Version: 20250316

From Differential Equations to Dynamical Systems

Konstantinos Efstathiou

Konstantinos Efstathiou
Duke Kunshan University
Kunshan, Jiangsu Province, China
k.efstathiou@dukekunshan.edu.cn

Contents

Preface

These notes came out of the 7-week undergraduate course *Ordinary Differential Equations and Dynamical Systems* that I have taught several times at Duke Kunshan University. The audience of the course consists mostly, but not exclusively, of sophomore and junior year students majoring in Mathematics. The course prerequisites are Multivariable Calculus and Linear Algebra — these notes make heavy use of them – and there is no assumption of any prior knowledge of differential equations and their solution methods.

Even though the notes spend a considerable time discussing solution methods, it is important to understand that most differential equations cannot be analytically solved. Then why spend a lot of effort discussing solution methods for specific types of differential equations? First, understanding the behavior of the solutions in simple systems can help us build an intuition for the behavior of the solutions in more complex systems. Second, the simple systems for which we obtain solutions appear either as subsystems or as local approximations of complex systems and, therefore, allow us to understand parts of the complex dynamics.

Traditionally, the study of differential equations has emphasized solution methods. We present the most commonly used solution methods but there are several more that we do not cover such as the Laplace transform, power series, and Green's method. We refer to other textbooks on differential equations, such as [3], for a discussion of these methods. On the contrary, the study of dynamical systems has deëmphasized solution methods and correspondingly emphasize qualitative methods since, after all, as mentioned earlier most differential equations cannot be analytically solved.

These notes alternate between discussing solution methods and qualitative approaches, often combining the two. In Chapter 1 we focus on solution methods for first-order equations. We switch to a qualitative, dynamical systems, point of view in Chapter 2 where we discuss how to understand the dynamics without solving the equations. In Chapter 3 we consider solution methods for linear second-order equations with constant coefficients that commonly appear in physical problems. In Chapter 4 we consider linear planar dynamical systems — even though we obtain explicit solutions, the focus is on the qualitative properties of the solutions and the classification of different types of dynamics. In Chapter 5 we discuss the general theory concerning solutions of linear systems — the emphasis is back on solution methods. Finally, in Chapter 6 we consider planar nonlinear dynamical systems — the approach is mostly qualitative.

I hope you will enjoy reading these notes as much as I enjoyed writing them.

Kunshan — March 2025

Typographical Conventions

We summarize here some typographical conventions used in the text.

- Vectors and vector valued functions are denoted by symbols in bold: $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{f}$, $\boldsymbol{g}$ etc.
- Vectors will always be considered as column vectors. For typographical convenience they are often represented as $\boldsymbol{x} = \langle x_1, x_2, \ldots, x_n \rangle$.
- The Euclidean length (norm) of a vector $\boldsymbol{x} \in \mathbf{R}^n$ is denoted by $\|\boldsymbol{x}\|$, that is, $\|\boldsymbol{x}\|^2 = \boldsymbol{x}\cdot\boldsymbol{x}$ and $\|\boldsymbol{x}\| = (x_1^2 + \cdots + x_n^2)^{1/2}$.
- If $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ are vectors in $\mathbf{R}^n$ then $U = [\boldsymbol{u}_1 | \ldots | \boldsymbol{u}_m]$ is the $n \times m$ matrix — that is, n rows and m columns — whose j-th column is the column vector $\boldsymbol{u}_j$.
- The end of proofs is marked by ☑, the end of remarks by ❞, and the end of examples by 🔥.

Chapter 1

First-Order Differential Equations

1.1 Differential Equations and Initial Value Problems

A *first-order ordinary differential equation* is an equation involving an unknown function $y(t)$, its first derivative $y'(t)$, and the variable t. The unknown function $y(t)$ is called the *dependent variable*; it depends on the single variable t which is called the *independent variable*. We commonly write y and y' instead of $y(t)$ and $y'(t)$ when it is clear which is the independent variable.

An equation involving t, y, and y' has the general form

$$g(t, y, y') = 0. \tag{1.1}$$

Here, g is a real-valued expression depending on t, y, and y'. It is convenient, and in many cases — but not always — possible, to use Eq. (1.1) to express y' in terms of t and y, that is, write the equation in the *standard form*

$$y' = f(t, y), \tag{1.2}$$

where f is a real-valued expression depending on t and y.

Example 1.1. The equation

$$yy' - 2t^3 = 0, \tag{1.3}$$

is a first-order ordinary differential equation in the form of Eq. (1.1). Algebraically solving the previous equation for y' we find

$$y' = \frac{2t^3}{y},$$

which is a first-order ordinary differential equation in the standard form, Eq. (1.2). It is not difficult to check that the function $y(t) = t^2$ satisfies Eq. (1.3), since

$$y(t)y'(t) - 2t^3 = t^2 \times 2t - 2t^3 = 0.$$

However, it is not the only possible solution. We can similarly check that $y(t) = -t^2$ is also a solution, and so is any function of the form $y(t) = \pm(t^4 + c)^{1/2}$ where c is a real number. In Section 1.2.1 we discuss in detail how to obtain the solution of equations such as Eq. (1.3).

The equations we consider in this chapter are called *first-order* because the highest derivative of $y(t)$ appearing in these equations is the first derivative. They are called *ordinary* to distinguish them from *partial* differential equations which involve an unknown function $u(s, t, \dots)$ — depending on more than one variable — and also involve the partial derivatives $\partial u/\partial s$, $\partial u/\partial t$, etc., of the unknown function.

It is a natural instinct when given an equation to try and solve it, that is, find the unknown function $y(t)$. Even though we will spend some time discussing solution methods for special types of ordinary differential equations, it turns out that the vast majority of differential equations cannot be analytically solved. For this reason, even though in this chapter we focus on solution methods, in the next chapter we consider qualitative methods for understanding the behavior of the solutions, without solving the equations.

Example 1.2 (Malthusian growth model). A simple model describing the growth of the size P of a population[1] as a function of time t is the Malthusian model

$$P' = kP, \tag{1.4}$$

where the independent variable is time t and k is a constant positive real number. The model makes sense only if the population has sufficient resources (space, food) to grow unconstrained.

Again, it is not difficult to check that any function of the form $P(t) = Ae^{kt}$, where A is a real number, solves Eq. (1.4). However, if we want to use this general form of the solution to determine the size of a population at a specific time t we also need to have a value for A — this picks one solution $P(t)$ out of the infinitely many solutions of the form Ae^{kt} by fixing A. Suppose that at time $t = t_0$ the size of the population has a known value $P(t_0) = P_0$. The condition $P(t_0) = P_0$ is called *initial condition.* Using that $P(t) = Ae^{kt}$ we find

$$P_0 = P(t_0) = Ae^{kt_0}.$$

Therefore, we find $A = P_0 e^{-kt_0}$ and this means that the specific solution of Eq. (1.4) that satisfies the condition $P(t_0) = P_0$ is given by

$$P(t) = Ae^{kt} = P_0 e^{-kt_0} e^{kt} = P_0 e^{k(t-t_0)}.$$

The collection of solutions $P(t) = Ae^{kt}$ is called the *general solution* of Eq. (1.4). The problem of finding the solution of Eq. (1.4) which satisfies the given initial condition $P(t_0) = P_0$ is called an *initial value problem.*

We summarize the introduced terminology in the following definition.

Definition 1.3. The *general solution* of a first-order ordinary differential equation is the collection of all functions $y(t)$ which satisfy the given equation. An *initial value problem* (IVP) consists of a first-order ordinary differential equation together with an *initial condition* of the form $y(t_0) = y_0$. A *solution to an initial value problem* satisfies both the given first-order ordinary differential equation and the given initial condition that comprise the initial value problem.

[1] The population in this model can be the population of people on Earth, or the population of bacteria in a Petri dish, or the amount of remaining radioactive material in an experiment. The validity of the model depends on the specifics of the problem, but after we have the equation, its analysis and solution can proceed with barely any reference to the problem being solved and therein lies much of the power of mathematics. Of course, after obtaining a solution it is important to check that it makes sense in the context of the problem.

In the rest of this chapter we first consider solution methods for some common first-order ordinary differential equations. Then we show that under certain mild conditions solutions of initial value problems are unique — we also discuss examples where this is not the case. Finally, we briefly discuss numerical methods for obtaining approximate solutions to initial value problems and how to use Mathematica to obtain and plot analytic solutions in some cases where such solutions can be computed.

1.2 Solution Methods for First-Order Differential Equations

There are two main types of first-order differential equations that can be solved with standard methods. These are the separable equations, and the linear equations. In this section we describe the solution methods for these two types of equations.

1.2.1 Separable First-Order Differential Equations

The first solution method concerns a very special type of first-order differential equation, described in the following definition.

Definition 1.4. A first-order ordinary differential equation $y' = f(t, y)$ is *separable* if $f(t, y)$ can be written in the form $p(y)g(t)$.

The separable equation

$$y' = \frac{\mathrm{d}y}{\mathrm{d}t} = p(y)g(t) \tag{1.5}$$

can be solved through a standard, almost algorithmic, procedure which follows the steps below. This procedure is justified later in this section.

Solution method for separable equations

Step 1. In Eq. (1.5) "separate" the terms which involve only the independent variable t from the terms which involve only the dependent variable y. This gives the equation

$$\frac{\mathrm{d}y}{p(y)} = g(t)\,\mathrm{d}t.$$

Step 2. Integrate both sides of the last equation — the left-hand side with respect to y and the right-hand side with respect to t. We get

$$\int \frac{\mathrm{d}y}{p(y)} = \int g(t)\,\mathrm{d}t + c,$$

where c is a constant real number called the *integration constant*. In particular, if $H(y)$ is an anti-derivative of $1/p(y)$ and $G(t)$ is an anti-derivative of $g(t)$ we write

$$H(y) = G(t) + c. \tag{1.6}$$

Equation (1.6) is an *implicit solution* of Eq. (1.5), that is, a relation that involves y and t but in a form that may not allow us to directly express y as a function of t.

Step 3. If $H(y)$ in Eq. (1.6) is not very complicated, we may be able to solve the equation for y in terms of t. This gives an *explicit solution* of Eq. (1.5). This step may not be always feasible but if it is then it should be done.

Step 4. If we are additionally given an initial condition $y(t_0) = y_0$ we can use either the implicit or the explicit solution to determine the value of c. In particular, from the implicit solution in Eq. (1.6) we find $c = H(y_0) - G(t_0)$.

A few examples can help to clarify this procedure.

Example 1.5. Consider the differential equation $yy' = 2t^3$ from Example 1.1. Writing

$$y\frac{\mathrm{d}y}{\mathrm{d}t} = 2t^3,$$

we can separate variables to get

$$y\,\mathrm{d}y = 2t^3\,\mathrm{d}t.$$

Integrating the two sides we find

$$\tfrac{1}{2}y^2 = \tfrac{1}{2}t^4 + c,$$

which, by defining $c_1 = 2c$, can be simplified to

$$y^2 = t^4 + c_1.$$

This is the implicit solution. To obtain an explicit solution, we solve for y and we obtain

$$y = \pm(t^4 + c_1)^{1/2},\ \ c_1 \in \mathbf{R}.$$

Example 1.6 (Revisiting the Malthusian growth model). Consider the differential equation $y' = ky$ of the Malthusian model discussed in Example 1.2 with initial condition $y(1) = 5$. The differential equation $y' = ky$ is separable with $p(y) = y$ and $g(t) = k$. Notice that other choices for p and g are also possible — for example, $p(y) = ky$ and $g(t) = 1$ — but the final result remains the same.

Separating variables we get

$$\frac{\mathrm{d}y}{y} = k\,\mathrm{d}t,$$

while integration gives

$$\int \frac{\mathrm{d}y}{y} = \int k\ \mathrm{d}t + c.$$

Computing the two integrals we find the implicit solution

$$\log|y| = kt + c,$$

where c is an arbitrary real constant. To obtain an explicit solution we solve for y and we find

$$y = (\pm e^c)e^{kt} = Ae^{kt},$$

where we defined $A = \pm e^c$. Since c is an arbitrary real number, A can be any positive or negative real number except zero.

However, $y = 0$ also solves the original equation and is not included in the solutions of the form Ae^{kt}; it is missing because when we separate variables we divide by y. Therefore, the general solution of the equation $y' = ky$ is the collection of solutions $y = Ae^{kt}$, $A \neq 0$, together with the solution $y = 0$. However, by allowing A to be any real number, the collection of functions $y = Ae^{kt}$ also includes the solution $y = 0$. Therefore, the general solution to the equation $y' = ky$ is

$$y = Ae^{kt},\ A \in \mathbf{R}.$$

Finally, we consider the initial condition $y(1) = 5$. From the general solution we find $5 = Ae^{k}$. Therefore, $A = 5e^{-k}$ and thus the solution to the given initial value problem is given by

$$y = 5e^{k(t-1)}.$$

Remark 1.7. The equation $y' = ky$ is by far the most important first-order differential equation — it appears again and again in these notes and in the study of dynamical systems. It is important to remember the form of its general solution $y = Ae^{kt}$ as well as the solution $y = y_0 e^{k(t-t_0)}$ to the initial value problem with $y(t_0) = y_0$.

Example 1.8. Consider the differential equation[2]

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{x}{t}.$$

Separating variables and integrating the two sides we get

$$\int \frac{\mathrm{d}x}{x} = \int \frac{\mathrm{d}t}{t} + c,$$

that is,[3]

$$\log|x| = \log|t| + c.$$

Solving for x in terms of t we find

$$x = (\pm e^{c})\, t.$$

Since $x = 0$ is also a solution, replace $\pm e^{c}$ by $a \in \mathbf{R}$. Then the general solution can be written as

$$x = at,\ a \in \mathbf{R}.$$

Justification of Separation of Variables. Steps 1 and 2 in the separation of variables procedure involve the manipulation of differentials. Even though the procedure is convenient, it is not properly justified. We now rectify the situation by proving the validity of Eq. (1.6).

Recall that we want to find a solution $y(t)$ to the equation

$$y' = \frac{\mathrm{d}y}{\mathrm{d}t} = p(y)g(t)$$

which we can also write as

$$\frac{1}{p(y)} y' = g(t).$$

[2] Here we use x to represent the dependent variable, that is, the unknown function $x(t)$ that we want to determine.

[3] We denote the natural logarithm by log.

For a function $y(t)$ to be a solution of the last equation, it must satisfy

$$\frac{1}{p(y(t))}y'(t) = g(t).$$

Suppose that $G(t)$ is an anti-derivative of $g(t)$ so that $G'(t) = g(t)$ and that $H(y)$ is an anti-derivative of $1/p(y)$ so that $H'(y) = 1/p(y)$. A remark on notation is in order here: G and H are functions of a single variable and $'$ denotes derivation with respect to that variable. Therefore,

$$\frac{1}{p(y(t))}y'(t) = H'(y(t))y'(t) = (H \circ y)'(t),$$

where the last equality is a direct application of the Chain Rule. Then we have

$$(H \circ y)'(t) = G'(t).$$

Since the two derivatives are equal, the corresponding functions differ by a constant, that is,

$$(H \circ y)(t) = G(t) + c.$$

Therefore, any solution $y(t)$ satisfies the implicit equation

$$H(y(t)) = G(t) + c,$$

which is exactly Eq. (1.6).

1.2.2 Linear First-Order Differential Equations

We now consider another type of first-order differential equations that can be solved using a standard procedure.

Definition 1.9. A first-order ordinary differential equation is *linear* if it has the form

$$a_1(t)y' + a_0(t)y = b(t), \tag{1.7}$$

where $a_0(t)$, $a_1(t)$, and $b(t)$ are continuous in an interval $I \subseteq \mathbf{R}$, and $a_1(t) \neq 0$ for all $t \in I$.

Given a linear equation in the form of Eq. (1.7), we bring it into *standard form* by dividing with $a_1(t)$. Defining $P(t) = a_0(t)/a_1(t)$ and $Q(t) = b(t)/a_1(t)$, Eq. (1.7) can be written in the standard form

$$y' + P(t)y = Q(t). \tag{1.8}$$

Remark 1.10. If $P(t) = 0$, we can directly integrate the equation $y' = Q(t)$ to obtain the solution $y(t) = \int Q(t)\ \mathrm{d}t + c$. ❞

We discuss two methods for solving a linear first-order equation such as Eq. (1.8): the method of *integrating factor* and the method of *variation of parameters*.

Integrating Factor

The main idea behind the method of integrating factor is that the left-hand side of the equation $y' + P(t)y = Q(t)$ may be expressed as the time-derivative of a product of two functions if we multiply the whole equation by an appropriate "integrating factor" $\mu(t)$.

Multiplying by the — unknown for now — function $\mu(t)$ we find

$$\mu(t)y' + \mu(t)P(t)y = \mu(t)Q(t).$$

We observe that if we would have μ' instead of μP at the left-hand side of the last equation then the whole right-hand side would equal $(\mu y)'$. In particular, we note that

$$(\mu(t)y)' = \mu(t)y' + \mu'(t)y,$$

and, if we demand that $\mu' = \mu P$, then we find

$$(\mu(t)y)' = \mu(t)y' + \mu(t)P(t)y.$$

Then the differential equation becomes

$$(\mu(t)y)' = \mu(t)Q(t),$$

which after integration gives

$$\mu(t)y = \int \mu(t)Q(t)\,\mathrm{d}t + c.$$

Solving for y we find the solution

$$y(t) = \frac{1}{\mu(t)}\left[\int \mu(t)Q(t)\,\mathrm{d}t + c\right]. \tag{1.9}$$

The only thing that remains is to find an expression for the integrating factor $\mu(t)$. The equation $\mu' = \mu P$ is a differential equation that must be solved for μ. It is a separable equation of the form

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = P(t)\,\mu.$$

A solution of the last equation is given by

$$\mu(t) = e^{\int P(t)\,\mathrm{d}t}. \tag{1.10}$$

This is not the general solution of $\mu' = P\mu$, but we do not *need* the general solution. *Any* function $\mu(t)$ that satisfies $\mu' = P\mu$ can be used.

In summary, a linear equation of the form in Eq. (1.8) can be solved by applying Eq. (1.9) where the integrating factor $\mu(t)$ is given by Eq. (1.10).

Exercise 1.1. Use the method of separation of variables discussed in Section 1.2.1 to show that the equation $\mu' = P\mu$ has the general solution $\mu(t) = A\exp\big(\int P(t)\,\mathrm{d}t\big)$ with $A \in \mathbf{R}$. Then, verify that Eq. (1.9) produces the same solution $y(t)$ for any choice of $A \neq 0$. Therefore, we are justified to take $A = 1$, giving Eq. (1.10).

We now apply the method of integrating factor to several examples.[4]

Example 1.11. We solve the equation

$$\frac{1}{t}y' - \frac{2}{t^2}y = t\cos t,\ t > 0,$$

applying the method of integrating factor. First, we multiply both sides of the equation by t so that the equation comes into the standard form

$$y' - \frac{2}{t}y = t^2\cos t,\ t > 0.$$

Here, $P(t) = -2/t$ and $Q(t) = t^2\cos t$.

Multiplying with an integrating factor μ we find

$$\mu y' - \mu\frac{2}{t}y = \mu t^2\cos t,\ t > 0.$$

For the left-hand side to be the derivative of a product, that is, $(\mu y)' = \mu y' + \mu' y$, μ must satisfy the equation

$$\mu' = -\mu\frac{2}{t}.$$

Separating variables and integrating we find

$$\log|\mu| = -2\log|t| = -2\log t,$$

where we ignore the integration constant. Solving for μ we obtain the solutions $\mu = \pm t^{-2}$, and since we need only one solution we can choose the integrating factor

$$\mu = \frac{1}{t^2}.$$

Multiplying both sides of the equation by the integrating factor we obtain

$$\left(\frac{y}{t^2}\right)' = \frac{1}{t^2}y' - \frac{2}{t^3}y = \cos t.$$

Then

$$\frac{1}{t^2}y = \sin t + c,\ c \in \mathbf{R},$$

and solving for y we find the general solution

$$y(t) = t^2(\sin t + c),\ c \in \mathbf{R}.$$

Example 1.12. We solve the equation

$$n' + kn = Ae^{-\lambda t},$$

applying the method of integrating factor. Here, k, λ, and A are positive constants with $k \neq \lambda$, and $n(t)$ is the unknown function.

[4] It is better to reproduce all steps of the method as shown in the following examples, instead of directly applying Eq. (1.9). This is the difference between blindly following a recipe and understanding what you are doing. Only in the latter case you can transfer your knowledge to other problems.

The given equation is already in standard form. The integrating factor must satisfy $\mu' = k\mu$ for which a solution is given by $\mu = e^{kt}$.

Multiplying both sides of the differential equation by $\mu = e^{kt}$ we find

$$(e^{kt}n)' = e^{kt}n' + ke^{kt}n = Ae^{(k-\lambda)t}.$$

Integrating gives

$$e^{kt}n = \frac{A}{k-\lambda}e^{(k-\lambda)t} + c,\ c \in \mathbf{R}.$$

Finally, solving for n we find

$$n(t) = \frac{A}{k-\lambda}e^{-\lambda t} + ce^{-kt},\ c \in \mathbf{R}.$$

Remark 1.13. In the previous example, if $k = \lambda$, we find $(e^{kt}n)' = A$ with solution $n(t) = e^{-kt}(At + c)$.

Variation of Parameters

Consider again the standard form of a linear first-order differential equation in Eq. (1.8), that is,

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

If the equation did not contain the term $Q(t)$, i.e., if $Q(t) \equiv 0$, then the resulting *homogeneous* equation $y' + P(t)y = 0$ is separable. Its general solution is

$$y = Ae^{-\int P(t)\ \mathrm{d}t},\ A \in \mathbf{R}.$$

Incidentally, notice the expression above can also be written as $y = A/\mu$, where μ is the integrating factor in Eq. (1.10).

We return now to the full, *non-homogeneous* equation which includes $Q(t)$. The premise behind the method of variation of parameters is that $Q(t)$ introduces only a time-dependent variation of the constant (parameter) A, that is, the solution of the original equation has the form

$$y = A(t)e^{-\int P(t)\ \mathrm{d}t},\ A(t) \text{ is a real-valued function.}$$

Substituting the last expression into the left-hand side of the original equation we get

$$y' + P(t)y = \left[A'(t)e^{-\int P(t)\ \mathrm{d}t} - A(t)P(t)e^{-\int P(t)\ \mathrm{d}t}\right] + P(t)A(t)e^{-\int P(t)\ \mathrm{d}t} = A'(t)e^{-\int P(t)\ \mathrm{d}t},$$

where the terms in the square bracket equal y'. Therefore,

$$A'(t)e^{-\int P(t)\ \mathrm{d}t} = Q(t),$$

and we obtain

$$A'(t) = Q(t)e^{\int P(t)\ \mathrm{d}t}.$$

Integrating, we find

$$A(t) = \int Q(t)e^{\int P(t)\ \mathrm{d}t}\ \mathrm{d}t + c,\ c \in \mathbf{R}.$$

This means that the solution is given by

$$y(t) = e^{-\int P(t)\ \mathrm{d}t}\Big[\int Q(t)e^{\int P(t)\ \mathrm{d}t}\ \mathrm{d}t + c\Big],\ c \in \mathbf{R}. \tag{1.11}$$

Remark 1.14. Using that the integrating factor is $\mu(t) = \exp\big(\int P(t)\ \mathrm{d}t\big)$, Eq. (1.11) can be rewritten as

$$y(t) = \frac{1}{\mu(t)}\Big[\int Q(t)\mu(t)\ \mathrm{d}t + c\Big],\ c \in \mathbf{R},$$

which is exactly Eq. (1.9).

Example 1.15. We consider again the linear equation

$$y' - \frac{2}{t}y = t^2\cos t,\ t > 0,$$

from Example 1.11, which we consider directly in its standard form. The corresponding homogeneous equation is

$$y' - \frac{2}{t}y = 0,\ t > 0,$$

with general solution

$$y(t) = Ae^{\int 2/t\ \mathrm{d}t} = Ae^{2\log t} = At^2,\ A \in \mathbf{R}.$$

Then consider solutions of the form

$$y(t) = A(t)\,t^2.$$

Substituting into the original equation we find

$$A't^2 + 2At - \frac{2}{t}At^2 = t^2\cos t,$$

giving $A' = \cos t$, and thus $A(t) = \sin t + c$, $c \in \mathbf{R}$. Therefore, the general solution is

$$y(t) = t^2(\sin t + c),\ c \in \mathbf{R}.$$

Exercise 1.2. Use the method of variation of parameters to solve the differential equation in Example 1.12.

This finishes our discussion of first order equations for which there are standard solution methods. This does not mean however that these are the only first order equations that can be solved. In Section 1.2.3 we explore a few other types of first-order equations that can be solved using a change of the independent or the dependent variable.

1.2.3 Changes of Variables

Often, a first order differential equation can be transformed to an equation that can be solved, defining a new dependent variable or a new independent variable. We first discuss how to simplify an equation using such changes of variables, and then we use a change of the dependent variable to solve a type of first order equations called *Bernoulli equations*.

Simplifying an Equation

The logistic equation was introduced by Pierre Verhulst in 1845 to model the growth of the size of the population when the environment can support a population with size up to some value $M > 0$ called the *carrying capacity*. The logistic equation is given by

$$P' = kP\Big(1 - \frac{P}{M}\Big),\ P \geq 0. \tag{1.12}$$

We can first simplify the equation by introducing a change of the dependent variable which eliminates the parameter M. Define

$$p(t) = \frac{P(t)}{M},$$

that is, express the population size as a fraction of the carrying capacity. Then we obtain the equation

$$p' = \frac{P'}{M} = k\frac{P}{M}\Big(1 - \frac{P}{M}\Big) = kp(1-p).$$

We can simplify the last equation even further by introducing a change of the independent variable which eliminates k. Define a new "time" $\tau = kt$ and the function $q(\tau) = p(\tau/k) = p(t)$. Then we have[5]

$$q'(\tau) = p'(t)\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{1}{k}kp(t)(1-p(t)) = q(\tau)(1-q(\tau)),$$

that is, $q' = q(1-q)$. We can solve thus equation for $q(\tau)$ using separation of variables.

Exercise 1.3. Show that the solution to the initial value problem $q' = q(1-q)$ with $q(\tau_0) = q_0$ is

$$q(\tau) = \frac{q_0}{q_0 + e^{-(\tau-\tau_0)}(1-q_0)}.$$

Finally, we have to give an expression for $P(t)$. Assume that the initial condition for the original problem is $P(t_0) = P_0$. This corresponds to $p(t_0) = P_0/M$ and $q(kt_0) = P_0/M$. Therefore, $\tau_0 = kt_0$ and $q_0 = P_0/M$. Then we find

$$q(\tau) = \frac{P_0}{P_0 + e^{-(\tau-kt_0)}(M-P_0)},$$

and tracing our steps backwards we get

$$p(t) = q(kt) = \frac{P_0}{P_0 + e^{-k(t-t_0)}(M-P_0)},$$

and finally

$$P(t) = Mp(t) = \frac{MP_0}{P_0 + e^{-k(t-t_0)}(M-P_0)}.$$

[5]It is very common to see this written as

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = \frac{\mathrm{d}p}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = p(1-p),$$

but notice that the "p" in the first and third expressions is a different function than the "p" in the second expression.

Bernoulli Equations

A Bernoulli equation has the form

$$y' + P(x)y = Q(x)y^\alpha, \quad \alpha \in \mathbf{R}. \tag{1.13}$$

There are two special cases for α in which Eq. (1.13) can be solved using the methods we have already seen in this chapter.

(i) If $\alpha = 0$, then $y' + Py = Q$ is a linear equation.
(ii) If $\alpha = 1$, then Eq. (1.13) becomes the separable equation $y' + (P - Q)y = 0$.

To solve Eq. (1.13) in the remaining cases $\alpha \neq 0, 1$ we introduce a new dependent variable

$$u = y^s,$$

where $s \in \mathbf{R}$ will be chosen in such a way so that u satisfies a linear differential equation.

We have

$$u' = sy^{s-1}y' = sy^{s-1}(-Py + Qy^\alpha) = -sPy^s + sQy^{s-1+\alpha} = -sPu + sQy^{s-1+\alpha}.$$

To ensure that the obtained equation is linear, it is sufficient to demand that $s = 1 - \alpha$. Then the equation for u becomes the linear equation

$$\frac{du}{dx} + (1-\alpha)Pu = (1-\alpha)Q. \tag{1.14}$$

After solving the last equation for u, we can determine y from the equation $u = y^{1-\alpha}$.

Example 1.16. Consider the equation

$$y' - 5y = -\frac{5}{2}xy^3.$$

This is a Bernoulli equation with $\alpha = 3$ and to solve it we make the transformation $u = 1/y^2$. Then we have

$$u' = -\frac{2}{y^3}y' = -\frac{2}{y^3}\left(5y - \frac{5}{2}xy^3\right) = -\frac{10}{y^2} + 5x = -10u + 5x.$$

The obtained equation for u is linear and we can solve it using one of the standard methods. The solution is

$$u(x) = ce^{-10x} + \frac{1}{20}(10x - 1),$$

where $c \in \mathbf{R}$. Therefore, the solution to the original problem is

$$y = \pm\frac{1}{\sqrt{ce^{-10x} + \frac{1}{20}(10x-1)}}.$$

Given that $y = 0$ is also a solution to the original problem we must also include it separately to the general solution. ♠

1.3 Existence and Uniqueness of Solutions

In Section 1.2 we discussed several solution methods for first-order ordinary differential equations. However, given a specific initial value problem, how do we know that it has some solution? And if we find a solution, how do we know that there are no other solutions? It turns out that there are indeed cases where an initial value problem has no solutions or it has more than one solutions — we see such examples later in this section. However, the general case is well-behaved, in a way that is described by the following fundamental theorem.

Theorem 1.17 (Existence and Uniqueness Theorem). *Consider the initial value problem*

$$y' = f(t, y), \;\; y(t_0) = y_0, \tag{1.15}$$

and assume that there are intervals $[a, b] \subseteq \mathbf{R}$ *and* $[c, d] \subseteq \mathbf{R}$ *such that:*

(i) $f(t, y)$ *is continuous for all* (t, y) *with* $t \in [a, b]$ *and* $y \in [c, d]$*;*
(ii) $\partial f(t, y)/\partial y$ *is continuous for all* (t, y) *with* $t \in [a, b]$ *and* $y \in [c, d]$*;*
(iii) $t_0 \in (a, b)$ *and* $y_0 \in (c, d)$*.*

Then there is an interval $[a', b'] \subseteq [a, b]$ *such that the initial value problem in Eq.* (1.15) *has a* unique *solution* $y(t)$ *defined for* $t \in [a', b']$*.*

Example 1.18 (Unique solution). Consider the initial value problem $y' = ky$ with $y(t_0) = y_0$. The condition of Theorem 1.17 hold on $\mathbf{R}^2$. Therefore, there is an interval $[a', b']$ containing t_0 such that the given initial value problem has a unique solution in this interval. In this case, Theorem 1.17 does not give the whole picture — the function $y(t) = y_0 e^{k(t-t_0)}$ is the unique solution to the given initial value problem defined for all $t \in \mathbf{R}$.

Example 1.19. Consider the equation

$$ty' = 2y, \;\; t \in \mathbf{R}, \;\; y(0) = y_0.$$

The given initial value problem does not satisfy the conditions of Theorem 1.17, since

$$f(t, y) = \frac{2y}{t}$$

is not defined at $t = 0$. We show that when $y_0 \neq 0$ there are no solutions to the given initial value problem, and that when $y_0 = 0$ there are infinitely many solutions.

If the given initial value problem has a solution, then this must extend to t near $t_0 = 0$. Suppose that this is the case and that at $t_1 \neq 0$ the value of the solution is $y(t_1) = y_1$. One can easily check that Theorem 1.17 applies to the initial value problem with initial condition $y(t_1) = y_1$ for $t_1 > 0$ and therefore there is a unique solution which satisfies this initial condition.

To find this unique solution we can use separation of variables to solve the given equation. Then we obtain the general solution $y(t) = At^2$. For the given initial value problem we have $y_1 = At_1^2$ and thus the unique solution

$$y(t) = \left(\frac{y_1}{t_1^2}\right) t^2.$$

Changing $t_1 \neq 0$ and y_1 we can obtain infinitely many distinct solutions but we observe that all these solutions satisfy $y(0) = 0$ as shown in Fig. 1.1. Therefore, there are no solutions that satisfy the initial condition $y(0) = y_0 \neq 0$, and there are infinitely many solutions that satisfy the initial condition $y(0) = y_0 = 0$.

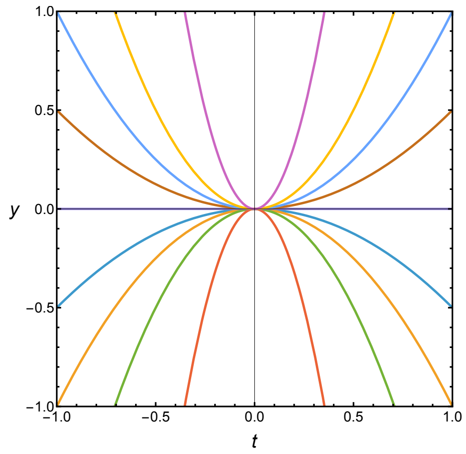


Figure 1.1: Graph of solutions of the form $y(t) = (y_1/t_1^2)t^2$. Notice that all solutions go through the origin.

Example 1.20. Consider the initial value problem

$$y' = 3(y^2)^{1/3}, \; y(1) = 0.$$

The function $f(t, y) = 3(y^2)^{1/3}$ is continuous for all $y \in \mathbf{R}$ but $\partial f/\partial y = \frac{2}{3}y^{-1/3}$ is not defined at $y = 0$. This means that the conditions of Theorem 1.17 are not satisfied for any choice of interval $[c, d]$ containing 0.

We can directly check that $y(t) = 0$ is a solution to the given initial value problem. It turns out however that there are also infinitely many other solutions. To find another solution, use the method of separation of variables to solve the given equation. We obtain the solution

$$y(t) = (t + c)^3, \; c \in \mathbf{R},$$

and taking into account the given initial condition we find $c = -1$, and thus another solution is

$$y(t) = (t - 1)^3.$$

To get infinitely many solutions we cut up in pieces the solution $(t + c)^3$ and then glue the pieces back together as follows. Choose any two numbers a, b such that $a \leq 1 \leq b$ and define the function

$$g_{a,b}(t) = \begin{cases} (t-a)^3, & t \leq a, \\ 0, & a \leq t \leq b, \\ (t-b)^3, & b \leq t, \end{cases} \tag{1.16}$$

see Fig. 1.2 for an example. One can easily check that each function $g_{a,b}$ is a solution of the equation $y' = 3(y^2)^{1/3}$ for $t \neq a, b$. It is also a solution at $t = a$ and $t = b$ where $g_{a,b}(t) = g'_{a,b}(t) = 0$. Moreover, it satisfies $g_{a,b}(1) = 0$. Therefore, each one of the infinitely many functions $g_{a,b}$ solves the given initial value problem.

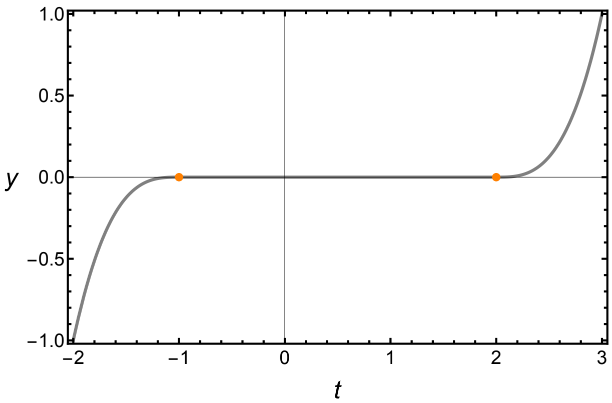


Figure 1.2: Graph of the function $g_{a,b}(t)$ in Eq. (1.16) with $a = -1$ and $b = 2$.

1.4 Numerical Approximation of Solutions

We close this chapter with a brief discussion of methods for the numerical approximation of the solution of an initial value problem such as

$$y' = f(t, y), \ y(t_0) = y_0. \tag{1.17}$$

To obtain the numerical approximation of the solution $y(t)$ we fix a small step size h and we define the sequence of times $t_k = t_0 + kh$ — notice that we can also have $h < 0$ meaning that we approximate the solution going backward in time. Denote by y_k the approximate value of the solution $y(t)$ at t_k. The methods we discuss produce an approximation for y_{k+1}, the approximate value of $y(t)$ at t_{k+1}, based on the value y_k. For more details on methods for the numerical approximation of solutions we refer to [3] or to books on numerical analysis, such as [4] or [2].

1.4.1 Euler Method

The simplest approximation is obtained by considering the solution of $y' = f(t, y)$ which satisfies $y(t_k) = y_k$. Taylor expanding $y(t)$ around $t = t_k$ up to first order terms and evaluating at $t = t_{k+1}$ gives

$$y(t_{k+1}) = y(t_k) + y'(t_k)(t_{k+1} - t_k) + \frac{1}{2}y''(\xi_k)h^2,$$

where $\xi_k \in [t_k, t_{k+1}]$. Using that $y(t_k) = y_k$, $y'(t_k) = f(t_k, y(t_k)) = f(t_k, y_k)$, and $t_{k+1} = t_k + h$, we obtain

$$y(t_{k+1}) = y_k + hf(t_k, y_k) + \frac{1}{2}y''(\xi_k)h^2.$$

If h is chosen sufficiently small, we can approximate $y(t_{k+1})$ by the value

$$y_{k+1} = y_k + hf(t_k, y_k), \tag{1.18}$$

that is, by ignoring the remainder term $\frac{1}{2}y''(\xi_k)h^2$ with size $O(h^2)$. Equation (1.18) defines the *Euler integration method.*

Applying the Euler method

To numerically approximate the solution of the initial value problem $y' = f(t, y)$, $y(t_0) = y_0$ in an interval $[t_0, t_0 + T]$ using the Euler method, follow these steps:

(i) Choose a step size h and an integer number N such that $T = Nh$.
(ii) Define the approximations $y_k \approx y(t_k)$, where $t_k = t_0 + kh$ with $k = 0, \dots, N$, by recursively applying Eq. (1.18).

The following function implements the Euler method in Mathematica. The function `euler` takes as input the function f (`f`), the initial time t_0, the initial state y_0 (`y0`), the step size h (`h`), and the number of steps N (`n`).

```
euler[f_ , t0_ , y0_ , h_ , n_] := Module[{step},
  step[{t_ , y_}] := {t + h, y + h f[t, y]};
  NestList[step, {t0, y0}, n]
  ]
```

The function `euler` can be used as follows to integrate the initial value problem $y' = y$, $y(0) = 1$ up to time $t = 1$ with step size $h = 10^{-2}$.

```
euSol = euler[Function[{t,y}, y], 0, 1, 0.01, 100];
```

The last element of `euSol` is `{1., 2.70481}`. The exact solution to the given problem is $y(t) = e^t$ and thus $y(1) = 2.71828\dots$. The absolute error in the solution is $\approx 1.3 \times 10^{-2}$.

1.4.2 Runge-Kutta Methods

To obtain better approximations than the Euler method we can consider higher-degree terms in the Taylor series giving $y(t_{k+1})$. This approach produces higher order, that is, more accurate approximation methods, called *Taylor methods*. The issue with such methods is that they involve the computation of higher-order derivatives $y^{(j)}(t_k)$ which can be cumbersome and computationally expensive.

An alternative is to combine the values of $f(t, y)$ at different points between t_k and t_{k+1} in such a way that the combination approximates $y(t_{k+1})$ with an accuracy of the same order as a Taylor method. Without going into the details of this construction, we directly give a Runge-Kutta method for which the remainder term is $O(h^5)$. This is the fourth-order Runge-Kutta method defined by the following iterative scheme:

$$\begin{aligned}
k_1 &= hf(t_k, y_k), \\
k_2 &= hf(t_k + \tfrac{1}{2}h, y_k + \tfrac{1}{2}k_1), \\
k_3 &= hf(t_k + \tfrac{1}{2}h, y_k + \tfrac{1}{2}k_2), \\
k_4 &= hf(t_k + h, y_k + k_3), \\
y_{k+1} &= \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).
\end{aligned} \tag{1.19}$$

Applying the Runge-Kutta method

To numerically approximate the solution of the initial value problem $y' = f(t, y)$, $y(t_0) = y_0$ in an interval $[t_0, t_0 + T]$ using the fourth-order Runge-Kutta method, follow these steps:

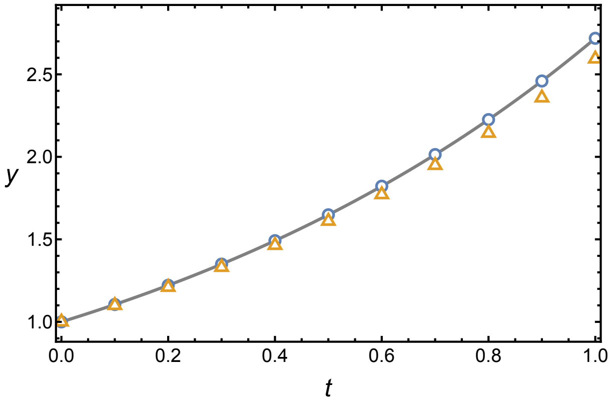


Figure 1.3: Comparison of numerically obtained solutions. The solid gray curve represents the solution computed using `NDSolveValue`, the orange triangles represent the values computed using `euler`, and the blue circles represent the values computed using `rk4`.

(i) Choose a step size h and an integer number N such that $T = Nh$.
(ii) Define the approximations $y_k \approx y(t_k)$, where $t_k = t_0 + kh$ with $k = 1, \ldots, N$, by recursively applying Eq. (1.19).

```
rk4[f_ , t0_ , y0_ , h_ , n_] := Module[{step},
  step[{t_ , y_}] := Module[{k1, k2, k3, k4},
    k1 = h f[t, y];
    k2 = h f[t + h/2, y + k1/2];
    k3 = h f[t + h/2, y + k2/2];
    k4 = h f[t + h, y + k3];
    {t + h, y + 1/6 (k1 + 2 k2 + 2 k3 + k4)}
  ];
  NestList[step, {t0, y0}, n]
  ]
```

```
rkSol = rk4[Function[{t,y}, y], 0, 1, 0.01, 100];
```

The absolute error in the solution is $\approx 2.2 \times 10^{-10}$. Choosing a different value of h will give another approximation with different error. For example, if $h = 10^{-1}$ then we find that the absolute error in the solution is $\approx 2 \times 10^{-6}$. The error remains reasonable for several practical purposes and the result can be obtained much faster since it takes only 10 time steps to go from $t_0 = 0$ to $t_0 + T = 1$. The code to obtain the solution with $h = 10^{-1}$ is the following.

```
rkSolLargeStep = rk4[Function[{t,y}, y], 0, 1, 0.1, 10];
```

1.4.3 Using Wolfram Mathematica

Mathematica can analytically solve a variety of differential equations. For example to solve the equation $y' = ky$ use

```
DSolveValue[y'[t] == k y[t], y[t], t]
```

The result is `E^(k t) C[1]`, representing the general solution $C_1 e^{kt}$ with C_1 a constant. To solve the initial value problem $y' = ky$, $y(t_0) = y_0$ use

```
DSolveValue[y'[t] == k y[t] && y[t0] == y0, y[t], t]
```

The result is `E^(k t - k t0) y0`, representing the solution $y(t) = y_0 e^{k(t-t_0)}$.

To numerically approximate the solution to an initial value problem in Mathematica we use the function `NDSolveValue`.

```
f[t_] = NDSolveValue[y'[t] == y[t] && y[0] == 1, y[t], {t, 0, 1}]
```

Notice that now we must specify the t interval in which we want to obtain an approximate solution and we can no longer have unspecified parameters such as k, t_0, y_0. When `NDSolveValue` is used like this, the function automatically chooses an integration method and a variable step size based on accuracy requirements. The result is an interpolating function that can be evaluated at all $t \in [0, 1]$. For example `f[1]` evaluates to 2.71828 and the absolute error is approximately 3.3×10^{-8}.

A comparison of the numerical solution obtained with `NDSolveValue` and the numerical approximations using the Euler method and the fourth-order Runge-Kutta method is shown in Fig. 1.3. The figure is produced using the following code.

```
euSolLargeStep = euler[Function[{t,y}, y], 0, 1, 0.1, 10];
rkSolLargeStep = rk4[Function[{t,y}, y], 0, 1, 0.1, 10];
f[t_] = NDSolveValue[y'[t] == y[t] && y[0] == 1, y[t], {t, 0, 1}];

Show[
  Plot[f[t], {t, 0, 1},
    PlotStyle -> Directive[AbsoluteThickness[3], Gray]],
  ListPlot[{rkSolLargeStep, euSolLargeStep},
    PlotMarkers -> {"OpenMarkers", 10}],
  PlotRange -> {{0, 1}, {0.9, 2.9}},
  PlotRangePadding -> Scaled[.01],
  AxesOrigin -> {0, 0},
  Frame -> True,
  FrameLabel -> {Style[t, FontSize -> 24], Style[y, FontSize -> 24]},
  RotateLabel -> False,
  FrameStyle -> Directive[Black, AbsoluteThickness[2], 18],
  ImageSize -> Large]
```

For more details on Mathematica we refer to the Wolfram Language & System Documentation Center [11].

Chapter 2

Dynamical Systems in One Dimension

In this chapter we mainly consider first-order differential equations of the form

$$x' = f(x), \tag{2.1}$$

where $x(t)$ denotes the dependent variable and the independent variable is t. Differential equations such as Eq. (2.1), where the right-hand side does not explicitly depend on the independent variable, are called *autonomous*. After discussing the properties of the solutions of autonomous equations, we give the abstract definition of a *dynamical system*, which reflects these properties. Then we introduce the concept of *phase line* and use it to qualitatively understand the solutions of Eq. (2.1) without analytically computing them. We particularly emphasize *equilibria* and their stability. Then we consider how the dynamics of the system change when the system parameters are varied and we discuss the concept of *bifurcations*. Finally, we consider one-dimensional *discrete-time dynamical systems* and some of their basic features, and show how one can obtain discrete-time dynamical systems from non-autonomous differential equations through *Poincaré maps*.

2.1 Autonomous First-Order Differential Equations

Consider the initial value problem for the autonomous equation given by Eq. (2.1) with initial condition $x(t_0) = x_0$. Since f in Eq. (2.1) does not explicitly depend on t, the Existence and Uniqueness Theorem (Theorem 1.17) applies if f and $f' = \mathrm{d}f/\mathrm{d}x$ are continuous for x in an interval $I = [c, d] \subseteq \mathbf{R}$ which contains x_0.

In this chapter we assume that f in Eq. (2.1) satisfies these conditions, that is, f is a continuously differentiable function in an interval I with $x_0 \in I$, and therefore there is a time interval $[a', b'] \subseteq \mathbf{R}$ such that the initial value problem given by Eq. (2.1) with initial condition $x(t_0) = x_0$ has a unique solution $x(t)$ defined for all $t \in [a', b']$.

Equation (2.1) is separable. If $G(x)$ is an anti-derivative of $1/f(x)$, then the method of separation of variables in Section 1.2.1 gives the implicit solution

$$G(x(t)) = t + c, \;\; c \in \mathbf{R}. \tag{2.2}$$

Provided that G is an invertible function,[1] we can express the solutions of Eq. (2.1) in the

[1] This holds if $f(x(t)) \neq 0$ along the solution $x(t)$, since then $f(x)$ and thus also $G'(x) = 1/f(x)$ is either strictly positive or strictly negative, and G is either strictly increasing or strictly decreasing.

explicit form

$$x(t) = G^{-1}(t+c),\ c \in \mathbf{R}, \tag{2.3}$$

where G^{-1} is the inverse function of G.

Given an initial condition $x(t_0) = x_0$, we have from Eq. (2.2) that $c = G(x_0) - t_0$, and therefore the explicit solution to the given initial value problem is

$$x(t) = G^{-1}(G(x_0) + (t - t_0)). \tag{2.4}$$

The right-hand side in the last equation depends on x_0 and on $t - t_0$. If we define a function $\phi(t, x)$ by

$$\phi(t, x) = G^{-1}(G(x) + t),$$

then the solution to the given initial value problem is written as

$$x(t) = \phi(t - t_0, x_0). \tag{2.5}$$

Example 2.1. Recall that the initial value problem $x' = kx$, $x(t_0) = x_0$, has solution $x(t) = x_0 e^{k(t-t_0)}$, defined for all $t \in \mathbf{R}$.

Example 2.2. The initial value problem $x' = x^2$, $x(t_0) = x_0$, has solution

$$x(t) = \frac{x_0}{1 - (t - t_0)x_0}.$$

Note that, differently from Example 2.1, the solution here is not defined for all $t \in \mathbf{R}$. For $x_0 > 0$ the solution is defined only for $t < t_c := t_0 + 1/x_0$, since $\lim_{t \to t_c^-} x(t) = \infty$. Note that in this example the function $f(x) = x^2$ is continuously differentiable for all $x \in \mathbf{R}$, and despite the fact that $f(x)$ does not depend explicitly on t, the (unique) solution is not defined for all $t \in \mathbf{R}$ since it diverges in finite time. In this case we talk about *finite time blowup* of the solution.

We now give an alternative proof of the fact that the solutions of the initial value problem in Eq. (2.1) have the form given in Eq. (2.5) based on general arguments that can also be used in higher dimensional systems such as those discussed in Chapters 4 and 6.

We introduce some notation. Consider the initial value problem $x' = f(x)$ with $x(t_0) = x_0$. The state of the system depends on the time t and it also depends on the values of t_0 and x_0 that determine the initial condition since, in general, if we change t_0 or x_0 we also change the obtained solution. For this reason, we denote the solution to the initial value problem $x' = f(x)$ with $x(t_0) = x_0$ by $h(t; t_0, x_0)$. The function h, being a solution to the given initial value problem, satisfies the differential equation

$$\frac{\partial h}{\partial t}(t; t_0, x_0) = f(h(t; t_0, x_0)). \tag{2.6}$$

Moreover, h satisfies the given initial condition, that is,

$$h(t_0; t_0, x_0) = x_0. \tag{2.7}$$

Remark 2.3. Why use a partial derivative $\partial/\partial t$ in Eq. (2.6) instead of the ordinary derivative $\mathrm{d}/\mathrm{d}t$? First, the ordinary derivative does not make sense in this case, since h depends on three variables. Second, to verify the validity of Eq. (2.6), fix t_0, x_0 and consider the solution

$x(t)$ satisfying the initial value problem $x' = f(x)$, $x(t_0) = x_0$. That is, $x(t) = h(t; t_0, x_0)$, and we have

$$\frac{\partial h}{\partial t}(t; t_0, x_0) = \lim_{\varepsilon \to 0} \frac{h(t+\varepsilon; t_0, x_0) - h(t; t_0, x_0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{x(t+\varepsilon) - x(t)}{\varepsilon} = x'(t).$$

Since $x'(t) = f(x(t)) = f(h(t; t_0, x_0))$, we obtain Eq. (2.6).

Proposition 2.4. *Let $h(t; t_0, x_0)$ be the unique solution to the initial value problem $x' = f(x)$, $x(t_0) = x_0$. Then there is a function ϕ such that $h(t; t_0, x_0) = \phi(t - t_0, x_0)$. Moreover, ϕ has the following properties:*

(i) $\phi(0, x_0) = x_0$*;*
(ii) $\partial\phi/\partial t \; (t, x_0) = f(\phi(t, x_0))$*.*

Proof. We show that for any s we have

$$h(t+s; t_0+s, x_0) = h(t; t_0, x_0). \tag{2.8}$$

Recall that $h(t; t_0, x_0)$ is, by definition, the solution to the initial value problem $x' = f(x)$, $x(t_0) = x_0$. We show that $y(t) = h(t+s, t_0+s, x_0)$ solves the same initial value problem and then the uniqueness of solutions implies that the two solutions coincide. For $t = t_0$, we have

$$y(t_0) = h(t_0+s; t_0+s, x_0) = x_0,$$

since, by definition, $h(t; t_0+s, x_0)$ solves the initial value problem $x' = f(x)$, $x(t_0+s) = x_0$. Moreover,

$$y'(t) = \frac{\partial h}{\partial t}(t+s, t_0+s, x_0) = f(h(t+s, t_0+s, x_0)) = f(y(t)).$$

We conclude that $y(t) = h(t+s, t_0+s, x_0)$ solves the same initial value problem as $h(t; t_0, x_0)$ and thus Eq. (2.8) holds.

Using $s = -t_0$ in Eq. (2.8) we find $h(t; t_0, x_0) = h(t - t_0; 0, x_0)$. Define

$$\phi(t, x_0) = h(t; 0, x_0),$$

that is, $\phi(t, x_0)$ is the solution to the initial value problem $x' = f(x)$, $x(0) = x_0$. Then,

$$h(t; t_0, x_0) = h(t - t_0; 0, x_0) = \phi(t - t_0, x_0).$$

Using $t = t_0$ in the last equation we also obtain

$$\phi(0, x_0) = h(0; 0, x_0) = x_0.$$

Finally, Eq. (2.6) implies that

$$\frac{\partial \phi}{\partial t}(t, x_0) = \frac{\partial h}{\partial t}(t; 0, x_0) = f(h(t; 0, x_0)) = f(\phi(t, x_0)).$$

Corollary 2.5. *If the initial value problem $x' = f(x)$, $x(0) = x_0$ has solution $\phi(t, x_0)$ then the initial value problem $x' = f(x)$, $x(t_0) = x_0$, has solution $\phi(t - t_0, x_0)$.*

The significance of Proposition 2.4 is that for autonomous equations it is sufficient to consider only initial value problems of the form $x(0) = x_0$. If we then have the initial value problem $x(t_0) = x_0$, its solution can be simply obtained by shifting time by t_0, as shown in Fig. 2.1.

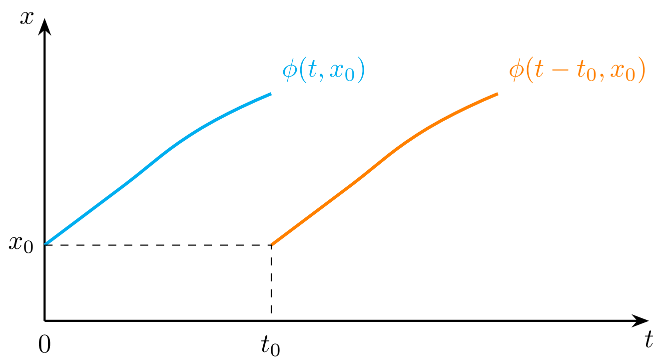


Figure 2.1: For autonomous equations $x' = f(x)$ the solution to the initial value problem $x(t_0) = x_0$ is a translation of the solution to the initial value problem $x(0) = x_0$.

Proposition 2.6. *If the initial value problem $x' = f(x)$, $x(0) = x_0$ has solution $\phi(t, x_0)$ then*

$$\phi(t, \phi(s, x_0)) = \phi(t + s, x_0). \tag{2.9}$$

Proof. Fix s and consider the functions $x(t) = \phi(t + s, x_0)$ and $y(t) = \phi(t, \phi(s, x_0))$, where only t changes. We show that these functions solve the same initial value problem $x' = f(x)$, $x(0) = \phi(s, x_0)$ and therefore they are identical. For $t = 0$ we have

$$y(0) = \phi(0, \phi(s, x_0)) = \phi(s, x_0) = x(0),$$

where for the second equality we used that $\phi(0, x_0) = x_0$ for all x_0, see Proposition 2.4. That is, $x(t)$ and $y(t)$ satisfy the same initial condition. Moreover,

$$x'(t) = \frac{\mathrm{d}}{\mathrm{d}t}[\phi(s + t, x_0)] = \frac{\partial \phi}{\partial t}(s + t, x_0) = f(\phi(s + t, x_0)) = f(x(t)),$$

and

$$y'(t) = \frac{\mathrm{d}}{\mathrm{d}t}[\phi(t, \phi(s, x_0))] = \frac{\partial \phi}{\partial t}(t, \phi(s, x_0)) = f(\phi(t, \phi(s, x_0)) = f(y(t)).$$

Therefore, the functions $x(t) = \phi(t + s, x_0)$ and $y(t) = \phi(t, \phi(s, x_0))$ solve the same initial value problem and from uniqueness of solutions we conclude that they are equal, that is, Eq. (2.9) holds. ☑

Remark 2.7. The discussion in this section can be easily generalized to *autonomous systems of differential equations* $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ in $\mathbf{R}^n$. Analogues of Propositions 2.4 and 2.6 hold in that case and exactly the same argument can be used only changing the one-dimensional x to the n-dimensional vector $\boldsymbol{x}$. ”

Despite the fact that we can always solve an autonomous initial value problem in dimension one,[2] it turns out that it is more productive in terms of understanding of the behavior of solutions to adopt a more qualitative and general approach. We do this in subsequent sections. Our starting point in the next section will be the following summary of properties of ϕ.

[2] Up to solving the implicit relation in Eq. (2.2) for $x(t)$.

Summary of properties of $\phi(t,x)$

The function $\phi(t,x)$ satisfies the following properties:

(i) $\phi(t,x_0)$ is the solution to the initial value problem $x' = f(x)$, $x(0) = x_0$.
(ii) $\phi(0,x_0) = x_0$.
(iii) $\phi(t,\phi(s,x_0)) = \phi(t+s,x_0)$.

Finally, we introduce the notion of orbits of the equation $x' = f(x)$.

Definition 2.8. Given $x_0 \in \mathbf{R}$ the *orbit* through x_0 is defined by

$$\mathcal{O}(x_0) = \{\phi(t,x_0) : t \in \mathbf{R}\}.$$

Proposition 2.9. *If $x_2 \in \mathcal{O}(x_1)$ then $x_1 \in \mathcal{O}(x_2)$ and $\mathcal{O}(x_2) = \mathcal{O}(x_1)$.*

Proof. If $x_2 \in \mathcal{O}(x_1)$, there is t_2 such that $x_2 = \phi(t_2,x_1)$. This implies

$$\phi(-t_2,x_2) = \phi(-t_2,\phi(t_2,x_1)) = \phi(-t_2+t_2,x_1) = \phi(0,x_1) = x_1.$$

Therefore, $x_1 \in \mathcal{O}(x_2)$.

We then show that $\mathcal{O}(x_1) \subseteq \mathcal{O}(x_2)$ and $\mathcal{O}(x_2) \subseteq \mathcal{O}(x_1)$. This will imply that $\mathcal{O}(x_1) = \mathcal{O}(x_2)$. First, consider any $x \in \mathcal{O}(x_1)$. Then there is t such that $x = \phi(t,x_1)$, and we have

$$x = \phi(t,\phi(-t_2,x_2)) = \phi(t-t_2,x_2),$$

that is, $x \in \mathcal{O}(x_1)$. This implies $\mathcal{O}(x_1) \subseteq \mathcal{O}(x_2)$. Finally, consider any $x \in \mathcal{O}(x_2)$. Then there is t such that $x = \phi(t,x_2)$, and we have

$$x = \phi(t,\phi(t_2,x_1)) = \phi(t+t_2,x_1),$$

that is, $x \in \mathcal{O}(x_2)$. This implies $\mathcal{O}(x_2) \subseteq \mathcal{O}(x_1)$ and concludes the proof. ☑

Proposition 2.9 implies that for any given points x_1, x_2 the orbits $\mathcal{O}(x_1)$, $\mathcal{O}(x_2)$ are either identical, that is, $\mathcal{O}(x_1) = \mathcal{O}(x_2)$, or they are disjoint, that is, $\mathcal{O}(x_1) \cap \mathcal{O}(x_2) = \emptyset$.

Remark 2.10. The notions of *equivalence relation* and *equivalence class* are important in mathematics. We will not directly use these notions here, but it is worth briefly discussing them and explaining their relation to orbits.

Given a set A, an *equivalence relation* in A is a binary relation $\sim$ which is reflexive ($a \sim a$ for all $a \in A$), symmetric (if $a \sim b$ for $a, b \in A$, then $b \sim a$), and transitive (if $a \sim b$ and $b \sim c$ for $a, b, c \in A$, then $a \sim b$). We read $a \sim b$ as "a is equivalent to b" Given an equivalence relation $\sim$ in A, the *equivalence class* of $a \in A$, is the set $[a]$ containing all the elements of A which are equivalent to a, that is, $[a] = \{b \in A : b \sim a\}$. It is not difficult to show that equivalent classes split A into disjoint subsets.

How is this related to orbits? Define a relation $\sim$ in $\mathbf{R}$ by $x_1 \sim x_2$ if $x_2 \in \mathcal{O}(x_1)$. This is an equivalence relation. It is clearly reflexive since $x_1 \in \mathcal{O}(x_1)$. It is symmetric because of Proposition 2.9, since $x_2 \in \mathcal{O}(x_1)$ implies $x_1 \in \mathcal{O}(x_2)$. And, it is transitive: if $x_2 \in \mathcal{O}(x_1)$ and $x_3 \in \mathcal{O}(x_2)$, implies $x_3 \in \mathcal{O}(x_1)$, since $\mathcal{O}(x_1) = \mathcal{O}(x_2)$. The equivalence class $[x_1]$ is the orbit $\mathcal{O}(x_1)$. Therefore, orbits split $\mathbf{R}$ into disjoint subsets. ❞

2.2 Dynamical Systems and Flows

We now slightly change our point of view, and define the notion of a continuous-time dynamical system in one dimension. Then, we discuss the relation between such dynamical systems and autonomous first-order differential equations.

We consider the space $\mathbf{R} \times \mathbf{R}$ with coordinates (t, x). The variable t represents time, while the variable x represents the (evolving) state of a system. For example, x could represent the position of a particle moving in one dimension, the population of bacteria on a Petri dish, or the average temperature in a room.

Definition 2.11. A *smooth dynamical system* is a continuously differentiable function $\phi : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ that satisfies

(i) $\phi(0, x_0) = x_0$;
(ii) $\phi(s + t, x_0) = \phi(t, \phi(s, x_0))$ for all $s, t \in \mathbf{R}$.

The function $\phi(t, x_0)$ is called the *flow* of the system — if a system is at the state x_0 at time 0, then $\phi(t, x_0)$ represents the state of the system after time t.

We also use the notation

$$\phi^t(x_0) = \phi(t, x_0),$$

where for each $t \in \mathbf{R}$, ϕ^t is a function $\phi^t : \mathbf{R} \to \mathbf{R}$. With this notation, the two properties of smooth dynamical systems in Definition 2.11 can be written as

$$\phi^0 = \text{id and } \phi^{s+t} = \phi^s \circ \phi^t. \tag{2.10}$$

The notation is reminiscent of the properties of exponentials — this is not accidental, as can be seen in Example 2.1. Each function ϕ^t is invertible with $(\phi^t)^{-1} = \phi^{-t}$ since

$$\phi^t \circ \phi^{-t} = \phi^{-t} \circ \phi^t = \phi^{-t+t} = \phi^0 = \text{id}.$$

Definition 2.11 states what is a dynamical system but it does not specify how to construct one. However, the discussion in Section 2.1 already establishes a close connection between dynamical systems in one dimension and autonomous first-order differential equation. In particular, Propositions 2.4 and 2.6 show that the function $\phi(t, x_0)$ which solves the initial value problem $x' = f(x)$, $x(0) = x_0$ defines a flow — provided that it is defined for all $t \in \mathbf{R}$. Because of the last condition, we restrict our attention now to the following type of differential equations.

Definition 2.12. The autonomous equation $x' = f(x)$ is *complete* if its solutions $x(t)$ are defined for all $t \in \mathbf{R}$.

Example 2.13. The equation $x' = kx$ is complete, while the equation $x' = x^2$ is not complete, see Examples 2.1 and 2.2.

We then have the following result.

Proposition 2.14. *Consider a complete autonomous differential equation $x' = f(x)$ and let $\phi(t, x_0)$ be the solution to the initial value problem $x' = f(x)$, $x(0) = x_0$. Then ϕ defines a smooth dynamical system.*

Proof. The fact that ϕ satisfies properties (i) and (ii) of Definition 2.11 is a direct consequence of Propositions 2.4 and 2.6. For the proof of the smooth dependence of ϕ on x_0 we refer to [1]. ☑

Proposition 2.14 shows that complete autonomous first-order differential equations define dynamical systems. The opposite is also true.

Proposition 2.15. *Given a smooth dynamical system ϕ, define*

$$f(x) = \frac{\partial \phi}{\partial t}(0, x).$$

Then $\phi(t, x_0)$ is the solution to the initial value problem $x' = f(x)$, $x(0) = x_0$.

Proof. Clearly, $\phi(0, x_0) = x_0$ so the initial condition is satisfied. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}[\phi(t, x_0)] = \frac{\partial \phi}{\partial t}(t, x_0),$$

and

$$\begin{aligned} f(\phi(t, x_0)) &= \frac{\partial \phi}{\partial t}(0, \phi(t, x_0)) \\ &= \lim_{\varepsilon \to 0} \frac{\phi(\varepsilon, \phi(t, x_0)) - \phi(0, \phi(t, x_0))}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{\phi(t + \varepsilon, x_0) - \phi(t, x_0)}{\varepsilon} \\ &= \frac{\partial \phi}{\partial t}(t, x_0). \end{aligned}$$

Therefore, $\phi(t, x_0)$ satisfies the equation $x' = f(x)$. ☑

Example 2.16. Consider again the initial value problem $x' = kx$, $x(0) = x_0$, with solution $x(t) = x_0 e^{kt}$. The corresponding flow is defined as $\phi(t, x_0) = x_0 e^{kt}$ and it can be checked that it satisfies Definition 2.11. Conversely, given the flow $\phi(t, x_0) = x_0 e^{kt}$ we can obtain the differential equation that produces ϕ by defining

$$f(x) = \frac{\partial \phi}{\partial t}(0, x) = \frac{\partial}{\partial t}\Big|_{t=0} (x e^{kt}) = (k x e^{kt})|_{t=0} = kx.$$

Example 2.17. Consider the initial value problem $x' = x^2$, $x(0) = x_0$. Recall from Example 2.2 that in this case the solution cannot be defined for all $t \in \mathbf{R}$. However, we can still define a function

$$\phi(t, x_0) = \frac{x_0}{1 - t x_0},$$

being cognizant of the fact that it does not satisfy all the properties of a smooth dynamical system. However, it clearly satisfies $\phi(0, x_0) = x_0$. Moreover,

$$\phi(t, \phi(s, x_0)) = \frac{\phi(s, x_0)}{1 - t\phi(s, x_0)} = \frac{x_0}{1 - s x_0 - t x_0} = \phi(t + s, \phi(s, x_0)).$$

If we are given ϕ, then we obtain

$$f(x) = \frac{\partial \phi}{\partial t}(0, x) = \frac{\partial}{\partial t}\Big|_{t=0} \left(\frac{x}{1 - tx}\right) = \left(\frac{x}{1 - tx}\right)^2\Big|_{t=0} = x^2.$$

2.3 Phase Line

We now develop a general methodology for qualitatively understanding the dynamics generated by autonomous first-order differential equations $x' = f(x)$, without solving the equations. The starting point here is the notion of equilibrium, which is the simplest type of solution that can exist in such a system. Equilibria are important because they act as organizing centers for the rest of the dynamics.

Definition 2.18. A point $x_e \in \mathbf{R}$ such that $f(x_e) = 0$ is called an *equilibrium* of the equation $x' = f(x)$.

An equilibrium is a solution to $x' = f(x)$ in the sense that the function $x(t) = x_e$ is the unique solution to $x' = f(x)$ that satisfies the initial condition $x(0) = x_e$. To check that $x(t) = x_e$ is indeed a solution, note that we have $x'(t) = (x_e)' = 0$ and $f(x(t)) = f(x_e) = 0$. Moreover, note that the equilibrium solution $x(t) = x_e$ is defined for all $t \in \mathbf{R}$. In terms of the flow ϕ we write $\phi(t, x_e) = x_e$ for all $t \in \mathbf{R}$. This implies $\mathcal{O}(x_e) = \{x_e\}$.

Example 2.19. Below are some examples of equations $x' = f(x)$ and their equilibria.

(i) The equation $x' = kx$ has the single equilibrium $x_e = 0$.
(ii) The equation $x' = kx(1 - x)$ has equilibria $x_e = 0$ and $x_e = 1$.
(iii) The equation $x' = x - x^3$ has equilibria $x_e = 0, \pm 1$.
(iv) The equation $x' = x^2$ has a single equilibrium $x_e = 0$.
(v) The equation $x' = x^2 + 1$ has no equilibria.
(vi) All integers $k \in \mathbf{Z}$ are equilibria of the equation $x' = \sin(\pi x)$.

In the following discussion we will assume that all the equilibria of $x' = f(x)$ are isolated, that is, for each equilibrium x_e there is an open interval U containing x_e and no other equilibria besides x_e.

Remark 2.20. This assumption excludes, for example, functions $f(x)$ which are zero on an interval. Moreover, it excludes functions $f(x)$ such as

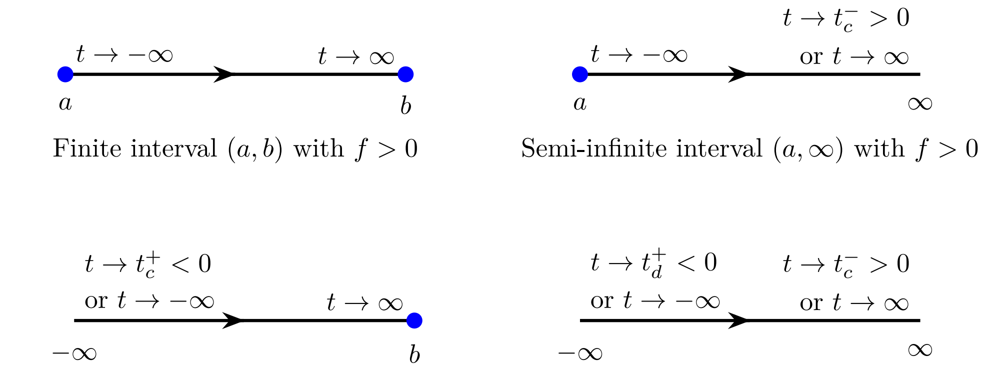
$$f(x) = \begin{cases} x \sin(\pi / x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

In this case, the equation $x' = f(x)$ has the equilibrium $x_0 = 0$ and infinitely many equilibria of the form $x_n = 1/n$, $n \in \mathbf{Z} \setminus \{0\}$. The equilibrium $x_0 = 0$ is not isolated since any open interval containing 0, must also contain infinitely many numbers of the form $1/n$.

Since equilibria are isolated, the real axis can be decomposed into open intervals I of the form $(-\infty, b)$, (a, b), (a, ∞), where the numbers a and b are equilibria of $x' = f(x)$ and each open interval I contains no equilibria. Since f is continuous, it must have constant sign on each open interval I, that is, either $f(x) > 0$ for all $x \in I$ or $f(x) < 0$ for all $x \in I$.

We now describe in detail what types of dynamics we can have in each such type of interval. To simplify the discussion we consider only the case with $f(x) > 0$ — similar conclusions can be drawn for $f(x) < 0$ by reversing the direction of time. The four possible types of intervals are shown in Fig. 2.2.

First, consider a finite interval (a, b) and let $x_0 \in (a, b)$. Assume that $f(x_0) > 0$ which implies that $f(x) > 0$ for all $x \in (a, b)$, while $f(a) = f(b) = 0$. Therefore for the solution $x(t)$ to the initial value problem $x' = f(x)$, $x(0) = x_0$, we have $x'(t) > 0$ as long as the



Finite interval (a, b) with $f > 0$ — Semi-infinite interval (a, ∞) with $f > 0$

Semi-infinite interval $(-\infty, b)$ with $f > 0$ — Infinite interval $(-\infty, \infty)$ with $f > 0$

Figure 2.2: Dynamics in an interval of type (a, b), (a, ∞), $(-\infty, b)$, or $(-\infty, \infty)$. We have assumed $f > 0$ and the right-going arrows indicate that the solution $x(t)$ is an increasing function.

solution remains in (a, b). This means that the solution $x(t)$ is a strictly increasing function of t. Moreover, the solution $x(t)$ will reach any $x_1 \in I$ in finite time which is given by

$$t_1 = \int_{x_0}^{x_1} \frac{\mathrm{d}s}{f(s)}.$$

However, the solution $x(t)$ cannot reach a in finite time since this would imply that $a \in \mathcal{O}(x_0)$ and thus $\mathcal{O}(x_0) = \mathcal{O}(a) = \{a\}$. For the same reason $x(t)$ cannot reach b in finite time. We have $\lim_{t\to\infty} x(t) = b$, $\lim_{t\to-\infty} x(t) = a$, and $\mathcal{O}(x_0) = (a, b)$.

Second, consider an interval (a, ∞) and let $x_0 \in (a, \infty)$. Assume $f(x_0) > 0$ so that $f(x) > 0$ for all $x \in (a, \infty)$. This implies again that $x(t)$ is an increasing function and that for any $x_1 \in (a, \infty)$ there is a finite time t_1 such that $x(t_1) = x_1$. Similar to the previous case we have $\lim_{t\to-\infty} x(t) = a$. However, going forward in time, it might occur that the solution is not defined for all $t > 0$ as the example of $x' = x^2$ shows. In particular, we distinguish two cases. In the first case, the solution $x(t)$ is defined for all $t \in \mathbf{R}$. In this case we have $\lim_{t\to\infty} x(t) = \infty$. In the second case, there is $t_c > 0$ such that $x(t)$ is defined only for $t < t_c$ and we have *finite time blowup* with $\lim_{t\to t_c^-} x(t) = \infty$, recall Example 2.2. The previous discussion shows that $\{x(t) : t \in T\} = (a, \infty)$, where T is either $\mathbf{R}$ or $(-\infty, t_c)$. This suggests to relax the definition of an orbit as

$$\mathcal{O}(x_0) = \{x(t) : \ t \in T\},$$

where T is the largest connected interval that contains 0 and such that the solution $x(t)$ of the initial value problem $x' = f(x)$, $x(0) = x_0$ is defined for all $t \in T$.

Exercise 2.1. Explain what is shown in Fig. 2.2 for the other two types of intervals $(-\infty, b)$ and $(-\infty, \infty)$. Then, draw the analogue of Fig. 2.2 for the case $f < 0$.

To represent the dynamics of $x' = f(x)$ it is convenient to make use of the *phase line*, that is, the real line together with information about the dynamics, including equilibria and the direction of motion along solutions.

To draw the phase line, we first mark on the real axis the equilibria of $x' = f(x)$. For each interval between two equilibria we check the sign of $f(x)$. If $f(x) > 0$, then $x(t)$ increases

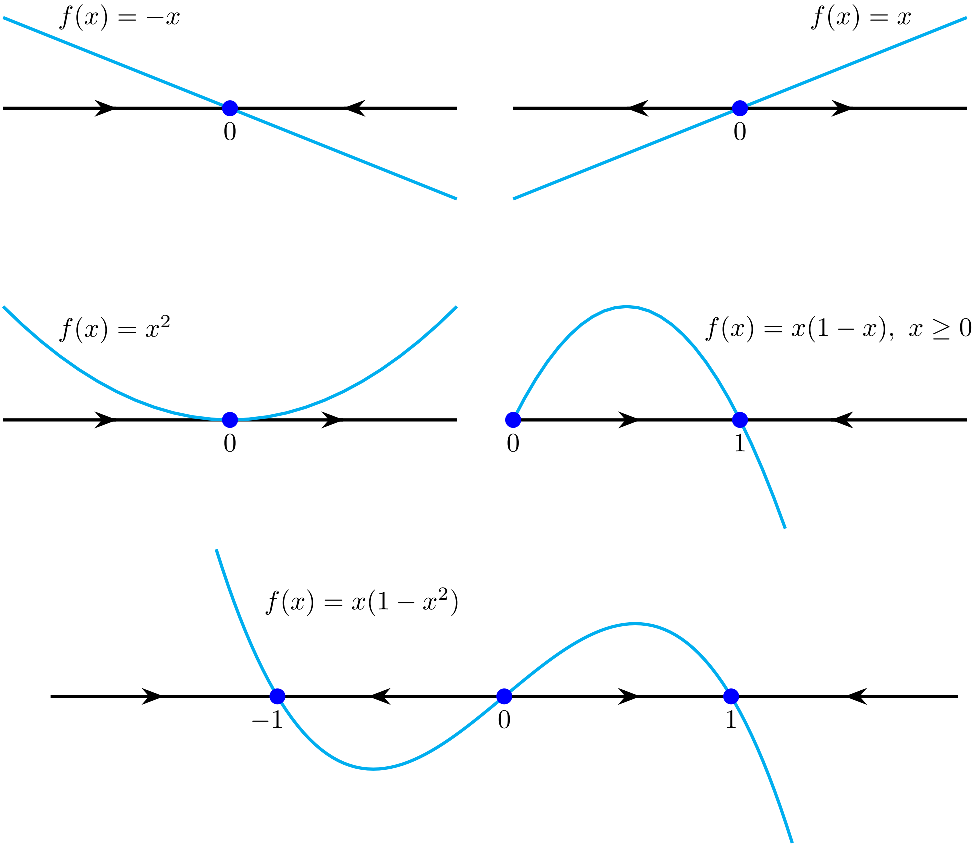


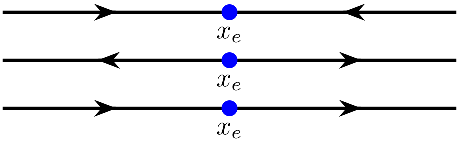
Figure 2.3: Phase lines for autonomous equations $x' = f(x)$ with different choices of $f(x)$. Equilibria are marked by blue points on the phase line, the direction of motion is indicated by arrows, and the graph of the function $f(x)$ is represented by the cyan line.

with time, and we draw a right pointing arrow in that interval. If $f(x) < 0$, then $x(t)$ decreases with time, and we draw a left pointing arrow.

Example 2.21. In Fig. 2.3, we have drawn the phase lines for different systems $x' = f(x)$. After finding and marking the equilibria of each system on the corresponding phase line, we indicate the direction of motion in each interval using a right-pointing arrow when $f(x) > 0$ and a left-pointing arrow when $f(x) < 0$. Notice the decomposition of each of the phase linee into orbits — equilibria and intervals — in Fig. 2.3.

2.4 Stability of Equilibria and Linearization

The discussion in the previous section shows that if x_e is an isolated equilibrium then the local dynamics is one of the following three types, depending on whether solutions in adjacent intervals move toward the equilibrium or away from it.



In the first of these cases we say that the equilibrium is *stable*. In the second and third cases we say that the equilibrium is *unstable*.

If $f'(x_e) \neq 0$ then the behavior near the equilibrium x_e is determined by the sign of $f'(x_e)$. First, suppose that $f'(x_e) > 0$. Then for $x > x_e$ (but close enough to x_e) we have that $f(x) > 0$ while for $x < x_e$ we have that $f(x) < 0$. In this case, which corresponds locally to $f(x) \approx \lambda(x - x_e)$ with $\lambda > 0$, the equilibrium is unstable, see also Fig. 2.3 for $f(x) = x$. If $f'(x_e) < 0$ then similar considerations show that the equilibrium is stable. This case corresponds locally to $f(x) \approx \lambda(x - x_e)$ with $\lambda < 0$, see Fig. 2.3 for $f(x) = -x$.

If $f'(x_e) = 0$ we can still determine the stability of the equilibrium by looking at the sign of $f(x_e)$ at the left and right of x_e. Suppose for example that $f(x) = x^2$. Then $x_e = 0$ is an equilibrium with $f'(x_e) = 0$ and $f(x) > 0$ for all $x \neq 0$. However, for $f(x) = x^3$, where $x_e = 0$ is again an equilibrium with $f'(x_e) = 0$, we have a behavior that is reminiscent of the case $f'(x_e) > 0$ and the equilibrium is unstable. For $f(x) = -x^3$ we find in a similar way that the equilibrium is asymptotically stable.

Exercise 2.2. Draw the phase lines for $x' = \pm x^3$ and confirm the claims in the previous paragraph.

2.4.1 Linearization

In the case where $f'(x_e) \neq 0$ we can also understand the local dynamical behavior near x_e through *linearization*. Let $y = x - x_e$ denote the relative position with respect to the equilibrium x_e. Then

$$y' = x' = f(x) = f(x_e + y).$$

Taylor expand at $y = 0$, and use that $f(x_e) = 0$ to get

$$y' = f'(x_e)y + \frac{1}{2}f''(x_e)y^2 + O(y^3).$$

In the linear approximation we may assume that y is small so that terms y^2 and higher can be neglected. This gives the differential equation

$$y' = f'(x_e)y = \lambda y,$$

where $\lambda = f'(x_e)$. This equation has the solution

$$y(t) = y(0)\exp(\lambda t),$$

which is an approximate solution to $y' = f(x_e + y)$ for y small enough.

The solution of the linearized equation shows that if $\lambda > 0$ then $|y|$ increases and therefore the equilibrium is unstable. If $\lambda < 0$ then $y \to 0$ as $t \to \infty$ and therefore the equilibrium is asymptotically stable. Note that the solutions we have obtained do not only tell us the stability of the equilibrium but also how fast the solutions approach the equilibrium or move away from it.

Remark 2.22. An equilibrium x_e where $f'(x_e) \neq 0$ is called *hyperbolic*. Hyperbolic equilibria have the property that a small change of f does not change the local dynamics near x_e. This property plays an important role in bifurcations, discussed in Section 2.5. ❞

Example 2.23. Consider the equation $x' = x - x^3$. The equilibria are $0, \pm 1$. We have

$$f'(x) = 1 - 3x^2.$$

Therefore, $f'(0) = 1 > 0$, implying that 0 is an unstable equilibrium. Moreover, $f'(\pm 1) = -2 < 0$, implying that ± 1 are stable equilibria. These stability results can also be checked through the phase line shown in Fig. 2.3 .

2.5 Bifurcations

Equilibria provide the organizing center for all the dynamics of the system. In systems that depend on parameters, the number and stability type of equilibria may change as the parameters vary. Then we say that a *bifurcation* takes place. In this section we consider equations of the form

$$x' = f(a, x), \tag{2.11}$$

where $x(t)$ is the dependent variable and $a \in \mathbf{R}$ is a parameter which does not change as the system evolves. We want to understand how the dynamics of the system changes if we choose different values for the parameter a.

Understanding under what conditions a bifurcation takes place and what is the dynamical behavior of the system before and after the bifurcation is one of the cornerstones of dynamical systems theory and research. In the same sense that equilibria are the organizing centers for all the dynamics of a system, bifurcations are the organizing centers for the dynamics of systems that depend on parameters and they are best studied in a space that is the product of the phase line and the parameter space. For more information on bifurcations we refer to [6, 8, 9].

One of the main tools in the study of bifurcations is the implicit function theorem which is typically discussed in advanced courses on multivariable calculus or calculus on manifolds [5, 10]. To state the theorem, consider the space $\mathbf{R}^n \times \mathbf{R}^m$ with coordinates $\boldsymbol{x} \in \mathbf{R}^n$, $\boldsymbol{y} \in \mathbf{R}^m$.

Given a function $\boldsymbol{g} : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$ we denote by $D_{\boldsymbol{y}}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ the derivative with respect to $\boldsymbol{y}$ of the function $\boldsymbol{g}$ at the point $(\boldsymbol{x}, \boldsymbol{y})$ which can be represented by the matrix

$$D_{\boldsymbol{y}}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \begin{bmatrix} \partial g_1/\partial y_1 & \cdots & \partial g_1/\partial y_m \\ \vdots & \ddots & \vdots \\ \partial g_m/\partial y_1 & \cdots & \partial g_m/\partial y_m \end{bmatrix}.$$

Then the following holds.

Theorem 2.24 (Implicit Function Theorem). *Consider a smooth function* $\boldsymbol{g} : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$ *such that* $\boldsymbol{g}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \mathbf{0}$. *If* $\det D_{\boldsymbol{y}}\boldsymbol{g}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$ *then there is* $\delta > 0$ *and a unique smooth function* $\boldsymbol{h} : B_\delta(\boldsymbol{x}_0) \to \mathbf{R}^m$ *such that* $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{h}(\boldsymbol{x})) = 0$ *and* $\boldsymbol{h}(\boldsymbol{x}_0) = \boldsymbol{y}_0$.

One way to view the implicit function theorem is that it gives conditions under which the system of m equations $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ can be solved to express the m quantities $\boldsymbol{y} = \langle y_1, \dots, y_m \rangle$ in terms of $\boldsymbol{x}$, near a point $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ which already satisfies $\boldsymbol{g}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \mathbf{0}$.

We now restate the implicit function theorem for the easiest case, of one equation involving two quantities x, y, where we want to express one quantity in terms of the other one.

Theorem 2.25 (Implicit Function Theorem for functions $\mathbf{R}^2 \to \mathbf{R}$). *Consider a smooth function* $g : \mathbf{R}^2 \to \mathbf{R}$ *such that* $g(x_0, y_0) = 0$. *If* $g_y(x_0, y_0) \neq 0$ *then there is* $\delta > 0$ *and a unique smooth function* $h(x)$ *defined in* $(x_0 - \delta, x_0 + \delta)$ *such that* $g(x, h(x)) = 0$ *and* $h(x_0) = y_0$.

Remark 2.26. In the statement of Theorem 2.25 and in what follows we denote partial derivatives by a subscript, that is, for a function $g(a, x)$ we write $g_a = \partial g/\partial a$, $g_x = \partial g/\partial x$, $g_{xx} = \partial^2 g/\partial x^2$ and so on. ❞

Remark 2.27. The way to read Theorem 2.25 is that if $g(x, y) = 0$ represents an equation involving x and y and there is x_0 for which we can find the corresponding y_0 that satisfies the equation, i.e., $g(x_0, y_0) = 0$, then we can extend this solution to an interval around x_0 and express the solution y for each x as $y = h(x)$ with $h(x_0) = y_0$. The condition for this to work is that $g_y(x_0, y_0) \neq 0$. Notice that the theorem talks about local (i.e., around x_0) existence of a solution. Furthermore, notice that we can restate the theorem for expressing x in terms of y. In particular, we can solve for $x = j(y)$ with $x_0 = j(y_0)$ provided that $g_x(x_0, y_0) \neq 0$. ❞

Example 2.28. Consider the equation $g(x, y) = x^2 + y^2 - 1 = 0$ representing the unit circle on the (x, y) plane. We check that the point $(0, 1)$ satisfies $g(0, 1) = 0$ and $g_y(0, 1) = 2 \neq 0$. Therefore, there is a function $y = h(x)$ defined near $x = 0$ which satisfies $x^2 + h(x)^2 - 1 = 0$ and $h(0) = 1$. In this case we can compute that $h(x) = \sqrt{1 - x^2}$. We should however be aware that in general it is not possible to find an analytic expression for the function h but only to assert its existence.

If we try to do the same at the point $(1, 0)$ which also satisfies $g(1, 0) = 0$ we find that $g_y(1, 0) = 0$ and therefore we cannot use the Implicit Function Theorem to assert the existence of a function $h(x)$ that has the required properties and, indeed, such a function does not exist. However, we can also check that $g_x(1, 0) = 2 \neq 0$ and therefore there is a unique function $x = j(y)$ such that $j(y)^2 + y^2 - 1 = 0$ and $j(0) = 1$. In this case we can compute that $j(y) = \sqrt{1 - y^2}$.

2.5.1 Persistence of Equilibria

Before discussing bifurcations we first consider when bifurcations do not take place. The first result is that equilibria of the system given by Eq. (2.11) persist — in general — when the parameter a changes. We have the following theorem.

Theorem 2.29 (Persistence of equilibria). *Assume that $x' = f(a,x)$ has an equilibrium x_0 for $a = a_0$, that is, $f(a_0,x_0) = 0$, and further assume that $f_x(a_0,x_0) \neq 0$. Then there is $\delta > 0$ and a unique smooth function $g(a)$ defined in $(a_0 - \delta, a_0 + \delta)$ such that $g(a)$ is an equilibrium of $x' = f(a,x)$, that is, $f(a,g(a)) = 0$, and $g(a_0) = x_0$.*

Proof. Apply Theorem 2.25 to the function $f(a,x)$.

Remark 2.30. Theorem 2.29 says that if Eq. (2.11) has an equilibrium x_0 and $f_x(a_0,x_0) \neq 0$, then the equilibrium persists under small changes of the parameter a. It also says that to have a bifurcation a necessary condition is $f_x(a_0,x_0) = 0$. Recall that the linear stability of an equilibrium (a_0,x_0) in one-dimensional systems is given by the sign of $f_x(a_0,x_0)$. Therefore, we expect that a bifurcation may be associated to a change of stability type.

2.5.2 Fold Bifurcation

The most typical bifurcation in 1-dimensional dynamical system is the fold bifurcation. We first discuss the fold bifurcation in an example.

Example 2.31. Consider the equation

$$x' = f(a,x) = a - x^2, \quad a \in \mathbf{R}.$$

For $a < 0$ there are no equilibria, for $a = 0$ there is one equilibrium at $x_0 = 0$, and for $a > 0$ there are two equilibria $x_1(a) = -\sqrt{a}$ (unstable) and $x_2(a) = \sqrt{a}$ (asymptotically stable). At $a = 0$ a bifurcation takes place, and we have the simultaneous creation of two equilibria, a stable and an unstable one. This is called a *fold bifurcation* or *saddle-node bifurcation.* We can summarize what is going on in the following picture, called a *bifurcation diagram*, where we show the position and stability of the equilibria of the equation as the parameter a changes. Note that the bifurcation takes place at $a_0 = 0$ and there the equilibrium is $x_0 = 0$. We have $f'_{a_0}(x_0) = -2x_0 = 0$ and therefore from the discussion in the previous section we do expect that a bifurcation may be taking place here.

Remark 2.32. In what follows we assume that a bifurcation takes place at $a_0 = 0$ and $x_0 = 0$. This is not restrictive, since if in some system a bifurcation takes place at $a_0 \neq 0$ or $x_0 \neq 0$, we can define a new parameter $b = a - a_0$ and a new coordinate $y = x - x_0$ so that in terms of (b,y) the bifurcation takes place at $b_0 = 0$, $y_0 = 0$.

We now give a theoretical result on fold bifurcations.

Theorem 2.33 (Fold Bifurcation Theorem). *Consider the autonomous one-dimensional system $x' = f(a,x)$ and assume that:*

(i) $f(0,0) = 0$;
(ii) $f_x(0,0) = 0$;
(iii) $f_{xx}(0,0) \neq 0$;
(iv) $f_a(0,0) \neq 0$.

Then there is $\delta > 0$ and a function $h : (-\delta,\delta) \to \mathbf{R}$ such that the equilibria of the system are parameterized by $a = h(x)$ and h satisfies $h(0) = h'(0) = 0$, $h''(0) \neq 0$, that is, $h(x)$ is

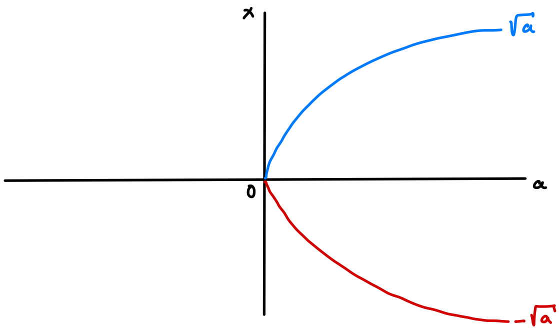


Figure 2.4: Bifurcation diagram for the equation $x' = a - x^2$. In this and all subsequent bifurcation diagrams blue curves correspond to stable equilibria and red curves to unstable ones.

approximately a parabola near $x = 0$*. Moreover, the stability of the equilibria changes when* x *changes sign.*

Proof. From assumption (i) we have that $f(0,0) = 0$. Because of assumption (iv) we can apply Theorem 2.25 to solve the equation $f(a,x) = 0$ for a. That is, there is $\delta > 0$ and a function $h : U \to \mathbf{R}$ such that $h(0) = 0$ and $f(h(x), x) = 0$. The function h parameterizes a curve $a = h(x)$ on the (a,x) plane.

Therefore, for any x_0 sufficiently close to 0 there is a unique $a_0 = h(x_0)$ such that x_0 is an equilibrium of $x' = f(a_0, x)$.

We now determine the shape of the graph of $h(x)$. Define the function $g : U \to \mathbf{R}$ by

$$g(x) = f(h(x), x) = 0.$$

Then

$$g'(x) = f_a(h(x), x)h'(x) + f_x(h(x), x) = 0.$$

Evaluating the last relation at $x = 0$, which implies $h(0) = 0$, we find

$$f_a(0,0)h'(0) + f_x(0,0) = 0.$$

Because of assumption (ii) the last term in the left-hand side vanishes, and therefore using again assumption (iv) we find

$$h'(0) = 0.$$

We can proceed in the same way for the second derivative of g. We find

$$g''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[f_a(h(x), x)\right] h'(x) + f_a(h(x), x)h''(x) + \frac{\mathrm{d}}{\mathrm{d}x}\left[f_x(h(x), x)\right] = 0.$$

Evaluating at $x = 0$ we find

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0}\left[f_a(h(x), x)\right] h'(0) + f_a(0,0)h''(0) + \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0}\left[f_x(h(x), x)\right] = 0.$$

Since $h'(0) = 0$ we further get

$$f_a(0,0)h''(0) + \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0} [f_x(h(x), x)] = 0.$$

To finish this computation, note that

$$\frac{\mathrm{d}}{\mathrm{d}x} [f_x(h(x), x)] = f_{ax}(h(x), x)h'(x) + f_{xx}(h(x), x),$$

which when evaluated at $x = 0$ gives

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0} [f_x(h(x), x)] = f_{ax}(0,0)h'(0) + f_{xx}(0,0)$$

or

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0} [f_x(h(x), x)] = f_{xx}(0,0).$$

Therefore, we find

$$f_a(0,0)h''(0) + f_{xx}(0,0) = 0,$$

and from here

$$h''(0) = -\frac{f_{xx}(0,0)}{f_a(0,0)} \neq 0.$$

Since $h(0) = h'(0) = 0$ and $h''(0) \neq 0$ this shows that the function $a = h(x)$ looks like the parabola $a = \frac{1}{2}h''(0)x^2$ near the origin.

To check the stability of the equilibria we need to compute $f_x(h(x), x)$ for x near the origin. Define $s(x) = f_x(h(x), x)$, that is, the sign of $s(x)$ determines the stability of the equilibrium at x. We have $s(0) = f_x(0,0) = 0$ and

$$s'(x) = f_{ax}(h(x), x)h'(x) + f_{xx}(h(x), x),$$

giving

$$s'(0) = f_{ax}(0,0)h'(0) + f_{xx}(0,0) = f_{xx}(0,0) \neq 0.$$

Therefore, $s(x)$ changes sign at $x = 0$ and we conclude that the branch of equilibria for $x < 0$ has the opposite stability from the branch of equilibria with $x > 0$. ☑

Remark 2.34. For $x > 0$, $h'(x)$ has the same sign as $h''(0)$, that is, the same sign as $-f_{xx}(0,0)/f_a(0,0)$. So, $f_x(h(x), x)$ has the same sign as $f_{xx}(0,0)$. Therefore, if $f_{xx}(0,0) > 0$ then the branch $x > 0$ is unstable. If $f_{xx}(0,0) < 0$ then the branch $x > 0$ is stable. In the example before we had $f(a, x) = a - x^2$. Therefore, $f_{xx}(0,0) = -2 < 0$. This implies that the branch with $x > 0$ is stable. ❞

Remark 2.35. The first two assumptions in Theorem 2.33 must always be satisfied for any bifurcation of equilibria to take place. From this point of view they do not provide any restrictions. Assumptions (iii) and (iv) mean that almost all functions $f(a, x)$ that satisfy assumptions (i) and (ii) go through a fold bifurcation. The idea here is that to have $f_{xx}(0,0) = 0$ or $f_a(0,0) = 0$ then we must choose f in a special way. A "randomly" chosen function f will not satisfy any of these properties and therefore it will satisfy assumptions (iii) and (iv). ❞

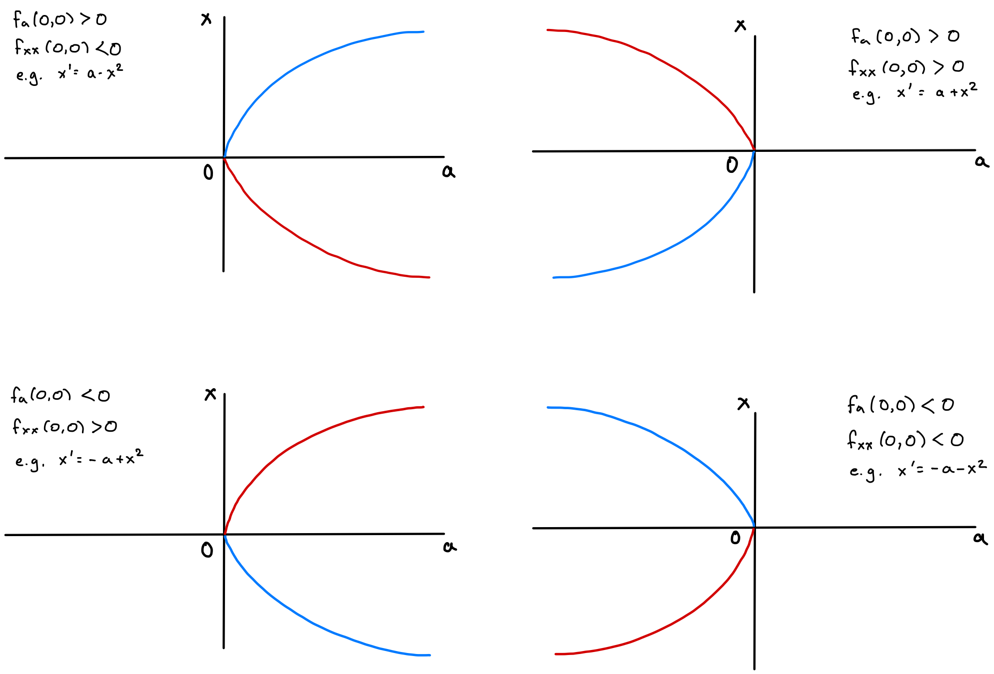


Figure 2.5: Fold bifurcation diagrams for different signs of $f_a(0,0)$ and $f_{xx}(0,0)$.

2.5.3 Cusp Bifurcation

We have seen that equilibria are organizing centers for the dynamics of a system, and bifurcations involving one parameter are organizing centers for the equilibria in their vicinity. It turns out that bifurcations involving two parameters are organizing centers for one-parameter bifurcations. We briefly discuss here the cusp bifurcation which is an example of a two-parameter bifurcation. For a more in depth discussion of the cusp bifurcation we refer to [6, 8].

Systems depending on two parameters can be written as

$$x' = f(a, b, x), \ (a, b) \in \mathbf{R}^2.$$

Since the function f is now a function from $\mathbf{R}^3$ to $\mathbf{R}$ we need the following special case of the general implicit function theorem (Theorem 2.24).

Theorem 2.36 (Implicit Function Theorem for functions $\mathbf{R}^3 \to \mathbf{R}$). *Consider a smooth function $f : \mathbf{R}^3 \to \mathbf{R}$ such that $f(a_0, b_0, x_0) = 0$. If $f_x(a_0, b_0, x_0) \neq 0$ then there is $\delta > 0$ and a unique smooth function $h(a, b)$ defined in $B_\delta(a_0, b_0) = \{(a, b) : (a - a_0)^2 + (b - b_0)^2 < \delta^2\}$ such that $h(a_0, b_0) = x_0$ and $f(a, b, h(a, b)) = 0$.*

The previous theorem asserts that if the system $x' = f(a_0, b_0, x)$ has an equilibrium at x_0 and $f_x(a_0, b_0, x_0) \neq 0$, then the equilibrium persists for (a, b) sufficiently close to (a_0, b_0). We discuss now an example where the condition $f_x(a_0, b_0, x_0) \neq 0$ is not satisfied. Consider the system

$$x' = f(a, b, x) = a + bx - x^3.$$

The equation $f(a, b, x) = 0$ defines a two-dimensional surface S in $\mathbf{R}^3$ with coordinates (a, b, x), shown in Fig. 2.6. Each point $(a_0, b_0, x_0) \in S$ corresponds to an equilibrium of the equation $x' = f(a_0, b_0, x)$. We observe that the surface S "folds" and this has as a consequence that there are values (a, b) for which the system has a single equilibrium and other values for which the system has three equilibria. In the boundary between these two regions the system has two equilibria. Finally, when $(a, b) = (0, 0)$ where $f(a, b, x) = -x^3$ there is exactly one equilibrium.

From Theorem 2.36 we expect that the number of equilibria changes only when $f(a, b, x) = f_x(a, b, x) = 0$, that is, only when the assumption of the theorem fails. The equations $f(a, b, x) = f_x(a, b, x) = 0$ define a curve C on the surface S which is given by

$$a + bx - x^3 = 0, \ b - 3x^2 = 0. \tag{2.12}$$

We find $b = 3x^2 \geq 0$ and $a = -2x^3$. Eliminating x from these equations we obtain a curve $\widehat{C}$ in the (a, b) plane shown in Fig. 2.6 and given by

$$27a^2 - 4b^3 = 0. \tag{2.13}$$

The curve $\widehat{C}$ in Fig. 2.6 has a cusp at the origin which is what gives this bifurcation its name.

Remark 2.37. The curve $\widehat{C}$ in Eq. (2.13) is the projection to the (a, b) plane of the curve C in the (a, b, x) space. The curve C is parameterized as $(a, b, x) = (-2x^3, 3x^2, x)$ and therefore it is a smooth curve. Observe that the cusp singularity of $\widehat{C}$ is an artifact of the projection of C. ❞

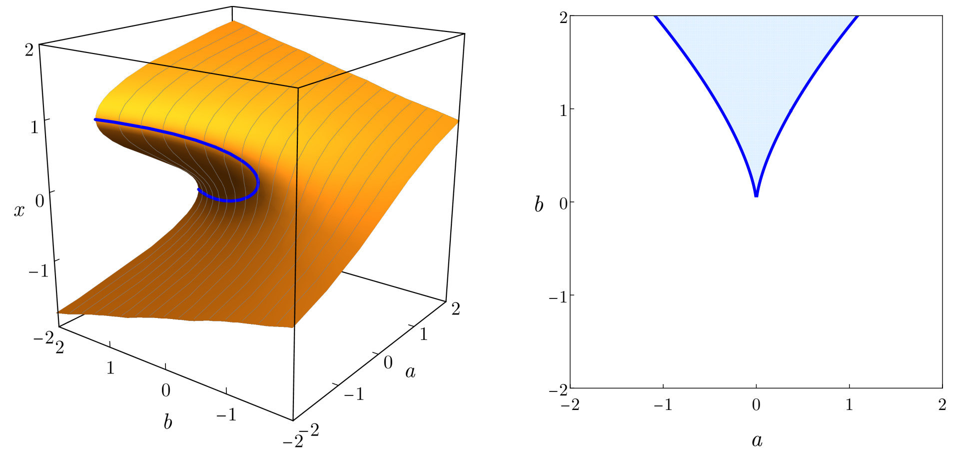


Figure 2.6: Left: Cusp bifurcation surface S, i.e., the solution set of the equation $f(a,b,x) = 0$. This picture shows for each value of the parameters (a,b) the positions x of the corresponding equilibria. The thick blue curve C shows the solution set of the simultaneous equations $f(a,b,x) = f_x(a,b,x) = 0$. Right: The projection of the curve C to the (a,b) parameter plane gives the thick blue curve $\widehat{C}$ having a cusp at the origin. The light blue area corresponds to values of (a,b) for which the system has three equilibria.

In the region on the (a,b) plane outside the cusp the function $f(a,b,x)$ has a single real root (and two complex roots which do not correspond to equilibria of the system). In the region inside the cusp the function has three real roots. At the origin the function has the triple root $x = 0$. Finally, along the two curves $27a^2 - 4b^3 = 0$ (excluding the origin), the function has two real roots, one of which repeats twice.

Proposition 2.38. *For each fixed $b_0 > 0$ the system $x' = f(a, b_0, x)$ goes through fold bifurcations at*

$$a_0 = \pm \frac{2}{3\sqrt{3}} b_0^{3/2}.$$

The content of Proposition 2.38 is shown in Fig. 2.7 showing the dependence of the equilibria of $x' = f(a,b,x)$ on a for some fixed $b > 0$.

Proof. Fix a value $b = b_0 > 0$ and let

$$a_0 = \pm \frac{2}{3\sqrt{3}} b_0^{3/2}.$$

For these values of a_0, b_0 the system $x' = f(a_0, b_0, x)$ has an equilibrium at $x_0 = \mp\sqrt{b_0/3}$ (i.e., for $a_0 < 0$ we have $x_0 > 0$ and vice versa). We check the conditions for the fold bifurcation theorem at (a_0, b_0, x_0). First, we have $f(a_0, b_0, x_0) = 0$ and $f_x(a_0, b_0, x_0) = 0$. Then we check

$$f_{xx}(a_0, b_0, x_0) = -6x_0 = \pm 2\sqrt{3b_0} \neq 0,$$

and

$$f_a(a_0, b_0, x_0) = 1 \neq 0.$$

Therefore, we conclude that there are indeed fold bifurcations at the given values of a_0. ☑

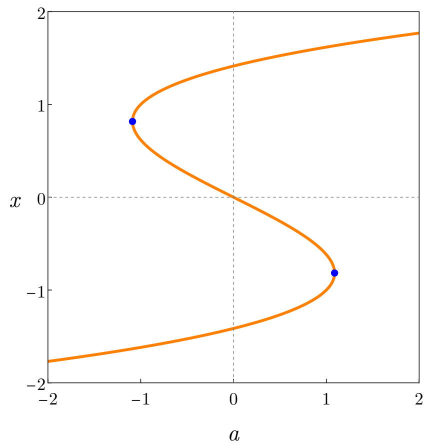


Figure 2.7: Equilibria of $x' = a + bx - x^3$ for fixed $b = 2$. The blue points mark the two fold bifurcations.

2.6 Discrete-time Dynamical Systems

We now consider a different type of dynamical system on an interval $I \subseteq \mathbf{R}$, where the time is not continuous but it changes in discrete steps, taking integer values. Given a function $p : I \to I$ we define dynamics on I through repeated applications of p. That is, given an initial point $x_0 \in I$ we define

$$x_k = p(x_{k-1}), \quad k = 1, 2, 3, \ldots$$

For notational convenience we denote by p^k the composition of k copies of p, that is,

$$p^k = \overbrace{p \circ p \circ \cdots \circ p}^{k \text{ times}},$$

with p^0 being defined as the identity function $p^0(x) = x$, $x \in I$. Using this notation, we can write

$$x_k = p^k(x_0), \quad k = 0, 1, 2, 3, \ldots$$

If the function p is invertible we define for negative integers k,

$$p^k = \overbrace{p^{-1} \circ p^{-1} \circ \cdots \circ p^{-1}}^{|k| \text{ times}},$$

and we extend the sequence $\{x_k\}$ also to negative k by defining

$$x_k = p^k(x_0), \quad k \in \mathbf{Z}.$$

Example 2.39. Consider the function $p : \mathbf{R} \to \mathbf{R}$ given by $p(x) = 2x$. For an initial point

x_0 we obtain for $x \geq 0$ that

$$
\begin{aligned}
x_1 &= 2x_0, \\
x_2 &= 2x_1 = 2^2 x_0, \\
x_3 &= 2x_2 = 2^3 x_0, \\
&\vdots \\
x_k &= 2^k x_0, \\
&\vdots
\end{aligned}
$$

The given function is invertible, with inverse $p^{-1}(x) = 2^{-1}x$. Therefore, we have

$$
\begin{aligned}
x_{-1} &= 2^{-1} x_0, \\
x_{-2} &= 2^{-1} x_1 = 2^{-2} x_0, \\
&\vdots \\
x_{-k} &= 2^{-k} x_0, \\
&\vdots
\end{aligned}
$$

Therefore, for all $k \in \mathbf{Z}$ we have

$$x_k = 2^k x_0.$$

2.6.1 Fixed Points and Periodic Points

Fixed points and periodic points in discrete-time dynamical systems play a similar role to equilibria in continuous-time dynamical systems, as they organize the dynamics for the whole space.

Definition 2.40. A point $x_0 \in I$ is a *fixed point* of the map $p : I \to I$ if $p(x_0) = x_0$ and it is a *period m point* of p if $p^m(x_0) = x_0$.

Notice that fixed points and period 1 points are the same. Moreover, period m points of p are fixed points of $p^m = p \circ \cdots \circ p$ (m times).

Example 2.41. Consider the map $p(x) = 2x$. Then the only fixed point, satisfying $p(x) = x$, is $x = 0$.

Example 2.42. Consider the map $p : [0, 2\pi]$ given by

$$p(x) = x + \frac{3}{4} \sin x.$$

The fixed points are obtained by solving $p(x) = x$, that is, they are the solutions of $\sin x = 0$ in $[0, 2\pi]$. Therefore, we find that the fixed points of p are $0, \pi, 2\pi$. This can be seen graphically, by considering the intersection points of the graph of p and the identity function x in $[0, 2\pi]$, see Fig. 2.8.

Example 2.43. Consider the *logistic map* $p : [0, 1]$ given by

$$p(x) = 4x(1 - x).$$

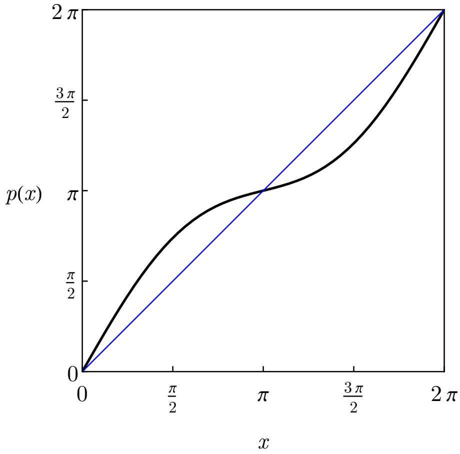


Figure 2.8: The graph of $p(x) = x + \frac{3}{4}\sin x$ in Example 2.42. The fixed points of p can be found by considering the intersections of the graph of p with the graph of the identity function.

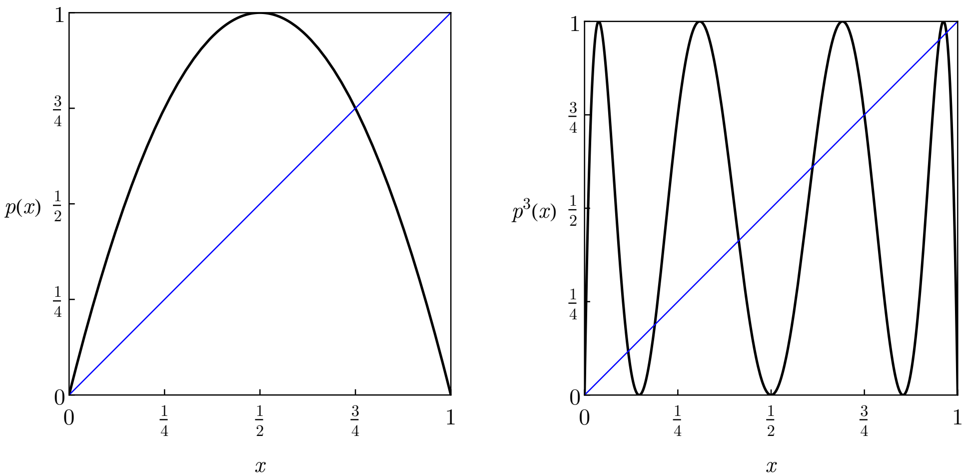


Figure 2.9: Left: the graph of $p(x) = 4x(1-x)$ in Example 2.43. Right: the graph of $p^3(x)$.

The fixed points are obtained by solving $p(x) = x$ in $[0, 1]$ and they are 0 and 3/4, see Fig. 2.9 (left). For this map, we consider also period 3 points. To graphically locate these points we draw the graph of $p^3(x)$ and find its intersections with the graph of the identity function, see Fig. 2.9 (right). We observe that there are 8 fixed points of p^3, that is, 8 period 3 points of p. Notice that 0 and 3/4 are fixed points of p and they are also period 3 points of p. This reflects the fact that if x_0 is a period m point, then it is also a period km point for all $k = 1, 2, 3, \dots$, since $p^{km}(x_0) = p^m \circ p^m \circ \cdots \circ p^m(x_0)$ (k times) $= x$.

We now turn our attention to the question of stability of a fixed point x_* of p. We want to understand if we start with an initial point x_0 near x_*, what will be the behavior of the sequence $\{x_k\}$. Assuming that p is differentiable we consider the *linearization* of the map p at x_*, that is, we approximate

$$p(x) \approx p(x_*) + p'(x_*)(x - x_*) = x_* + p'(x_*)(x - x_*).$$

Therefore,

$$x_k - x_* = p(x_{k-1}) - x_* \approx p'(x_*)(x_{k-1} - x_*).$$

If we denote $y_k = x_k - x_*$ and $\lambda = p'(x_*)$ we can write the last relation as $y_k = \lambda y_{k-1}$. This implies

$$y_k = \lambda^k y_0,$$

that is,

$$x_k - x_* = \lambda^k (x_0 - x_*),$$

and in terms of distance,

$$|x_k - x_*| = |\lambda|^k |x_0 - x_*|.$$

Therefore, if $|\lambda| < 1$ then $\lim_{k\to\infty} x_k = x_*$. If $|\lambda| > 1$ then the successive distances $|x_0 - x_*|, |x_1 - x_*|, |x_2 - x_*|, \dots$ will increase until the linear approximation is no longer appropriate.

Remark 2.44. If $|\lambda| = 1$ then we cannot conclude stability through the linearization. For example, the maps $p_+(x) = x + x^3$ and $p_-(x) = x - x^3$ have the unique fixed point $x_* = 0$ and $p'_\pm(0) = 1$. However, 0 is an unstable fixed point of p_+, while it is an asymptotically stable fixed point of p_-.

Remark 2.45. The result we obtained, characterizing the stability of a fixed point x_* in terms of $|p'(x_*)|$, also applies to a period m point x_0 of p by considering it as fixed point of p^m. Let $x_k = p^k(x_0)$ for $k = 0, 1, 2, \dots$. Then, $x_m = x_0$ and, in general, $x_{k+m} = x_k$. For each $k = 0, 1, \dots, m - 1$ we have

$$\begin{aligned}(p^m)'(x_0) &= (p \circ p \circ \cdots \circ p)'(x_k)\\ &= p'(p^{m-1}(x_0))p'(p^{m-2}(x_0)) \cdots p'(p(x_0))p'(x_0)\\ &= p'(x_{m-1})p'(x_{m-2}) \cdots p'(x_1)p'(x_0),\end{aligned}$$

We conclude that the fixed point x_0 of p^m is asymptotically stable if and only if

$$\lambda = p'(x_0)p'(x_1) \cdots p'(x_{m-2})p'(x_{m-1}),$$

satisfies $|\lambda| < 1$. Notice that λ is the product of the derivatives $p'(x_k)$, $k = 0, \dots, m - 1$, over all points of the period m orbit. It is not difficult to show, using the same argument, that

$$(p^m)'(x_k) = p'(x_{m-1})p'(x_{m-2}) \cdots p'(x_1)p'(x_0) = \lambda.$$

Therefore, the number λ, being the same for all x_k, $k = 0, \dots, m - 1$, characterizes the stability of the whole periodic orbit.

2.7 Periodic Forcing and Poincaré Maps

In this section we discuss how continuous-time dynamics can give rise to discrete-time dynamics. We consider first order equations $x' = f(t,x)$, where we now allow that f depends on t, that is, we consider non-autonomous equations.

For notational convenience we introduce the function $\phi(t;t_0,x_0)$ which solves the initial value problem $x' = f(t,x)$, $x(t_0) = x_0$, that is,

$$\frac{d}{dt}\big[\phi(t;t_0,x_0)\big] = f(t,\phi(t;t_0,x_0)), \qquad \phi(t_0;t_0,x_0) = x_0.$$

The function $\phi(t;t_0,x_0)$ generalizes the notion of the *flow* that we introduced for autonomous differential equations in Section 2.1.

Moreover, we assume that the dependence of f on t is periodic, that is, there is $T > 0$ such that

$$f(t+T,x) = f(t,x), \quad \text{for all } t \in \mathbf{R}.$$

Proposition 2.46. *If $x_1(t)$ solves the initial value problem $x' = f(t,x)$ with $x(0) = x_0$, then $x_2(t) = x_1(t - mT)$ solves the initial value problem $x' = f(t,x)$ with $x(mT) = x_0$, $m \in \mathbf{Z}$, that is,*

$$x_2(t) = \phi(t;mT,x_0) = \phi(t-mT;0,x_0) = x_1(t-mT).$$

Proof. Suppose that $x_1(t)$ solves the initial value problem $x' = f(t,x)$, $x(0) = x_0$, that is, $x_1'(t) = f(t,x_1(t))$ and $x_1(0) = x_0$. Let $x_2(t) = x_1(t-mT)$. Then

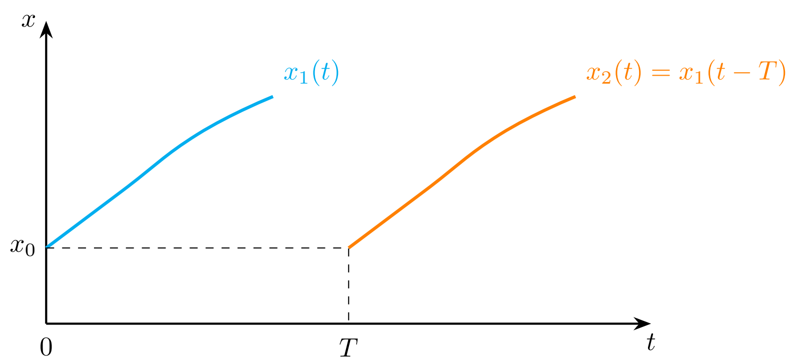
$$x_2'(t) = x_1'(t-mT) = f(t-mT,x_1(t-mT)) = f(t-mT,x_2(t)) = f(t,x_2(t)),$$

where in the last step we used the periodicity of f. Moreover,

$$x_2(mT) = x_1(mT-mT) = x_1(0) = x_0.$$

That is, $x_2(t)$ solves the initial value problem $x' = f(t,x)$, $x(mT) = x_0$. ☑

The relation between the solutions $x_1(t)$ and $x_2(t)$ is shown schematically in the picture below.



Definition 2.47. Consider the equation $x' = f(t,x)$, where $f(t+T,x) = f(t,x)$ for all $t \in \mathbf{R}$ and some $T > 0$. The associated *Poincaré map* $P : \mathbf{R} \to \mathbf{R}$ is defined by

$$P(x_0) = \phi(T;0,x_0).$$

In other words, for a given x_0 we define $P(x_0)$ by finding the solution $\phi(t; 0, x_0)$ of the initial value problem $x' = f(t, x)$, $x(0) = x_0$, and setting $P(x_0) = \phi(T; 0, x_0)$.

Proposition 2.48. *Consider the equation $x' = f(t, x)$, where $f(t + T, x) = f(t, x)$ for all $t \in \mathbf{R}$ and some $T > 0$ and its associated Poincaré map P. Then*

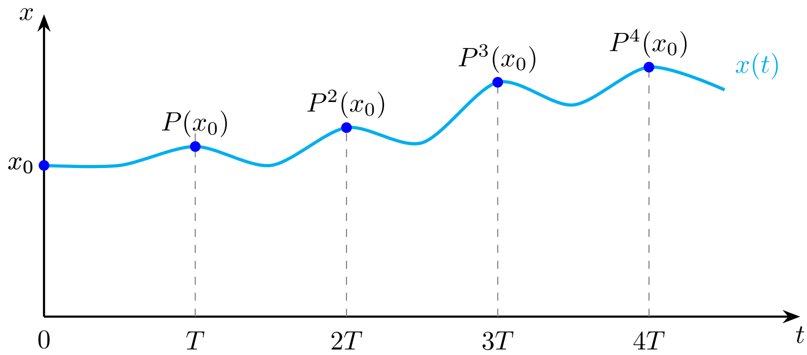
$$P^k(x_0) = \phi(kT; 0, x_0).$$

Proof. The given statement is true for $k = 1$, where it is just the definition of P. Assuming that it holds for some k, then we have

$$\begin{aligned} P^{k+1}(x_0) &= P(P^k(x_0)) \\ &= P(\phi(kT; 0, x_0)) \\ &= \phi(T; 0, \phi(kT; 0, x_0)) \\ &= \phi((k+1)T; kT, \phi(kT; 0, x_0)) \qquad \text{(by Proposition 2.46)} \\ &= \phi((k+1)T; 0, x_0). \end{aligned}$$

For the last step we used the fact that $\phi(t; kT, \phi(kT; 0, x_0))$ represents the solution that starts at time kT from the point $\phi(kT; 0, x_0)$. However, this solution is the same solution $\phi(t; 0, x_0)$ that started at time 0 from the point x_0 and in time kT has reached the point $\phi(kT; 0, x_0)$. Therefore, we have $\phi(t; kT, \phi(kT; 0, x_0)) = \phi(t; 0, x_0)$. ☑

Remark 2.49. Proposition 2.48 has a very intuitive interpretation. Suppose that we consider the solution $x(t)$ to the initial value problem $x' = f(t, x)$, $x(0) = x_0$, that is, $x(t) = \phi(t; 0, x_0)$. The sequence of points $x(kT)$, $k = 0, 1, 2, \ldots$, corresponds to directing a spotlight to the point $x(t)$ only at times $t_k = kT$. The Poincaré map is that spotlight. Every time we apply P is the same as waiting for time T and then turning on our spotlight. Because of this, the Poincaré map is also often called the *stroboscopic map*. This is shown in the picture below. ❞



Proposition 2.50. *A point x_0 is a fixed point of the Poincaré map P if and only if $\phi(t; 0, x_0)$ is a T periodic solution, that is, $\phi(t + T; 0, x_0) = \phi(t; 0, x_0)$ for all $t \in \mathbf{R}$. A point x_0 is a period m point of the Poincaré map P if and only if $\phi(t; 0, x_0)$ is a mT periodic solution.*

Proof. We only prove the part of the statement concerning fixed points. First, if $\phi(t + T; 0, x_0) = \phi(t; 0, x_0)$ then we have

$$P(x_0) = \phi(T; 0, x_0) = \phi(0; 0, x_0) = x_0.$$

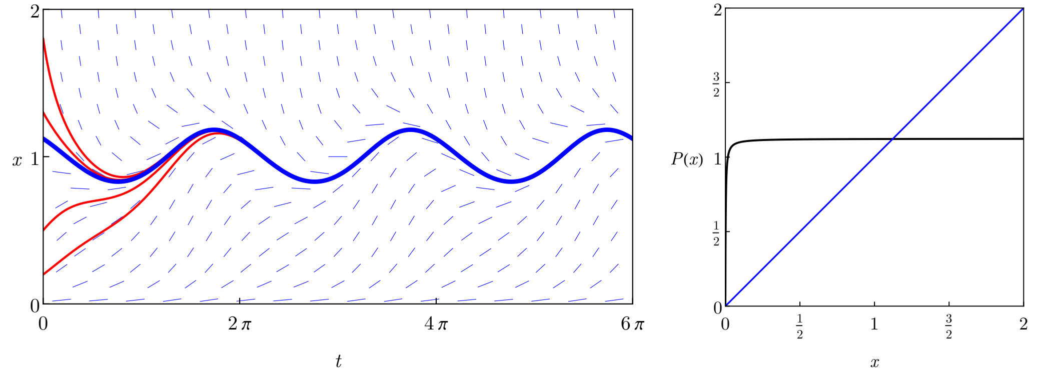


Figure 2.10: The graph of the Poincaré map for $h = 0.25$.

For the opposite direction, assume that x_0 is a fixed point of P, that is, $x_0 = P(x_0) = \phi(T; 0, x_0)$. Then for any t we have

$$\phi(t + T; 0, x_0) = \phi(t + T; T, \phi(T; 0, x_0)),$$

where we use here the same argument as at the end of the proof of Proposition 2.48. That is, we have $\phi(t; 0, x_0) = \phi(t; T, \phi(T; 0, x_0))$ since the solution that starts at time T from the point $\phi(T; 0, x_0)$ is the same solution that started at time 0 from the point x_0 and in time T has reached $\phi(T; 0, x_0)$. Then

$$\phi(t + T; T, \phi(T; 0, x_0)) = \phi(t + T; T, x_0) = \phi(t; 0, x_0),$$

where at the last step we used Proposition 2.46. Therefore, we conclude that

$$\phi(t + T; 0, x_0) = \phi(t; 0, x_0),$$

that is, the solution is periodic.

Example 2.51. We consider the logistic equation in population dynamics with the addition of periodic harvesting that is proportional to the population size. The resulting equation is

$$x' = x(1 - x) - hx \sin t,$$

where $h \geq 0$ is constant. For $h = 0$, the autonomous equation $x' = x(1-x)$ has the equilibria 0 and 1. When $h > 0$, $x(t) = 0$ remains an equilibrium solution of the equation. However, the other equilibrium solution, $x(t) = 1$, is no longer a valid solution, but is replaced by a periodic solution of period 2π.

Even though in this particular case we can find an analytic expression for the Poincaré map P, this is not in general possible, and we usually adopt a numerical approach to computing the Poincaré map, its fixed points, and their stability.

```
P[h_,x0_] := NDSolveValue[
  x'[t] == x[t](1-x[t]) - h x[t] Sin[t] && x[0] == x0,
  x[2 Pi], {t, 0, 2 Pi}]
```

From the numerically computed graph of the Poincaré map we observe that there are two fixed points: one at $x_0 = 0$, and the other one around $x_0 \approx 1.12$. The second fixed

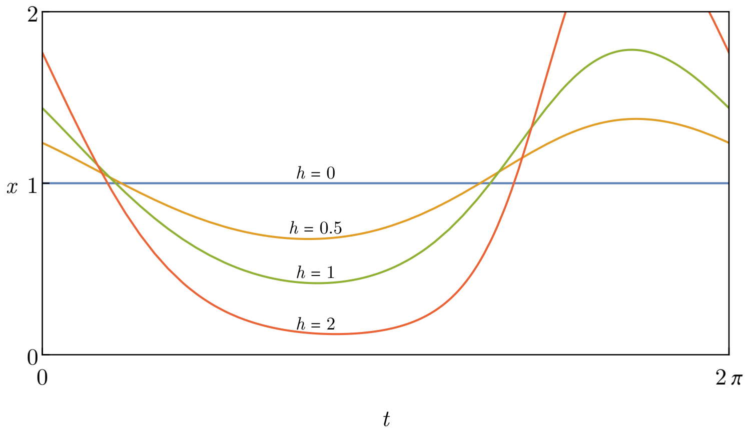


Figure 2.11: Periodic solutions of the logistic equation with harvesting for different values of the harvesting intensity h.

point corresponds to a periodic orbit of period 2π and it is asymptotically stable, since the derivative λ of the Poincaré map at $x_0 \approx 1.12$ can be seen in the plot to be close to 0 (actually, $\lambda \approx 0.0018$), that is, $|\lambda| < 1$. This means that if we start at an initial point close to the fixed point, we will approach the fixed point. This also means that the corresponding solution for the differential equation will approach the periodic solution and we conclude that the periodic solution is asymptotically stable.

To precisely locate the second fixed point we use the following code in Mathematica.

```
p[h_ , x_?NumericQ] := P[h, x]
FindRoot[p[0.25, x0] == x0, {x0, 1}]
```

which returns

```
{x0 -> 1.12154}
```

The value of $P'(x_0)$ can be approximated through a second order finite difference scheme as

$$P'(x_0) \approx \frac{P(x_0+h) - P(x_0-h)}{2h}.$$

Remark 2.52. The equation $x' = x(1-x) - hx\sin t$ is a Bernoulli equation, and therefore it can be solved — at least in principle. A somewhat lengthy computation shows that

$$P(x_0) = \phi(2\pi; 0, x_0) = \frac{x_0 e^{2\pi}}{1 + x_0 e^{2\pi} H(h)},$$

where

$$H(h) = e^{-h-2\pi} \int_0^{2\pi} e^{s+h\cos(s)} ds.$$

The function $H(h)$ has for $h = 0$ the value $H(0) = 1 - e^{-2\pi}$, and is decreasing but strictly positive for $h > 0$. In terms of this expression for the Poincaré map, we find that the fixed points of P are given by $x_0 = 0$ with $P'(0) = e^{2\pi}$ and $x_0 = (1-e^{-2\pi})/H(h)$ with $P'(x_0) = e^{-2\pi}$.

Remark 2.53. Notice that there is a seemingly counterintuitive fact here. The parameter h is related to the intensity of the periodic harvesting. However, we notice that, since $H(h)$ is a decreasing function, the fixed point corresponding to the stable oscillating population is increasing with h. This is an artifact of the fact that the Poincaré map catches the state of the system at a specific moment. If we compute the average of the periodic solution over a period 2π then we find that it always equals 1. Therefore, the intensity of the periodic harvesting does not change the average population, however, larger h lead to larger amplitude oscillations as shown in Fig. 2.11. ❞

Chapter 3

Linear Second-Order Differential Equations

In Chapter 1 we discussed first-order differential equations involving the dependent variable y and its first derivative y'. We now turn our attention to second-order differential equations, so called because they also involve the second derivative y'' of the dependent variable.

Consider the simplest second-order differential equation, $y'' = 0$, and denote by x the independent variable. Integrating the relation $y'' = 0$ we obtain $y' = c_1$, $c_1 \in \mathbf{R}$. Integrating once more we obtain the solution $y(x) = c_1 x + c_2$, $c_1, c_2 \in \mathbf{R}$. The two integrations used to solve the second-order equation introduced two parameters $c_1, c_2 \in \mathbf{R}$ in the general solution.

The appearance of the two parameters in the general solution implies that to specify one single solution we must be provided with two pieces of information for the solution. The most common choices are:

(i) The values of the solution $y(x)$ and its derivative $y'(x)$ at the same $x = x_0$; these give rise to *initial value problems*.
(ii) The values of the solution $y(x)$ at two different x; these give rise to *boundary value problems* which are briefly discussed in Section 3.4.

Definition 3.1 (Initial Value Problem). An initial value problem for a second-order differential equation consists of an equation $g(x, y, y', y'') = 0$ and the initial conditions $y(x_0) = \alpha_0$, $y'(x_0) = \alpha_1$.

Similarly to first-order equations one can state an Existence and Uniqueness Theorem similar for initial value problem for second-order differential equations. We give a special version of the Existence and Uniqueness Theorem as Theorem 3.3 in Section 3.1.

Increasing the order from one to two increases the difficulty of analytically solving the equation. To have any hope to obtain a general solution method we need to restrict our attention to the very limited, however very important, class of linear second-order differential equations discussed in the next section.

3.1 Linear Second-Order Differential Equations

Definition 3.2. A second-order differential equation is *linear* if it has the form

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x), \tag{3.1}$$

where $a_0(x)$, $a_1(x)$, $a_2(x)$, and $f(x)$ are continuous in an interval $I \subseteq \mathbf{R}$ and $a_2(x) \neq 0$ for all $x \in I$. Equation (3.1) is *homogeneous* when $f(x) = 0$ for all $x \in I$, and *non-homogeneous* otherwise.

The following theorem ensures the existence of a unique solution in the interval I specified in the previous definition.

Theorem 3.3 (Existence and Uniqueness). *Equation* (3.1) *has a unique solution* $y(x)$, $x \in I$, *satisfying given initial conditions* $y(x_0) = \alpha_0$, $y'(x_0) = \alpha_1$.

There are two properties of linear differential equations, especially homogeneous equations, that make them special. The first one is the superposition principle, stating that a linear combination of two solutions of a homogeneous equation is again a solution. The second one is the fact that the general solution of linear homogeneous equations is the collection of all linear combinations of two linearly independent solutions, discussed in detail in Sections 3.1.1 and 3.1.2. We start here with the superposition principle in a slightly generalized form.

Theorem 3.4 (Superposition principle). *Consider a solution* $y_1(x)$ *of the equation* $a_2(x)y'' + a_1(x)y' + a_0(x)y = f_1(x)$ *and a solution* $y_2(x)$ *of the equation* $a_2(x)y'' + a_1(x)y' + a_0(x)y = f_2(x)$. *Then*

$$y(x) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbf{R}$$

is a solution of

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = c_1 f_1(x) + c_2 f_2(x).$$

Proof. The proof is a straightforward computation:

$$\begin{aligned} a_2(x)y'' + a_1(x)y' + a_0(x)y &= a_2(x)(c_1 y_1'' + c_2 y_2'') + a_1(x)(c_1 y_1' + c_2 y_2') + a_0(x)(c_1 y_1 + c_2 y_2) \\ &= c_1(a_2(x)y_1'' + a_1(x)y_1' + a_0(x)y_1) + c_2(a_2(x)y_2'' + a_1(x)y_2' + a_0(x)y_2) \\ &= c_1 f_1(x) + c_2 f_2(x). \end{aligned}$$

☑

Corollary 3.5. *If* $y_1(x)$, $y_2(x)$ *are solutions of the homogeneous linear equation* $a_2(x)y'' + a_1(x)y' + a_0(x)y = 0$ *then any linear combination*

$$y(x) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbf{R}$$

is a solution of the same homogeneous linear equation.

3.1.1 Linear Independence

Linear equations have nice properties that significantly simplify the task of finding a general solution. To state the main result on the form of the general solution of a linear equation we first need to make a small detour and discuss the linear independence of functions.

Definition 3.6. Two functions $y_1(x), y_2(x)$ are *linearly independent* in an interval $I \subseteq \mathbf{R}$ if there are no $\lambda_1, \lambda_2 \in \mathbf{R}$ with $|\lambda_1| + |\lambda_2| \neq 0$ such that

$$\lambda_1 y_1(x) + \lambda_2 y_2(x) = 0,$$

for all $x \in I$.

Definition 3.7. The matrix

$$M[y_1, y_2](x) = \begin{bmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{bmatrix}$$

is called the *Wronskian matrix* of the functions $y_1(x)$, $y_2(x)$, and its determinant,

$$W[y_1, y_2](x) = y_1(x)y_2'(x) - y_1'(x)y_2(x),$$

is called the *Wronskian determinant.*

Proposition 3.8. *If y_1, y_2 are continuously differentiable functions on an interval I and $W[y_1, y_2](x) \neq 0$ for some $x \in I$, then y_1, y_2 are linearly independent on I.*

Proof. We will prove this statement by contradiction. Suppose that $W[y_1, y_2](x) \neq 0$ for some $x \in I$ but y_1, y_2 are linearly dependent on I. Then there are constant real numbers λ_1, λ_2, not both zero, such that $\lambda_1 y_1 + \lambda_2 y_2 \equiv 0$. Since at least one of λ_1, λ_2 is not zero we can assume that $\lambda_1 \neq 0$ — if $\lambda_1 = 0$ then choose λ_2 instead to continue the argument. Then $y_1 = -\lambda_2/\lambda_1 y_2 = \lambda y_2$ where we have defined $\lambda = -\lambda_2/\lambda_1$. Moreover, $y_1' = \lambda y_2'$. Then the Wronskian determinant is

$$W[y_1, y_2](x) = y_1(x)y_2'(x) - y_1'(x)y_2(x) = \lambda y_2(x)y_2'(x) - \lambda y_2'(x)y_2(x) = 0,$$

for all $x \in I$, contradicting that $W[y_1, y_2](x) \neq 0$ for some $x \in I$. ☑

Remark 3.9. The converse statement of Proposition 3.8 does not hold since one can find examples of linearly independent functions y_1, y_2 on an interval I for which $W[y_1, y_2](x) = 0$ for all $x \in I$. One such example is $y_1 = x^2$ and $y_2 = x|x|$ on $I = [-1, 1]$. It can be checked that y_1, y_2 are continuously differentiable on I with $y_1' = 2x$, $y_2' = 2|x|$, and thus $W[y_1, y_2](x) = 0$ for all $x \in I$. However, the functions are linearly independent since there are no (non-zero) constants λ_1, λ_2 with $\lambda_1 x^2 + \lambda_2 x|x| = 0$ for all $x \in I$. ”

Proposition 3.10. *If $y_1(x), y_2(x)$ are solutions of the homogeneous linear second-order equation $y'' + p(x)y' + q(x)y = 0$, then their Wronskian determinant $W(x) = y_1(x)y_2'(x) - y_1'(x)y_2(x)$ satisfies the differential equation $W' = -p(x)W$.*

Proof. We compute

$$W' = (y_1 y_2' - y_1' y_2)' = y_1 y_2'' - y_1'' y_2.$$

Since y_1, y_2 are solutions of the given homogeneous equation they satisfy $y_1'' = -p(x)y_1' - q(x)y_1$ and $y_2'' = -p(x)y_2' - q(x)y_2$. Therefore,

$$W' = -p(x)y_1y_2' - q(x)y_1y_2 + p(x)y_1'y_2 + q(x)y_1y_2 = -p(x)W.$$ ☑

Proposition 3.11. *If $y_1(x), y_2(x)$ are linearly independent solutions of the homogeneous linear second-order equation in Eq.* (3.1) *with $f(x) \equiv 0$ then $W[y_1, y_2](x) \neq 0$ for all $x \in I$.*

Proof. Dividing Eq. (3.1) with $f(x) \equiv 0$ by $a_1(x)$, the differential equation takes the form $y'' + p(x)y' + q(x)y = 0$. Therefore, the Wronskian determinant $W(x) = W[y_1, y_2](x)$ satisfies the differential equation $W' = -p(x)W$. If $x_0 \in I$, then solving the equation for $W(x)$ gives

$$W(x) = W(x_0) \exp\Big(-\int_{x_0}^{x} p(s)\,\mathrm{d}s\Big).$$

This means that either $W(x) = 0$ for all $x \in I$ or $W(x) \neq 0$ for all $x \in I$. ☑

3.1.2 General Solution

The form of the general solution of a homogeneous linear equation is described by the following theorem, the proof of which we defer to Section 5.5.

Theorem 3.12. *The general solution of a homogeneous second-order linear equation, given by Eq.* (3.1) *with $f(x) \equiv 0$, has the form*

$$y(x) = c_1 y_1(x) + c_2 y_2(x), \tag{3.2}$$

where $y_1(x)$ and $y_2(x)$ are two linearly independent solutions of Eq. (3.1) *and $c_1, c_2 \in \mathbf{R}$.*

Therefore, given Theorem 3.12 we now have a strategy for solving homogeneous second-order linear equations: find two linearly independent solutions and write the general solution as their linear combination.

Remark 3.13. We know from Corollary 3.5 that if y_1, y_2 are two solutions of a linear homogeneous second-order equation then any linear combination is also a solution. Theorem 3.12 adds to this that if y_1, y_2 are linearly independent then there is nothing more. *All* solutions are linear combinations of y_1, y_2.

For non-homogeneous equations we have the following result.

Theorem 3.14. *The general solution of a non-homogeneous second-order linear equation, given by Eq.* (3.1) *with $f(x) \not\equiv 0$, has the form*

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + y_p(x), \tag{3.3}$$

where $c_1 y_1(x) + c_2 y_2(x)$ is the general solution of the corresponding homogeneous equation with $f(x) \equiv 0$ and $y_p(x)$ is any solution of the non-homogeneous equation.

Proof. Suppose that $y(x)$ is any solution of the non-homogeneous equation and $y_p(x)$ is the obtained particular solution. Define $y_h(x) = y(x) - y_p(x)$. Then

$$\begin{aligned} a_2(x)y_h'' + a_1(x)y_h' + a_0(x)y_h &= \big[a_2(x)y'' + a_1(x)y' + a_0(x)y\big] \\ &\quad - \big[a_2(x)y_p'' + a_1(x)y_p' + a_0(x)y_p\big] \\ &= f(x) - f(x) = 0. \end{aligned}$$

This means that y_h satisfies the associated homogeneous equation and therefore, because of Theorem 3.12, it must have the form $y_h = c_1 y_1 + c_2 y_2$. Subsequently, y must have the form $y = c_1 y_1 + c_2 y_2 + y_p$.

Theorem 3.14 also suggests a strategy for solving non-homogeneous equations: first solve the associated homogeneous equation, and then find any solution $y_p(x)$ — called *particular solution* — of the non-homogeneous equation.

3.2 Solution Method for Second-Order Linear Differential Equations with Constant Coefficients

Linear second-order differential equations are still difficult to solve in their most general form. To simplify things further we consider *second-order linear differential equations with constant coefficients* of the form

$$ay'' + by' + cy = f(x), \ \ a, b, c \in \mathbf{R}, \ \ a \neq 0. \tag{3.4}$$

In this section we describe how to solve equations of the form given in Eq. (3.4) starting with the homogeneous case, and presenting afterwards the non-homogeneous case.

3.2.1 Homogeneous Case

The solution method for homogeneous second-order linear differential equations with constant coefficients is summarized below. Several examples and the explanation are given later in this section.

Solution method for homogeneous second-order linear differential equations with constant coefficients

To find the general solution of the equation

$$ax'' + bx' + cx = 0, \tag{3.5}$$

substitute the solution $x = e^{rt}$ into Eq. (3.5). This produces the *auxiliary* (or *characteristic*) equation

$$ar^2 + br + c = 0, \tag{3.6}$$

which can be solved to find values for r. There are three cases.

(i) If Eq. (3.6) has two distinct real solutions $r_1 \neq r_2$ then the general solution of Eq. (3.5) has the form

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x},\ c_1, c_2 \in \mathbf{R}.$$

(ii) If Eq. (3.6) has a double real solution r then the general solution of Eq. (3.5) has the form

$$y = c_1 e^{rx} + c_2 x e^{rx},\ c_1, c_2 \in \mathbf{R}.$$

(iii) If Eq. (3.6) has a pair of complex conjugate solutions $\alpha \pm i\beta$ then the general solution of Eq. (3.5) has the form

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

In the following four examples, Examples 3.15 to 3.18, we apply the solution method to obtain the general solution for the given linear homogeneous second-order differential equations with constant coefficients.

Example 3.15 (Distinct real roots). Consider the initial value problem

$$y'' + 2y' - 3y = 0, \quad y(0) = 0,\ y'(0) = 2.$$

The auxiliary equation is $r^2 + 2r - 3 = 0$ with distinct real roots -3 and 1. Therefore, the general solution is

$$y(x) = c_1 e^{-3x} + c_2 e^{x}.$$

For the given initial conditions we have $y(0) = c_1 + c_2 = 0$, and then $y(x) = c_1 e^{-3x} - c_1 e^{x}$. Computing the derivative we find

$$y'(x) = (c_1 e^{-3x} - c_1 e^{x})' = c_1(-3e^{-3x} - e^{x}),$$

so $y'(0) = -4c_1 = 2$, giving $c_1 = -1/2$. We conclude that the solution to the given initial value problem is

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

Example 3.16 (Distinct real roots). Consider the initial value problem

$$y'' + 2y' = 0, \quad y(0) = 0,\ y'(0) = 2.$$

The auxiliary equation is $r^2 + 2r = 0$ with distinct real roots -2 and 0. Therefore, the general solution is

$$y(x) = c_1e^{-2x} + c_2e^{0x} = c_1e^{-2x} + c_2.$$

For the given initial conditions we have $y(0) = c_1 + c_2 = 0$, and then $y(x) = c_1e^{-2x} - c_1$. Computing the derivative we find

$$y'(x) = (c_1e^{-2x} - c_1)' = -2c_1e^{-2x},$$

so $y'(0) = -2c_1 = 2$, giving $c_1 = -1$. We conclude that the solution to the given initial value problem is

$$y(x) = -e^{-2x} + 1.$$

Example 3.17 (Double root). Consider the initial value problem

$$y'' + 2y' + y = 0, \quad y(0) = 0,\ y'(0) = 2.$$

The auxiliary equation is $r^2 + 2r + 1 = 0$ with double root -1. Therefore, the general solution is

$$y(x) = c_1e^{-x} + c_2xe^{-x}.$$

For the given initial conditions we have $y(0) = c_1 = 0$, and then $y(x) = c_2xe^{-x}$. Computing the derivative we find

$$y'(x) = (c_2xe^{-x})' = c_2(e^{-x} - xe^{-x}),$$

so $y'(0) = c_2 = 2$. We conclude that the solution to the given initial value problem is

$$y(x) = 2xe^{-x}.$$

Example 3.18 (Complex conjugate roots). Consider the initial value problem

$$y'' + 2y' + 2y = 0, \quad y(0) = 0,\ y'(0) = 2.$$

The auxiliary equation is $r^2 + 2r + 2 = 0$ with roots $-1 \pm i$, that is, $\alpha = -1$ and $\beta = 1$. Therefore, the general solution is

$$y(x) = c_1e^{-x}\cos x + c_2e^{-x}\sin x.$$

For the given initial conditions we have $y(0) = c_1 = 0$, and then $y(x) = c_2e^{-x}\sin x$. Computing the derivative we find

$$y'(x) = (c_2e^{-x}\sin x)' = c_2(-e^{-x}\sin x + e^{-x}\cos x),$$

so $y'(0) = c_2 = 2$. We conclude that the solution to the given initial value problem is

$$y(x) = 2e^{-x}\sin x.$$

Example 3.19 (Complex conjugate roots). Consider the initial value problem

$$y'' + 4y = 0, \quad y(0) = 1,\ y'(0) = 2.$$

The auxiliary equation is $r^2 + 4 = 0$ with roots $\pm 2i$, that is, $\alpha = 0$ and $\beta = 2$. Therefore, the general solution is

$$y(x) = c_1 \cos 2x + c_2 \sin 2x.$$

For the given initial conditions we have $y(0) = c_1 = 1$. Computing the derivative we find

$$y'(x) = -2c_1 \sin 2x + 2c_2 \cos 2x,$$

so $y'(0) = 2c_2 = 2$. We conclude that the solution to the given initial value problem is

$$y(x) = \cos 2x + \sin 2x.$$

We are now ready to show the validity of the solution method for second-order homogeneous linear equations with constant coefficients.

Theorem 3.20. *The equation $ay'' + by' + cy = 0$, $a \neq 0$, has one of the following three types of general solutions depending on the sign of the discriminant $\Delta = b^2 - 4ac$ of the auxiliary equation $ar^2 + br + c = 0$.*

(i) If $\Delta > 0$, then the general solution is $y(x) = c_1 e^{r_1 x} + c_2 e^{r_2 x}$, where r_1, r_2 are the real and distinct roots of the auxiliary equation.
(ii) If $\Delta = 0$, then the general solution is $y(x) = c_1 e^{rx} + c_2 x e^{rx}$, where $r = -b/(2a)$ is the real double root of the auxiliary equation.
(iii) If $\Delta < 0$, then the general solution is $y(x) = c_1 e^{\alpha x} \cos(\beta x) + c_2 e^{\alpha x} \sin(\beta x)$, where $\alpha \pm i\beta$ are the two complex conjugate roots of the auxiliary equation.

Proof. Because of Theorem 3.12 we only need to find two linearly independent solutions $y_1(x)$, $y_2(x)$ of the given equation. First, notice that if r satisfies the auxiliary equation $ar^2 + br + c = 0$ then e^{rx} is a solution of the equation $ay'' + by' + cy = 0$, since substituting e^{rx} into the given equation we obtain

$$ar^2 e^{rx} + bre^{rx} + ce^{rx} = (ar^2 + br + c)e^{rx} = 0.$$

(i) In the case $\Delta > 0$, the auxiliary equation has two distinct real roots r_1, r_2. Therefore, $y_1 = e^{r_1 x}$ and $y_2 = e^{r_2 x}$ are solutions of the differential equation. These are linearly independent since

$$W[y_1, y_2](x) = (r_2 - r_1)e^{(r_1 + r_2)x} \neq 0.$$

Therefore, the general solution in this case is $y(x) = c_1 e^{r_1 x} + c_2 e^{r_2 x}$.

(ii) In the case $\Delta = 0$, the auxiliary equation has a double real root $r = -b/(2a)$. Therefore, $y_1 = e^{rx}$ is a solution of the differential equation. We show that $y_2 = xy_1 = xe^{rx}$ is also a solution. Since $y_2' = xy_1' + y_1$ and $y_2'' = xy_1'' + 2y_1'$, we have

$$\begin{aligned} ay_2'' + by_2' + cy_2 &= a(xy_1'' + 2y_1') + b(xy_1' + y_1) + cxy_1 \\ &= x(ay_1'' + by_1' + cy_1) + (2ay_1' + by_1). \end{aligned}$$

The term $ay_1'' + by_1' + cy_1 = 0$, since $y_1 = e^{rx}$ is a solution of the equation. For the second term, we notice that $y_1' = re^{rx} = ry_1$. Therefore,

$$ay_2'' + by_2' + cy_2 = 2ay_1' + by_1 = (2ar + b)y_1 = 0.$$

The solutions y_1 and y_2 are linearly independent. The Wronskian determinant is given by
$$W[y_1, y_2](x) = e^{2rx} \neq 0.$$
Therefore, the general solution in this case is $y(x) = c_1 e^{rx} + c_2 x e^{rx}$.

(iii) In the case $\Delta < 0$, the auxiliary equation has the complex conjugate roots $\alpha \pm i\beta$ with $\beta \neq 0$. If we look beyond the realm of real-valued functions, this implies that $z(x) = e^{(\alpha+i\beta)x}$ and its complex conjugate $\bar{z}(x) = e^{(\alpha-i\beta)x}$ are solutions of the homogeneous equation. Any linear combinations of z and $\bar{z}$ also solve the same equation. Therefore, the functions
$$y_1(x) = \frac{1}{2}(z(x) + \bar{z}(x)) = \mathrm{Re}(z(x)) = e^{\alpha x}\cos(\beta x),$$
and
$$y_2(x) = \frac{1}{2i}(z(x) - \bar{z}(x)) = \mathrm{Im}(z(x)) = e^{\alpha x}\sin(\beta x),$$
must also be solutions (check this directly!). To check that they are linearly independent we compute
$$W[y_1, y_2](x) = \beta e^{2\alpha x} \neq 0.$$
Therefore, the general solution in this case is
$$y(x) = c_1 e^{\alpha x}\cos(\beta x) + c_2 e^{\alpha x}\sin(\beta x).$$
☑

Alternative forms of the general solution in the case of complex roots

In the case of complex roots $\alpha \pm i\beta$ of the auxiliary equation the general solution is
$$y(x) = c_1 e^{\alpha x}\cos(\beta x) + c_2 e^{\alpha x}\sin(\beta x), \tag{3.7}$$
where $c_1, c_2 \in \mathbf{R}$. There are two other commonly used expressions for the general solution in this case, that can be useful in other contexts.

Complex exponentials First, recall that in this case $z(x) = e^{(\alpha+i\beta)x}$ and $\bar{z}(x) = e^{(\alpha-i\beta)x}$ are solutions of the equation. Looking back at the superposition principle, it is easy to see that it also applies if all the quantities involved are complex numbers or complex-valued functions. Therefore, for any numbers $a, b \in \mathbf{C}$, the linear combination
$$y(x) = az(x) + b\bar{z}(x),$$
is also a solution of the equation. It turns out that the solution $y(x)$ is real-valued it and only if $b = \bar{a}$, that is, we get the solution
$$y(x) = az(x) + \bar{a}\bar{z}(x) = ae^{(\alpha+i\beta)x} + \bar{a}e^{(\alpha-i\beta)x} = 2\,\mathrm{Re}(ae^{(\alpha+i\beta)x}), \tag{3.8}$$
where $a \in \mathbf{C}$.

Remark 3.21. To show that $y(x)$ is real-valued if and only if $b = \bar{a}$, recall that $y(x)$ is real-valued if and only if $y(x) - \overline{y(x)} = 0$. We compute
$$y(x) - \overline{y(x)} = az + b\bar{z} - \bar{a}\bar{z} - \bar{b}z = (a - \bar{b})z + (b - \bar{a})\bar{z}.$$
Clearly, if $b = \bar{a}$ then the last expression becomes zero. Conversely, if the last expression is zero, then $b = \bar{a}$. Otherwise, we could rearrange terms to get the relation
$$\frac{a - \bar{b}}{b - \bar{a}} = \frac{\bar{z}}{z} = e^{-2i\beta x},$$
which cannot be true since the right-hand side is not constant for $\beta \neq 0$. ❞

Example 3.22. Consider again the initial value problem

$$y'' + 4y = 0, \quad y(0) = 1, \ y'(0) = 2,$$

from Example 3.19. Since the solutions of the auxiliary equation are $\pm 2i$, the general solution of the differential equation is

$$y(x) = ae^{2ix} + \bar{a}e^{-2ix}.$$

The initial condition $y(0) = 1$ gives $a + \bar{a} = 1$, while the initial condition $y'(0) = 2$ gives $2ia - 2i\bar{a} = 2$, that is, $a - \bar{a} = -i$. Solving for a we obtain $a = \frac{1}{2}(1 - i)$. Therefore, the solution to the initial value problem is

$$y(x) = \frac{1}{2}(1-i)e^{2ix} + \frac{1}{2}(1+i)e^{-2ix} = \mathrm{Re}((1-i)e^{2ix}) = \cos(2x) + \sin(2x).$$

Exercise 3.1. Show that Eq. (3.7) and Eq. (3.8) describe the same solution if and only if $a = \frac{1}{2}(c_1 - ic_2)$. This shows that Eq. (3.8) is an alternative, but equivalent, form of the general solution of the given differential equation.

Single trigonometric function An alternative useful representation of the general solution is the form

$$y(x) = Ae^{\alpha x}\cos(\beta x - \phi), \tag{3.9}$$

where $A \in \mathbf{R}$, and $\phi \in (-\pi, \pi]$.

Using standard trigonometric identities, Eq. (3.9) can be rewritten as

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

Comparing with Equation (3.7) we find that

$$c_1 = A\cos(\phi), \quad c_2 = A\sin(\phi).$$

Therefore, if (c_1, c_2) represent Cartesian coordinates of a point in $\mathbf{R}^2$, then (A, ϕ) are its corresponding polar coordinates. In particular,

$$A = \sqrt{c_1^2 + c_2^2}, \quad \phi = \arg(c_1 + ic_2).$$

Exercise 3.2. The alternative trigonometric representation $y(x) = Ae^{\alpha x}\sin(\beta x + \phi)$ is commonly used in physics. Show that in this case $c_1 = A\sin(\phi)$ and $c_2 = A\cos(\phi)$, that is, (A, ϕ) are polar coordinates for the point with Cartesian coordinates (c_2, c_1).

Example 3.23. Consider again the initial value problem

$$y'' + 4y = 0, \quad y(0) = 1, \ y'(0) = 2,$$

from Example 3.19. Since the solutions of the auxiliary equation are $\pm 2i$, the general solution of the differential equation is

$$y(x) = A\cos(2x - \phi).$$

The initial condition $y(0) = 1$ gives $A\cos\phi = 1$, while the initial condition $y'(0) = 2$ gives $2A\sin\phi = 2$, that is, $A\sin\phi = 1$. That is, (A, ϕ) are polar coordinates for the point with Cartesian coordinates $(1, 1)$, giving $A = \sqrt{2}$ and $\phi = \pi/4$. Therefore, the solution to the initial value problem is

$$y(x) = \sqrt{2}\cos\left(2x - \frac{\pi}{4}\right) = \sqrt{2}\cos(2x)\cos\left(\frac{\pi}{4}\right) + \sqrt{2}\sin(2x)\sin\left(\frac{\pi}{4}\right) = \cos(2x) + \sin(2x).$$

3.2.2 Non-Homogeneous Case

We now consider the non-homogeneous second order linear equation with constant coefficients

$$ay'' + by' + cy = f(x). \tag{3.10}$$

The following is a direct consequence of Theorem 3.14.

Solution method for non-homogeneous second order linear differential equations with constant coefficients

To find the general solution of Equation (3.10):

(i) Find the general solution $c_1y_1 + c_2y_2$ of the associated homogeneous equation $ay'' + by' + cy = 0$.
(ii) Find any solution $y_p(x)$ of Equation (3.10), also called a *particular solution*.
(iii) Then, the general solution of Equation (3.10) is

$$y(x) = c_1y_1(x) + c_2y_2(x) + y_p(x).$$

In Section 3.2.1 we discussed how to obtain the general solution of the associated homogeneous equation. Therefore, to find the general solution to the non-homogeneous Eq. (3.10) we still need to find a particular solution $y_p(x)$. The method of undetermined coefficients provides a systematic, algebraic method for determining $y_p(x)$ when the right-hand side $f(x)$ is of a specific type.

Method of undetermined coefficients

The method of undetermined coefficients is based on the idea that the type of a particular solution $y_p(x)$ that satisfies $ay_p'' + by_p' + cy_p = f(x)$ can be determined from the type of $f(x)$. We discuss in detail how this works for different types of $f(x)$.

To streamline the presentation define the differential operator L which acts on functions $h(x)$ and gives

$$L[h] = ah'' + bh' + ch.$$

Finding a particular solution y_p of $ay_p'' + by_p' + cy_p = f(x)$ means to find a function that satisfies $L[y_p] = f$.

Polynomial $f(x)$. Suppose that $f(x)$ is a polynomial of degree n and write

$$f(x) = c_nx^n +_{n-1} x^{n-1} + \cdots + c_1x + c_0.$$

If Q is a polynomial of degree up to n then $L[Q]$ is also a polynomial of degree up to n and therefore by matching the coefficients of $L[Q]$ to those of f we can determine Q. To find the polynomial $y_p = Q$, write

$$y_p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

substitute into Eq. (3.10), and solve the resulting equations for the unknown coefficients, $a_0, \ldots, a_n$.

Example 3.24. Consider the equation $y'' - y = 2 - x^2$. Since $f(x) = 2 - x^2$ is a polynomial of degree 2, to find a particular solution y_p we write the most general degree-2 polynomial

$$y_p(x) = a_2 x^2 + a_1 x + a_0.$$

Substituting y_p into the equation $y'' - y = 2 - x^2$, we find

$$y_p'' - y_p = 2a_2 - (a_2 x^2 + a_1 x + a_0) = -a_2 x^2 - a_1 x + (2a_2 - a_0) = -x^2 + 2.$$

Equating coefficients of the same powers of x we obtain the linear system of equations $-a_2 = -1$, $-a_1 = 0$, and $2a_2 - a_0 = 2$, with solution $a_2 = 1$, $a_1 = a_0 = 0$. Therefore, a particular solution is $y_p(x) = x^2$.

The general solution of the associated homogeneous equation $y'' - y = 0$ is

$$y_h(t) = c_1 e^x + c_2 e^{-x}.$$

Therefore, the general solution of the non-homogeneous equation is

$$y(t) = c_1 e^x + c_2 e^{-x} + x^2.$$

Exponential $f(x)$. Suppose that $f(x) = P(x) e^{\lambda x}$, where $P(x)$ is a polynomial of degree n. This case also includes the case where $P(x)$ is a polynomial of degree 0, that is, a constant. If $Q(x)$ is a polynomial of degree n, then $L[Q e^{\lambda x}] = R Q e^{\lambda x}$, where $R(x)$ is a polynomial of degree up to n.

(i) If λ is not one of the roots of the auxiliary equation, try $y_p(x) = Q(x) e^{\lambda x}$ where $\deg Q = \deg P$.
(ii) If the auxiliary equation has two real, distinct roots and λ equals one of these roots try $y_p(x) = Q(x) x e^{\lambda x}$ where $\deg Q = \deg P$.
(iii) If the auxiliary equation has a double real root r and $\lambda = r$ try $y_p(x) = Q(x) x^2 e^{\lambda x}$ where $\deg Q = \deg P$.

Example 3.25. Consider the equation $y'' - y = 2e^{2x}$. The auxiliary equation has roots ± 1 and $\lambda = 2$ is not one of them. Moreover, $P(x) = 2$ with $\deg P = 0$. Therefore, we consider the particular solution $y_p(x) = A e^{2x}$.

Substituting into the equation we find

$$y_p'' - y_p = 4A e^{2x} - A e^{2x} = 3A e^{2x} = 2e^{2x}.$$

Therefore, $A = 2/3$ and the particular solution is

$$y_p(x) = \frac{2}{3} e^{2x},$$

and the general solution is

$$y(t) = c_1 e^x + c_2 e^{-x} + \frac{1}{2} e^{2x}.$$

Example 3.26. Consider the equation $y'' - y = 2e^x$. We want to find again a particular solution but we notice that $\lambda = 1$ is one of the roots of the auxiliary equation.

First, notice what goes wrong if we try a particular solution of the form $y_p = A e^x$, where since $\deg P = 0$ we take $Q(x)$ constant. Even without any computations it should be clear that since e^x is one of the solutions of the associated homogeneous equation, Ae^x cannot be a solution of the non-homogeneous equation. If we try to force the matter, we find that

$$y_p'' - y_p = Ae^x - Ae^x = 0 \neq 2e^x.$$

Instead, consider $y_p = Axe^x$. We have $y_p' = A(1+x)e^x$ and $y_p'' = A(2+x)e^x$. Therefore,

$$y_p'' - y_p = A(2+x)e^x - Axe^x = 2Ae^x = 2e^x,$$

implying $A = 1$. The particular solution is

$$y_p(x) = xe^x,$$

and the general solution is

$$y(t) = c_1 e^x + c_2 e^{-x} + xe^x.$$

Example 3.27. Consider the equation $y'' - 2y' + y = 2e^x$. The auxiliary equation has the double root $r = 1$ and we also have $\lambda = 1$. In this case, try $y_p = Ax^2e^x$. We compute $y_p' = A(2x + x^2)e^x$ and $y_p'' = A(2 + 4x + x^2)e^x$. Therefore,

$$y_p'' - 2y_p' + y_p = A(2 + 4x + x^2 - 4x - 2x^2 + x^2)e^x = 2Ae^x = 2e^x,$$

implying $A = 1$. The general solution is

$$y(t) = c_1 e^x + c_2 xe^x + x^2e^x = (x^2 + c_2 x + c_1)e^x.$$

Trigonometric $f(x)$. Suppose that $f(x) = P_1(x)\cos(\beta x) + P_2(x)\sin(\beta x)$ where $P_1(x)$, $P_2(x)$ are polynomials with degrees $n_1 = \deg P_1$ and $n_2 = \deg P_2$. Let $n = \max(n_1, n_2)$. We have two cases.

(i) If the auxiliary equation does not have imaginary roots $\pm i\beta$ then consider the particular solution
$$y_p(x) = Q_1(x)\cos(\beta x) + Q_2(x)\sin(\beta x),$$
where $Q_1(x)$, $Q_2(x)$ are polynomials with $\deg Q_1 = \deg Q_2 = n$.

(ii) If the auxiliary equation has imaginary roots $\pm i\beta$ then consider the particular solution
$$y_p(x) = Q_1(x)x\cos(\beta x) + Q_2(x)x\sin(\beta x),$$
where $Q_1(x)$, $Q_2(x)$ are polynomials with $\deg Q_1 = \deg Q_2 = n$.

Remark 3.28. If the auxiliary equation has imaginary roots $\pm i\beta$ then the general solution of the associated homogeneous equation is $c_1\cos(\beta x) + c_2\sin(\beta x)$.

Remark 3.29. It is useful to think here in terms of complex exponentials instead of sine and cosine. The functions $\cos(\beta x)$ and $\sin(\beta x)$ can be written as linear combinations of $e^{i\beta x}$ and $e^{-i\beta x}$. Then we are effectively in the case of "exponential" $f(x)$ discussed earlier. If the auxiliary equation has roots $\pm i\beta$ then the general solution of the associated homogeneous equation is also a linear combination of $e^{i\beta x}$ and $e^{-i\beta x}$ and therefore we multiply the standard $y_p(x)$ by x as we also did in the case of "exponential" $f(x)$.

Example 3.30. Consider $y'' + 2y' + 2y = 5\cos x$. The auxiliary equation $r^2 + 2r + 2$ has roots $-1 \pm i$ and here $\lambda = \pm i$ does not coincide with these roots. Therefore, we try the particular solution

$$y_p(x) = A\cos x + B\sin x.$$

We compute $y_p' = -A\sin x + B\cos x$, $y_p'' = -A\cos x - B\sin x$. Then

$$y_p'' + 2y_p' + 2y_p = \cos x(A + 2B) + \sin x(B - 2A) = 5\cos x.$$

Therefore, we have the equations $A + 2B = 5$ and $B - 2A = 0$ with solutions $A = 1$, $B = 2$. The general solution is

$$y(x) = c_1 e^{-x}\cos x + c_2 e^{-x}\sin x + \cos x + 2\sin x.$$

Combined Exponential and Trigonometric $\boldsymbol{f(x)}$. Suppose that

$$f(x) = e^{\alpha x}\big(P_1(x)\cos(\beta x) + P_2(x)\sin(\beta x)\big).$$

Observe that sine and cosine both depend on βx with the same β, and there is only one exponential $e^{\alpha x}$. In terms of complex exponentials we have here $e^{(\alpha\pm i\beta)x}$.

(i) If the auxiliary equation does not have complex roots $\alpha \pm i\beta$ then consider the particular solution

$$y_p(x) = e^{\alpha x}\big(Q_1(x)\cos(\beta x) + Q_2(x)\sin(\beta x)\big),$$

where $Q_1(x)$, $Q_2(x)$ are polynomials with $\deg Q_1 = \deg Q_2 = \max(p_1, p_2)$.

(ii) If the auxiliary equation has complex roots $\alpha \pm i\beta$ then consider the particular solution

$$y_p(x) = xe^{\alpha x}\big(Q_1(x)\cos(\beta x) + Q_2(x)\sin(\beta x)\big),$$

where $Q_1(x)$, $Q_2(x)$ are polynomials with $\deg Q_1 = \deg Q_2 = \max(p_1, p_2)$.

Linear combination $\boldsymbol{f(x) = k_1 f_1(x) + k_2 f_2(x)}$. Suppose that

$$f(x) = k_1 f_1(x) + k_2 f_2(x),\ \ k_1, k_2 \in \mathbf{R},$$

and that it is possible to use the techniques described until now in this section to find a particular solution $y_{p,1}$ for the problem $ay'' + by' + cy = f_1(x)$ and a particular solution $y_{p,2}$ for the problem $ay'' + by' + cy = f_2(x)$. Then the function

$$y_p = k_1 y_{p,1} + k_2 y_{p,2}$$

is a particular solution for the problem $ay'' + by' + cy = f(x)$.

Example 3.31. Consider the equation $y'' - y = 2 - x^2 + 4e^x$. In Example 3.24 we computed the particular solution $y_{p,1}(x) = x^2$ for $f_1(x) = 2 - x^2$. In Example 3.26 we computed the particular solution $y_{p,2}(x) = xe^x$ for $f_2(x) = 2e^x$. Since $f(x) = f_1(x) + 2f_2(x)$ the corresponding particular solution is

$$y_p(x) = y_{p,1}(x) + 2y_{p,2}(x) = x^2 + 2xe^x.$$

Example 3.32. Consider the equation

$$y'' - 2y' + 2y = 2e^x\cos x - 3xe^x\cos(2x).$$

Note that the two trigonometric functions correspond to different complex exponentials: $e^x\cos x$ corresponds to $e^{(1\pm i)x}$, while $e^x\cos x$ corresponds to $e^{(1\pm 2i)x}$. This means that we must combine two particular solutions of the type described for combined exponential and trigonometric $f(x)$.

First, we consider $f_1(x) = 2e^x \cos x$. Since the roots of the auxiliary equation are $1 \pm i$ we try the particular solution

$$y_{p,1}(x) = xA_1 e^x \cos x + xB_1 e^x \sin x.$$

Then, we consider $f_2(x) = 3xe^x \cos(2x)$ and we try the particular solution

$$y_{p,2}(x) = (A_2 + A_3 x)e^x \cos(2x) + (B_2 + B_3 x)e^x \sin(2x).$$

We can either find $y_{p,1}(x)$ and $y_{p,2}(x)$ independently and then combine them to obtain the particular solution $y_p(x) = y_{p,1}(x) - y_{p,2}(x)$ or we can directly consider the particular solution

$$y_p(x) = xA_1 e^x \cos x + xB_1 e^x \sin x - (A_2 + A_3 x)e^x \cos(2x) - (B_2 + B_3 x)e^x \sin(2x).$$

We proceed using the latter approach. A long computation gives $A_1 = 0$, $B_1 = 1$, $A_2 = 0$, $A_3 = -1$, $B_2 = 4/3$, $B_3 = 0$. The particular solution is

$$y_p(x) = xe^x \sin x + xe^x \cos(2x) - \frac{4}{3}e^x \sin(2x),$$

and the general solution is

$$y(x) = c_1 e^x \cos x + c_2 e^x \sin x + xe^x \sin x + xe^x \cos(2x) - \frac{4}{3}e^x \sin(2x).$$

3.3 Solution Method for Higher-Order Linear Differential Equations with Constant Coefficients

We now briefly consider n-th order linear homogeneous differential equations with constant coefficients, that is, equations of the form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0,$$

where $a_0, a_1, \ldots, a_n \in \mathbf{R}$ and $a_0 \neq 0$. The solution method for n-th order equations is based on the idea that the general solution is the linear combination of n linearly independent solutions which can be read off the roots of the auxiliary equation. The method is summarized below.

Solution method for n-th order linear homogeneous differential equations with constant coefficients

Consider the homogeneous linear n-th order differential equation with constant coefficients

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0, \quad a_0, \ldots, a_n \in \mathbf{R},\ a_n \neq 0.$$

Determine the roots of the auxiliary equation

$$a_n r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0 = 0,$$

together with their multiplicity.[a]

Then construct a set S of n linearly independent solutions $y_1, y_2, \ldots, y_n$ of the given equation in the following way.

(i) For each real root r of multiplicity $k \geq 1$, include in S the k linearly independent solutions $e^{rx}, xe^{rx}, \ldots x^{k-1}e^{rx}$.
(ii) For each pair of complex conjugate roots $\alpha \pm i\beta$ of multiplicity $k \geq 1$, include in S the $2k$ linearly independent solutions $x^j e^{\alpha x}\cos(\beta x)$, and $x^j e^{\alpha x}\sin(\beta x)$, with $j = 0, 1, \ldots, k-1$.

The general solution $y(x)$ is the linear combination of all solutions $y_1, y_2, \ldots, y_n$ contained in S, that is,

$$y(x) = \sum_{j=1}^{n} c_j y_j(x), \; c_j \in \mathbf{R}.$$

[a] Since the auxiliary equation is a polynomial equation of degree n, it has n roots counted with their multiplicity.

Example 3.33. Consider the equation

$$\begin{aligned} y^{(11)} - 6y^{(10)} + 11y^{(9)} + 14y^{(8)} - 117y^{(7)} + 302y^{(6)} - 487y^{(5)} \\ + 562y^{(4)} - 484y^{(3)} + 304y'' - 124y' + 24y = 0, \end{aligned}$$

with auxiliary equation

$$24 - 124r + 304r^2 - 484r^3 + 562r^4 - 487r^5 + 302r^6 - 117r^7 + 14r^8 + 11r^9 - 6r^{10} + r^{11} = 0.$$

The polynomial in the last equation factorizes as

$$(r-1)^3(r-2)(r+3)(r^2-2r+2)^2(r^2+1) = 0,$$

and the roots and corresponding solutions are:

(i) 1 with multiplicity 3; the corresponding solutions are e^x, xe^x, x^2e^x.
(ii) 2 with multiplicity 1; the corresponding solution is e^{2x}.
(iii) -3 with multiplicity 1; the corresponding solution is e^{-3x}.
(iv) $1 \pm i$ with multiplicity 2; the corresponding solutions are $e^x \cos x$, $e^x \sin x$, $xe^x \cos x$, $xe^x \sin x$.
(v) $\pm i$ with multiplicity 1; the corresponding solutions are $\cos x$, $\sin x$.

Therefore, the general solution is

$$\begin{aligned} y(x) &= c_1e^x + c_2xe^x + c_3x^2e^x + c_4e^{2x} + c_5e^{-3x} \\ &\quad + c_6e^x\cos x + c_7e^x\sin x + c_8xe^x\cos x + c_9xe^x\sin x \\ &\quad + c_{10}\cos x + c_{11}\sin x. \end{aligned}$$

3.4 Boundary Value Problems

In these notes we focus on initial value problems, since they are related to dynamical systems. However, there is another type of problems involving differential equations, called *boundary value problems*. Boundary value problems arise usually in the context of solving a partial differential equation.

In boundary value problems we find solutions to a differential equation that satisfy certain conditions at the boundary points of the interval in which we solve the equation. To make this discussion more concrete, suppose we want to solve a second order equation in an interval $[a, b]$, that is,

$$y'' = f(x, y, y'),\ a \le x \le b.$$

A common type of boundary conditions are called *Dirichlet boundary conditions*. In this case, we specify the values $y(a)$, $y(b)$ of the unknown function $y(x)$ at $x = a$ and $x = b$. In *Neumann boundary conditions* we specify the values of $y'(a)$ and $y'(b)$. In *mixed boundary conditions*, we specify the values of $y(a)$ and $y'(b)$, or the values of $y'(a)$ and $y(b)$. And, in *periodic boundary conditions* we do not specify the values of y and y' but we ask that $y(a) = y(b)$ and $y'(a) = y'(b)$ — this is what one should do if x is an angle and $a = 0$, $b = 2\pi$.

There is a rich theory concerning the solutions of boundary value problems. Here we want to give a flavor of some of the intricacies of boundary value problems by looking at a specific example motivated by an elementary problem in quantum mechanics.

Consider the second order equation

$$-y'' = Ey,\ 0 \le x \le 1, \tag{3.11}$$

with the Dirichlet boundary conditions $y(0) = y(1) = 0$ and the additional requirement that $y(x)$ is not identically zero for $x \in [0, 1]$.

Remark 3.34. Equation (3.11) is the Schrödinger equation for a particle constrained inside a one-dimensional box — the slightly more realistic three-dimensional box requires to solve the same equation three times, once for each coordinate axis. The constant E is proportional to the energy of the particle. ❞

Equation (3.11) is a linear second order equation with constant coefficients. The auxiliary equation is $r^2 = -E$.

For $E < 0$, the general solution is

$$y(x) = c_1 e^{\sqrt{-E}x} + c_2 e^{-\sqrt{-E}x}.$$

If we try to impose the boundary conditions $y(0) = y(1) = 0$ we find

$$c_1 + c_2 = 0,\ c_1 e^{2\sqrt{-E}} + c_2 = 0,$$

giving $c_1 = c_2 = 0$, which must be rejected.

For $E = 0$, the equation becomes $y'' = 0$ and the general solution is $y(x) = c_1 + c_2 x$. The boundary conditions $y(0) = y(1) = 0$ again give $c_1 = c_2 = 0$.

Finally, for $E > 0$, the general solution is

$$y(x) = c_1 \cos(\sqrt{E}x) + c_2 \sin(\sqrt{E}x).$$

In this case, the boundary condition $y(0) = 0$ gives $c_1 = 0$, but the second boundary condition $y(1) = 0$ gives $\sin(\sqrt{E}) = 0$, that is, $\sqrt{E} = n\pi$, $n = 1, 2, 3, \ldots$.

Reviewing the three cases for E analyzed above, we note that boundary value problems may not always have solutions that satisfy all the requirements. In particular, there are no solutions for $E \le 0$ and there are solutions for $E > 0$ only when $E = n^2\pi^2$, $n = 1, 2, 3, \ldots$. Moreover, in the case $E = n^2\pi^2$ the solution is not unique since $y(x) = c_2 \sin(n\pi x)$ is a solution to the given boundary value problem for any $c_2 \in \mathbf{R}$.

Remark 3.35. Uncovering that Eq. (3.11) admits solutions only when $E = n^2\pi^2$, $n = 1, 2, 3, \ldots$, is the easiest example of *quantization* in physical systems, that is, the observation that the energy of certain physical systems can take values only in a discrete set and is not always a continuous quantity. ❞

Chapter 4

Linear Dynamical Systems in Two Dimensions

In the previous chapter we have seen some examples of the dynamics of planar systems and some simple cases in which we can understand parts of the dynamics either by transforming the planar system to a second order equation or by solving the reduced equation. In this and subsequent chapters we will start a more systematic study of planar systems by looking at the main features of their dynamics. We start in this chapter with the study of planar linear systems and their classification.

4.1 Planar Linear Systems

Definition 4.1. A *planar linear system* has the form

$$\boldsymbol{x}' = A(t)\boldsymbol{x} + \boldsymbol{f}(t), \tag{4.1}$$

where the dependent variable $\boldsymbol{x}(t)$ is a function of the independent variable t and takes values in $\mathbf{R}^2$. Moreover, $A(t)$ is a 2×2 matrix and $\boldsymbol{f}(t)$ is a vector. Both $A(t)$ and $\boldsymbol{f}(t)$ depend continuously on t, for t in an interval $U \subseteq \mathbf{R}$.

In this chapter we only consider planar linear systems of the form

$$\boldsymbol{x}' = A\boldsymbol{x}, \tag{4.2}$$

where A is a constant matrix and $\boldsymbol{f}(t) \equiv 0$. Writing

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

Eq. (4.2) corresponds to the system of equations

$$\begin{aligned} x_1' &= ax_1 + bx_2, \\ x_2' &= cx_1 + dx_2, \end{aligned}$$

where the unknown functions are the components $x_1(t), x_2(t)$ of $\boldsymbol{x}(t)$.

Even though in this chapter we obtain explicit expressions for the solutions of the planar linear system $\boldsymbol{x}' = A\boldsymbol{x}$, our focus is on understanding the dynamics of the system. As t varies, each solution $\boldsymbol{x}(t)$ of the system $\boldsymbol{x}' = A\boldsymbol{x}$ traces a curve on $\mathbf{R}^2$. We want to understand the properties of these curves and to visualize them on the plane, obtaining a "portrait" of the dynamics of the system.

As we discuss in Remark 4.2, planar linear systems are closely connected to second-order linear differential equations with constant coefficients, the topic of Chapter 3. In particular, we can solve a second-order linear differential equation with constant coefficient by solving an associated planar linear system. However, our interest in understanding and classifying linear systems originates in the study of the dynamics of nonlinear systems near their equilibria. We will consider this problem in Chapter 6.

The general theory of linear systems that will be discussed in Chapter 5 shows that the general solution of Eq. (4.2) has the form

$$\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t),\ \ c_1, c_2 \in \mathbf{R},$$

where $\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)$ are vector-valued linearly independent solutions of Eq. (4.2). The linear independence of two solutions $\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)$ turns out to be equivalent to the condition that $\boldsymbol{x}_1(0), \boldsymbol{x}_2(0)$ are linearly independent vectors in $\mathbf{R}^2$. Therefore, to solve Eq. (4.2) it is sufficient to obtain two linearly independent solutions.

Then the main idea is to consider solutions of the form

$$\boldsymbol{x}(t) = e^{rt}\boldsymbol{u},$$

where $\boldsymbol{u}$ is a constant vector. Substituting into Eq. (4.2) we obtain the equation

$$re^{rt}\boldsymbol{u} = Ae^{rt}\boldsymbol{u}.$$

Since $e^{rt} \neq 0$, the last equation simplifies to

$$(A - rI)\boldsymbol{u} = \mathbf{0}, \tag{4.3}$$

where I is the 2×2 identity matrix. Equation (4.3) is satisfied for $\boldsymbol{u} \neq \mathbf{0}$ when the characteristic polynomial

$$p(r) = \det(A - rI), \tag{4.4}$$

vanishes. The characteristic polynomial for the 2×2 matrix A is given by

$$p(r) = (r - a)(r - d) - bc = r^2 - (a + d)r + (ad - bc) = r^2 - Tr + D, \tag{4.5}$$

where

$$T = a + d \text{ and } D = ad - bc$$

are respectively the trace of A and its determinant.

We distinguish three cases.

(i) If $p(r)$ has two distinct real roots r_1, r_2 then the corresponding eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2$ are real and linearly independent. Then we directly get the linearly independent solutions $\boldsymbol{x}_1(t) = e^{r_1 t}\boldsymbol{u}_1$, $\boldsymbol{x}_2(t) = e^{r_2 t}\boldsymbol{u}_2$.

(ii) If $p(r)$ has a pair of complex conjugate roots $r = \alpha + i\beta$ and $\bar{r} = \alpha - i\beta$ with $\beta \neq 0$ then the corresponding eigenvectors form the complex conjugate pair $\boldsymbol{u}$ and $\bar{\boldsymbol{u}}$. For this case, we discuss how to obtain two real linearly independent solutions by considering the real and imaginary parts of the complex valued solution $e^{rt}\boldsymbol{u}$.

(iii) If $p(r)$ has a double real root r and a corresponding eigenvector is $\boldsymbol{u}_1$, then one of the linearly independent solutions is $e^{rt}\boldsymbol{u}_1$. A second independent solution is given below and elaborated in Section 4.4.

We give below a summary of the solution method for planar linear systems.

Solution method for planar linear systems

The general solution of Equation (4.2) has the form

$$\boldsymbol{x}(t) = c_1 \boldsymbol{x}_1(t) + c_2 \boldsymbol{x}_2(t), \quad c_1, c_2 \in \mathbf{R}.$$

To determine $\boldsymbol{x}_1(t)$, $\boldsymbol{x}_2(t)$ consider the characteristic polynomial $p(r) = \det(A - rI)$ of the matrix A whose roots are the eigenvalues of A.

(i) If A has two real distinct eigenvalues r_1, r_2 with corresponding eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ then

$$\boldsymbol{x}_1(t) = e^{r_1 t}\boldsymbol{u}_1, \ \boldsymbol{x}_2(t) = e^{r_2 t}\boldsymbol{u}_2.$$

(ii) If A has two complex conjugate eigenvalues $r = \alpha + i\beta$ and $\overline{r} = \alpha - i\beta$, $\alpha, \beta \in \mathbf{R}, \beta \neq 0$, with corresponding eigenvectors $\boldsymbol{u} = \boldsymbol{a} + i\boldsymbol{b}$, $\overline{\boldsymbol{u}} = \boldsymbol{a} - i\boldsymbol{b}$, $\boldsymbol{a}, \boldsymbol{b} \in \mathbf{R}^2$, then

$$\begin{aligned}
\boldsymbol{x}_1(t) &= \mathrm{Re}(e^{rt}\boldsymbol{u}) = e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}), \\
\boldsymbol{x}_2(t) &= \mathrm{Im}(e^{rt}\boldsymbol{u}) = e^{\alpha t}(\cos(\beta t)\boldsymbol{b} + \sin(\beta t)\boldsymbol{a}).
\end{aligned}$$

(iii) If A has a double real eigenvalue r, then we consider two subcases.

(a) If A is a diagonal matrix of the form $A = rI$, then

$$\boldsymbol{x}_1(t) = e^{rt}\boldsymbol{e}_1, \ \boldsymbol{x}_2(t) = e^{rt}\boldsymbol{e}_2.$$

In particular, the general solution is $\boldsymbol{x}(t) = e^{rt}\boldsymbol{c}$, where $\boldsymbol{c} = \langle c_1, c_2 \rangle \in \mathbf{R}^2$.

(b) Otherwise, let $\boldsymbol{u}_1$ be an eigenvector of A, and $\boldsymbol{u}_2$ be *any* linearly independent vector in $\mathbf{R}^2$. Then $(A - rI)\boldsymbol{u}_2 = \lambda \boldsymbol{u}_1$ for some $\lambda \in \mathbf{R}$, and

$$\boldsymbol{x}_1(t) = e^{rt}\boldsymbol{u}_1, \ \boldsymbol{x}_2(t) = e^{rt}(\lambda t \boldsymbol{u}_1 + \boldsymbol{u}_2).$$

In subsequent sections, we show that the given expressions are indeed solutions but we do more than that. First, we obtain these solutions by defining new coordinates on $\mathbf{R}^2$ in which the system of differential equations $\boldsymbol{x}' = A\boldsymbol{x}$ splits to two independent first order differential equations that can be directly solved with the methods discussed in Chapter 1. Second, we visualize the dynamics in the new coordinates and then use the results to also visualize the dynamics in the original coordinates.

Remark 4.2. If the three cases and the discussion above remind the reader of the corresponding solution method for linear homogeneous second-order differential equations with constant coefficients this is because there is a close connection between them and the planar linear systems considered here. Specifically, every linear homogeneous second-order differential equations with constant coefficients can be written as a linear system of the form in Eq. (4.2). To see how this works, consider the second-order equation

$$y'' + py' + qy = 0, \ p, q \in \mathbf{R},$$

where y depends on the independent variable t, and define $x_1 = y$ and $x_2 = y'$. Then $x_1' = y' = x_2$ and $x_2' = y'' = -py' - qy = px_2 - qx_1$. Written in matrix form we get

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q & -p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Therefore, all linear homogeneous second-order differential equations with constant coefficients can be expressed as planar linear systems of the form in Eq. (4.2). Moreover, notice that the characteristic equation for the matrix in the last equation is $r^2+pr+q=0$, that is, it is exactly the auxiliary equation for the given second-order equation. We come back to the relation between second-order linear equations and planar linear systems in Section 5.5. ❞

4.2 Real Distinct Eigenvalues

When A has two real distinct eigenvalues $r_1 \neq r_2$, then the corresponding eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ are linearly independent. This means that the solutions $\boldsymbol{x}_1(t) = e^{r_1 t}\boldsymbol{u}_1$, $\boldsymbol{x}_2(t) = e^{r_2 t}\boldsymbol{u}_2$, are linearly independent since at $t=0$ we have $\boldsymbol{x}_1(0) = \boldsymbol{u}_1$, $\boldsymbol{x}_2(0) = \boldsymbol{u}_2$. Therefore, we can directly conclude that the general solution is

$$\boldsymbol{x}(t) = c_1 e^{r_1 t}\boldsymbol{u}_1 + c_2 e^{r_2 t}\boldsymbol{u}_2,\ c_1, c_2 \in \mathbf{R}. \tag{4.6}$$

We now obtain Eq. (4.6) without referring to the general theory that we develop in Chapter 5, but using only a combination of Linear Algebra and results from Chapter 1.

Given the linearly independent eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ of A, define the matrix

$$U = [\boldsymbol{u}_1 | \boldsymbol{u}_2],$$

having as columns the two eigenvectors. Since $\boldsymbol{u}_1, \boldsymbol{u}_2$ are linearly independent, the matrix U is non-singular, that is, $\det U \neq 0$.

Define new coordinates $\boldsymbol{z} = \langle z_1, z_2 \rangle$ on $\mathbf{R}^2$, in which the system of differential equations attains a simpler form, by

$$\boldsymbol{x} = U\boldsymbol{z} \Leftrightarrow \boldsymbol{z} = U^{-1}\boldsymbol{x}. \tag{4.7}$$

Then, $\boldsymbol{z}$ satisfies the differential equation

$$\boldsymbol{z}' = U^{-1}\boldsymbol{x}' = U^{-1}A\boldsymbol{x} = U^{-1}AU\boldsymbol{z} = R\boldsymbol{z}. \tag{4.8}$$

The reason that the equation $\boldsymbol{z}' = R\boldsymbol{z}$ is simpler than the original equation $\boldsymbol{x}' = A\boldsymbol{x}$ is that R is a diagonal matrix. In particular, we have the following standard result from Linear Algebra.

Proposition 4.3. $U^{-1}AU = R := \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}.$

Remark 4.4. Before giving the proof of Proposition 4.3 we recall a basic fact concerning the columns of a matrix that will be used frequently in computations below: if we denote by $\boldsymbol{e}_1 = \langle 1, 0 \rangle$ and $\boldsymbol{e}_2 = \langle 0, 1 \rangle$ the unit vectors along the coordinate axes, then for an arbitrary 2×2 matrix M, $M\boldsymbol{e}_j$ gives the j-th column of M for $j = 1, 2$. Then, the j-th column of U is

$$U\boldsymbol{e}_j = \boldsymbol{u}_j \Leftrightarrow \boldsymbol{e}_j = U^{-1}\boldsymbol{u}_j.$$

❞

Proof of Proposition 4.3. We compute that the j-th column of $U^{-1}AU$ is

$$U^{-1}AU\boldsymbol{e}_j = U^{-1}A\boldsymbol{u}_j = U^{-1}(r_j\boldsymbol{u}_j) = r_j U^{-1}\boldsymbol{u}_j = r_j\boldsymbol{e}_j.$$

This shows that the j-th column of $U^{-1}AU$ has the eigenvalue r_j in the j-th position and zeros everywhere else, that is, $U^{-1}AU$ is the diagonal matrix

$$R = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}.$$

☑

Equation (4.8), $\boldsymbol{z}' = R\boldsymbol{z}$, can be written as the two first order equations $z_1' = r_1 z_1$ and $z_2' = r_2 z_2$. These two equations have the general solutions $z_1(t) = c_1 e^{r_1 t}$ and $z_2(t) = c_2 e^{r_2 t}$. That is, the general solution of Eq. (4.8) is

$$\boldsymbol{z}(t) = \begin{bmatrix} c_1 e^{r_1 t} \\ c_2 e^{r_2 t} \end{bmatrix} = c_1 e^{r_1 t} \boldsymbol{e}_1 + c_2 e^{r_2 t} \boldsymbol{e}_2.$$

Expressing this solution in terms of the original coordinates $\boldsymbol{x} = \langle x_1, x_2 \rangle$ we obtain

$$\boldsymbol{x}(t) = U\boldsymbol{z}(t) = c_1 e^{r_1 t} U\boldsymbol{e}_1 + c_2 e^{r_2 t} U\boldsymbol{e}_2 = c_1 e^{r_1 t} \boldsymbol{u}_1 + c_2 e^{r_2 t} \boldsymbol{u}_2,$$

which is exactly Eq. (4.6).

Summarizing, using the new coordinates $\boldsymbol{z} = \langle z_1, z_2 \rangle$, the original system $\boldsymbol{x}' = A\boldsymbol{x}$ has been reduced to two independent linear differential equations — one for z_1 and one for z_2 — which can be easily solved to give the solution $\boldsymbol{z}(t)$ and thus the solution $\boldsymbol{x}(t) = U\boldsymbol{z}(t)$.

To understand and visualize the dynamics of the linear system, we consider different cases based on the signs of the eigenvalues r_1, r_2. We consider three cases: distinct positive eigenvalues, distinct negative eigenvalues, and eigenvalues with opposite signs. The solution in Eq. (4.6) applies also to the case where one of the eigenvalues is zero, but we do not analyze this case in subsequent sections. However, the methods we develop can also be applied to the case of one zero eigenvalue.

4.2.1 Distinct Positive Eigenvalues: Unstable Node

Consider first the case where the matrix A has two real distinct positive eigenvalues r_1, r_2. In this case, the equilibrium at the origin is called an *unstable node*. Without loss of generality assume that $0 < r_1 < r_2$.

We first consider the shape of the solution curves using the coordinates $\boldsymbol{z} = \langle z_1, z_2 \rangle$. Consider the parametric form of the solution curve

$$\boldsymbol{z}(t) = \begin{bmatrix} c_1 e^{r_1 t} \\ c_2 e^{r_2 t} \end{bmatrix}.$$

Since $0 < r_1 < r_2$, we find that for $c_1 \neq 0$ we have $\lim_{t\to\infty} |z_1(t)| = \infty$ and, similarly for $c_2 \neq 0$ that $\lim_{t\to\infty} |z_2(t)| = \infty$. Therefore, all solution curves — except the equilibrium solution at the origin — move away from the origin as t increases. We also have $\lim_{t\to-\infty} z_1(t) = \lim_{t\to-\infty} z_2(t) = 0$, therefore, all solution curves move toward the origin as t decreases. Additionally, as $t \to -\infty$ we have $r_2 t < r_1 t$ and thus for very negative t we get $e^{r_2 t} \ll e^{r_1 t}$, which implies that the solution curve $\boldsymbol{z}(t)$ comes closer to the z_1 axis than the z_2 axis and, therefore, it tends to become tangent to the z_1 axis as it approaches the origin.

For the solution curves in the original coordinates $\boldsymbol{x} = \langle x_1, x_2 \rangle$ we can use the equation $\boldsymbol{x} = U\boldsymbol{z}$ to transform the solution curves that we obtained in the $\boldsymbol{z}$ coordinates. Since $U\boldsymbol{e}_1 = \boldsymbol{u}_1$ and $U\boldsymbol{e}_2 = \boldsymbol{u}_2$, the solution curves become tangent to the direction of $\boldsymbol{u}_1$ as they approach the origin when $t \to -\infty$, and move away from the origin when $t \to \infty$.

How to draw the phase portrait for an unstable node

Assume that A has eigenvalues $0 < r_1 < r_2$.

(i) Find the corresponding eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.
(ii) Draw solution curves moving away from the origin along the straight half-lines in

the directions $\pm\boldsymbol{u}_1$ and $\pm\boldsymbol{u}_2$.

(iii) Draw solution curves that are tangent to the $\pm\boldsymbol{u}_1$ direction[a] at the origin and that move away from the origin.

[a] This is the direction corresponding to the eigenvalue with the smaller absolute value.

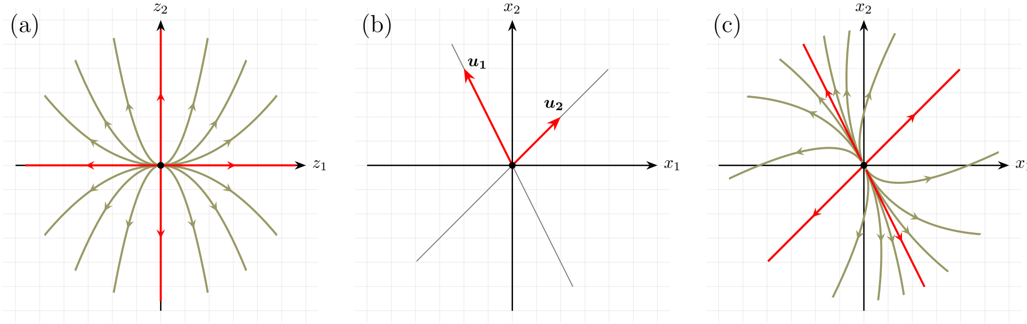


Figure 4.1: (a) Solution curves for an unstable node in the z-plane with $r_1 = 6$ and $r_2 = 3$. (b) Eigenvectors for the problem $\boldsymbol{x}' = A\boldsymbol{x}$ in Example 4.5. (c) Solution curves in the original coordinates $\boldsymbol{x}$ for Example 4.5. The solution curves shown in this picture are the same as the curves shown in (a), transformed using the matrix $U = [\boldsymbol{u}_1|\boldsymbol{u}_2]$.

Example 4.5. Consider the planar linear system $\boldsymbol{x}' = A\boldsymbol{x}$ with

$$A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}.$$

The matrix A has eigenvalues $r_1 = 3$ and $r_2 = 6$ with corresponding eigenvectors $\boldsymbol{u}_1 = \langle -1, 2\rangle$ and $\boldsymbol{u}_2 = \langle 1, 1\rangle$. The solution curves in the $\boldsymbol{z}$ plane are shown in Fig. 4.1(a), while Fig. 4.1(b) shows the two eigenvectors in the $\boldsymbol{x}$ plane. The solution curves in the $\boldsymbol{x}$ plane are shown in Fig. 4.1(c).

4.2.2 Distinct Negative Eigenvalues: Stable Node

Consider now the case where the matrix A has two real distinct negative eigenvalues r_1, r_2. In this case, the equilibrium at the origin is called a *stable node*. Without loss of generality assume that $r_2 < r_1 < 0$.

We consider the shape of the solution curves in the $\boldsymbol{z}$ plane. Consider the parametric form of the solution curve

$$\boldsymbol{z}(t) = \begin{bmatrix} c_1 e^{r_1 t} \\ c_2 e^{r_2 t} \end{bmatrix}.$$

Since $r_2 < r_1 < 0$, we find that for $c_1 \neq 0$ we have $\lim_{t\to-\infty} |z_1(t)| = \infty$ and, similarly for $c_2 \neq 0$ that $\lim_{t\to-\infty} |z_2(t)| = \infty$. Therefore, all solution curves — except the equilibrium solution at the origin — move away from the origin as t decreases. We also have $\lim_{t\to\infty} z_1(t) = \lim_{t\to\infty} z_2(t) = 0$, therefore, all solution curves move toward the origin as t increases. Additionally, as $t \to \infty$ we have $r_2 t < r_1 t$ and thus for very large positive t we get $e^{r_2 t} \ll e^{r_1 t}$, which implies that the solution curve $\boldsymbol{z}(t)$ comes closer to the z_1 axis than the z_2 axis and, therefore, it tends to become tangent to the z_1 axis as it approaches the origin.

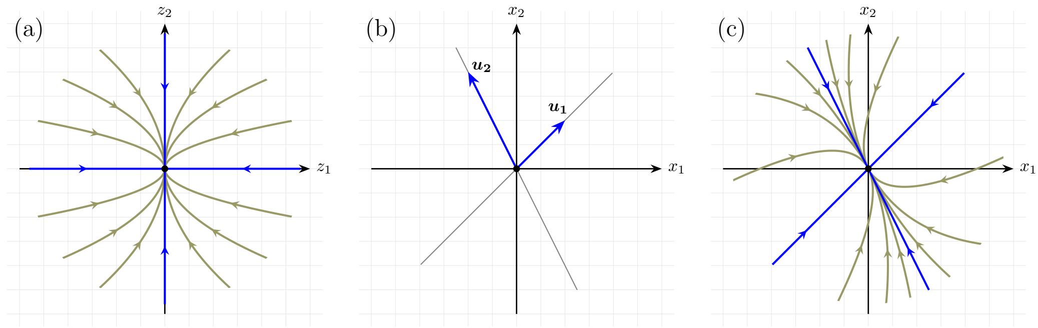


Figure 4.2: (a) Solution curves for an unstable node in the z-plane with $r_1 = 6$ and $r_2 = 3$. (b) Eigenvectors for the problem $\boldsymbol{x}' = A\boldsymbol{x}$ in Example 4.6. (c) Solution curves in the original coordinates $\boldsymbol{x}$ for Example 4.6. The solution curves shown in this picture are the same as the curves shown in (a), transformed using the matrix $U = [\boldsymbol{u}_1|\boldsymbol{u}_2]$.

For the solution curves in the original coordinates $\boldsymbol{x} = \langle x_1, x_2\rangle$ we can again use the equation $\boldsymbol{x} = U\boldsymbol{z}$ to transform the solution curves that we obtained in the $\boldsymbol{z}$ coordinates. Since $U\boldsymbol{e}_1 = \boldsymbol{u}_1$ and $U\boldsymbol{e}_2 = \boldsymbol{u}_2$, the solution curves become tangent to the direction of $\boldsymbol{u}_1$ as they approach the origin when $t \to \infty$, and move away from the origin when $t \to -\infty$.

How to draw the phase portrait for a stable node

Assume that A has eigenvalues $r_2 < r_1 < 0$.

(i) Find the corresponding eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.
(ii) Draw solution curves moving toward the origin along the straight half-lines in the directions $\pm\boldsymbol{u}_1$ and $\pm\boldsymbol{u}_2$.
(iii) Draw solution curves that are tangent to the $\pm\boldsymbol{u}_2$ direction[a] at the origin and that move toward the origin.

[a]This is the direction corresponding to the eigenvalue with the smaller absolute value.

Example 4.6. Consider the planar linear system $\boldsymbol{x}' = A\boldsymbol{x}$ with

$$A = \begin{bmatrix} -5 & -1 \\ -2 & -4 \end{bmatrix}.$$

The matrix A has eigenvalues $r_1 = -3$ and $r_2 = -6$ with corresponding eigenvectors $\boldsymbol{u}_1 = \langle -1, 2\rangle$ and $\boldsymbol{u}_2 = \langle 1, 1\rangle$. The solution curves in the $\boldsymbol{z}$ plane are shown in Fig. 4.1(a), while Fig. 4.1(b) shows the two eigenvectors in the $\boldsymbol{x}$ plane. The solution curves in the $\boldsymbol{x}$ plane are shown in Fig. 4.1(c).

4.2.3 Distinct Eigenvalues with Opposite Signs: Saddle

Consider finally the case where the matrix A has two real eigenvalues with opposite signs. In this case, the equilibrium at the origin is called a *saddle*. Without loss of generality assume that $r_1 < 0 < r_2$.

We consider the shape of the solution curves in the $\boldsymbol{z}$ plane. Consider the parametric form of the solution curve

$$\boldsymbol{z}(t) = \begin{bmatrix} c_1 e^{r_1 t} \\ c_2 e^{r_2 t} \end{bmatrix}.$$

For $t = 0$ we find $\boldsymbol{z}(0) = \boldsymbol{c}$. If the solution starts along one of the two coordinates axes on the $\boldsymbol{z}$ plane (corresponding to the directions of the eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ on the $\boldsymbol{x}$ plane), then we observe two different behaviors.

If the solution starts on the z_1 axis, that is, if $z_2(0) = 0$, then the solution is $\boldsymbol{z}(t) = \langle z_1(0)e^{r_1 t}, 0\rangle$. That is, the solution stays on the z_1 axis, and since $r_1 < 0$, it approaches the origin as t increases. However, if the solution starts on the z_2 axis, that is, if $z_1(0) = 0$, then the solution is $\boldsymbol{z}(t) = \langle 0, z_2(0)e^{r_2 t}\rangle$. That is, the solution stays on the z_2 axis, but since $r_2 < 0$, it moves away from the origin as t increases (and moves toward the origin as t decreases). Because of these distinct behaviors along the two axes, we call the z_1 axis the *stable direction* and the z_2 axis the *unstable direction*.

For all other initial conditions, having $z_1(0) \neq 0$ and $z_2(0) \neq 0$, the solution is a combination of a motion toward the z_2 axis and a motion away from the z_1 axis as t increases. Moreover, the function $g(z_1, z_2) = z_1^{r_2} z_2^{-r_1}$ remains constant along the solution curves since

$$\frac{\mathrm{d}}{\mathrm{d}t}[g(z_1(t), z_2(t))] = r_2 z_1^{r_2-1} z_1' z_2^{-r_1} - r_1 z_1^{r_2} z_2^{-r_1-1} z_2' = r_1 r_2 z_1^{r_2} z_2^{-r_1} - r_1 r_2 z_1^{r_2} z_2^{-r_1} = 0,$$

where we used that $z_1' = r_1 z_1$ and $z_2' = r_2 z_2$. Since $r_2 > 0$ and $-r_1 > 0$ the level sets of $z_1^{r_2} z_2^{-r_1}$ are generalized hyperbolas, as can be seen in Fig. 4.3(a). The z_1 and z_2 axes correspond to the two asymptotic directions of each of these generalized hyperbolas and the direction of motion along these curves is such that the motion becomes asymptotically aligned with the z_2 axis as t increases.

When the phase portrait is mapped to the $\boldsymbol{x}$ plane through the transformation $\boldsymbol{x} = U\boldsymbol{z}$, the stable direction is mapped to the direction defined by the eigenvector $\boldsymbol{u}_1$ with eigenvalue $r_1 < 0$ and the unstable direction is mapped to the direction defined by the eigenvector $\boldsymbol{u}_2$ with eigenvalue $r_2 > 0$. The generalized hyperbolas are mapped to "skewed" hyperbolas in the four regions between the stable and unstable directions, as can be seen in Fig. 4.3(c). The $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ axes correspond to the two asymptotic directions of each of these "skewed" hyperbolas and the direction of motion along these curves is such that the motion becomes asymptotically aligned with the $\boldsymbol{u}_2$ axis as t increases.

How to draw the phase portrait for a saddle

Assume that A has eigenvalues $r_1 < 0 < r_2$.

(i) Find the corresponding eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.
(ii) Draw solution curves moving toward the origin along the straight half-lines in the directions $\pm\boldsymbol{u}_1$ (corresponding to the negative eigenvalue) and solution curves moving away from the origin along the straight half-lines in the directions $\pm\boldsymbol{u}_2$ (corresponding to the positive eigenvalue).
(iii) Draw hyperbola-like solution curves that are asymptotic as t increases to one of the straight lines defined by $\pm\boldsymbol{u}_2$ and as t decreases to one of the straight half-lines defined by $\pm\boldsymbol{u}_1$.

Example 4.7. We consider the system $\boldsymbol{x}' = A\boldsymbol{x}$ where $\boldsymbol{x} \in \mathbf{R}^2$ and

$$A = \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix}.$$

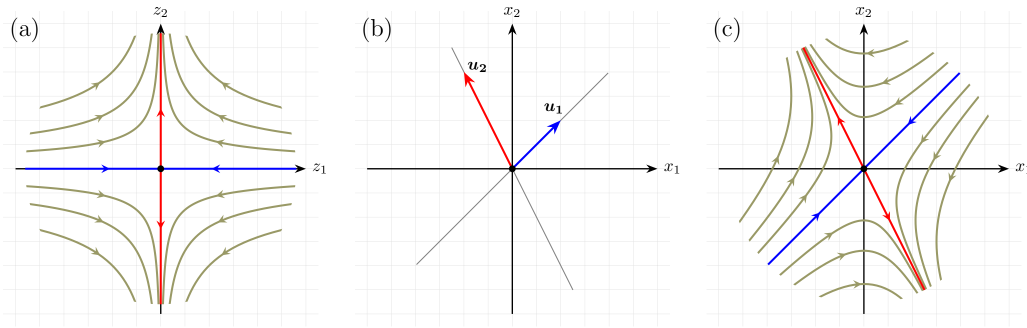


Figure 4.3: (a) Solution curves for a saddle in the z-plane with $r_1 = -2$ and $r_2 = 1$. (b) Eigenvectors for the problem $\boldsymbol{x}' = A\boldsymbol{x}$ in Example 4.7. (c) Solution curves in the original coordinates $\boldsymbol{x}$ for Example 4.7. The solution curves shown in this picture are the same as the curves shown in (a), transformed using the matrix $U = [\boldsymbol{u}_1|\boldsymbol{u}_2]$.

The eigenvalues are $r_1 = -2$ and $r_2 = 1$ with corresponding eigenvectors $\boldsymbol{u}_1 = \langle 1, 1\rangle$ and $\boldsymbol{u}_2 = \langle -1, 2\rangle$. The general solution is

$$\boldsymbol{x}(t) = c_1 e^{-2t}\boldsymbol{u}_1 + c_2 e^{t}\boldsymbol{u}_2 = c_1 e^{-2t}\begin{bmatrix}1\\1\end{bmatrix} + c_2 e^{t}\begin{bmatrix}-1\\2\end{bmatrix} = \begin{bmatrix}c_1 e^{-2t} - c_2 e^{t}\\ c_1 e^{-2t} + 2c_2 e^{t}\end{bmatrix}.$$

The solution curves in the $\boldsymbol{z}$ plane are shown in Fig. 4.3(a), while Fig. 4.3(b) shows the two eigenvectors in the $\boldsymbol{x}$ plane. The solution curves in the $\boldsymbol{x}$ plane are shown in Fig. 4.3(c).

4.3 Complex Conjugate Eigenvalues

We now consider the case of complex conjugate eigenvalues $r = \alpha + i\beta$, $\overline{r} = \alpha - i\beta$. We adopt the convention that $\beta > 0$ and denote the corresponding eigenvectors as $\boldsymbol{u} = \boldsymbol{a} + i\boldsymbol{b}$ and $\overline{\boldsymbol{u}} = \boldsymbol{a} - i\boldsymbol{b}$. The (real) vectors $\boldsymbol{a}$, $\boldsymbol{b}$ are linearly independent. Since

$$A(\boldsymbol{a} + i\boldsymbol{b}) = (\alpha + i\beta)(\boldsymbol{a} + i\boldsymbol{b}) = \alpha\boldsymbol{a} - \beta\boldsymbol{b} + i(\beta\boldsymbol{a} + \alpha\boldsymbol{b}),$$

we obtain that

$$A\boldsymbol{a} = \alpha\boldsymbol{a} - \beta\boldsymbol{b}, \quad A\boldsymbol{b} = \beta\boldsymbol{a} + \alpha\boldsymbol{b}. \tag{4.9}$$

The complex vector-valued functions

$$\boldsymbol{w}(t) = e^{rt}\boldsymbol{u} = e^{(\alpha+i\beta)t}(\boldsymbol{a} + i\boldsymbol{b}), \quad \overline{\boldsymbol{w}}(t) = e^{\overline{r}t}\overline{\boldsymbol{u}} = e^{(\alpha-i\beta)t}(\boldsymbol{a} - i\boldsymbol{b}),$$

are solutions of the given linear system $\boldsymbol{x}' = A\boldsymbol{x}$. However, since we want real vector-valued solutions we consider their linear combinations

$$\boldsymbol{x}_1(t) = \mathrm{Re}(\boldsymbol{w}(t)) = \frac{1}{2}(\boldsymbol{w}(t) + \overline{\boldsymbol{w}}(t)), \quad \boldsymbol{x}_2(t) = \mathrm{Im}(\boldsymbol{w}(t)) = \frac{1}{2i}(\boldsymbol{w}(t) - \overline{\boldsymbol{w}}(t)).$$

We compute

$$\begin{aligned}\boldsymbol{w}(t) &= e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))(\boldsymbol{a} + i\boldsymbol{b})\\ &= e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}) + ie^{\alpha t}(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b}).\end{aligned}$$

From the previous equation we can read that

$$\boldsymbol{x}_1(t) = e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}), \quad \boldsymbol{x}_2(t) = e^{\alpha t}(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b}). \tag{4.10}$$

Exercise 4.1. Verify directly that $\boldsymbol{x}_1(t)$, $\boldsymbol{x}_2(t)$ in Eq. (4.10) are solutions of $\boldsymbol{x}' = A\boldsymbol{x}$.

Moreover, $\boldsymbol{x}_1(0) = \boldsymbol{a}$, $\boldsymbol{x}_2(0) = \boldsymbol{b}$, implying that the solutions are linearly independent, and thus the general solution of $\boldsymbol{x}' = A\boldsymbol{x}$ is

$$\boldsymbol{x}(t) = e^{\alpha t}\big[c_1(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}) + c_2(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b})\big], \; c_1, c_2 \in \mathbf{R}. \tag{4.11}$$

We now obtain the solution in Eq. (4.11) using similar methods as in Section 4.2 and, again, without referring to the general theory. Define the matrix

$$U = [\boldsymbol{a}|\boldsymbol{b}],$$

and the change of coordinates $\boldsymbol{x} = U\boldsymbol{z}$. Then $\boldsymbol{z}' = R\boldsymbol{z}$, where $R = U^{-1}AU$. Unlike the case of distinct real eigenvalues, the matrix R here is not diagonal.[1]

Proposition 4.8. $U^{-1}AU = R := \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$.

Proof. The first column of U is $U\boldsymbol{e}_1 = \boldsymbol{a}$, and thus $\boldsymbol{e}_1 = U^{-1}\boldsymbol{a}$. Similarly, $\boldsymbol{e}_2 = U^{-1}\boldsymbol{b}$. We then compute that the first column of $R = U^{-1}AU$ is

$$R\boldsymbol{e}_1 = U^{-1}AU\boldsymbol{e}_1 = U^{-1}A\boldsymbol{a} = U^{-1}(\alpha\boldsymbol{a} - \beta\boldsymbol{b}) = \alpha\boldsymbol{e}_1 - \beta\boldsymbol{e}_2 = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix}.$$

Similarly, the second column is

$$R\boldsymbol{e}_2 = U^{-1}AU\boldsymbol{e}_2 = U^{-1}A\boldsymbol{b} = U^{-1}(\beta\boldsymbol{a} + \alpha\boldsymbol{b}) = \beta\boldsymbol{e}_1 + \alpha\boldsymbol{e}_2 = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

☑

As we saw in the discussion earlier, the dynamics in the case of complex eigenvalues is a combination of scaling and rotation. Rotation is often better described in polar coordinates. We first give a general result on how to write the dynamics in polar coordinates which will be again used in Section 6.5.1.

Polar coordinates

If $\boldsymbol{x} = \langle x_1, x_2\rangle$ are Cartesian coordinates on $\mathbf{R}^2$ then polar coordinates are defined by $x_1 = r\cos\theta$, $x_2 = r\sin\theta$. The radius $r \geq 0$ is defined uniquely and it satisfies $r^2 = x_1^2 + x_2^2$. The angle θ is defined up to the addition of integer multiples of 2π.

Proposition 4.9. *Consider a curve $\boldsymbol{x}(t) = \langle x_1(t), x_2(t)\rangle$ on $\mathbf{R}^2$ and denote by $r(t)$, $\theta(t)$ the polar coordinates parameterization of $\boldsymbol{x}(t)$. Then,*

$$r' = \frac{\boldsymbol{x}\cdot\boldsymbol{x}'}{\|\boldsymbol{x}\|} \text{ and } \theta' = \frac{x_1x_2' - x_1'x_2}{x_1^2 + x_2^2} = \frac{\boldsymbol{x}\times\boldsymbol{x}'}{\|\boldsymbol{x}\|^2}. \tag{4.12}$$

Remark 4.10. In Eq. (4.12) we have written $\boldsymbol{x}\times\boldsymbol{x}' = x_1x_2' - x_1'x_2$. In general, if $\boldsymbol{u}$, $\boldsymbol{v}$ are vectors on $\mathbf{R}^2$ we define their cross product as the real number $\boldsymbol{u}\times\boldsymbol{v} = u_1v_2 - u_2v_1 = \det[\boldsymbol{u}|\boldsymbol{v}]$.

Proof of Proposition 4.9. The easiest way to obtain r' is to observe that $r^2 = \|\boldsymbol{x}\|^2 = \boldsymbol{x}\cdot\boldsymbol{x}$. Then, taking the derivative of both sides of the equation with respect to t we

[1] We can diagonalize the matrix A using the complex transformation matrix $[\boldsymbol{u}|\overline{\boldsymbol{u}}]$, we prefer however to work only with real transformation matrices and real coordinates.

obtain

$$2rr' = 2\boldsymbol{x} \cdot \boldsymbol{x}',$$

which is the equation for r'. The correct way to obtain θ' is to take derivatives of both sides of the equations $x_1 = r\cos\theta$, $x_2 = r\sin\theta$. Then,

$$x_1' = r'\cos\theta - r\theta'\sin\theta, \quad x_2' = r'\sin\theta + r\theta'\cos\theta.$$

Multiplying the second equation by $r\cos\theta$ and the first equation by $r\sin\theta$ and subtracting we find

$$r^2\theta' = r\cos\theta x_2' - r\sin\theta x_1' = x_1x_2' - x_1'x_2 = \boldsymbol{x} \times \boldsymbol{x}'.$$ ☑

Define polar coordinates (ρ, ϕ) on the $\boldsymbol{z}$ plane by $z_1 = \rho\cos\phi$, $z_2 = \rho\sin\phi$. According to Proposition 4.9 we have

$$\rho' = \frac{\boldsymbol{z} \cdot \boldsymbol{z}'}{\|\boldsymbol{z}\|} \text{ and } \phi' = \frac{\boldsymbol{z} \times \boldsymbol{z}'}{\|\boldsymbol{z}\|^2}. \tag{4.13}$$

A direct computation for the system $\boldsymbol{z}' = R\boldsymbol{z}$ with $R = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ gives the two independent first order equations

$$\rho' = \alpha\rho, \ \ \phi' = -\beta.$$

The equation $\rho' = \alpha\rho$ has the general solution $\rho(t) = \rho_0 e^{\alpha t}$, $\rho_0 \geq 0$, and the equation $\phi' = -\beta$ has the general solution $\phi(t) = -\beta t + \phi_0$, $\phi_0 \in \mathbf{R}$. Therefore, in polar coordinates on the $\boldsymbol{z}$ plane the general solution has the form

$$\rho(t) = \rho_0 e^{\alpha t}, \ \ \phi(t) = -\beta t + \phi_0, \ \ \rho_0 \geq 0, \ \ \phi_0 \in \mathbf{R}. \tag{4.14}$$

This corresponds to a combination of two motions. The first is a *radial motion* which as t increases goes away from the origin if $\alpha > 0$, goes toward the origin if $\alpha < 0$, or stays at a fixed radius $\rho = \rho_0$ if $\alpha = 0$. The second is a *clockwise rotation* which is periodic with period $T = 2\pi/\beta$. Notice that the combined motion is periodic only if $\alpha = 0$.

We now write the solution in the coordinates z_1, z_2. We have

$$\begin{aligned} z_1(t) &= \rho_0 e^{\alpha t}\cos(-\beta t + \phi_0) = e^{\alpha t}\big[(\rho_0\cos\phi_0)\cos(\beta t) + (\rho_0\sin\phi_0)\sin(\beta t)\big], \\ z_2(t) &= \rho_0 e^{\alpha t}\sin(-\beta t + \phi_0) = e^{\alpha t}\big[(\rho_0\sin\phi_0)\cos(\beta t) - (\rho_0\cos\phi_0)\sin(\beta t)\big]. \end{aligned}$$

Defining $c_1 = \rho_0\cos\phi_0$, $c_2 = \rho_0\sin\phi_0$, we obtain

$$\boldsymbol{z}(t) = e^{\alpha t}\big[c_1(\cos(\beta t)\boldsymbol{e}_1 - \sin(\beta t)\boldsymbol{e}_2) + c_2(\sin(\beta t)\boldsymbol{e}_1 + \cos(\beta t)\boldsymbol{e}_2)\big]. \tag{4.15}$$

Equation (4.15) can be written as

$$\boldsymbol{z}(t) = e^{\alpha t}\begin{bmatrix} c_1\cos(\beta t) + c_2\sin(\beta t) \\ -c_1\sin(\beta t) + c_2\cos(\beta t) \end{bmatrix} = e^{\alpha t}\begin{bmatrix} \cos(\beta t) & \sin(\beta t) \\ -\sin(\beta t) & \cos(\beta t) \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{4.16}$$

Rotation matrices

Consider the matrix

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

If $\boldsymbol{v}$ is a vector on $\mathbf{R}^2$ then $R(\theta)\boldsymbol{v}$ is rotated counterclockwise by an angle θ with respect to $\boldsymbol{v}$. For this reason, matrices of the form $R(\theta)$ are called *rotation matrices.*

Observe that

$$\boldsymbol{z}(t) = e^{\alpha t} R(-\beta t)\boldsymbol{c}, \;\; \boldsymbol{c} = \langle c_1, c_2 \rangle \in \mathbf{R}^2.$$

The last expression represents a clockwise rotation of the vector $\boldsymbol{c}$ by an angle βt and a simultaneous scaling of the length of the vector by the factor $e^{\alpha t}$.

Finally, we express the solution in Eq. (4.15) in the original coordinates x_1, x_2. Since $U\boldsymbol{e}_1 = \boldsymbol{a}$, $U\boldsymbol{e}_2 = \boldsymbol{b}$, we find

$$\boldsymbol{x}(t) = U\boldsymbol{z}(t) = e^{\alpha t}\big[c_1(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}) + c_2(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b})\big],$$

which is Eq. (4.11).

In the rest of this section we analyze separately the cases $\alpha = 0$, $\alpha > 0$, and $\alpha < 0$.

4.3.1 Zero Real Part: Center

In the case where $\alpha = 0$ the dynamics in the $\boldsymbol{z}$ plane is given by $\boldsymbol{z}(t) = R(-\beta t)\boldsymbol{c}$, that is, it is clockwise rotation by angle βt. The length of the vector $\boldsymbol{z}(t)$ does not change, therefore the solutions $\boldsymbol{z}(t)$ trace circles on the $\boldsymbol{z}$ plane with radius determined by the length of $\boldsymbol{c}$. The phase portrait of a center in the $\boldsymbol{z}$ plane is shown in Fig. 4.4(a).

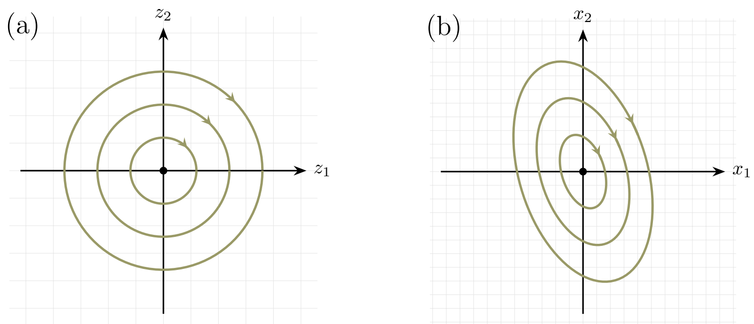


Figure 4.4: (a) Phase portrait for a center on the $\boldsymbol{z}$ plane. (b) Phase portrait for the center from Example 4.13 on the $\boldsymbol{x}$ plane.

The dynamics around a center is *periodic* with period

$$T = \frac{2\pi}{\beta},$$

since for all $k \in \mathbf{Z}$ we have

$$R(-\beta(t + kT)) = R(-\beta t + 2k\pi) = R(-\beta t).$$

We now want to understand how the corresponding phase portrait looks like in the $\boldsymbol{x}$ plane. The first thing to observe is that under the transformation $\boldsymbol{x} = U\boldsymbol{z}$ the circles in the $\boldsymbol{z}$ plane are transformed to ellipses in the $\boldsymbol{x}$ plane. We have the following proposition, see also Fig. 4.5.

Proposition 4.11. *Given a coordinate transformation $\boldsymbol{x} = U\boldsymbol{z}$, denote by $0 < \rho_1 \leq \rho_2$ the eigenvalues of the symmetric, positive definite matrix UU^t and by $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ the corresponding eigenvectors. Then the circle $\mathcal{C}_R = \{z_1^2 + z_2^2 = R^2\}$ of radius $R > 0$ on the $\boldsymbol{z}$ plane is mapped to an ellipse $\mathcal{E}_R$ on the $\boldsymbol{x}$ plane. The minor semiaxis of $\mathcal{E}_R$ is along the direction of $\boldsymbol{w}_1$ and has length $R\sqrt{\rho_1}$, while the major semiaxis of $\mathcal{E}_R$ is along the direction of $\boldsymbol{w}_2$ and has length $R\sqrt{\rho_2}$.*

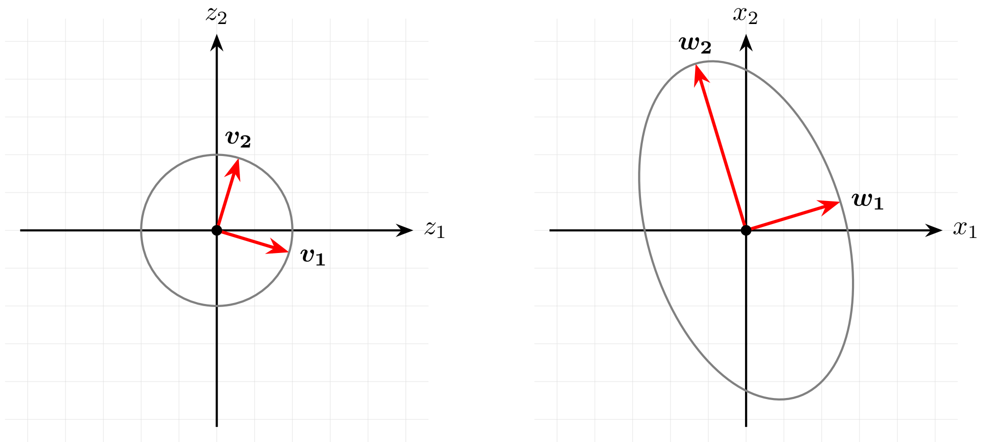


Figure 4.5: The unit circle $\mathcal{C}_1$ is shown on the $\boldsymbol{z}$ plane together with the eigenvectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ of U^tU normalized to 1. The ellipse $\mathcal{E}_1$ which is the image of $\mathcal{C}_1$ under the mapping $\boldsymbol{x} = U\boldsymbol{z}$ is shown on the $\boldsymbol{x}$ plane. The eigenvectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ of UU^t are normalized to $\sqrt{\rho_1}$, $\sqrt{\rho_2}$ respectively and mark the semiaxes of $\mathcal{E}_1$. The pictures correspond to Example 4.13.

Proof. Since UU^t is symmetric and positive definite it has real positive eigenvalues $0 < \rho_1 \leq \rho_2$. Moreover, the corresponding eigenvectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are orthogonal, that is, $\boldsymbol{w}_1 \cdot \boldsymbol{w}_2 = 0$, and they can be chosen to have lengths $\|\boldsymbol{w}_1\| = \sqrt{\rho_1}$ and $\|\boldsymbol{w}_2\| = \sqrt{\rho_2}$. Therefore, $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ form an orthogonal — but not necessarily orthonormal — basis for the $\boldsymbol{x}$ plane. An orthonormal basis is defined by $\hat{\boldsymbol{w}}_j = \boldsymbol{w}_j/\sqrt{\rho_j}$, $j = 1, 2$.

Define vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ on the $\boldsymbol{z}$ plane by $\boldsymbol{v}_j = U^{-1}\boldsymbol{w}_j$, $j = 1, 2$. Then,

$$U^tU\boldsymbol{v}_j = U^t\boldsymbol{w}_j = U^{-1}UU^t\boldsymbol{w}_j = U^{-1}(\rho_j\boldsymbol{w}_j) = \rho_j U^{-1}\boldsymbol{w}_j = \rho_j\boldsymbol{v}_j,$$

that is, $\boldsymbol{v}_j$ is an eigenvector of U^tU with eigenvalue ρ_j. We compute that

$$\boldsymbol{w}_i \cdot \boldsymbol{w}_j = (U\boldsymbol{v}_i) \cdot (U\boldsymbol{v}_j) = (U^tU\boldsymbol{v}_i) \cdot \boldsymbol{v}_j = \rho_i \boldsymbol{v}_i \cdot \boldsymbol{v}_j.$$

Therefore,

$$\|\boldsymbol{v}_1\|^2 = 1, \quad \|\boldsymbol{v}_2\|^2 = 1, \quad \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0.$$

This means that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ form an orthonormal basis for the $\boldsymbol{z}$ plane.

Consider $\boldsymbol{z} \in \mathcal{C}_R$. Then, there are unique z_1, z_2 such that $\boldsymbol{z} = z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2$ and we have

$$R^2 = \boldsymbol{z} \cdot \boldsymbol{z} = (z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2) \cdot (z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2) = z_1^2 + z_2^2.$$

In terms of the orthonormal basis $\hat{\boldsymbol{w}}_1, \hat{\boldsymbol{w}}_2$ on the $\boldsymbol{x}$ plane, a vector on $\mathcal{E}_R$ is written uniquely as $\boldsymbol{x} = x_1\hat{\boldsymbol{w}}_1 + x_2\hat{\boldsymbol{w}}_2$. However, since $\boldsymbol{x} \in \mathcal{E}_R$ there is $\boldsymbol{z} \in \mathcal{C}_R$ such that

$$\boldsymbol{x} = U\boldsymbol{z} = z_1 U\boldsymbol{v}_1 + z_2 U\boldsymbol{v}_2 = z_1\boldsymbol{w}_1 + z_2\boldsymbol{w}_2 = z_1\sqrt{\rho_1}\hat{\boldsymbol{w}}_1 + z_2\sqrt{\rho_2}\hat{\boldsymbol{w}}_2.$$

This implies that $x_j = z_j\sqrt{\rho_j}$, $j = 1, 2$ and since $z_1^2 + z_2^2 = R^2$ we obtain

$$\frac{x_1^2}{\rho_1} + \frac{x_2^2}{\rho_2} = R^2.$$

Since the basis $\hat{\boldsymbol{w}}_1, \hat{\boldsymbol{w}}_2$ is orthonormal, the last equation describes an ellipse with the orientation and size asserted in the proposition. ☑

There is one last detail. The direction of rotation on the $\boldsymbol{z}$ plane is always clockwise, but what about the direction of rotation on the $\boldsymbol{x}$ plane?

Proposition 4.12. *The direction of rotation on the $\boldsymbol{x}$ plane is clockwise if* $\det U > 0$ *and counterclockwise if* $\det U < 0$.

Proof. This proposition reflects the basic fact that a linear transformation U with $\det U < 0$ reverses the orientation of the coordinate system and this also reverses the direction of rotation.

We give now a detailed argument. A direct computation shows that if $\boldsymbol{v}$, $\boldsymbol{u}$ are vectors on $\mathbf{R}^2$ and U is a 2×2 matrix then

$$(U\boldsymbol{v}) \times (U\boldsymbol{u}) = \det U(\boldsymbol{v} \times \boldsymbol{u}). \tag{4.17}$$

Since $\boldsymbol{x} = U\boldsymbol{z}$, we have

$$\theta' = \frac{\boldsymbol{x} \times \boldsymbol{x}'}{\|\boldsymbol{x}\|^2} = \frac{(U\boldsymbol{z}) \times (U\boldsymbol{z}')}{\|\boldsymbol{x}\|^2} = \det U \frac{\boldsymbol{z} \times \boldsymbol{z}'}{\|\boldsymbol{x}\|^2} = \det U \frac{\|\boldsymbol{z}\|^2}{\|\boldsymbol{x}\|^2}\phi'. \tag{4.18}$$

Since $\phi' = -\beta$ we obtain

$$\theta' = -\beta \det U \frac{\|\boldsymbol{z}\|^2}{\|\boldsymbol{x}\|^2}. \tag{4.19}$$

Specifically, for $\det U > 0$ we have $\theta' < 0$ which implies clockwise rotation, while for $\det U < 0$ we have $\theta' > 0$ which implies counterclockwise rotation. ☑

How to draw the phase portrait for a center

Assume that A has eigenvalues $\pm i\beta$, $\beta > 0$.

(i) Find an eigenvector $\boldsymbol{u}$ for $i\beta$, write $\boldsymbol{a} = \mathrm{Re}(\boldsymbol{u})$, $\boldsymbol{b} = \mathrm{Im}(\boldsymbol{u})$, and define $U = [\boldsymbol{a}|\boldsymbol{b}]$.
(ii) Find the eigenvalues $0 < \rho_1 \leq \rho_2$ of the matrix UU^t and the corresponding eigenvectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$.
(iii) Draw several ellipses centered at the origin with major semiaxis in the $\boldsymbol{w}_2$ direction and minor semiaxis in the $\boldsymbol{w}_1$ direction. The ratio of the two semiaxes is $\sqrt{\rho_2/\rho_1}$.
(iv) Mark the direction of rotation by drawing arrows on the ellipses. If $\det U > 0$ then the rotation is clockwise; if $\det U < 0$ then the rotation is counterclockwise.

If we are not interested in the precise shape and orientation of the ellipses, step (ii) can be skipped and in step (iii) we can draw a family of ellipses without paying attention to their shape and orientation. The direction of rotation, however, is always important and it can be computed without first computing $\det U$: on the positive x_1 semiaxis, with $x_1 > 0$ and $x_2 = 0$, we have $x_2' = cx_1$ and, therefore, the rotation is clockwise if $c < 0$ and counterclockwise if $c > 0$.

Exercise 4.2. Show that if A has complex conjugate eigenvalues then $c \neq 0$ and $\det U = -\beta/c$. Therefore $\det U$ and c have opposite signs.

Example 4.13. Consider the linear system $\boldsymbol{x}' = A\boldsymbol{x}$ with

$$A = \begin{bmatrix} 1 & 2 \\ -5 & -1 \end{bmatrix}.$$

The eigenvalues of A are $\pm 3i$. An eigenvector corresponding to the eigenvalue $r = 3i$. Since $\mathrm{Re}(r) = 0$, the origin is a center. The solution curves on the $\boldsymbol{z}$ plane rotate along circles around the origin in the clockwise direction. The solution curves on the $\boldsymbol{z}$ plane rotate along ellipses around the origin. The rotation is also in the clockwise direction since $A_{21} = -5 < 0$.

For a more precise understanding of the shape of the ellipses we compute that an eigenvector of A corresponding to $r = 3i$ is $\boldsymbol{u} = \langle -1-3i, 5\rangle$, giving $\boldsymbol{a} = \langle -1, 5\rangle$, $\boldsymbol{b} = \langle -3, 0\rangle$. Applying Eq. (4.11) we can write the general solution as

$$\begin{aligned} \boldsymbol{x}(t) &= c_1 \left(\cos(3t) \begin{bmatrix} -1 \\ 5 \end{bmatrix} - \sin(3t) \begin{bmatrix} -3 \\ 0 \end{bmatrix} \right) + c_2 \left(\sin(3t) \begin{bmatrix} -1 \\ 5 \end{bmatrix} + \cos(3t) \begin{bmatrix} -3 \\ 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} -(c_1 + 3c_2)\cos(3t) + (3c_1 - c_2)\sin(3t) \\ 5c_1 \cos(3t) + 5c_2 \sin(3t) \end{bmatrix}. \end{aligned}$$

The transformation matrix from the $\boldsymbol{z}$ plane to the $\boldsymbol{x}$ plane is

$$U = \begin{bmatrix} -1 & -3 \\ 5 & 0 \end{bmatrix}.$$

We compute $\det U = 15 > 0$, showing again that the rotation is clockwise. Moreover,

$$UU^t = \begin{bmatrix} 10 & -5 \\ -5 & 25 \end{bmatrix},$$

with eigenvalues $\rho_1 \approx 8.486$ and $\rho_2 \approx 26.514$ giving that the ratio of the major to minor axes of the ellipses on the $\boldsymbol{x}$ plane is $\sqrt{\rho_2/\rho_1} \approx 1.768$. The direction of the minor semiaxis is given by $\boldsymbol{w}_1 = \langle 0.957, 0.290\rangle$ and that of the major semiaxis by $\boldsymbol{w}_2 = \langle -0.290, 0.957\rangle$, see Fig. 4.4 and Fig. 4.5 .

4.3.2 Positive Real Part: Unstable Spiral

We consider now the case where the eigenvalues of A are complex conjugate $\alpha \pm i\beta$ with $\alpha > 0$. In this case the dynamics in the $\boldsymbol{z}$ plane is given by

$$\boldsymbol{z}(t) = e^{\alpha t} R(-\beta t) \boldsymbol{z}_0,$$

and as we discussed earlier it combines the rotation $R(-\beta t)$ with the scaling $e^{\alpha t}$. Since $\alpha > 0$, as t increases the distance from the origin also increases. This implies that the solution curves trace a spiral in the $\boldsymbol{z}$ plane which rotates clockwise around the origin as it moves further away. In this case the equilibrium is called *unstable spiral*.

The shape of the spiral depends on the ratio α/β since the distance from the origin is multiplied by $e^{\alpha T} = e^{2\pi\alpha/\beta}$ when time T passes. Smaller values α/β make the spiral tighter and as we saw in the previous section $\alpha/\beta = 0$ kills the spiral and we get circles around a center.

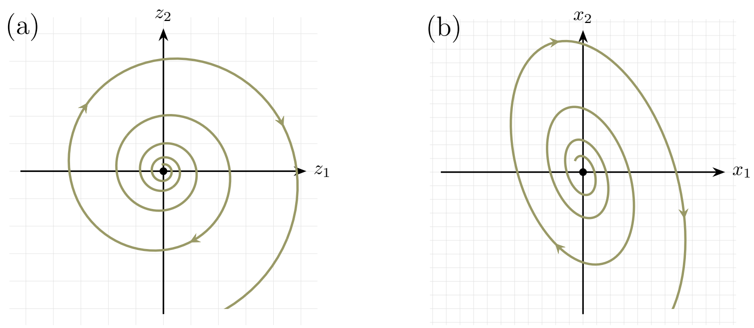


Figure 4.6: (a) Phase portrait for an unstable spiral on the $\boldsymbol{z}$ plane with $e^{2\pi\alpha/\beta} = e^{2\pi/9} \approx 2.01$. (b) Phase portrait for the unstable spiral from Example 4.14 on the $\boldsymbol{x}$ plane.

How to draw the phase portrait for an unstable spiral

Assume that A has eigenvalues $\alpha \pm i\beta$, with $\alpha > 0$ and $\beta > 0$.

(i) Find an eigenvector $\boldsymbol{u}$ for $i\beta$, write $\boldsymbol{a} = \mathrm{Re}(\boldsymbol{u})$, $\boldsymbol{b} = \mathrm{Im}(\boldsymbol{u})$, and define $U = [\boldsymbol{a}|\boldsymbol{b}]$.
(ii) Determine the direction of rotation. If $\det U > 0$ then the rotation is clockwise; if $\det U < 0$ then the rotation is counterclockwise. Alternatively, check the sign of c in the matrix U.
(iii) Find the eigenvalues $0 < \rho_1 \leq \rho_2$ of the matrix UU^t and the corresponding eigenvectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$.
(iv) Draw a spiral curve starting from near the origin and going outwards, elongated along the $\boldsymbol{w}_2$ direction and squeezed along the $\boldsymbol{w}_1$ direction.

Example 4.14. Consider the planar system $\boldsymbol{x}' = A\boldsymbol{x}$ with

$$A = \begin{bmatrix} 4 & 6 \\ -15 & -2 \end{bmatrix}.$$

The eigenvalues are $\alpha \pm i\beta = 1 \pm 9i$. Therefore, the solution curves on the $\boldsymbol{z}$ plane are spirals moving outwards and every time the solution curve makes one rotation around the origin, the distance is multiplied by $e^{2\pi\alpha/\beta} \approx 2.01$, as shown in Fig. 4.6(a).

An eigenvector of A corresponding to $r = 1 + 9i$ is $\boldsymbol{u} = \langle -1 - 3i, 5\rangle$, giving $\boldsymbol{a} = \langle -1, 5\rangle$, $\boldsymbol{b} = \langle -3, 0\rangle$. These are the same eigenvectors as in Example 4.13 and therefore we get the same transformation matrix U from the $\boldsymbol{z}$ plane to the $\boldsymbol{x}$ plane, given by

$$U = \begin{bmatrix} -1 & -3 \\ 5 & 0 \end{bmatrix}.$$

Moreover, applying Eq. (4.11) we can write the general solution as

$$\begin{aligned} \boldsymbol{x}(t) &= c_1 e^t \left(\cos(9t) \begin{bmatrix} -1 \\ 5 \end{bmatrix} - \sin(9t) \begin{bmatrix} -3 \\ 0 \end{bmatrix} \right) + c_2 e^t \left(\sin(9t) \begin{bmatrix} -1 \\ 5 \end{bmatrix} + \cos(9t) \begin{bmatrix} -3 \\ 0 \end{bmatrix} \right) \\ &= e^t \begin{bmatrix} -(c_1 + 3c_2)\cos(9t) + (3c_1 - c_2)\sin(9t) \\ 5c_1 \cos(9t) + 5c_2 \sin(9t) \end{bmatrix}. \end{aligned}$$

The phase portrait of the system is shown in Fig. 4.6(b).

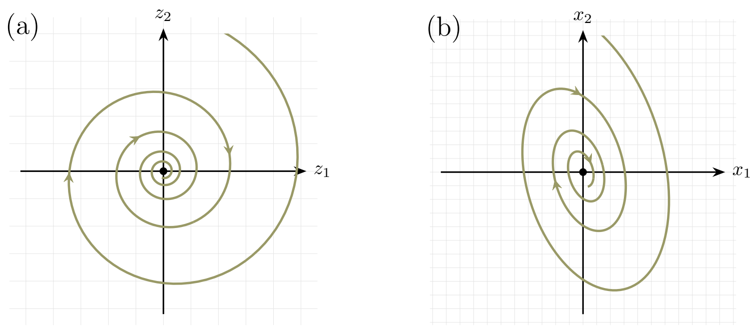


Figure 4.7: (a) Phase portrait for a stable spiral on the $\boldsymbol{z}$ plane. (b) Phase portrait for the stable spiral from Example 4.15 on the $\boldsymbol{x}$ plane.

4.3.3 Negative Real Part: Stable Spiral

We consider now the case where the eigenvalues of A are complex conjugate $\alpha \pm i\beta$ with $\alpha < 0$. Since $\alpha < 0$, as t increases the distance from the origin decreases and the solution curves approach the origin as $t \to \infty$. This implies that the solution curves trace a spiral in the $\boldsymbol{z}$ plane which rotates clockwise around the origin as it moves toward the origin. In this case the equilibrium is called *stable spiral*.

The shape of the spiral depends on the ratio α/β since the distance from the origin is multiplied by $e^{\alpha T} = e^{2\pi\alpha/\beta} < 1$ when time T passes. Smaller values $|\alpha|/\beta$ make the spiral tighter.

How to draw the phase portrait for a stable spiral

Assume that A has eigenvalues $\alpha \pm i\beta$, with $\alpha < 0$ and $\beta > 0$.

(i) Find an eigenvector $\boldsymbol{u}$ for $i\beta$, write $\boldsymbol{a} = \mathrm{Re}(\boldsymbol{u})$, $\boldsymbol{b} = \mathrm{Im}(\boldsymbol{u})$, and define $U = [\boldsymbol{a}|\boldsymbol{b}]$.
(ii) Determine the direction of rotation. If $\det U > 0$ then the rotation is clockwise; if $\det U < 0$ then the rotation is counterclockwise. Alternatively, check the sign of c in the matrix U.
(iii) Find the eigenvalues $0 < \rho_1 \leq \rho_2$ of the matrix UU^t and the corresponding eigenvectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$.
(iv) Draw a spiral curve starting from near the origin and going outwards, elongated along the $\boldsymbol{w}_2$ direction and squeezed along the $\boldsymbol{w}_1$ direction.

Example 4.15. Consider the planar system $\boldsymbol{x}' = A\boldsymbol{x}$ with

$$A = \begin{bmatrix} 2 & 6 \\ -15 & -4 \end{bmatrix}.$$

The eigenvalues are $\alpha \pm i\beta = -1 \pm 9i$. Therefore, the solution curves on the $\boldsymbol{z}$ plane are spirals moving inward and every time the solution curve makes one rotation around the origin, the distance is multiplied by $e^{2\pi\alpha/\beta} \approx 0.497$, as shown in Fig. 4.7(a).

An eigenvector of A corresponding to $r = -1 + 9i$ is $\boldsymbol{u} = \langle -1 - 3i, 5\rangle$, giving $\boldsymbol{a} = \langle -1, 5\rangle$, $\boldsymbol{b} = \langle -3, 0\rangle$. These are the same eigenvectors as in Example 4.13 and therefore we get the same transformation matrix U from the $\boldsymbol{z}$ plane to the $\boldsymbol{x}$ plane, given by

$$U = \begin{bmatrix} -1 & -3 \\ 5 & 0 \end{bmatrix}.$$

Moreover, applying Eq. (4.11) we can write the general solution as

$$\begin{aligned}\boldsymbol{x}(t) &= c_1 e^{-t}\left(\cos(9t)\begin{bmatrix}-1\\5\end{bmatrix} - \sin(9t)\begin{bmatrix}-3\\0\end{bmatrix}\right) + c_2 e^{-t}\left(\sin(9t)\begin{bmatrix}-1\\5\end{bmatrix} + \cos(9t)\begin{bmatrix}-3\\0\end{bmatrix}\right)\\ &= e^{-t}\begin{bmatrix}-(c_1+3c_2)\cos(9t) + (3c_1 - c_2)\sin(9t)\\ 5c_1\cos(9t) + 5c_2\sin(9t)\end{bmatrix}.\end{aligned}$$

The phase portrait of the system is shown in Fig. 4.7(b).

4.4 Real Repeated Eigenvalue

We turn now our attention to the case where A has a repeated eigenvalue $r_0 \in \mathbf{R}$. This means that the characteristic polynomial factorizes as

$$p(r) = r^2 - Tr + D = (r - r_0)^2,$$

and the Cayley-Hamilton theorem implies that

$$p(A) = (A - r_0 I)^2 = 0.$$

This means

$$(A - r_0 I)^2\boldsymbol{u} = \mathbf{0}, \text{ for all } \boldsymbol{u} \in \mathbf{R}^2.$$

Suppose that $\boldsymbol{u}_1$ is an eigenvector of A with eigenvalue r_0.

Proposition 4.16. *The matrix A has two linearly independent eigenvector $\boldsymbol{u}_1, \boldsymbol{u}_2$ with eigenvalue r_0 if and only if $A = r_0 I$.*

Proof. If $A = r_0 I$, clearly all non-zero vectors $\boldsymbol{u} \in \mathbf{R}^2$ are eigenvectors of A with eigenvalue r_0 and therefore we can find two linearly independent eigenvectors, for example, $\boldsymbol{u}_1 = \boldsymbol{e}_1$ and $\boldsymbol{u}_2 = \boldsymbol{e}_2$.

Conversely, suppose that $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ are two linearly independent eigenvectors. Then any $\boldsymbol{u} \in \mathbf{R}^2$ can be written as $\boldsymbol{u} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2$ and therefore

$$A\boldsymbol{u} = c_1 A\boldsymbol{u}_1 + c_2 A\boldsymbol{u}_2 = c_1 r_0 \boldsymbol{u}_1 + c_2 r_0 \boldsymbol{u}_2 = r_0\boldsymbol{u}.$$

This implies that $(A - r_0 I)\boldsymbol{u} = \mathbf{0}$ for all $\boldsymbol{u} \in \boldsymbol{R}^2$ and therefore $A - r_0 I = 0$, that is, $A = r_0 I$. ☑

Given the previous proposition we distinguish below the cases where A is diagonal with $A = r_0 I$ and where A is not diagonal.

4.4.1 Diagonal Matrix

When $A = r_0 I$ both coordinates x_1, x_2 on the $\boldsymbol{x}$ plane satisfy the same equation $x_j' = r_0 x_j$, $j = 1, 2$. Given initial condition $\boldsymbol{x}_0 = \boldsymbol{x}(0) = [x_1(0), x_2(0)]^t$ the solutions are $x_j(t) = e^{r_0 t} x_j(0)$. This implies that the solution in vector form is

$$\boldsymbol{x}(t) = e^{r_0 t}\boldsymbol{x}_0. \tag{4.20}$$

For $t \in \mathbf{R}$ the solution curve $\boldsymbol{x}(t)$ traces the straight half-line in the direction given by $\boldsymbol{x}_0$ excluding the origin. When $r_0 > 0$, we have $\lim_{t\to-\infty}\boldsymbol{x}(t) = \mathbf{0}$, while $\lim_{t\to\infty}\|\boldsymbol{x}(t)\| = \infty$. When $r_0 < 0$ the situation is reversed: $\lim_{t\to\infty}\boldsymbol{x}(t) = \mathbf{0}$, while $\lim_{t\to-\infty}\|\boldsymbol{x}(t)\| = \infty$.

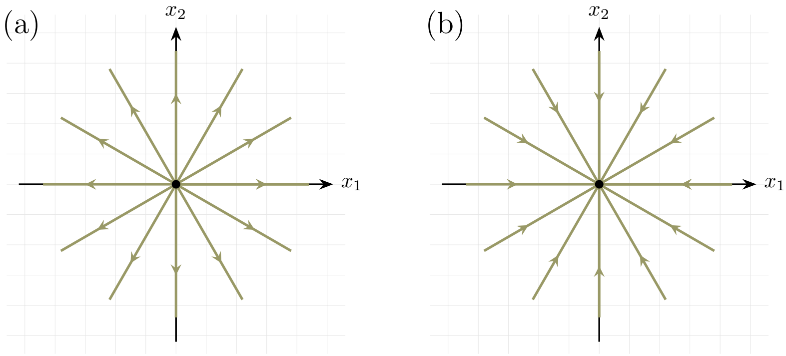


Figure 4.8: Phase portraits on the $\boldsymbol{x}$ plane for the case $A = r_0 I$: (a) $r_0 > 0$; (b) $r_0 < 0$.

4.4.2 Non-Diagonal Matrix

Consider now the case where $A \neq r_0 I$. This implies that A has only one eigenvector $\boldsymbol{u}_1$ with eigenvalue r_0. Take any non-zero vector $\boldsymbol{u}$ which is linearly independent from $\boldsymbol{u}_1$ and define $\boldsymbol{w} = (A - r_0 I)\boldsymbol{u}$. Then,

$$(A - r_0 I)\boldsymbol{w} = (A - r_0 I)^2 \boldsymbol{u} = \boldsymbol{0},$$

implying that $\boldsymbol{w}$ is an eigenvector of A with eigenvalue r_0 and therefore it must satisfy $\boldsymbol{w} = \lambda \boldsymbol{u}_1$ for some $\lambda \neq 0$. Define $\boldsymbol{u}_2 = \boldsymbol{u}/\lambda$. Then,

$$(A - r_0 I)\boldsymbol{u}_2 = \frac{1}{\lambda}(A - r_0 I)\boldsymbol{u} = \frac{1}{\lambda}\boldsymbol{w} = \boldsymbol{u}_1,$$

that is, $A\boldsymbol{u}_2 = \boldsymbol{u}_1 + r_0 \boldsymbol{u}_2$. Notice that the choice of $\boldsymbol{u}_2$ is not unique.

Define the matrix

$$U = [\boldsymbol{u}_1 | \boldsymbol{u}_2],$$

and the coordinate transformation $\boldsymbol{x} = A\boldsymbol{z}$. Like in previous cases we have $\boldsymbol{z}' = R\boldsymbol{z}$ with $R = U^{-1}AU$.

Proposition 4.17. $U^{-1}AU = R := \begin{bmatrix} r_0 & 1 \\ 0 & r_0 \end{bmatrix}$.

Proof. We have $A\boldsymbol{u}_1 = r_0 \boldsymbol{u}_1$ and $A\boldsymbol{u}_2 = \boldsymbol{u}_1 + r_0 \boldsymbol{u}_2$. The first column of R is given by

$$R\boldsymbol{e}_1 = (U^{-1}AU)\boldsymbol{e}_1 = U^{-1}A\boldsymbol{u}_1 = r_0 U^{-1}\boldsymbol{u}_1 = r_0 \boldsymbol{e}_1 = \begin{bmatrix} r_0 \\ 0 \end{bmatrix},$$

and the second column by

$$R\boldsymbol{e}_2 = (U^{-1}AU)\boldsymbol{e}_2 = U^{-1}A\boldsymbol{u}_2 = U^{-1}\boldsymbol{u}_1 + r_0 U^{-1}\boldsymbol{u}_2 = \boldsymbol{e}_1 + r_0 \boldsymbol{e}_2 = \begin{bmatrix} 1 \\ r_0 \end{bmatrix}.$$ ☑

This means that on the $\boldsymbol{z}$ plane with coordinates $\boldsymbol{z} = [z_1, z_2]^t$ we get the equations $z_1' = r_0 z_1 + z_2$ and $z_2' = r_0 z_2$. To solve this system of equations with initial conditions $\boldsymbol{z}_0 = \boldsymbol{z}(0) = [z_1(0), z_2(0)]^t$ we start with the second equation for which the solution is $z_2(t) = z_2(0)e^{r_0 t}$. Substituting this solution into the first equation we obtain a linear differential equation for $z_1(t)$ as

$$z_1' = r_0 z_1 + z_2(0)e^{r_0 t}.$$

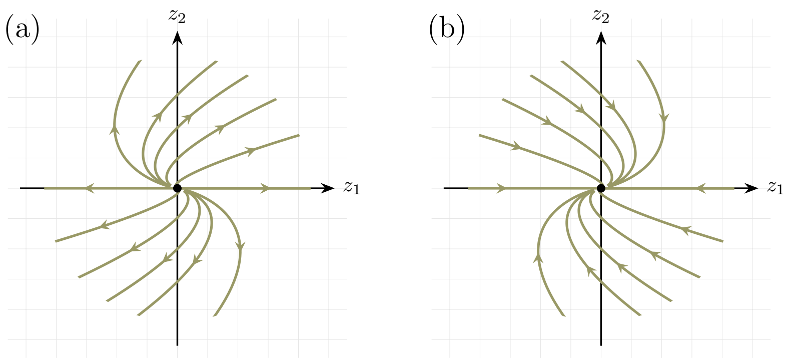


Figure 4.9: Phase portraits on the $\boldsymbol{z}$ plane for the case where A is not diagonal: (a) $r_0 > 0$; (b) $r_0 < 0$.

The last differential equation can be solved for $z_1(t)$ using either the method of integrating factor or the method of variation of parameters in Section 1.2.2. The general solution is

$$z_1(t) = ce^{r_0 t} + z_2(0)te^{r_0 t}.$$

Using the given initial condition for z_1 we obtain $c = z_1(0)$. Therefore, the solution is

$$z_1(t) = z_1(0)e^{r_0 t} + z_2(0)te^{r_0 t}.$$

Writing the solutions in vector form we have

$$\boldsymbol{z}(t) = e^{r_0 t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \boldsymbol{z}_0.$$

We can also write a closed form expression for the solution on the $\boldsymbol{x}$ plane. We have

$$\boldsymbol{x}(t) = U\boldsymbol{z}(t) = e^{r_0 t}U \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \boldsymbol{z}_0 = e^{r_0 t}U \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} U^{-1}\boldsymbol{x}_0.$$

However, we also have

$$\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} = I + (R - r_0 I)t,$$

which implies

$$U \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} U^{-1} = U(I + (R - r_0 I)t)U^{-1} = I + (URU^{-1} - r_0 I)t = I + (A - r_0 I)t.$$

Therefore,

$$\boldsymbol{x}(t) = e^{r_0 t}[I + (A - r_0 I)t]\boldsymbol{x}_0.$$

In the next chapter we discuss a more straightforward way to obtain the last solution.

4.5 The Trace-Determinant Plane

We close this chapter by summarizing the classification of the different types of linear dynamics through the trace-determinant plot shown in Fig. 4.10.

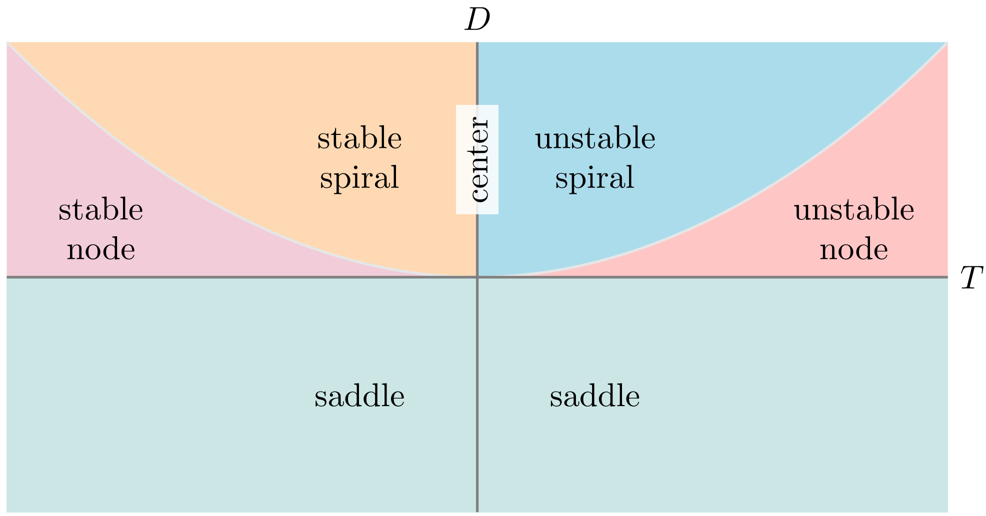


Figure 4.10: The trace-determinant plane.

The regions in the trace-determinant plot corresponding to different types of linear dynamics are obtained as follows. Recall that the characteristic polynomial of A is

$$p(r) = r^2 - Tr + D,$$

where T is the trace of A and D is its determinant, and its discriminant is $\Delta = T^2 - 4D$.

The case of two distinct real eigenvalues, discussed in Section 4.2, corresponds to $D < T^2/4$. In this case the eigenvalues are given by

$$r_1 = \frac{1}{2}(T + \sqrt{\Delta}) \text{ and } r_2 = \frac{1}{2}(T - \sqrt{\Delta}),$$

with

$$r_1 + r_2 = T \text{ and } r_1 r_2 = D.$$

Therefore, if $T > 0$ and $D > 0$ then $r_1, r_2 > 0$, corresponding to the case of unstable node. If $T < 0$ and $D > 0$ then $r_1, r_2 < 0$, corresponding to the case of stable node. If $D < 0$ then the two eigenvalues have opposite sign corresponding to the case of saddle.

The case of complex conjugate eigenvalues, discussed in Section 4.3, corresponds to $D > T^2/4$. In this case the eigenvalues are

$$\frac{1}{2}(T \pm i\sqrt{-\Delta}),$$

that is, the real part is $\alpha = T/2$. Therefore, $T > 0$ corresponds to unstable spiral, $T < 0$ to stable spiral, and $T = 0$ to center.

Finally, the case of repeated real eigenvalue, discussed in Section 4.4, corresponds to $D = T^2/4$, that is, to the boundary between the regions of complex eigenvalues and the regions of distinct real eigenvalues.

Remark 4.18. We have not discussed the case where one (or both) of the eigenvalues is zero. These correspond in the trace-determinant plot to the horizontal axis $D = 0$. ❞

Chapter 5

General Theory of Linear Dynamical Systems

5.1 Linear Systems

In this chapter we consider the general properties and solution methods for linear systems. In particular, we consider systems of the form

$$\boldsymbol{x}' = A(t)\boldsymbol{x} + \boldsymbol{f}(t), \tag{5.1}$$

where $\boldsymbol{x} \in \mathbf{R}^n$, $A(t)$ is a real $n \times n$ matrix, and $\boldsymbol{f} : \mathbf{R} \to \mathbf{R}^n$ is a function defined on $\mathbf{R}$ and taking values in $\mathbf{R}^n$. In components, we write

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \ A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{bmatrix}, \ \boldsymbol{f}(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

Solutions to Eq. (5.1) are vector valued functions $\boldsymbol{x}(t)$, with $\boldsymbol{x} : I \subseteq \mathbf{R} \to \mathbf{R}^n$. If $\boldsymbol{f}(t) \equiv \mathbf{0}$, the linear system in Eq. (5.1) is called *homogeneous*. Otherwise, it is called *non-homogeneous*. Linear systems such as Eq. (5.1) have unique solutions.

Theorem 5.1 (Existence and Uniqueness). *If $A(t)$ and $\boldsymbol{f}(t)$ are continuous on an open interval $I = (a, b) \subseteq \mathbf{R}$ and $t_0 \in I$, then for any initial vector $\boldsymbol{x}_0 \in \mathbf{R}^n$ the initial value problem*

$$\boldsymbol{x}' = A(t)\boldsymbol{x} + \boldsymbol{f}(t), \ \boldsymbol{x}(t_0) = \boldsymbol{x}_0,$$

has a a unique solution $\boldsymbol{x}(t)$ on I.

Corollary 5.2. *The only solution to the homogeneous initial value problem $\boldsymbol{x}' = A(t)\boldsymbol{x}$ with $\boldsymbol{x}(t_0) = \mathbf{0}$, where $A(t)$ is continuous on U and $t_0 \in I$, is the* zero solution *given by $\boldsymbol{x}(t) = \mathbf{0}$ for all $t \in I$.*

5.2 General Properties of Homogeneous Linear Systems

In this chapter we mostly consider homogeneous linear systems

$$\boldsymbol{x}' = A(t)\boldsymbol{x}, \tag{5.2}$$

and we assume that $A(t)$ is continuous on an interval $I \subseteq \mathbf{R}$. We return to a special type of non-homogeneous linear systems in Section 5.4.

Theorem 5.3 (Superposition Principle). *If $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_m(t)$ are solutions of Eq.* (5.2), *then any linear combination*

$$\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + \cdots + c_m\boldsymbol{x}_m(t), \; c_1, \ldots, c_m \in \mathbf{R}$$

is also a solution.

Proof. We directly check that

$$\begin{aligned} \boldsymbol{x}' &= (c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_m\boldsymbol{x}_m)' \\ &= c_1\boldsymbol{x}_1' + c_2\boldsymbol{x}_2' + \cdots + c_m\boldsymbol{x}_m' \\ &= c_1A(t)\boldsymbol{x}_1 + c_2A(t)\boldsymbol{x}_2 + \cdots + c_mA(t)\boldsymbol{x}_m \\ &= A(t)\big(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_m\boldsymbol{x}_m\big) \\ &= A(t)\boldsymbol{x}. \end{aligned}$$

☑

5.2.1 Linearly Independent Solutions and General Solution

Definition 5.4. The m vector valued functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are *linearly dependent* on an interval $I \subseteq \mathbf{R}$ if there exist $c_1, \ldots, c_m \in \mathbf{R}$, not all of them zero, such that on I we have

$$c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) + \cdots + c_m\boldsymbol{x}_m(t) \equiv \mathbf{0}.$$

The m vector valued functions $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$ are *linearly independent* on $I \subseteq \mathbf{R}$ if they are not linearly dependent on I, that is, if the relation

$$c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) + \cdots + c_m\boldsymbol{x}_m(t) \equiv \mathbf{0},$$

implies $c_1 = c_2 = \cdots = c_m = 0$.

Example 5.5. Consider the following three vector valued functions

$$\boldsymbol{x}_1(t) = \begin{bmatrix} e^t \\ 0 \\ e^t \end{bmatrix}, \quad \boldsymbol{x}_2(t) = \begin{bmatrix} 3e^t \\ 0 \\ 3e^t \end{bmatrix}, \quad \boldsymbol{x}_3(t) = \begin{bmatrix} t \\ 1 \\ 0 \end{bmatrix}.$$

These are linearly dependent on $\mathbf{R}$ since $3\boldsymbol{x}_1(t) + (-1)\boldsymbol{x}_2(t) + 0\boldsymbol{x}_3(t) = \mathbf{0}$ for all $t \in \mathbf{R}$.

The next theorem, Theorem 5.6, is important because it leads to the main result of this section, Theorem 5.9, asserting that the general solution of the homogeneous system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ is a linear combination of n linearly independent solutions.

Theorem 5.6. *Given solutions $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_m(t)$ of the linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ where $A(t)$ is continuous on an interval $I \subseteq \mathbf{R}$ the following three statements are equivalent:*

(i) The vector-valued functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are linearly independent on I.
(ii) The vectors $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_m(t)$ are independent for each $t \in I$.
(iii) The vectors $\boldsymbol{x}_1(t_0), \ldots, \boldsymbol{x}_m(t_0)$ are independent for some $t_0 \in I$.

The requirement in Theorem 5.6 that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are solutions of $\boldsymbol{x}' = A(t)\boldsymbol{x}$ is essential. In general, two arbitrary vector-valued functions $\boldsymbol{x}_1, \boldsymbol{x}_2$ can be linearly independent even though for each $t \in \mathbf{R}$ the vectors $\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)$ are linearly dependent, as in the following example.

Example 5.7. Let $\boldsymbol{x}_1(t) = \langle \cos t, 0\rangle$ and $\boldsymbol{x}_2(t) = \langle 1, 0\rangle$. Clearly, there is no choice of $c_1, c_2 \in \mathbf{R}$, not both zero, such that $c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) = \mathbf{0}$ for all $t \in \mathbf{R}$. However, for each $t \in \mathbf{R}$, the linear combination $\boldsymbol{x}_1(t) - \cos t\, \boldsymbol{x}_2(t) = 0$.

Proof of Theorem 5.6. We want to prove the following three equivalencies.

$$\begin{tikzcd}
& \text{(i) } \boldsymbol{x}_1, \dots, \boldsymbol{x}_m \text{ are linearly independent solutions on } I \arrow[dl, Leftrightarrow] \arrow[dr, Leftrightarrow] & \\
\text{(ii) } \boldsymbol{x}_1(t), \dots, \boldsymbol{x}_m(t) \text{ are linearly independent vectors for each } t \in I \arrow[rr, Leftrightarrow] & & \text{(iii) } \boldsymbol{x}_1(t_0), \dots, \boldsymbol{x}_m(t_0) \text{ are linearly independent vectors for some } t_0 \in I
\end{tikzcd}$$

Actually, it is sufficient to prove the following three implications.

$$\begin{tikzcd}
& \text{(i) } \boldsymbol{x}_1, \dots, \boldsymbol{x}_m \text{ are linearly independent solutions on } I \arrow[dl, Rightarrow] & \\
\text{(ii) } \boldsymbol{x}_1(t), \dots, \boldsymbol{x}_m(t) \text{ are linearly independent vectors for each } t \in I \arrow[rr, Rightarrow] & & \text{(iii) } \boldsymbol{x}_1(t_0), \dots, \boldsymbol{x}_m(t_0) \text{ are linearly independent vectors for some } t_0 \in I \arrow[ul, Rightarrow]
\end{tikzcd}$$

Then, for example, (iii) implies (ii) since (iii) implies (i) and (i) implies (ii). We now check the implications in the diagram above.

First, statement (iii) is a direct consequence of (ii): if the vectors $\boldsymbol{x}_1(t), \dots, \boldsymbol{x}_m(t)$ are linearly independent for each $t \in I$, then just pick any $t_0 \in I$ and (iii) holds there.

Second, (iii) implies (i). If $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are not linearly independent on I then there are $c_1, \dots, c_m$, not all zero, such that $c_1\boldsymbol{x}_1(t) + \dots + c_m\boldsymbol{x}_m(t) = \mathbf{0}$ for all $t \in I$. This implies that $c_1\boldsymbol{x}_1(t_0) + \dots + c_m\boldsymbol{x}_m(t_0) = \mathbf{0}$ contradicting that the vectors $\boldsymbol{x}_1(t_0), \dots, \boldsymbol{x}_m(t_0)$ are linearly independent. So, we can conclude that $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are linearly independent on I.

Finally, we close the loop by showing that (i) implies (ii) in Proposition 5.8. This is the only part where we use that $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are solutions of $\boldsymbol{x}' = A(t)\boldsymbol{x}$ and not arbitrary vector-valued functions.

Proposition 5.8. *If $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are linearly independent solutions on I of $\boldsymbol{x}' = A(t)\boldsymbol{x}$, where $A(t)$ is continuous on I, then the vectors $\boldsymbol{x}_1(t), \dots, \boldsymbol{x}_m(t)$ are linearly independent vectors for each $t \in I$.*

Proof. We prove the equivalent, contrapositive statement: if there is some $t_0 \in I$ such that the vectors $\boldsymbol{x}_1(t_0), \dots, \boldsymbol{x}_m(t_0)$ are linearly dependent, then the solutions $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are linearly dependent on I.

Since the vectors $\boldsymbol{x}_1(t_0), \ldots, \boldsymbol{x}_m(t_0)$ are assumed linearly dependent, there are $c_1, \ldots, c_m \in \mathbf{R}$, not all zero, such that

$$c_1\boldsymbol{x}_1(t_0) + \cdots + c_m\boldsymbol{x}_m(t_0) = \mathbf{0}.$$

Since $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are solutions of $\boldsymbol{x}' = A(t)\boldsymbol{x}$, the linear combination $\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + \cdots + c_m\boldsymbol{x}_m(t)$ is also a solution. Moreover, $\boldsymbol{x}(t)$ satisfies the initial condition $\boldsymbol{x}(t_0) = c_1\boldsymbol{x}_1(t_0) + \cdots + c_m\boldsymbol{x}_m(t_0) = \mathbf{0}$. Then Corollary 5.2 implies that $\boldsymbol{x}(t)$ is the zero solution, that is,

$$c_1\boldsymbol{x}_1(t) + \cdots + c_m\boldsymbol{x}_m(t) = \mathbf{0} \text{ for all } t \in \mathbf{R}.$$

Therefore, the solutions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are linearly dependent on I. ☑

We can now state and prove the main theorem on the general solution of the homogeneous linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$.

Theorem 5.9. *Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be linearly independent solutions of $\boldsymbol{x}' = A(t)\boldsymbol{x}$ on $I \subseteq \mathbf{R}$, and assume that $A(t)$ is continuous on I. If $\boldsymbol{x}(t)$ is any solution of the same equation $\boldsymbol{x}' = A(t)\boldsymbol{x}$ then there are unique $c_1, \ldots, c_n \in \mathbf{R}$ such that*

$$\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + \cdots + c_n\boldsymbol{x}_n(t).$$

Proof. Since the vector-valued functions $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ are linearly independent on I, there is some $t_0 \in I$ such that the vectors $\boldsymbol{x}_1(t_0), \ldots, \boldsymbol{x}_n(t_0)$ are linearly independent and therefore they form a basis for $\mathbf{R}^n$. This implies that there are unique numbers $c_1, \ldots, c_n$ such that

$$\boldsymbol{x}(t_0) = c_1\boldsymbol{x}_1(t_0) + \cdots + c_n\boldsymbol{x}_n(t_0).$$

Therefore, both $\boldsymbol{x}(t)$ and $c_1\boldsymbol{x}_1(t) + \cdots + c_n\boldsymbol{x}_n(t)$ satisfy the same initial value problem and so they must coincide. ☑

5.2.2 Wronskian Determinant

The linear independence of n solutions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is often expressed in terms of the following concept.

Definition 5.10. The *Wronskian determinant* of n solutions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ of the system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ is

$$W[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n](t) = \det[\boldsymbol{x}_1(t)| \ldots |\boldsymbol{x}_n(t)] = \det \begin{bmatrix} x_{1,1}(t) & x_{2,1}(t) & \ldots & x_{n,1}(t) \\ x_{1,2}(t) & x_{2,2}(t) & \ldots & x_{n,2}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,n}(t) & x_{2,n}(t) & \ldots & x_{n,n}(t) \end{bmatrix}.$$

Clearly, for each $t \in I$ the vectors $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ are linearly independent if and only if $W[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n](t) \neq 0$. Therefore, for the special case where $m = n$, we can restate Theorem 5.6 in the following form using the Wronskian determinant.

Theorem 5.11. *Given n solutions $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ of the linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$, $\boldsymbol{x}(t) \in \mathbf{R}^n$, where $A(t)$ is continuous on an interval $I \subseteq \mathbf{R}$, the following three statements are equivalent.*

(i) *The vector-valued functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent on I.*
(ii) *$W[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n](t) \neq 0$ for each $t \in I$.*
(iii) *$W[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n](t_0) \neq 0$ for some $t_0 \in I$.*

5.2.3 Fundamental Matrix

Theorem 5.9 shows that the general solution of the homogeneous linear equation $\boldsymbol{x}' = A(t)\boldsymbol{x}$ is the linear combination of any collection of n linearly independent solutions $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. It is useful to formalize the concept of collections of n linearly independent solutions and to study its properties.

Definition 5.12. A *fundamental solution* of the linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ is a collection of n linearly independent solutions $\boldsymbol{x}_1(t), \dots, \boldsymbol{x}_n(t)$. The associated *fundamental matrix* is

$$X(t) = [\boldsymbol{x}_1(t)| \dots |\boldsymbol{x}_n(t)],$$

that is, it is the matrix whose j-th column is the solution $\boldsymbol{x}_j$.

Notice that if $X(t) = [\boldsymbol{x}_1(t)| \dots |\boldsymbol{x}_n(t)]$ is a fundamental matrix, then its determinant is the Wronskian determinant of the n solutions. That is,

$$\det X(t) = W[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n](t).$$

Moreover, since the solutions $\boldsymbol{x}_1(t), \dots, \boldsymbol{x}_n(t)$ are linearly independent, $\det X(t) \neq 0$. Therefore, fundamental matrices are invertible.

We have seen in Theorem 5.9 that, given a fundamental solution $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ of the homogeneous linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$, any solution can be written as $\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + \dots + c_n\boldsymbol{x}_n(t)$. In matrix form,

$$\boldsymbol{x}(t) = X(t)\boldsymbol{c}, \text{ where } \boldsymbol{c} = \langle c_1, \dots, c_n \rangle.$$

Consider the initial value problem $\boldsymbol{x}' = A(t)\boldsymbol{x}$ with $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$. Then, $\boldsymbol{x}_0 = \boldsymbol{x}(t_0) = X(t_0)\boldsymbol{c}$, and we find $\boldsymbol{c} = X(t_0)^{-1}\boldsymbol{x}_0$. Therefore, if $X(t)$ is a fundamental matrix, then the solution to the initial value problem $\boldsymbol{x}' = A(t)\boldsymbol{x}$, $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is given by

$$\boldsymbol{x}(t) = X(t)X(t_0)^{-1}\boldsymbol{x}_0. \tag{5.3}$$

Definition 5.13. Let $M : \mathbf{R} \to \mathbf{R}^{n \times n}$ be a $n \times n$ matrix valued function of $t \in \mathbf{R}$ and denote the elements of $M(t)$ by $m_{ij}(t)$, $1 \leq i, j \leq n$. The matrix valued function M is *differentiable* at t if each function $m_{ij} : \mathbf{R} \to \mathbf{R}$ is differentiable at t. The *derivative* of M with respect to t is the matrix $M'(t)$ whose elements are the derivatives $m'_{ij}(t)$, $1 \leq i, j \leq n$.

We first establish some general properties of the derivatives of matrix valued functions.

Proposition 5.14. *If $X(t)$, $Y(t)$ are $n \times n$ matrix valued functions then the following statements hold.*

(i) If $X(t), Y(t)$ are differentiable at t, then $X(t)Y(t)$ is differentiable at t and

$$[X(t)Y(t)]' = X'(t)Y(t) + X(t)Y'(t).$$

(ii) If $X(t)$ is invertible and differentiable at t, then $X^{-1}(t)$ is differentiable at t and

$$(X^{-1})'(t) = -X^{-1}(t)X'(t)X^{-1}(t).$$

(iii) If $X(t)$ is differentiable at t, then $\det X(t)$ *is differentiable at t. If $X(t)$ is also invertible then*

$$[\det X(t)]' = \operatorname{tr}(X'(t)X^{-1}(t)) \det(X(t)).$$

Proof. (i) This is a direct consequence of the product rule for differentiation. Denote the elements of X and Y by x_{ij} and y_{ij} respectively — we do not show the dependence on t to keep the notation reasonable. The elements of the product XY are given by

$$(XY)_{ij} = \sum_{k=1}^{n} X_{ik} Y_{kj}.$$

Then,

$$(XY)'_{ij} = \sum_{k=1}^{n} X'_{ik} Y_{kj} + \sum_{k=1}^{n} X_{ik} Y'_{kj} = (X'Y + XY')_{ij}.$$

(ii) We have $X(t)X^{-1}(t) = I$. Assume first that $X^{-1}(t)$ is differentiable at t. Then taking derivatives of the relation $X(t)X^{-1}(t) = I$, and applying property (i) we find

$$X'(t)X^{-1}(t) + X(t)(X^{-1})'(t) = 0.$$

Solving for $(X^{-1})'(t)$ we get

$$(X^{-1})'(t) = -X^{-1}(t)X'(t)X^{-1}(t).$$

We now establish that if $X(t)$ is differentiable and invertible then $X^{-1}(t)$ is indeed differentiable. The determinant $\det X(t)$ can be expressed in terms of sums of products of the elements $x_{ij}(t)$. Therefore, if $X(t)$ is differentiable then $\det X(t)$ is differentiable. Moreover, since $X(t)$ is invertible, we have $\det X(t) \neq 0$, and thus $1/\det X(t)$ is also differentiable. Finally, each element of the inverse matrix $X^{-1}(t)$ is the product of $1/\det X(t)$ with a differentiable expression consisting of sums of products of the elements $x_{ij}(t)$, and therefore it is differentiable.

(iii) In the proof of property (ii) we have shown that if $X(t)$ is differentiable then $\det X(t)$ is also differentiable. Then we have

$$\begin{aligned}
[\det X(t)]' &= \lim_{h\to 0} \frac{1}{h}\big[\det X(t+h) - \det X(t)\big] \\
&= \det X(t) \lim_{h\to 0} \frac{1}{h}\big[\det(X(t+h)X^{-1}(t)) - \det(X(t)X^{-1}(t))\big] \\
&= \det X(t) \lim_{h\to 0} \frac{1}{h}\big[\det(X(t+h)X^{-1}(t)) - 1\big].
\end{aligned}$$

The Taylor theorem gives that

$$X(t+h) = X(t) + hX'(t) + O(h^2),$$

and thus

$$X(t+h)X^{-1}(t) = I + hX'(t)X^{-1}(t) + O(h^2).$$

Additionally, for any matrix B we have

$$\det(I + hB + O(h^2)) = 1 + h \operatorname{tr} B + O(h^2).$$

Therefore, applying the last equation with $B = X'(t)X^{-1}(t)$ we get

$$\begin{aligned}
[\det X(t)]' &= \det X(t) \lim_{h\to 0} \frac{1}{h}\big[h \operatorname{tr}(X'(t)X^{-1}(t)) + O(h^2)\big] \\
&= \det X(t) \operatorname{tr}(X'(t)X^{-1}(t)),
\end{aligned}$$

which is the claimed expression for $[\det X(t)]'$.

☑

Proposition 5.15. *If $X(t)$ is a fundamental solution of the system $\boldsymbol{x}' = A(t)\boldsymbol{x}$, then $X'(t) = A(t)X(t)$, that is, $X(t)$ satisfies the (matrix) differential equation $X' = A(t)X$.*

Proof. Since each column $\boldsymbol{x}_j(t)$ satisfies $\boldsymbol{x}_j'(t) = A(t)\boldsymbol{x}_j(t)$ we directly get that $X'(t) = A(t)X(t)$.

Proposition 5.16. *If $X(t)$ and $Y(t)$ are two fundamental matrices for the system $\boldsymbol{x}' = A(t)\boldsymbol{x}$, then there exists a constant invertible matrix C such that $Y(t) = X(t)C$.*

Proof. We first prove that $X^{-1}(t)Y(t)$ is a constant matrix. We compute

$$\begin{aligned}(X^{-1}(t)Y(t))' &= (X^{-1})'(t)Y(t) + X^{-1}(t)Y'(t)\\ &= -X^{-1}(t)X'(t)X^{-1}(t)Y(t) + X^{-1}(t)Y'(t).\end{aligned}$$

Since $X(t)$ and $Y(t)$ are fundamental matrices they satisfy $X'(t) = A(t)X(t)$ and $Y'(t) = A(t)Y(t)$. Therefore,

$$\begin{aligned}(X^{-1}(t)Y(t))' &= -X^{-1}(t)A(t)X(t)X^{-1}(t)Y(t) + X^{-1}(t)A(t)Y(t)\\ &= -X^{-1}(t)A(t)Y(t) + X^{-1}(t)A(t)Y(t)\\ &= 0,\end{aligned}$$

implying that $X^{-1}(t)Y(t) = C$ for some constant matrix C. Then $Y(t) = X(t)C$. Finally, since both $X(t)$ and $Y(t)$ are invertible, C is also invertible.

Moreover, if $X(t)$ is a fundamental matrix for the linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ and C is an invertible matrix, then the matrix $Y(t) = X(t)C$ is also a fundamental matrix for the same linear system, since

$$Y'(t) = X'(t)C = A(t)X(t)C = A(t)Y(t).$$

Therefore, the columns of $Y(t)$ are solutions of the system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ and they are linearly independent since $\det Y(t) = \det X(t) \det C \neq 0$.

Proposition 5.17 (Abel's formula). *Consider the linear system $\boldsymbol{x}' = A(t)\boldsymbol{x}$ in $\mathbf{R}^n$, let $X(t)$ be a fundamental solution and denote by $W(t) = \det X(t)$ the corresponding Wronskian determinant. Then*

$$W'(t) = (\operatorname{tr} A(t))\, W(t),$$

and, consequently,

$$W(t) = W(t_0) \exp\Big(\int_{t_0}^{t} \operatorname{tr} A(s)\, ds\Big).$$

Proof. Proposition 5.14 gives

$$W'(t) = [\det X(t)]' = \operatorname{tr}(X'(t)X^{-1}(t)) \det(X(t)) \operatorname{tr}(X'(t)X^{-1}(t)) W(t).$$

Since $X(t)$ is a fundamental matrix we have $X'(t) = A(t)X(t)$, and thus

$$W'(t) = \operatorname{tr}(A(t)X(t)X^{-1}(t))\, W(t) = \operatorname{tr}(A(t))\, W(t).$$

The equation $W' = (\operatorname{tr} A(t))W$ is a first-order separable differential equation for W and integrating it gives Abel's formula

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

5.3 Homogeneous Linear Systems with Constant Matrix

We now focus on the case where A is a constant $n \times n$ matrix and we want to find a fundamental matrix for

$$\boldsymbol{x}' = A\boldsymbol{x}.$$

The linear system with constant A is autonomous, and therefore it is sufficient to consider only initial conditions with $t_0 = 0$. Then given a fundamental matrix $X(t)$ and the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$, the solution to the given initial value problem can be written as

$$\boldsymbol{x}(t) = X(t)X^{-1}(0)\boldsymbol{x}_0.$$

5.3.1 Matrix Exponential

Suppose we compute a fundamental matrix $E(t)$ such that $E(0) = I$. Then $\boldsymbol{x}(t) = E(t)\boldsymbol{x}_0$. Therefore, applying the matrix $E(t)$ to the initial condition $\boldsymbol{x}_0$ directly gives the solution $\boldsymbol{x}(t)$ at time t. For this reason, the fundamental matrix $E(t)$ plays a central role in the solution methods for the linear system $\boldsymbol{x}' = A\boldsymbol{x}$, and we have the following definition.

Definition 5.18. Consider a constant matrix A, and the fundamental matrix $E(t)$ of the linear system $\boldsymbol{x}' = A\boldsymbol{x}$ satisfying $E(0) = I$, that is, the unique matrix satisfying the initial value problem $E' = AE$, $E(0) = I$. The *exponential of the matrix At* is defined as $\mathrm{e}^{At} = E(t)$. In particular, $e^{A} = E(1)$.

We have the following series expression for the matrix exponential.

Proposition 5.19. *If A is a $n \times n$ matrix then*

$$\mathrm{e}^{At} = I + tA + \frac{1}{2}t^2A^2 + \frac{1}{3!}t^3A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}t^kA^k. \tag{5.4}$$

Proof. Assuming that the series in Eq. (5.4) converges for all $t \in \mathbf{R}$, we compute its term-by-term derivative which will also converge. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=0}^{\infty} \frac{1}{k!}t^kA^k\right) = \sum_{k=1}^{\infty} \frac{k}{k!}t^{k-1}A^k = A\sum_{k=1}^{\infty} \frac{1}{(k-1)!}t^{k-1}A^{k-1} = A\sum_{k=0}^{\infty} \frac{1}{k!}t^kA^k.$$

Moreover, for $t = 0$ the series sums to the identity matrix. Therefore, the series satisfies the initial value problem $E' = AE$, $E(0) = I$ and, therefore, it equals e^{At}.

The proof that the series converges requires a discussion of matrix norms and their properties and we refer to standard Linear Algebra textbooks for establishing this fact. ☑

Remark 5.20. Equation (5.4) is identical to the Taylor series of the exponential function, given by $e^a = \sum_{k=0}^{\infty} \frac{1}{k!}a^k$. ❞

We collect in the following proposition, several properties of the matrix exponential.

Proposition 5.21. *Let A, B be $n \times n$ matrices, let I be the $n \times n$ identity matrix, and let O be the $n \times n$ zero matrix. Then the following properties hold:*

(i) $\mathrm{e}^{O} = I$.
(ii) $\mathrm{e}^{It} = \mathrm{e}^{t}I$.
(iii) *If $AB = BA$ then* $\mathrm{e}^{A+B} = \mathrm{e}^{A}\mathrm{e}^{B}$.

(iv) $e^{A(t+s)} = e^{At}e^{As}$.
(v) $(e^{At})^{-1} = e^{-At}$.
(vi) *If* $AB = BA$ *then* $e^{(A+B)t} = e^{At}e^{Bt}$.
(vii) *If* $A = U^{-1}BU$ *then* $e^{At} = U^{-1}e^{Bt}U$.
(viii) *If* A *is diagonal with* $A = \operatorname{diag}(r_1, \ldots, r_n)$, *then* $e^{At} = \operatorname{diag}(e^{r_1 t}, \ldots, e^{r_n t})$.

Proof. (i) The relation $e^O = I$ follows directly from Eq. (5.4). Alternatively, the initial value problem $E' = O$, $E(0) = I$, has the solution $E(t) = I$, and thus $e^O = E(1) = I$.

(ii) The relation $e^{It} = e^t I$ follows from Eq. (5.4), since

$$e^{It} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k I^k = \left(\sum_{k=0}^{\infty} \frac{1}{k!} t^k \right) I = e^t I.$$

Alternatively, it can be directly checked that $E(t) = e^t I$ is the solution to the initial value problem $E' = IE$, $E(0) = I$ and, by definition, $e^{It} = E(t) = e^t I$.

(iii) Using Eq. (5.4) we have

$$e^A e^B = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{n!\, m!} A^n B^m = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{j!(k-j)!} A^j B^{k-j},$$

where we transformed the double sum over with n, m taking values $0, 1, 2, \ldots$ to a double sum with $k = n + m$ taking values $0, 1, 2, \ldots$ and for each fixed value of k, we write $n = j$, $m = k - j$ with j taking values $0, 1, \ldots, k$. If $AB = BA$ then the following binomial formula holds:

$$(A + B)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j}.$$

Therefore,

$$e^A e^B = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j} = \sum_{k=0}^{\infty} \frac{1}{k!} (A + B)^k = e^{A+B}.$$

(iv) Since the matrices At and As commute, we get $e^{At+As} = e^{At}e^{As}$.

(v) We have $e^{At}e^{-At} = e^{A(t-t)} = e^O = I$. Therefore, the inverse of e^{At} is e^{-At}.

(vi) Since the matrices At and Bt commute we get $e^{(A+B)t} = e^{At}e^{Bt}$.

(vii) For each $k \geq 0$ we have $A^k = (U^{-1}BU)^k = U^{-1}B^kU$. Therefore,

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k = \sum_{k=0}^{\infty} \frac{1}{k!} t^k U^{-1} B^k U = U^{-1} \left(\sum_{k=0}^{\infty} \frac{1}{k!} t^k B^k \right) U = U^{-1} e^{Bt} U.$$

Alternatively, let $e^{Bt} = E_B(t)$ be the solution to the initial value problem $E' = BE$, $E(0) = I$. Then $E_A(t) = U^{-1}E_B(t)U$ satisfies

$$E_A'(t) = U^{-1}E_B'(t)U = U^{-1}BE_B(t)U = AU^{-1}E_B(t)U = AE_A(t),$$

and $E_A(0) = U^{-1}E_B(0)U = U^{-1}IU = I$. Therefore, $E_A(t)$ is the solution to the initial value problem $E' = AE$, $E(0) = I$, that is, it equals e^{At}. We conclude that $e^{At} = U^{-1}e^{Bt}U$.

(viii) If $A = \operatorname{diag}(r_1, \ldots, r_n)$ then for each $k \geq 0$ we have $A^k = \operatorname{diag}(r_1^k, \ldots, r_n^k)$. Therefore,

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \operatorname{diag}(r_1^k, \ldots, r_n^k) = \operatorname{diag}\left(\sum_{k=0}^{\infty} \frac{1}{k!} t^k r_1^k, \ldots, \sum_{k=0}^{\infty} \frac{1}{k!} t^k r_n^k \right),$$

finally giving

$$e^{At} = \operatorname{diag}(e^{r_1 t}, \ldots, e^{r_n t}).$$

Alternatively, define $E(t) = \operatorname{diag}(e^{r_1 t}, \ldots, e^{r_n t})$ and compute that

$$E'(t) = \operatorname{diag}(r_1 e^{r_1 t}, \ldots, r_n e^{r_n t}) = AE(t)$$

and $E(0) = I$. That is, $E(t)$ satisfies the initial value problem $E' = AE$ with $E(0) = I$, and therefore, $e^{At} = E(t) = \operatorname{diag}(e^{r_1 t}, \ldots, e^{r_n t})$. ☑

We are not much closer to computing $E(t) = e^{At}$ for an arbitrary matrix A but consider the following innocent observation.

Proposition 5.22. *If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent vectors in $\mathbf{R}^n$ then the collection*

$$\{e^{At}\boldsymbol{v}_1, \ldots, e^{At}\boldsymbol{v}_n\}$$

is a fundamental solution of the linear system $\boldsymbol{x}' = A\boldsymbol{x}$.

Proof. Let $\boldsymbol{x}_j(t) = e^{At}\boldsymbol{v}_j$. Then $\boldsymbol{x}_j' = Ae^{At}\boldsymbol{v}_j = A\boldsymbol{x}_j$. Therefore, each $\boldsymbol{x}_j$ is a solution of the linear system $\boldsymbol{x}' = A\boldsymbol{x}$. Moreover, at $t_0 = 0$, the n vectors $\boldsymbol{x}_j(0) = \boldsymbol{v}_j$ are linearly independent and therefore the corresponding solutions $\boldsymbol{x}_j(t)$ are linearly independent on $\mathbf{R}$. ☑

The previous proposition does not appear to be significantly simplifying our work of finding e^{At}. However, a judicious choice of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ may allow us to more easily compute the n linearly independent solutions $e^{At}\boldsymbol{v}_1, \ldots, e^{At}\boldsymbol{v}_n$ defining a fundamental matrix $X(t)$. Then we have

$$e^{At} = X(t)X(0)^{-1},$$

where

$$X(0) = V = [\boldsymbol{v}_1 | \ldots | \boldsymbol{v}_n].$$

What do we mean by judicious choice? Consider, for example, the case where $\boldsymbol{u}$ is an eigenvector of A with eigenvalue r. Then we have the following result which we further explore in the next section.

Proposition 5.23. *If $\boldsymbol{u}$ is an eigenvector of A with eigenvalue r (either real or complex) then*

$$e^{At}\boldsymbol{u} = e^{rt}\boldsymbol{u}.$$

Proof. We give two proofs of this fact. The first proof directly uses the series form for e^{At}, while the second proof introduces a technique that we use again later.

For the first proof, we have

$$e^{At}\boldsymbol{u} = I\boldsymbol{u} + tA\boldsymbol{u} + \frac{1}{2!}t^2A^2\boldsymbol{u} + \frac{1}{3!}t^3A^3\boldsymbol{u} + \cdots .$$

We observe that $A^k\boldsymbol{u} = r^k\boldsymbol{u}$ for $k = 1, 2, 3, \ldots$, and therefore

$$\begin{aligned}\exp(At)\boldsymbol{u} &= \boldsymbol{u} + tr\boldsymbol{u} + \frac{1}{2!}t^2r^2\boldsymbol{u} + \frac{1}{3!}t^3r^3\boldsymbol{u} + \cdots \\ &= \left(1 + tr + \frac{1}{2!}t^2r^2 + \frac{1}{3!}t^3r^3 + \cdots\right)\boldsymbol{u} \\ &= e^{rt}\boldsymbol{u}.\end{aligned}$$

For the second proof, write

$$e^{At} = e^{Irt}e^{(A-rI)t} = e^{rt}e^{(A-rI)t}.$$

Then

$$e^{(A-rI)t}\boldsymbol{u} = I\boldsymbol{u} + t(A - rI)\boldsymbol{u} + \frac{1}{2!}t^2(A - rI)^2\boldsymbol{u} + \cdots .$$

Since $\boldsymbol{u}$ is an eigenvector, we have $(A - rI)\boldsymbol{u} = \boldsymbol{0}$ and therefore $(A - rI)^k\boldsymbol{u} = \boldsymbol{0}$ for all $k = 1, 2, 3, \ldots$. The only term that remains from the series is $e^{(A-rI)t}\boldsymbol{u} = \boldsymbol{u}$. We conclude again that

$$e^{At}\boldsymbol{u} = e^{rt}e^{(A-r\mathrm{I})t}\boldsymbol{u} = e^{rt}\boldsymbol{u}.$$

☑

5.3.2 Distinct eigenvalues

If A has n real distinct eigenvalues $r_1, \ldots, r_n$ then the corresponding eigenvectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ are linearly independent. This means that the n solutions $\boldsymbol{x}_j(t) = e^{r_jt}\boldsymbol{u}_j$ are linearly independent and they form the fundamental matrix $X(t)$ with $X(0) = U = [\boldsymbol{u}_1|\ldots|\boldsymbol{u}_n]$. Therefore,

$$e^{At} = X(t)U^{-1}.$$

If A also has complex eigenvalues then they must come in complex conjugate pairs since the characteristic polynomial $p(r)$ is real. Then suppose that $r = \alpha + i\beta$ and $\bar{r} = \alpha - i\beta$ are complex conjugate eigenvalues and $\boldsymbol{w} = \boldsymbol{a} + i\boldsymbol{b}$, $\overline{\boldsymbol{w}} = \boldsymbol{a} - i\boldsymbol{b}$ are the corresponding eigenvectors with $\boldsymbol{a}$, $\boldsymbol{b}$ in $\mathbf{R}^n$. Notice that $\boldsymbol{a}$, $\boldsymbol{b}$ are linearly independent. Then we have

$$e^{At}\boldsymbol{w} = e^{At}\boldsymbol{a} + ie^{At}\boldsymbol{b},$$

and also

$$e^{At}\boldsymbol{w} = e^{rt}\boldsymbol{w} = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))(\boldsymbol{a} + i\boldsymbol{b}),$$

which gives

$$e^{At}\boldsymbol{w} = e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}) + ie^{\alpha t}(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b}).$$

Comparing the last equations we find

$$e^{At}\boldsymbol{a} = e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}), \text{ and } e^{At}\boldsymbol{b} = e^{\alpha t}(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b}).$$

Since $\boldsymbol{a}$, $\boldsymbol{b}$ are linearly independent vectors we conclude that $e^{At}\boldsymbol{a}$, $e^{At}\boldsymbol{b}$ are linearly independent solutions.

If all the n eigenvalues are distinct, either real or complex, then for each complex conjugate pair of eigenvalues $\alpha_1 \pm i\beta_1, \ldots, \alpha_k \pm i\beta_k$ we consider the corresponding eigenvectors $\boldsymbol{a}_1 \pm$

$i\boldsymbol{b}_1, \dots, \boldsymbol{a}_k \pm i\boldsymbol{b}_k$, and for each real eigenvalue $r_1, \dots, r_{n-2k}$ we consider the corresponding real eigenvectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-2k}$. The vectors $\boldsymbol{a}_1, \boldsymbol{b}_1, \dots, \boldsymbol{a}_k, \boldsymbol{b}_k, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-2k}$ are linearly independent. Then, the matrix

$$X(t) = [e^{At}\boldsymbol{a}_1 | e^{At}\boldsymbol{b}_1 | \dots | e^{At}\boldsymbol{a}_k | e^{At}\boldsymbol{b}_k | e^{At}\boldsymbol{u}_1 | \dots | e^{At}\boldsymbol{u}_{n-2k}],$$

is a fundamental matrix for which each column is easily computed. Finally, $e^{At} = X(t)X^{-1}(0)$. In particular, in this case the general solution of the system $\boldsymbol{x}' = A\boldsymbol{x}$ is

$$\begin{aligned} \boldsymbol{x}(t) =& \sum_{j=1}^{k} c_{2j-1} e^{\alpha_j t}(\cos(\beta_j t)\boldsymbol{a}_j - \sin(\beta_j t)\boldsymbol{b}_j) + c_{2j} e^{\alpha_j t}(\sin(\beta_j t)\boldsymbol{a}_j + \cos(\beta_j t)\boldsymbol{b}_j) \\ &+ \sum_{j=1}^{n-2k} c_{j+2k} e^{r_j t}\boldsymbol{u}_j, \end{aligned}$$

where $c_1, \dots, c_n \in \mathbf{R}$.

Exercise 5.1. Use the ideas discussed here to directly show that the general solution for a planar linear system with real distinct eigenvalues r_1, r_2 is

$$\boldsymbol{x}(t) = c_1 e^{r_1 t}\boldsymbol{u}_1 + c_2 e^{r_2 t}\boldsymbol{u}_2.$$

Exercise 5.2. Use the ideas discussed here to directly show that the general solution for a planar linear system with a complex conjugate pair of eigenvalues $\alpha \pm i\beta$ is

$$\boldsymbol{x}(t) = c_1 e^{\alpha t}(\cos(\beta t)\boldsymbol{a} - \sin(\beta t)\boldsymbol{b}) + c_2 e^{\alpha t}(\sin(\beta t)\boldsymbol{a} + \cos(\beta t)\boldsymbol{b}).$$

Example 5.24. Consider the system $\boldsymbol{x}' = A\boldsymbol{x}$ where

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

We have seen that A has eigenvalues $1 \pm 2i$ with corresponding eigevectors $[\pm i, 1]$. Therefore, we can take $\alpha = 1$, $\beta = 2$, $\boldsymbol{a} = [0, 1]$, $\boldsymbol{b} = [1, 0]$. Then we have the linearly independent solutions

$$\begin{aligned} \operatorname{Re}\left(e^{(1+2i)t}\begin{bmatrix} i \\ 1 \end{bmatrix}\right) &= e^t\left(\cos 2t \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \sin 2t \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = e^t \begin{bmatrix} -\sin 2t \\ \cos 2t \end{bmatrix}, \\ \operatorname{Im}\left(e^{(1+2i)t}\begin{bmatrix} i \\ 1 \end{bmatrix}\right) &= e^t\left(\sin 2t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \cos 2t \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = e^t \begin{bmatrix} \cos 2t \\ \sin 2t \end{bmatrix}. \end{aligned}$$

The corresponding fundamental matrix is

$$X(t) = e^t \begin{bmatrix} -\sin 2t & \cos 2t \\ \cos 2t & \sin 2t \end{bmatrix}.$$

5.3.3 Repeated eigenvalues

We consider now the case where the matrix A has repeated (real or complex) eigenvalues. It turns out that in this case the matrix may not have n linearly independent eigenvectors but it always has n linearly independent generalized eigenvectors. We show in Proposition 5.28 how to compute the solution $e^{At}\boldsymbol{u}$ when $\boldsymbol{u}$ is a generalized eigenvector of A and thus obtain a fundamental solution of the system $\boldsymbol{x}' = A\boldsymbol{x}$.

Definition 5.25. A non-zero vector $\boldsymbol{u}$ that satisfies $(A - r\mathsf{I})^m \boldsymbol{u} = \mathbf{0}$ for $r \in R$ and positive integer m, is called a *generalized eigenvector* of A associated with r.

Remark 5.26. The number r must be an eigenvalue of A. To see this suppose that m is the smallest positive integer such that $(A - r\mathsf{I})^m \boldsymbol{u} = \mathbf{0}$. Then

$$(A - r\mathsf{I}) \cdot \left[(A - r\mathsf{I})^{m-1} \boldsymbol{u}\right] = \mathbf{0},$$

which shows that $\boldsymbol{v} = (A - r\mathsf{I})^{m-1}\boldsymbol{u} \neq \mathbf{0}$ is an eigenvector of A with eigenvalue r. ”

Theorem 5.27. *Consider a matrix A that has k distinct eigenvalues (real or complex) $r_1, \dots, r_k$ with multiplicities $m_1, \dots, m_k$ satisfying $m_1 + \dots + m_k = n$, that is, the characteristic polynomial factors as*

$$p(r) = (r - r_1)^{m_1}(r - r_2)^{m_2} \cdots (r - r_k)^{m_k}.$$

Then for each $j = 1, \dots, k$ there exist m_j linearly independent generalized eigenvectors $\boldsymbol{u}_1^{(j)}, \dots, \boldsymbol{u}_{m_j}^{(j)}$ with

$$(A - r_j\mathsf{I})^{m_j} \boldsymbol{u}_i^{(j)} = \mathbf{0}, \quad i = 1, 2, \dots, m_j.$$

The n generalized eigenvectors $\{\boldsymbol{u}_i^{(j)}\}$ with $j = 1, \dots, k$ and $i = 1, \dots, m_j$ are linearly independent. Moreover, if the generalized eigenvectors for a complex eigenvalue r_j have been chosen as $\boldsymbol{a}_1^{(j)} + i\boldsymbol{b}_1^{(j)}, \dots, \boldsymbol{a}_{m_j}^{(j)} + i\boldsymbol{b}_{m_j}^{(j)}$ so that they are linearly independent, then the $2m_j$ vectors $\boldsymbol{a}_1^{(j)}, \boldsymbol{b}_1^{(j)}, \dots, \boldsymbol{a}_{m_j}^{(j)}, \boldsymbol{b}_{m_j}^{(j)}$ are linearly independent, and the collection of all vectors found in this way are also linearly independent.

The reason that one can use the generalized eigenvectors to compute a fundamental matrix is based on the following result that shows that e^{At} can be expressed as a finite sum.

Proposition 5.28. *If $\boldsymbol{u}$ is a generalized eigenvector of A that satisfies $(A - r\mathsf{I})^m \boldsymbol{u} = \mathbf{0}$ then*

$$\exp(At)\boldsymbol{u} = e^{rt}\left(\boldsymbol{u} + t(A - r\mathsf{I})\boldsymbol{u} + \frac{1}{2!}t^2(A - r\mathsf{I})^2\boldsymbol{u} + \dots + \frac{1}{(m-1)!}t^{m-1}(A - r\mathsf{I})^{m-1}\boldsymbol{u}\right).$$

Proof. Write

$$\exp(At)\boldsymbol{u} = \exp(rIt)\exp((A - rI)t)\boldsymbol{u} = e^{rt}\exp((A - rI)t)\boldsymbol{u}.$$

Then

$$\exp((A - rI)t)\boldsymbol{u} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k (A - rI)^k \boldsymbol{u}.$$

Since $(A - rI)^m \boldsymbol{u} = \mathbf{0}$, we also have $(A - rI)^k \boldsymbol{u} = \mathbf{0}$ for all $k \geq m$, and therefore

$$\exp(At)\boldsymbol{u} = e^{rt} \sum_{k=0}^{m-1} \frac{1}{k!} t^k (A - rI)^k \boldsymbol{u}.$$

☑

Example 5.29. We compute a fundamental matrix for the equation $\boldsymbol{x}' = A\boldsymbol{x}$ where

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

The characteristic polynomial is $p(r) = (r-1)^2$ with repeated eigenvalue $r = 1$. We compute

$$(A - I)^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore, all $\boldsymbol{u} \in \mathbf{R}^2$ are generalized eigenvectors of A. Choose the standard basis of $\mathbf{R}^2$ as linearly independent generalized eigenvectors, that is, $\boldsymbol{u}_1 = \boldsymbol{e}_1 = \langle 1, 0 \rangle$ and $\boldsymbol{u}_2 = \boldsymbol{e}_2 = \langle 0, 1 \rangle$. We compute that $(A-I)\boldsymbol{u}_1 = \mathbf{0}$ and $(A-I)\boldsymbol{u}_2 = \boldsymbol{e}_1$. Then we have the linearly independent solutions

$$\begin{aligned} \boldsymbol{x}_1(t) &= e^{At}\boldsymbol{u}_1 = e^t \left(\boldsymbol{u}_1 + t(A - I)\boldsymbol{u}_1\right) = e^t \left(\boldsymbol{u}_1 + t\mathbf{0}\right) = \langle e^t, 0 \rangle, \\ \boldsymbol{x}_2(t) &= e^{At}\boldsymbol{u}_2 = e^t \left(\boldsymbol{u}_2 + t(A - I)\boldsymbol{u}_2\right) = e^t \left(\boldsymbol{e}_2 + t\boldsymbol{e}_1\right) = \langle te^t, e^t \rangle. \end{aligned}$$

The corresponding fundamental matrix is given by

$$X(t) = [\boldsymbol{x}_1(t) | \boldsymbol{x}_2(t)] = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix}.$$

Additionally, observe that $X(0) = I$, and therefore the computed fundamental matrix $X(t)$ is the exponential e^{At}. This is the result of choosing $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ to be the standard basis in $\mathbf{R}^2$.

Exercise 5.3. Use the method of generalized eigenvectors to obtain the general solution of the system $\boldsymbol{x}' = A\boldsymbol{x}$, where

$$A = \begin{bmatrix} r_0 & \lambda \\ 0 & r_0 \end{bmatrix}.$$

Example 5.30. We compute a fundamental matrix $X(t)$ for the system $\boldsymbol{x}' = A\boldsymbol{x}$, where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

and then we compute e^{At}. The characteristic polynomial of A is

$$p(r) = \det(A - rI) = -(r-1)^2(r-3),$$

and therefore the eigenvalues are $r_1 = r_2 = 1$ and $r_3 = 3$.

Generalized eigenvectors corresponding to the double eigenvalue $r_1 = r_2 = 1$ are solutions of $(A - I)^2\boldsymbol{u} = \mathbf{0}$. Writing $\boldsymbol{u} = \langle a_1, a_2, a_3 \rangle$ we have the equation

$$(A - I)^2\boldsymbol{u} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}^2 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 4 & 0 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We find that $a_1 + 2a_2 = 0$ and that a_3 arbitrary. Choose the linearly independent generalized eigenvectors $\boldsymbol{u}_1 = \langle 0, 0, 1 \rangle$ and $\boldsymbol{u}_2 = \langle -2, 1, 0 \rangle$, and compute that $(A - I)\boldsymbol{u}_1 = \langle 0, 0, 0 \rangle$, $(A - I)\boldsymbol{u}_2 = \langle 0, 0, 1 \rangle$. Notice that $\boldsymbol{u}_1$ is an eigenvector of A, this is not, however, important — we could have chosen any two linearly independent generalized eigenvectors with neither being an eigenvector. With this choice of $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ we obtain the solutions

$$\begin{aligned} \boldsymbol{x}_1 &= e^{At}\boldsymbol{u}_1 = e^t \left(\boldsymbol{u}_1 + t(A - \mathsf{I})\boldsymbol{u}_1\right) = \langle 0, 0, e^t \rangle, \\ \boldsymbol{x}_2 &= e^{At}\boldsymbol{u}_2 = e^t \left(\boldsymbol{u}_2 + t(A - \mathsf{I})\boldsymbol{u}_2\right) = \langle -2e^t, e^t, te^t \rangle. \end{aligned}$$

The eigenvalue $r_3 = 3$ is not repeated and we compute a standard eigenvector $\boldsymbol{u}_3$. The computation gives $\boldsymbol{u}_3 = \langle 0, 2, 1\rangle$, and the corresponding solution is

$$\boldsymbol{x}_3 = e^{At}\boldsymbol{u}_3 = e^{3t}\boldsymbol{u}_3 = \langle 0, 2e^{3t}, e^{3t}\rangle.$$

Therefore, a fundamental matrix is

$$X(t) = [\boldsymbol{x}_1|\boldsymbol{x}_2|\boldsymbol{x}_3] = \begin{bmatrix} 0 & -2e^t & 0 \\ 0 & e^t & 2e^{3t} \\ e^t & te^t & e^{3t} \end{bmatrix}.$$

We compute

$$X(0)^{-1} = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}^{-1} = \frac{1}{4}\begin{bmatrix} -1 & -2 & 4 \\ -2 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}.$$

Then, e^{At} is obtained as

$$e^{At} = X(t)X(0)^{-1} = \frac{1}{4}\begin{bmatrix} 0 & -2e^t & 0 \\ 0 & e^t & 2e^{3t} \\ e^t & te^t & e^{3t} \end{bmatrix}\begin{bmatrix} -1 & -2 & 4 \\ -2 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix},$$

and the final result is

$$e^{At} = \frac{1}{4}\begin{bmatrix} 4e^t & 0 & 0 \\ 2e^{3t} - 2e^t & 4e^{3t} & 0 \\ e^{3t} - e^t - 2te^t & 2e^{3t} - 2e^t & 4e^t \end{bmatrix}.$$

Remark 5.31. After a long computation it is good practice to check the result for easy to catch mistakes. A check that should always be made is to verify that for $t = 0$ the computed expression for e^{At} gives the identity matrix. In the final result above we can directly check that this holds. If not, then there must have been a mistake in the computation of the inverse $X(0)^{-1}$ or the computation of the product $X(t)X(0)^{-1}$.

5.4 Non-homogeneous Linear Systems with Constant Matrix

Consider the equation

$$\boldsymbol{x}' = A\boldsymbol{x} + \boldsymbol{f}(t).$$

To solve this system we will use the method of variation of parameters.

Recall that the general solution of the homogeneous system $\boldsymbol{x}' = A\boldsymbol{x}$ is $\boldsymbol{x}(t) = e^{At}\boldsymbol{c}$. To find a solution for the non-homogeneous system we replace the constant $\boldsymbol{c}$ by a function $\boldsymbol{h}(t)$. Then, substituting the solution $\boldsymbol{x}(t) = e^{At}\boldsymbol{h}(t)$ into the given equation we obtain

$$Ae^{At}\boldsymbol{h}(t) + e^{At}\boldsymbol{h}'(t) = Ae^{At}\boldsymbol{h}(t) + \boldsymbol{f}(t).$$

That is,

$$\boldsymbol{h}'(t) = e^{-At}\boldsymbol{f}(t),$$

and we can integrate to get

$$\boldsymbol{h}(t) = \int e^{-At}\boldsymbol{f}(t)\,\mathrm{d}t + \boldsymbol{c}.$$

Therefore, the general solution is

$$\boldsymbol{x}(t) = e^{At}\boldsymbol{c} + e^{At}\int e^{-At}\boldsymbol{f}(t)\,\mathrm{d}t.$$

Example 5.32. Consider the second order non-homogeneous equation $y'' + y = \cos t$. The method of undetermined coefficients gives the general solution $y(t) = c_1 \cos t + c_2 \sin t + \frac{1}{2} t \sin t$. We use the result of this section to obtain the same result.

Define $x_1 = y$ and $x_2 = y'$. Then we find $x_1' = x_2$ and $x_2' = -x_1 + \cos t$. Therefore, the corresponding linear system is $\boldsymbol{x}' = A\boldsymbol{x} + \boldsymbol{f}(t)$, where

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \ \boldsymbol{f}(t) = \begin{bmatrix} 0 \\ \cos t \end{bmatrix}.$$

We first compute e^{At}. The characteristic polynomial of A is $p(r) = r^2 + 1$ with roots $r = \pm i$. An eigenvector for $r = i$ is $\boldsymbol{u} = \langle 1, i \rangle$. Therefore, we have

$$e^{it} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} \cos t + i \sin t \\ -\sin t + i \cos t \end{bmatrix} = \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} + i \begin{bmatrix} \sin t \\ \cos t \end{bmatrix}.$$

This implies that a fundamental matrix is

$$X(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix},$$

and since $X(0)$ is the identity we also find $e^{At} = X(t)$. Moreover,

$$e^{-At} = \begin{bmatrix} \cos(-t) & \sin(-t) \\ -\sin(-t) & \cos(-t) \end{bmatrix} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix},$$

and therefore

$$e^{-At} \boldsymbol{f}(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} 0 \\ \cos t \end{bmatrix} = \begin{bmatrix} -\sin t \cos t \\ \cos^2 t \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -\sin(2t) \\ 1 + \cos(2t) \end{bmatrix}.$$

Integrating we obtain

$$\int e^{-At} \boldsymbol{f}(t) dt = \frac{1}{2} \int \begin{bmatrix} -\sin(2t) \\ 1 + \cos(2t) \end{bmatrix} dt = \frac{1}{4} \begin{bmatrix} \cos(2t) \\ 2t + \sin(2t) \end{bmatrix}.$$

Then,

$$e^{At} \int e^{-At} \boldsymbol{f}(t) dt = \frac{1}{4} \begin{bmatrix} 2t \sin t + \cos t \cos(2t) + \sin t \sin(2t) \\ 2t \cos t - \sin t \cos(2t) + \cos t \sin(2t) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2t \sin t + \cos t \\ 2t \cos t + \sin t \end{bmatrix}.$$

Recalling that $\cos a \cos b + \sin a \sin b = \cos(a - b)$ and $\sin a \cos b - \cos a \sin b = \sin(a - b)$ we find

$$e^{At} \int e^{-At} \boldsymbol{f}(t) dt = \frac{1}{4} \begin{bmatrix} 2t \sin t + \cos t \\ 2t \cos t + \sin t \end{bmatrix}.$$

Therefore, the solution of the linear system is

$$\boldsymbol{x}(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 2t \sin t + \cos t \\ 2t \cos t + \sin t \end{bmatrix}.$$

Writing only the solution for $y(t) = x_1(t)$ we get

$$y(t) = (c_1 + \tfrac{1}{4}) \cos t + c_2 \sin t + \tfrac{1}{2} t \sin t,$$

and by defining $c_1 + 1/4 = b_1$ and $c_2 = b_2$ we find the general solution

$$y(t) = b_1 \cos t + b_2 \sin t + \tfrac{1}{2} t \sin t,$$

which agrees with the solution given by the method of undetermined coefficients.

Remark 5.33. This was a long computation for something that can be much more easily obtained using the method of undetermined coefficients. However, the method of variation of parameters used here has the advantage that it can be applied to cases where the non-homogeneous term of the equation is not one of the special forms that we considered in the method of undetermined coefficients. Moreover, it has the advantage that it gives an expression for the solution that can be then further analyzed. ❞

5.5 Linear Differential Equations as Linear Systems

Consider the n-th order linear differential equation

$$y^{(n)} + p_{n-1}(t)y^{(n-1)} + p_{n-2}(t)y^{(n-2)} + \cdots + p_1(t)y' + p_0(t)y = g(t).$$

This equation can be expressed as an equivalent linear system by defining $x_k = y^{(k-1)}$ for $k = 1, 2, \ldots, n$. This implies $x_k' = y^{(k)}$ and therefore we get the system of equations

$$\begin{aligned}
&x_1' = y' = x_2, \\
&x_2' = y'' = x_3, \\
&\vdots \\
&x_{n-1}' = y^{(n-1)} = x_n, \\
&x_n' = (y^{(n-1)})' = y^{(n)} = -p_{n-1}(t)y^{(n-1)} - p_{n-2}(t)y^{(n-2)} - \cdots - p_1(t)y' - p_0(t)y + g(t).
\end{aligned}$$

This system can be written in matrix form as $\boldsymbol{x}' = A(t)\boldsymbol{x} + \boldsymbol{f}(t)$, where

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y \\ y' \\ y'' \\ \vdots \\ y^{(n-2)} \\ y^{(n-1)} \end{bmatrix}, \ A(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-2} & -p_{n-1} \end{bmatrix}, \ \boldsymbol{f}(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ g(t) \end{bmatrix}.$$

Consider now specifically the homogeneous case with $n = 2$, that is, a linear homogeneous second order equation

$$y'' + p(t)y' + q(t)y = 0,$$

and the corresponding homogeneous planar linear system

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q(t) & -p(t) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

with $x_1 = y$, $x_2 = y'$. Notice that any solution $\boldsymbol{x}(t) = \langle x_1(t), x_2(t) \rangle$ of the linear system satisfies $x_2(t) = x_1'(t)$. Therefore, the solution $\boldsymbol{x}(t)$ can be completely determined by giving its first component $x_1(t) = x(t)$ and defining $x_2(t) = x'(t)$.

Proposition 5.34. *The vector valued function $\boldsymbol{x}(t) = \langle x(t), x'(t) \rangle$ is a solution of the linear system if and only if $x(t)$ is a solution of the linear equation.*

Proof. If $x(t)$ solves the linear equation then $x'' + p(t)x' + q(t)x = 0$ and therefore $x_1' = x' = x_2$ and

$$x_2' = x'' = -q(t)x - p(t)x' = -q(t)x_1 - p(t)x_2,$$

implying that $\langle x(t), x'(t)\rangle$ solves the linear system. Conversely, if $\langle x(t), x'(t)\rangle$ solves the linear system then

$$x'' + p(t)x' + q(t)x = x_2' + p(t)x_2 + q(t)x_1 = 0,$$

implying that $x(t)$ solves the linear equation. ☑

Proposition 5.35. *The vector valued functions $\boldsymbol{x}_1(t) = \langle x_1(t), x_1'(t)\rangle$ and $\boldsymbol{x}_2(t) = \langle x_2(t), x_2'(t)\rangle$ are linearly dependent on $I \subseteq \mathbf{R}$ if and only if the functions $x_1(t)$, $x_2(t)$ are linearly dependent on $I \subseteq \mathbf{R}$.*

Proof. If $\boldsymbol{x}_1(t)$, $\boldsymbol{x}_2(t)$ are linearly dependent on $I \subseteq \mathbf{R}$, then there are $c_1, c_2 \in \mathbf{R}$, not both zero, such that $c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) = \mathbf{0}$ for all $t \in I$. This implies $c_1x_1(t) + c_2x_2(t) = 0$ for all $t \in I$ and thus the functions $x_1(t)$, $x_2(t)$ are linearly dependent on $I \subseteq \mathbf{R}$.

Conversely, if the functions $x_1(t)$, $x_2(t)$ are linearly dependent on $I \subseteq \mathbf{R}$ then there are $c_1, c_2 \in \mathbf{R}$, not both zero, such that $c_1x_1(t) + c_2x_2(t) = 0$ for all $t \in I$. Differentiating the last relation we also obtain $c_1x_1'(t) + c_2x_2'(t) = 0$, and therefore $c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) = \mathbf{0}$ for all $t \in I$. This means that $\boldsymbol{x}_1(t)$, $\boldsymbol{x}_2(t)$ are linearly dependent on $I \subseteq \mathbf{R}$. ☑

Corollary 5.36. *The vector valued functions $\boldsymbol{x}_1(t) = \langle x_1(t), x_1'(t)\rangle$ and $\boldsymbol{x}_2(t) = \langle x_2(t), x_2'(t)\rangle$ are linearly independent on $I \subseteq \mathbf{R}$ if and only if the functions $x_1(t)$, $x_2(t)$ are linearly independent on $I \subseteq \mathbf{R}$.*

Let $\boldsymbol{x}_1(t) = \langle x_1(t), x_1'(t)\rangle$ and $\boldsymbol{x}_2(t) = \langle x_2(t), x_2'(t)\rangle$ be two linearly independent solutions of the linear system, which implies that the functions $x_1(t)$, $x_2(t)$ are also linearly independent. The general theory developed in this chapter asserts that the general solution of the linear system is

$$\begin{bmatrix} x(t) \\ x'(t) \end{bmatrix} = \begin{bmatrix} c_1x_1(t) + c_2x_2(t) \\ c_1x_1'(t) + c_2x_2'(t) \end{bmatrix},$$

with $c_1, c_2 \in \mathbf{R}$. If $y(t)$ is a solution of the linear equation then $\langle y(t), y'(t)\rangle$ is a solution of the linear system and thus there are $c_1, c_2 \in \mathbf{R}$ such that

$$y(t) = c_1x_1(t) + c_2x_2(t).$$

We conclude that the general solution of the linear equation is the collection of linear combinations of the linearly independent functions $x_1(t)$, $x_2(t)$, as was claimed in Theorem 3.12.

Chapter 6

Nonlinear Dynamical Systems in Two Dimensions

In a sense, everything we discussed in the previous parts of these notes was a preparation for the topic of this chapter: planar nonlinear dynamical systems.

6.1 Planar Dynamical Systems

In this chapter we consider systems of differential equations of the form $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$. The dependent variable is $\boldsymbol{x} = \langle x_1, x_2 \rangle$ taking values in $\mathbf{R}^2$. It is a function of the independent variable that we typically denote by t. The function $\boldsymbol{f} = \langle f_1, f_2 \rangle$ is a function from $\mathbf{R}^2$ to $\mathbf{R}^2$. Occasionally, we denote the components of $\boldsymbol{x}$ as $\langle x, y \rangle$ and the components of $\boldsymbol{f}$ as $\langle f, g \rangle$. Notice that we consider only the case where $\boldsymbol{f}$ does not explicitly depend on t, that is, we assume that the system of differential equations is *autonomous*.

Similarly to Chapter 2, all solutions are obtained as translations of solutions with initial condition of the form $\boldsymbol{x}(0) = \boldsymbol{x}_0$. In particular, if $\boldsymbol{x}(t)$ is the solution satisfying the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$, then the solution satisfying the initial condition $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is $\boldsymbol{x}_{t_0}(t) := \boldsymbol{x}(t - t_0)$. Moreover, we can define the *flow* of the system as the function $\phi(t, \boldsymbol{x}_0)$ which gives the value at time t of the solution $\boldsymbol{x}(t)$ with initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$.

6.1.1 Examples

In Chapter 4 we saw our first examples of planar systems of differential equations: linear systems of the form $\boldsymbol{x}' = A\boldsymbol{x}$ where A is a constant matrix. Here we discuss a few more examples. For each of these examples we also draw the *phase portrait* of the system, showing several solution curves.

Example 6.1 (Lotka-Volterra model). The Lotka-Volterra model is the archetypal example of a model for predator-prey dynamics. Suppose that in a closed habitat you have two populations. The population of rabbits (the prey) at time t is $x(t)$, while the population of foxes (the predators) at time t is $y(t)$.

In this model the dynamics of these two interacting populations is given by the planar system

$$\begin{aligned} x' &= k_1 x - p_1 xy, \\ y' &= -k_2 y + p_2 xy, \end{aligned} \tag{6.1}$$

where k_1, p_1, k_2, p_2 are strictly positive parameters.

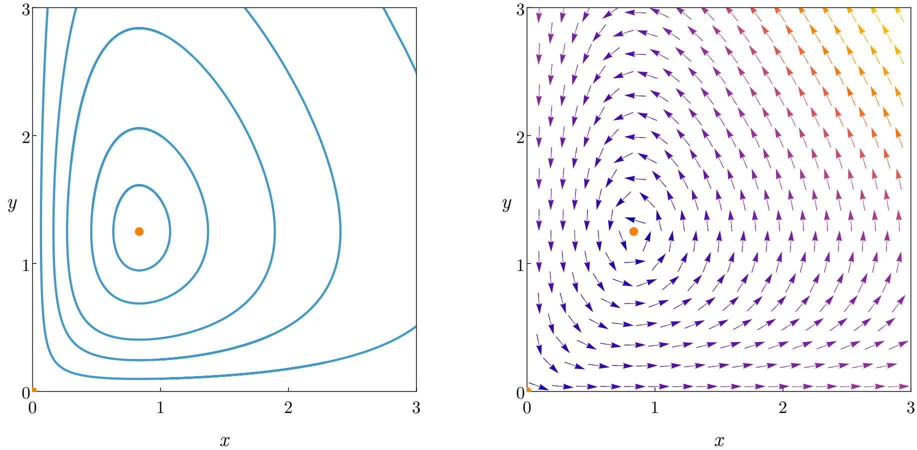


Figure 6.1: Phase portrait and vector fields for the Lotka-Volterra system.

Before analyzing the dynamics through the reduced equation, we consider some special cases. First, notice that $(0,0)$ is an equilibrium. If both populations are zero at some time t_0, then the solution is $x(t) = y(t) = 0$ for all $t \in \mathbf{R}$. Moreover, if either $x(t_0) = 0$ or $y(t_0) = 0$ then we have, respectively, $x(t) = 0$ or $y(t) = 0$ for all $t \in \mathbf{R}$.

In particular, $x(t) = 0$ solves the equation for x and gives the equation $y' = -k_2 y$ with solution $y(t) = y(t_0)e^{-k_2(t-t_0)}$ which asymptotically approaches zero as t increases. Similarly, $y(t) = 0$ solves the equation for y and gives the equation $x' = k_1 x$ with solution $x(t) = x(t_0)e^{k_1(t-t_0)}$ which increases exponentially fast with t.

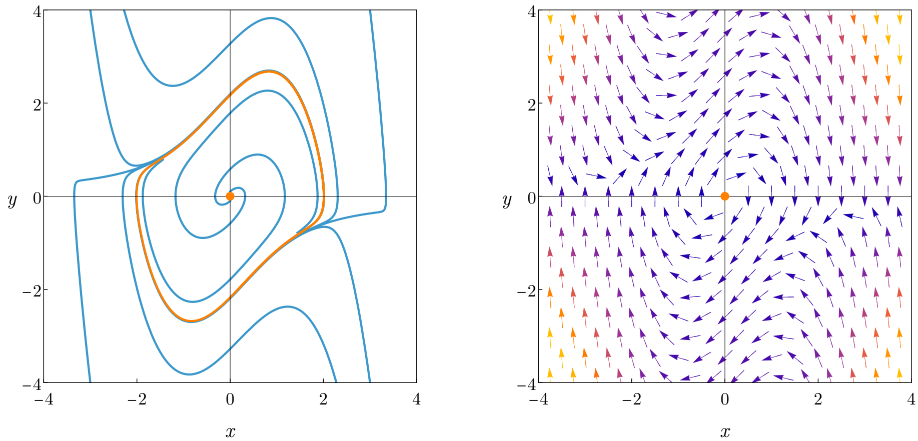


Figure 6.2: Phase portrait and vector field for the van der Pol oscillator with $\mu = 1$.

Example 6.2 (Van der Pol oscillator). The van der Pol oscillator is given by the second order nonlinear equation $x'' + x + \mu(x^2 - 1)x' = 0$, where $\mu \geq 0$. Written as a planar system, the equation for the van der Pol oscillator becomes the system

$$\begin{aligned} x' &= y, \\ y' &= -x - \mu(x^2 - 1)y. \end{aligned} \tag{6.2}$$

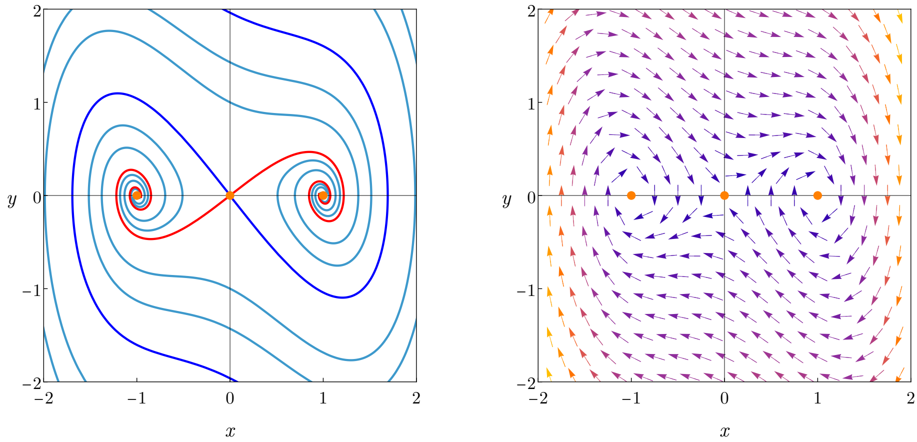


Figure 6.3: Phase portrait and vector field for the Duffing oscillator without external forcing and $b = 1/2$.

Example 6.3 (Duffing oscillator). Consider the planar system

$$\begin{aligned} x' &= y, \\ y' &= -by + x - x^3, \ \ b > 0. \end{aligned} \tag{6.3}$$

The phase portrait of the system is shown in Fig. 6.3.

6.1.2 Vector Fields and Flows

Consider the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$. Each solution $\boldsymbol{x}(t)$ traces a curve on $\mathbf{R}^2$ which is parameterized by t. The velocity vector for the curve defined by $\boldsymbol{x}(t)$ at t is $\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t))$ and it is tangent to the curve at the point $\boldsymbol{x}(t)$. Therefore, $\boldsymbol{f}(\boldsymbol{x})$ is the velocity vector for the solution curve of the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ that passes through the point $\boldsymbol{x}$. From this point of view, we can think of $\boldsymbol{f}$ as a *vector field*, that is, a function that assigns to each point $\boldsymbol{x} \in \mathbf{R}^2$ the vector $\boldsymbol{f}(\boldsymbol{x})$, and think of the solution curves of the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ as those curves on the plane for which at each point $\boldsymbol{x}$ their velocity vector is given by $\boldsymbol{f}(\boldsymbol{x})$.

In summary, we can think of systems of differential equations as vector fields and vice versa. In the context of the study of vector fields the solution curves of the corresponding differential equation $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ are called *integral curves*. In Figures 6.1 to 6.3 we show examples of the vector fields for the Lotka-Volterra model, the van der Pol oscillator, and the Duffing oscillator.

Similarly, to the discussion in Section 2.2 the *flow* associated to the continuous vector field $\boldsymbol{f} : \mathbf{R}^2 \to \mathbf{R}^2$ is the smooth function $\phi : \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2$ where $\phi(t, \boldsymbol{x}_0)$ is defined to be the unique solution to the initial value problem $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$. That is, ϕ satisfies

$$\frac{\partial \phi}{\partial t}(t, \boldsymbol{x}_0) = \boldsymbol{f}(\phi(t, \boldsymbol{x}_0)),$$

and $\phi(0, \boldsymbol{x}_0) = \boldsymbol{x}_0$. Moreover, it can be shown using similar arguments as in Section 2.2 that ϕ satisfies $\phi(t, \phi(s, \boldsymbol{x}_0)) = \phi(s + t, \boldsymbol{x}_0)$.

The reason for calling ϕ the flow associated to the vector field $\boldsymbol{f}$ is that we can think of the plane $\mathbf{R}^2$ as being covered by a thin layer of fluid and each particle of the fluid has velocity $\boldsymbol{f}(\boldsymbol{x})$ when at the point $\boldsymbol{x}$. If a particle at time $t = 0$ is at the point $\boldsymbol{x}_0$ then its motion is determined by the solution of the initial value problem $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ and therefore it is given by $\phi(t, \boldsymbol{x}_0)$.

6.1.3 Reduction to One Dimension

Consider the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$, where we write $\boldsymbol{x} = \langle x, y\rangle$ and $\boldsymbol{f} = \langle X, Y\rangle$, and a solution $\boldsymbol{x}(t) = \langle x(t), y(t)\rangle$. As we discussed earlier, each solution traces a curve on $\mathbf{R}^2$. In this section we discuss how to obtain a first-order equation that determines the *shape* of the solution curve, at the cost of throwing away the information about how the solution depends on t. The hope is that the obtained first-order equation will be solvable with some of the methods we have discussed in Section 1.2.

Assume that the solution curve is a *graph over x*. That is, the shape of this curve can be expressed as $y = h(x)$ for some — unknown for now — function h. Therefore, along the solution we have

$$y(t) = h(x(t)).$$

Differentiating both sides of this equation we find

$$y'(t) = h'(x(t))x'(t),$$

and thus

$$Y(x(t), y(t)) = h'(x(t))X(x(t), y(t)),$$

finally giving

$$Y(x(t), h(x(t))) = h'(x(t))X(x(t), h(x(t))).$$

Even though different solutions trace different curves on $\mathbf{R}^2$ and thus correspond to different functions h, all such functions are solutions of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{Y(x, y)}{X(x, y)}, \tag{6.4}$$

since any solution $y = h(x)$ of Eq. (6.4) satisfies

$$h'(x) = \frac{Y(x, h(x))}{X(x, h(x))}.$$

Equation (6.4) is called the *reduced equation* for the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$.

Remark 6.4. The method for obtaining the reduced equation is formal but not straightforward. For an easier to remember method notice that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}x/\mathrm{d}t}{\mathrm{d}y/\mathrm{d}t} = \frac{Y(x, y)}{X(x, y)}.$$

❞

Remark 6.5. If we assume that the solution curve is a graph over y, that is, if we can write $x = h(y)$, then we obtain the reduced equation

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{X(x, y)}{Y(x, y)}.$$

Very often, solution curves are, at least locally, graphs both over x and over y and therefore, we can choose the reduced equation that is easier to solve. ❞

Example 6.6 (Harmonic oscillator). The second-order equation $x'' = -\omega^2 x$ describes the motion of a mass attached to a spring and moving with no friction and no external forcing. Using the methods in Section 3.2.1 we find that the general solution is

$$x(t) = c_1 \cos(\omega t) + c_2 \sin(\omega t).$$

We transform the equation $x'' = -\omega^2 x$ to a linear system by defining $y = x'$. Then, we get the system

$$\begin{aligned} x' &= y, \\ y' &= -\omega^2 x, \end{aligned}$$

which can be solved using the methods in Chapter 4. It turns out that the system dynamics is that of a center, and therefore the solution curves are ellipses.

Here we approach the same equation by reducing to a first-order equation. The equation is the separable equation

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

Separating variables we find $y\,\mathrm{d}y = -\omega^2 x\,\mathrm{d}x$, and integrating both sides we get the implicit solution

$$\frac{1}{2}y^2 = -\frac{1}{2}\omega^2 x^2 + c.$$

We can now solve for y to get the explicit solutions

$$y = \pm\sqrt{2c - \omega^2 x^2}.$$

However, it is more enlightening to solve for c. We find

$$c = \frac{1}{2}y^2 + \frac{1}{2}\omega^2 x^2.$$

This implies that the value of the function

$$E(x, y) = \frac{1}{2}y^2 + \frac{1}{2}\omega^2 x^2,$$

is constant along each solution, and thus each solution is restricted on a level set of $E(x, y) = c$. Which level set? Clearly, the one on which the solution started. If the initial condition is $x(t_0) = x_0$, $y(t_0) = y_0$, then $c = E(x_0, y_0)$.

For each $c > 0$, the level set $E(x, y) = c$ is an ellipse whose axes are aligned with the coordinate axes x and y and the sizes of the two semiaxes are $\sqrt{2c}/\omega$ in the x direction and $\sqrt{2c}$ in the y direction. For $c = 0$, the level set consists of a single point, the origin.

Functions such as $E(x, y)$ with the property that they remain constant along solutions are called *conserved quantities* or *integrals of motion*. In this example, we see that knowing a conserved quantity allows us to understand the shape of the solutions. We come back to this idea in Section 6.3.

Example 6.7 (Lotka-Volterra model). Even though we cannot solve analytically Eq. (6.1), we can still attempt to work with the reduced equation. We find

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{(-k_2 + p_2 x)y}{(k_1 - p_1 y)x}.$$

Separating variables we find

$$\Big(\frac{k_1}{y} - p_1\Big)\,\mathrm{d}y = \Big(\frac{-k_2}{x} + p_2\Big)\,\mathrm{d}x,$$

and integrating we get

$$k_1 \log y - p_1 y = -k_2 \log x + p_2 x + c.$$

Solving for c we finally find

$$L(x, y) \coloneqq k_1 \log y - p_1 y + k_2 \log x - p_2 x = c.$$

This shows that the quantity $L(x, y)$ is conserved and therefore we can understand the shape of the solution curves of the Lotka-Volterra system through the level sets of L. The level sets of L coincide with the solution curves shown in Fig. 6.1.

6.2 Equilibria and Stability

Equilibria are important because they are organizing centers for the dynamics around them in a way that we make precise in this section.

Definition 6.8. A point $\boldsymbol{x}_e \in \mathbf{R}^2$ is an *equilibrium* of the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ if $\boldsymbol{f}(\boldsymbol{x}_e) = \mathbf{0}$. In this case, the constant function $\boldsymbol{x}(t) = \boldsymbol{x}_e$ is called *equilibrium solution.*

Example 6.9. The equilibria of the Lotka-Volterra system are found by solving

$$\begin{aligned} k_1 x - p_1 xy &= p_1 x\Big(\frac{k_1}{p_1} - y\Big) = 0, \\ -k_2 y + p_2 xy &= p_2 y\Big(-\frac{k_2}{p_2} + x\Big) = 0. \end{aligned}$$

Therefore, there are two equilibria. The origin $\langle x, y\rangle = \langle 0, 0\rangle$, and the point

$$\langle x_*, y_*\rangle = \Big\langle \frac{k_2}{p_2}, \frac{k_1}{p_1} \Big\rangle.$$

The main idea behind the definition of stable and unstable equilibria is to distinguish those equilibria for which all nearby solution curves stay near it from those for which solution there are nearby solution curves that wander away. Definition 6.11, due to Aleksandr Lyapunov, makes precise this intuitive idea. We first need to introduce the notion of an open ball around a point in $\mathbf{R}^2$.

Definition 6.10. The *open ball* with radius $\delta > 0$ centered at $\boldsymbol{x}_0 \in \mathbf{R}^2$ is the set

$$B_\delta(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbf{R}^2 : \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta\},$$

that is, the set of all points in $\mathbf{R}^2$ whose distance from $\boldsymbol{x}_0$ is less than δ.

Then the notions of stable, asymptotically stable, and unstable equilibria are defined as follows.

Definition 6.11. Consider an equilibrium $\boldsymbol{x}_e$ of the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ and denote by $\phi(t, \boldsymbol{x}_0)$ the solution of the planar system satisfying the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$.

(i) The equilibrium $\boldsymbol{x}_e$ is *stable* if for every open ball $B_\varepsilon(\boldsymbol{x}_e)$ there is a (smaller) open ball $B_\delta(\boldsymbol{x}_e)$ such that $\phi(t, \boldsymbol{x}_0) \in B_\varepsilon(\boldsymbol{x}_e)$ for all $t \geq 0$ and for all $\boldsymbol{x}_0 \in B_\delta(\boldsymbol{x}_e)$.

(ii) The equilibrium $\boldsymbol{x}_e$ is *asymptotically stable* if it is stable and if there is an open ball $B_\eta(\boldsymbol{x}_e)$ such that $\lim_{t\to\infty} \phi(t, \boldsymbol{x}_0) = \boldsymbol{x}_e$ for all $\boldsymbol{x}_0 \in B_\eta(\boldsymbol{x}_e)$.

(iii) The equilibrium $\boldsymbol{x}_e$ is *unstable* if it is not stable. That is, the equilibrium $\boldsymbol{x}_e$ is unstable if there is an open ball $B_{\varepsilon_0}(\boldsymbol{x}_e)$ such that for all smaller open balls $B_\delta(\boldsymbol{x}_e)$ there is some $\boldsymbol{x}_\delta \in B_\delta(\boldsymbol{x}_e)$ and some $t_\delta > 0$ for which $\phi(t_\delta, \boldsymbol{x}_\delta) \notin B_{\varepsilon_0}(\boldsymbol{x}_e)$.

Remark 6.12. In Chapter 4 we studied in detail and classified the different types of dynamics of planar linear systems. You should be able to check that centers are stable but not asymptotically stable, stable nodes and stable spirals are asymptotically stable, while unstable nodes, unstable spirals, and saddles are unstable.

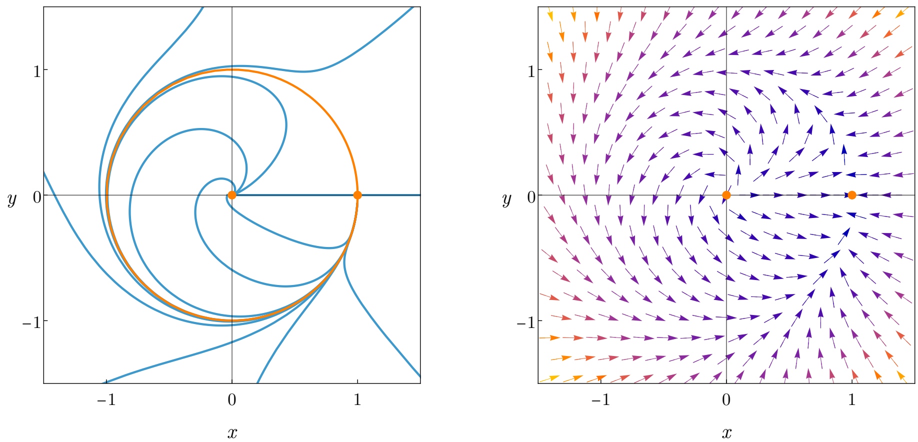


Figure 6.4: Phase portrait and vector field for the system in Example 6.13 which has an equilibrium at $(1, 0)$ that is globally attracting but unstable.

Example 6.13. The definition of asymptotically stable equilibria imposes first the requirement that the equilibrium is stable before asking that all nearby orbits approach the equilibrium as t increases. At first glance, the requirement that the equilibrium appears unnecessary however it is there to guard against the possibility that a solution curve goes away from the equilibrium before eventually coming back. The following example shows that this possibility is real.

Consider the planar system which in polar coordinates has the form $r' = r(1 - r^2)$, $\theta' = 1 - \cos\theta$. The system has two equilibria, the origin with $r = 0$, and the point $(r, \theta) = (1, 0)$ corresponding to Cartesian coordinates $(x, y) = (1, 0)$. The phase portrait of this system is shown in Fig. 6.4 — we discuss later in this chapter how to piece together the dynamics, but for now we rely on numerically integrating the differential equations with Mathematica.

We observe that if we consider the open ball $B_{1/2}(1, 0)$ then for every smaller open ball $B_\delta(1, 0)$ there is a point on the circle $r = 1$ with $0 < \theta < \pi/2$ which lies inside $B_\delta(1, 0)$ but after some time will get out of $B_{1/2}(1, 0)$. Therefore, the equilibrium is unstable.

However — and this is where it gets interesting — all solution curves of the system, except the equilibrium solution at the origin, approach $(1, 0)$ as t increases! Despite this, we do not call $(1, 0)$ stable or asymptotically stable. One should think of stability as a local property. If, locally, there is a single solution curves that goes away from the equilibrium then we characterize the equilibrium as unstable, even if (much) later the same solution curve comes back toward the equilibrium.

6.2.1 Linearized Dynamics

Consider a planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ for which $\boldsymbol{x}_e$ is an equilibrium. We linearize the dynamics near $\boldsymbol{x}_e$ so as to obtain an approximation of the solution curves. For this purpose let $\boldsymbol{x} = \boldsymbol{x}_e + \boldsymbol{\xi}$. Then

$$\boldsymbol{\xi}' = \boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_e + \boldsymbol{\xi}).$$

We consider now the Taylor polynomial of $\boldsymbol{f}(\boldsymbol{x}_e + \boldsymbol{\xi})$ at $\boldsymbol{\xi} = \boldsymbol{0}$.

Derivatives and Taylor Polynomials

We recall some basic background material from Multivariable Calculus and we refer to [5] for details. Given a function $\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}^n$, the derivative $D\boldsymbol{f}(\boldsymbol{x})$ of $\boldsymbol{f}$ at $\boldsymbol{x}$ can be represented — when it exists — by the $n \times n$ matrix

$$D\boldsymbol{f}(\boldsymbol{x}) = \left[\frac{\partial f_i}{\partial x_j}\right]_{i,j}(\boldsymbol{x}).$$

There are here subtle questions on the existence of the derivative but we bypass them by assuming that all partial derivatives $\partial f_i/\partial x_j$ are continuous in an open ball centered at $\boldsymbol{x}$. When $n = 2$ the expression for the derivative reads

$$D\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \partial f_1/\partial x_1\ (\boldsymbol{x}) & \partial f_1/\partial x_2\ (\boldsymbol{x}) \\ \partial f_2/\partial x_1\ (\boldsymbol{x}) & \partial f_2/\partial x_1\ (\boldsymbol{x}) \end{bmatrix}.$$

For each component function f_j of $\boldsymbol{f}$ the corresponding first Taylor polynomial of $\boldsymbol{f}$ around a point $\boldsymbol{x}$ with its remainder term is

$$f_j(\boldsymbol{x} + \boldsymbol{\xi}) = f_j(\boldsymbol{x}) + \sum_{k=1}^{n} \frac{\partial f_j}{\partial x_k}(\boldsymbol{x})\xi_k + \frac{1}{2}\sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial^2 f_j}{\partial x_k\, \partial x_l}(\boldsymbol{x} + c\,\boldsymbol{\xi})\xi_k\xi_l,$$

for some $c \in (0, 1)$. Then we can write

$$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{\xi}) = \boldsymbol{f}(\boldsymbol{x}) + D\boldsymbol{f}(\boldsymbol{x})\boldsymbol{\xi} + O(\|\boldsymbol{\xi}\|^2). \tag{6.5}$$

Assuming that $\boldsymbol{f}$ is twice continuously differentiable and applying Eq. (6.5) at $\boldsymbol{x}_e$ we obtain

$$\boldsymbol{f}(\boldsymbol{x}_e + \boldsymbol{\xi}) = \boldsymbol{f}(\boldsymbol{x}_e) + D\boldsymbol{f}(\boldsymbol{x}_e)\boldsymbol{\xi} + O(\|\boldsymbol{\xi}\|^2).$$

Since $\boldsymbol{x}_e$ is an equilibrium we have $\boldsymbol{f}(\boldsymbol{x}_e) = \boldsymbol{0}$. Ignoring the $O(\|\boldsymbol{\xi}\|^2)$ terms and combining everything together we obtain the *linearized* approximate equation

$$\boldsymbol{\xi}' = D\boldsymbol{f}(\boldsymbol{x}_e)\boldsymbol{\xi}. \tag{6.6}$$

Equation (6.6) describes a planar linear system with constant matrix $A = D\boldsymbol{f}(\boldsymbol{x}_e)$. We emphasize that this is an approximate equation. In Section 6.2.2 we discuss under which conditions the linearized dynamics described by Eq. (6.6) can be applied to the original dynamics near $\boldsymbol{x}_e$.

Example 6.14 (Duffing oscillator). Consider the planar system

$$\begin{aligned} x' &= y, \\ y' &= -by + x - x^3,\ b > 0. \end{aligned} \tag{6.7}$$

The equilibria of the Duffing oscillator are $(0,0)$, $(-1,0)$, and $(1,0)$.

To determine the linearized dynamics at each equilibrium we first compute the derivative $D\boldsymbol{f}(\boldsymbol{x})$ at arbitrary $\boldsymbol{x} \in \mathbf{R}^2$ — here, $\boldsymbol{x} = \langle x, y\rangle$, $\boldsymbol{f} = \langle y, -by + x - x^3\rangle$. We compute

$$D\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} 0 & 1 \\ 1-3x^2 & -b \end{bmatrix}.$$

For the linearized dynamics at $\boldsymbol{x}_1 = (0,0)$, with $\boldsymbol{\xi} = \boldsymbol{x}$ we have $\boldsymbol{\xi}' = A_1\boldsymbol{\xi}$ with

$$A_1 \coloneqq D\boldsymbol{f}(\boldsymbol{x}_1) = \begin{bmatrix} 0 & 1 \\ 1 & -b \end{bmatrix}.$$

We have $\operatorname{tr} A_1 = -b < 0$ and $\det A_1 = -1 < 0$. Therefore, the linearized dynamics is that of a saddle. The stable and unstable directions are given by the corresponding eigenvectors of A_1.

For the linearized dynamics at $\boldsymbol{x}_2 = (1,0)$, with $\boldsymbol{\xi} = \boldsymbol{x} - \boldsymbol{x}_2 = \langle x-1, y\rangle$, we have $\boldsymbol{\xi}' = A_2\boldsymbol{\xi}$ with

$$A_2 \coloneqq D\boldsymbol{f}(\boldsymbol{x}_2) = \begin{bmatrix} 0 & 1 \\ -2 & -b \end{bmatrix}.$$

We have $\operatorname{tr} A_2 = -b < 0$ and $\det A_1 = 2 > 0$. Since $\operatorname{tr} A_2 < 0$, the linearized dynamics is stable. To distinguish between stable node and stable spiral we check the sign of $(\operatorname{tr} A_2)^2 - 4\det A_2 = b^2 - 8$. When $b^2 > 8$ then the linearized dynamics is that of a stable node and when $b^2 < 8$ it is that of stable spiral. In the case of a spiral the rotation is clockwise since the bottom left element of A_2 is $-2 < 0$.

Finally, for the linearized dynamics at $\boldsymbol{x}_3 = (-1,0)$, with $\boldsymbol{\xi} = \boldsymbol{x} - \boldsymbol{x}_3 = \langle x+1, y\rangle$, we have $\boldsymbol{\xi}' = A_3\boldsymbol{\xi}$ with $A_3 = A_2$. Therefore, the linearized dynamics near $\boldsymbol{x}_3$ is the same as the linearized dynamics near $\boldsymbol{x}_2$.

Example 6.15 (Simple pendulum). We consider a simple pendulum hanging from a point through a massless rod with length ℓ. We denote by θ the angle between the direction of the rod and the downward vertical direction. Then the equation describing the pendulum dynamics is

$$\theta'' = -\gamma^2 \sin\theta, \quad \gamma^2 = \frac{g}{\ell}.$$

To transform this equation into a planar system, let $x = \theta$, $y = \theta'$. Then

$$\begin{aligned} x' &= y, \\ y' &= -\gamma^2 \sin x. \end{aligned} \tag{6.8}$$

The equilibria of the system in Equation (6.8) are given by solving $y = 0$ and $\sin x = 0$, that is, the equilibria form the set $\{(k\pi, 0), k \in \mathbf{Z}\} \subseteq \mathbf{R}^2$. Notice, however, that since x is an angle, $x + 2n\pi \equiv x$. This means that there are two distinct equilibria. The origin $(0,0)$ corresponding to the pendulum being at rest in its natural downward position, and the point $(\pi,0) \equiv (-\pi,0)$ corresponding to the pendulum being in the uppermost position, which from our experience is unstable. We analyze the linear dynamics of the pendulum near these two equilibria. The matrix $D\boldsymbol{f}$ at an arbitrary point (x,y) is:

$$D\boldsymbol{f}(x,y) = \begin{bmatrix} 0 & 1 \\ -\gamma^2\cos x & 0 \end{bmatrix}.$$

For the equilibrium at $(0, 0)$ we find

$$A_0 := D\boldsymbol{f}(0,0) = \begin{bmatrix} 0 & 1 \\ -\gamma^2 & 0 \end{bmatrix},$$

with eigenvalues $r = \pm i\gamma$. Therefore, the linearized equations correspond to a *center*. Since the equilibrium is not hyperbolic we cannot apply Theorem 6.17. In particular, we cannot deduce that the equilibrium is also a center for the original dynamics given by 6.8. In Section 6.3, we will however find that when we have a linear center in a planar system with a conserved quantity then the equilibrium is also a center in the original nonlinear system.

For the equilibrium at $(\pi, 0)$ we find

$$A_1 := D\boldsymbol{f}(\pi,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

with eigenvalues $r = \pm 1$. Therefore, the linearized equilibrium is a saddle and Theorem 6.17 ensures that it is a saddle also for the nonlinear dynamics in Eq. (6.8).

Example 6.16 (Weakly unstable spiral). Consider the system

$$\begin{aligned} x' &= -y + x(x^2 + y^2), \\ y' &= x + y(x^2 + y^2). \end{aligned}$$

The origin $\boldsymbol{x}_0 = \boldsymbol{0}$ is an equilibrium for the given system. We compute

$$A = D\boldsymbol{f}(\boldsymbol{0}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The matrix A has eigenvalues $\pm i$ (alternatively, check that $\operatorname{tr} A = 0$ and $\det A > 0$), and therefore the origin is a center for the linear system.

What about the dynamics for the full system? This is easiest to study in polar coordinates where we have

$$rr' = xx' + yy' = (x^2 + y^2)^2 = r^4, \; r^2\theta' = xy' - x'y = x^2 + y^2 = r^2.$$

Therefore, $r' = r^3$, $\theta' = 1$. Since for $r \neq 0$ we have $r' > 0$, $r(t)$ increases with t and therefore the origin is weakly unstable. Here, the characterization *weakly unstable* means that the solution curves initially move away from the origin slower than exponentially — even though the solution blows up in finite time, here we are interested in what happens very close to the equilibrium.

6.2.2 Hartman-Grobman Theorem

Under what conditions we can be sure that the approximate dynamics in the linearized system offer a good representation of the actual dynamics of the system near the equilibrium?

Theorem 6.17 (Hartman-Grobman). *Consider the planar linear system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$, where $\boldsymbol{f}$ has continuous partial derivatives, and assume that $\boldsymbol{f}(\boldsymbol{0}) = \boldsymbol{0}$ and that the matrix $A = D\boldsymbol{f}(\boldsymbol{0})$ has no eigenvalue with zero real part. Then there exists a homeomorphism H of an open set U containing the origin onto an open set V containing the origin such that for each $\boldsymbol{x}_0 \in U$ there is an open interval $I_0 \subseteq \mathbf{R}$ containing zero such that for all $\boldsymbol{x}_0 \in U$ and $t \in I_0$*

$$H \circ \varphi^t_{\boldsymbol{f}}(\boldsymbol{x}_0) = e^{At} H(\boldsymbol{x}_0), \tag{6.9}$$

i.e., H maps solution curves of $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ near the origin onto solution curves of the linear system $\boldsymbol{x}' = A\boldsymbol{x}$ near the origin and preserves the parameterization by time. If, additionally, $\boldsymbol{f}$ has continuous second partial derivatives then H can be chosen to be a diffeomorphism.

Remark 6.18. Theorem 6.17 generalizes to systems $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ except for the last statement: there are examples of systems in $\mathbf{R}^3$ with polynomial $\boldsymbol{f}$ for which H cannot be chosen to be a diffeomorphism, see [7]. ❞

6.3 Conservative Systems and Energy Method

Second order equations of the form $x'' = f(x)$ commonly appear in physical problems when the force exerted on a point particle depends only on the particle's position. Writing $x' = y$ we obtain the linear system

$$\begin{aligned} x' &= y, \\ y' &= f(x). \end{aligned} \tag{6.10}$$

Given their origin in problems from classical mechanics, such systems are often called *mechanical.*

Either by integrating the separable reduced equation $\mathrm{d}y/\mathrm{d}x = f(x)/y$, or through a direct computation, we can check that the quantity

$$E(x,y) = \frac{1}{2}y^2 + U(x), \tag{6.11}$$

is a conserved quantity, where $U(x)$ is an anti-derivative of $-f(x)$.

In the context of classical mechanics, $U(x)$ is called *potential energy*, $K(y) := \frac{1}{2}y^2$ is called *kinetic energy*, and their sum $E(x,y) = K(y) + U(x)$ is called the *total energy* or *mechanical energy* of the system.

We describe a systematic, graphical, method for finding the level sets of the mechanical energy $E(x,y)$ for arbitrary potential energy $U(x)$. This allows us to draw the phase portraits of planar mechanical systems and understand their dynamics without computations.

Suppose that we want to draw the level set of $E(x,y)$ for a given value h. That is, we want to find the points (x,y) that satisfy

$$\frac{1}{2}y^2 + U(x) = h.$$

Solving for y we find

$$y = \pm v_h(x) := \pm\sqrt{2(h - U(x))}. \tag{6.12}$$

We observe the following.

(i) The expression in Equation (6.12) is defined only when $U(x) \le h$. Therefore, to draw the corresponding level set $E(x,y) = h$ we first identify the $x \in \mathbf{R}$ for which $U(x) \le h$.

(ii) If there is a point x_0 where $U(x_0) = h$, then $v_h(x_0) = 0$ and thus such x_0 corresponds to the point $(x_0, 0)$ in the xy plane.

(iii) If for some point x we have $U(x) < h$, then $v_h(x) > 0$ and thus such x correspond to two points $(x, \pm v_h(x))$ on the xy plane, which are symmetric with respect to reflection through the horizontal axis.

(iv) The value of $v_h(x)$ increases when the value of $|h - U(x)|$ increases. Notice that $|h - U(x)|$ gives the vertical distance at x between the graph of U and the horizontal curve $U = h$.

We apply new these observations to deduce the level sets of the energy $E(x, y)$ for several examples of mechanical systems.

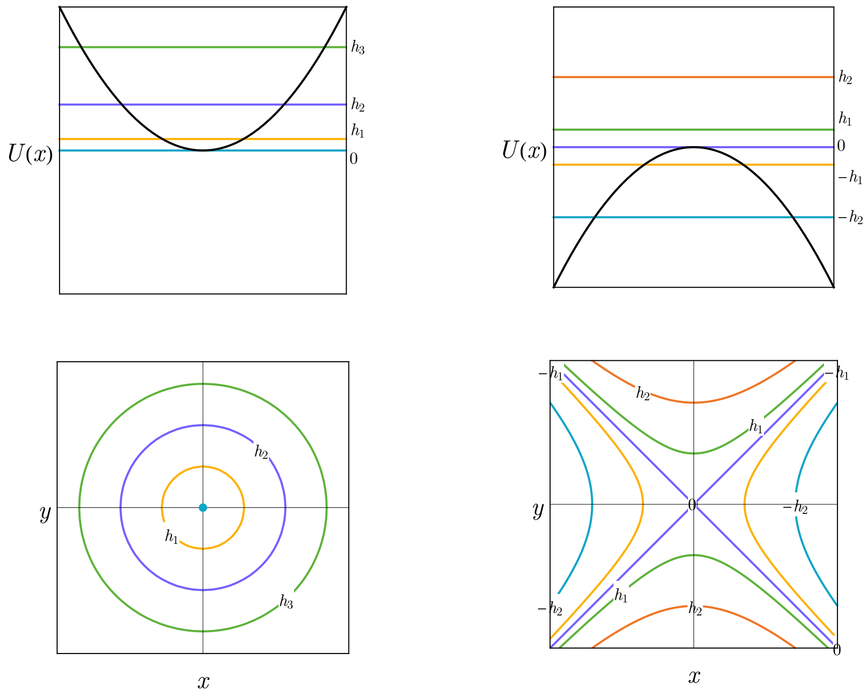


Figure 6.5: Level sets of the mechanical energy $E(x, y) = \frac{1}{2}y^2 + U(x)$. The left two panels correspond to $U(x) = \frac{1}{2}x^2$, Example 6.19. The right two panels correspond to $U(x) = -\frac{1}{2}x^2$, Example 6.20. For each system, the top panels show the graph of $U(x)$ in black and several energy levels. The energy method can be used to produce the corresponding level sets of $E(x, y)$ shown in the bottom panels.

Example 6.19 (Harmonic oscillator $U(x) = \frac{1}{2}\omega^2 x^2$). Consider the potential function

$$U(x) = \frac{1}{2}\omega^2 x^2,$$

corresponding to the second order equation $x'' = -\omega^2 x$. The standard physical motivation for the study of this equation is that it describes the motion of a mass m attached to a spring with Hooke constant k, and $\omega^2 := k/m > 0$. The importance of this equation is much larger though: it is a local approximation near non-degenerate local minima of the potential energy of mechanical systems.

Since the energy function is $E(x, y) = \frac{1}{2}y^2 + \frac{1}{2}\omega^2 x^2$ its level sets are ellipses when $E(x, y) = h > 0$ while the level set $E(x, y) = 0$ consists only of the origin.

Alternatively, the energy method can be used to determine the main qualitative features of the level sets without directly using the algebraic expression for $E(x, y)$ and using only the shape of the graph of the potential $U(x)$. For example, for $h = 0$ the horizontal line $U = 0$ intersects the graph of $U(x)$ only at $x = 0$. Therefore, the corresponding level set is the point $(0, 0)$. For $h > 0$, the horizontal line $U = h$ intersects the graph of $U(x)$ at two points $x_1 < 0 < x_2$. Only for $x \in [x_1, x_2]$ we have $U(x) \leq h_1$. Therefore, one part of the corresponding level set at the upper half of the xy plane can be drawn as a curve that starts at $(x_1, 0)$ and meets again the horizontal axis at $(x_2, 0)$, while the vertical distance of the curve from the horizontal axis is larger for larger vertical distance of $U(x)$ from h_1 — as x moves from x_1 to x_2, the value of y first increases until it reaches a maximum value and the decreases again until it becomes zero when $x = x_2$. The second part of the curve is then the reflection of the first part through the horizontal axis. Together, the two parts produce a closed curve that looks like an ellipse. Actually, in this example the level sets for $h > 0$ are exactly ellipses.

Example 6.20 (Repulsive potential $U(x) = -\frac{1}{2}\alpha^2 x^2$). Consider the potential function

$$U(x) = -\frac{1}{2}\alpha^2 x^2.$$

This example is important since it is a local approximation near non-degenerate local maxima of the potential energy of mechanical systems.

For fixed $h \neq 0$ the level sets $E(x, y) = h$ are hyperbolas defined by $y^2 - \alpha^2 x^2 = 2h$. For $h = 0$, the level set consists of the two lines $y = \pm\alpha x$ intersecting at the origin.

Alternatively, one can use the energy method to produce a qualitative picture of the level sets. We leave it to the reader to explain how the energy method produces the level sets of the system with $U(x) = -\frac{1}{2}x^2$ shown in Fig. 6.5.

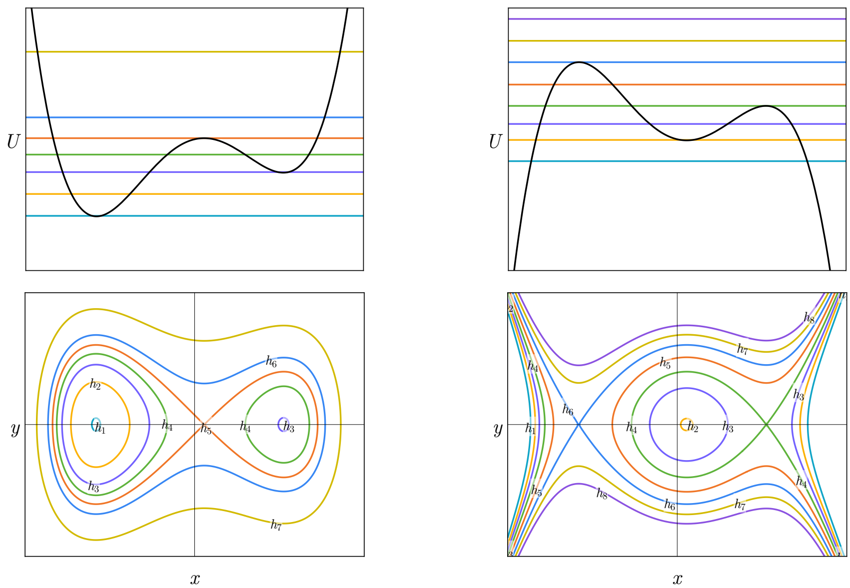


Figure 6.6: Level sets of the mechanical energy $E(x, y) = \frac{1}{2}y^2 + U_\pm(x)$ in Example 6.21. The left two panels correspond to $U_+(x)$ and the right two panels correspond to $U_-(x)$. For each system, the top panels show the graph of $U(x)$ in black and several energy levels. The energy method can be used to produce the corresponding level sets of $E(x, y)$ shown in the bottom panels.

Example 6.21. We consider two more examples, where the potential energy is given by

$$U_{\pm}(x) = \pm\Big(-\frac{1}{2}x^2 + \tfrac{1}{4}x^4 + \tfrac{1}{10}x\Big).$$

The level sets of the corresponding energy $E_{\pm}(x,y) = \frac{1}{2}y^2 + U_{\pm}(x)$ is shown in Fig. 6.6.

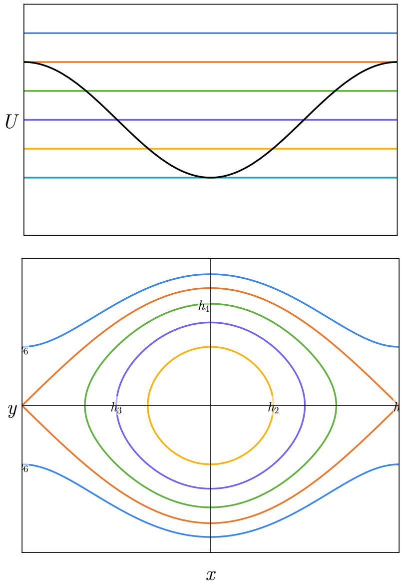


Figure 6.7: Level sets of the mechanical energy of the pendulum $E(x,y) = \frac{1}{2}y^2 - \gamma^2\cos x$ in Example 6.22.

Example 6.22 (Simple pendulum $U(x) = -\gamma^2\cos x$). We finish these examples with another important physical system. The simple pendulum is given by the equation $x'' = -\gamma^2\sin x$, with $\gamma = \sqrt{g/\ell}$, where g is the gravitational constant and ℓ is the length of the rod to which the pendulum is attached. Therefore, we can choose $U(x) = -\gamma^2\cos x$ and the mechanical energy is

$$E(x,y) = \frac{1}{2}y^2 - \gamma^2\cos x.$$

The level sets of $E(x,y)$ are shown in Fig. 6.7. Recall that in Example 6.15 we found that the linearized dynamics near $(0,0)$ is that of a center, but we could not use the Hartman-Grobman theorem to deduce something about the nonlinear dynamics. Figure 6.7 shows that the nonlinear dynamics near the origin is also that of a center.

Exercise 6.1. Describe the pendulum motion for each of the different level sets shown in Fig. 6.7.

We now consider equilibria of mechanical systems. Equilibria satisfy $x' = 0$ and $y' = 0$, giving the equations $y = 0$ and $f(x) = 0$. Since $U'(x) = -f(x)$, equilibria correspond to points x_e where $U'(x_e) = 0$. An equilibrium $(x_e, 0)$ is called *non-degenerate* if $U''(x_e) \neq 0$.

Let $(x_e, 0)$ be an equilibrium of a mechanical system. The corresponding linearization is given by $\boldsymbol{x}i' = A\boldsymbol{\xi}$ where

$$A = \begin{bmatrix} 0 & 1 \\ f'(x_e) & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -U''(x_e) & 0 \end{bmatrix}.$$

Assuming that the equilibrium $(x_e, 0)$ is non-degenerate we consider two cases.

$\boldsymbol{U''(x_e) < 0}$**.** In this case the potential energy has a local maximum at x_e. We define $\alpha > 0$ by $U''(x_e) = -\alpha^2$. The matrix A has eigenvalues $\pm\alpha$, that is, the equilibrium is a saddle. The dynamics near the equilibrium is the same as the dynamics discussed in Example 6.20.

$\boldsymbol{U''(x_e) > 0}$**.** In this case the potential energy has a local minimum at x_e. We define $\omega > 0$ by $U''(x_e) = \omega^2$. The matrix A has eigenvalues $\pm i\omega$, that is, the linearized dynamics at the equilibrium is that of a center. We cannot use the Hartman-Grobman theorem to deduce anything about the nonlinear dynamics near the equilibrium. However, the existence of the conserved quantity

$$E(x, y) = \frac{1}{2}y^2 + U(x) = \frac{1}{2}y^2 + \frac{1}{2}\omega^2 x^2 + O(x^3),$$

implies that solution curves near the equilibrium are closed and therefore the nonlinear dynamics is also that of a center.

6.4 Lyapunov's Direct Method

Luapunov's results on stability concern isolated equilibria, which are formally defined as follows.

Definition 6.23. An equilibrium $\boldsymbol{x}_e$ of the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ is *isolated* if there is an open disk $B_\varepsilon(\boldsymbol{x}_e)$ of radius $\varepsilon > 0$ containing no other equilibria of the planar system.

Definition 6.24. Let $W : \mathbf{R}^n \to \mathbf{R}$ be a function which is continuous for all $\boldsymbol{x}$ in some disk, D, centered at $\mathbf{0}$, with $W(\mathbf{0}) = 0$. Let D_* denote the punctured disk $D \setminus \{\mathbf{0}\}$.

(i) The function $W(\boldsymbol{x})$ is *positive definite* on D if $W(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in D_*$.
(ii) The function $W(\boldsymbol{x})$ is *positive semidefinite* on D if $W(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in D$.
(iii) The function $W(\boldsymbol{x})$ is *negative definite* on D if $W(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \in D_*$.
(iv) The function $W(\boldsymbol{x})$ is *negative semidefinite* on D if $W(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x} \in D$.

Consider a planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ and suppose that $\boldsymbol{x}(t)$ is a solution curve. Moreover, consider a differentiable function $V(\boldsymbol{x})$. We want to find how fast the value of V changes as we move along the solution curve, that is, we want to evaluate the derivative with respect to t of the function composition $V \circ \boldsymbol{x}$. We have

$$\begin{aligned} (V \circ \boldsymbol{x})'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\big[V(\boldsymbol{x}(t))\big] &= \frac{\partial V}{\partial x_1}(\boldsymbol{x}(t))\frac{\mathrm{d}x_1(t)}{\mathrm{d}t} + \frac{\partial V}{\partial x_2}(\boldsymbol{x}(t))\frac{\mathrm{d}x_2(t)}{\mathrm{d}t} \\ &= \boldsymbol{x}'(t) \cdot \nabla V(\boldsymbol{x}(t)) = \boldsymbol{f}(\boldsymbol{x}(t)) \cdot \nabla V(\boldsymbol{x}(t)). \end{aligned} \tag{6.13}$$

We can express the last result more concisely in the following way. Given a function $V : \mathbf{R}^n \to \mathbf{R}$ and the system of differential equations $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ with $\boldsymbol{f} : \mathbf{R}^n \to \mathbf{R}^n$, define a new function $\dot{V} : \mathbf{R}^n \to \mathbf{R}$ by

$$\dot{V}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) \cdot \nabla V(\boldsymbol{x}).$$

Then, Eq. (6.13) implies that $(V \circ \boldsymbol{x})' = \dot{V} \circ \boldsymbol{x}$, that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[V(\boldsymbol{x}(t))\big] = \dot{V}(\boldsymbol{x}(t)).$$

Therefore, when $\dot{V}$ is evaluated at a point $\boldsymbol{x}(t)$ along a solution curve, it gives the rate of change of V along the solution curve at the point $\boldsymbol{x}(t)$.

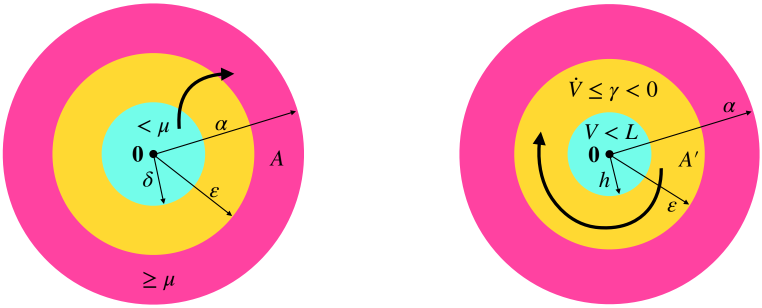


Figure 6.8: Proof of Theorem 6.26. Left: To show that the origin is stable, we construct two regions, an inner region with $V < \mu$ and an outer region with $V \geq \mu$. Since $\dot{V} \leq 0$, a solution curve cannot go from the inner region to the outer region. Right: To show that the origin is asymptotically stable, we first assume the opposite and construct a region in which solution curves remain trapped for all $t \geq 0$ and such that $\dot{V} \leq \gamma < 0$. Then, for sufficiently large t the value of V should become negative, leading to a contradiction.

Remark 6.25. The expression $\boldsymbol{f}(\boldsymbol{x}) \cdot \nabla V(\boldsymbol{x})$ is the directional derivative of V at $\boldsymbol{x}$ along the direction of the vector $\boldsymbol{f}(\boldsymbol{x})$. ❞

We now have all the terminology in place to state Lyapunov's stability theorem.

Theorem 6.26 (Lyapunov's stability theorem). *Let $V(\boldsymbol{x})$ be a positive definite function on an open set D which contains $\mathbf{0}$ and assume that $\mathbf{0}$ is an isolated equilibrium of the planar system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$.*

(i) If the function $\dot{V}(\boldsymbol{x})$ is negative semidefinite on D, then $\mathbf{0}$ is stable.
(ii) If the function $\dot{V}(\boldsymbol{x})$ is negative definite on D, then $\mathbf{0}$ is asymptotically stable.

Remark 6.27. A function V satisfying the conditions of (either part of) Theorem 6.26 is often called a *Lyapunov function*. ❞

Proof of Theorem 6.26. (i) Assume that V satisfies the conditions of the theorem and that $\dot{V}$ is negative semidefinite on D. To show that the origin is stable we need to show that for any given $B_\varepsilon(\mathbf{0})$ there is a neighborhood $B_\delta(\mathbf{0})$ such that $\boldsymbol{x}(0) \in B_\delta(\mathbf{0})$ implies $\boldsymbol{x}(t) \in B_\varepsilon(\mathbf{0})$ for all $t \geq 0$, that is, if $\|\boldsymbol{x}(0)\| < \delta$ then $\|\boldsymbol{x}(t)\| < \varepsilon$ for all $t \geq 0$.

Consider a closed disk $\overline{B}_\alpha(\mathbf{0}) = \{\boldsymbol{x} \in \mathbf{R}^2 : \|\boldsymbol{x}\| \leq \alpha\}$ of radius $\alpha > 0$ such that $\overline{B}_\alpha(\mathbf{0}) \subseteq D$.

Then, for any ε with $0 < \varepsilon < \alpha$ consider the annulus

$$A = \{\boldsymbol{x} \in \mathbf{R}^2 : \varepsilon \leq \|\boldsymbol{x}\| \leq \alpha\}.$$

The set A is a closed and bounded subset of $\mathbf{R}^2$. Since the function V is continuous on A, and A is closed and bounded, V attains its minimum value on A, that is, there is some point $\boldsymbol{x}_m \in A$ such that $\mu = \min_{\boldsymbol{x} \in A} V(\boldsymbol{x}) = V(\boldsymbol{x}_m)$. Since V is positive definite and $\boldsymbol{x}_m \neq \mathbf{0}$ we have $\mu > 0$. This implies that for all $\boldsymbol{x} \in A$ we have $V(\boldsymbol{x}) \geq \mu > 0$.

Select δ such that $0 < \delta < \varepsilon$ and $V(\boldsymbol{x}) < \mu$ for all $\boldsymbol{x} \in B_\delta(\mathbf{0})$. Such δ exists since $V(\mathbf{0}) = 0$ and V is continuous.

Suppose that $\boldsymbol{x}(t)$ is a solution with $\boldsymbol{x}(0) \in B_\delta(\mathbf{0})$ such that $\boldsymbol{x}(t)$ does not stay in $B_\varepsilon(\mathbf{0})$ for all $t \geq 0$, that is, there is t_1 such that $\|\boldsymbol{x}(t_1)\| = \varepsilon$, which implies, $\boldsymbol{x}(t_1) \in A$. Since

$\boldsymbol{x}(0) \in B_\delta(\mathbf{0})$ we have $V(\boldsymbol{x}(0)) < \mu$ and since $\boldsymbol{x}(t_1) \in A$ we have $V(\boldsymbol{x}(t_1)) \geq \mu$. In particular, $V(\boldsymbol{x}(0)) < V(\boldsymbol{x}(t_1))$.

We observe that for any $t_1 \geq t_0 \geq 0$ $\boldsymbol{x}(t_0), \boldsymbol{x}(t_1) \in D$ we have

$$V(\boldsymbol{x}(t_1)) - V(\boldsymbol{x}(t_0)) = \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t}[V(\boldsymbol{x}(t))]\,\mathrm{d}t = \int_{t_0}^{t_1} \dot{V}(\boldsymbol{x}(t))\,\mathrm{d}t.$$

Since $\dot{V}$ is negative semidefinite, applying the last relation with $t_0 = 0$ we find

$$V(\boldsymbol{x}(t_1)) - V(\boldsymbol{x}(0)) = \int_0^{t_1} \dot{V}(\boldsymbol{x}(t))\,\mathrm{d}t \leq 0,$$

and thus $V(\boldsymbol{x}(t_1)) \leq V(\boldsymbol{x}(0))$, contradicting that $V(\boldsymbol{x}(0)) < V(\boldsymbol{x}(t_1))$. Therefore, we conclude that a solution $\boldsymbol{x}(t)$ with $\boldsymbol{x}(0) \in B_\delta(\mathbf{0})$ stays in $B_\varepsilon(\mathbf{0})$ for all $t \geq 0$.

(ii) Assume now that $\dot{V}$ is negative definite on D. Of course, being negative definite also implies that it is negative semidefinite and thus $\mathbf{0}$ is stable. To show that $\mathbf{0}$ is asymptotically stable we want to show that there is $\eta > 0$ such that if $\boldsymbol{x}(0) \in B_\eta(\mathbf{0})$ then $\|\boldsymbol{x}(t)\| \to 0$ as $t \to \infty$. Since V is continuous and positive definite we have that $\|vvx(t)\| \to 0$ as $t \to \infty$ if and only if $V(\boldsymbol{x}(t)) \to 0$ as $t \to \infty$. Therefore, we need that find $\eta > 0$ such that if $\boldsymbol{x}(0) \in B_\eta(\mathbf{0})$ then $V(\boldsymbol{x}(t)) \to 0$ as $t \to \infty$.

Following the notation of the first part of the proof, let $\varepsilon = \alpha/2$, choose $0 < \delta < \varepsilon$ be in the same way as earlier, so that for $\boldsymbol{x}(0) \in B_\delta(\mathbf{0})$ we have $\boldsymbol{x}(t) \in B_\varepsilon(\mathbf{0})$ for all $t \geq 0$, and then choose $\eta = \delta$.

Since $\dot{V}$ is negative definite the value of $V(\boldsymbol{x}(t))$ is strictly decreasing with t, since for $t_1 > t_0$ we have

$$V(\boldsymbol{x}(t_1)) - V(\boldsymbol{x}(t_0)) = \int_{t_0}^{t_1} \dot{V}(\boldsymbol{x}(t))\,\mathrm{d}t < 0.$$

If we assume that $V(\boldsymbol{x}(t))$ does not approach 0 as $t \to \infty$, then there is $L > 0$ such that $V(\boldsymbol{x}(t)) \geq L$ for all $t \geq 0$.

Since V is continuous and $V(\mathbf{0}) = 0$ there is $h > 0$ such that $V(\boldsymbol{x}) < L$ for all $\boldsymbol{x} \in B_h(\mathbf{0})$. Therefore, for all $t \geq 0$ we have

$$\boldsymbol{x} \in A' = \{\boldsymbol{x} \in \mathbf{R}^2 : h \leq \|\boldsymbol{x}\| \leq \varepsilon\}.$$

Let $\gamma = \max_{\boldsymbol{x} \in A'} \dot{V}(\boldsymbol{x})$. Since $\dot{V}$ is negative definite in A' and $\mathbf{0} \notin A'$ we have $\dot{V} < 0$ on A' and thus $\gamma < 0$.

Then

$$V(\boldsymbol{x}(t)) - V(\boldsymbol{x}(0)) = \int_0^t \dot{V}(\boldsymbol{x}(s))\,\mathrm{d}s \leq \gamma t.$$

This gives $V(\boldsymbol{x}(t)) \leq V(\boldsymbol{x}(0)) + \gamma t$ and therefore for large enough t we find $V(\boldsymbol{x}(t)) < 0$ which contradicts that V is positive definite. Therefore, $V(\boldsymbol{x}(t))$ must approach 0 as $t \to \infty$ and thus $\mathbf{0}$ is asymptotically stable. ☑

Remark 6.28. We can apply Theorem 6.26 to any equilibrium $\boldsymbol{x}_e$ by translating the coordinates on the plane so that $\boldsymbol{x}_e$ moves to the origin, that is, by defining $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{x}_e$ and considering the equation of motion for $\boldsymbol{y}$ which is $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y} + \boldsymbol{x}_e)$. ❞

Example 6.29. Consider the system

$$\begin{aligned} x' &= y - xy^2 - x^3, \\ y' &= -x - x^2y - y^3. \end{aligned}$$

Then $\boldsymbol{x}_0 = (0,0)$ is an isolated equilibrium. Let $V(x,y) = x^2 + y^2$. The function V is continuous and positive definite on $\mathbf{R}^2$, and it satisfies $V(0,0) = 0$. We compute

$$\begin{aligned} \dot{V}(x,y) &= \frac{\partial V}{\partial x}(y - xy^2 - x^3) + \frac{\partial V}{\partial y}(-x - x^2y - y^3) \\ &= 2x(y - xy^2 - x^3) + 2y(-x - x^2y - y^3) \\ &= -2(x^4 + y^4). \end{aligned}$$

The function $\dot{V}$ is negative definite on $\mathbf{R}^2$ and therefore we conclude that the origin is asymptotically stable.

Example 6.30 (Damped pendulum). Consider the pendulum under the influence of a friction force which is proportional to the velocity. This can be modeled by the system

$$\begin{aligned} x' &= y, \\ y' &= -by - \sin x, \; b > 0. \end{aligned} \tag{6.14}$$

The equilibria of this system are $(k\pi, 0)$, $k \in \mathbf{Z}$. Since the equations are 2π periodic in x, there are effectively two distinct equilibria $(0,0)$ and $(\pi, 0)$.

Let us first understand the local behavior near $(\pi, 0)$ — remember that this corresponds to the unstable upper position of the pendulum. Linearization at $(\pi, 0)$ produces the matrix

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & -b \end{bmatrix},$$

with $\det A_1 = -1 < 0$. Therefore, this equilibrium is a saddle for all values of b.

The equilibrium $(0,0)$ is a center for $b = 0$. Linearization at $(0,0)$ for $b > 0$ produces the matrix

$$A_0 = \begin{bmatrix} 0 & 1 \\ -1 & -b \end{bmatrix},$$

with $\det A_0 = 1 < 0$ and $\operatorname{tr} A_0 = -b < 0$. Therefore, the equilibrium is a stable spiral for small b and a stable node for larger b.

We will now use Lyapunov's method to examine the stability of $(0,0)$. One question is which function to choose as a Lyapunov function V. However, since the term $-by$ in the equation for y' causes energy dissipation, a natural choice seems to be to choose V as the energy

$$E(x,y) = \frac{1}{2}y^2 - \cos x.$$

There is only one, easily fixable, issue. We have $E(0,0) = -1$, and to make $E(x,y)$ positive definite near $(0,0)$ we define

$$V(x,y) = E(x,y) + 1 = \frac{1}{2}y^2 + (1 - \cos x). \tag{6.15}$$

Then $V(0,0) = 0$. We have $V(x,y) > 0$ for $y \neq 0$ and $\cos x \neq 2kn, \quad k \in \mathbb{Z}$, so $V(x,y)$ is positive definite is a domain containing $(0,0)$. We have

$$\dot{V}(x,y) = \frac{\partial V}{\partial x} f + \frac{\partial V}{\partial y} g = \sin x \cdot y + y \cdot (-by - \sin x) = -by^2 \leq 0$$

since $b > 0$. Here $\dot{V}$ is negative semidefinite so we can conclude that $(0,0)$ is stable.

Even though the linear stability analysis shows that $(0,0)$ is asymptotically stable, Lyapunov's method using V in Equation (6.15) only gives us that $(0,0)$ is stable. This is a case where linear stability analysis is easier *and* more powerful than Lyapunov's method.

Example 6.31. We will show that the origin is a stable equilibrium for the system

$$\begin{aligned} x' &= -2y^3, \\ y' &= x - 3y^3, \end{aligned}$$

using Lyapunov's method with Lyapunov function $V(x,y) = x^2 + y^4$.

The point $(0,0)$ is the only equilibrium of the given system, and thus it is isolated. Clearly, $V(x,y)$ is positive definite on $\mathbf{R}^2$, since V is continuous, $V(0,0) = 0$ and $V(x,y) > 0$ for all $(x,y) \in \mathbf{R}^2 \setminus \{(0,0)\}$. Then we compute

$$\dot{V}(x,y) = 2x(-2y^3) + 4y^3(x - 3y^3) = 4xy^3 + 4xy^3 - 12y^6 = -12y^6.$$

Therefore, $\dot{V}(x,y)$ is negative semi-definite on $\mathbf{R}^2$ since it is continuous with $\dot{V}(0,0) = 0$ and $\dot{V}(x,y) \leq 0$ for all $(x,y) \in \mathbf{R}^2$. These properties of V and $\dot{V}$ imply that the origin $(0,0)$ is stable.

6.5 Periodic Solutions

We have already seen several examples of planar dynamical systems having periodic solutions, that is, solutions $\boldsymbol{x}(t)$ for which there is some $T > 0$ such that $\boldsymbol{x}(t+T) = \boldsymbol{x}(t)$ for all $t \in \mathbf{R}$. The number T is called a period.

Equilibrium solutions have the property that $\boldsymbol{x}(t+T) = \boldsymbol{x}(t)$ for all $t \in \mathbf{R}$ and *all* $T \in \mathbf{R}$. Therefore, they are also periodic solutions, but in a trivial manner. We are interested in *nontrivial periodic solutions*, that is, solutions for which there exists a *minimal* $T > 0$ satisfying $\boldsymbol{x}(t+T) = \boldsymbol{x}(t)$ for all $t \in \mathbf{R}$. Nontrivial periodic solutions correspond to simple closed curves in $\mathbf{R}^2$, that is, closed curves without any self-intersections.

In this section we discuss general properties of periodic solutions and limit cycles, starting with a simple example, and then we discuss Bendixson's negative criterion establishing the non-existence of periodic solutions in a simply connected domain, and the Poincaré-Bendixson theorem establishing the existence of periodic solutions in a region in $\mathbf{R}^2$.

6.5.1 Limit cycles

We have seen examples of systems, such as the simple pendulum or the Lotka-Volterra system which have several periodic solutions. However, there are systems, such as the van der Pol oscillator, which have isolated periodic solutions. In this part we consider the latter type of systems.

We consider a simple example of a planar system with an isolated periodic solution. The system is given by

$$\begin{aligned} x' &= x - y - x(x^2 + y^2), \\ y' &= x + y - y(x^2 + y^2). \end{aligned} \tag{6.16}$$

The system described by Eq. (6.16) has a single equilibrium, the origin $x = y = 0$. The linearization of the system at the origin gives the matrix

$$A_0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

with eigenvalues $1 \pm i$. Therefore, the origin is an unstable spiral.

We now want to understand what happens further away from the equilibrium. Written in this form, Eq. (6.16) does not appear amenable to our previous methods. However, passing to polar coordinates, the radial and angular dynamics decouple and we can understand the dynamics of the whole system.

Applying Eq. (4.12) to the system in Eq. (6.16) we find

$$\begin{aligned} rr' &= x\left(x - y - x(x^2 + y^2)\right) + y\left(x + y - y(x^2 + y^2)\right) = r^2 - r^4, \\ r^2\theta' &= x\left(x + y - y(x^2 + y^2)\right) - y\left(x - y - x(x^2 + y^2)\right) = r^2, \end{aligned}$$

therefore,

$$r' = r(1 - r^2), \qquad \theta' = 1.$$

The equation $\theta' = 1$ shows that for $r \neq 0$, the angle θ changes at a constant rate, and the solution is $\theta(t) = \theta_0 + t$, where $\theta(0) = \theta_0$.

The radial equation $r' = r(1 - r^2)$ is a first-order separable equation. Moreover, we can observe that with the substitution $u = r^2$ we find $u' = 2u(1 - u)$, that is, the evolution of u is described by the logistic equation which we have solved earlier. Even without solving the equation $r' = r(1 - r^2)$, we can understand the radial dynamics through the phase half-line $r \geq 0$.

In particular, the radial dynamics has two equilibria: one at $r = 0$ and one at $r = 1$. The value $r = 0$ corresponds to the origin in the xy plane which we have earlier determined to be an equilibrium. The value $r = 1$, however, corresponds to the unit circle. Since we also have $\theta' = 1$, the solution moves counterclockwise on the unit circle and returns to the initial point after every 2π period. This is an example of an isolated periodic solution.

Exercise 6.2. Show with a direct computation that $x(t) = \cos(t + \theta_0)$, $y(t) = \sin(t + \theta_0)$, solves Eq. (6.16).

Since $r = 1$ is an asymptotically stable point for the radial dynamics, for an initial point with $0 < r_0 < 1$ the corresponding solution $r(t)$ for the radial dynamics has the property that $r(t) < 1$ with

$$\lim_{t\to\infty} r(t) = 1.$$

Combined with the fact that the solution on the xy plane always rotates counterclockwise with $\theta' = 1$ we can conclude that the solution spirals while approaching the unit circle from inside. Similarly, for an initial point with $r_0 > 1$ we again have

$$\lim_{t\to\infty} r(t) = 1,$$

but with $r(t) > 1$. Therefore, we can conclude that the solution spirals while approaching the unit circle from outside.

Combining the information we obtained about the radial and angular dynamics of the system, we can now draw its phase portrait, shown in Figure 6.9.

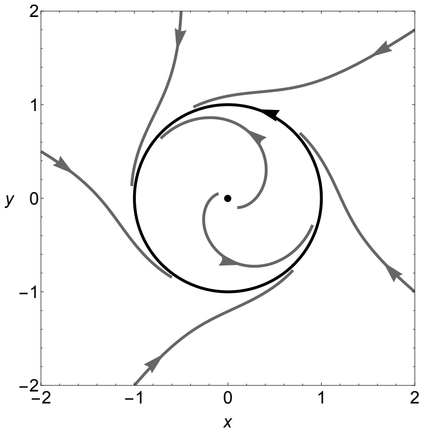


Figure 6.9: Phase portrait for the system in Equation (6.16).

The periodic orbit at $r = 1$ in the previous section is an example of a limit cycle.

Definition 6.32. A nontrivial periodic solution with at least one other solution curve spiraling into it as $t \to \infty$ or as $t \to -\infty$ is called a *limit cycle*.

Remark 6.33. The example in the previous section is rather atypical in the sense that the limit cycle is a circle. In general, a limit cycle does not need to be a circle. ❞

An easy source of examples of limit cycles, both asymptotically stable and unstable, is provided by considering planar systems of the form

$$\begin{aligned} x' &= x - y - xf(r), \\ y' &= x + y - yf(r). \end{aligned} \tag{6.17}$$

Notice, that Eq. (6.16) has exactly the form in Eq. (6.17) with $f(r) = r^2$. We can check that $\theta' = 1$ and that

$$r' = r(1 - f(r)).$$

Therefore, the specific form of $f(r)$ in Eq. (6.17) can be used to control the number and stability of periodic orbits of the system.

Example 6.34. We consider the planar system given by

$$\begin{aligned} x' &= x - y - x(3r - r^2 - 1), \\ y' &= x + y - y(3r - r^2 - 1). \end{aligned} \tag{6.18}$$

Then we find

$$r' = r(r^2 - 3r + 2) = r(r - 1)(r - 2).$$

The radial dynamics has equilibria at $r = 0$, $r = 1$, and $r = 2$. The values $r = 1$ and $r = 2$ correspond to periodic solutions in the planar system. Since $r = 1$ is an asymptotically stable

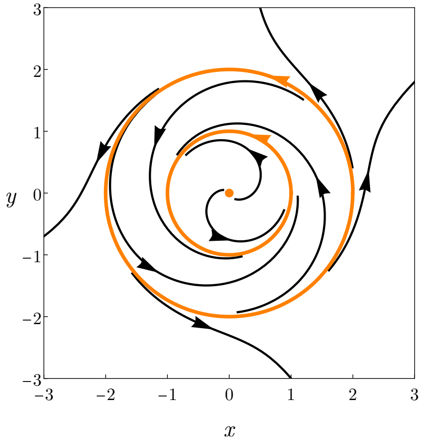


Figure 6.10: Phase portrait for the system in Eq. (6.18). The limit cycle at $r = 1$ is asymptotically stable; the limit cycle at $r = 2$ is unstable.

equilibrium for the radial dynamics we conclude that the corresponding periodic solution is also asymptotically stable. Similarly, since $r = 2$ is an unstable equilibrium for the radial dynamics we conclude that the corresponding periodic orbit is also unstable. We can now draw the phase portrait, shown in Fig. 6.10.

Limit cycles impose restrictions on the number and type of equilibria that they may enclose. Without going into any detail, we give the following basic result.

Theorem 6.35. *A limit cycle on the plane must enclose at least one equilibrium. If a limit cycle encloses exactly one equilibrium then this equilibrium cannot be a saddle.*

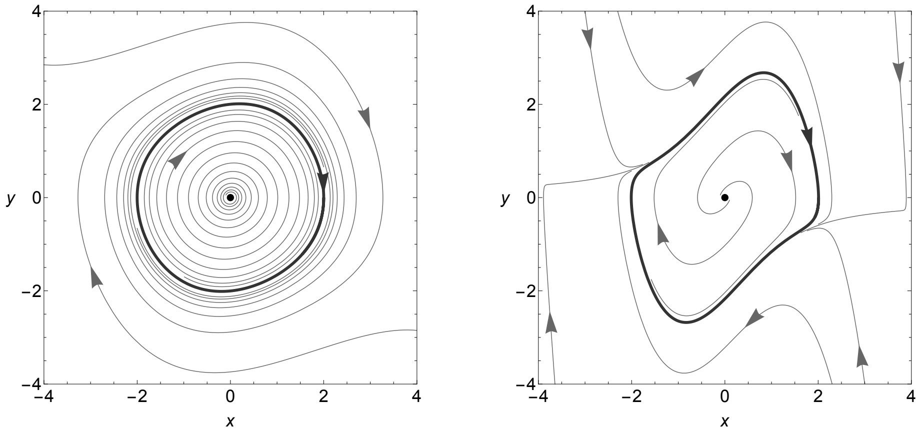


Figure 6.11: Numerically computed phase portrait for the van der Pol oscillator in Equation (6.19) with $\mu = 0.1$ (left) and $\mu = 1$ (right).

Example 6.36 (Van der Pol oscillator). Consider the van der Pol oscillator $x'' + x + \mu(x^2 - 1)x' = 0$, where $\mu \geq 0$. Written as a planar system, the equation for the van der Pol oscillator becomes

$$
\begin{aligned}
x' &= y, \\
y' &= -x - \mu(x^2 - 1)y.
\end{aligned}
\tag{6.19}
$$

For $\mu = 0$, the system is linear, and it is not difficult to see that the origin is a center. As μ increases, a numerical computation of the phase portrait reveals the appearance of an asymptotically stable limit cycle. The phase portraits for $\mu = 0.1$ and $\mu = 1$ are shown in Figure 6.11.

6.5.2 Bendixson's Negative Criterion

The following theorem gives us conditions under which we can deduce that the system does not have a nontrivial periodic solution in a simply connected domain D.

Theorem 6.37 (Bendixson's Negative Criterion). *Assume that $\boldsymbol{f} : \mathbf{R}^2 \to \mathbf{R}^2$ has continuous first partial derivatives in a simply connected domain $D \subseteq \mathbf{R}^2$ and that $\nabla \boldsymbol{f}(\boldsymbol{x})$ is either strictly positive or strictly negative on D. Then the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ does not have any non-trivial periodic orbits that lie entirely in D.*

Proof. Suppose that there is a non-trivial periodic orbit $\boldsymbol{\gamma}(t) = (x_1(t), x_2(t))$ in D with period T. Let $\Gamma = \{\gamma(t) : 0 \leq t \leq T\}$ and denote by Ω the subdomain of D enclosed by Γ. Notice that since D is simply connected, Ω is also simply connected.

Applying Green's theorem to $\boldsymbol{f}$ on the simply connected domain Ω we find that

$$
\int_{\boldsymbol{\gamma}} f_1 \, \mathrm{d}x_2 - f_2 \, \mathrm{d}x_1 = \int_{\Omega} \nabla \boldsymbol{f} \, \mathrm{d}x_1 \, \mathrm{d}x_2.
$$

For the left hand side of the above equation we have

$$
\begin{aligned}
\int_{\boldsymbol{\gamma}} f_1 \, \mathrm{d}x_2 - f_2 \, \mathrm{d}x_1 &= \int_{\boldsymbol{\gamma}} f_1 \, \mathrm{d}x_2 - \int_{\boldsymbol{\gamma}} f_2 \, \mathrm{d}x_1 \\
&= \int_0^T f_1(\boldsymbol{\gamma}(t)) \frac{\mathrm{d}x_2}{\mathrm{d}t} \, \mathrm{d}t - \int_0^T f_2(\boldsymbol{\gamma}(t)) \frac{\mathrm{d}x_1}{\mathrm{d}t} \, \mathrm{d}t \\
&= \int_0^T \left[f_1(\boldsymbol{\gamma}(t)) f_2(\boldsymbol{\gamma}(t)) - f_2(\boldsymbol{\gamma}(t)) f_1(\boldsymbol{\gamma}(t)) \right] \mathrm{d}t \\
&= 0.
\end{aligned}
$$

However, for the right hand side we have that since $\nabla \boldsymbol{f}$ is either strictly positive or strictly negative on D and thus also on $\Omega \subseteq D$, then the integral $\int_\Omega \nabla \boldsymbol{f}(x, y) \, dx \, dy$ is also either strictly positive or strictly negative. In particular, this integral is never zero, leading to a contradiction, and we conclude that the system cannot have a non-trivial periodic orbit lying entirely in D.

Example 6.38. Consider the planar system on $\mathbf{R}^2$ given by

$$
\begin{aligned}
x' &= y, \\
y' &= -g(x)y + f(x),
\end{aligned}
$$

where f, g are continuously differentiable on $\mathbf{R}$ and $g(x) > 0$ for all $x \in \mathbf{R}$. We have

$$
\nabla \boldsymbol{f} = \frac{\partial y}{\partial x} + \frac{\partial(-g(x)y - f(x))}{\partial y} = -g(x) < 0.
$$

Therefore, the given system does not have any nontrivial periodic solutions on $\mathbf{R}^2$.

Exercise 6.3. Use Bendixson's negative criterion to show that planar linear systems $\boldsymbol{x}' = A\boldsymbol{x}$ with $\operatorname{tr} A \neq 0$ cannot have nontrivial periodic solutions.

Exercise 6.4. Use Bendixson's negative criterion to show that if $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ is a planar system with $\boldsymbol{f}(\boldsymbol{0}) = \boldsymbol{0}$ and $\operatorname{tr} D\boldsymbol{f}(\boldsymbol{0}) \neq 0$, then there is a neighborhood D of $\boldsymbol{0}$ which does not contain any nontrivial periodic solutions.

Remark 6.39. The simple pendulum with damping has the form considered in Example 6.38, with $f(x) = -\sin x$ and $g(x) = b > 0$. Therefore, the simple pendulum with damping does not have any nontrivial periodic solutions.

6.5.3 Poincaré-Bendixson Theorem

The following theorem gives a criterion to prove the existence of a periodic solution.

Theorem 6.40 (Poincaré-Bendixson). *Assume that $\boldsymbol{f}$ has continuous first partial derivatives in a closed bounded region R and the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ has no equilibria in R. Then any solution of the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ which stays in R for all t larger than some $t_0 \in R$ is either periodic or it approaches a limit cycle in R.*

Example 6.41. Consider the planar system

$$\begin{aligned} x' &= y, \\ y' &= -x^3 - (4x^2 + y^2 - 4)y. \end{aligned}$$

The system has only one equilibrium, the origin. To show that the system has a periodic orbit we will find a region R in which we can apply Theorem 6.40. Such region R should have the property that solution curves stay in R and that it contains no equilibria.

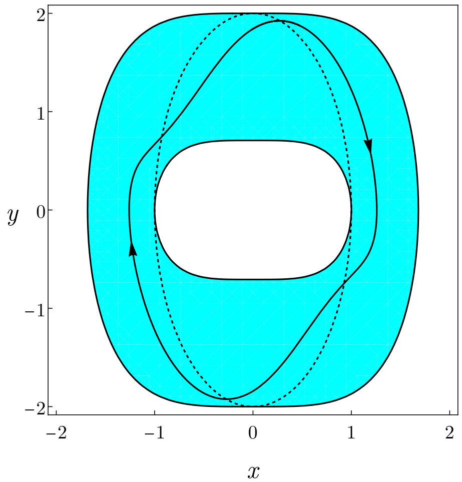


Figure 6.12: Trapping region (in cyan) and numerically computed limit cycle for Example 6.41. The dashed line represents the ellipse $4x^2 + y^2 = 4$.

To construct R we use an idea from Lyapunov's direct method for stability. Consider the positive definite function

$$V(x, y) = \frac{1}{2}y^2 + \frac{1}{4}x^4.$$

Then we compute

$$\dot{V}(x,y) = \boldsymbol{f} \cdot \nabla V = \frac{\partial V}{\partial x} f_1 + \frac{\partial V}{\partial y} f_2 = -(4x^2 + y^2 - 4)y^2.$$

Therefore, $\dot{V}(x,y) \geq 0$ on $P = \{(x,y) : 4x^2 + y^2 \leq 4\}$ and $\dot{V}(x,y) \leq 0$ on $N = \{(x,y) : 4x^2 + y^2 \geq 4\}$.

In particular, this shows that if there is a level set of V that is contained entirely inside P then a solution curve cannot cross it toward level sets with smaller values of V since on P we have $\dot{V} \geq 0$. Similarly, if there is a level set of V that is contained entirely inside N then a solution curve cannot cross it toward level sets with larger values of V since on N we have $\dot{V} \leq 0$.

We define the level set C_1 by $V(x,y) = 1/4$ and the level set C_2 by $V(x,y) = 2$. Then we can show that C_1 is in P, while C_2 is in N. To see this, consider a point (x,y) with $V(x,y) = R$. Then $y^2 = 2R - x^4/2$, and we have

$$-(4x^2 + y^2 - 4) = \frac{1}{2}x^4 - 4x^2 + 4 - 2R =: c(x^2).$$

To ensure that $\dot{V}(x,y)$ does not change sign on the level set $V(x,y) = R$, it is sufficient to ensure that the quadratic function $c(\xi) = \xi^2/2 - 4\xi + 4 - 2R$ does not change sign for $0 \leq \xi \leq 2\sqrt{R}$. The roots of $c(\xi) = 0$ are $\xi_1 = 4 - 2\sqrt{R+2}$, $\xi_2 = 4 + 2\sqrt{R+2}$. For $R \geq 2$ we have $\xi_1 \leq 0$, $\xi_2 \geq 2\sqrt{R}$, and $c(0) = 2(2-R) \leq 0$. Therefore, for $R \geq 2$ we have $c(\xi) \leq 0$ and correspondingly $\dot{V} \leq 0$. For $R \leq 1/4$ we have $2\sqrt{R} \leq \xi_1 < \xi_2$, and $c(0) = 2(2-R) \geq 0$. Therefore, for $R \leq 1/4$ we have $c(\xi) \geq 0$ and correspondingly $\dot{V} \geq 0$.

The annulus R, between the level sets C_1 and C_2, defined by

$$R = \left\{(x,y) : \frac{1}{4} \leq V(x,y) \leq 2\right\},$$

is a closed bounded region and if a solution curve is in R then it cannot escape. Moreover, since R does not contain the origin, it does not contains any equilibria. Therefore, the Poincaré-Bendixson theorem asserts that the system has at least one non-trivial periodic orbit in R.

Example 6.42. Consider the system

$$\begin{aligned} r' &= r(1-r^2) + \mu r \cos\theta, \\ \theta' &= 1. \end{aligned}$$

Recall that for $\mu = 0$ the system has an asymptotically stable limit cycle at $r = 1$. We will use the Poincaré-Bendixson theorem to show that the system has a limit cycle for sufficiently small $\mu > 0$.

Consider the function $V = r$, so that

$$\dot{V} = r(1-r^2) + \mu r \cos\theta.$$

Therefore,

$$r(1 - r^2 - \mu) \leq \dot{V} \leq r(1 - r^2 + \mu).$$

Choose a level set $V = r_1$ such that $\dot{V} > 0$. For this it is sufficient that $1 - r_1^2 - \mu > 0$, i.e, $r_1^2 < 1 - \mu$, provided that $0 < \mu < 1$. Similarly, choose a level $V = r_2$ such that $\dot{V} < 0$. For this it is sufficient that $1 - r_2^2 + \mu < 0$, i.e., $r_2^2 > 1 + \mu$.

The annulus R defined by $r_1 \leq r \leq r_2$ is a closed bounded region with the property that if a solution curve is in R then it cannot escape. Moreover, R contains no equilibria. Therefore, the Poincaré-Bendixson theorem asserts that the system has at least one non-trivial periodic orbit in R.

Example 6.43. Consider the system

$$\begin{aligned} x' &= y, \\ y' &= 2 - y - \sin x, \end{aligned}$$

where x is an angle and $y \in \mathbf{R}$. Since x is an angle, the correct representation of the phase of the system is not as the plane $\mathbf{R}^2$ but as a cylinder $S^1 \times \mathbf{R}$ with $x \in S^1$ and $y \in \mathbf{R}$.

Consider the function $V = y$. Then

$$\dot{V} = 2 - y - \sin x.$$

Therefore,

$$1 - y \leq \dot{V} \leq 3 - y.$$

Choose the level set $V = y = 0.99$ so that $\dot{V} > 0$ and the level set $V = y = 3.01$ so that $\dot{V} < 0$. The region R defined by $0.99 \leq y \leq 3.01$ is closed and bounded — notice that $R = S^1 \times [0.99, 3.01]$. Moreover, any solution curves in R cannot escape, and R contains no equilibria. Therefore, the Poincaré-Bendixson theorem asserts that the system has at least one non-trivial periodic orbit in R.

The periodic orbit in this case is a closed curve only when we think of the phase space as a cylinder.

Remark 6.44. The system in Example 6.43 appears to have the same form as the system studied in Example 6.38, with $f(x) = 2 - \sin x$ and $g(x) = 1 > 0$. But in Example 6.38 we showed that such systems have no periodic orbits while in Example 6.43 we have just seen that the system has a periodic orbit. Is there a contradiction?

The answer is no. If, on the one hand, we consider the system as a system on $\mathbf{R}^2$ then it indeed does not have any closed orbits. What appears as a periodic orbit on $S^1 \times \mathbf{R}$, appears as an unbounded orbit on $\mathbf{R}^2$. If, on the other hand, we consider the system on $S^1 \times \mathbf{R}$ then we note that $S^1 \times \mathbf{R}$ is not simply connected — a closed curve going around the cylinder cannot be shrunk to a point. Therefore, one cannot apply Bendixson's negative criterion in this case and the conclusion we reached in Example 6.38 is not valid.

6.6 Stability of Periodic Solutions

To understand the stability of a periodic orbit we reduce the problem to checking the stability of a fixed point of an appropriately defined map. Consider a non-trivial periodic orbit parameterized as $\boldsymbol{\gamma}(t)$, $0 \leq t \leq T$, and tracing a closed curve Γ on the plane. Then consider a line segment S intersecting the curve Γ transversally at $\boldsymbol{p}_0 := \boldsymbol{\gamma}(0)$. We can define a map in a neighborhood of $\boldsymbol{p}_0$ on S in the following way. Consider a point $\boldsymbol{p}$ on S. Since the solution curve starting at $\boldsymbol{p}_0$ reaches S again at the same point $\boldsymbol{p}_0$ after time T, we must have that if $\boldsymbol{p}$ is sufficiently close to $\boldsymbol{p}_0$, then the corresponding solution curve starting at $\boldsymbol{p}$ will reach S again after some time close to T and at a point $\boldsymbol{p}'$. This construction gives rise to a map $F : S \to S$ given by $F(\boldsymbol{p}) = \boldsymbol{p}'$. The map F is also called a Poincaré map.

Notice that $\boldsymbol{p}_0$ is a fixed point of F. If the solution curve starting at the point $\boldsymbol{p} \in S$ asymptotically approaches the periodic orbit Γ, then the Poincaré map iterates $F^n(\boldsymbol{p})$ asymptotically approach $\boldsymbol{p}_0$ as n increases. Similarly, if the solution curve starting at the point $\boldsymbol{p} \in S$ moves away the periodic orbit Γ, then the Poincaré map iterates $F^n(\boldsymbol{p})$ move away from $\boldsymbol{p}_0$ as n increases. The two cases correspond to $\boldsymbol{p}_0$ being an asymptotically stable or an unstable fixed point of F respectively.

Therefore, analyzing the stability of the fixed point of F allows us to determine the behavior of solution curves near the periodic orbit.

To make the study more quantitative we can consider a coordinate system in an annulus containing S. Let s denote a coordinate along S, with $s = 0$ at $\boldsymbol{p}_0$, and extend it so that $s = 0$ along the periodic orbit. Let ϕ denote a coordinate that is periodic and increases along the periodic solution from 0 to 2π. Then in terms of s and ϕ we can write equations of motion and then obtain the reduced equation

$$\frac{\mathrm{d}s}{\mathrm{d}\phi} = h(s, \phi).$$

Since the periodic solution is $s = 0$ we must have $h(0, \phi) = 0$. Consider now a small deviation s from $s = 0$. We have

$$\frac{\mathrm{d}s}{\mathrm{d}\phi} = h(s, \phi) = h(0, \phi) + \frac{\partial h}{\partial s}(0, \phi)s + \cdots .$$

Keeping only linear terms we find

$$\frac{\mathrm{d}s}{\mathrm{d}\phi} = \frac{\partial h}{\partial s}(0, \phi)\, s =: g(\phi)\, s.$$

The last equation can be integrated using separation of variables. We find that after ϕ changes by 2π we have

$$s(2\pi) = s(0)e^{\int_0^{2\pi} g(\phi)\,\mathrm{d}\phi}.$$

The last expression gives the linearization of the Poincaré map F at the fixed point and thus determines the linear stability of the latter.

Example 6.45. Consider the system $r' = r(1 - r^2)$, $\theta' = 1$ which we explored at the beginning of this section. We can solve the equations explicitly and understand the dynamics, however, we want to find the linearization of the Poincaré map. Here choose S to be the semiaxis Ox, and let $s = r - 1$, and $\phi = \theta$. Then the dynamics is given by $s' = -s(s+1)(s+2)$, $\phi' = 1$. The reduced equation is

$$\frac{\mathrm{d}s}{\mathrm{d}\phi} = -s(s + 1)(s + 2) = -s^3 - 3s^2 - 2s.$$

Therefore, with $h(s, \phi) = -s^3 - 3s^2 - 2s$ we have $h_s(s, \phi) = -3s^2 - 6s - 2$ and $h_s(0, \phi) = -2$. This means that $s(2\pi) = s(0)e^{-4\pi} \approx 3.5 \times 10^{-6}s(0)$ and shows that the fixed point is (strongly) asymptotically stable.

Example 6.46. We revisit the van der Pol oscillator, Equation (6.19). To define the Poincaré map we consider the half-line $S = \{(0, y) : y \geq 0\}$. Then for $y \geq 0$ we define the one-dimensional Poincaré map $P_\mu : S \to S$ by considering the solution curve with initial condition $(0, y) \in S$ until it reaches again S at a point $(0, P_\mu(y))$.

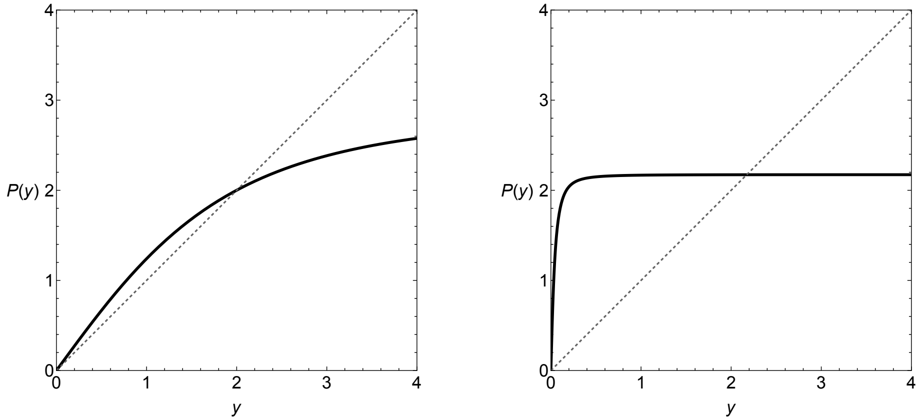


Figure 6.13: Numerically computed Poincaré map for the van der Pol oscillator in Equation (6.19) with $\mu = 0.1$ (left) and $\mu = 1$ (right). The dashed line represents the identity map.

The following Mathematica code numerically computes the Poincaré map. For a given value of μ, the function `poincareMap[mu]` takes as input a number $y \geq 0$ and returns the value of the Poincaré map $P_\mu(y)$. To compute $P_\mu(y)$, the function integrates the solution curve starting at the initial point $(0, y) \in S$ until it reaches again S after some time T. Since the value of T is not known a priori, the integration continues until the conditions $x = 0$ with $y \geq 0$ are satisfied, and then the value of y at that moment is returned.

```
poincareMap[mu_][y0_?NumericQ] := Module[{stopTime, x, y},
  NDSolveValue[
    x'[t] == y[t]
    && y'[t] == - x[t] - mu*(x[t]^2-1)*y[t]
    && x[0] == 0
    && y[0] == y0
    && WhenEvent[x[t] == 0 && y[t] >= 0, stopTime = t; "StopIntegration"],
   y[stopTime],
   {t, 0, Infinity}]]
```

With the function `poincareMap[mu]` in hand, we can compute the graph of $P_\mu(y)$ for different values of μ. Figure 6.13 shows the graph of $P_\mu(y)$ for $\mu = 0.1$ and $\mu = 1$ and it may be instructive to compare it with the phase portraits in Figure 6.11.

Moreover, we can numerically compute the non-zero fixed point y_μ^* of the Poincaré map P_μ, which correspond to the point where the limit cycle of the van der Pol oscillator intersects the half-line S. The non-zero fixed point y_μ^* is asymptotically stable for $\mu > 0$. In particular, for $\mu = 0.1$ we numerically compute that $P_\mu'(y_\mu^*) \approx 0.5330$, while for $\mu = 1$ the corresponding value is $P_\mu'(y_\mu^*) \approx 0.93 \times 10^{-3}$. This implies that the corresponding limit cycle is also asymptotically stable.

6.7 Hopf Bifurcation

Consider a planar dynamical system that depends on parameter $a \in \mathbf{R}$ and is given by

$$\boldsymbol{x}' = \boldsymbol{f}(a, \boldsymbol{x}). \tag{6.20}$$

Suppose that when the parameter is a_0 the system in Eq. (6.20) has an equilibrium at $\boldsymbol{x}_0$.

The implicit function theorem, stated in its general form as Theorem 2.24, has the following form for functions from $\mathbf{R}^3$ to $\mathbf{R}^2$.

Theorem 6.47 (Implicit Function Theorem for functions $\mathbf{R}^3 \to \mathbf{R}^2$). *Consider a smooth function $\boldsymbol{f} : \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2$ such that $\boldsymbol{f}(a_0, \boldsymbol{x}_0) = \mathbf{0}$. If $\det D_{\boldsymbol{x}} \boldsymbol{f}(a_0, \boldsymbol{x}_0) \neq 0$ then there is $\delta > 0$ and a unique smooth function $\boldsymbol{h} : (a_0 - \delta, a_0 + \delta) \to \mathbf{R}^2$ such that $\boldsymbol{f}(\boldsymbol{a}, \boldsymbol{h}(a)) = 0$ and $\boldsymbol{h}(a_0) = \boldsymbol{x}_0$.*

Theorem 6.47 asserts that equilibria of Eq. (6.20) persist under variations of a unless $\det A_0 = 0$ where $A_0 = D_{\boldsymbol{x}} \boldsymbol{f}(a_0, \boldsymbol{x}_0)$. Notice that A_0 is the matrix of the linearization of the dynamics at the equilibrium $\boldsymbol{x}_0$. Recalling the trace-determinant plane in Section 4.5, this means that an equilibrium may disappear or give rise to new equilibria only when A_0 crosses the horizontal axis in the trace-determinant plane as a varies. This can alternatively be characterized in terms of the eigenvalues of A_0 since the condition $\det A_0 = 0$ is equivalent to the condition that at least one of the eigenvalues of A_0 vanishes. We do not discuss this type of bifurcations further in these notes. Instead, we want to turn our attention to a different type of bifurcation that occurs when the equilibrium changes stability from stable to unstable (or vice versa) as the parameter a varies.

Suppose that the system in Eq. (6.20) has for each a in some interval $I \subseteq \mathbf{R}$ an equilibrium $\boldsymbol{x}_0(a)$ depending smoothly on a. Then the matrix of the linearization at the equilibrium $\boldsymbol{x}_0(a)$ is

$$A(a) := D_{\boldsymbol{x}} \boldsymbol{f}(a, \boldsymbol{x}_0(a)).$$

The matrix $A(a)$, its determinant $\det A(a)$ and trace $\operatorname{tr} A(a)$, and its eigenvalues $r_1(a)$, $r_2(a)$ all depend smoothly on a. Assume that for all $a \in I$, $\det A(a) > 0$. Then the equilibrium $\boldsymbol{x}_0(a)$ is stable when $\operatorname{tr} A(a) < 0$ and unstable when $\operatorname{tr} A(a) > 0$. For the equilibrium to change stability type there must be a parameter $a_0 \in I$ such that $\operatorname{tr} A(a_0) = 0$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}a}\right|_{a=a_0} \operatorname{tr} A(a) \neq 0.$$

Recall that when $\det A(a) > 0$ and $\operatorname{tr} A(a)$ is sufficiently close to zero, the eigenvalues of $A(a)$ are complex conjugate $\alpha(a) \pm i\beta(a)$ with $\alpha(a) = \frac{1}{2} \operatorname{tr} A(a)$. Therefore, the condition for the equilibrium $\boldsymbol{x}_0(a)$ to change stability type at a_0 is

$$\alpha(a_0) = 0 \quad \text{and} \quad \alpha'(a_0) \neq 0.$$

When the equilibrium changes stability type we have another interesting phenomenon that cannot be detected in the linearized system. To understand what happens we consider an example before briefly discussing the general theory.

Consider the planar system

$$\begin{aligned} x' &= ax - y - kx(x^2 + y^2), \\ y' &= x + ay - ky(x^2 + y^2), \end{aligned} \tag{6.21}$$

depending on a parameter $a \in \mathbf{R}$. The number k takes only the values ± 1 and distinguishes between two different bifurcation scenarios.

The origin is an equilibrium of the system in Eq. (6.21). The linearized dynamics at the origin is given by the matrix

$$A = \begin{bmatrix} a & -1 \\ 1 & a \end{bmatrix},$$

with complex eigenvalues $a \pm i$, that is $\alpha(a) = a$. Therefore, at $a = 0$ we have $\alpha(0) = 0$ and $\alpha'(0) = 1$, and thus the equilibrium changes its stability type from stable spiral to unstable spiral as a changes sign from negative to positive.

The complete dynamics of the system can be analyzed by using polar coordinates (r, θ) with $x = r\cos\theta$ and $y = r\sin\theta$. In polar coordinates the dynamics is given by the equations

$$r' = r(a - kr^2), \tag{6.22a}$$

$$\theta' = 1. \tag{6.22b}$$

The polar dynamics in Eq. (6.22a) always has an equilibrium at $r = 0$ corresponding to the origin in the xy plane, which is an equilibrium of the system in Eq. (6.21). If Eq. (6.22a) has additional equilibria with $r > 0$ then these correspond to periodic solutions of Eq. (6.21) with period $T = 2\pi$, since $\theta' = 1$. In particular, if $r_0 > 0$ is an equilibrium of Eq. (6.22a) then the solution with initial conditions $r(0) = r_0$, $\theta(0) = \theta_0$ is given in polar coordinates by

$$r(t) = r_0, \ \theta(t) = \theta_0 + t,$$

and in cartesian coordinates by

$$x(t) = r_0 \cos(\theta_0 + t), \ y(t) = r_0 \sin(\theta_0 + t).$$

Therefore, this solution curve traces the circle with radius r_0 and makes one counterclockwise rotation in time 2π.

For $k = 1$, such periodic solutions are obtained by solving $a - r^2 = 0$. The last equation has the solution $r(a) = \sqrt{a} > 0$ when $a > 0$ and it has no solutions when $a < 0$. To determine the stability type of the periodic solution created for $a > 0$ we consider again the polar dynamics. Writing $f(r) = r(a - kr^2)$, we compute $f'(r) = a - 3kr^2$. For $k = 1$, $r = \sqrt{a}$, $a > 0$ we find $f'(\sqrt{a}) = -2a < 0$. Therefore, $r(a) = \sqrt{a}$ is a stable equilibrium of the polar dynamics and thus the corresponding periodic solution is a stable limit cycle. In particular, nearby solution curves approach the limit cycle as $t \to \infty$, while they move away from the limit cycle as t decreases. Therefore, in this case, when a changes sign from negative to positive, the origin changes stability type from stable spiral to unstable spiral and at the same time a stable limit cycle is created with constant radius given by $r(a) = \sqrt{a}$ for $a > 0$. This type of bifurcation, involving a stable limit cycle, is called a *supercritical Hopf bifurcation.*

For $k = -1$, the periodic solutions are obtained by solving $a + r^2 = 0$. The last equation has the solution $r(a) = \sqrt{-a} > 0$ when $a < 0$ and it has no solutions when $a > 0$. For $k = -1$, $r = \sqrt{-a}$, $a < 0$ we find $f'(\sqrt{a}) = -2a > 0$. Therefore, $r(a) = \sqrt{-a}$ is an unstable equilibrium of the polar dynamics and thus the corresponding periodic solution is an unstable limit cycle. In particular, nearby solution curves move away from the limit cycle as t increases, while they approach the limit cycle as $t \to -\infty$. Therefore, in this case, when a changes sign from negative to positive, the origin changes stability type from stable spiral to unstable spiral and at the same time the unstable periodic solution which exists for $a < 0$ with radius $r(a) = \sqrt{-a}$ disappears. This type of bifurcation, where an unstable stable limit cycle is created, is called a *supercritical Hopf bifurcation.*

Therefore, $r(a) = \sqrt{-a}$ is an unstable equilibrium of the polar dynamics and thus the corresponding periodic solution is unstable, meaning that nearby solutions move away from the periodic solution as t increases, while they approach the periodic solution as $t \to -\infty$. This type of bifurcation, involving an unstable limit cycle, is called a *subcritical Hopf bifurcation.*

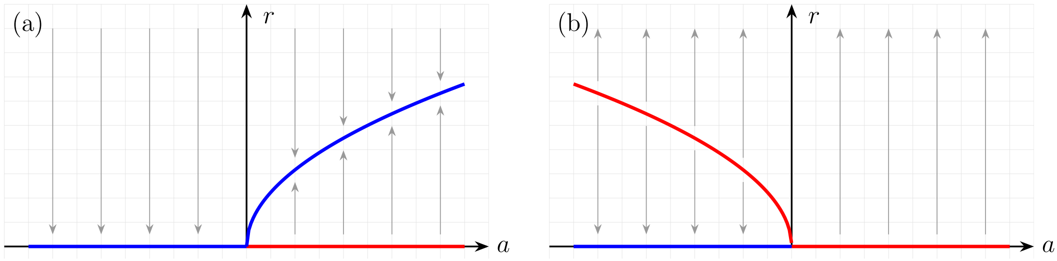


Figure 6.14: Bifurcation diagrams for the two types of Hopf bifurcations: (a) supercritical Hopf bifurcation; (b) subcritical Hopf bifurcation.

Remark 6.48. We emphasize here that what distinguishes supercritical from subcritical Hopf bifurcations is the stability of the limit cycle (unstable in the subcritical case, stable in the supercritical case), and not whether the limit cycle exists for $a > 0$ or $a < 0$. Additionally, notice that the limit cycle, when it exists, has the opposite stability type from the equilibrium. ❞

Bibliography

[1] V. I. Arnol'd. *Ordinary Differential Equations*. 3rd ed. Springer, 1992.
[2] K. Atkinson and W. Han. *Elementary Numerical Analysis*. 3rd ed. Wiley, 2004.
[3] W. E. Boyce and R. C. DiPrima. *Elementary Differential Equations and Boundary Value Problems*. 11th ed. Wiley, 2017.
[4] R. L. Burden, D. J. Faires, and A. M. Burden. *Numerical Analysis*. 10th ed. Cengage, 2016.
[5] J. J. Duistermaat and J. A. C. Kolk. *Multidimensional Real Analysis I: Differentiation*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2004.
[6] J. K. Hale and H. Koçak. *Dynamics and Bifurcations*. Texts in Applied Mathematics 3. Springer, 1991.
[7] P. Hartman. "A lemma in the theory of structural stability of differential equations". In: *Proceedings of the American Mathematical Society* 11.4 (1960), pp. 610–620. URL: http://www.jstor.org/stable/2034720.
[8] Y. A. Kuznetsov. *Elements of Applied Bifurcation Theory*. 3rd ed. Vol. 112. Applied Mathematical Sciences. Springer, 2004. DOI: 10.1007/978-1-4757-3978-7.
[9] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. 2nd ed. CRC Press, 2015.
[10] L. W. Tu. *An introduction to manifolds*. 2nd ed. Springer, 2011. DOI: 10.1007/978-1-4419-7400-6.
[11] Wolfram Research. *Wolfram Language & System Documentation Center*. URL: https://reference.wolfram.com/language/.