:

$$9y^{(5)} - 6y^{(4)} + y^{(3)} = 0$$

solution:
The characteristic equation inspirational

$$9\lambda^5 - 6\lambda^4 + \lambda^3 = \lambda^3(9\lambda^2 - 6\lambda + 1) = \lambda^3(3\lambda - 1)^2 = 0$$

It has triple root $\lambda = 0$ and the double root $\lambda = \frac{1}{3}$. The triple root $\lambda = 0$ contributes

$$c_1 e^{0x} + c_2 x e^{0x} + c_3 x^3 e^{0x} = c_1 + c_2 x + c_3 x^2$$

to the solution, while the double root $\lambda = \frac{1}{3}$ contributes $c_4 e^{x/3} + c_5 x e^{x/3}$. Hence a general solution of the given differential equation is

$$y(x) = c_1 + c_2 x + c_3 x^2 + c_4 e^{x/3} + c_5 x e^{x/3}$$

Example 3.3.2: Solve the equation

$$(D^3 + 1)x = 0$$

solution:
Since $-1 = 1 \cdot e^{-i\pi}$ $(r = 1, \theta = -\pi)$ we see that the characteristic equation $\lambda^3 + 1 = 0$ has solution

$$\lambda_k = \sqrt[n]{r}\exp\left(i\frac{\theta + 2k\pi}{n}\right) = \sqrt[3]{1}\exp\left(i\frac{-\pi + 2k\pi}{3}\right) = e^{i\frac{\pi}{3}}, e^{i\pi}, e^{i\frac{5\pi}{3}},\ k = 1, 2, 3$$

It other words,

$$\lambda_1 = \frac{1}{2} + i\frac{\sqrt{3}}{2}, \lambda_2 = -1, \lambda_3 = \frac{1}{2} - i\frac{\sqrt{3}}{2}$$

The general solution is then $x = c_1 e^{-t} + e^{t/2}[c_2 \cos(\sqrt{3}t/2) + c_3 \sin(\sqrt{3}t/2)]$

Example 3.3.3: The characteristic equation of the differential equation

$$y^{(3)} + y' - 10y = 0$$

is the cubic equation

$$r^3 + r - 10 = 0$$

To find the roots of the above polynomial, we fist observe that a third order polynomial of the form

$$P(x) = (x-a)(x-b)(x-c) = x^3 - (a+b+c)x^2 + (ab+bc+ca)x - abc$$

hints us that the roots of the polynomial have to be the factors of the constant term. We see that the only possible rational roots are the factors ± 1, ± 2, ± 5, and ± 10 of the constant term 10. By trial and error (if not by inspection) we discover the root 2. We then divide the polynomial by $r-2$ to get

$$r^3 + r - 10 = (r-2)(r^2 + 2r + 5) = (r-2)[(r+1)^2 + 4] = 0$$

We get the roots

$$r_1 = 2, r_2 = -1 + 2i, r_3 = -1 - 2i$$

Therefore, the general solution is given by

$$y(x) = c_1 e^2 x + c_2 e^{-1t} \cos 2t + c_3 e^{-1t} \sin 2t$$

For the case where the characteristic equation (3.12) has **complex repeated roots**, say $\lambda = \alpha + i\beta$ is a root of multiplicity k, then $\bar{\lambda} = \alpha - i\beta$ is also a root of multiplicity k. Similar to the unrepeated complex case, we can derive $2k$ real solutions

$$e^{\alpha t} \cos \beta t, te^{\alpha t} \cos \beta t, \cdots, t^{k-1} e^{\alpha t} \cos \beta t$$

and

$$e^{\alpha t} \sin \beta t, te^{\alpha t} \sin \beta t, \cdots, t^{k-1} e^{\alpha t} \sin \beta t$$

Example 3.3.4: The roots of the characteristic equation of a certain differential equation are $3, -5, 0, 0, 0, 0, -5, 2 \pm 3i$, and $2 \pm 3i$. Write a general solution of this homogeneous differential equation.
solution:
The solution can be read directly from the list of roots. It is

$$\begin{aligned} y(x) = & c_1 + c_2 x + c_3 x^2 + c_4 x^3 + c_5 x^{3x} + c_6 e^{-5x} + c_7 x e^{-5x} \\ & e^{2x}(c_8 \cos 3x + c_9 \sin 3x) + x e^{2x}(c_{10} \cos 3x + c_{11} \sin 3x) \end{aligned}$$

Lemma 3.3.2. Suppose the characteristic equation (3.12) has roots $\lambda_1, \cdots, \lambda_m$ of multiplicity $k_1, \cdots, k_m$ respectively ($k_i \geqslant 1$, and unrepeated root λ_j has $k_j = 1$), and $k_1 + \cdots + k_m = n$, $\lambda_i \neq \lambda_j$ when $i \neq j$, then the equation $L[x] = 0$ has solutions

$$\begin{cases} e^{\lambda_1 t}, te^{\lambda_1 t} \cdots, t^{k_1 - 1} e^{\lambda_1 t} \\ e^{\lambda_2 t}, te^{\lambda_2 t} \cdots, t^{k_1 - 1} e^{\lambda_2 t} \\ \cdots\cdots \\ e^{\lambda_m t}, te^{\lambda_1 t} \cdots, t^{k_m - 1} e^{\lambda_m t} \end{cases}$$

and we need to show that these solutions are linearly independent.

Proof. BWOC, suppose these functions are linearly dependent. Then,

$$\sum_{r=1}^{m} (A_0^{(r)} + A_1^{(r)} t + \cdots + A_{k_r - 1}^{(r)} t^{k_r - 1}) e^{\lambda_r t} := \sum_{r=1}^{m} P_r(t) e^{\lambda_r t} = 0 \tag{3.14}$$

where $A_j^{(r)}$ are constants that are not all zero. WLOG, we let the polynomial $P_m(t)$ have at least a non-zero coefficient. Hence $P_m(t) \neq 0$. Divide eq. (3.14) by $e^{\lambda_1 t}$ and differentiate k_1 times with respect to t to get

$$\sum_{r=1}^{m} D^{k_1}[P_r(t) e^{(\lambda_r - \lambda_1)t}] = \sum_{r=2}^{m} Q_r(t) e^{(\lambda_r - \lambda_1)t} = 0 \tag{3.15}$$

where $Q_r(t) = (\lambda_r - \lambda_1)^{k_1} P_r(t) + S_r(t)$, with $S_r(t)$ a polynomial or a lower order than $P_r(t)$, so that $P_r(t)$ and $Q_r(t)$ have the same order and $Q_m(t) \neq 0$. Eq. (3.15) and eq. (3.14) are similar but has less terms. If we do the same process as eq. (3.14) to eq. (3.15) (i.e. divide eq. (3.15) by $e^{(\lambda_2 - \lambda_1)t}$ and differentiate it by k_2 times), then we will get an equation with fewer items. If we keep doing this, after $m-1$ times, we will get the following equation

$$R_m(t) e^{(\lambda_m - \lambda_{m-1})t} = 0$$

which is impossible by $e^x > 0$, $R_m(t) \neq 0$ and the fact that $R_m(t)$ and $P_m(t)$ have the same order. It is not hard to calculate that

$$R_m(t) = (\lambda_m - \lambda_1)^{k_1}(\lambda_m - \lambda_2)^{k_2} \cdots (\lambda_m - \lambda_{m-1})^{k_{m-1}} P_m(t) + W_m(t)$$

with $W_m(t)$ a polynomial or lower order than $P_m(t)$. □

We ended this subsection with a notable example.

Example 3.3.5 (Euler–Cauchy Equation)**:** We call the ODE of the following form

$$(x^n D^n + a_1 x^n - 1D^{n-1} + \cdots + a_{n-1}xD + a_n)y = \sum_{i=1}^{n} a_{n-1} D^i y = 0$$

the **Euler–Cauchy Equation**. Notice that we can always assume $a_0 = 1$ becasue otherwise we can divide the whole eq. by a_0. This eq. is solvable by method of substitution to transform it into a homogeneous linear ODE. In fact, introduce*

$$x = e^t, t = \ln x$$

One may see the procedure on wikipedia for second order. We only give the conclusion below. The characteristic equation

$$K(K-1)\cdots(K-n+1) + a_1 K(K-1)\cdots(K-n+2) + \cdots + a_n = 0$$

determines the roots. For each root $K = K_0$ of multiplicity m, we have m solutions of the equation

$$x^{K_0}, x^{K_0} \ln|x|, x^{K_0} \ln^2|x|, \cdots, x^{K_0} \ln^{m-1}|x|$$

If $K = K_0 = \alpha + i\beta$, then we get $2m$ real solutions

$$x^\alpha \cos(\beta \ln|x|), x^\alpha \ln|x| \cos(\beta \ln|x|), \cdots, x^\alpha \ln^{m-1}|x| \cos(\beta \ln|x|)$$

$$x^\alpha \sin(\beta \ln|x|), x^\alpha \ln|x| \sin(\beta \ln|x|), \cdots, x^\alpha \ln^{m-1}|x| \sin(\beta \ln|x|)$$

3.3.2 Nonhomogeneous Linear ODEs

We have noted before that to solve nonhomogeneous linear ODE one first solves its homogeneous counterpart and then using method of variation of parameters to find a particular solution. However, there are some special forms of nonhomogeneous linear ODE which we can solve directly in a neater way. We will talk about the method of undetermined coefficients and Laplace Transform method.

Method of Undetermined Coefficients

Given a nonhomogeneous linear ODE of n-th order.

$$L[x] := \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_{n-1}\frac{dx}{dt} + a_n x = f(t)$$

whose characteristic equation is

$$F(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0$$

We by the type of $f(t)$ consider two types that can be applied with the **Method of Undetermined Coefficients**.

I . Polynomial & exponential form: $f(t) = (b_0 t^m + b_1 t^{m-1} + \cdots + b_{m-1}t + b_m)e^{\lambda_0 t}$

II . Triangular form: $f(t) = [A(t)\cos\beta t + B(t)\cos\beta t]e^{\alpha t}$

*if $x < 0$ then use $x = -e^t$ to get the same result. For convenience, we assume $x > 0$ and plug in $t = \ln|x|$ at the end.

Type I:
For $f(t) = (b_0 t^m + b_1 t^{m-1} + \cdots + b_{m-1} t + b_m)e^{\lambda_0 t}$, where λ_0 can be 0, we have a particular solution of the form

$$\widetilde{x} = t^k (B_0 t^m + B_1 t^{m-1} + \cdots + B_{m-1} + B_m)e^{\lambda_0 t}$$

where k is the multiplicity of the root λ_0 of the characteristic equation $F(\lambda) = 0$ (when λ_0 is a distinct root, take $k = 1$; when λ_0 is not a root of the characteristic equation, take $k = 0$), and $B_0, B_1, \cdots, B_m$ are coefficients to be determined (plugging the corresponded form back into the original equation). For the case $\lambda_0 \neq 0$, we can do the transformation $x = ye^{\lambda_0 t}$ and transform the solution back after solving the transformed equation.

Example 3.3.6: Solve the ODE

$$(D^2 - 2D - 3)x = 3t + 1$$

The first step is always finding the general solution of the homogeneous counterpart

$$(D^2 - 2D - 3)x = 0$$

The characteristic eq. $F(\lambda) = \lambda^2 - 2\lambda - 3 = 0$ has solutions $\lambda_1 = 3, \lambda_2 = -1$. Hence the general solution of the homogeneous eq. is $x_c = c_1 e^{3t} + c_2 e^{-t}$. We then need to find a particular solution of the original. Notice that $\lambda_0 = 0$ is not a root of the characteristic eq. Thus, we set a particular solution a polynomial of the same order as $f(t) = 3t + 1$

$$x_p = B_0 t + B_1$$

Plugging it into the original results in

$$-2B_0 - 3B_0 t - 3B_1 = 3t + 1 \Rightarrow -3B_0 = 3, -2B_0 - 3B_1 = 1 \Rightarrow B_0 = -1 B_1 = \frac{1}{3}$$

Thus, $x_p = -t + \frac{1}{3}$ and the general solution of the ODE $(D^2 - 2D - 3)x = 3t + 1$ is

$$x(t) = x_c + x_p = c_1 e^{3t} + c_2 e^{-t} - t + \frac{1}{3}$$

Example 3.3.7: Solve the ODE

$$(D^2 - 2D - 3)x = e^{-2t}$$

We from example 3.2.9 get the general solution of the corresponding homogeneous eq.

$$x_c = c_1 e^{3t} + c_2 e^{-t}$$

We now seek a particular solution of the equation. Since $f(t) = e^{-2t}$ and $\lambda_0 = -2$ is not a root of the characteristic equation $\lambda^2 - 2\lambda - 3 = 0$, we see that a particular solution is $\widetilde{x} = Ae^{-2t}$, which, plugged into the original eq, results in

$$4Ae^{-2t} + 4Ae^{-2t} - 3Ae^{-2t} = e^{-2t} \Rightarrow 4A + 4A - 3A = 1 \Rightarrow A = 1/5, x_p = \frac{1}{5}e^{-2t}$$

The general solution is then

$$x(t) = x_c + x_p = c_1 e^{3t} + c_2 e^{-t} + \frac{1}{5}e^{-2t}$$

Example 3.3.8: Solve the ODE

$$(D^3 + 3D^2 + 3D + 1)x = e^{-t}(t - 5)$$

Its characteristic equation $\lambda^3 + 3\lambda^2 + 3\lambda + 1 = (\lambda + 1)^3 = 0$ has $\lambda_{1,2,3} = -1$ as a root of multiplicity 3. The general solution of the homogeneous equation is then

$$x_c = (c_1 + c_2 t + c_3 t^2)e^{-t}$$

A particular solution of the equation is given by

$$\widetilde{x} = t^3(A + Bt)e^{-t}$$

which, plugged into the original eq, results in

$$(6A + 24Bt)e^{-t} = e^{-t}(t - 5) \Rightarrow A = -\frac{5}{6}, B = \frac{1}{24} \Rightarrow x_p = \frac{1}{24}t^3(t - 20)e^{-t}$$

The general solution is then

$$x(t) = x_c + x_p = (c_1 + c_2t + c_3t^2)e^{-t} + \frac{1}{24}t^3(t - 20)e^{-t}$$

Type II:
For $f(t) = [A(t)\cos\beta t + B(t)\cos\beta t]e^{\alpha t}$, where α, β are constants, $A(t), B(t)$ are real-coefficient polynomials with their highest order denoted as m, we have particular solution

$$\widetilde{x} = t^k[P(t)\cos\beta t + Q(t)\sin\beta t]e^{\alpha t}$$

where k is the multiplicity of the root $\alpha + i\beta$ of the characteristic equqation $F(\lambda) = 0$, and $P(t), Q(t)$ are real-coefficient polynomials of order lower than m to be determined by plugging in. Note that if $\alpha + i\beta$ is a root of a polynomial then the conjugate $\alpha + i\beta$ is also a root (Complex Conjugate Root Theorem).

Example 3.3.9: Solve the ODE

$$(D^2 + 4D + 4)x = \cos 2t$$

Its characteristic equation $\lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0$ has $\lambda_{1,2} = -2$ as a root of multiplicity 2. The general solution of the homogeneous equation is then

$$x_c = (c_1 + c_2t)e^{-2t}$$

Notice that

$$A(t) = 1, B(t) = 0, m = 0, \beta = 2, \alpha = 0$$

Since $\pm 2i$ is not a root of $F(\lambda) = 0$, A particular solution of the equation is then given by

$$\widetilde{x} = A\cos 2t + B\cos 2t$$

which, plugged into the original eq, results in

$$8B\cos 2t - 8A\sin 2t = \cos 2t \Rightarrow A = 0, B = \frac{1}{8} \Rightarrow x_p = \frac{1}{8}\sin 2t$$

The general solution is then

$$x(t) = x_c + x_p = (c_1 + c_2t)e^{-2t} + \frac{1}{8}\sin 2t$$

Example 3.3.10: Determine the appropriate form for a particular solution of the fifth-order equation

$$(D - 2)^3(D^2 + 9)x = t^2e^{2t} + t\sin 3t$$

Its characteristic equation $(\lambda - 2)^3(\lambda^2 + 9) = 0$ has roots $\lambda = 2, 2, 2, +3i, -3i$. The general solution of the homogeneous equation is then

$$x_c = (c_1 + c_2t + c_3t^2)e^{2t} + c_4\cos 3t + c_5\sin 3t$$

We get a particular solution of the form

$$\widetilde{x} = t^3[At^2 + Bt + C]e^{2t} + t[(Dt + E)\cos 3t + (Ft + G)\sin 3t]$$

Exercise* 3.3.1*: *1. Solve the following ODEs:*

(1)

$$x'' + x = \sin t - \cos 2t$$

(2)

$$x'' - 4x' + 4x = e^t + e^{2t} + 1$$

(3)

$$x'' - 2x' + 2x = te^t\cos t$$

Laplace Transform Method

Definition 3.3.3 (Laplace Transform)**.** Given a function $f(t)$ defined for all $t \geq 0$, the **Laplace transform** of f is the function F defined as follows:

$$F : D \subseteq \mathbb{C} \to \mathbb{C}$$
$$s \mapsto F(s) = \mathscr{L}[f(t)](s) := \int_0^{+\infty} f(t)e^{-st}\mathrm{d}t$$

where D is the domain of the transform where the integral converges for $s \in D$.

In fact, the domain D always takes the form $\mathrm{Re}(s) > \sigma$ for which $|f(t)| < Me^{\sigma t}$ with some positive constants σ and M. $|\cdot|$ stands for the norm (modulus) in $\mathbb{C}$. Below is a table of frequently used Laplace transforms.

function $f(t) = \mathscr{L}^{-1}[F(s)](t)$	transform $\mathscr{L}[f(t)](s)$	Domain D
unit step: $u(t) = \mathbb{1}_{[0,+\infty)}$	$1/s$	$\mathrm{Re}(s) > 0$
delayed unit step: $u_c(t) = u(t-c) = \mathbb{1}_{[c,+\infty)}$	e^{-cs}/s	$\mathrm{Re}(s) > 0$
unit impulse (Dirac function): $\delta(t)$	1	all s
delayed unit impulse: $\delta_c(t) = \delta(t-c)$	e^{-cs}	$\mathrm{Re}(s) > 0$
t	$1/s^2$	$\mathrm{Re}(s) > 0$
$t^n(n = 0, 1, \cdots)$	$n!/s^{n+1}$	$\mathrm{Re}(s) > 0$
$t^q(\mathrm{Re}(q) > -1)$	$\Gamma(q+1)/s^{q+1}$	$\mathrm{Re}(s) > 0$
e^{zt}	$1/(s-z)$	$\mathrm{Re}(s) > \mathrm{Re}(z)$
te^{zt}	$1/(s-z)^2$	$\mathrm{Re}(s) > \mathrm{Re}(z)$
$t^n e^{zt}(n > -1)$	$n!/(s-z)^{n+1}$	$\mathrm{Re}(s) > \mathrm{Re}(z)$
$\sin \omega t$	$\omega/(s^2+\omega^2)$	$\mathrm{Re}(s) > 0$
$\cos \omega t$	$s/(s^2+\omega^2)$	$\mathrm{Re}(s) > 0$
$\sinh \omega t$	$\omega/(s^2-\omega^2)$	$\mathrm{Re}(s) > \|\omega\|$
$\cosh \omega t$	$s/(s^2-\omega^2)$	$\mathrm{Re}(s) < \|\omega\|$
$t \sin \omega t$	$2s\omega/(s^2+\omega^2)^2$	$\mathrm{Re}(s) > 0$
$t \cos \omega t$	$(s^2-\omega^2)/(s^2+\omega^2)^2$	$\mathrm{Re}(s) > 0$
$e^{\lambda t} \sin \omega t$	$\omega/(s-\lambda)^2+\omega^2$	$\mathrm{Re}(s) > \lambda$
$e^{\lambda t} \cos \omega t$	$(s-\lambda)/(s-\lambda)^2+\omega^2$	$\mathrm{Re}(s) > \lambda$
$te^{\lambda t} \sin \omega t$	$2\omega(s-\lambda)/[(s-\lambda)^2+\omega^2]^2$	$\mathrm{Re}(s) > \lambda$
$te^{\lambda t} \cos \omega t$	$(s-\lambda)^2-\omega^2/[(s-\lambda)^2+\omega^2]^2$	$\mathrm{Re}(s) > \lambda$

Now we explain how can Laplace transform help us solve initial value problem:

suppose we have an ODE

$$L[x] = \sum_{i=0}^{n} a_i \frac{\mathrm{d}^i x}{\mathrm{d}t^i} = a_n \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}x}{\mathrm{d}t} + a_0 = f(t), a_n = 1$$

with initial values

$$x(0) = x_0, x'(0) = x_0', \cdots, x^{(n-1)}(0) = x_0^{(n-1)}$$

and $f(t)$ is continuous, defined on $t \in [0, \infty)$ with $|f(t)| < Me^{\sigma t}$. Denote

$$F(s) = \mathscr{L}[f(t)](s) \equiv \int_0^{+\infty} f(t)e^{-st}\mathrm{d}t$$
$$X(s) = \mathscr{L}[x(t)](s) \equiv \int_0^{+\infty} x(t)e^{-st}\mathrm{d}t$$

Then

$$\begin{aligned}\mathscr{L}[x'(t)](s) &\equiv \int_0^{+\infty} x'(t)e^{-st}\mathrm{d}t = \int_0^{+\infty} e^{-st}\mathrm{d}(x(t)) \\ &= \left[e^{-st}x(t)\right]_0^{\infty} - \int_0^{+\infty}(-s)x(t)e^{-st}\mathrm{d}t \\ &= 0 - x(0) + s\int_0^{+\infty} x(t)e^{-st}\mathrm{d}t = sX(s) - x_0\end{aligned}$$

By induction (assume $\mathscr{L}[x^{(k)}(t)](s) = s^k X(s) - s^{k-1}x_0 - s^{k-2}x_0' - \cdots - x_0^{(k-1)}$) we see

$$\mathscr{L}[x^{(n)}(t)](s) = s^n X(s) - \sum_{i=0}^{n-1} x_0^{(i)} s^{n-1-i}, n \geq 1$$

because

$$\begin{aligned}
\mathscr{L}[x^{(k+1)}(t)](s) \equiv \int_0^{+\infty} x^{(k+1)}(t)e^{-st}\mathrm{d}t &= \int_0^{+\infty} e^{-st}\mathrm{d}(x^{(k)}(t)) \\
&= \left[e^{-st}x^{(k)}(t)\right]_0^{\infty} + s\int_0^{+\infty} x^{(k)}(t)e^{-st}\mathrm{d}t \\
&= -x^{(k)}(0) + s[x^{(k)}(t)](s) \\
&= s\left(s^k X(s) - \sum_{i=0}^{k-1} x_0^{(i)} s^{k-1-i}\right) - x_0^{(k)} \\
&= s^{k+1}X(s) - \sum_{i=0}^{k} x_0^{(i)} s^{k-i}
\end{aligned}$$

Therefore, transform two sides of the equation $L[x] = f(t)$ gives

$$\begin{aligned}
\mathscr{L}\left[\sum_{i=0}^{n} a_i \frac{\mathrm{d}^i x}{\mathrm{d}t^i}\right] =& a_0 X(s) + \sum_{i=1}^{n} a_i \left[s^{(i)}X(s) - \sum_{j=0}^{i-1} x_0^{(i)} s^{i-1-j}\right] \\
=& a_0 X(s) + a_1\left(sX(s) - x_0\right) \\
&+ a_2\left(s^2 X(s) - s x_0 - x_0'\right) \\
&+ \cdots \\
&+ a_n\left(s^n X(s) - s^{n-1}x_0 - s^{n-2}x_0' - \cdots - x_0^{(n-1)}\right) = F(s)
\end{aligned}$$

Let

$$A(s) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0 = \sum_{i=0}^{n} a_i s^i$$

$$B(s) = (a_1 + a_2 s + \cdots + a_n s^{n-1})x_0 + (a_2 + \cdots + a_n s^{n-2})x_0' + \cdots + a_n x_0^{(n-1)}$$

Then $L[x] = f(t)$ transforms to

$$A(s)X(s) = F(s) + B(s)$$

Since $A(s), B(s)$, and $F(s)$ are known (computable) polynomials we have

$$X(s) = \frac{F(s) + B(s)}{A(s)} \Rightarrow x(t) = \mathscr{L}^{-1}[X(s)](t) = \mathscr{L}^{-1}\left[\frac{F(s) + B(s)}{A(s)}\right](t)$$

as the solution of the original ODE.

Before we explore some of the properties of the Laplace transform, we shall first see examples to which it can be applied.

Example 3.3.11: solve the following IVP

$$\frac{\mathrm{d}x}{\mathrm{d}t} - x = e^{2t}, \text{ i.v. } x(0) = 0$$

solution:

$$F(s) = \mathscr{L}[e^{2t}](s) = \frac{1}{s-2}, s > 2$$

$$A(s) = s - 1$$

$$B(s) = x_0 = 0$$

Thus

$$X(s)=\frac{F(s)+B(s)}{A(s)}=\frac{\frac{1}{s-2}}{s-1}=\frac{1}{(s-1)(s-2)}=\frac{1}{s-2}-\frac{1}{s-1}$$
$$\Rightarrow x(t)=\mathscr{L}^{-1}\left[\frac{1}{s-2}-\frac{1}{s-1}\right](t)=\mathscr{L}^{-1}\left[\frac{1}{s-2}\right](t)-\mathscr{L}^{-1}\left[\frac{1}{s-1}\right](t)=e^{2t}-e^{t}$$

is the solution of IVP.

Example 3.3.12: solve the following IVP

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+2\frac{\mathrm{d}y}{\mathrm{d}x}+y=e^{-x},\ \text{ i.v. } y(1)=y'(1)=0$$

solution: Notice that we need to have initial values at $x=0$. Hence, we do the translation $z=\varphi(x)=x-1$ or $x=\varphi^{-1}(z)=\phi(z)=z+1$ so that

$$y(x)=y(\phi(z))$$
$$\frac{\mathrm{d}y(\phi(z))}{\mathrm{d}z}=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot 1=\frac{\mathrm{d}y}{\mathrm{d}x}$$

$$\begin{aligned}\frac{\mathrm{d}^2y(\phi(z))}{\mathrm{d}z^2}&=\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}y(\phi(z))}{\mathrm{d}z}\right)=\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}\right)\\&=\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}y}{\mathrm{d}x}(\phi(z))\right)\cdot\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}+\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}\right)\\&=\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\cdot\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}\right]\cdot\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}+\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}\right)\\&=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\left(\frac{\mathrm{d}\phi(z)}{\mathrm{d}z}\right)^2+\frac{\mathrm{d}^2\phi(z)}{\mathrm{d}z^2}\\&=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\cdot 1^2+0=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\end{aligned}$$

Therefore the IVP becomes

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+2\frac{\mathrm{d}y}{\mathrm{d}x}+y=e^{-z-1}=e^{-1}e^{-z},\ \text{ i.v. } y(0)=y'(0)=0$$

we then perform the Laplace transform method:

$$F(s)=\mathscr{L}[e^{-1}e^{-z}](s)=\frac{1}{e}\cdot\frac{1}{s+1},s>-1$$
$$A(s)=s^2+2s+1$$
$$B(s)=(1s+2)y_0+y_0'=0$$

Thus

$$Y(s)=\frac{F(s)+B(s)}{A(s)}=\frac{\frac{1}{e}\cdot\frac{1}{s+1}}{s^2+2x+1}=\frac{1}{e}\frac{1}{(s+1)^3}=\frac{1}{2e}\frac{2!}{(s+1)^3}$$
$$\Rightarrow y(z)=\mathscr{L}^{-1}\left[\frac{1}{2e}\frac{2!}{(s+1)^3}\right](z)=\frac{1}{2}\cdot z^2e^{-z-1}$$

However, one should carefully interpret this result. In particular, writing $y(x)=y(\phi(z))$ will give an incorrect answer because this notation misunderstands y as a mapping. A safer way of writing this can be

$$y=g(z)=\frac{1}{2}\cdot z^2e^{-z-1}$$

and we want to write y in terms of x by combining the above result with relationship between x and z.

$$y=g(z)=g(\varphi(x))=g(x-1)=\frac{1}{2}\cdot(x-1)^2e^{-x}$$

which is the solution of IVP.

We then list some properties of the Laplace transform.

some common total derivatives:

- $$x\mathrm{d}y + y\mathrm{d}x = d(xy)$$
- $$\frac{y\mathrm{d}x - x\mathrm{d}y}{y^2} = \mathrm{d}\left(\frac{x}{y}\right)$$
- $$\frac{-y\mathrm{d}x + x\mathrm{d}y}{x^2} = \mathrm{d}\left(\frac{y}{x}\right)$$
- $$\frac{y\mathrm{d}x - x\mathrm{d}y}{xy} = \mathrm{d}\left(\ln\left|\frac{x}{y}\right|\right)$$
- $$\frac{y\mathrm{d}x - x\mathrm{d}y}{x^2 + y^2} = \mathrm{d}\left(\arctan\frac{x}{y}\right)$$
- $$\frac{y\mathrm{d}x - x\mathrm{d}y}{x^2 - y^2} = \frac{1}{2}\mathrm{d}\left(\ln\left|\frac{x-y}{x+y}\right|\right)$$

***Exercise** 3.3.2:* *1. Given fundamental pairs $\{x_1, x_2\}$, solve the following ODEs:*

(1) $x_1 = e^t, x_2 = e^{-t}$.
$$x'' - x = \cos t$$

(2) $x_1 = t, x_2 = e^t$.
$$x'' + \frac{t}{1-t}x' - \frac{1}{1-t}x = t - 1$$

(3) $x_1 = t^2, x_2 = t^3$.
$$t^2x'' - 4tx' + 6x = 36\frac{\ln t}{t}$$

(4) $x_1 = t^2\cos(2\ln t), x_2 = t^2\sin 2\ln t$.
$$t^2x'' - 3tx' = 8x = 18t^2\sin(\ln t)$$

2. let $x_i(t)(i = 1, 2, \cdots, n)$ be any n solutions of the the homogeneous linear ODE (3.2), and let $W(t)$ be their Wronskian determinant. Prove that $W(t)$ satisfies the following first order linear ODE:
$$W' + a_1(t)W = 0$$
and thus
$$W(t) = W(t_0)t_0\exp\left(-\int_{t_0}^{t} a_1(s)ds\right), t \in (a, b)$$

3.4 Systems of Constant Coefficient Linear ODEs

We now present general theory of systems of linear ODEs. We first look an example.

Example 3.4.1: We consider salt and brine tanks, and water flows from one to the other and back. Let the tanks be evenly mixed.

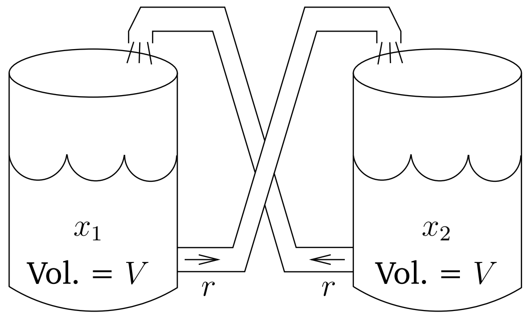

Figure 3.1: A closed system of two brine tanks.

Suppose we have two tanks, each containing volume V liters of salt brine. The amount of salt in the first tank is x_1 grams, and the amount of salt in the second tank is x_2 grams. The liquid is perfectly mixed and flows at the rate r liters per second out of each tank into the other. See 3.1.

The rate of change of x_1, that is x_1', is the rate of salt coming in minus the rate going out. The rate coming in is the density of the salt in tank 2, that is $\frac{x_2}{V}$, times the rate r. The rate coming out is the density of the salt in tank 1, that is $\frac{x_1}{V}$, times the rate r. In other words it is

$$x_1' = \frac{x_2}{V}r - \frac{x_1}{V}r = \frac{r}{V}x_2 - \frac{r}{V}x_1 = \frac{r}{V}(x_2 - x_1).$$

Similarly we find the rate x_2', where the roles of x_1 and x_2 are reversed. All in all, the system of ODEs for this problem is

$$\begin{aligned} x_1' &= \frac{r}{V}(x_2 - x_1), \\ x_2' &= \frac{r}{V}(x_1 - x_2). \end{aligned}$$

In this system we cannot solve for x_1 or x_2 separately. We must solve for both x_1 and x_2 at once, which is intuitively clear since the amount of salt in one tank affects the amount in the other. We can't know x_1 before we know x_2, and vice versa.

We don't yet know how to find all the solutions, but intuitively we can at least find some solutions. Suppose we know that initially the tanks have the same amount of salt. That is, we have an initial condition such as $x_1(0) = x_2(0) = C$. Then clearly the amount of salt coming and out of each tank is the same, so the amounts are not changing. In other words, $x_1 = C$ and $x_2 = C$ (the constant functions) is a solution: $x_1' = x_2' = 0$, and $x_2 - x_1 = x_1 - x_2 = 0$, so the equations are satisfied.

Let us think about the setup a little bit more without solving it. Suppose the initial conditions are $x_1(0) = A$ and $x_2(0) = B$, for two different constants A and B. Since no salt is coming in or out of this closed system, the total amount of salt is constant. That is, $x_1 + x_2$ is constant, and so it equals $A + B$. Intuitively if A is bigger than B, then more salt will flow out of tank one than into it. Eventually, after a long time we would then expect the amount of salt in each tank to equalize. In other words, the solutions of both x_1 and x_2 should tend towards $\frac{A+B}{2}$. Once you know how to solve systems you will find out that this really is so.

3.4.1 The Matrix Exponential

Consider the following autonomous linear first-order system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0 \tag{3.16}$$

where A is an n by n matrix. If we perform the Picard iteration we obtain

$$\begin{aligned} x_0(t) &= x_0 \\ x_1(t) &= x_0 + \int_0^t A x_0(s) ds = x_0 + A x_0 \int_0^t ds = x_0 + tAx_0 \\ x_2(t) &= x_0 + \int_0^t A x_1(s) ds = x_0 + A x_0 \int_0^t ds + A^2 x_0 \int_0^t s ds \\ &= x_0 + tAx_0 + \frac{t^2}{2} A^2 x_0 \end{aligned}$$

and hence by induction

$$x_m(t) = \sum_{j=0}^{m} \frac{t^j}{j!} A^j x_0. \tag{3.17}$$

The limit as $m \to \infty$ is given by

$$x(t) = \lim_{m \to \infty} x_m(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j x_0. \tag{3.18}$$

In the one dimensional case $(n = 1)$ this series is just the usual exponential and hence we will write

$$x(t) = \exp(tA) x_0, \tag{3.19}$$

where we define the **matrix exponential** by

$$\exp(A) = \sum_{j=0}^{\infty} \frac{1}{j!} A^j. \tag{3.20}$$

Hence, in order to understand our original problem, we have to understand the matrix exponential! The Picard iteration ensures convergence of $\exp(A)x_0$ for every vector x_0 and choosing the canonical basis vectors of $\mathbb{R}^n$ we see that all matrix elements converge. However, for later use we want to introduce a suitable norm for matrices and give a direct proof for convergence of the above series in this norm.

We will use $\mathbb{C}^n$ rather than $\mathbb{R}^n$ as underlying vector space since $\mathbb{C}$ is algebraically closed (which will be important later on, when we compute the matrix exponential with the help of the Jordan canonical form). So let A be a complex matrix acting on $\mathbb{C}^n$ and introduce the **matrix norm**

$$\|A\| = \sup_{x:|x|=1} |Ax| \tag{3.21}$$

It is not hard to see that the vector space of n by n matrices $\mathbb{C}^{n \times n}$ becomes a Banach space with this norm (Problem 3.4.1). Moreover, using (Problem 3.4.2)

$$\left\|A^j\right\| \leq \|A\|^j$$

convergence of the series (3.20) follows from convergence of $\sum_{j=0}^{\infty} \frac{\|A\|^j}{j!} = \exp(\|A\|)$.

Exercise* 3.4.1*: *Show that the space of n by n matrices $\mathbb{C}^{n \times n}$ together with the matrix norm is a Banach space. In particular, show that a sequence of matrices converges if and only if all matrix entries converge. (Hint: Show that the matrix entries a_{jk} of A satisfy $\max_{j,k} |a_{jk}| \leq \|A\|$ and $\|A\| \leq n \max_{j,k} |a_{jk}|$.)*

Exercise* 3.4.2*: *Show that the matrix norm satisfies*

$$\|AB\| \leq \|A\| \|B\|.$$

(This shows that $\mathbb{C}^{n \times n}$ is even a Banach algebra.) Conclude $\left\|A^j\right\| \leq \|A\|^j$.

However, note that in general $\exp(A+B) \neq \exp(A)\exp(B)$ unless A and B commute, that is, unless the **commutator**

$$[A, B] = AB - BA$$

vanishes. In this case you can mimic the proof of the one dimensional case to obtain

Lemma 3.4.1. Suppose A and B commute. Then

$$\exp(A+B) = \exp(A)\exp(B), \quad [A, B] = 0.$$

If we perform a linear change of coordinates,

$$y = U^{-1}x,$$

then the matrix exponential in the new coordinates is given by

$$U^{-1}\exp(A)U = \exp\left(U^{-1}AU\right).$$

This follows from (3.20) by using $U^{-1}A^jU = \left(U^{-1}AU\right)^j$ together with continuity of the matrix product. Hence in order to compute $\exp(A)$ we need a coordinate transform which renders A as simple as possible:

Theorem 3.4.1 (Jordan canonical form)**.** *Let A be an n by n matrix. Then there exists a linear change of coordinates U such that A transforms into a block matrix,*

$$U^{-1}AU = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$$

with each block of the form

$$J = \alpha\mathbb{I} + N = \begin{pmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \alpha & \ddots & \\ & & & \ddots & 1 \\ & & & & \alpha \end{pmatrix} \tag{3.22}$$

Here N is a matrix with ones in the first diagonal above the main diagonal and zeros elsewhere.

The numbers α are the eigenvalues of A and the new basis vectors u_j (the columns of U) consist of generalized eigenvectors of A. The general procedure of finding the Jordan canonical form is quite cumbersome and hence and we defer the details to Appendix. In particular, since most computer algebra systems can easily do this job for us!

Example 3.4.2: Let

In[1]:= A = $\begin{pmatrix} -11 & -35 & -24 \\ -1 & -1 & -2 \\ 8 & 22 & 17 \end{pmatrix}$;

Then the command

```
In[2]:= {U, J} = JordanDecomposition[A];
```

gives us the transformation matrix U plus the Jordan canonical form $J = U^{-1}AU$.

```
In[3]:= J // MatrixForm
```

Out[3]//MatrixForm= $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

If you don't trust me (or *Mathematica*), you can also check it:

```
In[4]:= A == U.J.Inverse[U]
Out[4]= True
```

To compute the exponential we observe

$$\exp\left(U^{-1}AU\right)=\begin{pmatrix}\exp(J_1) & & \\ & \ddots & \\ & & \exp(J_m)\end{pmatrix}$$

and hence it remains to compute the exponential of a single Jordan block $J=\alpha\mathbb{I}+N$ as in (3.22). Since $\alpha\mathbb{I}$ commutes with N, we infer from Lemma 3.4.1 that

$$\exp(J)=\exp(\alpha\mathbb{I})\exp(N)=\mathrm{e}^{\alpha}\sum_{j=0}^{k-1}\frac{1}{j!}N^{j}.$$

The series for $\exp(N)$ terminates after k terms, where k is the size of N. In fact, it is not hard to see that N^j is a matrix with ones in the j 'th diagonal above the main diagonal and vanishes once j reaches the size of J :

$$N=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{pmatrix}, N^2=\begin{pmatrix}0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}, N^3=\begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

and $N^4=0$. In summary, $\exp(J)$ explicitly reads

$$\exp(J)=\mathrm{e}^{\alpha}\begin{pmatrix}1&1&\frac{1}{2!}&\cdots&\frac{1}{(k-1)!}\\ &1&1&\ddots&\vdots\\ & &1&\ddots&\frac{1}{2!}\\ & & &\ddots&1\\ & & & &1\end{pmatrix}.$$

Note that if A is in Jordan canonical form, then it is not hard to see that

$$\det(\exp(A))=\exp(\operatorname{tr}(A))$$

Since both the determinant and the trace are invariant under linear transformations, the formula also holds for arbitrary matrices. In fact, we even have (exercise):

Lemma 3.4.2. A vector u is an eigenvector of A corresponding to the eigenvalue α if and only if u is an eigenvector of $\exp(A)$ corresponding to the eigenvalue e^{α}.

Moreover, the Jordan structure of A and $\exp(A)$ are the same. In particular, both the geometric and algebraic multiplicities of α and e^{α} are the same.

Clearly Mathematica can also compute the exponential for us: `In [5] := MatrixExp[J]// MatrixForm` `Out [5]//MatrixForm=`

$$\begin{pmatrix}\mathrm{e}&0&0\\0&\mathrm{e}^2&\mathrm{e}^2\\0&0&\mathrm{e}^2\end{pmatrix}$$

To end this section let me emphasize, that both the eigenvalues and generalized eigenvectors can be complex even if the matrix A has only real entries. However, in many applications only real solutions are of interest. For such a case there is also a **real Jordan canonical form** which we want to mention briefly.

So suppose the matrix A has only real entries. If an eigenvalue α is real, both real and imaginary parts of a generalized eigenvector are again generalized eigenvectors. In particular, they can be chosen real and there is nothing else to do for such an eigenvalue.

If α is nonreal, there must be a corresponding complex conjugate block $J^* = \alpha^* \mathbb{I} + N$ and the corresponding generalized eigenvectors can be assumed to be the complex conjugates of our original ones. Therefore we can replace the pairs u_j, u_j^* in our basis by $\operatorname{Re}(u_j)$ and $\operatorname{Im}(u_j)$. In this new basis the block $J \oplus J^*$ is replaced by

$$\begin{pmatrix} R & \mathbb{I} & & & \\ & R & \mathbb{I} & & \\ & & R & \ddots & \\ & & & \ddots & \mathbb{I} \\ & & & & R \end{pmatrix}$$

where

$$R = \begin{pmatrix} \operatorname{Re}(\alpha) & \operatorname{Im}(\alpha) \\ -\operatorname{Im}(\alpha) & \operatorname{Re}(\alpha) \end{pmatrix} \quad \text{and} \quad \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

commute, the exponential is given by

$$\begin{pmatrix} \exp(R) & \exp(R) & \exp(R)\frac{1}{2!} & \cdots & \exp(R)\frac{1}{(n-1)!} \\ \exp(R) & \exp(R) & \ddots & \vdots & \\ & & \exp(R) & \ddots & \exp(R)\frac{1}{2!} \\ & & \ddots & \exp(R) & \\ & & & \exp(R) & \end{pmatrix}$$

where

$$\exp(R) = \mathrm{e}^{\operatorname{Re}(\alpha)} \begin{pmatrix} \cos(\operatorname{Im}(\alpha)) & \sin(\operatorname{Im}(\alpha)) \\ -\sin(\operatorname{Im}(\alpha)) & \cos(\operatorname{Im}(\alpha)) \end{pmatrix}$$

3.4.2 Linear Autonomous First-Order Systems

In the previous section we have seen that the solution of the autonomous linear first-order system (3.16) is given by

$$x(t) = \exp(tA)x_0. \tag{3.23}$$

In particular, the map $\exp(tA)$ provides an isomorphism between all initial conditions x_0 and all solutions. Hence the set of all solutions is a vector space isomorphic to $\mathbb{R}^n$ (respectively $\mathbb{C}^n$ if we allow complex initial values).

In order to understand the dynamics of the system (3.16), we need to understand the properties of the function $\exp(tA)$. We will start with the case of two dimensions which covers all prototypical cases. Furthermore, we will assume A as well as x_0 to be real-valued.

In this situation there are two eigenvalues, α_1 and α_2, which are either both real or otherwise complex conjugates of each other. We begin with the generic case where A is diagonalizable and hence there are two linearly independent eigenvectors, u_1 and u_2, which form the columns of U. In particular,

$$U^{-1}AU = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

and the solution (3.23) is given by

$$x(t) = U \exp\left(tU^{-1}AU\right) U^{-1}x_0 = U \begin{pmatrix} \mathrm{e}^{\alpha_1 t} & 0 \\ 0 & \mathrm{e}^{\alpha_2 t} \end{pmatrix} U^{-1}x_0.$$

Abbreviating $y_0 = U^{-1}x_0 = (y_{0,1}, y_{0,2})$ we obtain

$$x(t) = y_{0,1}\mathrm{e}^{\alpha_1 t}u_1 + y_{0,2}\mathrm{e}^{\alpha_2 t}u_2. \tag{3.24}$$

In the case where both eigenvalues are real, all quantities in (3.24) are real. Otherwise we have $\alpha_2 = \alpha_1^*$ and we can assume $u_2 = u_1^*$ without loss of generality. Let us write $\alpha_1 \equiv \alpha = \lambda + \mathrm{i}\omega$ and $\alpha_2 \equiv \alpha^* = \lambda - \mathrm{i}\omega$. Then **Euler's formula**

$$\mathrm{e}^{\mathrm{i}\omega} = \cos(\omega) + \mathrm{i}\sin(\omega)$$

implies

$$\mathrm{e}^{\alpha t} = \mathrm{e}^{\lambda t}(\cos(\omega t) + \mathrm{i}\sin(\omega t)), \quad \alpha = \lambda + \mathrm{i}\omega. \tag{3.25}$$

Moreover, $x_0^* = x_0$ implies $y_{0,1}u_1 + y_{0,2}u_2 = y_{0,1}^*u_2 + y_{0,2}^*u_1$ which shows $y_{0,1}^* = y_{0,2}$. Hence, both terms in (3.24) are complex conjugates of each other implying

$$\begin{aligned} x(t) &= 2\,\mathrm{Re}\left(y_{0,1}\mathrm{e}^{\alpha_1 t}u_1\right) \\ &= 2\cos(\omega t)\mathrm{e}^{\lambda t}\,\mathrm{Re}\left(y_{0,1}u_1\right) - 2\sin(\omega t)\mathrm{e}^{\lambda t}\,\mathrm{Im}\left(y_{0,1}u_1\right). \end{aligned}$$

This finishes the case where A is diagonalizable.

If A is not diagonalizable, both eigenvalues must be equal $\alpha_1 = \alpha_2 \equiv \alpha$. The columns u_1 and u_2 of the matrix U are the eigenvector and generalized eigenvector of A, respectively. Hence

$$U^{-1}AU = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$$

and with a similar computation as before the solution is given by

$$x(t) = (y_{0,1} + y_{0,2}t)\,\mathrm{e}^{\alpha t}u_1 + y_{0,2}\mathrm{e}^{\alpha t}u_2.$$

This finishes the case where A is not diagonalizable. Next, let us try to understand the qualitative behavior for large t. For this we need to understand the function $\exp(\alpha t)$. From (3.25) we can read off that $\exp(\alpha t)$ will converge to 0 as $t \to \infty$ if $\lambda = \mathrm{Re}(\alpha) < 0$ and grow exponentially if $\lambda = \mathrm{Re}(\alpha) > 0$. It remains to discuss the possible cases according to the respective signs of $\mathrm{Re}(\alpha_1)$ and $\mathrm{Re}(\alpha_2)$.

Case 1. Suppose that the eigenvalues of A are real and positive. We find two corresponding eigenvectors and plot them in the plane. For example, take the matrix $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right]$. The eigenvalues are 1 and 2 and corresponding eigenvectors are $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. See 3.2.

Let (x, y) be a point on the line determined by an eigenvector $\vec{v}$ for an eigenvalue λ. That is, $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \alpha\vec{v}$ for some scalar α. Then

$$\begin{bmatrix} x \\ y \end{bmatrix}' = A\begin{bmatrix} x \\ y \end{bmatrix} = A(\alpha\vec{v}) = \alpha(P\vec{v}) = \alpha\lambda\vec{v}.$$

The derivative is a multiple of $\vec{v}$ and hence points along the line determined by $\vec{v}$. As $\lambda > 0$, the derivative points in the direction of $\vec{v}$ when α is positive and in the opposite direction when α is negative. We draw the lines determined by the eigenvectors, and we draw arrows on the lines to indicate the directions. See 3.3.

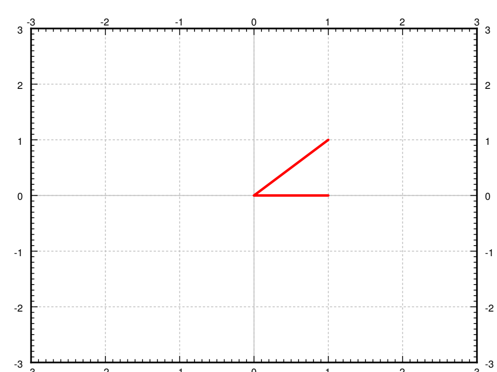

Figure 3.2: Eigenvectors of A.

We fill in the rest of the arrows for the vector field and we also draw a few solutions. See 3.4. The picture looks like a source with arrows coming out from the origin. Hence we call this type of picture a **source** or sometimes an **unstable node**.

Case 2. Suppose both eigenvalues are negative. For example, take the negation of the matrix in case 1, $\left[\begin{smallmatrix} -1 & -1 \\ 0 & -2 \end{smallmatrix}\right]$. The eigenvalues are -1 and -2 and corresponding eigenvectors are the same, $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. The calculation and the picture are almost the same. The only difference is that the eigenvalues are negative and hence all arrows are reversed. We get the picture in 3.5. We call this kind of picture a **sink** or a **stable node**.

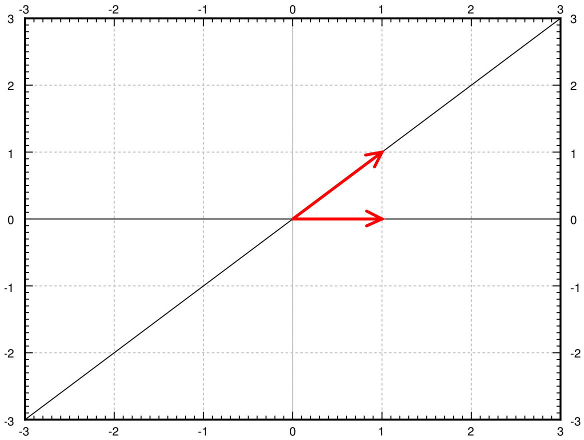

Figure 3.3: Eigenvectors of A with directions.

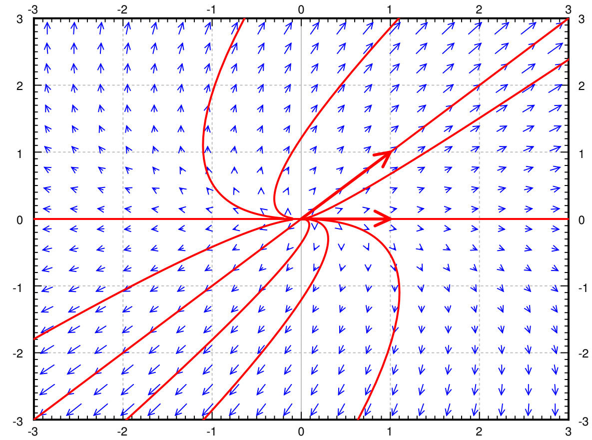

Figure 3.4: Example source vector field with eigenvectors and solutions.

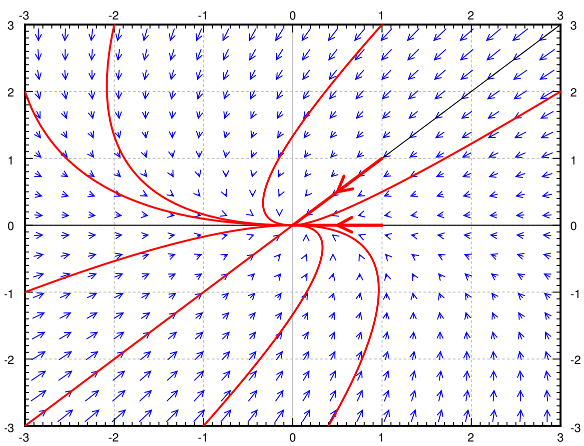

Figure 3.5: Example sink vector field with eigenvectors and solutions.

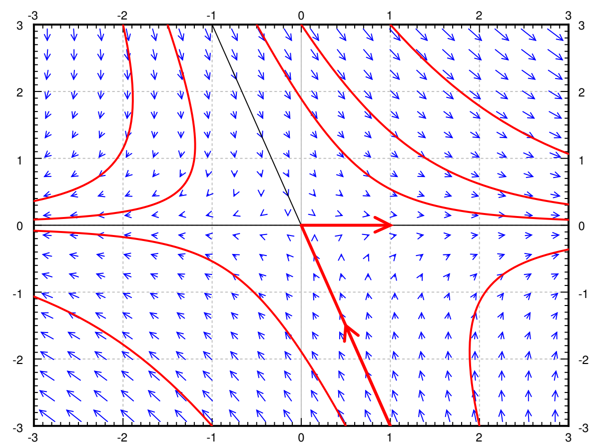

Figure 3.6: Example saddle vector field with eigenvectors and solutions.

Case 3. Suppose one eigenvalue is positive and one is negative. For example the matrix $\left[\begin{smallmatrix}1 & 1\\ 0 & -2\end{smallmatrix}\right]$. The eigenvalues are 1 and -2 and corresponding eigenvectors are $\left[\begin{smallmatrix}1\\ 0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}1\\ -3\end{smallmatrix}\right]$. We reverse the arrows on one line (corresponding to the negative eigenvalue) and we obtain the picture in 3.6. We call this picture a **saddle point**.

For the next three cases we will assume the eigenvalues are complex. In this case the eigenvectors are also complex and we cannot just plot them in the plane.

Case 4. Suppose the eigenvalues are purely imaginary, that is, $\pm ib$. For example, let $A = \left[\begin{smallmatrix}0 & 1\\ -4 & 0\end{smallmatrix}\right]$. The eigenvalues are $\pm 2i$ and corresponding eigenvectors are $\left[\begin{smallmatrix}1\\ 2i\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}1\\ -2i\end{smallmatrix}\right]$. Consider the eigenvalue $2i$ and its eigenvector $\left[\begin{smallmatrix}1\\ 2i\end{smallmatrix}\right]$. The real and imaginary parts of $\vec{v}e^{2it}$ are

$$\operatorname{Re}\begin{bmatrix}1\\ 2i\end{bmatrix} e^{2it} = \begin{bmatrix}\cos(2t)\\ -2\sin(2t)\end{bmatrix}, \qquad \operatorname{Im}\begin{bmatrix}1\\ 2i\end{bmatrix} e^{2it} = \begin{bmatrix}\sin(2t)\\ 2\cos(2t)\end{bmatrix}.$$

We can take any linear combination of them to get other solutions, which one we take depends on the initial conditions. Now note that the real part is a parametric equation for an ellipse. Same with the imaginary part and in fact any linear combination of the two. This is what happens in general when the eigenvalues are purely imaginary. So when the eigenvalues are purely imaginary, we get **ellipses** for the solutions. This type of picture is sometimes called a **center**. See 3.7.

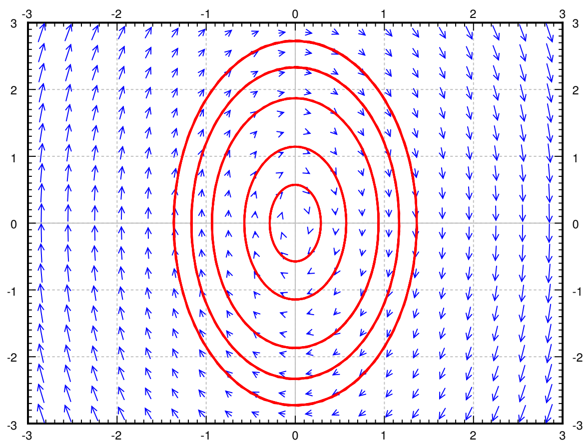

Figure 3.7: Example center vector field.

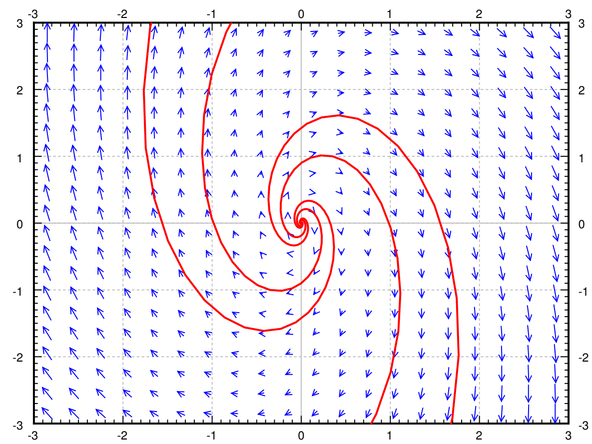

Figure 3.8: Example spiral source vector field.

Case 5. Now suppose the complex eigenvalues have a positive real part. That is, suppose the eigenvalues are $a \pm ib$ for some $a > 0$. For example, let $A = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix}$. The eigenvalues turn out to be $1 \pm 2i$ and eigenvectors are $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2i \end{bmatrix}$. We take $1 + 2i$ and its eigenvector $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$ and find the real and imaginary parts of $\vec{v}e^{(1+2i)t}$ are

$$\operatorname{Re} \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(1+2i)t} = e^t \begin{bmatrix} \cos(2t) \\ -2\sin(2t) \end{bmatrix}, \qquad \operatorname{Im} \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(1+2i)t} = e^t \begin{bmatrix} \sin(2t) \\ 2\cos(2t) \end{bmatrix}.$$

Note the e^t in front of the solutions. The solutions grow in magnitude while spinning around the origin. Hence we get a **spiral source**. See 3.8.

Case 6. Finally suppose the complex eigenvalues have a negative real part. That is, suppose the eigenvalues are $-a \pm ib$ for some $a > 0$. For example, let $A = \begin{bmatrix} -1 & -1 \\ 4 & -1 \end{bmatrix}$. The eigenvalues turn out to be $-1 \pm 2i$ and eigenvectors are $\begin{bmatrix} 1 \\ -2i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$. We take $-1 - 2i$ and its eigenvector $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$ and find the real and imaginary parts of $\vec{v}e^{(-1-2i)t}$ are

$$\operatorname{Re} \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(-1-2i)t} = e^{-t} \begin{bmatrix} \cos(2t) \\ 2\sin(2t) \end{bmatrix}, \qquad \operatorname{Im} \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(-1-2i)t} = e^{-t} \begin{bmatrix} -\sin(2t) \\ 2\cos(2t) \end{bmatrix}.$$

Note the e^{-t} in front of the solutions. The solutions shrink in magnitude while spinning around the origin. Hence we get a **spiral sink**. See 3.9.

We summarize the behavior of linear homogeneous two-dimensional systems given by a nonsingular matrix in 3.1. Systems where one of the eigenvalues is zero (the matrix is singular) come up in practice from time to time, see 3.4.1, and the pictures are somewhat different (simpler in a way). See the exercises.

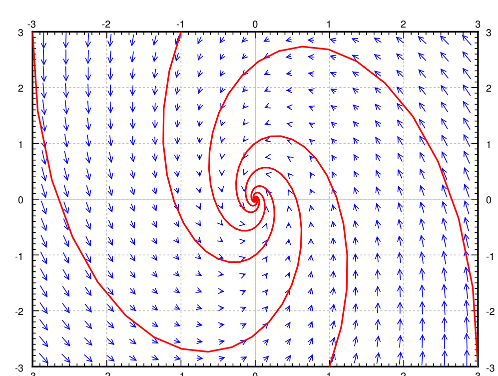

Figure 3.9: Example spiral sink vector field.

Exercise 3.4.3: *Take the equation $mx'' + cx' + kx = 0$, with $m > 0$, $c \geq 0$, $k > 0$ for the mass-spring system.*

a) Convert this to a system of first order equations.
b) Classify for what m, c, k do you get which behavior.
c) Explain from physical intuition why you do not get all the different kinds of behavior here?

Eigenvalues	Behavior
real and both positive	source / unstable node
real and both negative	sink / stable node
real and opposite signs	saddle
purely imaginary	center point / ellipses
complex with positive real part	spiral source
complex with negative real part	spiral sink

Table 3.1: Summary of behavior of linear homogeneous two-dimensional systems.

***Exercise* 3.4.4:** *What happens in the case when $A = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$? In this case the eigenvalue is repeated and there is only one independent eigenvector. What picture does this look like?*

***Exercise* 3.4.5:** *What happens in the case when $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$? Does this look like any of the pictures we have drawn?*

***Exercise* 3.4.6:** *Which behaviors are possible if A is diagonal, that is $A = \left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$? You can assume that a and b are not zero.*

***Exercise* 3.4.7:** *Take the system from 3.4.1, $x_1' = \frac{r}{V}(x_2 - x_1)$, $x_2' = \frac{r}{V}(x_1 - x_2)$. As we said, one of the eigenvalues is zero. What is the other eigenvalue, how does the picture look like and what happens when t goes to infinity.*

***Exercise* 3.4.101:** *Describe the behavior of the following systems without solving:*

a) $x' = x + y, \quad y' = x - y.$

b) $x_1' = x_1 + x_2, \quad x_2' = 2x_2.$

c) $x_1' = -2x_2, \quad x_2' = 2x_1.$

d) $x' = x + 3y, \quad y' = -2x - 4y.$

e) $x' = x - 4y, \quad y' = -4x + y.$

***Exercise* 3.4.102:** *Suppose that $\vec{x}\,' = A\vec{x}$ where A is a 2 by 2 matrix with eigenvalues $2 \pm i$. Describe the behavior.*

***Exercise* 3.4.103:** *Take $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]' = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$. Draw the vector field and describe the behavior. Is it one of the behaviors that we have seen before?*

Now we come to the general case. As before, the considerations of the previous section show that it suffices to consider the case of one Jordan block

$$\exp(tJ) = \mathrm{e}^{\alpha t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ & 1 & t & \ddots & \vdots \\ & & 1 & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix}.$$

In particular, every solution is a linear combination of terms of the type $t^j \exp(\alpha t)$. Since $\exp(\alpha t)$ decays faster than any polynomial, our entire Jordan block converges to zero if $\lambda = \operatorname{Re}(\alpha) < 0$. If $\lambda = 0, \exp(\alpha t) = \exp(\mathrm{i}\omega t)$ will remain at least bounded, but the polynomial terms will diverge. However, if we start in the direction of the eigenvector $(1, 0, \ldots, 0)$, we won't see the polynomial terms. In summary,

Theorem 3.4.2. *A solution of the linear system (3.16) converges to 0 as $t \to \infty$ if the initial condition x_0 lies in the subspace spanned by the generalized eigenspaces corresponding to eigenvalues with negative real part. It will remain bounded if x_0 lies in the subspace spanned by the generalized eigenspaces corresponding to eigenvalues with negative real part plus the eigenspaces corresponding to eigenvalues with vanishing real part.*

Note that to get the behavior as $t \to -\infty$, you just need to replace negative by positive.

A linear system (not necessarily autonomous) is called **stable** if all solutions remain bounded as $t \to \infty$ and **asymptotically stable** if all solutions converge to 0 as $t \to \infty$.

Theorem 3.4.1. *The linear system (3.16) is stable if all eigenvalues α of A satisfy $\operatorname{Re}(\alpha) \leq 0$ and for all eigenvalues with $\operatorname{Re}(\alpha) = 0$ the algebraic and geometric multiplicities are equal.*

The linear system (3.16) is asymptotically stable if all eigenvalues α of A satisfy $\operatorname{Re}(\alpha) < 0$. Finally, observe that the solution of the inhomogeneous equation

$$\dot{x}(t) = Ax(t) + g(t), \quad x(0) = x_0$$

is given by

$$x(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)g(s)ds,$$

which can be verified by a straightforward computation (however, we will in fact prove a more general result in Theorem **??** below). As always for linear equations, note that the solutions consists of the general solution of the linear equation plus a particular solution of the inhomogeneous equation.

3.5 Systems of Non-constant Coefficient Linear ODEs

In this section we want to consider the case of linear systems, where the coefficient matrix can depend on t. As a preparation let me remark that a matrix $A(t)$ is called differentiable with respect to t if all coefficients are. In this case we will denote by $\frac{d}{dt}A(t) \equiv \dot{A}(t)$ the matrix, whose coefficients are the derivatives of the coefficients of $A(t)$. The usual rules of calculus hold in this case as long as one takes noncommutativity of matrices into account. For example we have the product rule

$$\frac{d}{dt}A(t)B(t) = \dot{A}(t)B(t) + A(t)\dot{B}(t)$$

and, if $\det(A(t)) \neq 0$,

$$\frac{d}{dt}A(t)^{-1} = -A(t)^{-1}\dot{A}(t)A(t)^{-1}$$

(exercise. Hint: $AA^{-1} = \mathbb{I}$). Note that the order is important! We now turn to the general linear first-order system

$$\dot{x}(t) = A(t)x(t), \tag{3.26}$$

where $A \in C\left(I, \mathbb{R}^{n \times n}\right)$. Clearly our theory from chapter 1 applies:

Theorem 3.5.1. *The linear first-order system (3.26) has a unique solution satisfying the initial condition $x\left(t_0\right) = x_0$. Moreover, this solution is defined for all $t \in I$.*

Proof. This follows directly from Theorem 1.0.5 since we can choose $L(T) = \max_{[0,T]} \|A(t)\|$ for every $T \in I$. □

It seems tempting to suspect that the solution is given by the formula $x(t) = \exp\left(\int_{t_0}^t A(s)ds\right)x_0$. However, as soon as you try to verify this guess, noncommutativity of matrices will get into your way. In fact, this formula only solves our initial value problem if $[A(t), A(s)] = 0$ for all $t, s \in \mathbb{R}$. Hence it is of little use. So we still need to find the right generalization of $\exp\left(\left(t - t_0\right)A\right)$.

We start by observing that linear combinations of solutions are again solutions. Hence the set of all solutions forms a vector space. This is often referred to as **superposition principle**. In particular, the solution corresponding to the initial condition $x\left(t_0\right) = x_0$ can be written as

$$\phi\left(t, t_0, x_0\right) = \sum_{j=1}^n \phi\left(t, t_0, \delta_j\right) x_{0,j},$$

where δ_j are the canonical basis vectors, (i.e, $\delta_{j,k} = 1$ if $j = k$ and $\delta_{j,k} = 0$ if $j \neq k$) and $x_{0,j}$ are the components of x_0 (i.e., $x_0 = \sum_{j=1}^n \delta_j x_{0,j}$). Using the solutions $\phi\left(t, t_0, \delta_j\right)$ as columns of a matrix

$$\Pi\left(t, t_0\right) = \left(\phi\left(t, t_0, \delta_1\right), \ldots, \phi\left(t, t_0, \delta_n\right)\right),$$

we see that there is a linear mapping $x_0 \mapsto \phi(t, t_0, x_0)$ given by

$$\phi(t, t_0, x_0) = \Pi(t, t_0)\, x_0.$$

The matrix $\Pi(t, t_0)$ is called **principal matrix solution** (at t_0) and it solves the matrix valued initial value problem

$$\dot{\Pi}(t, t_0) = A(t)\Pi(t, t_0), \quad \Pi(t_0, t_0) = \mathbb{I}. \tag{3.27}$$

Again observe that our basic existence and uniqueness result applies. In fact, it is easy to check, that a matrix $X(t)$ satisfies $\dot{X} = A(t)X$ if and only if every column satisfies (3.26). In particular, $X(t)c$ solves (3.26) for every constant vector c in this case. In summary,

Theorem 3.5.2. *The solutions of the system (3.26) form an n dimensional vector space. Moreover, there exists a matrix-valued solution $\Pi(t, t_0)$ such that the solution satisfying the initial condition $x(t_0) = x_0$ is given by $\Pi(t, t_0)\, x_0$.*

Example 3.5.1: In the simplest case, where $A(t) \equiv A$ is constant, we of course have $\Pi(t, t_0) = \mathrm{e}^{(t-t_0)A}$.

Example 3.5.2: Consider the system

$$\dot{x} = \begin{pmatrix} 1 & t \\ 0 & 2 \end{pmatrix} x$$

which explicitly reads

$$\dot{x}_1 = x_1 + t x_2, \quad \dot{x}_2 = 2x_2.$$

We need to find the solution corresponding to the initial conditions $x(t_0) = \delta_1 = (1,0)$ respectively $x(t_0) = \delta_2 = (0,1)$. In the first case $x(t_0) = \delta_1$, the second equation gives $x_2(t) = 0$ and plugging this into the first equation shows $x_1(t) = \mathrm{e}^{t-t_0}$, that is, $\phi(t, t_0, \delta_1) = (\mathrm{e}^{t-t_0}, 0)$. Similarly, in the second case $x(t_0) = (0,1)$, the second equation gives $x_2(t) = \mathrm{e}^{2(t-t_0)}$ and plugging this into the first equation shows $x_1(t) = \mathrm{e}^{2(t-t_0)}(t-1) - \mathrm{e}^{t-t_0}(t_0 - 1)$, that is, $\phi(t, t_0, \delta_2) = \left(\mathrm{e}^{2(t-t_0)}(t-1) - \mathrm{e}^{t-t_0}(t_0-1), \mathrm{e}^{2(t-t_0)}\right)$. Putting everything together we obtain

$$\Pi(t, t_0) = \begin{pmatrix} \mathrm{e}^{t-t_0} & \mathrm{e}^{2(t-t_0)}(t-1) - \mathrm{e}^{t-t_0}(t_0-1) \\ 0 & \mathrm{e}^{2(t-t_0)} \end{pmatrix}.$$

Furthermore, $\Pi(t, t_0)$ satisfies

$$\Pi(t, t_1)\, \Pi(t_1, t_0) = \Pi(t, t_0) \tag{3.28}$$

since both sides solve $\dot{\Pi} = A(t)\Pi$ and coincide for $t = t_1$. In particular, choosing $t = t_0$, we see that $\Pi(t, t_0)$ is an isomorphism with inverse $\Pi(t, t_0)^{-1} = \Pi(t_0, t)$.

More generally, taking n solutions $\phi_1, \ldots, \phi_n$ we obtain a matrix solution $U(t) = (\phi_1(t), \ldots, \phi_n(t))$. The determinant of $U(t)$ is called Wronski determinant

$$W(t) = \det(\phi_1(t), \ldots, \phi_n(t)).$$

If $\det U(t) \neq 0$, the matrix solution $U(t)$ is called a **fundamental matrix solution**. Moreover, if $U(t)$ is a matrix solution, so is $U(t)C$, where C is a constant matrix. Hence, given two fundamental matrix solutions $U(t)$ and $V(t)$ we always have $V(t) = U(t)U(t_0)^{-1} V(t_0)$, since a matrix solution is uniquely determined by an initial condition. In particular, the principal matrix solution can be obtained from any fundamental matrix solution via $\Pi(t, t_0) = U(t)U(t_0)^{-1}$.

The following lemma shows that it suffices to check $\det U(t) \neq 0$ for one $t \in \mathbb{R}$.

Lemma 3.5.1. The Wronskian determinant of n solutions satisfies

$$W(t) = W(t_0) \exp\left(\int_{t_0}^{t} \operatorname{tr}(A(s)) ds\right). \tag{3.29}$$

This is known as **Abel's identity** or **Liouville's formula**.

Proof. By (3.27) we have

$$\Pi(t+\varepsilon, t)=\mathbb{I}+A(t)\varepsilon+o(\varepsilon)$$

and using $U(t+\varepsilon)=\Pi(t+\varepsilon, t)U(t)$ we obtain (exercise)

$$W(t+\varepsilon)=\det(\mathbb{I}+A(t)\varepsilon+o(\varepsilon))W(t)=(1+\operatorname{tr}(A(t))\varepsilon+o(\varepsilon))W(t)$$

implying

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

This equation is separable and the solution is given by (3.29). □

Now let us turn to the inhomogeneous system

$$\dot{x}=A(t)x+g(t), \quad x\left(t_0\right)=x_0, \tag{3.30}$$

where $A \in C\left(I, \mathbb{R}^n \times \mathbb{R}^n\right)$ and $g \in C\left(I, \mathbb{R}^n\right)$. Since the difference of two solutions of the inhomogeneous system (3.30) satisfies the corresponding homogeneous system (3.26), it suffices to find one particular solution. This can be done using the following ansatz

$$x(t)=\Pi\left(t, t_0\right) c(t), \quad c\left(t_0\right)=x_0,$$

which is known as **variation of constants** (also **variation of parameters**). Differentiating this ansatz we see

$$\dot{x}(t)=A(t)x(t)+\Pi\left(t, t_0\right)\dot{c}(t)$$

and comparison with (3.30) yields

$$\dot{c}(t)=\Pi\left(t_0, t\right) g(t).$$

Integrating this equation shows

$$c(t)=x_0+\int_{t_0}^t \Pi\left(t_0, s\right) g(s) ds$$

and we obtain (using (3.28))

Theorem 3.5.3. *The solution of the inhomogeneous system corresponding to the initial condition $x\left(t_0\right)=x_0$ is given by*

$$x(t)=\Pi\left(t, t_0\right) x_0+\int_{t_0}^t \Pi(t, s) g(s) ds,$$

where $\Pi\left(t, t_0\right)$ is the principal matrix solution of the corresponding homogeneous system.

To end this section, let me emphasize that there is no general way of solving linear systems except for the trivial case $n=1$. However, if one solution is known, one can reduce the order by one (see the following exercise).

Exercise* 3.5.1** (Reduction of order (d'Alembert)):*** *Suppose one solution $x_0(t)$ of the 2×2 system $\dot{x}=A(t)x$ is known and make the change of coordinates*

$$x(t)=X(t)y(t), \quad \text{where} \quad X(t)=\begin{pmatrix} x_{0,1}(t) & 0 \\ x_{0,2}(t) & 1 \end{pmatrix}.$$

Show that, if $x_{0,1}(t) \neq 0$, the differential equation for the new coordinates $y(t)$ reads

$$\dot{y}=X(t)^{-1}A(t)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} y.$$

In particular, the right hand side does not involve y_1. Hence this system can be solved by first solving the second component (which involves only y_2) and then integrating the first component. Generalize this result to $n \times n$ systems.

Exercise* 3.5.2** (Periodic linear system):*** *We may consider the case $A(t+T) = A(t)$ with period T in (3.26), which implies that $x(t+T)$ is again a solution if $x(t)$ is. Show the following lemma and theorem*

***Lemma* 3.5.2.** Suppose $A(t)$ is periodic with period T. Then the principal matrix solution satisfies

$$\Pi\left(t+T, t_0+T\right)=\Pi\left(t, t_0\right).$$

***Theorem* 3.5.4** (Floquet)**.** *Suppose $A(t)$ is periodic. Then the principal matrix solution of the corresponding linear system has the form*

$$\Pi\left(t, t_0\right)=P\left(t, t_0\right) \exp \left(\left(t-t_0\right) Q\left(t_0\right)\right),$$

where $P\left(., t_0\right)$ has the same period as $A(.) and P\left(t_0, t_0\right)=\mathbb{I}$.

(see [11] section 3.5 for periodic linear system).

Chapter 4

Boundary Value Problems

T chapter 5 and perhaps Richard Haberman Applied partial differential equations section 5.3-5.10. In essense, S and L consider a second order ode with bc that comes from a class of pdes due to some transformation.

Chapter 5

Dynamical Systems

Need to include [10] flow on line and other qualitative analyses.

Chapter 6

Chaos

Do as [10] did to incorporate T chapter 9 under chaos theory.

Part II

Complex Systems

Chapter 7

Differential Equations in Complex Domains

T chapter 4

Chapter 8

Riemann Surfaces

[5] chapter 1 and 2. [1] and [4].

Chapter 9

Appendix

9.1 Inverse and Implicit Function Theorem

Proposition 9.1.1. Suppose $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open subsets and $F : U \to V$ is a diffeomorphism. Then $m = n$, and for each $a \in U$, the total derivative $DF(a)$ is invertible, with $DF(a)^{-1} = D\left(F^{-1}\right)(F(a))$.

Proof. Because $F^{-1} \circ F = \mathrm{Id}_U$, the chain rule implies that for each $a \in U$,

$$\mathrm{Id}_{\mathbb{R}^n} = D\left(\mathrm{Id}_U\right)(a) = D\left(F^{-1} \circ F\right)(a) = D\left(F^{-1}\right)(F(a)) \circ DF(a).$$

Similarly, $F \circ F^{-1} = \mathrm{Id}_V$ implies that $DF(a) \circ D\left(F^{-1}\right)(F(a))$ is the identity on $\mathbb{R}^m$. This implies that $DF(a)$ is invertible with inverse $D\left(F^{-1}\right)(F(a))$, and therefore $m = n$. □

Next we study the relationship between total and partial derivatives. Suppose $U \subseteq \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^m$ is differentiable at $a \in U$. As a linear map between Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, $DF(a)$ can be identified with an $m \times n$ matrix. The next proposition identifies that matrix as the Jacobian of F.

Proposition 9.1.2. Let $U \subseteq \mathbb{R}^n$ be open, and suppose $F : U \to \mathbb{R}^m$ is differentiable at $a \in U$. Then all of the partial derivatives of F at a exist, and $DF(a)$ is the linear map whose matrix is the Jacobian of F at a :

$$DF(a) = \left(\frac{\partial F^j}{\partial x^i}(a)\right).$$

Proof. Let $B = DF(a)$, and for $v \in \mathbb{R}^n$ small enough that $a + v \in U$, let $R(v) = F(a+v) - F(a) - Bv$. The fact that F is differentiable at a implies that each component of the vector-valued function $R(v)/|v|$ goes to zero as $v \to 0$. The i th partial derivative of F^j at a, if it exists, is

$$\begin{aligned} \frac{\partial F^j}{\partial x^i}(a) &= \lim_{t \to 0} \frac{F^j\left(a + te_i\right) - F^j(a)}{t} = \lim_{t \to 0} \frac{B_i^j t + R^j\left(te_i\right)}{t} \\ &= B_i^j + \lim_{t \to 0} \frac{R^j\left(te_i\right)}{t}. \end{aligned}$$

The norm of the quotient on the right above is $\left|R^j\left(te_i\right)\right| / \left|te_i\right|$, which approaches zero as $t \to 0$. It follows that $\partial F^j / \partial x^i(a)$ exists and is equal to B_i^j as claimed. □

Theorem 9.1.1 (Inverse Function Theorem)**.** *Suppose U and V are open subsets of $\mathbb{R}^n$, and $F : U \to V$ is a smooth function. If $DF(a)$ is invertible at some point $a \in U$, then there exist connected neighborhoods $U_0 \subseteq U$ of a and $V_0 \subseteq V$ of $F(a)$ such that $\left.F\right|_{U_0} : U_0 \to V_0$ is a diffeomorphism.*

The proof of this theorem is based on an elementary result about metric spaces, which we describe first.

Let X be a metric space. A map $G : X \to X$ is said to be a **contraction** if there is a constant $\lambda \in (0,1)$ such that $d(G(x), G(y)) \leq \lambda d(x, y)$ for all $x, y \in X$. Clearly, every contraction is continuous. A **fixed point** of a map $G : X \to X$ is a point $x \in X$ such that $G(x) = x$.

Lemma 9.1.3 (Contraction Lemma)**.** Let X be a nonempty complete metric space. Every contraction $G : X \to X$ has a unique fixed point.

Proof. Uniqueness is immediate, for if x and x' are both fixed points of G, the contraction property implies $d(x, x') = d(G(x), G(x')) \leq \lambda d(x, x')$, which is possible only if $x = x'$.

To prove the existence of a fixed point, let x_0 be an arbitrary point in X, and define a sequence $(x_n)_{n=0}^{\infty}$ inductively by $x_{n+1} = G(x_n)$. For any $i \geq 1$ we have $d(x_i, x_{i+1}) = d(G(x_{i-1}), G(x_i)) \leq \lambda d(x_{i-1}, x_i)$, and therefore by induction

$$d(x_i, x_{i+1}) \leq \lambda^i d(x_0, x_1).$$

If N is a positive integer and $j \geq i \geq N$,

$$\begin{aligned} d(x_i, x_j) &\leq d(x_i, x_{i+1}) + d(x_{i+1}, x_{i+2}) + \cdots + d(x_{j-1}, x_j) \\ &\leq \left(\lambda^i + \cdots + \lambda^{j-1}\right) d(x_0, x_1) \\ &\leq \lambda^i \left(\sum_{n=0}^{\infty} \lambda^n\right) d(x_0, x_1) \\ &\leq \lambda^N \frac{1}{1-\lambda} d(x_0, x_1). \end{aligned}$$

Since this last expression can be made as small as desired by choosing N large, the sequence (x_n) is Cauchy and therefore converges to a limit $x \in X$. Because G is continuous,

$$G(x) = G\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} G(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

so x is the desired fixed point. □

Lemma 9.1.4 (Lipschitz Estimate for $\boldsymbol{C^1}$ Functions)**.** Let $U \subseteq \mathbb{R}^n$ be an open subset, and suppose $F : U \to \mathbb{R}^m$ is of class C^1. Then F is Lipschitz continuous on every compact convex subset $K \subseteq U$. The Lipschitz constant can be taken to be $\sup_{x \in K} |DF(x)|$ (Frobenius norm).

Theorem 9.1.1. *If $U \subseteq \mathbb{R}^n$ is an open subset and $F : U \to \mathbb{R}^m$ is of class C^1, then f is locally Lipschitz continuous.*

Proof. Each point of U is contained in a ball whose closure is contained in U, and Lemma 9.1.4 shows that the restriction of F to such a ball is Lipschitz continuous. □

Proof of the inverse function theorem. We begin by making some simple modifications to the function F to streamline the proof. First, the function F_1 defined by

$$F_1(x) = F(x + a) - F(a)$$

is smooth on a neighborhood of 0 and satisfies $F_1(0) = 0$ and $DF_1(0) = DF(a)$; clearly, F is a diffeomorphism on a connected neighborhood of a if and only if F_1 is a diffeomorphism on a connected neighborhood of 0 . Second, the function $F_2 = DF_1(0)^{-1} \circ F_1$ is smooth on the same neighborhood of 0 and satisfies $F_2(0) = 0$ and $DF_2(0) = I_n$; and if F_2 is a diffeomorphism in a neighborhood of 0 , then so is F_1 and therefore also F. Henceforth, replacing F by F_2, we assume that F is defined in a neighborhood U of $0, F(0) = 0$, and $DF(0) = I_n$. Because the determinant of $DF(x)$ is a continuous function of x, by shrinking U if necessary, we may assume that $DF(x)$ is invertible for each $x \in U$.

Let $H(x) = x - F(x)$ for $x \in U$. Then $DH(0) = I_n - I_n = 0$. Because the matrix entries of $DH(x)$ are continuous functions of x, there is a number $\delta > 0$ such that $B_\delta(0) \subseteq U$ and for all $x \in \bar{B}_\delta(0)$, we have $|DH(x)| \leq \frac{1}{2}$. If $x, x' \in \bar{B}_\delta(0)$, the Lipschitz estimate for smooth functions (Lemma 9.1.4) implies

$$|H(x') - H(x)| \leq \frac{1}{2} |x' - x|. \tag{9.1}$$

In particular, taking $x' = 0$, this implies

$$|H(x)| \leq \frac{1}{2}|x|. \tag{9.2}$$

Since $x' - x = F(x') - F(x) + H(x') - H(x)$, it follows that

$$|x' - x| \leq |F(x') - F(x)| + |H(x') - H(x)| \leq |F(x') - F(x)| + \frac{1}{2}|x' - x|,$$

and rearranging gives

$$|x' - x| \leq 2|F(x') - F(x)| \tag{9.3}$$

for all $x, x' \in \bar{B}_\delta(0)$. In particular, this shows that F is injective on $\bar{B}_\delta(0)$. Now let $y \in B_{\delta/2}(0)$ be arbitrary. We will show that there exists a unique point $x \in B_\delta(0)$ such that $F(x) = y$. Let $G(x) = y + H(x) = y + x - F(x)$, so that $G(x) = x$ if and only if $F(x) = y$. If $|x| \leq \delta$, (9.2) implies

$$|G(x)| \leq |y| + |H(x)| < \frac{\delta}{2} + \frac{1}{2}|x| \leq \delta, \tag{9.4}$$

so G maps $\bar{B}_\delta(0)$ to itself. It then follows from (9.1) that $|G(x) - G(x')| = |H(x) - H(x')| \leq \frac{1}{2}|x - x'|$, so G is a contraction. Since $\bar{B}_\delta(0)$ is a complete metric space, the contraction lemma implies that G has a unique fixed point $x \in \bar{B}_\delta(0)$. From (9.4), $|x| = |G(x)| < \delta$, so in fact $x \in B_\delta(0)$, thus proving the claim.

Let $V_0 = B_{\delta/2}(0)$ and $U_0 = B_\delta(0) \cap F^{-1}(V_0)$. Then U_0 is open in $\mathbb{R}^n$, and the argument above shows that $F : U_0 \to V_0$ is bijective, so $F^{-1} : V_0 \to U_0$ exists. Substituting $x = F^{-1}(y)$ and $x' = F^{-1}(y')$ into (9.3) shows that F^{-1} is continuous. Thus $F : U_0 \to V_0$ is a homeomorphism, and it follows that U_0 is connected because V_0 is.

The only thing that remains to be proved is that F^{-1} is smooth. If we knew it were smooth, Proposition 9.1.1 would imply that $D(F^{-1})(y) = DF(x)^{-1}$, where $x = F^{-1}(y)$. We begin by showing that F^{-1} is differentiable at each point of V_0, with total derivative given by this formula.

Let $y \in V_0$ be arbitrary, and set $x = F^{-1}(y)$ and $L = DF(x)$. We need to show that

$$\lim_{y' \to y} \frac{F^{-1}(y') - F^{-1}(y) - L^{-1}(y' - y)}{|y' - y|} = 0.$$

Given $y' \in V_0 \backslash \{y\}$, write $x' = F^{-1}(y') \in U_0 \backslash \{x\}$. Then

$$\begin{aligned}
&\frac{F^{-1}(y') - F^{-1}(y) - L^{-1}(y' - y)}{|y' - y|} \\
&= L^{-1}\left(\frac{L(x' - x) - (y' - y)}{|y' - y|}\right) \\
&= \frac{|x' - x|}{|y' - y|} L^{-1}\left(-\frac{F(x') - F(x) - L(x' - x)}{|x' - x|}\right).
\end{aligned}$$

The factor $|x' - x| / |y' - y|$ above is bounded thanks to (C.17), and because L^{-1} is linear and therefore bounded, the norm of the second factor is bounded by a constant multiple of

$$\frac{|F(x') - F(x) - L(x' - x)|}{|x' - x|}.$$

As $y' \to y$, it follows that $x' \to x$ by continuity of F^{-1}, and then (**??**) goes to zero because $L = DF(x)$ and F is differentiable. This completes the proof that F^{-1} is differentiable.

By Proposition 9.1.2, the partial derivatives of F^{-1} are defined at each point $y \in V_0$. Observe that the formula $D(F^{-1})(y) = DF(F^{-1}(y))^{-1}$ implies that the matrix-valued function $y \mapsto D(F^{-1})(y)$ can be written as the composition

$$y \xmapsto{F^{-1}} F^{-1}(x) \xmapsto{DF} DF(F^{-1}(y)) \xmapsto{i} DF(F^{-1}(y))^{-1}, \tag{9.5}$$

where i is matrix inversion. In this composition, F^{-1} is continuous; DF is smooth because its component functions are the partial derivatives of F; and i is smooth because Cramer's rule expresses the entries of an inverse matrix as rational functions of the entries of the matrix. Because $D(F^{-1})$ is a composition of continuous functions, it is continuous. Thus the partial derivatives of F^{-1} are continuous, so F^{-1} is of class C^1.

Now assume by induction that we have shown that F^{-1} is of class C^k. This means that each of the functions in (9.5) is of class C^k. Because $D\left(F^{-1}\right)$ is a composition of C^k functions, it is itself C^k; this implies that the partial derivatives of F^{-1} are of class C^k, so F^{-1} itself is of class C^{k+1}. Continuing by induction, we conclude that F^{-1} is smooth. □

Theorem 9.1.2. *Suppose $U \subseteq \mathbb{R}^n$ is an open subset, and $F : U \to \mathbb{R}^n$ is a smooth function whose Jacobian determinant is nonzero at every point in U.*

(a) F is an open map.

(b) If F is injective, then $F : U \to F(U)$ is a diffeomorphism.

Proof. For each $a \in U$, the fact that the Jacobian determinant of F is nonzero implies that $DF(a)$ is invertible, so the inverse function theorem implies that there exist open subsets $U_a \subseteq U$ containing a and $V_a \subseteq F(U)$ containing $F(a)$ such that F restricts to a diffeomorphism $\left.F\right|_{U_a} : U_a \to V_a$. In particular, this means that each point of $F(U)$ has a neighborhood contained in $F(U)$, so $F(U)$ is open. If $U_0 \subseteq U$ is an arbitrary open subset, the same argument with U replaced by U_0 shows that $F\left(U_0\right)$ is also open; this proves (a). If in addition F is injective, then the inverse map $F^{-1} : F(U) \to U$ exists for set-theoretic reasons; on a neighborhood of each point $F(a) \in F(U)$ it is equal to the inverse of $\left.F\right|_{U_a}$, so it is smooth. □

The next result is one of the most important consequences of the inverse function theorem. It gives conditions under which a level set of a smooth function is locally the graph of a smooth function.

Theorem 9.1.2 (Implicit Function Theorem)**.** *Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^k$ be an open subset, and let $(x, y) = \left(x^1, \ldots, x^n, y^1, \ldots, y^k\right)$ denote the standard coordinates on U. Suppose $\Phi : U \to \mathbb{R}^k$ is a smooth function, $(a, b) \in U$, and $c = \Phi(a, b)$. If the $k \times k$ matrix*

$$\left(\frac{\partial \Phi^i}{\partial y^j}(a, b)\right)$$

is nonsingular, then there exist neighborhoods $V_0 \subseteq \mathbb{R}^n$ of a and $W_0 \subseteq \mathbb{R}^k$ of b and a smooth function $F : V_0 \to W_0$ such that $\Phi^{-1}(c) \cap \left(V_0 \times W_0\right)$ is the graph of F, that is, $\Phi(x, y) = c$ for $(x, y) \in V_0 \times W_0$ if and only if $y = F(x)$.

Proof. Consider the smooth function $\Psi : U \to \mathbb{R}^n \times \mathbb{R}^k$ defined by $\Psi(x, y) = (x, \Phi(x, y))$. Its total derivative at (a, b) is

$$D\Psi(a, b) = \begin{pmatrix} I_n & 0 \\ \frac{\partial \Phi^i}{\partial x^j}(a, b) & \frac{\partial \Phi^i}{\partial y^j}(a, b) \end{pmatrix},$$

which is nonsingular because it is block lower triangular and the two blocks on the main diagonal are nonsingular. Thus by the inverse function theorem there exist connected neighborhoods U_0 of (a, b) and Y_0 of (a, c) such that $\Psi : U_0 \to Y_0$ is a diffeomorphism. Shrinking U_0 and Y_0 if necessary, we may assume that $U_0 = V \times W$ is a product neighborhood.

Writing $\Psi^{-1}(x, y) = (A(x, y), B(x, y))$ for some smooth functions A and B, we compute

$$\begin{aligned} (x, y) &= \Psi\left(\Psi^{-1}(x, y)\right) = \Psi(A(x, y), B(x, y)) \\ &= (A(x, y), \Phi(A(x, y), B(x, y))). \end{aligned} \tag{9.6}$$

Comparing the first components in this equation, we find that $A(x, y) = x$, so Ψ^{-1} has the form $\Psi^{-1}(x, y) = (x, B(x, y))$.

Now let $V_0 = \{x \in V : (x, c) \in Y_0\}$ and $W_0 = W$, and define $F : V_0 \to W_0$ by $F(x) = B(x, c)$. Comparing the second components in (9.6) yields

$$c = \Phi(x, B(x, c)) = \Phi(x, F(x))$$

whenever $x \in V_0$, so the graph of F is contained in $\Phi^{-1}(c)$. Conversely, suppose $(x, y) \in V_0 \times W_0$ and $\Phi(x, y) = c$. Then $\Psi(x, y) = (x, \Phi(x, y)) = (x, c)$, so

$$(x, y) = \Psi^{-1}(x, c) = (x, B(x, c)) = (x, F(x)),$$

which implies that $y = F(x)$. This completes the proof. □

9.2 Jordan Canonical Form

In this section we want to review some further facts on the Jordan canonical form. In addition, we want to draw some further consequences to be used later on.

Consider a decomposition $\mathbb{C}^n = V_1 \oplus V_2$. Such a decomposition is said to **reduce** A if both subspaces V_1 and V_2 are **invariant** under A, that is, $AV_j \subseteq V_j, j = 1, 2$. Changing to a new basis $u_1, \ldots, u_n$ such that $u_1, \ldots, u_m$ is a basis for V_1 and $u_{m+1}, \ldots, u_n$ is a basis for V_2, implies that A is transformed to the block form

$$U^{-1}AU = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

in these new coordinates. Moreover, we even have

$$U^{-1} \exp(A)U = \exp\left(U^{-1}AU\right) = \begin{pmatrix} \exp(A_1) & 0 \\ 0 & \exp(A_2) \end{pmatrix}.$$

Hence we need to find some invariant subspaces which reduce A. If we look at one-dimensional subspaces we must have

$$Ax = \alpha x, \quad x \neq 0, \tag{9.7}$$

for some $\alpha \in \mathbb{C}$. If (9.7) holds, α is called an **eigenvalue** of A and x is called **eigenvector**. In particular, α is an eigenvalue if and only if $\operatorname{Ker}(A - \alpha) \neq \{0\}$ and hence $\operatorname{Ker}(A - \alpha)$ is called the **eigenspac** of α in this case. Since $\operatorname{Ker}(A - \alpha) \neq \{0\}$ implies that $A - \alpha$ is not invertible, the eigenvalues are the zeros of the characteristic polynomial of A,

$$\chi_A(z) = \prod_{j=1}^{m} (z - \alpha_j)^{a_j} = \det(z\mathbb{I} - A),$$

where $\alpha_i \neq \alpha_j$. The number a_j is called **algebraic multiplicity** of α_j and $g_j = \dim \operatorname{Ker}(A - \alpha_j)$ is called **geometric multiplicity** of α_j. The set of all eigenvalues of A is called the **spectrum** of A,

$$\sigma(A) = \{\alpha \in \mathbb{C} \mid \operatorname{Ker}(A - \alpha) \neq \{0\}\}.$$

If the algebraic and geometric multiplicities of all eigenvalues happen to be the same, we can find a basis consisting only of eigenvectors and $U^{-1}AU$ is a diagonal matrix with the eigenvalues as diagonal entries. Moreover, $U^{-1} \exp(A)U$ is again diagonal with the exponentials of the eigenvalues as diagonal entries.

However, life is not that simple and we only have $g_j \leq a_j$ in general. It turns out that the right objects to look at are the **generalized eigenspaces**

$$V_j = \operatorname{Ker}(A - \alpha_j)^{a_j}.$$

Lemma 9.2.1. Let A be an n by n matrix and let $V_j = \operatorname{Ker}(A - \alpha_j)^{a_j}$. Then the V_j 's are invariant subspaces and $\mathbb{C}^n$ can be written as a direct sum

$$\mathbb{C}^n = V_1 \oplus \cdots \oplus V_m.$$

As a consequence we obtain

Theorem 9.2.1 (Cayley-Hamilton)**.** *Every matrix satisfies its own characteristic equation*

$$\chi_A(A) = 0.$$

So, if we choose a basis u_j of generalized eigenvectors, the matrix $U = (u_1, \dots, u_n)$ transforms A to a block structure

$$U^{-1}AU = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix},$$

where each matrix A_j has only the eigenvalue α_j. Hence it suffices to restrict our attention to this case.

A vector $u \in \mathbb{C}^n$ is called a **cyclic vector** for A if the vectors $A^j u$, $0 \le j \le n-1$ span $\mathbb{C}^n$, that is,

$$\mathbb{C}^n = \left\{ \sum_{j=0}^{n-1} a_j A^j u \mid a_j \in \mathbb{C} \right\}.$$

The case where A has only one eigenvalue and where there exists a cyclic vector u is quite simple. Take

$$U = \left(u, (A-\alpha)u, \dots, (A-\alpha)^{n-1}u\right),$$

then U transforms A to

$$J = U^{-1}AU = \begin{pmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \alpha & \ddots & \\ & & & \ddots & 1 \\ & & & & \alpha \end{pmatrix}, \tag{9.8}$$

since $\chi_A(A) = (A-\alpha)^n = 0$ by the Cayley-Hamilton theorem. The matrix (9.8) is called a **Jordan block**. It is of the form $\alpha\mathbb{I} + N$, where N is **nilpotent**, that is, $N^n = 0$.

Hence, we need to find a decomposition of the spaces V_j into a direct sum of spaces V_{jk}, each of which has a cyclic vector u_{jk}.

We again restrict our attention to the case where A has only one eigenvalue α and set

$$K_j = \operatorname{Ker}(A-\alpha)^j.$$

In the cyclic case we have $K_j = \oplus_{k=1}^j \operatorname{span}\left\{(A-\alpha)^{n-k}\right\}$. In the general case, using $K_j \subseteq K_{j+1}$, we can find L_k such that

$$K_j = \bigoplus_{k=1}^{j} L_k.$$

In the cyclic case $L_n = \operatorname{span}\{u\}$ and we would work our way down to L_1 by applying $A-\alpha$ recursively. Mimicking this, we set $M_n = L_n$ and since $(A-\alpha)L_{j+1} \subseteq L_j$ we have $L_{n-1} = (A-\alpha)L_n \oplus M_{n-1}$. Proceeding like this we can find M_l such that

$$L_k = \bigoplus_{l=k}^{n} (A-\alpha)^{n-l} M_l$$

Now choose a basis u_j for $M_1 \oplus \cdots \oplus M_n$, where each u_j lies in some M_l. Let V_j be the subspace generated by $(A-\alpha)^l u_j$. Then $V = V_1 \oplus \cdots \oplus V_m$ by construction of the sets M_k and each V_j has a cyclic vector u_j. In summary, we get

Theorem 9.2.2 (Jordan canonical form)**.** *Let A be an n by n matrix. Then there exists a basis for $\mathbb{C}^n$, such that A is of block form with each block as in (9.8).*

In addition, to the matrix exponential we will also need its inverse. That is, given a matrix A we want to find a matrix B such that

$$A = \exp(B).$$

In this case we will call $B = \ln(A)$ a **matrix logarithm** of A. Clearly, by (**??**) this can only work if $\det(A) \neq 0$. Hence suppose that $\det(A) \neq 0$. It is no restriction to assume that A is in Jordan canonical form and to consider the case of only one Jordan block, $A = \alpha\mathbb{I} + N$. Motivated by the power series for the logarithm,

$$\ln(1+x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} x^j, \quad |x| < 1,$$

we set

$$\begin{aligned} B &= \ln(\alpha)\mathbb{I} + \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j\alpha^j} N^j \\ &= \begin{pmatrix} \ln(\alpha) & \frac{1}{\alpha} & \frac{-1}{2\alpha^2} & \cdots & \frac{(-1)^n}{(n-1)\alpha^{n-1}} \\ & \ln(\alpha) & \frac{1}{\alpha} & \ddots & \vdots \\ & & \ln(\alpha) & \ddots & \frac{-1}{2\alpha^2} \\ & & & \ddots & \frac{1}{\alpha} \\ & & & & \ln(\alpha) \end{pmatrix} \end{aligned}$$

By construction we have $\exp(B) = A$. Note that B is not unique since different branches of $\ln(\alpha)$ will give different matrices. Moreover, it might be complex even if A is real. In fact, if A has a negative eigenvalue, then $\ln(\alpha) = \ln(|\alpha|) + \mathrm{i}\pi$ implies that $\ln(A)$ will be complex. We can avoid this situation by taking the logarithm of A^2.

Lemma 9.2.2. A matrix A has a logarithm if and only if $\det(A) \neq 0$. Moreover, if A is real and all real eigenvalues are positive, then there is a real logarithm. In particular, if A is real we can find a real logarithm for A^2.

Proof. Since the eigenvalues of A^2 are the squares of the eigenvalues of A (show this), it remains to show that B is real if all real eigenvalues are positive.

In these only the Jordan block corresponding to complex eigenvalues could cause problems. We consider the real Jordan canonical form (**??**) and note that for

$$R = \begin{pmatrix} \operatorname{Re}(\alpha) & \operatorname{Im}(\alpha) \\ -\operatorname{Im}(\alpha) & \operatorname{Re}(\alpha) \end{pmatrix} = r \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad \alpha = r\mathrm{e}^{\mathrm{i}\varphi},$$

the logarithm is given by

$$\ln(R) = \ln(r)\mathbb{I} + \begin{pmatrix} 0 & -\varphi \\ \varphi & 0 \end{pmatrix}.$$

Now write the real Jordan block $R\mathbb{I} + N$ as $R\left(\mathbb{I} + R^{-1}N\right)$. Then one can check that

$$\log(R\mathbb{I} + N) = \log(R)\mathbb{I} + \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j} R^{-j} N^j$$

is the required logarithm. □

Similarly, note that the resolvent $(A-z)^{-1}$ can also be easily computed in Jordan canonical form, since for a Jordan block we have

$$(J-z)^{-1} = \frac{1}{\alpha - z} \sum_{j=0}^{n-1} \frac{1}{(z-\alpha)^j} N^j.$$

In particular, note that the resolvent has a pole at each eigenvalue with the residue being the projector onto the corresponding generalized eigenspace. For later use we also introduce the subspaces

$$E^{\pm}(A) = \bigoplus_{|\alpha_j|^{\pm 1} < 1} \operatorname{Ker}\left(A - \alpha_j\right)^{a_j},$$

$$E^{0}(A) = \bigoplus_{|\alpha_j| = 1} \operatorname{Ker}\left(A - \alpha_j\right)^{a_j},$$

where α_j are the eigenvalues of A and a_j are the corresponding algebraic multiplicities. The subspaces $E^+(A), E^-(A), E^0(A)$ are called **contracting**, **expanding**, **unitary subspace** of A, respectively. The restriction of A to these subspaces is denoted by A_+, A_-, A_0, respectively.

Exercise* 9.2.1*: *Denote by* $r(A) = \max_j \{|\alpha_j|\}$ *the* ***spectral radius*** *of* A*. Show that for every* $\varepsilon > 0$ *there is a norm* $\|.\|_\varepsilon$ *such that*

$$\|A\|_\varepsilon = \sup_{x:\|x\|_\varepsilon=1} \|Ax\|_\varepsilon \leq r(A) + \varepsilon.$$

Hint: It suffices to prove the claim for a Jordan block $J = \alpha\mathbb{I} + N$ *(why?). Now choose a diagonal matrix* $Q = \operatorname{diag}(1, \varepsilon, \ldots, \varepsilon^n)$ *and observe* $Q^{-1}JQ = \alpha\mathbb{I} + \varepsilon N$*.*

Exercise* 9.2.2*: *Suppose* $A(\lambda)$ *is* C^k *and has no unitary subspace. Then the projectors* $P^\pm(A(\lambda))$ *onto the contracting, expanding subspace are* C^k*.*

Hint: Use the formulas

$$P^-(A(\lambda)) = \frac{1}{2\pi \mathrm{i}} \int_{|z|=1} \frac{dz}{z - A(\lambda)}, \quad P^+(A(\lambda)) = \mathbb{I} - P^-(A(\lambda)).$$

Bibliography

[BD] William E. Boyce and Richard C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, 11th edition, John Wiley & Sons Inc., New York, NY, 2017.

[1] De Faria, Edson and De Melo, Welington. *Mathematical Tools for One-Dimensional Dynamics*, Cambridge University Press, vol. 115, 2008.

[EP] C.H. Edwards and D.E. Penney, *Differential Equations and Boundary Value Problems: Computing and Modeling*, 5th edition, Pearson, 2014.

[2] Farkas, Hershel M. and Kra, Irwin. *Riemann Surfaces*, Springer, 1992.

[3] Forster, Otto. *Lectures on Riemann Surfaces*, Springer Science & Business Media, vol. 81, 2012.

[4] McMullen, Curt. *Complex Analysis on Riemann Surfaces* (Lecture Note)

[5] Milnor, John. *Dynamics in One Complex Variable: Introductory Lectures*, Springer Science & Business Media, 2013.

[6] Morosawa, Shunsuke. *Holomorphic Dynamics*, Cambridge University Press, vol. 66, 2000.

[7] Pesin, Yakov and Climenhaga, Vaughn. *Lectures on Fractal Geometry and Dynamical Systems*, Student Mathematical Library, vol. 52, 2009.

[8] Rudin, Walter. *Principles of Mathematical Analysis*, McGraw-Hill New York, vol. 3, 1964.

[9] Rudin. Walter. *Real and Complex Analysis, Third Edition*, McGraw-Hill, 1987.

[10] Strogatz, Steven H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, CRC Press, 2018.

[11] Teschl, Gerald. *Ordinary Differential Equations and Dynamical Systems*, American Mathematical Soc., vol. 140, 2012.